European Residential Rebuilding Costs Index, a First Approach

Rients VAN WIJNGAARDEN, Michiel JELLEMA, Marien DE BAKKER, the Netherlands

Key words: Rebuilding cost, Econometric model, Machine Learning, Real estate, LiDAR

SUMMARY

Accurate estimation of the rebuilding costs of is essential in the property insurance process both for the customer and the insurance company. For the Netherlands Infofolio has developed an objective, unambiguous and robust econometric model to estimate the rebuilding cost of all home. With the growing availability of data in Europe, both open and commercial, the question arises whether it is possible to construct, renew and publish a European rebuilding cost index using the Netherlands as a baseline to estimate the rebuilding cost of homes across Europe.

Internet research shows that many datasets containing geo-spatial and administrative data exist for Germany, the UK and Sweden, however, most of these data sources cover either a specific area or a specific time period. The combination of 3D laser scanning (LiDAR) data and building and/or cadastre data is most promising as it is widely available, fairly uniform and can be used to construct detailed 3D models of homes. By combining features extracted from these 3D models and available European construction cost data a European rebuilding cost index can be created.

Further research is needed to determine whether the proposed approach to collecting data and constructing a European rebuilding cost index yield sufficiently accurate estimations of rebuilding costs. We propose to apply the proposed method to four middle-sized cities in Europe and review and assess the results with international insurance partners.

European Residential Rebuilding Costs Index, a First Approach

Rients VAN WIJNGAARDEN, Michiel JELLEMA, Marien DE BAKKER, the Netherlands

1. INTRODUCTION

Property insurance is a major part of the European insurance industry, accounting for 105 billion euros in premiums in 2018 (Insurance Europe, 2019). An accurate estimation of the rebuilding cost of a home, i.e. the cost of completely rebuilding a home if it is destroyed beyond repair, is an integral part of the insurance policy. On the one hand insurers need a reliable and up-to-date estimation of the rebuilding cost to make sure consumers are paying the proper premium as property insurance is often either mandatory or required to apply for a mortgage. On the other hand, Solvency II, a risk-based compliance framework put in place in 2016, mandates reporting on over- and underinsurance and the size, spread and cumulation of risk, among other things (Jellema, 2017a). This reporting cannot be done properly without an accurate estimation of rebuilding costs.

Infofolio delivers digital information products for the insurance and has developed an objective, unambiguous and robust econometric model to estimate the rebuilding cost for all homes in the Netherlands (Jellema, 2014). The output of this model is widely used by insurers, representatives and intermediaries and enables them to streamline the offering process, periodical checks and renewal of policies, risk management and reinsurance. The model is developed in collaboration with market parties using reference data of representative houses and is constantly updated using the information gathered in the Infofolio data platform, a nationwide, high quality database with detailed information about the build environment in the Netherlands. The Infofolio data platform is based on data from national, provincial and municipal authorities and market parties. The model uses the physical dimensions of the building as well as architectural, utilization, economic, demographic and geographical dimensions.

Although similar models for other countries exist (Corlier, 1986), given the level of accuracy, objectiveness and market acceptance of the Dutch model a logical next step is to try to apply this model in other European countries. However, in earlier research by Infofolio it has become clear that despite the goals set forth in the Inspire Directive (European Commission - Joint Research Centre, 2019) regional and national differences in the availability and quality of spatial data make a common approach to gathering, assessing and applying this data currently infeasible. Furthermore, reference data as used in the development of this model is limited to the Netherlands and obtaining accurate, correct and up-to-date reference sets outside of the Netherlands would be very time consuming and labour intensive.

For the reasons laid out above we attempt to assess what data is available outside of the Netherlands and how this data can be collected, validated and enriched regionally, based on

availability. Furthermore, we will formulate an approach to build a first model for estimating rebuilding costs based on the minimal set of parameters available and assess whether data on European construction costs is available which we can combine with this model to create a European rebuilding cost index, using the Netherlands as reference. This European rebuilding cost index can be used to correct for relative differences in rebuilding costs between countries and the index will be updated and published periodically on https://www.infofolio.eu. Ultimately the European rebuilding cost index will enable us to estimate rebuilding costs for homes in Europe as long as the appropriate data is available.

2. METHOD

The availability of potentially useful geospatial data is assessed through internet research. In this research we take quality aspects like consistency, completeness, timeliness, accuracy and precision into account and both public and commercial sources are assessed (Jellema, 2017b). In order to keep the research concise and focused we limit ourselves to assessing data availability in three European countries apart from the Netherlands. Germany is chosen as it is the most populous country within the EU and recently passed a law mandating that research data is made publicly available (European Data Portal, 2017). The UK is chosen because a number of insurers working together with Infofolio are active in both the Netherlands and the UK. Finally, Sweden is chosen as it has a good track record regarding open government data and is one of the most innovative countries in the world (Brouwers, 2019).

