Working with Shapefiles and Geopackages in QGIS

Maia Williams
mw@maia.id.au

February 2023

Contents

1 Setup

1.1 Installation

As at March 2023 the stable version of QGIS is 3.22 and is the recommended version to use. The screenshots in this document are from an older version so may look slightly different to the 3.22 interface (but the tools work the same way).

QGIS can be downloaded from here: https://qgis.org/en/site/forusers/download.html. Accept all the defaults during installation.

1.2 Shapefiles

Shapefiles are spatial vector filetype made of 4 – 6 components (.shp, .dbf, .prj, .shx and more). These component files must all be kept together and have the same name, otherwise the data will become corrupted. If a QGIS project file (.qgz) is moved (rather than using 'Save As') or if shapefile is moved, then the link between the QGIS project file and shapefile will be broken. In this case a new path can be specified.

To load a shapefile to a QGIS project drag and drop the .shp file to the Layers window within QGIS.

1.3 Geopackage files

A geopackage is a spatial vector data type stored in one file with extension .gpkg. A gpkg can contain multiple layers but data downloaded from data providers usually only contains one layer.

A gpkg file can be loaded to a QGIS project by dragging and dropping the gpkg file to the Layers window.

1.4 Base layers

You can add Nearmaps jpg images (or other georeferenced datasets) to QGIS by dragging and dropping the file into the Layers window. Take care to keep .jgw files alongside .jpg files to store the associated coordinate system.

To add streamed base layers go to 'XYZ Tiles' or 'WMS/WMTS' in the Browser window. Right click to add a new connection. Paste in one of the service URLs from below, give the connection a name and Okay. Drag and drop the resulting layer into the Layers window.

1.4.1 XYZ Tiles

Google hybrid

http://mt0.google.com/vt/lyrs=y&hl=en&x={x}&y={y}&z={z}

Google Satellite

http://mt0.google.com/vt/lyrs=s&hl=en&x={x}&y={y}&z={z}

Google maps

http://mt0.google.com/vt/lyrs=m&hl=en&x={x}&y={y}&z={z}

OpenStreetMap

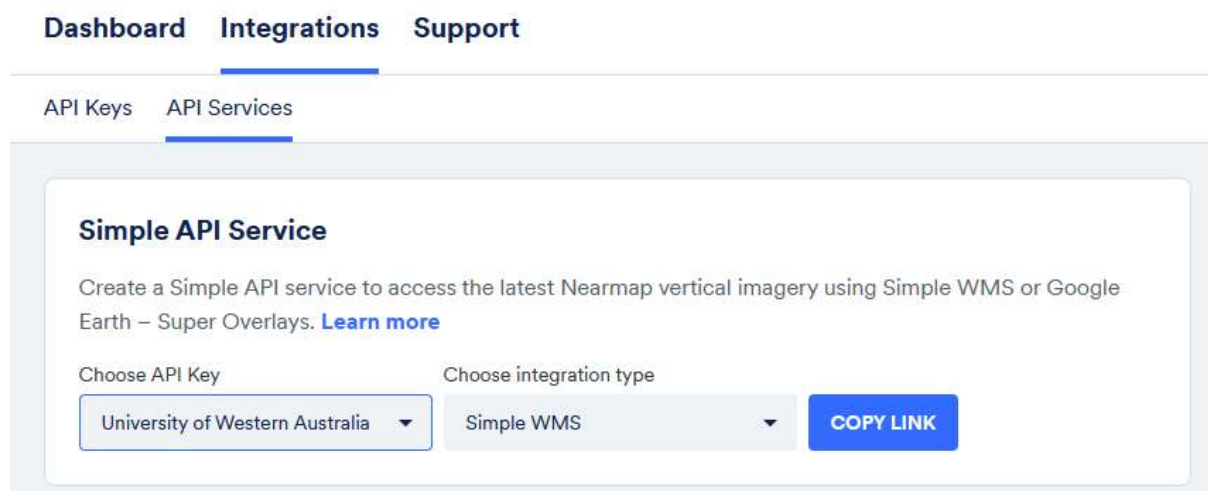https://tile.openstreetmap.org/{z}/{x}/{y}.png

1.4.2 WMS layers

Nearmap

Nearmap images can be added as a streamed Web Map Service (WMS) in QGIS.

To get the WMS URL, sign in to Nearmap (using your UWA email address) and go to Integrations>API Services. Select the API Key as the 'University of Western Australia' and the integration type as 'Simple WMS'. Copy the link and then add as a new WMS service in QGIS.

SLIP WMS layers

All public layers with WMS services are listed here:
https://catalogue.data.wa.gov.au/group/?q=SLIP+Public

All subscription layers with WMS services are listed here:
https://catalogue.data.wa.gov.au/group/?q=SLIP+Subscription

> Eg: Environmental layers
> https://services.slip.wa.gov.au/public/services/SLIP_Public_Services/Environment/MapServer/WMSServer
>
> Boundary layers
> https://services.slip.wa.gov.au/public/services/SLIP_Public_Services/Boundaries/MapServer/WMSServer
>
> Roads
> https://services.slip.wa.gov.au/arcgis/services/Landgate_Subscription_Services/Roads/MapServer/WMSServer

1.5 Coordinate Reference Systems (CRS)

GISs differ from CAD applications because they can accurately overlay datasets with different coordinate systems. This means it is easy to combine datasets from multiple sources. It is important, however, to specify the coordinate system for the QGIS project and when exporting to dxf or other formats.

Tip: Datasets exported to dxf will need to be given the same coordinate system as the base feature survey (or other base data sources) to be used in AutoCAD. If the coordinate systems don't match then layers will not overlay (even if manually moved).

The common coordinate systems used in Perth and their codes are:

- Perth Coastal Grid (GDA2020 PCG2020): EPSG 8031
- (Perth Coastal Grid (GDA94 PCG94): EPSG 102216)
- GDA2020 MGA50: EPSG 7850
- (GDA94 MGA50: EPSG 28350)

Other coordinate systems used in Western Australia are:

- GDA2020 MGA51: EPSG 7851
- (GDA94 MGA51: EPSG 28351)

- GDA2020 MGA52: EPSG 7852
- (GDA94 MGA52: EPSG 28352)

GDA2020 (Geocentric Datum of Australia 2020) is replacing GDA94, so over the next years GDA2020 will become more common than GDA94.

Please note there are many more coordinate systems than the ones listed above. The list of Western Australian towns with their own local coordinate systems is here: https://www0.landgate.wa.gov.au/business-and-government/specialist-services/geodetic/project-grids

Most of these local town grids do not have their coordinate systems defined in QGIS so cannot be selected from the list offered. They can be manually added using custom CRS definitions.

1.6 Colours

A colour palette can be set up and saved as a .gpl file. This file can then be imported to any computer running QGIS. The file only needs to be loaded once on a computer and will then be available to all QGIS projects.

To set up a colour palette go to Settings>Options>Colours. Use the '-' button to remove the default colours. Use the '+' button to specify new colours. You can add colours using RGB values or HTML codes (eg #ff0001) (but not CYMK values). You can give the colours names by double clicking in the 'Label' column. After you have added all the colours in your palette click the "disk" button (Export Colours) in the RHS menu. Save this .gpl file somewhere.

To load the gpl file onto another computer, open QGIS and go to Settings>Options>Colours>RHS menu "folder" icon (Import colours from file). These colours will be available as Standard Colours through all Symbology menus.

QGIS currently does not allow for the definition of colours using the CYMK system. It also does not allow for defining a colour space (RGB or CYMK) when exporting images.

2 Data conversions

2.1 From QGIS layer to AutoCAD

QGIS can export vector data (from shapefiles or geopackages) to dxf format. This is a good way to get shapefile data into AutoCAD. If the dataset has elevation values (eg. contour lines), then the data needs to be converted to a 3D shapefile before being exported to dxf. Polygon datasets will export to dxf as hatch areas, so it can be useful to convert them to lines so they export to polylines instead. If datasets are large it can be useful to clip them to your area of interest before processing, but it is not absolutely necessary.
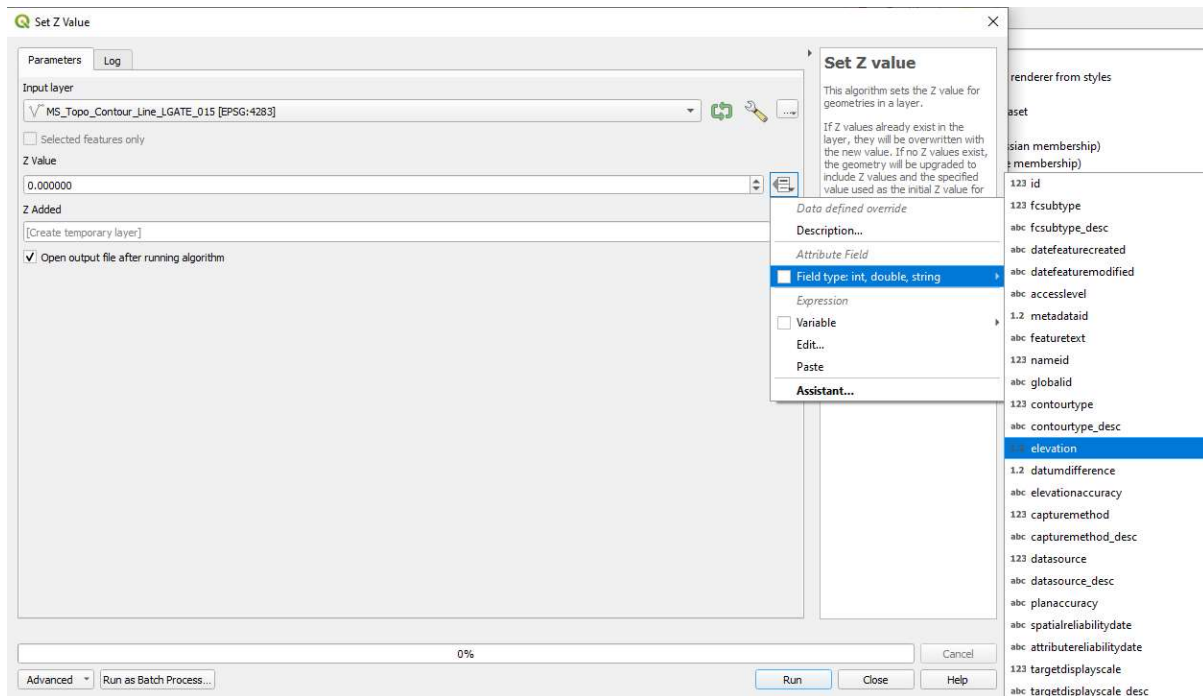
2.1.1 Pre-processing

