AGRICULTURE AND FOOD
www.csiro.au

PAT - Precision Agriculture Tools Plugin for QGIS

User Manual

Christina Ratcliff, David Gobbett, Rob Bramley

January 2019 PAT v0.1.1

Wine Australia

Citation

Ratcliff C, Gobbett D, Bramley R (2019) Precision Agriculture Tools (PAT) plugin v0.1.1 for QGIS: User Manual. CSIRO, Australia.

Copyright

© Commonwealth Scientific and Industrial Research Organisation 2018. To the extent permitted by law, all rights are reserved and no part of this publication covered by copyright may be reproduced or copied in any form or by any means except with the written permission of CSIRO.

Important disclaimer

The notes and software described here are based on a number of technical, circumstantial or otherwise specified assumptions and parameters. The user must make their own assessment of the suitability for use of the information or material contained in or generated from these notes. To the extent permitted by law, CSIRO excludes all liability to any party for expenses, losses, damages and costs arising directly or indirectly from using these notes and software tools.

Enquiries

Contact PAT@csiro.au for queries about the tools described in this manual.

Contents

PAT - Precision Agriculture Tools 4
1 Installing, Upgrading and Uninstalling 7
1.1 License 7
1.2 Install PAT 7
1.3 Update **PAT** 11
1.4 Uninstall PAT 12
2 Individual Tools 13
2.1 Create Block Grid 13
2.2 Clean, Trim, Normalise Point Data 14
2.3 Run Kriging Using VESPER 18
2.4 Import VESPER Results 22
2.5 Create Polygon from On-The-Go GPS Point Trail Data 23
2.6 Rescale or Normalise a Raster 27
2.7 Generate Random Pixel Selection 28
2.8 Extract Pixel Statistics for Points 29
2.9 Calculate Image Indices for Blocks 32
2.10 Resample Image Band for Blocks 35
2.11 Create Zones with *k*-means Clusters 38
2.12 Create Strip Trial Points 40
2.13 Whole-of-block Analysis 42
2.14 Settings 46
3 Technical Notes 47
4 How-To's 48
4.1 Create a block boundary polygon from a CSV of GPS collected points 48
4.2 Loading PAT Symbols into QGIS 50
4.3 Analysing strip trial data using the moving window t-test 52
5 Contributors 53
6 Acknowledgments 53
7 References 54

PAT - Precision Agriculture Tools

The Precision Agriculture Tools (PAT) plugin is a suite of open source tools developed by CSIRO for Precision Agriculture (PA) data analysis. The tools run within Quantum Geographic Information System (QGIS), a free and open-source desktop geographic information system that supports viewing, editing, and analysis of geospatial data. PAT aims to provide an easy-to-use interface for processing PA data through an established workflow developed for constructing maps using on-the-go data (e.g. from yield monitors or EM38 soil surveys) as shown in Figure 1 and Table 1 (Bramley and Williams 2001; Taylor *et al.* 2007; Bramley *et al.* 2008; Bramley and Jensen 2014). It also includes tools for analysis of remotely sensed imagery and some on farm experiments and for clustering map layers to create potential management zones. Over time more tools and 'how-to' instructions will be added to expand the functionality and usefulness for both practical and research purposes.

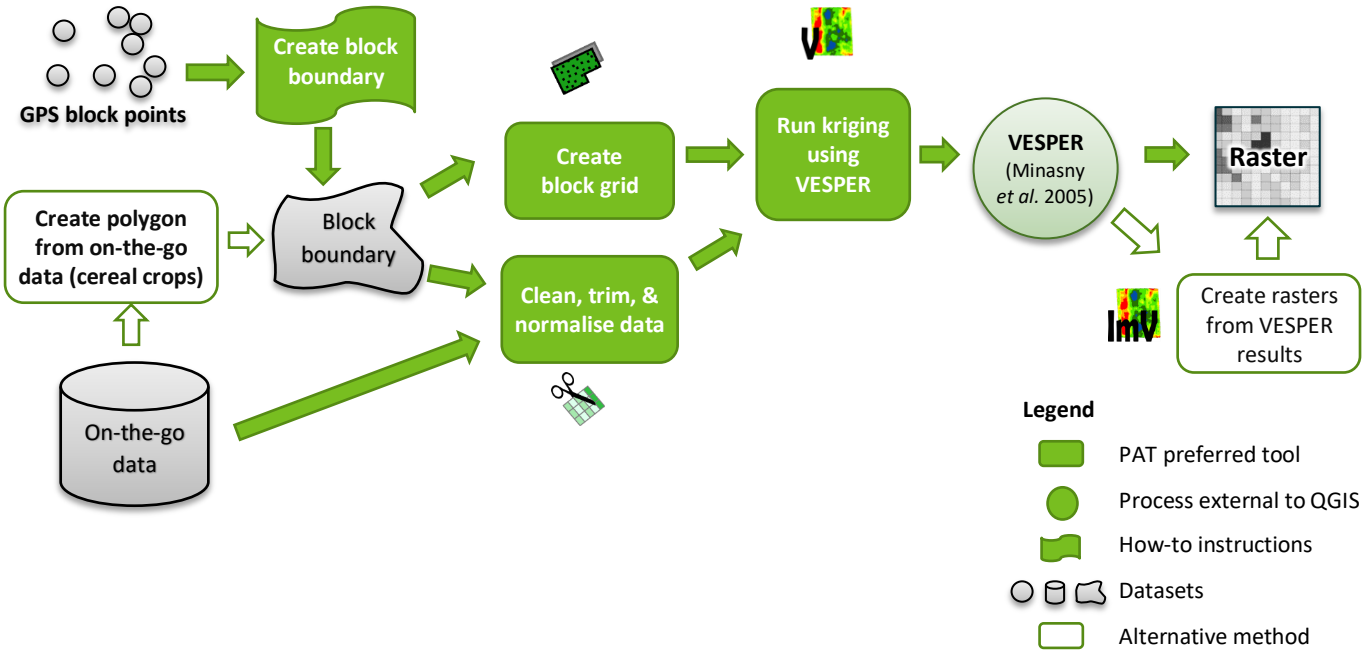


Figure 1 The basic map production workflow used in PAT and some of the existing tools available for processing on-the-go data.

Table 1 Brief description of tools available in PAT

	PAT TOOLS	DESCRIPTION
	Create Block Grid	Convert polygon features (e.g. a block boundary – see page 48) to • a TIFF raster and • a VESPER text grid file of X,Y point values used by VESPER.
	Clean, Trim, Normalise Point Data	Process an on-the-go data file (e.g. from yield monitors) by applying clipping, cleaning and filtering rules and output as a CSV file.
	Run Kriging Using VESPER	Create a VESPER control file and data files and run VESPER kriging.
	Import VESPER Results	Convert VESPER outputs to raster TIFF format.
	Create Polygon from On-The-Go GPS Point Trail Data	Generate a polygon block boundary from on-the-go data (e.g. from yield monitors) containing GPS points. The GPS points must cover the entire block (not just be a set of points around the boundary - see page 48for instructions on creating a boundary from a set of surveyed data points)
	Rescale or Normalise a Raster	Create rasters by • rescaling (standardising) values between a fixed range (i.e. 0-1, or 0-255) • normalising values to a mean of 0 and standard deviation of 1
	Generate Random Pixel Selection	Generate a selection of random pixels from a raster and save to a points Shapefile.
	Extract Pixel Statistics for Points	Extract pixel statistics using a square neighbourhood footprint from multiple rasters at set locations.
	Calculate Image Indices for Blocks	Resample and smooth imagery to a larger pixel size, as well as calculate indices such as PCD and NDVI.
	Resample Image Band for Blocks	Clip imagery to one or more blocks and resample to a different pixel size.
	Create Zones with *k*-means Cluster	Create zones with *k*-means clustering
	Create Strip Trial Points	Tool for analysing simple strip-based experiments. It generates pairs of points from a central line within a strip and offset at a distance outset the strip. These points can be used for the comparison of performance along and adjacent to a strip trial (such as an N-rich strip) using the Lawes and Bramley (2012) moving window t-test method.
	Whole-of-block	Analyse whole-of-block experiments where 2 or 3 treatments are applied across a block or paddock.
	Error! Reference source not found.	User settings for PAT. Provides the ability to set data directories and the location of VESPER (if installed).
	About	About Precision Agriculture Tools (PAT). This provides information on the currently installed release, and relevant open source licences.
	Help	Display the PAT user manual.

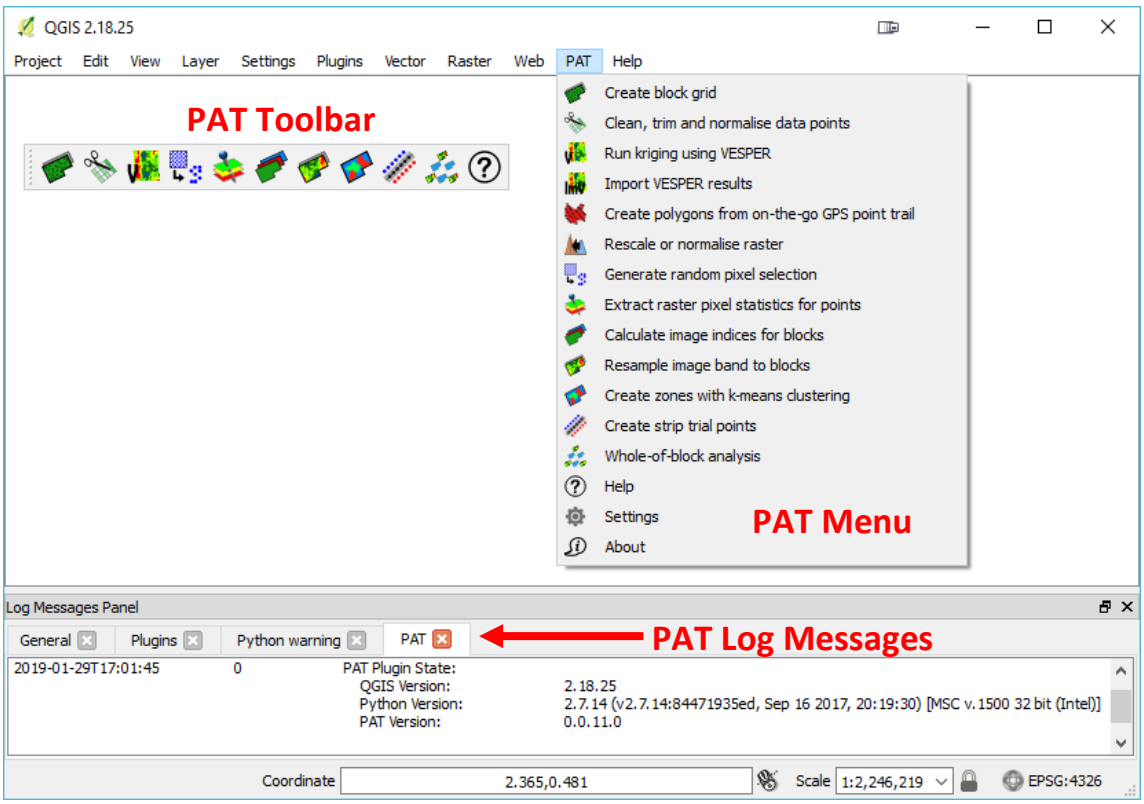


Figure 2 PAT toolbar, menu and log messages panel within QGIS

1 Installing, Upgrading and Uninstalling

1.1 License

The PAT plugin for QGIS is free to download and use under a GPL V3 Open Source license. PAT license details can also be found in the ***About*** tool under PAT menu. The companion Python library 'pyprecag' is also freely available and is released under the CSIRO Open Source Software License Agreement (variation of the BSD / MIT License). All source code and license files are available from https://github.com/CSIRO-Precision-Agriculture.

1.2 Install PAT

Requirements

External Applications

VESPER: To use the ***Run Kriging Using VESPER*** tool requires installation of VESPER (note: only available for Microsoft Windows operating systems). VESPER is a kriging (map interpolation) program (Minasny *et al.* 2005). It needs to be installed independently if you would like to undertake kriging of maps using VESPER. VESPER is **not** distributed with this plugin but is the recommended map interpolation tool. To download or view more information on VESPER visit https://sydney.edu.au/agriculture/pal/software/vesper.shtml.

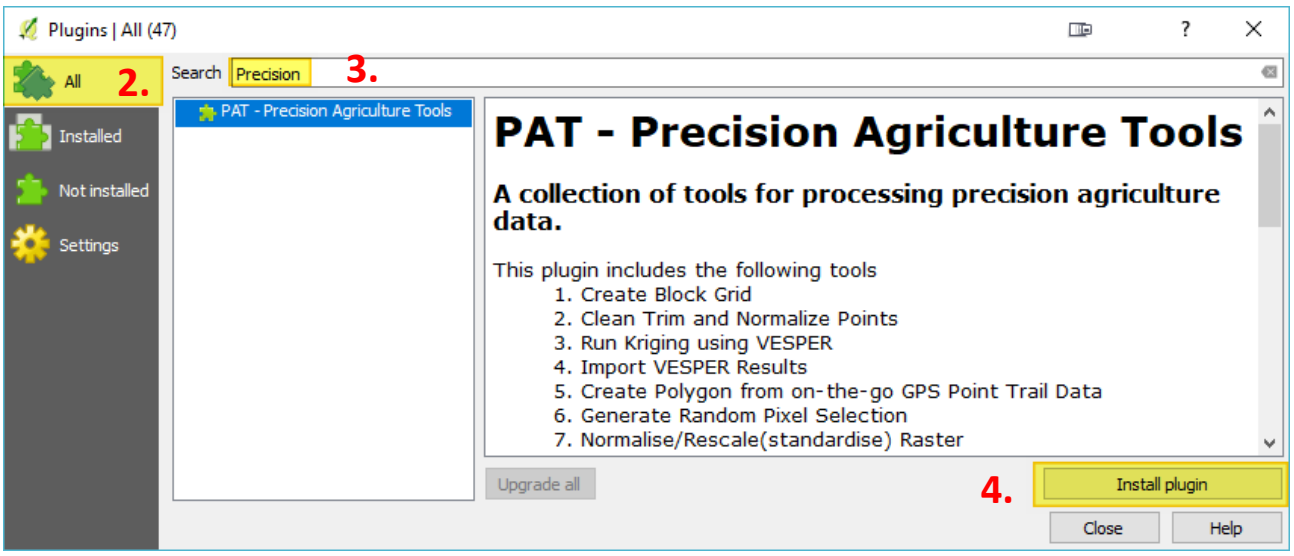
R 3.5.1+: The R statistical language is required to carry out the analytical processing for the ***Whole-of-block Analysis*** (Section 2.13). R is **not** distributed with this plugin. If ***Whole-of-block analysis*** is required then you must download and install R in addition to QGIS. The minimum R version suitable for this is v3.5.1. See the **Configuring QGIS to use R** section on page 8.

QGIS LTR 2.18.12+: Download and install the QGIS standalone long term release (version 2.18.12 or newer, but not version 3) from the QGIS download page. **Note that the PAT tools do not currently work with QGIS version 3 or later.**

Additional Python Packages: QGIS includes Python and numerous Python packages. However, PAT requires the following additional packages: ***pyprecag, fiona and rasterio***. The ***fiona and rasterio*** packages have been bundled with the PAT plugin and instructions on how to install them are included.

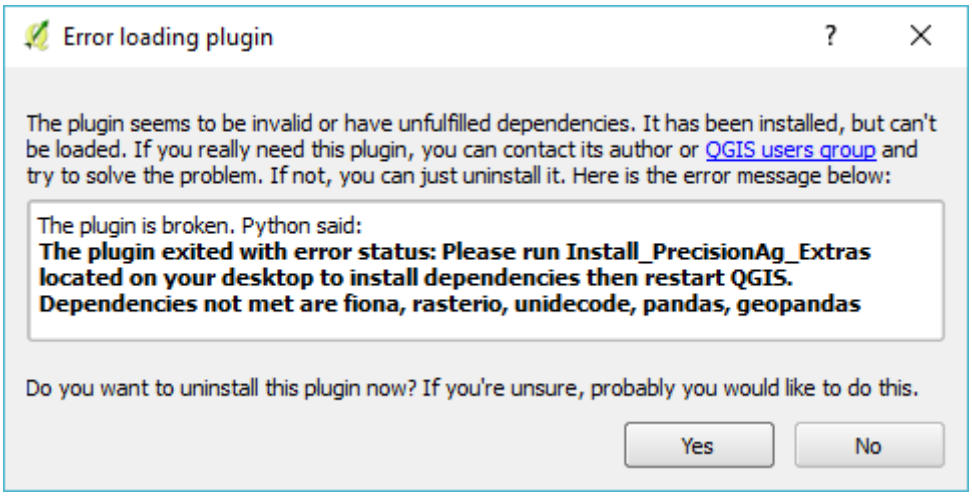
Install the Precision Agriculture Tools (PAT) Plugin.

1. In QGIS open the plugin manager.
   Plugins Menu → Manage and Install Plugins
2. Select the ***All*** or ***Not installed*** section.
3. Search for and select **Precision Agriculture Tools (PAT)**.
   Click ***Install plugin***.



Installing or Upgrading PAT Python Dependencies

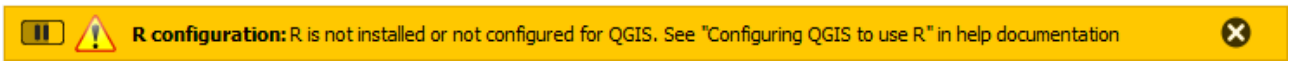
While installing or upgrading PAT, a check will be completed to ensure the required Python packages are installed on your system. If this check fails, a message box and an ***Error loading plugin*** dialog (as displayed) will appear warning you that the plugin is broken. This is normal. If this occurs. Click **No** to dismiss this box and **Quit** QGIS.



