

ACT Project Number:
271498

Project acronym:
ELEGANCY

Project full title:
Enabling a Low-Carbon Economy via Hydrogen and CCS

ERA-Net ACT project

Starting date: 2017-08-31
Duration: 36 months

D4.5.2
User Documentation – GIS Tools
Issue 1

Actual DELIVERY date: 2020-10-19

Organization name of lead participant for this deliverable:
Imperial College London

ACT ELEGANCY, Project No 271498, has received funding from DETEC (CH), BMWi (DE), RVO (NL), Gassnova (NO), BEIS (UK), Gassco, Equinor and Total, and is cofunded by the European Commission under the Horizon 2020 programme, ACT Grant Agreement No 691712.		
Dissemination Level		
PU	Public	X
CO	Confidential, only for members of the consortium (including the Commission Services)	

Deliverable number:	D4.5.2
Deliverable title:	User Documentation – GIS Tools
Work package:	WP 4
Lead participant:	Imperial College London (ICL)

Authors		
Name	Organisation	E-mail
Nixon Sunny*	ICL	nixon@imperial.ac.uk
Niall Mac Dowell	ICL	niall@imperial.ac.uk
Nilay Shah	ICL	n.shah@imperial.ac.uk

* Corresponding author

Executive Summary
<p>The purpose of this document is to provide a potential user with a reference guide for use in conjunction with the Geographical Information Systems (GIS) input and output tools as part of the design mode in work package (WP) 4 of the ERA-Net ACT ELEGANCY project. The aim of these tools is to automate the pre/post processing procedures that are required to convert relevant input/output features into a format that can be read using the mathematical modelling software and visualised with ease. The tool can be effectively used to eliminate any manual/ human errors that may occur when translating the data using other means. This issue of the user guide accompanies a toolkit release, which can be applied by the relevant WP 5 case study teams to generate the input data needed for network design. It is not necessary to use the tools for case study analysis, but it may be useful if the case study teams have collated input data with geographical coordinates. The document contains detailed step-by-step instructions on interaction with QGIS (free GIS visualisation/ analysis software) to enable the usage of the developed tools. The version of QGIS used for generating this user guide is 2.18.22. All version 2 releases should be compatible with these tools and if there are any issues, please forward your comments/ enquiries to the following email address: nixon@imperial.ac.uk</p> <p>This issue contains additional details on adding user-defined inputs to create vector layers from data shaped in the form of delimited text files. This will enable users to create their individual vector layers from data and an accompanying coordinate reference system rather than have shapefiles and their dependencies.</p>

TABLE OF CONTENTS

	Page No.
1	REQUIREMENTS2
2	QGIS PROJECT CREATION/ VISUALISATION3
2.1	Create a new project3
2.2	Save the new project4
2.3	Load up the data layers5
2.4	Loading up delimited text-based layers6
2.5	Saving layers using a common CRS7
2.6	Install processing toolbox8
2.7	Add Input/ Output Tool script9
3	INPUT TOOL DETAILS10
3.1	Shapefile path10
3.2	Land use raster10
3.3	Complete region shapefile11
3.4	Desired Number of Cells14
3.5	Smallest grids14
3.6	Subsequent options15
4	OUTPUT TOOL DETAILS17
4.1	Outputs path17
4.2	File names17
4.3	Gridded shapefile layer17
4.4	Remaining options (Optional)18
5	ACKNOWLEDGEMENTS20

1 REQUIREMENTS

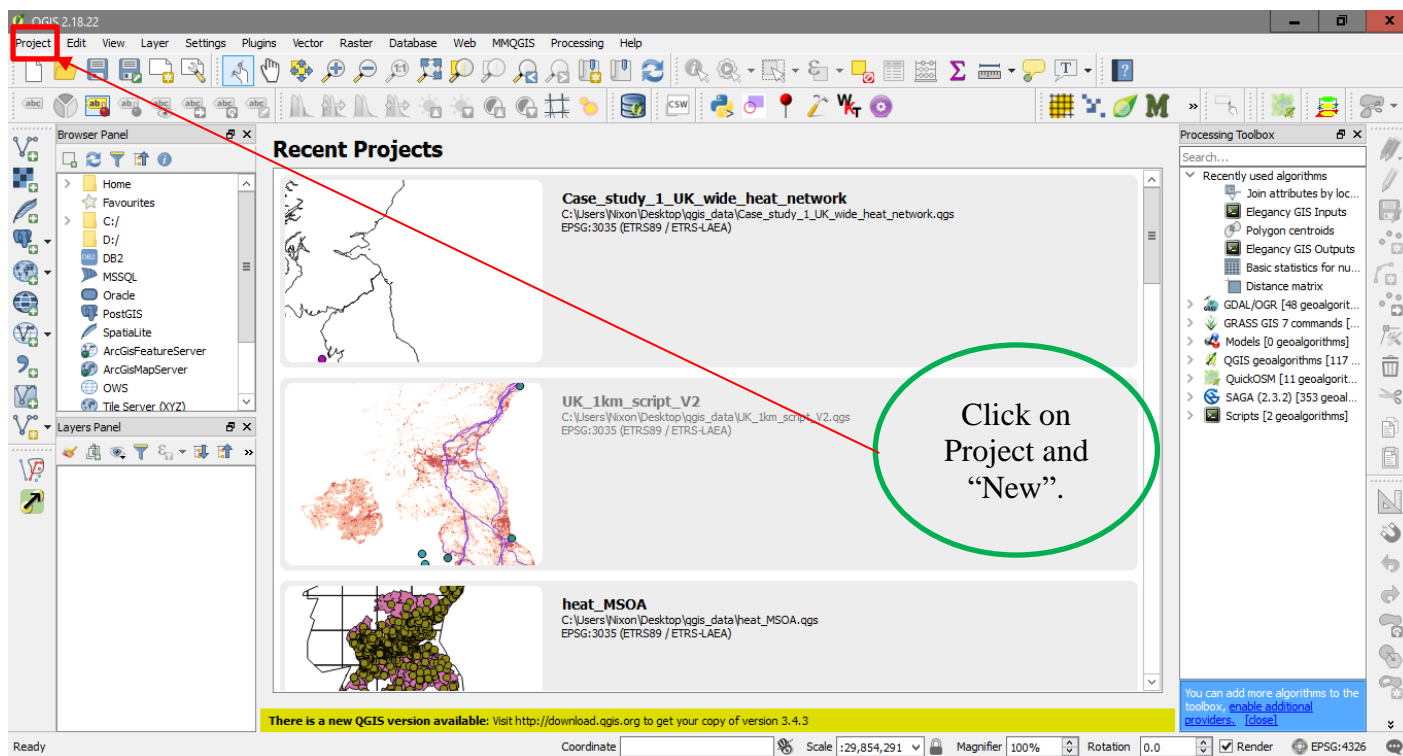
A few of the requirements for using the tools are detailed below:

- A working version of QGIS – version 2 is recommended as QGIS version 3 uses Python 3, where many of the inherent functions are not available and the scripts for the tools will require modification before use. The latest releases of QGIS can be downloaded for free via the following link: <https://qgis.org/en/site/forusers/download.html>. In particular, the core functions were tested using version 2.14.21. We recommend using the long-term release repository of version 2, which is listed under standalone installers.
- A case/ project where you have visually and spatially mapped the data. The input tool should only be applied after you have been able to project the data spatially and visualise it in a map format. Thus, you must have a set of spatially dependent data that you wish to visualise. You should have the various spatial datasets saved to the same Coordinate Reference System (CRS). Typical examples of such data: transportation/ heating demands, H₂ demands, CO₂/ H₂ geological storage sites, etc.
- Input tool should ideally be used when you have data available at high spatial resolutions. The tool works on the principle of data aggregation. Thus, it is not entirely relevant when you have data at relatively coarse resolutions which can be split into the respective “grids” (discussed later) using visual inspection alone.
- All spatial layers should be kept in the same directory/ folder so that the processing functions can easily retrieve them when running algorithms to modify the data. This is the case for both the input and the output tool.

2 QGIS PROJECT CREATION/ VISUALISATION

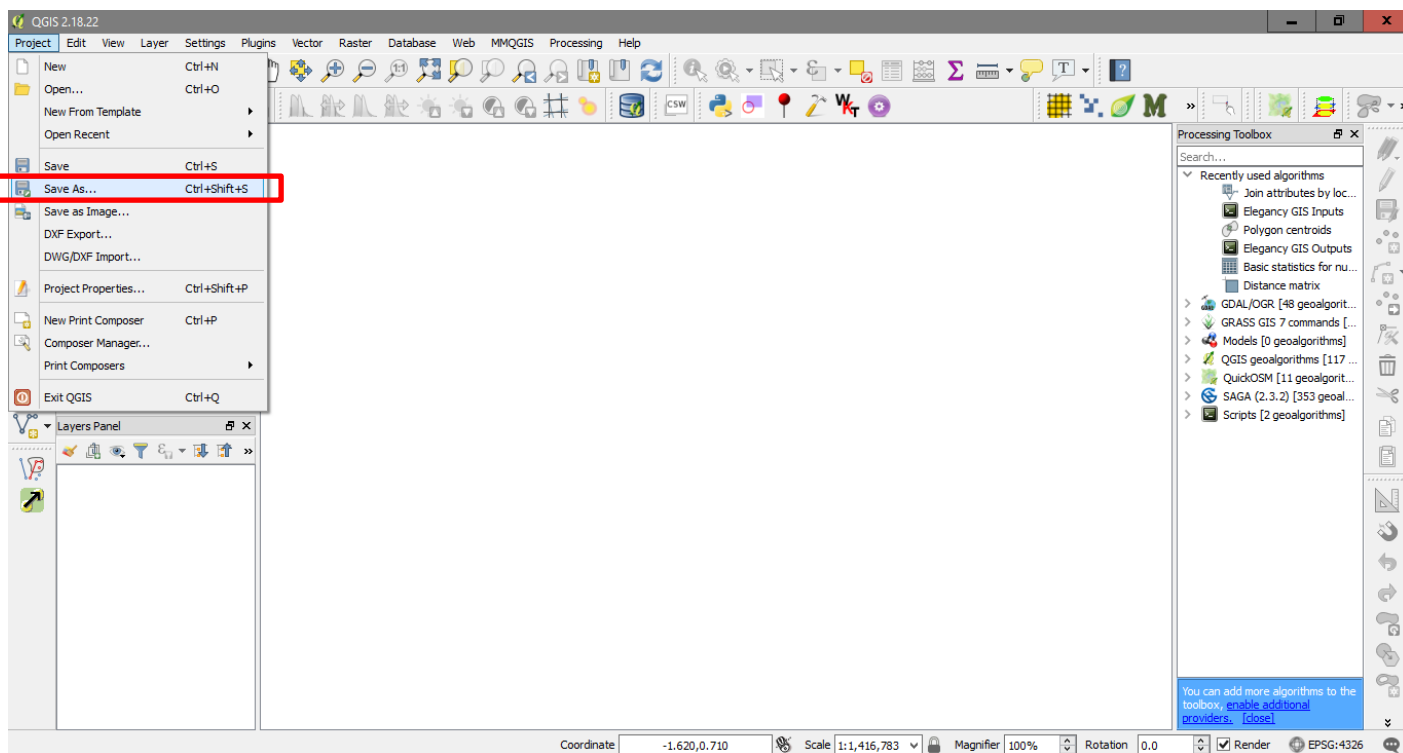
Upon installation of QGIS, a user should first create a new project and ensure that some useful plugins are installed and activated. Instructions on how to create a new project and install plugins are detailed here with further indications on visualising data sets.