To clip shapefiles/geopackages:

1. Create a shapefile to mark the boundary of your area of interest. Go to Layer>Create Layer>New Shapefile Layer. Specify the location of the file and the name. Specify the file to be a polygon and give it the same coordinate system as your project has.
2. Highlight the new layer then click on the 'Toggle Editing' pencil button in the top toolbar. Click on the 'Add Polygon Feature' star button. Click to form the corners of the bounding area and right click to end the shape.
3. Click the 'Toggle Editing' pencil again and select Save edits.
4. In the Processing Toolbox window type "Clip" into the search window. Open the tool called 'Clip'. Choose the original dataset as the first layer and the newly created area of interest shapefile as the second. Specify the location and name of the output shapefile.

To add Z values to contour data:

1. In the Processing Toolbox window type "Z". Open the tool called 'Set Z value'.
2. Specify the contour shapefile. Under Z Value click on the 'Data defined override box' and choose the attribute (field) containing the contour elevation value. Specify the location and name of the output file.

To reproject layers to the coordinate system being used in AutoCAD (eg GDA94 MGA50)

1. In the Processing Toolbox window type in "reproject". Open the 'Reproject layer' tool.
2. Specify the dataset to reproject and choose the output coordinate reference system. Choose the same as the desired coordinate system for the AutoCAD file.

To convert polygons to lines:

1. In the Processing Toolbox window type in "polygon". Open the 'Polygon to line' tool.
2. Specify the layer to convert and the output location and file name.

2.1.2   Convert to dxf Method 1 (for multiple layers without z values)

1. Make sure all layers to convert are prepared (converted to lines, reprojected to the coordinate system of the destination AutoCAD file (eg GDA94 MGA50)).
2. Zoom to your area of interest (to define the extent of exported data if you have not clipped it already).
3. Go to Project>Import/Export>Export Project to DXF.
4. Make sure the correct layers are selected. Specify the correct coordinate system. Specify the output file location and name. Select 'Feature Symbology'. Set the symbology scale to 10000 (or appropriate).
5. Tick 'Use layer title as name if set', ' Export features intersecting the current map extent' and 'Export labels as MTEXT elements' (if the data is labelled). Leave all other defaults.

Tip: If, when opened in AutoCAD, the exported dxf has an error when using COPYBASE then save the file as a 2013 dwg and reopen. There are not usually errors with Method 1.

Note that contour data (with z values defined) exported using Method 1 will convert to 3D polyline layers which can be annoying in AuotCAD. Method 2 is the best method for exporting contour data to 2D polylines with Elevation values.
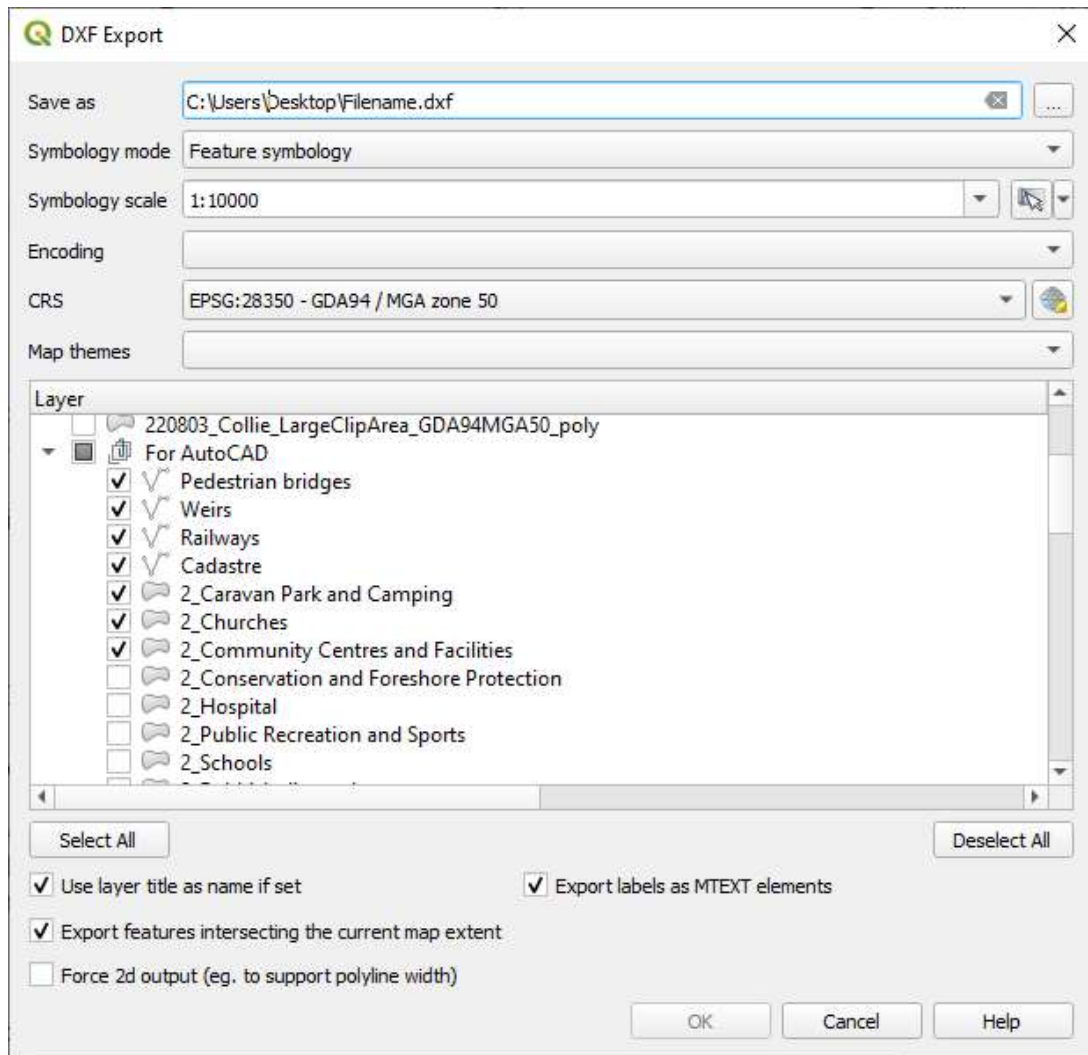
2.1.3   Convert to dxf Method 2 (for layers with z values)

1. Select the line layer with z values specified (see Pre-processing section, above). Right click and choose Save Features As.
2. Specify the location and name of the output file, and specify the format as dxf. Specify the output coordinate system (to match other data already in AutoCAD). Choose 'Linestring' as geometry type. Tick 'Include z-dimension'.
3. Open the resultant dxf in AutoCAD.

This method of export doesn't allow for exporting labels. In AutoCAD the z values will show in the layer Properties as elevation values.
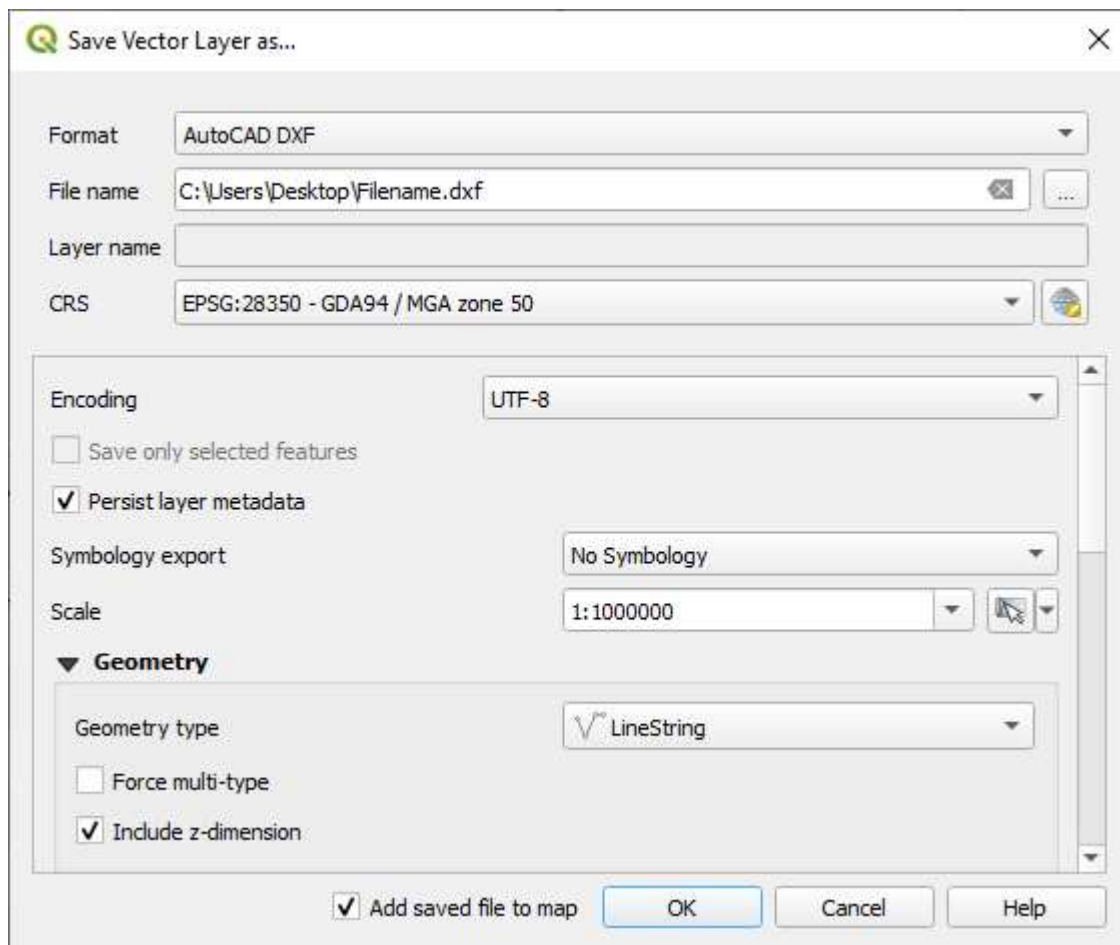
## 2.2 From QGIS layers to Illustrator layers