For the three chosen countries data is gathered and assessed in three distinct steps (Hox, 2005). First potentially useful data sources are located and ownership and data format are determined. Next the accessibility of the potential data sources is assessed by determining whether the data is freely available or behind a paywall. Third the quality of the potential data sources is determined based on currentness, precision, and completeness (Veregin, 1999). Currentness refers to the degree to which a data source is up-to-date. Precision refers to the level of detail within the data source and is often referred to as resolution in the context of geospatial data. Completeness in the context of this paper refers feature completeness i.e. to the extent to which the features available defined in a data source are filled.

After samples from the data sources found to be available and of sufficient quality for our purpose are collected an additional assessment is done. The goal of this additional assessment is to establish whether the collected data sources contain enough features to enable the development of a new model to estimate rebuilding costs. Based on our experience with our current model, physical dimensions of the building as well as utilization and geographical dimensions are minimum requirements to develop a feasible model.

Finally, we assess whether data on European construction costs is available and sufficiently detailed to combine with the proposed model to create a European rebuilding cost index.

3. RESULTS

We found that in the three countries included in our search many different data sources are available, both public and commercial. However, most of these data sources cover either a specific area or a specific time period, making them unsuitable for our purpose. In some cases, complete datasets containing the required features were available commercially but prohibitively expensive again making them unsuitable for our purpose.

The most uniformly and easily available type of data turned out to be 3D laser scanning (LiDAR) data in the form of point clouds. Currently point cloud data for 72% of England is available while in Germany almost the entire country is covered. The coverage in Sweden is currently much lower, however the entire country will be covered by 2022. More importantly LiDAR data is also available for the entire Netherlands which will provide a reliable baseline in the development of the simplified rebuilding cost model. We were also able to determine that building and cadastre data is available in all countries included in our search.

LiDAR data can be used to construct 3D models of buildings (Gooding, 2015, Wu, 2019). In fact for many countries this has already been done and the results are available. However, these are models with a level of detail (LOD) of 1 (see figure 1), not detailed enough to accurately calculate volumes and floorspaces. In order to get to the level of detail needed we combined the LiDAR point clouds with building polygon data from national registers and used available tooling to extract roof forms. In this way we ended up with detailed 3D models of buildings which enables us to calculate floor space, volume and determine the type of home. Since the LiDAR data is combined with building and cadastre data, addresses and locations are either available or can be added by geocoding and intersecting separate address files. The results of this approach are shown in figure 2.

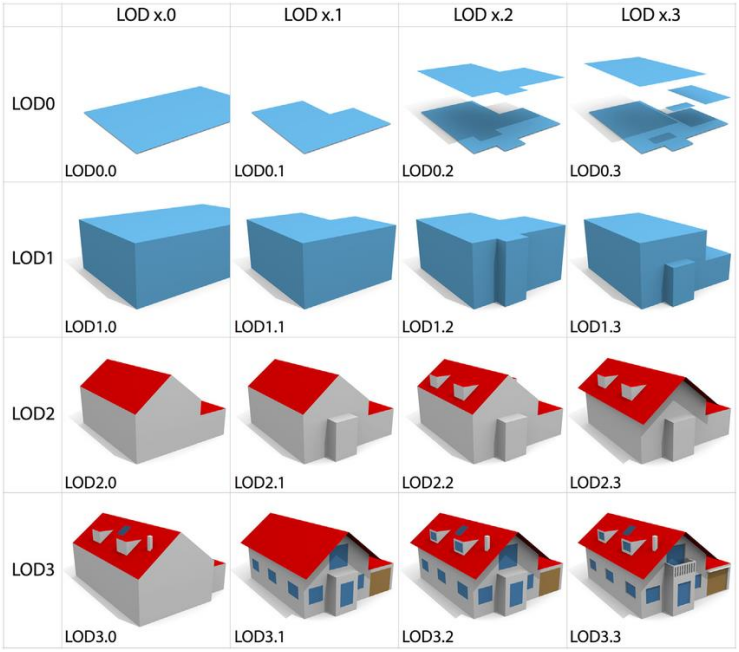

Figure 1: An improved LOD specification for 3D building models (Biljecki, 2016).

Figure 2: 3D model of the municipality of Zeist in the Netherlands

Finally, several sources were found with detailed information regarding construction costs in several European countries (e.g. http://constructioncosts.eu/, Eurostat, 2020), including the countries covered in this research. By using Econometric and Machine Learning techniques these sources can be combined with the proposed model to calculate relative differences between countries in the form of an index.

4. DISCUSSION

Regional and national differences in the availability and quality of spatial data make a uniform approach to gathering, assessing and applying this data similar to the approach currently used by Infofolio in the Netherlands unfeasible at this time. However, with the methods and results described in this paper we have shown that we are able to gather data in such a way that the development of a first model for estimating rebuilding costs based on internationally available data is possible. We have also shown that sufficient data is available to develop and periodically publish a European rebuilding cost index. Given these results it seems feasible to ultimately calculate a rebuilding cost for the majority of houses is Europe.

