**TUDelft**

Delft University of Technology

# Estimating building age with 3D GIS

Biljecki, Filip; Sindram, M.

**Important note**
To cite this publication, please use the final published version (if applicable).
Please check the document version above.

# ESTIMATING BUILDING AGE WITH 3D GIS

F. Biljecki [a,b*], M. Sindram [c]

[a] National University of Singapore, Singapore
[b] Delft University of Technology, Netherlands
[c] Technical University of Munich, Germany

**KEY WORDS:** 3D city models, building age, year of construction, CityGML, machine learning, random forest regression

**ABSTRACT:**

Building datasets (e.g. footprints in OpenStreetMap and 3D city models) are becoming increasingly available worldwide. However, the thematic (attribute) aspect is not always given attention, as many of such datasets are lacking in completeness of attributes. A prominent attribute of buildings is the year of construction, which is useful for some applications, but its availability may be scarce. This paper explores the potential of estimating the year of construction (or age) of buildings from other attributes using random forest regression. The developed method has a two-fold benefit: enriching datasets and quality control (verification of existing attributes). Experiments are carried out on a semantically rich LOD1 dataset of Rotterdam in the Netherlands using 9 attributes. The results are mixed: the accuracy in the estimation of building age depends on the available information used in the regression model. In the best scenario we have achieved predictions with an RMSE of 11 years, but in more realistic situations with limited knowledge about buildings the error is much larger (RMSE = 26 years). Hence the main conclusion of the paper is that inferring building age with 3D city models is possible to a certain extent because it reveals the approximate period of construction, but precise estimations remain a difficult task.

## 1. INTRODUCTION

Nowadays many places around the world are covered with GIS building data (2D footprints and 3D city models). While the completeness of building datasets is continuously improving, attributes are often lacking, and they have to either be integrated from other sources or estimated (Kunze and Hecht, 2015; Hecht et al., 2013; Barron et al., 2013). The issue is of both homogeneous and heterogeneous nature: (1) a certain attribute may not collected at all due to the data specification not requiring it (e.g. not uncommon in official data (Hecht et al., 2015)); and (2) the completeness is varying across a dataset, which is a case inherent to volunteered geoinformation (Camboim et al., 2015).

While most 3D GIS analyses rely primarily on the building geometry (e.g. line of sight analyses), there is a sizeable number of analyses requiring particular attributes (Biljecki et al., 2015). For example, energy demand estimation is such a case relying on several attributes about buildings (Nouvel et al., 2015). Furthermore, there are also some analyses in which attributes are not essential but they provide added value to certain application domains. As an example, population estimation is possible using only the geometry of a building. However, the knowledge of the type of buildings significantly improves the results because it enables filtering out non-residential structures (Ural et al., 2011; Biljecki et al., 2016a).

There have been a few studies investigating the thematic completeness of GIS datasets (Mobasheri et al., 2017). For example, Fan et al. (2014) analyse the attribute completeness of buildings in OpenStreetMap (OSM) in Germany. The study reveals that only 8.5% of buildings have the information on their type, and less than 0.1% the number of storeys. In the only study we could find that includes the year of construction of buildings, Agugiaro (2016b) suggests that even relatively populated datasets such as

the official dataset of Vienna (cf. 86% building use and 84% number of storeys), have lower levels of completeness (76%) of the year of construction. While OSM has a tag for noting the year of construction of a building (Goetz and Zipf, 2012), we did not find studies investigating its completeness in practice. A quick inspection of OSM data covering different spatial extents revealed that the completeness of the year of construction of a building is scarce in most cases. This is not surprising considering the challenges to collect such information.

The goal of this paper is to investigate whether it is possible to infer the age (i.e. year of construction) of buildings with 3D GIS and machine learning. This work has multiple benefits, primarily: (1) enriching datasets with this additional attribute (or supplementing datasets that have partial completeness of such information); and (2) applications in data quality, i.e. verifying the values of existing attributes, a topical subject since thematic quality of building data is not very good in practice (Agugiaro, 2016a). Because building age is a standard attribute in many building data specifications, this work has benefits for both 2D (building footprints) and 3D GIS (e.g. CityGML).

Our work builds on other papers dealing with classifying and estimating building characteristics (Section 2). To the extent of our knowledge, the estimation of year of construction of a building has not been investigated, and in Section 3 we present our method to infer it. We perform experiments on data representing buildings in Rotterdam in the Netherlands (Section 4).

