LiDAR and Multispectral Remote Sensing for the Murray Darling Basin Sustainable Rivers Audit

Vegetation and Physical Form.

10 November 2010

Published by the Murray-Darling Basin Authority

Postal Address: GPO Box 1801, Canberra ACT 2601

Telephone: (02) 6279 0100 international + 61 2 6279 0100
Facsimile: (02) 6248 8053 international + 61 2 6248 8053
Email: info@mdba.gov.au
Internet: http://www.mdba.gov.au

post PO Box 729 Fortitude Valley Queensland Australia 4006 phone 07 3257 1011 fax 07 3257 1037 email info@terranean.com.au www.terranean.com.au ABN 52 113 485 475

terranean mappingtechnologies

Executive Summary

This report present the results of a LiDAR and Multispectral image project to collect information upon which to base two of the five themes of the Sustainable Rivers Audit (SRA); 1) Physical Form (of the river channels), and 2) Vegetation (the distribution of riparian foliage in three dimensions).

At each of one thousand six hundred and ten (1610) river sites randomly stratified across the Murray Darling Basin, full waveform LiDAR was collected at a minimum density of 4 outgoing pulses per square metre and multi-spectral Vexcel Imagery was captured with a ground pixel spacing of 30cm. Field survey was conducted to verify the accuracy of the processed remotely sensed data.

Innovative methods were developed in the course of the project in a process of collaboration between the SRA, the Independent Sustainable Rivers Audit Group (ISRAG) and Terranean Mapping Technologies. These methods were implemented as algorithms and software tools to extract measurements of Physical Form and Vegetation provide information on the health of the river sites with respect to these themes.

The project ran from December 2009 through October 2010, with aerial survey completing in June. Flooding caused delays in the early stages of the project and on a number subsequent periods but the project was completed successfully within 4 months of the schedule.

In summary, the project involved four stages, each generating progressively more refined information. These stages are:

1. Data collection; aerial and field survey
2. Primary processing of the LiDAR and imagery
3. Development of Variable Extraction software
4. Secondary processing to extract channel features (centreline, top banks, bottom banks, transect profiles and riparian buffer zones).
5. Data measurements of the physical dimensions of channel profiles and centrelines, also vegetation measurements based on the vertical distribution of the LiDAR point cloud.

The project was logistically complex as field survey and two aerial surveys had to be coordinated such that field survey progressed in advance of aerial survey, the Vexcel and LiDAR surveys occurred within two weeks of each other, and both aerial surveys provided data at a rate sufficient to maintain continuity of data processing.

Despite delays due to flooding and the development of novel methodologies and software tools during the course of the project, the project completed to specification and the data has been verified as being of a consistently high standard of quality. More than 200,000 spatial layers of information and 2,000,000 measurements were generated as outputs from the project. Much of this information, although extremely useful, has not previously been available for environmental assessment because methods for extracting it from full waveform LiDAR had not been developed. The original theoretical basis of the project plan was vindicated by the richness and quality of the information generated.

Separate reports are provided by 1) Atlass for the LiDAR survey, 2) Aerometrex for the Vexcel imagery survey and 3) RPS for the field Survey. A User Manual for the Variable Extraction Tool is also provided as a separate document.

Table of Contents

Methodology

DATA COLLECTION

A stratified sampling design was employed, generating 70 sites for each of the 23 valleys in the Basin. Within each valley the sites were stratified against elevation zone and broad vegetation type. A total of 1610 sampling sites were selected. An additional 109 check sites were also selected to measure the spatial accuracies achieved in different types of vegetation and topography.

Each sampling site was defined as a 2000 metre by 700 metre rectangle aligned to the primary river channel. Within the rectangular site, a one kilometre stretch of river channel and adjacent riparian zone were analysed.

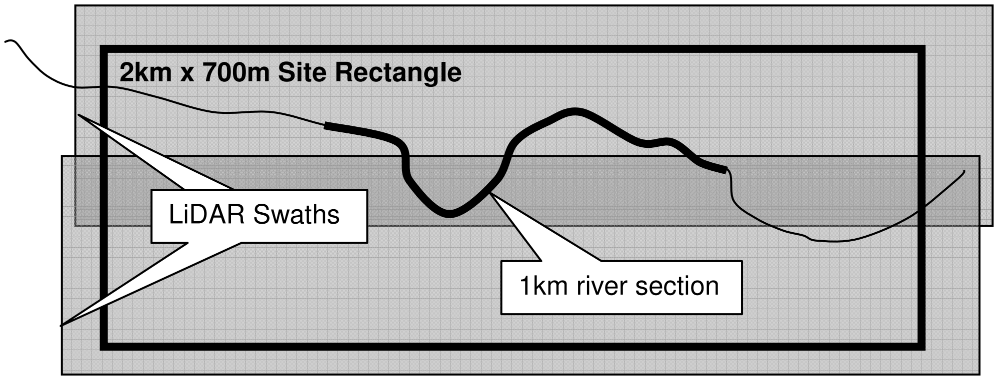

Figure 1. Each 2000 m x 700 m site is covered by two 577 m wide LiDAR swathes with 35% overlap. A 1km section of river within each site is analysed.

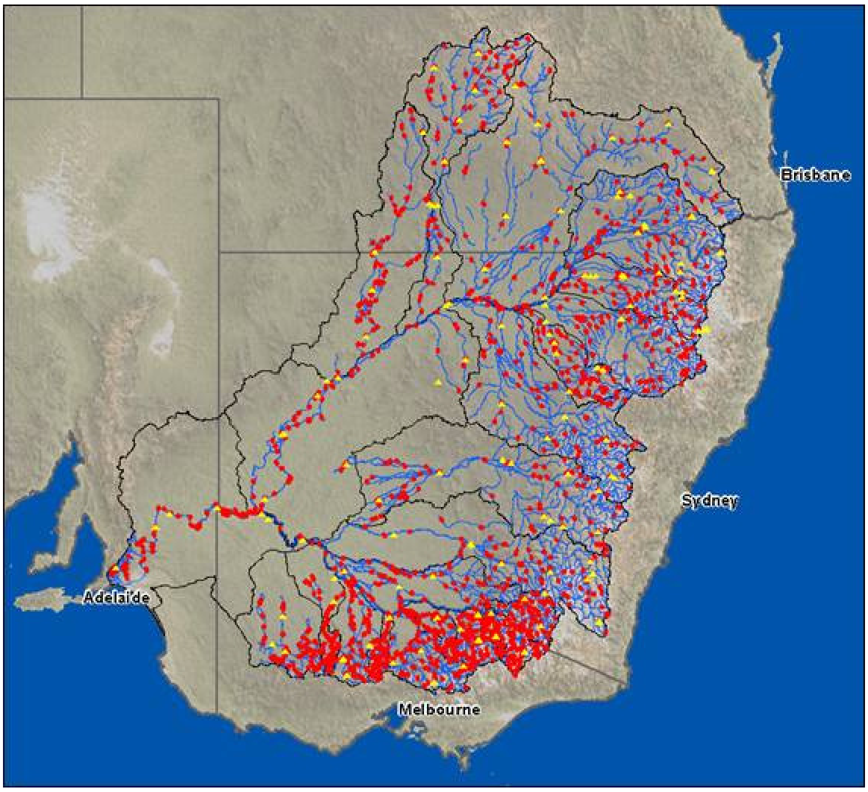

Figure 2. Channel sites (red) and field check sites (yellow triangles) overlaid on the 23 major river valleys that make up the Murray-Darling Basin.

