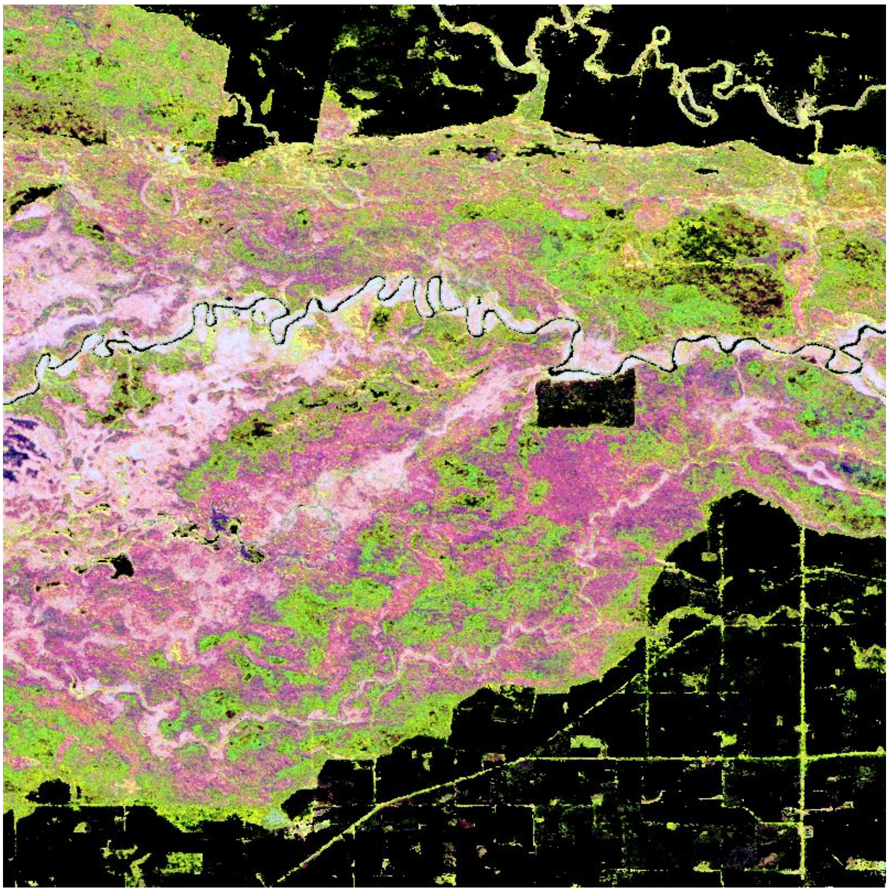
DEVELOPMENT OF A STAND CONDTION MONITORING TOOL FOR THE MURRAY-DARLING BASIN



Graeme Newell, Peter Griffioen and Matt White

Final Milestone Report to the Murray-Darling Basin Authority as part of Contract MD003805

Arthur Rylah Institute for environmental research

Ecoinformatics

Graeme Newell * and Matt White

Arthur Rylah Institute for Environmental Research,

Victorian Department of Environment, Land, Water and Planning, Heidelberg, VIC 3084

Peter Griffioen

Ecoinformatics Pty. Ltd., 20 Alexander Ave, Montmorency, VIC 3094

* Corresponding author: Tel.: +61 3 9450 8600

E-mail address: graeme.newell@delwp.vic.gov.au

This report should be cited as: Newell, G, White M and Griffioen P (2017). Development of a Stand Condition Monitoring Tool for the Murray Darling Basin. Murray-Darling Basin Authority, Canberra.

Executive Summary

Context
The condition of the floodplain and riverine forests of the Murray Darling Basin (the Basin) has been declining over recent decades. These forests provide and support nationally important social, environmental, cultural and material resources. Governments and the community have invested significant resources towards arresting this decline and improving the ecological health of the Basin. Reliable and repeatable monitoring and reporting mechanisms are required to provide evidence to support ongoing investment and decision making.

The Murray Darling Basin Authority MDBA has developed a robust approach to monitoring the condition of riverine and floodplain forest that uses both field data and contemporaneous summaries of satellite imagery. Such modelling allows the MDBA to report on the condition of the basins flood dependent forests on a regular basis. Regulated river basin systems are dynamic and therefore models need to be updated periodically to incorporate new data. This can include the sampling of novel (or previously un-encountered) hydrological events, and/or where new field data has been acquired from previously un-sampled floodplain regions of the basin.

Aims
The primary aim of this project is to deliver an updated 'Stand Condition Monitoring Tool (the Tool) for use by the Murray-Darling Basin Authority (MDBA).

The purpose of this report is to provide:

1. Details of updates for the algorithmic models of stand condition based upon Landsat satellite imagery data and recently acquired field data.
2. Details of the incorporation of these models to deliver the updated Stand Condition Monitoring Tool.
3. Recommendations on the prospects for future improvements to stand condition models, particularly relating to ancillary remote-sensed data, the revision of the base layers used, and improvements to field data collection.

Importantly, this report needs to be read and considered in conjunction with the 'Stand Condition Monitoring Tool Users Guide', which provides details how to install and use the Tool.

Implications
The Tool allows the MDBA to make up-to-date appraisals of the condition of floodplain and riparian-associated native vegetation forms across the Basin. Newly acquired Landsat imagery that are compiled using similar protocols used for this and earlier reports can be incorporated into this Tool. These new spatial data will enable the production of updated stand condition modelled outputs in future years.

Introduction

Notable declines in the condition of the floodplain forests and woodlands have been evident across the Murray-Darling Basin over many decades (Cunningham et al., 2009b). These detrimental changes are associated with river regulation, water extraction for agriculture declines in rainfall across the Basin (Cunningham *et al.* in press).

In 2002 the then Murray-Darling Basin Commission (MDBC) instituted 'The Living Murray' (TLM) program which aimed to restore the health of the Basin by returning water to many of the natural floodplains across the it (MDBC, 2002). The 'TLM' program comprised a variety of activities at a series of Icon Sites including the construction and development of infrastructure supporting positive environmental effects through water recovery, environmental watering and monitoring. The physical and geographic scale of the region led to the decision in 2008 to undertake monitoring of the changes in the environmental condition of forests and woodlands across the Basin through using remote sensing technologies.

Previous stand condition modelling

Several approaches have been taken to assess and report on the condition or quality of native vegetation across the Basin. The initial models and maps of stand condition related to river redgum and black box stands across TLM Icon Sites, by using a combination of field data (175 reference sites) and Landsat satellite imagery (Cunningham, *et al.,* 2009). This work suggested that approximately 79% of these vegetation communities were in a stressed state. These models were successfully developed using an artificial neural network modelling framework, using structural data from the remote sensed imagery and field data (R^2 = 0.68). When these models were applied retrospectively to data for the Icon Sites from 2003 to 2008 using historic Landsat imagery, there was a discernible trajectory of increasing stress on these ecosystems. Importantly, it was recognised that this general approach was capable of reporting on condition states over both time and space. Furthermore, it was possible to detect and document decreased levels of vegetation stress for regions associated with environmental watering events between 2003 and 2009, as well as a continued decline for regions across the Murray River floodplain where water was more restricted (Cunningham *et al.* 2009a).

