ALGORITHM FOR CONSTRUCTIONAL CHARACTERISTICS DATA CLEANSING OF LARGE-SCALE PUBLIC BUILDINGS DATABASE

HRVOJE KRSTIC & MIHAELA TENI
Faculty of Civil Engineering Osijek, Josip Juraj Strossmayer University of Osijek, Croatia

ABSTRACT

Research presented in this paper utilizes public-sector buildings database obtained from the Croatian Energy Management Information System (EMIS) which comprises over 3,500 public sector buildings. EMIS provides a transparent oversight and control of energy consumption, making itself an inevitable tool for systematic energy management. The EMIS database holds static technical data of each facility, including general, constructional data, energy performance data and dynamic energy usage data. But there are a lot of variables in a database with data values that are impossible, i.e. have values that are not logical or outside of possible, acceptable ranges, and they are probably the consequence of user input errors. Besides this, there are also cases with missing data. As previously stated, this raises the question: Is it possible to make an algorithm for data cleansing and find a way to calculate the missing data? To use the obtained database for further, more complex, analysis like clustering, machine learning and neural network applications, it is necessary to remove extreme values from the database. Research presented in this paper deals with this problem with an emphasis on buildings constructional characteristics and proposes a cleansing algorithm. As a result a possible range of variables and procedure for replacement of invalid input values is proposed. Research results and findings can be used in similar buildings databases to optimize the datasets and exclude variables with extreme values which can significantly impact modelling process. Further, the proposed algorithm can be useful when making decisions for energy refurbishment and building maintenance since it eliminates cases from the database that have misleading data. The presented results show that in some cases there are more than 80% of missing or excluded data. Findings can also be implemented in EMIS or a similar system to avoid further entering of unacceptable data values.
Keywords: Energy Management Information System, public sector buildings database, building characteristics, building maintenance, building energy refurbishment.

1 INTRODUCTION

Buildings have been recognized as a key pathway and setting for the reduction of energy and carbon emissions worldwide [1]. The growing concern over the worldwide increase in energy consumption and greenhouse gas emissions by buildings has resulted in huge efforts to improve building energy performance [2]. The domestic housing sector accounts for over a quarter (27%) of overall energy consumption in the UK and more than half (53%) of the carbon emissions are associated with energy use for space heating [3]. In Croatia buildings handle 40% of energy use and 70% of that energy is used for space heating and preparation of hot water [4], [5]. Comparable situation can be found in public buildings as well, especially because residential and non-residential building stock is mostly over 30-year-old in Croatia an as well in the rest of the Europe.

Directive 2010/31/EU on the energy performance of buildings introduced the concept of the "nearly zero-energy building". In this directive, the member states agreed to ensure that, after 31 December 2018, new buildings occupied and owned by public authorities will be nearly zero-energy buildings and that and by 31 December 2020 all new buildings will be nearly zero-energy buildings [6].

But what about existing, old buildings? Will they continue to waste massive amounts of energy?

The stock of residential buildings constructed before the 1970s, which have low performance regarding energy saving makes up more than 3/4 of the total existing residential buildings in European countries [7]. One of the first steps towards dealing with issue of existing buildings with poor energy performance was proscribing obligatory energy certification of buildings with final goal – ensuring energy efficient refurbishment.

Energy Performance Building Directive (EPBD) states that the energy efficiency of buildings has to be calculated in the member states and they are implementing the EPBD at the national level by taking into account local climate and conditions, requirements for indoor climate and cost efficiency [8]. To ensure successful implementation, measures were at first directed towards public buildings.

Energy Management Information System – EMIS, in Croatia, is a web application for energy sources consumption monitoring and analysis in public sector buildings. It provides a transparent oversight and control of energy consumption, making itself an inevitable tool for systematic energy management [9]. Systematic energy management includes strategic planning and sustainable management of energy resources. EMIS was developed as a part of a common project between the Ministry of Economy, Labour and Entrepreneurship and United nation developed program and with support of The Environmental Protection and Energy Efficiency Fund [10].