A shortcut will appear on your desktop named ***Install_PAT_Extras***. Once you have quit QGIS, **run the shortcut** to install the missing components choosing **YES** to any other messages which may appear. Restart QGIS and the check will run again to ensure the installation occurred correctly. If the PAT menu and toolbar are not present then reinstall and/or check/activate it using the plugin manager (see above). QGIS should now contain the PAT menu, toolbar and log panel similar to Figure 2 on page 6.

Configuring QGIS to use R

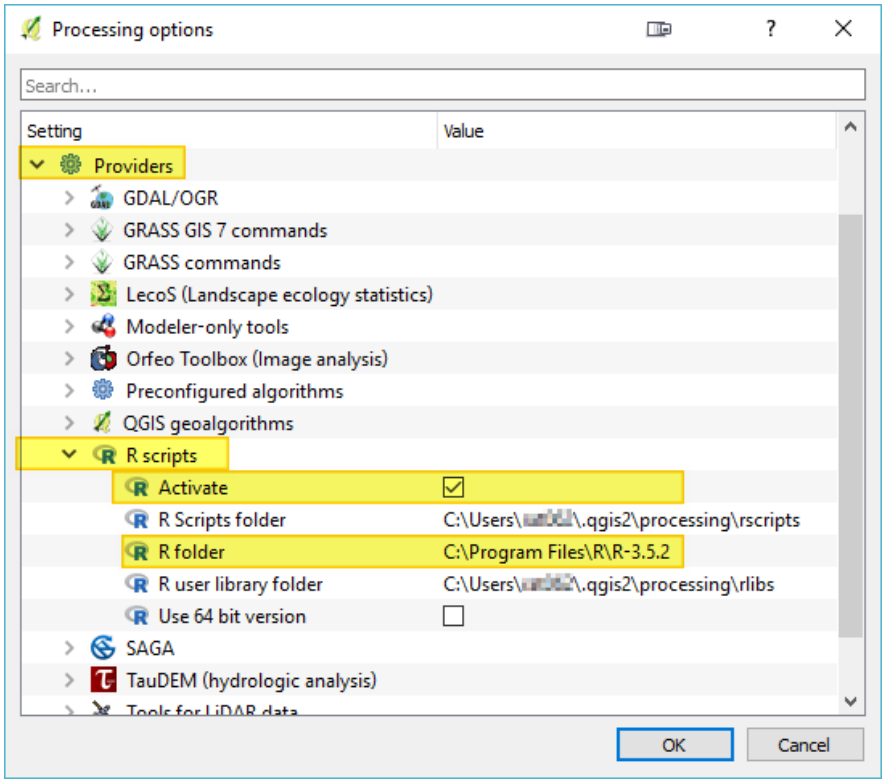
If R is not installed and you launch the ***Whole-of-block analysis tools*** via the menu or toolbar you will a warning message as shown.



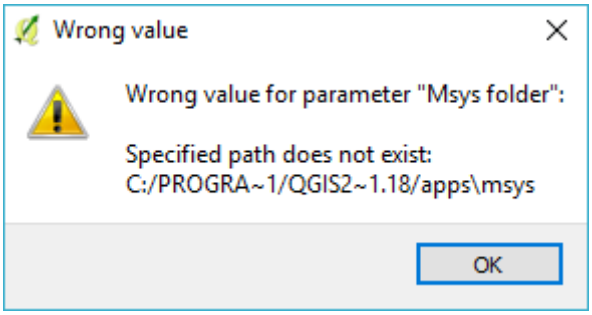
To resolve this, follow the following steps to configure R and QGIS

1. Download R from Windows version can be downloaded from http://r-project.org/ . The minimum R version suitable for this is v3.5.1.
2. Install using all the default options
3. If required, activate the **Processing** plugin in QGIS plugin manager.
    a. In QGIS open the plugin manager.
    Plugins Menu → Manage and Install Plugins
    b. Select the ***Installed*** *section*
    c. Find **Processing** and ensure the adjacent checkbox is ticked.
    d. Close the Plugins dialog.
    This will give you a processing menu and the processing side panel.

4. From the ***Processing*** menu select ***Options***.
    a. Expand ***Providers***
    b. Expand ***R scripts***
    c. Tick on ***Activate***
    d. Ensure that the R installation folder is shown against ***R folder***
    e. Click OK.

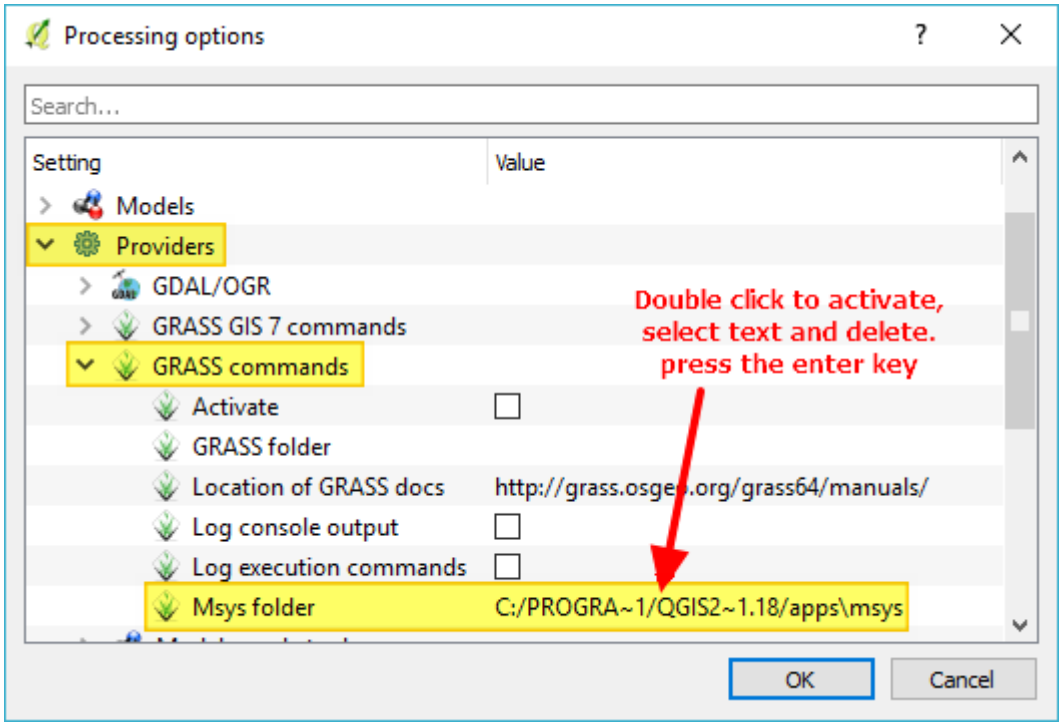


When clicking OK you may get a ***Wrong value*** error message like this



and your mouse cursor will remain a circle. To resolve treat the mouse circle as an arrow and

1. Under ***Providers***, expand ***GRASS commands***
2. Double click the text to the right of Msys folder, select and delete it and press the enter key.



3. Click OK
4. Close and re-open QGIS to reset your mouse cursor.
   It should now all work.

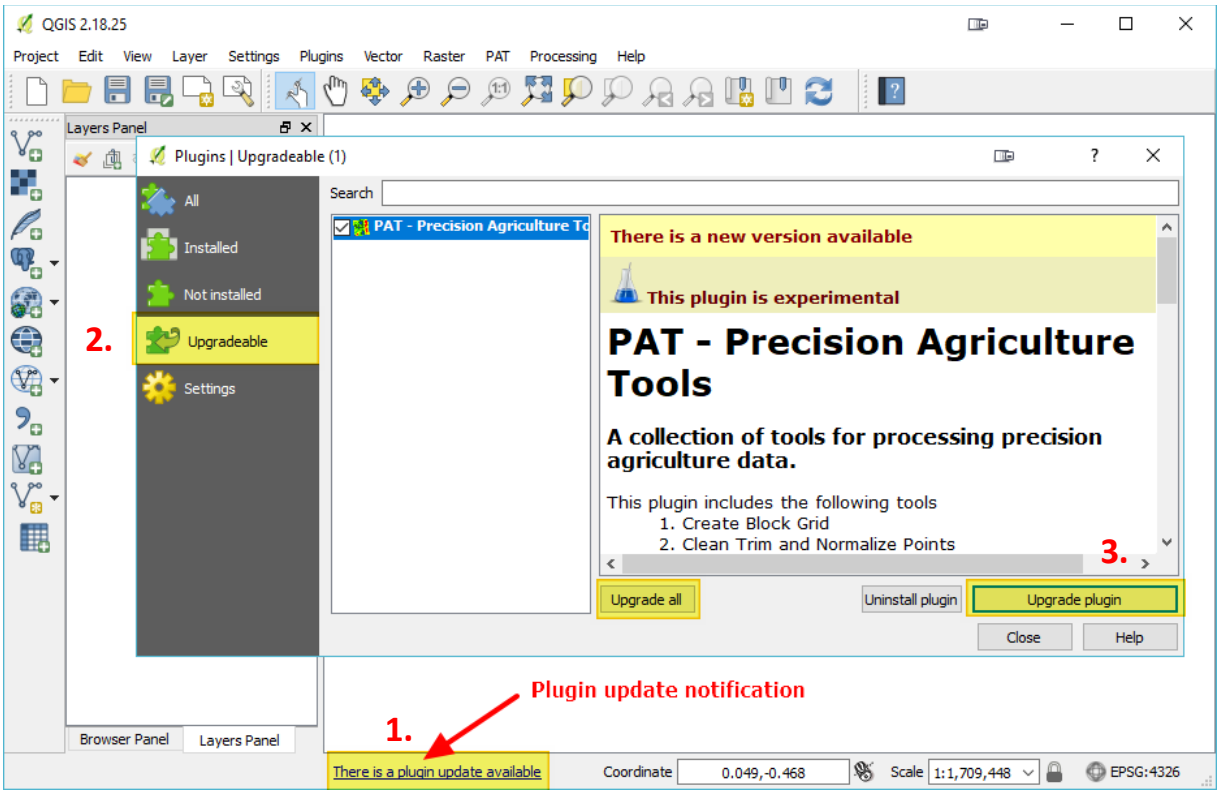
1.3 Update **PAT**

If the ***Check for Updates*** option is checked on in the plugin manager, QGIS will advise the user if new plugins are available, or if installed plugins have been updated. This notification is displayed in the QGIS interface's status bar as shown below. It is recommended that ***Check for Updates*** is checked.

To Update

1. In QGIS open the Plugin Manager by clicking on the link in the status bar or via ***Plugins. Menu → Manage and Install Plugins***
2. Either select the ***Upgradeable*** left side tab or search for your plugin.
3. Upgrade by selecting the PAT plugin and click ***Upgrade plugin***

A check will be run to ensure the Python packages required by PAT are installed and are of the correct version. If this check fails follow the instructions in *Installing or Upgrading PAT Python Dependencies* to upgrade the dependencies.

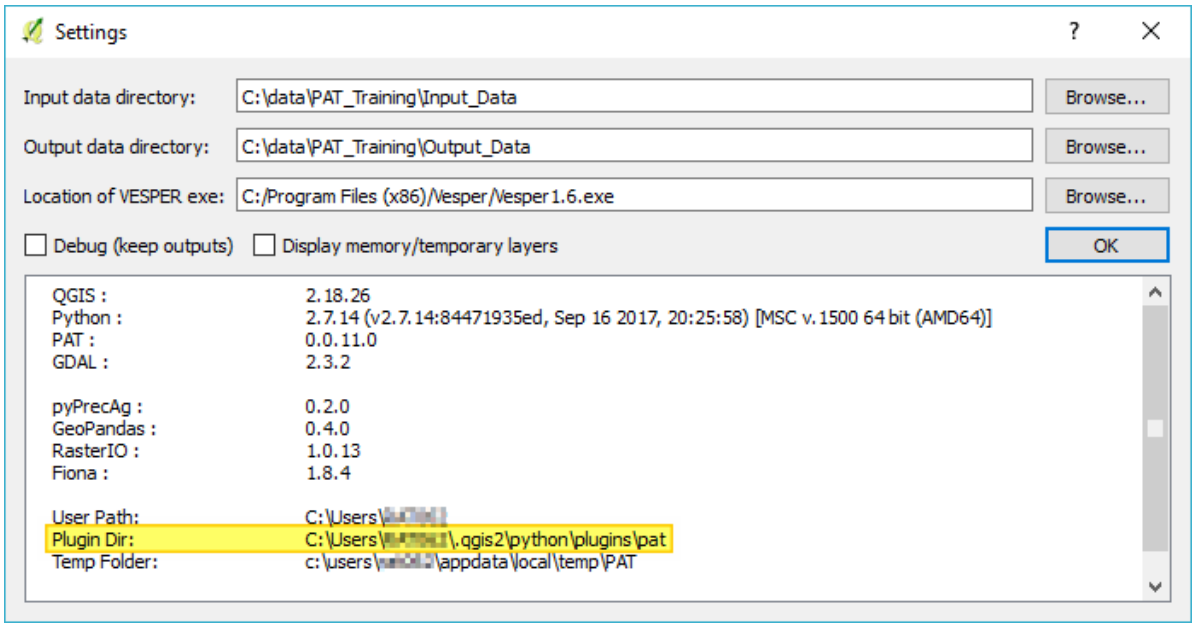


1.4 Uninstall PAT

If you wish to uninstall PAT and its Python dependencies you must first uninstall the Python dependencies, followed by the plugin as instructed below.

1.4.1 Uninstall PAT Python dependencies for all users.

1. Navigate to the current user's PAT installation directory. This folder is listed in as **Plugin Dir** in the ***settings dialog*** as shown highlighted in yellow below.



2. Find the ***UnInstall_PAT_Extras.bat*** file in the python_packages folder.
3. **Right click** the file and select ***Run as Administrator*** and choose ***Yes*** for any messages which appear.

1.4.2 Uninstall PAT Plugin.

PAT Python dependencies should be uninstalled PRIOR to uninstalling the PAT plugin. The plugin can be uninstalled in one of two ways.

1. **Via the QGIS Plugin Manager**
    a. In QGIS, open the plugin manager, and find the PAT plugin.
    b. Click Uninstall plugin.
2. **Via Windows Explorer.**
    a. Navigate to the current users PAT installation directory as show in 1.4.1.1 above.
    b. Delete the entire pat folder.

2 Individual Tools

2.1 Create Block Grid

Summary

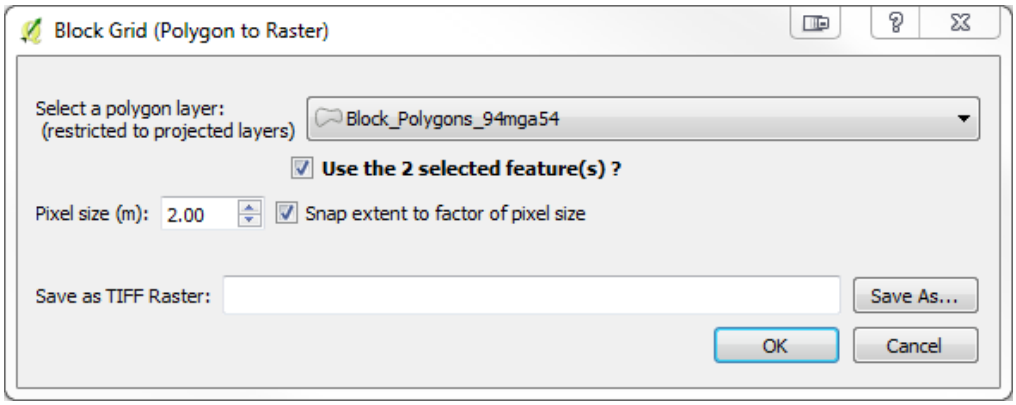
This tool converts polygon features such as a block boundary to a raster grid using a set pixel size. This is a critical part of setting up the analysis environment for creating maps from on-the-go data because it generates the base grid onto which maps will be interpolated, by using the outer most extent of the block boundary as the 'frame' for the grid. Page 48 details how to generate a block boundary if one is not already available.

The raster outputs created by the Block Grid tool are:

- a TIFF raster and
- a grid file of X,Y point values used by VESPER for kriging.

Areas inside the polygon boundary will be assigned a value of 1 while areas outside will be assigned a no-data value of -9999.

Dialog



Block Grid (Polygon to Raster)		
Select a polygon layer		A layer containing the block boundary polygon features in a projected coordinate system to be converted to the raster outputs.
Use selected features	Default is unchecked	If checked, will only use the selected features, if unchecked, all the features will be used. Useful in creating a grid for a single field from a boundary layer comprising more than one field.
Pixel size (m)	0.00m to 6 km Default is 2 m	The pixel size to assign to the raster outputs. This is expressed in metres. Recommended Values: Viticulture: 2 m Sugarcane: 2 m Broadacre grains: < 200 ha 5 m > 200 ha 10 m
Snap extent to factor of pixel Size	Default is checked	Snap the output raster extent to a factor of the pixel size. This will ensure adjacent rasters use a common origin which is important for future analysis.
Save as TIFF raster	Default derived from input layer	The output raster TIFF file to be created. The VESPER grid file will have a suffix of _v.txt

2.2 Clean, Trim, Normalise Point Data

Summary

This tool processes on-the-go data files (e.g. from a yield monitor or EM38 soil survey) containing GPS coordinates recorded as latitude and longitude in decimal degrees, to output point values in a projected coordinate system. The tool also applies cleaning and filtering rules.