A follow-up study in 2010 using an updated field data and similar modelling approaches displayed poorer model performance (R^2 = 0.58; Cunningham *et al.,* 2011), which was attributed to imbalances in stratification of the field based data, where the extremes of the condition states (both good and poor condition) were not widely surveyed, and that majority of the data (77%) related to sites in poor to moderate condition. The effect of the distribution of training data was to 'flatten' the model, decreasing model performance at the 'tails', and this was addressed statistically by the linear transformation of the predictions (Cunningham *et al*. 2014). This scaled and enforced a direct relationship between the stand condition and full range of condition states observed in the field and improved the statistical performance of the models.

A subsequent modelling investigation during 2013 altered the approach by using RapidEye imagery, following the demise of the Landsat 5 satellite. This was accompanied by delays in field data acquisition in response to extensive floods, and time allowed for ecological responses to this natural event. This modelling study coincided with the development of the original Basin-wide Stand Condition Modelling Tool (Cunningham *et al*. 2013a), and therefore necessitated the need to re-model stand condition for the three preceding years to ensure consistent model performance within the tool. These stand condition models provided relatively strong model fit for TLM Icon Sites (R^2 = 0.75 and 0.61; 2009 and 2010 respectively). Building a multi-year model from surveys recorded during two drought years, and the year following extensive floods provided substantial improvements for the predictions of condition (R^2 = 0.87), when compared with models based on individual years (R^2 = 0.60-0.75). The Stand Condition Tool built from the multi-year model provided strong predictions (R^2 = 0.84) for a survey of 50 sites not used for modelling stand condition. Together these results suggested that the Stand Condition Tool would be able to predict stand condition under a range of environmental setting and conditions. The combination of these studies demonstrated that the stand condition modelling approach provided a robust framework for assessing, understanding and reporting on stand condition over time, and across extensive spatial extents.

This current report details the approaches taken to update the models of stand condition, and the software tools that enable the MDBA to develop mapped outputs of stand condition across the Murray-Darling Basin. This was achieved by incorporating additional field observations recently acquired across the Basin in 2014, 2015 and late 2016 / early 2017, in conjunction with updates to the library of remote-sensed data available to develop models. In contrast to previous documents, this report does not report on current stand condition, but on the production of a software tool that enables the MDBA to produce up-to-date appraisals of stand condition on an as-needs basis, and therefore provides the ability to monitor stand condition over time. This monitoring tool will provide useful expressions of stand condition, until the models can be revised with new field data in the future. This software is provided with an installation manual, and an additional tool that allows users to view input imagery and modelled outputs.

Methods

Study area

The Murray-Darling Basin covers a substantial portion of the Australian continent (1,0589,000 sq km), and the current study covers the full extent of the Basin, coincident with Landsat data.

This large study area and spatial data sets contains many other vegetation and land cover classes beyond the woody vegetation forms of the riverine and floodplain systems. Consequently, the model will make predictions outside the extent of these target systems, and careful consideration will need to be given by the users of the tool on how the models are constrained or limited to predicting. This issue will be addressed later in this document, and in the Stand Condition Monitoring Tool User's Guide.

Floodplain vegetation types

The floodplains of the Murray-Darling Basin contain a range of plant communities. Many of these vegetative forms can be difficult to accurately differentiate using data on the reflected or emitted radiation as detected by satellites. Broad vegetation types were previously modelled from vegetation plot data and remotely sensed imagery across the entire basin as part of a pilot study (Table 1; Cunningham et al 2013c). These base maps of forest types aided the modeling of stand condition, and are also important in identifying the relevant floodplain forests and woodlands where the model will most reliably apply.

Floodplain vegetation type	Dominant taxa
River red gum forest and woodlands	*Eucalyptus camaldulensis*
Black box woodlands	*Eucalyptus largiflorens*
Coolibah woodlands	*Eucalyptus coolabah*
River Cooba woodlands	*Acacia stenophylla*
River Oak Forest	*Casaurina cunninghamia*
Lignum shrublands	*Muehlenbeckia florulenta & Muehlenbeckia horrida*
Grasslands	Poaceae
Wetlands	Species tolerant of > 6 months inundation

Table 1 Dominant taxa that defined the target floodplain vegetation types. (Note: The current Condition Monitoring Tool focuses on vegetation types dominated by tree species, and does not consider those types displayed in grey)

Field observations (dependent data)

In 2016, the MDBA commissioned several ecological consultancies to collect contemporary field data from 605 sites across the Basin (shown as pink crosses in Figure 1). This data was acquired in the latter part of 2016 and the first few months of 2017, with two sites (CMN02 and CMN03) visited in both 2016 and 2017. Of these site assessments 175 sites were repeat evaluations, while the remainder (430 sites) were the first surveys at new locations, and therefore provide data from new environmental settings unknown to previous model versions.

The data from the 430 new sites has been incorporated into a database which now contains a total of 1754 stand condition survey observations for modelling. In summary, these observations relate to 912 unique sites, where observations have been made between 2009 and 2017. A summary of the additions of observations to the database are provided in Table 2, and the spatial distributions of the recent observations are displayed in Figure 1.

Year of Observations	Site Count
2009	175
2010	175
2013	172
2014	475
2015	150
2016	175
2017	432

Table 2 Sequence of data accessions to the stand condition database across years

The field assessments are devised to record data, and to report on estimates of the live basal area of the trees at a site (LBA), the crown extent (CE), and an index of the total plant area (PAI). The field methods have been previously documented in Cunningham *et al.* (2009, 2011), and were updated in Cunningham *et al.* (2013).

LBA assessments are scaled from 0-10, while CE is scored on a 0-5 scale, and the PAI is scored on a 0-2 scale. The scores assigned to each site measurement are combined into an overall site Condition Score which has a potential range from 0-10, where the Condition Score is the re-scaled average of the three indices:

$$\textbf{Condition Score} = (CE * 2 + PAI * 5 + LBA * 1) / 3$$

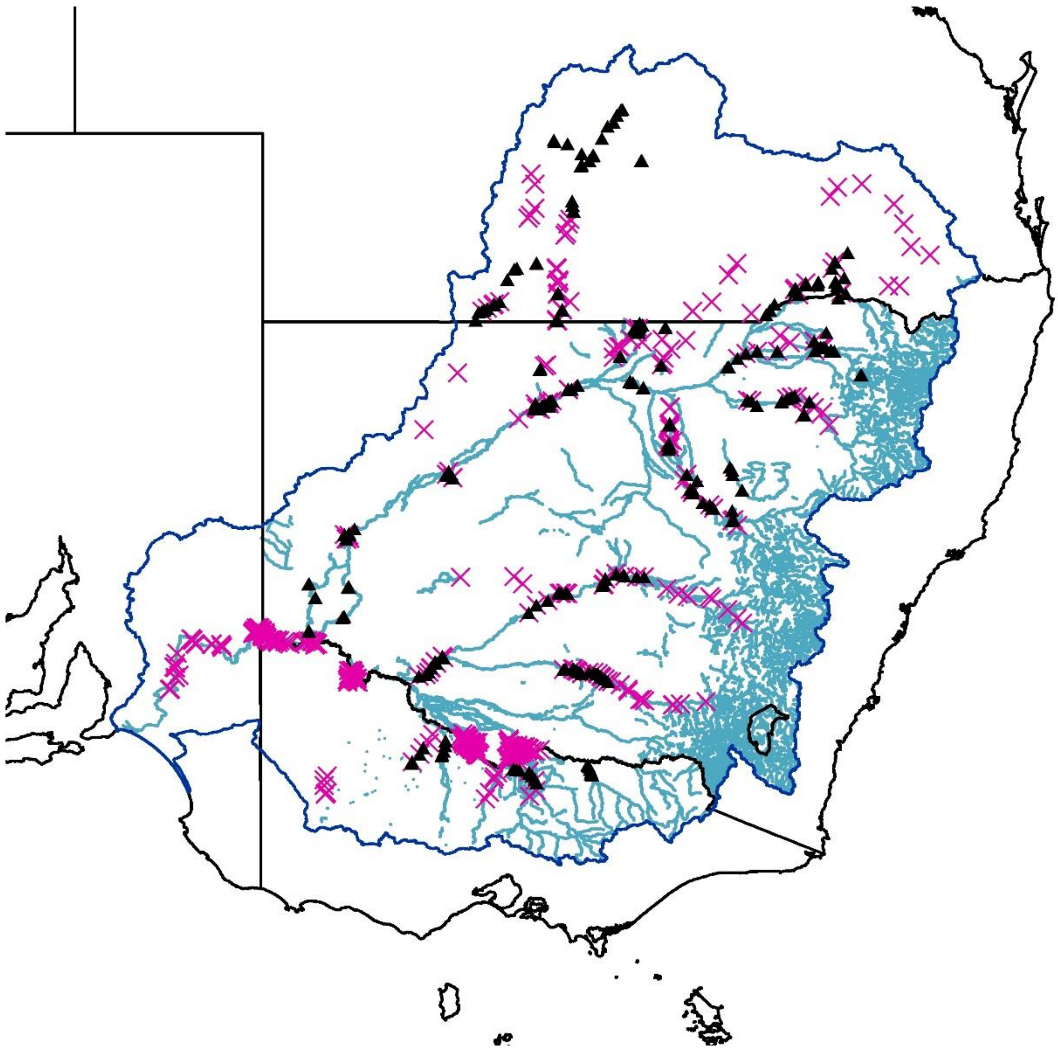


Figure 1 Location of the field observations.

Black triangles indicate site visits in 2014, while the pink crosses display sampling locations during the 20I6 / 2017 campaign.

Remote-sensed data (independent data)

The approach used in this current study for independent data, followed the structure described by Cunningham *et al* (2013b), with some notable differences. The source and therefore the specifications of the remote-sensed independent data have varied over the series of stand condition models within the Murray-Darling Basin (Cunningham, *et al.* 2009b,c, 2011, 2013b), and these processes undoubtedly influence the current inputs and processes. Models were initially developed using Landsat imagery across several years (2003, 2008, 2009, and 2010), as this platform provided a low cost yet effective option for data with reasonable grain size (30m pixel). However, these models only covered the southern section of the basin largely defined by the mid and lowers reaches of the Murray River (Cunningham *et al.* 2013a).

During 2012 errors began to appear in Landsat imagery, and it appeared that the reliability of the satellite was waning. In 2013 the opportunity arose to use RapidEye imagery to develop stand condition models across the entire Murray-Darling Basin with a finer spatial resolution (5m pixels), and with similar spectral characteristics. There were many pros and cons to using data of this finer resolution (Cunningham 2013b). Apart from the increased costs of data acquisition, a further negative aspect was the increase in data volume by a factor of nine, and the commensurate increase in the number of individual imagery tiles that were required to assemble and mosaic together to cover the Basin. While these data were masked to regions of extant floodplain vegetation, this added substantially to the data preparation required prior to modelling, and to the model application processes used to generate the mapped outputs, as file sizes were also nine-times larger than those derived from Landsat data. The modelling tool that resulted from that study was found to be useful and accurate (Cunningham 2013b), however the data preparation requirement and processing load were costly.

In 2014, Landsat imagery regained reliability with the introduction of Landsat 8. This current study uses the method developed in 2012 (Cunningham *et al.* 2013a). Landsat imagery data for the current project were commissioned and supplied by Geoscience Australia.

Temporal layers

The independent data used to train and project the models are statistical summaries of the full complement of satellite images across each calendar year. This has been done to counter the effect of variable cloud cover over any single image. Percentiles are calculated for each Landsat reflectance band from the approximately 23 separate Landsat image capture events across any year after each capture has been subjected to a cloud and cloud shadow removal process. Data are summarised at the pixel scale as median values (50th percentile) and the upper (75th percentile) and lower (25th percentile) quartiles. As such, the number of samples at each pixel used to calculate the percentiles across a year will vary with the amount of obscuring cloud cover on the Landsat capture days in that year (i.e. Landsat passes over the MD Basin once every 16 days). In combination, the composite images provide a stable indication of the central tendency of the reflectance data at any location across the Basin within a year.

The three sets of imagery were supplied as a series of one-degree tiles by Geoscience Australia. These data were mosaiced together to form imagery for two larger tiles; the southern, and the northern basin areas (as displayed in Figure 2). While these two composites can be mosaiced to cover the entire basin, the file size for the complete mosaic can be problematic for data processing in many software packages.

In summary, the entire basin is generally covered by 127 individual one-degree tiles. The combination of all data from all satellite-years provides a tally of **2286 tiles that are available for modelling** (i.e. 127 Landsat tiles x 6 years (2009, 2010, 2013, 2014, 2015, 2016) x 3 quartiles).