Because the age of a building and its year of construction are both used in literature, in this paper we use both terms interchangeably. The year of construction is the attribute that is usually stored in databases, from which age is often calculated in order to use them in spatial analyses.

---

[*] Corresponding author at filip@nus.edu.sg

## 2. BACKGROUND

### 2.1 Building age / year of construction

The attribute is conceptually featured in many papers in the field (Isikdag and Zlatanova, 2009), and it is often collected by authorities of different countries (Dalmau et al., 2014; Nouvel et al., 2017). In the CityGML 2.0 standard of the Open Geospatial Consortium (2012) the year of construction is one of the key attributes for buildings and other features such as tunnels (Gröger and Plümer, 2012).

While the year of construction is regularly part of many different specifications, the attribute is in practice not always available, as the introduction of this paper suggests. A possible reason for such omission is that the year of construction of a building is difficult to procure, in comparison with other usual (and 'visually obvious') attributes such as building type, number of storeys, and roof type.

Furthermore, there are not many papers demonstrating the attribute's use and value in practice. For example, the information about the year of construction of a building is useful in applications such as urban planning (Dalmau et al., 2014). Most prominently, it is relied heavily on in energy demand estimation (Mastrucci et al., 2014; Nouvel et al., 2015; Agugiaro, 2016a; Krüger and Kolbe, 2012), as it may serve as a proxy for their energy efficiency. Some researchers also include the year of refurbishment in their simulations (Agugiaro, 2016a). Another possible application of the attribute of building age is in procedural modelling (Smelik et al., 2014), where the procedures may be context-aware in such a way that a certain architectural style typical for an era may be used for a building depending on its year of construction.

Besides infrequent literature, a hindrance is that the papers do not mention what is the acceptable level of quality of this attribute. Hence, the results that we obtain in this paper cannot be evaluated with respect to their usability in practice (e.g. it is uncertain whether an estimation error of 10 years would still be acceptable).

### 2.2 Related work

There have been many research papers in GIS and remote sensing on inferring additional knowledge of features, such as semantics of surfaces in 3D city models (Rook et al., 2016), roof type from 2D data (Henn et al., 2013), and building type (Neidhart and Sester, 2004; Meinel et al., 2009).

Among other research efforts, our paper is influenced by the following related research in 3D GIS. Henn et al. (2012) use supervised learning (support vector machines — SVM) automatically classifying the type of building from LOD1 models containing a few building attributes (e.g. number of storeys) and several attributes about the surrounding context (e.g. distance to nearest school). The year of construction is however not mentioned in the paper. Their work suggests that the building geometry (incl. height) gives a hint at the building type, thus we investigate in this paper whether age could also be successfully estimated in a similar fashion. A similar work was carried out by Hecht et al. (2015) using a Random Forest classifier and by integrating multiple data sources. In a regression problem, Biljecki et al. (2017) use several building attributes (e.g. type, age) for predicting their heights, in order to generate LOD1 models without elevation measurements. Some of our work is inversely based on this paper, because among

other attributes we use the building height to predict age. Important takeaways from the paper are: (1) building age is moderately associated to storey height (older buildings in general tend to have taller ceiling height; such a relation was also noticed by Kaden and Kolbe (2013)); and (2) the complexity of the shape of the 2D footprint may reveal additional information about the building.

## 3. METHODOLOGY

### 3.1 Random forest regression

In our work we use random forest (RF), which is an ensemble learning method that can be used for regression (Breiman, 2001). RF fits a large number of decision trees on subsets of the dataset. Advantages of RF are that it controls over-fitting and enables computing measures to evaluate the importance of different variables (i.e. feature importance) used for the predictions (Grömping, 2009).

For the implementation we use Python and Scikit-learn (Pedregosa et al., 2011).

