POLITECNICO DI MILANO
GEOlab, Como Campus





A GRASS-based automated procedure to compare OpenStreetMap and authoritative road network datasets

Maria Antonia Brovelli, Marco Minghini, Monia Elisa Molinari

Politecnico di Milano, Como Campus, DICA, via Valleggio 11, 22100 Como (Italy)

Motivation of the work – VGI & OSM quality

- Increasing popularity of OpenStreetMap (OSM) as today's most notable Volunteered Geographic Information (VGI) project on the Internet
- Increasing concern on VGI (and OSM) data quality:
  - spatial accuracy
  - temporal accuracy
  - semantic accuracy
  - up-to-dateness
  - consistency
  - fitness-for-use & fitness-for-purpose
  - ...
- Increasing availability of open data from NMAs and CSC that can be used as a source of comparison for VGI (and OSM) data:
  - comparing two spatial datasets against each other is a challenging geocomputation problem!

Motivation of the work – OSM comparisons

- Literature provides plenty of works assessing or comparing OSM quality against that of authoritative datasets:
  - strongly focused on road network
  - mainly investigating OSM positional accuracy
  - OSM compared to data from NMA (UK Ordnance Survey, French NMA, USGS TNM/TIGER, etc.) and CSC (Navteq, TeleAtlas, etc.)
  - semi- or fully-automated
  - results from poor to very good
- Comparison techniques are very strong and fit for purpose, but mostly application and dataset specific:
  - hard to replicate
  - difficult to extend to other dataset comparisons

Our methodology

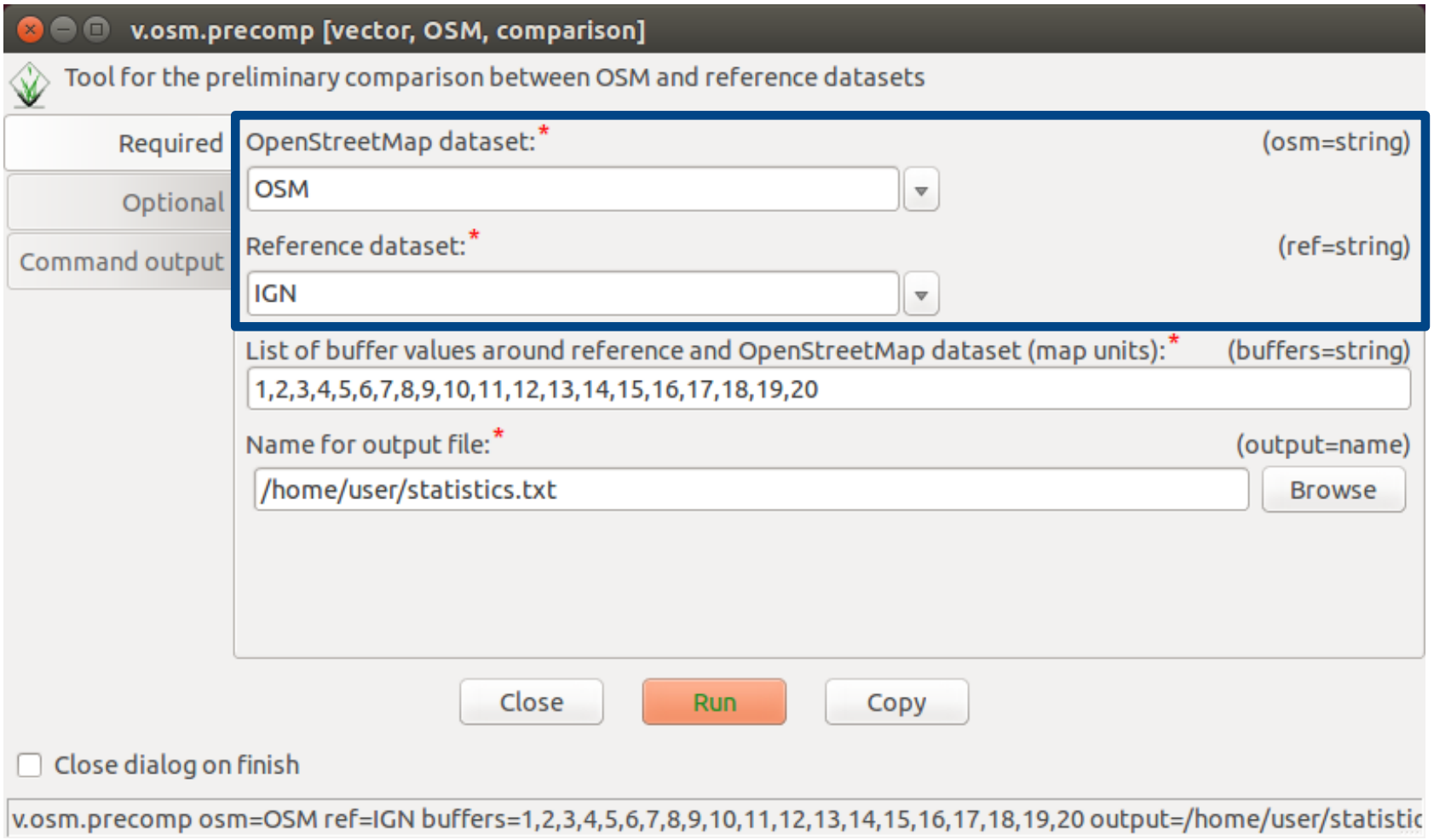
- ✔ Novel methodology to compare OSM and authoritative road datasets:
  - ➔ fully automated
  - ➔ focused on spatial accuracy and completeness
  - ➔ flexible, i.e. not developed for a specific dataset
    - ✗ made of required and optional operations
    - ✗ users can define the value of the parameters involved to adapt the procedure to their specific authoritative datasets
    - ✗ users are supposed to be familiar with the authoritative dataset used as reference
  - ➔ built with FOSS4G (Free and Open Source Software for Geospatial)
    - ✗ reusable and extensible in case of need

Our methodology – Overview

- Currently developed as 3 GRASS GIS modules:
  - written in Python
  - available with a Graphical User Interface (GUI)
- Comparison between OSM and reference (IGN) road network datasets composed of 3 consecutive steps:
  - 1. Preliminary comparison of the datasets and computation of global statistics
  - 2. Geometric preprocessing of the OSM dataset to extract a subset which is fully comparable with the IGN dataset
  - 3. Evaluation of OSM spatial accuracy using a grid-based approach

Step 1: Preliminary comparison of the datasets

- ✔ Import and select the OSM and IGN datasets [*required*]

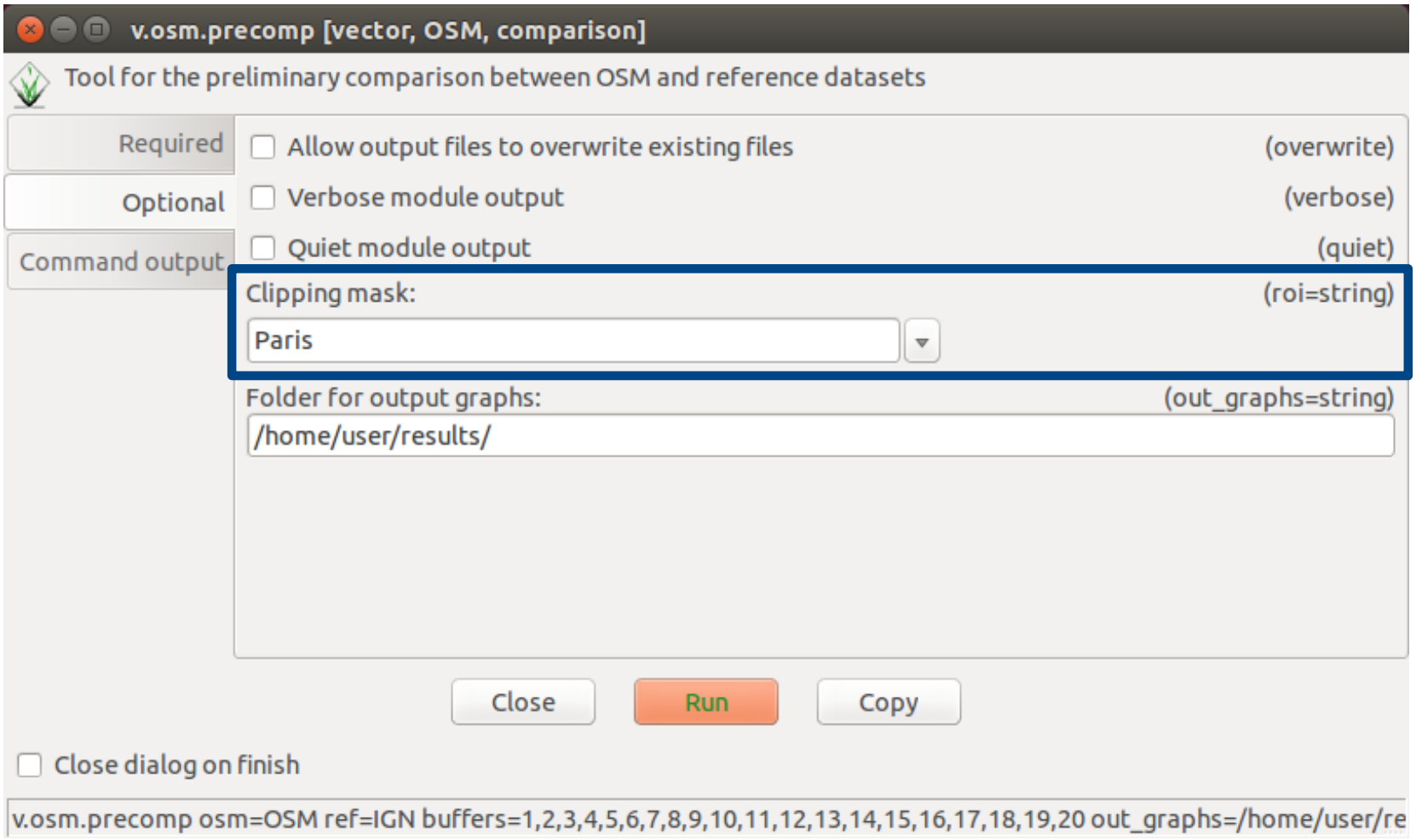


Step 1: Preliminary comparison of the datasets

- ✔ Import and select the OSM and IGN datasets [*required*]

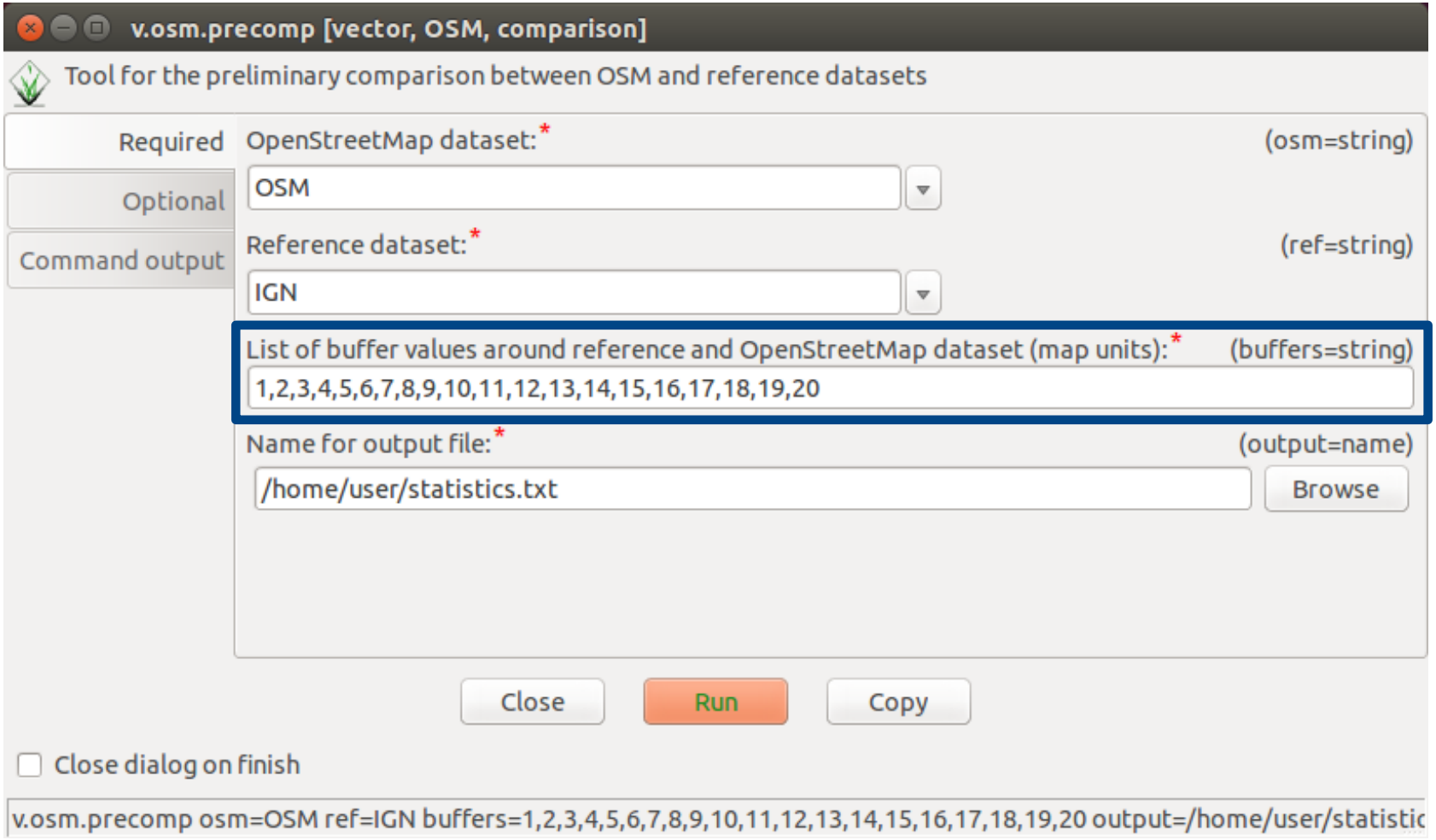
Step 1: Preliminary comparison of the datasets

- ✔ Import and select the OSM and IGN datasets [*required*]
- ✔ If the extent of the OSM and/or IGN datasets is larger than the one of interest, import a vector layer to be used as clipping mask [*optional*]



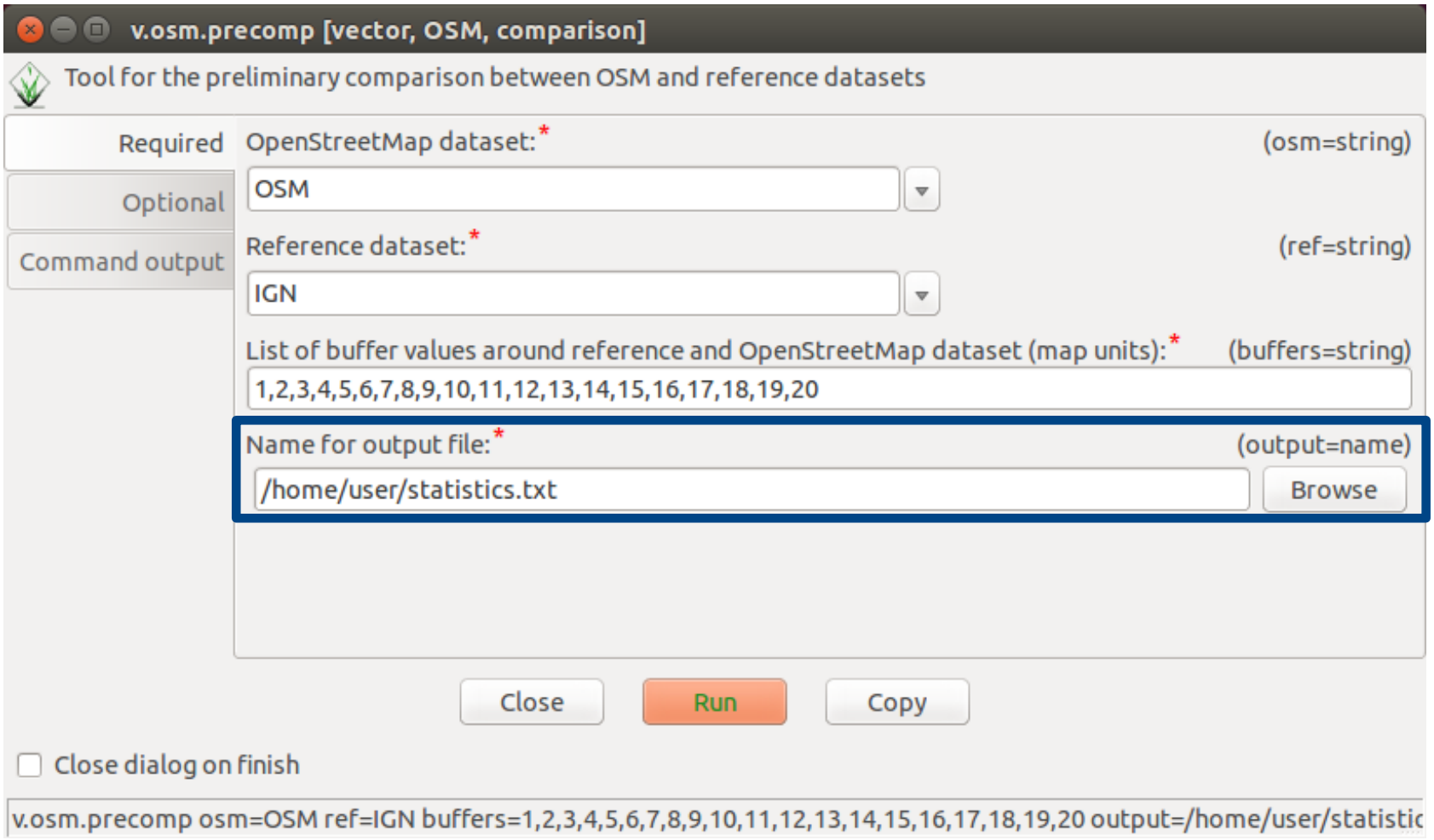
Step 1: Preliminary comparison of the datasets

- Apply a set of buffers of user-specified width around both the IGN and OSM datasets, to compute the length and the length percentage of the OSM and IGN datasets included in the buffer [*required*]



Step 1: Preliminary comparison of the datasets

- ✔ Compute also the total length of OSM and IGN datasets and their length difference, both in map units and percentage [*required*]
  - ➔ output values are returned in a text file



Step 1: Preliminary comparison of the datasets

- Compute also the total length of OSM and IGN datasets and their length difference, both in map units and percentage [*required*]
  - output values are returned in a text file

```
REF length: 2686373.1 m
OSM length: 3124627.0 m
REF-OSM difference: -438253.9 m (-16.3%)

BUFFER(m)|OSM_IN(m)|OSM_IN(%%)|OSM_OUT(m)|OSM_OUT(%%)|REF_IN(m)|REF_IN(%%)|REF_OUT(m)|REF_OUT(%%)
1.0|1374755.9|44.0|1749871.2|56.0|1366471.0|50.9|1319902.1|49.1
2.0|2014259.9|64.5|1110367.2|35.5|1982713.7|73.8|703659.4|26.2
3.0|2298072.4|73.5|826554.6|26.5|2223153.5|82.8|463219.6|17.2
4.0|2464185.0|78.9|660442.0|21.1|2329270.3|86.7|357102.8|13.3
5.0|2582784.2|82.7|541842.9|17.3|2387687.7|88.9|298685.4|11.1
6.0|2671758.8|85.5|452868.2|14.5|2424463.5|90.3|261909.6|9.7
7.0|2738327.0|87.6|386300.0|12.4|2451476.9|91.3|234896.2|8.7
8.0|2792053.8|89.4|332573.2|10.6|2471557.1|92.0|214816.0|8.0
9.0|2828903.0|90.5|295724.1|9.5|2488514.1|92.6|197859.0|7.4
10.0|2859512.1|91.5|265114.9|8.5|2501974.7|93.1|184398.4|6.9
11.0|2886190.1|92.4|238436.9|7.6|2513592.9|93.6|172780.2|6.4
12.0|2908071.9|93.1|216555.1|6.9|2523138.5|93.9|163234.6|6.1
13.0|2925602.0|93.6|199025.1|6.4|2532070.5|94.3|154302.6|5.7
14.0|2941922.8|94.2|182704.2|5.8|2540322.9|94.6|146050.2|5.4
15.0|2956112.7|94.6|168514.3|5.4|2548274.0|94.9|138099.1|5.1
16.0|2967813.5|95.0|156813.5|5.0|2555431.5|95.1|130941.6|4.9
17.0|2977318.7|95.3|147308.3|4.7|2562238.1|95.4|124135.0|4.6
18.0|2986371.8|95.6|138255.2|4.4|2568276.5|95.6|118096.6|4.4
19.0|2994833.4|95.8|129793.7|4.2|2574052.2|95.8|112320.9|4.2
20.0|3001796.0|96.1|122831.1|3.9|2579434.1|96.0|106939.0|4.0
```

Step 1: Preliminary comparison of the datasets

- ✔ Compute also the total length of OSM and IGN datasets and their length difference, both in map units and percentage [*required*]
  - ➔ output values are returned in a text file

```
REF length: 2686373.1 m
OSM length: 3124627.0 m
REF-OSM difference: -438253.9 m (-16.3%)
```

✗ ≅450 km more in OSM than IGN dataset!

```
BUFFER(m)|OSM_IN(m)|OSM_IN(%%)|OSM_OUT(m)|OSM_OUT(%%)|REF_IN(m)|REF_IN(%%)|REF_OUT(m)|REF_OUT(%%)
1.0|1374755.9|44.0|1749871.2|56.0|1366471.0|50.9|1319902.1|49.1
2.0|2014259.9|64.5|1110367.2|35.5|1982713.7|73.8|703659.4|26.2
3.0|2298072.4|73.5|826554.6|26.5|2223153.5|82.8|463219.6|17.2
4.0|2464185.0|78.9|660442.0|21.1|2329270.3|86.7|357102.8|13.3
5.0|2582784.2|82.7|541842.9|17.3|2387687.7|88.9|298685.4|11.1
6.0|2671758.8|85.5|452868.2|14.5|2424463.5|90.3|261909.6|9.7
7.0|2738327.0|87.6|386300.0|12.4|2451476.9|91.3|234896.2|8.7
8.0|2792053.8|89.4|332573.2|10.6|2471557.1|92.0|214816.0|8.0
9.0|2828903.0|90.5|295724.1|9.5|2488514.1|92.6|197859.0|7.4
10.0|2859512.1|91.5|265114.9|8.5|2501974.7|93.1|184398.4|6.9
11.0|2886190.1|92.4|238436.9|7.6|2513592.9|93.6|172780.2|6.4
12.0|2908071.9|93.1|216555.1|6.9|2523138.5|93.9|163234.6|6.1
13.0|2925602.0|93.6|199025.1|6.4|2532070.5|94.3|154302.6|5.7
14.0|2941922.8|94.2|182704.2|5.8|2540322.9|94.6|146050.2|5.4
15.0|2956112.7|94.6|168514.3|5.4|2548274.0|94.9|138099.1|5.1
16.0|2967813.5|95.0|156813.5|5.0|2555431.5|95.1|130941.6|4.9
17.0|2977318.7|95.3|147308.3|4.7|2562238.1|95.4|124135.0|4.6
18.0|2986371.8|95.6|138255.2|4.4|2568276.5|95.6|118096.6|4.4
19.0|2994833.4|95.8|129793.7|4.2|2574052.2|95.8|112320.9|4.2
20.0|3001796.0|96.1|122831.1|3.9|2579434.1|96.0|106939.0|4.0
```