Data in shapefile/geopackage format is vector data. If exported as a vector pdf from QGIS, the data can be opened and edited in Illustrator. By exporting multiple layers using the same print layout, each pdf (of a single vector dataset) will overlay correctly in Illustrator. These are the steps for exporting data to 1:10 000 at A3, for example. Some extra information about using Print Composer is here:
https://docs.qgis.org/3.22/en/docs/user_manual/print_composer/overview_composer.html

1. Set the QGIS project coordinate reference system to the one you want to use for the Illustrator file, via Project>Project Properties>CRS.
2. Start a new print layout (Project>New Print Layout) and set the page dimensions(right click>Page Properties).
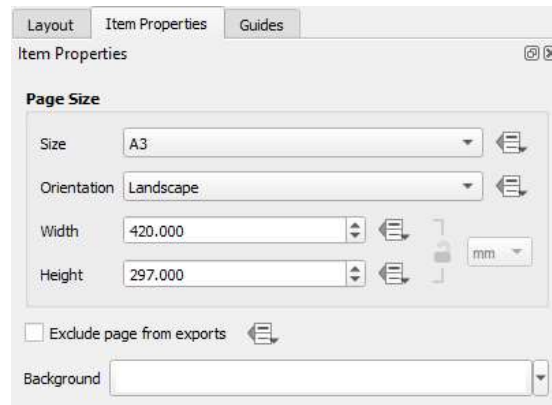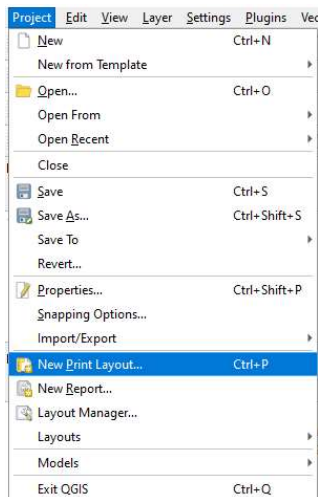
3. Set the page background to be transparent.
4. Add a map to the page (Add item>Add map).
5. Zoom and pan to the project area. Specify the scale as 10000 (with the map item selected>Item Properties).
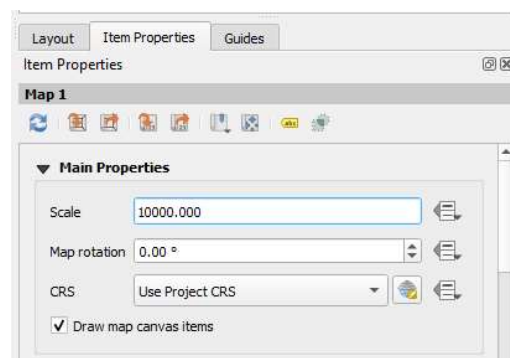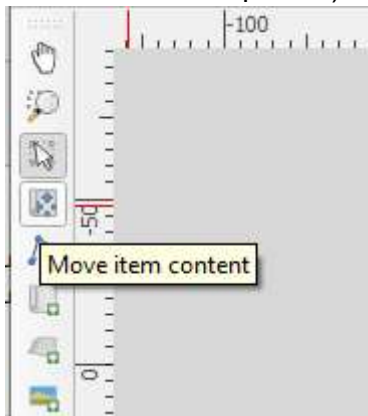
6. In the main QGIS map view turn on only one layer and set the symbology (and labels).
7. Back in the print layout view go to the Layout tab (next to the Item Properties tab) and tick on the 'Always export as vectors' option.
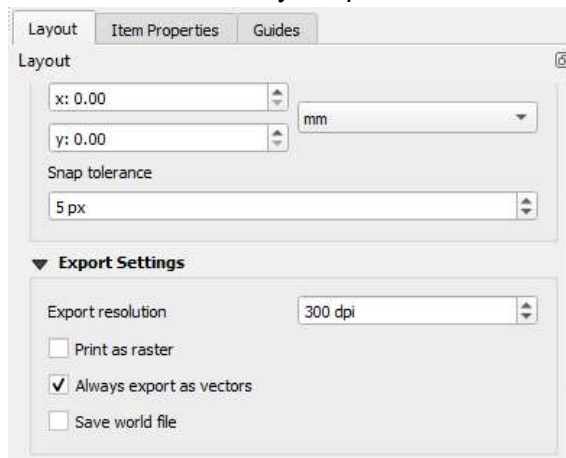
8. Go to Layout>Export as PDF. Accept the defaults and export the print layout.
9. Repeat 7 and 8 for all QGIS layers you need in Illustrator. Be careful not to pan or zoom the map area in print layout view between exporting layers.

10. Open the pdfs in Illustrator and use Paste in Place to overlay them as layers in a single Illustrator document. Provided they were all exported from the same QGIS print layout then they will overlay correctly in Illustrator.

3   Working with vector data

3.1   Attribute table

A shapefile/geopackage comprises feature data and feature geometry. In the map view you see the geometry of the features - each feature has a representative point, line or polygon. In the attribute table you can view and interrogate the data values associated with those features.

