



Chapitre d'actes

2019

Accepted version

Open Access

This is an author manuscript post-peer-reviewing (accepted version) of the original publication. The layout of the published version may differ .

Identifying key variables and interactions that explain hydraulic imbalance - application of regularization-based variable selection methods

Cho, Haein; Cabrera Santelices, Jose Daniel; Patel, Martin

How to cite

CHO, Haein, CABRERA SANTELICES, Jose Daniel, PATEL, Martin. Identifying key variables and interactions that explain hydraulic imbalance - application of regularization-based variable selection methods. In: CISBAT 2019 - Journal of Physics: Conference Series. [s.l.] : [s.n.], 2019. doi: 10.1088/1742-6596/1343/1/012043

This publication URL: <https://archive-ouverte.unige.ch/unige:136362>

Publication DOI: [10.1088/1742-6596/1343/1/012043](https://doi.org/10.1088/1742-6596/1343/1/012043)

PAPER • OPEN ACCESS

Identifying key variables and interactions that explain hydraulic imbalance - application of regularization-based variable selection methods

To cite this article: Haein Cho *et al* 2019 *J. Phys.: Conf. Ser.* **1343** 012043

View the [article online](#) for updates and enhancements.



IOP | ebooks™

Bringing together innovative digital publishing with leading authors from the global scientific community.

Start exploring the collection—download the first chapter of every title for free.

Identifying key variables and interactions that explain hydraulic imbalance - application of regularization-based variable selection methods

Haerin Cho^{a,*}, Daniel Cabrera^a, Martin K. Patel^a

^aChair in Energy Efficiency, Institute for Environmental Sciences (ISE) and Department F.-A. Forel for Environmental and Aquatic Sciences (DEFSE), Faculty of Science, University of Geneva, 1211 Geneva 4, Switzerland

haerin.cho@unige.ch

Abstract. Heating systems must be subjected to hydraulic balancing in order to ensure proper operation. When the heating system is hydraulically imbalanced, heat is unevenly distributed across dwellings resulting in a large temperature spread, overheating, and consequently, waste of energy. In this study, we investigate the extent to which hydraulic imbalance affects the thermal energy consumption in buildings. Furthermore, key variables and interactions that influence the thermal performance in buildings are identified. Results show that higher variation in indoor temperature lead to higher energy consumption. Compared to new buildings, old buildings that have large boiler capacity and small heated floor area are more likely to be hydraulically imbalanced and to consume more energy.

1. Introduction

In Europe, the building sector is responsible for 41% (27% for residential and 14% for tertiary) of the total energy consumption [1]. While energy efficiency of new buildings has been improved as a consequence of more stringent building codes, the main challenge is related to old existing buildings that represent more than 90% of the existing building stock in Europe. Since 2010, the Energy Performance of Buildings Directive (EPBD) has required EU member states' commitment to adopt renovation strategies to improve energy efficiency of old existing buildings [2].

For residential buildings, space heating is the predominant energy service and represents around 70% of the final energy use [3]. The thermal performance of buildings can be improved in three ways, i.e. 1) by enhancing the thermal performance of the envelope through improved insulation strategies, 2) by replacing the existing heating system by a more efficient one or 3) by optimizing the operation of the existing heating system. The first two options require relatively high investment costs and have a long payback period [4]. The benefits of the first two options are reduced if the existing heating system does not function well. It is therefore important to inspect and optimize the performance of existing heating system in buildings [5].

Focusing on the optimization measures, this study investigates hydraulic balancing, which is one of the main approaches applied for heating systems [6]. When the heating system is hydraulically imbalanced, the heating medium (typically hot water) that carries thermal energy will be unevenly distributed across dwellings, causing overheating, a large temperature spread, and consequently energy losses [7]. Energy savings potentials to be harvested from balancing the hydraulic system are estimated



Content from this work may be used under the terms of the [Creative Commons Attribution 3.0 licence](https://creativecommons.org/licenses/by/3.0/). Any further distribution of this work must maintain attribution to the author(s) and the title of the work, journal citation and DOI.

in the range of 2-23% through field experiments and simulations [8]–[12]. The estimation requires long-term monitoring of thermal energy consumption in combination with measurement of temperature [9], [10], as well as of the operational data of the heating system such as mass flow rate, supply water temperature or temperature difference between supply and return temperature to quantify heat losses. [13]. However, if we could correlate the existence of hydraulic imbalance with known physical building properties, we would be able to pinpoint candidate buildings more easily and improve the hydraulic performances with reduced effort. To our knowledge, there is no study that explored the explanatory power of building characteristics in order to establish the need for hydraulic balancing.