For this purpose for each building of the public sector (buildings owned by cities, counties, and the Government, such as administrative buildings, hospitals, schools, kindergartens, etc.), experts responsible for energy management gather and enter relevant data and information in EMIS [11].

But why it is important to monitor and analyse energy and water consumption and why is this especially important for public buildings? The answer is complex, but the main reason for this is the fact that maintenance and operations costs of public facilities are mostly covered by public finance [12]. It is therefore extremely essential to plan and manage above mentioned costs and to take into consideration some peculiarities of public buildings [13] like the:

- Particular nature of public works projects as social rather than investment capital;
- "Cradle to grave" (life cycle) or long periods of analysis;
- Low or zero income/revenue flows; and
- Sustainability performance (with a particular emphasis on environmental and societal impacts).

Research presented in this paper deals with EMIS data – missing data and data values that are impossible, i.e. have values that are not logical or outside of the possible range, and they are probable consequence of user input error. This error can be the consequence of negligence, unintentional input errors or simply the consequence of lack of required entry information. Presented research gives systematic approach concerning possible range of buildings constructional characteristics. Possible range of some variables is defined by regulations or can be found in existing literature, but for most of them different approach is required. There are research concerning big-data analysis but covers only information on energy consumption analysis, building type and building location [14], [15]. There was no literature systematically covering issues regarding possible values of building constructional characteristics especially variables like building shape factor, share of windows surfaces or thermal transmittance coefficient (U-value). An algorithm for constructional characteristics data cleansing of large-scale public buildings database is proposed and analysed in this paper.

Result of this research is also a proposal of procedure for replacement of invalid and/or missing input values.

2 USE OF EMIS FOR MODELING BUILDING ENERGY PERFORMANCE AND PREDICTING ENERGY CONSUMPTION

The EMIS database contains static technical data of each facility, including general, constructional characteristics and energy performance data for subject building, as well as dynamic energy resources usage data. Dynamic energy data contains all the data that are present on energy and water usage bills, which are provided on a monthly basis from the energy and water suppliers.

Some specific functionalities of EMIS system are as follows [16]:

- Managing basic data about buildings;
- Automatic and manual collection of energy and water consumption on a monthly, weekly, daily, or even real-time basis (monthly bills and/or meter reading);
- Easy access to information about the total amount of consumed energy and water, methods and places at which energy is consumed;
- Calculations and analysis in order to observe the unwanted, excessive and irrational energy and water usage and to identify opportunities for achieving energy and financial savings;
- Verification of achieved energy and water savings;
- Calculation of different energy consumption indicators;
- Automated alerts on critical events and malfunctions;
- Different user interfaces for each user role;
- Assortment of different building types; and
- Collection of automatic energy usage readings.

Predicting annual energy consumption is becoming a significant issue as utility costs become closely monitored and energy becomes a regulated and controlled resource in the built environment [17].

Predicted energy use is important to: set performance targets in new buildings; develop a business case for renovation measures; set budgets for utility costs; claim tax rebates and right size renewable energy equipment to meet a specific target [17]. Energy consumption modelling of buildings seeks to quantify energy requirements as a function of input parameters [18]. Models may be used for a variety of reasons, the most common being the determination of regional or national energy supply requirements (macro-scale) and the change in energy consumption of a particular dwelling due to an upgrade or addition of technology (micro-scale) [18]. In case of using real datasets in domain of energy consumption (like in EMIS), a modeller needs to deal with the problems of high-dimensionality, missing data, outliers, and inter-correlations, and use appropriate methods regarding the nature of data [19].

Real dataset can also be used for when making decisions for energy refurbishment and building maintenance since it eliminates cases from database that have misleading data. Buildings refurbishment not only decreases energy consumption but also improves whole condition of the building: its exploitation, noise insulation conditions, exterior, and comfort; prolongs buildings life cycle, increases value of the buildings, reduces negative impact to environment and guarantees healthy living and working conditions – satisfaction of these requirements is obligatory in sustainable refurbishment provision [20].

