A brief introduction to QGIS

Ivan Viveros Santos

2019-09-23

Contents

Chapter 1

Introduction

QGIS is an Open Source Geographic Information System (GIS) licensed under the GNU General Public License.

The aim of this document is to present some recipes about importing and exporting vectors, and treatment and exploration of spatial data. I plan to add more recipes on vector styling and map composition.

This is a brief introduction to QGIS, as such, it does not replace in any maner the QGIS User Guide. I also recommend you visiting the StackExchange chanel, where you can get support on questions related to QGIS.

The data used to illustrate the recipes presented in this document come from the Portail données ouvertes Montréal; particularly, the selected data from this site is under the license Creative Commons Attribution 4.0 International.

To follow the recipes presented in this document, please download the accompanying files from my GitHub page.

Chapter 2

QGIS installation and Graphical User Interface

2.1 QGIS installation

It is possible to install QGIS on Windows, Mac OS X, Linux, BSD, and Android operating systems. The installers can be downloaded from this website, and this link provides detailed installation instructions.

By the time of writing this document, the last version of QGIS is `3.8.2 Zanzibar`, which was released on October 16, 2019.

2.2 QGIS Graphical User Interface

The QGIS graphical user interface (GUI) is composed of:

- **Menu bar**: gives access to the main functionalities of QGIS.
- **Toolbars**: give a quick access to QGIS functionalities.
- **Panels**: they provide several functionalities, for instance managing layers, and browsing spatial data.
- **Map display**: shows the spatial data of the current project.

It is possible to customize the appearance of QGIS by navigating to **View** from the **Menu bar**. From here, you can select the panels and toolbars you want to display. I would recommend toggling the **Processing Toolbox** from the panel because it allows to quickly access to the main operations available in QGIS.

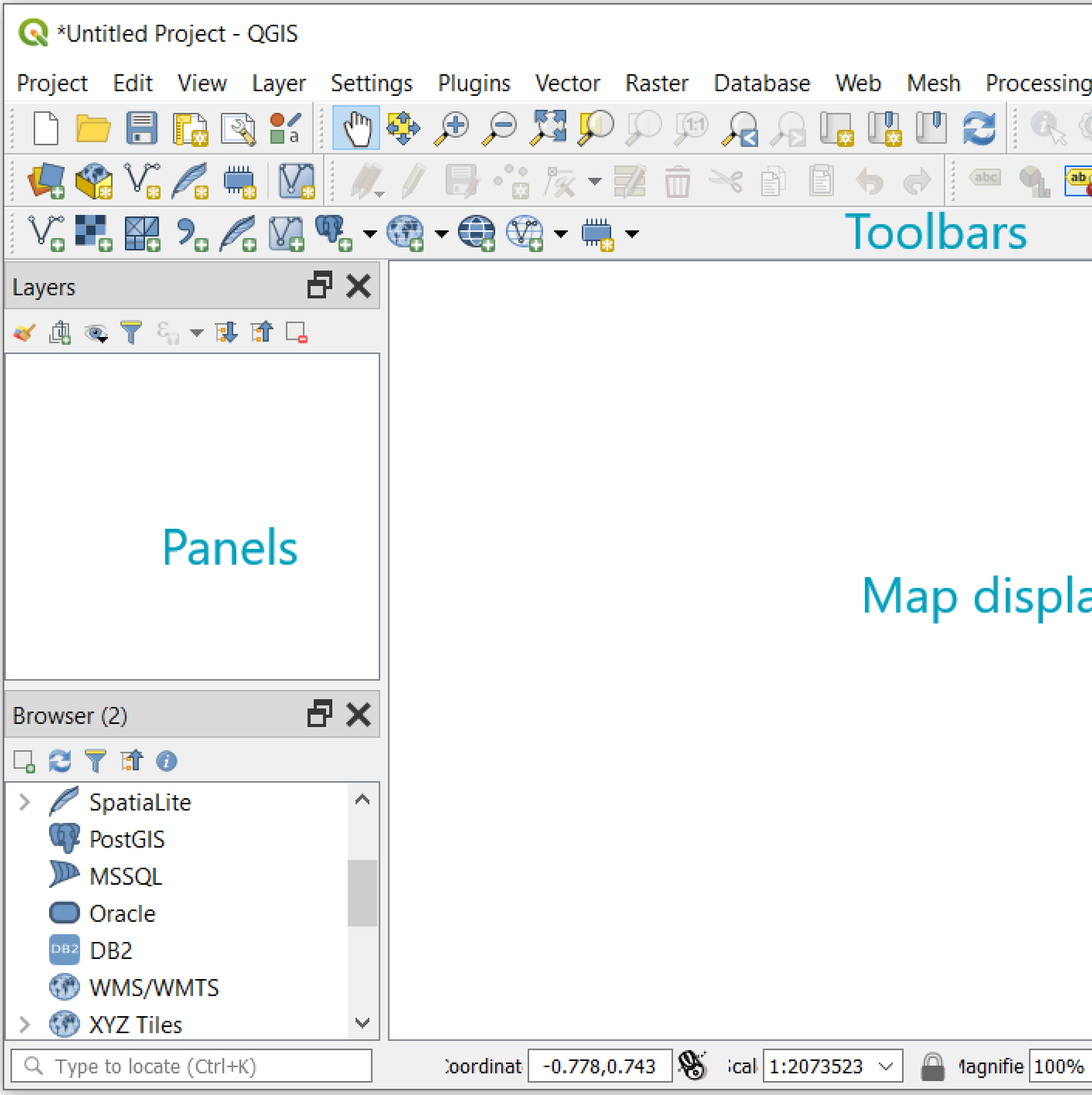

Figure 2.1: QGIS GUI

Chapter 3

Loading and Exporting data

The quickest way to load data into QGIS is the technique drag-and-drop. It only requires locating the data you want to import into QGIS, select the file, and drop it on the **Map display** or on the **layer panel**.

Another option for importing data is to use the browser panel. Navigate through the browser, locate the layer or data to be imported, right-click on it, and finally select Add Layer to Project. However, using the **Manage Layers Toolbar** or **Add Layer** from the **Layer** menu give more control on the importation.

3.1 Importing data from text files

This recipe is illustrated with the `coordonnees-stations-rsqa.csv` file, which reports the locations of the stations of the Air Quality Monitoring Network (RSQA) set up in Montreal. The aim of the RSQA is to monitor the atmospheric concentration of criteria pollutants.

1. Navigate to the **Layer** menu and select **Add delimited text layer**. The following dialog will pop up:
2. In the **File Name** field, indicate the path to the `coordonnees-stations-rsqa.csv` file.
3. In the **Geometry Definition** section, indicate that point coordinates are being imported. Then indicate the corresponding fields for longitude and latitude. According to the sources of the layer being imported, the coordinate system is WGS 84, so we don't need to change it. When you want to import an attribute table, you can select the option **No geometry**.
4. Click on **Add**. You will see the following set of points, probably not in the same colour. Optionally, you can add a base map or another layer to give some spatial context, but for now we have successfully imported the layer corresponding to the stations of the RSQA network.

   In section 6.2 we describe how to add a basemap to add spatial context.
5. **Optional step**: save the imported layer as a new shapefile (Section 3.4).

3.2 Importing data from KML files

QGIS supports importing Keyhole Markup Language (KML) files. KML is the format used by Google Earth to show spatial data.

The importation of a KML is illustrated with the `grandparcs.kml` file which contains polygons corresponding to the parks of Montreal.

