

MSc. Geomatics Thesis P5

Chengzhi Rao 5841089

Supervisors: Dr. Ken Arroyo Ochori
Dr. Stelios Vitalis

Co-reader: Dr.ir. Martijn Meijers

***Reconstructing a high-detailed 2D areal representation
of road network based on OSM data***

Contents

▪ Introduction & Background:

- Motivation
- Background
- Research questions
- Research objective

▪ Methodology & Results:

- Top-level of pipeline
- 3 Stages/phases
- Overview of results
- Detailed results

▪ Conclusion:

- Answers of research question and subquestions

Key words:

- ✓ OpenStreetMap, Road, Intersection,
- ✓ Road network model, Lane-level network, Linear representation, Areal representation,
- ✓ Topology, Graph, Geometry, Road centerline, Semantics, Traffic modes.

Introduction & Background

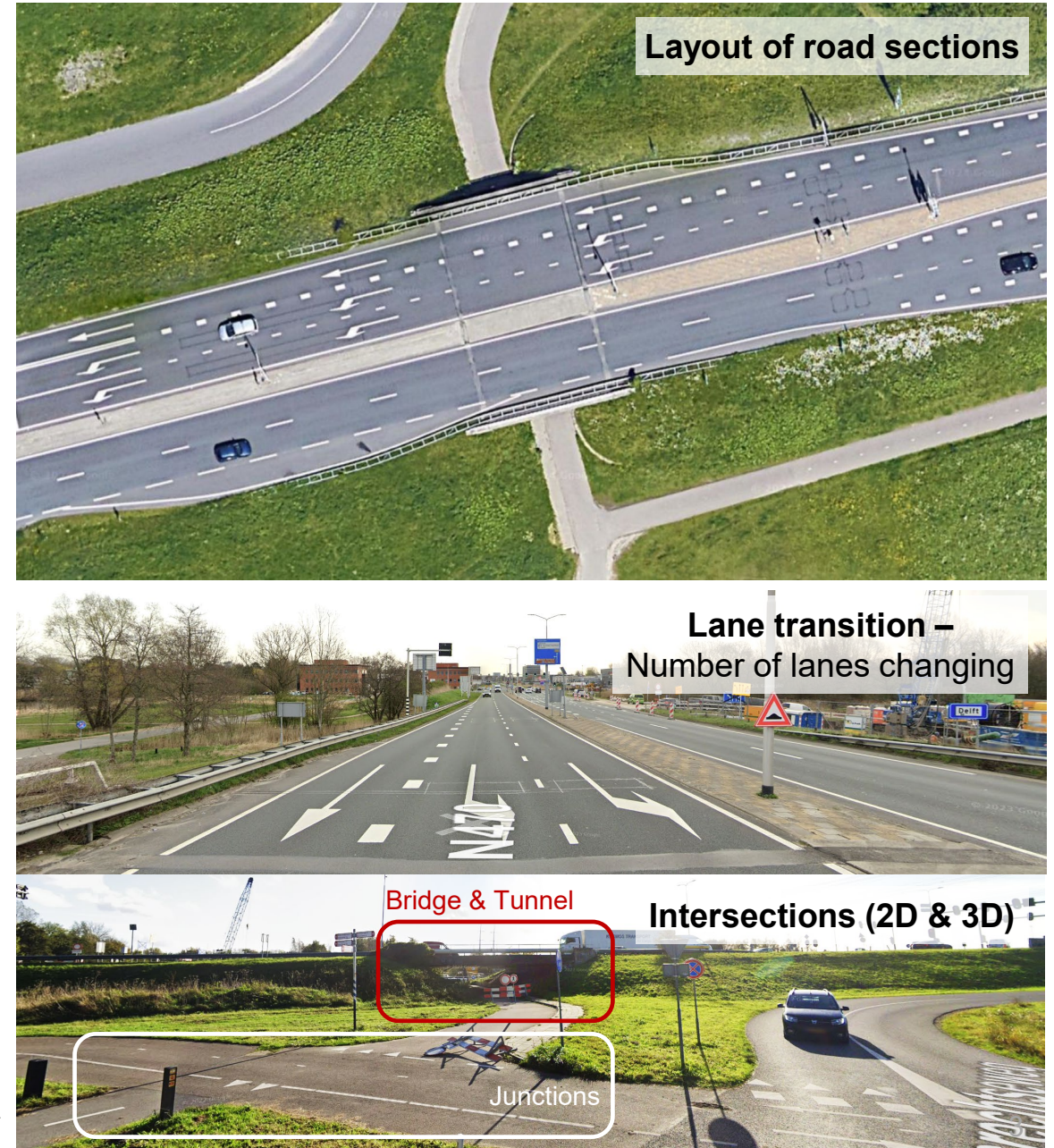
Motivation:

● Road network in real world

Physical road networks are complex structured system composed of **a large number of road segments and intersections**, shaping the **layout of cities** and built environment;

vs. Digital road network

A **virtual representation** of real-world road infrastructure: it encompasses the **geometry**, **topology**, and **attributes** of roads, intersections, etc., stored in a **digital format**.

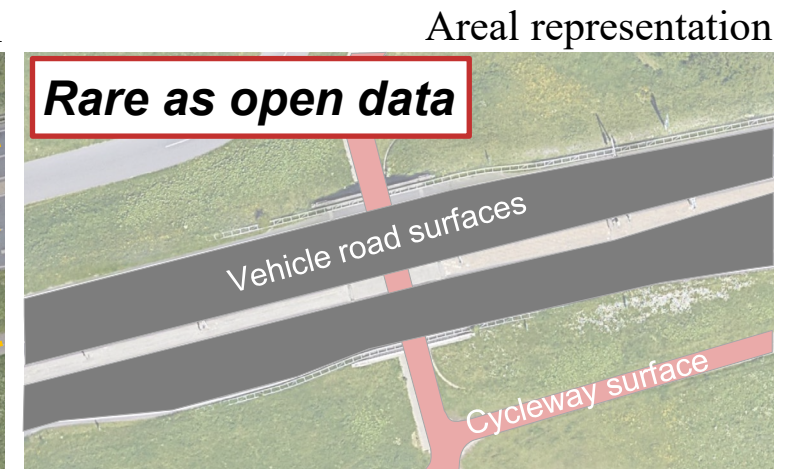
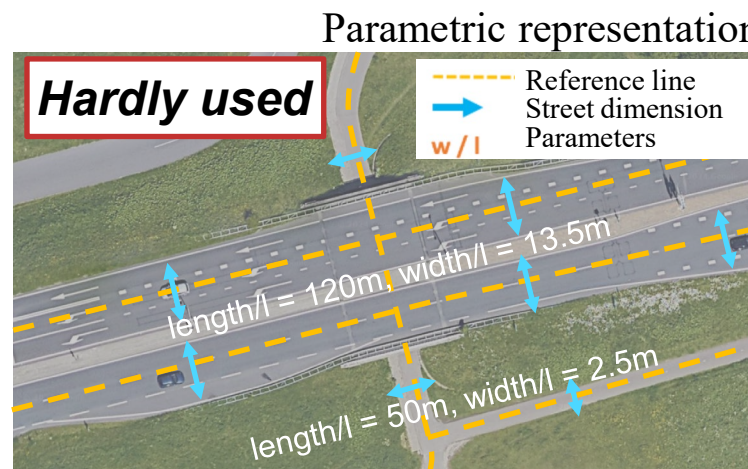
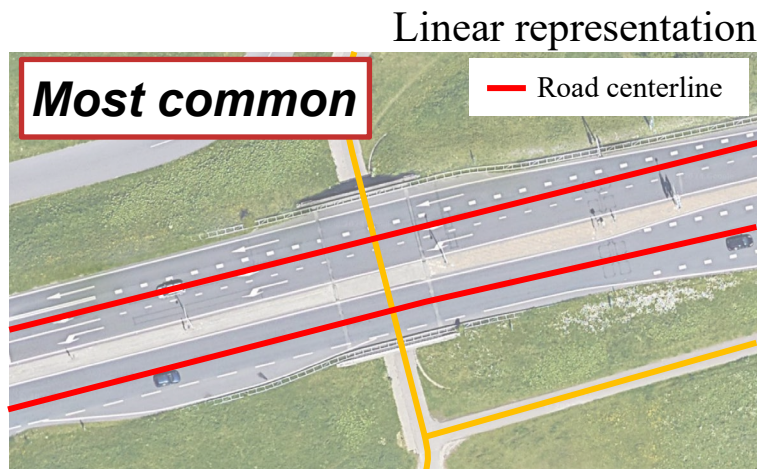
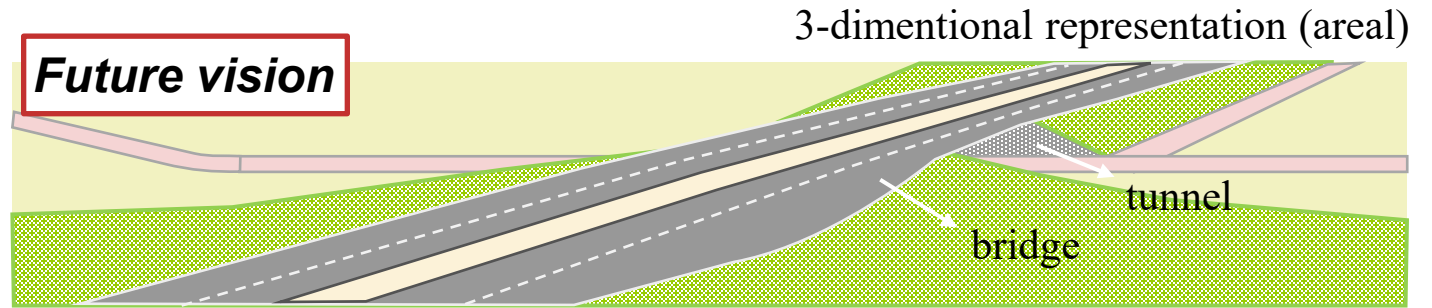


Introduction & Background

Motivation:

- **Types of road network representation**

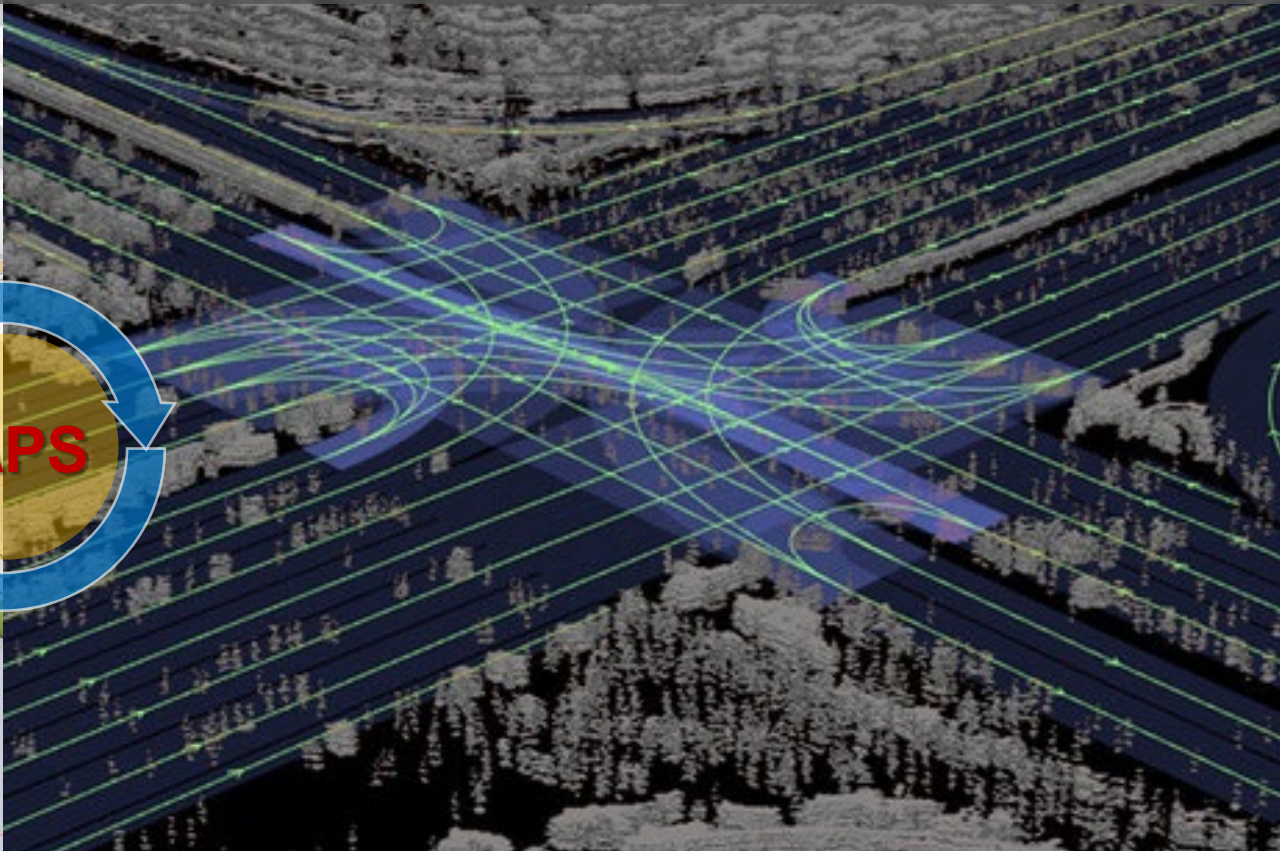
1. Linear representation
2. Parametric representation
3. Areal representation
4. 3D representation (linear, areal, volumetric)



Why reconstructing a high-detailed 2D areal representation of road network still matters?



3D building, terrain and vegetation models



Linear representation and driving trajectories

Introduction & Background

Motivation:

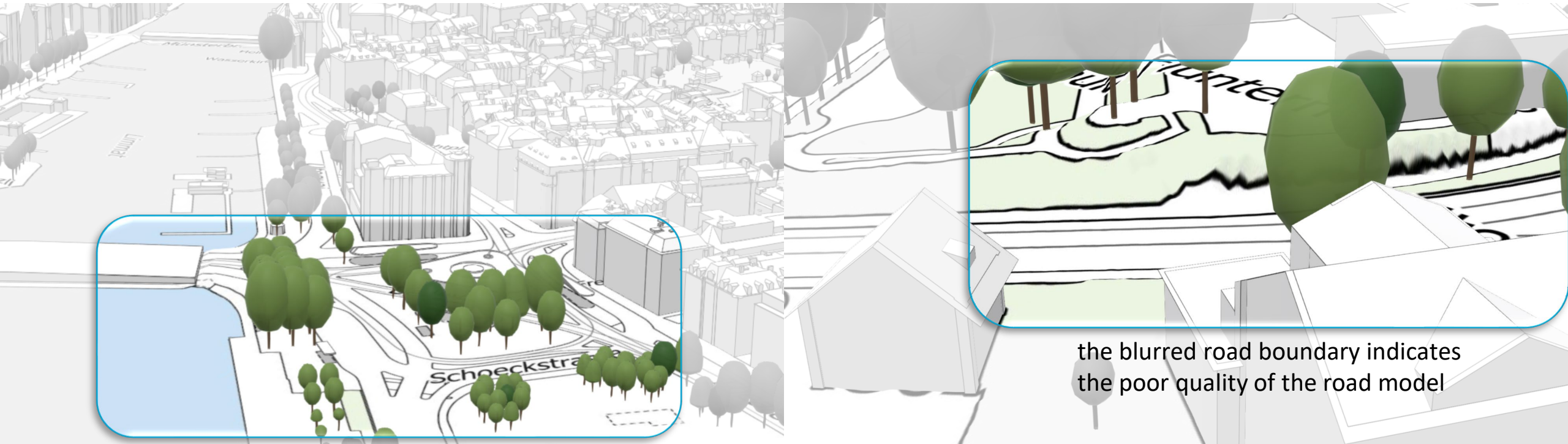
- Start-of-the-art 3D city model



Introduction & Background

Motivation:

- Research gaps among current studies



- 1. Not compact 2D polygon**
- 2. Cannot connect to the real 2.5D or 3D objects**

Introduction & Background

Motivation:

- A high-detailed 2D areal representation and its **importance**

Types	Focus	Topology	Geometry	Semantic
Linear representation	Polyline as a single reference line	✓	?	*
Areal representation	Polygon as the road shape boundary	?	✓	*
3D representation	Road shape/centerline with elevation	From 2D*	2D shape*	From 2D*

The **shift** from linear to 2D areal representations -----

- **A more accurate description:**
Road widths, boundaries, and surface characteristics by translating linear geometries into polygons.
- **Crucial for applications:**
Navigation, road safety management, parking space calculations, urban heat island effects, etc.
- **Integration with 3D data:**
Ensure that digital road models align seamlessly with other urban elements like buildings, water bodies, and terrain.



Changes of the road width

Introduction & Background

Research question:

“How can we achieve a high-detailed areal representation of road network model with **topological** and **geometric** correctness, and enrich its **semantic** information by only using the OpenStreetMap dataset?”



Introduction & Background

Background:

- A high-detailed 2D areal representation and its **Level of Details (LoDs)**
- Variation and changes of roads:
 1. Road / Carriageway / Lane
 2. Driving direction(s): one-way road, forward, backward
 3. Traffic modes: vehicle, cycling...
 4. Road/lane width: symmetric/asymmetric road section
 5. Intersection shapes

Lane level centerline →
Lane level polygon →
Complete road polygon

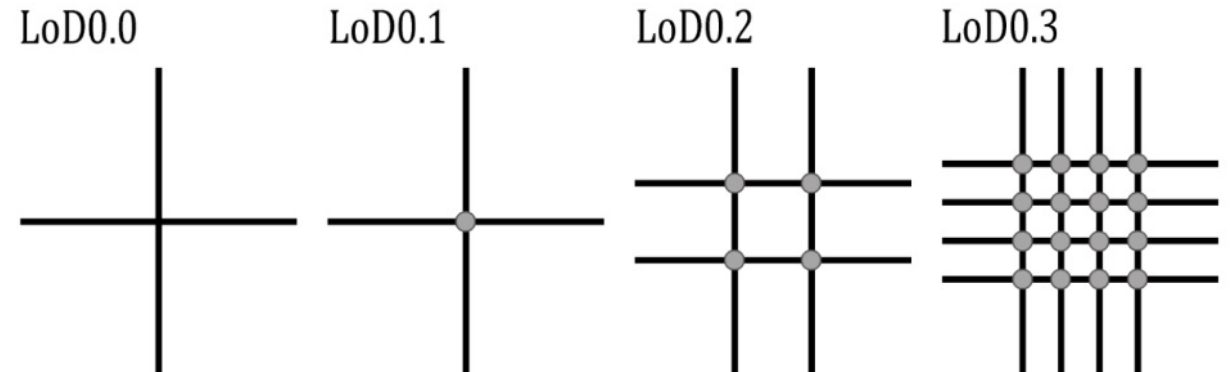


Figure 1. Refinement of the network/**linear** representation of roads in CityGML

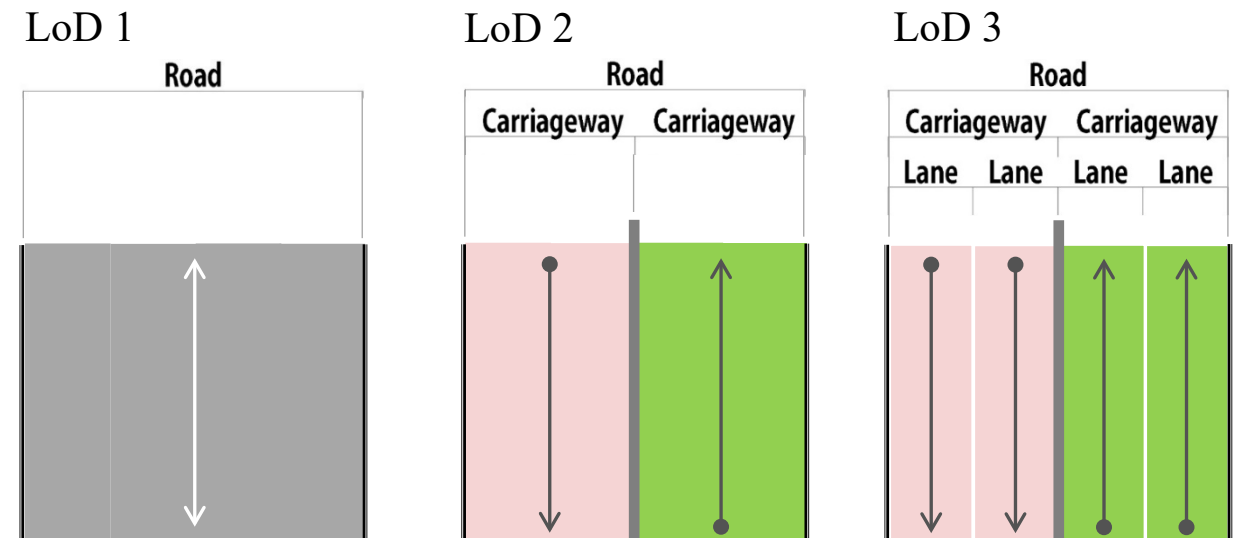
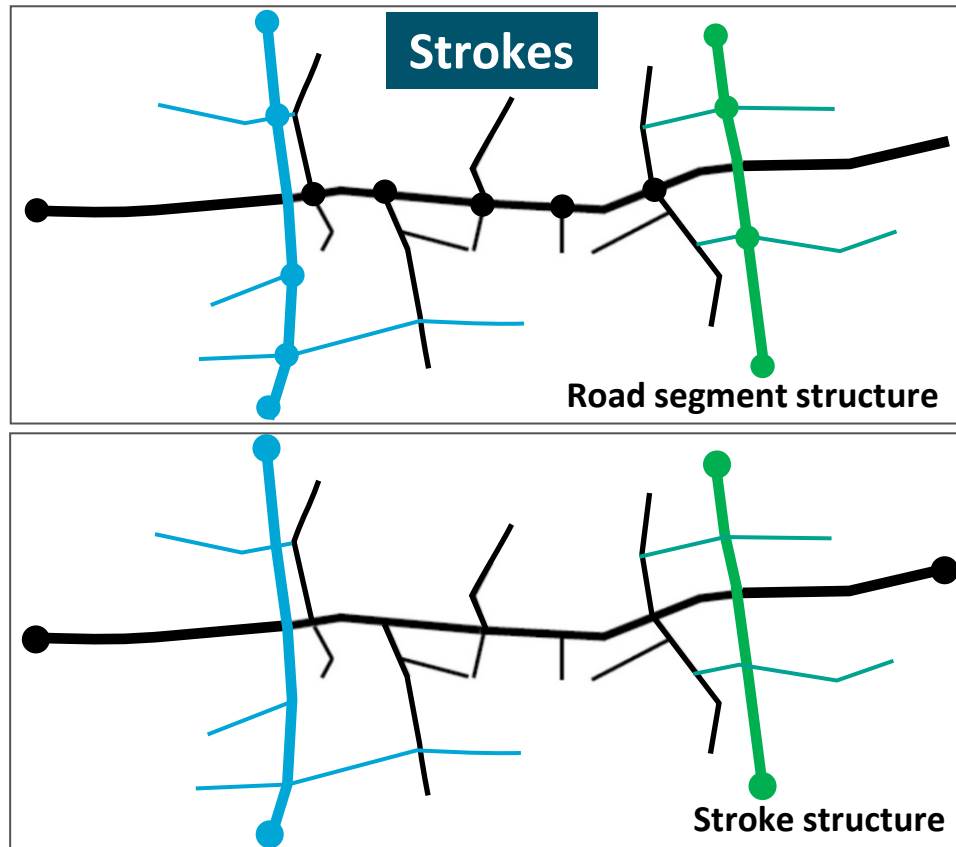


Figure 2. Refinement of the network/**areal** representation of roads in CityGML

Introduction & Background

Background:

- Geometric design rules/civil engineering principles



Network level: continuity of roads in network
Segment level: layout/arrangement of lanes

When observing a road network, natural linear elements will be seen which **extend through junctions** and maintain the **topology and geometry connectivity**, the aggregation of these elements is the “stroke”:

“a set of one or more segments in a non-branching and connected chain”.

- *its hierarchical characteristics*
- *the same road types*
- *no abrupt change in direction, or they intersect at a small angle.*

Introduction & Background

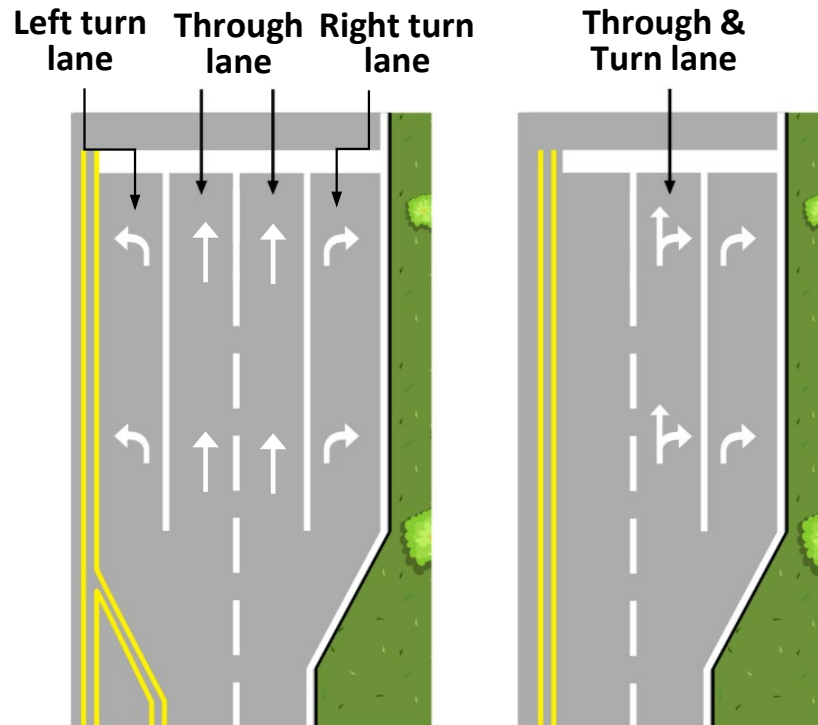
Background:

- Geometric design rules/civil engineering principles

Network level: continuity of roads in network

Segment level: layout/arrangement/relative location of lanes

Auxiliary lane & 'through' lane



Auxiliary lanes that form the transition between **the main route and the ramp/slip roads**.