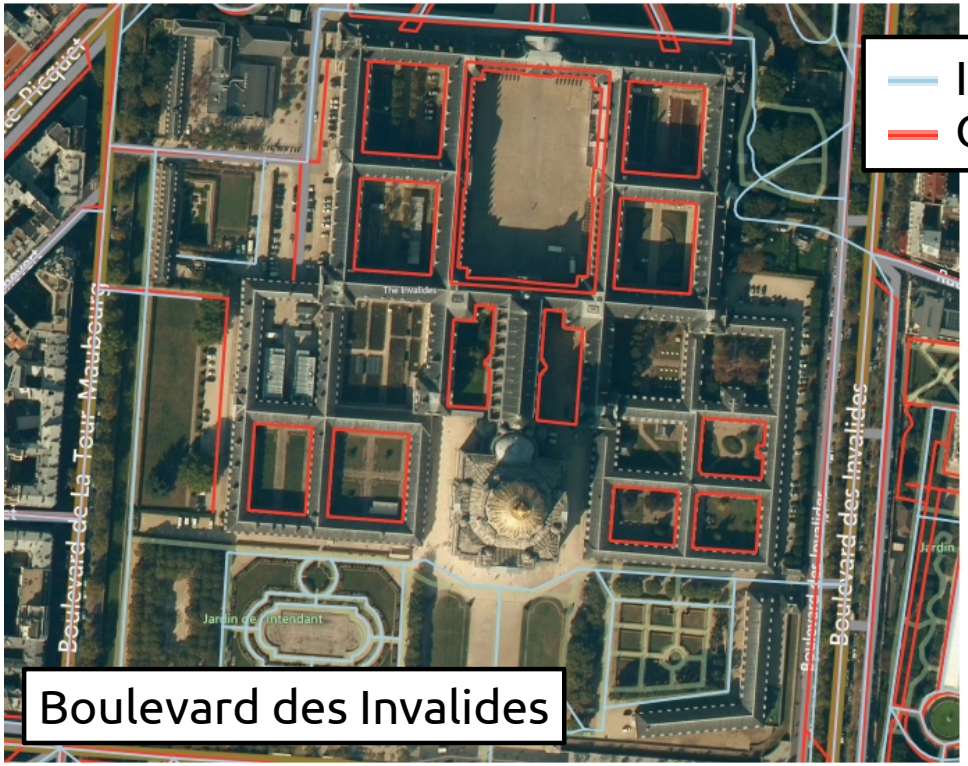
Step 1: Preliminary comparison of the datasets

- ✔ Compute also the total length of OSM and IGN datasets and their length difference, both in map units and percentage [*required*]
  - ➔ output values are returned in a text file

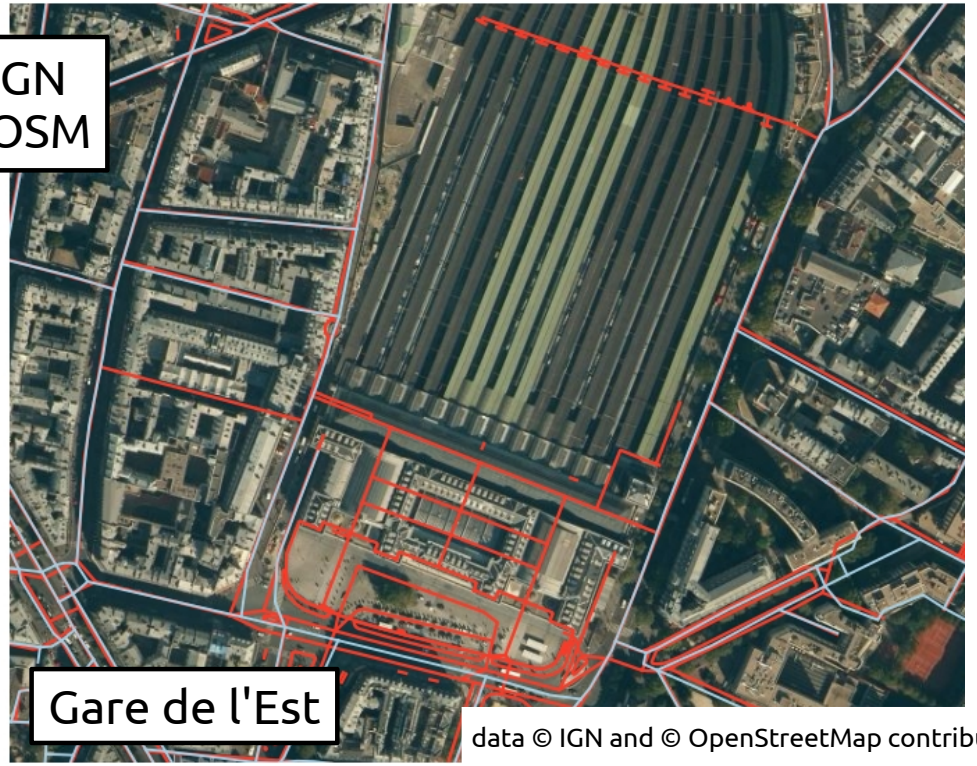
```
REF length: 2686373.1 m
OSM length: 3124627.0 m
REF-OSM difference: -438253.9 m (-16.3%)
```

✗ ≅450 km more in OSM than IGN dataset!

  - ➔ more footways and pedestrian routes mapped in OSM



Boulevard des Invalides



Gare de l'Est

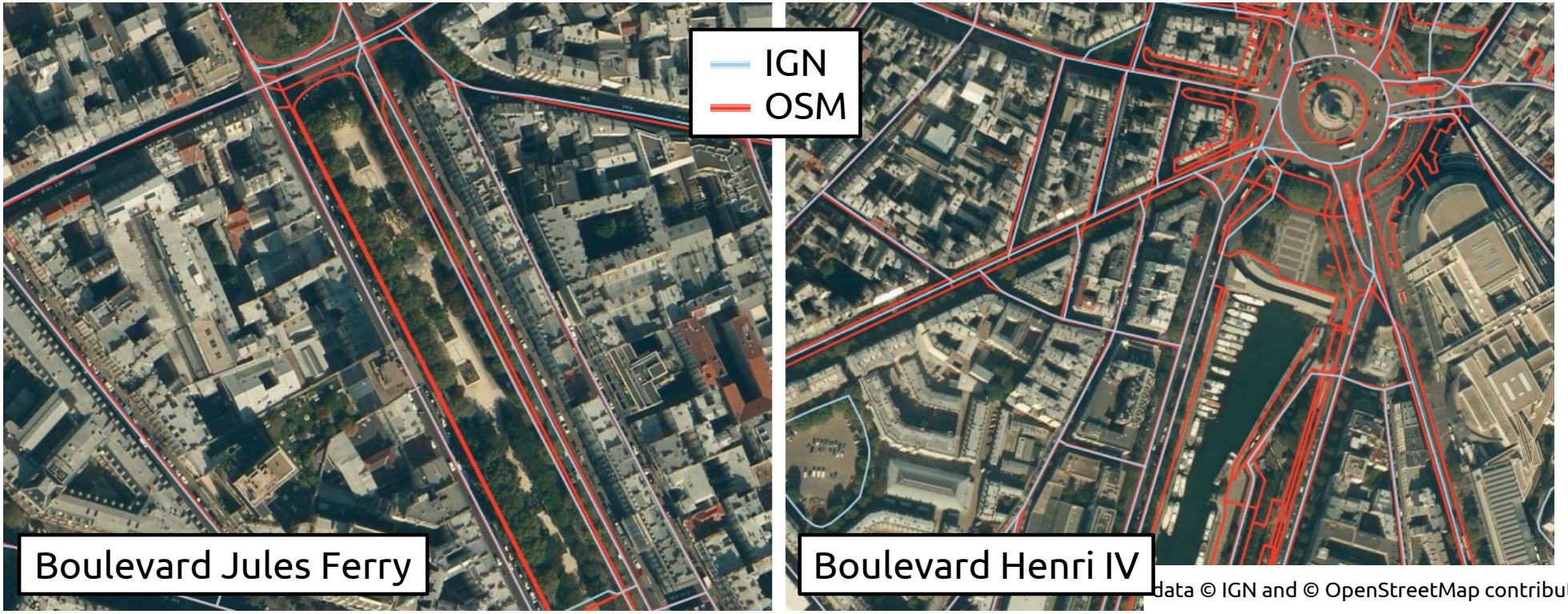
Step 1: Preliminary comparison of the datasets

- ✔ Compute also the total length of OSM and IGN datasets and their length difference, both in map units and percentage [*required*]
  - ➔ output values are returned in a text file

```
REF length: 2686373.1 m
OSM length: 3124627.0 m
REF-OSM difference: -438253.9 m (-16.3%)
```

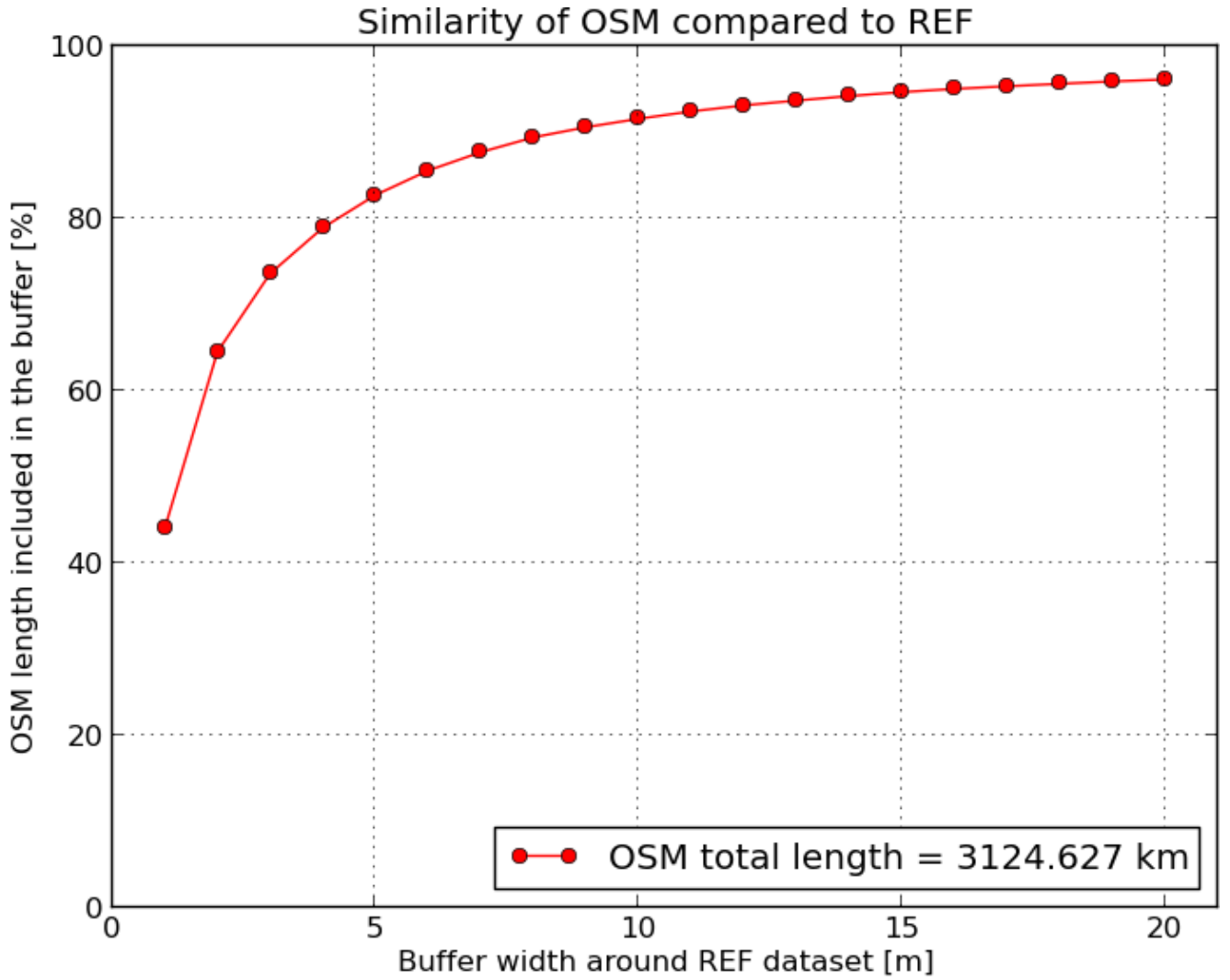
✗ ≅450 km more in OSM than IGN dataset!

  - ➔ cycleways and carriageways mapped as separate highways in OSM



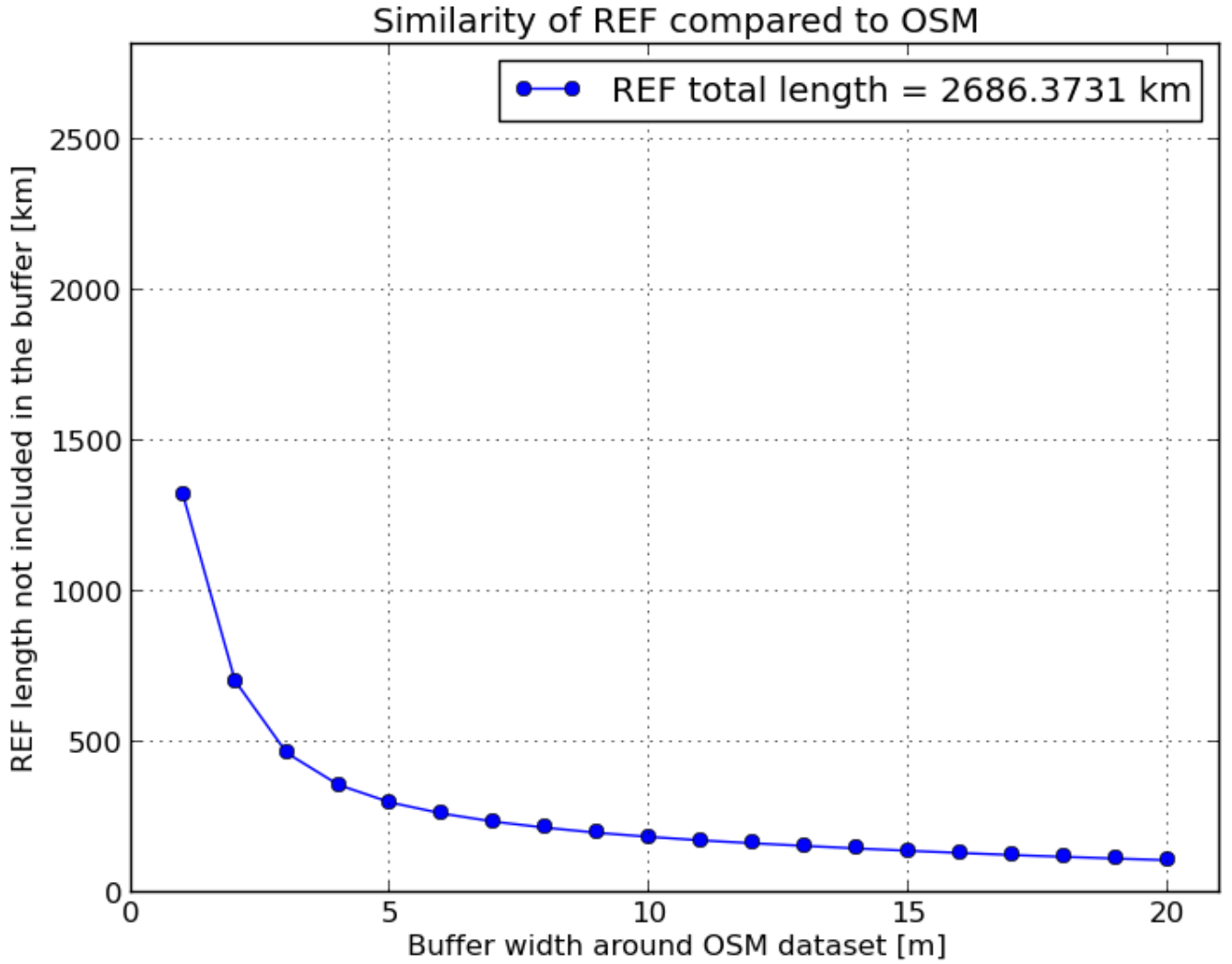
Step 1: Preliminary comparison of the datasets

- Outputs from Step 1 can be used to perform further analysis:
  - sensitivity analysis on the buffer width



Step 1: Preliminary comparison of the datasets

- Outputs from Step 1 can be used to perform further analysis:
  - sensitivity analysis on the buffer width

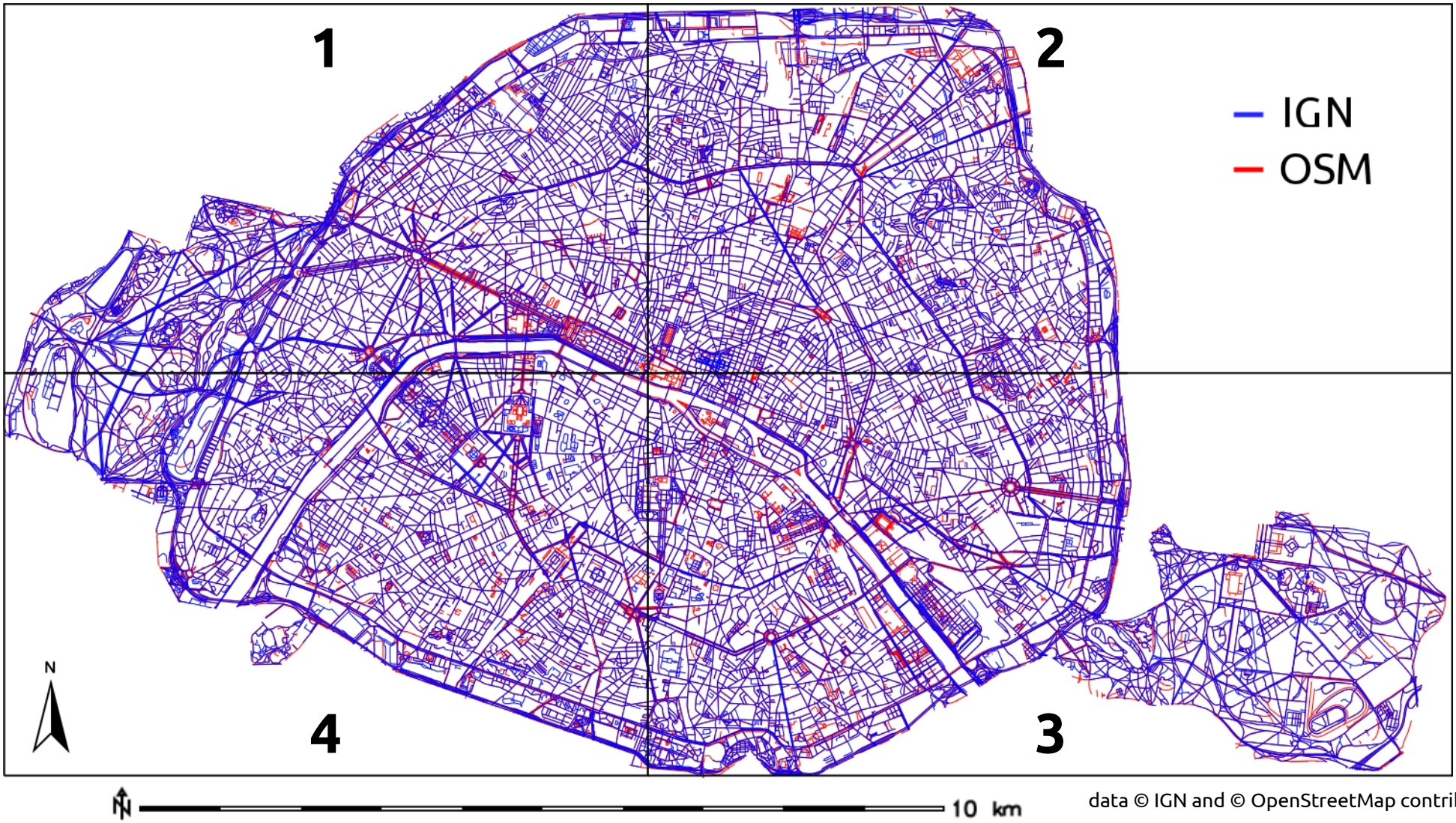


Step 2: preprocessing of the OSM dataset

- ✔ Cleaning of OSM dataset to make it comparable with IGN dataset

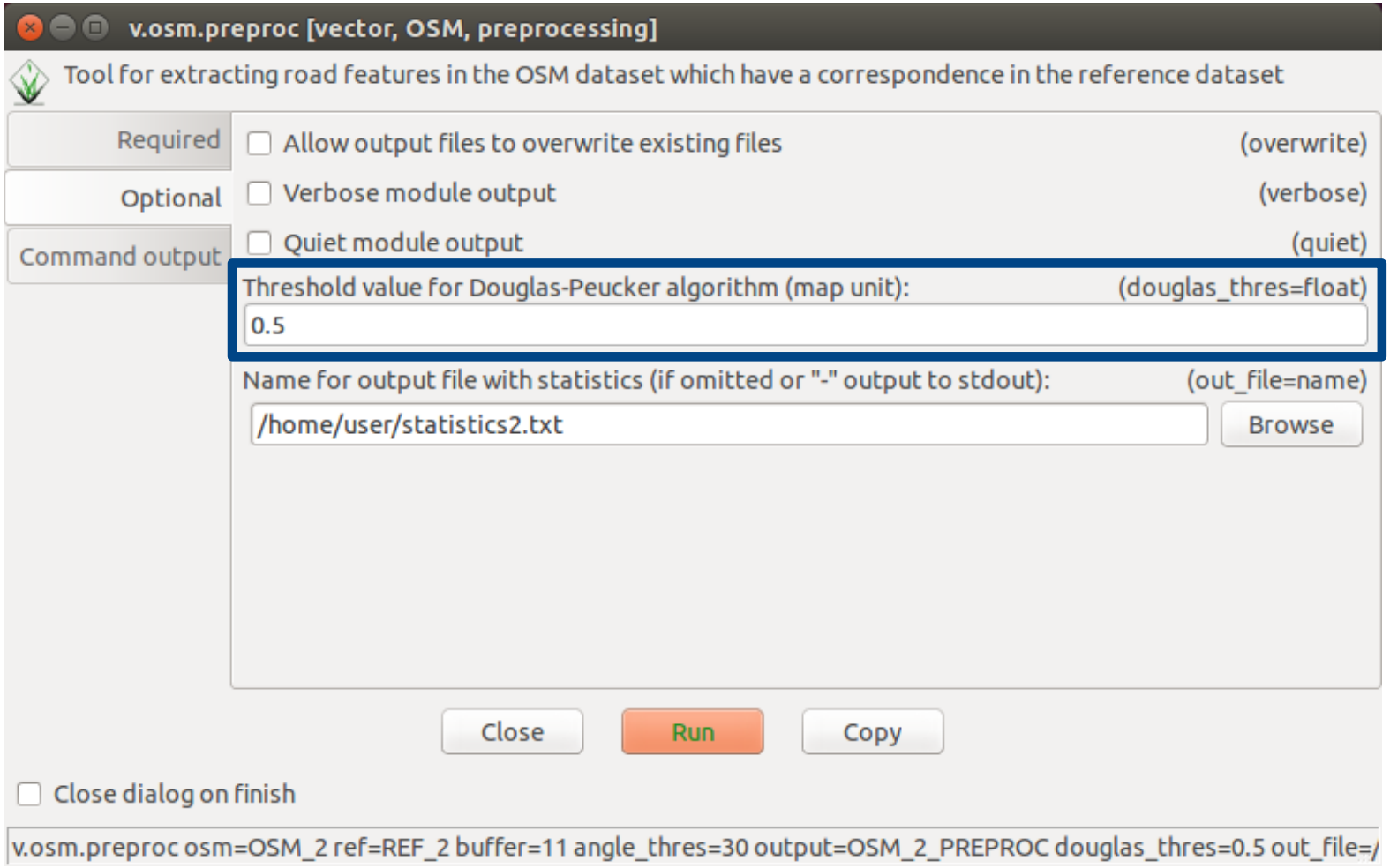
Step 2: preprocessing of the OSM dataset