Mickaityte et al. [20] emphasize the following integrated sustainable refurbishment principles:

- Improvement of living conditions and provision of user-friendly apartments, increasing flexibility of the whole building concept and its parts, according to the current and future needs of the inhabitants;
- Decrease energy use and related building operational expenses;
- Increased use of environment-friendly materials and renewable energy sources;
- Economically favourable and innovative planning, building and using measures.

The main results expected from public buildings refurbishment are [20]:

- Energy savings;
- Increase of comfort;
- Healthy working environment assurance;
- Extension of building life cycle;
- Economized exploitation;
- Environmental protection.

Measures to improve energy efficiency are defined as group of actions that generally lead to verifiable and measurable or estimable energy efficiency improvement and reduction of energy and water consumption [21]. Energy efficient measures applied to existing buildings during minor/major retrofits in order to reduce their energy consumption can be grouped into three categories [22]:

- Building envelopes – thermal insulation, thermal mass, windows/glazing (including daylighting) and reflective/green roofs;
- Internal conditions – indoor design conditions and internal heat loads (due to electric lighting and equipment/appliances); and
- Building services systems – HVAC (heating, ventilation and air conditioning), electrical services (including lighting) and vertical transportation (lifts and escalators).

Final goal of presented research is to use EMIS database for further, more complex, analysis like clustering, machine learning and neural network for modeling building energy performance and predicting energy consumption – in this case natural gas for heating purposes. To do so, it is necessary to remove extreme, junk, values from database and substitute missing data when it's possible.

3 DATABASE AND PROPOSITION OF ALGORITHM FOR DATA CLEANSING

Data presented in this paper are obtained from EMIS system during the year 2017 for purposes of project titled Methodological Framework for Efficient Energy Management by Intelligent Data Analytics. Although EMIS contains 17,554 public buildings for most of them data is not collected yet and there are no records except buildings names [23].

This paper utilizes 3,659 public sector buildings from EMIS and variables concerning constructional and general characteristics – their age, year of restauration, climatic data, building size (in terms of surfaces and volumes), number of users, energy needs, thermal insulation and thermal transmittance coefficient (U-value) of building parts. Descriptive statistics of building characteristics for buildings dataset is presented in detail in [24]. Summarized findings are described below.

Statistical analysis revealed that there are 53.70% buildings in system that are older than 47 years. Most of those buildings do not meet current requirements regarding thermal requirements proscribed in The ordinance on energy audits of construction works and energy certification of buildings [25], since most of them are D energy rating (where A+ is the best and G is the worst energy rating). Currently for new buildings and refurbishment of old buildings law prescribes mostly energy rating B. This is especially important because most of those buildings (46.19%) are educational buildings and beside poor thermal properties problem there is also possibility of bad indoor climate. Nevertheless, the analysis also gives a great insight in the potential and possibilities of energy refurbishment of buildings in the future, which will result in a great reduction of heat energy demands and reduction of CO_2 emissions.

Algorithm for database cleansing is proposed based on database containing 3,659 buildings and 25 variables. The first step of data cleansing was development of descriptive algorithm for each variable in order to determine an acceptable range for each variable. Description of proposed range is summarized in Table 1. Table 1 presents analysed variables with their name, description, number of entries, possible range, number of entries outside of proposed range and percentage of excluded data.