This tool:

- retains all columns (except coordinate columns) from the original file.
- converts coordinate columns to a projected coordinate system and renames them to Easting and Northing. An additional column (EN_EPSG) will be created and assigned the EPSG number for the projected coordinate system used to reproject data. An EPSG number is an international numbering system for coordinate systems (see http://www.epsg.org/)
- optionally saves a Shapefile version matching the output CSV file. A second Shapefile will also be saved containing the string *_removed* in its filename. This Shapefile will contain all the points the filter *discards* and will be attributed by filter type. The description of the values can be found in Table 2

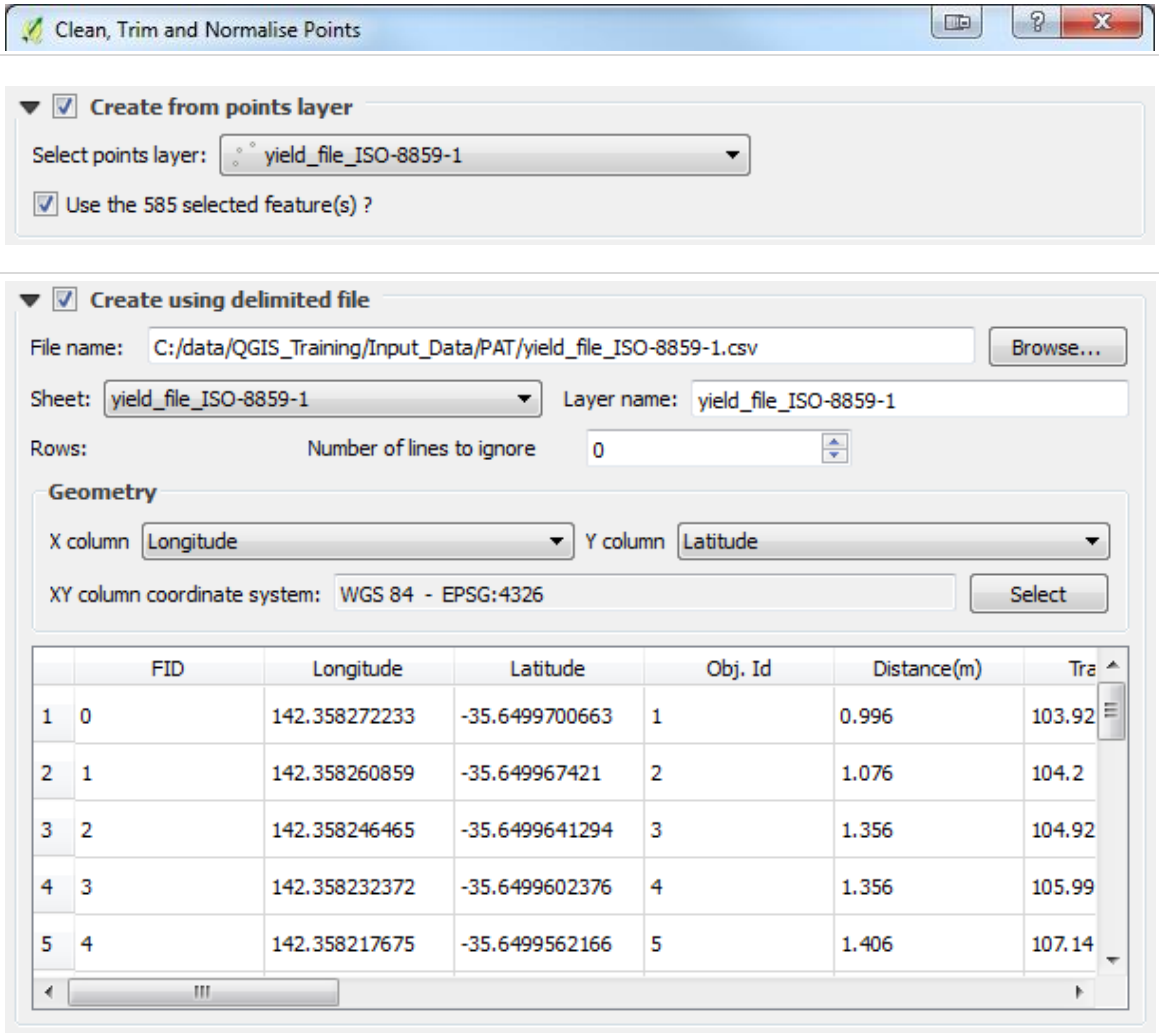
This tool filters data by:

- optionally clipping the data to a Shapefile boundary (i.e. removes points that lie outside the boundary)
- optionally removing values from a given column which are null (missing) or are less than or equal to zero
- creating a normalised column using a prefix of *nrm_* and calculating normalised values for the data such that their mean is zero and standard deviation is 1. The normalised value of column Z is calculated as (Z - mean(col Z))/(s.d.(col Z))
- optionally trimming normalised outliers based on a set number of standard deviations. This trim can optionally be performed iteratively with the normalised value re-calculated for each iteration. It is recommended that, by default, normalised data are trimmed to +/- 3 standard deviations.
- optionally thinning the data by removing points that are closer together than a set distance. For data that have been georeferenced using a differential GPS, we recommend removing points that are closer than 1 m apart.

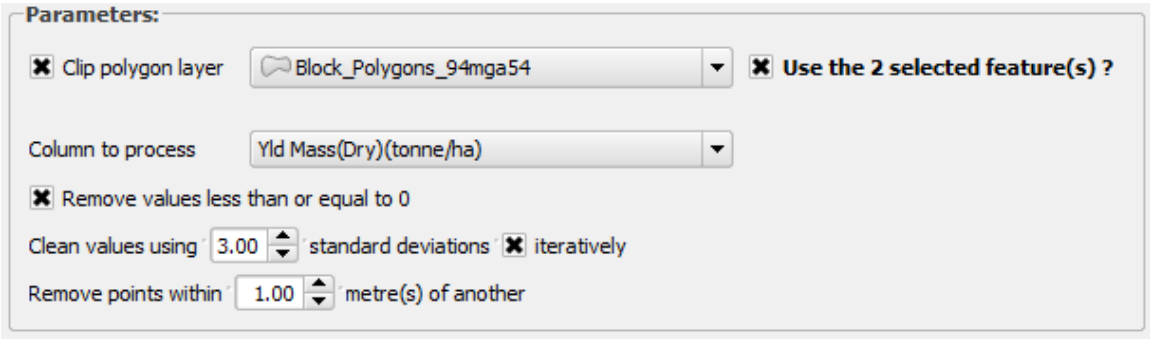
As part of the filtering process, the tool may rename some data columns to adhere to Shapefile column name limitations. The new names are displayed as a PAT log message and written to the log file. For more information on the location of the saved log file and temporary folder refer to the Technical Notes section on page 47.

Once processing is complete, the results of filtering are shown as a PAT log message. Only filters which remove points are listed. When iteratively filtering, results of all iterations are shown. An explanation of these is shown in Table 2.

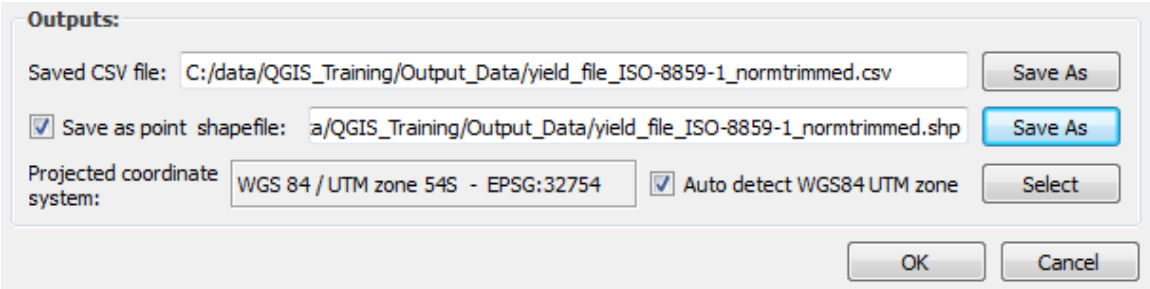
Dialog



Choose whether to use GPS data currently loaded in QGIS as a Points Layer		
Select points layer:		A point layer currently loaded in QGIS.
Use selected features	Default is unchecked	If checked, use only the selected features to generate a polygon If unchecked, all the features will be used
Or alternatively using a delimited text file		
Filename	Options include XLS, TXT, CSV	The filename of the table file.
Sheet		The sheet containing data. If the file type doesn't support sheets, then the filename will be used.
Layer name	Default is the sheet name.	The QGIS layer name for the output point and/or polygon files.
Number of lines to ignore	Default is 0	The number of lines at top of file to ignore. Can be used for skipping header text.
X column Y column	The tool will display a guess of these columns	The columns representing the X and Y coordinates within the input file.
XY columns coordinate system		The coordinate system of the XY columns. Note: GPS most commonly record data in WGS 1984 (EPSG: 4326)
Clip polygon layer	Default is unchecked	A layer containing polygon features used to clip the points data. Any points that fall outside the polygon features will be removed.



Use selected features	Default is unchecked	If checked, use the selected features to clip the input points. If unchecked, all the features will be used.
Column to process		The column to normalise, trim and clean.
Remove values less than or equal to 0	Default is checked	If checked, rows from the chosen column will be removed where the value is null or less than or equal to zero.
Clean values using X standard deviations	0.00 to 5.00 Default is 3 Use 0 to omit	Points where the normalised value is outside a set number of standard deviations will be removed.
Iteratively	Default is checked	Clean the points iteratively using the set number of standard deviations until all values are within +/- the set standard deviation.
Remove points within X metres of another	0.00 to 20.00 Default is 1 Use 0 to omit	Any points spaced at less than this distance apart will be removed.



Save CSV file	Derived from layer name	The name and path of the output CSV file containing the final data points.
Save as point Shapefile	Default is unchecked. Derived from layer name	If checked, save a Shapefile version of the output CSV file. A second point Shapefile will be saved with a suffix of ***_removed*** containing the points discarded and attributed by a filter column.
Projected coordinate system		The output projected coordinate system of the Shapefile.
Auto detect WGS84 UTM zone	If Auto detect is checked the following rules will apply: - Will try to match the input if it is a projected coordinate system - If the inputs are in Geographic's (lat/long) it will calculate the relevant UTM Zone Note: If in Australia then MGA GDA 1994 zones, will be used – see Figure 5 on page 47.	

Notes

Table 2 A description of the filter type used when reporting filtering results.

ORDER	FILTER TYPE	DESCRIPTION
01	clip	Points removed when clipping by a polygon.
02	zero	Nulls and/or Zeros removed.
03	3.0 std iter 1	Points removed during the standard deviation (std) iteration (iter) number 1 and uses 3 standard deviations.
04	3.0 std iter 2	Points removed during the standard deviation (std) iteration (iter) number 2 and uses 3 standard deviations.
05	3.0 std iter 3	Points removed during the standard deviation (std) iteration (iter) number 3 and uses 3 standard deviations.
06	pointXY (1.0 m)	Points removed which are spaced at less than the specified distance apart, in this case **1m**
07	pointX (1.0 m)	Points removed after sorting by the X coordinate, which are spaced at less than the specified distance apart, in this case **1 m**
08	pointY (1.0 m)	Points removed after sorting by the Y coordinate, which are spaced at less than the specified distance apart, in this case **1 m**
	Pts remaining	The number of points left after filtering occurs.
	Total	The total number of points in the dataset

- The following notification occurs when a delimited text file contains invalid characters in a column name. A copy of the file with corrected column names will be made in the PAT temporary folder (location shown in PAT log message). For more information on the location of the saved log file and temporary folder see the Technical Notes section on page 47.

2.3 Run Kriging Using VESPER

Summary

This tool is used for mapping on-the-go data, such as are collected by a yield monitor or an EM38 soil survey, following established protocols (Bramley and Williams 2001; Taylor *et al.* 2007; Bramley *et al.* 2008; Bramley and Jensen 2014) using local block kriging. It will create a VESPER control file and collate the files required for kriging. The following files will be created in a VESPER sub-folder located in the specified output folder.

- the VESPER control file. The control filename will be used as a base to derive other VESPER output files like the kriged map result.
- a subset of data to krige. All non-required columns are deleted.
- a Windows batch file (*Do_VESPER.bat)* which can be used to launch VESPER processing for all control files in the VESPER sub-folder. This process can be run outside of QGIS and the Python/pyPrecAg environment.

Any files matching the control file name will be deleted from the VESPER sub-folder including VESPER kriged results, to ensure that the files remaining belong to the newly create control file. If you want to retain old versions, use a different name for the control file.

The coordinate system of the CSV file will be determined by interrogating the EN_EPSG or ENEPSG columns generated as a by-product of running the ***Clean, Trim, Normalise Point*** tool. If the resulting value cannot be found or is incorrect, the coordinate system can be set manually. The resulting EPSG number will be stored within the control file and later used when converting VESPER outputs into raster TIFF format.

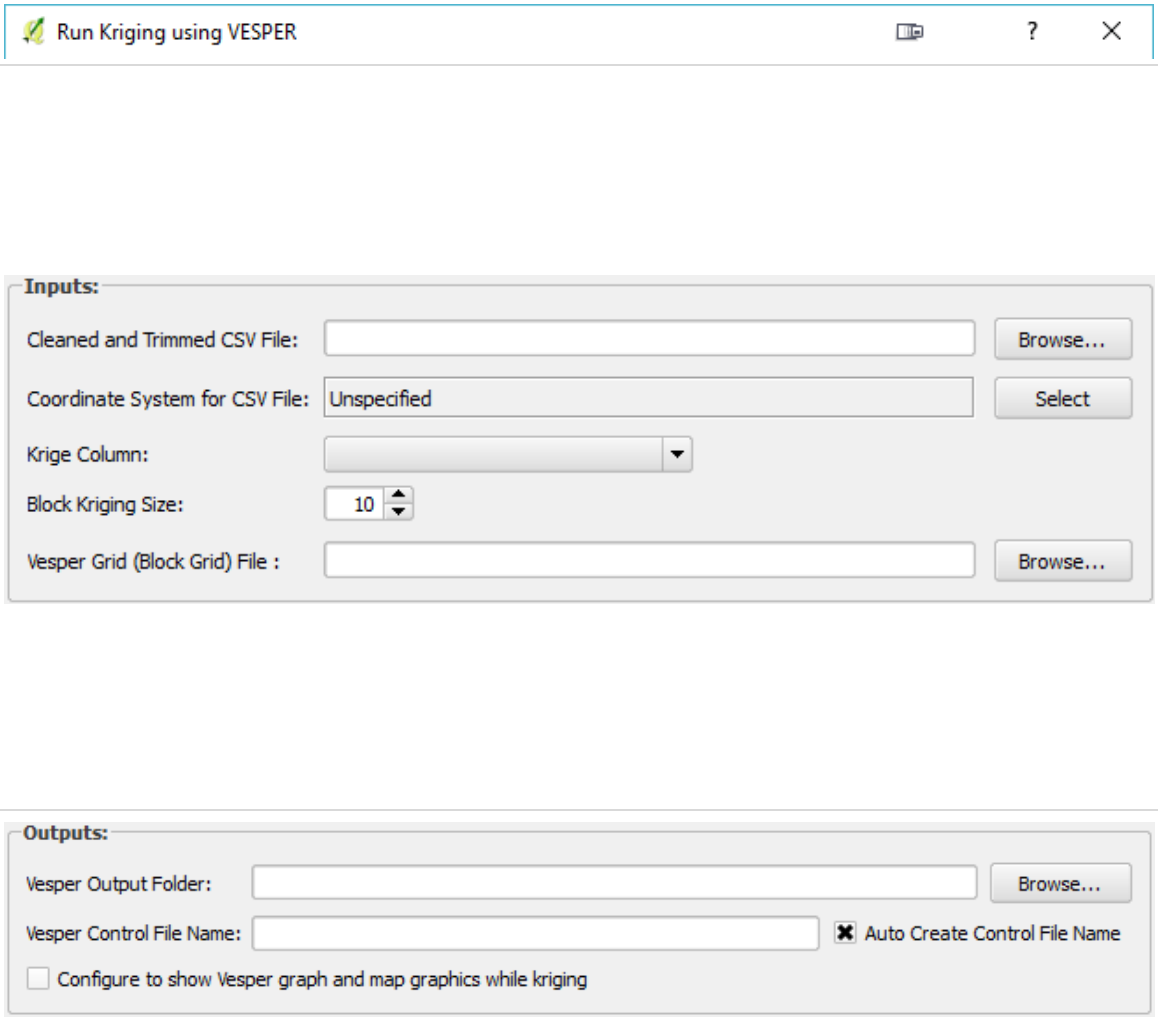
To krige (run) VESPER control files from within QGIS, VESPER must already be installed on the PC and configured in the PAT settings (see ***Settings Tool*** on page 46).

If the **Run VESPER Kriging Now** option is checked the output control file will be launched in VESPER and kriged immediately. Running this tool multiple times will add each output control file to the VESPER queue and run consecutively. The QGIS status bar is used to manage the queue; currently the queue can only be displayed and cleared. On the completion of each VESPER run, if the **Convert VESPER Files to Raster and Load in QGIS** option is checked the kriged results will be import to TIFF files and loaded into QGIS.

If the **Run VESPER Kriging Now** option is unchecked the tool will create the VESPER control, data and batch files in the selected folder allowing the user to run VESPER in their own time.

When **Run VESPER Kriging Now** or **Convert VESPER Files to Raster and Load in QGIS** options are unchecked the ***Import VESPER Results*** (post VESPER) tool can be used to import VESPER kriged results to TIFFs then loaded into QGIS.

Dialog



Cleaned and trimmed CSV file	A CSV file containing the data to krige. Use the ***Clean, Trim, Normalise Point*** tool to generate this input.	
Coordinate system for CSV file:	This is the coordinate system of the coordinate columns in the CSV file. By default the value of the EN_EPSG or ENEPSG column in the input CSV file will be used.	
Krige column	The column containing values to krige.	
Block kriging size	0 to 1000 Default is 10 Use 0 for Punctual Kriging	The value in meters to set as the block size when using block kriging in VESPER. This value should be approximately 5 times the pixel value. Recommended Values: Viticulture: 10 m Sugar: 10 m Broadacre grains: < 200 ha 30m > 200 ha 60 m
VESPER grid (block grid) file		The VESPER grid file representing the area to be kriged. Use the ***Create Block*** Grid ***Tool*** to generate this input.
VESPER output folder		The output folder for the VESPER files. A sub-folder called VESPER will be created if required.
VESPER control filename		The name to use as the VESPER control file. This will be appended to the output folder.