However, there are still some hurdles. First of all, the LiDAR data used in the construction of 3D models is not currently available for every country or region, limiting the coverage of our proposed methodology. Second, we do not know if we will be able to roll out our proposed methodology on scale, especially considering that the constructing of 3D models of buildings is computationally very intensive. Third, we do not know if the accuracy and explanatory power of the proposed model for calculating rebuilding costs will be sufficient. Last of all it is an open question whether the proposed European rebuilding cost index will be detailed enough to calculate relative differences between countries and regions within countries. In order to address these hurdles further research is required.

5. FUTURE STEPS

In this paper we have laid out a method for calculating rebuilding costs for houses in Europe and we have shown that this method is feasible. In order to further develop and test this method, the following steps are proposed. First, we will gather data for four average sized cities in Europe using the method described in this paper. Then we will develop the proposed model for calculating rebuilding costs using the features available in the collected datasets. Next, the developed model will be tested for accuracy and explanatory power using the Netherlands as a baseline. The developed model will then be combined with European construction cost data to create a detailed European rebuilding cost index expressing relative differences in rebuilding costs between countries and/or regions. Finally, the European rebuilding cost index will be used to calculate a preliminary rebuilding cost for every house in the previously gathered data. These results will be reviewed and assessed in collaboration with international insurance partners.

REFERENCES

- Biljecki, F. et al, 2016, An improved LOD specification for 3D building models, Computers, Environment and Urban Systems, volume 59, pp 25-37, Amsterdam, Elsevier
- Brouwers, B., 2019, Europa domineert de lijst van meest innovatieve landen: 12 landen in top-20, Nederland zakt naar plek 4, https://innovationorigins.com/
- Corlier, F. et al, 1986, A Model to Evaluate the Rebuilding Costs of Dwellings, Astin Bulletin, Volume 16 No. 1, pp 45-62, Cambridge, Cambridge University Press
- European Commission - Joint Research Centre, 2019, INSPIRE Geoportal, https://inspire.ec.europa.eu/
- European Constuction Costs, http://constructioncosts.eu/
- European Data Portal, 2017, New Open Data Act in Germany - European Data Portal - European Data Portal, https://www.europeandataportal.eu/en/news/new-open-data-act-german
- Eurostat, 2020, Construction cost (or producer prices), new residential buildings - quarterly data, https://ec.europa.eu/eurostat
- Gooding, J. et al, 2015, Modelling of roof geometries from low-resolution LiDAR data for city-scale solar energy applications using a neighbouring buildings method, Applied Energy, volume 148 issue C, pp 93-104, Amsterdam, Elsevier
- Hox, J. et al, 2005, Encyclopedia of social measurement, Amsterdam, Elsevier
- Insurance Europe, 2019, European Insurance in Figures – 2018 data, https://www.insuranceeurope.eu
- Jellema, M. et al, 2014, Hermes-Model: Universal Model to Estimate the Rebuilding Costs of Houses, FIG proceedings 2015, Sofia, Bulgaria
- Jellema, M. et al, 2017a, Compliance-raamwerk borgt de datakwaliteit, Infopaper nummer 2, pp 1-6, Zeist, Infofolio B.V.
- Jellema, M. et al, 2017b, Quality Assessment of Big Data With GIS, proceedings of AGILE conference 2017, Wageningen, the Netherlands
- Veregin, H., 1999. Data quality parameters, Geographical information systems, Volume 1, pp 177-189, Hoboken, John Wiley & Sons
- Wu, Y. et al, 2019, City-wide building height determination using light detection and ranging data, Environment and Planning B: Urban Analytics and City Science, volume 46 issue 9, Thousand Oaks, SAGE Journals

BIOGRAPHICAL NOTES

Rients van Wijngaarden has a master's degree in Artificial Intelligence and has over 10 years of experience in Data Science and Business Intelligence in several positions. Since 2019 he is manager Data Science at Infofolio B.V.

Michiel Jellema has over 35 years of experience in the geo-information sector. He did this in different positions and roles but always bridging organizational, Geo-ICT and other aspects of

making geo-information widely accessible. Since 2007 he is owner and director of Infofolio B.V.. In 2013 he earned his Phd-degree at the Delft University of Technology

Marinus de Bakker is senior lecturer Geo Media Design, HAS University of Applied Sciences and program manager Datalab Agrifood. He has been active in the field of Geo-information applications for over 35 years. His research interests are data quality, education for different users and the organization of geo-information in non profit and governmental organizations.

CONTACTS

Rients P. T. van Wijngaarden, MSc.
Infofolio B.V.
P.O. Box 72
3700 AB Zeist
THE NETHERLANDS
Tel. +3130-6911010
Email: r.vanwijngaarden@infofolio.nl
Web site: www.infofolio.nl

Dr. ir. Michiel Jellema
Infofolio B.V.
P.O. Box 72
3700 AB Zeist
THE NETHERLANDS
Tel. +3130-6911010
Email: m.jellema@infofolio.nl
Web site: www.infofolio.nl

Drs. Marien de Bakker
HAS University of Applied Sciences
Onderwijsboulevard 221
5223 DE ’s-Hertogenbosch
THE NETHERLANDS
Tel. +3188-8903356
Email: M.deBakker@has.nl
Web site: www.has.nl/nl