1. Navigate to the **Layer** menu and select **Add vector layer**. The following dialog will pop up:

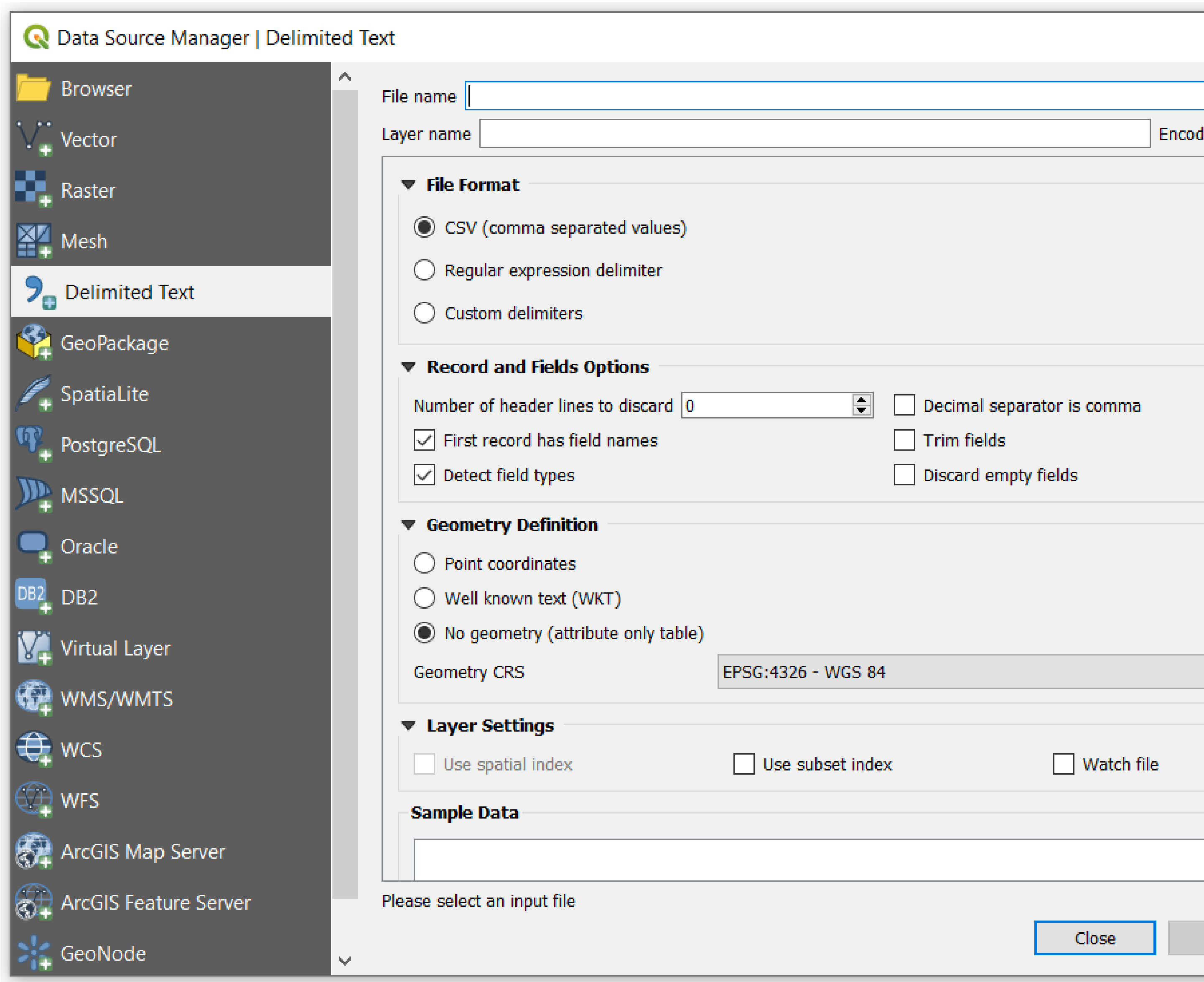

Figure 3.1: Add delimited text layer window

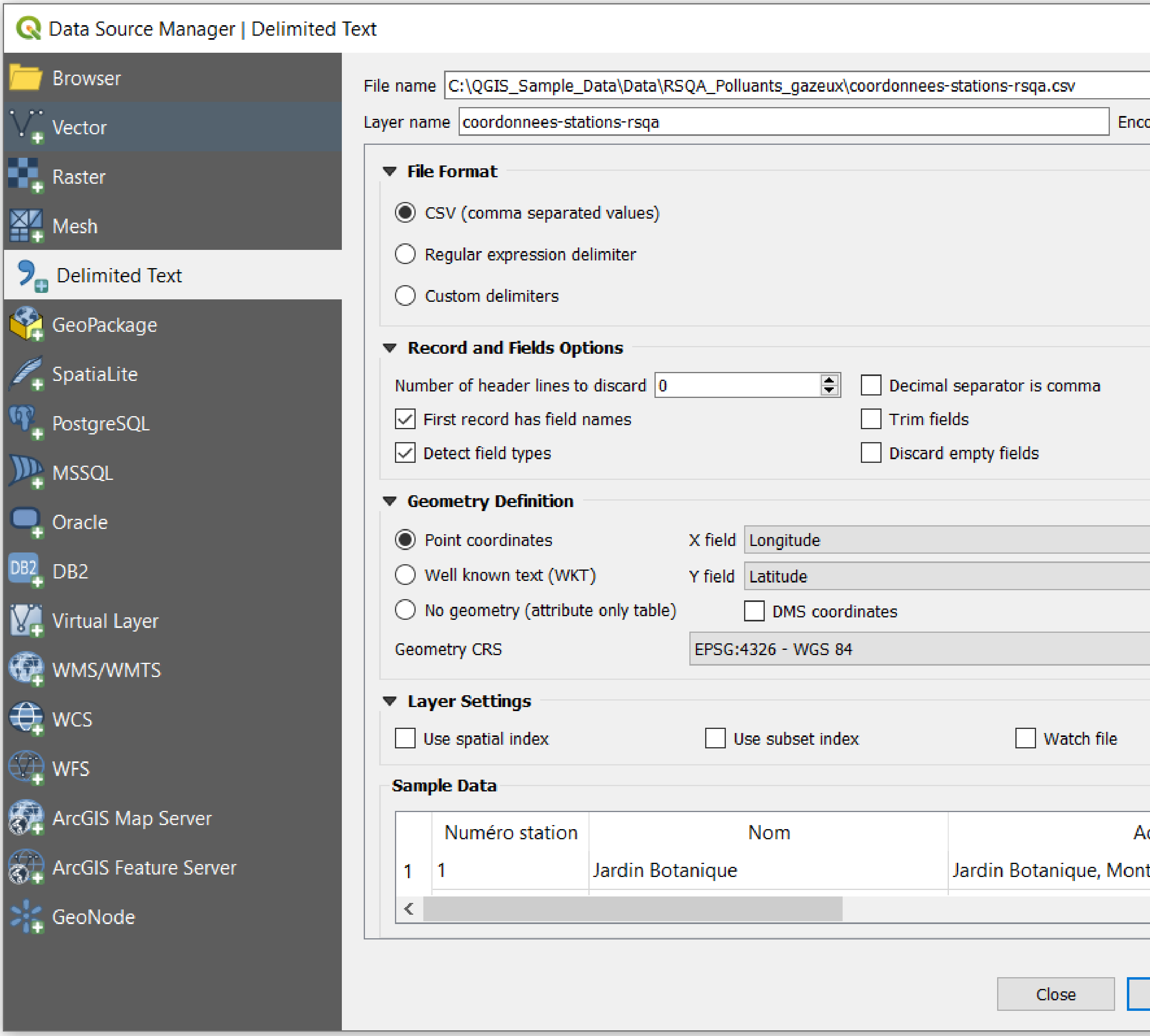

Figure 3.2: Indicating the path to the text file

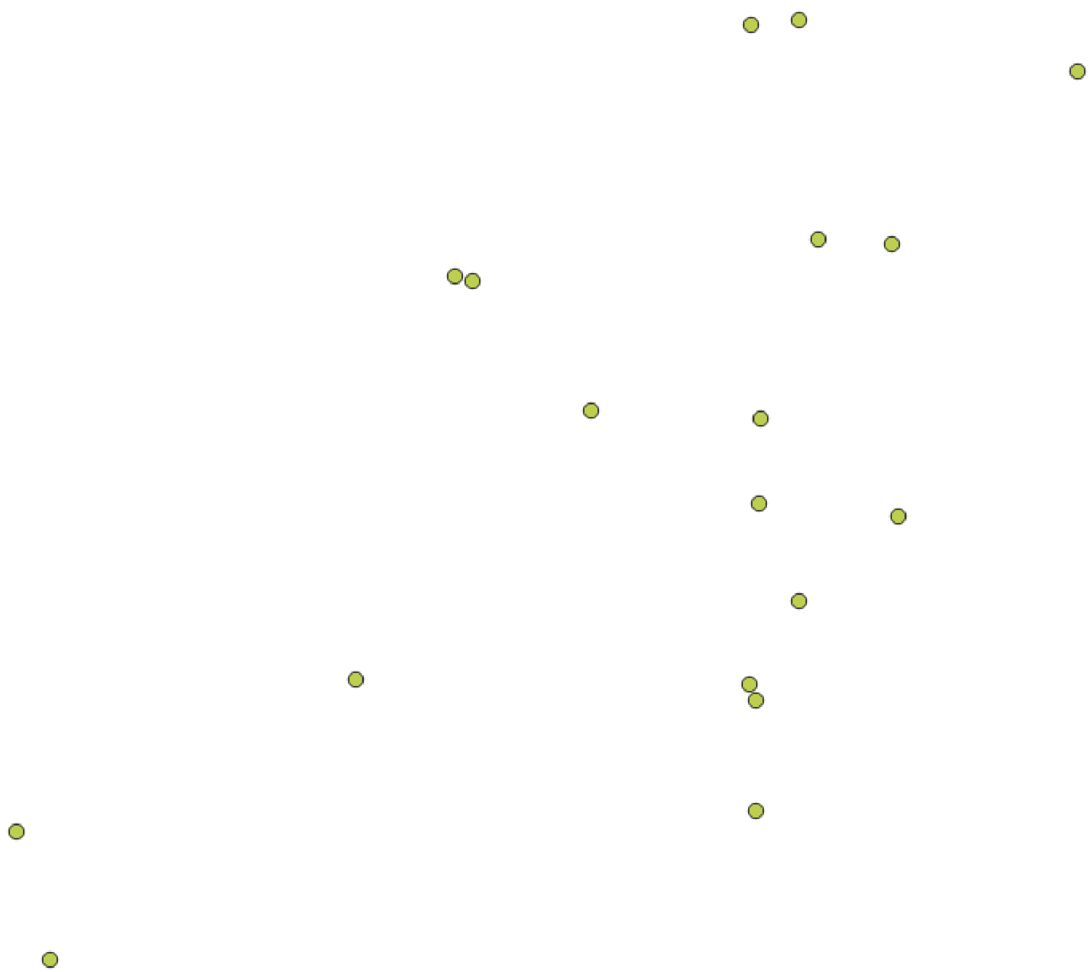

Figure 3.3: Montreal's Air Quality Monitoring Network

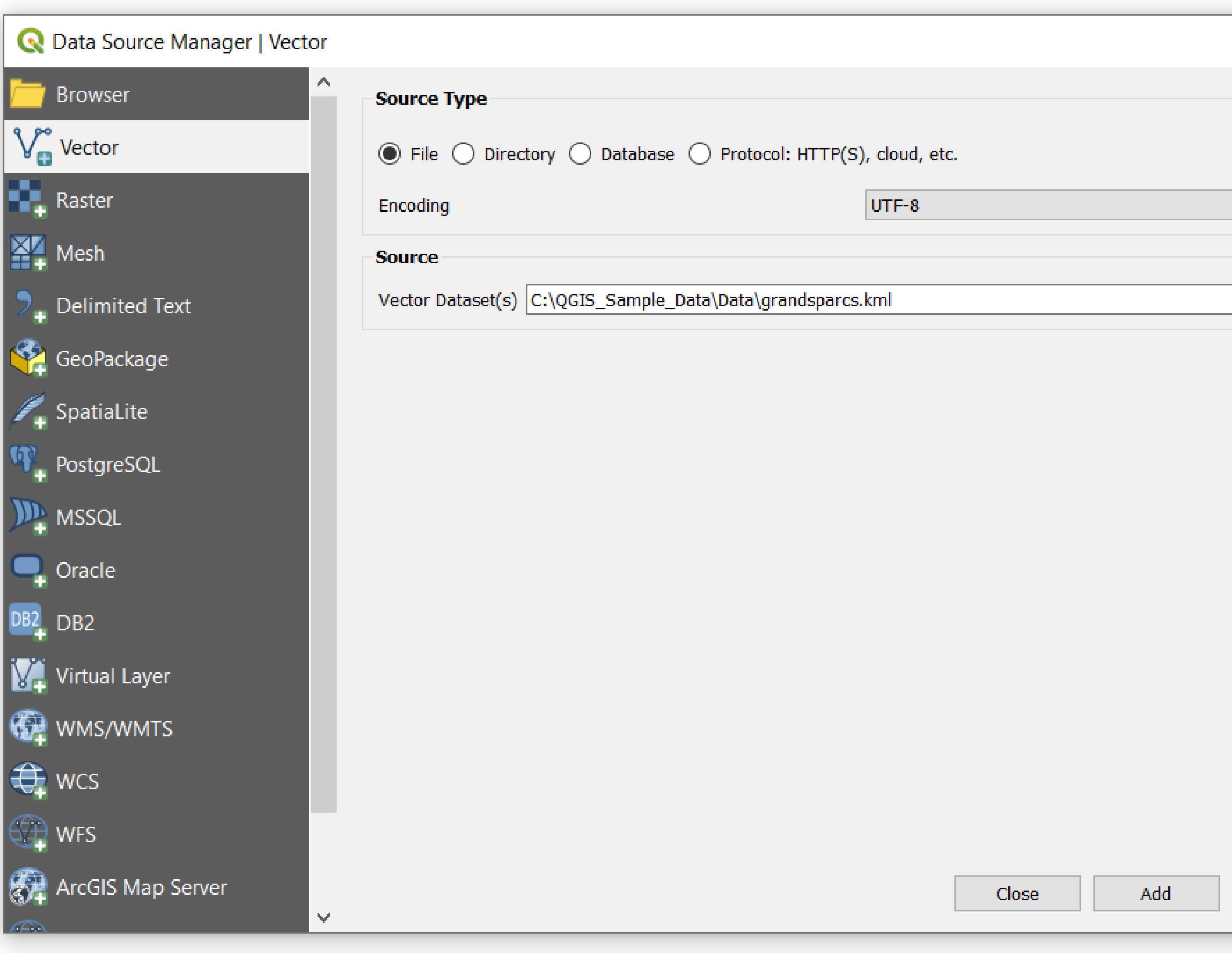

Figure 3.4: Add vector layer window

Figure 3.5: Montreal's parks

2. In the **File Name** field, indicate the path to the `grandparcs.kml` file. The following layer will be displayed in QGIS:

3.3 Importing GeoJSON files

GeoJSON is another frequently used format for storing and representing geographical attributes. This format is based on the JavaScript Object Notation (JSON).

The importation of a GeoJSON layer is illustrated with the `ilotschaleur.json` file. This file contains the urban heat islands (UHI) of Montreal. UHI correspond to urban areas characterized by higher summer temperatures than the immediate environment with differences between 5 and 10°C.

1. Navigate to the **Layer** menu and select **Add vector layer**. The following dialog will pop up:
2. In the **File Name** field, indicate the path to the `ilotschaleur.json` file. The following layer will be displayed in QGIS:

3.4 Saving a layer

We use the previously imported `coordonnees-stations-rsqa.csv` file to illustrate how to export a layer in a different format.

1. Right-click on the name of the layer, select **Export**, then **Save features As...** The following dialog window will pop up.
2. Select the format in which you want to export the layer. In this case, we have selected the widely used ESRI shapefile, we will use this layer in future recipes. Another commonly used format is GeoJson

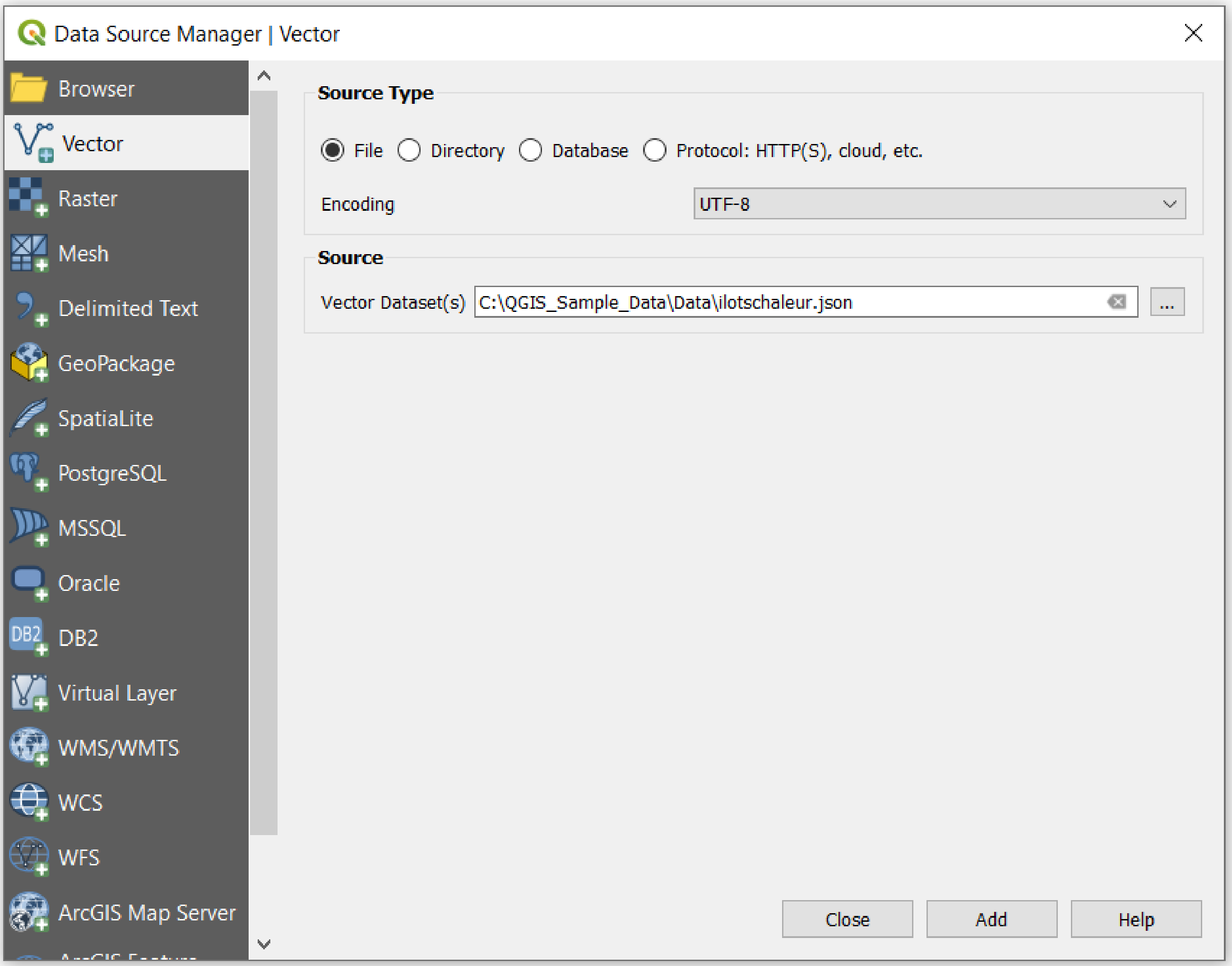

Figure 3.6: Add vector layer window

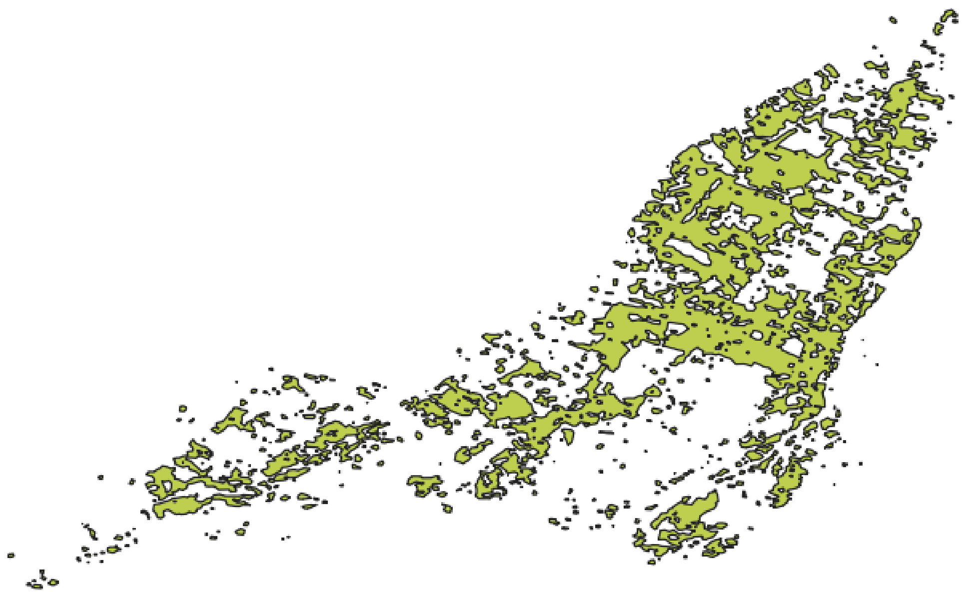

Figure 3.7: Urban heat islands (UHI) of Montreal

since it is compatible with web base applications.

3. Indicate the path where the layer will be stored and give it a name. Indicate whether you want to add the saved file to the current project, and finally click on OK.

3.5 Reprojecting a layer

Most of the time, the layers are not in a CRS that is more convenient for the operation in hand. QGIS offers on-the-fly reprojections for rendering the layers. However, when executing operations like spatial analysis, it is required that all layers be in the same CRS. This recipe is illustrated with the `LIMADMIN.shp` file that corresponds to the administrative limits of Montreal's boroughs.

Before executing a spatial analysis, it is recommended to reproject the layers to the most convenient CRS.

1. Right-click on the name of the layer, select **Export**, then **Save features As...**. The following dialog window will pop up.
2. Indicate the path where the layer will be stored and give it a name. If you tick the box **Add saved file to map**, the recently exported layer will be saved and added to the current project.
3. Click on the little globe to select the CRS in which the new layer will be projected. The following window will pop up:
4. You can filter the CRS. In this case we will export the new layer in **EPSG:6622**, since it is the most accurate for Quebec. Click **OK** to confirm the selection of the CRS and click again **OK** to save the layer.