2.1 Create a new project



If the user had recently installed QGIS and have never used it before, the recent projects section will remain blank.

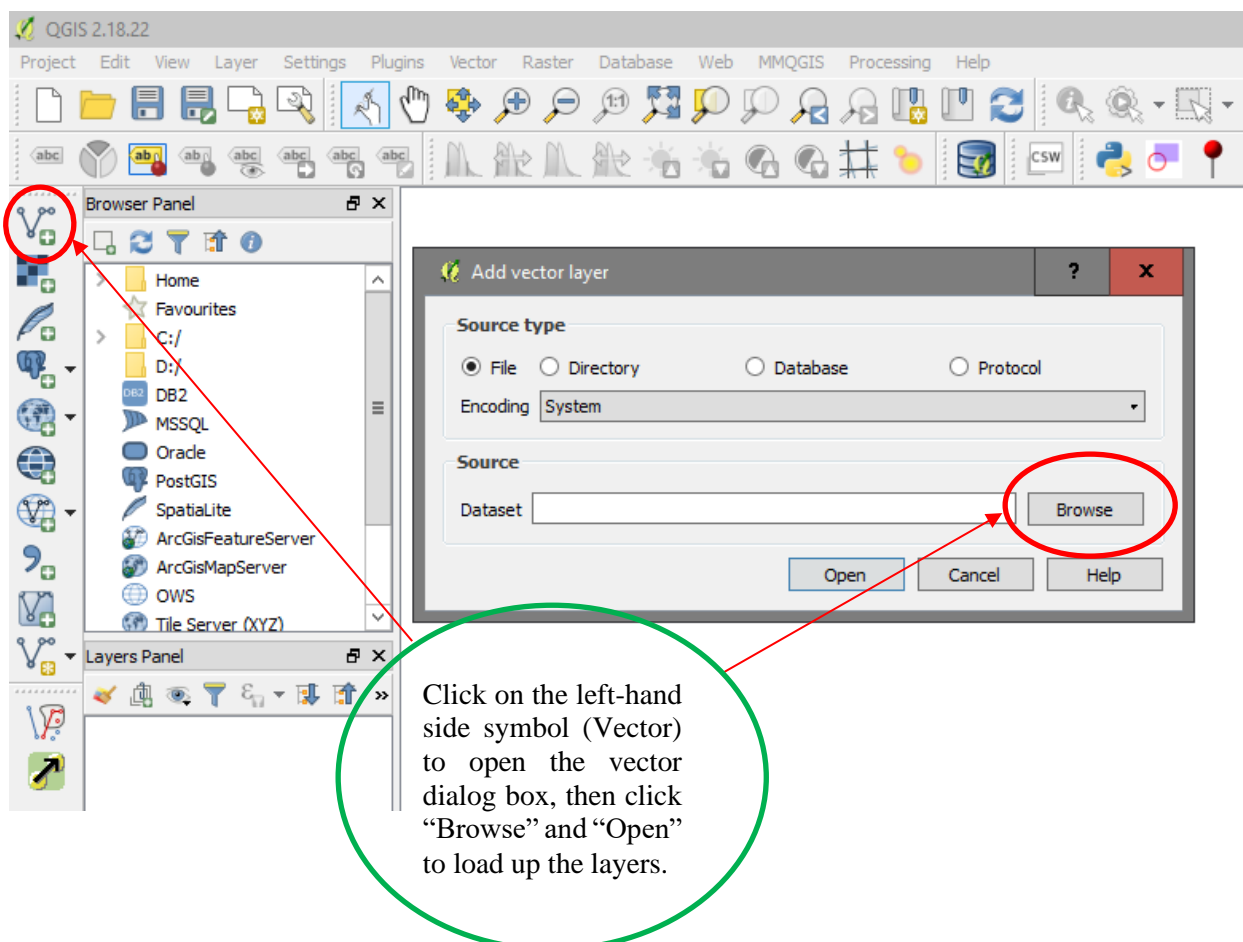
2.2 Save the new project



It is usually a good idea to save the project outside of the main program directory of QGIS. The path to the directory can be linked with QGIS to load and save any saved projects.

2.3 Load up the data layers

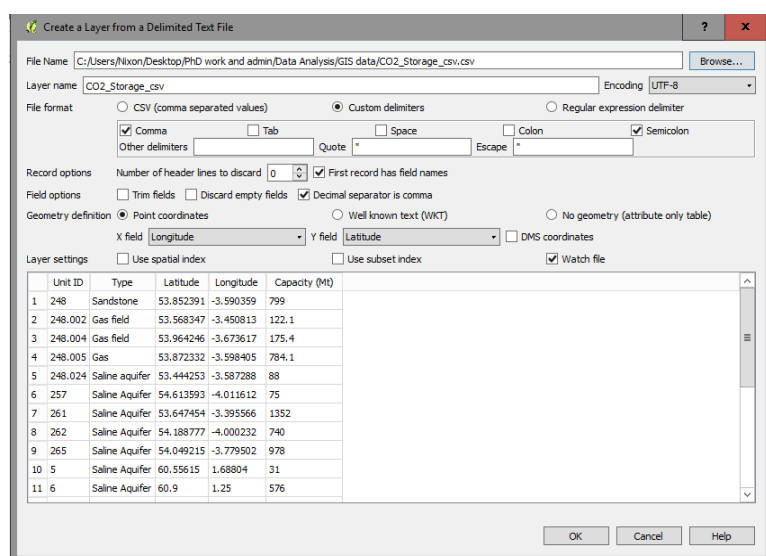
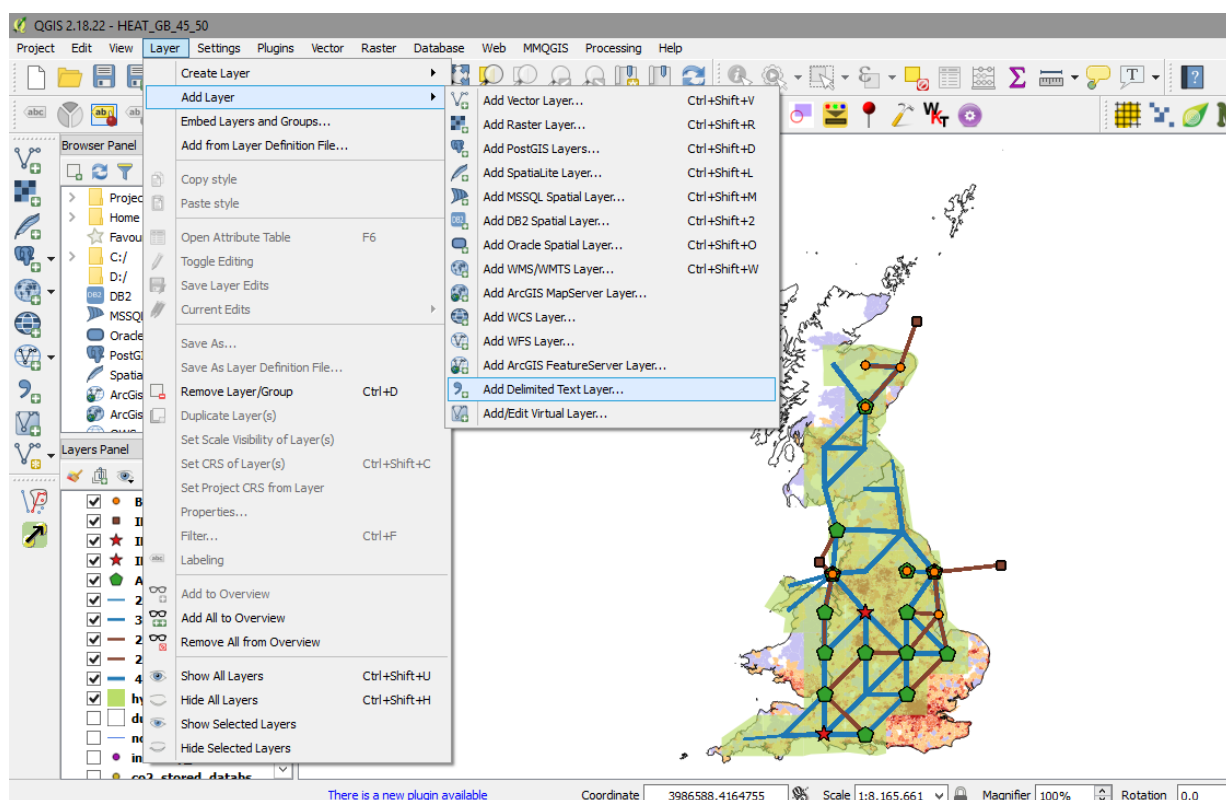
To load the layers, users must navigate to the directory with the shapefiles. As most of the data will be in a point or a line format, they are expected to be in the form of vector layers. The file endings for these types of data are .shp. Other data sets such as land use layers can be provided in a raster format as commonly done by data providers.