Therefore, with the goal of identifying the key variables that influence hydraulic imbalance and consequently, thermal energy consumption, we conduct a two stage-analysis. First, we assess the extent to which hydraulic imbalance affects energy consumption in buildings. Because the total energy consumption is determined by temperature spread, we aim to understand the influence that the variation in indoor temperature across the flats have on thermal energy consumption. Secondly, we investigate the effects of building characteristics on the thermal performance of buildings by applying advanced regularization method which allows to address the issue of multicollinearity among variables

2. Method

We study 49 multifamily buildings built before 2015 in Geneva (Switzerland). The buildings are heated with natural gas and fuel oil-fired boilers for which no hydraulic balancing had been conducted. We group the buildings that share one central boiler which resulted in 13 groups. We collected hourly indoor temperature between 2016 and 2018 from all 656 flats within the 13 groups of buildings, and annual final energy consumption for heating between 2016 and 2018. We correlate standard deviation and mean indoor temperature against annual final energy consumption for heating for each group. We then test to which extent various variables (number of floors, number of apartments, average dwelling size, height, age, boiler capacity and total heated area) explain the standard deviation of indoor temperature and mean indoor temperature. Age is introduced as a categorical variable (buildings constructed before 1980, between 1981 and 2000, between 2000 and 2015) in the analysis. Table 1 lists the considered variables.

Table 1. Explanation of variables.

Category	Name of variables	Additional explanation
Building	Number of floors	-
	Number of apartments	-
	Average dwelling size	Dwelling size (m ²)
	Height	Building height (m)
	Age	Building's construction year
Heating	Total SRE (Surface de Référence Énergétique)	Sum of all heated floor areas (m ²)
	Boiler capacity	Boiler capacity (kW)
Energy	Average IDC (Indice de Dépense de Chaleur)	Annual final energy consumption for heating (MJ/m ² /year)

First, we apply linear ordinary least squares (OLS) regression to understand the impact of temperature on building energy consumption. Then, we verify multicollinearity based on a variance inflation factor (VIF) and apply a prominent shrinkage and selection method for linear regression models, called 'Least Absolute Shrinkage and Selection Operator (LASSO)' [14]. LASSO is one of the commonly used regularization methods that correct overfitting which is caused when the model is too complex. By adding a penalty, which is the sum of the absolute value of the coefficients, LASSO shrinks

some parameters, with minor contribution to outcome, to zero. We subsequently identify a set of variables with the highest explanatory power. This technique has been applied to many different domains, as it simplifies the model to make it interpretable [15]. The main reasons for using the ‘LASSO’ technique are to select a small subset of variables that possess high explanatory power (thereby addressing the issues of multicollinearity) and to lower computational complexity. A final model is developed only with the selected variables for which we repeat OLS regression.

3. Results and discussion

At first, we investigate the correlation between indoor temperature and energy consumption, as shown in figure 1. We find that each age group forms a distinctive cluster in the scatter plot. Comparing the relationship between mean and standard deviation of indoor temperature on the one hand with annual final energy consumption for heating purposes on the other, we observe that the old buildings and new ones demonstrate an opposite pattern. The old buildings show low mean indoor temperature and high standard deviation of indoor temperature, whereas the new buildings show high indoor temperature and low standard deviation. Buildings of medium age take an intermediate position. The higher the mean and the standard deviation of indoor temperature is, the more energy is consumed.

