Section : . Data Acquisition

Module : . Topology Maintenance

Vector topology in Context

"Topology expresses the spatial relationships between connecting or adjacent vector features (points, polylines, and polygons) in a GIS."

Vector data has to conform to the topological model e.g. two lines in a roads vector layer that intersect should be joined by the same node. Topology is necessary for carrying out some types of spatial analysis, such as network analysis." In this module, we will inspect vector data for topology errors and explore ways in which to clean the dataset so that it conforms to the vector topology model.

The vector topology model stipulates that:

- Boundaries should not cross each other (i.e. boundaries which would cross must be split at their intersection to form distinct boundaries). Lines can cross each other e.g. bridges over rivers.
- Lines and boundaries share nodes only if their endpoints are identical.
- Common area boundaries should appear only once.
- Areas must be explicitly closed.
- It is recommended that lines, points and polygon should be stored separately.



You try:

Goal: To learn the basics of maintaining vector topology and how to clean data so that it adheres to the vector topology model.

- Load the raster layer and vector layer into QGIS. Ensure that the vector layer is rendered on top of the raster.
- Symbolize the vector layer using the Roads `style` options specified. What can you infer about the layer ?
- Open the properties of the vector layer and choose *layer rendering* under the *Symbology tab*.
- On *blending* mode select *feature blending* and choose *multiply* from the drop-down.
- Zoom to a scale of 2500. What do you see from the roads ?
- Are there duplicates from the roads layer ?. How would you define a duplicate?
- Open the attributes and note the number of features that are available from the layer.

Part 1 - Removing Duplicates

- In the *Processing* Menu select toolbox and search for the algorithm *Delete Duplicates geometries* and run it.
- Open the attributes of the new layer generated above and note the number of features available. How many features have been deleted from the roads_unclean layer?
- For the roads, you thought were duplicates, have they been cleaned up ?. Was your definition of duplicates concise?
- Use the **topology checker** to validate against duplicates. How many duplicates are still visible?
- Remove the roads_unclean layer from QGIS.

Part 2 - Checking other topology errors

- Using the **topology checker** plugin configure it to check for the following errors `dangling nodes`, `invalid geometries`, `multi-part geometries` and `pseudos`
- You can configure the tool to check for each error separately or all of them at the same time.
- After running `Validate All` navigate to each error by clicking each row in the results panel.
- What can you say about the topological errors that have been identified?
- Click on configure in the topology checker panel and delete all the rules defined. Run validate all to clean up the results panel

Part 3 - Cleaning up

- Use the processing tool to `Reproject` the road layer to the specified CRS.
- What other way exists to reproject (change the CRS) vector layers in QGIS and why is it important
- Remove the unprotected layer from QGIS.
- In the processing toolbox use the algorithm `v.clean`.
- Fill in the tool appropriately selecting the cleaning tool - `snap`
- Use the threshold of 20 to experiment with what the tool does. What are the units of this threshold?

NB: Run the tool iteratively using the cleaning tool `snap` and the threshold with small values that increment on each iteration.
Use the output of each iteration as input into the next iteration.