An auxiliary lane is a section of the roadway adjacent to the through lanes, designed for purposes such as **turning, storage for turning, that support the movement of 'through' traffic**.

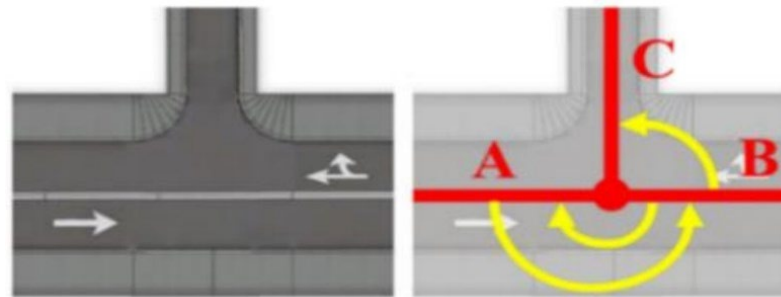
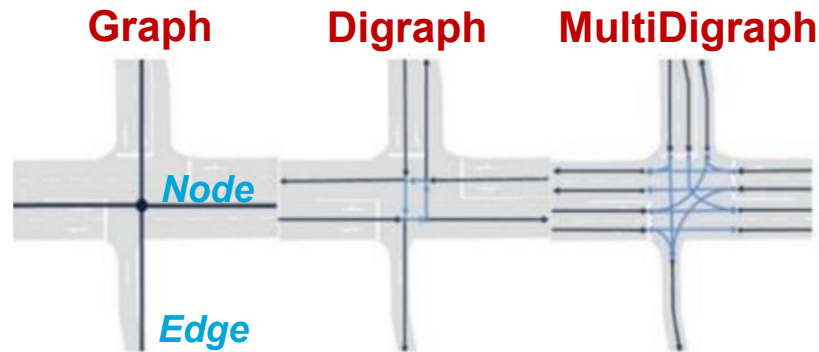
Introduction & Background

Background:

Graph theory

Topological characteristics in linear representation:

- **Node**
- **Edge**
- Edge with direction (adjacency)
- Node has degree as the number of converged edges

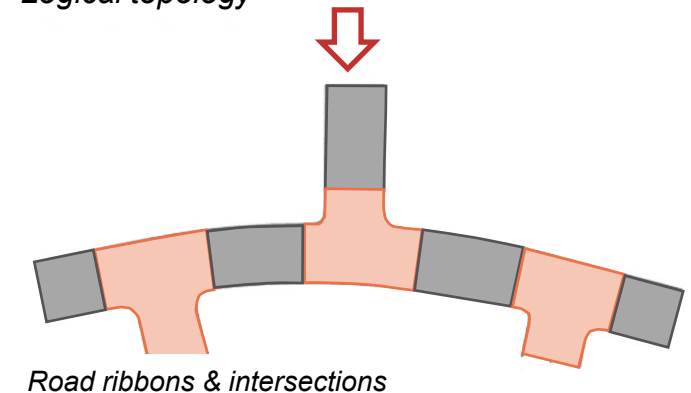
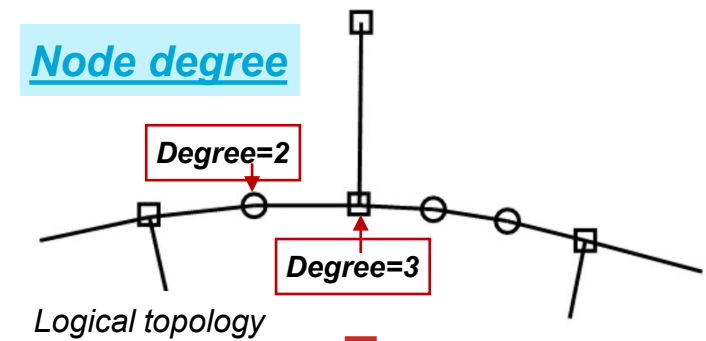
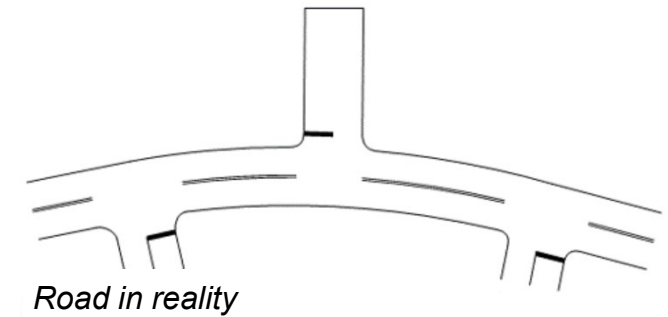


Adjacency

	Predecessor	Successor
A	B	B
B	A	A, C
C	B	-

Road network as Digraph and distinguish the adjacent connection

From linear to areal: inherit topology

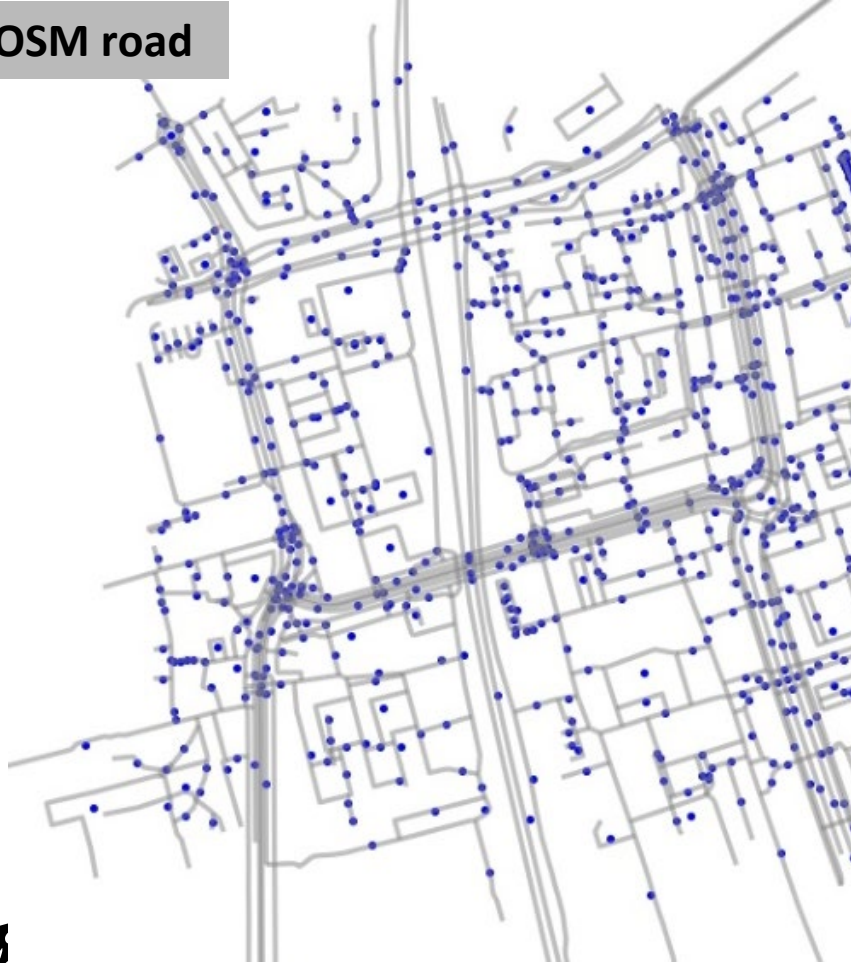


Node degree and road-level topology
↔
road polygon representation and generation

Introduction & Background

Background:

OSM road



OSM road data characteristics

OSM road data as vector, with the various tags and hidden information;

- Additional road/lane;
- Multiple lanes;
- Specific junction...

Polyline

+

Tags

*highway =**
*cycleway =**
*name =**
*lanes =**
*oneway =**
*junction =**
*turn =**
*bridge =**
*tunnel =**

Objective & Questions

Objectives:

**A high-detailed
areal representation**

Only using OSM data:

- *Topology correctness*
- *Geometry accuracy*
- *Semantic enrichment*

Subquestions:

1. How can the **information and attributes contained in OSM** road data be utilized to generate and optimize 2D linear and areal representations of road network models?
2. Is it possible to **enhance the resolution and LoDs** of a road network model from the original road-level to a high-detailed lane-level network using only OSM data?
3. Can the topology relationships, **lane-to-lane adjacency** for representing traffic flow movements, be preserved through OSM data processing and detailed generation methods?
4. How can we **generate 2D road polygons as errorless geometries to accurately capture road shape details**, including changes in road width and intersection variations?

Methodology

Integration of Goals, Gaps, Backgrounds

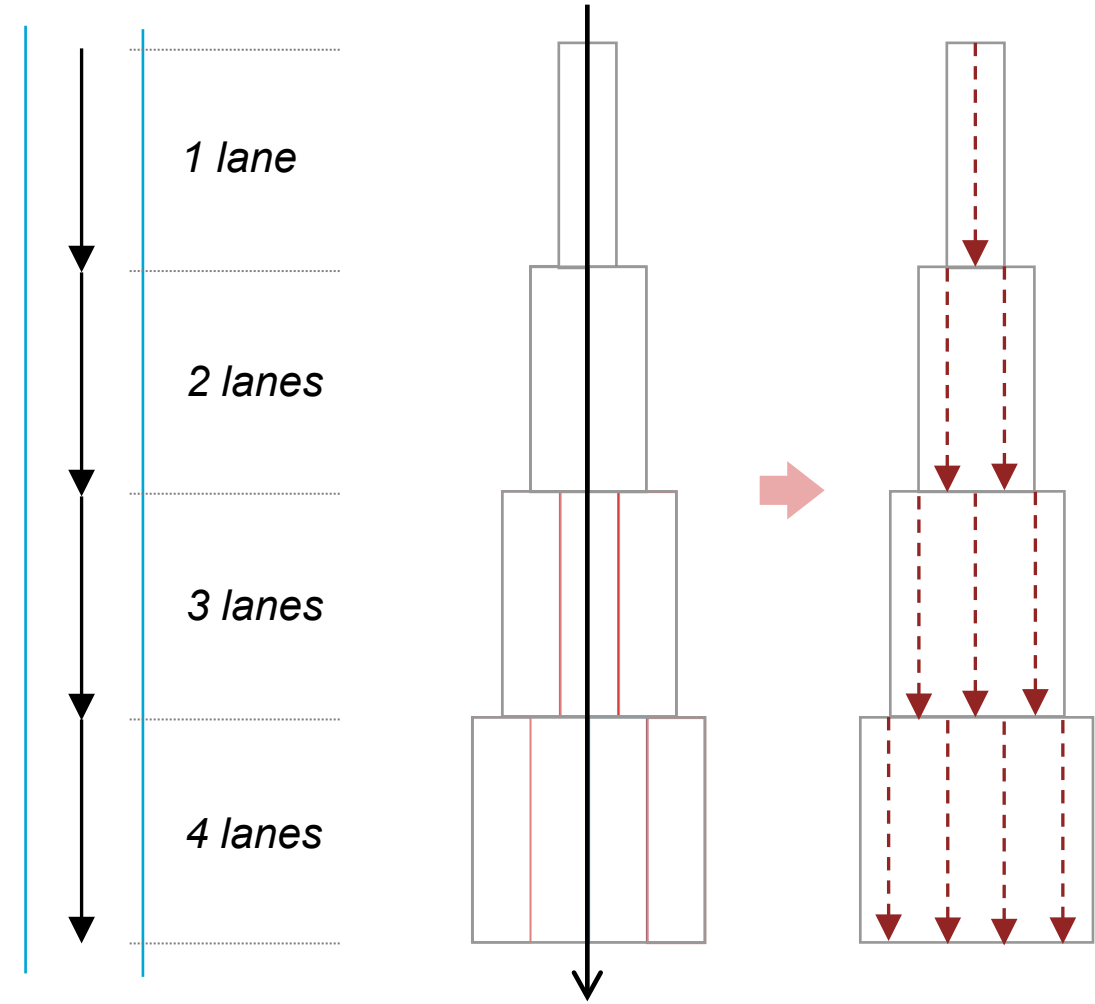
■ Relation of linear center and areal center:

- Road **centerline**
- **Center axis** of road polygon

- 1 continuous road/stroke
- 4 road centerlines as sections
- **Unknown** road **width & shape** changes

!!! Lane connectivity is not guaranteed

1. Centerline is the **center axis** of road polygon boundary

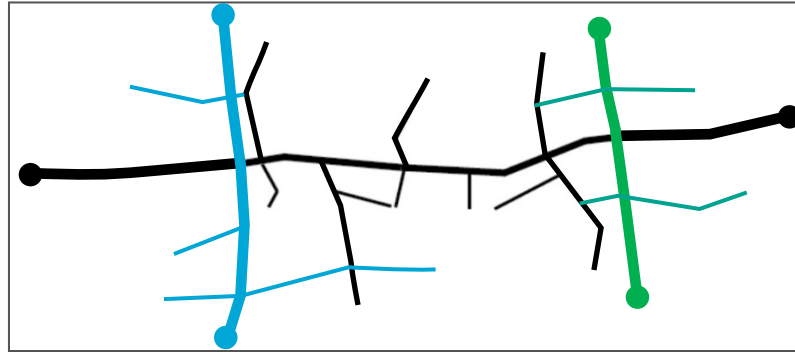


Wrong lane connectivity topology ⇔
Unrealistic road boundary

Methodology

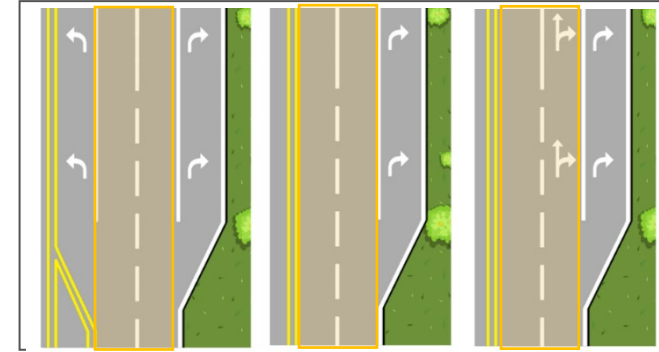
Integration of Goals, Gaps, Backgrounds

Strokes



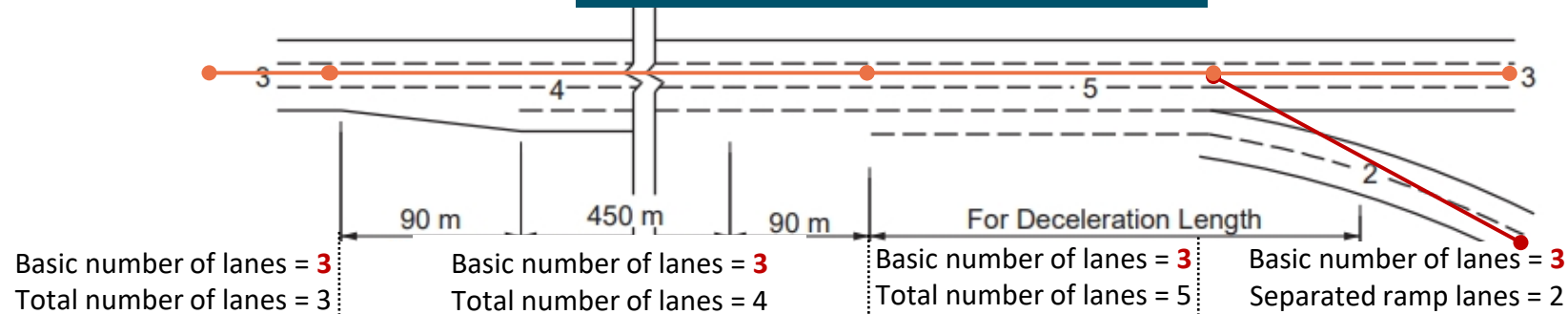
Main routes with 'Good connectivity' in the whole road network

Auxiliary lane & 'through' lane



Traffic entering or exiting the main route of a roadway.

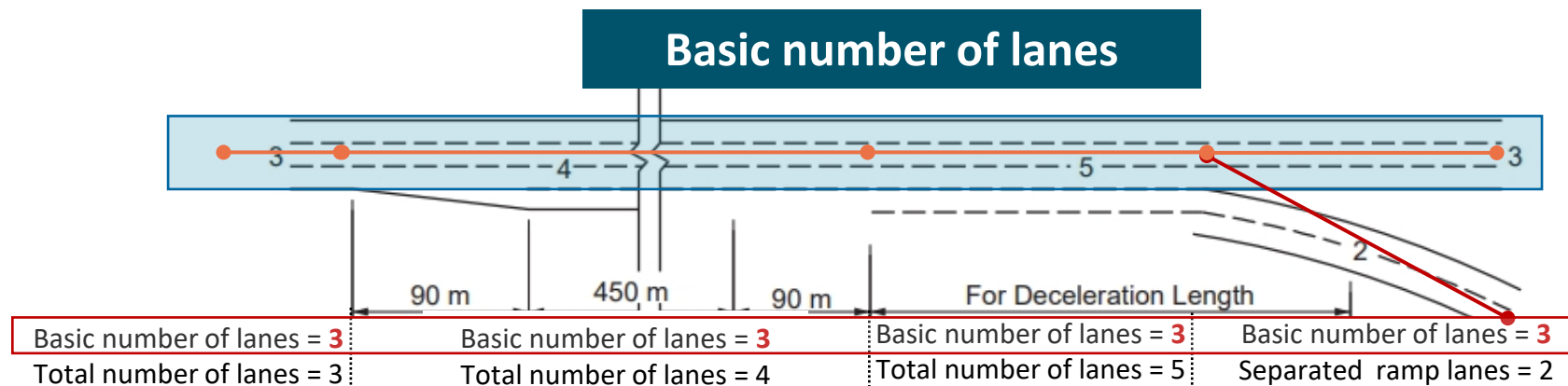
Basic number of lanes



Methodology

Integration of Goals, Gaps, Backgrounds

- It represents a **minimum and constant** number of lanes that are designated and sustained over a considerable length of the route, regardless of changes in traffic volume; **excluding any auxiliary lanes**.
- Maintain **consistency** in the number of lanes along any arterial route.
- Crucial for the **lane level connectivity**.



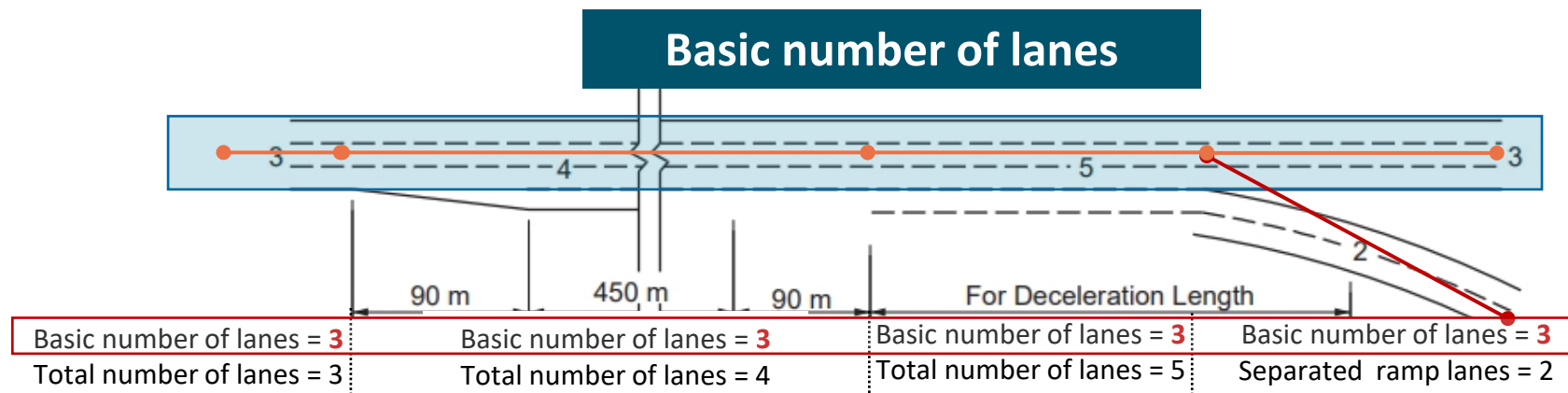
Methodology

Integration of Goals, Gaps, Backgrounds

Basic number of lanes = **3**
in this section/stroke



3 through lanes have
'good connectivity' in lane level



Methodology

Integration of Goals, Gaps, Backgrounds

■ Relation of linear center and areal center:

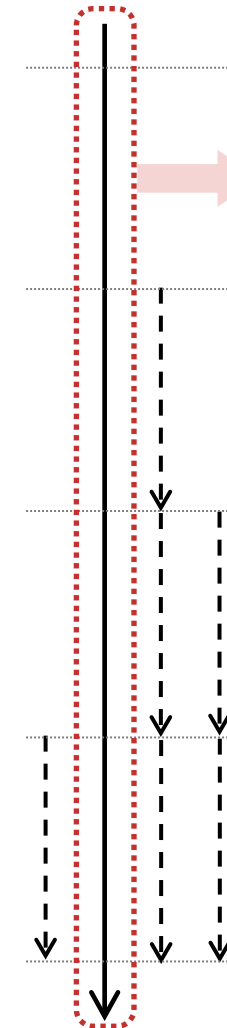
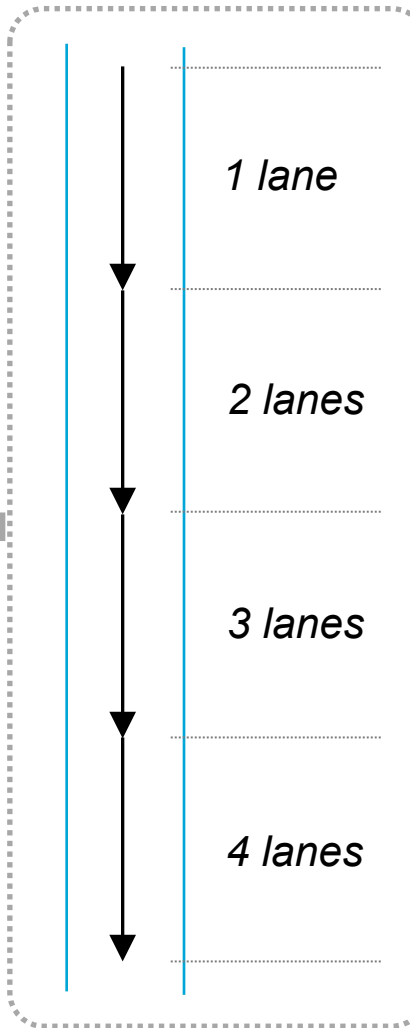
- Road **centerline**
- **Center axis** of road polygon



Road segments within the same grouped stroke S



2. Centerline as **movement shape of road** (through lane) representation



'Basic number of lane'
S_{min} of the grouped stroke represent the lanes that allow the traffic move forward without changing maneuvering

Correct lane connectivity topology

Methodology

Integration of Goals, Gaps, Backgrounds

■ Relation of linear center and areal center:

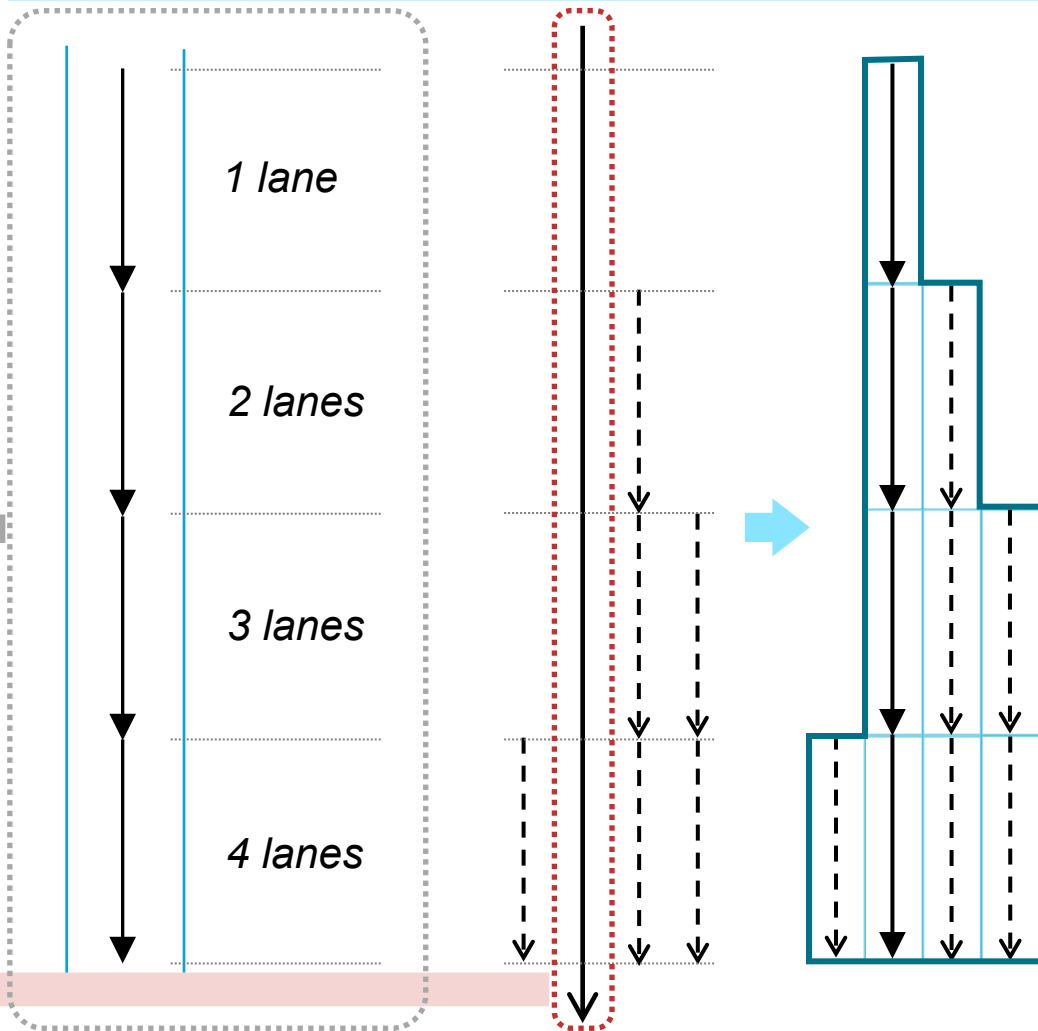
- Road **centerline**
- **Center axis** of road polygon



