Exploring the automatic Level of Detail inference for the validation of buildings in 3D city models

MSc thesis presentation
Geomatics for the Built Environment

Balázs Dukai

Supervisors:

Dr. Filip Biljecki

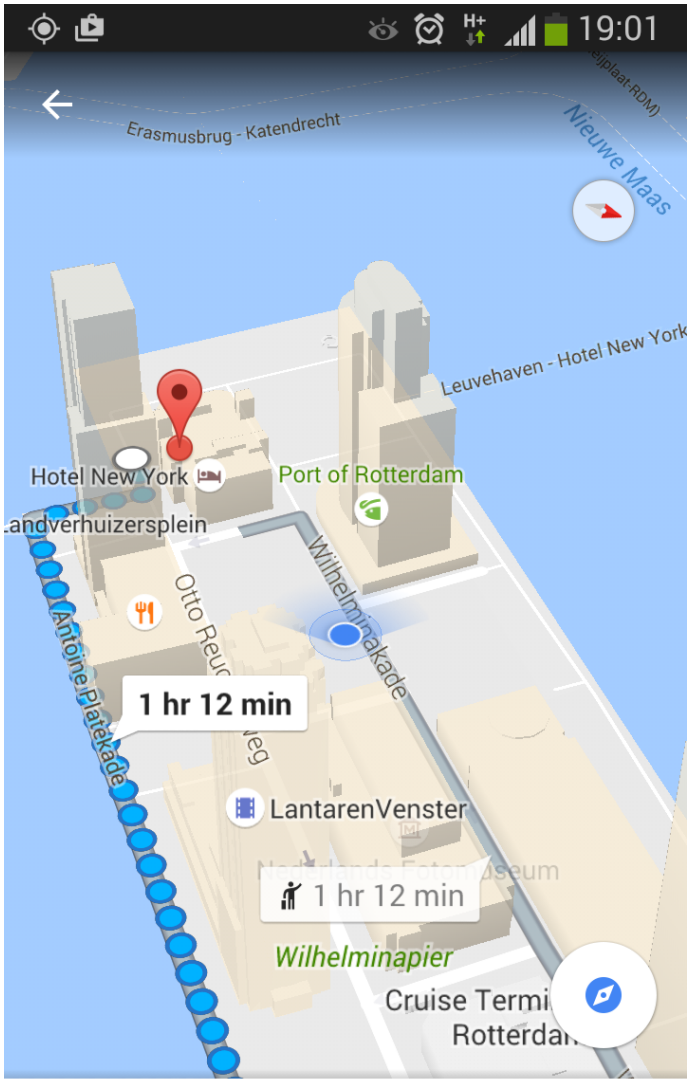
Dr. Hugo Ledoux

Anna Labetski

Co-reader: Dr. Jorge Lopez Gil

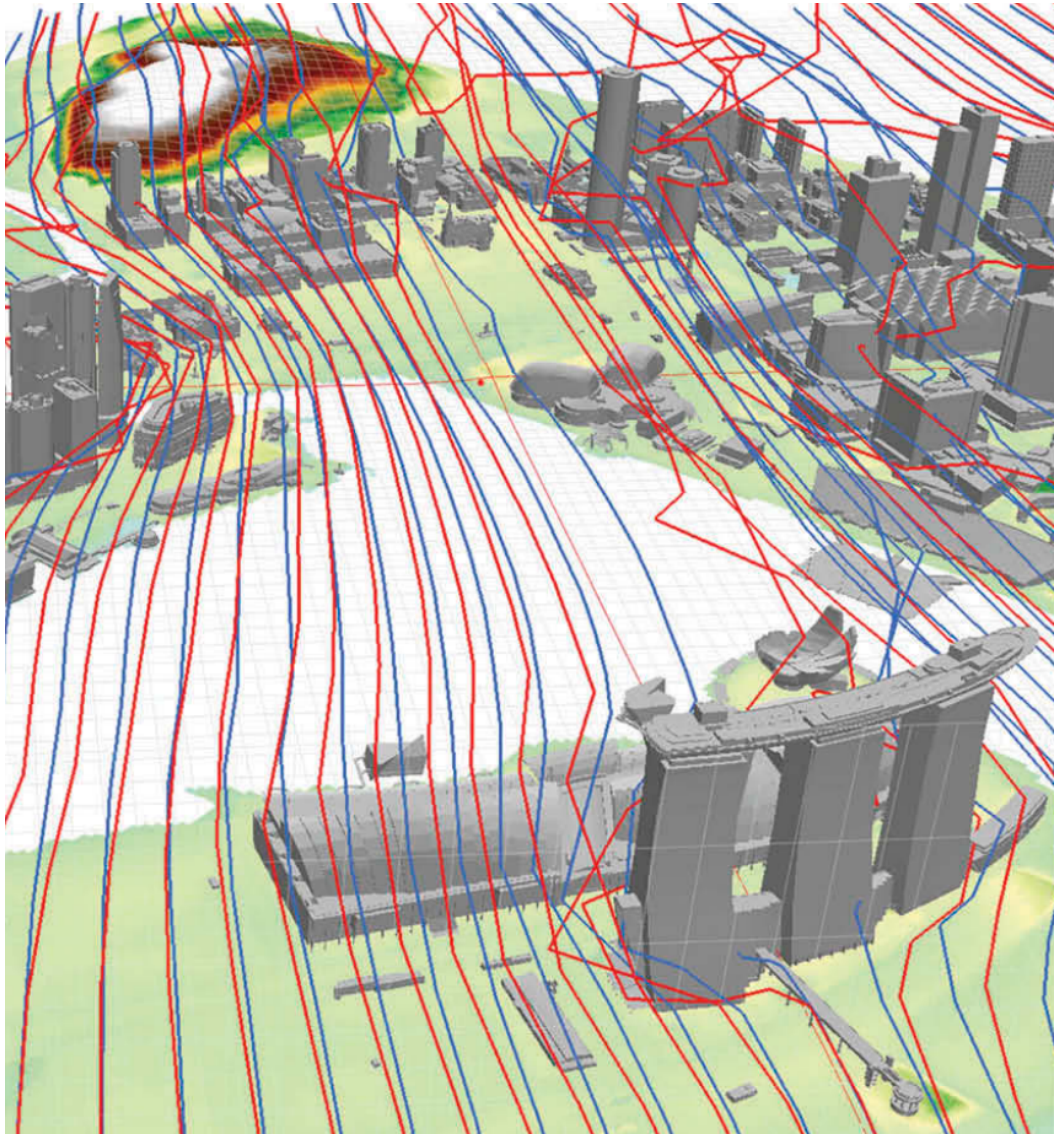
30.01.2018

TU Delft

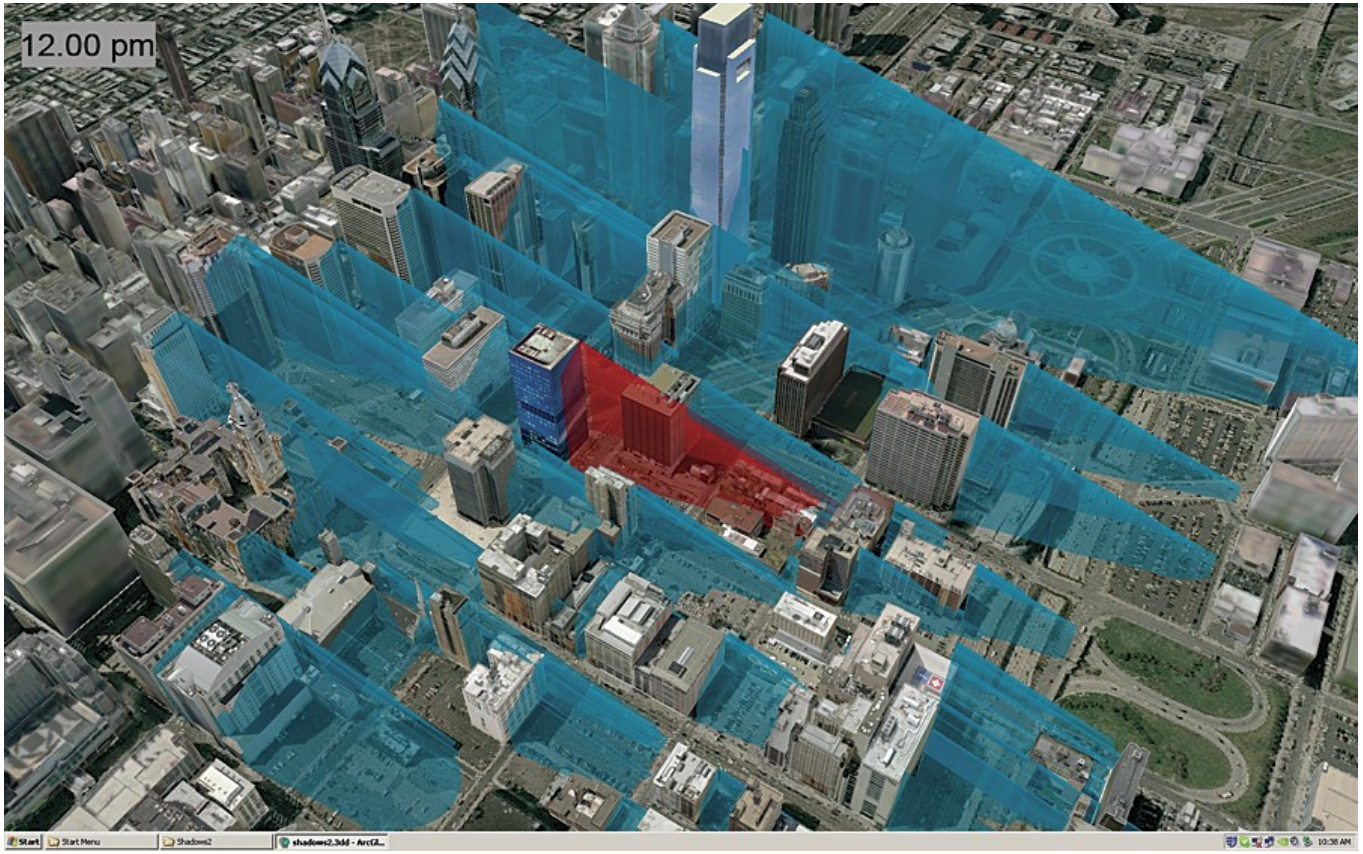


6 › › › 9 **1 hr 12 min**

Source: Google Maps Mobile Application

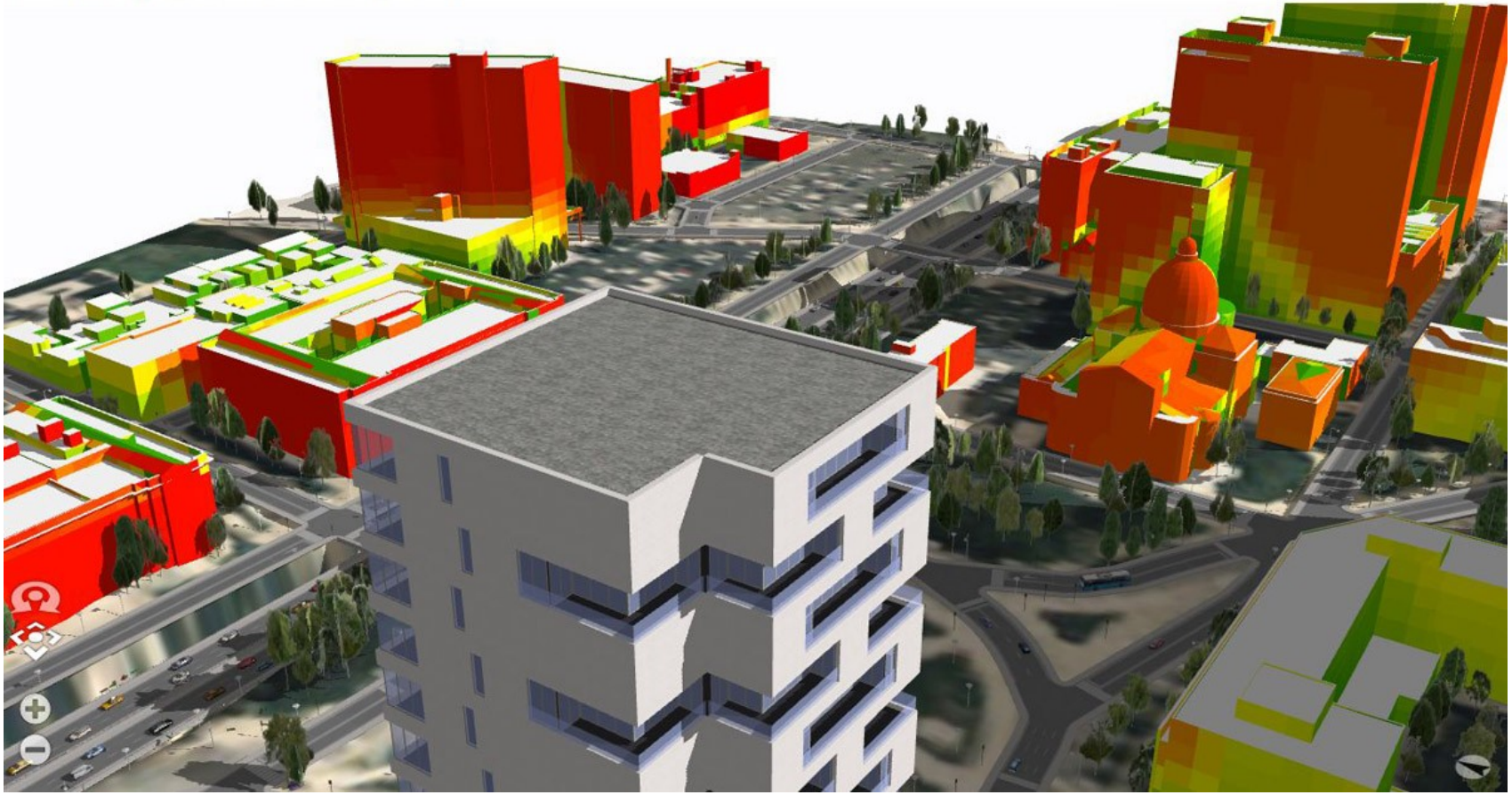


Source: [1]