### 3.2 Study area, data, and attributes

For the experiments we use a 3D city model generated from the cadastral (2D) dataset of the City of Rotterdam and the Dutch Kadaster obtained in 2015 (Figure 1), and a dense point cloud from the national elevation dataset of the Netherlands (AHN). The two datasets were combined with the open-source software 3dfier[1] creating an extruded LOD1 model. An illustration of 3D data generated from open resources of Dutch authorities is shown in Figure 2.
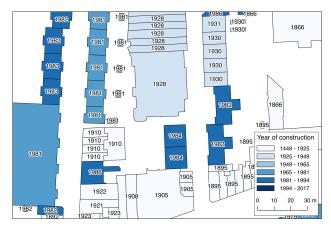


Figure 1. Building footprints with the values of the year of construction in the cadastral database of the City of Rotterdam and Dutch Kadaster.

For each building we have the following attributes: (0) *year of construction* (Figure 3 shows the distribution of values in Rotterdam), which we use for training and testing the regression model; (1) *building use* (residential, mixed, non-residential); (2) *building height* (which was calculated as the 90th percentile of the elevation of lidar points within the footprint, in order to avoid outliers and vegetation (Biljecki et al., 2016b)); (3) *footprint area*; (4) *number of storeys above ground*; and (5) *building volume*. We also calculate (6) the *ceiling height* by dividing the height of the

---

[1] https://github.com/tudelft3d/3dfier

Figure 2. 3D city model of Rotterdam (LOD1). The model was generated with 3dfier (open source software by TU Delft).

building with the number of storeys (see the discussion in Section 2); (7) *number of neighbouring buildings* (in a buffer of 30 m around each building, in accordance with related work discussed in Section 2); (8) *the complexity of the shape of the 2D footprint* (using a formula defined by Angel et al. (2010), which indicates the compactness of the shape by comparing the ratio of the perimeter of the equal-area circle with the perimeter of the shape in question); and following the evident observation in Figure 1 that in many cases adjacent buildings have similar year of construction we also calculate (9) *the average year of construction* of neighbouring buildings (using the same buffer of 30 m).
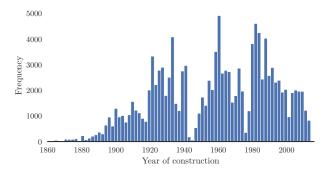


Figure 3. Distribution of buildings in Rotterdam by year of construction. The share of buildings constructed before 1860 is negligible.

Because RF is a supervised learning method, a training dataset is required. Therefore we split the dataset into two halves: a training (50%) and a test (50%) dataset.

### 3.3 Scenarios of completeness

As hinted at in the introduction, there are different situations of attribute completeness in practice. For simplicity, we describe three general situations. An attribute may: (1) not be available since it is not supposed to be collected or its availability may be scarce (it is supported but rarely acquired hence practically non-existent, e.g. number of storeys in OSM is an example from practice); (2) it might have heterogeneous completeness (e.g. it is available for many but not all features); and (3) it might reach full completeness (i.e. our dataset, which we use for training and testing the method). Such scenarios are important to consider when designing the predictive models, because in the case (2) it would be possible to use the attribute where available to enrich buildings where such is not available (e.g. calculating the average

age of buildings in the immediate neighbourhood and applying it to other buildings, which would not be possible in the situation (1) because of the lack of required information).

In parallel, because we rely on other attributes to infer the age of a building, we should also consider different scenarios of their completeness. For example, some building datasets might contain the building use, but not the number of floors. This is especially the case in OSM.

Hence, we design and train 8 different prediction models representing different scenarios of completeness in practice, and we evaluate their performance. The models will be given in the next section in a tabular form (Table 1), and here a rationale behind their design is explained. The model 0 takes advantage of the building height and number of floors to calculate the ceiling height, a potentially useful predictor of the building age as discussed in recent literature (Biljecki et al., 2017). Model 1 adds building use, while model 2 includes also the geometric attributes: footprint area, number of neighbours, volume, and shape complexity of the footprint. Model 3 relies only on the five geometric attributes that are always available in 3D city models, including in LOD1 models without semantics. The remaining models focus on the surrounding context of a building: the average age of other buildings in the buffer around the building since they tend to have equal or similar year of construction (Figure 1). In practice, such an attribute would not be always available: this prediction model would be applicable only in situations of heterogeneous completeness, being useful to improve the completeness of existing datasets, filling missing attributes of other buildings. Model 4 is therefore composed only of using the average year of construction (no 3D GIS operation is actually used here), while model 5 adds ceiling height, building height, and the number of floors. Model 6 further upgrades model 5 with building use. Finally, model 7 includes all 9 attributes. Model 7 is also useful to investigate how different variables interact with each other, and to determine what are the most useful attributes for predicting the year of construction of a building.