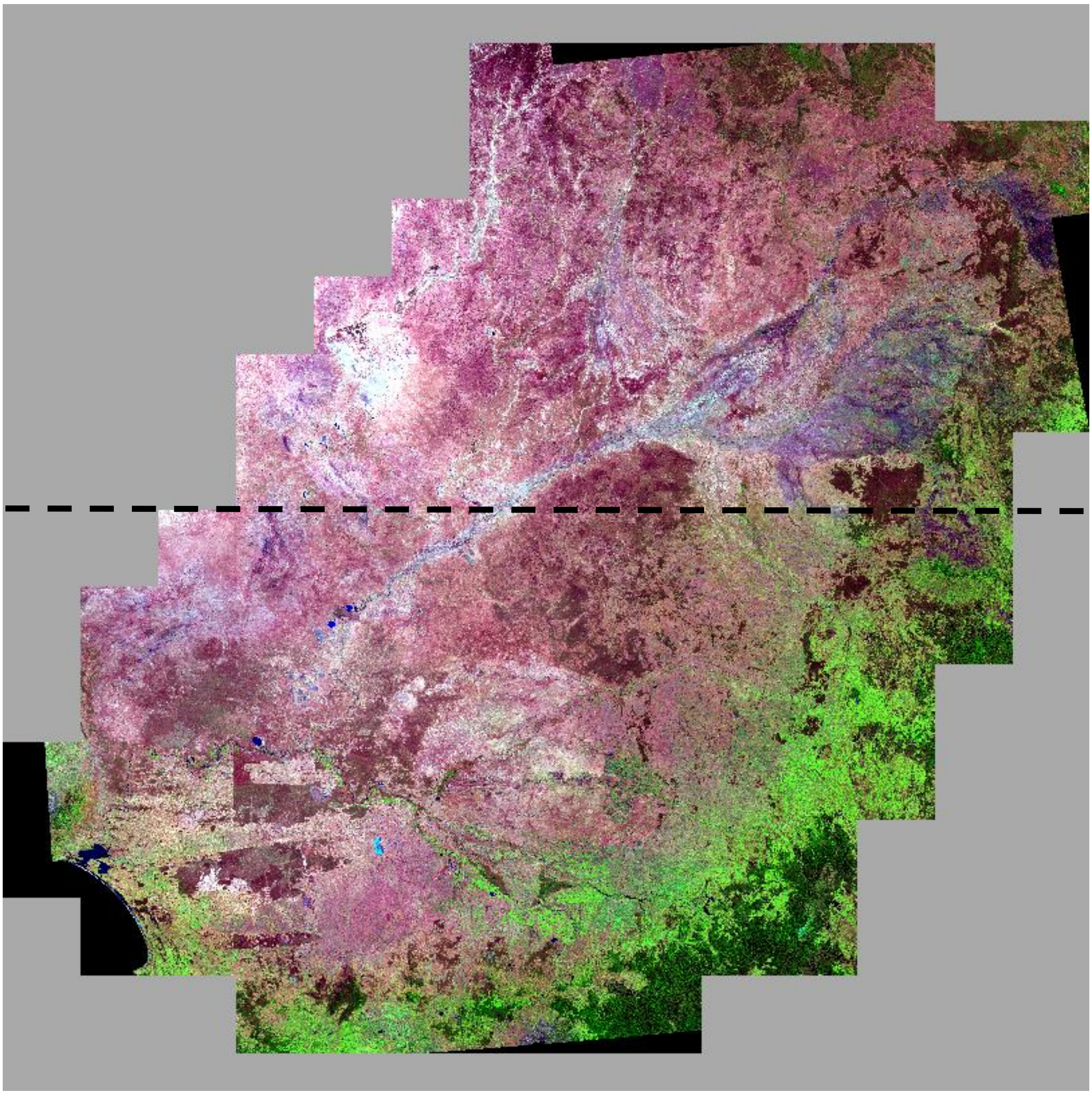


Figure 2 Full extent of Landsat imagery across the Murray-Darling Basin displayed as complete mosaic. Dashed line indicates the boundary between the northern and southern mosaics.

Constant data layers

In addition to the satellite-year summary imagery, several 'base' or constant independent data are also supplied to each model. In contrast to the annual data, these data provide a broad but static view of vegetation type and extent that is relevant to forest stand condition.

The three constant layers were:

MDBAVegNative:	This is a 4-layer interleaved spatial data file which provides model likelihoods for the presence of each of four vegetation (i.e. landcover) classes (unclassified, native, non-native and water). These data were developed for the RapidEye-based vegetation model produced and developed for MDBA in 2013.
MDBAVegSpp:	This is a nine-layer interleaved spatial file which provides model likelihoods for the presence of each of the following vegetation classes: woody vegetation, wetlands, drylands, River Redgum, Blackbox, Lignum, River Cooba, Coolibah and River Oak. Similarly, these data were developed for the RapidEye-based vegetation model produced and developed for MDBA in 2013.
AutumnLandsat:	This is a median Landsat image comprising six bands of derived Landsat imagery covering the entire Murray-Darling Basin. The median values are calculated from imagery across the autumnal quarter from each of the years between 2008 to 2012. These data provided a stable spectral foundation to the models of LBA, CE, and PAI.

General background to modelling

Condition models are effectively mathematical relationships between the Condition Score and the component indices, and a series of remote-acquired data from satellite imagery and several base data layers relating to native vegetation. These models could be developed using an enormous array of available statistical / numeric approaches. Over the last decade 'machine-learning' (ML) algorithms have progressed to the point where these methods have been developed to be very rapid and accurate for developing models, even for ecological applications (e.g. Kocev *et al.* 2007). The general approach to developing these types of models is to provide a 'train' dataset to develop the model, and a further exclusive dataset to 'test' the data. ML methods learn the patterns from one sub-set of data, and test the model performance with the 'hold-out' dataset. Importantly, this is not a single process, but is re-applied iteratively, and where the 'test' and 'train' data are continually randomised. Using this approach, a full dataset can be used to develop models iteratively, and to gain a collective appreciation of model performance.

In other words, while additional, newly acquired data could be sourced and used to test the models, model performance can be internally evaluated by partitioning the data across many iterations to understand the predictive power of the model for the dataset. Therefore, it needs to be recognised that when these models are applied to new or novel datasets (e.g. new geographic regions that may have little or no training data), the performance statistics can only be indicative at best, as the model cannot 'know' about these environments, and statistical confidence in these models cannot be known or imputed.

It is for this reason that it is important that the sampling design for field assessments attempt to adequately survey the complete study area prior to undertaking modelling. For a region as extensive as the Murray Darling Basin with a diversity of different condition states across a variety of environmental and geographic settings at any single point in time, it is recognised that a complete and adequate stratification is probably not practicable. This issue is likely to be compounded by additional issues of access / tenure and distance between stratified sampling locations. For this reason, a prudent strategy may be to aggregate and curate field data from across longer time periods and spatial extents, along with the appropriate remote sensed data, with a view to building a long-term data library of field and imagery data across a range of condition states. This data library can then provide a robust approach to developing models, and these are likely to become increasingly useful for prediction as more climatic and hydrological states are sampled.

Modelling process

The modelling process requires data from field observations (i.e. dependent data), and from the appropriate satellite-year image as well as the constant layers. These data are supplied to the machine-learning modelling software as a data array collated into training and test data.

The initial version of the stand condition tool developed for MDBA which covered the Murray River Icon sites was based upon Landsat imagery and used 'neural-network' type ML models developed using *Statistica* software. Neural networks can be complex and time-consuming to develop as they require substantial computational time to develop the network model, and then to use this model to produce a mapped output. The modelling approach used in this current study was first used with the 2014 Rapideye Stand Condition Tool, and uses another ML approach called Classification And Regression Trees (CART). Recent experience has identified that CART models provide similar or superior performance to neural networks, are more rapid to develop, and have the added advantage of being easier to interpret and communicate to stakeholders and end-users. These models were developed in the open source Clus software package.