Road segments
& Stroke **S**

'Basic number
of lane' **S_{min}** &
'through' lanes

2. Centerline as **movement shape of road** (through lane) representation

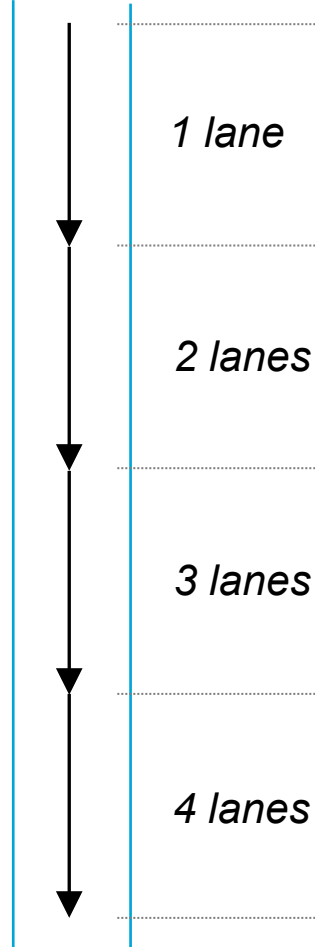


Methodology

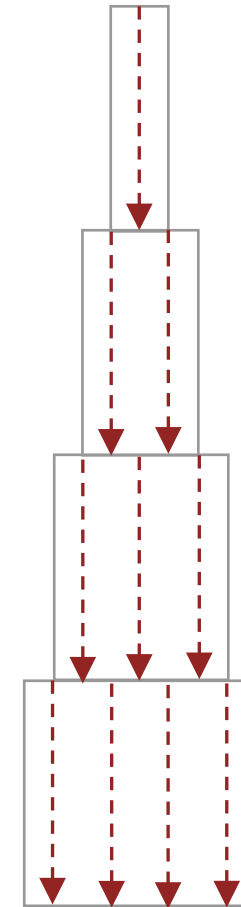
Integration of Goals, Gaps, Backgrounds

- Road level centerline is not always the center axis of road polygon

Strokes
Basic number of lanes
Lane topology

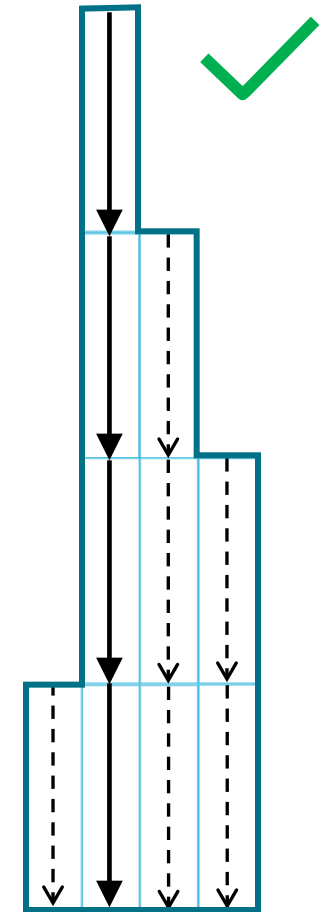


Traditional method



1. Centerline vertices
at the center axis of
road polygon boundary

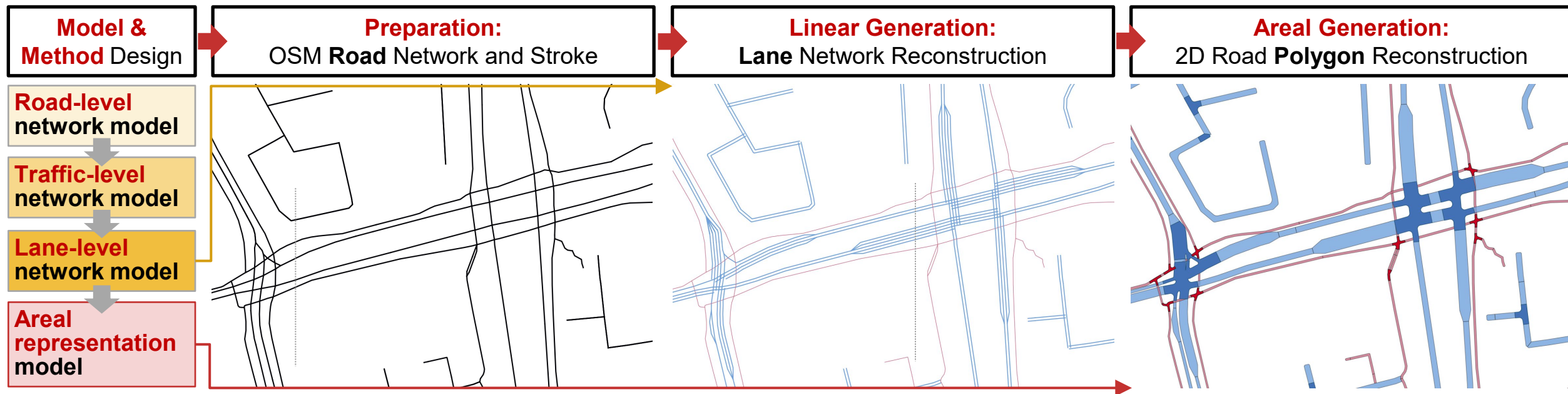
Our approach



2. Centerline as movement
shape representation

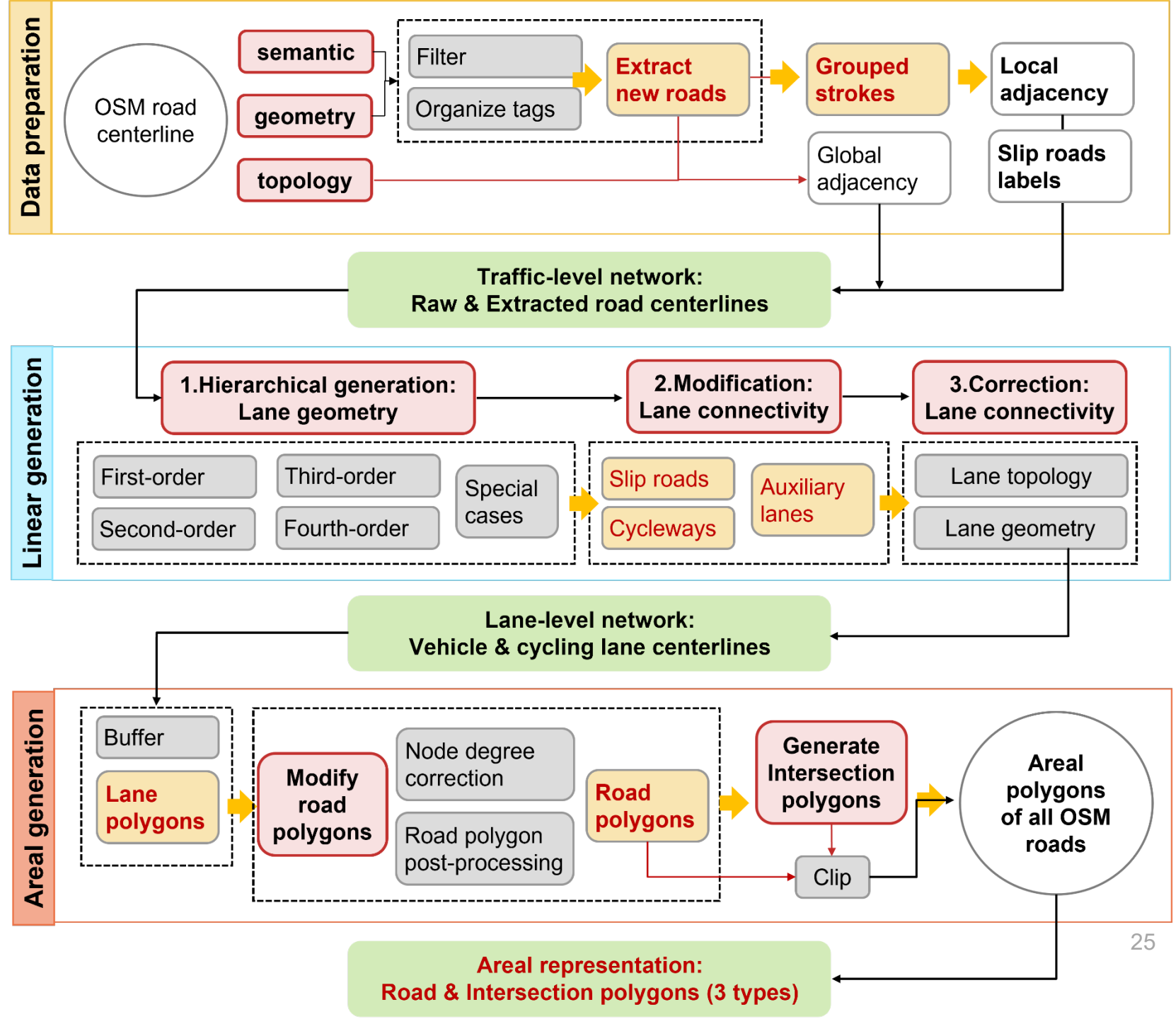
Methodology

Top-level pipeline:



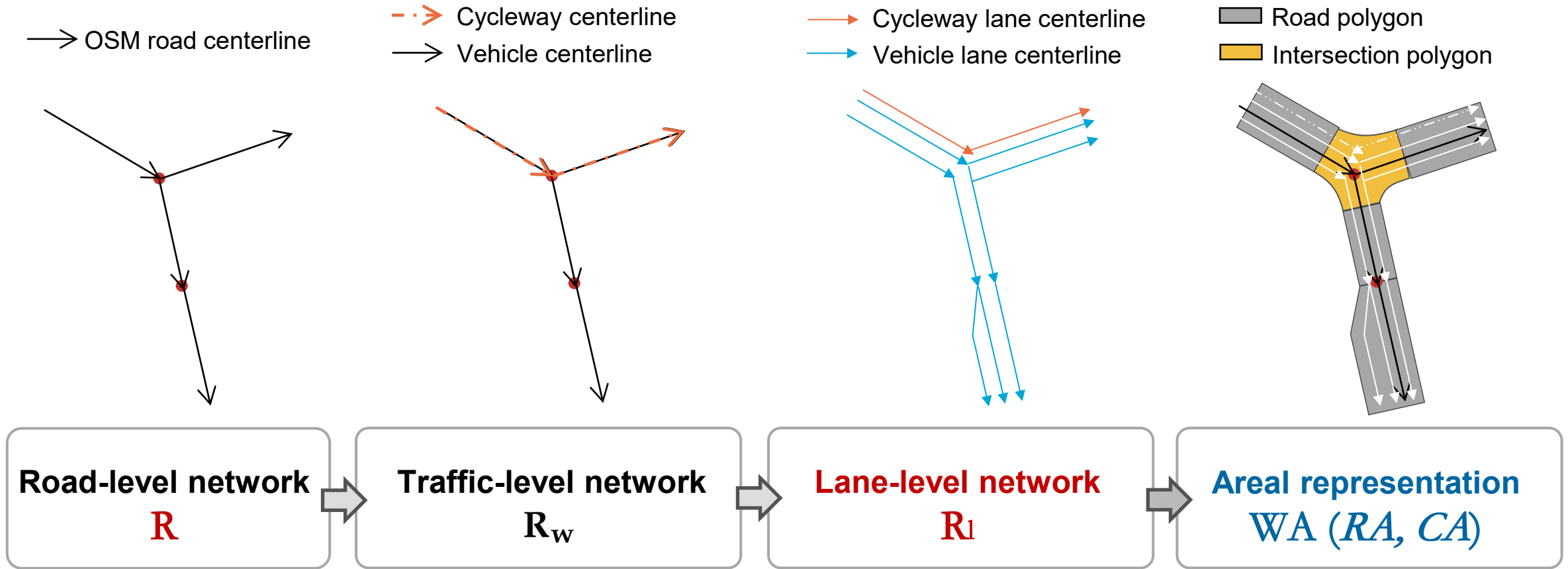
Methodology

Detailed workflow:



Methodology

Model design: *Procedural method*



Methodology

Model design:

Complete lanes as traffic flows in the network.

Turning attributes for the auxiliary lanes are essential for the reconstruction and alignment for 'main route and branches' in a stroke.

Shift linear to areal:
Reused the topology connection and node degree to form the road ribbons and intersections.

Road-level network
 R

Traffic-level network
 R_w

Forward vehicle roads

R_{wv}^f

Backward vehicle roads

R_{wv}^b

Cycleways & Bike lanes

R_{wb}

Lane-level network
 R_l

'through' vehicle lanes

'left_turn' vehicle lanes

'right_turn' vehicle lanes

'left_side' bike lanes

'right_side' bike lanes

Separate cycleways
(one-way & two-ways)

Areal representation
 $WA(RA, CA)$

Lane polygons

Road polygons

+

True node degree



Polygons of road ribbons

Polygons of all types of intersections

Methodology & Results

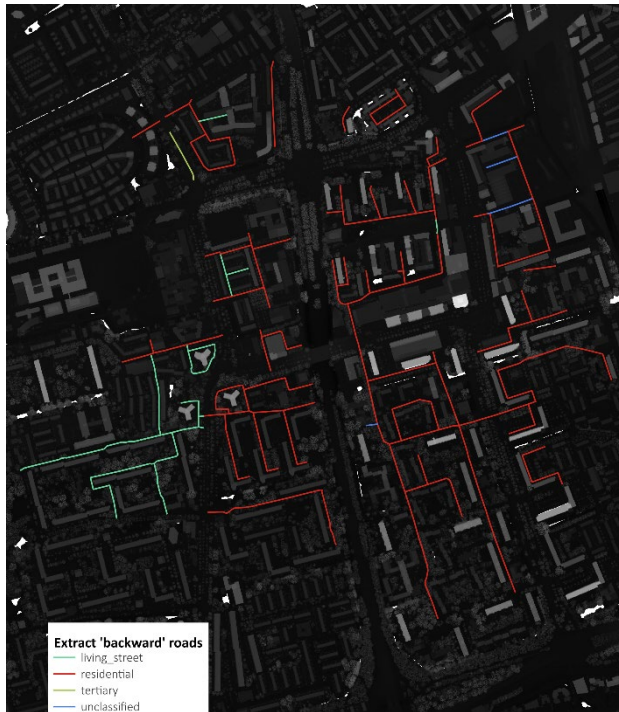
Phase 1 — Preparation: OSM Road Network and Stroke

➡ Subquestion 1 & 2

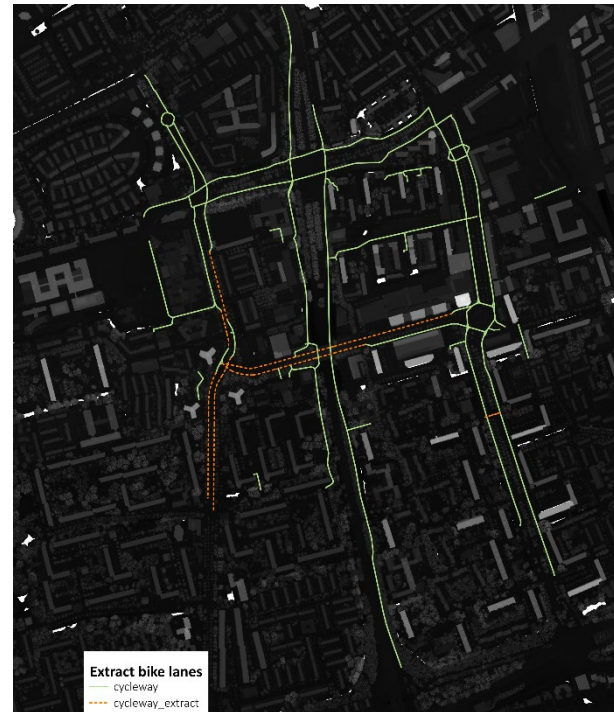
Semantic information and attributes in OSM road data tags



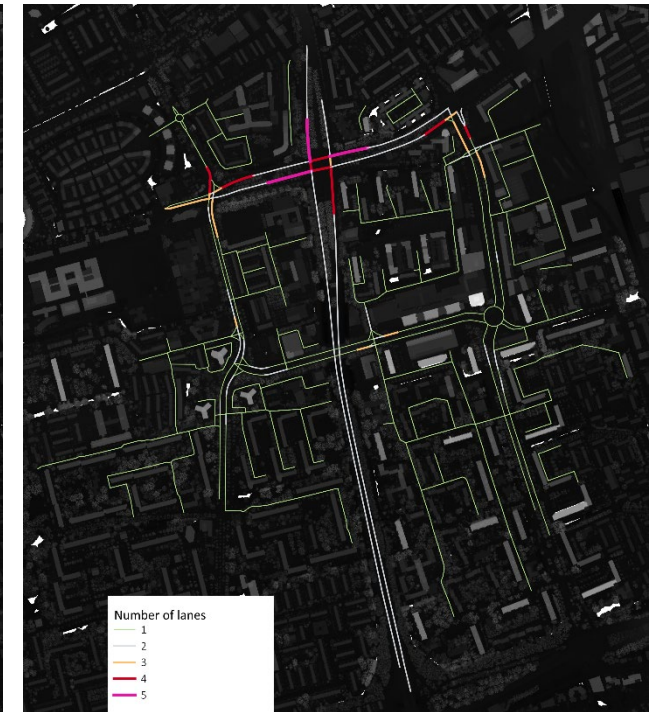
Raw data has **393** elements as the road centerlines



Identified **120** two-way roads, so added 120 backward road centerlines with the reversed vector direction



Extracted **41** bike lanes, including 3 left bike lanes, and 38 right bike lanes



Total number of elements after data pre-processing is **554**, as the traffic-level road network. The number of lanes of all roads are from **1 to 5**.

Methodology & Results

Phase 1 — Preparation: OSM Road Network and
Semantic information and attributes in OSM road data tags

Tags ---

*highway =**
*cycleway =**
*name =**

Raw data has **393** elements as
the road centerlines.
Region: Delft.



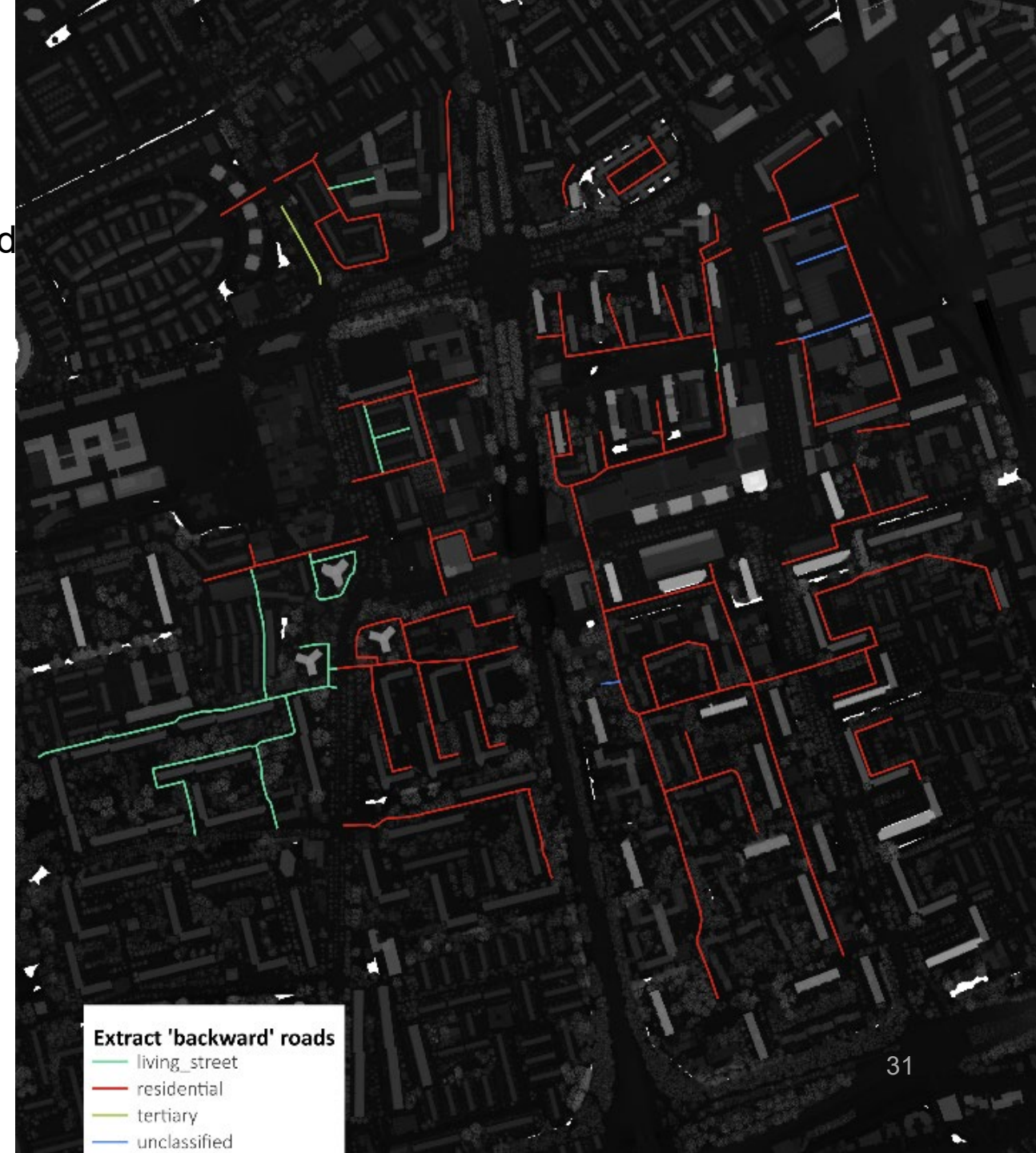
Methodology & Results

Phase 1 — Preparation: OSM Road Network and
Semantic information and attributes in OSM road data tags

Tags ---

oneway = *
cycleway = *

Identified **120** two-way roads,
so added **120 'backward'** road
centerlines with the **reversed**
vector direction.



Extract 'backward' roads

— living_street
— residential
— tertiary
— unclassified

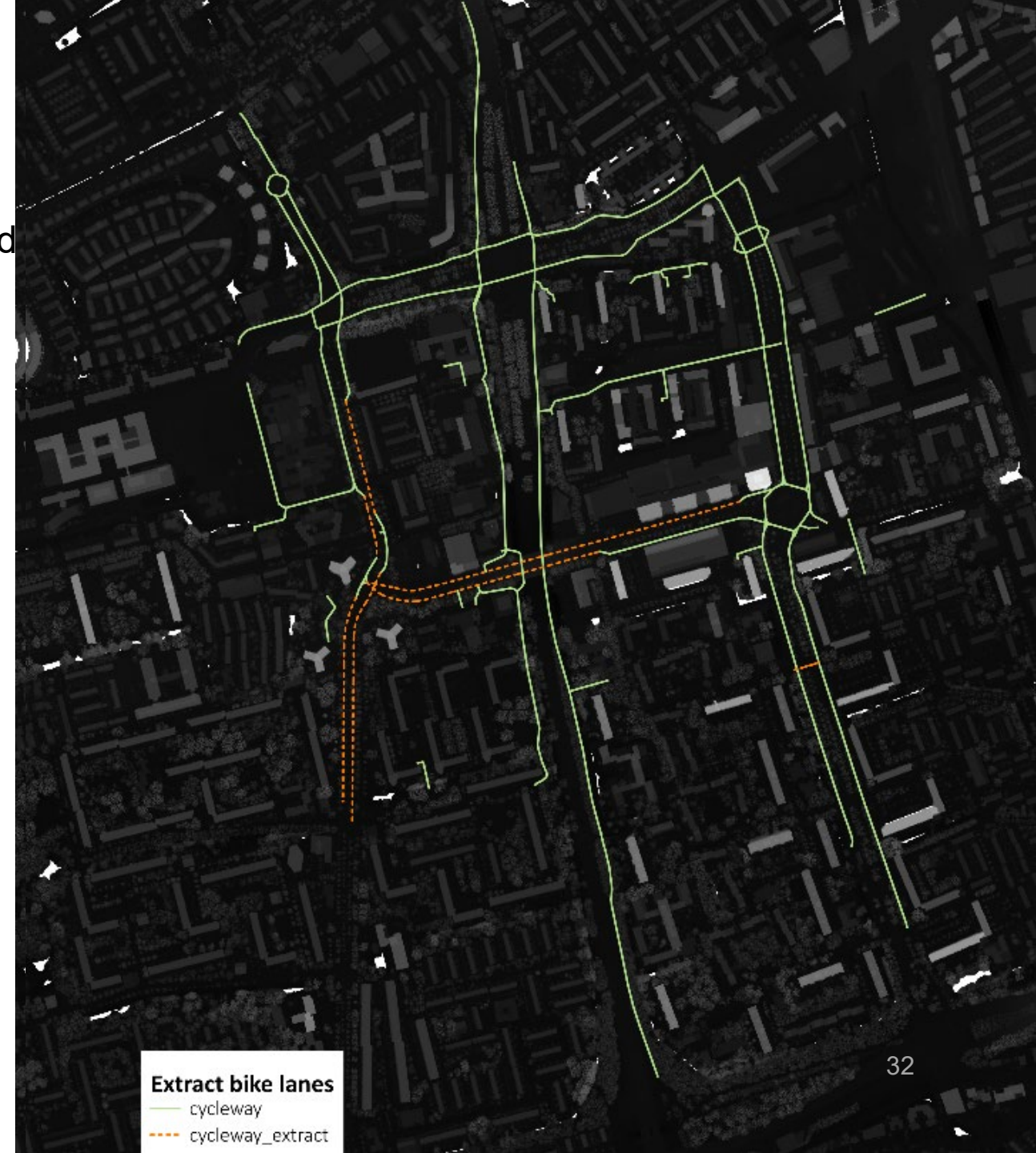
Methodology & Results

Phase 1 — Preparation: OSM Road Network and
Semantic information and attributes in OSM road data tags

Tags ---

`cycleway = *`

Extracted **41** bike lanes,
including 3 left bike lanes,
and 38 right bike lanes.