In the models 4–7 using data on the average year of construction of adjacent buildings, about 1.5% of buildings were not considered because they did not have neighbours that are close enough to consider them.

## 4. RESULTS AND DISCUSSION

### 4.1 General results

The errors are presented in both the mean absolute error (MAE) and root mean square error (RMSE) following the recommendation of Chai and Draxler (2014). Furthermore, in order to obtain a better understanding of the accuracy, for each model the percentile rank of a score was calculated. That is, the share of predictions that is below a certain value, in our case 15 years (this quantity was chosen arbitrarily). For example, the value 60 indicates that 60% of estimations have an error smaller or equal to 15 years.

The performance of the eight considered models is presented in Table 1. The RMSE errors range from 14 to 28 years, while the MAE errors are from 6 to 20 years. The percentile rank values vary from 53% to 91%, meaning that in the worst performing model a bit more than half of the estimations have an estimation error within 15 years. The distribution of errors of the eight

prediction models is also depicted in Figure 4, giving a visual impression of their performance.

The different statistical models give different insights in what can be achieved by having a certain combination of building attributes. In the continuation of the section we will elaborate on the performance of each model.
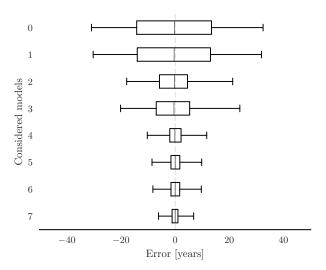


Figure 4. Distribution of errors for all models as box plots. The boxes span the interquartile range (the middle 50%). The line inside the rectangle indicates the median, while the whiskers cover 80% of the errors (10th and 90th percentile).

Model 0, targeted at taking advantage of the association of the ceiling height and building age, without considering other attributes, performed poorly with a mean absolute error of two decades. Adding building use (model 1) improved the results only marginally. While such an accuracy might still be acceptable for some users, it should be noted that the errors span a large range that in practice is not far from the lifespan of most buildings (Figure 4). The results much improved by including the number of floors and geometric attributes: the MAE dropped to 11.6 years (model 2).

Model 3 considers a basic set of attributes which, being derived from the geometry, are always available in LOD1 models. The results are similar to those of the previous model (model 2), which contains the building use and number of storeys. This means that the added set of semantics helps, but in this particular combination their value is not paramount.

The remaining models are suited for datasets that have partial completeness, in which the average age of the neighbouring buildings may be available. Only by considering the age of surrounding buildings the results are much better than the previous four models: the MAE drops to 7 years (model 4). Combining this attribute with the attributes from model 0 yields predictions with a mean absolute error of 6 years (model 5), while adding building use (model 6) brings a minuscule improvement. Finally, in model 7 all available attributes are combined. The MAE of the estimations is less than 5 years, which would probably be acceptable in applications such as energy demand estimation. However, it should be noted that such an estimation model requires a dataset that is rich in semantics. The final model is interesting also because it enables examining the relative contributions of the variables (explained in the next section).

## 4.2 Feature importance

Table 2 overviews the importance of each variable in each model. The two most important variables in each model are emphasised in the table.

Ceiling height appears to be valuable for such predictions, since in most models where it is used it was found to be among the two most important features. However, a disadvantage of this attribute is that it requires the information on the number of storeys, which is not readily available, especially not in volunteered geoinformation. Furthermore, calculating ceiling height is subject to errors and inconsistencies, as a fair share of buildings does not have a constant ceiling height (e.g. the height of the lobby of a hotel may be larger than the height of upper storeys).

In model 3, which features only geometric attributes, the building height and shape complexity have the highest importance. The latter is dependent on the architecture of the building. Hence a possible explanation is that the geometry of the building footprint is characterised by a certain period of construction of a building, which in turn indicates its approximate age.

Models 4–7 involve the knowledge of the age of adjacent buildings, which has the predominant feature importance. When such information is available, all other attributes have marginal feature importance.

## 4.3 A closer look on model 2

Because of the large number of models and results, we rather consider one model in more detail. Model 2 is selected as a representative model for a closer inspection because it has an average performance, but also because it is the model that would likely be used in practice: it contains a standard set of attributes in datasets that do not have information on the year of construction of buildings.