Random Forests (RF) were used as the specific form of CART models. This method learns the selection of relevant environmental variables, and the interactions between these variables through multiple model iterations. Additionally, RF overcome the inherent inaccuracies of seeking a single parsimonious model by constructing an ensemble model from the multiple model iterations. RF models are well suited to large sets with numerous independent variables, many of which may be highly correlated. This modelling technique creates a forest of regression trees. The algorithm randomly selects a small number of independent variables at each branch of the tree from all available variables, and creates the node on the basis of the variable(s) that minimise the model error. This contrasts with the neural network models used in our previous study (Cunningham *et al.*, 2009a) which must consider all independent variables supplied simultaneously. While over-fitting is often seen as a problem in statistical modelling, predictions using regression trees for independent data sets are not compromised by using many predictor variables, and are generally superior to many other methods (e.g. Generalised Linear Models, Generalised Additive Models and Multivariate Adaptive Regression Splines; Elith *et al.* 2006).

Model data and methods

Field sampling locations were matched temporally and spatially with the satellite imagery for the appropriate year and the constant layers. Data was extracted from the coincident loci and assembled into a large data array. This array was supplied to the *Clus* to develop the CART models. Training data were extracted at the coordinate supplied for each site or each model, and for the eight surrounding pixels using a -25 m, 0m and +25m offset east and north. This allowed for minor year to year variations in pixel registration / alignment. This provided nine samples from the images to train the models. 80% of the available site data as allocated to 'Train data', while the remaining 20% of the site data was withheld and used as Test/Validation dataset (Table 3). 'Test' and 'Train' data were stratified by site name rather than randomly from the whole dataset. Hence, once a site has been selected as a member of the 'Train' or 'Test' set it held this assignation across multiple or repeat visit sites over time. Only the centre pixel was used to assess and validate the models (i.e. none of the surrounding eight pixels were used for testing model performance).

	Year Summary	
Satellite Year	**Test**	**Train**
2009	35	140
2010	35	140
2013	33	139
2014	93	383
2015	30	120
2016 *	34	141
2016 #	85	347

Table 3 Allocation of the number of train and test data for models for each year.

* Repeat site visits, # New sites

Three different models were developed using different variations of the possible range of model inputs to examine the prospects for maximising model performance.

Model 1: **'Overall model'**. This model used observations from each year matched with all the relevant contemporary summaries of Landsat data. These data were provided within a single large training dataset. This model used 3 constant grids and the 3 satellite summary grids (upper, median and lower) for each of the 6 years. This model was generalised to provide the most stable results for any given year, given the Landsat composites (i.e. upper, median and lower quartiles).

Model 2: 'Simple model'. This model used the 3 constant grids and only the median values from the six, yearly summary Landsat data layers. This model did not use the upper and lower quartile images supplied by Geosciences Australia. This model was encoded into the Stand Condition Monitoring Tool to allow the production of any existing or future single stand condition map from the corresponding Landsat images, without the need for extensive processing to produce the quartile images (i.e. when upper and lower quantile data are unavailable). However, the preference would be to use the 'Overall Model', where possible.

Model 3: **'2016 Model'.** This model was created using only the 2016 and 2017 observations and the 2016 median, upper and lower quartile images, along with the three constant layers. The purpose of this model was to examine if this bespoke 2016 model would outperform the generalised 'Overall Model' (Model 1).

Comparison metrics

Model performance is most readily compared for each model by examining the correlation coefficients (r) between the observed and predicted values for each of the condition components that compose the stand condition score. These correlations were calculated using either *a)* the 'test' dataset which was withheld from model development process to assess model performance with data that had not been used to develop the model, as well as for *b)* the 'train' data that was initially used to develop the model. As would be expected, the 'fit' of the model using the 'test' data is generally lower than when the model is evaluated with 'train' data used in model development.

Results

Generalised model performance

Although three different forms of models were developed, only the 'Overall Model' and the 'Simple Model' can be directly compared, as both approaches use training data taken from across multiple years. In contrast, the 2016 Model is based upon a more limited dataset, which negates a cross-comparison with the other models.

Both the 'Overall' and 'Simple' models were shown to have similar levels of performance across the complete time series of the data, as displayed by correlation coefficients between observed and predicted condition values (Table 4). This is an interesting result considering that the 'Overall Model' had much more complex data array with which to form the model (i.e. upper and lower quantile data), than for the 'Simple Model'.

Model	Condition Score Correlation	Crown Extent Correlation	Live BA Correlation	PAI Correlation
'Overall Model'	0.739	0.646	0.678	0.792
'Simple Model'	0.717	0.627	0.640	0.793

Table 4 Overall model performance for the 'Overall' and 'Simple' model, expressed as correlation coefficients for the Condition Score, and for component indices.

When these two multi-year models were compared directly to each other, it was clear that there was a very high degree of congruence between these approaches (Table 5), indicating that that these two models were closely aligned. This is a useful result, suggesting that the 'Simple Model' may be used with some confidence to produce similar mapped outputs of condition values, even when upper, median and lower quartile imagery data are unavailable.

Cross Model Comparison	Condition Score Correlation	Crown Extent Correlation	Live BA Correlation	PAI Correlation
'Overall Model' to 'Simple Model'	0.969	0.951	0.929	0.987

Table 5 Correlation between the two multi-year modelling methods for the Condition Score, and for component indices.

Year to year consistency

In addition to examining the multi-year performance for the 'Overall Model', correlation coefficients can also be used to examine the performance on a yearly basis through examining the model fit from both observed and predicted test and train data. As indicated previously, it would be expected that the performance statistics would be higher for train datasets, than for test data that were withheld from the model development process (Table 6).

These results indicated that the 'Overall Model' performed reliably well for the five modelled years. Model performance for 2016/17 was lower than for previous years, whether examined for fit with 'test' or 'train' data. Performance was particularly low for current period when evaluated with the test data that was withheld from the model development.

Year	Condition Score Correlation	Crown Extent Correlation	Live BA Correlation	PAI Correlation
Test	**0.794**	**0.701**	**0.648**	**0.863**
2009	0.864	0.777	0.706	0.942
2010	0.849	0.803	0.602	0.913
2013	0.790	0.669	0.553	0.871
2014	0.814	0.707	0.671	0.858
2015	0.813	0.708	0.611	0.929
2016	0.636	0.541	0.744	0.662
Train	**0.918**	**0.874**	**0.898**	**0.890**
2009	0.938	0.894	0.923	0.910
2010	0.958	0.923	0.905	0.936
2013	0.943	0.926	0.923	0.867
2014	0.913	0.835	0.888	0.889
2015	0.957	0.924	0.943	0.933
2016	0.823	0.772	0.825	0.827

Table 6 Correlation coefficients of predictive performance for the within-year models developed with the Overall Model approach.

Results are shown for the Condition Score, and for component indices.

Which model best represents the current stand condition?

When considering the predictions of the most recent year 2016/17, all three modelling methods may be compared to each other. The table of r values for these comparisons (Table 7) indicates that the 'Overall Model' performed the best within 2016/17, even outperforming the bespoke '2016/17 Model' (Table 7). However, the differences in model performance between the three modelling methods are reasonably small. Collectively these statistics indicate that the model performance for

2016/17 is relatively poor. The 'Overall Model' (Model 1) developed using the median, upper and lower quartile imagery has the highest performance for each year, including 2016/7.