The LiDAR was initially recorded using a Trimble Harrier 56 LiDAR instrument. A Harrier 68 LiDAR system was later deployed to accelerate the aerial survey after initial delays due to flooding. Both systems were operated with the following parameters:

- Flying Height 500 m above ground.
- Flying Speed: 205 km/hr
- Scanning Angle 60 degrees
- Overlap: 35%
- Swath Width: 577 metres
- Scan Rate: 76 Hz
- Pulse Rate: 200 kHz
- Point Spacing:
  - Along Track 0.5 m
  - Across Track 0.5 m
- Ratio: 1 : 1 (along track : across track)
- Capture Point Density: 4.05 per square metre within swath
- Average Point Density: 6.89 per square metre
- Spot Footprint: 0.25 m

The 0.15 metres accuracy that can be achieved using PPP post processing (and CORS where available) is well within the specified absolute accuracy (relative to ausgeoid98) of 0.5 metres. For each sortie LiDAR was recorded over a ‘boot control’ site containing at least four horizontal control points on vertical structures, such as building eaves, and six vertical control points on hard bare flat surfaces, such as roads. Surveyed points, measured to an accuracy of better than 10 centimetres, were used as a gross error check during the data processing phase. Independent field data will be statistically analysed subsequently to determine the means and standard deviations of the LiDAR spatial errors in different vegetation types.

The aerial survey also included the capture of Vexcel multi-spectral imagery that was ortho-rectified against LiDAR terrain surfaces. It was delivered as 4-band multi-spectral GeoTIFF images, with false colour infra-red and natural colour enhancements in JP2000 format.

The aerial surveys were constrained to ensure that data was not captured when water was overflowing the river channels and the LiDAR and Vexcel imagery were captured within two weeks of each other. These constraints and the need to survey field check points ahead of the aerial survey introduced some logistical complexities that were exacerbated by extensive flooding that occurred early in the project (December 2009 through March 2010).

PRIMARY PROCESSING

The raw instrument data are converted from the temporal / angular domain to spatial coordinates in three dimensions by reference to differential GPS and Inertial Motion Unit (IMU). This process, performed using the Riegl program RiANALYZE™, also converts the waveform signal to points based on a number of parameters, including a signal amplitude threshold which determines the signal noise ratio of the point cloud and the proportion of the returned signal contained in all the points generated from each pulse.

After performing a gross error check for each sortie against the ‘boot control’ points, the two overlapping strips for each site are ‘levelled’, combined and trimmed to the 2000 m x 700 m site polygon.

TerraScan software was used to automatically classify the LIDAR to a preliminary ground/non-ground classification that was refined through manual editing and quality checking. The final classification includes the classes:

1. Unclassified,
2. Ground,
3. Low vegetation (0.1m - 1.0m),
4. Medium vegetation (1m - 3m),
5. High vegetation (> 3m),
6. Building,
7. Low point (noise),
8. Water,
9. Bridge,
10. High point (noise),
11. Transmission line/structure

An automated batch process to produce the primary datasets was developed in the object oriented TNTmips Spatial Modelling Language (SML). The batch tool automatically sorts batches of LAS files associating each file spatially with a site and assigns a name based on the Valley, Site, UTM Zone and Capture Date. The tool generates all the raster surfaces, foliage density layers, contours and other primary datasets, outputs these to the required formats of formats and re-imports to TNTmips for verification. The single batch process produced the primary datasets listed in Table 1, from the classified LAS files:

The TNTmips batch script also generates ANZLIC compliant metadata, obtaining information, such as capture date and extent of coverage, from the input data and other GIS layers.

An example of the file naming and directory structure can be seen below. This refers to the classified LAS file (product 1b) in Valley 16, Campaspe, Site 7446. The data is in projection MGA94 zone 55.

\MDB\16-CMP\site_74446\1b_Classified_LAS\CMP_74446_z55.las

The ability to automate the generation of the primary datasets, produce metadata and perform internal checks within a single process contributed significantly to the efficiency of the project. The only labour intensive processes were the levelling of the LiDAR strips, and the manual classification and checking of the LIDAR point cloud.

The Vexcel imagery was captured, ortho-rectified against the LiDAR DEM, mosaicked, enhanced and written to standard files by separate processes implemented by Aerometrex Pty Ltd.

Table 1. The primary data sets.

No.	Name	Description	format
1a	Raw LiDAR	One file per swath, named according to Valley, Site ID, capture date, sortie, and UTM zone	LAS
1b	Classified LiDAR	One file per site, trimmed to site polygon, named according to Valley, site, zone.	LAS
2	Ground Points	Ground points without built structures	ascii text
3	Ground + Building Points	Ground points with built structures	ascii text
4	Vegetation Points	Vegetation Points	ascii text
5	DEM	Raster terrain surface without built structures – 1 metre grid spacing	Arc Grid
5	DTM	Raster terrain surface with built structures – 1 metre grid spacing	Arc Grid
7	CEM	Vegetation height above ground – 1 metre grid spacing	Arc Grid
8	PLR	Percent LiDAR returns by strata. The percentage of LiDAR returns within 17 height ranges above the ground: • 0.0 – 0.1 m • 0.1 – 0.5 m • 0.5 – 1.0 m • 1 – 2 m • 2 – 5 m • 5 – 10 m • 5 – 12 m • 10 – 20 m • 20 – 35 m • > 2 m • > 3 m • > 5 m • >10 m • > 12 m • > 20 m • > 35 m	Arc Grid
9	Contour	0.25 m interval contour	ESRI Shape
10	AOI	Site rectangle polygon extracted from GIS layer of all site rectangles	ESRI Shape

ACCURACY ASSESSMENT

Residual errors were calculated as the vertical difference between surveyed check points and the 1 metre DEM generated from the LiDAR. The Root Mean Squared Error and the standard deviations of the errors were calculated for each vegetation type. Table xx shows the results of this accuracy assessment.