Figure 5 shows the histogram of errors of the estimations. The errors have a mean around zero, and while the mean absolute error is 11.6 years (Table 1), there is a share of predictions with errors larger than 40 years.
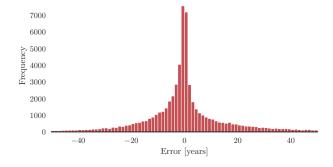


Figure 5. Histogram of errors for model 2.

Figure 6 offers more insight related to the percentile rank of score presented in Table 1. For example, almost half (48.6%) of the estimations are within 5 years, while 85.5% of the estimations are within a quarter of century. Such performance might still be acceptable for applications in which only the approximate period of construction is needed. However, because no paper (to the extent of our knowledge) discusses the required quality for the

| Attribute(s) | Models | | | | | | | |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| | 0 | 1 | 2 | 3 | 4 | 5 | 6 | 7 |
| Ceiling height | ● | ● | ● | | | ● | ● | ● |
| Building height | ● | ● | ● | ● | | ● | ● | ● |
| Number of storeys | ● | ● | ● | | | ● | ● | ● |
| Building use | | ● | ● | | | | ● | ● |
| Footprint area | | | ● | ● | | | | ● |
| No. neighbours | | | ● | ● | | | | ● |
| Volume | | | ● | ● | | | | ● |
| Shape complexity | | | ● | ● | | | | ● |
| Avg. neigh. age | | | | | ● | ● | ● | ● |
| RMSE [years] | 26.5 | 26.3 | 19.4 | 21.0 | 14.6 | 12.9 | 12.8 | 11.3 |
| MAE [years] | 19.4 | 19.1 | 11.6 | 12.9 | 7.3 | 6.1 | 6.0 | 4.9 |
| Percentile rank (score = 15 y) | 52.9 | 53.6 | 73.8 | 70.6 | 85.4 | 87.9 | 88.2 | 91.0 |

Table 1. Combinations of attributes and the performance (accuracy) of the predictive models.

| Attribute(s) | Feature importance per model/attribute | | | | | | | |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| | 0 | 1 | 2 | 3 | 4 | 5 | 6 | 7 |
| Ceiling height | **0.63** | **0.59** | **0.32** | | | 0.05 | 0.05 | **0.03** |
| Building height | **0.36** | **0.35** | **0.17** | **0.38** | | **0.05** | **0.05** | 0.02 |
| Number of storeys | 0.01 | 0.01 | 0.01 | | | 0.01 | 0.01 | 0.00 |
| Building use | | 0.05 | 0.01 | | | | 0.01 | 0.00 |
| Footprint area | | | 0.13 | 0.18 | | | | 0.03 |
| No. neighbours | | | 0.09 | 0.11 | | | | 0.02 |
| Volume | | | 0.11 | 0.12 | | | | 0.02 |
| Shape complexity | | | 0.16 | **0.20** | | | | 0.02 |
| Avg. neigh. age | | | | | **1.00** | 0.89 | 0.89 | 0.85 |

Table 2. Feature importance in the eight considered scenarios. The two most important variables in each model are emphasised.
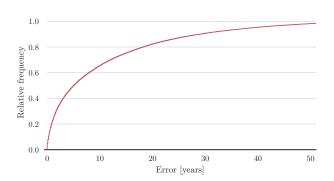


Figure 6. Cumulative histogram of estimation errors for model 2.

year of construction we can only speculate about the usability of values with a specific degree of uncertainty.

Figure 7 illustrates the errors decomposed by decade of construction. The classified boxplots reveal that the predictions exhibit different behaviour depending on the age of the building, in which the age of older buildings (i.e. from the 19th century) tends to be overestimated. The overestimations are large. However, considering that the share of such buildings is low (see the right side of the plot, or Figure 3), such an error does not significantly affect the overall results (see the rather narrow joint box plot at the bottom). Such a finding is also observable in the observed vs. estimated scatter plot given in Figure 8. In future work it would be beneficial to investigate whether this behaviour can be compensated reducing the errors.

Thanks to the availability of the attribute on building use, we have also classified errors depending on it (Figure 9). It appears that residential buildings have better estimates, indicating that building type plays a role in the estimations.