The file endings for raster data layers are .tif/ .tiff. To add raster layers, select the option under the vector layer which opens a raster dialog box from which a user can "Browse" and add the required layers in a similar manner to vector layers.

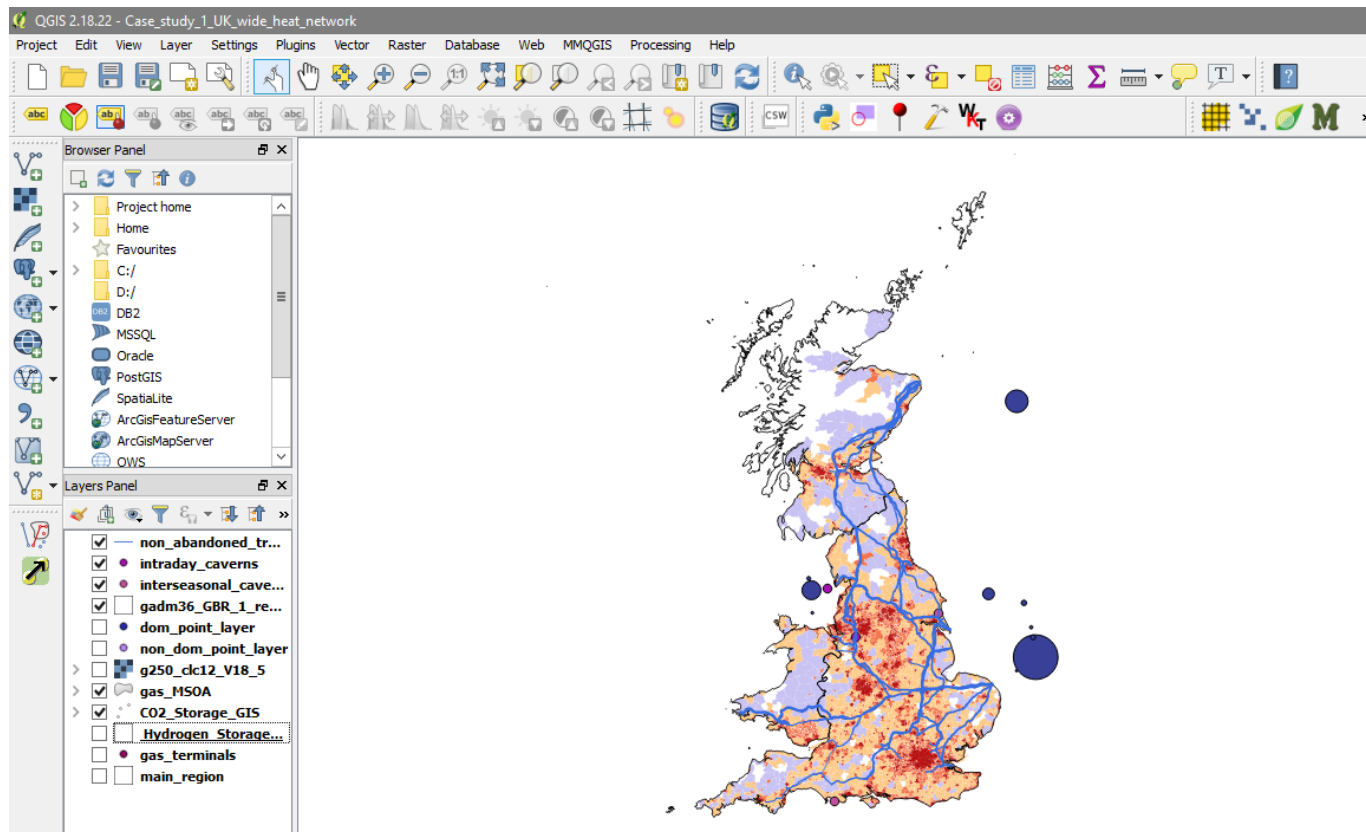
2.4 Loading up delimited text-based layers

To load vector layers on the basis of user-defined data, the user must have access to a .CSV file where at maximum you should have these features: unique identifier (numerical preferred), x-coordinate, y-coordinate, attribute values (numericals are essential). Alongside this information, users must also be aware of the coordinate projection that they are using for evaluating the x and y coordinates. This is essential for ensuring that the datasets are visualised in an accurate manner. You may add delimited text layers by navigating to the Layers tab, followed by “Add Layer”, followed by “Add delimited Text Layer” as shown in the image below.



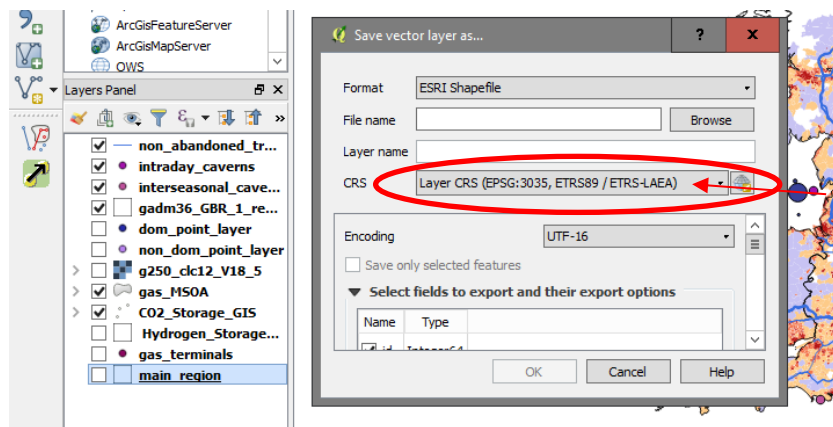
Please zoom in to the image on the left-hand side to see an example of a user-defined input file.

2.5 Saving layers using a common CRS



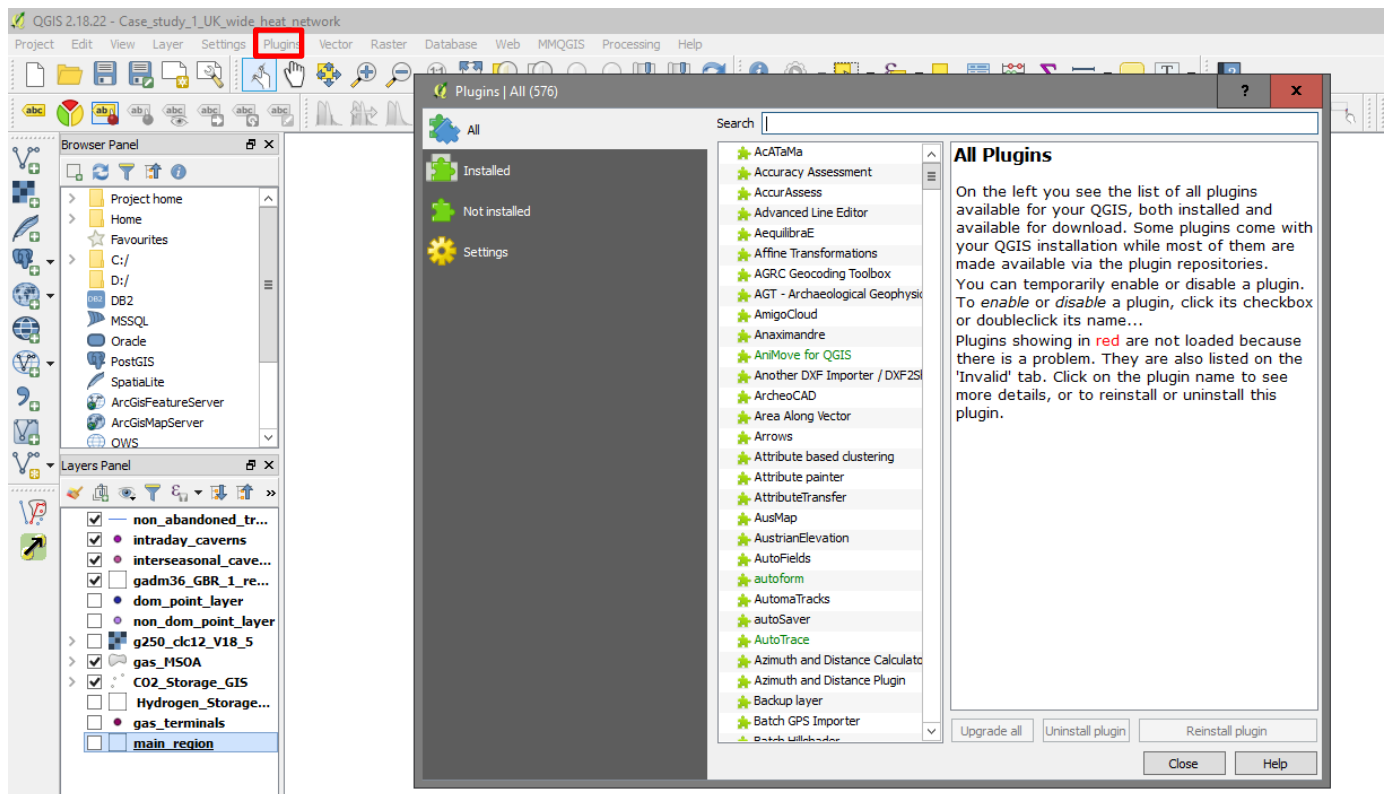
Once the data has been loaded, one should be able to visualise the data on the main page and the layers panel on the left hand-side should also be populated with the layers that were added.