Methodology & Results

Phase 1 — Preparation: OSM Road Network and
Semantic information and attributes in OSM road data tags

Tags ---

cycleway =*
lanes =*
oneway =*
turn =*

Total number of elements after data pre-processing is **554**, as the traffic-level road network. The number of lanes of all roads are from **1 to 5**.

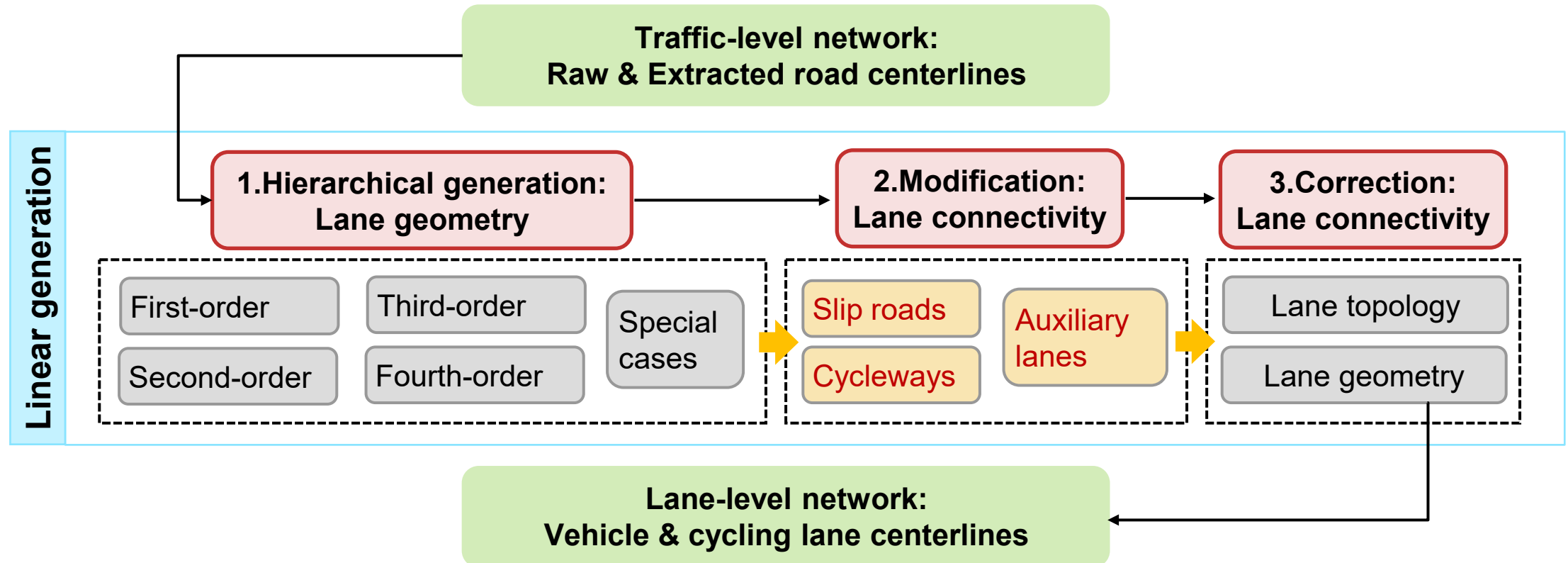


Methodology

Can the topology relationships, **lane-to-lane adjacency** for representing traffic flow movements, be preserved through OSM data processing and detailed generation methods?

Phase 2 — **Linear Generation:** Lane Network Reconstruction

➡ **Subquestion 3**



Methodology

Phase 2 — Linear Generation: Lane Network

Hierarchical generation: Lane geometry

- Lane topology correctness:
'Lane connectivity'

Road-level:

One-way roads:

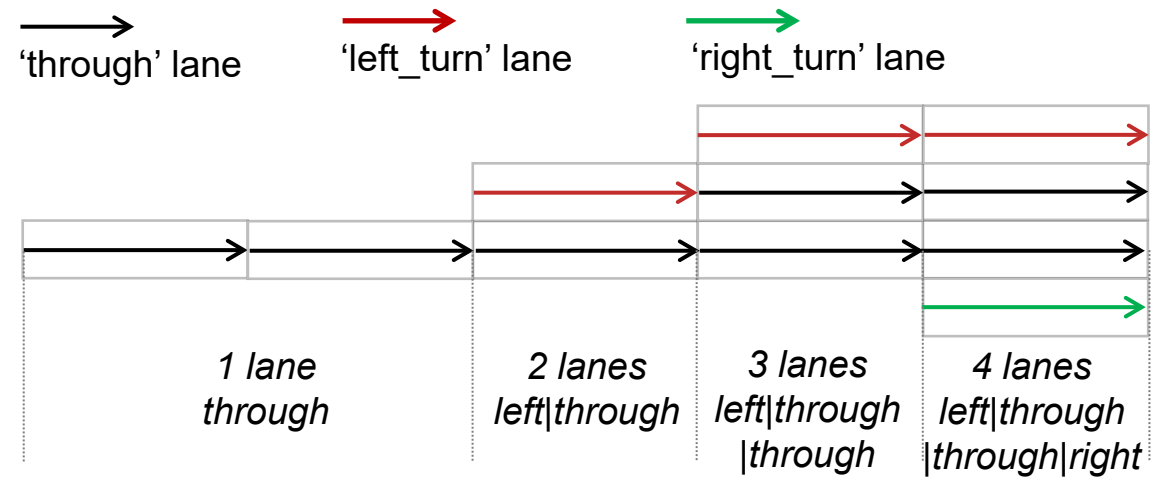
1st -3rd steps
for all the one-way roads

'basic number of lanes'
within each 'grouped
stroke'

Hierarchical orders of
all the roads

Lane-matching method

Hierarchical roads and lane-level stroke



Methodology

Phase 2 — Linear Generation: Lane Network

Hierarchical generation: Lane geometry

- Lane topology correctness:
'Lane connectivity'

Road-level:

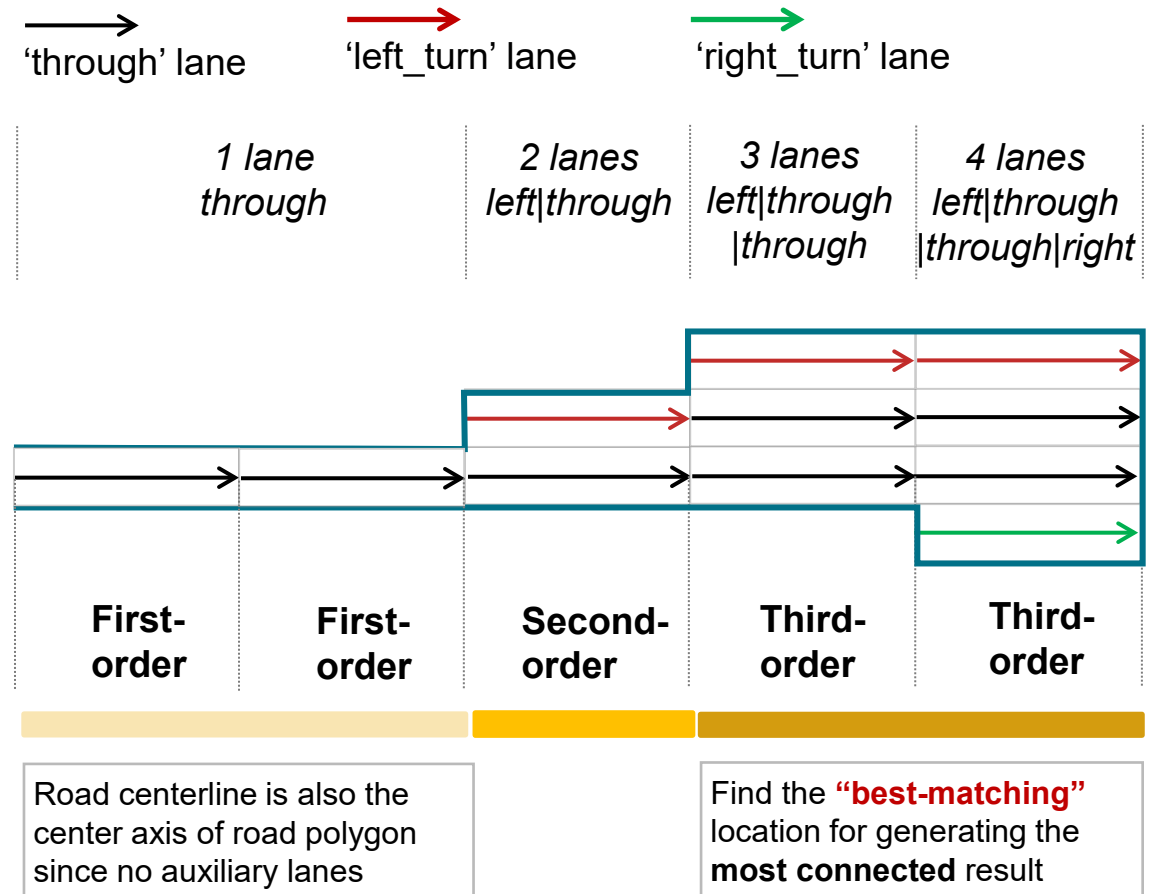
One-way roads:

1st -3rd steps
for all the one-way roads

'basic number of lanes'
within each 'grouped
stroke'

Hierarchical orders of
all the roads

Lane-matching method

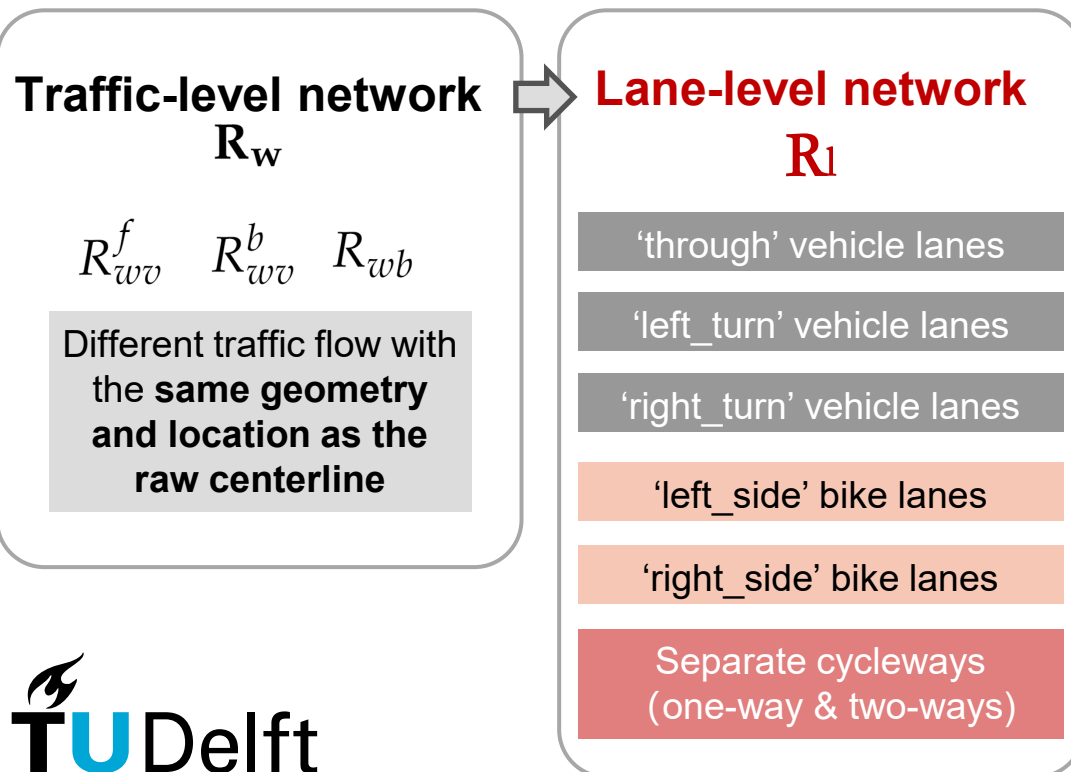


Methodology

Phase 2 — Linear Generation: Lane Network Reconstruction

Hierarchical generation: Lane geometry

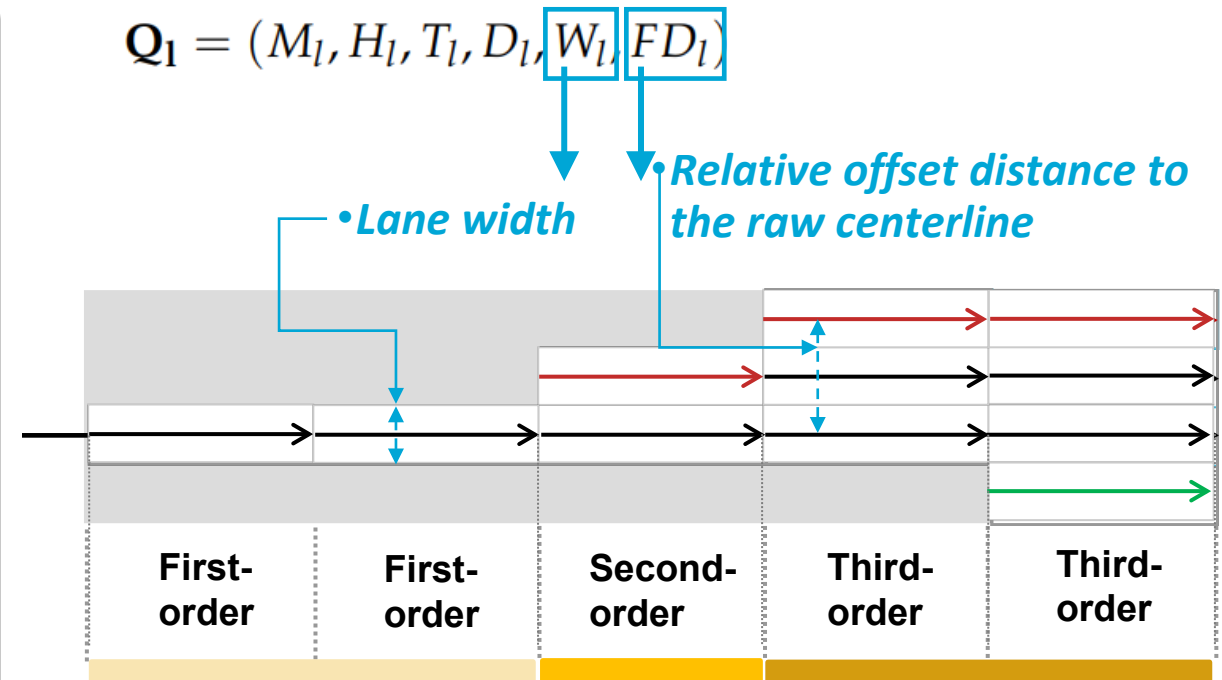
- Lane topology correctness:
'Lane connectivity'



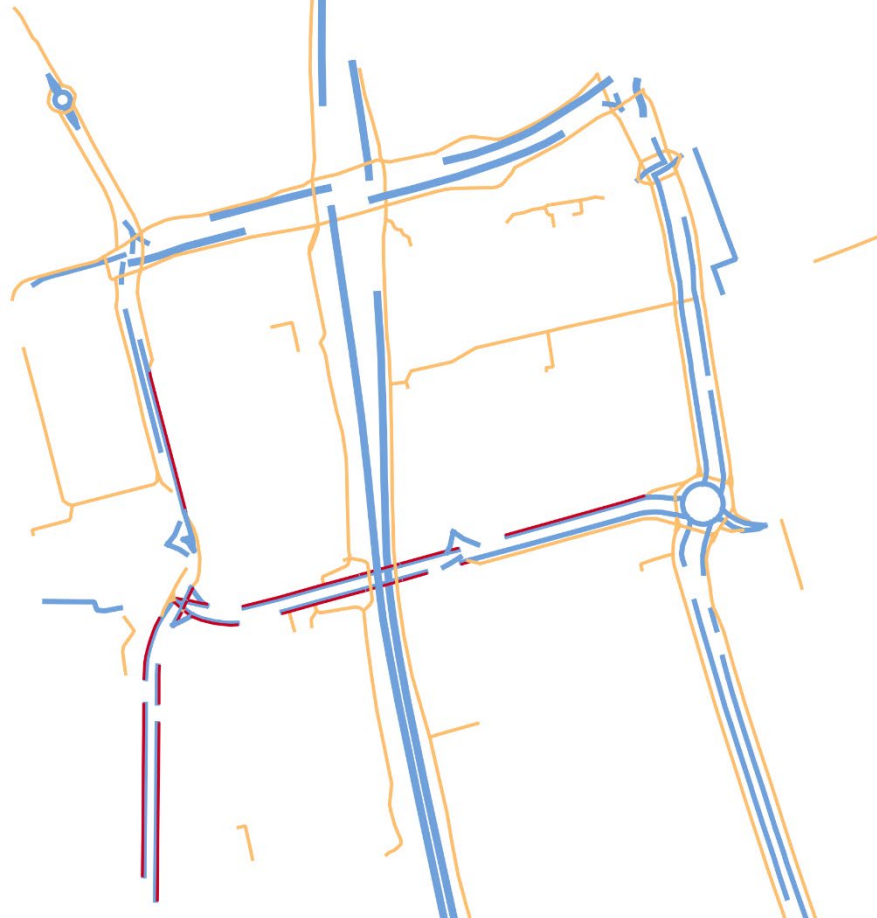
$$r_l = L_r = \{l_i\}_{i=1}^a$$

$$l = (I_r, I_w, V_l, VS_l, VE_l, Q_l)$$

$$Q_l = (M_l, H_l, T_l, D_l, W_l, FD_l)$$



First-order roads and generated lanes



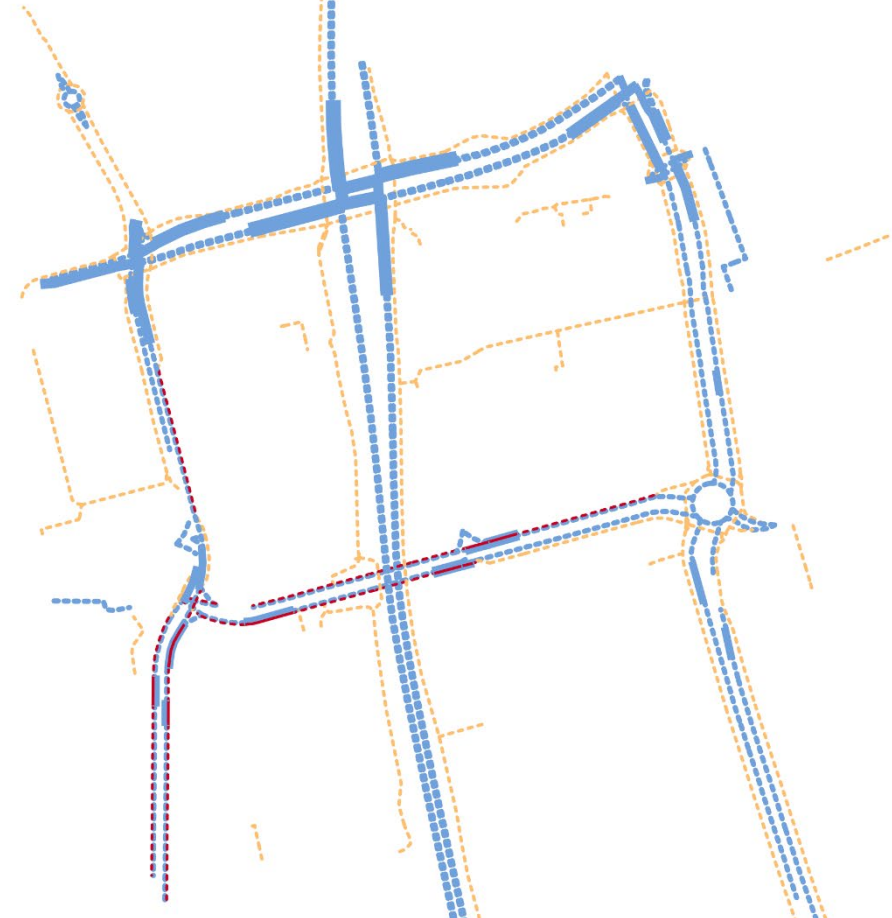
1st-order_road/lane_centerline

- cycleway
- road
- bike_lane

1st roads characteristics:

- The road centerline is regarded as the **shape of movement**;
- also the **center axis** of the road polygon

Second-order roads and generated lanes



2nd-order_road/lane_centerline

- road
- bike_lane

1st-order_road/lane_results

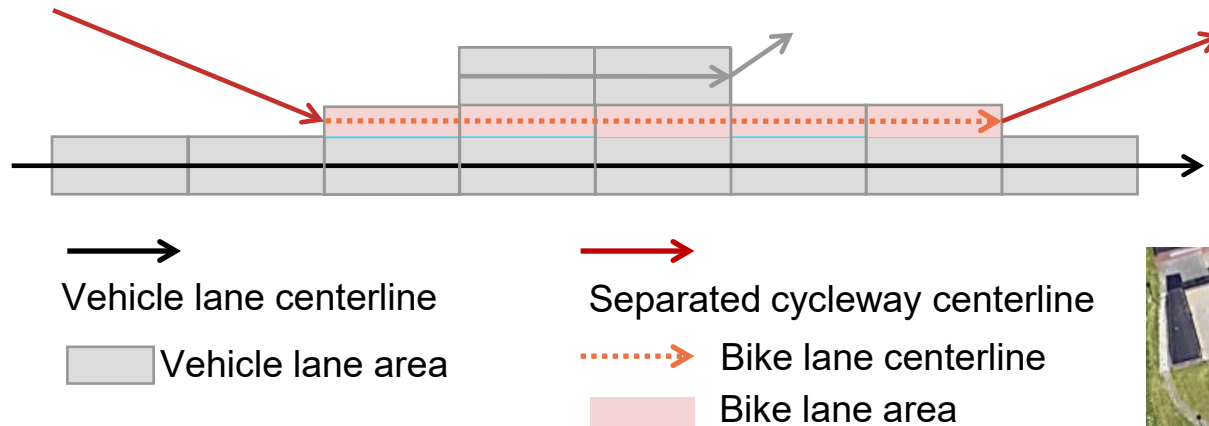
- road
- cycleway
- bike_lane

2nd roads characteristics:

- **Number of ‘through’ lanes** in 2nd roads equal to the “basic number of lanes”;
- contains **auxiliary lanes**

Methodology & Results

Phase 2 — **Linear Generation:** Lane Network Reconstruction



Lane-matching method in 2nd road step:

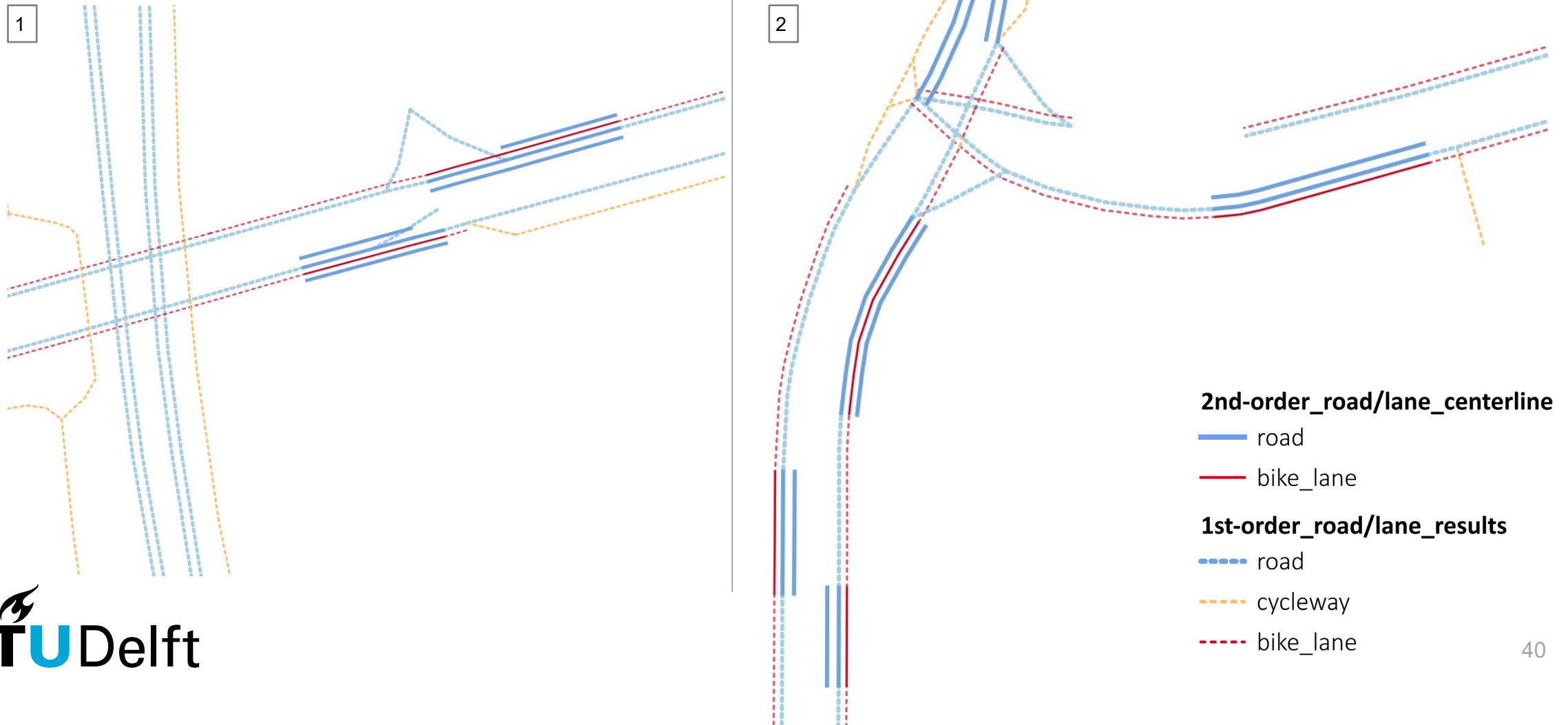
- If the adjacent 1st road has generated **“bike lane”**, then set the matching priority as: match the adjacent bike lanes firstly.
- Otherwise, match the ‘through’ lanes.