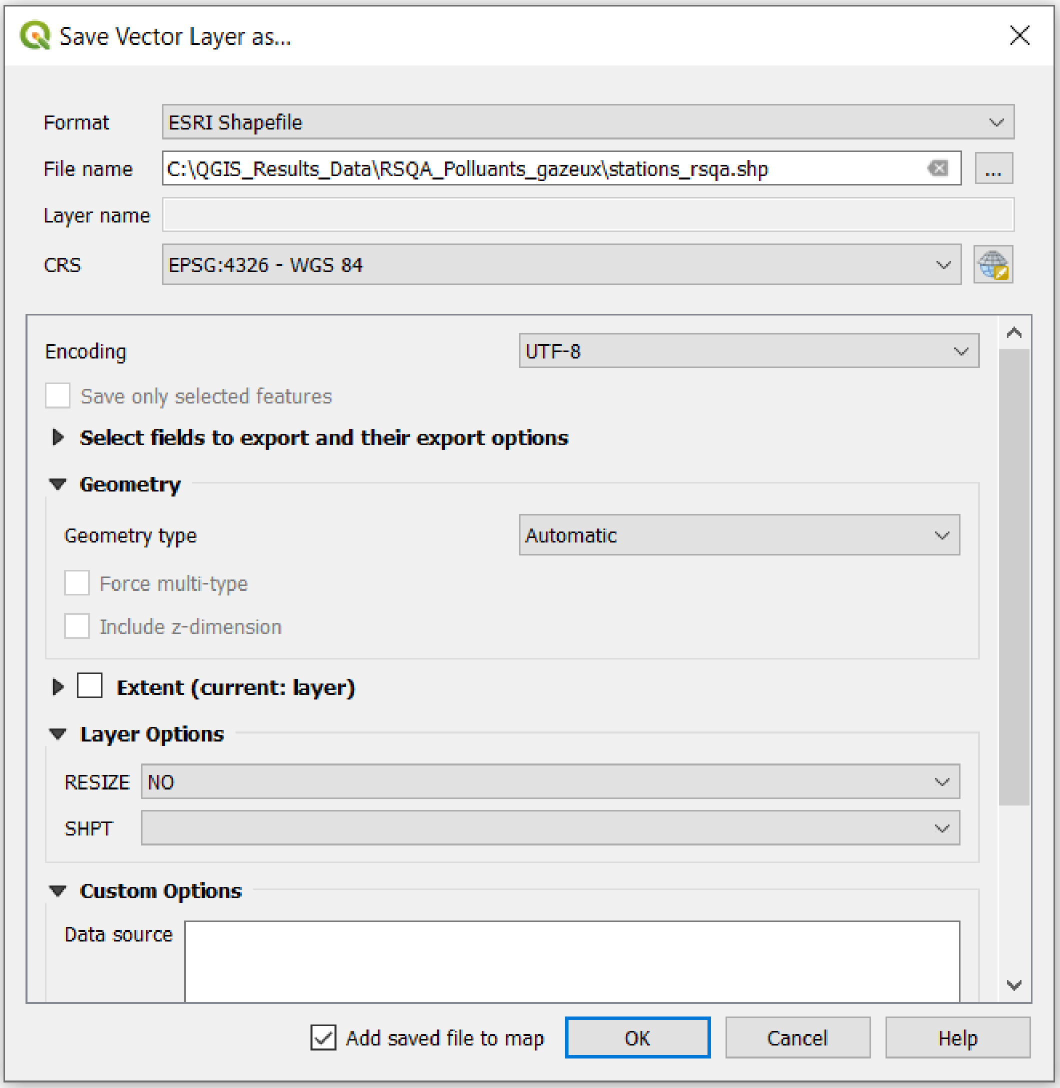

Figure 3.8: Save features as... window

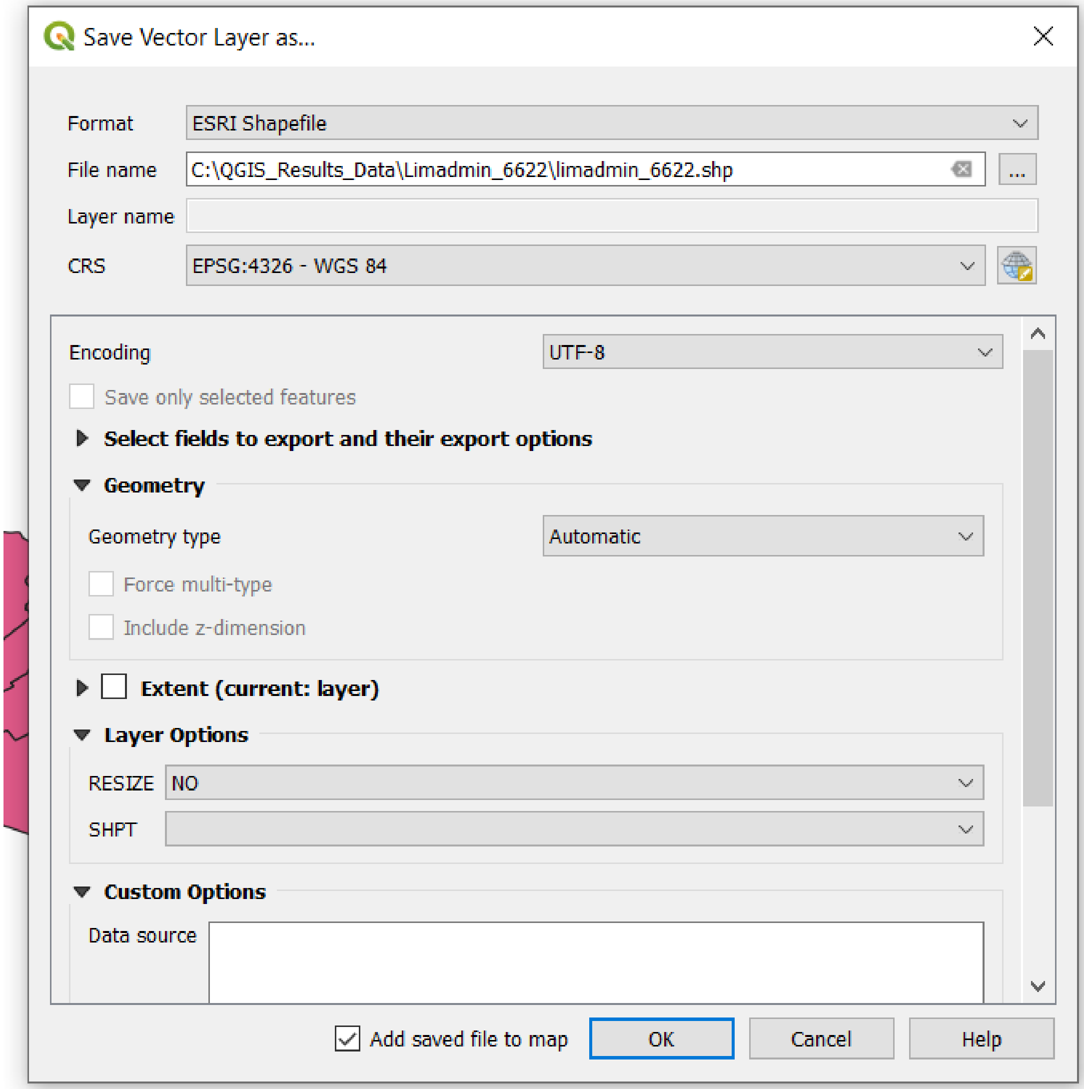

Figure 3.9: Reprojecting a layer

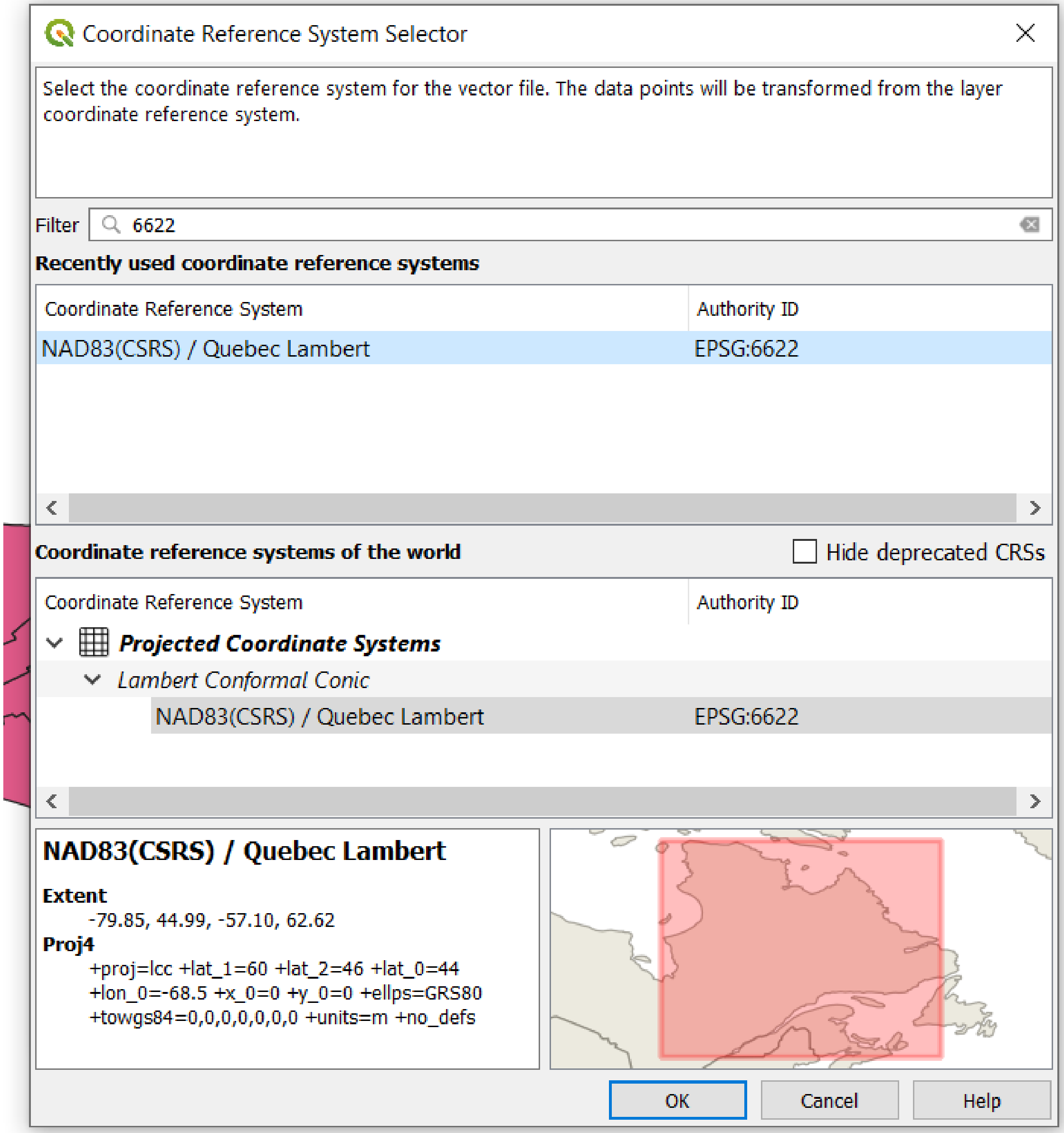

Figure 3.10: Selecting another CRS

Chapter 4

Data treatment

4.1 Joining a layer data

Another frequent task before executing spatial analysis with QGIS is to join an attribute table to a layer.

To illustrate this task, we refer to the `stations_rsqa.shp` file generated in section 1 and the file `pollutants_average_12_31_2019_13H.csv` file. When we load the shapefile into QGIS we identify that this layer only contains information on the identification and location of stations that compose the Air Quality Monitoring Network (RSQA).

We realize another problem, the name (nom) and the address of the stations are not correctly displayed. In order to fix it, right-click on the name of the layer, and select properties. A window will pop up. Go to **Source** and change the **Data Source Encoding** to **UTF-8**. Finally, click on Apply to accept the changes.

Now, import the `pollutants_average_12_31_2019_13H.csv` file. One kick method is to drag and drop the file from the file into QGIS. This works fine for this file; however, you also can use **Add Delimited Text Layer** to have more control on the importation.

The `pollutants_average_12_31_2019_13H.csv` file reports the average concentration of criteria pollutants from December 23, 2013, at 12h. The units of concentration are indicated in the following table.

Pollutant	Unit
CO	ppm
H2S	ppb
NO	ppb
NO2	ppb
SO2	$\mu g/m^3$
PM10	$\mu g/m^3$
PM2.5	$\mu g/m^3$
O3	ppb

In order to join the `pollutants_average_12_31_2019_13H` attribute table to the `stations_rsqa` layer, follow these steps:

1. Right-click on the name of `stations_rsqa` layer, select **Properties**, then **Joins** from the dialog window.
2. Click on the green + sign. The following window will pop up:

In this case, since we have imported only one attribute table, QGIS has already selected the Join layer.

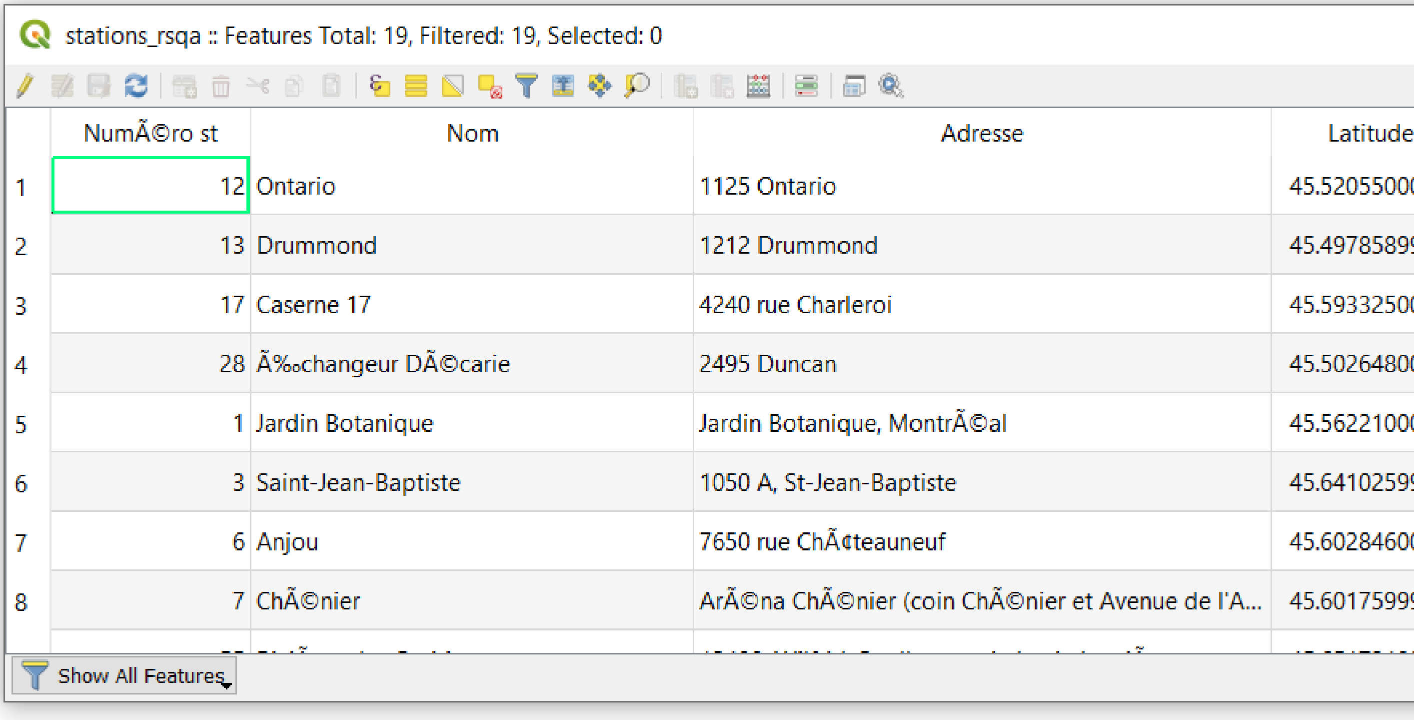

Figure 4.1: Attribute table of stations

3. Specify the **Join field** and the **Target field**, which correspond to the keys that relate the shapefile layer and the data layer. In this example, **NO_POSTE** is the identifier of the stations in the data layer and **Numéro st** is the identifier of the stations in the shapefile layer. Furthermore, it is possible to select the fields that will be joined, and the prefix that will be used. Since there are no repeated columns, we just deleted the default prefix, which corresponds to the name of the layer.

4. Click on **OK**, then on **Apply** to finish the joining.

   To verify if the join has worked, you can open the attribute table of the shapefile. To make this change permanent, you need to export the layer. After the join, the layer was exported as `stations_rsqa_12_31_2013.shp`.

4.2 Cleaning up the attribute table

Sometimes, data imported into QGIS is not in the correct format, the name of columns is not self-explanatory, or we simply want to discard the columns that will not be used during the task in hand.

The **Refactor fields** algorithm simplifies removing, renaming and converting the format of dbf tables in QGIS. This algorithm ca be accessed through the **Processing Toolbox**. The use of **Refactor fields** is illustrated with the `stations_rsqa_12_31_2013.shp` file that was generated in the previous section.

1. Import `stations_rsqa_12_31_2013.shp` file
2. Launch the **Refactor field**. It will detect the available layer in the current project. In this window we can identify the name of columns, the type of information they store, their length and precision. The columns corresponding to concentrations, such as CO, H2S, and NO, are currently stored as text (data type is string). This is not convenient, since we cannot do arithmetic with text. Furthermore, we may want to change the name of **Numéro st** field for **Station**.