Table 2.

Vegetation Type	RMSE	Std Dev	Check Points	Sites
Hard bare surface	0.206	0.213	2116	95
Open Grassland	0.210	0.232	203	13
Dense Vegetation	0.148	0.081	151	12
Grassland, low bushes	0.099	0.086	1156	30
Open Forests	0.138	0.182	183	14

The results show that the LiDAR accuracy is within the 0.5 metres specified in the contract in all vegetation types.

LiDAR was collected over some check sites on multiple sorties and it was observed that the residual errors were very consistent for all the sorties on a site. This confirms other lines of evidence that a significant portion of the residual errors results from the surveyed control sites and differences in the geoidal models used to calculate the LiDAR and control point stations. In Victoria for example, residual errors of up to 0.4m were observed between check points connected to permanent survey marks that make up the geodetic network, and the LiDAR which used an airborne GPS solution based on the CORS GPSnet. The CORS solution calculates elevations relative to the ellipsoid, which are then adjusted to the AusGeoid98 using the Geoscience Australia N-surface which is known to contain errors of up to 0.5m.

The accuracy of the LiDAR is therefore greater than suggested by Root Mean Squared Errors listed in table xx.

The higher accuracies achieved in different types of vegetation compared to Hard Bare Surfaces are not easily accounted for. Hypotheses could be proposed, taking into account the variations in the geodetic network and vegetation across the Basin, but such speculation is of little value without empirical evidence to support it. The data for individual valleys in Appendix XX shows significant discrepancies between the valleys.

VARIABLE EXTRACTION TOOL

A ‘Variable Extraction Toolkit’ was developed in the TNTmips object oriented Spatial Modelling Language to map channel and riparian features (secondary datasets) and also measure channel and riparian vegetation attributes (measurement datasets). The Variable Extraction Toolkit is integrated within a single a graphical user interface, and contains a number of interactive and automated feature extraction tools organised into a simple workflow.

The Variable Extraction Tool and operating manual are provided to the MDBA SRA. Two training sessions in the use of the Variable Extraction Tool were provided to the SRA remote sensing team.

The Variable Extraction Tool was used by Terranean to produce the secondary datasets and measurement matrices. Terranean has 8 TNTmips licences and was therefore able to process a number of valleys simultaneously.

SECONDARY PROCESSING

The primary datasets were used as the inputs to secondary processes that extracted river channel features. Two steps were involved: 1) mapping of water surfaces in the river channel, and 2) extraction of channel features and site zones.

Water surfaces within the river channels were mapped by manual interpretation of the Vexcel imagery and LiDAR point cloud, using TNTmips geospatial editor. In most sites exposed water is clearly visible in the false colour Vexcel imagery. Where the water surfaces are obscured by overhanging vegetation, the LiDAR ground points can be used to map the edges of water surfaces. This is because the infra-red laser beam is reflected strongly by water. Where the water surface is smooth the laser is reflected away from the LiDAR instrument and no signal is returned from ground level. In turbulent water the laser beam may be scattered, returning a signal to the LiDAR instrument; however, even in this situation, a contrast in the density of ground points usually allows water to be distinguished from ground.

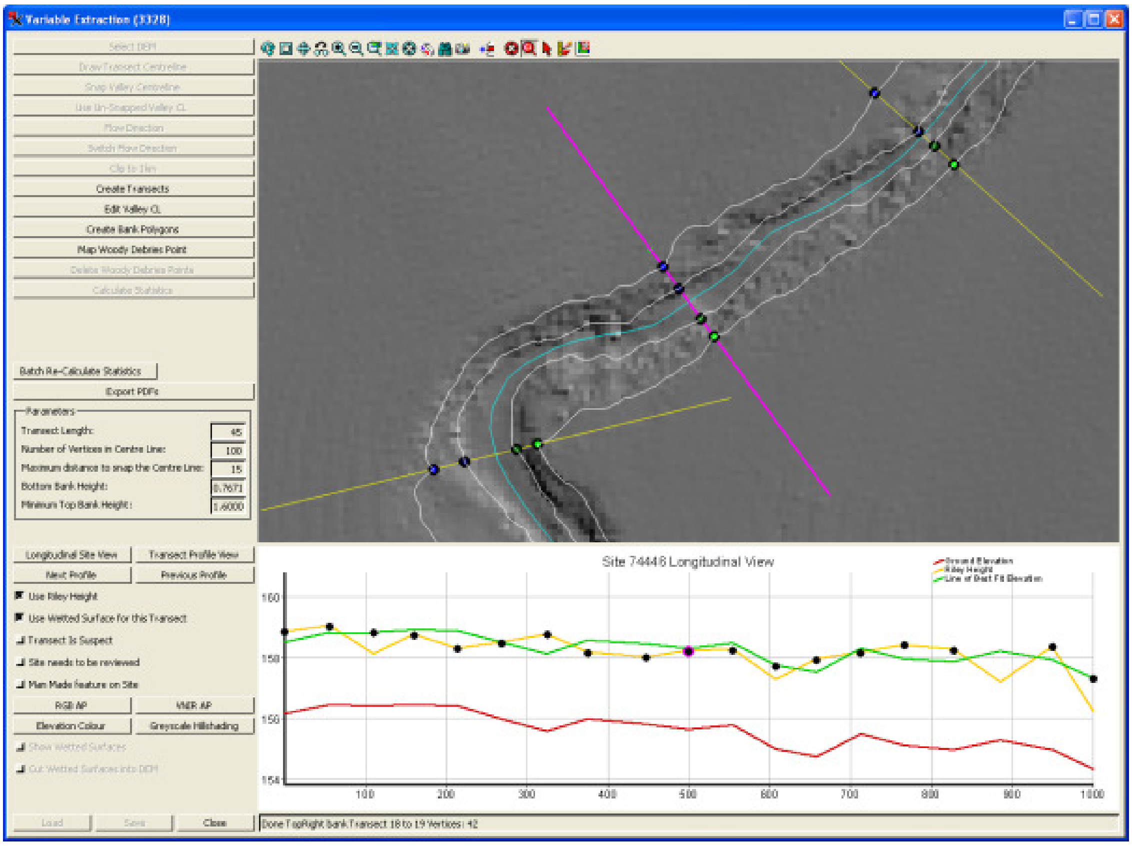

Figure 3. The Variable Extraction Toolkit interface.