	Auto create control file name	If checked, the name will be generated from the first 20 characters of the input CSV file, the string *'control'* and the Krige column. Uncheck this option to prevent manual edits from being overwritten when changing source CSV file or Krige column.	
	Configure to show VESPER graph and map graphics while kriging	Default is unchecked	Checking this option will display the map and graph while VESPER is running but will increase the processing time.
Run Vesper Kriging Now Convert Vesper Files to Raster and Load in QGIS OK Cancel	**Run VESPER kriging now**	Default is checked	Add the control file to the VESPER Queue for processing.
	Convert VESPER Files to Raster and Load in QGIS	Default is checked	On completion of VESPER processing, convert the outputs to raster TIFF formats.

File Naming Conventions

- <> denotes an existing element or input
- non-alphanumeric characters are removed with the exception of hyphens (-) and underscores (_).

DESCRIPTION	FILENAME	EXTENSION	EXAMPLE
*Files **required** by VESPER and copied to the VESPER output folder*			
VESPER control file	*<first 20 characters CSV file>_< first 10 characters of krige column>*_control	.txt	swblock_YldMassDry_control.txt
VESPER grid file	*<control file name>* where control is replaced by vespergrid	.txt	swblock_YldMassDry_vespergrid.txt
VESPER data file	*<control file name>* where control is replaced by vesperdata	.csv	swblock_YldMassDry_vesperdata.csv
*Files **created** by VESPER and saved to the VESPER output folder*			
VESPER Krige file (contains Prediction and SE)	*<control file name>* where control is replaced by kriged	.txt	swblock_YldMassDry_kriged.txt
VESPER parameter file	*<control file name>* where control is replaced by parameter	.txt	swblock_YldMassDry_parameter.txt
VESPER report file	*<control file name>* where control is replaced by report	.txt	swblock_YldMassDry_report.txt
*If **converting VESPER files to raster** is checked or the **Create Rasters from VESPER Results** tool is used*			
Kriged Prediction TIFF	*<control file name>* where kriged is replaced by PRED	.tif	swblock_YldMassDry_PRED.tif

source: VESPER Krige result above				
Standard Error TIFF source: VESPER Krige result above	<*control file name*>	where kriged is replaced by SE	.tif	swblock_YldMassDry_SE.tif
Confidence Interval (CI) metadata text file	<*control file name*>	where control is replaced by CI	.txt	swblock_YldMassDry_CI.txt

2.4 Import VESPER Results

Summary

This tool converts the output files generated by VESPER into raster TIFFs.

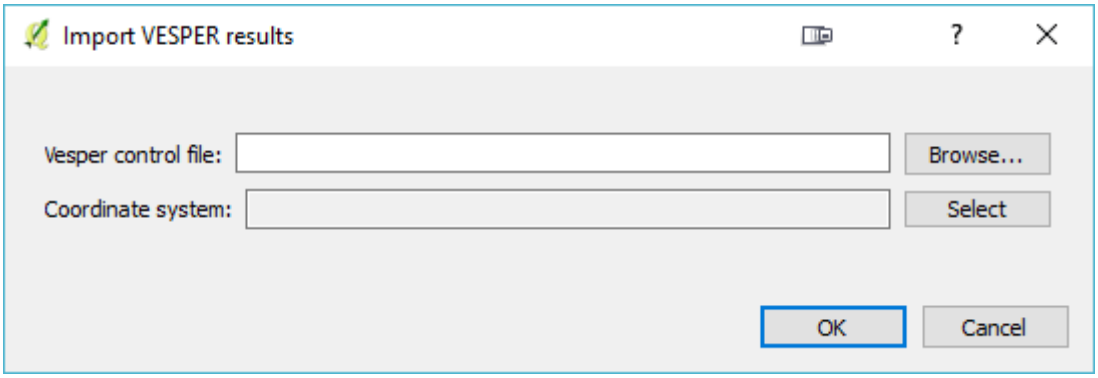
The output filenames will be created using the VESPER control file name as a base and demonstrated in the ***Prepare Data for Kriging using VESPER Tool - File Naming Conventions*** section on page 20.

The files created include:

- a TIFF file representing the predicted (kriged) value; these are the values used to generate the resulting map.
- a TIFF file representing the standard error (SE) of the kriging prediction.
- a text file containing the calculated median kriging prediction SE and the 95% confidence interval for the map. This is important information for testing the significance of the difference between different zones when different map layers are clustered to generate potential management zones.

The coordinate system to be assigned to the rasters will be extracted from the value stored within the control file. If the coordinate system cannot be found or is incorrect, it can be selected manually.

Dialog



Import VESPER Results		
VESPER Control File:		A VESPER control file used to identify VESPER results files and convert to rasters.
Coordinate System for CSV File:	By default it is extracted from the input control file.	This is the coordinate system of the VESPER results.

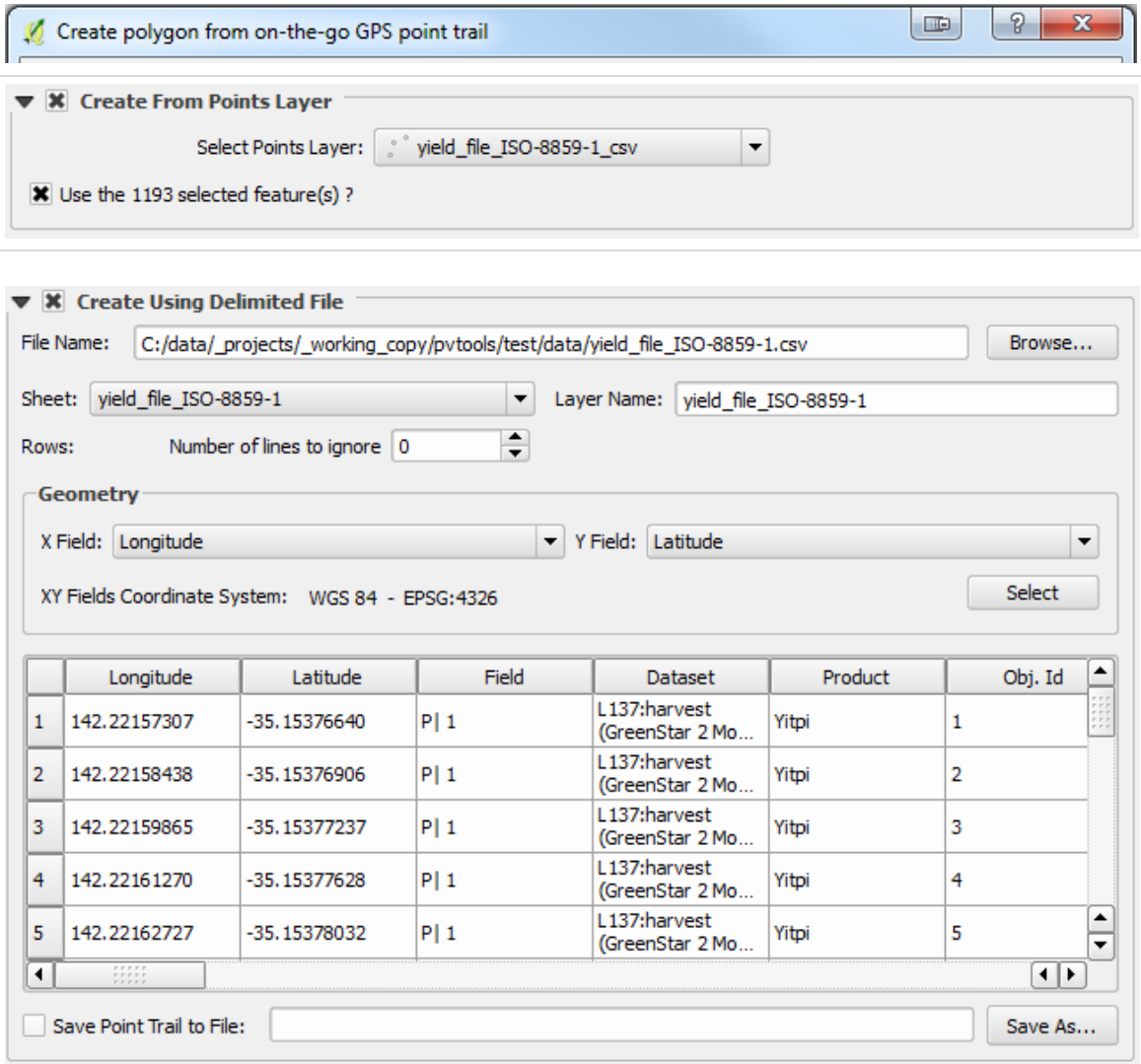
2.5 Create Polygon from On-The-Go GPS Point Trail Data

Summary

Having a block boundary polygon is central to many of the PAT processing steps. Boundary polygons are used to constrain data to a fixed extent. It is preferably that a boundary polygon is created by collecting accurate GPS points around the block and editing them in QGIS to create polygons (refer to ***How-To - Create a block boundary polygon from a CSV of GPS collected points*** on page 48 for instructions on this method). However, if accurate GPS data for the boundary is not available, then a less accurate block boundary polygon can be created using this tool, based upon a file of on-the-go GPS points (i.e. from a yield monitor or EM38 survey data). Note that this method is not preferred, but the tool is provided as a 'quick fix' for those wishing to generate maps in the absence of surveyed block or field boundaries.

As the process involves a dot-to-dot approach, it is critical that the input file of points are in order (i.e. sorted by an increasing time sequence - as would be normal for data collected by an on-the-go sensor such as yield monitor). For efficiency, points can be thinned by removing points closer than a set distance apart as justified by the accuracy of the GPS. Resulting points will be connected to form lines and then converted to polygons.

Dialog

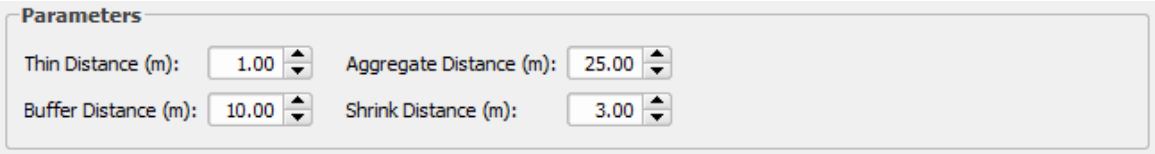


Choose whether to use GPS data currently loaded as a QGIS Points Layer

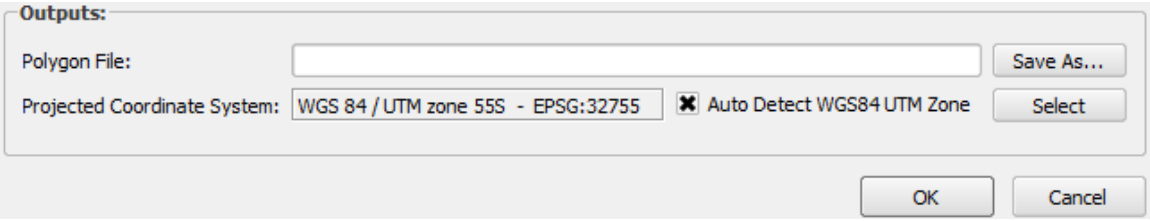
Use selected features		If checked, only uses the selected features to generate a polygon.

Or alternatively using a delimited text file

Filename	Options include XLS, TXT, CSV	The filename of the table file.
Sheet		The sheet containing data. If the file type doesn't support sheets, then the filename will be used.
Layer name	Default is the sheet name.	The name used when displaying QGIS layers, and as the default value for the output point and/or polygon files.
Number of lines to ignore	Default is 0	The number of lines at top of file to ignore. Can be used for skipping header text.
X column Y column	By default will attempt to match to X,Y, longitude, latitude; lon, lat etc.	The columns representing the X and Y coordinates within the input file.
XY columns coordinate system		The coordinate system of the XY Columns. Note: GPS normally records data in WGS 1984 (EPSG: 4326)
Save point trail to file	Default is unchecked	If checked, save the input delimited file as a points Shapefile in the output coordinate system.



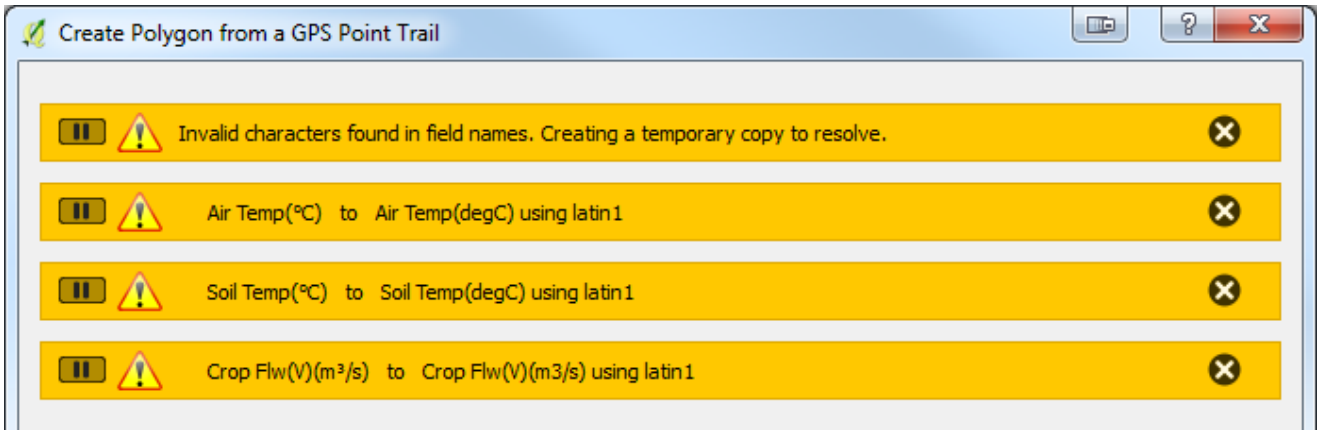
Thin distance (m)	0.00 to 100.00 Default is 1 Use 0 to omit	Any points spaced at less than the minimum distance apart will be removed.
Aggregate distance (m)	0.00 to 100.00 Default is 25	The maximum distance between points used to separate strings of connected points into individual lines. Typically this is slightly larger than the row/swath width.
Buffer distance (m)	0.00 to 100.00 Default is 10	A distance in meters used to buffer the lines to create overlapping polygons. Typically this is half the row/swath width.
Shrink distance (m)	0.00 to 100.00 Default is 3	A distance in meters used to shrink the overlapping polygons after dissolving. Typically this is slightly less than the buffer distance.



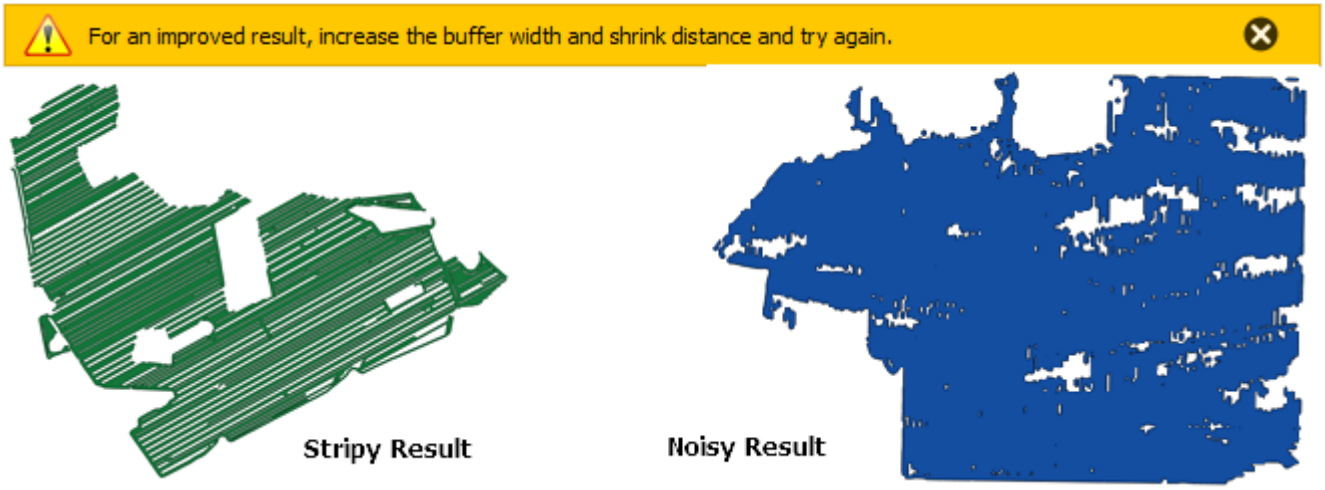
Polygon file	Derived from layer name	The output filename for the polygon Shapefile.
Projected coordinate system		The output projected coordinate system of the Shapefile.
Auto detect WGS84 UTM zone	If Auto detect is checked the following rules will apply: - Will try to match the input if it is a projected coordinate system - If the inputs are in Geographic's (lat/long) it will calculate the relevant WGS84 UTM Zone Note: If in Australia then MGA GDA 1994 zones, will be used – see Technical Notes on page 47.	

Notes

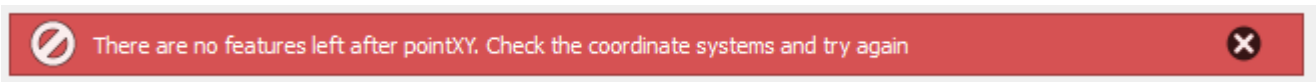
- The following notification occurs when a delimited text file contains invalid characters in the column name. A copy of the file will be made in the PAT temporary folder containing corrected column names as specified in the PAT log message and to file. For more information on the location of the saved log file and temporary folder see the Technical Notes section.



- A warning is triggered when the resulting polygon is considered stripy or noisy. This can usually be corrected by increasing the buffer and shrink distances.