- ✔ Cleaning of OSM dataset to make it comparable with IGN dataset
  - ➔ computationally intensive – work area divided in 4 sub-areas



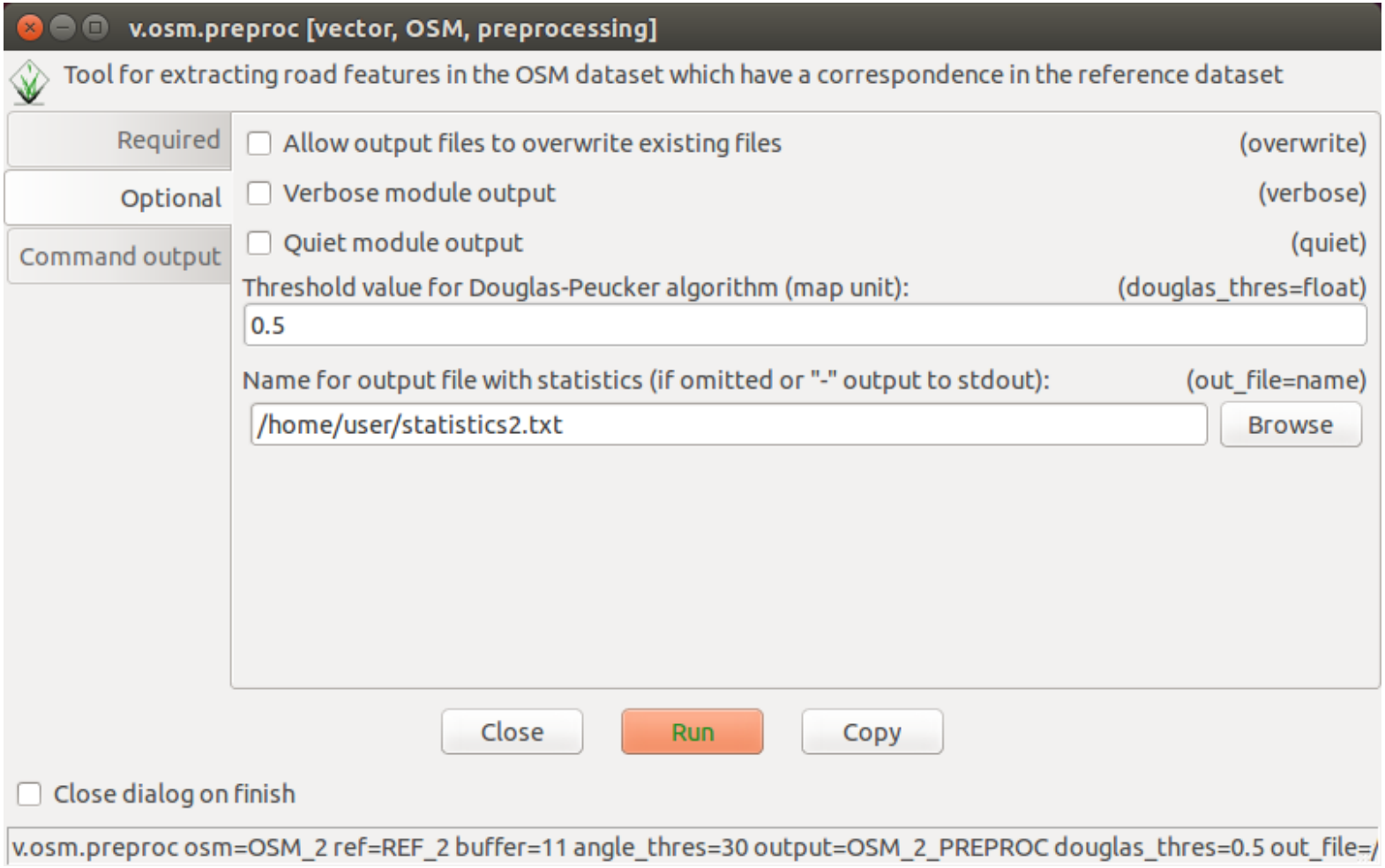
Step 2: preprocessing of the OSM dataset

- ✔ Generalize the IGN dataset with the Douglas-Peucker algorithm [*optional*]
  - ➔ users have to enter the threshold for the algorithm



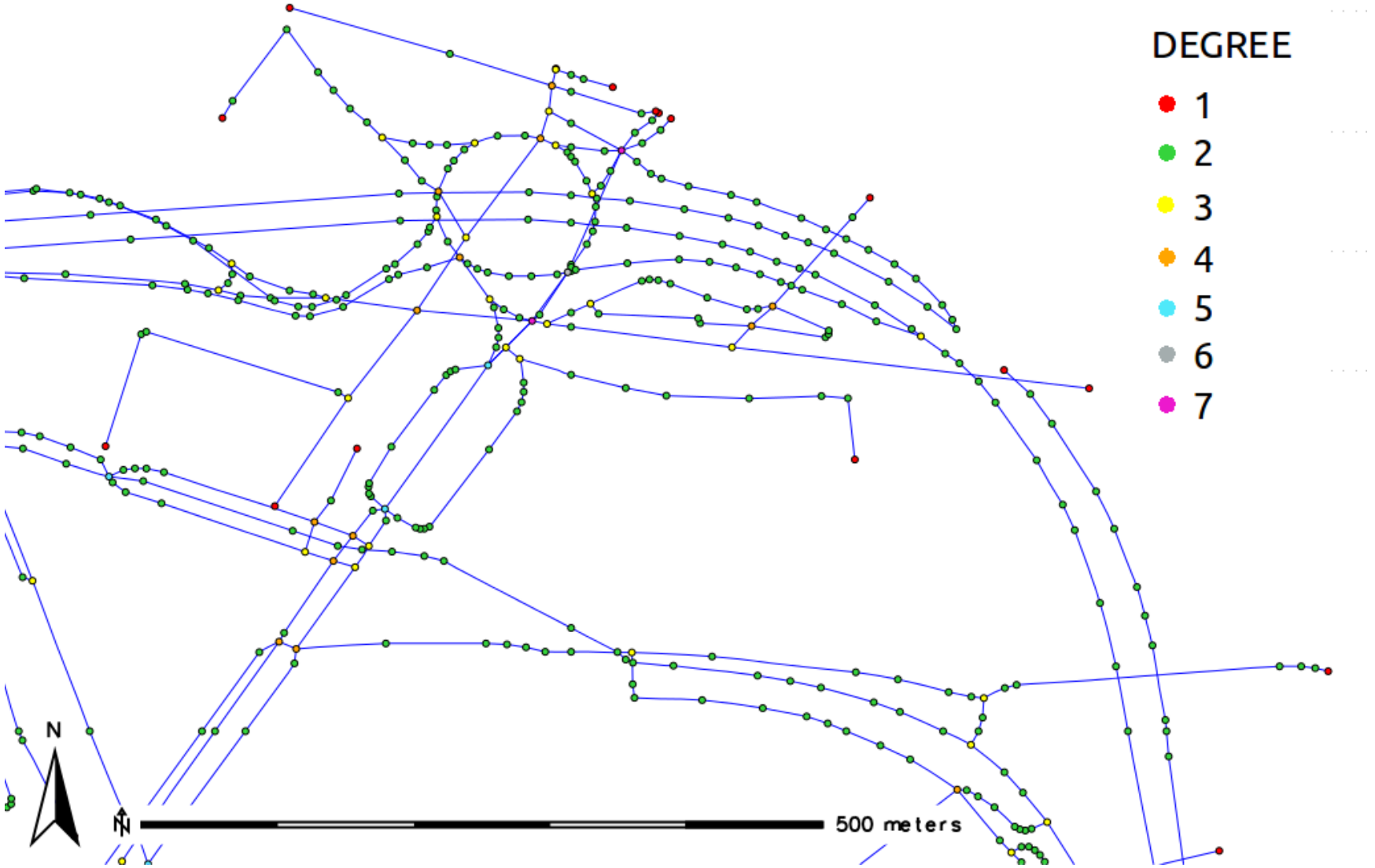
Step 2: preprocessing of the OSM dataset

- ✔ Generalize the IGN dataset with the Douglas-Peucker algorithm [*optional*]
  - ➔ users have to enter the threshold for the algorithm
- ✔ Split the line features of the datasets into segments [*required*]



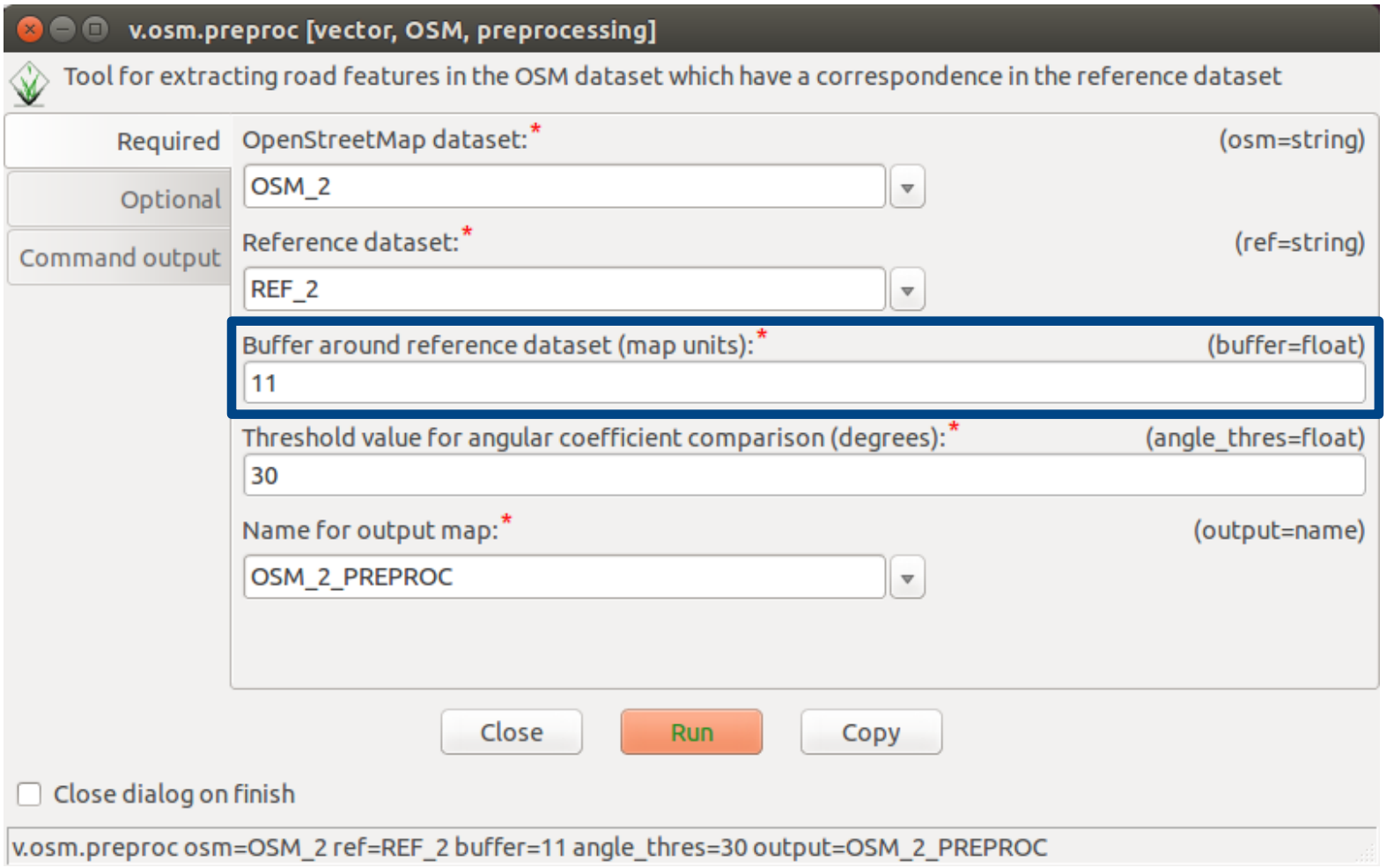
Step 2: preprocessing of the OSM dataset

- ✔ Compute a measure of degree for the nodes of IGN dataset [*required*]
  - ➔ identify the terminal nodes (degree = 1)



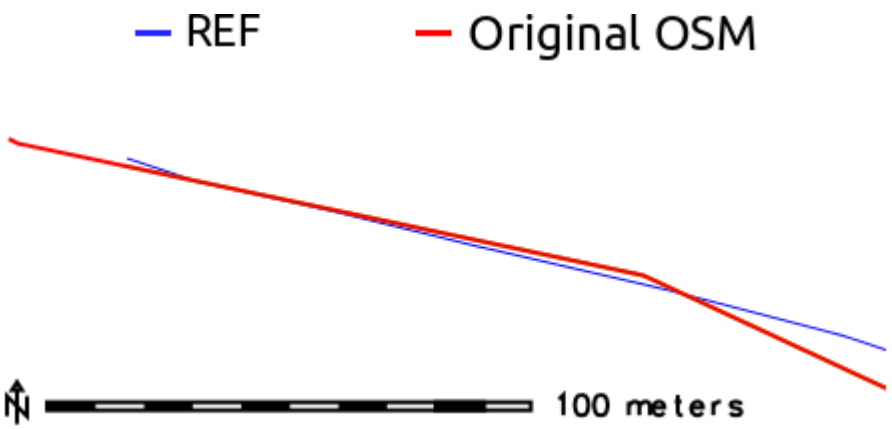
Step 2: preprocessing of the OSM dataset

- ✔ Apply a buffer of user-specified width around the IGN dataset [*required*]
  - ➔ suitable buffer width derived from Step 1
  - ➔ delete all the OSM roads falling outside the buffer



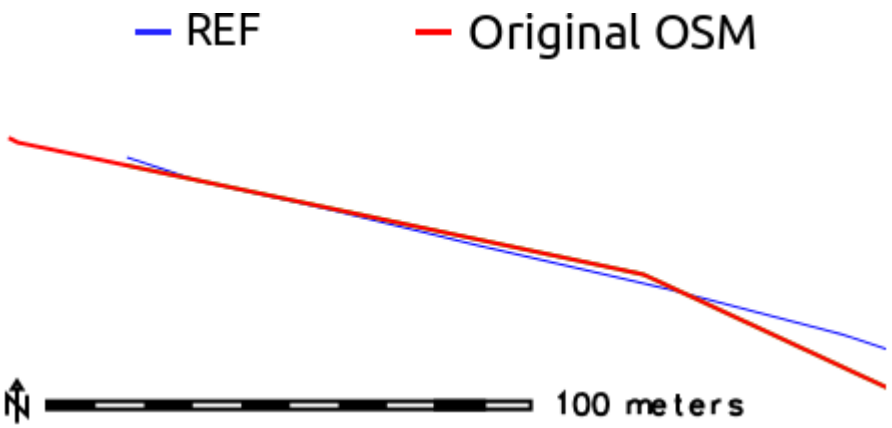
Step 2: preprocessing of the OSM dataset

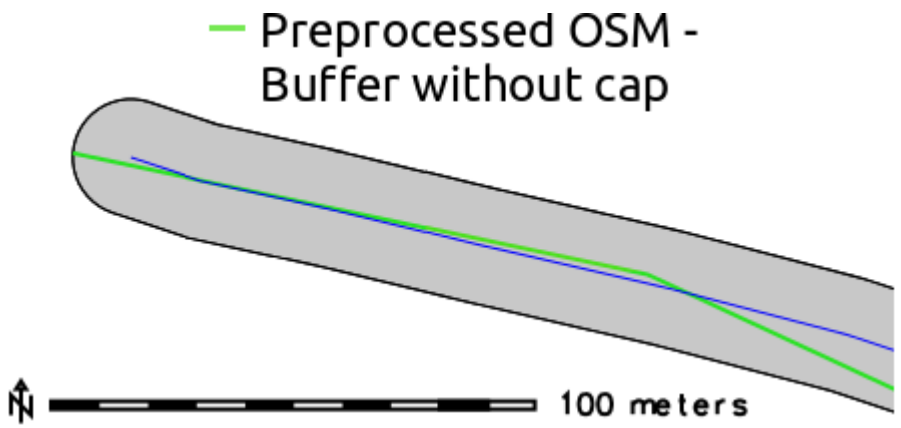
- ✔ Apply a buffer of user-specified width around the IGN dataset [*required*]
  - ➔ suitable buffer width derived from Step 1
  - ➔ delete all the OSM roads falling outside the buffer
  - ➔ buffer is applied without cap around the terminal nodes



Step 2: preprocessing of the OSM dataset

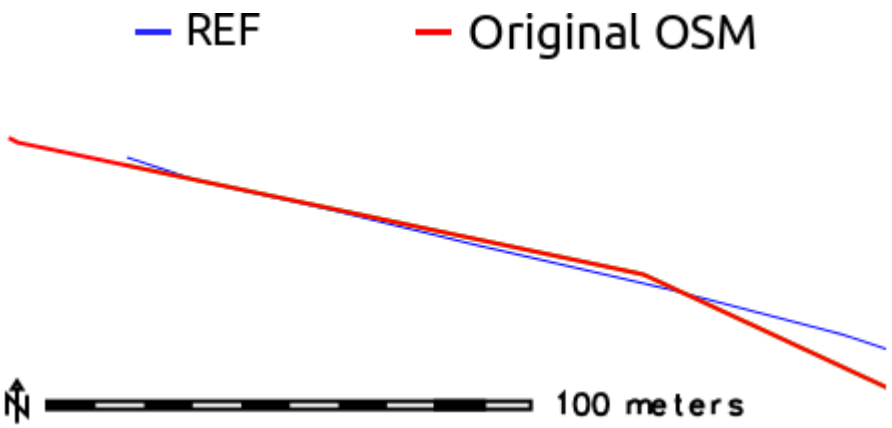
- ✔ Apply a buffer of user-specified width around the IGN dataset [*required*]
  - ➔ suitable buffer width derived from Step 1
  - ➔ delete all the OSM roads falling outside the buffer
  - ➔ buffer is applied without cap around the terminal nodes

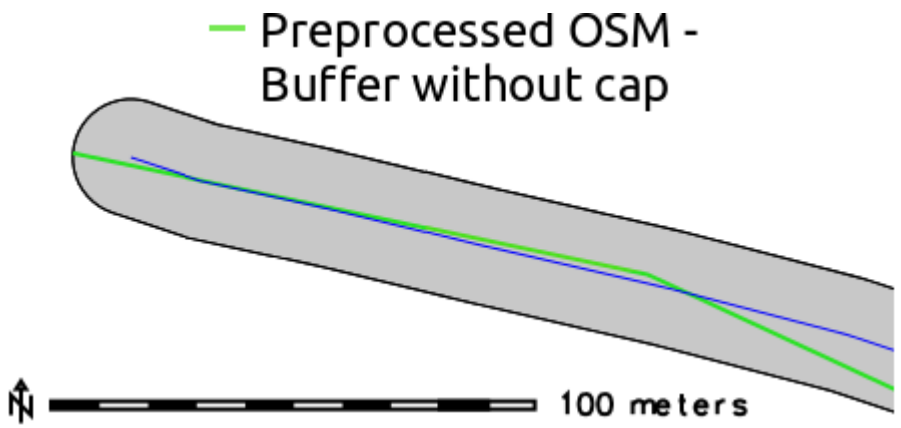


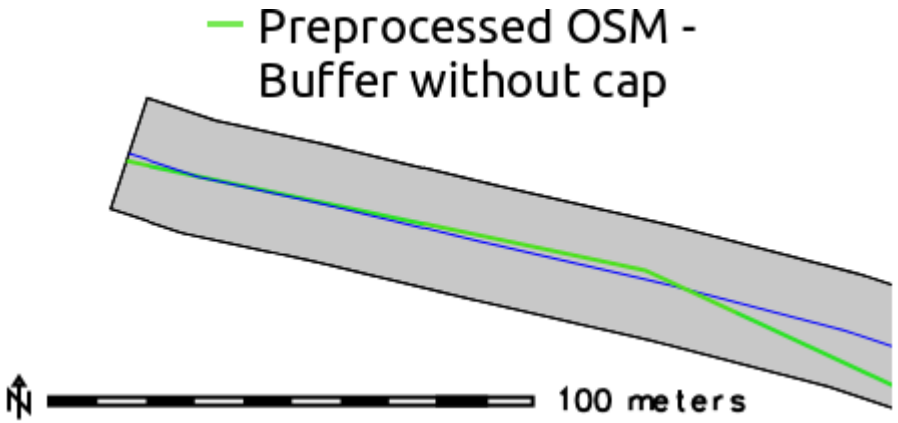


Step 2: preprocessing of the OSM dataset

- ✔ Apply a buffer of user-specified width around the IGN dataset [*required*]
  - ➔ suitable buffer width derived from Step 1
  - ➔ delete all the OSM roads falling outside the buffer
  - ➔ buffer is applied without cap around the terminal nodes