Model	Condition Correlation 2016	Crown Extent Correlation 2016	Live BA Correlation 2016	PAI Correlation 2016
'Overall Model' (Model 1)	0.636	0.541	0.744	0.662
'Simple Model' (Model 2)	0.608	0.505	0.703	0.672
'2016/7 Model' (Model 3)	0.596	0.483	0.655	0.673

Table 7 Correlation coefficients of predictive performance for 2016 model using all three modelling approaches. Results are shown for the Condition Score, and for component indices.

Observations statistics

Currently there is a total of 1754 stand condition surveys compiled in the stand condition database. These observations relate to 912 individual sites/ locations. Each observation for each site provides estimates for native woody vegetation of Live Basal Area (LBA), Crown Extent (CE) and Plant Area Index (PAI). These measures were made in accordance with the methods detailed in (Cunningham *et al.*, 2007, 2009a).

Due to localised flooding during the recent 2016/2017 field sampling season, a small number of sites targeted for resampling were not accessible. It appears that for some of these sites the sampling may have occurred at a new geographic location (if the GPS location is accurate), while retaining the original site name. Details of these sites are provided in Table 8, and it will be important to be mindful for future re-visits to sample sites that the actual location will need to be recorded, and not simply the site name. Models can then be based upon the actual position rather than the implied position from the site code or number.

Site Name	Site Id	Years visited
KP170	86	6
MB127	150	4
BN123	163	6
KP133	88	6
GB34A	106	6
KP165	96	6
GB41A	107	6
GB10A	103	6
BFN85	161	6
GB101A	108	6

Table 8 Sites surveyed in 2016/2017 with apparent shifts in geographic location.

Discussion

Government agencies that are responsible for managing the environment, need to understand the current condition and the dynamics of natural systems over time. This understanding is not only essential for improving management of these systems, but also to reliably and consistently report on the status of important natural assets like the Murray Darling Basin.

This current project had several purposes. The first aim was to incorporate new data from field-based stand condition assessments and contemporary remote-sensed data to create revised models of riverine and floodplain forest condition across the Murray Darling basin. These products provide the MDBA with a framework to monitor stand condition annually into the future and back-cast condition into any previous year post 1987.

The second aim was to develop a Stand Condition Monitoring Tool (SCMT), which incorporated these machine learning models into software that could enable the MDBA to produce revised spatial models of stand condition as new Landsat imagery becomes available. Additional features were added to the SCMT software that enable users to select a complete directory of Landsat imagery of individual tiles, rather than a compiled mosaic. This feature offers the Agency considerable flexibility, and even enables the model to be applied to individual one degree Landsat tiles. This could enable the Agency to easily / rapidly monitor individual sites or locations over time, where required.

The results relating to the first objective of updating the stand condition model(s) indicated that the ‘Overall Model’ (Model 1) developed using the median, upper and lower quartile imagery had the strongest performance when applied for each year. However, it is important to note that the performance statistics for the ‘Simple Model’ were not greatly dissimilar from the ‘Overall Model’, despite the ‘Overall Model’ having access to a quantile data, which would result in more complex CART models.

The interpretation of this result is not entirely clear. A simplistic appraisal could form the view that model parsimony (i.e. not having the variance data), does not detract from model performance. Alternatively, this result could be interpreted that we are currently at the limits of the usefulness of Landsat reflectance in modelling stand condition. This could be tested in future by incorporating other inputs such as preceding rainfall, high resolution satellite-borne radar data, groundwater data as it becomes available. It is anticipated that this data would augment significant predictive power to the current models, however this is currently unverified.

As a result, it is recommended that the ‘Overall Model’ process should be used to generate spatial models or ‘maps’ of stand condition for all years where the quantile data has been compiled for the Landsat imagery. However, when only a single Landsat image or median image is available (i.e. without accompanying upper and lower quartile data), the ‘Simple Model’ (Model 2) can be used to generate mapped outputs of stand condition, and these maps are likely to be generally comparable to those made using the ‘Overall Model’.

One of the curious results from this study was the decline in model performance for the 2016 within-year model. There are numerous possible reasons for this result, with its deviation from the pattern of high model performance observed for earlier years, however it will probably not be possible to dis-entangle which factor may have been the primary contributor. One possible explanation is that in contrast to all the other years, 2016 saw a return to the high winter and spring rains more typical of rainfall patterns across the eastern and southern regions of the basin prior to the onset of the millennial drought around 1996. When considered in the context of the last decade rainfall, the recent winter/spring period across SE Australia maybe considered an outlier year.

Close examination of the Landsat data did not suggest that there are data quality issues with the final supply of the Landsat data that was shipped from Geoscience Australia. The second possibility of the base layers may have presented some issues for model performance. The models prepared used the base models relating to landcover and dominant species, and these were represented as probabilistic models. These base layers are the same models as supplied in the 2014 RapidEye-based modelling tool, and were derived from extensive Victorian datasets, but with considerably less data from SA, NSW and QLD. As indicated earlier, models are likely to encounter performance issues where they make predictions into domains / regions where they have little or no information. This of course is untested presently without field observations, but if these landcover and dominant models fit poorly, then the performance of the stand condition models will also be impaired. This issue is worthy of further investigation, including accessing existing biodiversity datasets to reinforce and improve the base layers.

The third set of potential issues relate to field data, including aspects of the field data sampling methods, and with subsequent data management processes. As regular end users of ecological field data acquired from various field observers /sources / taxa / ecosystems for modelling purposes, the authors are keenly aware from personal experience how simple and inadvertent errors in data handling, particularly in the use of MS Excel tables, can lead to serious subsequent declines in model performance. Additionally, this can be very difficult to retrospectively track down.

Possible future improvements to the predictive power of stand condition models

The prospects for future improvements to stand condition models could arise from advances across four areas. These include *i)* the modelling approach and the algorithms that are used, *ii)* the remote-sensed data, *iii)* the base layers that regulate the model development and model expression (i.e. ‘masks’ to constrain the model from regions with little or no field data), and the *iv)* the field data.

Mathematical and ML models have rapidly become a cornerstone of many modern-day technologies, particularly in the way that that data is acquired, assimilated with other relevant information, and used by computers and smart devices. The technologies and algorithms of ‘learning’ patterns from simple or complex datasets (as opposed to conventional statistical methods) has surprisingly not been widely used in ecological contexts, but this is changing rapidly as the power and utility of these methods becomes more appreciated. While the performance of these methods will undoubtedly increase with the emergence of new algorithms, the greatest potential for future improvements arguably will come from the routine application of ‘deep learning’ algorithms.

Deep learning methods are currently employed for example in facial and voice recognition where they evaluate a large body of data to make a rapid prediction or decision. In terms of how this would relate to modelling stand condition, this could mean that contextual information associated with remotely sensed data sources relevant to the field sample could be used for modelling, rather than the current data provided as a single 30m pixel. Contextual information on the surrounding environment (e.g. adjacency to other forest, woodland, agriculture, surface water, preceding climatic conditions, etc.), will have a major influence on the local conditions, and how data is captured and reported on stand condition. Early trials by the authors have been successful in the development and early testing of deep learning methods (currently unpublished data), however they are not yet fully operational. As part of this study we partially considered generalised forms of contextual data by including summarised median values of a 3x3 array of imagery data around the sample location. This provides a means of dealing with inconsistencies in pixel registration in the imagery over time, and errors in field GPS location data. It also adds more training data as slight variations in survey position provides varying views of the remote sensed data that are equally valid to the central position, although they will be spatially auto-correlated by definition.