### 4.4 Limitations

A limitation of our work is that in order to derive building attributes (in this case age), we need other building attributes, which in practice may suffer from the same completeness issues (e.g. according to the German study of Fan et al. (2014) only 0.06% of buildings in OSM contain the information on the number of floors). For that reason we have evaluated different scenarios of availability of predictors. One of these scenarios (model 3) uses only features that are always available in LOD1 models because they are derived from the geometry and no attributes are used.

Another limitation is also related to completeness: supervised learning requires training data including the year of construction of a building. Training is inhibited by poor completeness, as it might be not possible to obtain data in the area where the predictions are carried out. Hence it might be required to train the models in one area and apply them to another area, which may exhibit different relationships. However, we did not run experiments using data from other spatial extents, thus we cannot hypothesise on the performance of the method in another city or country. It would be interesting to investigate this aspect in future work.
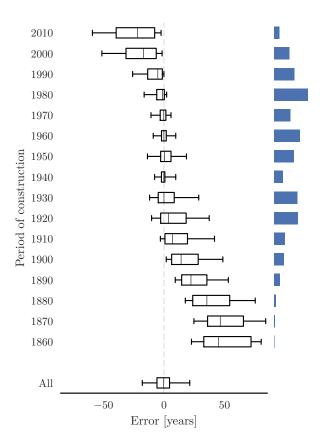
Figure 7. Errors in model 2 classified by period (decade) of construction with the indication of the quantity of the relative sample size (histogram on the right).
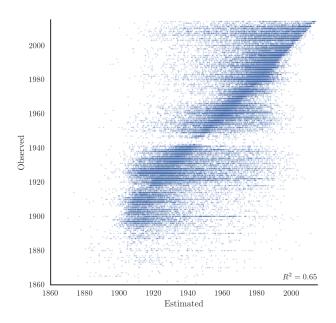


Figure 8. Observed vs. estimated plot (model 2) of the year of construction of a building.

## 5. CONCLUSIONS AND FUTURE WORK

While it is one of the key building attributes in standards such as CityGML, the year of construction of a building has not been
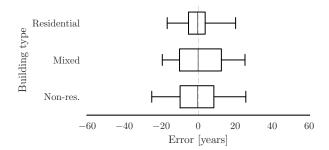


Figure 9. Errors by type of building in model 2. The decomposed errors reveal varying accuracy depending on the type of the building.

given much attention in practice and research. In fact, a literature review did not reveal much use of the attribute (with the exception of energy demand estimation), and mentions of the attribute are mostly found in papers related to standardisation. One possible reason for indifference towards this attribute is that even in governmental datasets in many developed countries the year of construction is either unavailable or only partially available. The cause for the frequent omission of the year of construction of a building is that it is difficult to obtain it in comparison to other attributes such as building use or the number of storeys.

In this paper we made an attempt to fill such void by developing a method that would automatically enrich 2D and 3D building datasets with the value of the building year of construction. We implemented and trained a random forest regression model to infer the building age from a set of geometric and non-geometric attributes commonly and some of them less frequently found in 3D city models. We hope that this paper will raise awareness of such omission, and that it will help towards the increased availability of this attribute.

While most of the results that we have achieved are not very accurate, at least they approximately indicate the period (i.e. decade) when the building was constructed. Some of the prediction models did yield a fairly good estimation of the year of construction of a building (e.g. model 7 predicted 91% with an error smaller than 15 years), but they also exhibit a certain share of gross errors (Figure 5).

Another interesting finding is that while attributes derived from 3D city models may help in determining the age of a building, the knowledge of the age of neighbouring buildings is by far the most useful indicator because adjacent buildings usually are constructed in the same period. However, that information is rarely available.

In this paper we have used LOD1 models. Perhaps in future we would obtain better results by using more detailed datasets. For example, LOD2 models contain a generalised roof structure, as opposed to flat top surfaces intrinsic to LOD1 models. The shape of the roof would be known in such a case potentially providing additional cue, and also a more accurate height of the building could be obtained. Considering that in our method we have used a percentile value of the lidar point cloud, the height of the building may be prone to errors not always indicating the top of the roof (ridge height). Such an error may propagate to an error in determining the ceiling height, a key piece of information for the prediction of the building age, as we have demonstrated in Section 4. With even more detailed models (LOD3), it might be

possible to analyse patterns in the geometry identifying the architectural style of a building (see the paper of Dehbi et al. (2016) for related work in 2D), which would perhaps hint at its period of construction.