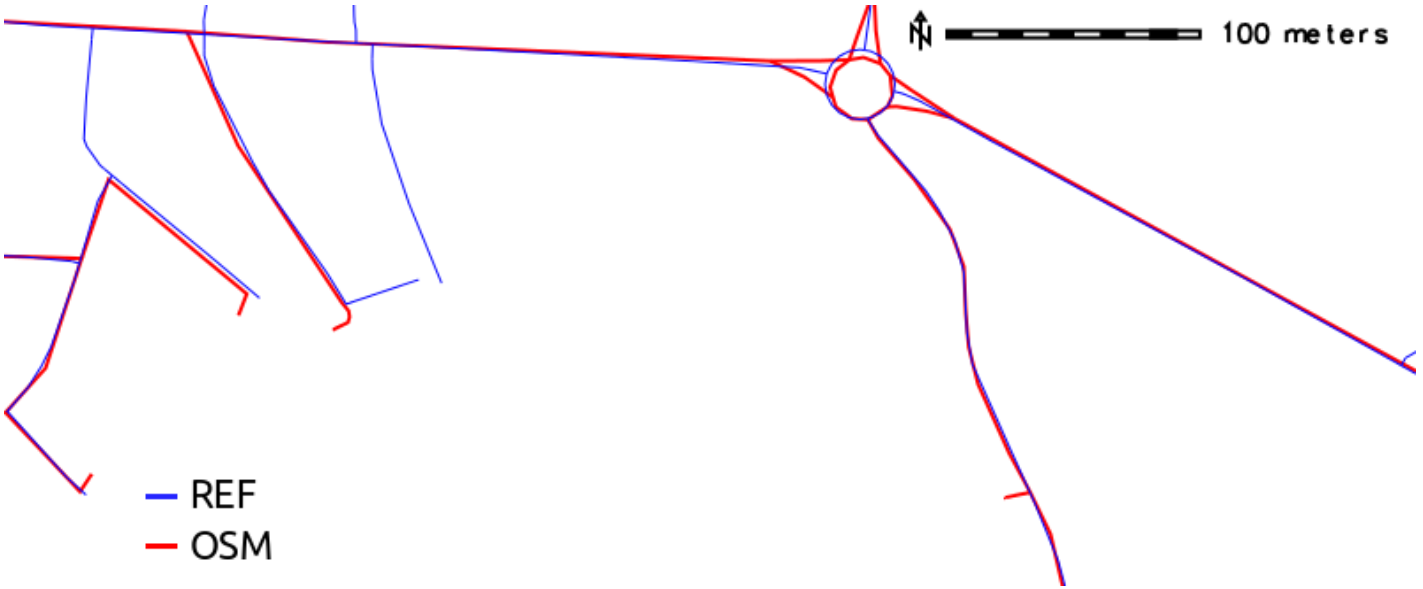






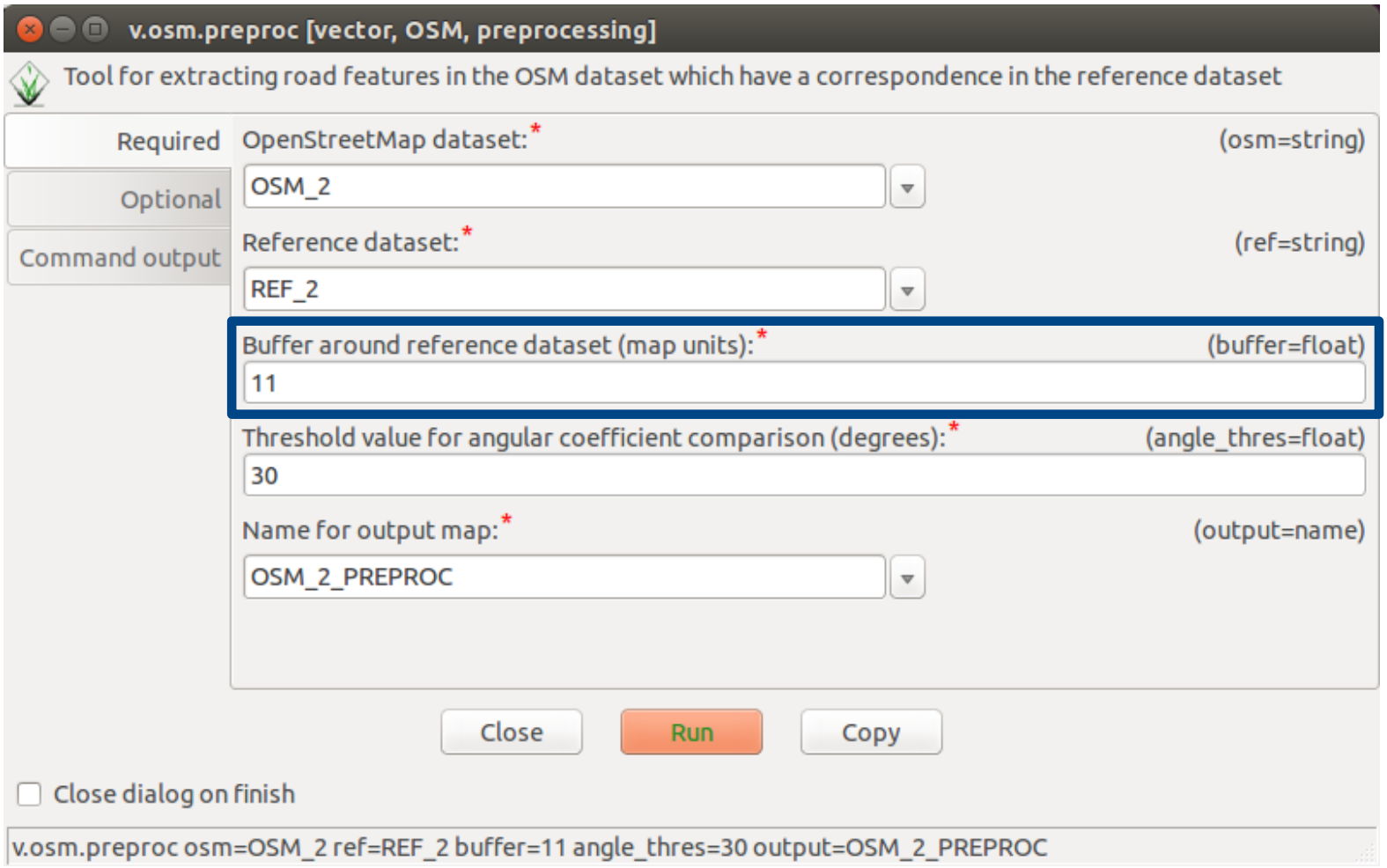
Step 2: preprocessing of the OSM dataset

- ✔ Further clean the OSM dataset [*required*]:



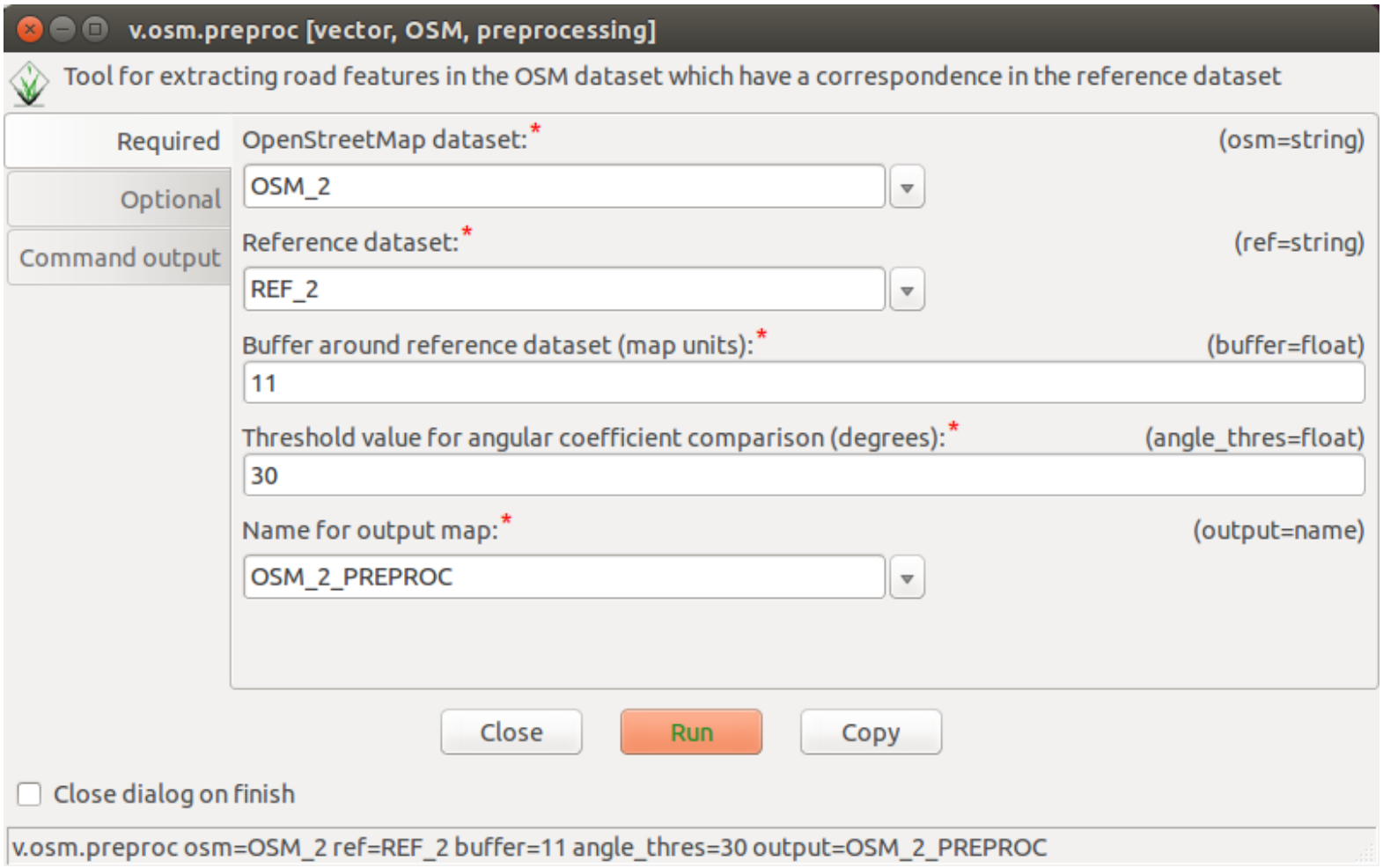
Step 2: preprocessing of the OSM dataset

- Further clean the OSM dataset [*required*]:
  - apply a buffer of user-specified width around each IGN segment



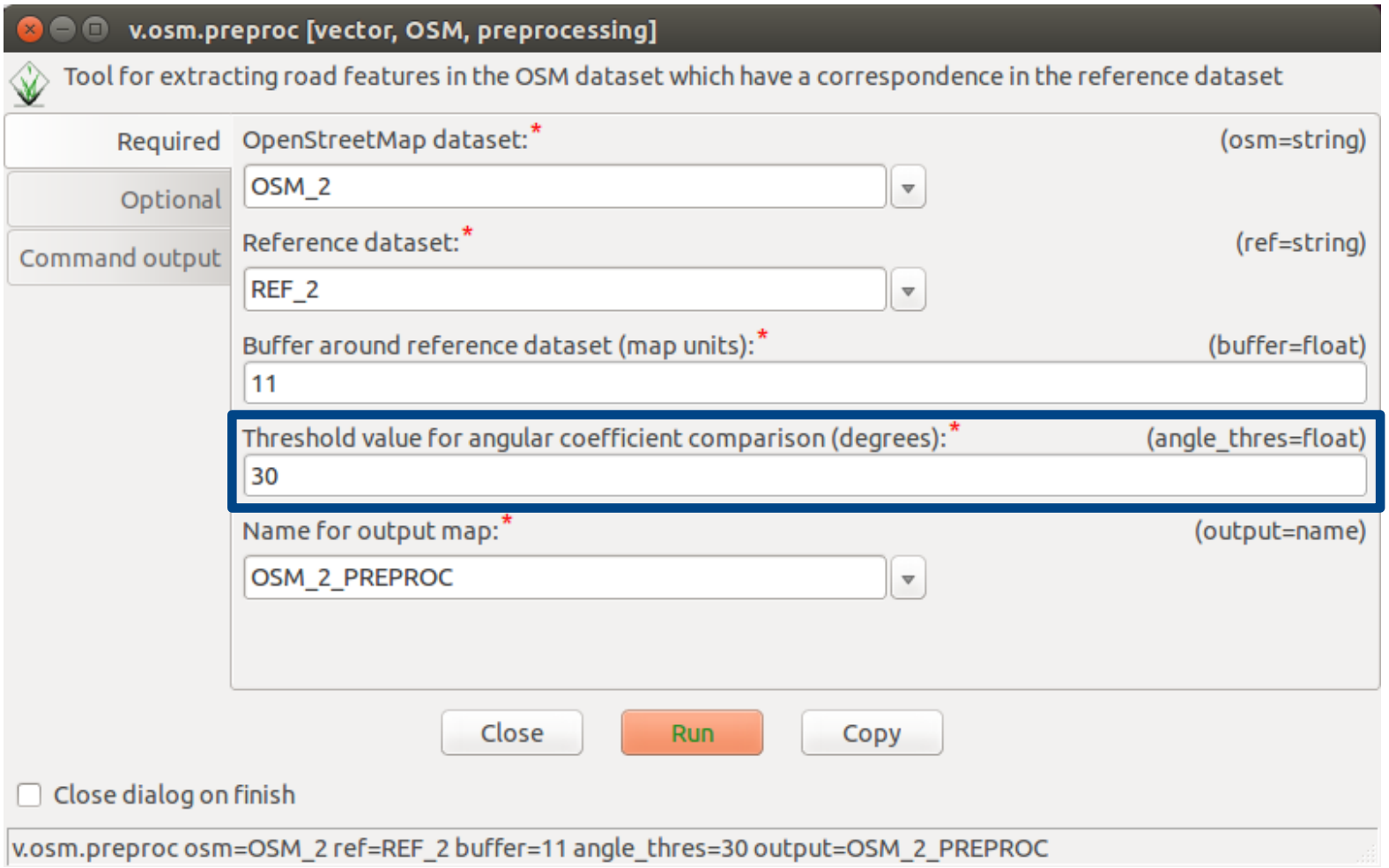
Step 2: preprocessing of the OSM dataset

- Further clean the OSM dataset [*required*]:
  - compute the angular coefficient of each IGN segment and all the OSM segments included in the buffer around it



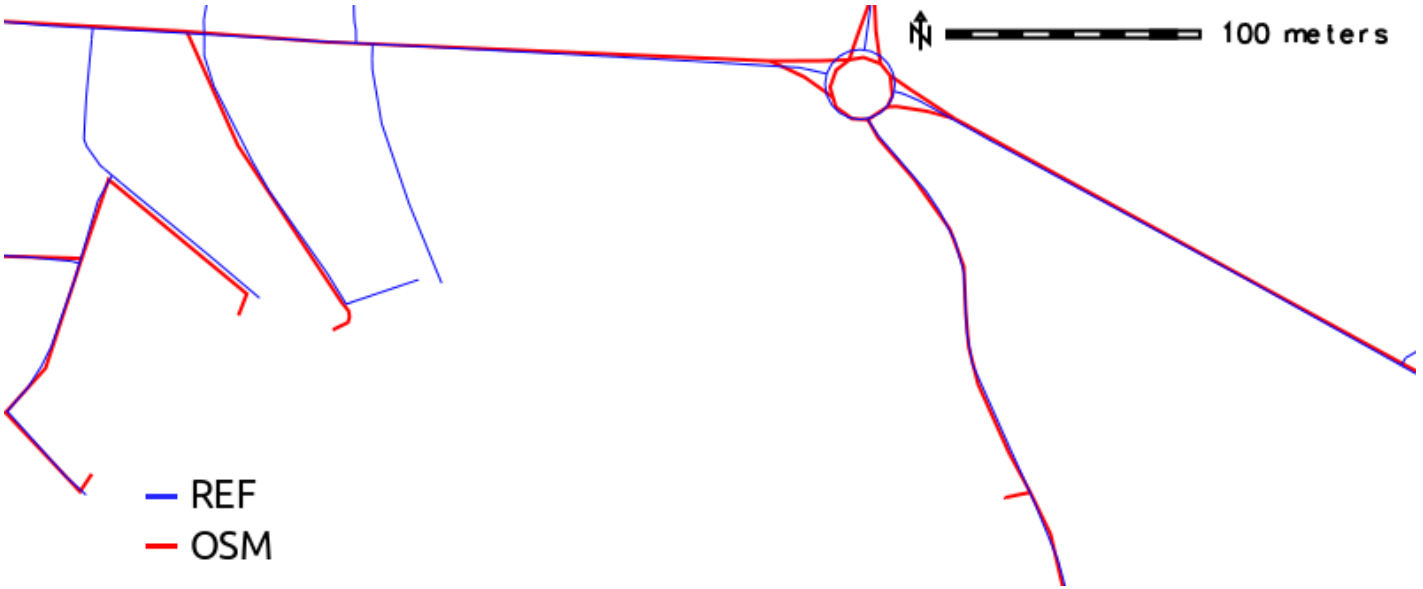
Step 2: preprocessing of the OSM dataset

- ✔ Further clean the OSM dataset [*required*]:
  - ➔ compare the difference between IGN and OSM angular coefficients with a user-specified threshold



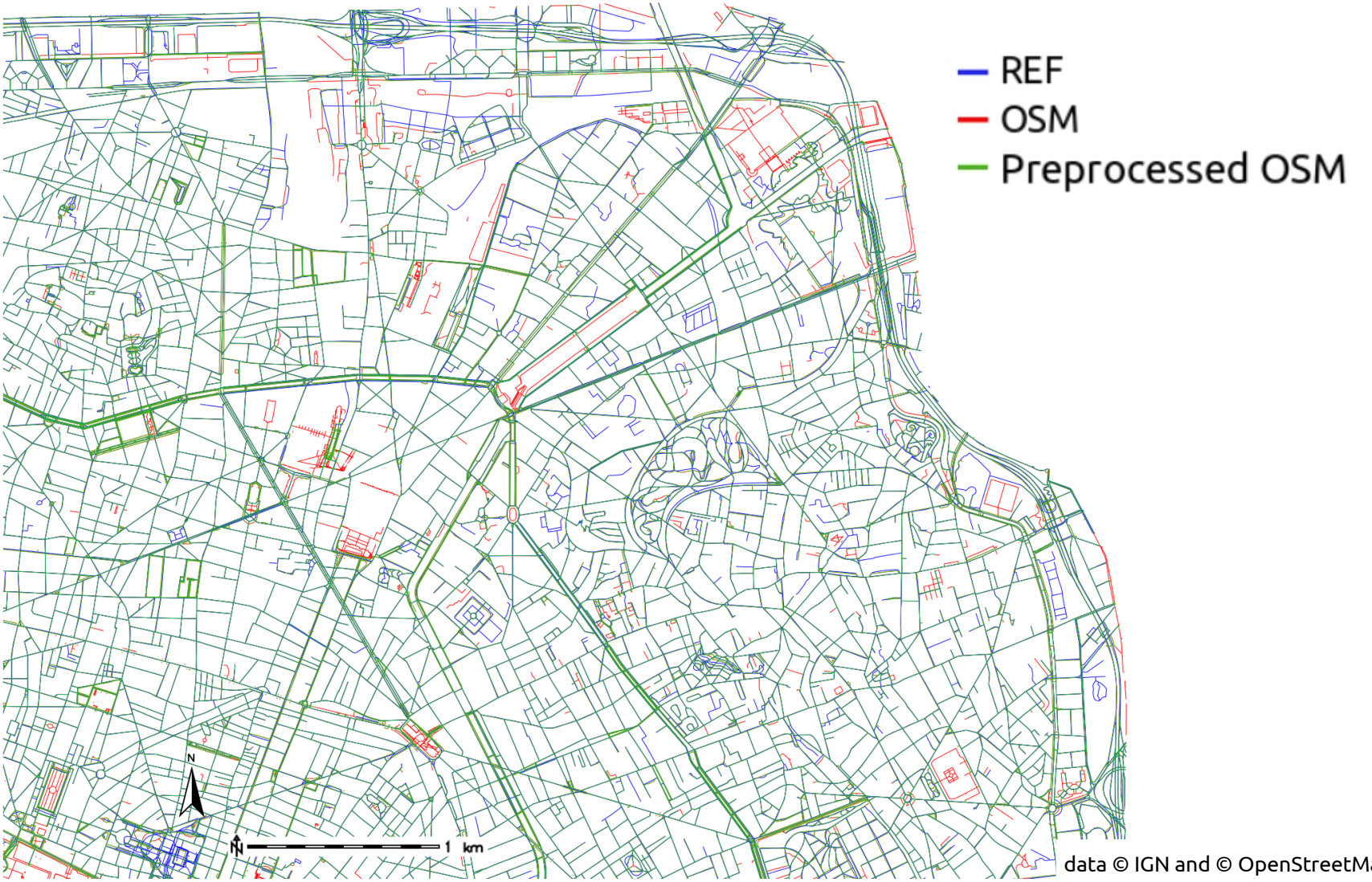
Step 2: preprocessing of the OSM dataset

- ✔ Further clean the OSM dataset [*required*]:



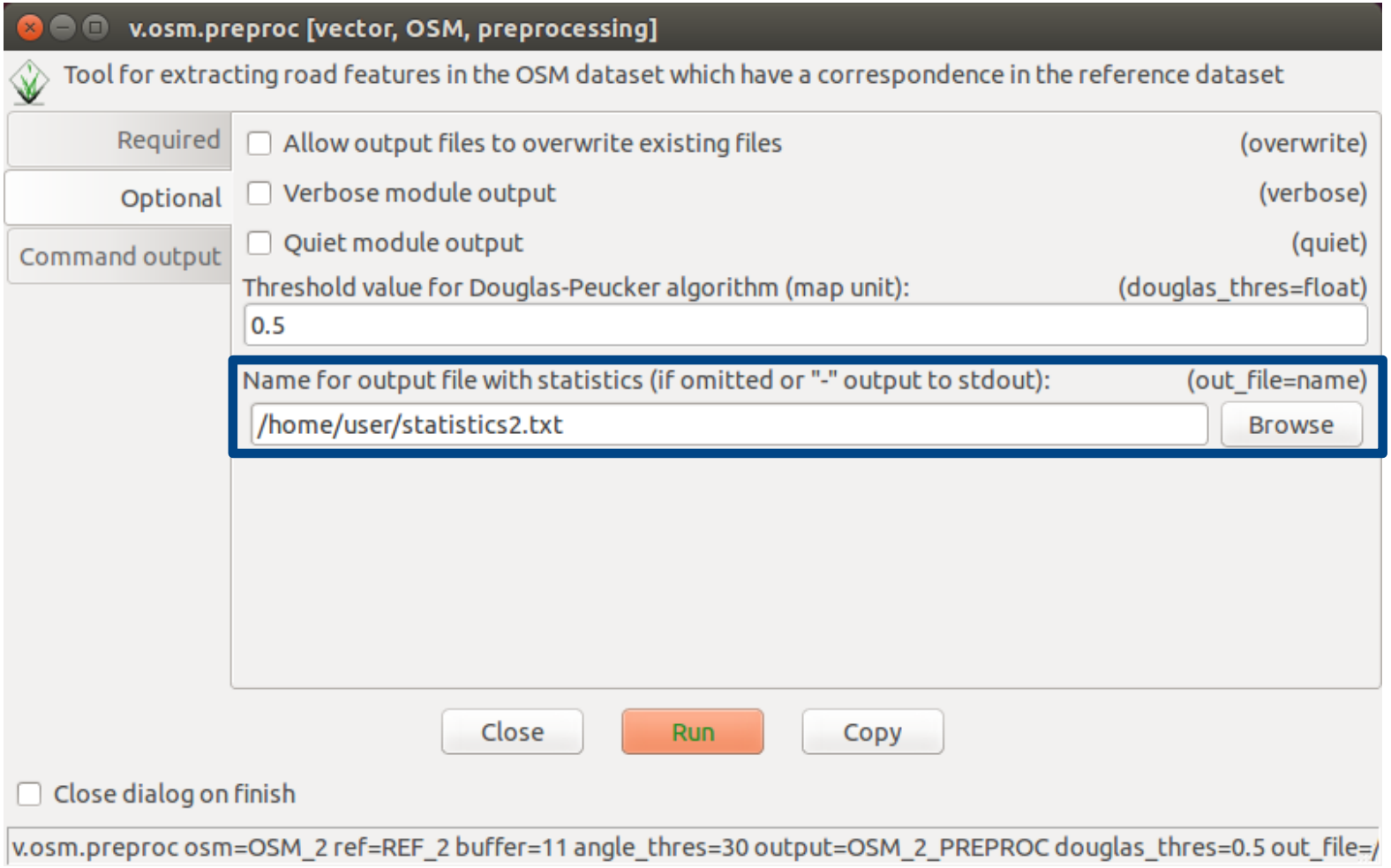
Step 2: preprocessing of the OSM dataset

- Outputs from Step 2 are saved and can be used for further analysis:
  - sensitivity analysis on the parameters involved



Step 2: preprocessing of the OSM dataset

- Outputs from Step 2 are saved and can be used for further analysis:
  - sensitivity analysis on the parameters involved



Step 2: preprocessing of the OSM dataset

- Outputs from Step 2 are saved and can be used for further analysis:
  - Area 2: generalization threshold = 0.5 m, buffer = 11 m

```
REF dataset length: 599282.7 m
Original OSM dataset length: 703604.0 m
Processed OSM dataset length: 651320.6 m
Difference between OSM original and processed datasets length: 52283.4 m (7.4%)
Difference between REF dataset and processed OSM dataset length: -52037.9 m (-8.7%)
```