The first step of the feature extraction process is to map the channel centreline. The approximate path of the river channel is digitised quickly over the relief shaded DEM. The variable extraction tool then maps the channel centreline as a line of best fit through the lowest part of the channel. Where water bodies have been mapped in the river channel, the channel centreline passes equidistantly though the centre of each water polygon. The channel centre line is automatically trimmed to a length of one kilometre.

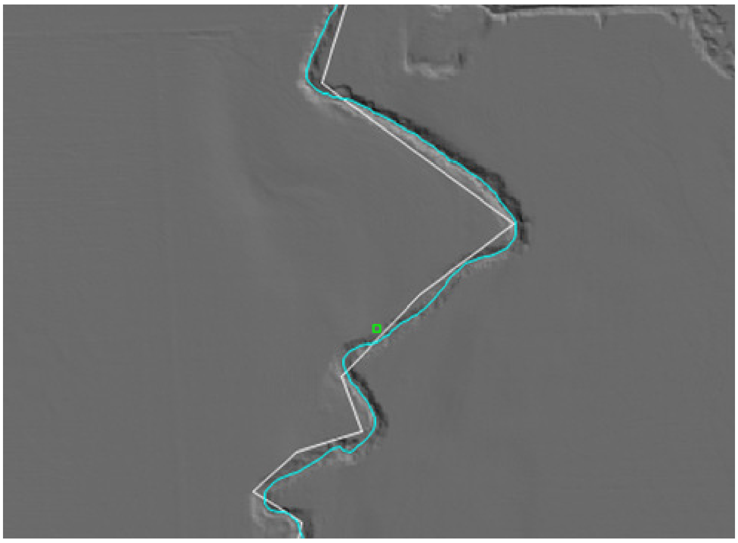

Figure 4. Mark a general channel route, then auto-generate a 1km channel centre line.

Flow direction is determined visually by inspection of the longitudinal profile of the channel centreline and if necessary by reference to a basin-wide drainage network. The channel must be oriented correctly in order to assign the left and right banks relative to the flow direction.

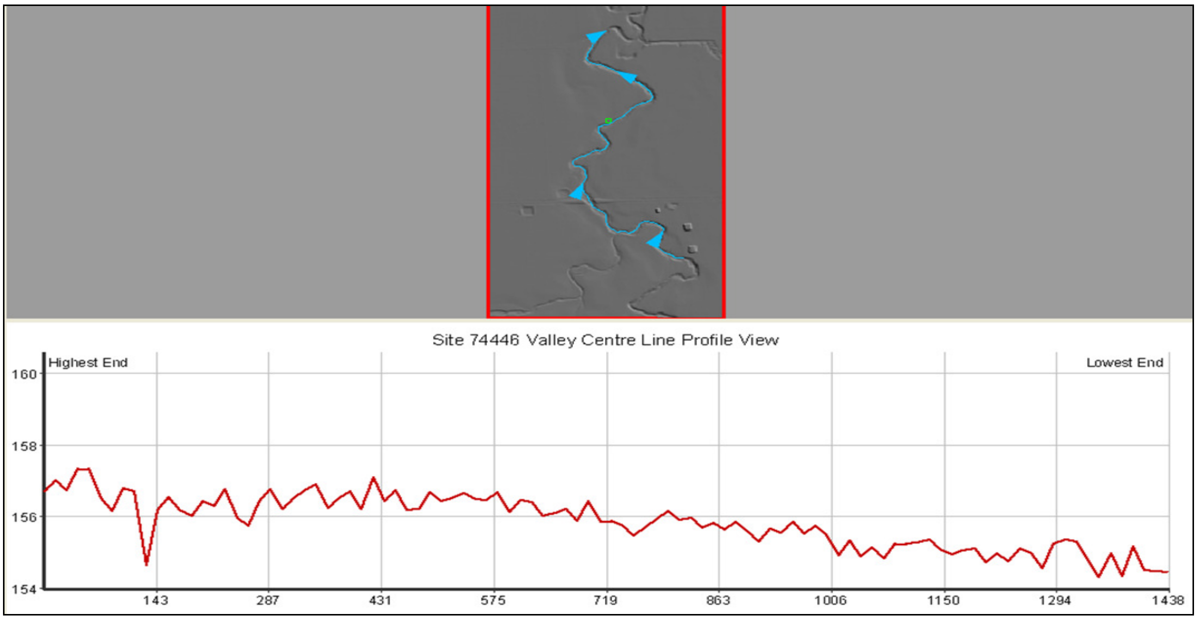

Figure 5. Set flow direction.

Nineteen transects are semi-randomly generated at right angles to the channel centreline.