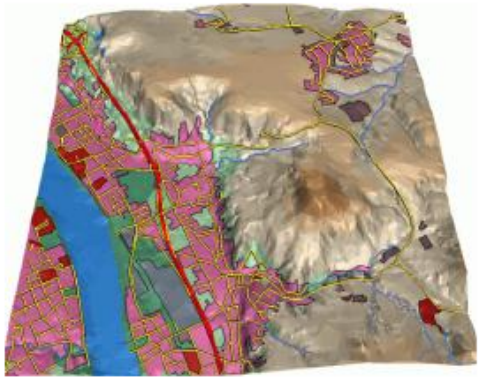
Source: http://www.esri.com/news/arcnews/summer10articles/summer10gifs/p13p1-lg.jpg

Philadelphia Redevelopment

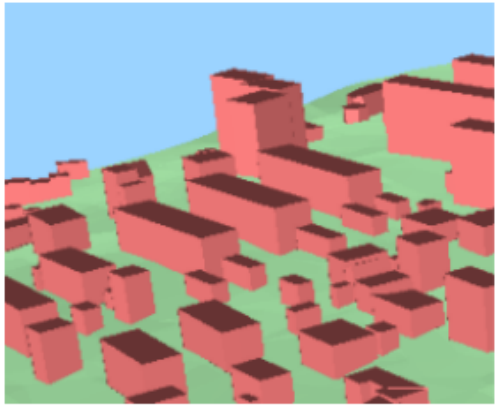


Source: https://informedinfrastructure.com/wp-content/uploads/2014/09/CityEngine_Philadelphia.jpg

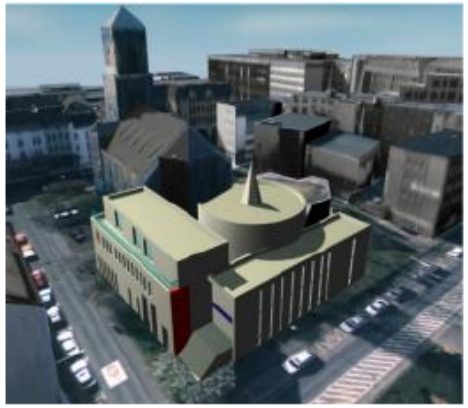
Level of Detail (LoD)



LOD0

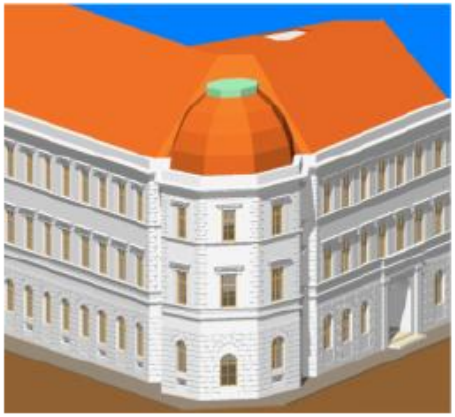


LOD1



LOD2

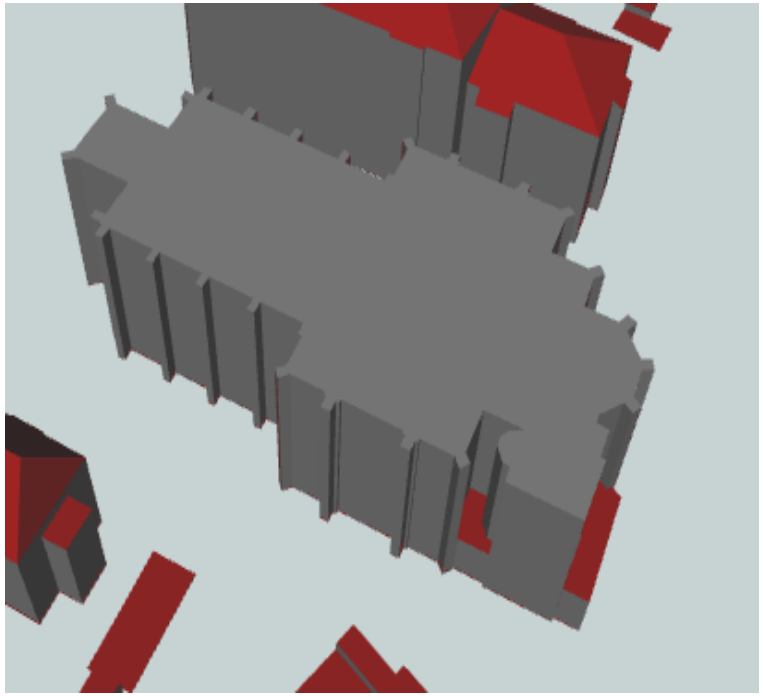
LOD3



LOD4

Source: [2]

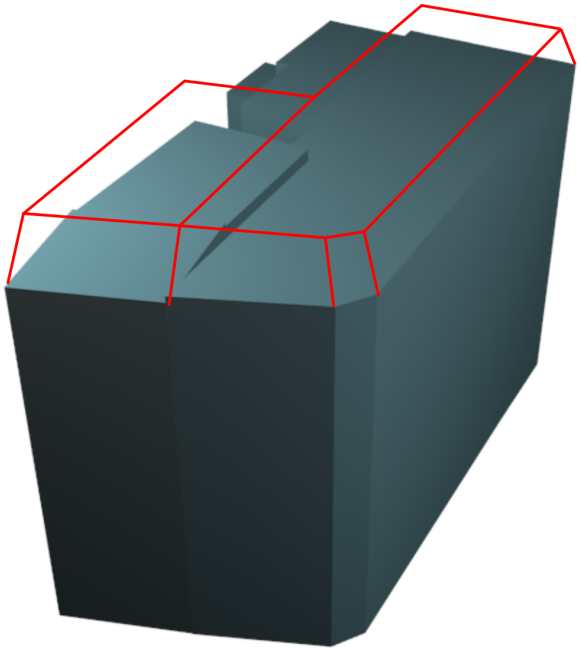
Motivation



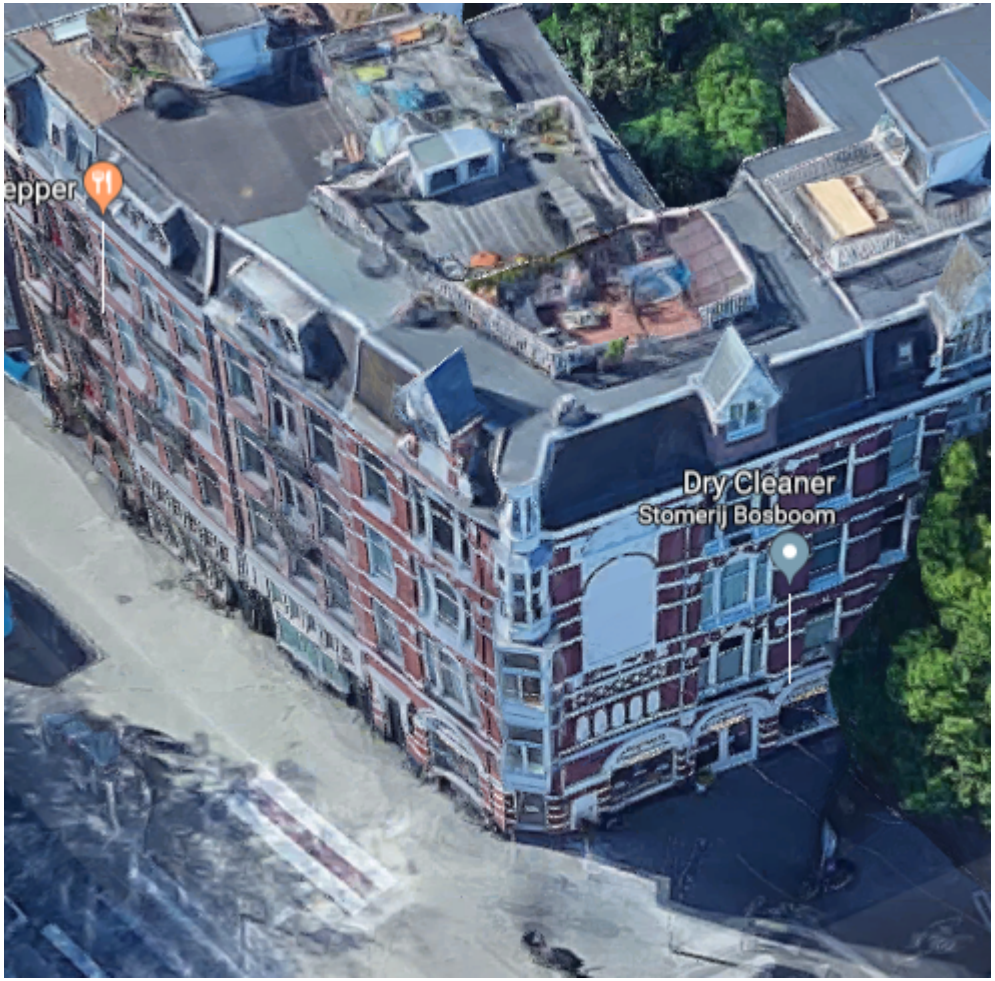
Source: LoD2 model of Bad Godesberg, NRW, Germany