Step 2: preprocessing of the OSM dataset

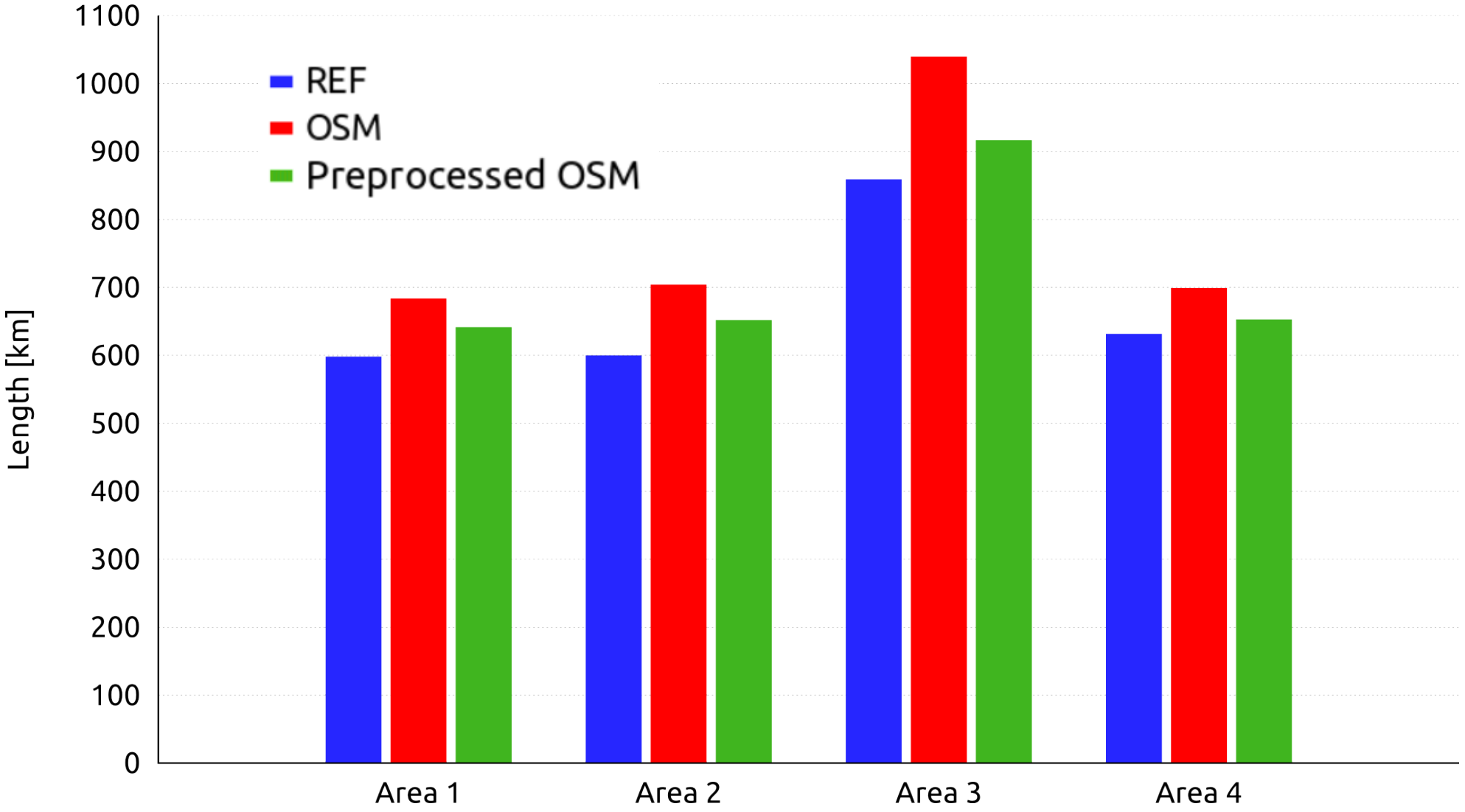
- ✔ Outputs from Step 2 are saved and can be used for further analysis:
  - ➔ Area 2: generalization threshold = 0.5 m, buffer = 11 m

```
REF dataset length: 599282.7 m
Original OSM dataset length: 703604.0 m
Processed OSM dataset length: 651320.6 m
Difference between OSM original and processed datasets length: 52283.4 m (7.4%)
Difference between REF dataset and processed OSM dataset length: -52037.9 m (-8.7%)
```

  - ✗ preprocessed OSM has ≅50 km less than original OSM
  - ✗ preprocessed OSM has still ≅50 km more than IGN

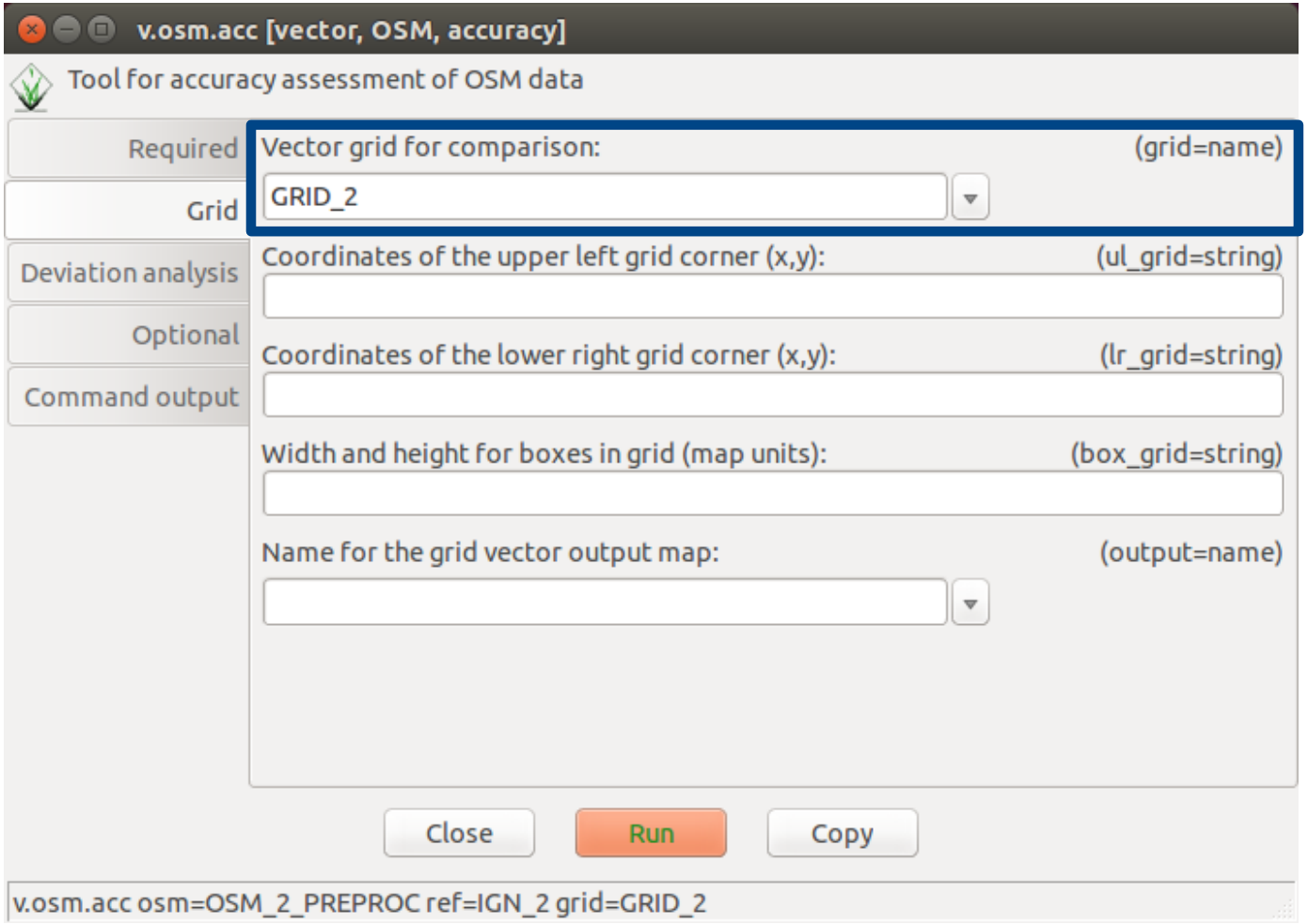
Step 2: preprocessing of the OSM dataset

- Outputs from Step 2 are saved and can be used for further analysis:



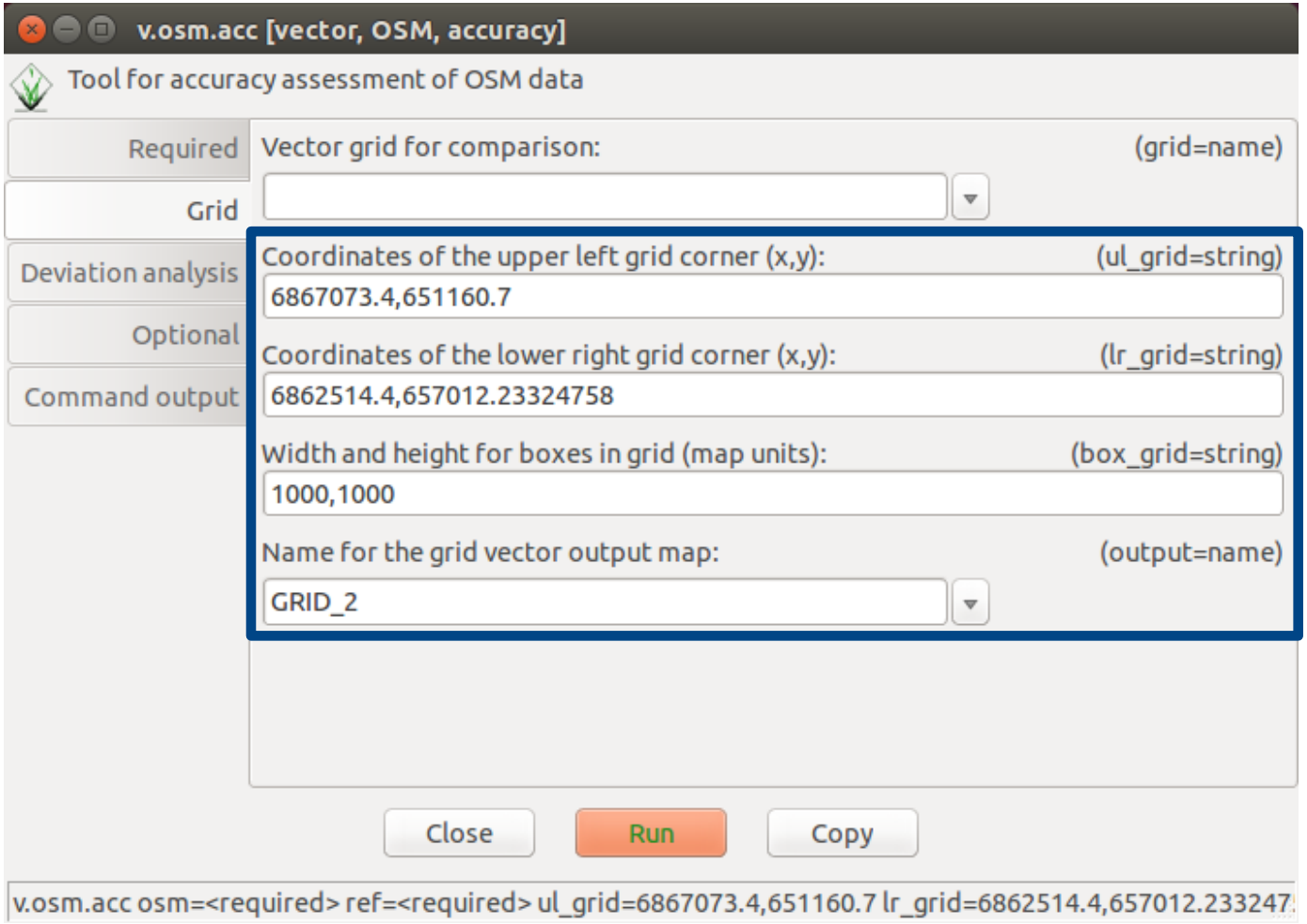
Step 3: grid-based evaluation of OSM accuracy

- Use a grid to take into account OSM heterogeneous nature [*optional*]:
  - import a vector layer to be used as grid



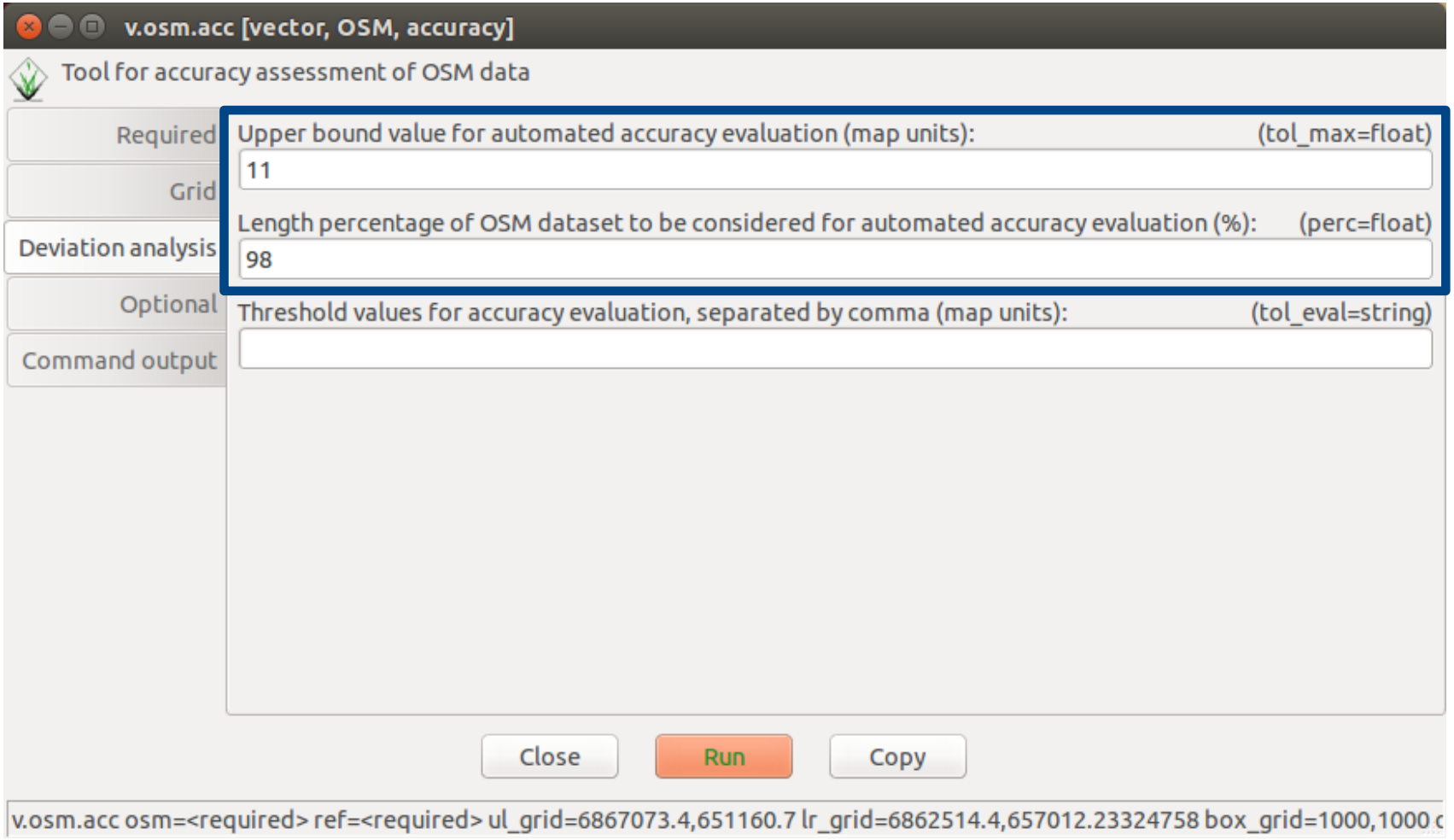
Step 3: grid-based evaluation of OSM accuracy

- Use a grid to take into account OSM heterogeneous nature [*optional*]:
  - import a vector layer to be used as grid
  - manually create a grid



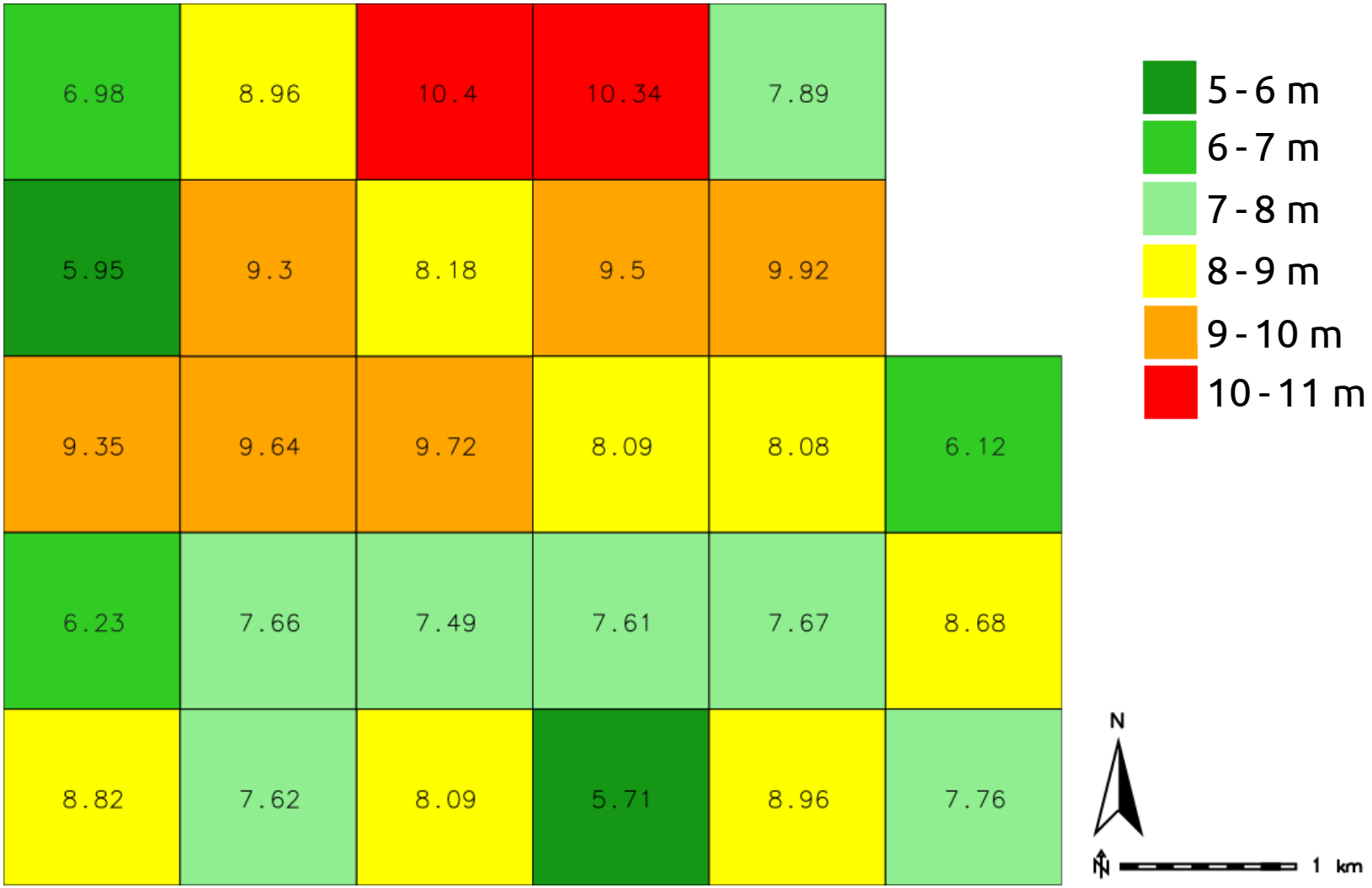
Step 3: grid-based evaluation of OSM accuracy

- For each grid cell, find the OSM maximum deviation from IGN [*optional*]:
  - enter an upper bound value for the deviation, and the percentage of OSM road length to be considered (to take into account outliers)



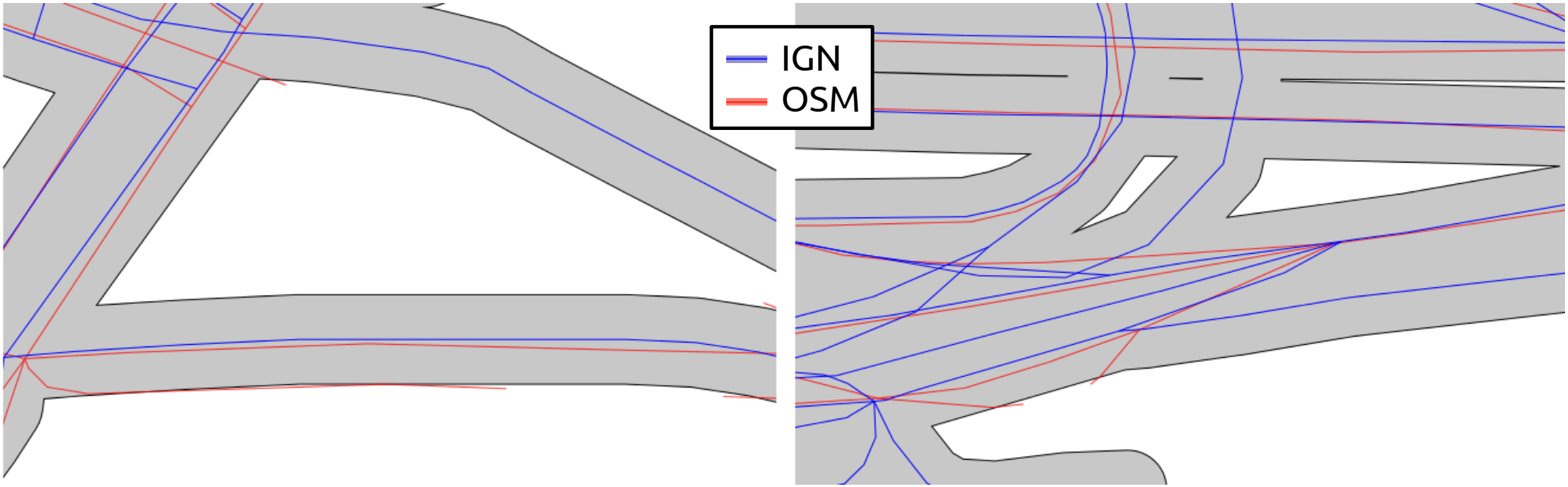
Step 3: grid-based evaluation of OSM accuracy

- For each grid cell, find the OSM maximum deviation from IGN [*optional*]:
  - Area 2: generalization threshold = 0.5 m, buffer = 11 m