- The following error is caused when the wrong coordinate system is applied to the input source file or layer. Thinning and/or clipping will result all points being removed leaving no points for further processing



2.6 Rescale or Normalise a Raster

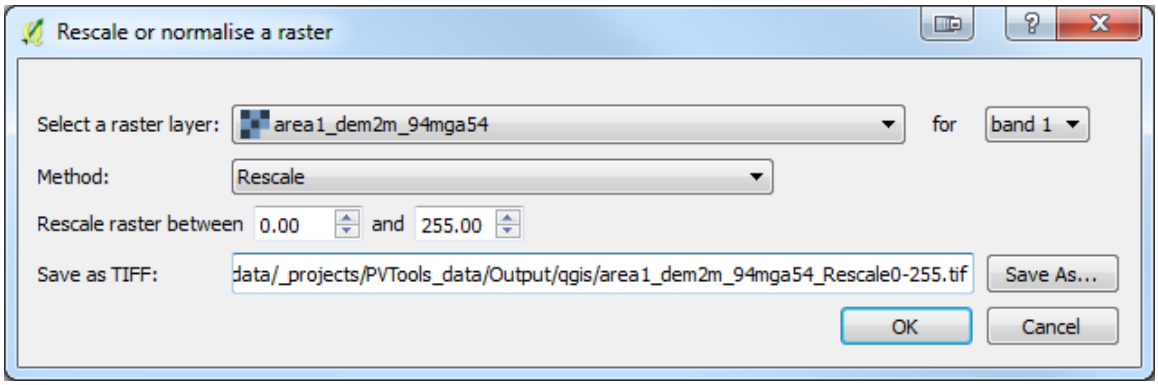
Summary

This tool rescales or normalises a raster and outputs to a new TIFF file. Existing no-data values will be ignored in any calculation.

- Rescale will adjust the raster between the specified values (e.g. output values all in the range 0 to 1)
- Normalise will adjust the raster to a mean of zero and standard deviation of one

Such re-scaling is useful when you wish to compare several map layers, for example, yield maps for several years.

Dialog



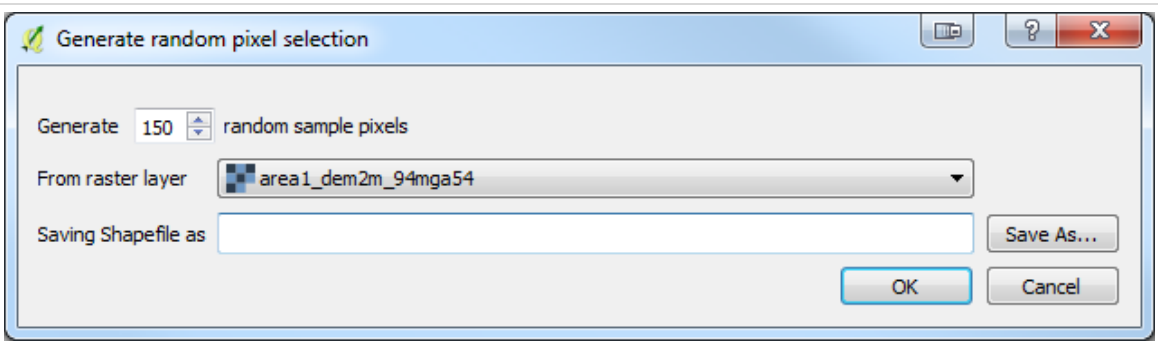
Rescale or normalise a raster		
Select a raster layer for band X	Default is band 1	The raster layer and band to rescale or normalise.
Method	Default is rescale	Options: • **Rescale** – adjust values to a fixed range • **Normalise** – adjust values to a mean of zero and a standard deviation of one
Rescale between (when selected)	Default 0 to 255	The range of values used with rescaling.
Save as Shapefile	Default will be derived from the input layer name	The output Shapefile to be created.

2.7 Generate Random Pixel Selection

Summary

This tool is used to select randomly distributed pixel locations from an existing raster and save them to a point Shapefile. Pixels will be selected from areas inside the raster extent which contain valid data. The resulting points will be located in the centre of the chosen pixel and the output shapefile will contain columns representing the X and Y coordinates. This tool provides the key location input for the ***Extract Pixel Statistics for Points*** tool described on the following page.

Dialog



Generate random pixel selection		
Generate random sample pixels	Default is band 1	The number of pixels to randomly select
Select a raster layer		The raster to use as the base for selecting pixels. No-data pixels will be exempt from selection.
Save as Shapefile	Default will be derived from the input layer name	The output Shapefile to be created.

2.8 Extract Pixel Statistics for Points

Summary

Extract pixel statistics for points is used to extract pixel statistics from multiple rasters using a pre-defined set of points. These points could be identified using the ***Generate Random Pixel Selection*** tool described on the previous page. Statistics are calculated on pixel values within a square neighbourhood and extracted to a CSV file. This tool is useful when relationships are being sought between different maps layers – for example, a regression or correlation analysis between yield and soil map data.

Applying a neighbourhood filter to rasters is useful for removing (smoothing) small anomalies introduced from instrument inaccuracies or on-the-go movement. The neighbourhood consists of a centre pixel and a number of pixels forming a square around the central pixel, as shown in Figure 3.

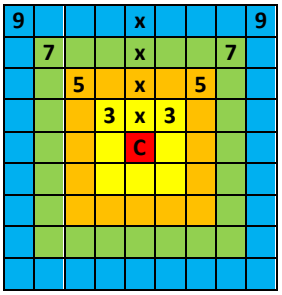


Figure 3. A representation of the 3x3, 5x5, 7x7, 9x9 neighbourhood size around a central pixel (C) as is used when calculating filtered statistics.

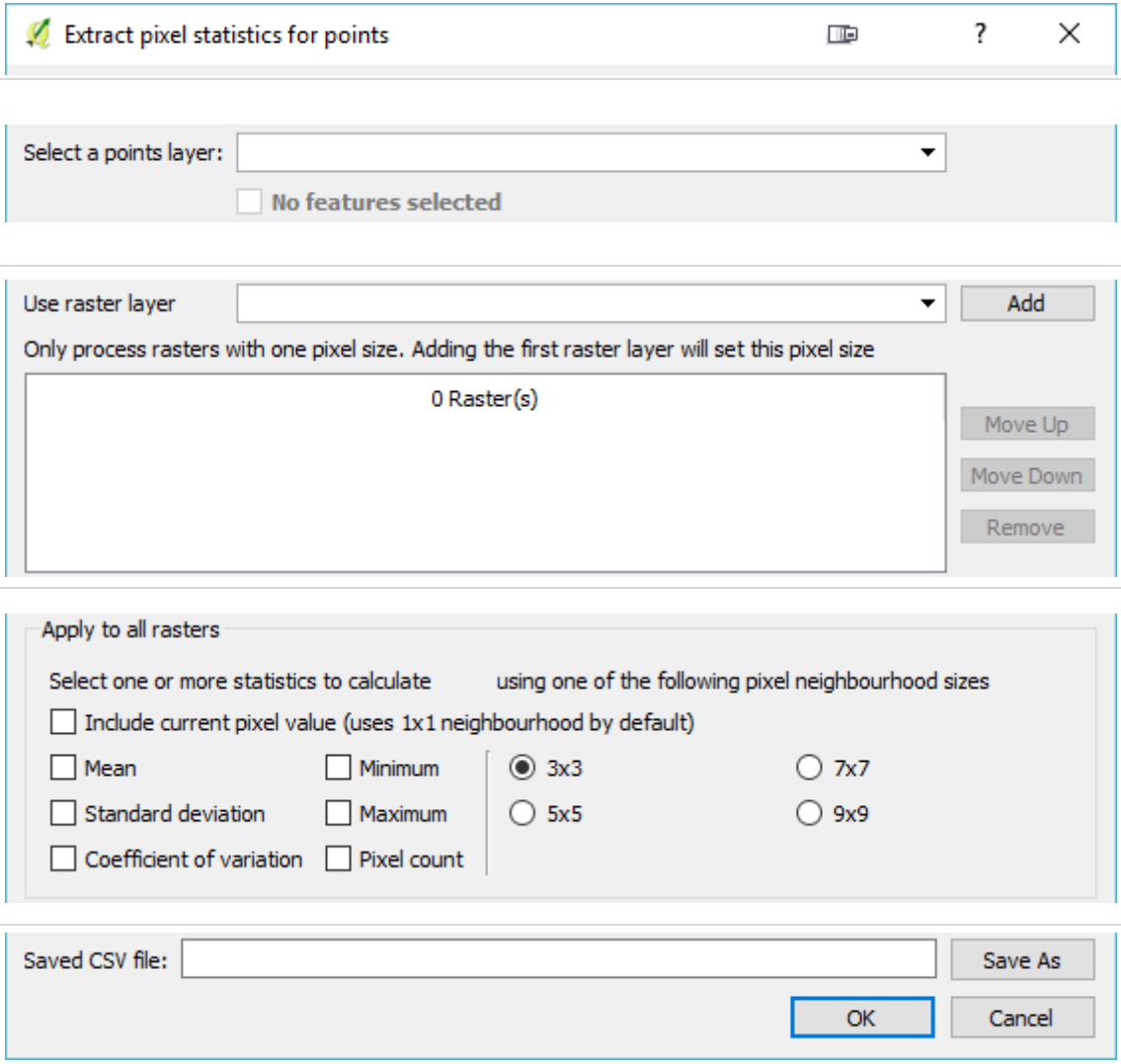
For example, a mean statistic for a 3x3 neighbourhood on 2 m pixels, will calculate the mean of the central pixel (red) and the surrounding 3x3 area of 8 pixels (yellow) equating to 36 m^2 on ground.

Pixels designated no-data will be excluded from the statistical calculation. However, a central no-data pixel may be assigned a value if at least one pixel in the neighbourhood has a valid value.

Currently, the statistical methods supported by this tool are: mean, standard deviation, co-efficient of variation (CV), minimum, maximum and a count of pixels contributing to the statistical calculation.

The output values are saved to a CSV data file. Column names for each raster and statistic combination are explained below.

Dialog



Select point layer		A point layer in QGIS used to extract pixel statistics values.
Use selected features	Default is unchecked	If **checked**, extract pixel values only for the selected points. If **unchecked**, all the points will be used
Use raster layer		The raster layer to extract values from. Only raster layers of the same pixel size can be analysed at a time and this is set by the first raster layer added. The *Add, Delete, Move Up, Move Down,* buttons can be used to manage the list of rasters. The order of rasters is retained in the output CSV file with separate columns for each raster and statistic combination.
Calculate pixel neighbourhood statistics		Select the required statistical methods for each raster in the list. **Current pixel** just extracts the pixel value and/or **Statistics** to be calculated using the set neighbourhood size. The output column naming format is explained below.
Neighbourhood size (pixels)	Default is 3x3	The size of the neighbourhood footprint to use while calculating statistics.
Save CSV file	Derived from layer name	The name and path of the output CSV file containing the final data points.

File and Column Naming Conventions

- <> denotes an existing element or input
- non-alphanumeric characters are removed with the exception of hyphens (-) and underscores (_).

Output Filenames:

CSV filename	*<QGIS points layer name>*_pixelvals.csv

Column Names:

NAMING RULE	EXAMPLE	EXPLANATION
<statistic><size>_<raster file>	mean3x3_swblock_YldMassDry_PRED	The column containing values from the mean 3x3 neighbourhood filter for the raster swblock_YldMassDry_PRED.tif.
	pixel_swblock_2009YldMassDry_PRED	The column containing the pixel values for the raster swblock_2009YldMassDry_PRED
	cv5x5_seblock_2014YldMassDry_PRED	The column containing values for coefficient of variation (CV) 5x5 neighbourhood filter for the raster seblock_2014YldMassDry_PRED.tif
	pixelcount7x7_swblock_2024YldMassDry_PRED	Column containing values representing the count of pixels used for statistical calculations with a 7x7 neighbourhood filter for the raster seblock_2014YldMassDry_PRED.tif

2.9 Calculate Image Indices for Blocks

Summary

Many providers of remotely sensed imagery deliver data to clients as multi band files. The Calculate Image Indices for Blocks tool is used to calculate indices for a multi band image and aligns the output to the extent and pixel size of a block grid. Indices currently supported are Normalised Difference Vegetation Index (NDVI); the 'simple ratio' or Plant Cell Density index (PCD); Green Normalised Difference Vegetation Index (GNDVI); Chlorophyll Red-Edge index (CHLRE); and Normalised Difference Red-Edge index (NDRE).

Table 3 describes how these indices are calculated.

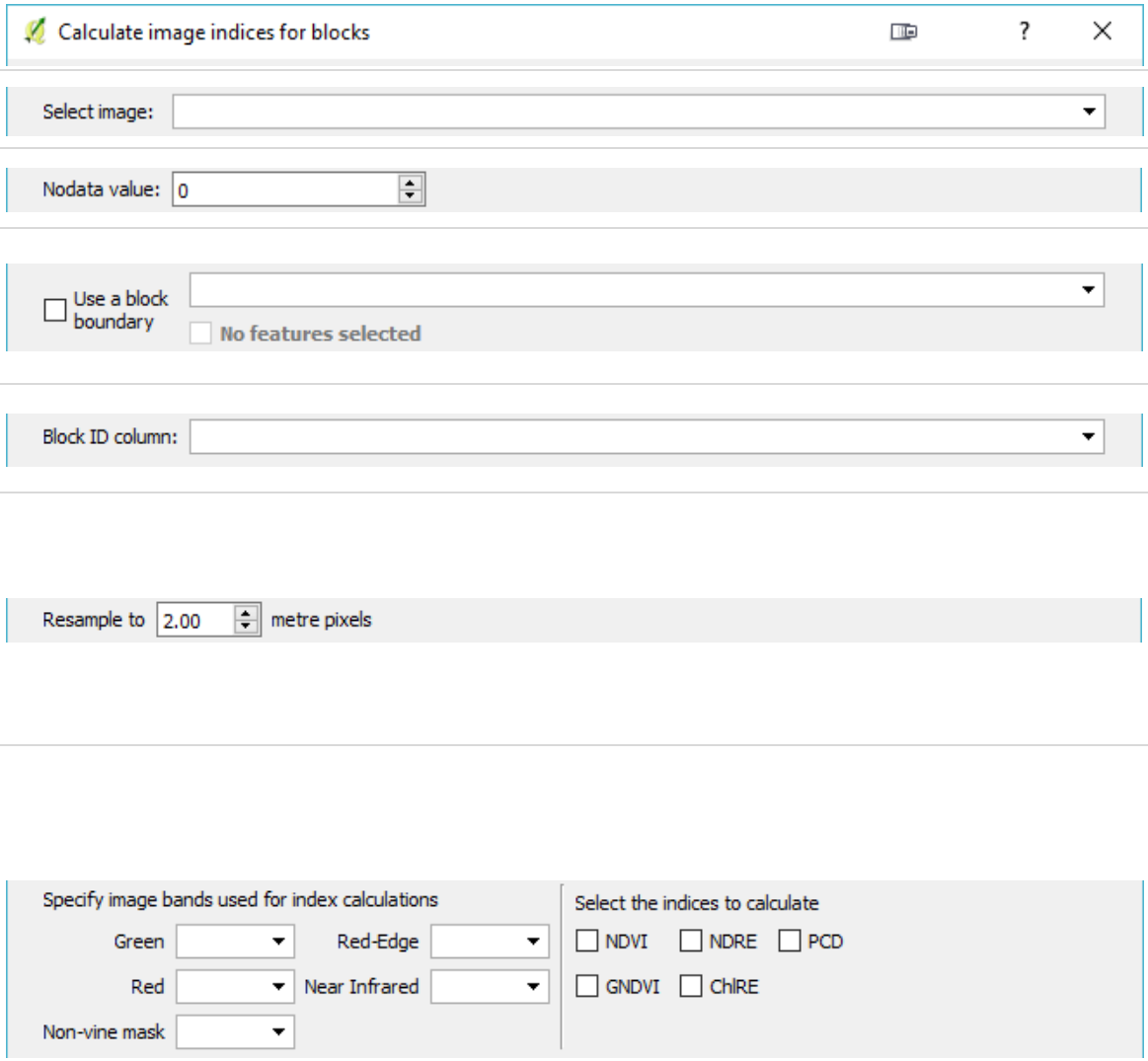
Each band in the image file is mapped to a spectral band e.g. Red, Red-edge, Near Infrared and is then used to calculate relevant image indices. In the case of vineyard data, if a non-vine mask is present in an existing band of the image, it can be used to remove non-vine signals prior to the resampling and alignment to the block grid. This will ensure that spectral signatures relating to ground cover are excluded from the resulting image outputs. This non-vine masking is not relevant for broadacre crops.

An optional block boundary polygon layer, and a column containing the block name or ID, can be used to separate the resulting images into individual blocks. By default, if no column is specified, then all polygons are assumed to be from the one block and will be processed accordingly. If no block boundary layer is specified, then a single polygon outlining the image (excluding no-data) will be used. The output from this tool will be aligned to a block grid that will be identical to the one created separately using the ***Create Block Grid*** on page 13.

Table 3 the list of supported indices

Normalized Difference Vegetation Index (NDVI)	$\text{NDVI} = \frac{(NIR - Red)}{(NIR + Red)}$
Green Normalized Difference Vegetation Index (GNDVI)	$\text{GNDVI} = \frac{(NIR - Green)}{(NIR + Green)}$
Normalised Difference Red-Edge Index (NDRE)	$\text{NDRE} = \frac{(NIR - Red\ Edge)}{(NIR + Red\ Edge)}$
Plant Cell Density Index (PCD)	$\text{PCD} = \frac{NIR}{Red}$
Chlorophyll Red Edge Index (ChlRE) (Gitelson 2004; Gitelson *et al.* 2005)	$\text{Chl}_{red\ edge} = \left(\frac{NIR}{Red\ Edge}\right) - 1$

Dialog



Select Image	The image layer containing the appropriate bands required for calculate indices.	
Nodata value	Default from image.	Can be used to specify a different no-data value.
Use a block boundary	Default is unchecked	A layer in QGIS containing polygon(s) representing blocks.
Use selected features	Default is unchecked	If **checked**, only the selected polygons will be used.
Block ID column	A column containing the block id or name. This will be used treat multiple polygons with the same block id or name as one.	
Resample pixel size (m)	0.00m to 6 km Default is 2 m	The pixel size to apply to the raster outputs. This is expressed in meters. Recommended Values: Viticulture: 2 m Sugarcane: 2 m Broadacre grains: < 200 ha 5 m > 200 ha 10 m
Specify image bands used for index calculations	Mapping of band numbers to band types. The mapped bands will enable/disable indices based on individual index requirements. For most images Green is band 2, Red is band 3, Near Infrared is band 4 Non-vine mask is a band where pixels <u>not</u> containing vine signals are set to no-data. Not required for broadacre data.	
Select the indices to calculate	The indices to calculate. If an index is disabled (greyed out) the bands it requires haven't been mapped. For index acronym and equation, see table above.	