Satellite imagery has become much more widely accessible over the last decade in terms of the breadth of data, the systems to deliver and manipulate it as end users, as well as lower costs. It is likely that these trends will continue, and an even broader suite of data will become available, and at finer resolutions. For example, Copernicus Sentinel-2 data from the European Union provides a useful example of an instrument that is commencing to supply data at nine times the resolution of the Landsat platform. Additionally, this platform provides additional sensors (e.g. radar data capable of providing information on vegetation structure), and with a more frequent re-sampling rate. This data has the potential to substantial enhance the ability of the MDBA and other agencies to monitor large geographic domains at finer resolutions in the future, without abandoning legacy Landsat 5 data series. As a result, the general mission for considering potential improvements for remote sensed data for the MBDA's needs would ideally consider developing a simple framework. This framework would map out how historic, current and emerging data are compiled together in robust ways, with a view to support monitoring of various NRM activities, and how the impact of the increased data size / complexity / computational requirements can be managed effectively in the future.

The other area where substantial improvements could be made is with the base layers, and with the masks that could be applied to model outputs. As indicated earlier, these layers were originally developed for an earlier vegetation mapping project (Cunningham *et al.* 2013c), and these were based heavily on Victorian and other data. Ideally a broader suite of curated botanical data would inform revisions of these products in the future.

Substantial improvements to data processing and modelling systems have taken place since these products were developed. For example, ARI recently developed a land cover model for the South Australian Government which covered the whole of the State. The product was a model which included a hierarchy of 'native' or non-native land cover, water sources, structural vegetation forms, etc. In total, there were more than 50 land cover 'ideas' produced as spatial models. Data such as these could provide significant predictive power to future models, and could provide very useful mechanisms for constraining the stand condition model to report only to relevant features across the Basin. This latter feature would minimise the risks that currently exist for the MDBA of the SCMT reporting to regions landscapes where the models should not apply (i.e. there is no native, woody vegetation present). Improving these 'masks' or limits to the current model are perhaps the most direct and immediate approach that the MDBA could take to minimising the risks of model commission.

Of all areas listed for potential improvement arguably the best opportunities for improving stand condition models will come from improving the quality of the data collected in the field. Mathematical models of any form or kind rely heavily upon the quality of the dependent data to accurately represent the feature(s) that a model is attempting to represent. If the data provide a poor representation of the environmental and geographic domains of the region of interest, or do not fully describe the range and expressions of the features that are to be modelled, then the outputs are less likely to be representative of the features, and be less useful. Without this breadth in the data the model will struggle with its ability to accurately predict the training data, and to interpolate the environmental and geographic domains between field sites. Additionally, this will limit the abilities for extrapolating into novel domains that have not previously been sampled with field data.

The recent field campaign has made significant improvements in increasing the breadth and geographic extent of field. With additional data in the future, including the sampling of some geographic / environmental domains that remain un-sampled, and repeat visits to at least a subs-set of those recently visited, further improvements in model performance and reliability would be expected.

Machine Learning methods, as with any modelling approach, rely upon data quality. Within the context of the Stand Condition Monitoring Tool, it is important to consider that while sampling may be associated with a single sampling campaign, the data will hopefully be contributing to larger, coherent library of exemplars across the Basin. Over time, these exemplars will provide a solid representation of the extremes of woody vegetation condition (both poor and good), as well as the geographic and environmental settings that riverine and floodplain ecosystem occupy. In other words, a strategic longer-term strategy could be to develop a library of training data that represents the Basin across its full extent and expression of stand condition. This 'meta-experiment' approach to data sampling and aggregation are at the core of why ML models perform well on making useful predictions for complex systems. This approach, even philosophy, is a key strength to the basic premise of the Stand Condition Monitoring Tool, and its process of continual improvement.

Other issues for consideration that may lead to improvements in model performance and the utility of the stand condition monitoring tool include the methods used for field assessments, and the use of ecological indices. The field method that is currently used for assessing stand condition was developed to assess the condition of redgum and black box communities at TLM sites. The spectral and structural characteristics of the riparian and floodplain communities that are dominated by these species are likely to have particular spectral signatures, and these are likely to differ from vegetation communities dominated by Coolibah, or *Acacia stenophylla* or *Casaurina cunninghamia* that dominate the northern regions of the Basin. The current field methods would not be appropriate for use with other native vegetation forms such as Lignum shrublands, grasslands and wetlands. It may be productive to consider reviewing the current field assessment method so that it can fairly and consistently represent the all ecological communities that are being assessed, while still aligning with the current assessment method to retain the legacy of the data collected in 2016-2017 and earlier campaigns.

Ecological indices can represent opportunities for rapid field assessments, however there can be false economy in their use if the indices are relatively subjective, subject to observer variation, or coarsely scaled and insensitive (e.g. Gorrod and Keith 2009, McCarthy *et al.* 2003). For example, where crown extent data are recorded in categories of 0%, 1-20%, 20-40%, etc, the CE index can only represent a relative small number of possible states, and cannot provide finer scale discrimination once data is encoded in that form. The scale and extent of the Murray-Darling Basin means that there will be subtle differences in the structure and floristics of native woody vegetation relevant to the local conditions, and these need to be captured, where possible. This is particularly important where the remote-sensed platform can detect these differences, but these are not captured on the ground. These types of data mis-match can contribute significantly to 'noise' in statistical or ML models. To be clear, the use of ecological indices by themselves is not a problem and they have a positive role in data simplification and communication, as ecological data can be complex and difficult to represent.

Emerging technologies and instrumentation mean that hopefully new quantitative field based assessment methods will soon emerge that will provide more detailed information than previously possible. For example, new laser scanning technologies that can rapidly provide quantitative data on vegetation structure (e.g. Leica BLK 360), and tools such as these will begin to address issues of subjectivity and observer variation for some (but not all) field based data collection methods. Whatever future improvements are made, it will be important for MDBA stand condition assessments that the methods and data are devised with care so data can be retrofitted to align with previous indices, thereby maintaining the continuity of the program while increasing capability.

In conclusion, this study and the report provide details of updates to models of stand condition across the Murray Darling Basin, and the development and delivery of the Stand Condition Monitoring Tool. This tool will support the MDBA in being able to form regularly updated views on stand condition, as new Landsat imagery becomes available. The power of this tool is that the model is stable, and has been based upon many years of field data collection under a range of seasonal conditions. This version of the tool will be suitable for use until the next campaign of field data sampling, at which time revised versions of the software can be efficiently incorporated into the modelling tool.