Description of the proposed range for each variable and justification of proposed range is as it follows below, where the order number corresponds with the number of the variable presented in Table 1:

1. Value 1 is for Adriatic part of Croatia, 2 is for Continental part of Croatia. The analysis did not find variables outside of this range.
2. Some facilities are placed inside buildings which are not covered with EMIS system. 16 buildings had data outside of the proposed range.
3. Value 1 is for buildings that are part of cultural heritage and 0 is for all other buildings. Analysis did not find variables outside of this range.
4. Buildings older than 200 years have undergone some sort of restoration and therefore year of last restoration should be used as relevant. This acceptable value range is subjective. 443 buildings are with year of construction equal to 0, 57 is constructed before 1818, 289 is without this data.
5. Follows the explanation for variable V4. First criterium is carried out by all buildings, the second one does not fulfil 115 buildings.
6. It was defined by using number of persons and useful surface area of building (V12). Number of employees should be larger than 0 and less than 1 person per 1 m². 314 buildings do not have employees and 48 has more than 48 persons per m².
7. It was defined by using number of persons and useful surface area of the building (V12). Number of users should be greater than 0 since those are public buildings, maximum is not limited since they are occasionally in the building. 372 public buildings without users.
8. Value 0 is for building that aren't being used in time of collecting data. Analysis revealed 3.53% of data outside of possible range.
9. Value 0 for building that aren't being used in time of collecting data. Analysis revealed 4.35% of data outside of possible range.
10. Value 0 for building that aren't being used in time of collecting data. Analysis revealed 0.64% of data outside of possible range.
11. Building with the largest surface area in Croatia is located in the capital, Zagreb, and has 100,000.00 m². 456 has surface area less than 1 and there are 19 buildings without data regarding surface area.

Table 1: Analysed variables and analysis results presented as percentage of excluded data.

Variable name	Variable description	Number of entries	Proposed range	No. of entries outside of a range	Percentage of excluded data
V1	Climatic data	3,766	1 or 2	0	0.00%
V2	Share of use of total building area (%)	3,766	$0<V2\leq100$	16	0.42%
V3	Cultural heritage building	3,766	0 or 1	0	0.00%
V4	Year of construction	3,476	$1818<V4<2018$	789	22.70%
V5	Year of last restoration	3,022	1818-2018 and $V5\geq V4$	115	3.81%
V6	Number of employees	3,749	Defined by using number per surface area	362	9.66%
V7	Number of users	3,747		372	9.93%
V8	Number of working days per week	3,766	0–7	133	3.53%
V9	Number of working days per year	3,766	0–365	164	4.35%
V10	Number of working hours per work day	3,766	0–24	24	0.64%
V11	Gross surface area of building (m^2)	3,747	$1<V11<100000$	475	12.68%
V12	Useful surface area of building, A_k (m^2)	3,738	$0<V12<100000$ and $0.70\text{x}V11<V12<0.90\text{x}V11$	1,524	40.77%
V13	Heated surface area of the building (m^2)	3,273	$V13\leq12$	3,039	92.85%
V14	Heated volume area of the building, V_e (m^3)	3,738	$V14>V12\text{x}2.20$	238	6.37%
V15	Number of floors	3,766	0–25	0	0.00%
V16	Internal project temperature (°C)	823	$16<V16<28$	9	1.09%
V17	Building shape factor, f_0 (m^{-1})	3,766	$0.05<V17<1.86$	603	16.01%
V18	Share of windows surfaces (%)	3,766	$0<V18<100$	1,529	40.60%
V19	Specific annual energy needs for heating for referential climatic data for non-residential buildings, $Q''_{H,nd,ref}$ ($kWh/(m^2a)$) – calculated value	3,766	$15<V19<700$	3,232	85.82%
V20	Maximum allowed annual energy needs for heating for referential climatic data for non-residential buildings, $Q'_{H,nd,max}$ ($kWh/(m^3a)$)	3,766	$0<V20<130.89$	854	22.68%
V21	Relative annual energy demand for heating for referential climatic data of non-residential buildings, $Q_{H,nd,rel}$ (%) – calculated value	3,766	$1<V21<1000$	3,149	83.62%
V22	Thickness of thermal insulation (cm)	981	$1\leq V22<150$	0	0.00%
V23	Thickness of construction element (cm)	2,244	$6.0\leq V23<150$	0	0.00%
V24	Thermal transmittance coefficient (U-value) (W/m^2K)	5,795	$0.02<V24<6.0$	0	0.00%
V25	Maximum thermal transmittance coefficient (U-value) (W/m^2K)	3,721	$0.25<V25<3.0$	0	0.00%

12. 21 building is with surface area equal to 0, 28 buildings are without this data, 48 buildings have area larger than 100,000.00 m^2, 350 has the useful area larger than gross area and 1,150 buildings have useful area equal to gross area.