3. We change the name of the first three columns according to the following figure, and the type of

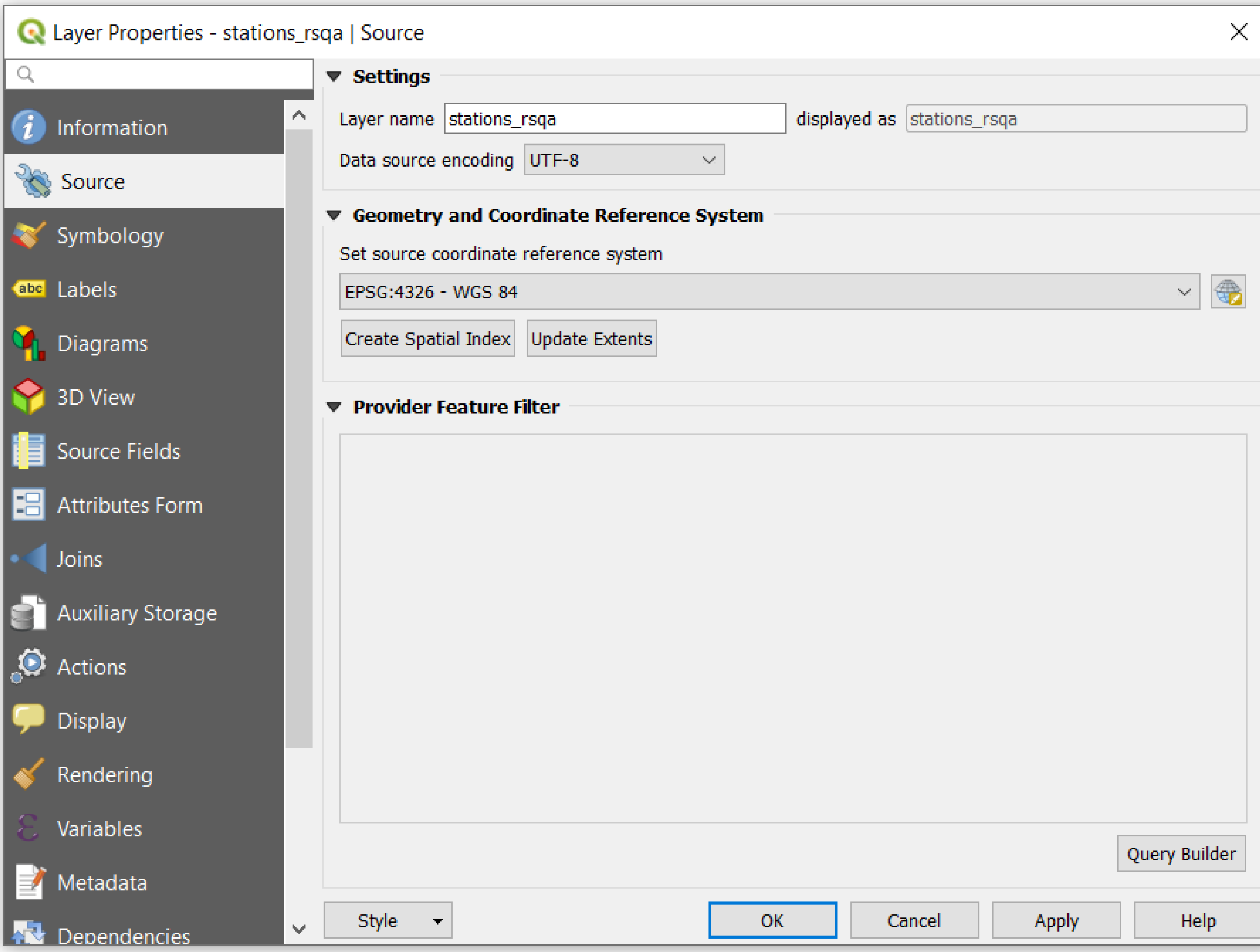

Figure 4.2: Changing encoding

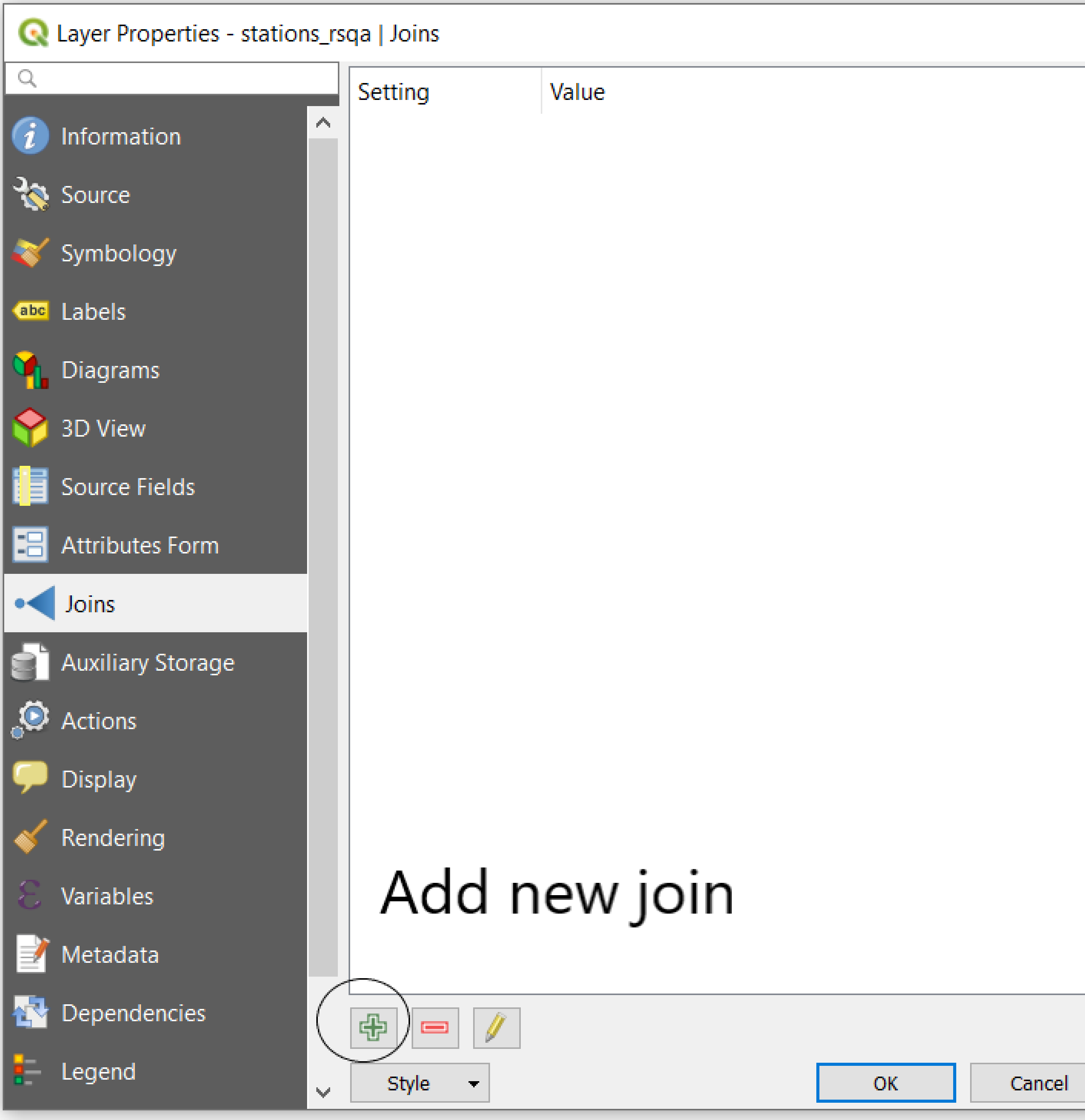

Figure 4.3: Layer properties window

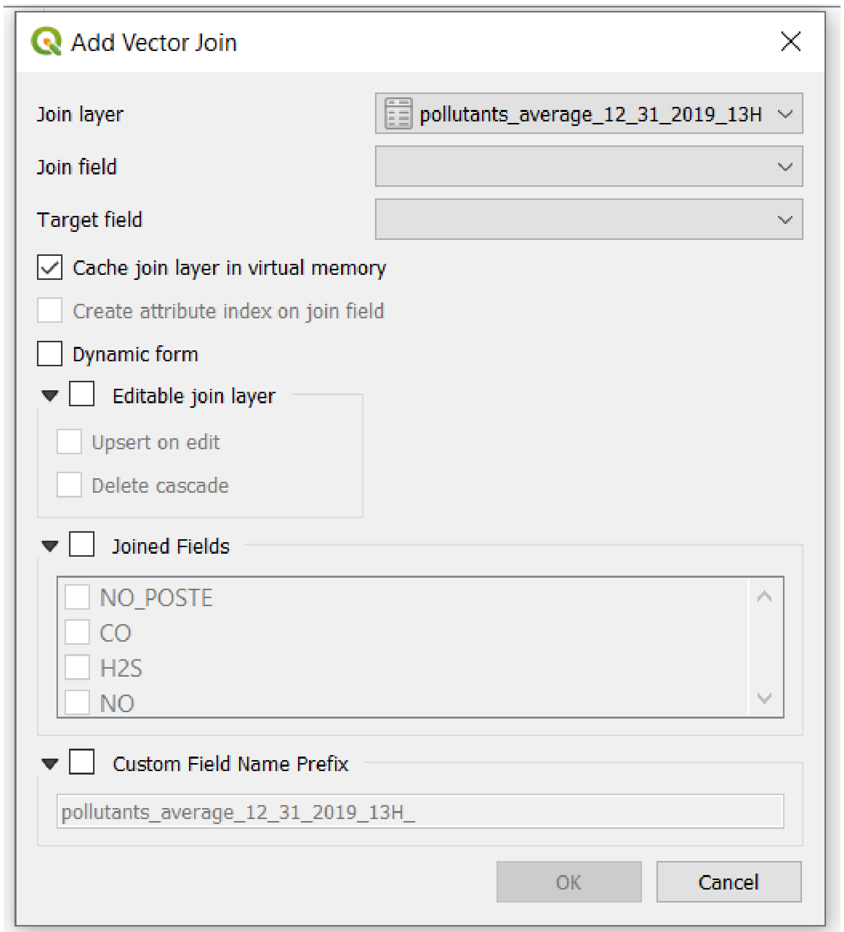

Figure 4.4: Joining a layer table

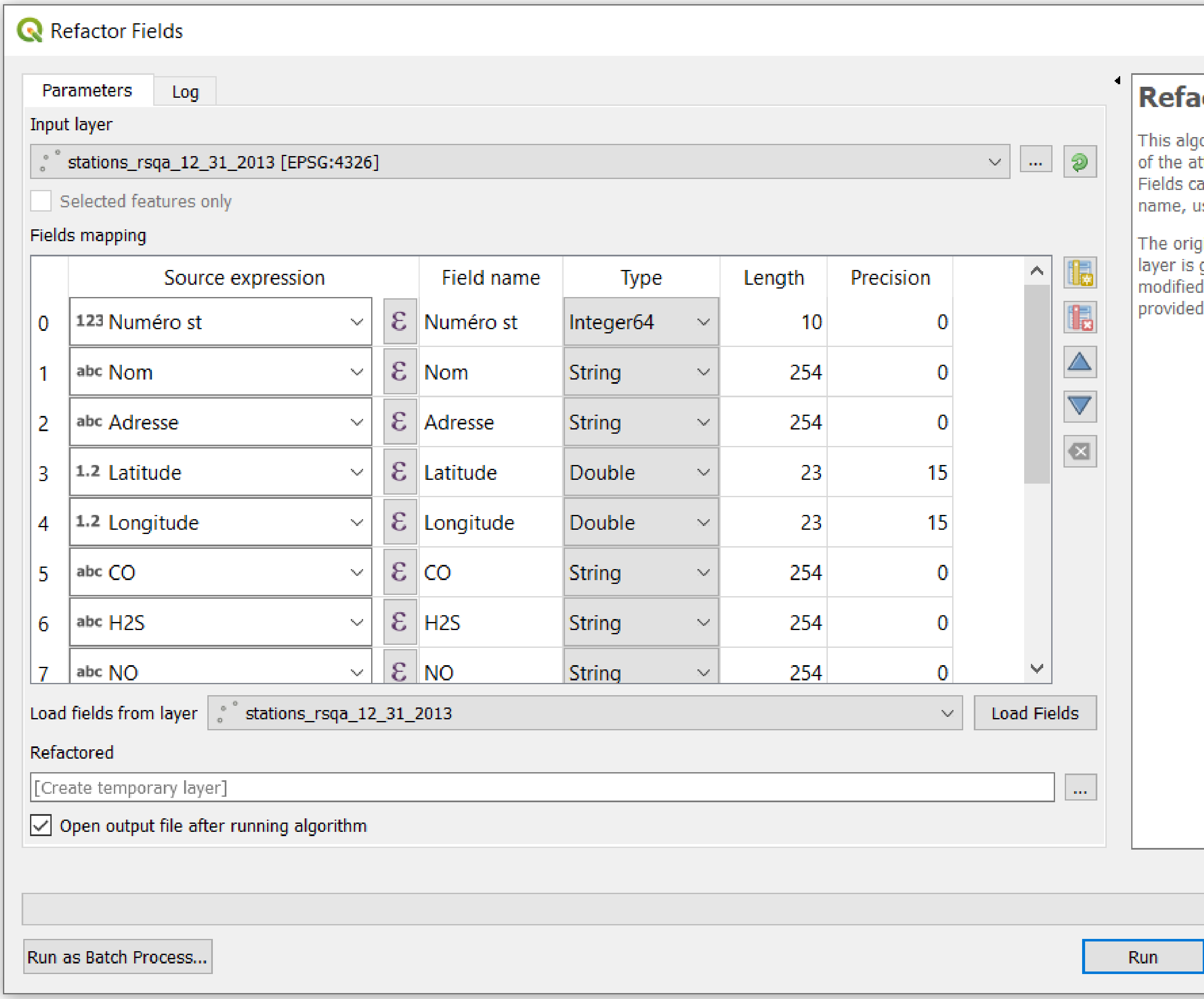

Figure 4.5: Refactor field window

columns corresponding to concentrations was set as **Double** (real number) with a length of 23 and a precision of 3. Click on **Run** to generate a new layer.

4. The generated layer is named by default Refactored, of course, you can change its name at your convenience.

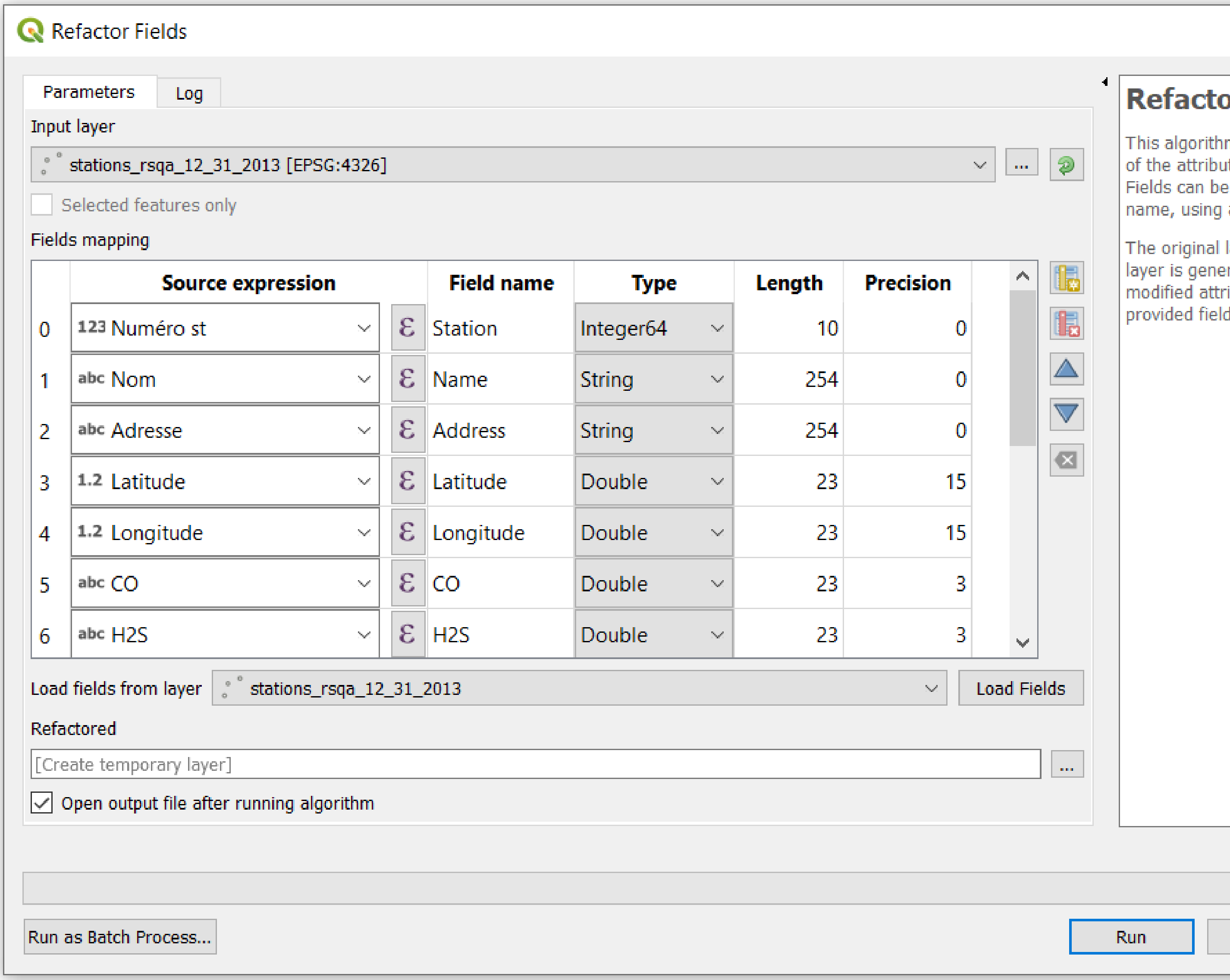

Figure 4.6: Using the refactor field algorithm

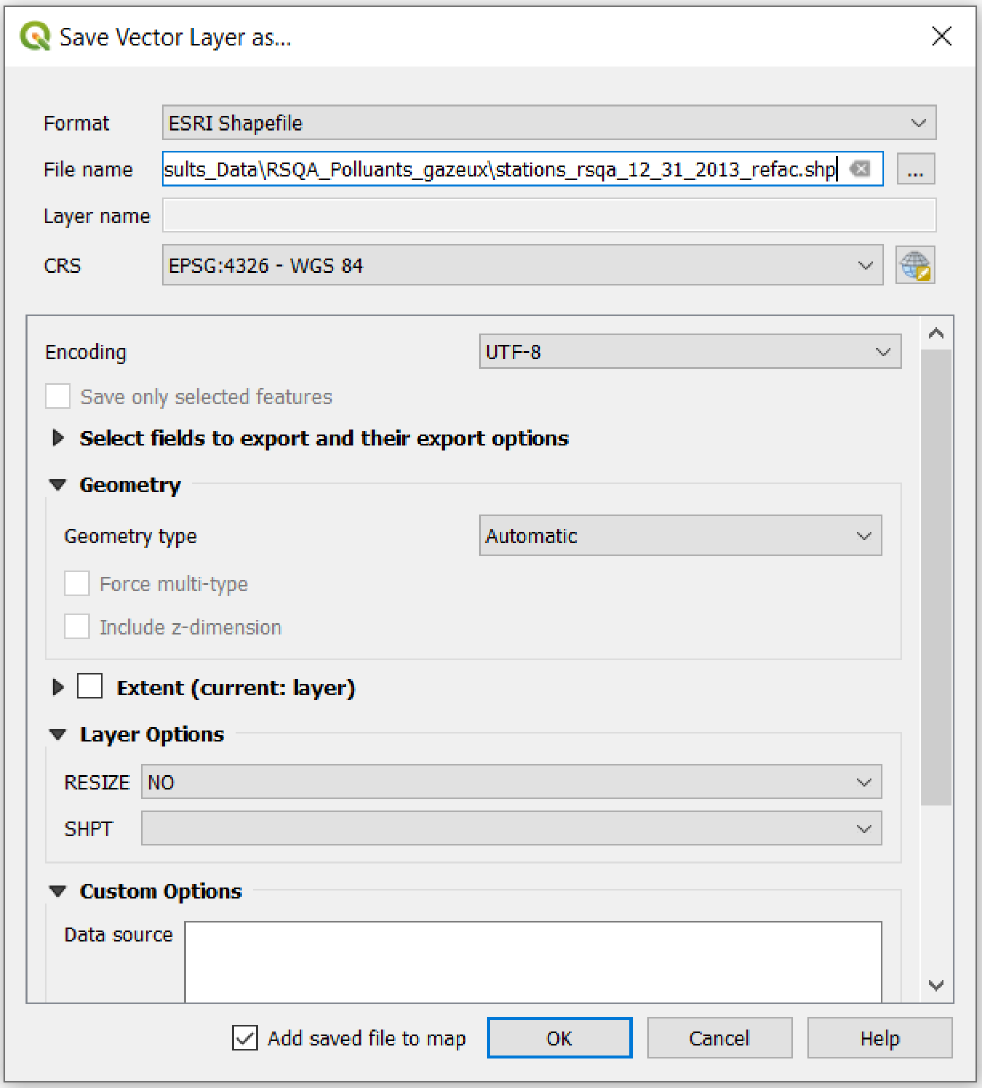

Figure 4.7: Saving the refactored layer

Chapter 5

Data preprocessing steps

5.1 Clipping vectors

In some cases, it is necessary than a layer covers only an area of interest. For this purpose, we can use a layer setting the extent we want to keep for a section of a layer. To accomplish this task, we use **Clip** from the **Geoprocessing Tool**.

To illustrate the operation of clipping vectors, we will use `terre_shp` and `LIMADMIN.shp` files.

1. Load both layers into QGIS. `LIMADMIN.shp` corresponds to the administrative limits of Montreal's boroughs (last update in 2013); however, the polygons extent beyond the terrestrial limits; whereas the `terre_shp` file corresponds to the terrestrial limits of Montreal Island.
2. Navigate to **Layer**, then to **Geoprocessing Tool**, and select **Clip**. The following window will pop up. `LIMADMIN.shp` layer corresponds to the Input layer, since it is the layer we want to cut, whereas the `quartier_limite.shp` is the Overlay layer, the layer we will use to set the limits we want to keep.
4. Click on Run to generate a new layer. Save the clipped layer as `limadmin_clipped.shp`.

5.2 Intersecting vectors

The **Intersection** algorithm extracts the properties from the input layer that overlap features in the overlay layer.