Top tip: When loading layers, it is likely that there will be a mixture of coordinate reference systems depending on the data compiler that was used to generate the input data set. It is a good idea to convert the coordinate reference system (CRS) to a common CRS across all layers to avoid any potential issues when aggregating data using the input tool. Right clicking on the desired layer on the left-hand side and selecting “Save as”, allows any layer to be saved in a different CRS.

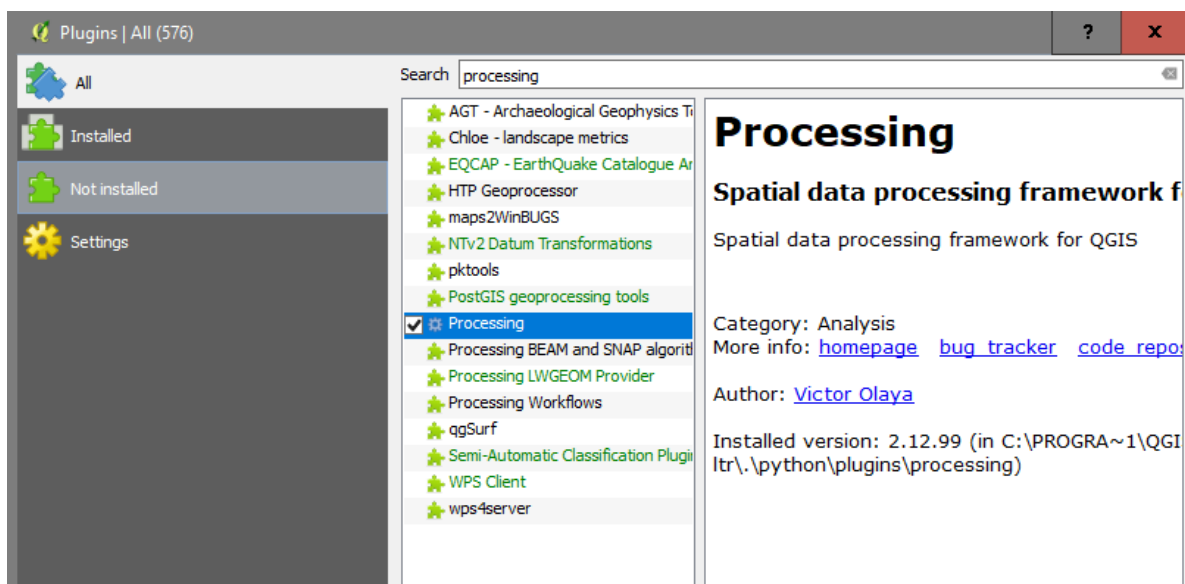


Select the CRS of interest and save the layer with a new filename with an “ESRI Shapefile” format.

2.6 Install processing toolbox

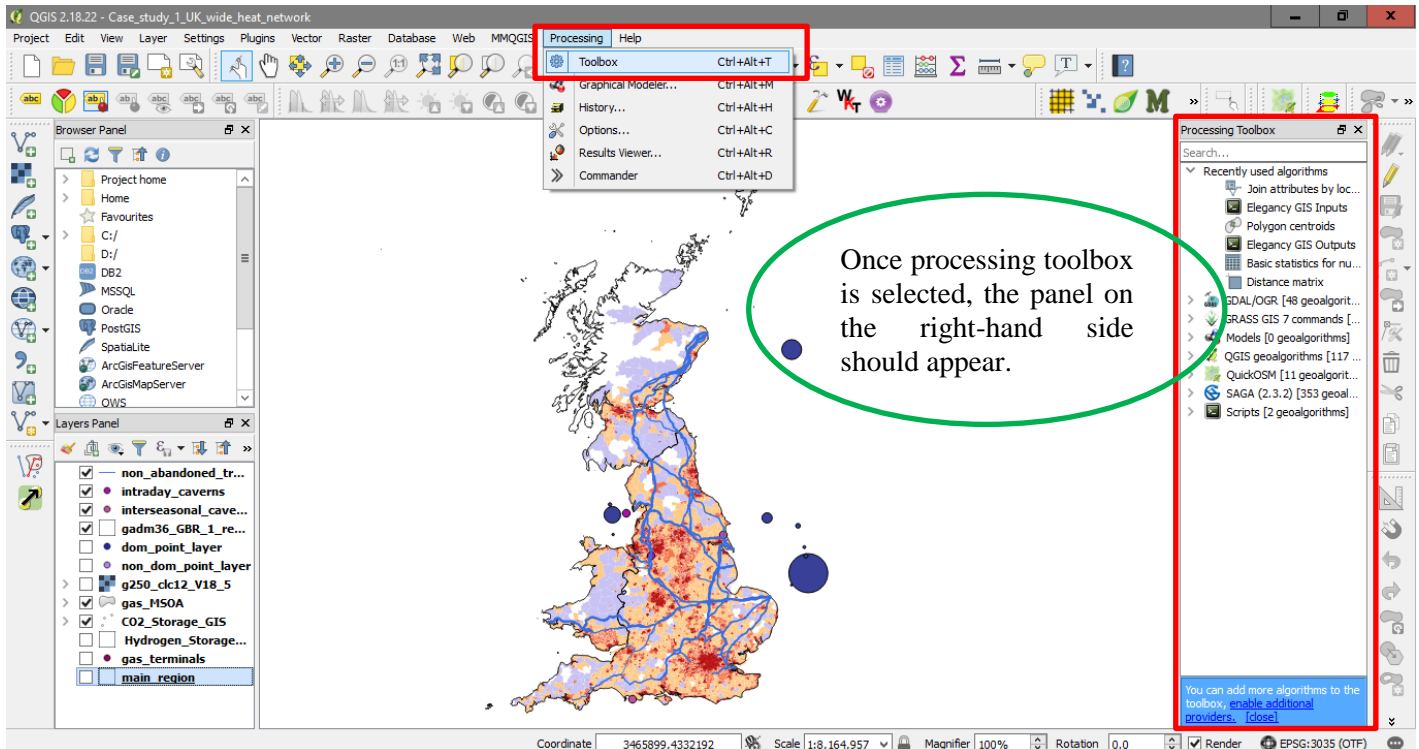


After clicking on the Plugins tab, select the “Manage and Install Plugins” option to load up the screenshot above. In the search bar, please type “processing” and ensure that the processing toolbox is installed and activated.

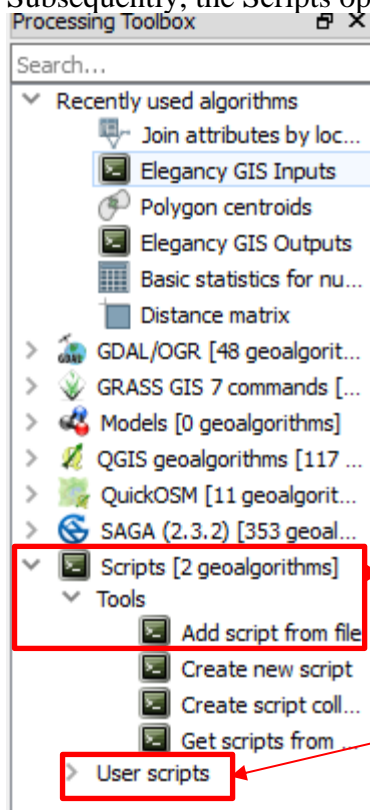


2.7 Add Input/ Output Tool script

For adding the script, the processing toolbox can be used.