Source: Google Maps

Motivation

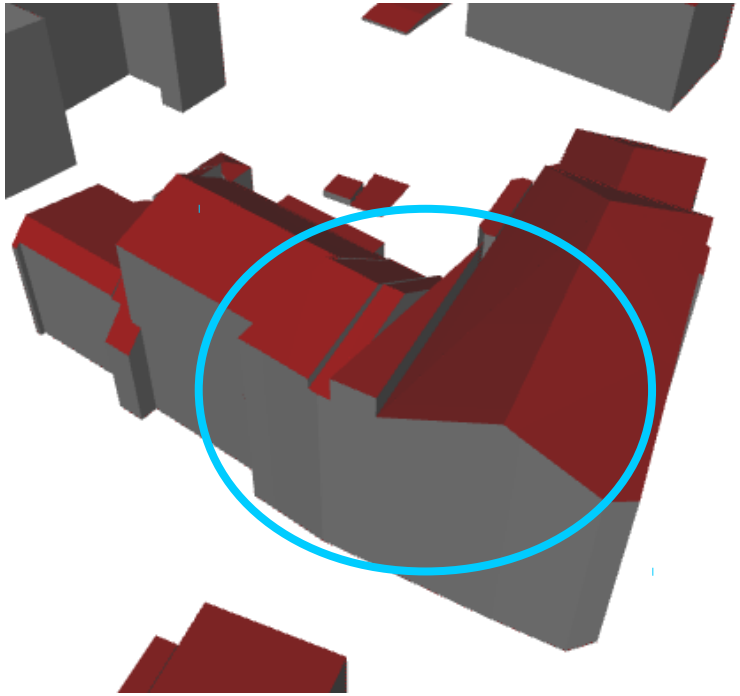


Source: LoD2 model of Amsterdam, virtualcitySystems

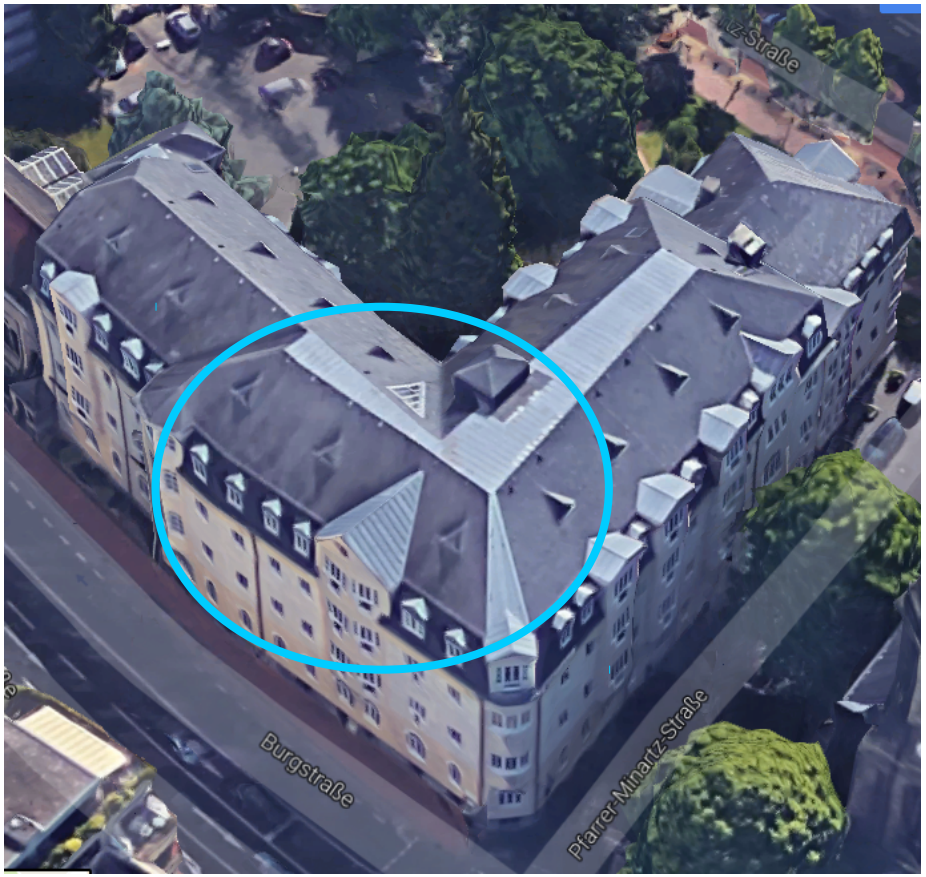


Source: Google Maps

Motivation



Source: LoD2 model of Bad Godesberg, NRW, Germany



Source: Google Maps

Motivation

- Knowing the accurate LoD is important for analysis and maintenance
- CityGML 2.0 is not clear on LoD, CityGML 3.0 will probably complicate things
- Roof reconstruction (>LoD2) fails occasionally
- Heterogenous LoD
- CityGML has no explicit LoD attribute per building, non-semantic formats have no tag at all

Research questions (paraphrased)

How to determine the geometric LoD automatically?

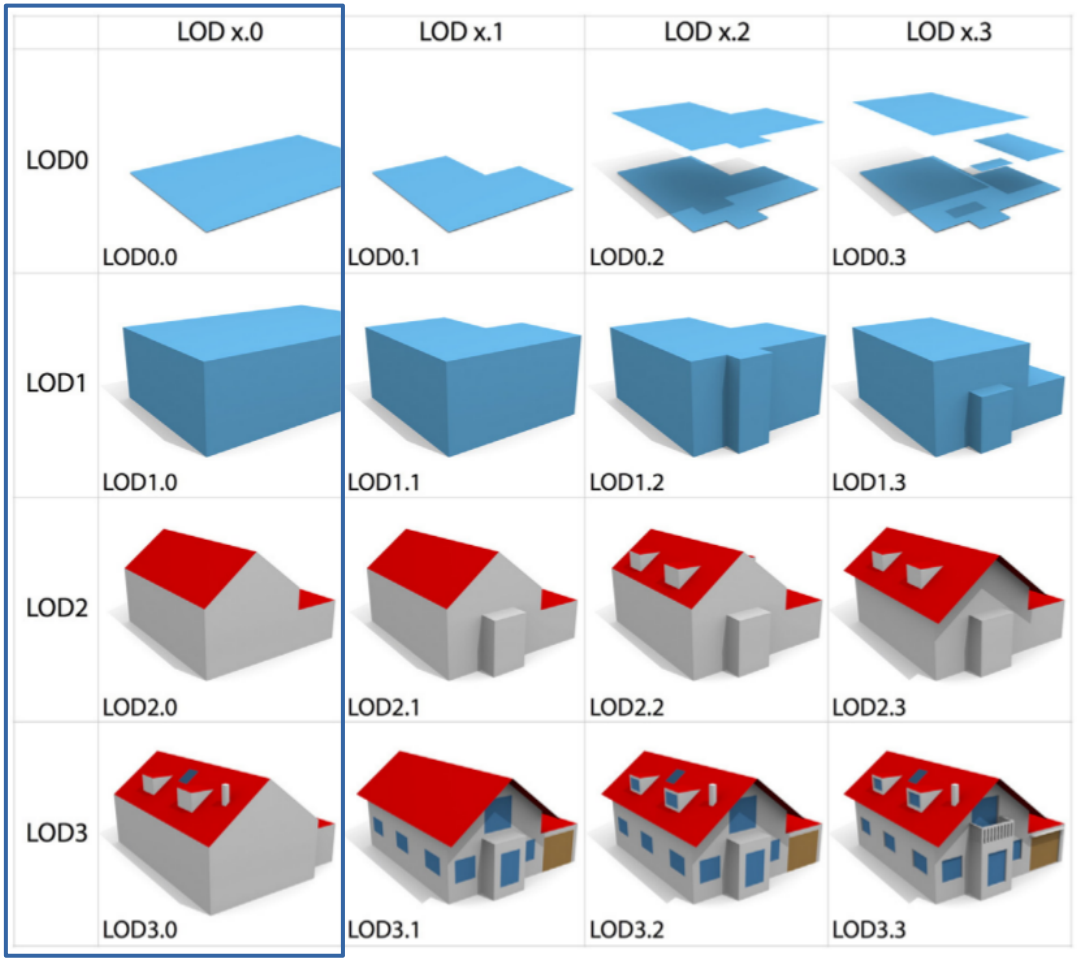
- How to classify the geometry of 3D building models (in terms of LoD)?
  - How to describe the geometry of a building model for the classification?

Research questions (paraphrased)

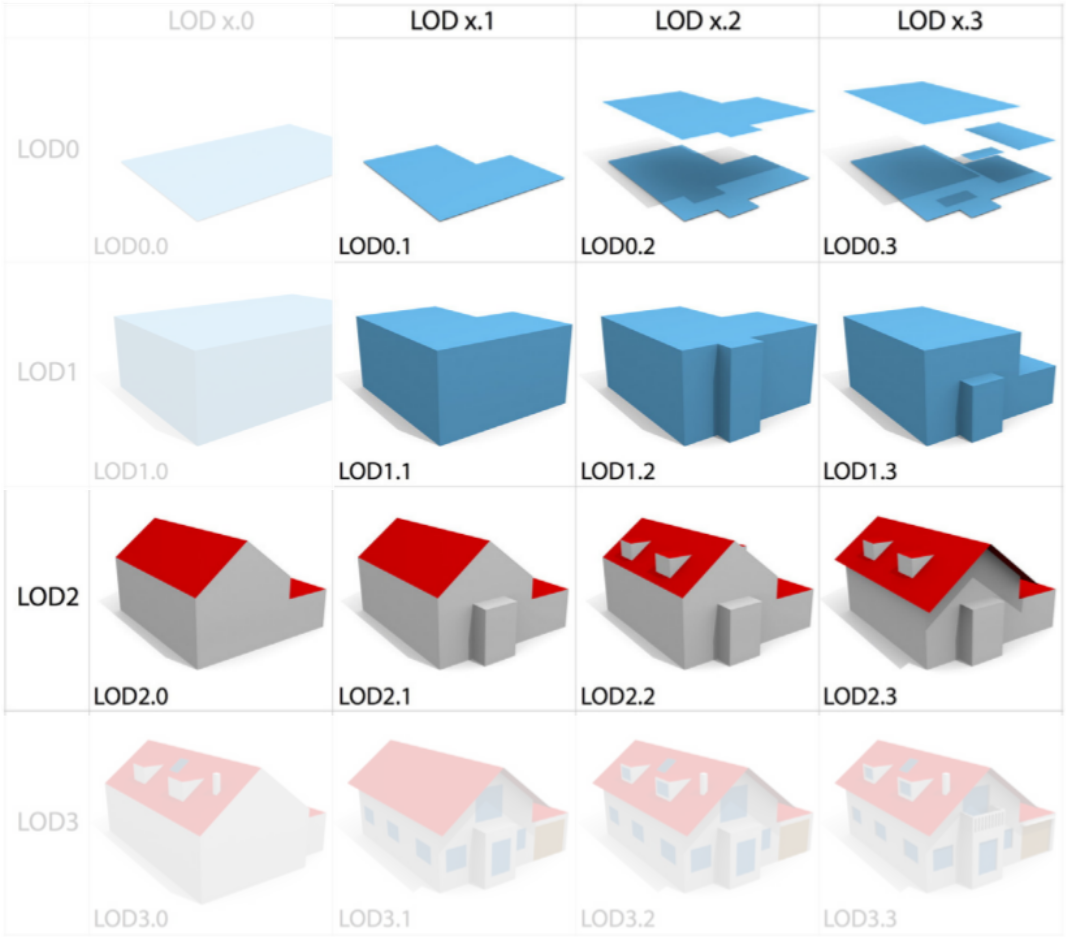
- How to validate the LoD automatically?
  - Without comparing to a reference data set?
  - By comparison with a reference data set?