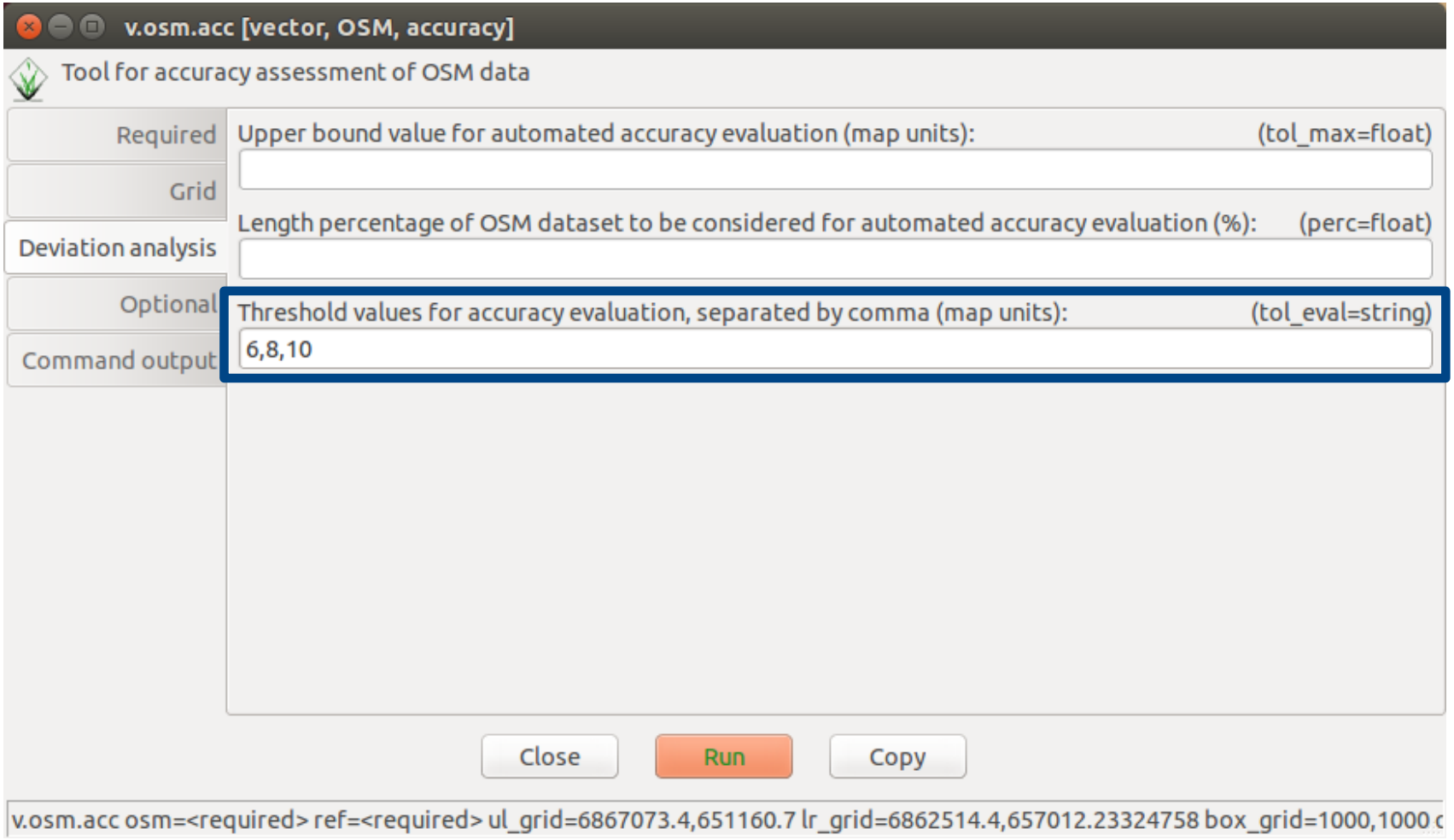
Step 3: grid-based evaluation of OSM accuracy

- ✔ For each grid cell, find the OSM maximum deviation from IGN [*optional*]:
  - ➔ Area 2: generalization threshold = 0.5 m, buffer = 11 m
  - ➔ worst results are mainly due to:
    - ✗ presence of 2 or more OSM roads for a single IGN road
    - ✗ inherent complexity of the road network



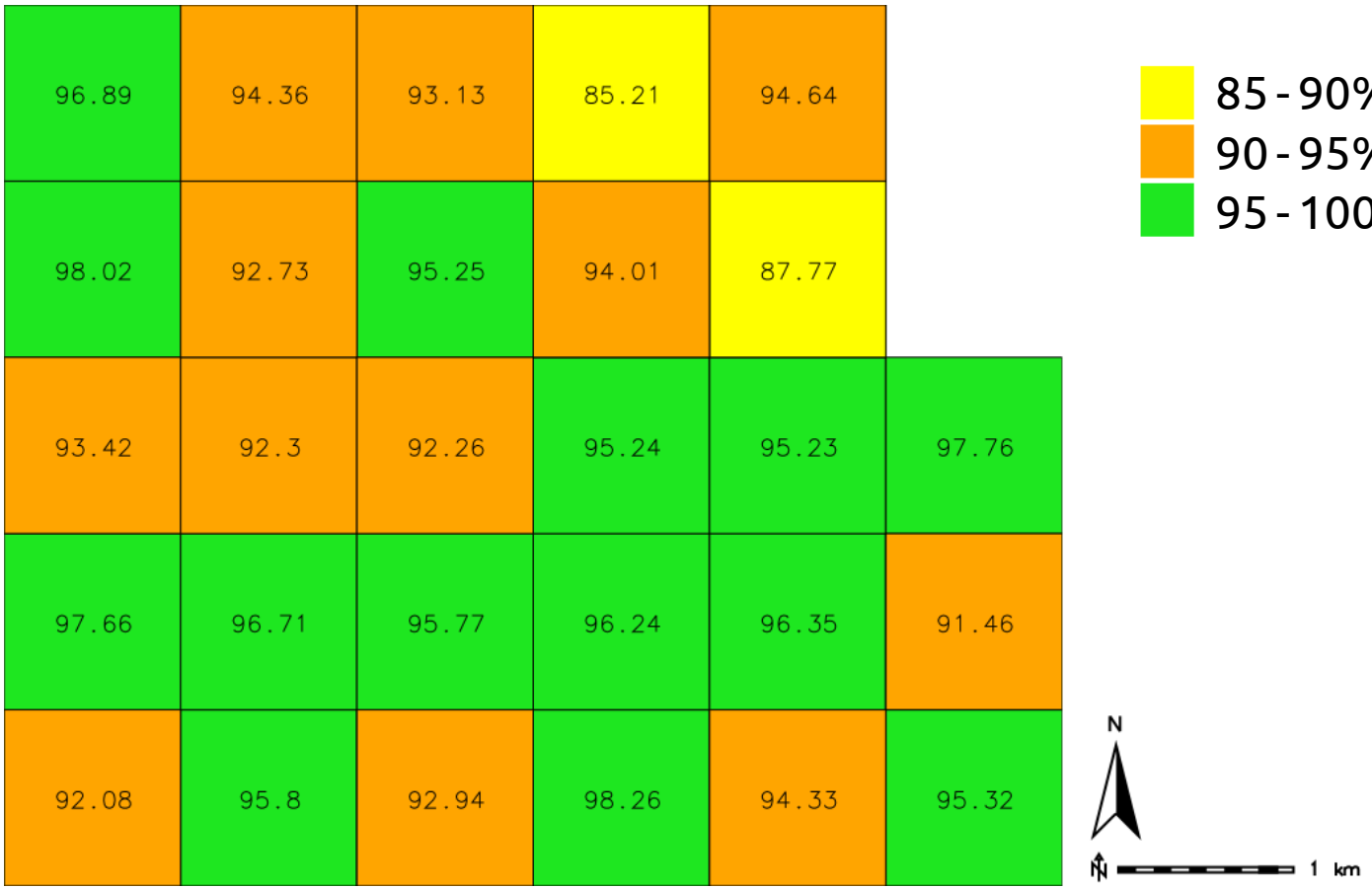
Step 3: grid-based evaluation of OSM accuracy

- For each grid cell, evaluate OSM accuracy against one or more threshold values of OSM deviation from IGN [*optional*]:
  - users have to enter one or more thresholds for deviation



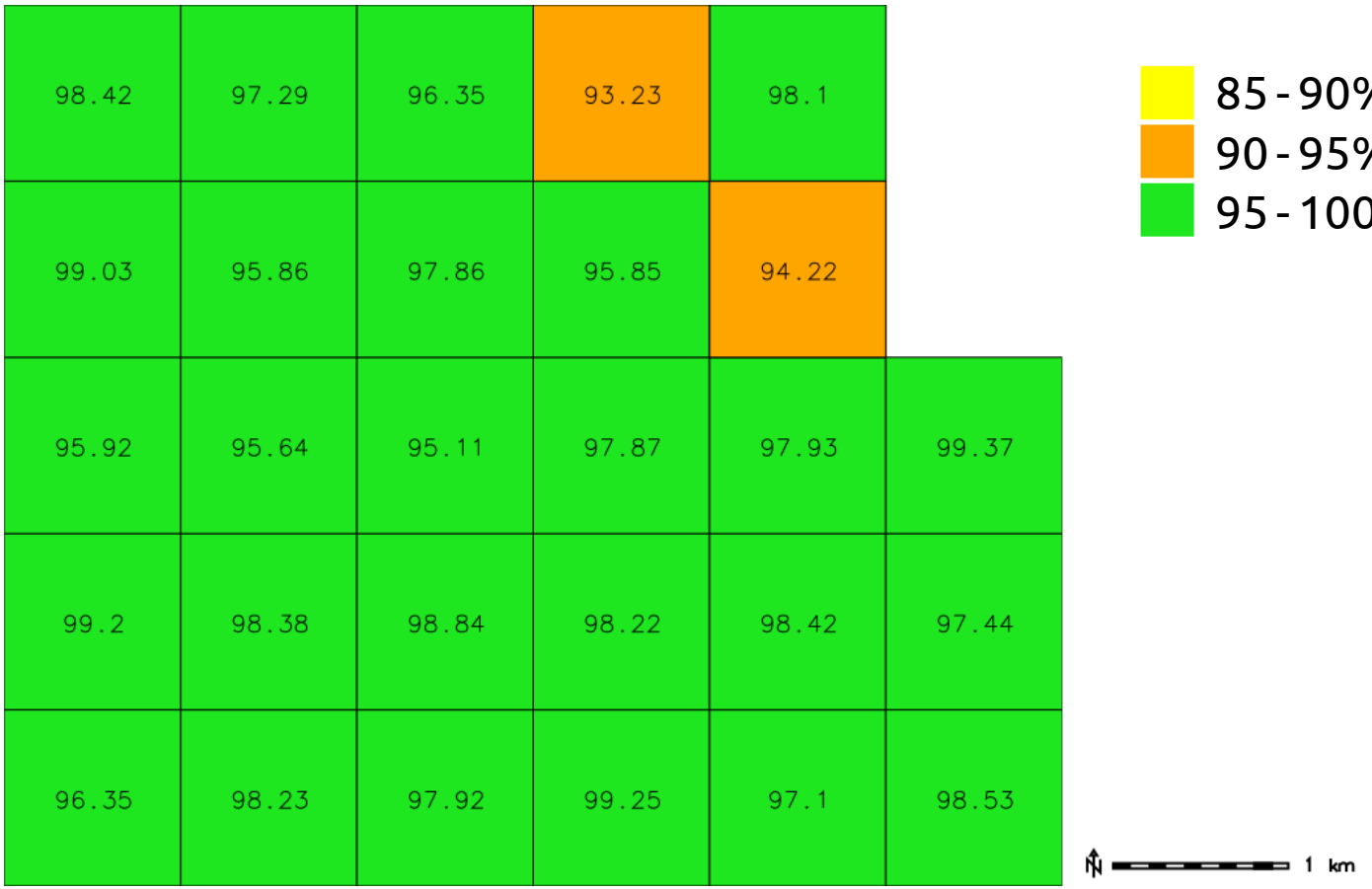
Step 3: grid-based evaluation of OSM accuracy

- For each grid cell, evaluate OSM accuracy against one or more threshold values of OSM deviation from IGN [*optional*]:
  - length percentage of OSM roads included in the threshold buffer
  - Area 2: threshold buffer = 6 m



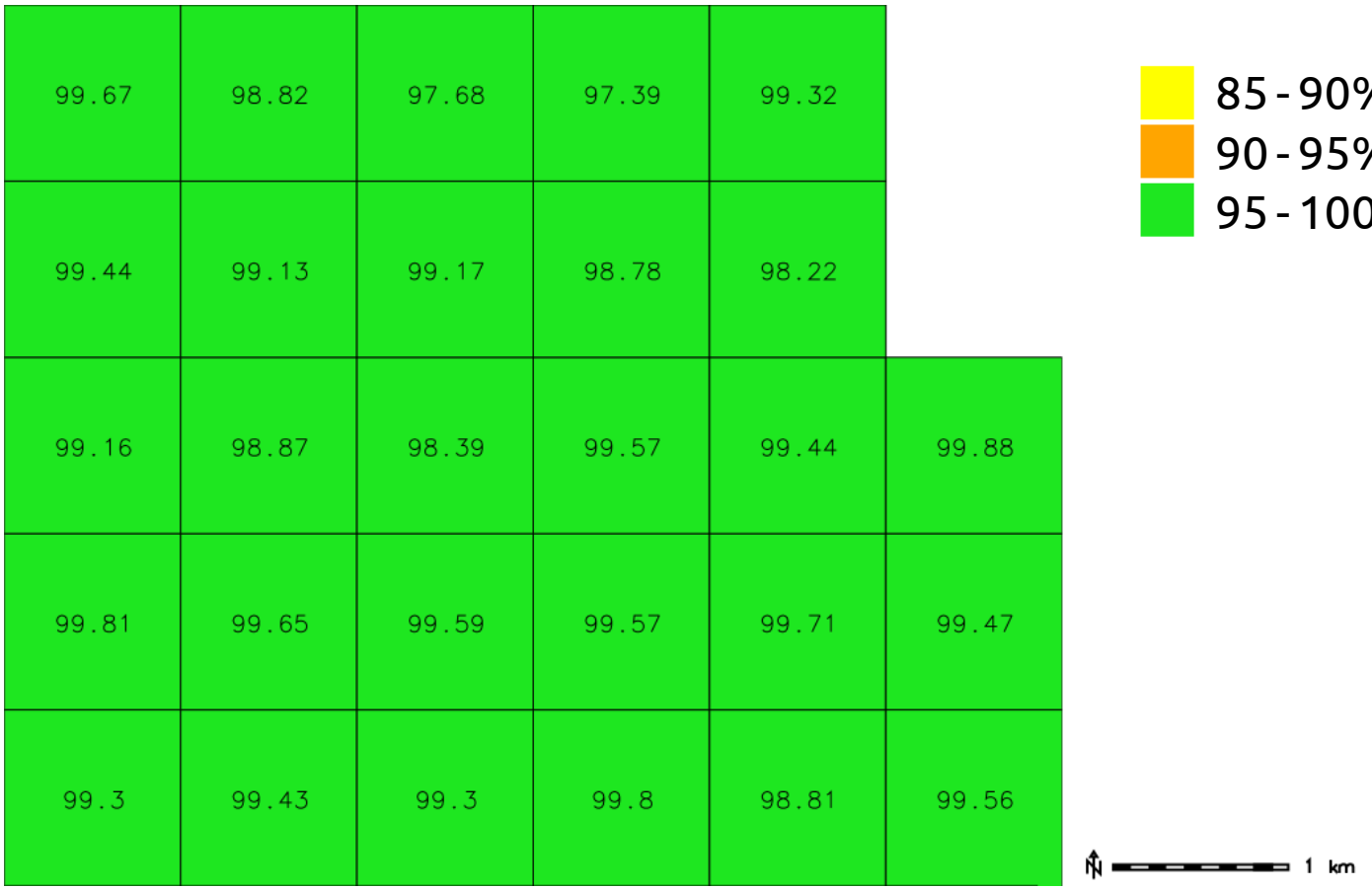
Step 3: grid-based evaluation of OSM accuracy

- For each grid cell, evaluate OSM accuracy against one or more threshold values of OSM deviation from IGN [*optional*]:
  - length percentage of OSM roads included in the threshold buffer
  - Area 2: threshold buffer = 8 m



Step 3: grid-based evaluation of OSM accuracy

- For each grid cell, evaluate OSM accuracy against one or more threshold values of OSM deviation from IGN [*optional*]:
  - length percentage of OSM roads included in the threshold buffer
  - Area 2: threshold buffer = 10 m



Step 3: grid-based evaluation of OSM accuracy

- ✔ Sensitivity analysis on the Douglas-Peucker generalization threshold:
  - ➔ number of grid cells where the percentage of OSM length satisfying the given accuracy is > 95%

		0 m	0.5 m	1 m
6 m	Area 1	10	10	10
	Area 2	14	14	14
	Area 3	7	7	7
	Area 4	8	8	8
8 m	Area 1	22	22	22
	Area 2	26	26	26
	Area 3	37	37	37
	Area 4	26	26	26
10 m	Area 1	28	28	28
	Area 2	28	28	28
	Area 3	44	44	44
	Area 4	30	30	30

Step 3: grid-based evaluation of OSM accuracy

- Sensitivity analysis on the Douglas-Peucker generalization threshold:
  - number of grid cells where the percentage of OSM length satisfying the given accuracy is > 95%
  - no difference at all!

		0 m	0.5 m	1 m
6 m	Area 1	10	10	10
	Area 2	14	14	14
	Area 3	7	7	7
	Area 4	8	8	8
8 m	Area 1	22	22	22
	Area 2	26	26	26
	Area 3	37	37	37
	Area 4	26	26	26
10 m	Area 1	28	28	28
	Area 2	28	28	28
	Area 3	44	44	44
	Area 4	30	30	30

Step 3: grid-based evaluation of OSM accuracy

- ✔ Sensitivity analysis on the Douglas-Peucker generalization threshold:
  - ➔ statistics on the differences between the percentages of OSM length satisfying each given accuracy, for the generalization thresholds of 0 m and 1 m (Area 1)

	media	dev. st.	min(abs)	max(abs)
6 m	-0.010	0.114	0.001	0.510
8 m	0.003	0.080	0.001	0.276
10 m	0.013	0.059	0.001	0.208

Step 3: grid-based evaluation of OSM accuracy

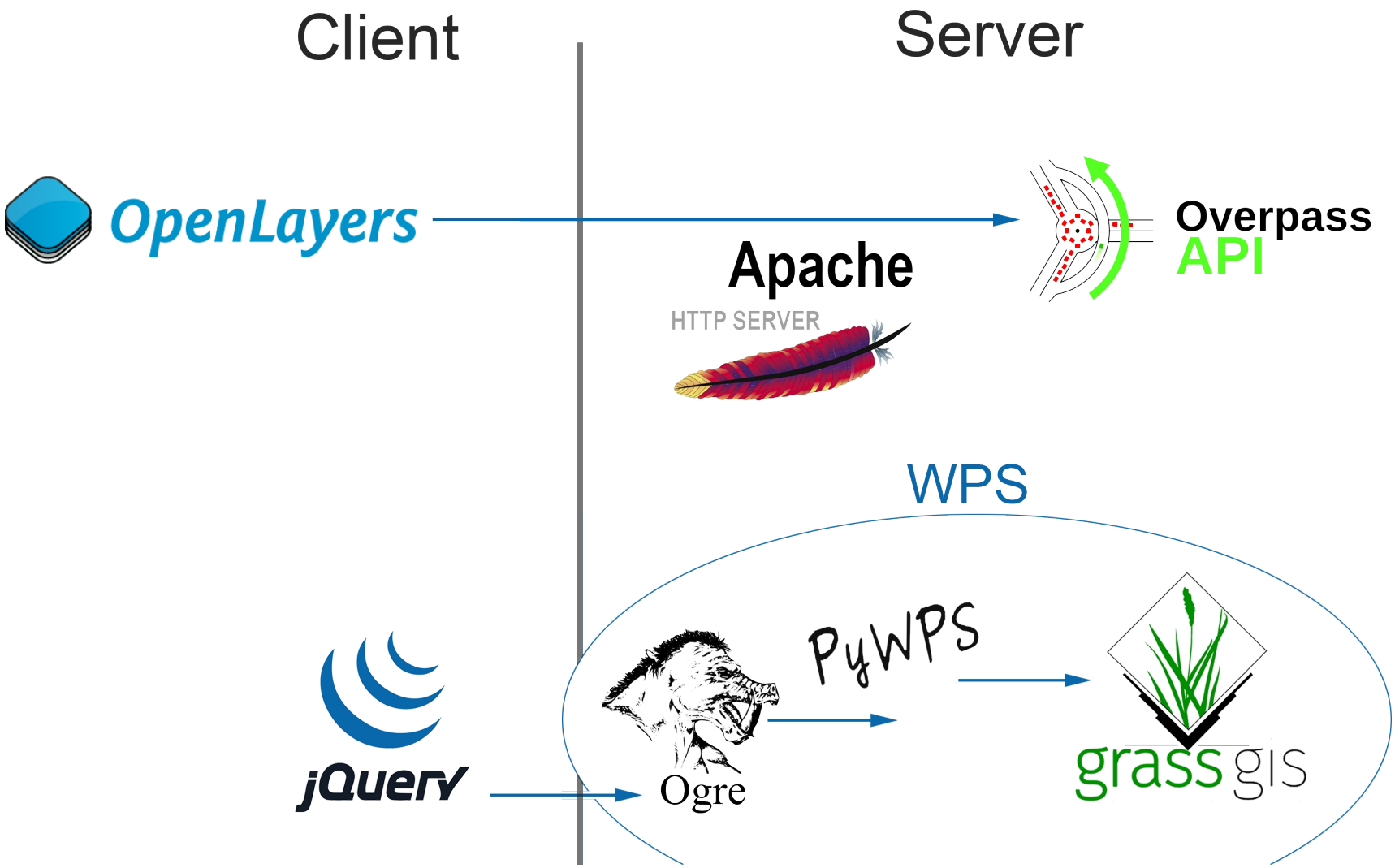
- ✔ Sensitivity analysis on the Douglas-Peucker generalization threshold:
  - ➔ statistics on the differences between the percentages of OSM length satisfying each given accuracy, for the generalization thresholds of 0 m and 1 m (Area 1)

	media	dev. st.	min(abs)	max(abs)
6 m	-0.010	0.114	0.001	0.510
8 m	0.003	0.080	0.001	0.276
10 m	0.013	0.059	0.001	0.208

  - ➔ generalization (within the nominal accuracy of the dataset) does not influence accuracy evaluation results – and allows to save much time!

Transposition of the procedure as a WPS

- Work in progress, currently available just for Step 1
  - available at http://131.175.143.84/WPS



Transposition of the procedure as a WPS

- ✔ User instructions on how to use the tool

Comparison of OpenStreetMap (OSM) with authoritative road datasets (REF)

The procedure allows to import customized road datasets from the OSM database and an authoritative source and returns some measures of their spatial similarity.