1. To open the attribute table, right click on the layer and click on 'Open Attribute Table'.
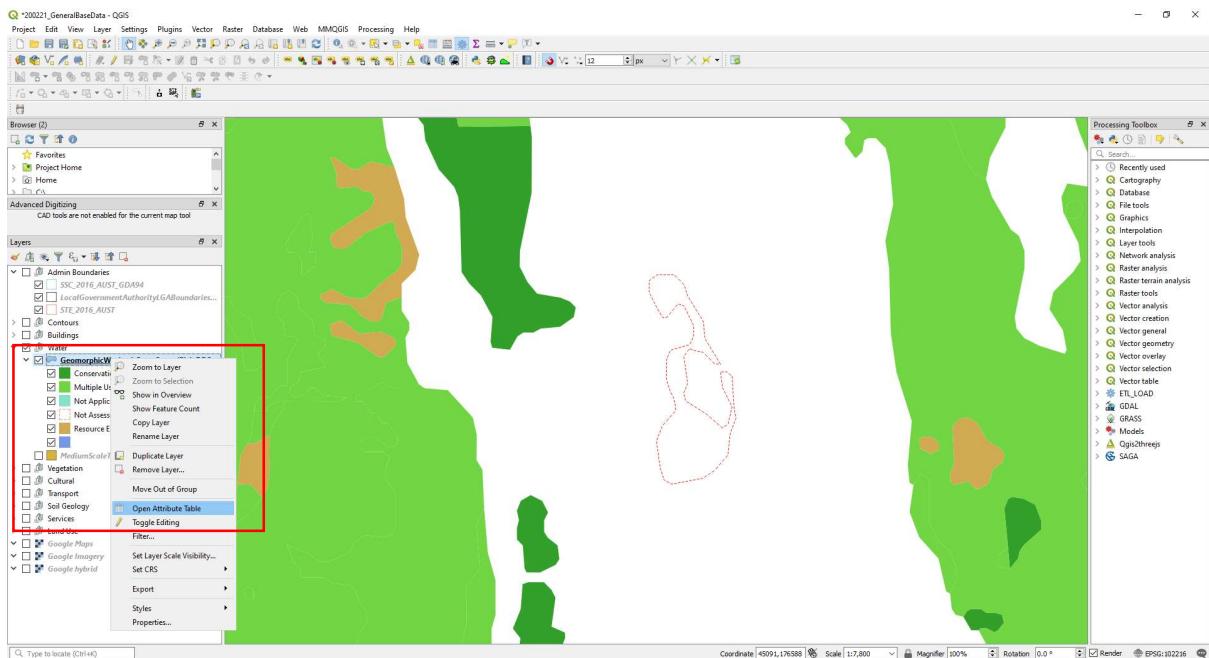
2. This is what an attribute table looks like. The attribute (column) names are limited to 10 characters in a shapefile, so they are often a little cryptic (there are no restrictions in geopackages). Some datasets from data.wa.gov.au (and other agencies) come with data dictionary/metadata documents which explains what each attribute is.
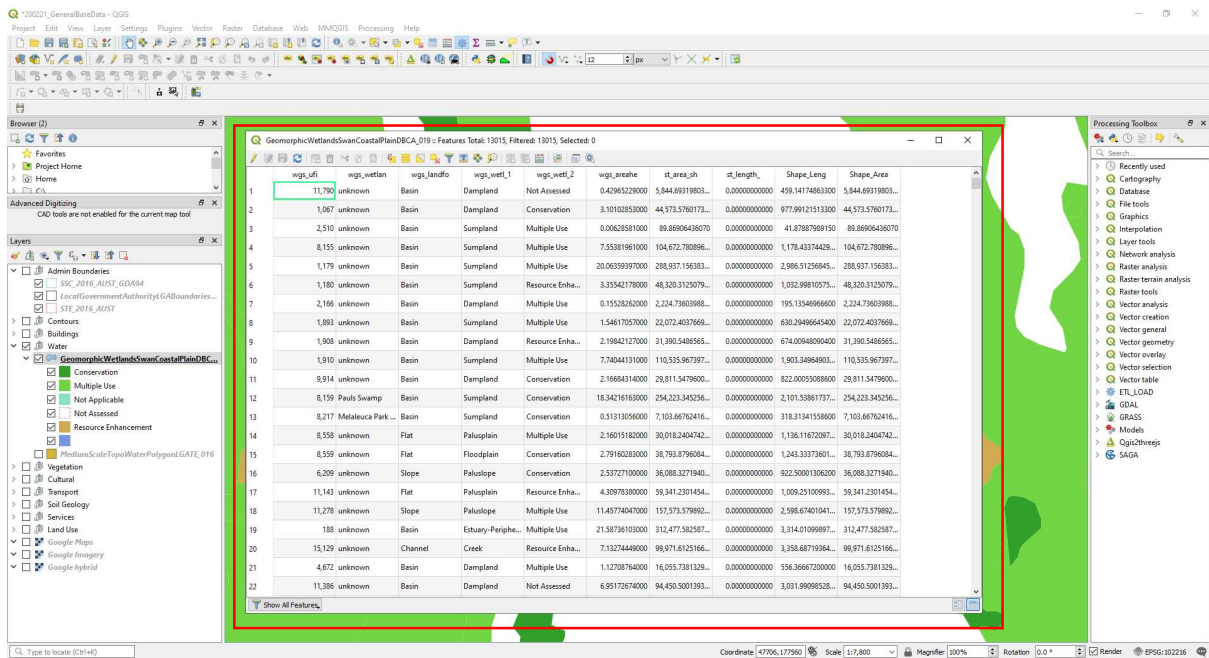
3. You can sort the columns by clicking on the attribute name. You can highlight features (which will also highlight the corresponding point, line or polygon) by clicking on the row number.
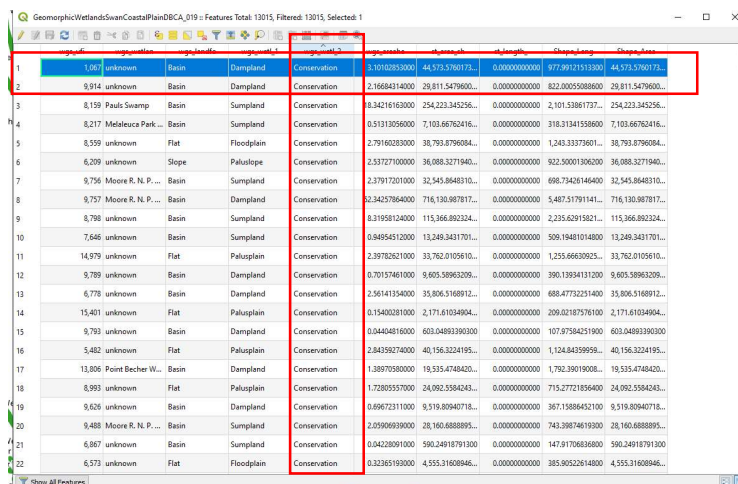

4. You can dock the attribute table window and choose to 'Show Selected Features' (bottom left button). Then if you use the Select Features tool (from the main toolbar) to click on a feature, its data will be displayed in the