LoD[3] revisited

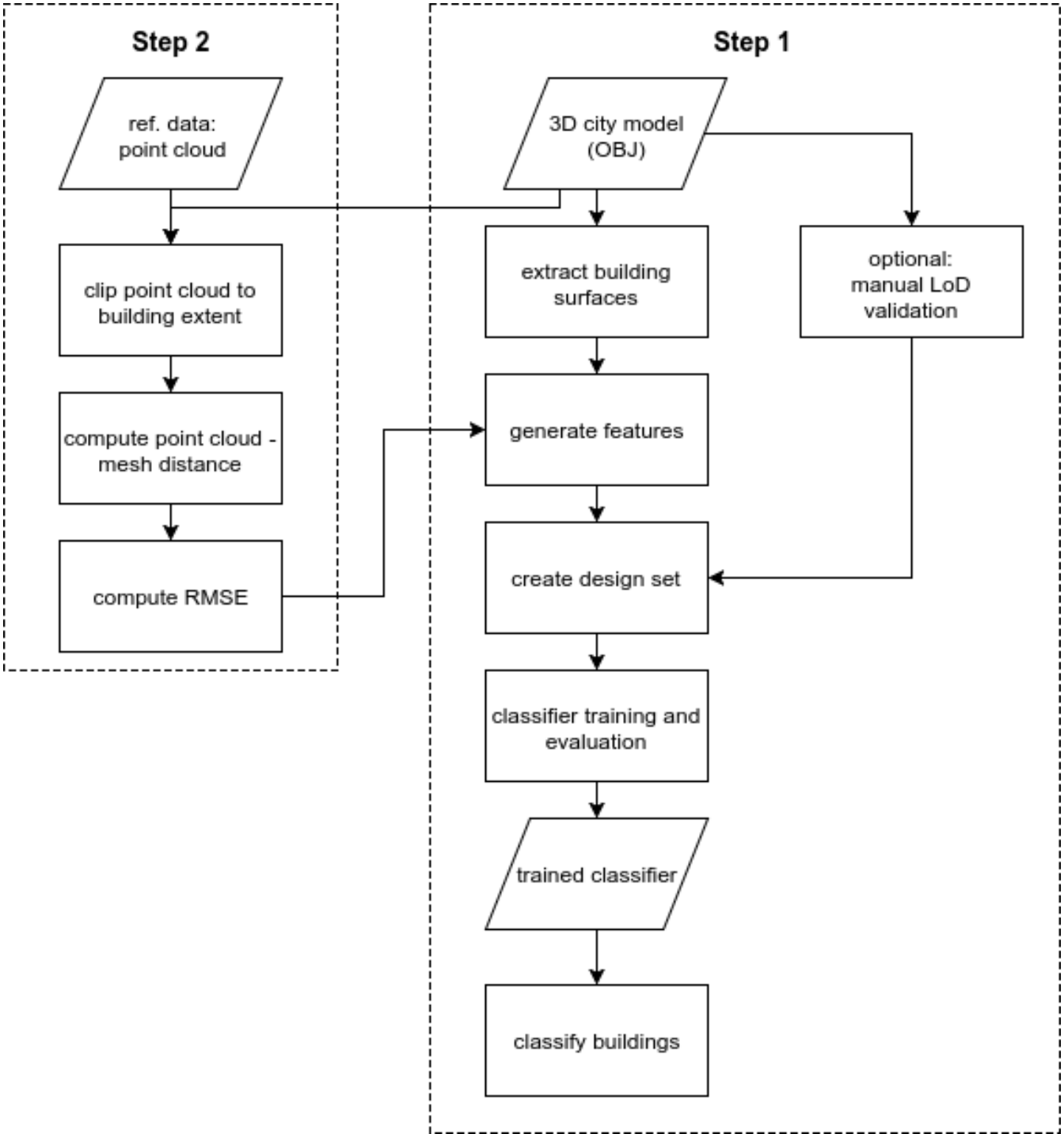
CityGML2.0



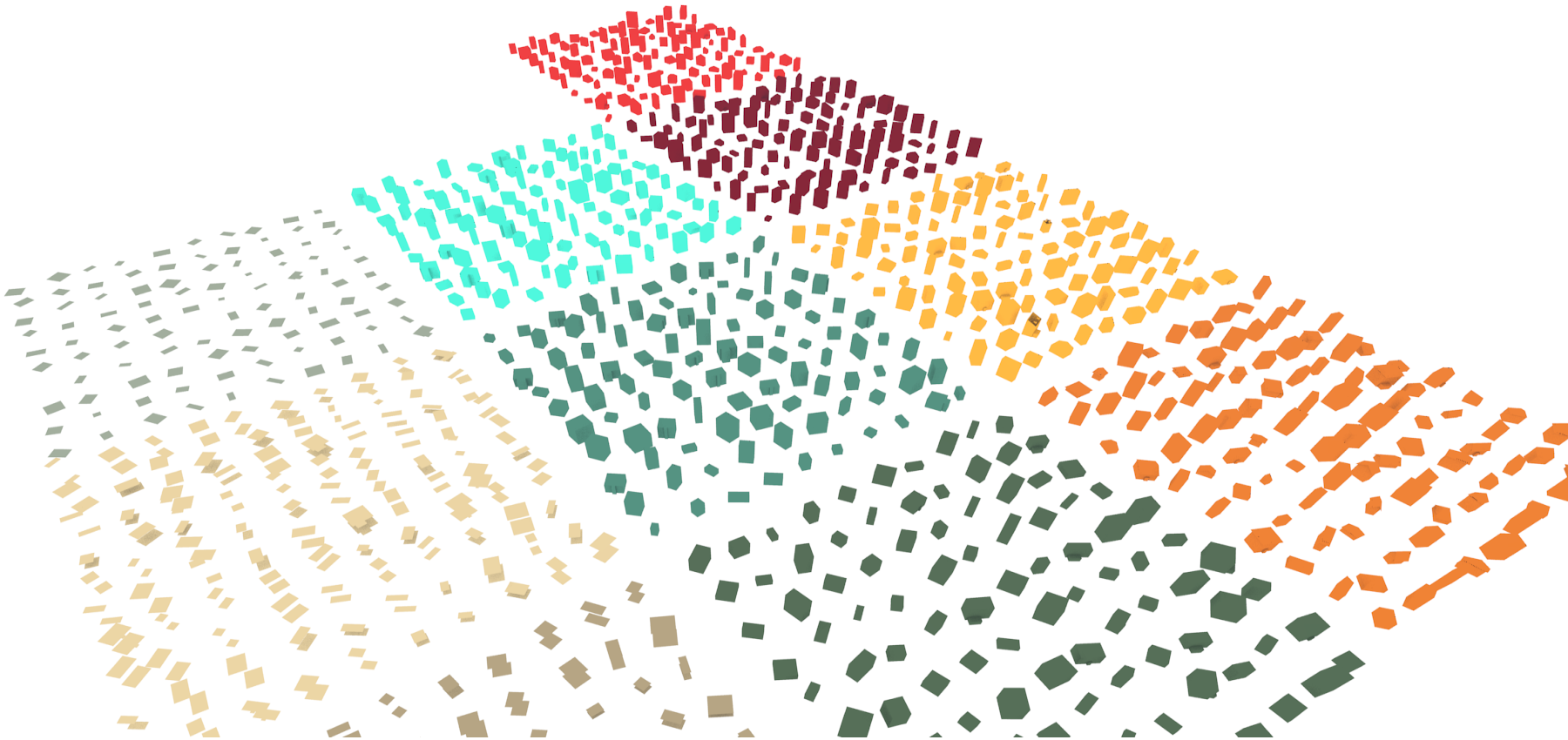
LoD0.1-0.3, 1.1-2.3



Method


Step 2
ref. data: point cloud
clip point cloud to building extent
compute point cloud - mesh distance
compute RMSE
Step 1
3D city model (OBJ)
extract building surfaces
optional: manual LoD validation
generate features
create design set
classifier training and evaluation
trained classifier
classify buildings