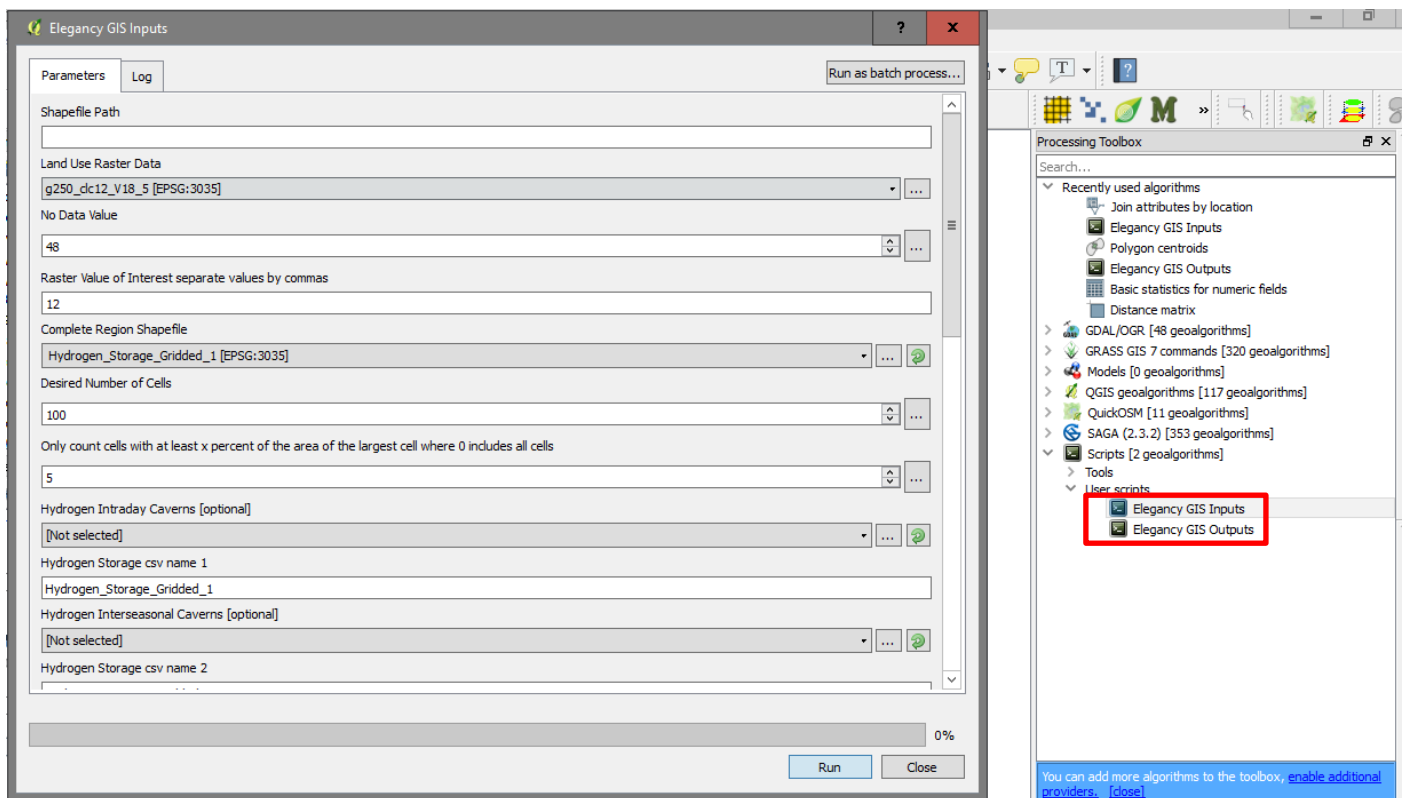


Subsequently, the Scripts option within the processing toolbox can be used to add any scripts.



Click on “Scripts” and then, “Tools”. Once all options are open, choose “Add script from file”. This will open a dialog box through which one has to “Open” the script. Upon opening the script, it should be listed under the User Scripts tab.

3 INPUT TOOL DETAILS



3.1 Shapefile path

This option requires a user entry which highlights the directory/ location of the spatial layers. As a reminder, it is assumed that all layers are stored within the same directory/ sub-directory. For example, a potential path may look similar to the following:

Example: C:\Users\JohnDoe\Desktop\Optimisation_Files

3.2 Land use raster

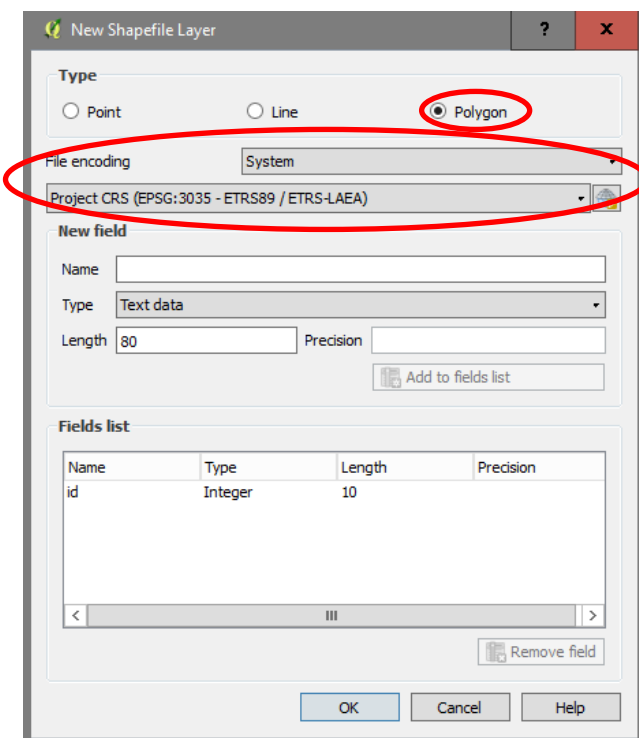
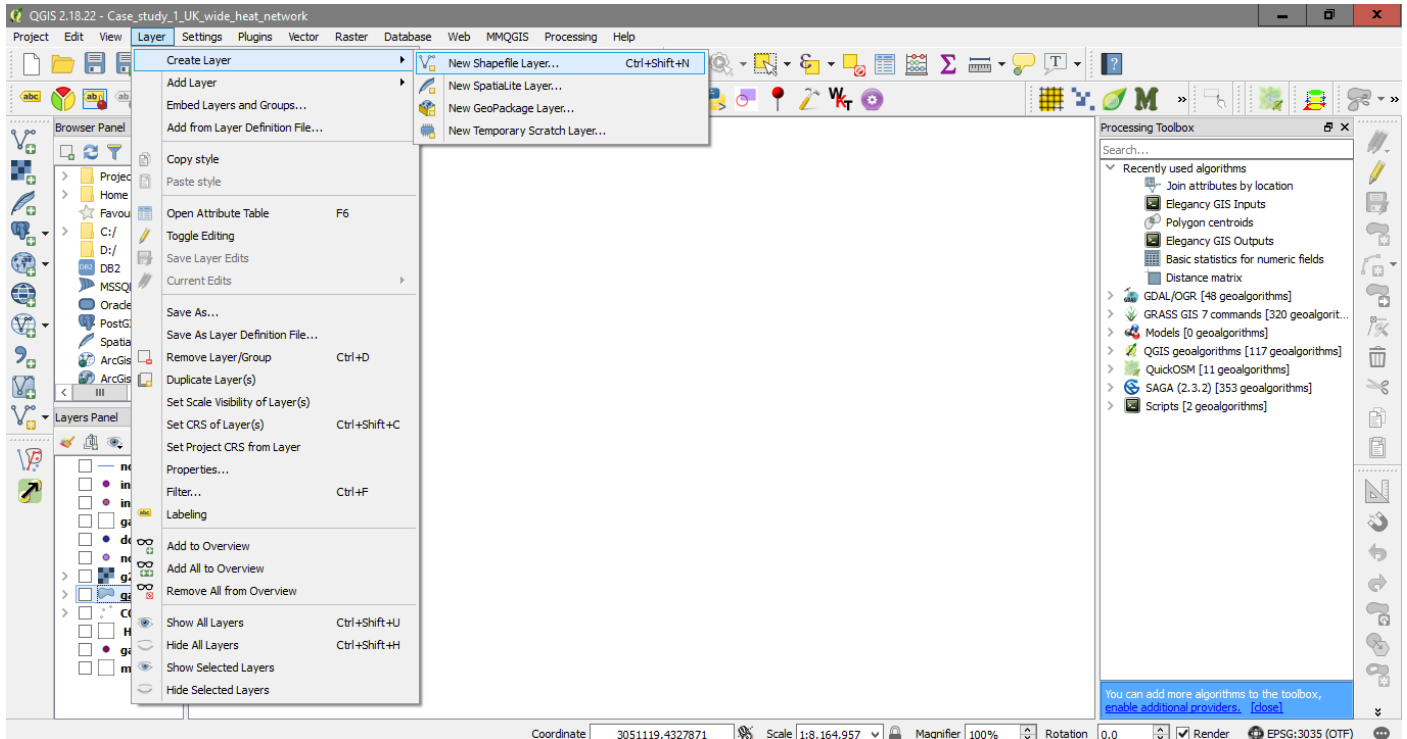
The next three inputs are associated with the land use raster. As all the case studies are within the EU geographical region, an EU land use raster is provided in tandem with the scripts. Once this layer is added onto the map, it can be selected under the “Land Use Raster Data” column.

The number of data values are defined by default and it is representative of the number of raster bands in the default land use raster database. Alternatively, if users prefer to use their own raster datasets, they are expected to enter the number of raster values present in the raster layer in this column.

Raster values of interest are used to generate an aggregate output of the areas of interest. Therefore, users should aim to utilise all the specific raster values corresponding to the areas of interest where the values are separated by commas. The raster values and the corresponding area types are provided in a text file along with the default land use raster.