attribute table. The highlighted feature is shown in yellow. You have to click on the layer name in the Layers list before you can highlight features from that shapefile/geopackage.
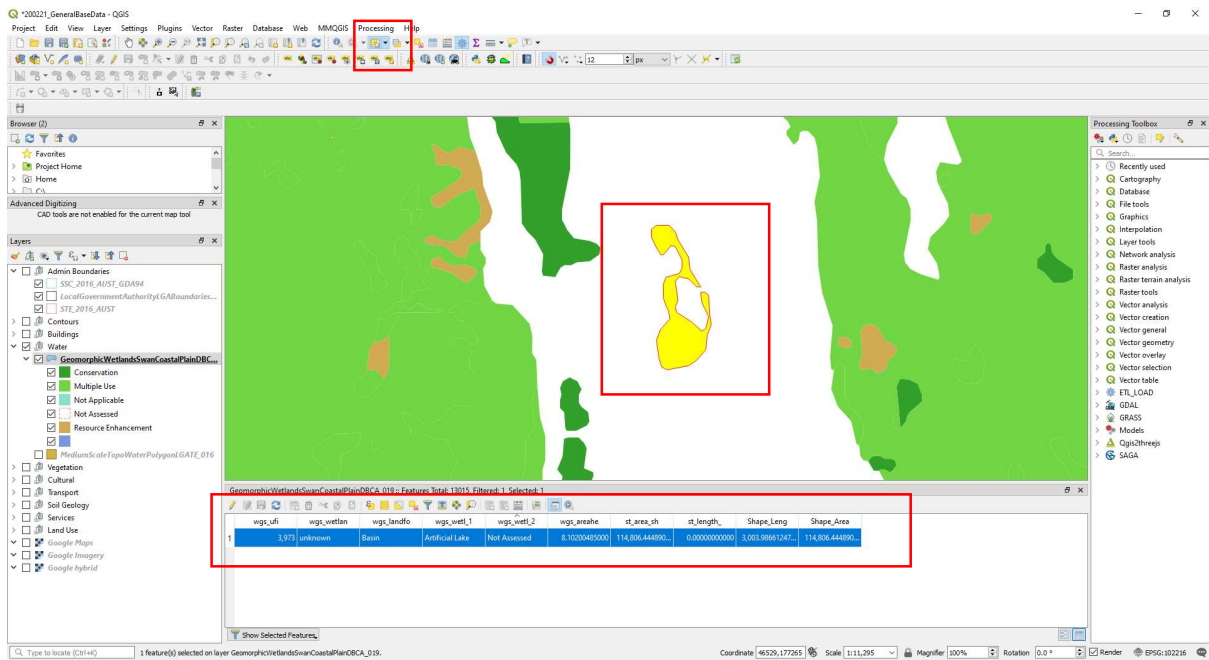
5. An alternative way to see the attribute data for a feature is to use the Identify tool (from the main toolbar) to click on a feature. This tool will pop up a data snapshot in the Identify Results window and mark the identified feature in red. You have to click on a layer name in the Layers list before you can identify features from that shapefile/geopackage.
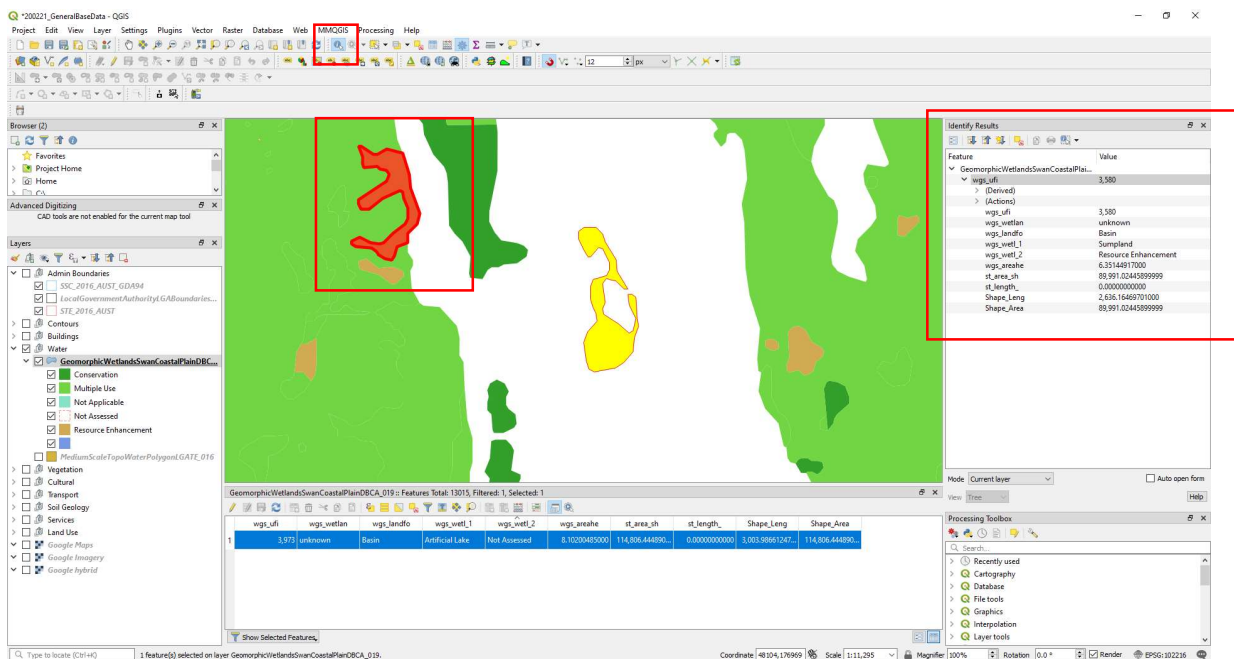
6. To edit a shapefile/geopackage (data or geometry) you click on the Toggle Editing icon in the menu bar. You can then add attributes, edit values, and add or modify feature geometry. You need to click the Toggle Editing button again to save edits and end the editing session once you're finished.
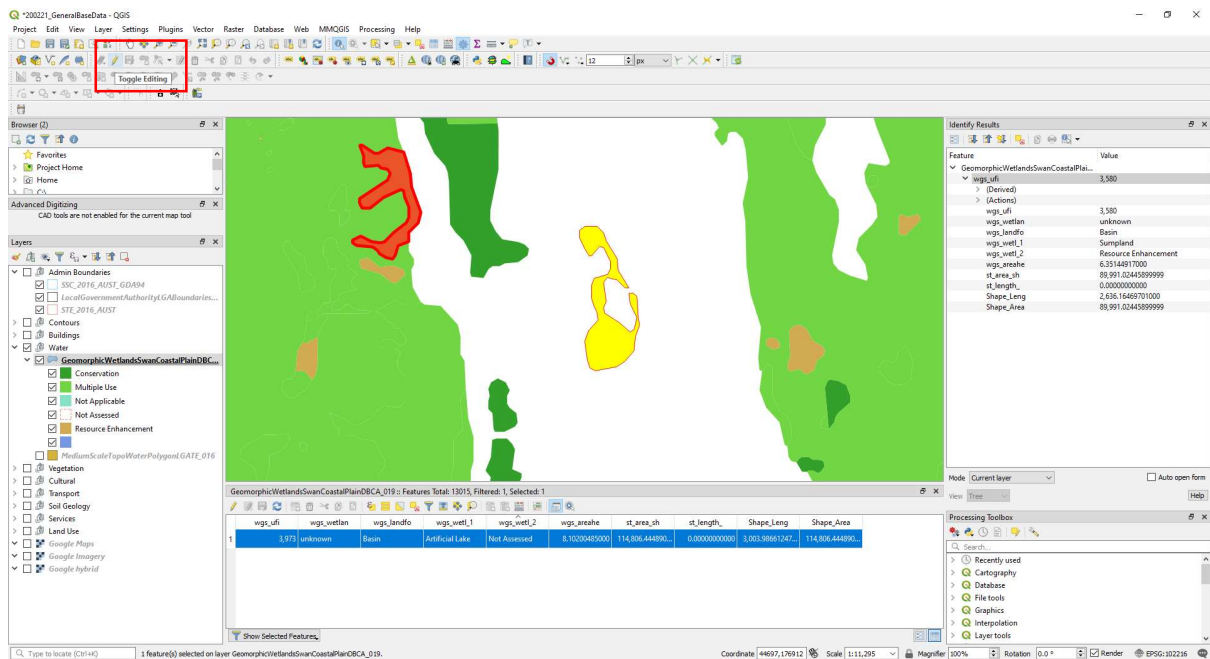
3.2 Symbology