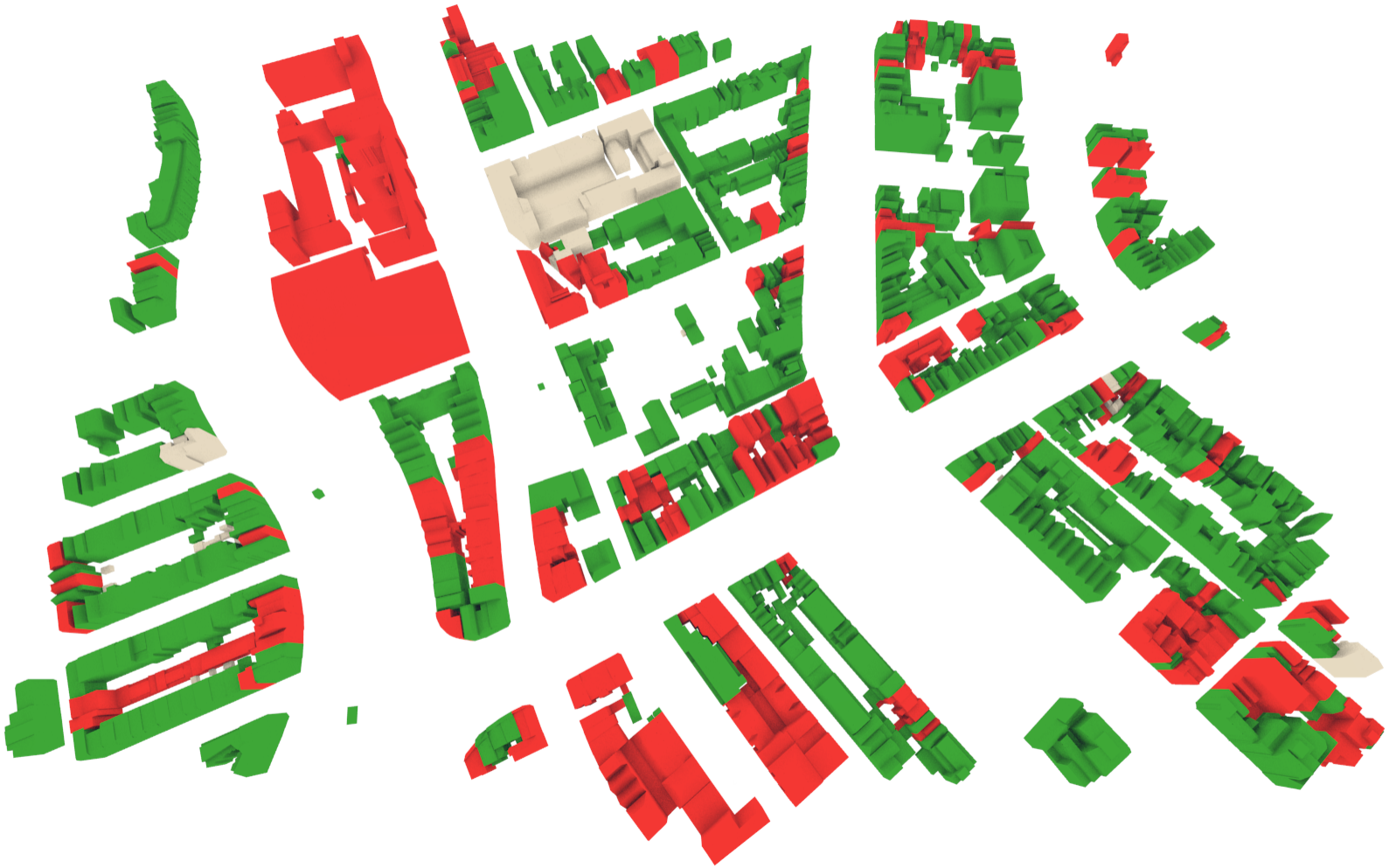
Synthetic data – LoD0.1-0.3, 1.1-2.3



1000 buildings
100 per class

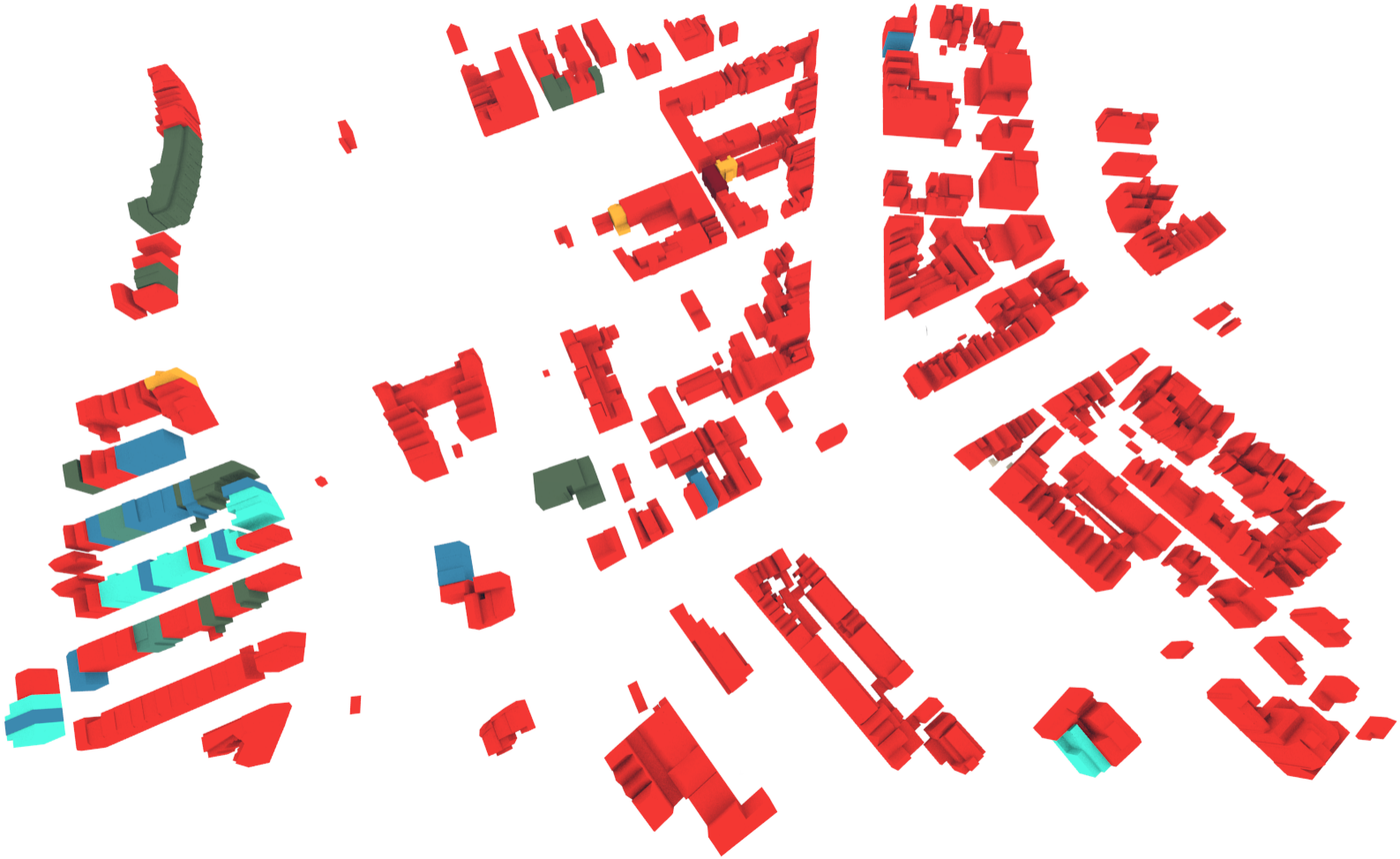
TU Delft

Amsterdam data – LoD1.2, LoD2



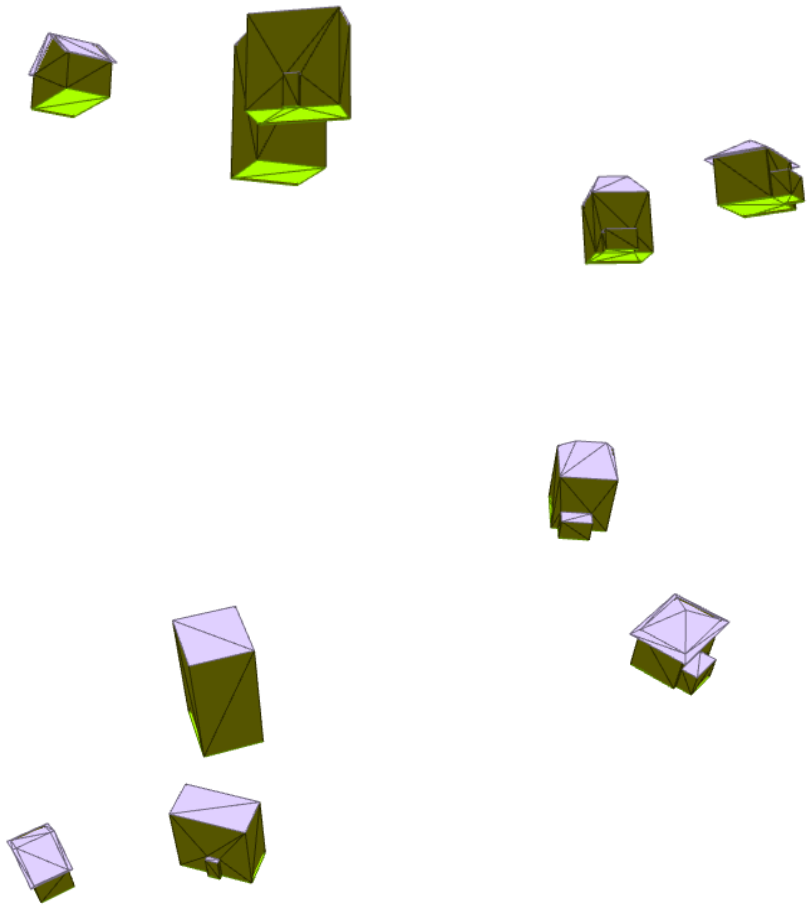
482 valid buildings (green)

Amsterdam data – LoD2 (and LoD1)



Imbalanced LoD classes – LoD2 (reds), LoD1 (blues)

Extract building surfaces

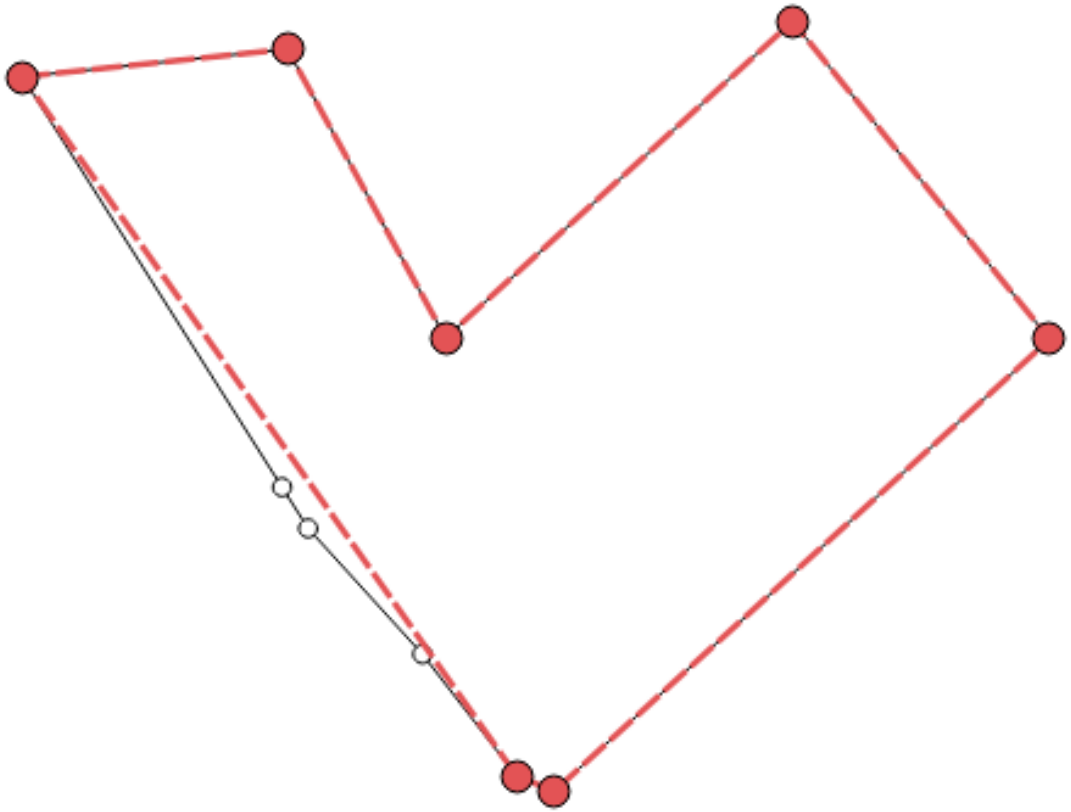


Generate features

Geometry	Feature	Related LoD requirement	Relevant LoD
2D footprint	Number of Shape Characterising Points (NSCP)	none	all
	Shape Characterising Lengths (SCL)	Size of building parts	≥ 0.1
	Footprint Area	Size of building parts	≥ 0.1
	Building Part Footprint Area	Size of building parts	≥ 0.1
3D solid	Building Volume	none	all
3D surface	Roof Type	Roof representation	≥ 1
	Median Roof Gap	Top surface (Single / Multi)	0.2-1.3
	Roof Overhangs	Explicit roof overhangs (if o.2m)	≥ 2.3
	Footprint-Roof Triangle Ratio	Roof superstructures	≥ 2.2
	Walls	Presence of walls	0
3D solid, Point Cloud	RMSE of PC-Model distance	(LoD validity)	all