Methodology & Results

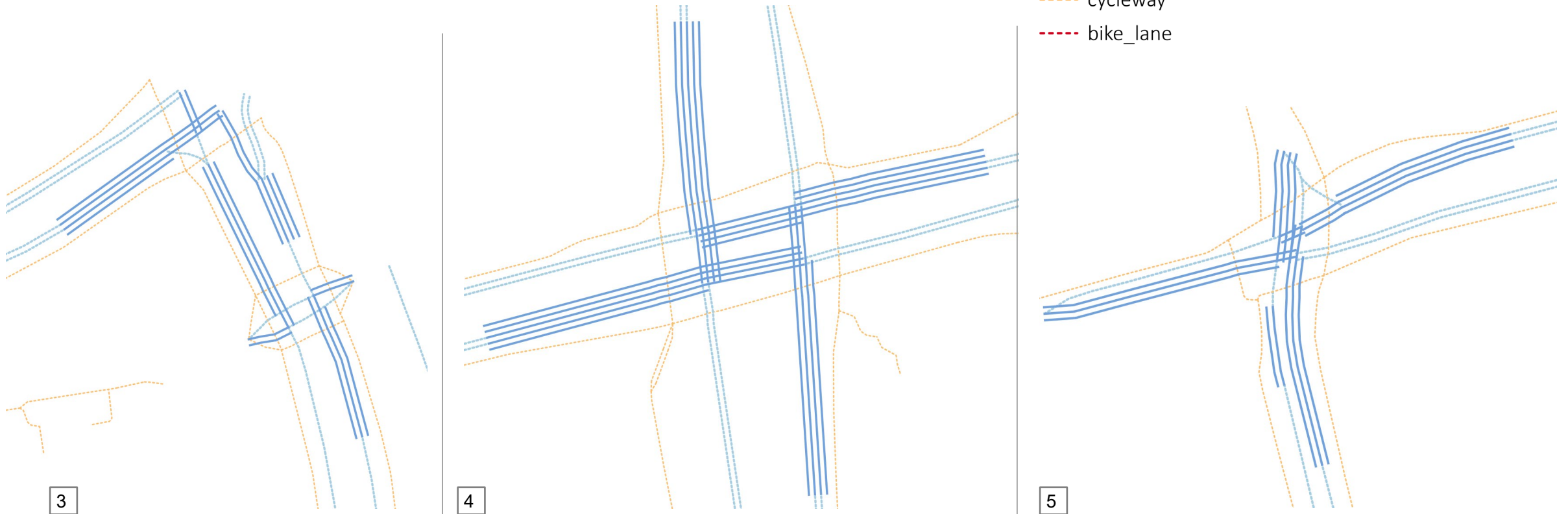
Phase 2 — **Linear Generation:** Lane Network Reconstruction

Second-order roads and generated lanes



Methodology & Results

Phase 2 — **Linear Generation:** Lane Network Reconstruction



Methodology & Results

Phase 2 — **Linear Generation:** Lane Network Reconstruction

Third-order roads and generated lanes

3rd roads characteristics:

- **Number of 'through' lanes** greater than "basic number of lanes";
- contains **auxiliary lanes** comparing to the adjacent 1st or 2nd roads.

3rd-order_road/lane_centerline

- road
- bike_lanes

2nd-order_road/lane_results

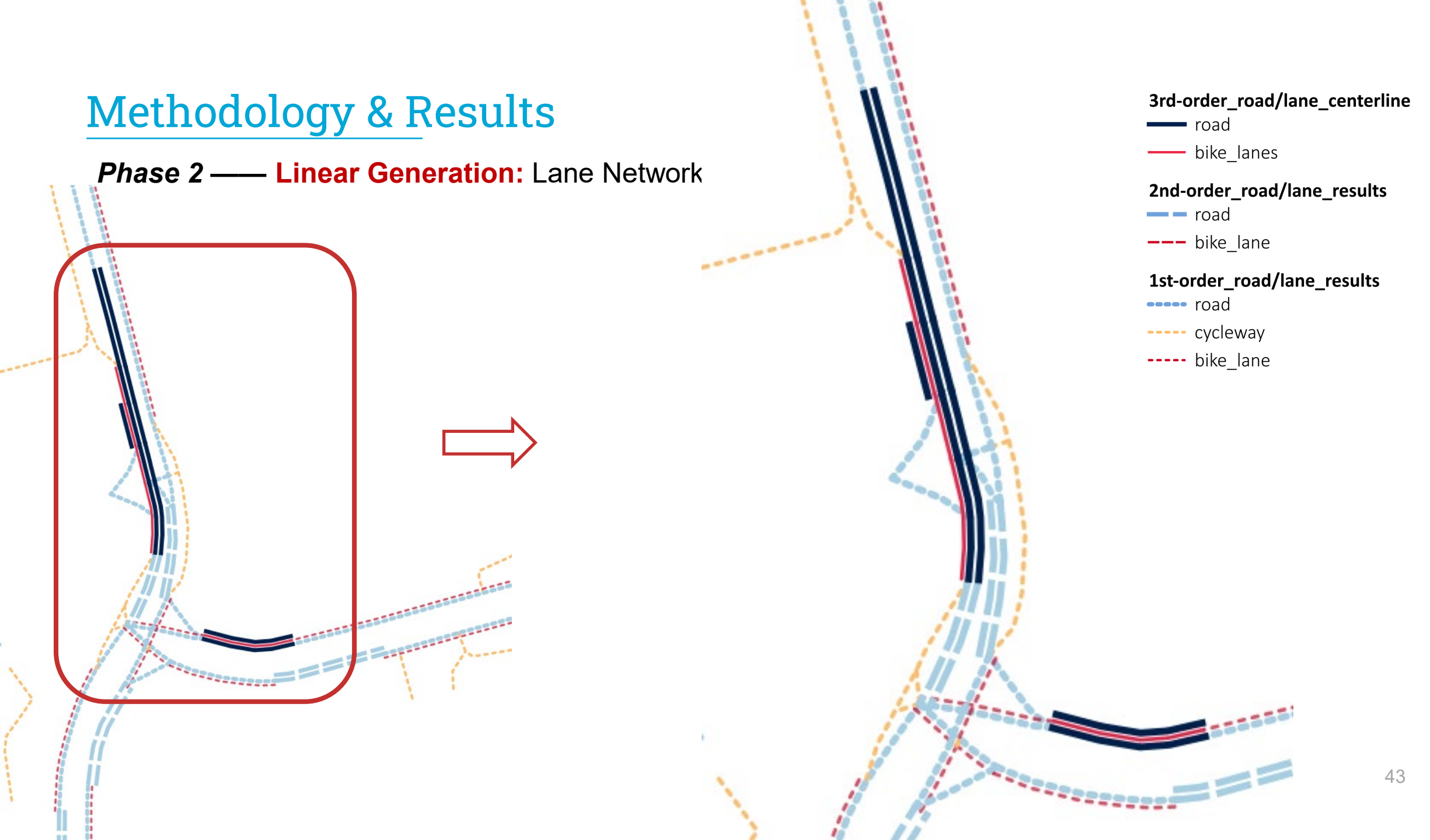
- - road
- - bike_lane

1st-order_road/lane_results

- ... road
- ... cycleway
- ... bike_lane

Methodology & Results

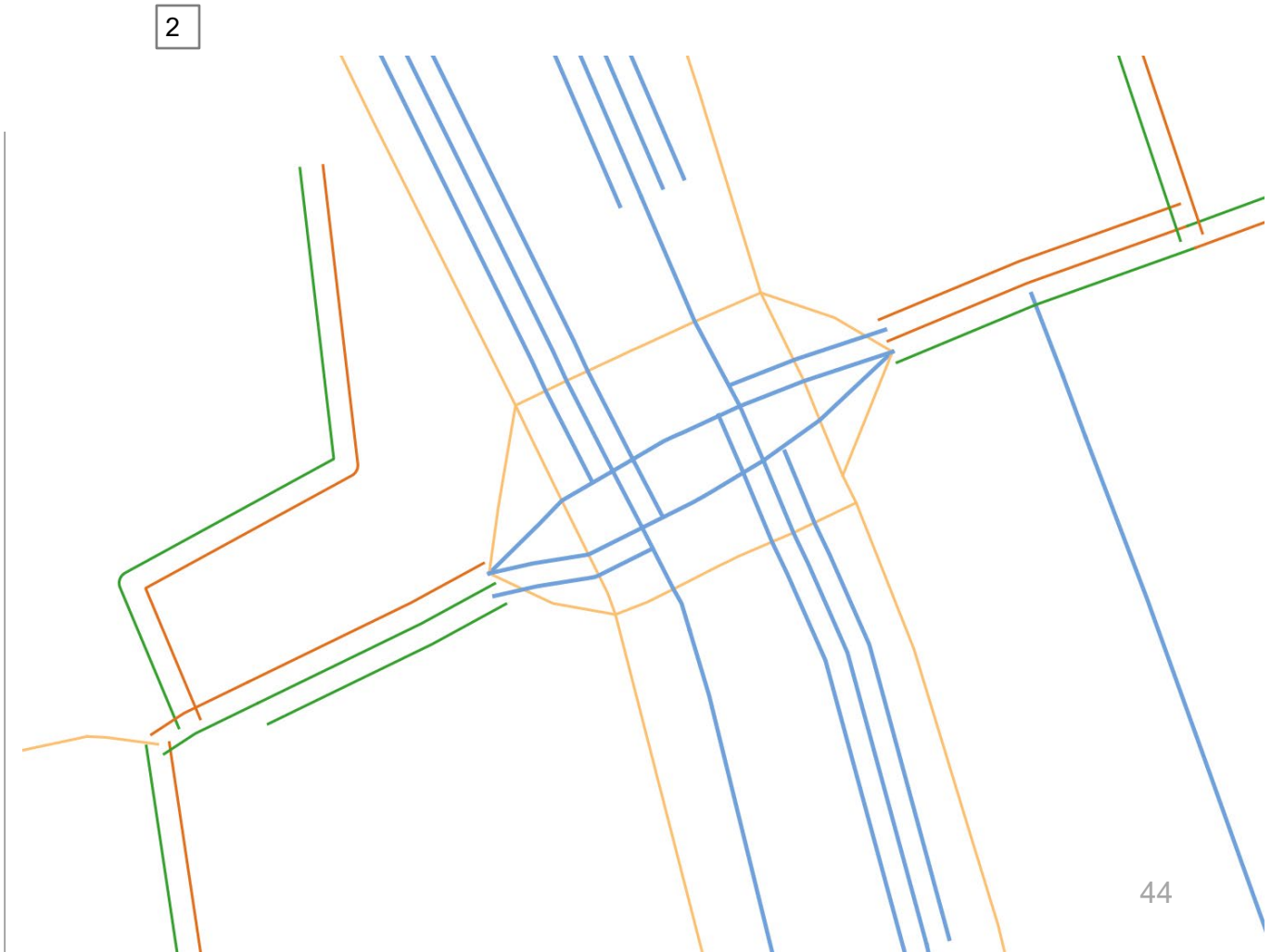
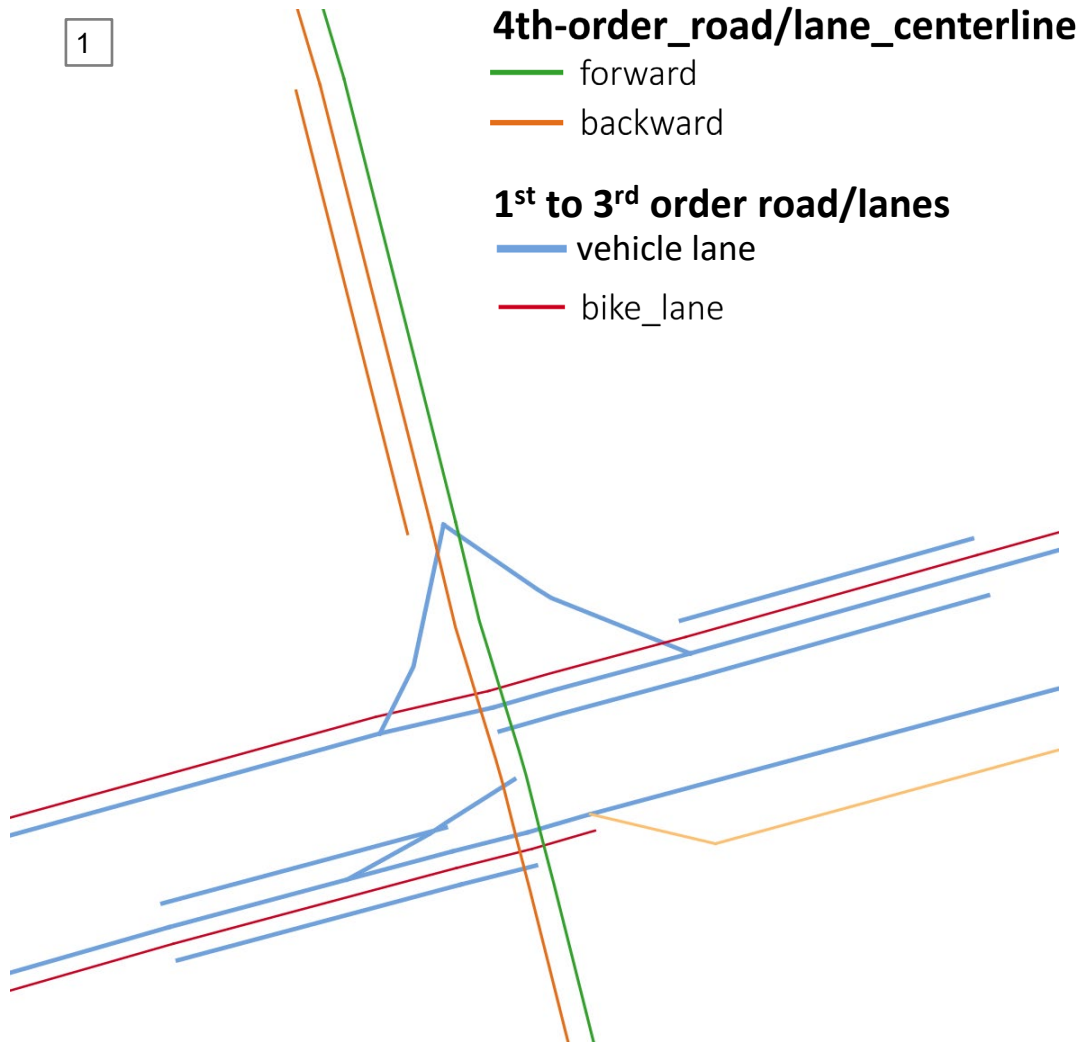
Phase 2 — Linear Generation: Lane Network



Methodology & Results

Phase 2 — **Linear Generation:** Lane Network Reconstruction

Fourth-order roads and generated lanes



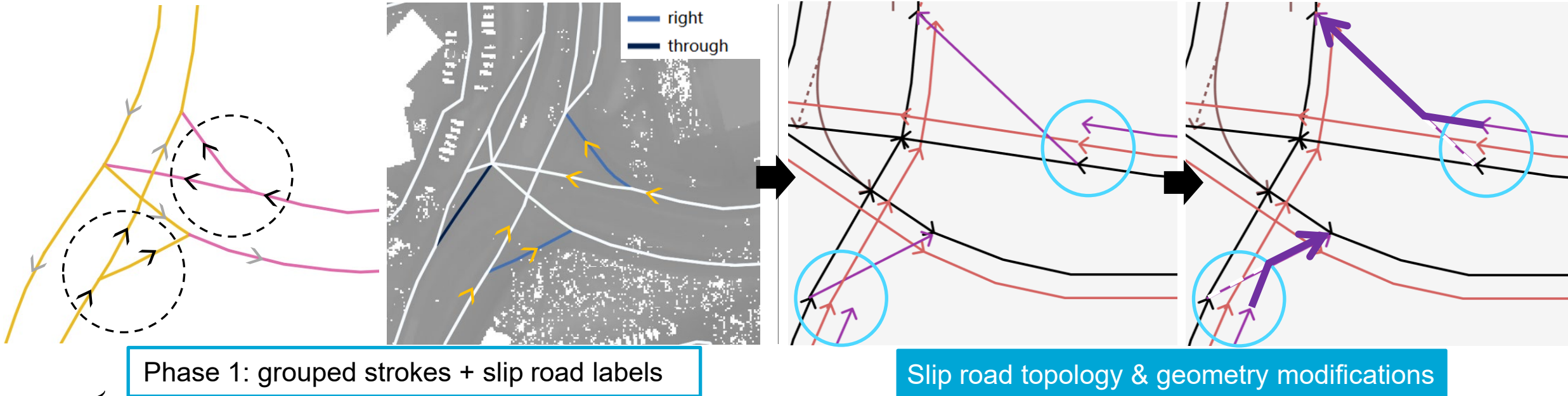
Methodology & Results

Phase 2 — **Linear Generation:** Lane Network Reconstruction

Lane-level Network Connectivity Modification

- Lane topology correctness: '**Lane connectivity**'

Slip roads: • 'slip road' label + Global and local adjacency



Methodology & Results

Phase 2 — **Linear Generation:** Lane Network Reconstruction

Lane-level Network Connectivity Modification

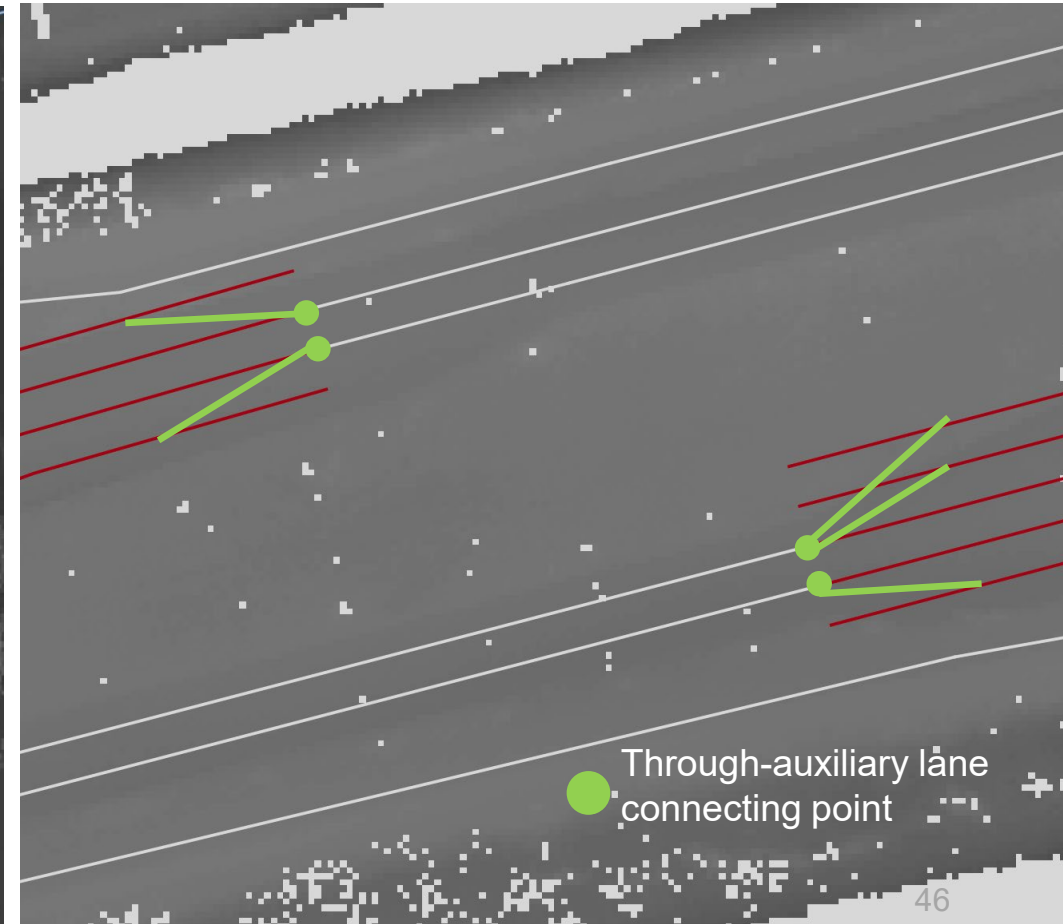
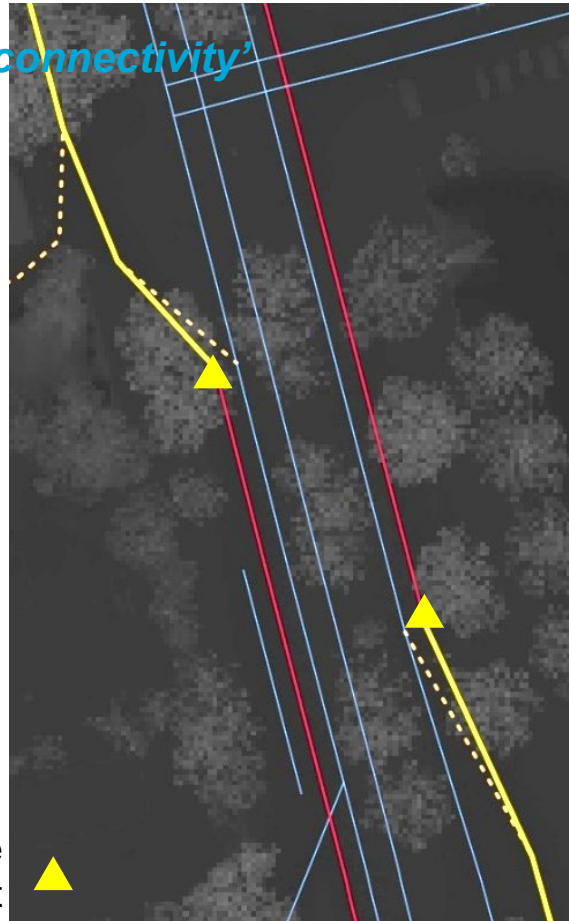
- Lane topology correctness: '**Lane connectivity**'

Separated cycleways:

- 'highway=cycleway/bike lane' label
- Global and local adjacency

Auxiliary lanes:

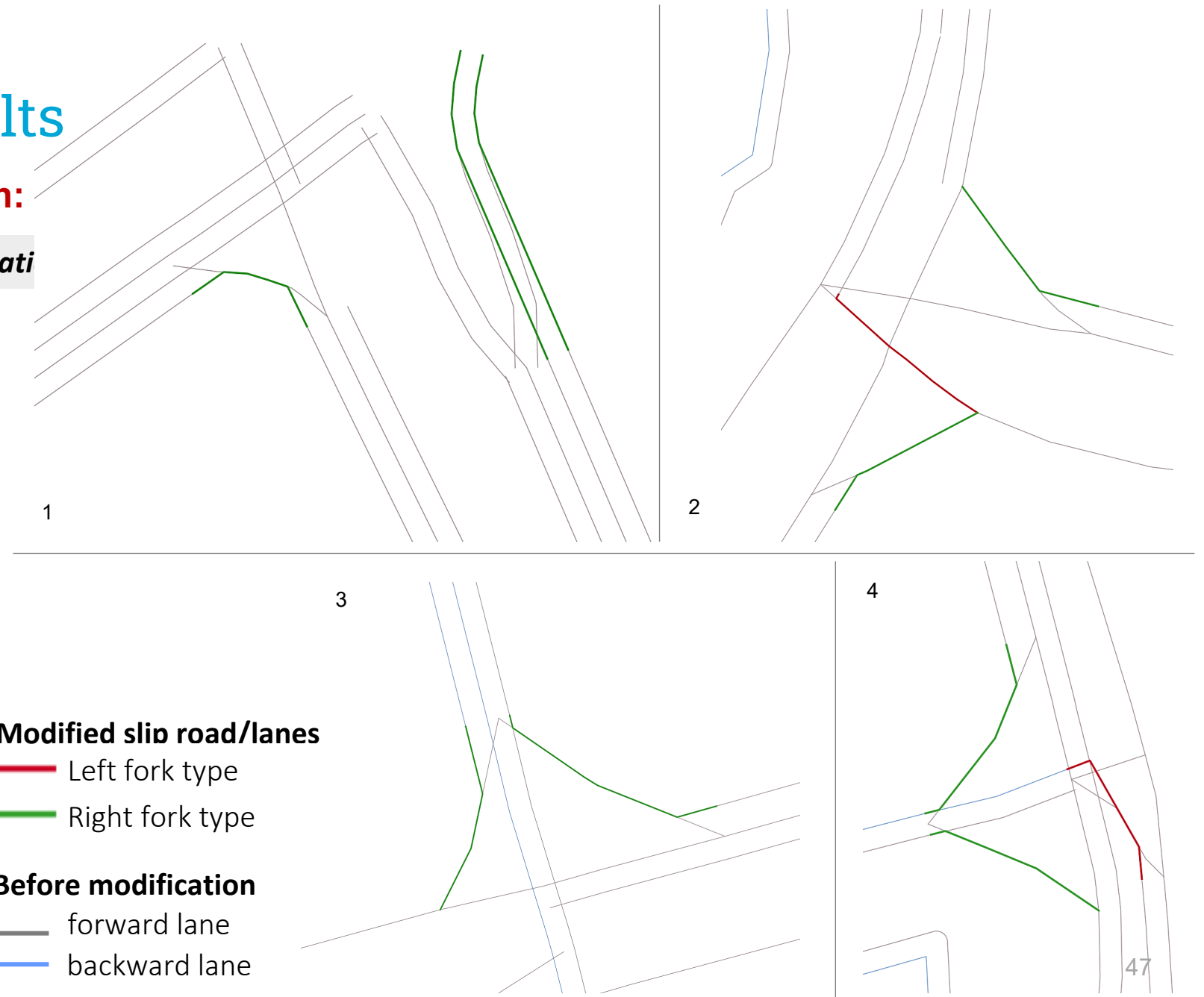
- 'turn' label
- Local adjacency



Methodology & Results

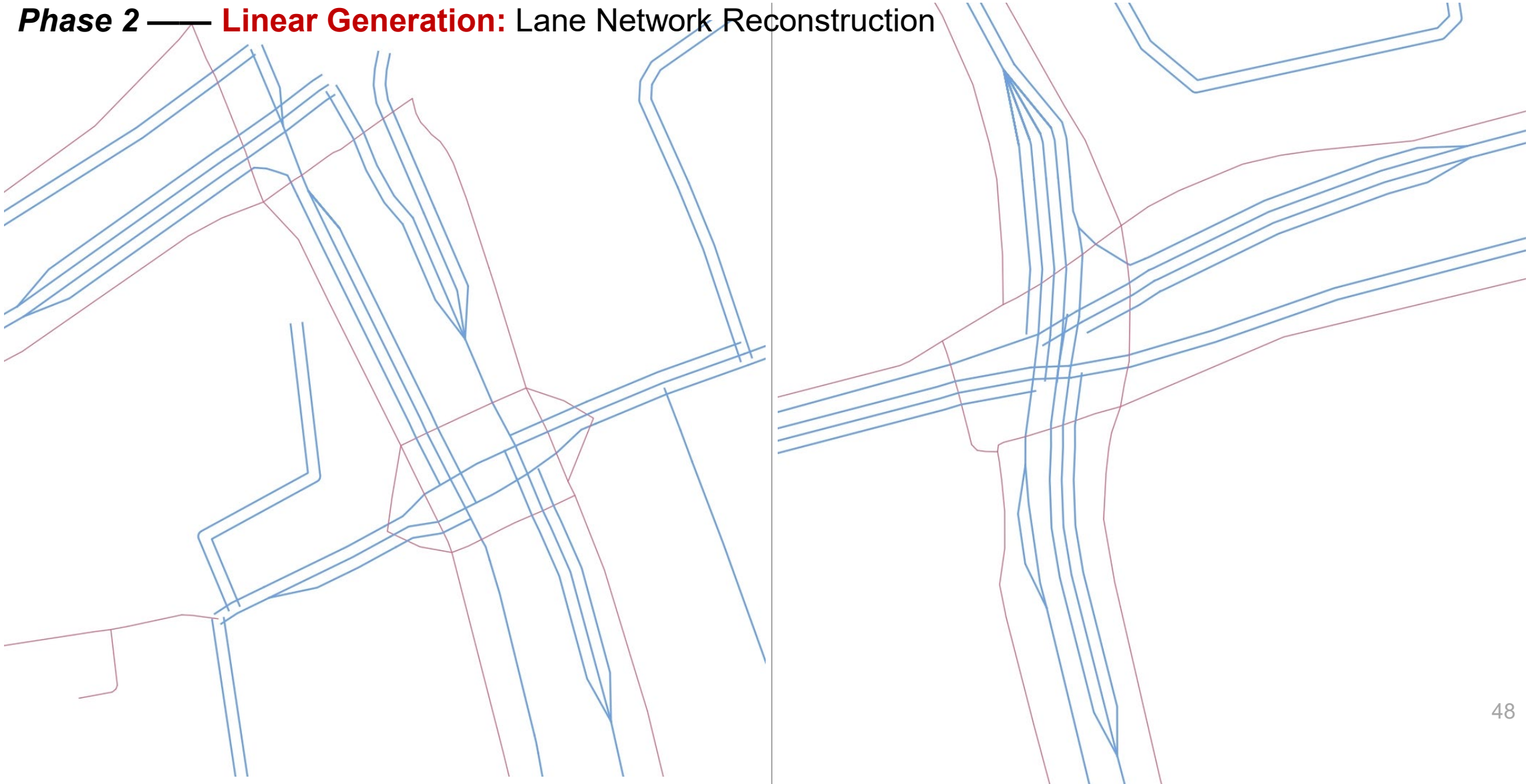
Phase 2 — Linear Generation:

Lane-level Network Connectivity Modification



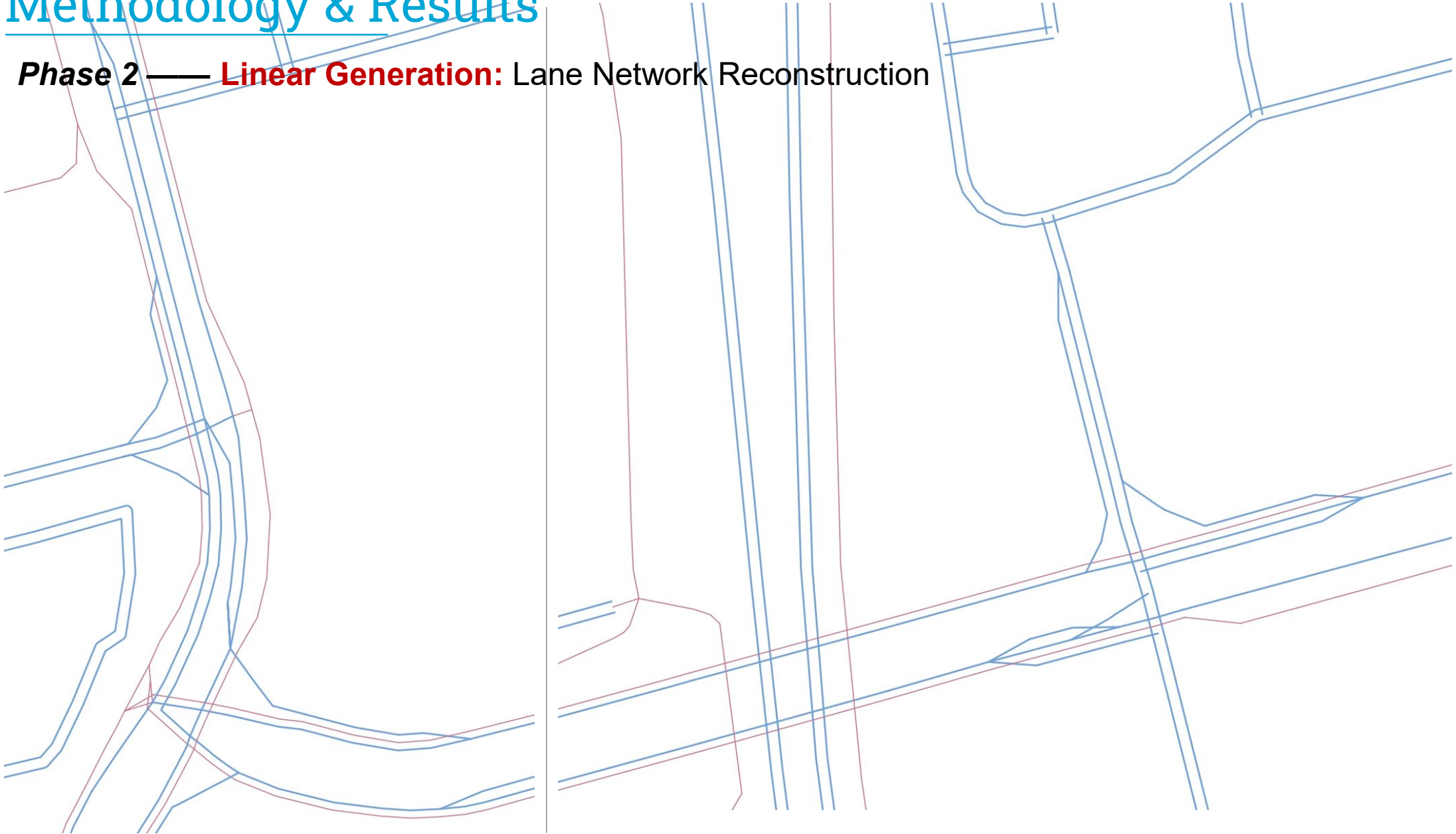
Methodology & Results

Phase 2 — **Linear Generation:** Lane Network Reconstruction



Methodology & Results

Phase 2 — **Linear Generation:** Lane Network Reconstruction



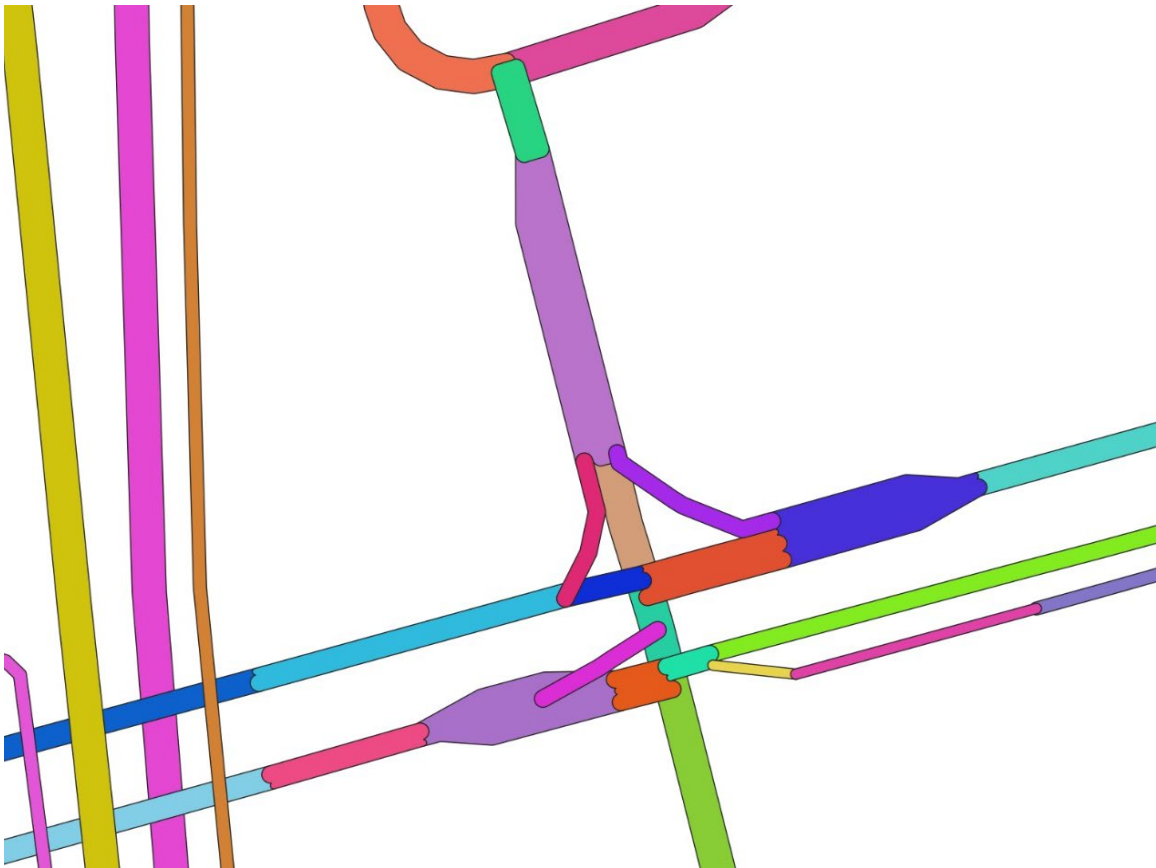
Methodology & Results

How can we generate 2D road polygons as **errorless** geometries to accurately capture **road shape details**, including changes in road width and **intersection variations**?

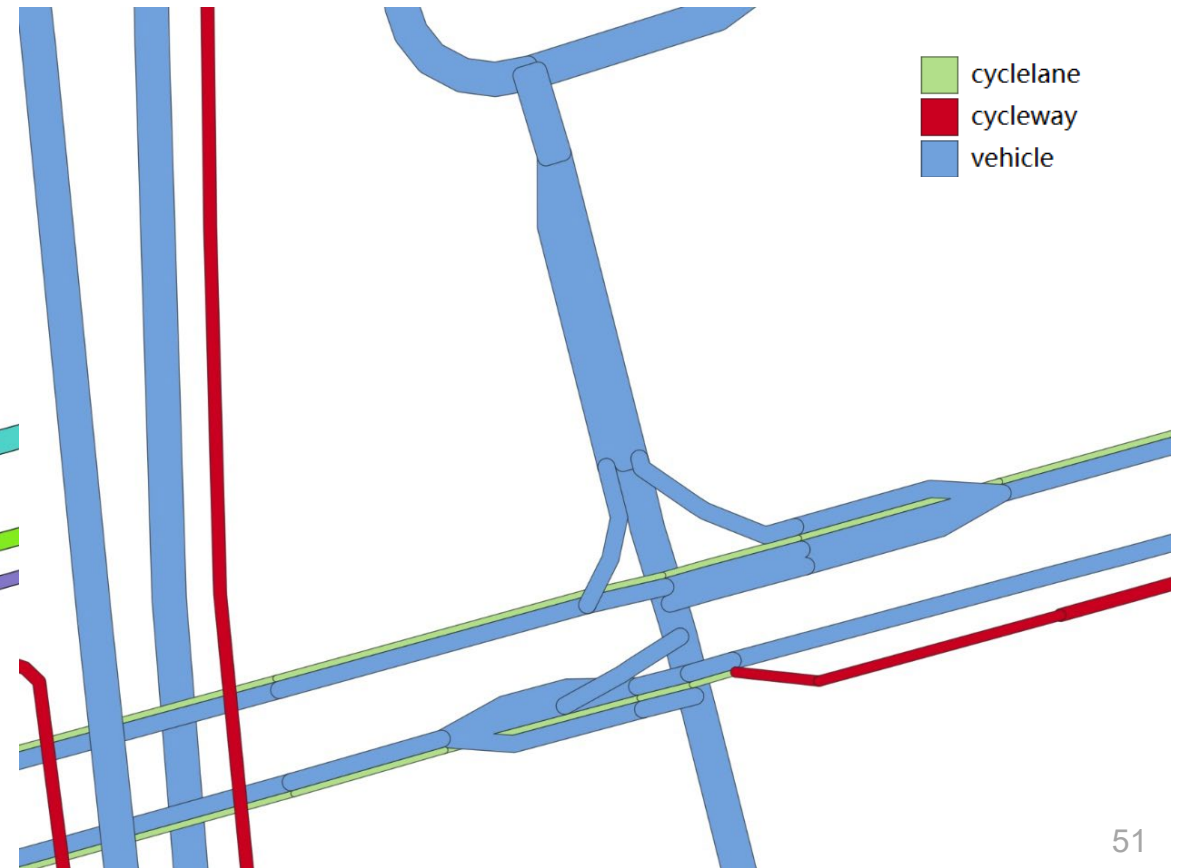
Phase 3 — **Areal Generation**: 2D Road Polygon Reconstruction

➡ **Subquestion 4**

- *Road polygon merged by road indices*



- *Road polygon merged by traffic mode types*



Methodology & Results

Phase 3 — **Areal Generation:** 2D Road Polygon Reconstruction

Node degree-based method for road polygon post-processing

- errorless geometries**

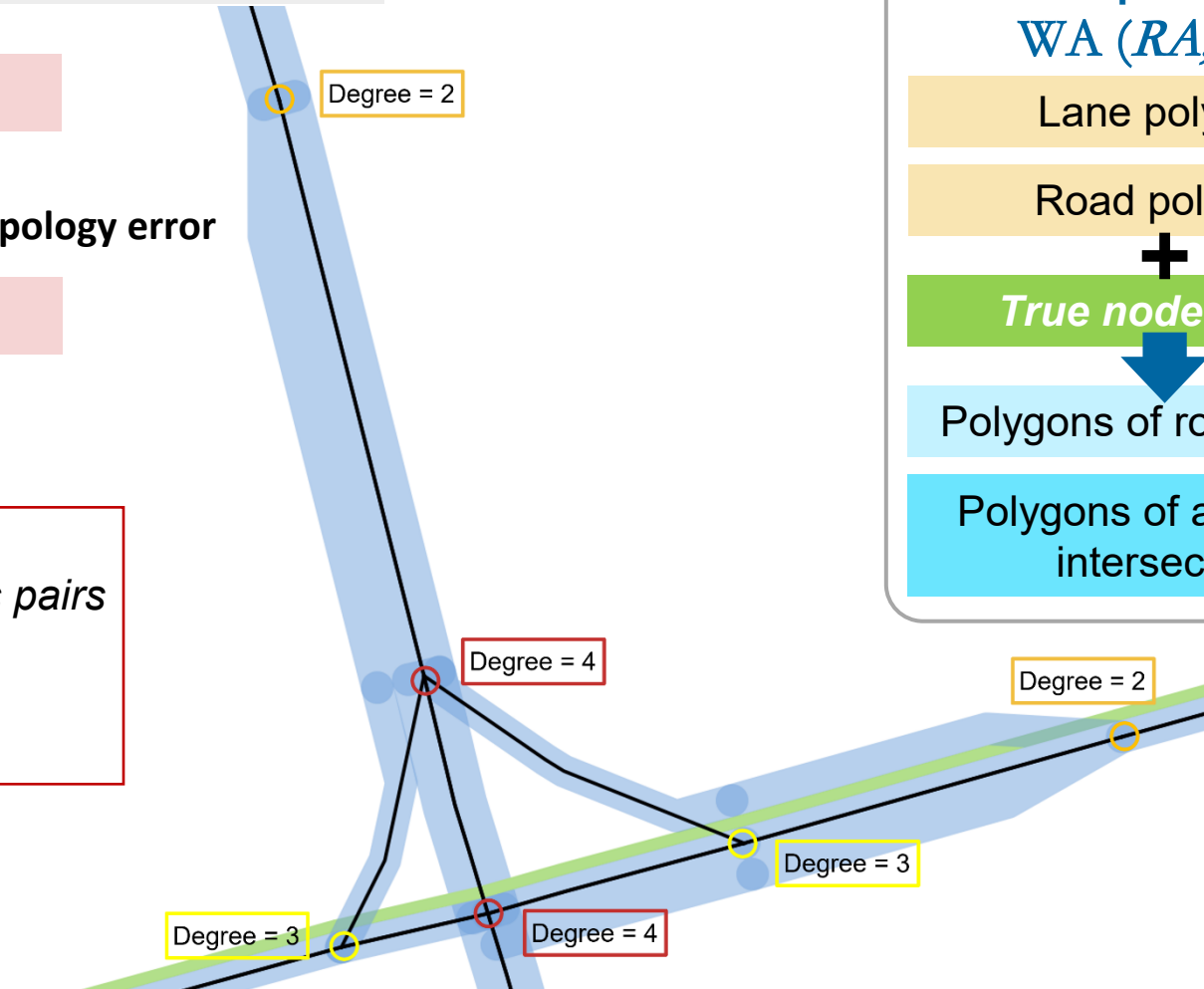
→ Road polygons without overlapping;
Trim the error parts caused by lane topology error

- intersection variations**

→ Identify all the intersections;
Generate intersection polygons.

If “true node degree is 2” →
trim the overlapping road polygon as pairs

Elif “true node degree > 2” →
generate intersection polygon



→ Subquestion 4

Areal representation

$WA(RA, CA)$

Lane polygons

Road polygons

+

True node degree



Polygons of road ribbons

Polygons of all types of intersections

Methodology & Results

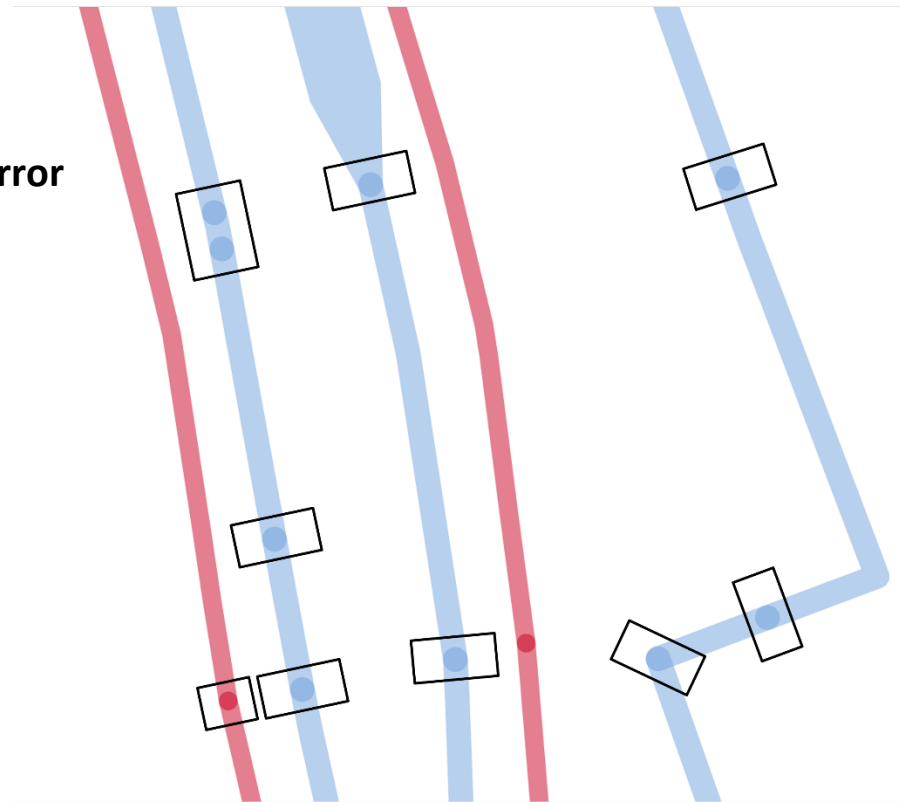
Phase 3 — **Areal Generation:** 2D Road Polygon Reconstruction

Node degree-based method for road polygon post-processing

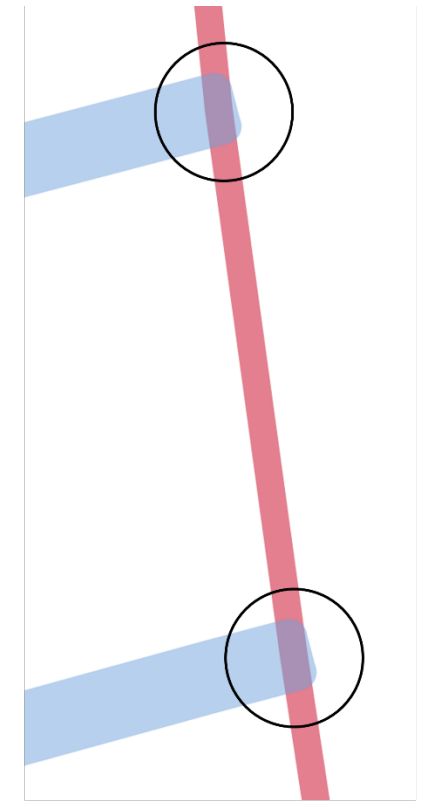
- **errorless** geometries

→ Road polygons without overlapping;
Trim the error parts caused by lane topology error

If “true node degree is 2” →
trim the overlapping road polygon as pairs



a) Node degree = 2:
remove overlapping part from adjacent roads



b) Node degree greater than 2,
and road width not equal.⁵³
remove protruding part from road

Methodology & Results

Phase 3 — **Areal Generation:** 2D Road Polygon Reconstruction

Node degree-based method for road polygon post-processing

- **errorless** geometries



Methodology & Results

Phase 3 — **Areal Generation:** 2D Road Polygon Reconstruction

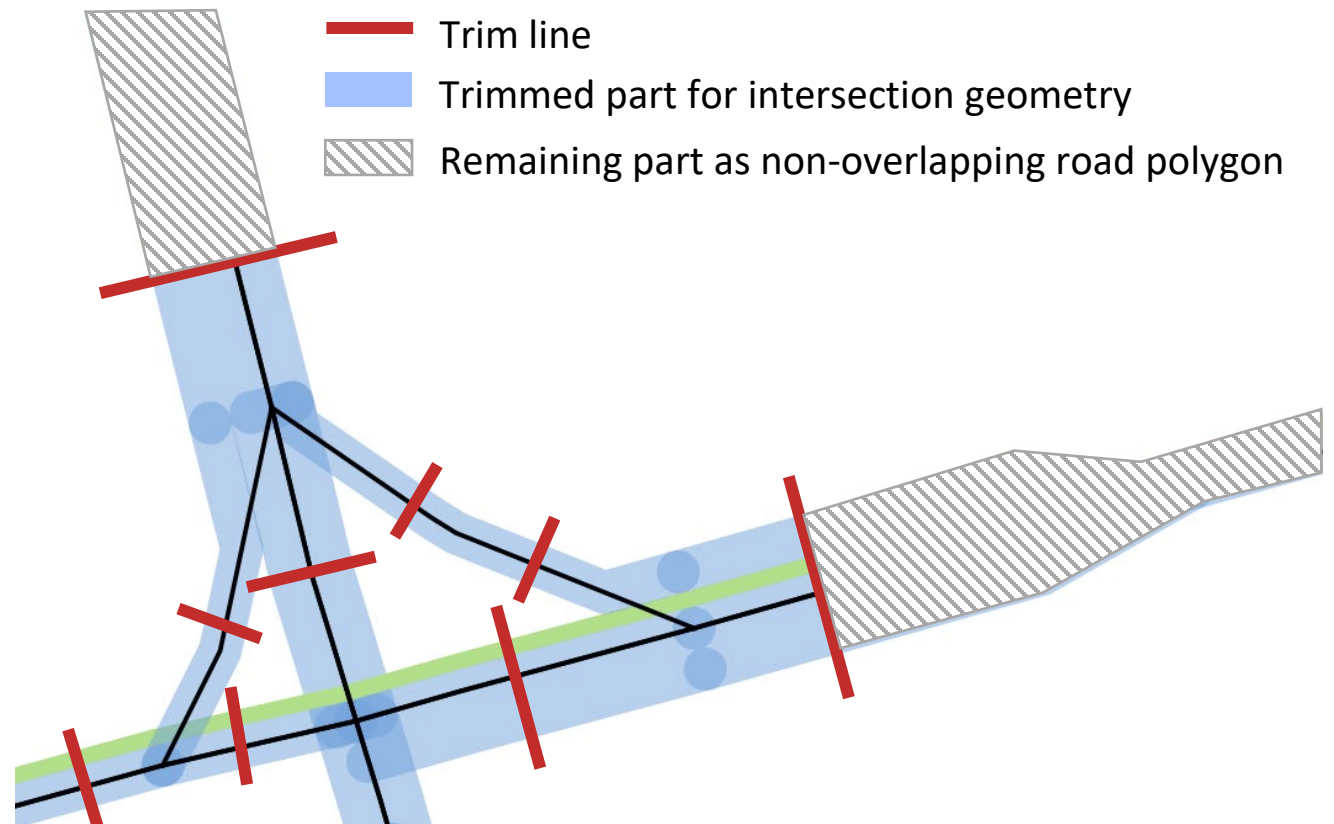
Node degree-based method for road polygon post-processing

• **intersection variations**

Identify all the intersections;

→ Generate intersection polygons using a “one-size-fits-all” solution.