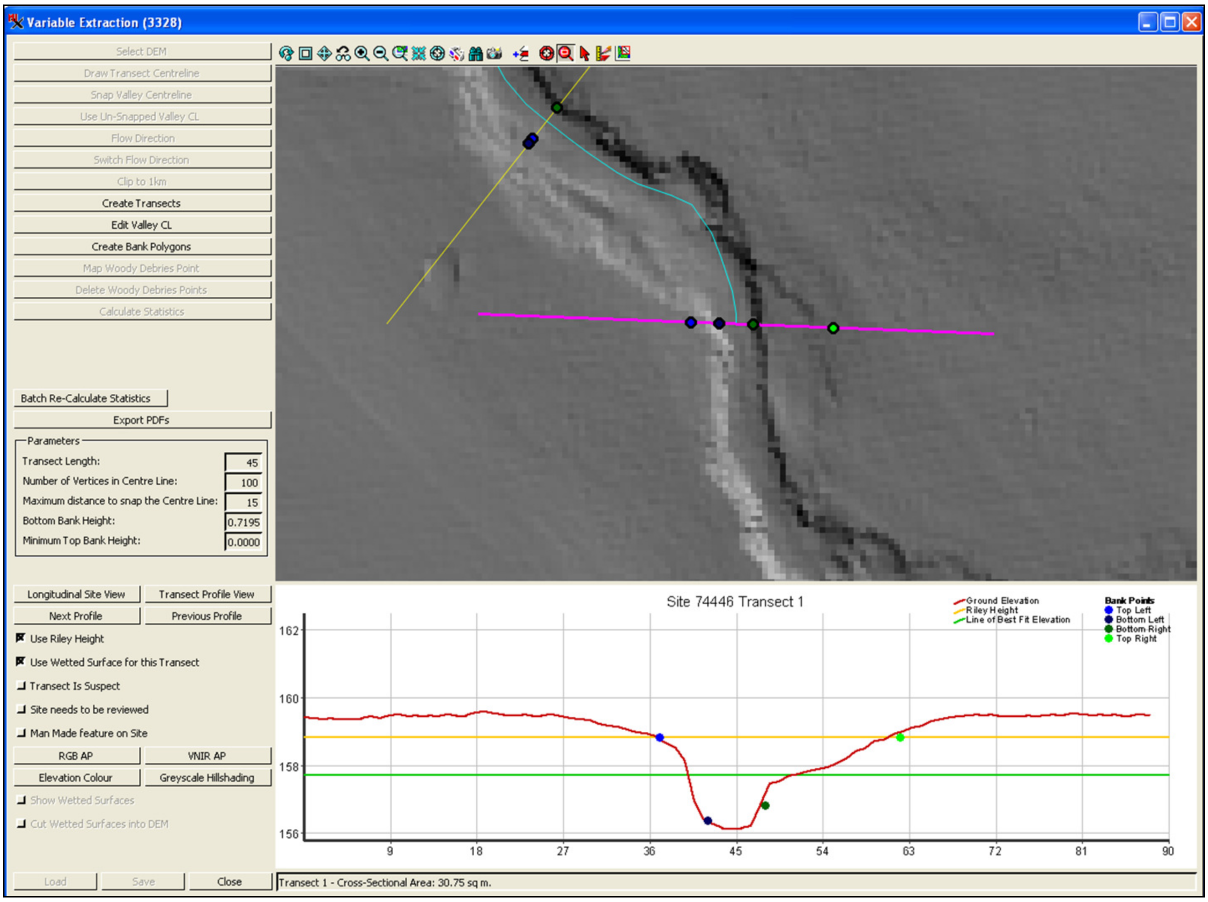

Figure 6. Generate transects and top-bank bottom-bank points.

On each transect, points representing the tentative location of the left and right top bank and left and right bottom bank are generated. The top bank points are initially located at the first Riley Bench Index maxima. The Riley Bench Index is calculated as the change in channel width over the change in channel slope. Maxima in the Riley Index occur on the inside edge of horizontal surfaces (Pickup, 1976).

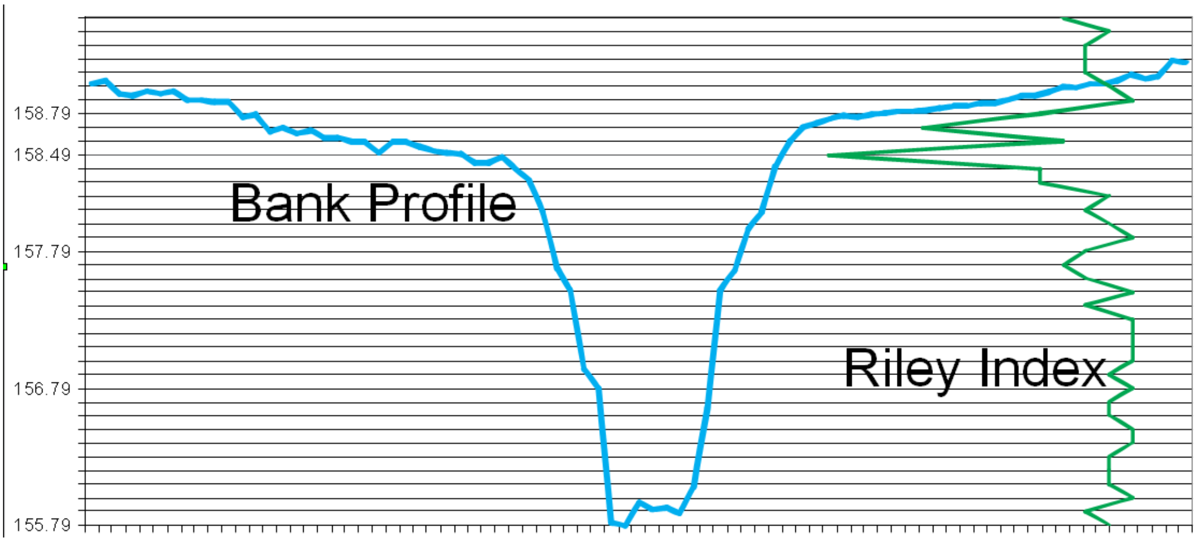

Figure 7. Riley Index plotted with bank profile.

The aim is to find the top and bottom of the active channel, which must be distinguished from minor channels within the active channel and wider incised channels. It was necessary to provide sufficient flexibility to enable the operator to guide the program to choose the correct channel, while minimising the amount of subjectivity that could be introduced. This was achieved by: 1) choosing between Riley Bench Index and Line of Best Fit (of the bank height above the channel centreline) as the criteria for selecting the bank points on the profile, 2) allowing the user to specify a minimum bank height for all the transects on a site, 3) by rejecting transects or sites that are unsuitable for the analysis.