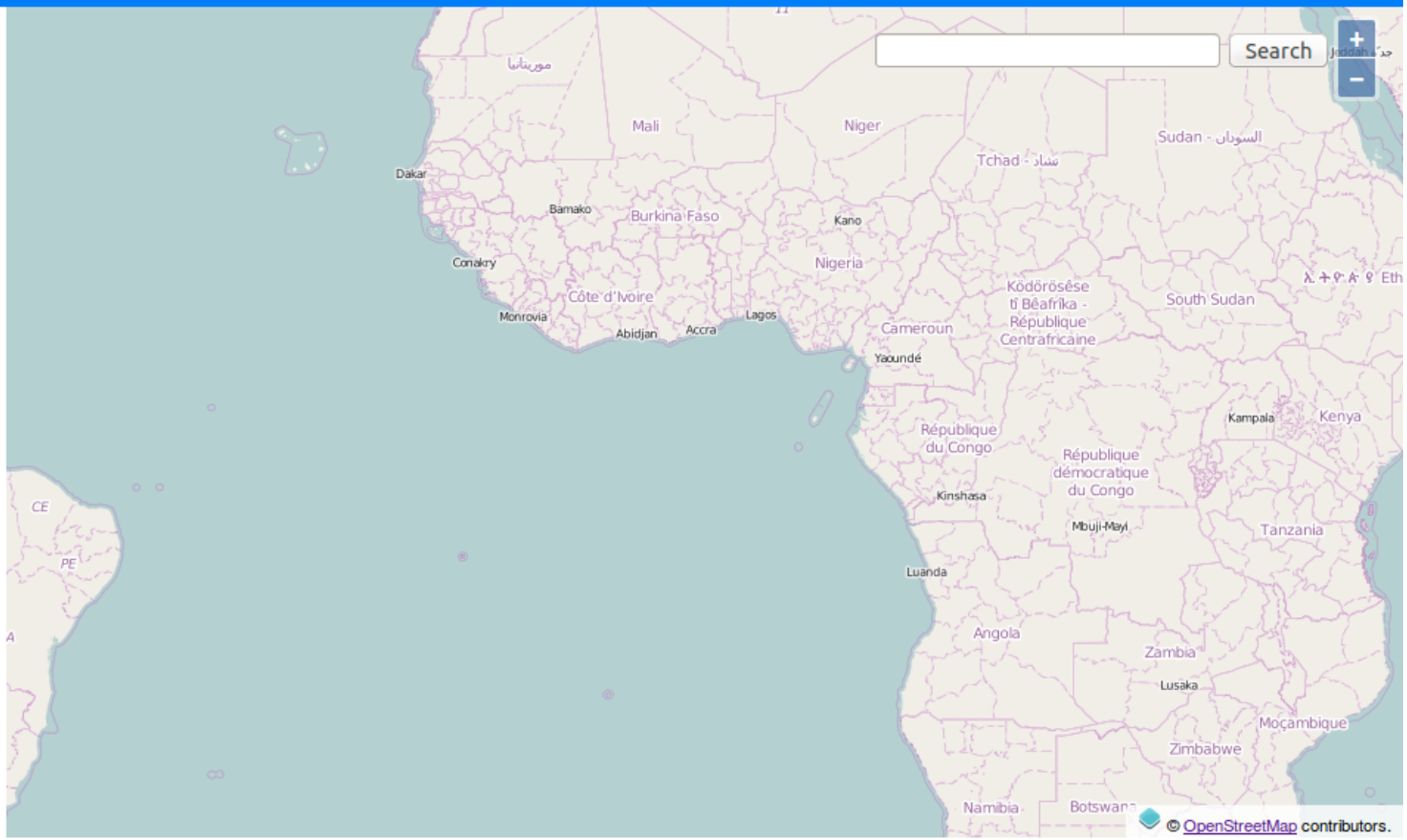
How to use the tool:

Road network datasets must (for REF) and can (for OSM) be uploaded as a compressed ESRI shapefile. Click the **Browse** button, select the **.zip** file from your system and adjust its color using the **Layer Color** selector. Enter also the reference system of the dataset.

The OSM road network dataset can be also retrieved (using the Overpass API) by clicking the **Retrieve OSM data** button. This will retrieve OSM data for the current map view; alternatively, hold down the shift key to select a rectangle on the map view and then click the **Selection Retrieve** button.

When at least two datasets have been uploaded, click the **Compare menu** button. This will open a menu where you can specify the buffer width for the comparison, the datasets to be used as REF and OSM and (optionally) the mask for clipping the previous datasets - to be previously uploaded as a compressed shapefile. Then click the **Start comparison** button to run the algorithm and download some statistics on the similarity between the datasets in a PDF file.

Start Tool



Transposition of the procedure as a WPS

- ✔ Geocoding service to move the map to a specified location

Comparison of OpenStreetMap (OSM) with authoritative road datasets (REF)

The procedure allows to import customized road datasets from the OSM database and an authoritative source and returns some measures of their spatial similarity.

How to use the tool:

Road network datasets must (for REF) and can (for OSM) be uploaded as a compressed ESRI shapfile. Click the **Browse** button, select the **.zip** file from your system and adjust its color using the **Layer Color** selector. Enter also the reference system of the dataset.

The OSM road network dataset can be also retrieved (using the Overpass API) by clicking the **Retrieve OSM data** button. This will retrieve OSM data for the current map view; alternatively, hold down the shift key to select a rectangle on the map view and then click the **Selection Retrieve** button.

When at least two datasets have been uploaded, click the **Compare menu** button. This will open a menu where you can specify the buffer width for the comparison, the datasets to be used as REF and OSM and (optionally) the mask for clipping the previous datasets - to be previously uploaded as a compressed shapefile. Then click the **Start comparison** button to run the algorithm and download some statistics on the similarity between the datasets in a PDF file.

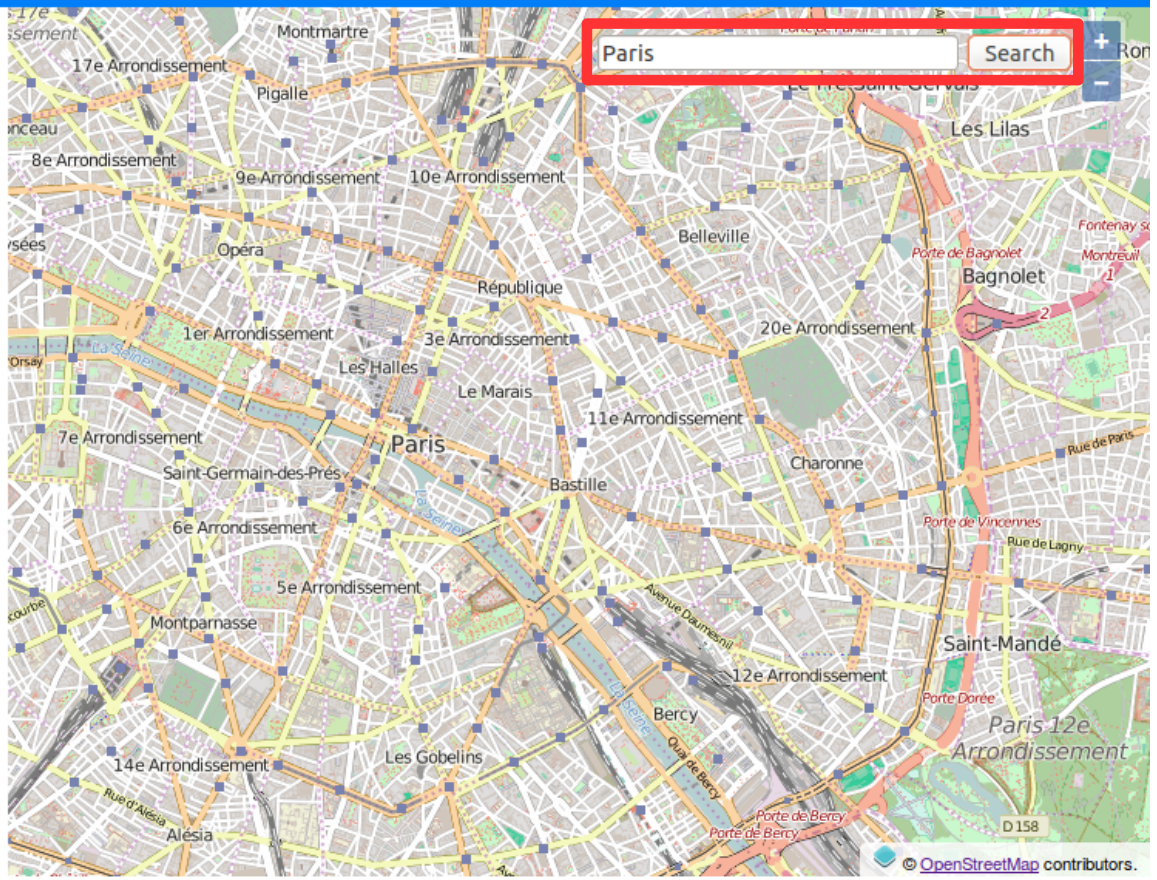
Start Tool

Menu:

Upload

Retrieve OSM data

Compare menu



Transposition of the procedure as a WPS

- ✔ Upload of IGN road network dataset

Comparison of OpenStreetMap (OSM) with authoritative road datasets (REF)

The procedure allows to import customized road datasets from the OSM database and an authoritative source and returns some measures of their spatial similarity.

How to use the tool:

Road network datasets must (for REF) and can (for OSM) be uploaded as a compressed ESRI shapfile. Click the **Browse** button, select the **.zip** file from your system and adjust its color using the **Layer Color** selector. Enter also the reference system of the dataset.

The OSM road network dataset can be also retrieved (using the Overpass API) by clicking the **Retrieve OSM data** button. This will retrieve OSM data for the current map view; alternatively, hold down the shift key to select a rectangle on the map view and then click the **Selection Retrieve** button.

When at least two datasets have been uploaded, click the **Compare menu** button. This will open a menu where you can specify the buffer width for the comparison, the datasets to be used as REF and OSM and (optionally) the mask for clipping the previous datasets - to be previously uploaded as a compressed shapefile. Then click the **Start comparison** button to run the algorithm and download some statistics on the similarity between the datasets in a PDF file.

Start Tool

Menu:

Upload

Retrieve OSM data

Compare menu

Select File

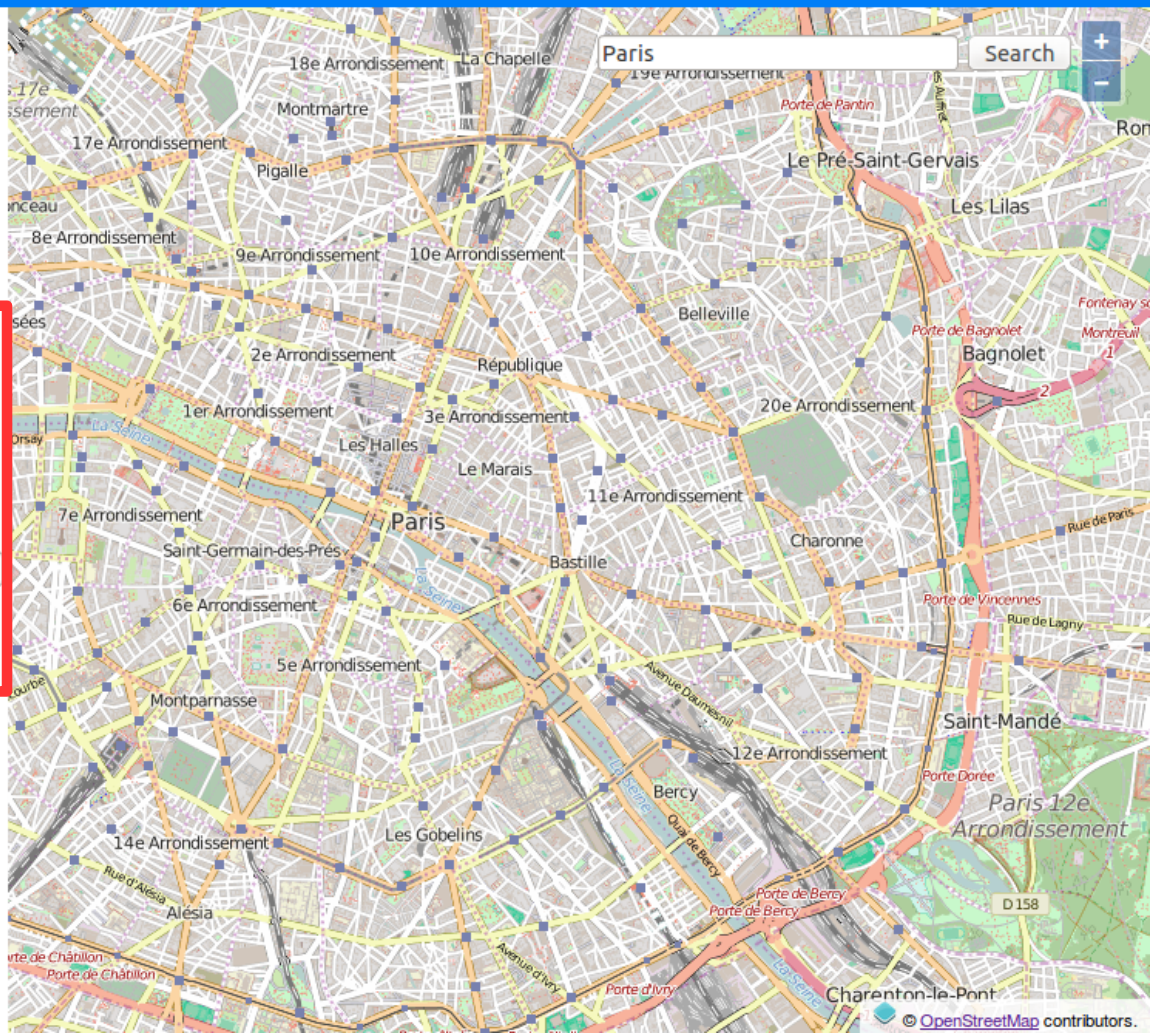
Browse... IGN_3.zip

Layer Color

Projection Source

EPSG:2154

Upload



Transposition of the procedure as a WPS

- ✔ Visualization of IGN road network dataset

Comparison of OpenStreetMap (OSM) with authoritative road datasets (REF)

The procedure allows to import customized road datasets from the OSM database and an authoritative source and returns some measures of their spatial similarity.

How to use the tool:

Road network datasets must (for REF) and can (for OSM) be uploaded as a compressed ESRI shapefile. Click the **Browse** button, select the **.zip** file from your system and adjust its color using the **Layer Color** selector. Enter also the reference system of the dataset.

The OSM road network dataset can be also retrieved (using the Overpass API) by clicking the **Retrieve OSM data** button. This will retrieve OSM data for the current map view; alternatively, hold down the shift key to select a rectangle on the map view and then click the **Selection Retrieve** button.

When at least two datasets have been uploaded, click the **Compare menu** button. This will open a menu where you can specify the buffer width for the comparison, the datasets to be used as REF and OSM and (optionally) the mask for clipping the previous datasets - to be previously uploaded as a compressed shapefile. Then click the **Start comparison** button to run the algorithm and download some statistics on the similarity between the datasets in a PDF file.

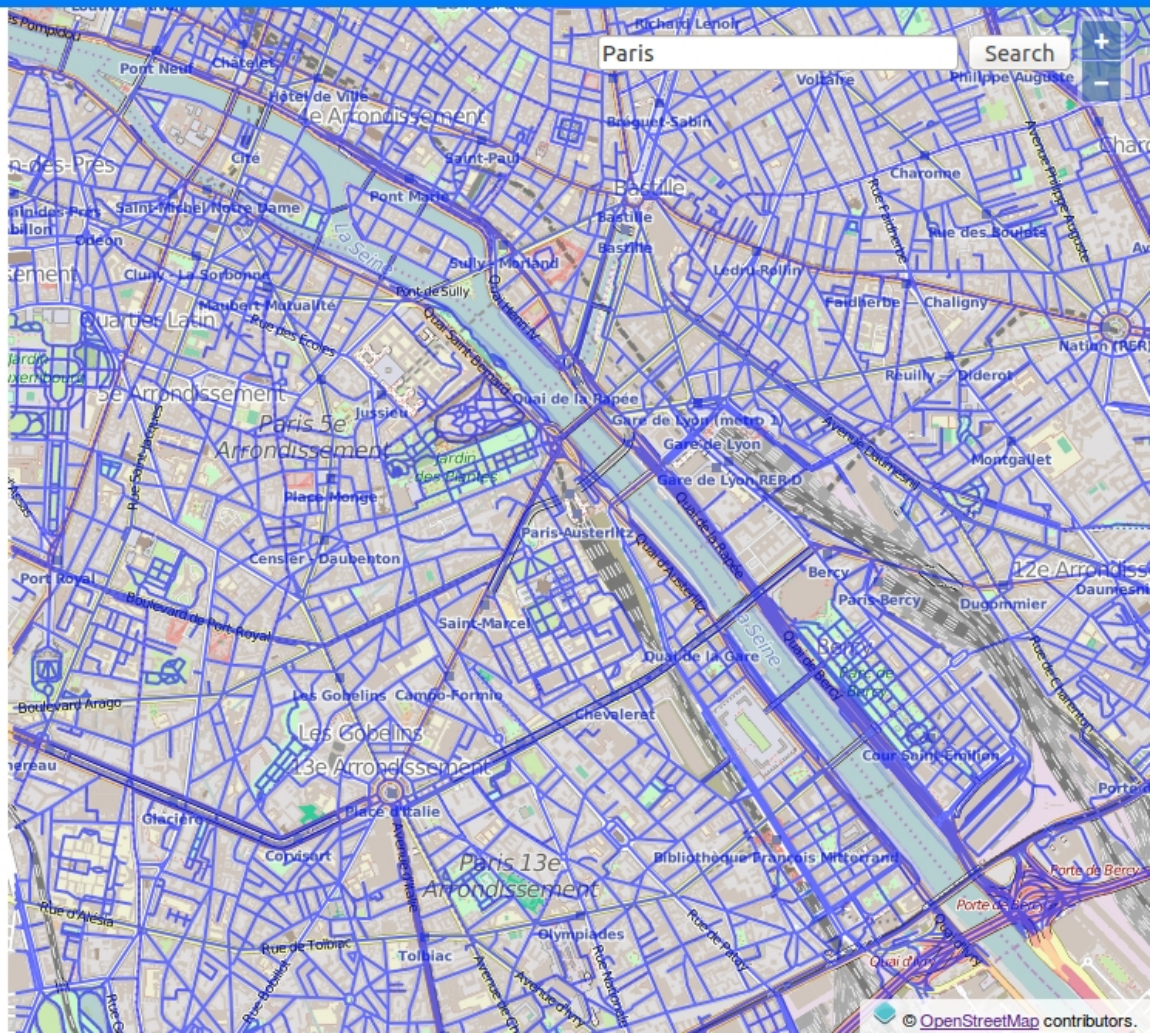
Start Tool

Menu:

Upload

Retrieve OSM data

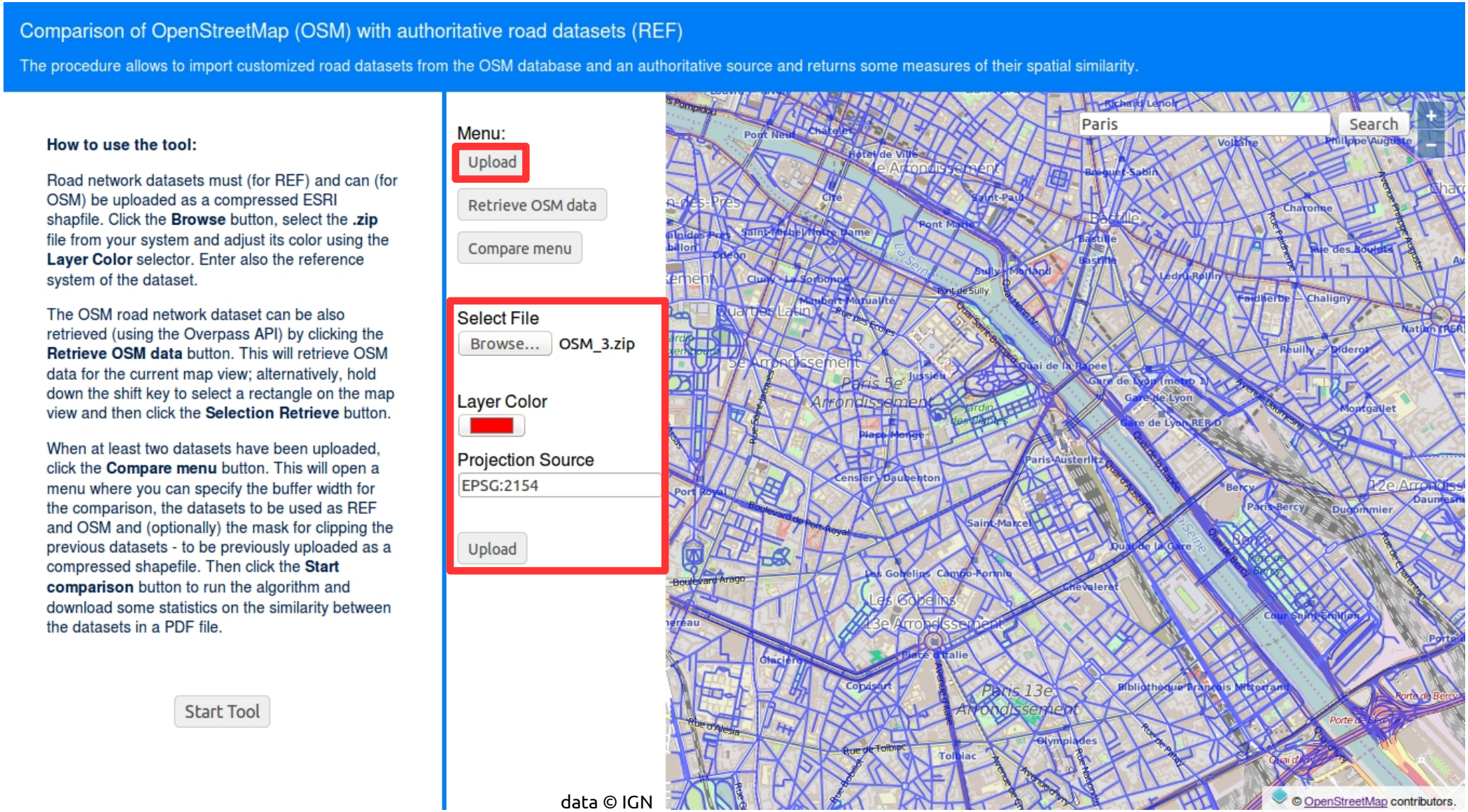
Compare menu



data © IGN

Transposition of the procedure as a WPS

- ✔ Upload of OSM road network dataset



Transposition of the procedure as a WPS

- ✔ Visualization of OSM road network dataset

Comparison of OpenStreetMap (OSM) with authoritative road datasets (REF)

The procedure allows to import customized road datasets from the OSM database and an authoritative source and returns some measures of their spatial similarity.

How to use the tool:

Road network datasets must (for REF) and can (for OSM) be uploaded as a compressed ESRI shapfile. Click the **Browse** button, select the **.zip** file from your system and adjust its color using the **Layer Color** selector. Enter also the reference system of the dataset.

The OSM road network dataset can be also retrieved (using the Overpass API) by clicking the **Retrieve OSM data** button. This will retrieve OSM data for the current map view; alternatively, hold down the shift key to select a rectangle on the map view and then click the **Selection Retrieve** button.

When at least two datasets have been uploaded, click the **Compare menu** button. This will open a menu where you can specify the buffer width for the comparison, the datasets to be used as REF and OSM and (optionally) the mask for clipping the previous datasets - to be previously uploaded as a compressed shapefile. Then click the **Start comparison** button to run the algorithm and download some statistics on the similarity between the datasets in a PDF file.