Elif “true node degree > 2” →
generate intersection polygon



Methodology & Results

Phase 3 — **Areal Generation:** 2D Road Polygon Reconstruction

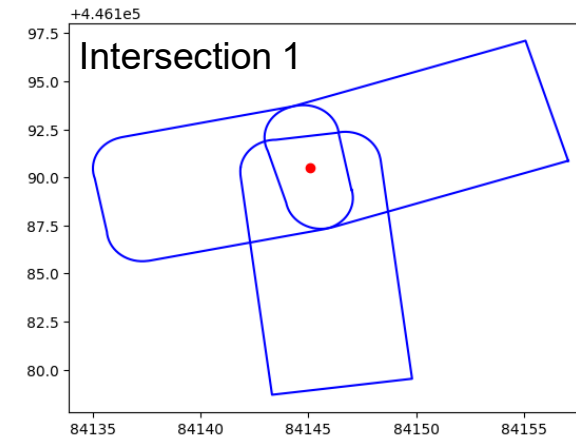
Node degree-based method for road polygon post-processing

• **intersection variations**

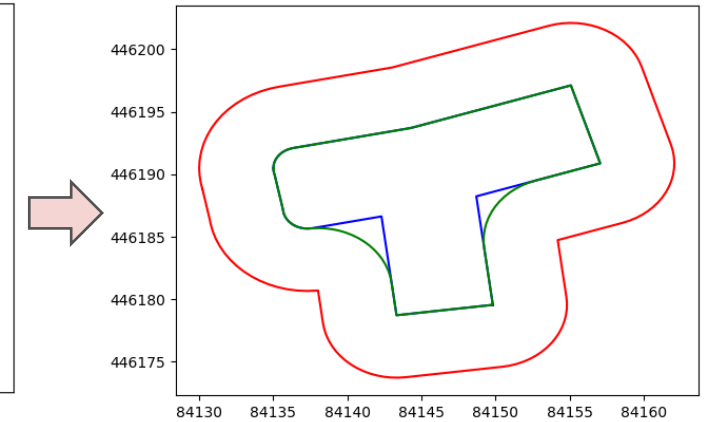
Identify all the intersections;

→ Generate intersection polygons using a “one-size-fits-all” solution.

Elif “true node degree > 2” →
generate intersection polygon



□ Trimmed converged road polygons

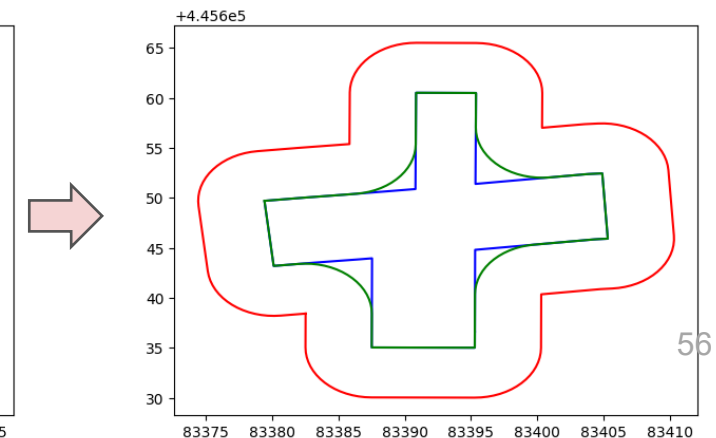
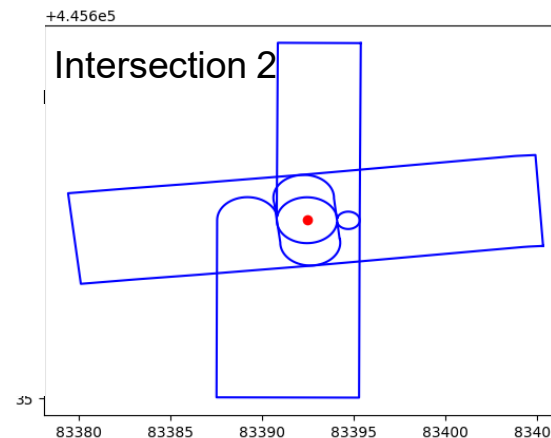


□ Merged converged road polygons

□ Polygon after ‘positive’ buffer

□ Polygon after ‘negative’ buffer

----- **Final result**



Methodology & Results

Phase 3 — **Areal Generation:** 2D Road Polygon Reconstruction

Node degree-based method for road polygon post-processing

- **intersection variations**

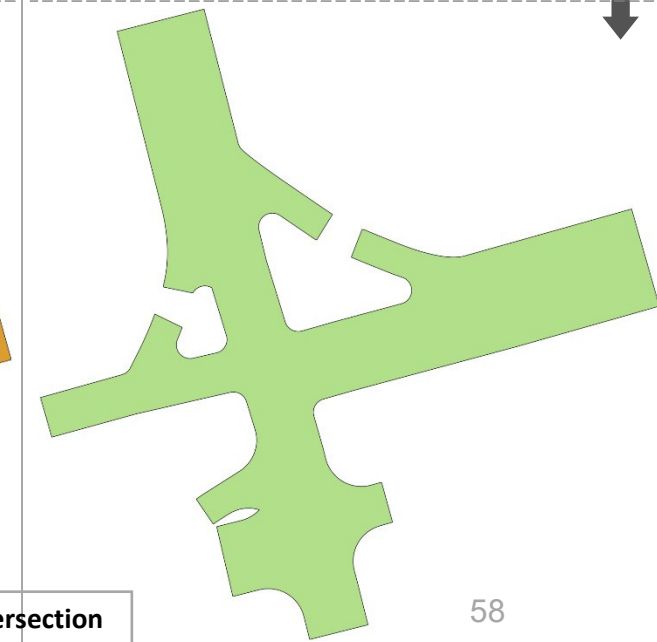
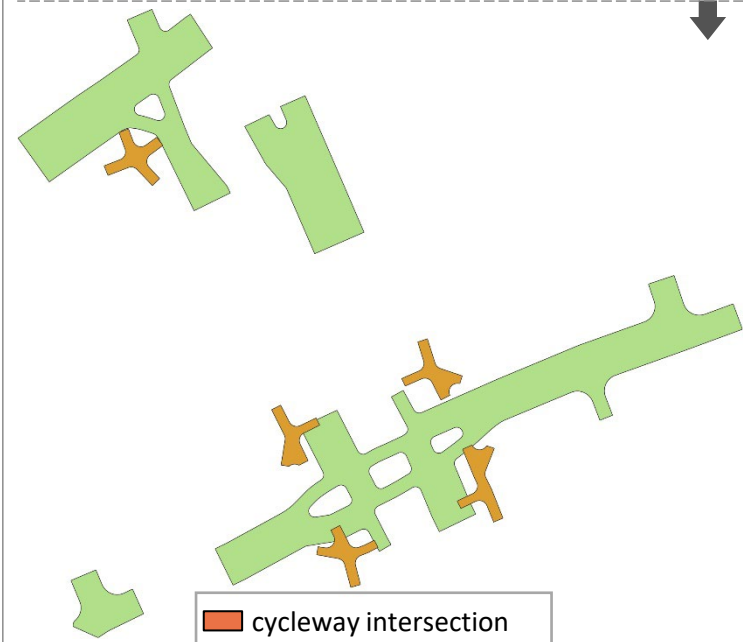
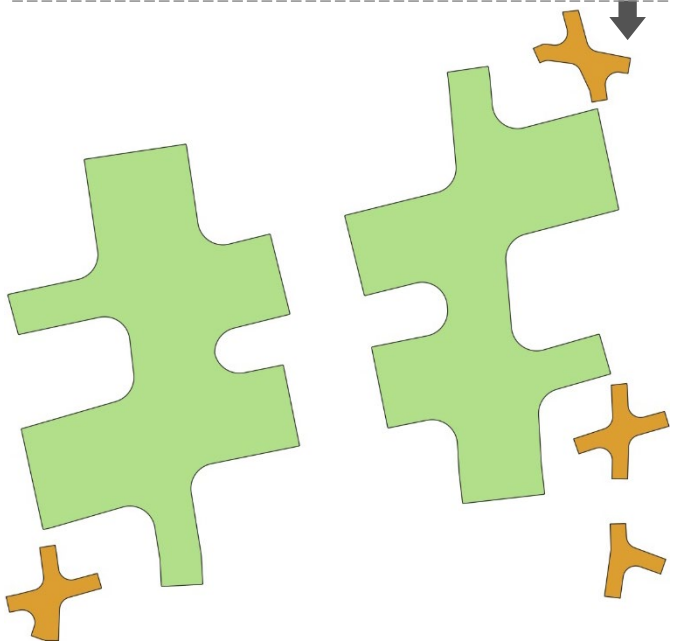
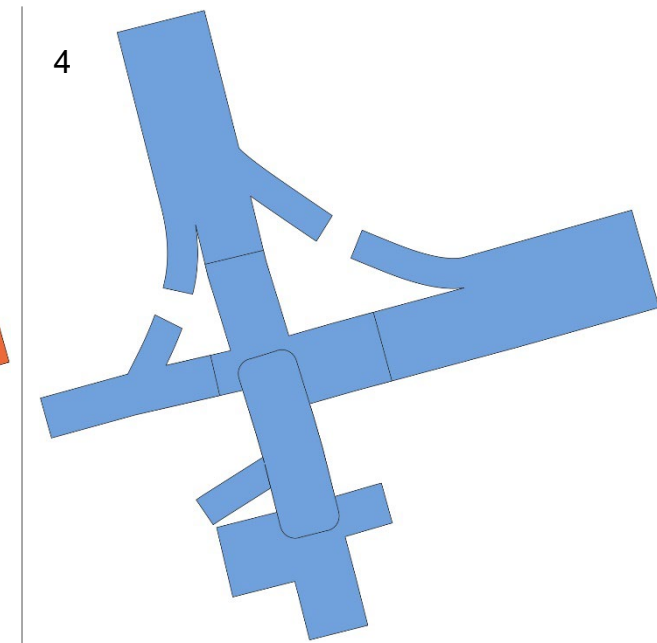
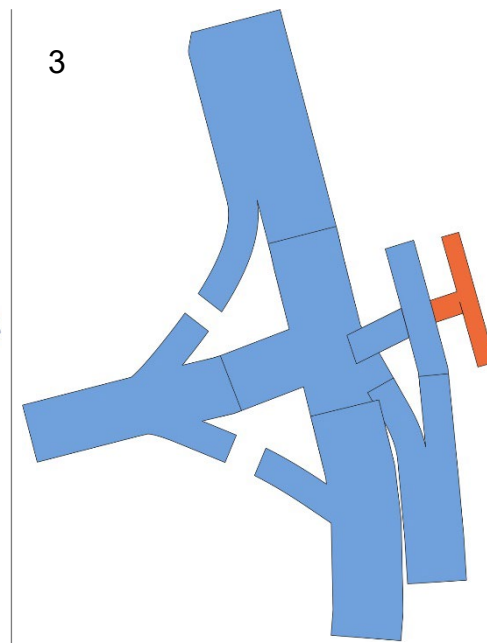
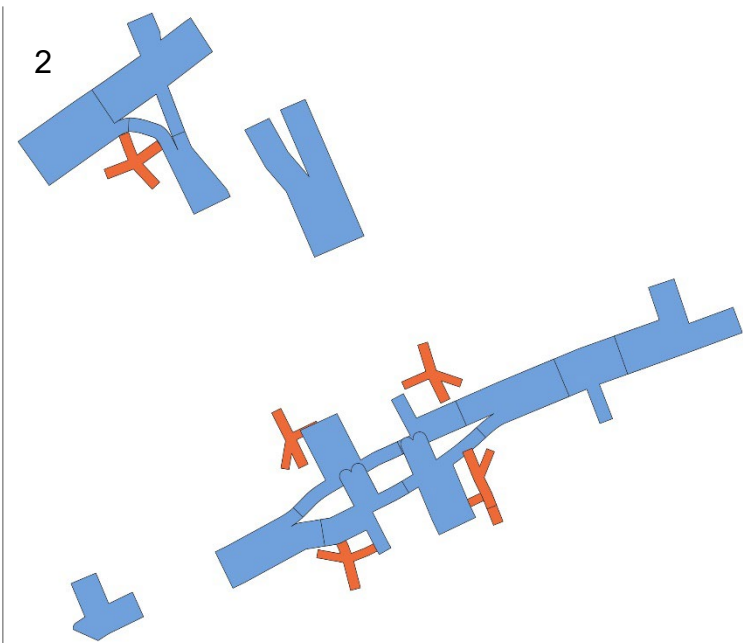
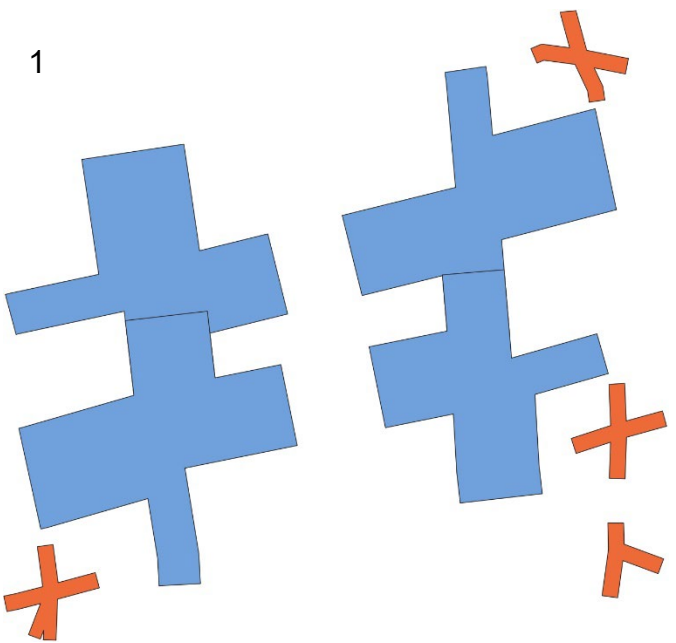
According to their shapes, the common types are:

- T/Y intersection;
- cross intersection;
- three ramps intersection;
- asterisk-shape intersection.

Finally, after ‘double-merge’ implementation, the number of intersections is **108**, of which **66** motorized intersections, **2** roundabouts, and **40** cycleway intersections.

Semantic information is preserved in areal representation





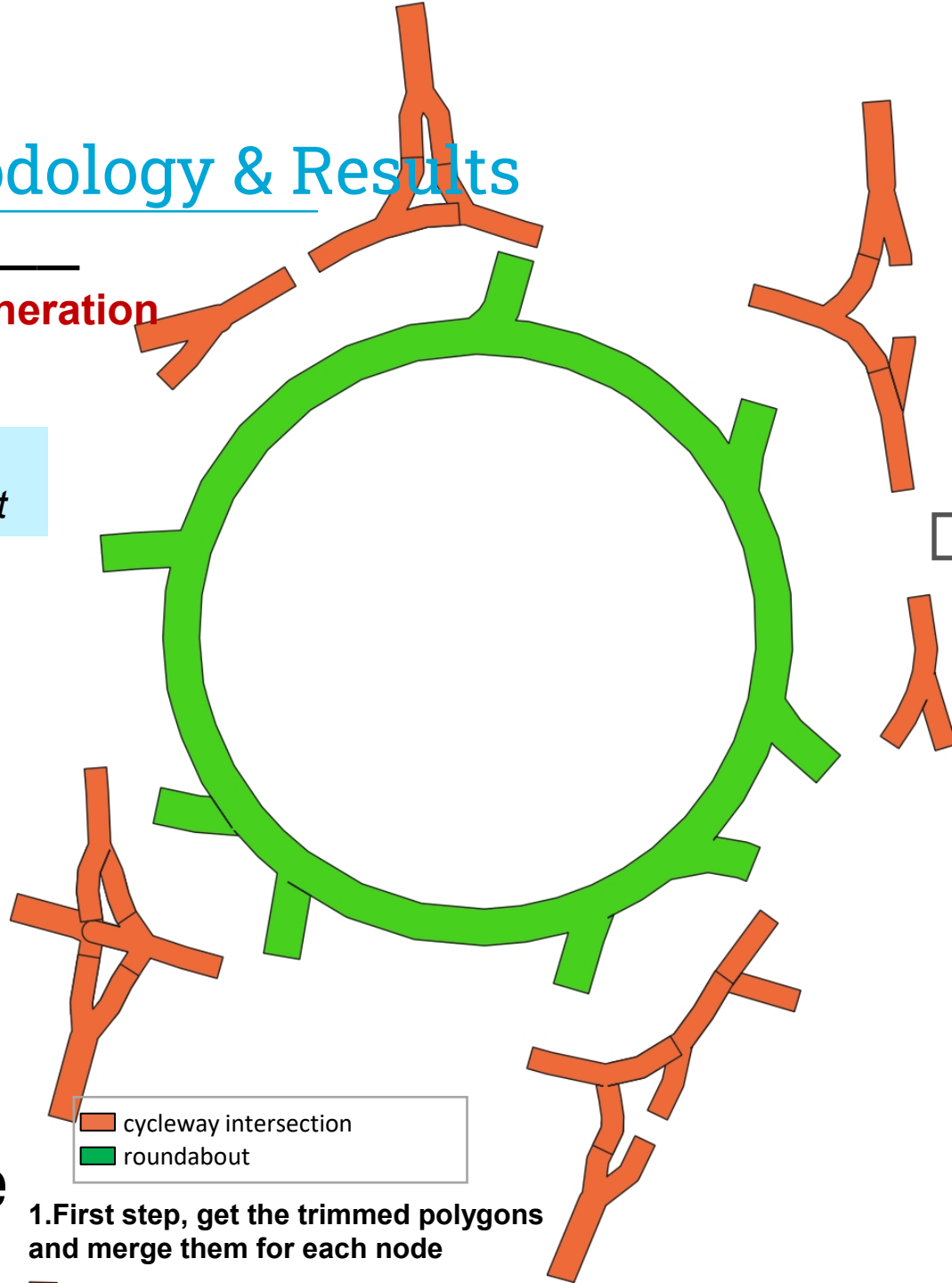
 cycleway intersection
 motorized intersection

 Buffer and Merged cycleway intersection
 Buffer and Merged motorized intersection