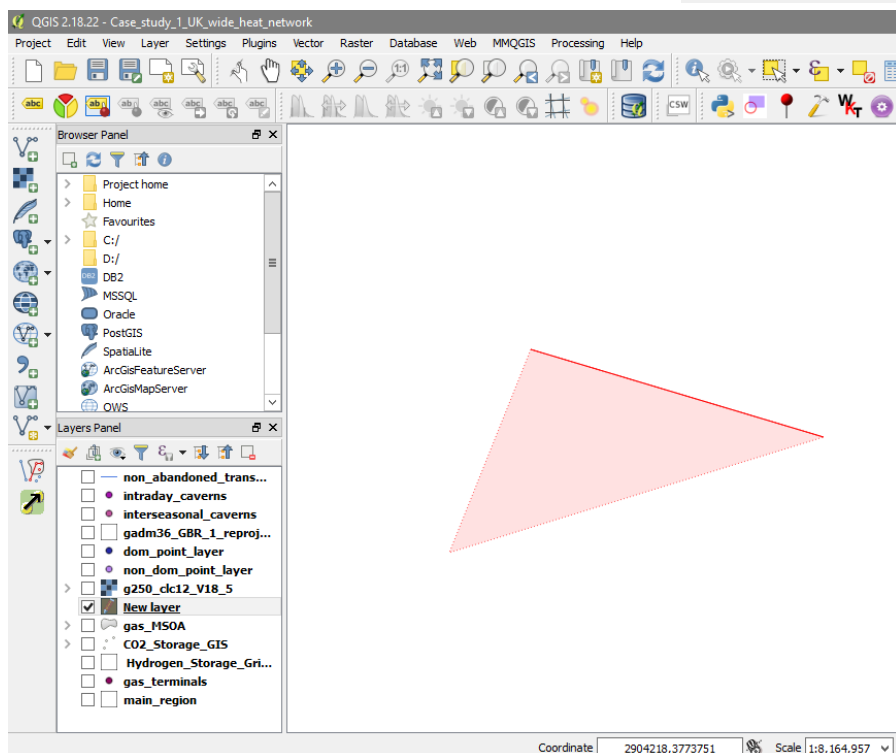
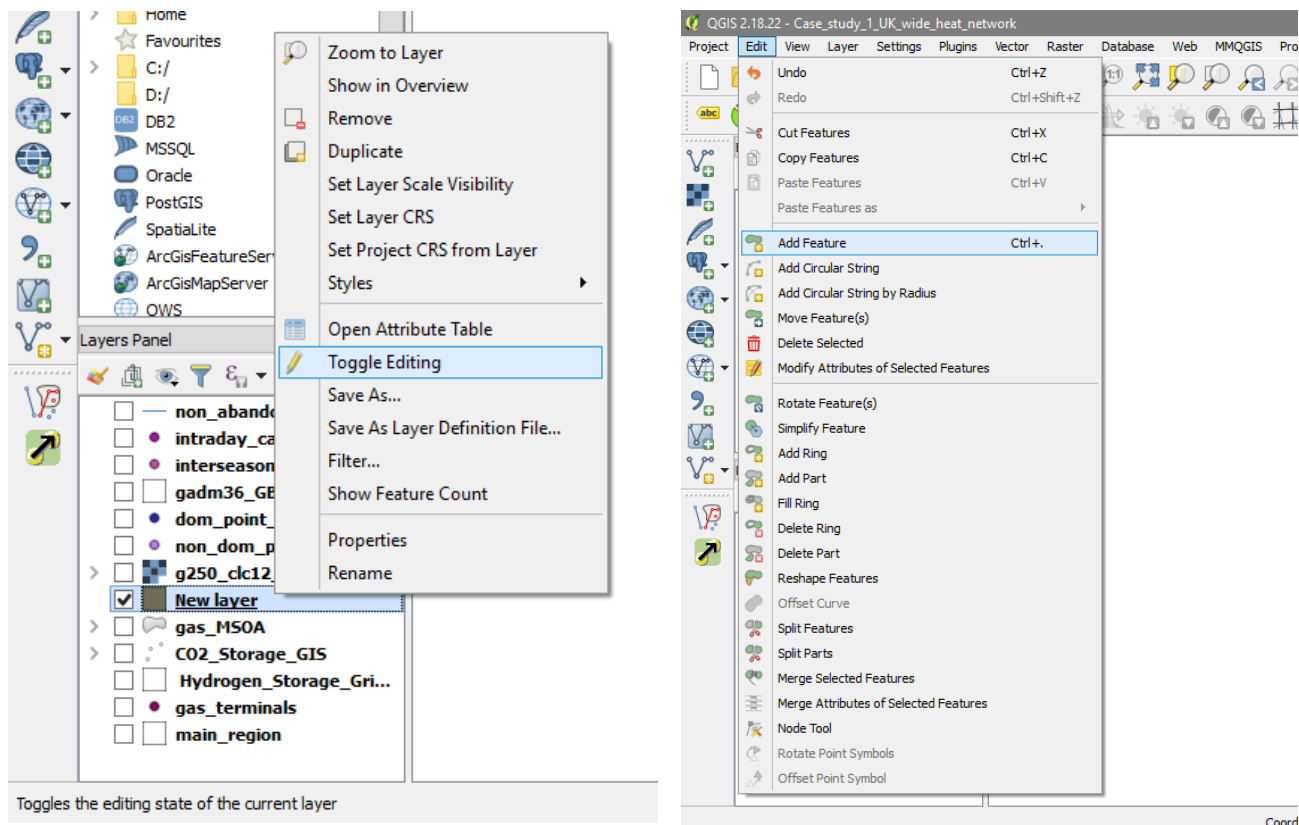
3.3 Complete region shapefile

In order to provide a region shapefile, users must create their own shapefile which overlays all the necessary features that will enable them to run the optimisation ensuring that they include everything that is necessary. This is a necessary pre-cursor for using the input tool.



Once the “New Shapefile layer” has been selected, the dialog box on the left-hand side opens. In this dialog box, choose “Polygon” as the Type with System encoding, and the common CRS used specific to the project. Then, press “OK”, users are now asked to name this layer and save it. Please make sure that this layer is saved in the same directory as the other input shapefiles.

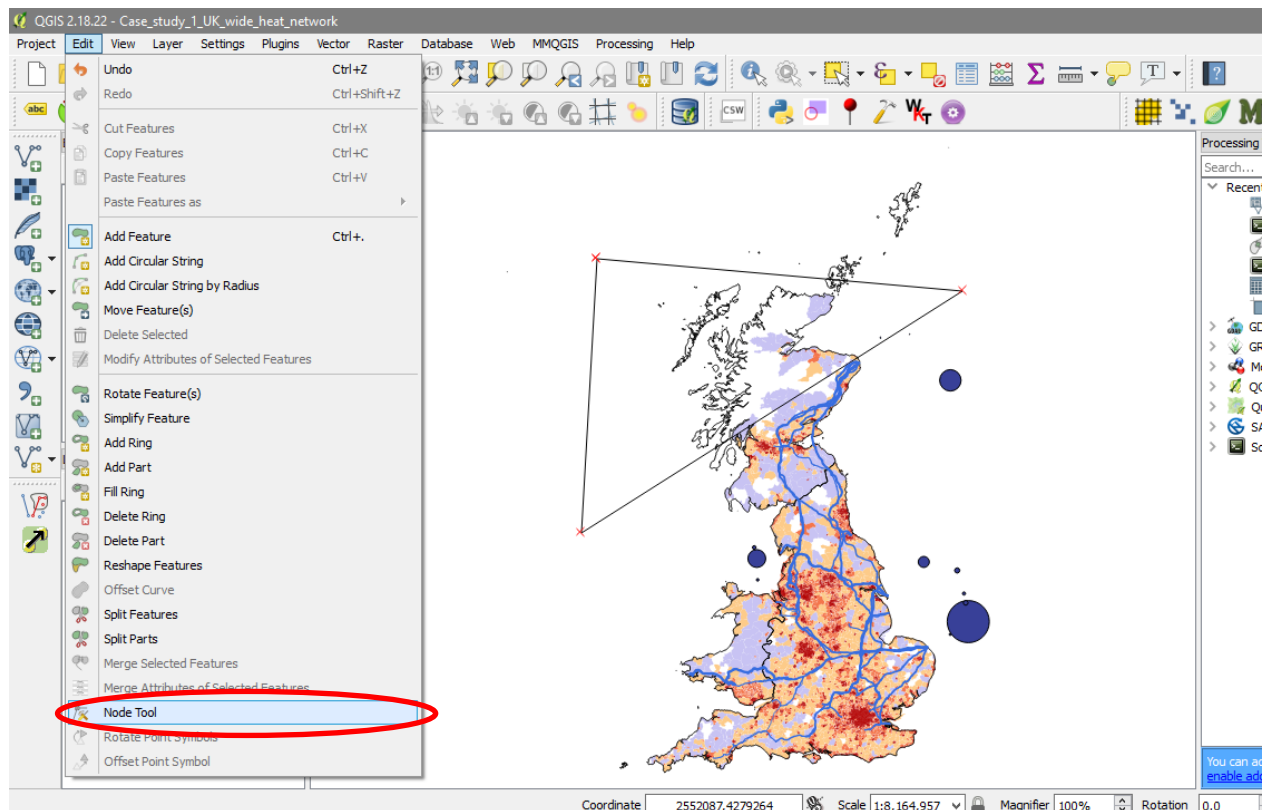
This layer is now available in the layers panel, but it will not be visible on the map as there are no features in the layer. Features can be added by right-clicking on the layer in the layers panel and selecting “Toggle editing”. Then, click on “Edit” and then select “Add Feature”.



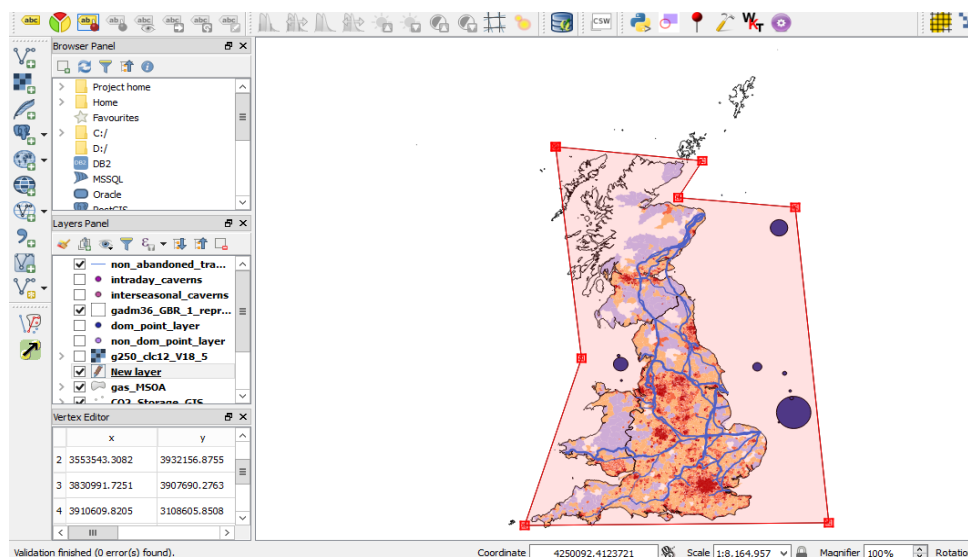
Once the “Add feature” is selected, users are expected to draw a polygon. Users may initially draw a triangle as shown here by left-clicking at three different locations on the map and then right-clicking. This process opens a dialog box where users are prompted to enter the feature id. Simply enter 1.

Once users begin to display the other layers, they will note that the triangle is not an accurate nor a good representation of the

spatial region for case studies. Luckily, they will find that they are not limited to regions with three vertices in QGIS. QGIS contains a tool called “Node Tool” which can be utilised in order to modify the polygon layer which has recently been created.