To illustrate this recipe, consider the `terre.shp` and `LIMADMIN.shp` files. In this case, we will create a layer resulting from the intersection between `terre.shp` and `LIMADMIN.shp`. Therefore, the `terre.shp` layer corresponds to the **Input Layer**, whereas `LIMADMIN.shp` to the **Overlay Layer**. We can add an overlay index to identify the features that were intersected from the overlay layer in Intersection layer generated

5.3 Check validity and fix geometries

In some cases, when clipping and intersecting vectors, errors may arise because of invalid geometries. Fortunately, QGIS allows us to check the validity of layers, and even more importantly to fix them. The following algorithms can be accessed through the **Processing Toolbox**:

- **Check validity**: The algorithm performs a validity check on the geometries of a vector layer. The geometries are classified in three groups (valid, invalid and error).
- **Fix geometries**: The algorithm attempts to create a valid representation of an invalid geometry without losing any of the input vertices.

Figure 5.1: Montreal's boroughs

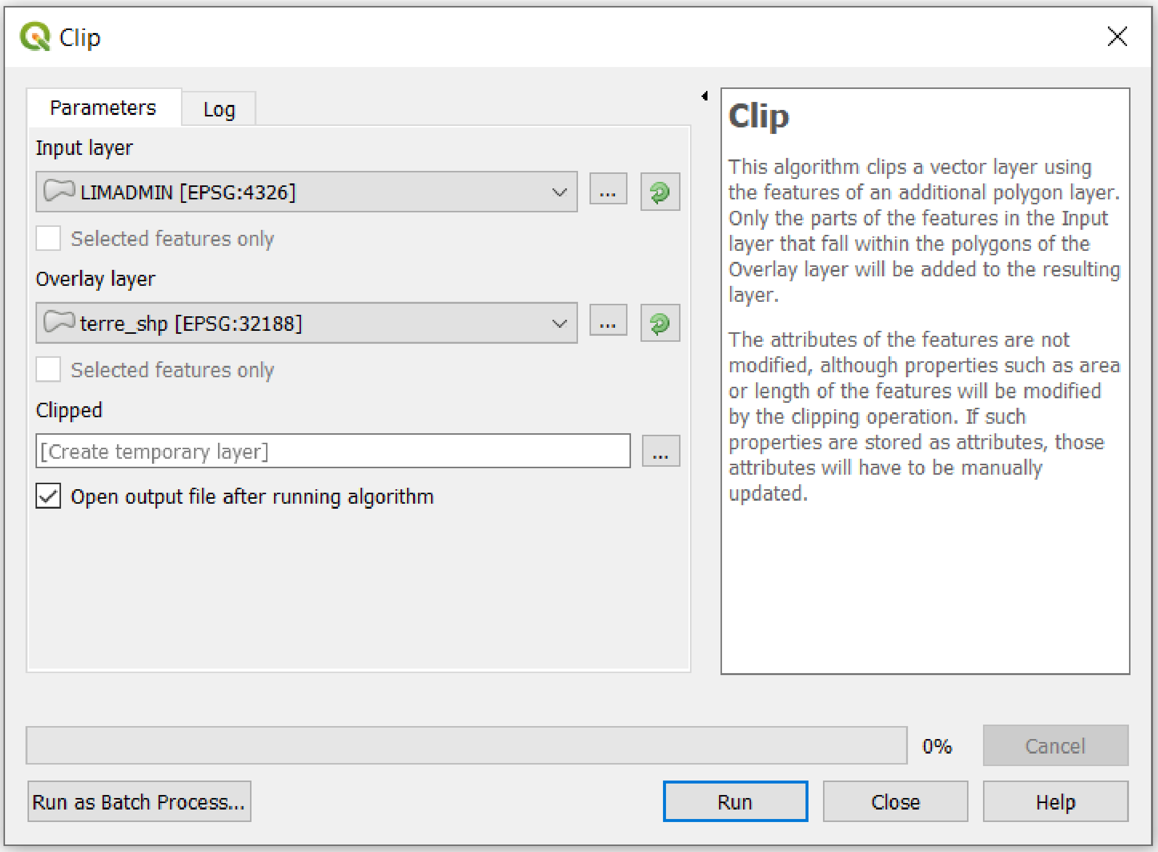

Figure 5.2: Using the clip algorithm

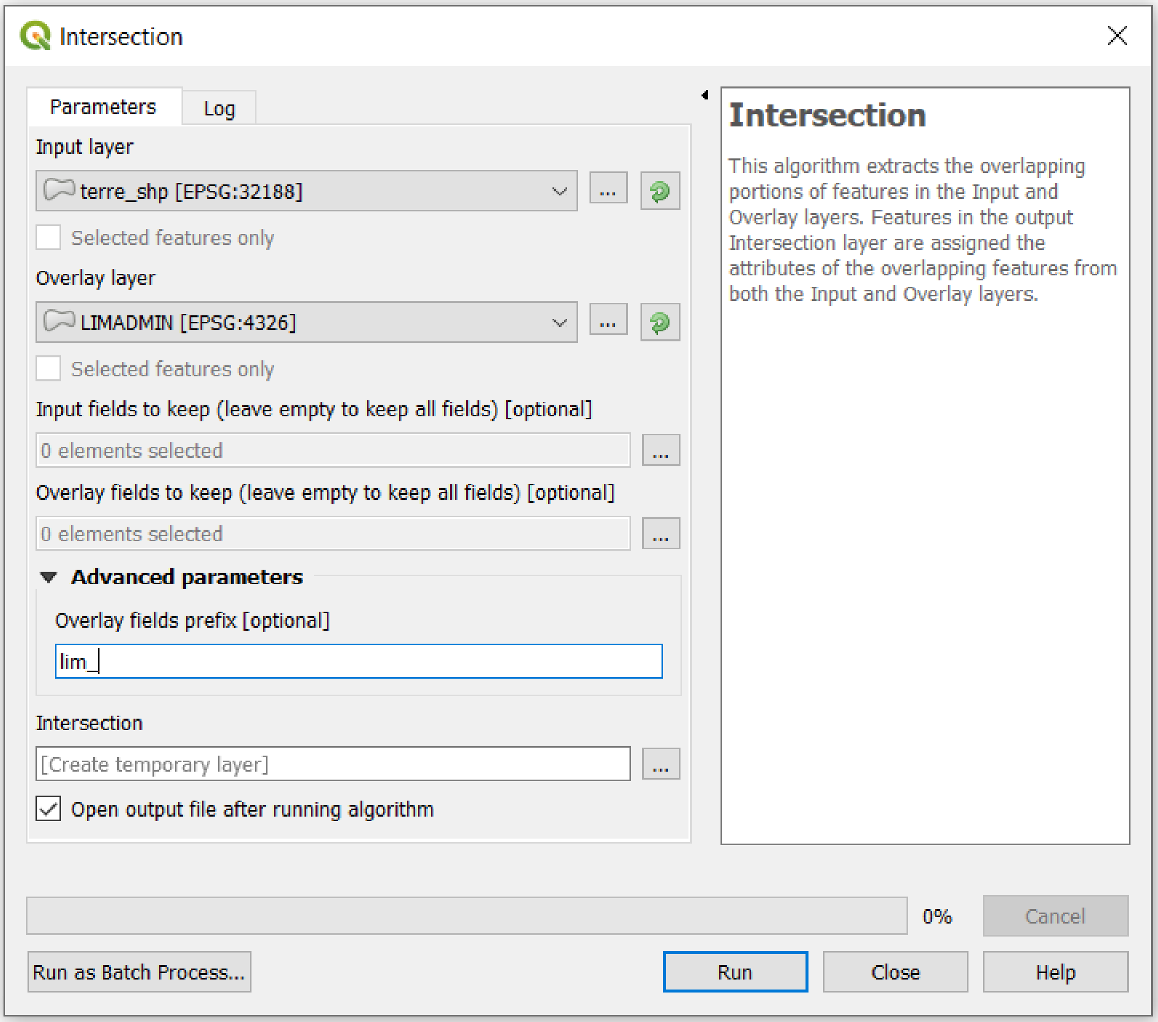

Figure 5.3: Intersection of vectors

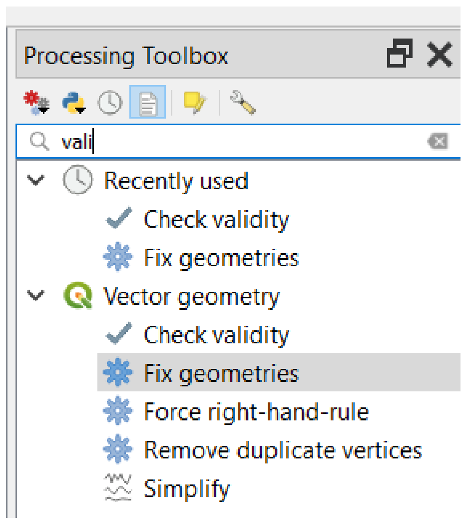

Figure 5.4: Check validity and fix geometries

Chapter 6

Data exploration

6.1 Listing unique values in a column

The **List Unique values** algorithm, from the **Processing Toolbox**, generates a table and a HTLM report of the unique values of a given layer's field.

To illustrate the use of the **List Unique values** algorithm, we will import the `grandparcs.kml` file. The field **Generique2** stores the type of parks located in Montreal, and we would like to know the unique values without having to open the attribute table.

1. Launch the **List Unique values** algorithm from the **Processing Toolbox**. It will identify the available layer in the current project.
2. Click on **...** from the **Target Field(s)** and select **Generique2**. Click on Run to generate a temporary layer and a HTLM report.

6.2 Loading BaseMaps

When we import layers into QGIS, sometimes it is difficult to identify what the points, lines or polygons of a layer correspond to. In these situations, it is very helpful to add a base map to give some spatial context.

One plugging that comes in hand to add spatial context is **QuickMapServices**. If it is not already available in your QGIS Desktop, navigate to **Plugins**, then to **Manage and Install Plugins**. In the window that has been displayed search **QuickMapServices** and click on **Install Plugin**.

To illustrate the use of **QuickMapServices**, load the `stations_rsqa_12_31_2013.shp`. Navigate to **Web** from the menu bar, select **QuickMapServices**, go to **OSM**, and select **OSM_Standard**. You will see the following set of air quality monitoring stations located in Montreal.