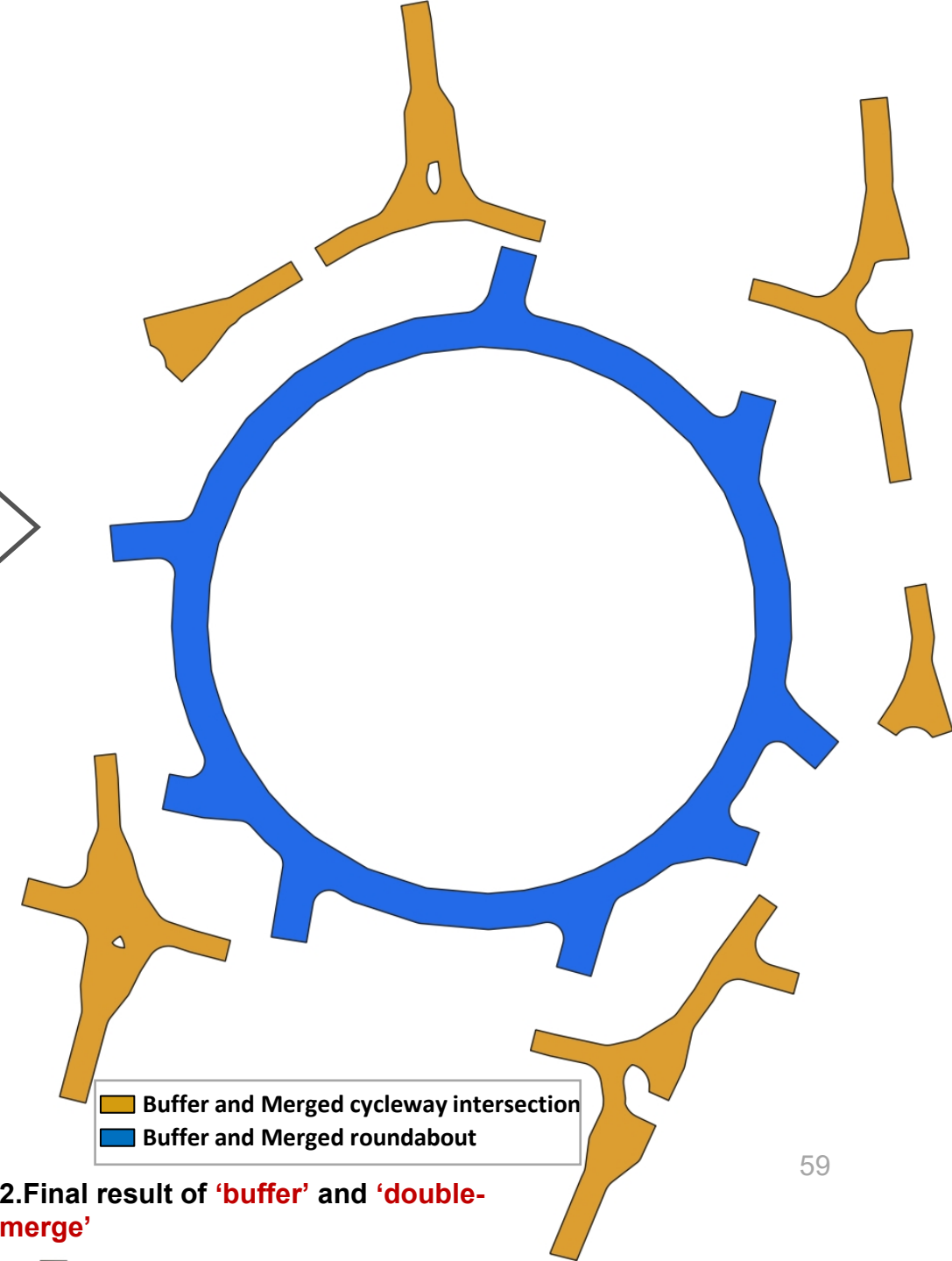
Methodology & Results

Phase 3 — Areal Generation

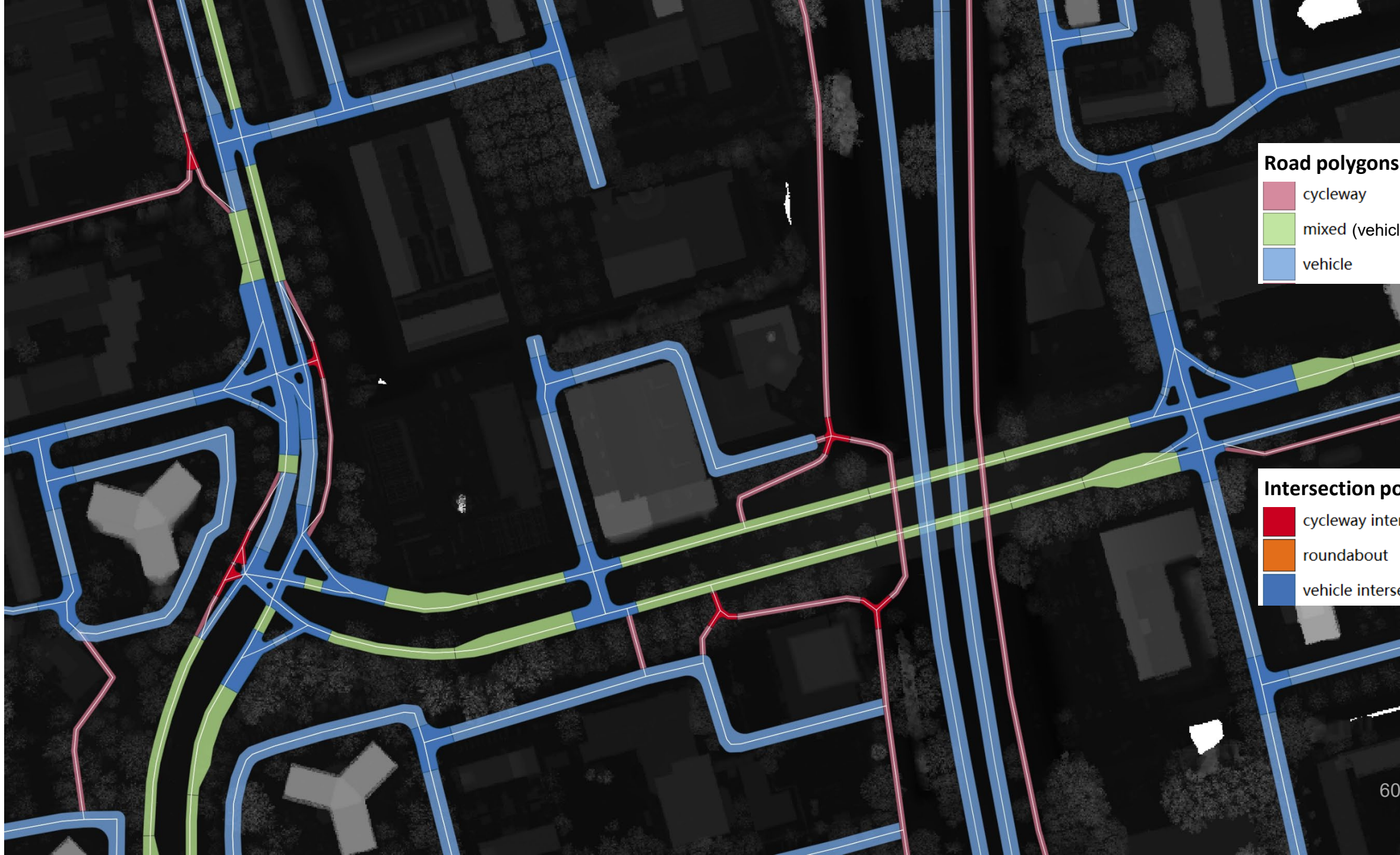
*junction =
roundabout*



1. First step, get the trimmed polygons and merge them for each node



2. Final result of 'buffer' and 'double-merge'

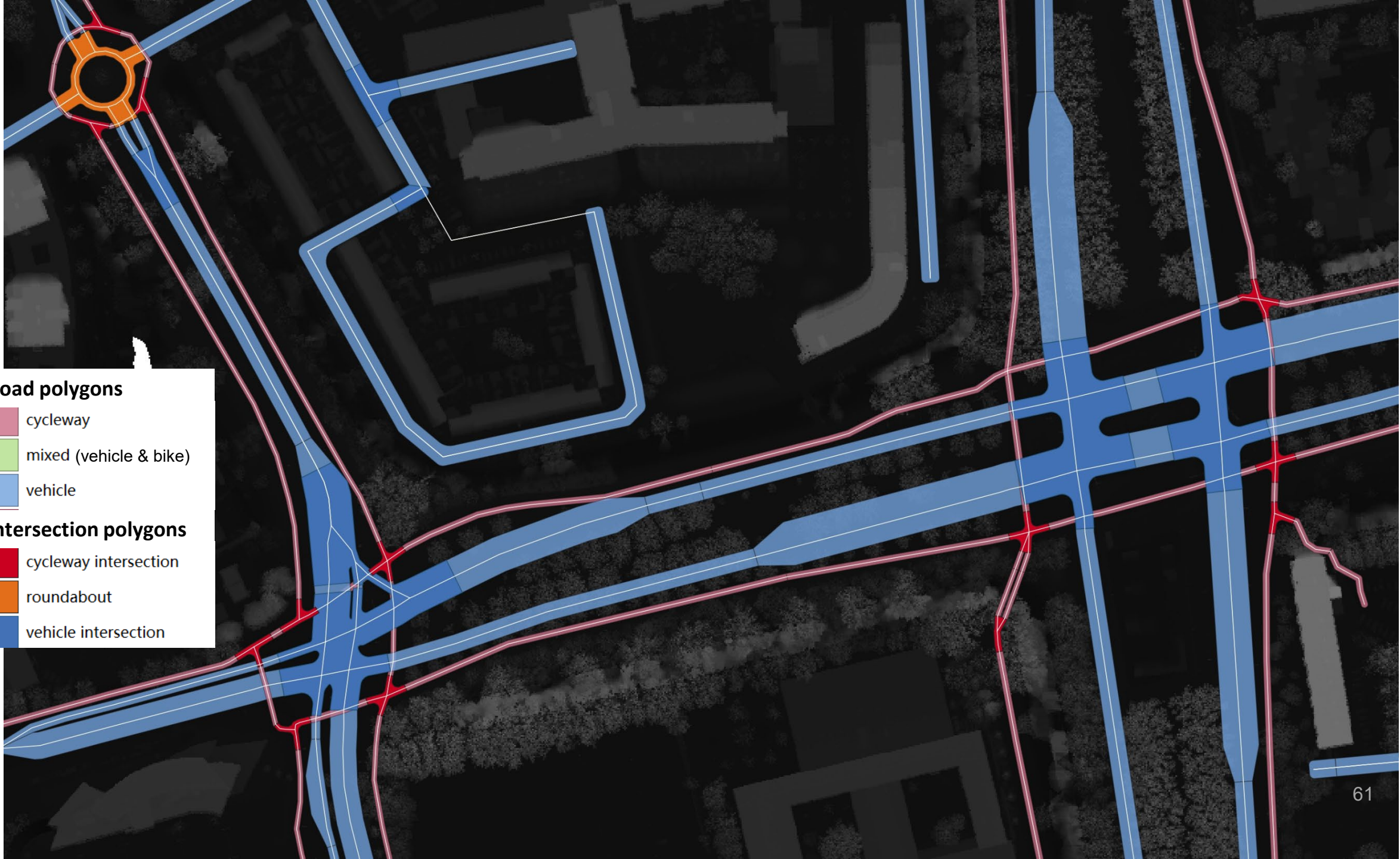


Road polygons

- cycleway
- mixed (vehicle & bike)
- vehicle

Intersection polygons

- cycleway intersection
- roundabout
- vehicle intersection



Road polygons

-  cycleway
-  mixed (vehicle & bike)
-  vehicle

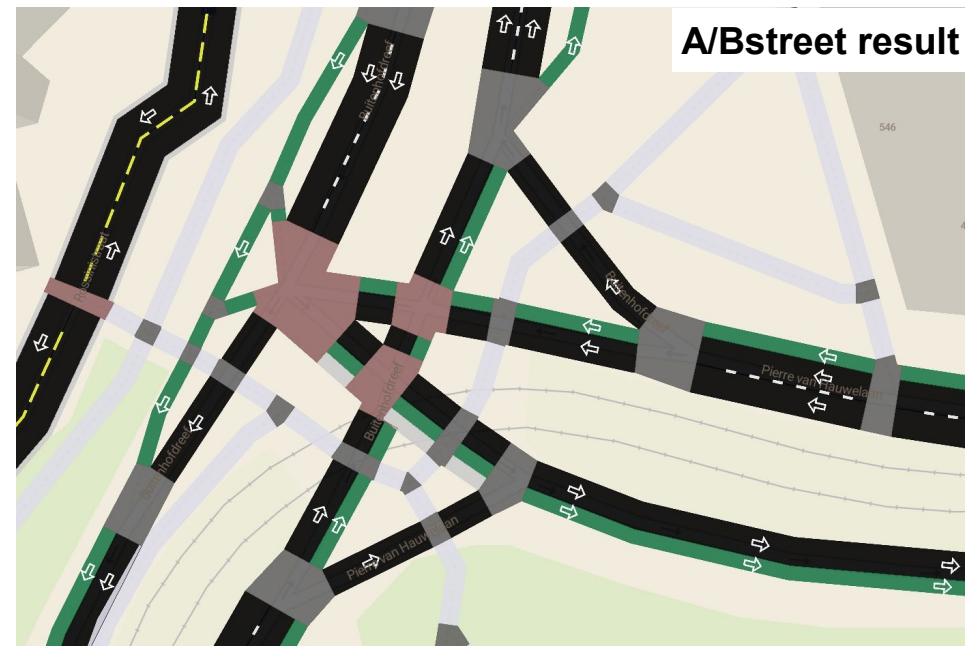
Intersection polygons

-  cycleway intersection
-  roundabout
-  vehicle intersection

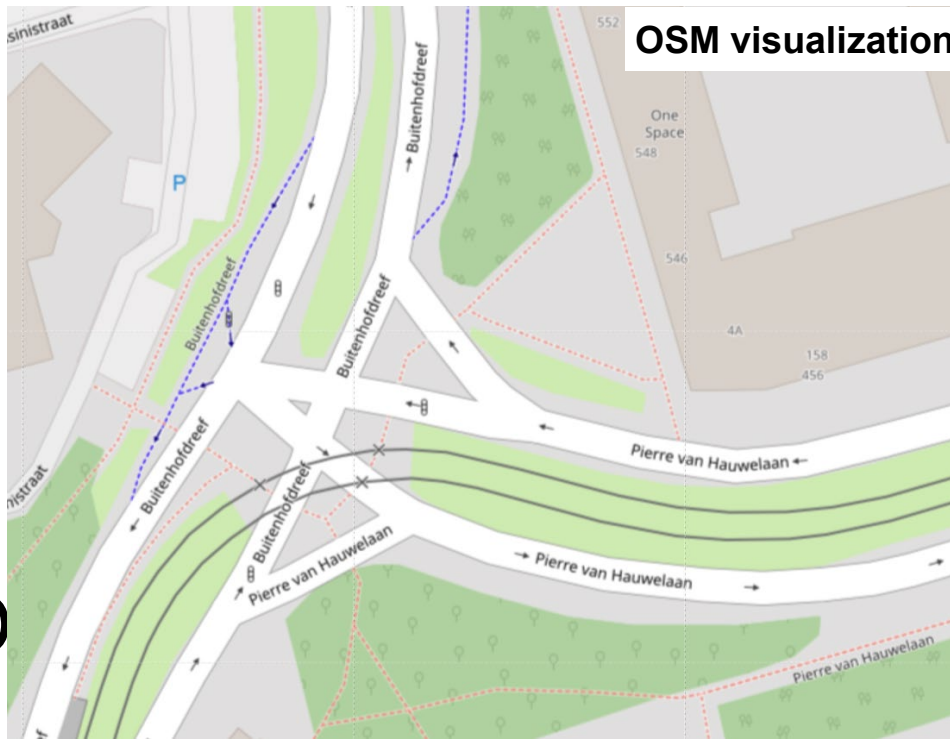
Our final result



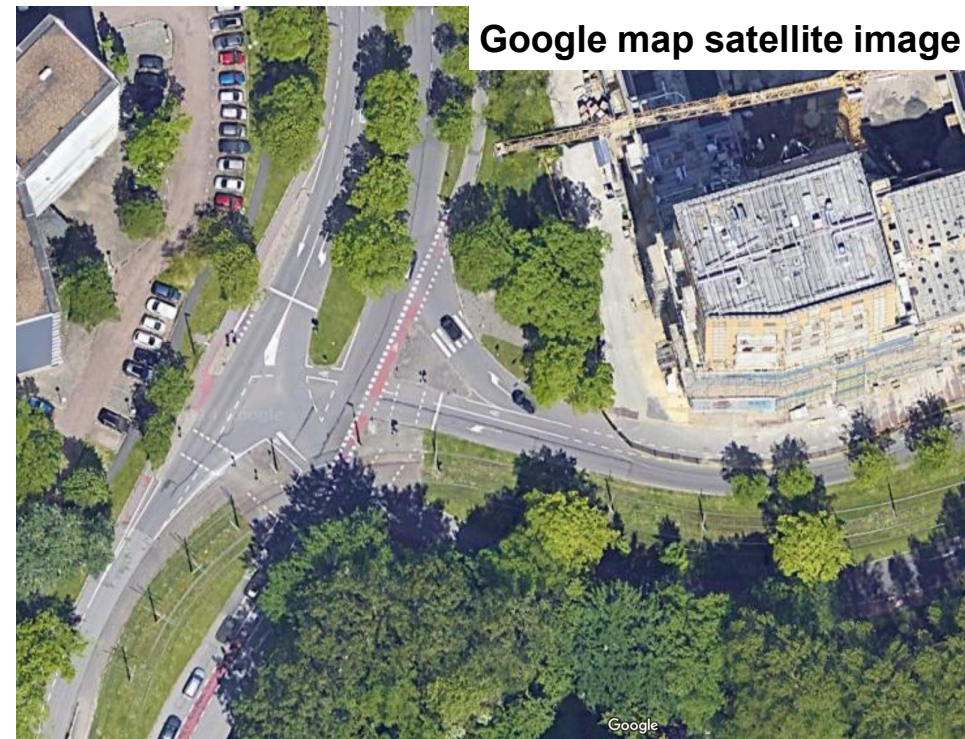
A/Bstreet result



OSM visualization



Google map satellite image



Conclusion

Subquestion 1 & 2: OSM data utilization and LoDs improvements

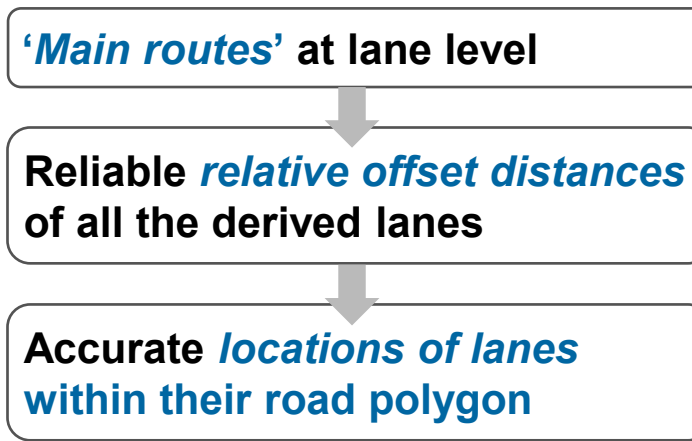
Phases/Tags	<i>highway</i>	<i>cycleway</i>	<i>name</i>	<i>oneway</i>	<i>lanes</i>	<i>turn</i>	<i>junction</i>	<i>bridge</i>	<i>tunnel</i>
Road-level (R)	x	x	-	x	x	x	x	x	x
Traffic-level (R_w)	x	x	x	x	x	-	-	-	-
Lane-level (R_l)	x	x	-	x	x	x	-	-	-
Grouped strokes (S)	x	x	x	x	-	-	-	-	-
Road polygon (RA)	x	x	-	x	x	-	-	-	-
Intersection polygon (C, CA)	x	x	-	-	x	-	x	x	x

Table 7.1: Usages of the tags and corresponding values.

Conclusion

Subquestion 3: Raw topology relationships, lane-to-lane adjacency as the lane-level network

- **Hierarchical generation**



The first phase of lane generation **identifies reliable lanes** that serve as references for building the topology structure.

By establishing both **global and local adjacency (grouped strokes)**, the method accurately maintains **lane-to-lane** relationships.

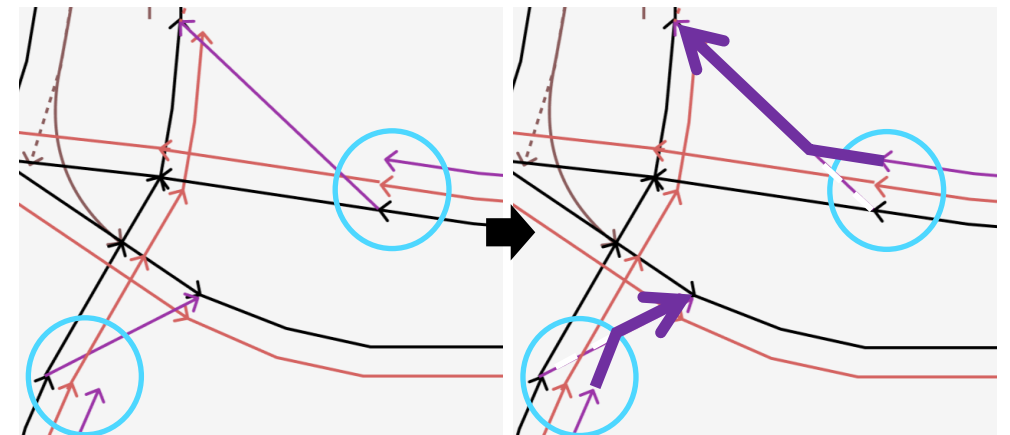
Conclusion

Subquestion 3: Raw topology relationships, lane-to-lane adjacency as the lane-level network

• Lane modification and correction

Several strategies to ensure correct lane adjacency:

- 1. Separate generation and modification of vehicle lanes and cycling lanes to prioritize the consistency of correct traffic flows and achieve proper lane matching.
- 2. Slip road labels (F_r) were used to identify turning lanes, ensuring topological correctness through proper connections.
- 3. Turn attributes (T_r and T_l) helped to determine the target lanes for slip road modifications.
- 4. Finally, comparing T values and relative offset distances (FD_l) allowed auxiliary lanes to find their corresponding 'through' lanes, completing the wedge-shaped modification of lane geometry.



Conclusion

Subquestion 4: Error-less areal representation with details information and changes

- *Node degree-based approach*

'True' node degree correction

Error types identification:

- Overlapping
- Protruding

Intersection detection

- *Geometric method for post-processing*

Error-free road ribbons

- *Intersection generation*

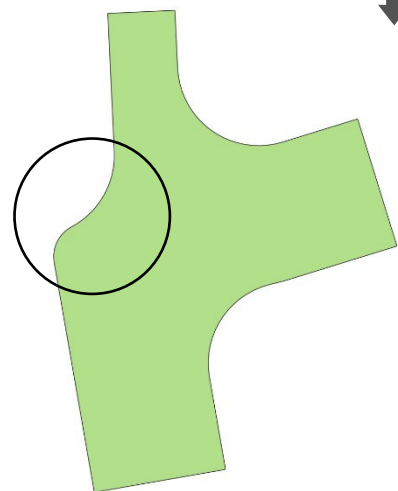
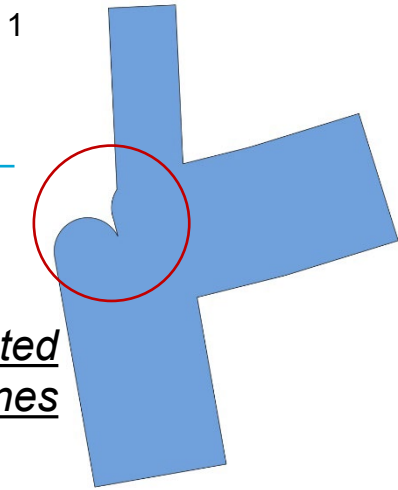
Several strength:

- 1. Resolving complex overlapping issues by merging trimmed polygons of converging roads, ensuring a smooth, cohesive shape.
- 2. Incorporating turning corner curvatures into intersection polygons, which optimizes the areal representation and follows transportation geometric design principles.
- 3. Addressing lane connectivity errors—especially those caused by missing road information—by generating compact, reliable intersection polygons that reflect real-world road areas.

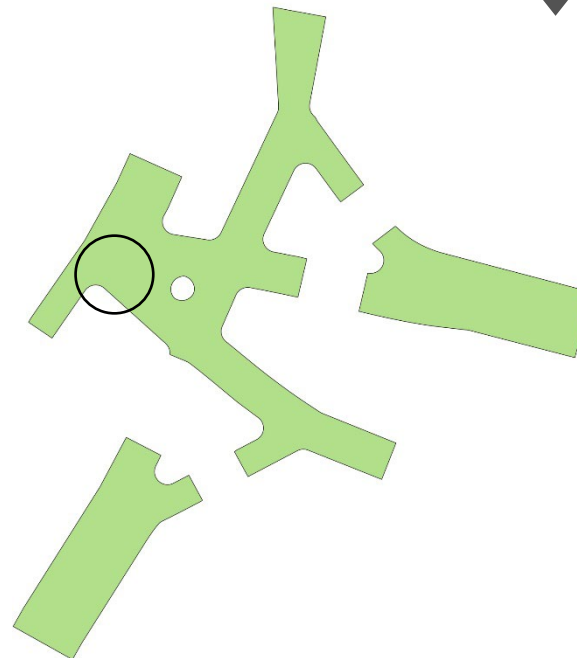
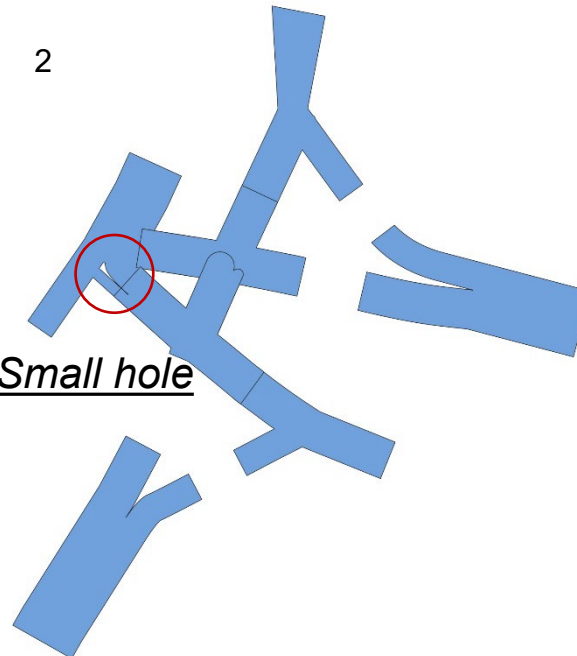
Conclusion

Subquestion 4:

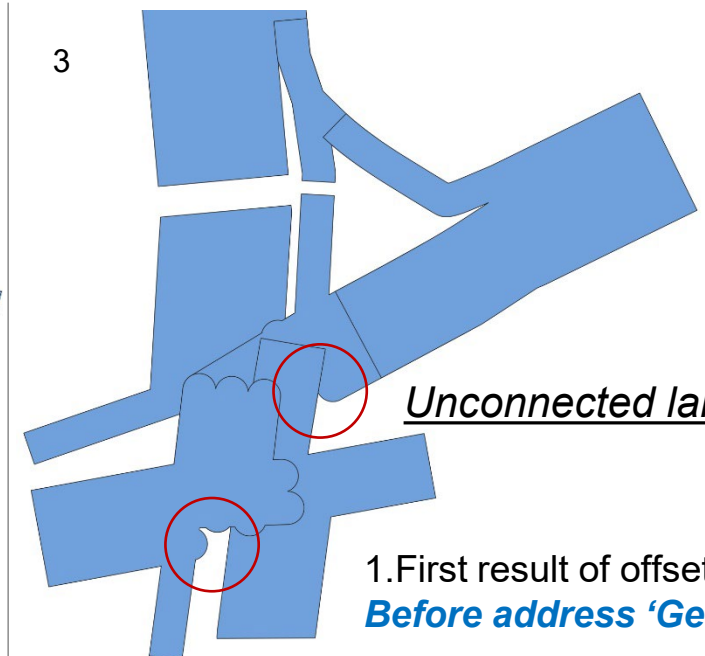
Unconnected lanes



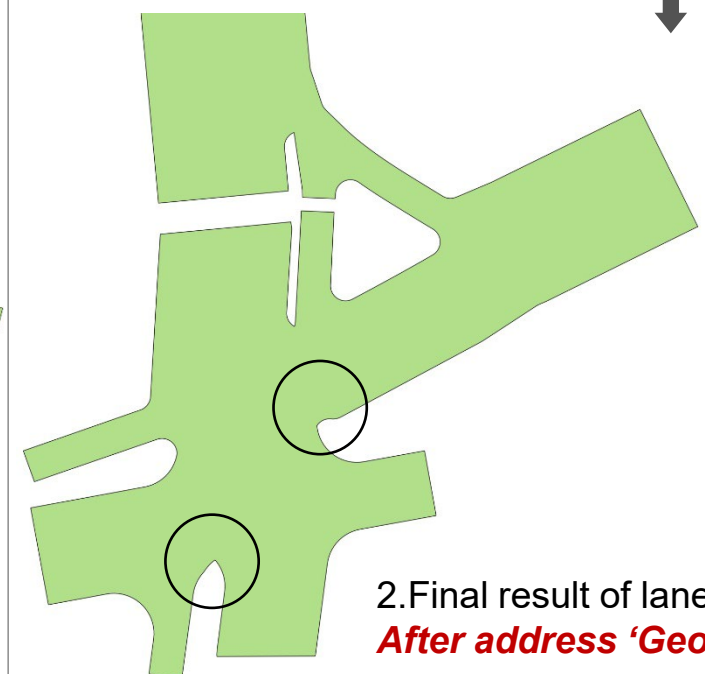
Small hole



Unconnected lanes



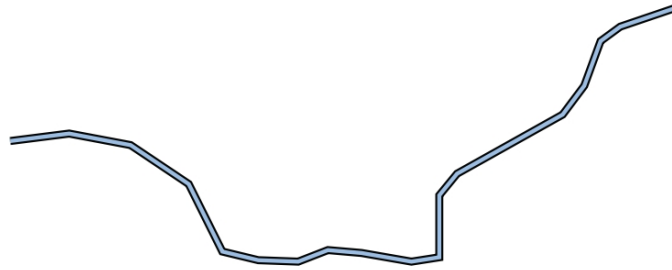
1. First result of offsetting lanes
Before address 'Geometric error'



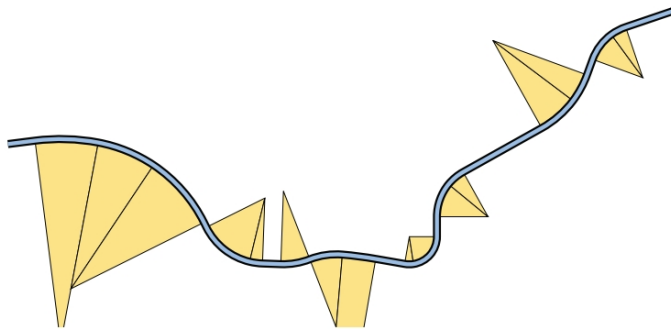
2. Final result of lane generation
After address 'Geometric error'

Future work

- ❑ Optimize the road centerlines to restore **C1 continuity** before lane generation to satisfy C1 continuity curve for road/lane centerlines and road boundaries.



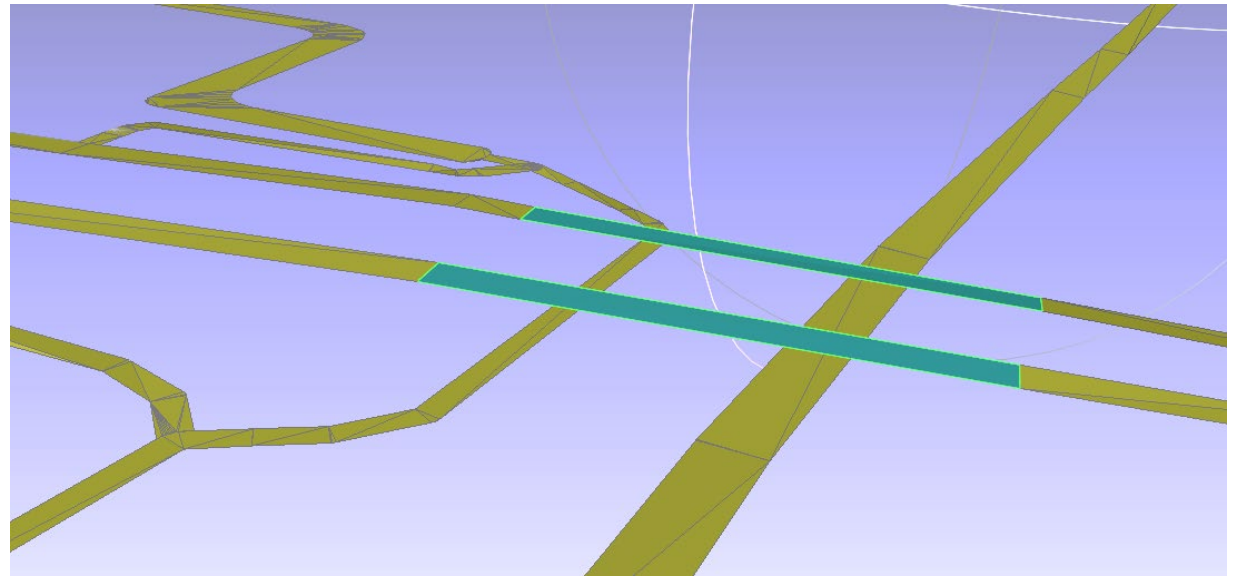
(a) Polyline road geometry (imported from the TIGER[®] database [4] via OpenStreetMap [5])



(b) An *arc road* derived from the above polyline. The orange arcs show the center and radius of each arc used to give the road its smooth appearance.

- ❑ Utilize the intersection-related tags and the areal model, to generate **2.5D/3D model** for bridges, tunnels, overpasses, etc.

junction = *
bridge = *
tunnel = *





Questions ?

Thanks for your attention

Chengzhi Rao