Symbology refers to the colours and style of the displayed data layers (shapefiles or raster images). There are some notes about Symbology here:
https://docs.qgis.org/3.22/en/docs/user_manual/working_with_vector/vector_properties.html#symbology-properties

Here are some steps to setting basic style options.

1. To access the Symbology menu, right click on a data layer and go to Properties>Symbology. From here you can choose to symbolise all features the same (Single Symbol) or use different symbols/colours depending on the value of the feature for a specified attribute (Categorized).
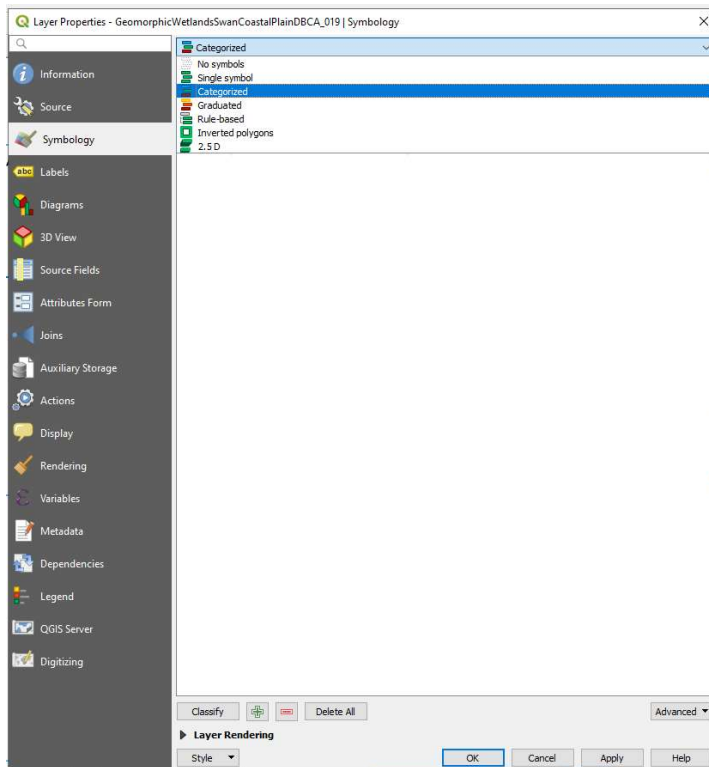
2. If you choose the Categorized option then choose the attribute containing the values you want to delineate. For example, Column = wgs_wetl_2. This will show different types of wetlands as different colours.
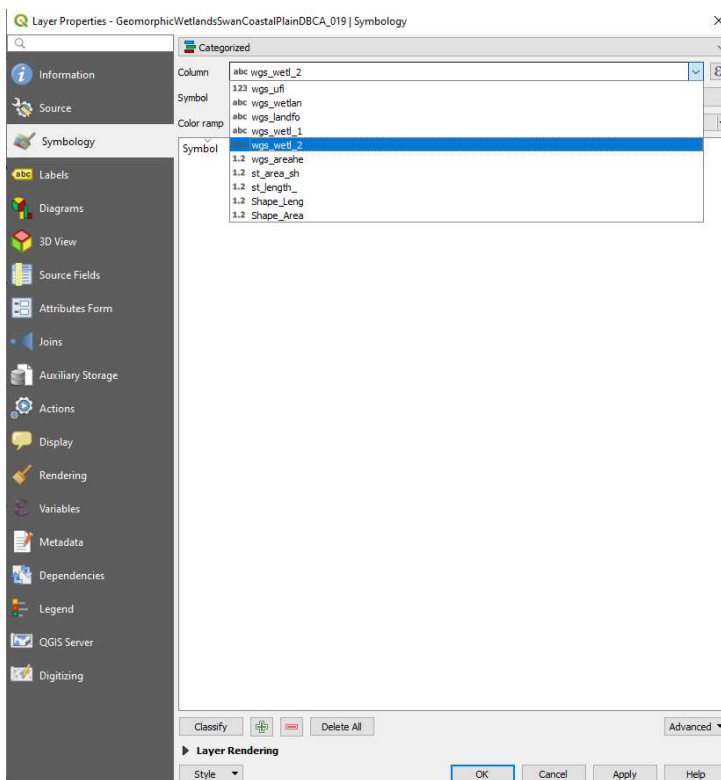
3. After choosing the attribute/column press the 'Classify' button. You will see all the values for the attribute listed. The bottom blank value captures any data that has a NULL value.
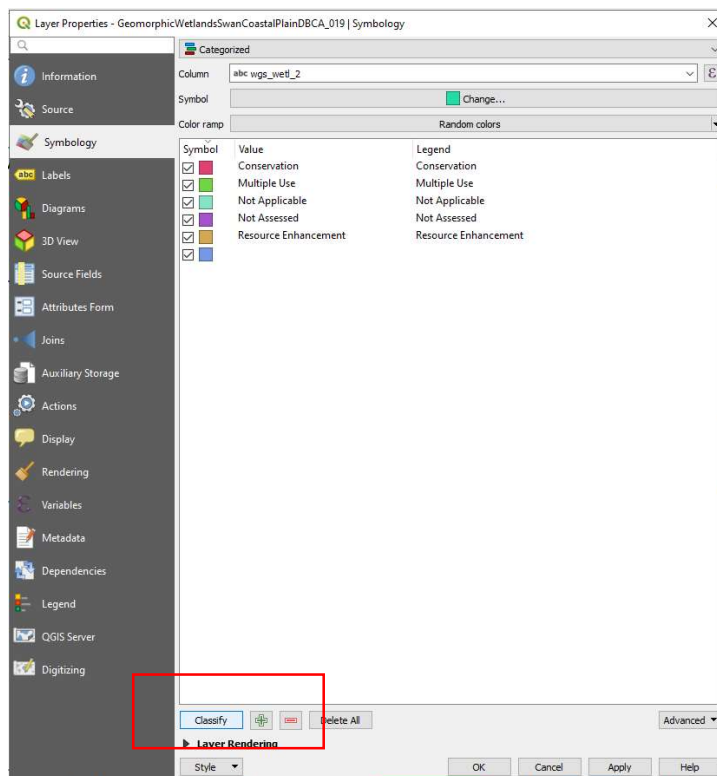
4. To make changes to all the symbols (eg remove border lines) you click on the button that says 'Change…'. Click on where it says 'Simple Fill' then next to 'Stroke Colour' select 'Transparent Stroke' to remove the stroke. Click okay to get back to the previous menu.
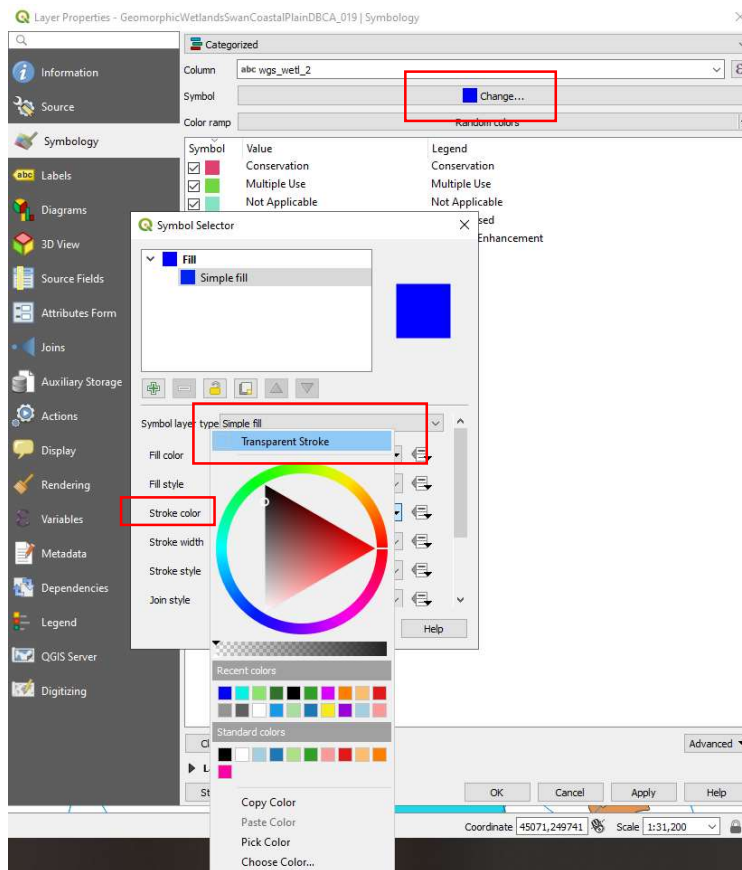
5. To set the colour for each value, double click on the 'Symbol' swatches. Click on the down arrow next to the colour sample where it says 'Colour'. Choose your colour from the list of 'Recent colours', 'Standard colours' or click 'Choose colour' to select a new colour.
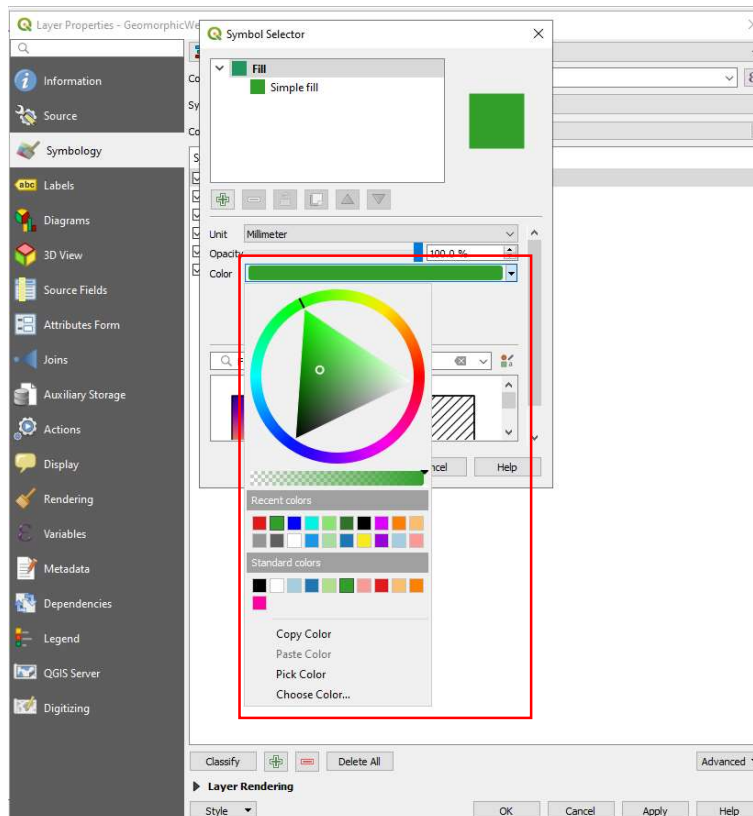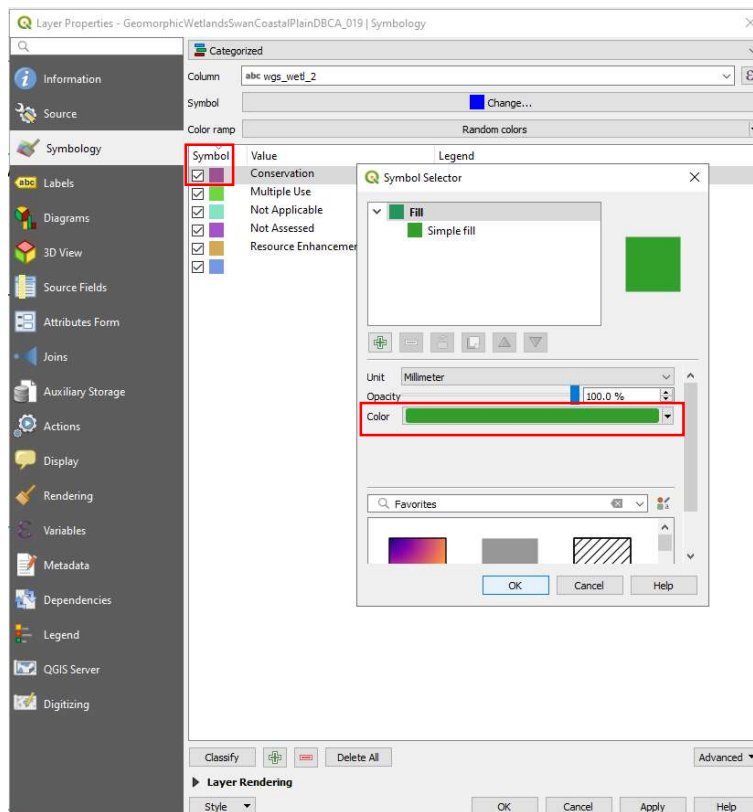
6. You can also choose to make the fill colour transparent and set a boundary colour/stroke style. For example: a red dashed outline means 'Fill colour' =