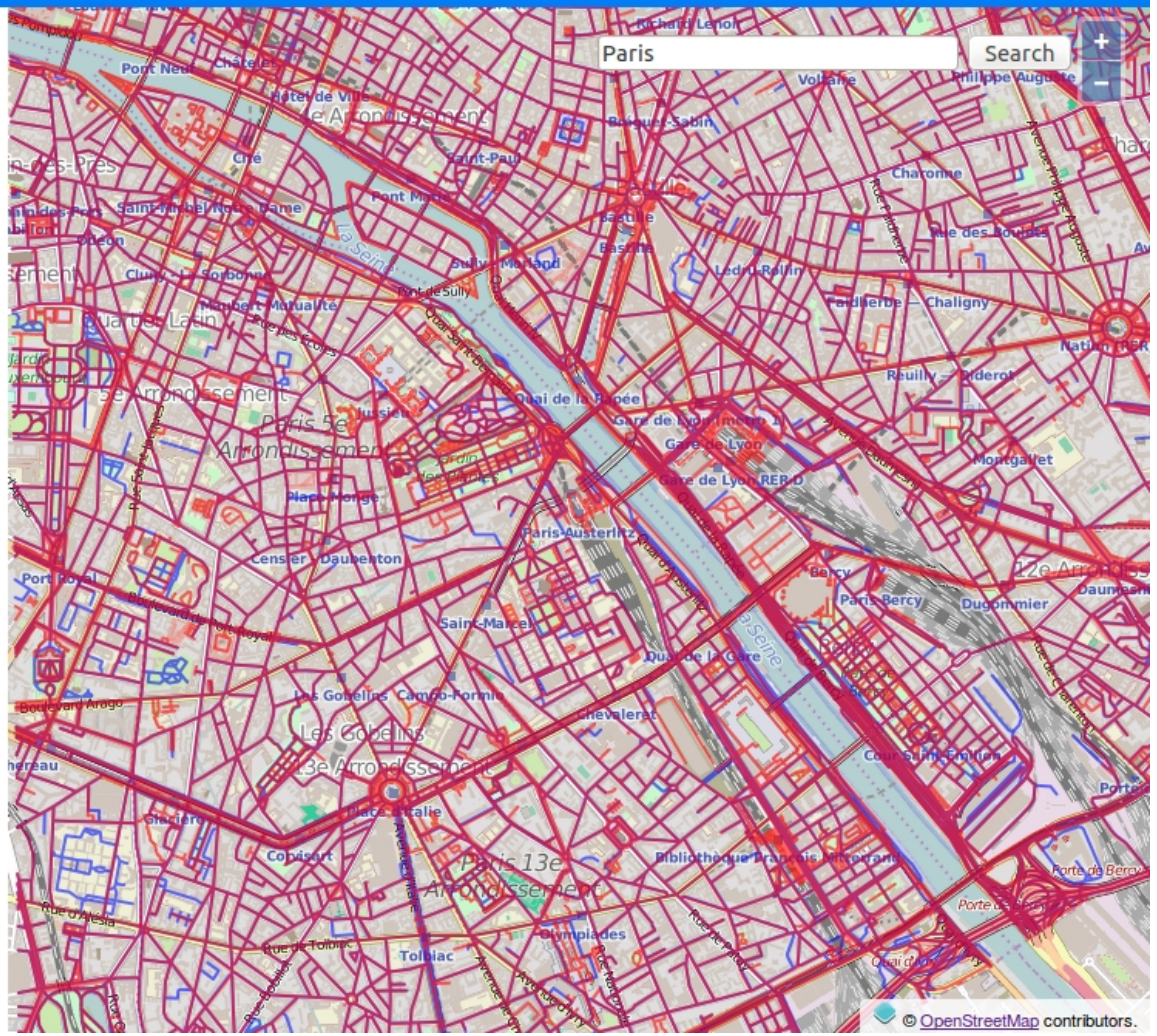
Start Tool

Menu:

Upload

Retrieve OSM data

Compare menu



data © IGN

Transposition of the procedure as a WPS

- ✔ Definition of layers and buffer value for comparison, a PDF is generated

Comparison of OpenStreetMap (OSM) with authoritative road datasets (REF)

The procedure allows to import customized road datasets from the OSM database and an authoritative source and returns some measures of their spatial similarity.

How to use the tool:

Road network datasets must (for REF) and can (for OSM) be uploaded as a compressed ESRI shapfile. Click the **Browse** button, select the **.zip** file from your system and adjust its color using the **Layer Color** selector. Enter also the reference system of the dataset.

The OSM road network dataset can be also retrieved (using the Overpass API) by clicking the **Retrieve OSM data** button. This will retrieve OSM data for the current map view; alternatively, hold down the shift key to select a rectangle on the map view and then click the **Selection Retrieve** button.

When at least two datasets have been uploaded, click the **Compare menu** button. This will open a menu where you can specify the buffer width for the comparison, the datasets to be used as REF and OSM and (optionally) the mask for clipping the previous datasets - to be previously uploaded as a compressed shapefile. Then click the **Start comparison** button to run the algorithm and download some statistics on the similarity between the datasets in a PDF file.

Start Tool

Menu:

Upload

Retrieve OSM data

Compare menu

Buffer (map units)

10

Reference Layer

IGN_3.zip

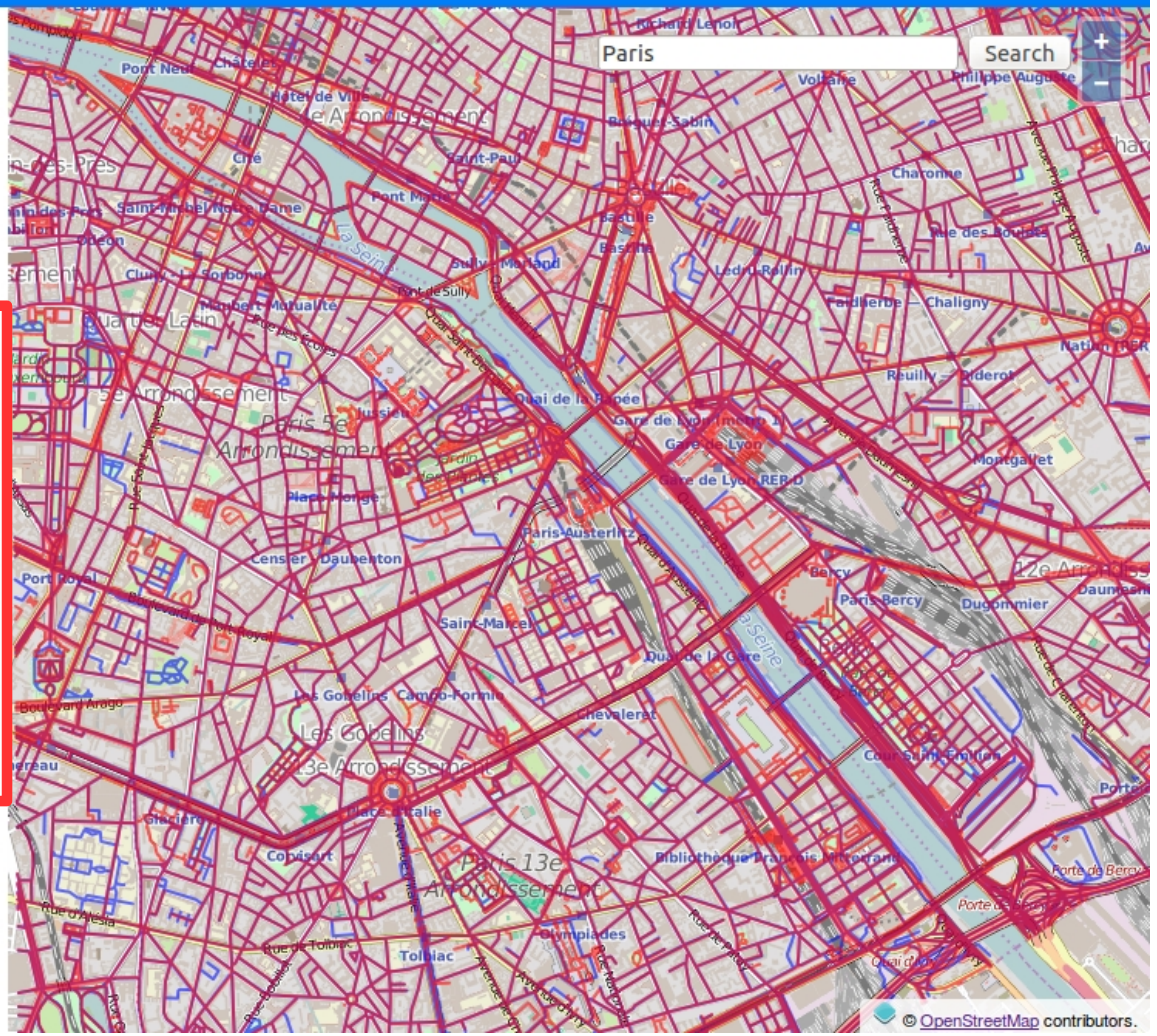
OSM Layer

OSM_3.zip

Mask Layer

Select

Start comparison



Transposition of the procedure as a WPS

- ✔ Retrieval of OSM road network dataset from the current map view

Comparison of OpenStreetMap (OSM) with authoritative road datasets (REF)

The procedure allows to import customized road datasets from the OSM database and an authoritative source and returns some measures of their spatial similarity.

How to use the tool:

Road network datasets must (for REF) and can (for OSM) be uploaded as a compressed ESRI shapfile. Click the **Browse** button, select the **.zip** file from your system and adjust its color using the **Layer Color** selector. Enter also the reference system of the dataset.

The OSM road network dataset can be also retrieved (using the Overpass API) by clicking the **Retrieve OSM data** button. This will retrieve OSM data for the current map view; alternatively, hold down the shift key to select a rectangle on the map view and then click the **Selection Retrieve** button.

When at least two datasets have been uploaded, click the **Compare menu** button. This will open a menu where you can specify the buffer width for the comparison, the datasets to be used as REF and OSM and (optionally) the mask for clipping the previous datasets - to be previously uploaded as a compressed shapefile. Then click the **Start comparison** button to run the algorithm and download some statistics on the similarity between the datasets in a PDF file.

Start Tool

Menu:

Upload

Retrieve OSM data

Compare menu

Layer Color

Layer Name

OSM

Retrieve data from the current map view

View Retrieve

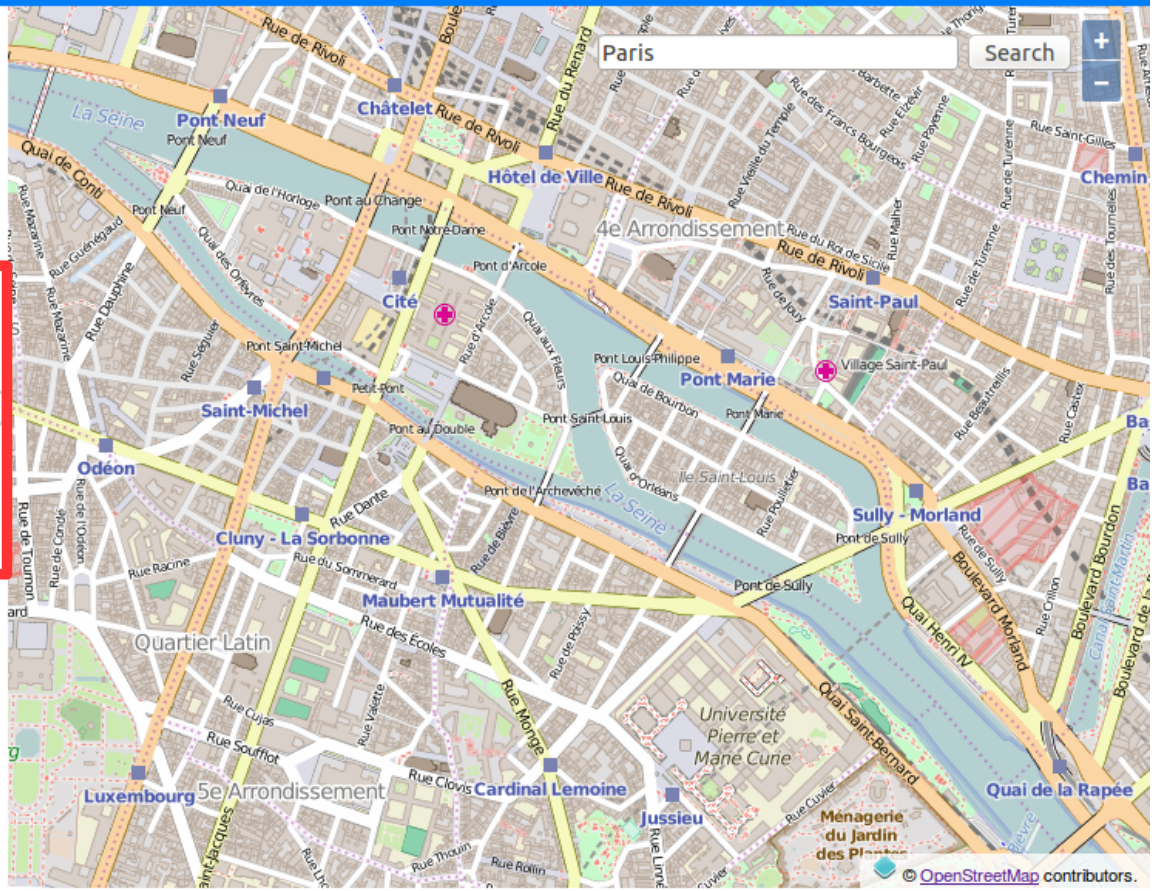
Select an Area (holding *shift key* and dragging the *mouse pointer*) and Retrieve data from the selected area

Selection Retrieve

Select a layer and get data from its bounding box

Select

Bounding box Retrieve



Transposition of the procedure as a WPS

- ✔ Retrieval of OSM road network dataset from the current map view

Comparison of OpenStreetMap (OSM) with authoritative road datasets (REF)

The procedure allows to import customized road datasets from the OSM database and an authoritative source and returns some measures of their spatial similarity.

How to use the tool:

Road network datasets must (for REF) and can (for OSM) be uploaded as a compressed ESRI shapfile. Click the **Browse** button, select the **.zip** file from your system and adjust its color using the **Layer Color** selector. Enter also the reference system of the dataset.

The OSM road network dataset can be also retrieved (using the Overpass API) by clicking the **Retrieve OSM data** button. This will retrieve OSM data for the current map view; alternatively, hold down the shift key to select a rectangle on the map view and then click the **Selection Retrieve** button.

When at least two datasets have been uploaded, click the **Compare menu** button. This will open a menu where you can specify the buffer width for the comparison, the datasets to be used as REF and OSM and (optionally) the mask for clipping the previous datasets - to be previously uploaded as a compressed shapefile. Then click the **Start comparison** button to run the algorithm and download some statistics on the similarity between the datasets in a PDF file.

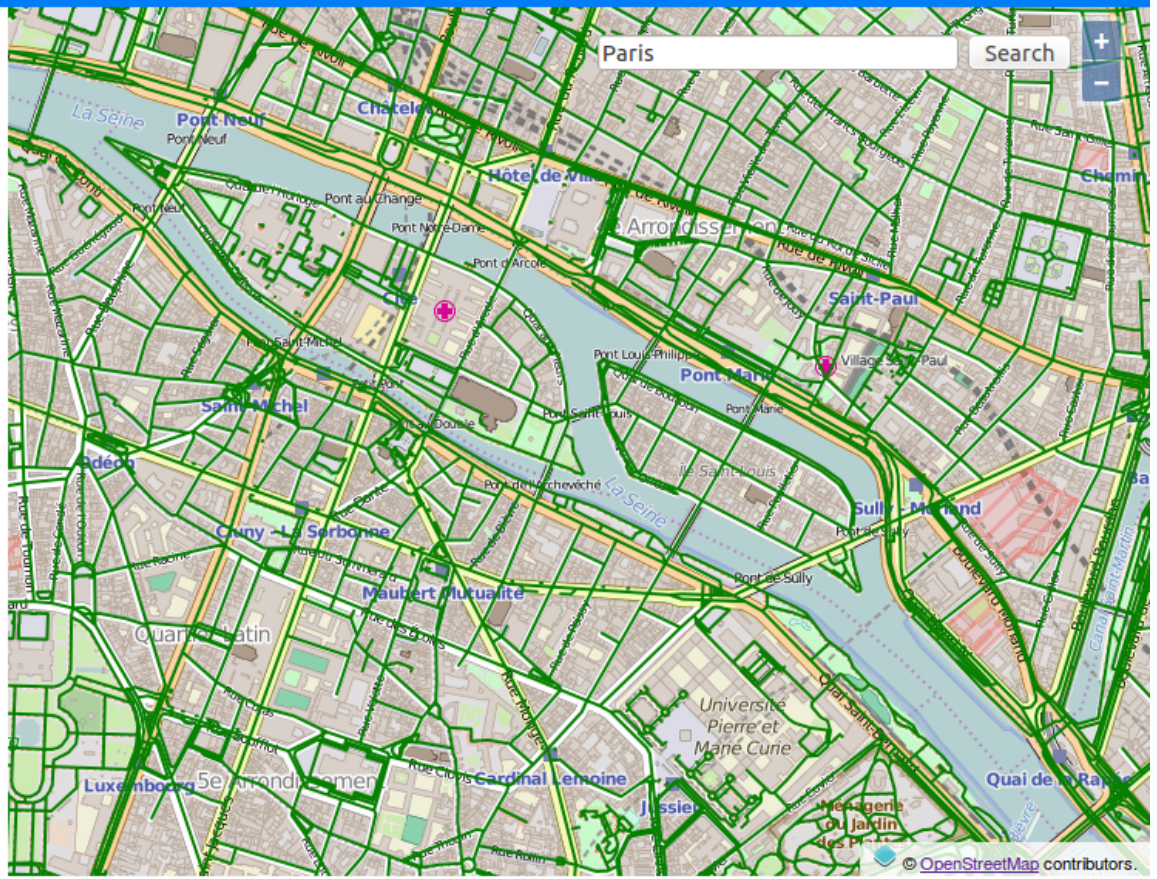
Start Tool

Menu:

Upload

Retrieve OSM data

Compare menu



Transposition of the procedure as a WPS

- Retrieval of OSM road network dataset from a rectangle drawn on the map or from the bounding box of an uploaded layer

Comparison of OpenStreetMap (OSM) with authoritative road datasets (REF)

The procedure allows to import customized road datasets from the OSM database and an authoritative source and returns some measures of their spatial similarity.

How to use the tool:

Road network datasets must (for REF) and can (for OSM) be uploaded as a compressed ESRI shapfile. Click the **Browse** button, select the **.zip** file from your system and adjust its color using the **Layer Color** selector. Enter also the reference system of the dataset.

The OSM road network dataset can be also retrieved (using the Overpass API) by clicking the **Retrieve OSM data** button. This will retrieve OSM data for the current map view; alternatively, hold down the shift key to select a rectangle on the map view and then click the **Selection Retrieve** button.

When at least two datasets have been uploaded, click the **Compare menu** button. This will open a menu where you can specify the buffer width for the comparison, the datasets to be used as REF and OSM and (optionally) the mask for clipping the previous datasets - to be previously uploaded as a compressed shapefile. Then click the **Start comparison** button to run the algorithm and download some statistics on the similarity between the datasets in a PDF file.

Start Tool

Menu:

Upload

Retrieve OSM data

Compare menu

Layer Color

Layer Name

OSM

Retrieve data from the current map view

View Retrieve

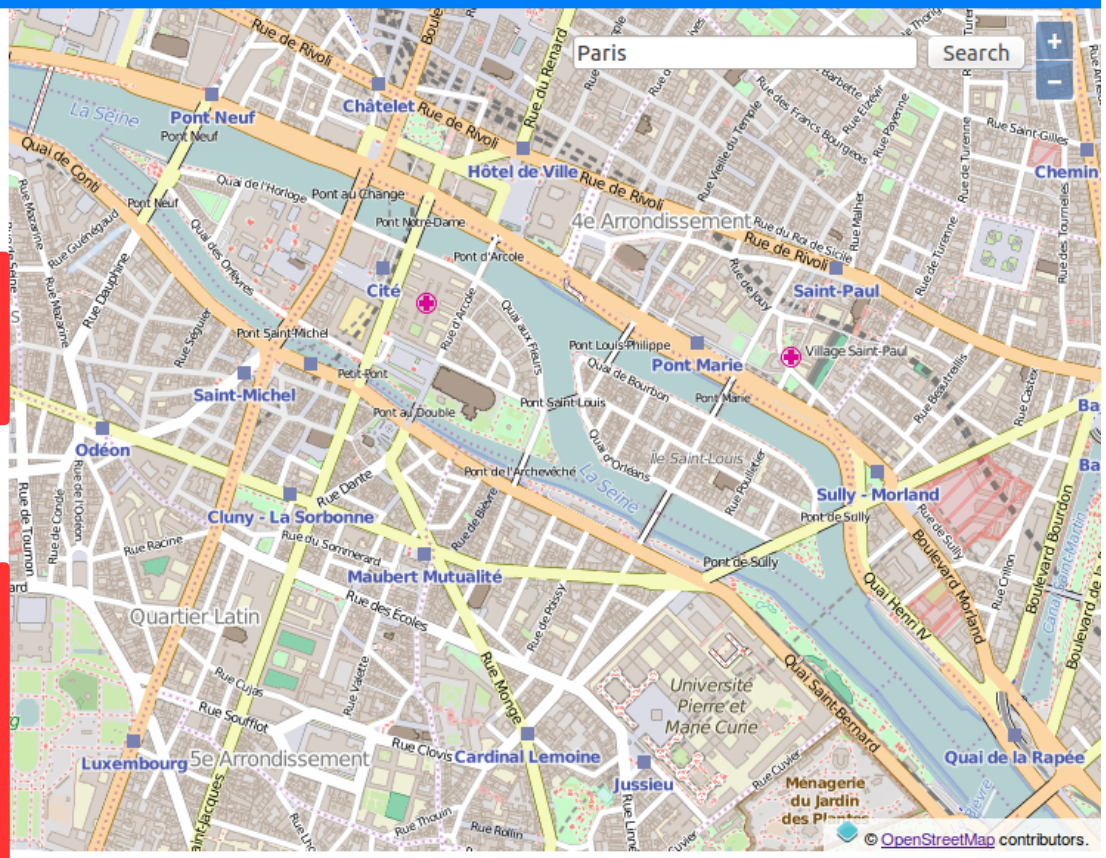
Select an Area (holding *shift key* and dragging the *mouse pointer*) and Retrieve data from the selected area

Selection Retrieve

Select a layer and get data from its bounding box

Select

Bounding box Retrieve



Conclusions

- Methodology for comparing OSM and authoritative road datasets:
  - (hopefully) fills a gap in literature
  - not tied to a specific reference dataset
  - generic, flexible and adaptable to any reference dataset
  - users have a key role in driving the procedure
  - parameter values should reflect the characteristic of the reference dataset involved (e.g. nominal scale and accuracy)
- Future work:
  - understand the influence of parameters through sensitivity analysis
  - reduce computational time (especially for Step 2)
  - increase usability through a WPS implementation (also for Steps 2 & 3)
  - extend the procedure to also compare attributes
  - test on different datasets (any test dataset is welcome)!

Links & publications

- Links:
  - source code: https://github.com/MoniaMolinari/OSM-roads-comparison
  - WPS client: http://131.175.143.84/WPS

- Related publications:
  - Brovelli M. A., Minghini M., Molinari M. and Mooney P (in press). Towards an automated comparison of OpenStreetMap with authoritative road datasets. *Transactions in GIS*.
  - Antunes F., Fonte C. C., Brovelli M. A., Minghini M., Molinari M. and Mooney P. (2015) Assessing OSM Road Positional Quality with Authoritative Data. *Proceedings of the VIII Conferência Nacional de Cartografia e Geodesia*, Lisbon (Portugal), October 29-30, 2015.
  - Brovelli M. A., Minghini M., Molinari M. and Mooney P. (2015) A FOSS4G-based procedure to compare OpenStreetMap and authoritative road network datasets. *Geomatics Workbooks* 12, 235-238, ISSN 1591-092X.

Contacts

Thanks for your attention!

Marco Minghini

Politecnico di Milano
GEOlab – Polo Territoriale di Como
Via Valleggio 11, 22100 Como (Italy)

marco.minghini@polimi.it

@MarcoMinghini