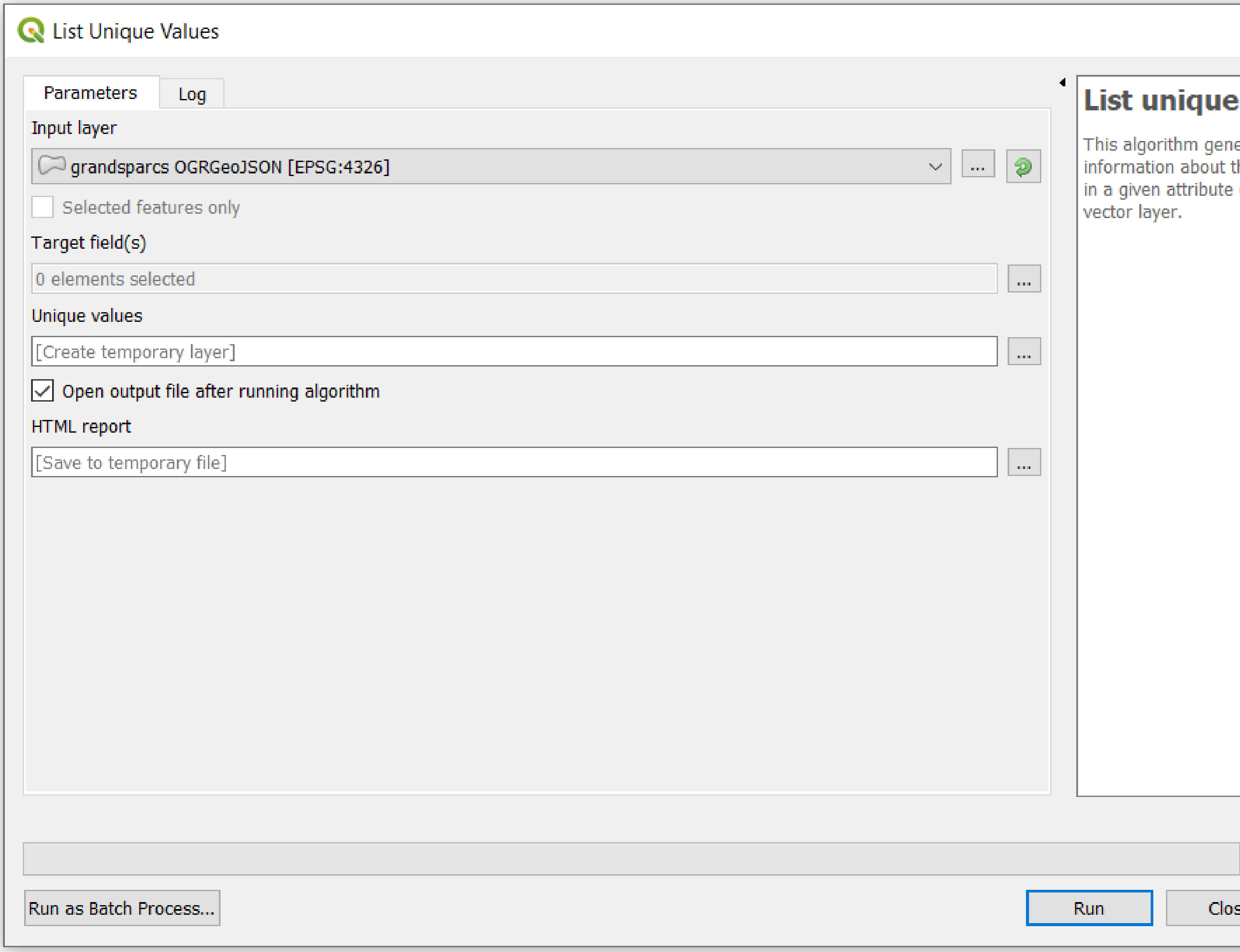

Figure 6.1: List unique values window

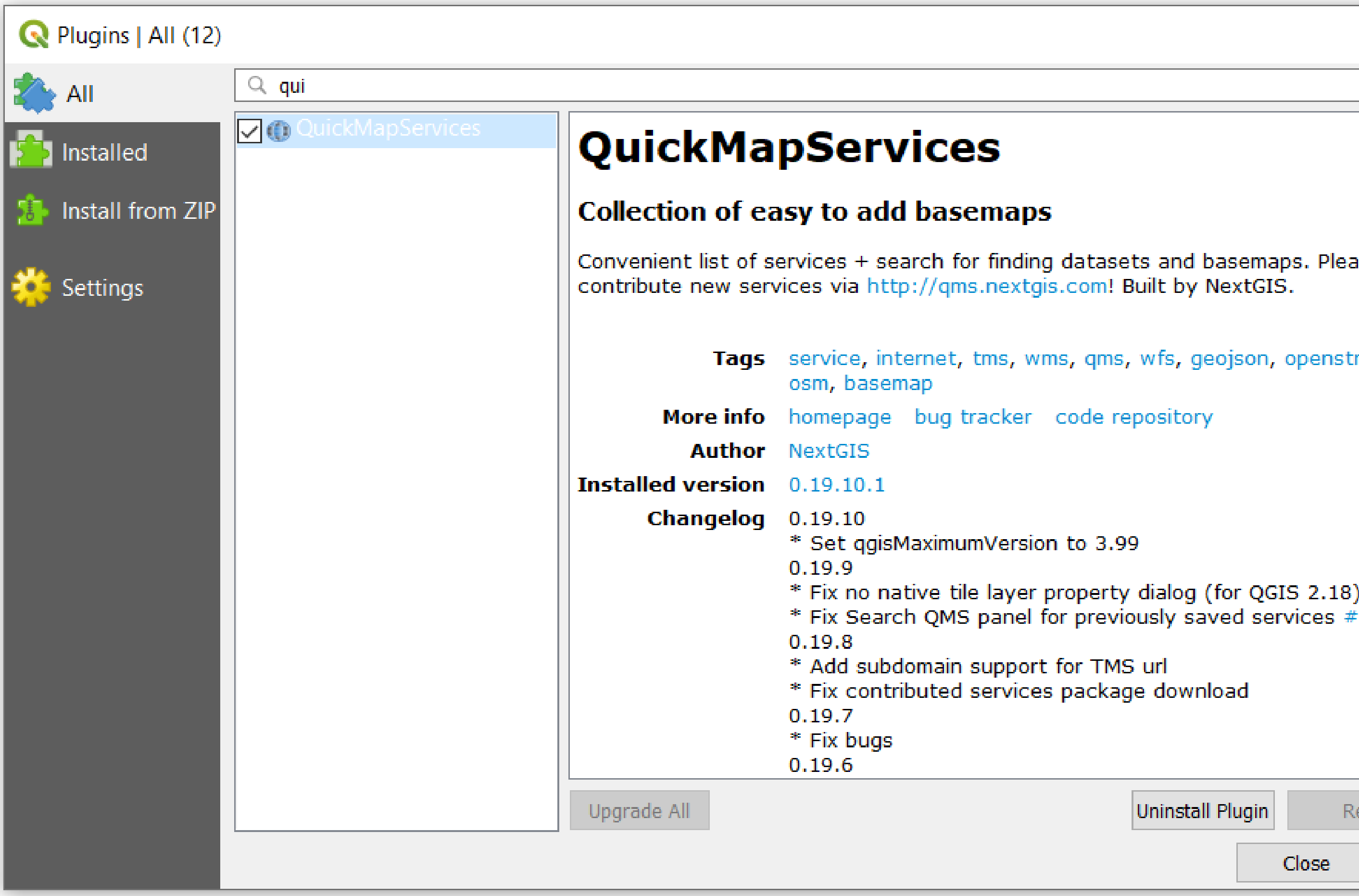

Figure 6.2: Manage and install plugins in QGIS

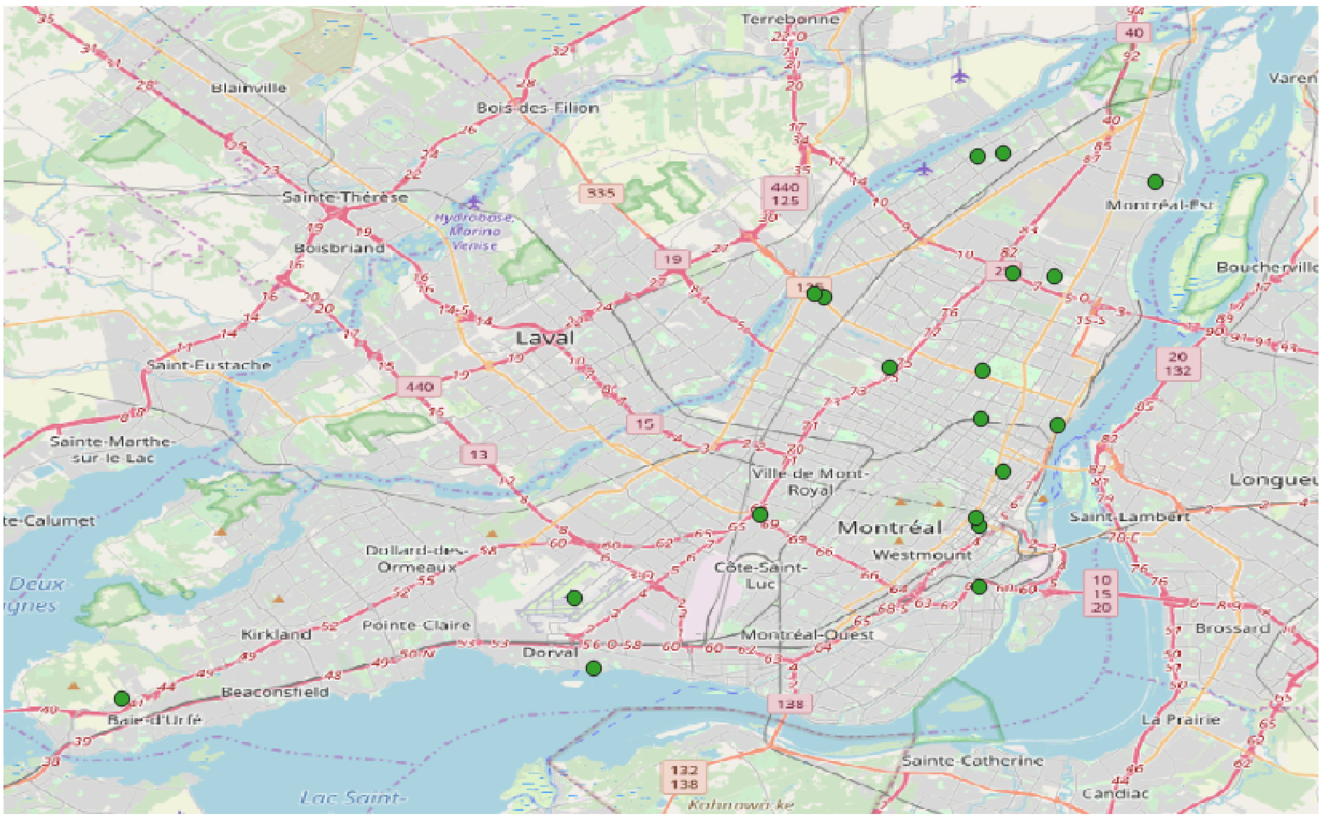

Figure 6.3: Montreal's Air Quality Monitoring Network

Chapter 7

Exercises

7.1 Exercise 1: Determine the area fraction of urban heat islands by boroughs of Montreal

The aim of this exercise is to generate a choropleth map showing the area fraction of urban heat islands by boroughs of Montreal.

1. Import `ilotschaleur.json` and `limadmim_clipped.shp`.

2. Export `ilotschaleur.json` as a shapefile and save it as `ilotschaleur.shp` (Section 3.4).

3. Intersect the recently created layer `ilotschaleur.shp` with `limadmim_clipped.shp`. Navigate to **Vector**, select **Geoprocessing Tools**, and the **Intersection** (Section 5.2). Set the following parameters:

- **Input layer**: `ilotschaleur.shp`
- **Overlay layer**: `limadmim_clipped.shp`

The aim is to generate a new layer in which the urban heat islands are divided according to Montreal's boroughs.

Click Run to execute the algorithm. However, it will stop since there are some invalid geometries in the input layer.

4. We will use the **Fix Geometries** algorithm from the **Processing Toolbox** to fix the geometries. Run the algorithm according to the following settings.

A temporary layer will be generated after running the algorithm. Right click on layer `Fixed geometries` and rename it `ilotschaleur_fixed`.

5. Run again the **Intersection** algorithm, but this time use the fixed layer of urban heat islands. Set the following parameters:

- **Input layer**: `ilotschaleur_fixed.shp`
- **Overlay layer**: `limadmim_clipped.shp`

A temporary layer `Intersection` will be generated after running the algorithm. To verify that the algorithm has properly worked, open the attribute table and you will notice that the `Intersection` layer has 584 attributes, whereas the `ilotschaleur_fixed` has 498.

6. Dissolve the `Intersection` layer by NOM. The aim is to combine the polygons corresponding to the same borough, which is given by the NOM field. In the **Dissolve field(s)**, click on the … and select NOM.