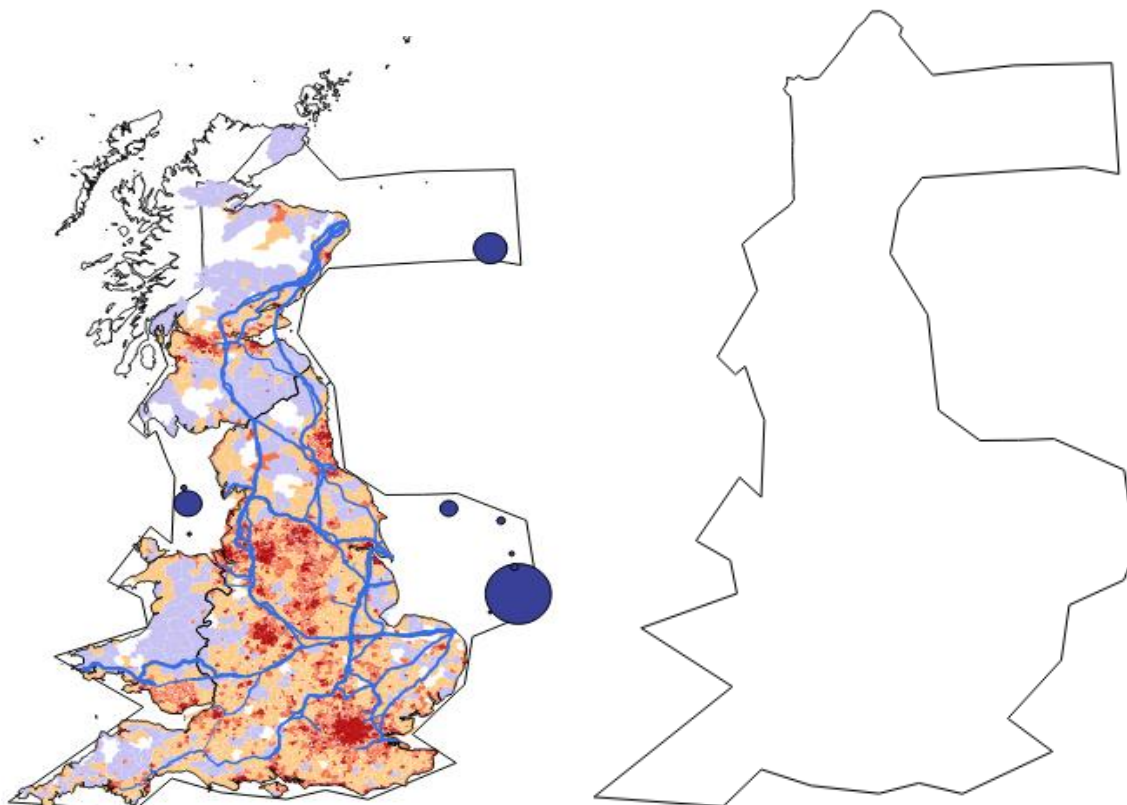
Upon selecting “Node Tool”, click on the triangle and the polygon will become shaded. Double clicking on any of the boundary lines will create a new node which can then be used to modify the spatial layout of the newly created layer. It is possible to create as many nodes as preferred to ensure that the layer captures all of the necessary information.



Once users are comfortable that the layer overlays enough of the spatial region, right-click on the layer once again and click “Toggle editing”, at which point they are prompted to save the edits.

ACT ELEGANCY, Project No 271498, has received funding from DETEC (CH), BMWi (DE), RVO (NL), Gassnova (NO), BEIS (UK), Gassco, Equinor and Total, and is cofunded by the European Commission under the Horizon 2020 programme, ACT Grant Agreement No 691712.

Please save these modifications and ensure that the region shapefile covers only the areas that are necessary. The key aspects to cover are CO₂ sequestration sites, demand areas, etc. An example of the final region file for an example considering the UK is shown below.



Here, the complete region shapefile is the region wrapping around the other areas of interest. This is subsequently used in the input tool as one of the key layers needed for analysis.

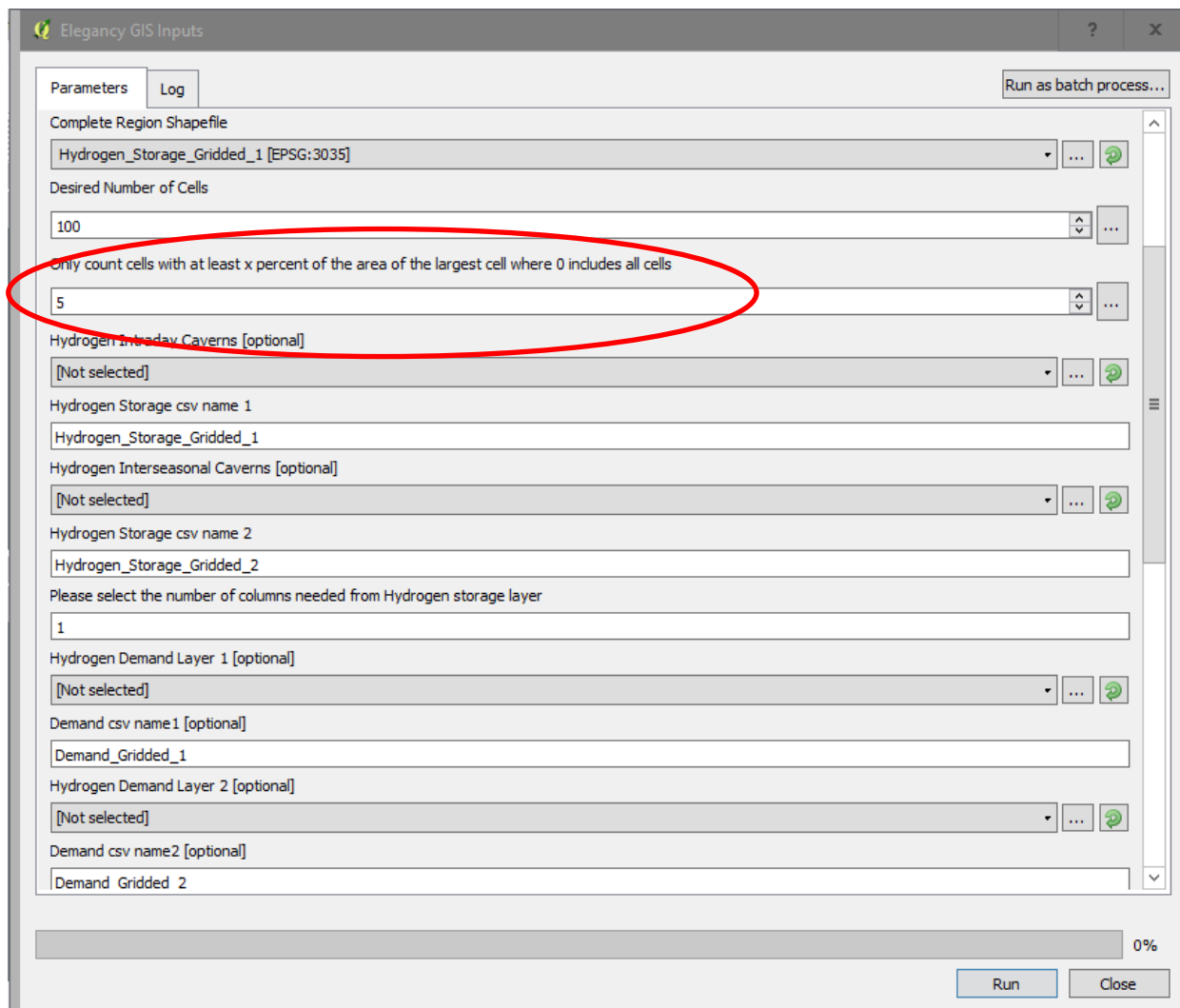
3.4 Desired Number of Cells

This option prompts the user to enter the required number of grid cells that they intend to use for optimisation. It should be noted here that there is a trade-off between accuracy and computational tractability in choosing the right number of cells. A large number of grid cells (i.e., higher spatial resolution) ensures a higher degree of accuracy in representing the data, provided that the data itself is of very high resolution. However, this increases the computational complexities associated with solving any instance of the optimisation model and often renders large numerical problems unable to solve. Thus, an appropriate number of grid cells must be chosen with a compromise between the accuracy and computational resources at hand. As a first estimate, 70-100 grid cells may be sufficient to represent large scale network design problem with reasonable accuracy.

3.5 Smallest grids

When splitting the complete region shapefile into the desired number of grids, some grids may be very small and constitute of very little area. For the purposes of optimisation, they may not be particularly important, and a user may choose to omit these small grids to improve the computational speed of the problem instance as mentioned above. There is an option in the input tool which enables a user to state the size of the cells that they choose to ignore. As the absolute area of any grid cell is dependent on the case study and the spatial region of interest, the user is

not asked to specify a specific area to remove but rather state the relative area fraction of the smallest possible grid cell when compared with the standard/ largest grid cells.