Furthermore, an interesting variable to investigate would be the energy consumption. Building age is associated with energy efficiency (Nouvel et al., 2015; Kaden and Kolbe, 2013), thus it might be useful to investigate whether the data on energy consumption may be used to improve the predictions. Finally, it would be interesting to research whether the age could be inferred with remote sensing techniques (e.g. analysing roof reflections).

## ACKNOWLEDGEMENTS

## References

Agugiaro, G., 2016a. Energy planning tools and CityGML-based 3D virtual city models: experiences from Trento (Italy). *Applied Geomatics* 8(1), pp. 41–56.

Agugiaro, G., 2016b. First steps towards an integrated CityGML-based 3D model of Vienna. *ISPRS Ann. Photogramm. Remote Sens. Spatial Inf. Sci.* III-4, pp. 139–146.

Angel, S., Parent, J. and Civco, D. L., 2010. Ten compactness properties of circles: measuring shape in geography. *The Canadian Geographer* 54(4), pp. 441–461.

Barron, C., Neis, P. and Zipf, A., 2013. A Comprehensive Framework for Intrinsic OpenStreetMap Quality Analysis. *Transactions in GIS* 18(6), pp. 877–895.

Biljecki, F., Arroyo Ohori, K., Ledoux, H., Peters, R. and Stoter, J., 2016a. Population Estimation Using a 3D City Model: A Multi-Scale Country-Wide Study in the Netherlands. *PLOS ONE* 11(6), pp. e0156808.

Biljecki, F., Ledoux, H. and Stoter, J., 2017. Generating 3D city models without elevation data. *Computers, Environment and Urban Systems* 64, pp. 1–18.

Biljecki, F., Ledoux, H., Stoter, J. and Vosselman, G., 2016b. The variants of an LOD of a 3D building model and their influence on spatial analyses. *ISPRS Journal of Photogrammetry and Remote Sensing* 116, pp. 42–54.

Biljecki, F., Stoter, J., Ledoux, H., Zlatanova, S. and Çöltekin, A., 2015. Applications of 3D City Models: State of the Art Review. *ISPRS International Journal of Geo-Information* 4(4), pp. 2842–2889.

Breiman, L., 2001. Random Forests. *Machine Learning* 45(1), pp. 5–32.

Camboim, S., Bravo, J. and Sluter, C., 2015. An Investigation into the Completeness of, and the Updates to, OpenStreetMap Data in a Heterogeneous Area in Brazil. *ISPRS International Journal of Geo-Information* 4(3), pp. 1366–1388.

Chai, T. and Draxler, R. R., 2014. Root mean square error (RMSE) or mean absolute error (MAE)? – Arguments against avoiding RMSE in the literature. *Geoscientific Model Development* 7(3), pp. 1247–1250.

Dalmau, F., Garcia-Almirall, P., Domínguez, E. and Escudero, D., 2014. From Raw Data to Meaningful Information: A Representational Approach to Cadastral Databases in Relation to Urban Planning. *Future Internet* 6(4), pp. 612–639.

Dehbi, Y., Gröger, G. and Plümer, L., 2016. Identification and Modelling of Translational and Axial Symmetries and their Hierarchical Structures in Building Footprints by Formal Grammars. *Transactions in GIS* 20(5), pp. 645–663.

Fan, H., Zipf, A., Fu, Q. and Neis, P., 2014. Quality assessment for building footprints data on OpenStreetMap. *International Journal of Geographical Information Science* 28(4), pp. 700–719.

Goetz, M. and Zipf, A., 2012. Towards Defining a Framework for the Automatic Derivation of 3D CityGML Models from Volunteered Geographic Information. *International Journal of 3-D Information Modeling* 1(2), pp. 1–16.

Gröger, G. and Plümer, L., 2012. CityGML – Interoperable semantic 3D city models. *ISPRS Journal of Photogrammetry and Remote Sensing* 71, pp. 12–33.

Grömping, U., 2009. Variable Importance Assessment in Regression: Linear Regression versus Random Forest. *The American Statistician* 63(4), pp. 308–319.

Hecht, R., Kunze, C. and Hahmann, S., 2013. Measuring Completeness of Building Footprints in OpenStreetMap over Space and Time. *ISPRS International Journal of Geo-Information* 2(4), pp. 1066–1091.