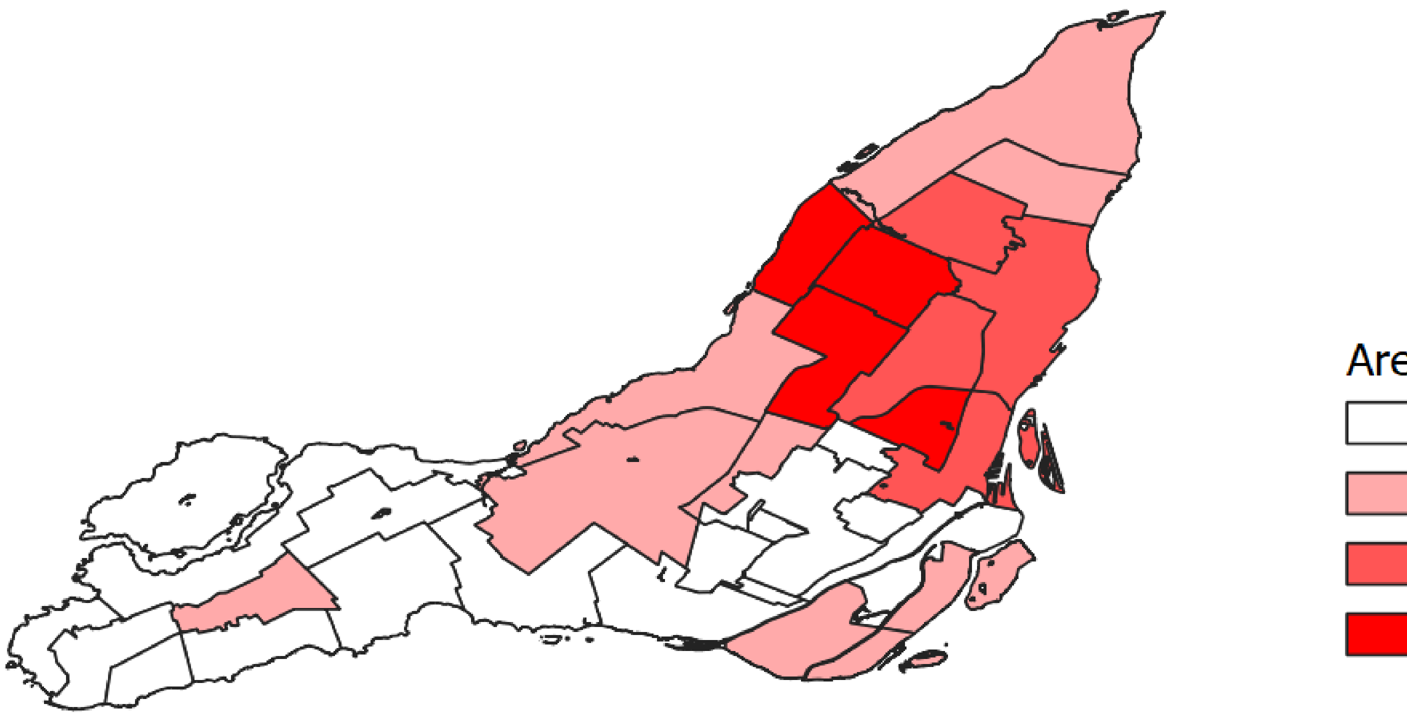

Figure 7.1: Area fraction of UHI by boroughs

Figure 7.2: UHI and boroughs of Montreal

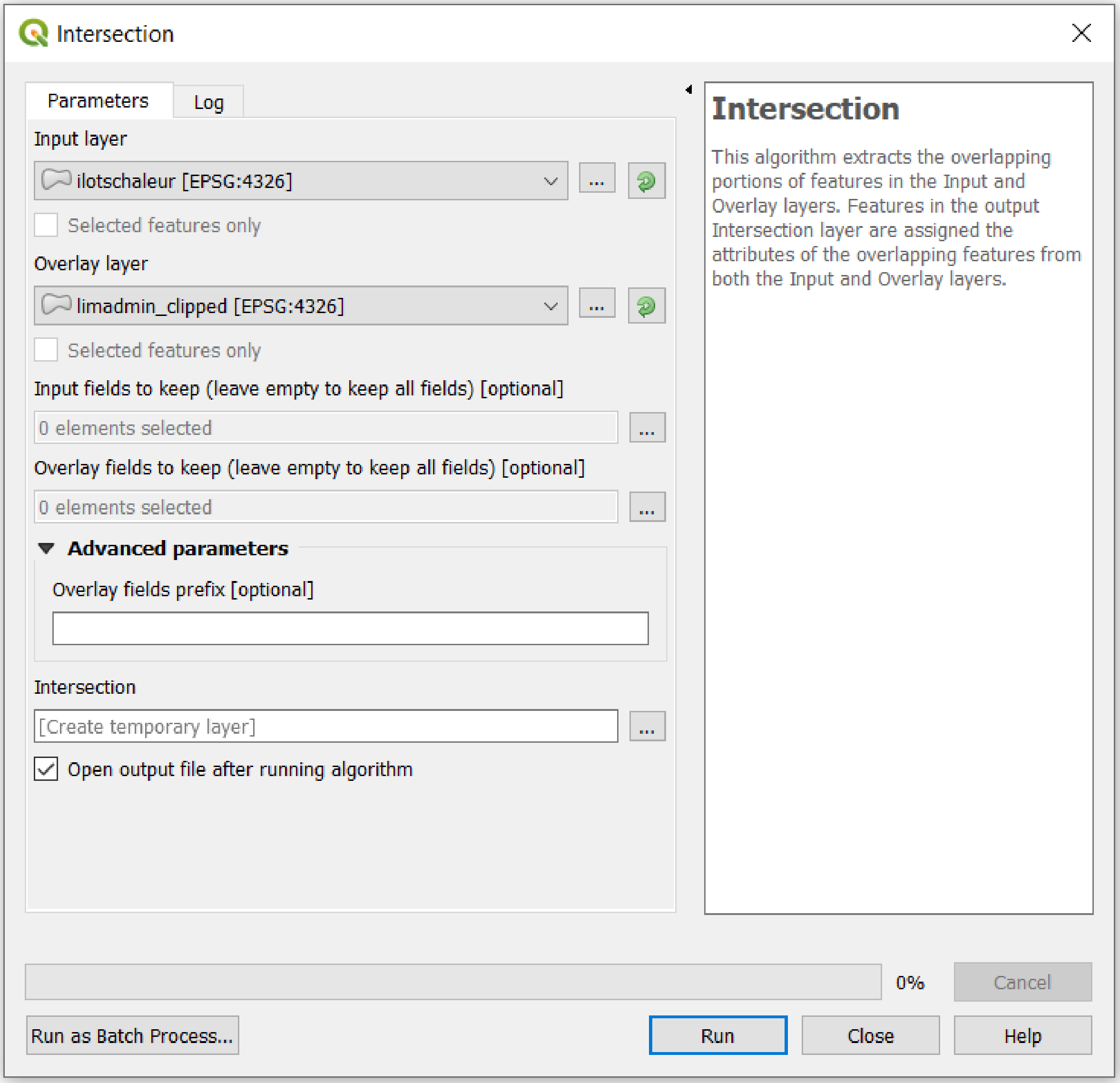

Figure 7.3: Intersection of layers

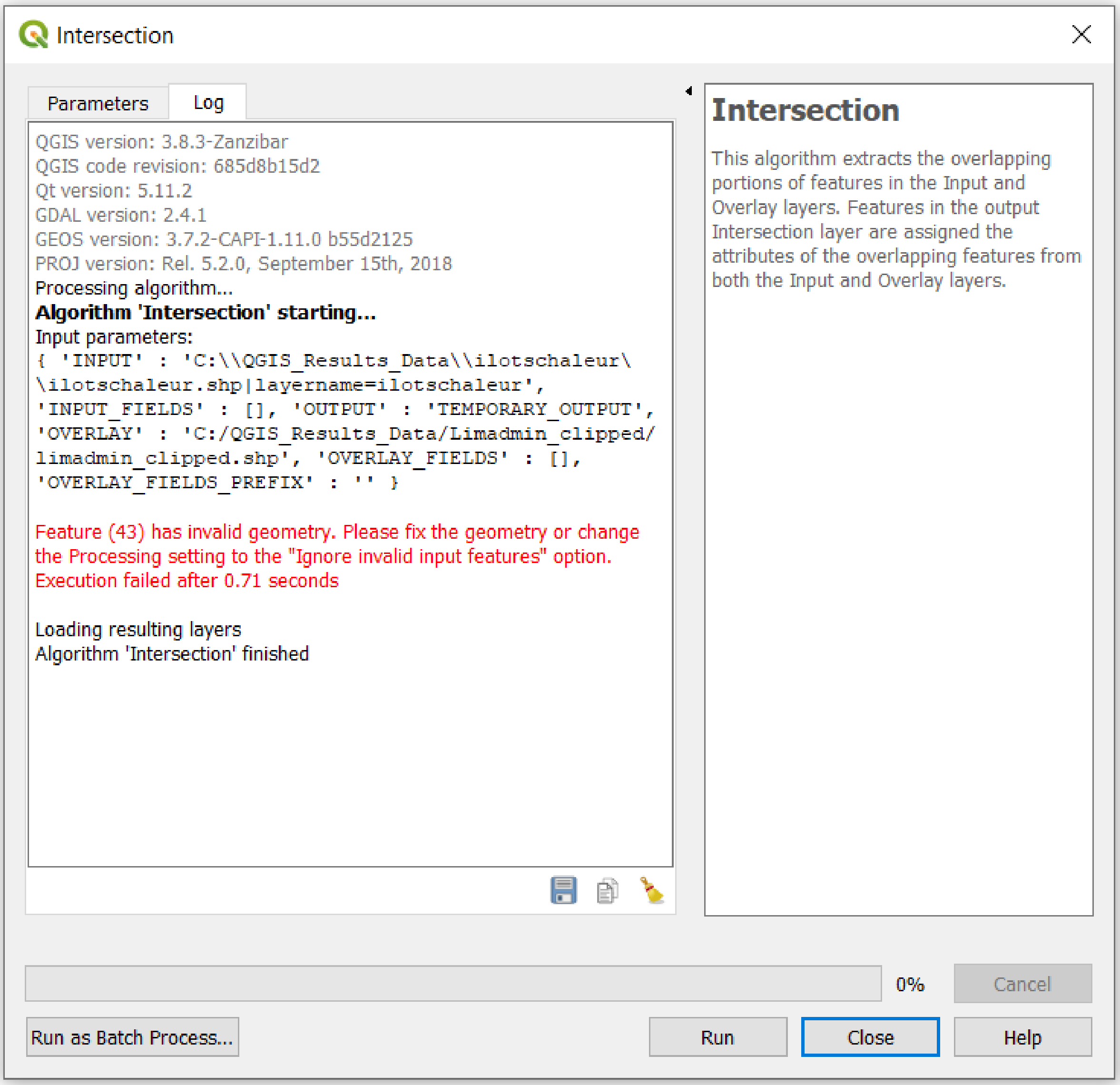

Figure 7.4: Error: Intersection of layers

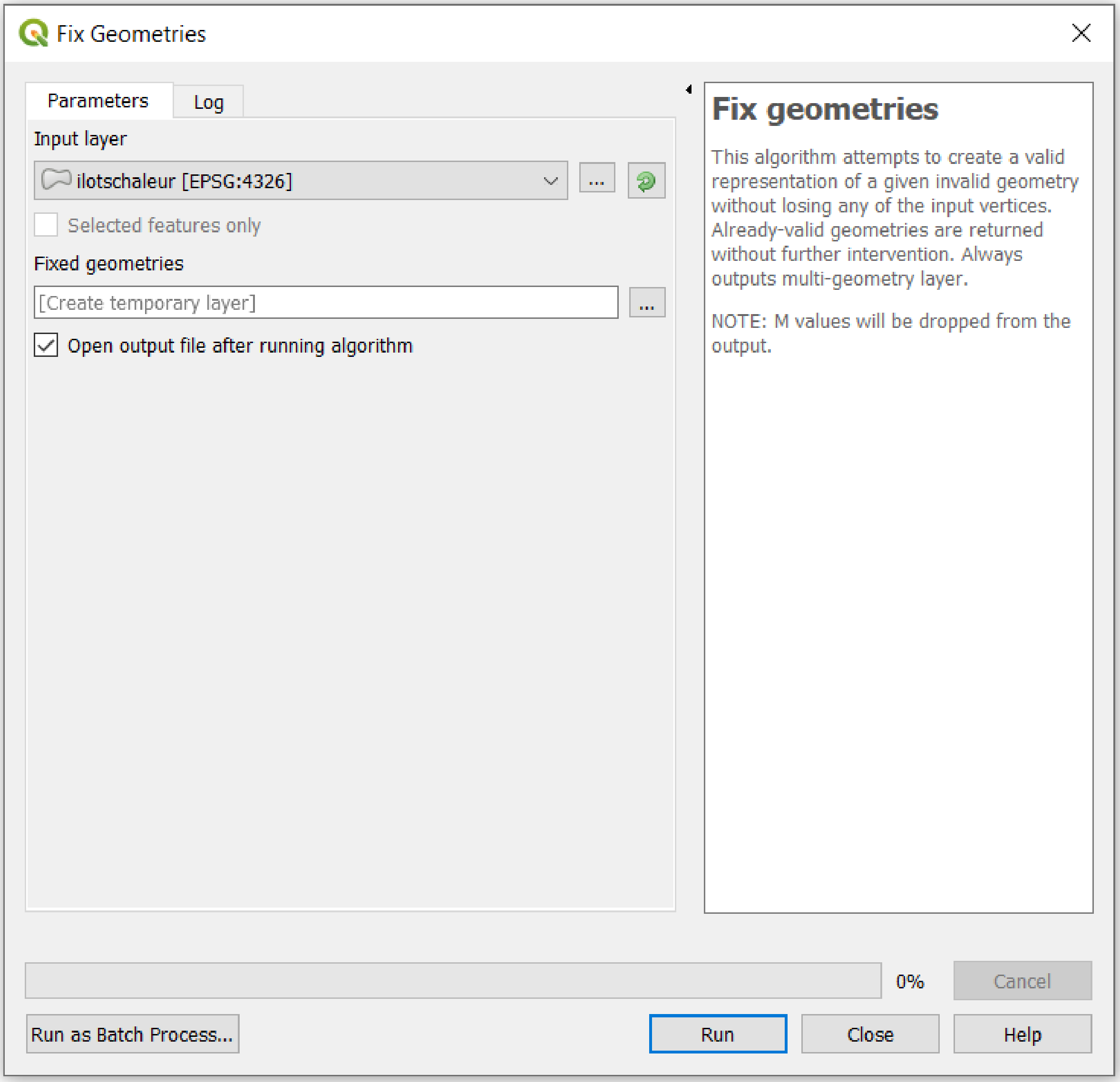

Figure 7.5: Fix geometries algorithm

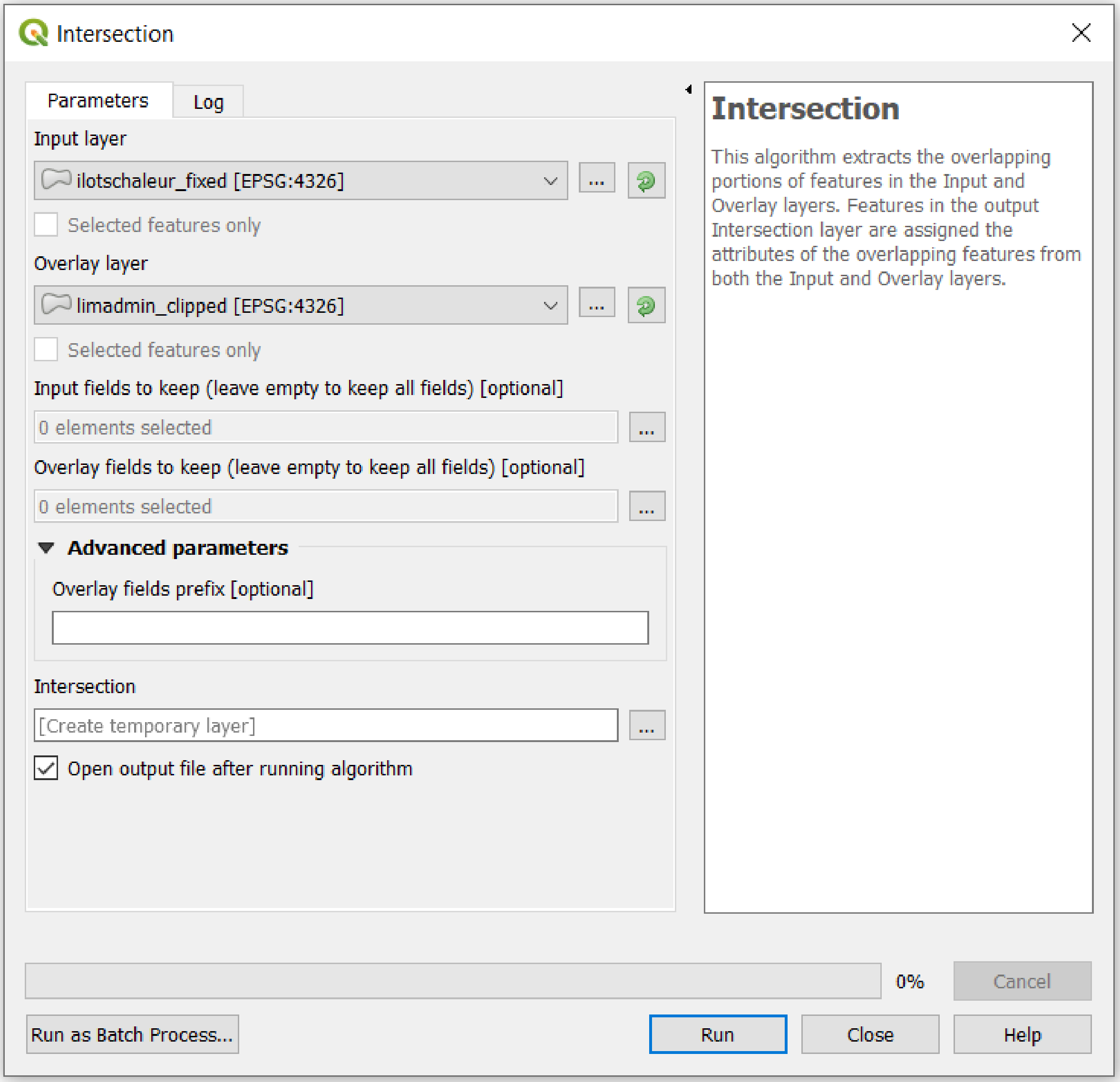

Figure 7.6: Intersection of layers

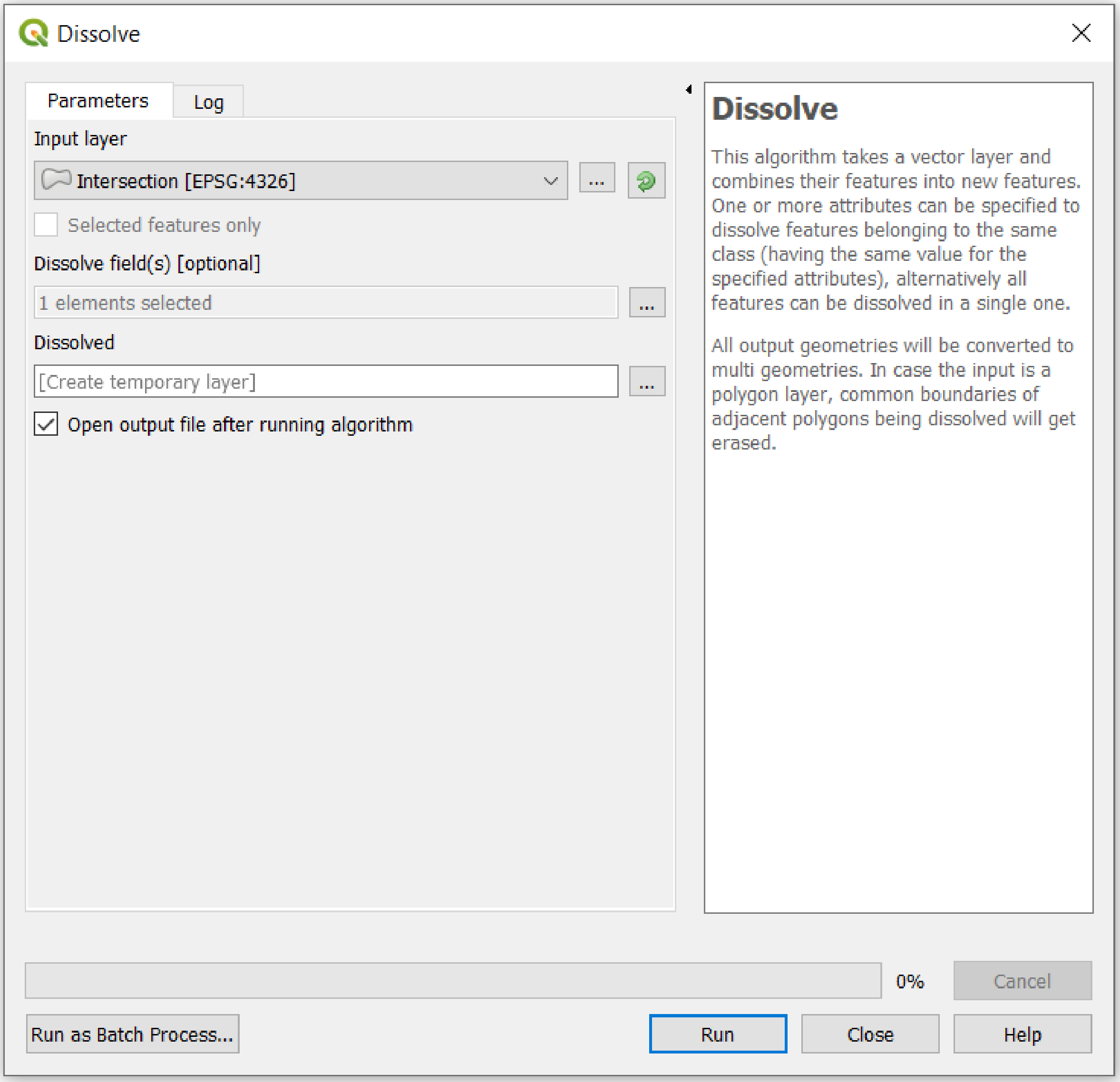

Figure 7.7: Dissolve Intersection layer by NOM

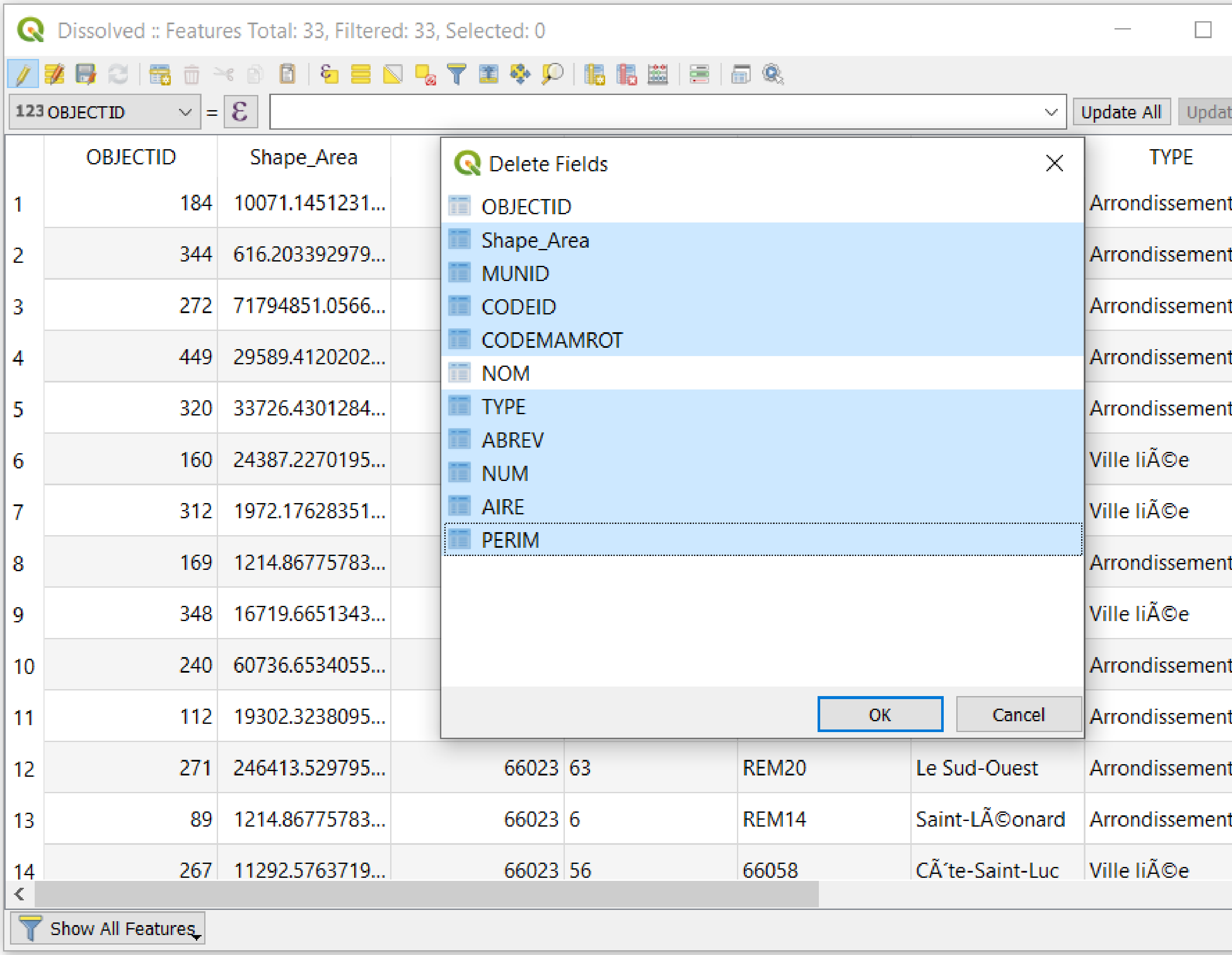

Figure 7.8: Delete some fields of Dissolved layer

After running the **Dissolve** algorithm, a temporary layer `Dissolved` will be generated. Open the attribute table of this layer and verify that it has 33 attributes; whereas the `Intersection` layer has 584. The task of combining the polygons has been accomplished.

7. Calculate the area of the urban heat islands by boroughs.

First of all, we will delete the fields of the `Dissolved` layer that will not be used. Open the attribute table, then click on the pencil shown in the left corner (**Toggle editing mode**) to allow editing the attribute table. Click on **Delete** field and select the fields shown in the following figure:

We will now calculate the area of the urban heat islands by borough. Click on **New field**. In the window that will display, select **Create a new field**, indicate *AREA* in the **Output field name**, set *Decimal number (real)* as the **Output field type**. Double click on *$area* from the center panel.

8. Join the AREA of the urban heat islands (`Dissolved` layer) to the respective borough in the `limadmim_clipped.shp` layer. The prefix **UHA__ was** set to distinguish the fields from the joined