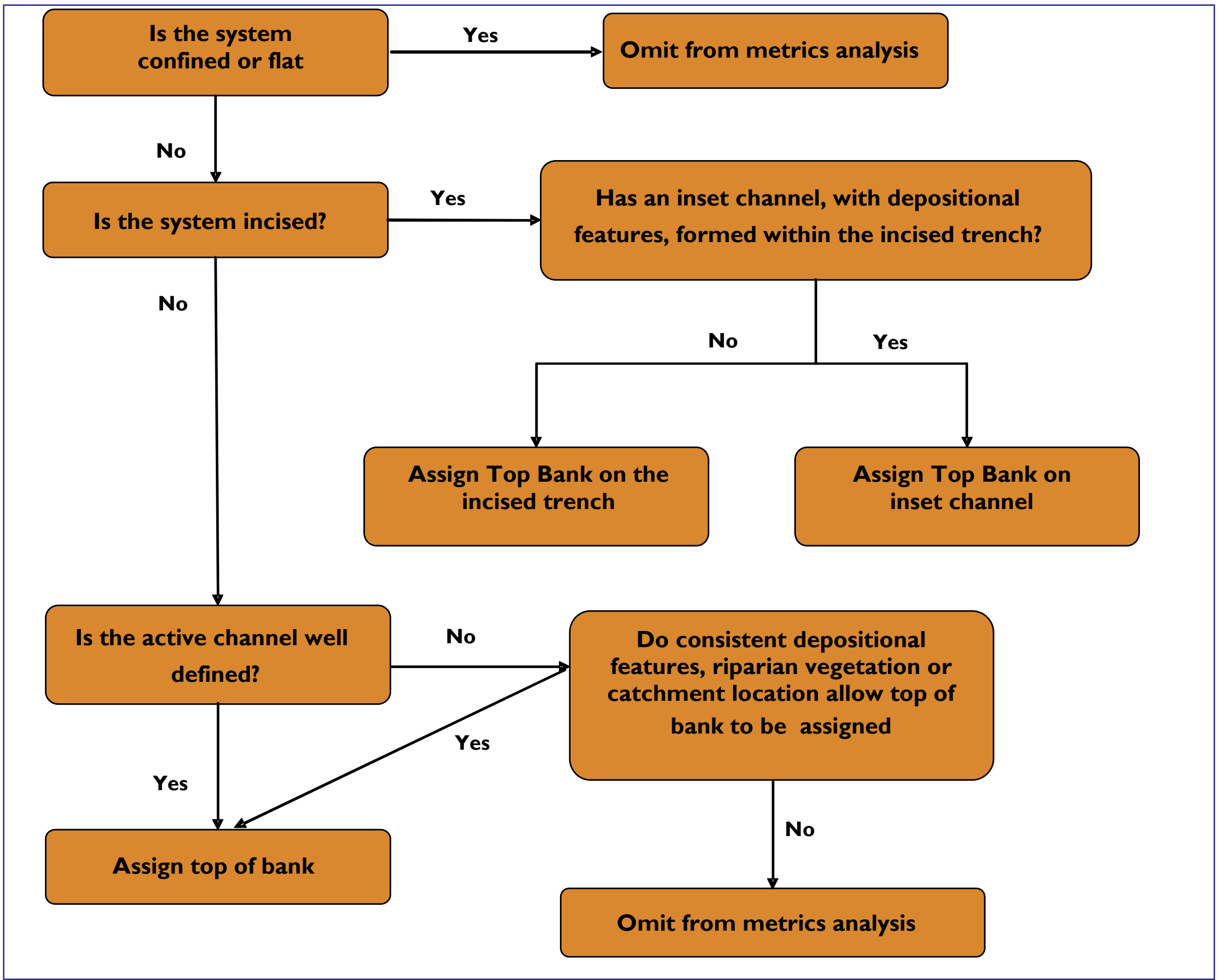

Figure 8. The decision process for mapping active channel banks.

The operators view the channel in 3D using anaglyph glasses and can switch between two different visualisations of the DEM surface and the false colour and natural colour Vexcel images of the site. Thus information, such as the presence of exposed sediments and vegetation, can be taken into account. As the operator discards transects as appropriate (e.g. at confluences) and switches bank points between the Riley Index and Line of Best Fit, the Line of Best Fit generally converges with the Riley Index as viewed in longitudinal profile.

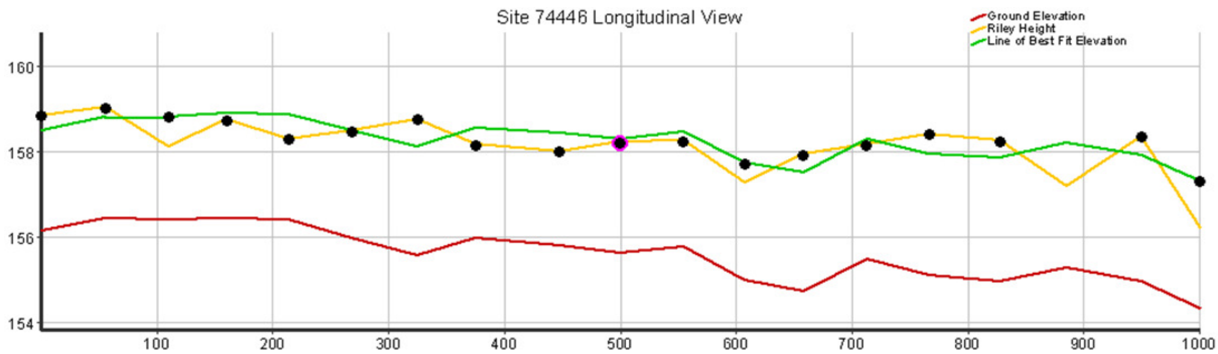

Figure 9. Longitudinal profile of a channel site, showing the Riley Index (yellow), The Line of Best Fit (green) and the channel centreline (red).

After the bottom bank and top bank points have been generated for each of the 19 transects, top and bottom bank lines are interpolated between the points using a path following algorithm that seeks to optimise the bank lines by applying, in order of priority, the following criteria:

1. Minimise change of slope along the bank line
2. Minimise the length of the bank line
3. Follow the zone of maximum terrain surface inflexion.