13. Heated surface area of the building cannot be smaller than useful surface area. There were 3,039 buildings that did not match this criterium, including building without data.

14. According to Croatian regulations minimal room height is 2.20 m. Based on that and value of surface area maximum, values of volume were determined. 238 buildings, including building without data did not match this criterion.

15. The tallest building in Croatia has 25 floors. Analysis did not find variables outside of this range

16. Highest internal design temperature is 28°C for closed swimming pools and 16°C for storerooms. 8 buildings are with temperature less than 16°C, 1 with temperature higher than 28°C. There are 2,942 buildings without this data.

17. Building shape factor is the ratio between the external surface area and heated volume. Accepted interval is consistent with the one obtained from another database containing 1,500 buildings where the factor ranged from 0.11 to 1.86. 562 entries have 0 value, 6 less than 0.05 m^{-1} and 35 bigger than 1.86 m^{-1}.

18. There are no buildings with 100% transparent façade and public buildings without windows (0%). 964 analysed buildings have a share of windows surfaces equal to 0, 565 is higher than 100.

19. Maximum values are hard to determine. For this purpose, another database was used that consist of 1,500 buildings in order to limit those maximum values. By using another database redundancy is avoided. Maximum value was 670 kWh/m^2 and therefore here 700 kWh/m^2 was accepted as maximum value. 15 kWh/m^2 is the minimum value since it is a value for passive building. 3,045 of analysed buildings have entered value of 0, 4 entries higher than 700 and 183 less than 15.

20. Maximum values are proscribed in the Technical Regulation on the Rational Use of Energy and Thermal Insulation in Buildings [25] depending on climatic zone. 854 buildings have this value equal to 0.

21. This variable achieves values higher than 0 but maximum values depends on ratio of specific annual energy needs for heating for referential climatic data for non-residential buildings, $Q'_{H,nd,ref}$ (kWh/(m^{3}a)) and Maximum allowed annual energy needs for heating for referential climatic data for non-residential buildings, $Q'_{H,nd,max}$ (kWh/(m^{3}a)). 3,145 of buildings have this value less than 1 and 4 buildings higher than 1,000.

22. Maximum values are based on a subjective assessment, according to construction practice and taking into account alternative thermal insulation materials like straw and sheep wool. Determination of the minimum value is based on a fact that any type of thermal insulation is thicker than 1 cm. The analysis did not find variables outside of this range.

23. If the construction element exists, its thickness should be at least 6,5 cm according to Eurocode 6 and HRN EN 1996-1-1/NA. 10 cm is proscribed for bearing reinforced concrete elements (20 cm for masonry walls). Minimum of 6.5 cm is taken since partition walls are usually thinner and 6.5 cm is the thinnest dimension of clay brick. Maximum value is subjective assessment and in this research is 150 cm, dimension that can be found only in old masonry and stone buildings. The analysis did not find variables outside of this range.

24. For thinnest wall defined for variable V28 maximum U-value is 5.72 W/m^{2}K and minimum for thickest wall and thermal insulation (V26 and V28) is 0.02 W/m^{2}K.

Maximum value for windows 5.90 W/m²K. The analysis did not find variables outside of this range.

25.   Maximum values are proscribed in the Technical Regulation on the Rational Use of Energy and Thermal Insulation in Buildings depending on climatic zone. The analysis did not find variables outside of this range.