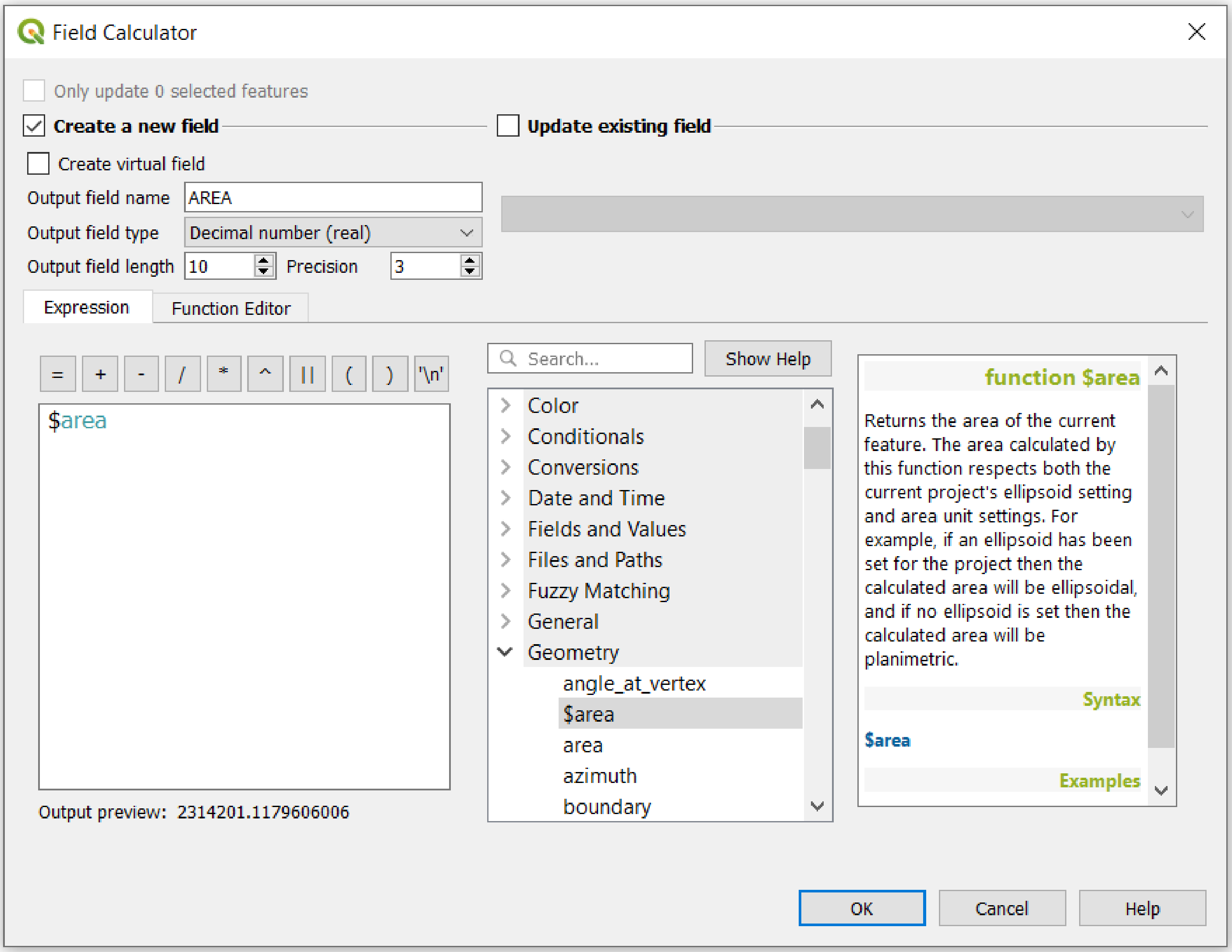

Figure 7.9: Calculate the area of UHI

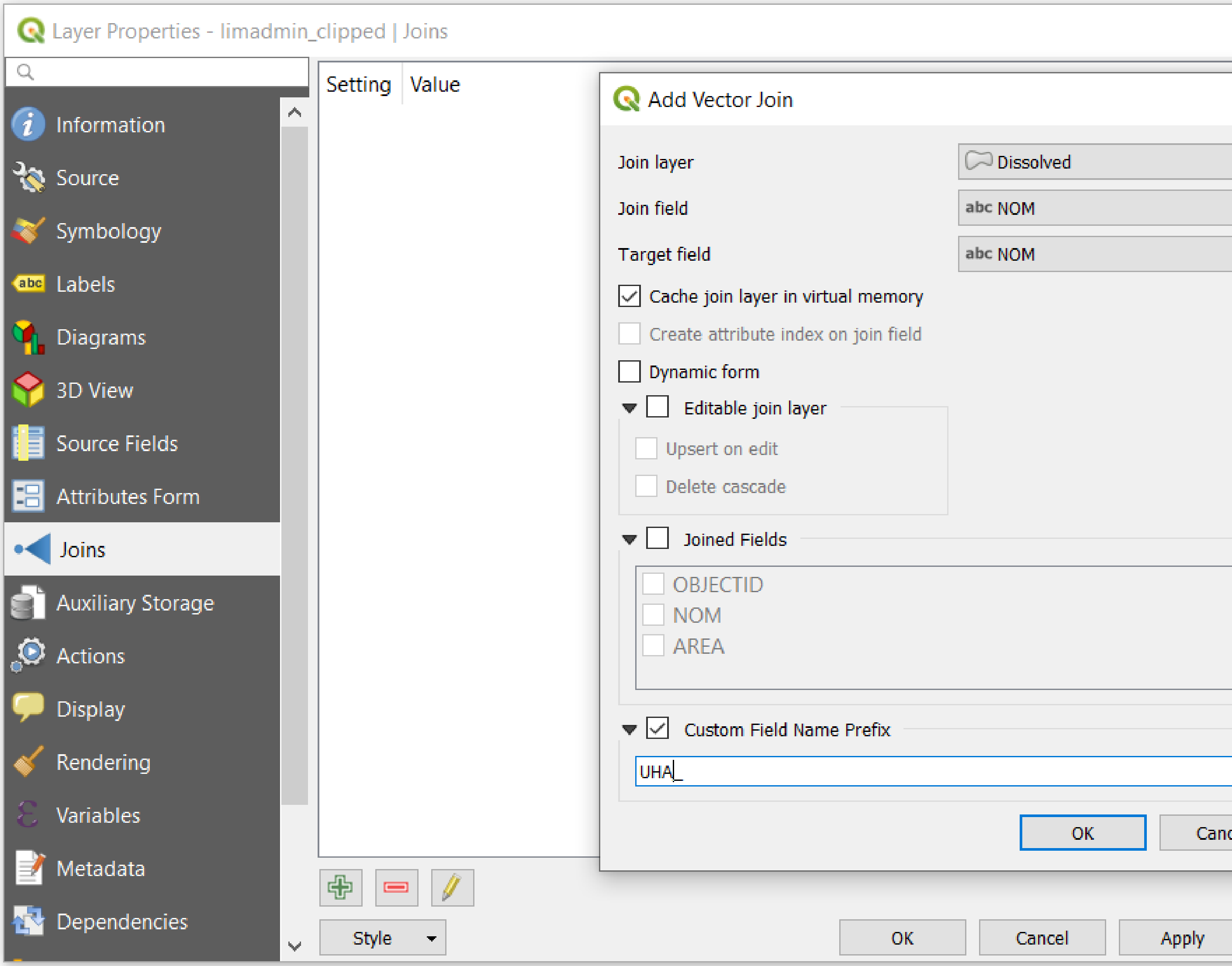

Figure 7.10: Join layers

layer.

9. Calculate the fraction area of urban heat islands (UHI) by boroughs.

First, calculate the AREA of each borough from the `limadmin_clipped.shp` layer. Click on **New field**. In the window that will display, select **Create a new field**, indicate AREA in the Output field name, set Decimal number (real) as the Output field type. Double click on $area from the center panel.

Finally, calculate the area fraction of UHI. In the center panel, go to **Fields and Values**, double click to select the involved fields in the computation of the new one. In this case, the area fraction is given by the expression: FRAC_UHA = UHA_AREA/AREA

Note: UHI was misspelled. So UHA stand for UHI.

10. Style the `limadmin_clipped.shp` layer so that it shows the area fraction of UHI in four classes of equal interval.

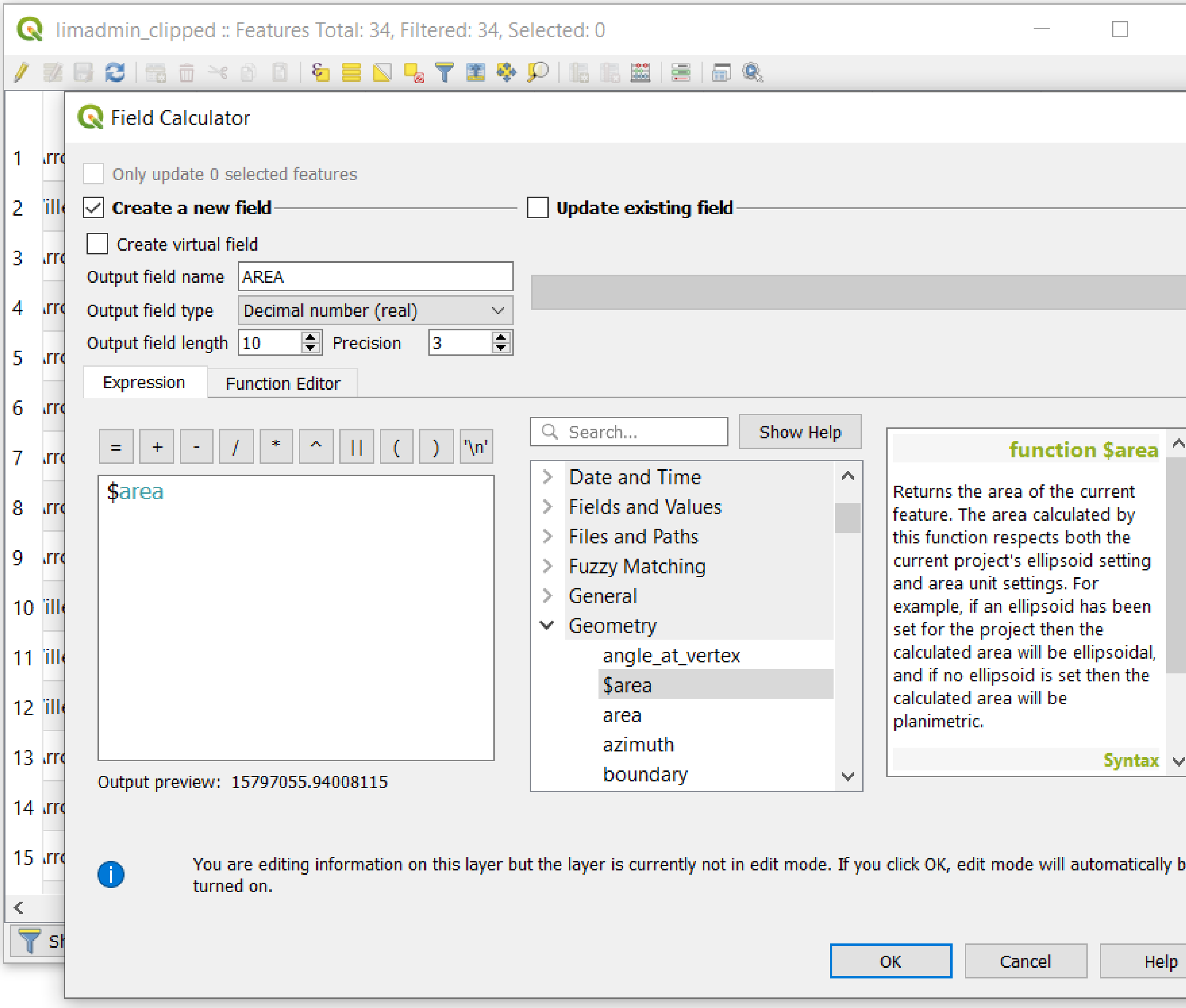

Figure 7.11: Calculate the area of Montreal's boroughs

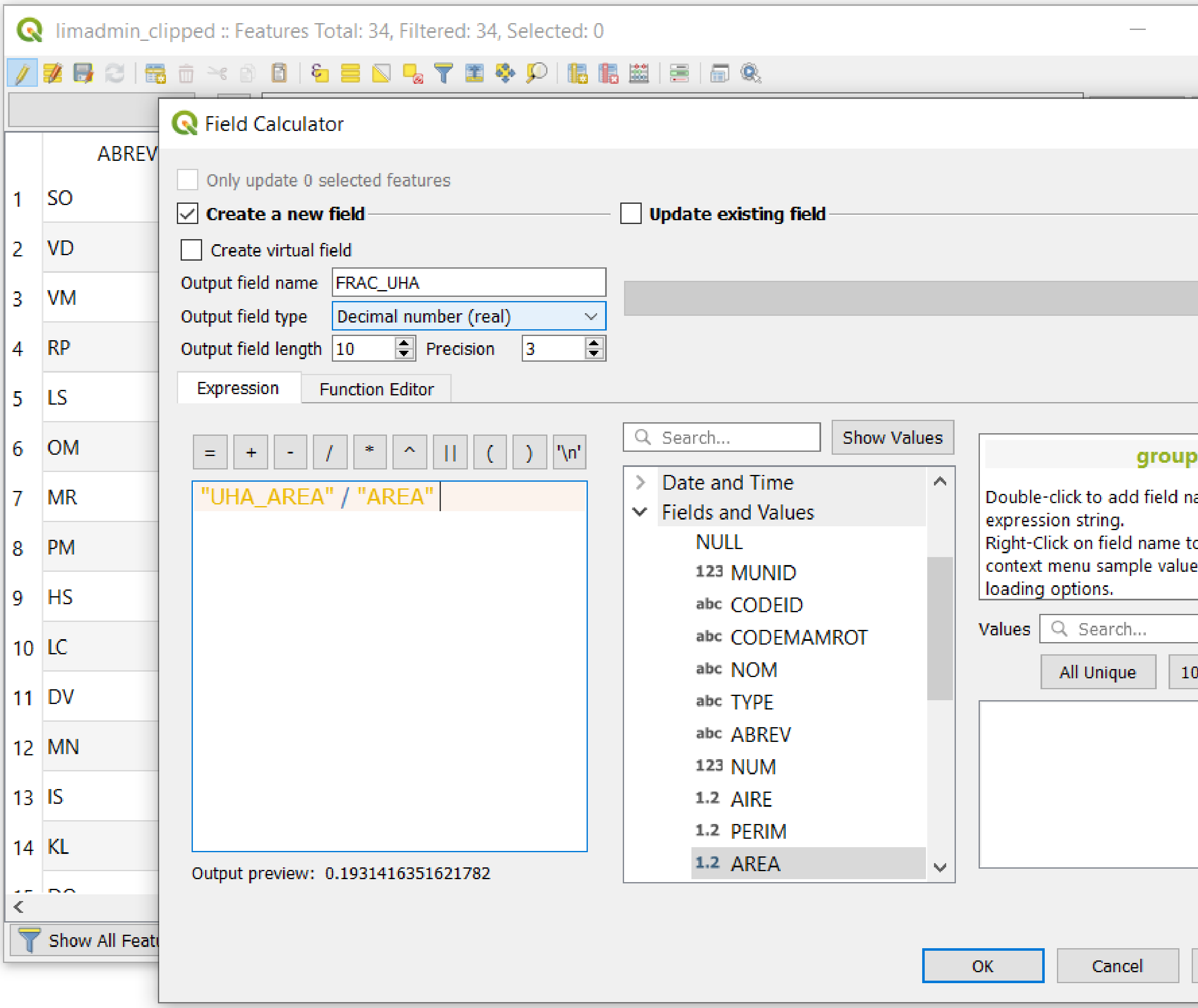

Figure 7.12: Calculate the area fraction of UHI

Right click on `limadmim_clipped.shp`, select **Properties**, then **Symbology**, and set the parameters according to the following figure.

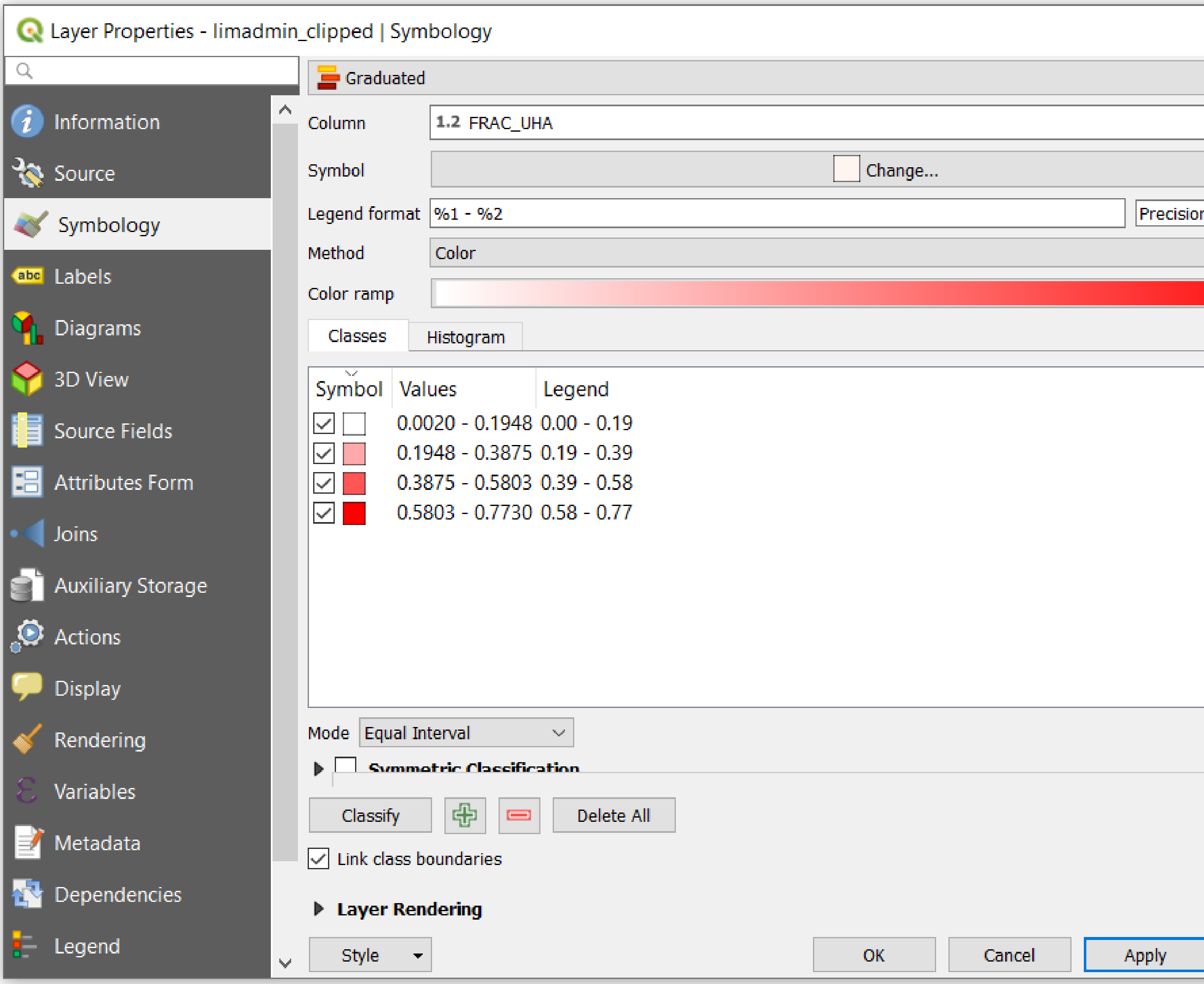

Figure 7.13: Changing the symbology of a layer

Chapter 8

References

Data	File name	Source
RSQA - liste des stations	coordonnees-stations-rsqa.csv	Portail données ouvertes Montréal
RSQA - polluants gazeux 2013-07-01 à 2013-12-31	pollutants_average_12_31_2019_13H.csv	Portail données ouvertes Montréal
Grands parcs	grandparcs.kml	Portail données ouvertes Montréal
Îlots de chaleur	ilotschaleur.json	Portail données ouvertes Montréal
Limite administrative de l'agglomération de Montréal (Arrondissement et Ville liée)	LIMADMIN.shp	Portail données ouvertes Montréal
Limite terrestre	terre.shp	Portail données ouvertes Montréal

Data	File name	Source