Acknowledgments

This research project was funded by the Murray-Darling Basin Authority (MD3805). We thank our collaborators at various junctures from the Murray-Darling Basin Authority (Inoka Wellhange, Katrina Phillips, David Stevens, Ben Seddon, Mathew Maliel and Michael Wilson) and Geosciences Australia (Leo Lymburner, Rachel Melrose, Jeff Kingwell and Medhavy Thankappan). We thank Geoscience Australia for their efforts in generating a consistent Landsat composite with quantile data across the Basin. We also thank the various ecological contractors who were engaged by MDBA to make field assessments. These models and the spatial products would not be possible without the combination of field and remotely-acquired data.

The authors deeply regret the passing of Shaun Cunningham in 2016, and wish to acknowledge the legacy of his contribution to ecological science and natural resource management of floodplain ecosystems across the Basin.

Relevant References

Bren, L.J., O'Neill, I.C. & Gibbs, N.L. (1987) Flooding in the Barmah Forest and its relationship to flow in the Murray-Edward River system. *Australian Forest Research*, **17**, 127-144.

CSIRO & Australian Bureau of Meteorology (2007) *Climate change in Australia*. CSIRO, Melbourne.

Cunningham, S.C., Griffioen, P., White, M. & Mac Nally, R. (In press). Assessment of ecosystems: a system for rigorous and rapid mapping of floodplain forest condition for Australia's most important river. *Land Degradation & Development.*

Cunningham, S.C., Read, J., Baker, P.J. & Mac Nally, R. (2007) Quantitative assessment of stand condition and its relationship to physiological stress in stands of *Eucalyptus camaldulensis* (Myrtaceae) in southeastern Australia. *Australian Journal of Botany*, **55**, 692-699.

Cunningham, S.C., Mac Nally, R., Griffioen, P. & White, M. (2009a) *Mapping the Condition of River Red Gum and Black Box Stands in The Living Murray Icon Sites. Stand Condition Report 2009 (with modelled results for 2003 and 2008)*. Murray-Darling Basin Authority, Canberra.

Cunningham S.C., Mac Nally R., Read J., Baker P.J., White M., Thomson J.R., & Griffioen P. (2009b) A robust technique for mapping vegetation condition across a major river system. Ecosystems, 12, 207-219.

Cunningham SC, Mac Nally R, Griffioen P and White M (2009c) Mapping the Condition of River Red Gum and Black Box Stands in The Living Murray Icon Sites. A Milestone Report to the Murray-Darling Basin Authority as part of Contract MD1114. Murray-Darling Basin Authority, Canberra

Cunningham, S.C., Griffioen, P., White, M. & Mac Nally, R. (2011) *Mapping the Condition of River Red Gum (Eucalyptus camaldulensis Dehnh.) and Black Box (Eucalyptus largiflorens F.Muell.) Stands in The Living Murray Icon Sites. Stand Condition Report 2010*. Murray-Darling Basin Authority, Canberra.

Cunningham, S.C., Griffioen, P., White, M. & Mac Nally, R. (2013a) *Mapping the Condition of River Red Gum (Eucalyptus camaldulensis Dehnh.) and Black Box (Eucalyptus largiflorens F.Muell.) Stands in The Living Murray Icon Sites. Comparison of the predictive power of Landsat and Rapideye imagery, and validation of future predictions based on imagery only.* Murray-Darling Basin Authority, Canberra.

Cunningham, S.C., Griffioen, P., White, M. & Mac Nally, R. (2013b) *Mapping the Condition of River Red Gum (Eucalyptus camaldulensis Dehnh.) and Black Box (Eucalyptus largiflorens F.Muell.) Stands in The Living Murray Icon Sites. Stand Condition Report 2012*. Murray-Darling Basin Authority, Canberra.

Cunningham, S.C., White, M. Griffioen, P., Newell, G and Mac Nally, R. (2013c). Mapping Floodplain vegetation types across the Murray-Darling Basin using remote sensing. Milestone Report to the MDBA as part of Contract MD2245.

Cunningham, S., Griffioen, P., White, M. & Mac Nally, R., 2014. A Tool for Mapping Stand Condition across the Floodplain Forests of The Living Murray Icon Sites, Canberra: Murray–Darling Basin Authority.

Elith, J., Graham, C.H., Anderson, R.P., Dudík, M., Ferrier, S., Guisan, A., Hijmans, R.J., Huettmann, F., Leathwick, J.R., Lehmann, A., Li, J., Lohmann, L.G., Loiselle, B.A., Manion, G., Moritz, C., Nakamura, M., Nakazawa, Y., Overton, J.M.M., Peterson, A.T., Phillips, S.J., Richardson, K., Scachetti-Pereira, R., Schapire, R.E., Soberón, J., Williams, S., Wisz, M.S. & Zimmermann, N.E. (2006) Novel methods improve prediction of species' distributions from occurrence data. *Ecography*, **29**, 129–151

Friedman, J.H. (2001) Greedy function approximation: a gradient boosting machine. Annals of Statistics, 29, 1189–1232.

Gorrod, E. J. and Keith, D. A. (2009), Observer variation in field assessments of vegetation condition: Implications for biodiversity conservation. Ecological Management & Restoration, 10: 31–40. doi:10.1111/j.1442-8903.2009.00437.x

Kocev, D., Vens, C., Struyf, J. & Džeroski, S. (2007) Ensembles of multi-objective decision trees. Machine Learning: ECML 2007. Proceedings of the 18th European Conference on Machine Learning, Warsaw, Poland, September 17-21, 2007 (ed. by J. Kok, J. Koronacki, R. De Mántaras, S. Matwin, D. Mladenić and A. Skowron), pp. 624-631. Springer, Berlin.

Margules & Partners (1990) Riparian Vegetation of the River Murray. Report prepared by Margules and Partners Pty. Ltd., P. & J. Smith Ecological Consultants and Department of Conservation Forests and Lands. Murray-Darling Basin Commission, Canberra.

MDBC (2002) The Living Murray: a Discussion Paper on Restoring the Health of the River Murray. In, p. 94. Murray-Darling Basin Commission, Canberra.

MDBC (2005) The Living Murray Foundation Report on the significant ecological assets targeted in the First Step Decision. Murray-Darling Basin Commission, Canberra.

Roberts, J. (2004) Floodplain Forests and Woodlands in the southern Murray-Darling Basin. In. Report JR 03/2004, Canberra.

Rokach, L. & Maimon, O. (2008) Data mining with decision trees: theory and applications. World Scientific Publishing, Singapore.

Rumelhart, D.E., Hinton, G.E. & Williams, R.J. (1986) Learning representations by back-propagating errors. Nature, 323, 533–536.

Souter, N., Cunningham, S., Little, S., Wallace, T., McCarthy, B., Henderson, M. & Bennets, K. (2012). Ground-based survey methods for The Living Murray assessment of condition of river red gum and black box populations. Murray-Darling Basin Authority, Canberra.

Stein, J.L. (2006) A Continental Landscape Framework for Systematic Conservation Planning for Australian Rivers and Streams. Available at http://hdl.handle.net/1885/49406. Australian National University, Canberra.

ter Steege, H. (1996) WINPHOT 5.0: a programme to analyze vegetation indices, light and 541 light quality from hemispherical photographs. In. Tropenbos Guyana Programme, Report 542 95-2., Tropenbos, Guyana.