Following analysis in this research was directed towards analysis of variables with a percentage of excluded data higher than 10%, those where variables V4, V11, V12, V13, V17, V18, V19, V20 & V21 – graphically presented and highlighted in Fig. 1.

This analysis was undertaken to reduce the number of rejected cases (buildings) and variables. Procedure is proposed for calculation of missing or wrongly entered data. This can be seen as a correction procedure or as a response for user input errors. The proposed procedure is presented in Table 2.

Based on the research results presented in Table 1, Table 2 and proposed variable range described below Table 1, flowchart describing the algorithm for constructional characteristics data cleansing of large-scale public buildings database is presented in Fig. 2.

4 CONCLUSION

Research results and findings and presented in this paper can be used in similar buildings databases to optimize dataset and exclude variables values with extreme values which can significantly impact modelling process. Proposed Algorithm for data cleansing significantly improves modelling procedure for future research and application of clustering, machine learning and neural networks with final goal – creation of model for predicting energy rating of buildings. The modelling process on data which are not possible or logical delivers models that are not applicable outside of dataset used for modelling. Further proposed algorithm can help when making decisions for energy refurbishment and building maintenance since it eliminates cases from the database that have misleading data. The proposed algorithm can easily be implemented in any statistical software's to eliminate values outside of possible range according to proposed range in Table 1 or to calculate missing values from existing ones according to proposed procedure presented in Table 2. Findings can also be implemented in EMIS or any other big data system containing analysed variables and apply
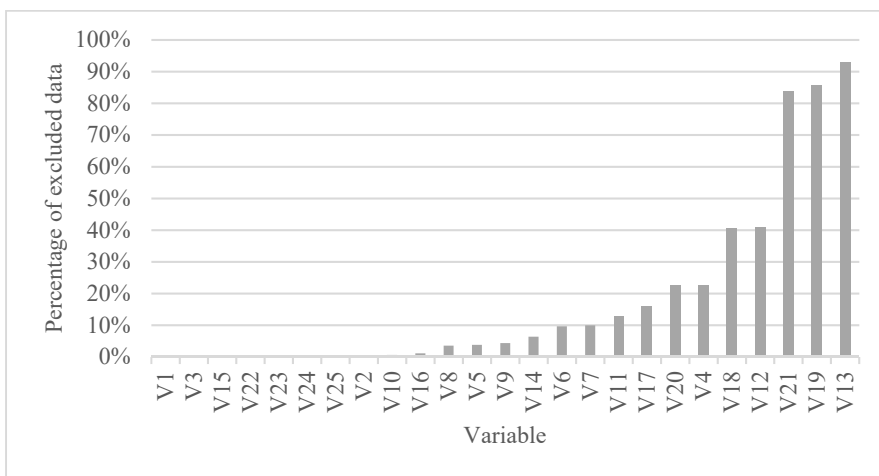

Figure 1: Graphical presentation of percentage of excluded data.

Table 2: Proposed procedure for replacement of invalid input values or missing values.

Variable name	Short variable description	Proposed procedure for replacement
V4	Year of construction	Replacement of year of construction with 7 construction periods according to [26] – before 1919, 1919–1945, 1946–1970, 1971–1980, 1981–1990, 1990–2000, 2000. Missing year could be replaced with 8th category.
V11	Gross surface area of building	Missing values can be calculated according to formula: V12/0.8 [27].
V12	Useful surface area of building	Missing values can be calculated according to formula: V14x0.32 [10].
V13	Heated surface area of the building	If missing or outside of the range it is advisable to use same value as V12 value.
V17	Building shape factor	If missing or outside of the range it is advisable to use arithmetic mean of the selected range – 0.955.
V18	Share of windows surfaces	If missing or outside of the range this variable should be omitted from further analysis since it cannot be calculated.
V19	Specific annual energy needs for heating – calculated value	If missing or outside of the range this variable should be omitted from further analysis since it cannot be calculated.
V20	Maximum allowed annual energy needs for heating	If missing or outside of the range this variable should be replaced with values proscribed in national regulations.
V21	Relative annual energy demand for heating – calculated value	If missing or outside of the range this variable should be omitted from further analysis since it cannot be calculated.