Projected coordinate system: Unspecified Select	Projected coordinate system	The projected coordinate system to apply to the output files. By default it will calculate the relevant coordinate system from the input image or block boundary's coordinate system. Note: If in Australia then MGA GDA 1994 zones, will be used – see Figure 5 on page 47.	
Output folder: Browse	Output folder	The folder to save output TIFF files. A folder based on the input image name will be created and all TIFF files will be saved here. See table below for *File Naming Conventions* used for output files.	
Display results OK Cancel	Display results	**Default is unchecked**	If checked all resulting TIFF files will be loaded into QGIS.

File Naming Conventions

- <> denotes an existing element or input
- non-alphanumeric characters are removed from strings with the exception of hyphens (-) and underscores (_).

Filenames:

NAMING RULE	EXAMPLE	EXPLANATION
Output Folder:		
<***output_folder***>\<***image_name***>	C:\data\vineyard\rgbi_jan_50cm_84sutm54_tif	A new folder based on the image name is created in the output folder and all created images are saved here. In this example the image **rgbi_jan_50cm_84sutm54.tif** is used to create a new folder called **rgbi_jan_50cm_84sutm54_tif**.
Image Names:		
<***block_id***>_<***index***>_<***pixel_size***>.*tif*	B1_NDVI_2m.tif	The TIFF file resampled to **2m** pixels for the Normalised Difference Vegetation Index (**NDVI**) created for block id/name of **B1**.
	PCD_250cm.tif	The TIFF file resampled **250cm** pixels for the Plant Cell Density (**PCD**) Index created **without** specifying a block id column.

2.10 Resample Image Band for Blocks

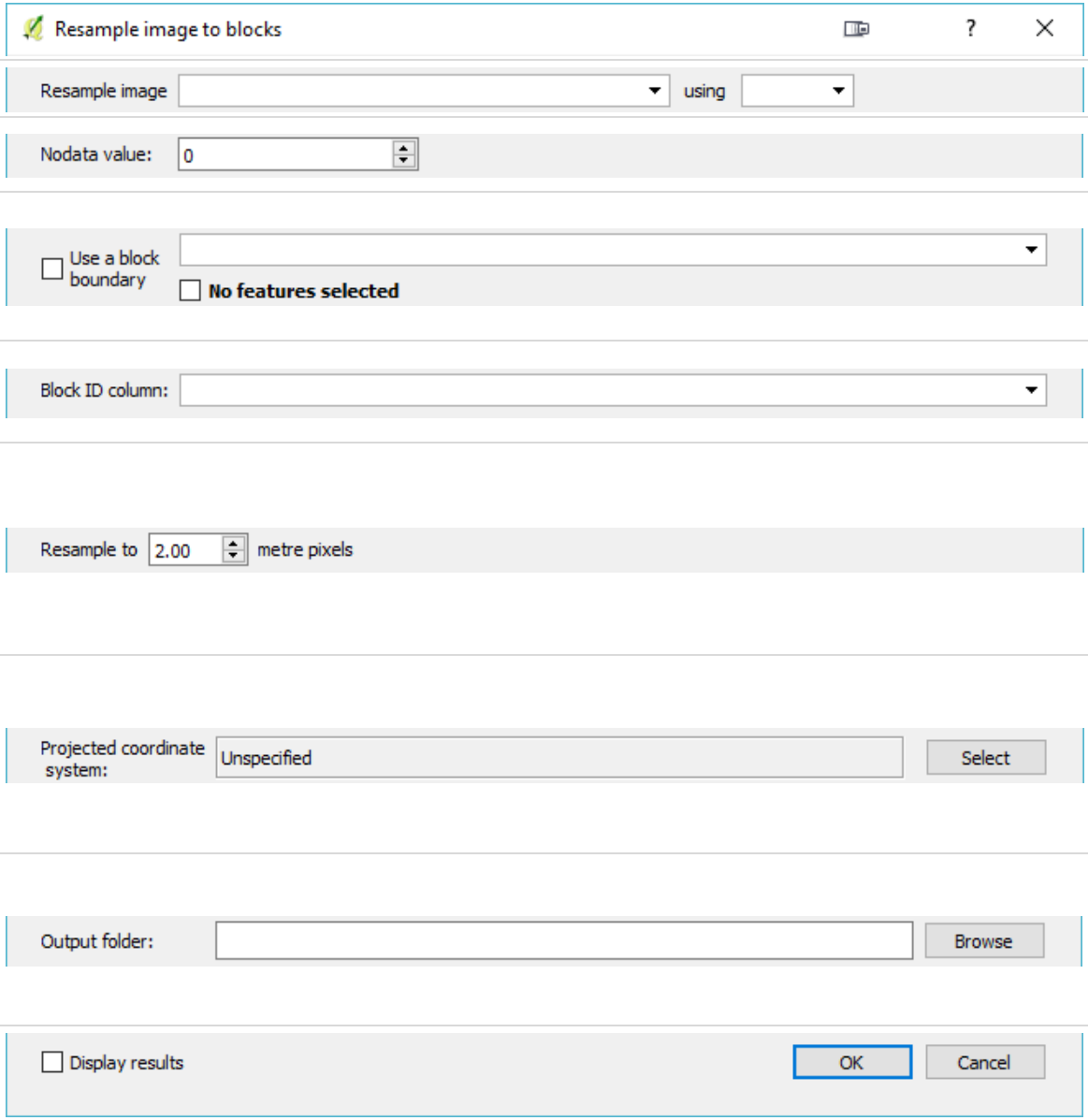
Summary

Resample Image Band for Blocks is used to resample, align and smooth an existing band of an image to match a block grid.

An optional block boundary polygon layer and a column containing the block name or ID, can be used to separate the resulting images into individual blocks. By default, if no column is specified, then all polygons are assumed to be from the one block and will be processed accordingly. If no block boundary layer is specified, then a single polygon outlining the input image will be used.

The output from this tool will be aligned to a block grid that will be identical to the one created separately using the ***Create Block Grid*** on page 13.

Dialog



Select Image	The image layer containing the band to resample.	
No-data value	Default from image.	Can be used to specify a different no data value.
Use a block boundary	Default is unchecked	A layer in QGIS containing polygon(s) representing blocks.
Use selected features	Default is unchecked	If **checked**, only the selected polygons will be used.
Block ID column	A column containing the block id or name. This will be used treat multiple polygons with the same block id or name as one.	
Resample pixel size (m)	0.00m to 6 km Default is 2 m	The pixel size to apply to the raster outputs. This is expressed in meters. Recommended Values: Viticulture: 2 m Sugarcane: 2 m Broadacre grains: 5 m for fields < 200 ha; 10 m for larger fields.
Projected coordinate system	The projected coordinate system to apply to the output files. By default it will calculate the relevant coordinate system from the input image or block boundary's coordinate system and extent coordinate system. Note: If in Australia then MGA GDA 1994 zones, will be used – see Figure 5 on page 47.	
Output folder	The folder to save output TIFF files. A folder based on the input image name will be created and all TIFF files will be saved here. See table below for *File Naming Conventions* used for output files.	
Display results	Default is unchecked	If checked all resulting TIFF files will be loaded into QGIS.

File Naming Conventions

- <> denotes an existing element or input
- non-alphanumeric characters are removed from strings with the exception of hyphens (-) and underscores (_).

Filenames:

NAMING RULE	EXAMPLE	EXPLANATION
Output Folder:		
<***output_folder***>\<***image_name***>	C:\data\vineyard\area1_rgbi_jan_50cm_84sutm54_tif	A new folder based on the image name is created in the output folder and all created images are saved here. In this example the image **area1_rgbi_jan_50cm_84sutm54.tif** is used to create a new folder called **area1_rgbi_jan_50cm_84sutm54_tif**.
Image Names:		
<***block_id***>_<***Band***>_<***pixel_size***>.*tif*	B1_**Band6**_2m.tif	The TIFF file where **Band 6** is resampled to **2m** for block id/name of **B1**.
	Band7_250cm.tif	The TIFF file where **Band 7** is resampled **250cm** pixels **without** specifying a block id column.

2.11 Create Zones with *k*-means Clusters

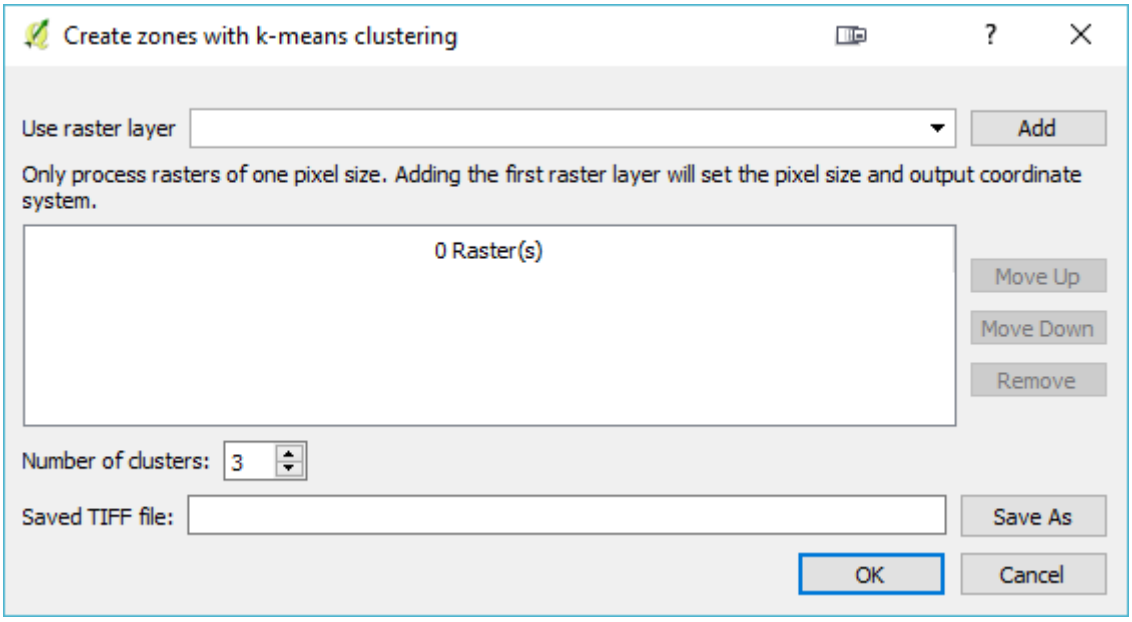
Summary

This tool allows zones to be created by performing *k*-means clustering on multiple rasters (i.e. map layers) to create clusters of similarity by minimising variability within clusters while maximising variability between clusters. If significant differences between clusters are observed then the clustered results can be used as potential management zones.

Raster files with the same single pixel size, and coordinate systems are used as inputs and the common area of overlap will be used to generate an output TIFF containing the clustered result.

On completion of *k*-means clustering, the mean and standard deviation for each zone/cluster and source raster combination will be calculated and written to a CSV File alongside the output TIFF file as well as being displayed in PAT's log messages panel.

Dialog



Create zones with *k*-means clustering		
Use raster layer:	The raster layer to extract values from. Only raster layers of the same pixel size can be analysed at a time. The first raster layer added sets both the pixel size and the output coordinate system. The *Add, Delete, Move Up, Move Down,* buttons can be used to manage and order the list of rasters. The order of rasters is retained in the output CSV file with separate columns for each raster and statistic combination.	
Number of clusters:	Default is 3	The number of clusters/zones to create. Aside from trying 3, it is recommended that users also try 2 clusters. 4, 5 or an even larger number of clusters may also be appropriate.
Save TIFF file	The name and path of the output TIFF file representing the zones. In addition, a statistics CSV file will be written to disk along with the TIFF, and results printed to PAT's log messages panel.	

File and Column Naming Conventions

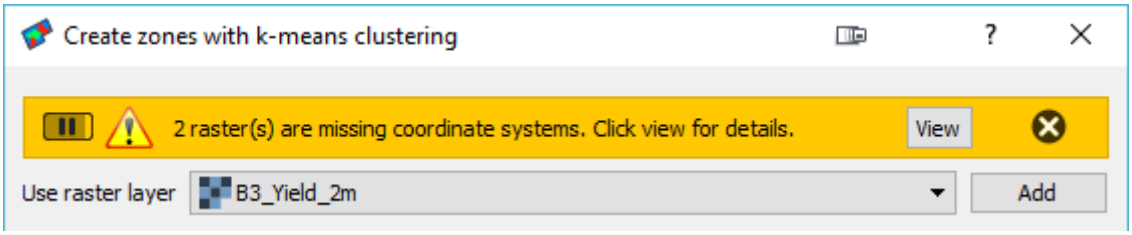
- <> denotes an existing element or input
- non-alphanumeric characters are removed from strings with the exception of hyphens (-) and underscores (_).

FILENAMES:		
Output TIFF		
	Naming Rule:	k-means_<*n_clusters*>*clusters*_<*n_rasters*>*rasters*_<*pixel_size*>*.tif*
	Example:	k-means_3clusters_5rasters_2m.tif
	Description:	A 2m pixel TIFF file for 3 k-means clusters using 5 input rasters
Output Statistics CSV	**<** output TIFF name**>** where **.tif** is replaced by **_statistics.csv**	

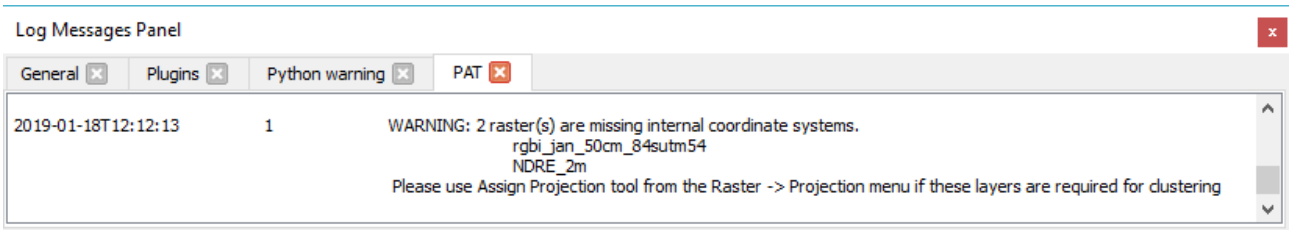
CSV COLUMN NAMES:	
Naming Rule:	**<*raster_file*> _<*statistic*>**
Example:	swblock_2024YldMassDry_PRED_**std**
Description:	The **standard deviation** value for swblock_2024YldMassDry_PRED.tif for the corresponding cluster/zone.

Notes

- To successfully use this tool, all input raster files must contain the coordinate system internal to the file. When a file without a coordinate system is loaded into QGIS, an external coordinate system will be applied based on your QGIS settings. These files cannot be used by this tool. The user will be notified of this when launching the tool as follows.



Clicking on ***view*** will open the PAT Log Messages Panel and provide a list of those images. Users can then use the *Assign Projection* Tool from the *Raster -> Projection* menu to assign an internal coordinate system to the files. If the *Assign Projection* tool is unavailaalbe activate the GDALTools plugin .



2.12 Create Strip Trial Points

Summary

This tool is used to generate the data needed to implement the moving window t-test method of Lawes and Bramley (2012) for analysis of simple strip trials. It may be useful to users who, for example, have implemented an N-rich strip in their cereal paddock, or who want to compare a row of vines that have been pruned differently to the rest of a vineyard. The tool uses a line representing the centre of a strip trial and creates points along and at an offset to the line; the offset points are those for comparison with the strip points. In this current version of PAT, use of this tool needs to be followed by using the strip and offset points generated as input to the ***Extract Pixel Statistics for Points tool*** to extract values from map raster(s) relevant to the strip experiment (e.g. a yield map) which can then be analysed outside of PAT. In a future version, the entire analysis will be provided in PAT.

Points are placed along the line using the nominated ***distance between points***. In addition, two sets of points are placed at the ***offset distance*** either side of the line. This allows the strip data, to be compared to data from the neighbouring (control) area.

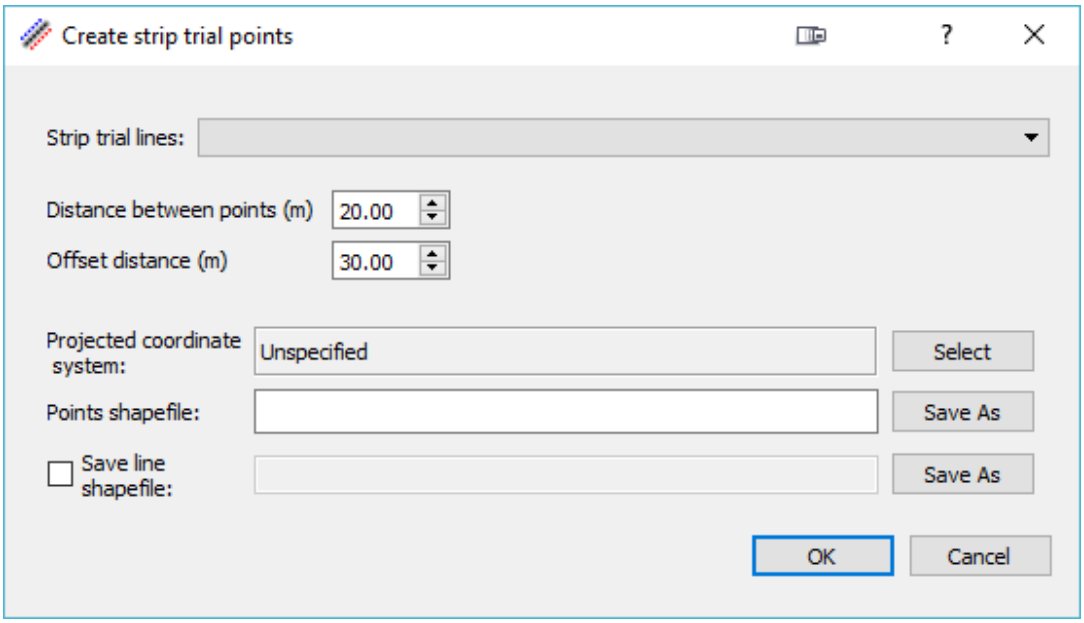
Each point is assigned a Side, PointID and line_dist values which are saved in the shapefile. The side attribute represents one of the following:

C – centre line;
R – Right offset to centre line
L – Left offset to centre Line.