transparent, 'Stroke colour' = red, 'Stroke style' = dashed line and you can change stroke width.
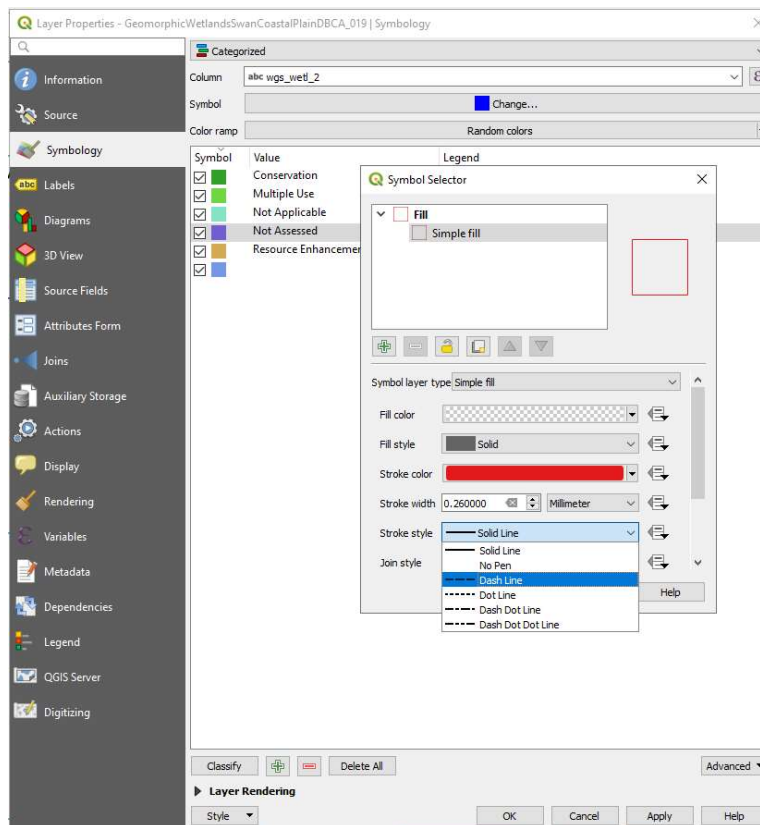
7. Once you have set all the colours and styles you click Okay to get back to the map view. You can still edit the styles from here by double clicking on the colour swatches in the Layers list.
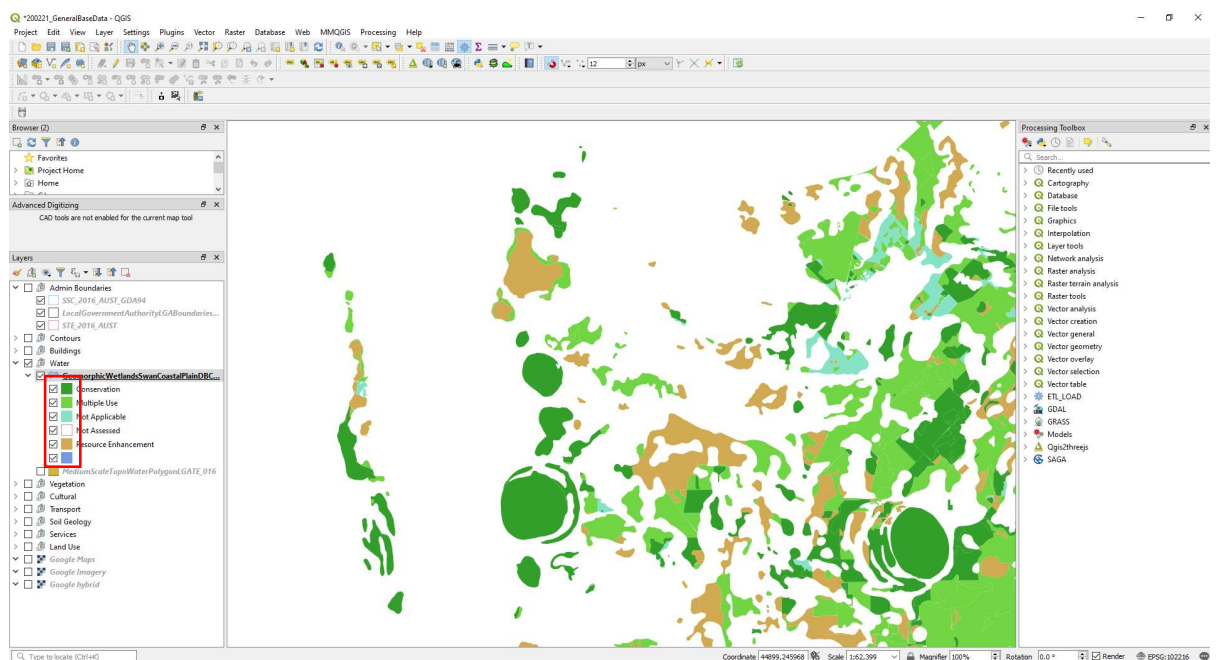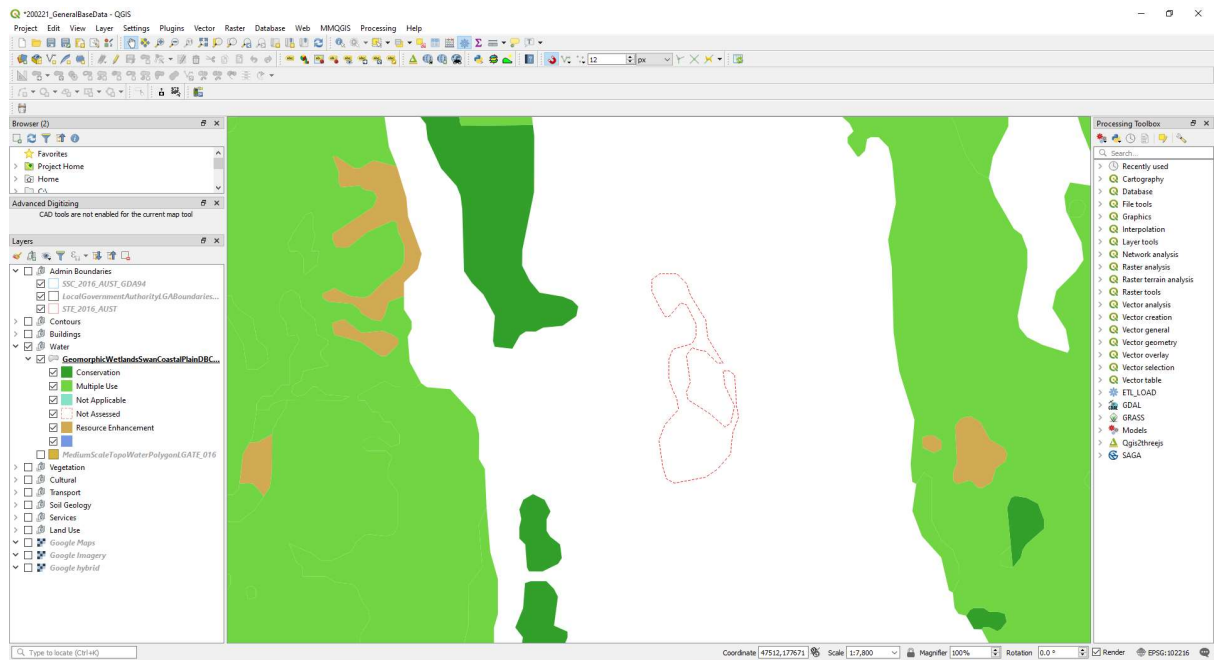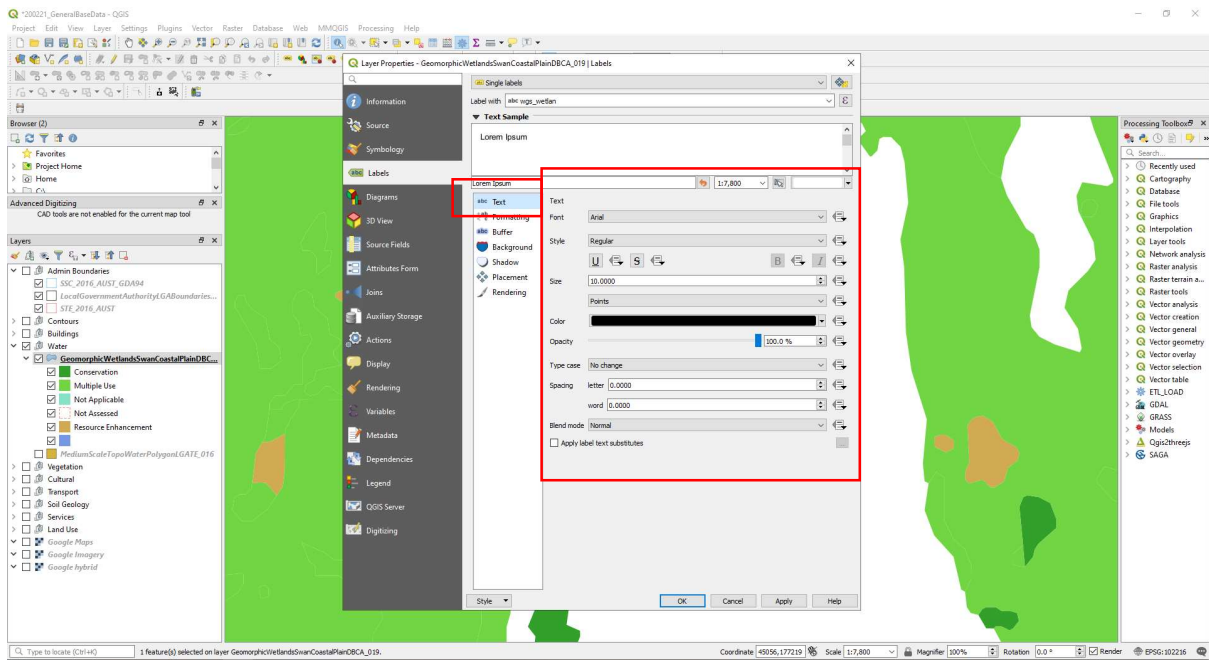
3.3 Labels

You can add labels to your map view (and therefore print layout) using values from the attribute table. You can manually move labels if the automatic label positions are not good.

Some notes are here:
https://docs.qgis.org/3.22/en/docs/user_manual/working_with_vector/vector_properties.html#labels-properties

These are steps to getting quick basic labels.

1. Right click on a layer and go to Properties>Labels. The most common is to choose 'Single Labels' from the labelling options.

2. From the 'Label with' drop down menu, choose the attribute with the labels you want to display.

3. Then set the font type, colour and size in the 'Text' section.

4. You may like to put a buffer around the font to make a visible on top of the map. You can adjust the width and colour of the buffer.
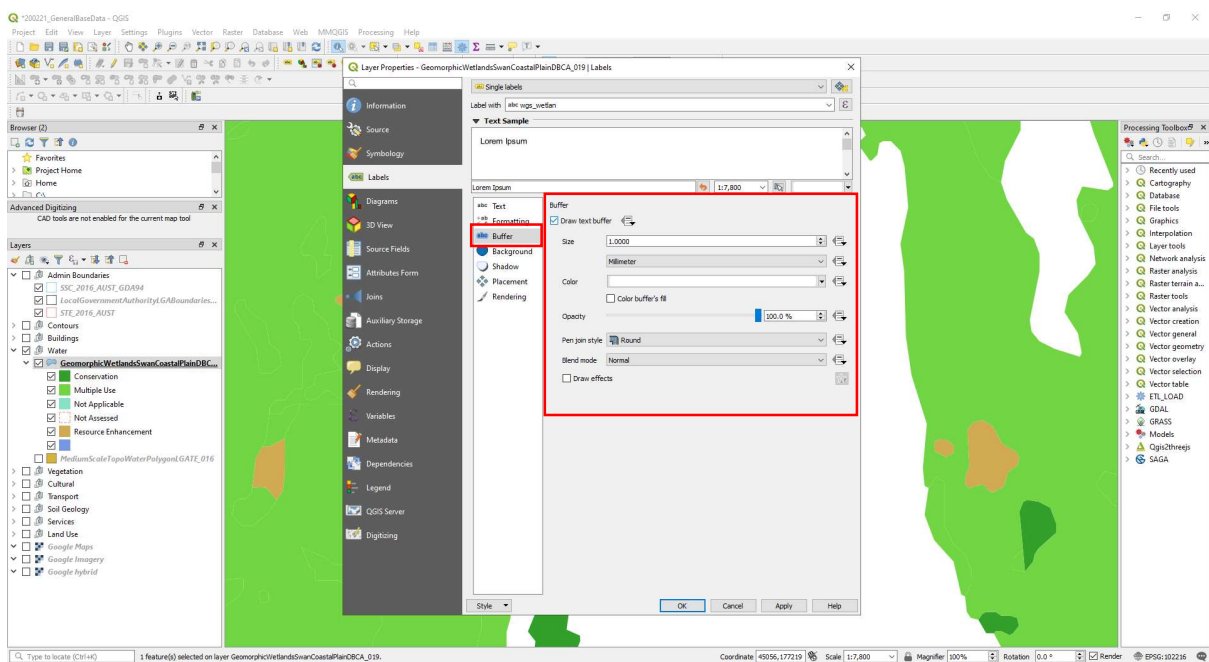