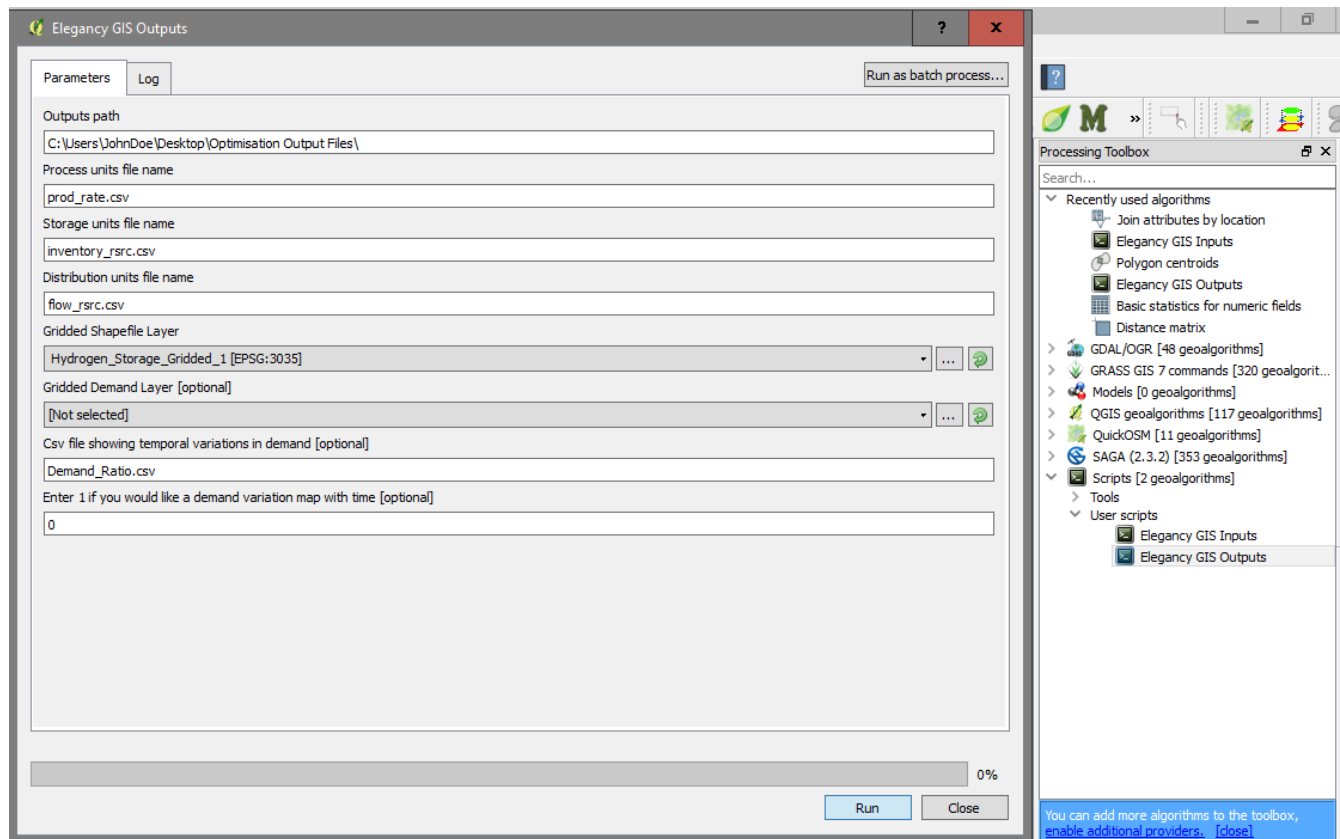
As a default value, the area proportion of the smallest grid in comparison to the standard sized grids is assumed to be 5%. This implies that any grids which contain an aggregate area which is less than 5% of the nominal grid cell area will be omitted from the output.

3.6 Subsequent options

The remaining options are self-explanatory and many of these are optional. This includes the H₂ storage layer, demand layer, CO₂ storage layer, natural gas transmission layer, etc. Alongside this, users are asked to provide the CSV name for these inputs and they may choose to delete the default names and input their own labels. After all the inputs have been entered, click “Run” and it may take a few moments/ minutes to run the script depending on the size of the data/ number of cells, etc. Following which, users should see multiple new layers in the original directory along with comma delimited text files. These text files are used as the inputs for the optimisation model. Please note that visually several temporary layers will be created and added to the layers panel, and they will be lost if you close the session without explicitly saving them. The input data is separated as text-based documents to enable edits to the data separately from the model. The nature

of the data files to use in these shapefile layers along with the type of data that are extracted are shown through examples that are contained in the .zip file accompanying this guide. Users must adhere to this data format to ensure that they do not have any unforeseen difficulties in processing data through the tools. Alternatively, they could modify the scripts themselves in order to process different data types. It is important to convert the data format to numerical to enable the functions to aggregate the contents of the shapefile layers. The input tool will not convert the format of the data into a numerical for subsequent aggregation.

4 OUTPUT TOOL DETAILS



4.1 Outputs path

Similarly, to the input tool, the output tool requires a path to identify the data files that must be processed for visualising results. The only difference is that the path is followed by a trailing backslash (\) in the output tool as opposed to the input tool where this is not present.

4.2 File names

The next three options require the user to provide the names of the CSV files that are to be used to extract the data and create the different visualisation layers. If users did not modify the CSV file names of the outputs from GAMS/ Pyomo, then they should keep the default names that are shown in the figure above.

4.3 Gridded shapefile layer

This prompts the user to provide a layer which contains the gridded cells. The purpose of this layer is to provide the coordinate references for the output layers by using the features within the gridded layers to generate point coordinates. The input tool should have generated gridded shapefiles along with the CSVs that were used for optimisation. Thus, inputting any of these generated gridded layers as the “Gridded shapefile layer” would suffice. The remaining prompts are optional and they are principally concerned with generating an animated map of demand variations with time. Thus, these options can be left empty if this is not of primary concern.

4.4 Remaining options (Optional)

A user can optionally input a demand layer that was created with the input tool, to generate a modified demand layer which incorporates temporal variations along with the spatial distribution of demand. This is achieved through the usage of a CSV file which details the temporal variations at each minor time period during the optimisation process. An example data file is provided in the .zip file with the title “Demand_Ratio.csv” as a reference. Following this, simply clicking “Run” creates new layers for each output technology and these layers can be visualised to show results.

A quick note on the data files: please make sure that the relevant metrics and values of interest associated with a particular spatial entity is listed in the final column of the layer attributes. As an example, please see the figures below. In the first figure, only the final attribute will be aggregated and provided in the input gridded layers. This corresponds to the demand for H₂ in a particular location inherent within the spatial layer.

dom_point_layer :: Features total: 8366, filtered: 8366, selected: 0

	geo_code	geo_label	geo_labelw	label	name	domestic_M	domestic_C
1	E02002873	Nottingham 006		E92000001E060...	Nottingham 006	E02002873	39805317.00000
2	E02000793	Richmond upon T...		E92000001E090...	Richmond upon T...	E02000793	45662476.00000
3	E02002661	Kingston upon H...		E92000001E060...	Kingston upon H...	E02002661	39683996.00000
4	E02000077	Bexley 013		E92000001E090...	Bexley 013	E02000077	39421563.00000
5	E02000818	Southwark 012		E92000001E090...	Southwark 012	E02000818	24279076.00000
6	E02004696	Basingstoke and ...		E92000001E070...	Basingstoke and ...	E02004696	26004761.00000
7	E02001910	Birmingham 084		E92000001E080...	Birmingham 084	E02001910	29487014.00000
8	E02006656	Wiltshire 014		E92000001E060...	Wiltshire 014	E02006656	29230047.00000
9	E02005761	Harrogate 001		E92000001E070...	Harrogate 001	E02005761	8843097.00000
10	E02002412	Leeds 083		E92000001E080...	Leeds 083	E02002412	37334176.00000
11	E02000724	Newham 011		E92000001E090...	Newham 011	E02000724	33909238.00000
12	E02002647	Blackpool 015		E92000001E060...	Blackpool 015	E02002647	41227445.00000
13	E02003113	South Gloucester...		E92000001E060...	South Gloucester...	E02003113	27944262.00000
14	E02001245	Tameside 017		E92000001E080...	Tameside 017	E02001245	39835752.00000
15	E02005933	Cherwell 013		E92000001E070...	Cherwell 013	E02005933	35806462.00000
16	E02004817	Test Valley 004		E92000001E070...	Test Valley 004	E02004817	35836120.00000
17	E02006147	Lichfield 002		E92000001E070...	Lichfield 002	E02006147	46075577.00000
18	E02003086	North Somerset ...		E92000001E060...	North Somerset ...	E02003086	33564675.00000
19	E02002844	Leicester 018		E92000001E060...	Leicester 018	E02002844	28607234.00000
20	E02003928	Cornwall 071		E92000001E410...	Cornwall 071	E02003928	41428250.00000
21	W02000174	Swansea 007		W92000004W06...	Swansea 007	W02000174	51552896.00000

Show All Features

CO2_Storage_GIS :: Features total: 12, filtered: 12, selected: 0

	Location	X-coordina	Y-coordina	Type	Capacity_j
1	Bunter 36	3805797	3487236	Yes	252.0000000000...
2	Viking fields	3819016	3442621	Yes	155.0000000000...
3	Captain X	3657078	3938349	Yes	60.0000000000...
4	Forties 5	3792577	3855728	Yes	1021.000000000...
5	Hamilton	3418787	3469051	Yes	129.0000000000...
6	Goldeneye	3713261	3944959	Yes	37.0000000000...
7	Hewett	3792236	3363295	Yes	110.0000000000...
8	Endurance	3741010	3503752	Yes	530.0000000000...
9	Morecambe North	3412177	3531843	Yes	180.0000000000...
10	Morecambe South	3417134	3510361	Yes	850.0000000000...
11	Bunter Closure 9	3827278	3388091	Yes	2000.000000000...
12	Bunter Closure 3	3822321	3424444	Yes	230.0000000000...

Similarly, only the final attribute, which in this case corresponds to geological storage capacity in Mt for CO₂, is aggregated and reported. It is important to note that the metric of interest must be listed under the last column in the attribute table in a integer/ floating point numeric format as opposed to a string.

5 ACKNOWLEDGEMENTS

The authors would like to greatly acknowledge the support received from Dr. Nathaniel Cooper of Department of Chemical Engineering, Imperial College London for the development of the GIS input tool. In particular, the input tool uses some of the same inherent functions (spatial mesh creation, etc.) as those found in the “Resource Mapping Tool” which can be found via the following link on GitHub: <https://github.com/nathanielcooper/Resource-Mapping-Tool>. The source code for the “Resource Mapping Tool” was used where applicable and discarded where irrelevant for the goals of this project. This reduced version of the source code was used as the backbone for the development of the preliminary version of the input tool for this project.