proposed algorithm presented in Fig. 2, to avoid further entering of impossible data values in the system. Doing so would prevent, or at least, warn users to check their values. Presented results show that in some cases (V13, V19 and V21) there were more than 80% of missing or excluded data. Findings can also be implemented in EMIS or similar system to avoid further entering of unacceptable data values. Future development of proposed algorithm will include analysis and control of variables about real energy consumption since presented analysis of energy consumption is only about calculated values obtained from an energy certification procedure. Since majority of public buildings was obliged to go through an energy certification procedure one conclusion is that users did not entered the data from energy certificates and that they should be alerting to correct this.
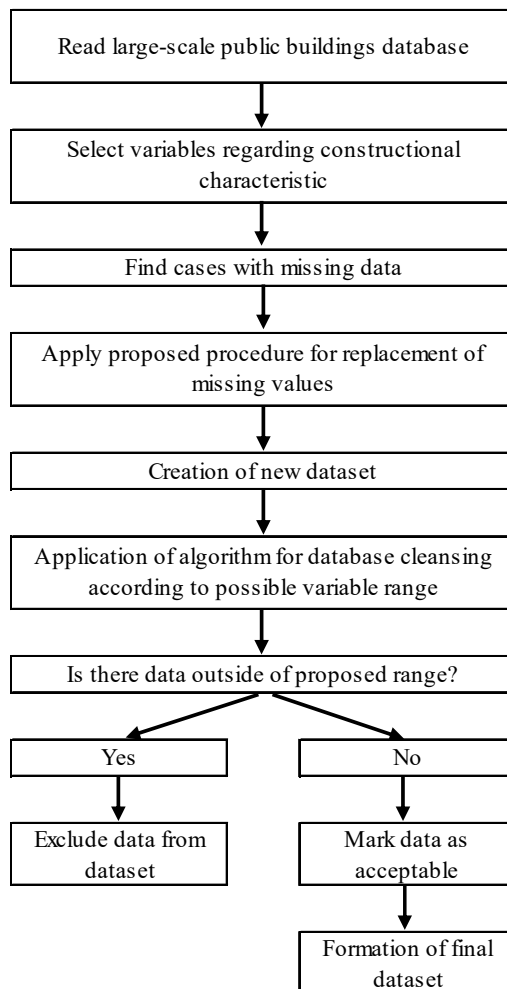
Figure 2: Flowchart of proposed algorithm for constructional characteristics data cleansing.

REFERENCES

[1]   Hsu, D., How much information disclosure of building energy performance is necessary? *Energy Policy*, **64**, pp. 263–272, 2014.

[2]   Lu, T., Lü, X. & Viljanen, M., A new method for modeling energy performance in buildings. *Energy Procedia*, **75**, pp. 1825–1831, 2015.

[3]   Pan, W., Relationships between air-tightness and its influencing factors of post-2006 new-build dwellings in the UK. *Building and Environment*, **45**(11), pp. 2387–2399, 2010.

[4]   Čulo, K. & Krstić, H., Cost benefit analysis of energy efficient family houses. *Second International Conference on Harmonisation Between Architecture and Nature, Eco-Architecture*, Algarve, Portugal, 2008.

[5]   Borković, Ž.H. et al., *Pilot projekt povecanja EE u zgradarstvu* (*The Pilot Project of Increasing EE in Buildings*). Energetski institut Hrvoje Požar, MZOPU, 2003.

[6]   Directive 2010/31 EU of the European Parliament and the Council of 19 May 2010 on the energy performance of buildings. *Official Journal of the European Communities*. L 153/21-2.

[7]   Ghazi Wakili, K. et al., Efficiency verification of a combination of high performance and conventional insulation layers in retrofitting a 130-year old building. *Energy and Buildings*, **82**, pp. 237–242, 2014.