Left and right values are assigned based on the original drawing direction of the line not how they appear on screen. The same PointID assigned to the centre line is also applied to the relevant L/R point.

The line_dist represents the distance in metres of the point from the start of respective line. The distances for the L/R sides are calculated from the start of the offset line, not the centre line.

Dialog



Create strip trial points		
Strip trial lines:	The layer containing the strip trial lines.	
Distance between points (m):	1 – 500.00 Default is 20.0	The distance used to space points along the line. In this instance, users are advised to select a value appropriate to their trial rather than accepting the default.
Offset distance (m)	1 – 500.00 Default is 30.0	The distance used to offset the points from the line.
Projected coordinate system:	The projected coordinate system to apply to the output files. By default it will calculate the relevant coordinate system from input layer's coordinate system. Note: If in Australia then MGA GDA 1994 zones, will be used – see Technical Notes on page 47.	
Points shapefile:	Derived from layer name	The name and path of the output points shapefile.
Save line shapefile:	Default is unchecked. Derived from layer name.	The name and path of the output points shapefile.

2.13 Whole-of-block Analysis

Summary

Whole-of-block experimentation, also known as spatially distributed experimentation, involves on-farm experiments where treatments are applied in a highly replicated design over an entire block or land management unit, or a large section of it. Given their scale, they are typically laid down using the farmer's normal equipment. These experiments can be thought of as a three-dimensional version of the strip trial, discussed in the previous section ***Create Strip Trial Points***. Users interested in this experimental approach may wish to read the book chapter by Bramley et al. (2013) and perhaps the references listed in that chapter. In brief, the objective of whole-of-block experimentation is to recognise the possible effect of underlying spatial variability in the field in which the experiment is located, on the crop response to experimental treatments, and so use this variability as an experimental tool to guide decisions about variable rate or targeted management. Whole-of-block experimentation therefore moves away from the classical approach to agronomic experimentation which seeks to identify whether treatment A is better than treatment B and typically uses small plots. Instead, it is recognised that treatments A and B may both deliver benefit, albeit in different parts of the block or field; the whole-of-block approach identifies these different areas.

An example is given in Figure 4. Here, some different vineyard floor management treatments have been evaluated in an experiment aimed at improving the availability of water to the vines. The cereal/legume treatment delivers a significant benefit over the other treatments, but this is expensive and so the vineyard manager would only want to use this in those areas where it is beneficial. The whole-of-block analysis tool enables users to generate maps such as those shown in Figure 4 – maps showing the response to the different treatments over the entire block, and maps indicating the significance of differences between these treatments.

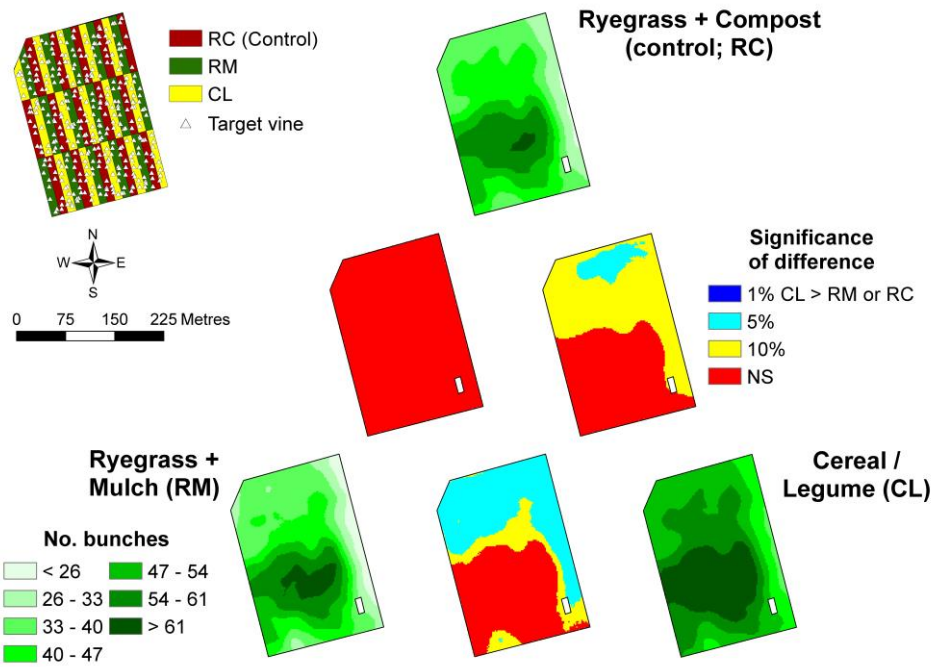


Figure 4 An example of a whole-of-block experiment conducted in a 4.8 ha Clare Valley vineyard. Panten and Bramley (2011) provide further information

The purpose of the ***Whole-of-Block analysis tool*** is to predict treatment responses and their differences over the entire experimentation area as illustrated by Figure 4. Similar to the use of VESPER for kriging, this tool interpolates point data (low or high density) to the area of the entire block. Low density data would typically include up to a few hundred points, perhaps collected through hand sampling, dispersed across a block, whereas high density data includes a few thousand or more points, typically collected using a sensor such as a yield monitor. The tool will handle either two or three treatments, and will identify areas of the block where the treatments are significantly different from each other. To address the technical and computational challenges in this analysis, the tool uses fast and efficient global and local cokriging techniques.

Use of R

This tool is somewhat different from the other PAT tools in that it uses the R statistical language (http://r-project.org/) to carry out the analytical processing. This means that R must be installed on your PC for this to work – see Configuring QGIS to use R. The minimum R version suitable for this is v3.5.1.

Algorithm for the Local Cokriging System – a brief commentary on the data radius used

A spatially-varying local co-kriging method is used to balance the trade-off between the computation efficiency and prediction accuracy. We use a neighbourhood radius r(s) to identify neighbours for location s.

For subregions with lower density observations, such as at the corners, or edges of a farm field, the neighbourhood radius required by local cokriging will be relatively bigger to ensure the stability of the estimated cross-variograms and balance the number of observations from different treatments within the neighbourhood.

For high density datasets we will use a relatively small neighbourhood radius, and ensure it will change relatively smoothly across the field. This adaption is achieved by two specific algorithm choices. Firstly, an initial neighbourhood radius is selected automatically based on the observation density and achieving a reasonable balance in the number of observations from different treatments. This neighbourhood radius increases iteratively to include more observations if the current radius is not large enough to meet prediction quality criteria. Secondly, a spatial-varying neighbourhood radius surface is estimated based on neighbourhood radius values from a subset of locations. This exploits the spatial-smoothness of the neighbourhood radius and provides a better default neighbourhood radius from which to optimise .

Dialog

Whole Of Block Analysis		
Input Points Layer	Input points layer	A text format data file (CSV format) containing the experimental data that has been loaded into QGIS. The file should be a minimum of four columns containing, the following:
Easting Northing	Easting and Northing	The names of columns in the CSV data file containing the projected Easting and Northing point coordinates
Treatment Column	Treatment column	The name of column in the CSV data file containing the treatment applied at each point. The treatments can be represented by numbers (1,2,3), short characters ("AM","BL","CS") or text 10 characters or less **without** spaces like ("Control","Mulch","Cover_crop").
Data Column	Data column	The name of column in the CSV data file containing the measured treatment response (e.g. Yield) at each point.
Input Block Grid	Input Block Grid or Prediction Grid	The grid onto which treatment predictions, differences and standard errors are to be interpolated by the tool. The block grid created by PAT Create Block Grid tool is ideal for this, or it can be a raster (TIF) file with the same coordinate system as the Input Points Layer.
Model Global	Model	Default is Global	There are two model options. • Global cokriging (default) is useful for small datasets (a few hundred points) • Local cokriging is useful for moderate or larger datasets (thousands of points), particularly with local heterogeneity. This option initiates a starting radius for localised cokriging and uses an iterative algorithm to optimise the local cross-covariance structure. The cross-variogram is fitted using a Linear Model of Coregionalisation (LMC). The current version of the software defines 20 lag classes, giving 20 points to the cross-variogram. The maximum distance used in this process (i.e. largest lag class) is set as 1/3 of the maximum distance between points in the input dataset.
Covariance Model Exponential	Covariance Model	Default is Exponential	Exponential (default); users may also select Spherical and Gaussian models

User Defined Neighbourhood for Local CoKriging Input Neighbourhood Size in Metre 30	User Defined Neighbourhood radius for Local Cokriging	Default is unchecked Default is 30m or 15 times the pixel size	Only applicable to the Local Cokriging method, and allows a user to define an initial neighbourhood radius. If this option is ticked then user should define the "Input Neighbourhood Size in Metres". The default is 30 m. A suggested starting value can be calculated as 15 times the pixel size of the block grid. However, if "User Defined Neighbourhood for Local Cokriging" is not ticked then the initial neighbourhood radius will be chosen as a function of total number of data points and corresponding number of treatments. This is the recommended starting point for using this tool.
Save Output [optional]	Save output	Folder in which to save the output files. See File Naming Conventions (below) for an explanation of the different output files that will be created in this folder.	

File Naming Conventions

- <> denotes an existing element or input. For example <Data Col> refers to the name of the Data Column selected in the input parameters (above).
- non-alphanumeric characters are removed from strings with the exception of hyphens (-) and underscores (_).

Filenames:

NAMING RULE	EXAMPLE	EXPLANATION
Output file names:		
<Data Col>_dump_file_<Model><Covariance Model>.tif	Bunch_wt_dump_file_GlobalExp_.log	Logged information mainly useful for debugging purposes
<Data Col>_model_parameters_<Model><Covariance Model>.txt	Bunch_wt_model_parameters_GlobalExp_.txt	Summary of parameters used during processing
<Data Col>_list_of_tif_files_<Model><Covariance Model>.txt	Bunch_wt_list_of_tif_files_GlobalExp_.txt	List of TIF files generated as output
<Data Col>_tr_<Tmt1>_<Model><Covariance Model>_.tif	Bunch_wt_tr_KC_GlobalExp_.tif	Treatment inference in "Input Block Grid or Prediction Grid".
<Data Col>_tr_<Tmt1>_var_<Model><Covariance Model>_.tif	Bunch_wt_tr_KC_GlobalExp_.tif	Corresponding variances of the treatments.
<Data Col>_ tr_diff_<Tmt1>_<Tmt2>_<Model><Covariance Model>_.tif	Bunch_wt_p_val_KC_SP_GlobalExp_.tif	Predictions of treatment response differences between treatments
<Data Col>_tr_<Tmt1>_var_<Model><Covariance Model>_.tif	Bunch_wt_tr_KC_var_GlobalExp_.tif	Corresponding covariances of the treatment differences.
<Data Col>_p_val_<Tmt1>_<Tmt2>_<Model><Covariance Model>_.tif	Bunch_wt_p_val_KC_SP_GlobalExp_.tif	p-values of the treatment differences between pairs of treatments.
<Data Col>_z_<Tmt1>_<Tmt2>_<Model><Covariance Model>_.tif	Bunch_wt_z_KC_SP_GlobalExp_.tif	Z-statistics of the treatment differences.

2.14 Settings

Summary

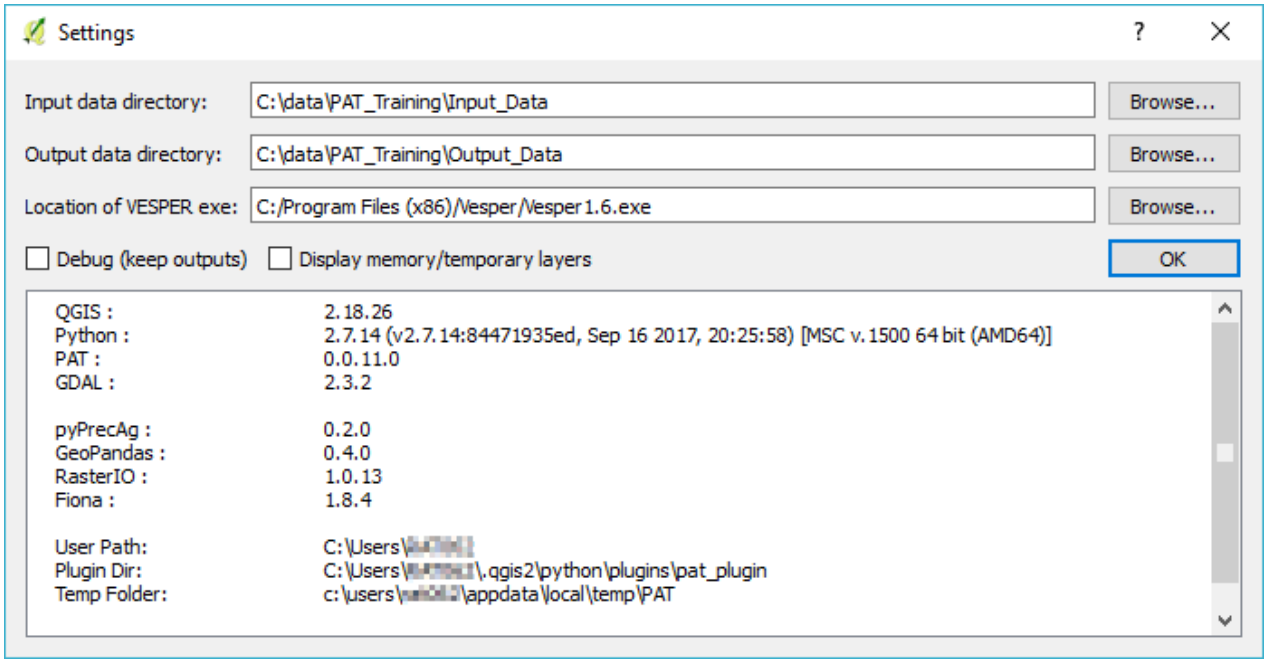
This tool is used to display and edit PAT settings.

The input and output data directories set here will be used to set the default paths by the *browse for file/folder* functionality for all tools. Each tool will store and access its own values after first time use.

Checking the ***Debug (keep outputs)*** box will save intermediate files created while processing data to file. It should be noted that this will slow down the time taken to run tools, but can be a useful diagnostic tool.

The ***Display Memory/Temporary Layers*** checkbox can be used to add the intermediate files to QGIS along with any in-memory or virtual layers which are used but not saved to disk.

Dialog



Input Data Directory	A folder containing input data.*
Output Data Directory	This is the location of the output file.
Location of VESPER Exe	Set the location of the VESPER executable. If this is not found or specified, VESPER cannot be run, but relevant VESPER input files can still be created.
Debug (Keep Outputs)	When checked, various intermediate files will be written to disk to assist with error/debugging analysis. However, this will slow down processing.
Display Memory/Temporary Layers	If checked, any files written to disk, or created in memory will be loaded into QGIS into a grouped layer labelled DEBUG and can be used for error/debugging analysis.

* By default this will be the PAT sub-folder in the user's home directory. To quickly navigate here type *%homepath%/PAT* in the address bar of Windows Explorer.

3 Technical Notes

- PAT makes use of the CSIRO-developed pyPrecAg Python module which is an open source Python package containing a range of specialised analysis functions.

- All intermediate files created while processing are located in the PrecisionAg folder of the user's temporary folder. To quickly navigate to the temporary folder, type %temp%/PrecisionAg in the address bar of windows explorer. This folder is deleted when QGIS exits.

- All progress, messages and errors are displayed in the PAT tab of the Log Panel as shown in Figure 2 and are saved to a log file located in the PrecisionAg folder of the user's temporary folder. A list of important paths including the location of temp and the user's plugin folder.

- A Users QGIS Plugin folder can be found by typing %homepath%/.qgis2/python/plugins into the address bar of Windows Explorer.

- When a projected coordinate system is predicted from the source data, and the data is identified as being in Australia, the Map Grid of Australia (MGA) reference system is used.
  The MGA zones align closely with the SUTM zones, but use the 1994 Geocentric Datum of Australia. Figure 5 shows the position of these the MGA zones.

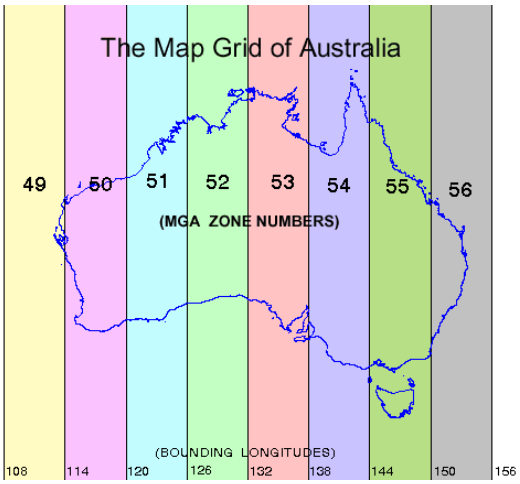
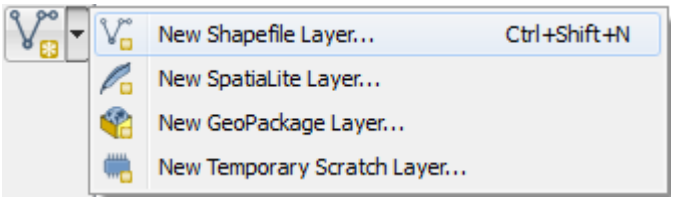


Figure 5 The Map Grid of Australia. (www.environment.gov.au/erin/tools/mga2geo-gda.html).