Hecht, R., Meinel, G. and Buchroithner, M., 2015. Automatic identification of building types based on topographic databases – a comparison of different data sources. *International Journal of Cartography* 1(1), pp. 18–31.

Henn, A., Gröger, G., Stroh, V. and Plümer, L., 2013. Model driven reconstruction of roofs from sparse LIDAR point clouds. *ISPRS Journal of Photogrammetry and Remote Sensing* 76, pp. 17–29.

Henn, A., Römer, C., Gröger, G. and Plümer, L., 2012. Automatic classification of building types in 3D city models. *Geoinformatica* 16(2), pp. 281–306.

Isikdag, U. and Zlatanova, S., 2009. Towards Defining a Framework for Automatic Generation of Buildings in CityGML Using Building Information Models. In: *3D Geo-Information Sciences*, Springer Berlin Heidelberg, pp. 79–96.

Kaden, R. and Kolbe, T. H., 2013. City-wide total energy demand estimation of buildings using semantic 3D city models and statistical data. *ISPRS Ann. Photogramm. Remote Sens. Spatial Inf. Sci.* II-2/W1, pp. 163–171.

Krüger, A. and Kolbe, T. H., 2012. Building analysis for urban energy planning using key indicators on virtual 3D city models—The energy atlas of Berlin. *Int. Arch. Photogramm. Remote Sens. Spatial Inf. Sci.* XXXIX-B2, pp. 145–150.

Kunze, C. and Hecht, R., 2015. Semantic enrichment of building data with volunteered geographic information to improve mappings of dwelling units and population. *Computers, Environment and Urban Systems* 53, pp. 4–18.

Mastrucci, A., Baume, O., Stazi, F. and Leopold, U., 2014. Estimating energy savings for the residential building stock of an entire city: A GIS-based statistical downscaling approach applied to Rotterdam. *Energy and Buildings* 75, pp. 358–367.

Meinel, G., Hecht, R. and Herold, H., 2009. Analyzing building stock using topographic maps and GIS. *Building Research & Information* 37(5-6), pp. 468–482.

Mobasheri, A., Sun, Y., Loos, L. and Ali, A. L., 2017. Are Crowdsourced Datasets Suitable for Specialized Routing Services? Case Study of OpenStreetMap for Routing of People with Limited Mobility. *Sustainability* 9(6), pp. 997.

Neidhart, H. and Sester, M., 2004. Identifying building types and building clusters using 3-D laser scanning and GIS-data. *Int. Arch. Photogramm. Remote Sens. Spatial Inf. Sci.* XXXV/B4, pp. 715–720.

Nouvel, R., Mastrucci, A., Leopold, U., Baume, O., Coors, V. and Eicker, U., 2015. Combining GIS-based statistical and engineering urban heat consumption models: Towards a new framework for multi-scale policy support. *Energy and Buildings* 107, pp. 204–212.

Nouvel, R., Zirak, M., Coors, V. and Eicker, U., 2017. The influence of data quality on urban heating demand modeling using 3D city models. *Computers, Environment and Urban Systems* 64, pp. 68–80.

Open Geospatial Consortium, 2012. OGC City Geography Markup Language (CityGML) Encoding Standard 2.0.0. Technical report.

Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., Blondel, M., Prettenhofer, P., Weiss, R., Dubourg, V., Vanderplas, J., Passos, A., Cournapeau, D., Brucher, M., Perrot, M. and Duchesnay, É., 2011. Scikit-learn: Machine Learning in Python. *Journal of Machine Learning Research* 12(Oct), pp. 2825–2830.

Rook, M., Biljecki, F. and Diakité, A. A., 2016. Towards automatic semantic labelling of 3D city models. *ISPRS Ann. Photogramm. Remote Sens. Spatial Inf. Sci.* IV-2/W1, pp. 23–30.

Smelik, R. M., Tutenel, T., Bidarra, R. and Benes, B., 2014. A Survey on Procedural Modelling for Virtual Worlds. *Computer Graphics Forum* 33(6), pp. 31–50.

Ural, S., Hussain, E. and Shan, J., 2011. Building population mapping with aerial imagery and GIS data. *International Journal of Applied Earth Observation and Geoinformation* 13(6), pp. 841–852.