NSCP & SCL

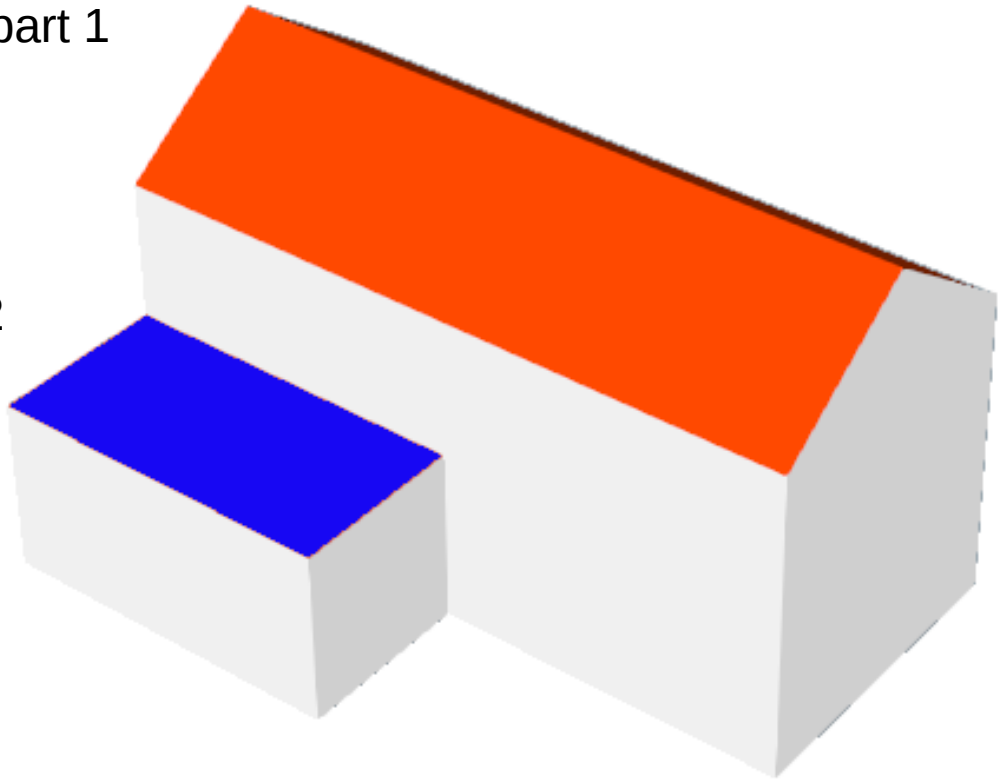
Inner angle < 160°



Building part area

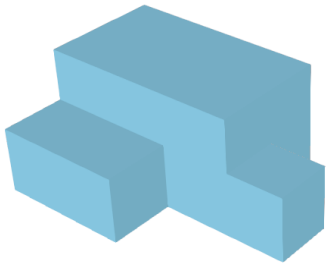
Building part 1

Building part 2

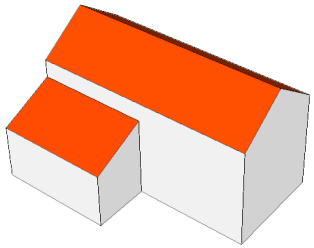


Roof type

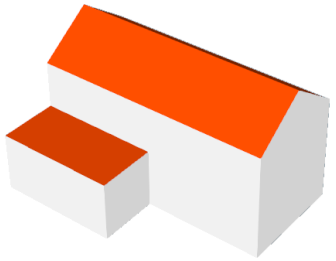
planar



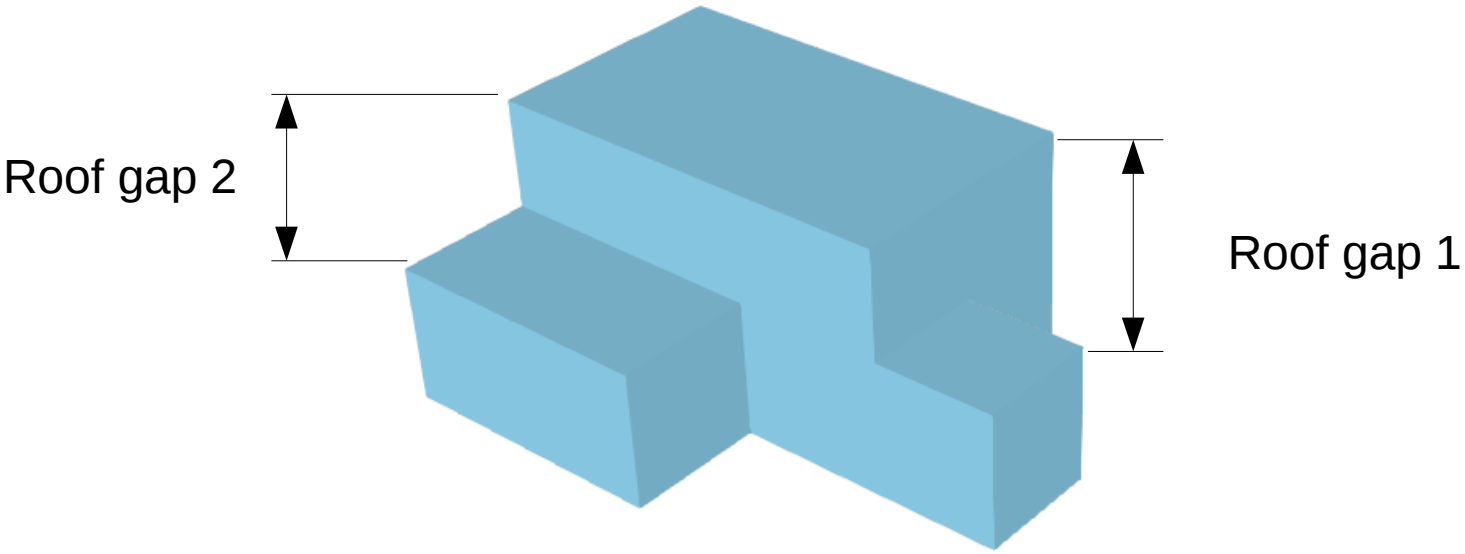
non-planar



mixed

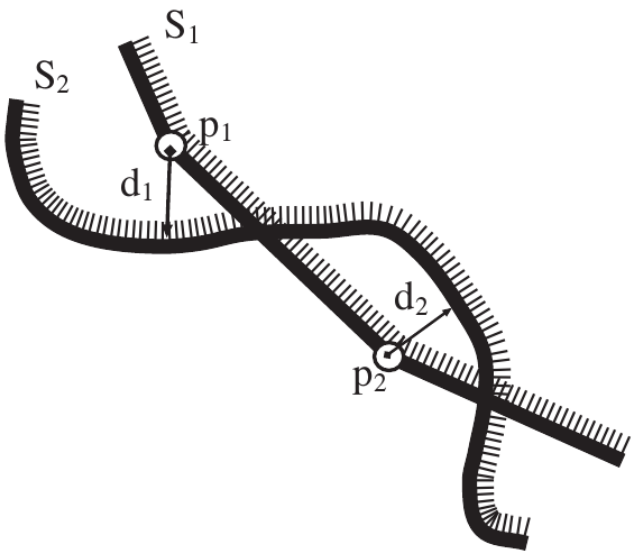


Median roof gap



RMSE

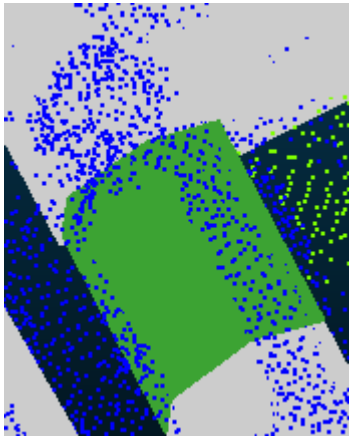
- Signed distance from point cloud to mesh
- With CloudCompare, per building

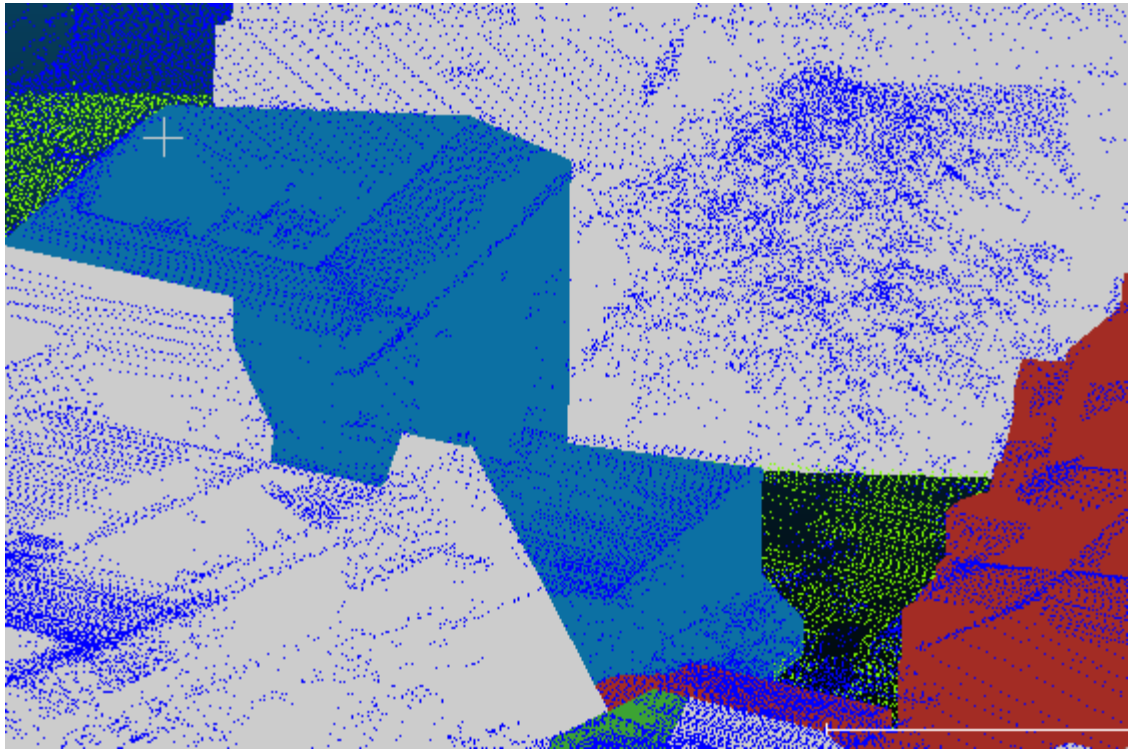


[2]

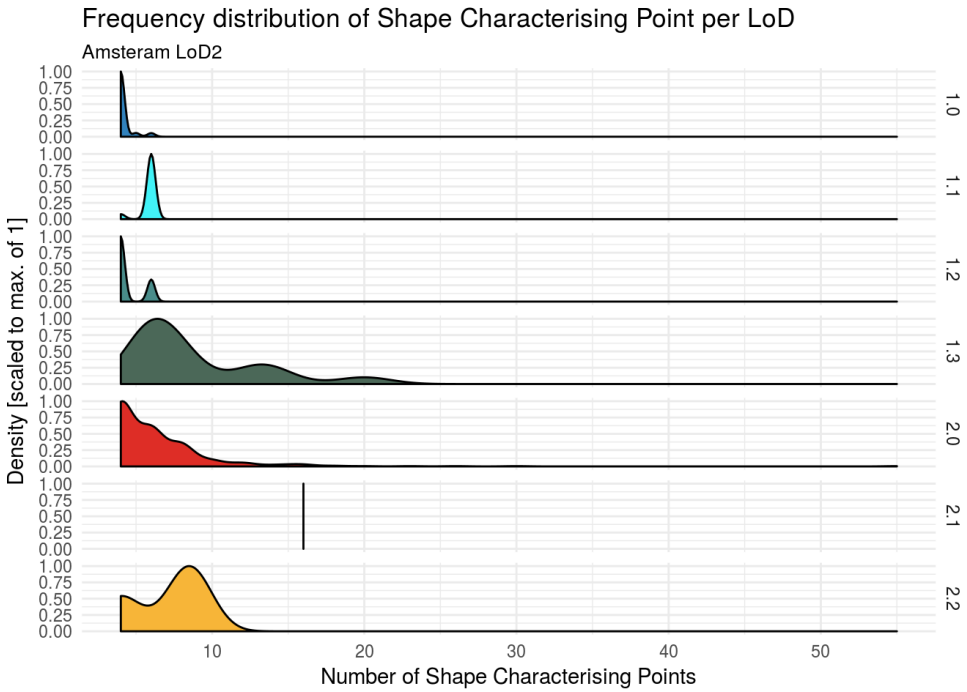
Figure 2: *Signed distance evaluation; distance is positive in p_1 and negative in p_2 (S_1 is the sampled curve).*

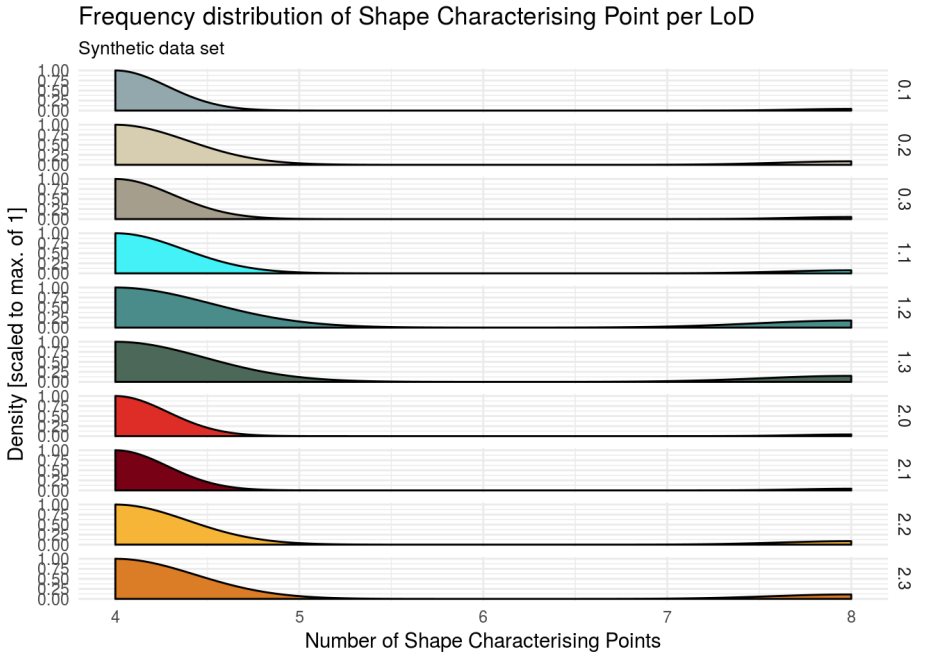
RMSE



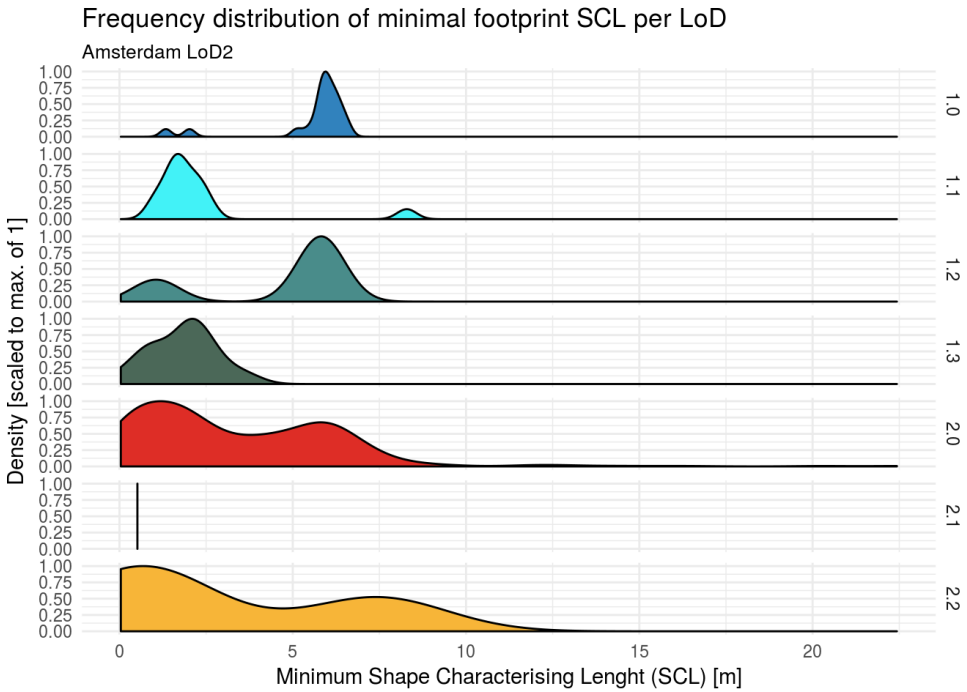


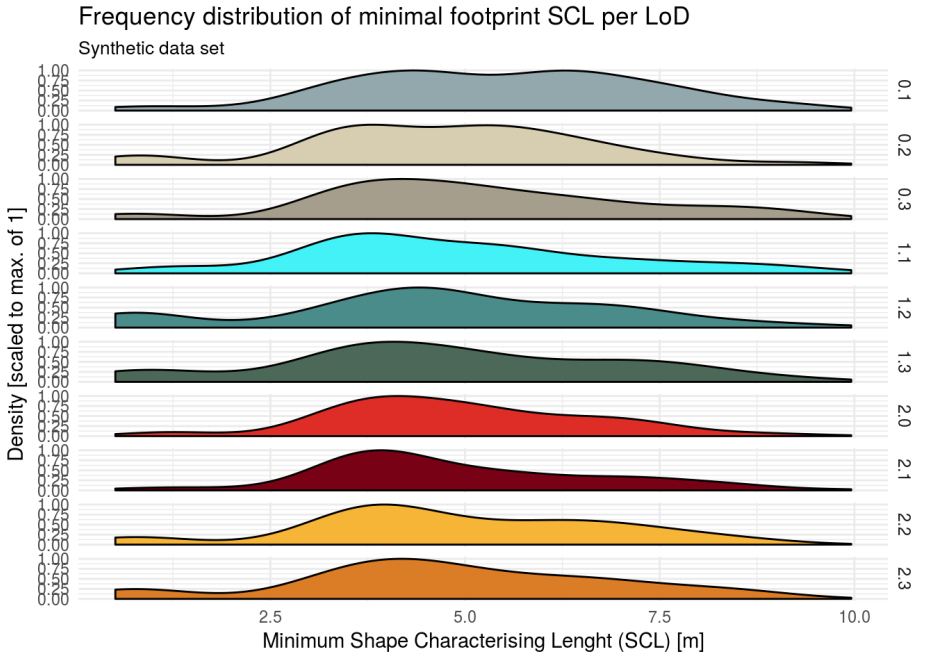
NSCP


Frequency distribution of Shape Characterising Point per LoD
Amsteram LoD2
Density [scaled to max. of 1]
1.00
0.75
0.50
0.25
0.00
1.0
1.1
1.2
1.3
2.0
2.1
2.2
10
20
30
40
50
Number of Shape Characterising Points