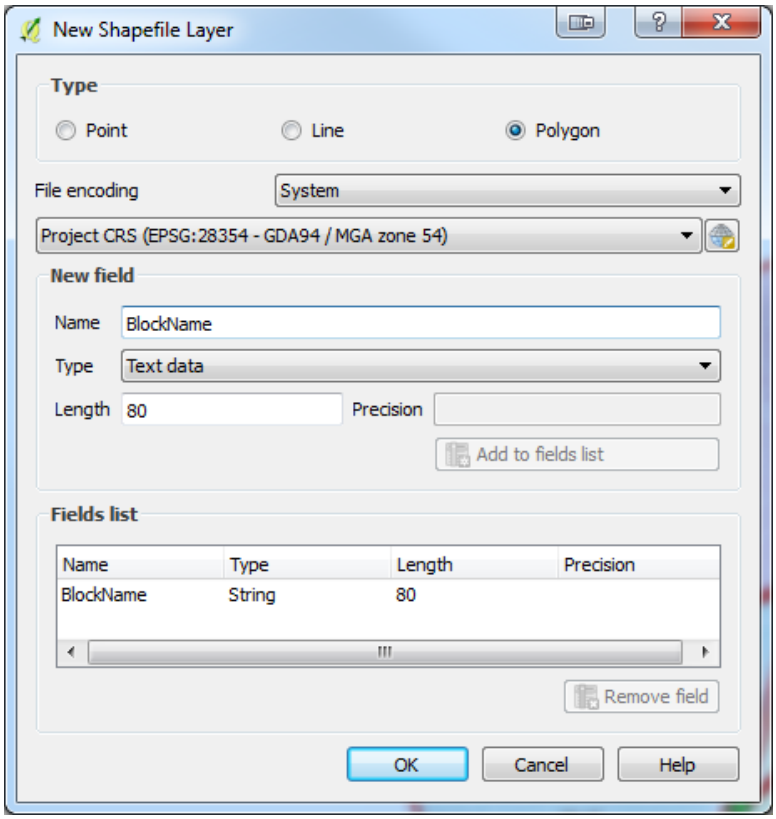
4 How-To's

4.1 Create a block boundary polygon from a CSV of GPS collected points



1. Using the ***New Shapefile Layer*** tool from the ***manage layers toolbar*** *(or Layer menu -> Create Layer -> New Shapefile Layer*) create a new polygon shapefile adding the relevant coordinate system* and attribute fields you require. Clicking OK will prompt you for the location to save the shapefile.

More information on coordinate systems can be found at



2. Set a style and labelling to the polygon layer. A hatching polygon fill works well for editing.

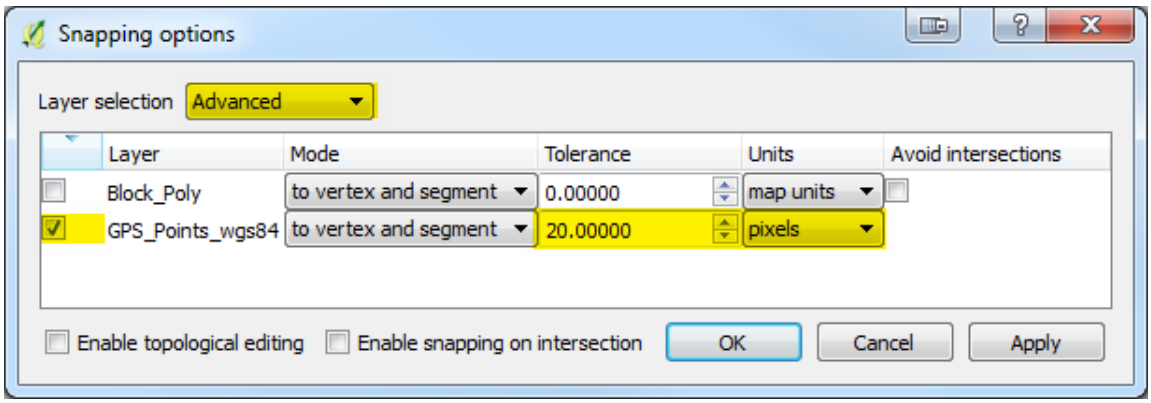
3. Launch the ***Add Delimited Text Layer*** tool from the ***manage layers toolbar*** *(or Layer menu-> Add Layer -> Add Delimited Text Layer)* and load your GPS CSV file as a layer into QGIS.

   Hint: Your coordinate system is probably WGS 84.

4. If required, load other vector or raster data, like imagery, which can be used as reference.

5. **Setup your snapping environment**.
   a. Open ***snapping*** options (*Settings menu -> Snapping Options*)
   b. Change ***layer selection*** to ***advanced***.
   c. **Tick** the layer containing the **point layer** loaded in step 3.
   d. Change the ***tolerance*** to **20** and set **units to pixels**

*More information on Map Grid of Australia can be found in Figure 5. For further information on coordinate systems: https://docs.qgis.org/2.18/en/docs/gentle_gis_introduction/coordinate_reference_systems.html



6. Select/activate the **polygon** layer in the layers panel and click the ***toggle editing*** icon from the ***digitizing toolbar***.

7. **Add new features** by using the ***add features*** tool from the ***digitizing toolbar***.

8. As you move the mouse close to a point the point will change to show a magenta cross hairs (+), this means the mouse has snapped to this point. Clicking the mouse will add this point as a vertex in the polygon. Continue following around the points to form a polygon. Right-mouse-click to finish a polygon.

9. When you finish a polygon a dialog will open to allow you to enter attributes. Click **OK** to add attributes and finalise polygon.

   Save your edits using the ***save layer edits*** icon on ***digitizing toolbar*** and ***toggle editing*** off when complete.

To add, move or delete a vertex, toggle to node mode using the ***node tool*** . Click the polygon to edit. Nodes/Vertex will appear as red squares.

- **Double click** to **add** new vertex.
- **Single click** to select existing vertex. The square will turn blue. Use the **DEL** key to delete
- **Click** and **drag** a vertex to **move**.

To add a hole (donut) to a polygon use the ***add ring*** tool from the ***advanced digitizing toolbar*** and sketch your polygon as described in step 8.

To delete a hole (donut) in a polygon use the ***delete ring*** tool from the ***advanced digitizing toolbar*** and click in the hole.

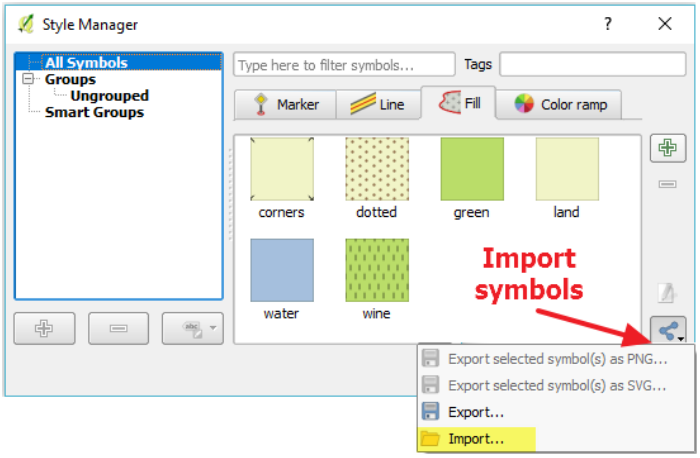
To split a polygon use the ***split features*** tool from the ***advanced digitizing toolbar*** and sketch the path to split. Multiple polygons will be created having the same attribution.

Useful editing shortcut keys.

Add new feature	Ctrl+.	**Zoom in**	Scroll wheel or	Ctrl ++
Delete last vertex	Del	**Zoom out**	Scroll wheel or	Ctrl +-
Undo	Ctrl+Z	**Zoom full**	Ctrl+Shift+F	
Cancel edit	Escape	**Pan**	Middle mouse or Spacebar (while adding a feature only) Note: spacebar also turns active layer visibility on/off	

4.2 Loading PAT Symbols into QGIS

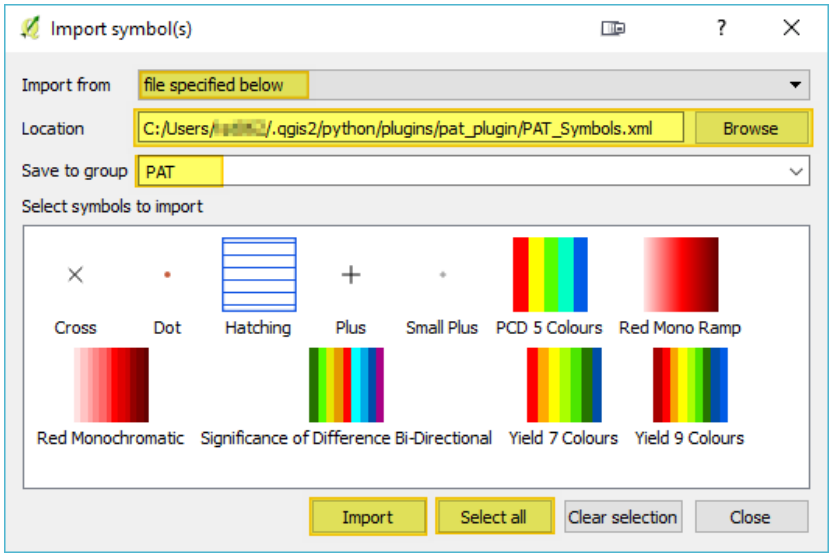
The PAT plugin includes a pre-defined set of symbols and colour ramps for use with datasets derived while using the plugin.



1. In QGIS, launch the **Style Manager**

   Settings menu -> Style Manager

2. From the lower right corner of the dialog, select ***Import*** dialog
3. Set ***Import from*** to ***file specified below.*** Browse to the users QGIS plugin folder and find the **PAT_symbols.xml** file in the pat folder (see 1.4 Uninstall PAT for help finding this folder).
4. Enter PAT as the ***Save to group.***
5. Select symbols to import or click ***Select all***.
6. Click ***Import***. If symbols with the same name are already loaded, will be notified and given the option to overwrite.



4.3 Analysing strip trial data using the moving window t-test

Following use of the ***Create Strip Trial Points*** and ***Extract pixel statistics for points*** tools, the user should have an output dataset which contains a minimum of five columns: Columns of x and y coordinates corresponding to locations along the trial strip and three columns of data corresponding to the extracted pixel values from the strip 'centre line' and from right and left offsets to that centre line (see page 40). The data are the input to the moving window t-test of Lawes and Bramley (2012). These data can then be analysed in a spreadsheet or statistics package. For users of Excel, this may be done as described below. Note that in this example, we demonstrate testing of the difference between the strip and one offset line only and use a simple case in which no prior information (e.g. location of different management zones) is available. In practice, users may which to compare the strip with both of the offset lines or alternative use the mean of the two offset lines for comparison with the strip.

1. Data setup:

Column A	Record number
Column B	x coordinate or Easting
Column C	y coordinate or Northing
Column D	Data from the trial strip
Column E	Data from the offset comparison line

It is assumed here that the column headings are in row 1 with the data starting in row 2.

2. Calculate a treatment difference

In Column F, insert a formula to calculate a treatment difference, e.g.:

= D2 – E2

Copy this formula down to the last row the dataset.

3. Select the size of your moving window.

The choice of this will need to be considered in relation to the grid pixel size, the length of the trial strip, and the distance setting used in the Create Strip Trial Points tool. In the example given in Lawes and Bramley (2012), they extracted values from a yield map at approximately 5 m intervals along their strip, which was approximately 700 m in length, and then used a window of 5 data values (i.e. approximately 25 m) for the analysis.

4. Calculate the treatment difference over the moving window

This will be calculated in Column G. Assuming that the size of the window is 5 data points, enter values of 0 in the first 4 rows of the column. In cell G6, calculate the average treatment difference from the difference data in Column F (step 2):

= Average(F2:F6)

Copy this formula down to the last row the dataset.

5. **Calculate the significance (P value) of the difference calculated at step 4 using the two-tail paired student t-Test.**

This will be calculated in Column H. In the cell of H6, enter the following formula:

=TTEST(D2:D6,E2:E6,2,1)

Copy this formula down to the last row the dataset.

The data generated by steps 4 and 5 enable the kinds of graphs displayed in Lawes and Bramley (2012) to be produced showing how the response to the treatment in the strip differs from the areas outside the strip, and the statistical significance of that difference, along the strip length.

5 Contributors

The following individuals have contributed the production of PAT and the underpinning pyprecag Python package, through coding, configuration and expertise in PA, and in the case of the Whole-of-block analysis tool, the development of statistical methodology.

Core Project Team: **Christina Ratcliff** (lead developer), **Rob Bramley** (project leader) and **David Gobbett** (technical lead)

Whole-of-block analysis tool: **Shuvo Bakar**, **Warren Jin** and **Brent Henderson**

pyprecag packaging and publishing: **Andrew Spiers**, **Adrian D'Alessandro** and **Daniel Collins**

6 Acknowledgments

CSIRO's development of PAT draws on many years of Precision Agriculture RDE across a range of cropping sectors and variously supported by CSIRO, Wine Australia, the former Cooperative Research Centre for Viticulture, Sugar Research Australia and the Grains Research and Development Corporation. However, the production of the PAT toolset has been sponsored and supported by Wine Australia (Project No. CSA1603) through funding from the Australian Federal Department of Agriculture and Water Resources as part of its Rural R&D for Profit program (Project No. 15-02-018 - Digital Technologies for dynamic management of disease, stress and yield). The input of Andrew Spiers, Adrian D'Alessandro and Daniel Collins was funded through a CSIRO IMT eResearch Collaboration Project. The support of these funders is gratefully acknowledged, as is the assistance of the project advisory panel [Dr Kathy Evans (TIA/University of Tasmania), Colin Hinze (Bird in Hand Winery) and Hans Loder (Penley Estate)].

7 References

Bramley RGV, Jensen TA. 2014. Sugarcane yield monitoring: A protocol for yield map interpolation and key considerations in the collection of yield data. International Sugar Journal 116, 1–12. doi:http://www.scopus.com/inward/record.url?eid=2-s2.0-84916603197&partnerID=40&md5=f5f44e347b0ac4a1e0d35d95e64ddb27.

Bramley RGV, Kleinlagel B, Ouzman J. 2008. A Protocol for the Construction of Yield Maps From Data Collected Using Commercially Available Grape Yield Monitors - Supplement No. 2. Precision viticulture Cooperative Research Centre for Viticulture, Adelaide 1–4.

Bramley RGV, Lawes RA, Cook SE. 2013. Spatially distributed experimentation: tools for the optimization of targeted management. Chapter 12 in: Oliver MA, Bishop TFA, Marchant BM. (Eds). Precision Agriculture for Sustainability and Environmental Protection. Earthscan, Food and Agriculture Series. Routledge, Abingdon, UK. pp. 205-218.

Bramley RGV, Williams S. 2001. A Protocol for the Construction of Yield Maps From Data Collected Using Commercially Available Grape Yield Monitors. Precision viticulture Cooperative Research Centre for Viticulture, Adelaide 1–4.

Gitelson AA. 2004 Wide Dynamic Range Vegetation Index for Remote Quantification of Biophysical Characteristics of Vegetation. Journal of Plant Physiology 161, 165–173. doi:10.1078/0176-1617-01176.

Gitelson AA, Viña A, Ciganda V, Rundquist DC, Arkebauer TJ. 2005. Remote estimation of canopy chlorophyll content in crops. Geophysical Research Letters 32, L08403. doi:10.1029/2005GL022688.

Lawes RA, Bramley RGV. 2012. A Simple Method for the Analysis of On-Farm Strip Trials. Agronomy Journal 104, 371-377.

Minasny B, McBratney AB, Whelan BM. 2005. VESPER version 1.62. Aust. Cent. Precis. Agric. McMillan Build. A 5,. https://sydney.edu.au/agriculture/pal/software/vesper.shtml.

Taylor JA, McBratney AB, Whelan BM. 2007. Establishing management classes for broadacre agricultural production. Agronomy Journal 99, 1366–1376. doi:10.2134/agronj2007.0070.

Panten K, Bramley RGV. 2011. Viticultural experimentation using whole blocks: Evaluation of three floor management options. Australian Journal of Grape and Wine Research 17 136-146.

CONTACT US

t 1300 363 400
+61 3 9545 2176
e csiroenquiries@csiro.au
w www.csiro.au
w https://www.csiro.au/en/Research/AF

AT CSIRO, WE DO THE EXTRAORDINARY EVERY DAY

We innovate for tomorrow and help improve today – for our customers, all Australians and the world.

Our innovations contribute billions of dollars to the Australian economy every year. As the largest patent holder in the nation, our vast wealth of intellectual property has led to more than 150 spin-off companies.

With more than 5,000 experts and a burning desire to get things done, we are Australia's catalyst for innovation.

CSIRO. WE IMAGINE. WE COLLABORATE. WE INNOVATE.

FOR FURTHER INFORMATION

CSIRO Agriculture and Food
Christina Ratcliff
t +61 8 8273 8157
e christina.ratcliff@csiro.au
w https://people-my.csiro.au/r/c/christina-ratcliff

CSIRO Agriculture and Food
David Gobbett
t +61 8 8303 8741
e david.gobbett@csiro.au
w https://people.csiro.au/g/d/david-gobbett

CSIRO Agriculture and Food
Dr Rob Bramley
t +61 8 8303 8594
e rob.bramley@csiro.au
w https://people.csiro.au/B/R/Rob-Bramley