Part 4 - Tracing

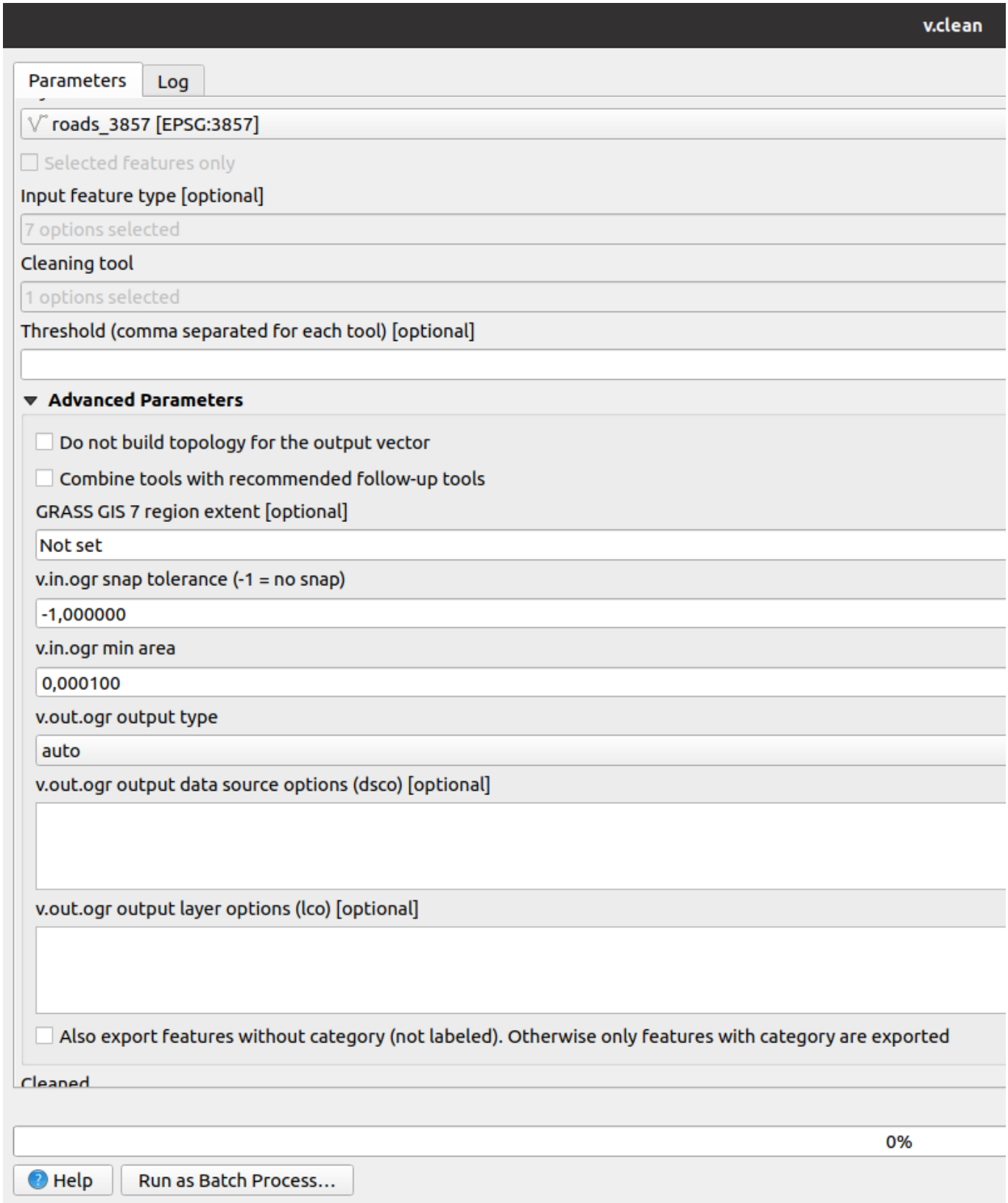
- Use the layer creation options specified in `Roads Tracing` to create a new vector layer.
- Enable the Snapping toolbar.
- On the Snapping toolbar icon click enable snapping.
- Click Advanced configuration on the snapping toolbar icon. (The drop-down icons next to the Magnet)
- Select `0pen snapping options`.
- Activate snapping for both layers (the newly created one and the reprojected vector layer).
- Set the tolerance to values between 40-60 and the Units to pixels.
- Enable `Topological Editing`, `Snapping on intersection` and close the dialog.
- Toggle edit the layer you have recently created.
- Click on add line and start digitizing following the outline of the roads layer and fixing areas that do not look nice.

- On the Advanced toolbar select `Enable tracing`
- Trace the outlines of the existing roads and click on the endpoints

Bonus Section

- Investigate other plugins or tools native to QGIS which can be used to clean the vector layer.
- What is the difference between using polygons with the topology checker.

Name	Expectation
Background Layer	Google Satellite or topology_cleaning/ 2821AD_21_2014_747_RGB_RECT.tif
data	roads_unclean
Validity tools (Plugins)	Topology checker, Geometry validity
CRS	EPSG:3857
Cleaning tools	v.clean
Roads Style	Simple line, Colour - #e31a1c, Stroke width - 0.86
Roads Tracing - Geometry	Multiline
Roads Tracing - CRS	EPSG:3857



More about vector topology

The vector topology model stipulates that:

- Boundaries should not cross each other (i.e. boundaries which would cross must be split at their intersection to form distinct boundaries). Lines can cross each other e.g. bridges over rivers.
- Lines and boundaries share nodes only if their endpoints are identical.
- Common area boundaries should appear only once.
- Areas must be explicitly closed.
- It is recommended that lines, points and polygon should be stored separately.

Data that does not adhere to the topological model inhibits spatial analysis to be done and will always give erroneous results. Typical errors are:

- Undershoots - occurs when digitised roads do not intersect.
- Overshoots - occurs if a line ends beyond the line it should.
- Slivers - occurs when the vertices of two or more polygons do not match up on their borders.

GRASS GIS is a fully topological GIS package and using the GRASS algorithms to check and clean data helps in generating clean vector data which can be used in an analysis.

Check your knowledge:

1. Why is vector topology important:
   a. *It enhances spatial analysis because distances across a road network can be accurate*
   b. *It ensures that the symbology of the layers is accurate*
   c. *It allows vector data to be converted into different formats i.e. KML*
2. Which type of data should be checked for topology errors:
   a. *A satellite image that contains roads*
   b. *All layers in QGIS should checked*
   c. *A GIS vector dataset where each polygon represents a type of building*
3. When cleaning a polygon is it important to activate snapping:
   a. *True*
   b. *False*

Module video tutorial:

- https://youtu.be/RdQQp2Rzgrs

Further reading:

- Topology: https://docs.qgis.org/2.14/en/docs/gentle_gis_introduction/topology.html
- Plugins_topology_checker: http://docs.qgis.org/2.14/pl/docs/user_manual/plugins/plugins_topology_checker.html

Download the sample data for the lesson from http://changelog.qgis.org/media/images/lesson/worksheet/external_data/e3d74596f51c26a98abe017d37c481251b2ace35.zip.