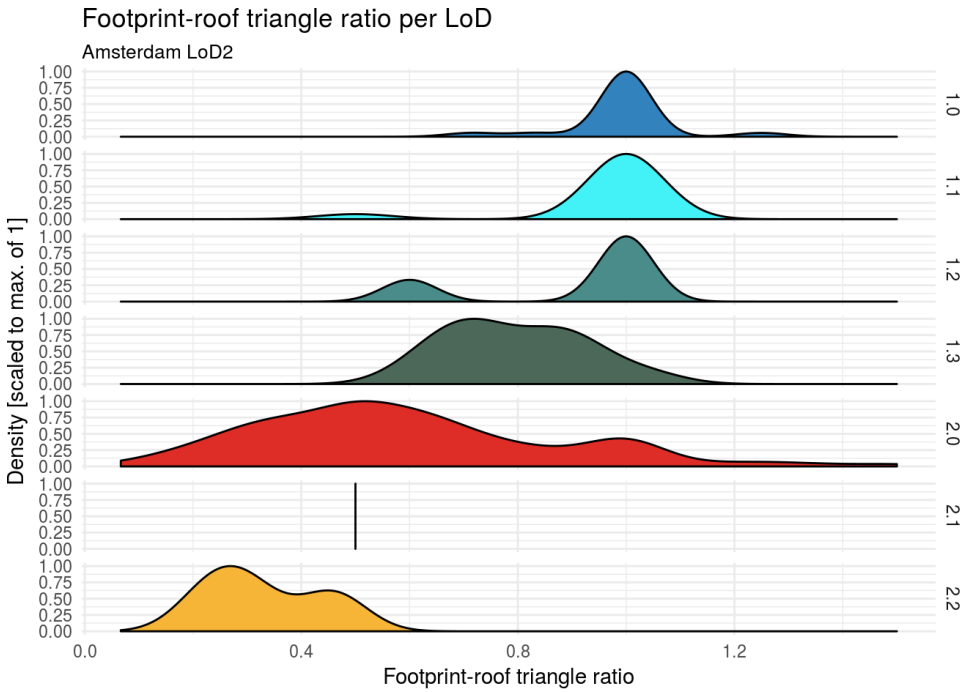

Frequency distribution of Shape Characterising Point per LoD
Synthetic data set
Density [scaled to max. of 1]
1.00
0.75
0.50
0.25
0.00
0.1
0.2
0.3
1.1
1.2
1.3
2.0
2.1
2.2
2.3
4
5
6
7
8
Number of Shape Characterising Points

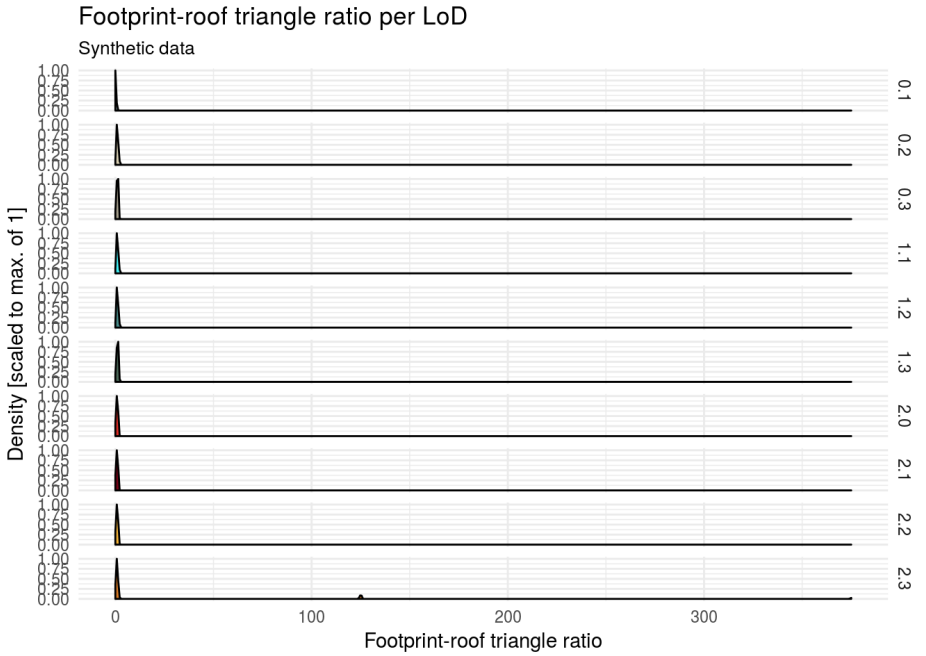
Min. SCL


Frequency distribution of minimal footprint SCL per LoD
Amsterdam LoD2
Density [scaled to max. of 1]
1.00
0.75
0.50
0.25
0.00
1.0
1.1
1.2
1.3
2.0
2.1
2.2
0
5
10
15
20
Minimum Shape Characterising Lenght (SCL) [m]


Frequency distribution of minimal footprint SCL per LoD
Synthetic data set
Density [scaled to max. of 1]
0.1
0.2
0.3
1.1
1.2
1.3
2.0
2.1
2.2
2.3
2.5
5.0
7.5
10.0
Minimum Shape Characterising Lenght (SCL) [m]

Footprint-roof ratio





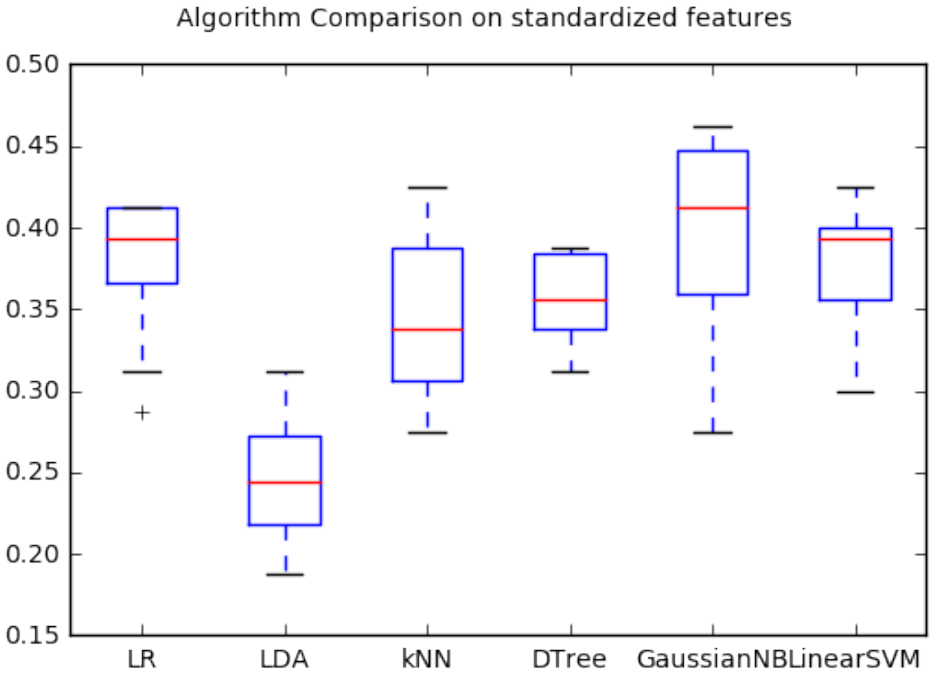
Classification

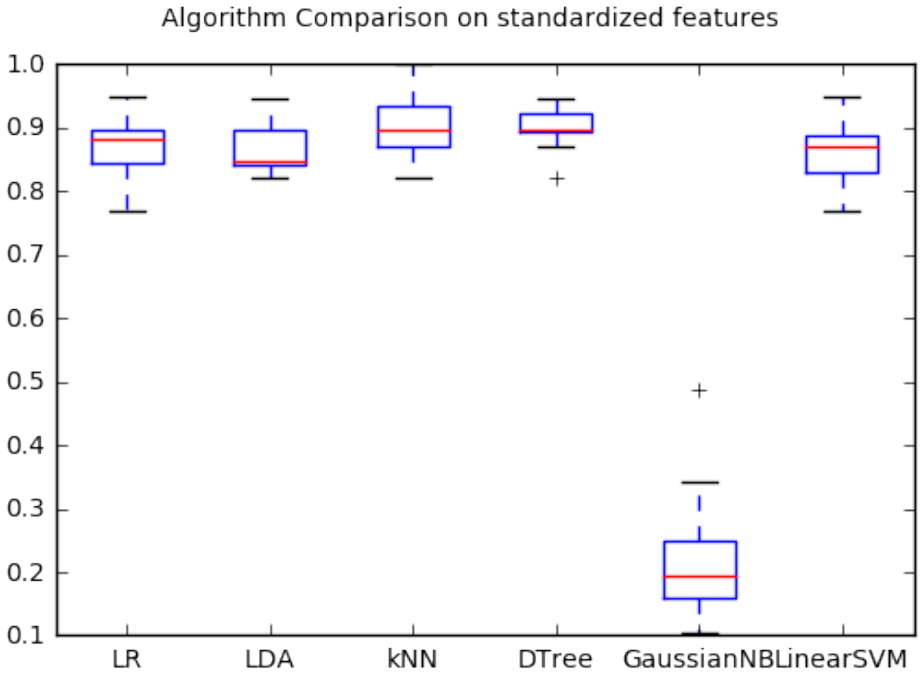
- Logistic Regression
- Linear Discriminant Analysis
- K Nearest Neighbours
- Decision Tree
- Gaussian Naive Bayes
- Support Vector Machine

Experiment 1&2

- Not / Standardized features
- Train and test in the same data
- Cross-validation and prediction

Experiment 1&2 – Raw and standardized features





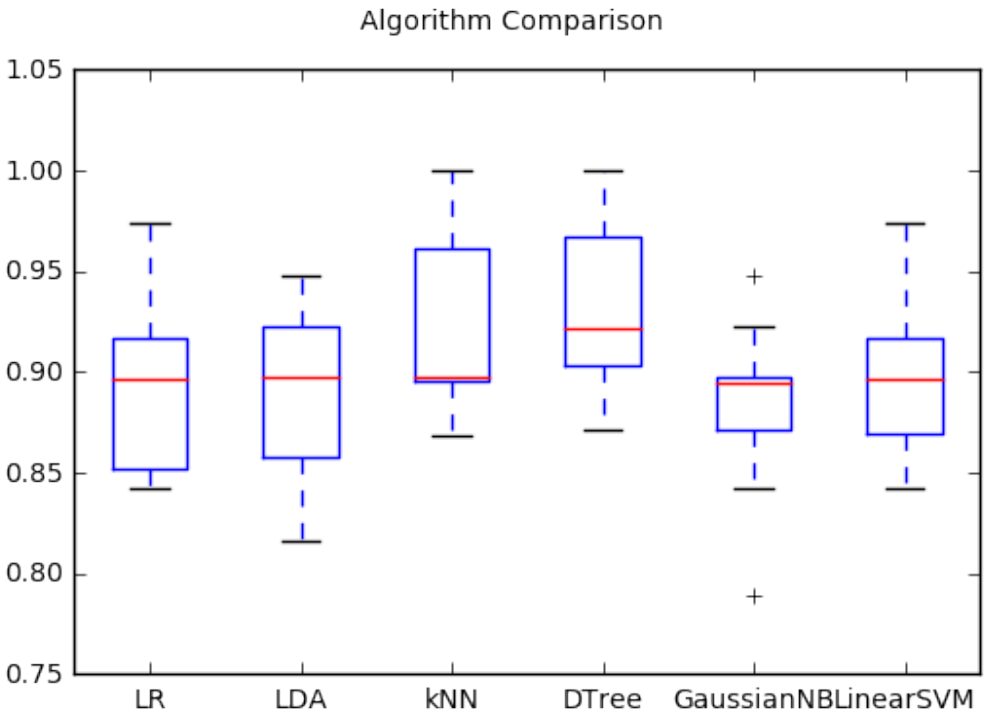
Synthetic data
LR prediction:
42.5%

Amsterdam data
DTree prediction:
88.6%

Experiment 4

- Standardized features
- Train and test in Amsterdam data
- Include RMSE
- Binary classes (LoD2 or not)