5. In the 'Placement' tab you can choose Offset, Around centroid or Horizontal (slow) as the best label placement options. Slide the 'Priority' bar to High to make sure your labels are not hidden by other data layers.
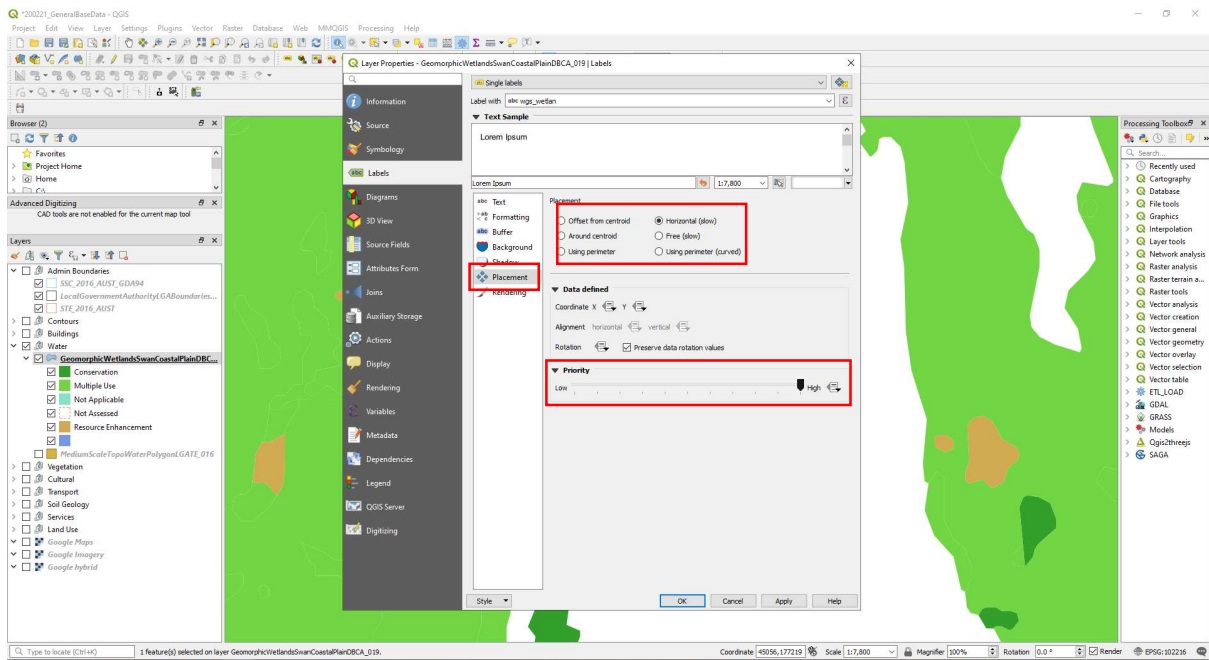
6. Click Okay to display the labels, as below.
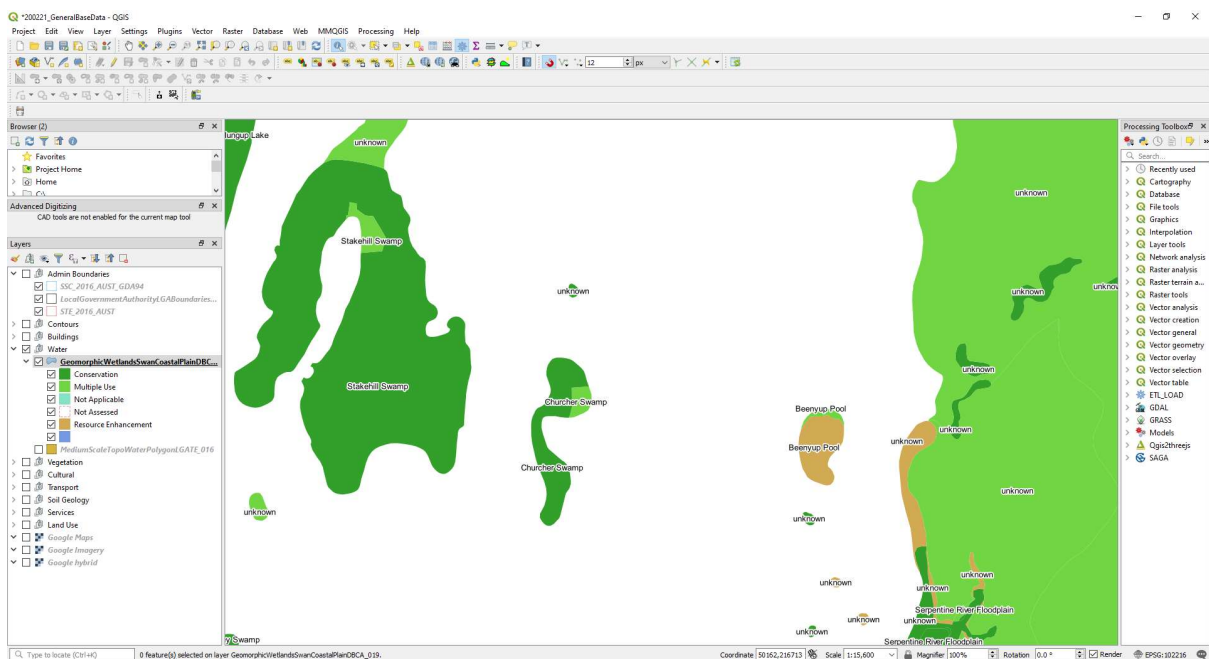
7. You can manually move labels that have not been placed in a good location. First open the 'Labels Toolbar' by right clicking in the main toolbar area. Then tick on the Move Labels tool within the Labels Toolbar. Click on the label you want to move. You will be asked which field to store the label position data against. You must choose an attribute that contains a unique identifier for the features so that you can move each label independently. Drag and drop labels. Take note of the scale you're working at because the manually moved labels will need to be specifically placed to work at a given print layout scale.
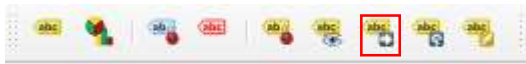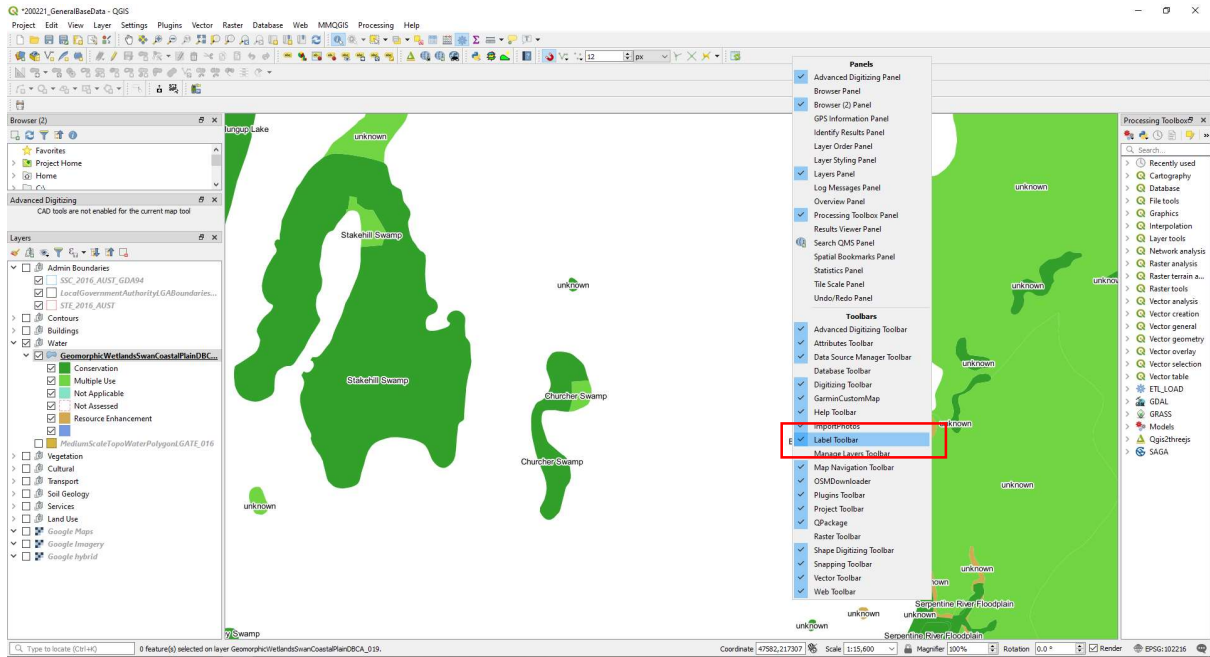
8. If you are labelling data where many features have the same attribute value then you may end up with labels like below. The work around for this situation is to dissolve the shapefile and use this dissolved version to produce the labels (and reduce duplicates).