[8]   Jokisalo, J. & Kurnitski, J., Performance of EN ISO 13790 utilisation factor heat demand calculation method in a cold climate. *Energy and Buildings*, **39**(2), pp. 236–247, 2007.

[9]   APN, Basic Information – Energy Management Information System (EMIS). www.enu.fzoeu.hr/isge. Accessed on: 1 Oct. 2017.

[10]   UNDP, *Priručnik za energetske savjetnike* (*Handbook for Energy Advisers*), Program UN-a za razvoj Zagreb, 2008.

[11]   APN, System structure – Energy management. www.enu.fzoeu.hr/isge/struktura-isge-sustava. Accessed on 1 Oct. 2017.

[12]   Krstić, H. & Marenjak, S., Maintenance and operation costs model for university buildings. *Technical Gazette*, **24**, 2017.

[13]   *Life Cycle Costing (LCC) as a Contribution to Sustainable Construction*, Final Guidance: London, 2007.

[14]   Mathew, P.A. et al., Big-data for building energy performance: Lessons from assembling a very large national database of building energy use. *Applied Energy*, **140**, pp. 85–93, 2015.

[15]   Steadman, P., Bruhns, H.R. & Rickaby, P.A., An introduction to the national non-domestic building stock database. *Environment and Planning B: Planning and Design*, **27**(1), pp. 3–10, 2000.

[16]   APN, Functionality – Energy Management Information System – EMIS. www.enu.fzoeu.hr/isge/osnovne-funkcije-isge-sustava. Accessed on: 1 Oct. 2017.

[17]   Collins, L., Predicting annual energy consumption with thermal simulation: A UK perspective on mitigation of risks in estimation and operation. *Building Simulation*, **5**(2), pp. 117–125, 2012.

[18]   Swan, L.G. & Ugursal, V.I., Modeling of end-use energy consumption in the residential sector: A review of modeling techniques. *Renewable and Sustainable Energy Reviews*, **13**(8), pp. 1819–1835, 2009.

[19]   Zekić-Sušac, M., Machine learning in energy consumption management. *The 14th International Symposium on Operations Research in Slovenia*, Bled, Slovenia, 2017.

[20]   Mickaityte, A. et al., The concept model of sustainable buildings refurbishment. *International Journal of Strategic Property Management*, **12**(1), pp. 53–68, 2008.

[21]   Basic concepts. www.mgipu.hr/default.aspx?id=14521. Accessed on 3 Mar. 2015.

[22]   Pyloudi, E., Papantoniou, S. & Kolokotsa, D., Retrofitting an office building towards a net zero energy building. *Advances in Building Energy Research*, pp. 1–14, 2014.

[23]   Zekić-Sušac, M., Modeliranje potrošnje prirodnog plina javnih zgrada inteligentnom podatkovnom analitikom. *16th Natural Gas, Heat And Water Conference And 9th International Natural Gas, Heat And Water Conference*, Osijek, Croatia, 2017.

[24]   Krstić, H. & Teni, M., Analysis of energy performance and buildings characteristics obtained from Croatian energy management information system. *3rd International Conference on Architecture, Materials and Construction (ICAMC 2017)*, Amsterdam, Netherlands, 2017.

[25] *Pravilnik o energetskom pregledu zgrade i energetskom certificiranju* (*Ordinance on energy audits of construction works and energy certification of buildings*). Narodne novine broj 48/14, 150/14, 133/15, 22/16, 49/16.

[26] *Housing Statistics in European Union,* T.H.M.o.t.I.a.K. Relations, Editor. OTB Research Institute for the Built Environment & Delft University of Technology: Hague, 2010.

[27] *Regulations for the Calculation of Areas and Volumes of the Facilities*, 118/08, 74/10, 86/10, 24/12 and 121/12, Editor. Official Gazette of Republika Srpska, 2013.