Experiment 4



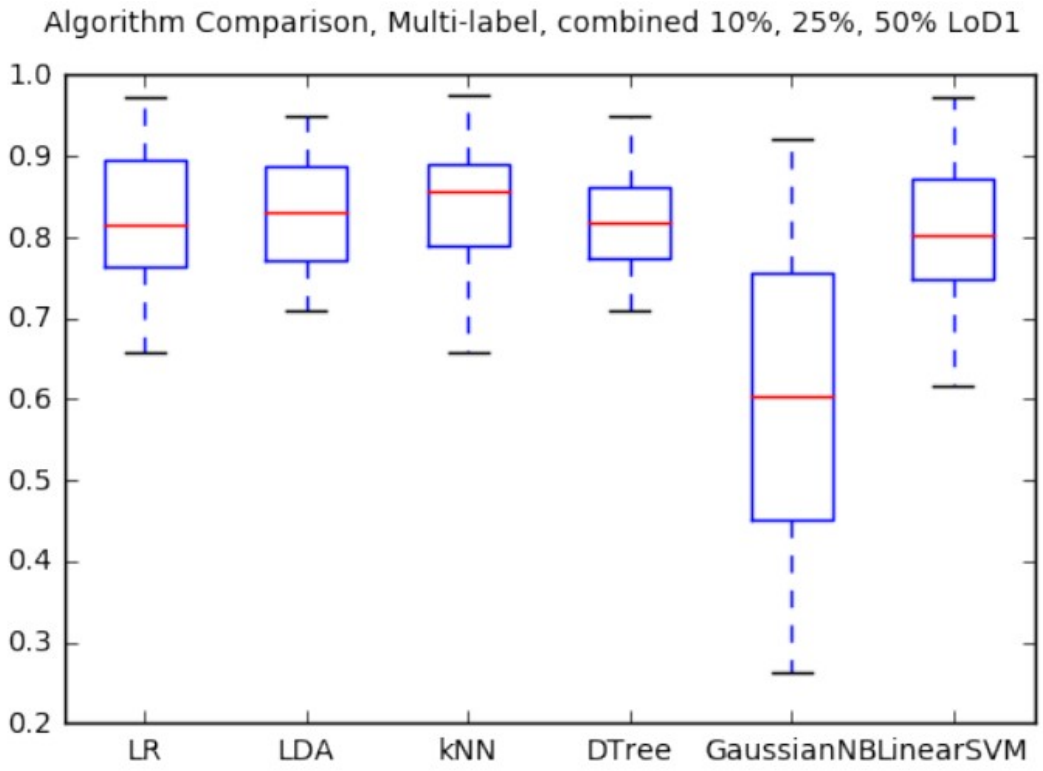
Dtree prediction 92.5% but:

	Not LoD2	LoD2
Not LoD2	7	5
LoD2	2	83

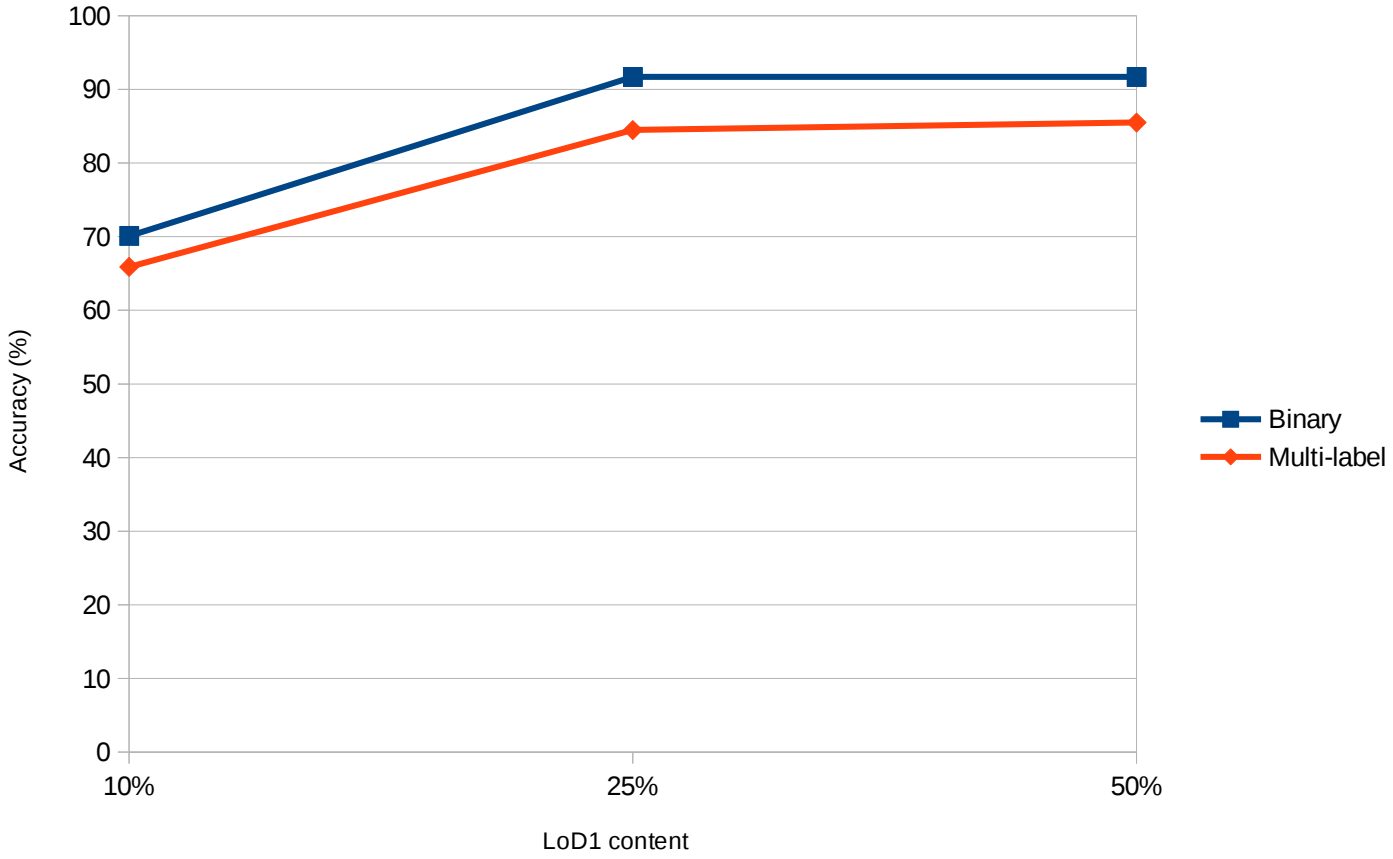
Experiment 5

- Standardized features
- Train and test in Amsterdam data
- Replace 10, 25, 50 of LoD2 with LoD1
- Include RMSE
- Multi-class and Binary classes (LoD2 or not)

Experiment 5 – mixed LoD1&2



Experiment 5 – kNN



Experiment 3&6

- Standardized features
- Train in synthetic and test in Amsterdam
- Replace 10, 25, 50 of LoD2 with LoD1
- Include RMSE
- Multi-class

Experiment 3&6

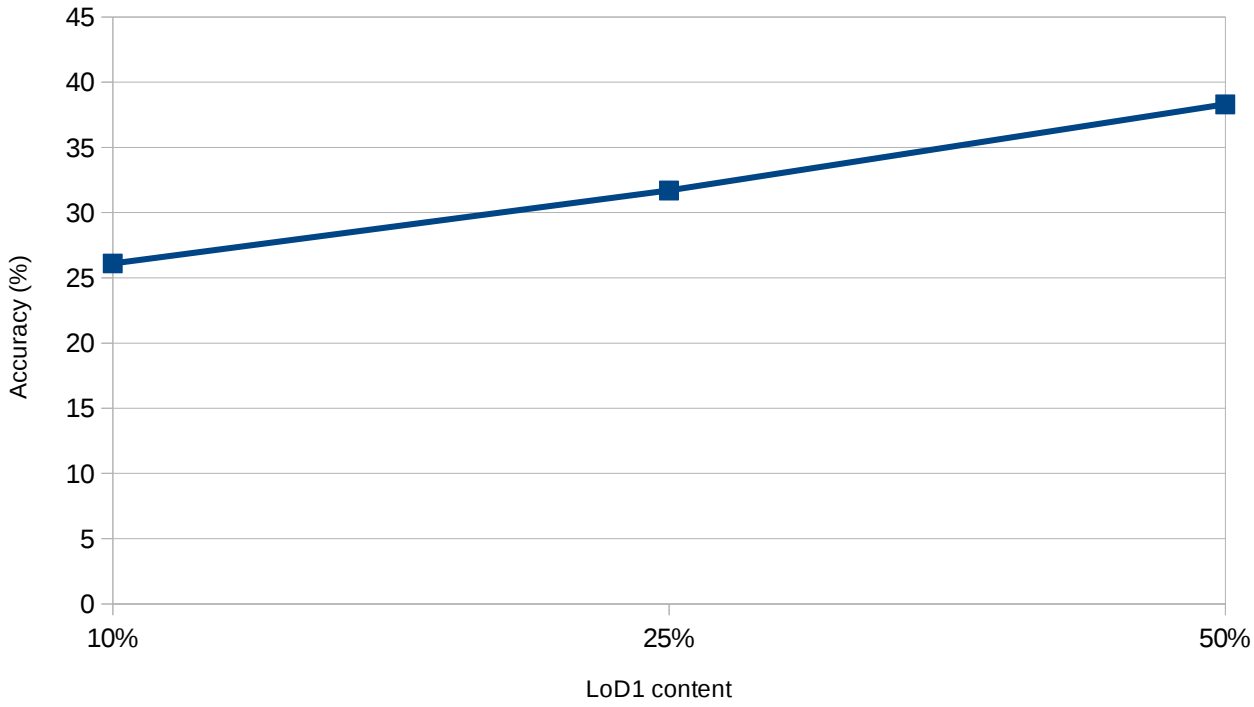
Experiment 3

With LoD0

LR	DTree	NB
7.4%	3.7%	0.0%

Experiment 6

No LoD0



Conclusions

- Synthetic data is not suitable as design set
  - Representative data set
- Features seem to be OK, but are there better?
- 42%, 88%, binary classes 92%
- Class imbalance is an open problem
- Issues with noisy point cloud, distances are not reliable
  - Other reference data?
  - RMSE might be too coarse
- LoD *inference* and *validation*

References

[1] Biljecki, F.; Heuvelink, G. B.; Ledoux, H. & Stoter, J. The effect of acquisition error and level of detail on the accuracy of spatial analyses Cartography and Geographic Information Science, Taylor & Francis, 2018, 45, 156-176

[2] Gröger, G.; Kolbe, T. H.; Nagel, C. & Häfele, K.-H. OGC City Geography Markup Language (CityGML) Encoding Standard, Version 2.0 Open Geospatial Consortium, Open Geospatial Consortium, 2012, 344

[3] Biljecki, F., Ledoux, H., Stoter, J., 2016. An improved LOD specification for 3D building models. Computers, Environment and Urban Systems 59, 25–37. doi:10.1016/j.compenvurbsys.2016.04.005

[4] Cignoni, P., Rocchini, C., Scopigno, R., 1998. Metro: Measuring Error on Simplified Surfaces. Computer Graphics Forum 17, 167–174. https://doi.org/10.1111/1467-8659.00236