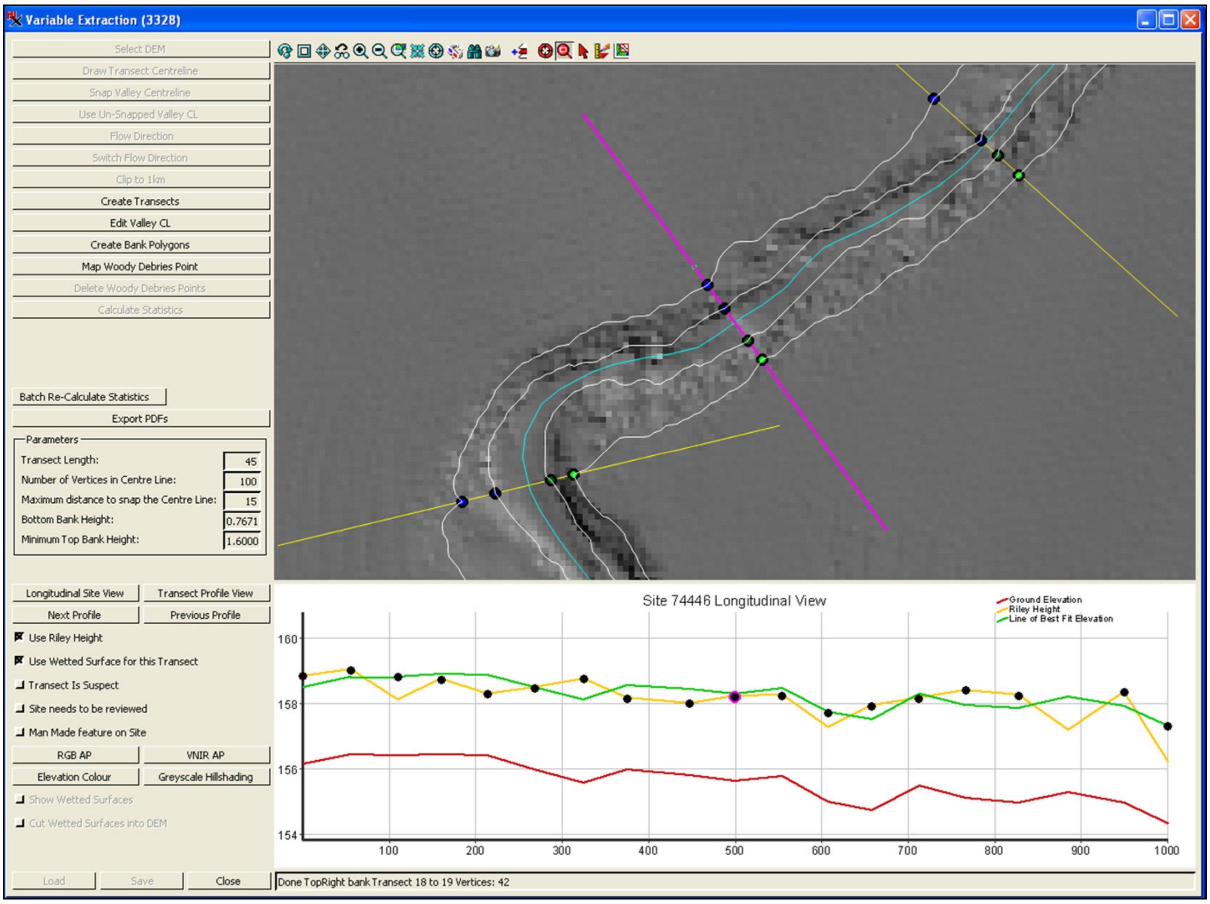

Figure 10. Interpolate top and bottom bank lines.

The accuracy of the bank lines is less critical, for the purposes of this project, than the accurate placement of bank points on transects. This is because the physical form measurements apply directly to the transects, while the bank lines are used only for segmenting the site into channel, banks and floodplains for the purpose of measuring vegetation structure in these 'Site Zones'. Therefore no interactive control was provided for the mapping of bank lines.

After generating the bank lines, the site is segmented into the Site Zones listed in Table 3 and illustrated in Figure 11.

Table 3. Site zone polygons for vegetation measurements

Zone	**Description**
Left bed	Between channel centreline and left bottom bank
Right bed	Between channel centreline and right bottom bank
Left bank	Between left bottom bank and left top bank
Right bank	Between right bottom bank and right top bank
25LB	25 metre buffer from left top bank
25RB	25 metre buffer from right top bank
50LB	Between 50 metre buffer from left top bank and 25 metre buffer from left top bank
50RB	Between 50 metre buffer from right top bank and 25 metre buffer from right top bank
50LP	From 50LP to site boundary
50RP	From 50RP to site boundary

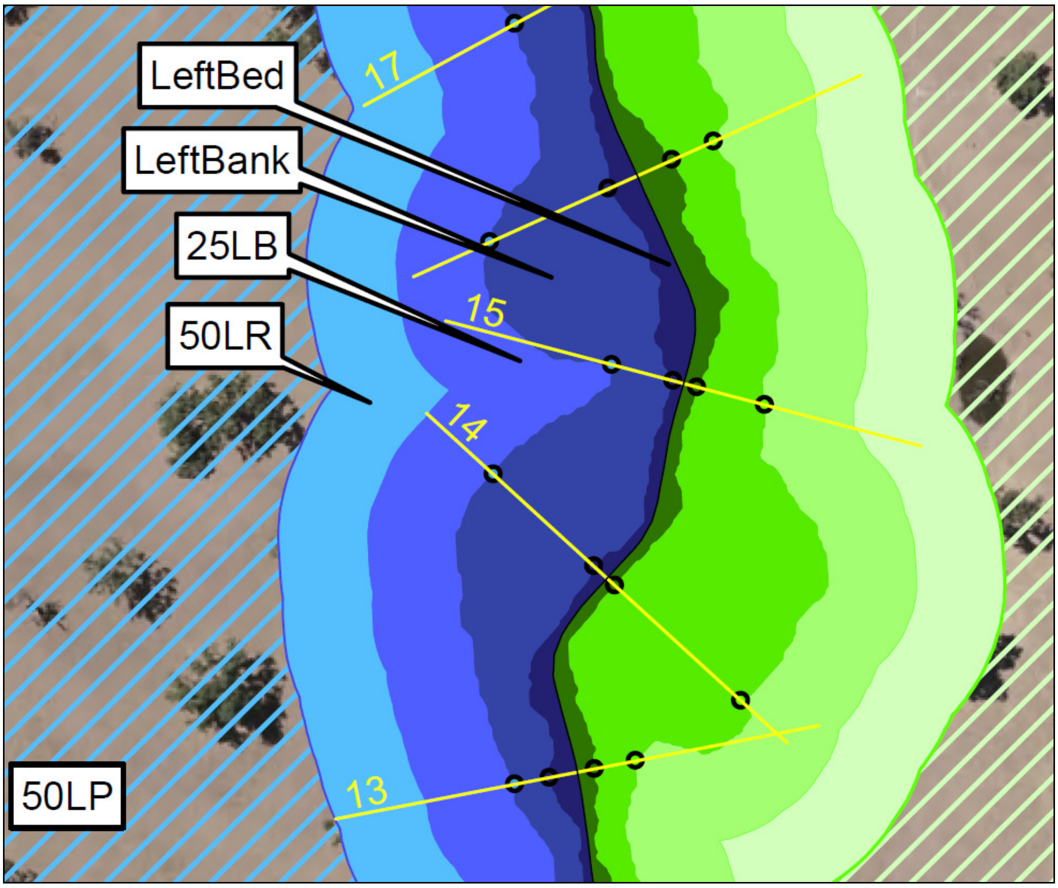

Figure 11. Site polygons segmenting channel and adjacent riparian zone.

The secondary data sets produced by these methods are listed in Table 3. Each data set includes ANZLIC compliant metadata.

Table 4 Site zone polygons for vegetation measurements

No.	Name	Description	Format
1	ChannelCL	Computed channel centreline -lowest path or mid path through water.	Shape + TNTmips
2	Transects	Maximum 19 transects for a site	Shape + TNTmips
3	BankPts	Left and right top and bottom points for all transects on a site	Shape + TNTmips
4	BankLines	Interpolated left and right, top and bottom bank lines for a site.	Shape + TNTmips
5	BankPolygons	Polygons generated from bank lines and perpendicular end line.	Shape + TNTmips
6	SiteZones	Site zone polygons described in table 2.	Shape + TNTmips

DATA MEASUREMENT

The final data generated for the project consist of two data matrices: 1) a Vegetation measurement matrix, and 2) a Physical Form measurement matrix. These matrices are to be used as inputs for subsequent statistical analyses.

The spatial units for the vegetation measurements were produced by intersecting the Site Zone polygons with vegetation polygons from the best available healthy vegetation cover mapping, to mimic a pre-1750 vegetation map that had been compiled for the MDBA for this purpose. For each resulting polygon, representing the intersection of a site zone polygon with a vegetation polygon, the mean of the Percentage LiDAR Returns for each of the 17 strata listed in Table 1 are calculated as polygon attributes. These data represent the vertical distribution and density of foliage by vegetation type across the zones associated with each river channel.

Intersecting the vegetation polygons with the Site Zone polygons, calculating the mean PLR for each stratum as polygon attributes and writing the results to a data matrix in CSV format is performed as a single process within the Variable Extraction Tool.

Similarly, a range of Physical Form measurements is generated from the transect profiles, bank points and channel centreline. These attributes, listed in Table 4, are calculated as attributes of the GIS line features and are exported to a data matrix for each site as an automated process.

Table 4. Physical Form Measurements

Input Variable	**Description**
ChnlCent	Channel centreline; length in metres
ValCent	Valley centre; length in metres
MaxElev	Maximum elevation of channel centreline
MinElev	Minimum elevation of channel centreline
ElevDiff	Elevation Range of the channel centreline
TranMin	Minimum elevation of transect
TBnkAng(L&R)	Left and right bank top angles
TBnkHt(L&R)	Left and right bank top height
BBnkHt(L&R)	Left and right bank bottom height
TBnkArea	Cross sectional area below top bank level
AngLn(L&R)	Length of profile from bottom bank to top bank – left and right
SAngLnR	Length of profile from spill height to bottom bank
Convex(L&R)	Cross sectional area of the bank profile above the bank angle line.
Concave(L&R)	Cross sectional area of the bank profile below bank angle line.
Inflect(L&R)	Number of times bank lines crosses bank angle line
Length(L&R)	Length of bank lines for the site – left and right banks
WetBed	Does transect intersect water? Yes or No
WetWdth	Width of water at transect from wetted area layer
BedWdth	Distance between left and right bottom banks
ChnlWdth	Distance between left and right top banks
ChnlDpth	Channel Depth; height of spill level above bottom bank
AreaChan	Channel area – between left and right top banks
AreaBnk(L&R)	Area between top and bottom banks for the site – left and right
StrmPow	Stream Power; (max elev minus min elev) / valley centreline length
ManMade	Evidence of manmade channel or features, or water control

Appendix A. Detailed Accuracy Assessment

			Hard bare surface	Open Grassland	Dense Vegetation	Grassland, low bushes	Open Forests
1	Gwydir	RMSE	**0.165**	**0.233**	**0.193**	**0.123**	**0.236**
		Std. Dev.	0.155	0.068	0.181	0.058	0.011
		N	*128*	*66*	*12*	*30*	*3*
2	Castlereagh	RMSE	**0.142**				**0.223**
		Std. Dev.	0.164				0.070
		N	*37*				*40*
3	Border Rivers	RMSE	**0.182**	**0.205**			**0.207**
		Std. Dev.	0.100	0.047			0.035
		N	*86*	*7*			*6*
4	Warrego	RMSE	**0.079**	**0.096**		**0.080**	
		Std. Dev.	0.098	0.072		0.108	
		N	*196*	*7*		*19*	
5	Paroo	RMSE	**0.106**			**0.141**	
		Std. Dev.	0.138			0.140	
		N	*49*			*27*	
6	Lower Murray	RMSE	**0.150**			**0.222**	**0.033**
		Std. Dev.	0.156			0.181	0.027
		N	*78*			*67*	*10*
7	Darling	RMSE	**0.149**		**0.401**	**0.149**	**0.135**
		Std. Dev.	0.156		0.218	0.028	0.029
		N	*81*		*30*	*10*	*5*
8	Macquarie	RMSE	**0.263**		**0.222**		
		Std. Dev.	0.229		0.118		
		N	*62*		*32*		
9	Condamine	RMSE	**0.190**	**0.295**		**0.311**	
		Std. Dev.	0.117	0.039		0.049	
		N	*138*	*10*		*15*	
10	Lachlan	RMSE	**0.196**	**0.156**		**0.138**	
		Std. Dev.	0.211	0.027		0.055	
		N	*138*	*12*		*10*	
11	Murrumbidgee	RMSE	**0.174**	**0.029**		**0.236**	
		Std. Dev.	0.141	0.032		0.073	
		N	*160*	*12*		*11*	
12	Central Murray	RMSE	**0.411**			**0.243**	
		Std. Dev.	0.042			0.078	
		N	*106*			*15*	
13	Wimmera	RMSE	**0.311**	**0.279**		**0.261**	**0.267**
		Std. Dev.	0.047	0.035		0.028	0.038
		N	89	65		12	6

			Hard bare surface	Open Grassland	Dense Vegetation	Grassland, low bushes	Open Forests
14	Avoca	RMSE	**0.446**				**0.026**
		Std. Dev.	0.016				0.030
		N	*15*				*10*
15	Loddon	RMSE	**0.267**			**0.120**	**0.142**
		Std. Dev.	0.091			0.071	0.260
		N	*479*			*144*	*21*
16	Campaspe	RMSE	**0.384**			**0.298**	
		Std. Dev.	0.051			0.015	
		N	*43*			*12*	
17	Broken	RMSE	**0.010**	**0.022**		**0.017**	
		Std. Dev.	0.013	0.020		0.017	
		N	*15*	*12*		*11*	
18	Goulburn	RMSE	**0.169**		**0.140**	**0.199**	**0.187**
		Std. Dev.	0.116		0.151	0.114	0.095
		N	*75*		*23*	*104*	*24*
19	Ovens	RMSE	**0.118**		**0.113**	**0.146**	**0.066**
		Std. Dev.	0.109		0.164	0.119	0.068
		N	*68*		*31*	*634*	*38*
20	Kiewa	RMSE	**0.094**			**0.069**	**0.066**
		Std. Dev.	0.101			0.011	0.083
		N	*14*			*12*	*9*
21	Mitta Mitta	RMSE	**0.102**		**0.235**	**0.140**	
		Std. Dev.	0.037		0.086	0.032	
		N	*19*		*11*	*12*	
22	Upper Murray	RMSE	**0.142**	**0.135**	**0.128**	**0.026**	**0.096**
		Std. Dev.	0.082	0.060	0.073	0.030	0.076
		N	*40*	*12*	*12*	*11*	*11*