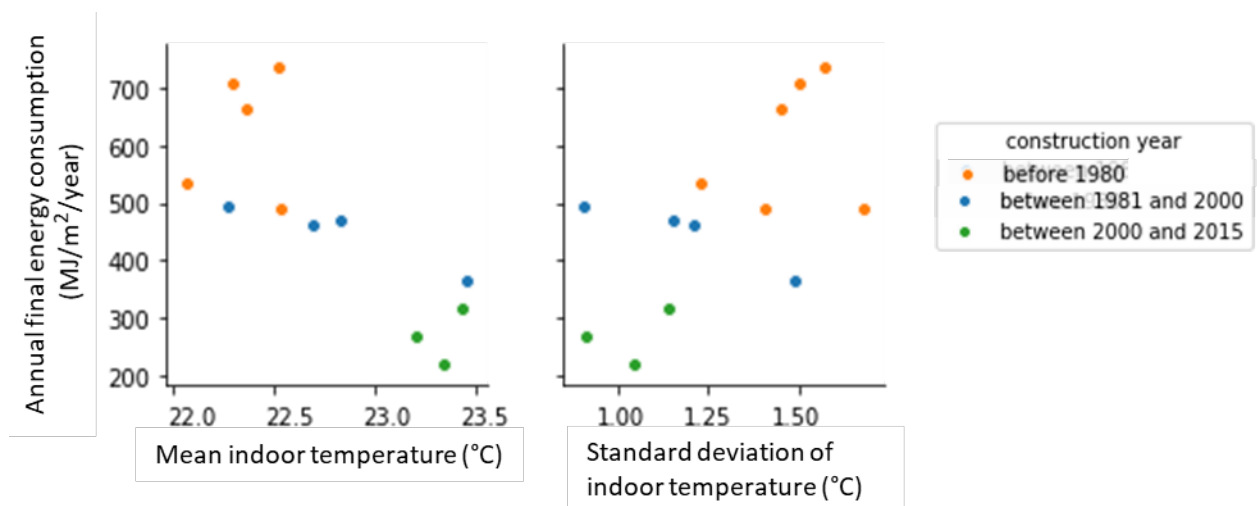


Figure 1. Scatter plot of annual final energy consumption as a function of mean and standard deviation (indicated as ‘sd indoor temp’) of indoor temperature between 2016 and 2018

In the next step, we combine all variables and select variables that have the highest explanatory power. We examine the variance-inflation factors (VIF) that indicate the level of collinearity among variables within a multiple regression [16]. A VIF value of 1 means that there is no multicollinearity and a VIF value higher than 1 indicates multicollinearity. As applied in an earlier study [17], we use a VIF value of 3.3 as threshold beyond which we carry out LASSO regression. LASSO regression effectively selects a subset of variables by applying a fitting procedure and setting the coefficients of irrelevant variables to zero. Tables 2 and 3 show the coefficients of the LASSO regression and the subsequent OLS regression when the standard deviation and average of indoor temperature are taken as a dependant variable, respectively.

Two variables that have a non-zero coefficient (β_L) in both cases (table 2 and 3) after LASSO regression, i.e. boiler capacity and total SRE (total heated floor area), are selected as key variables. The final models developed using these parameters as independent variables result in a coefficient of determination R^2 of i) 11.3% for standard deviation of indoor temperature as dependent variable and of ii) 49.6% for the average of indoor temperature as dependent variable.

Table 2. LASSO regression results for standard deviation of indoor temperature as a dependent variable (β_L represents the coefficients of the Lasso regression, B_{OLS} stands for the unstandardized coefficients of the reduced OLS regression, SE_{OLS} is standard error of unstandardized coefficients)

Name of variables	β_L	B_{OLS}	SE_{OLS}
Number of floors	0.000000	NA	NA
Number of apartments	0.000000	NA	NA
Average dwelling size	0.000000	NA	NA
Boiler capacity	-0.002176	-0.0023	0.001
Total heated floor area	9.410450e-05	8.74e-05	3.71e-05
Height	0.000000	NA	NA
Age	0.000000	NA	NA

Table 3. LASSO regression results for average of indoor temperature as a dependent variable (see Table 2 for abbreviations)

Name of variables	β_L	B_{OLS}	SE_{OLS}
Number of floors	0.000000	NA	NA
Number of apartments	0.000000	NA	NA
Average dwelling size	0.000000	NA	NA
Boiler capacity	-0.002176	-0.0023	0.001
Total heated floor area	9.410450e-05	8.74e-05	3.71e-05
Height	0.000000	NA	NA
Age	0.000000	NA	NA

Figure 2 is a scatter plot matrix that shows all possible pairwise scatter plots of the selected variables in one figure. We observe that boiler capacity and total heated area are related to the extent to which the building is hydraulically imbalanced and to the amount of final energy consumed for heating purposes. As the colored zones (orange and green) indicate, old buildings (orange) and new buildings (green) are placed in distinctively opposite zones. The new buildings are characterized by small boiler capacity and large total heated area, while the opposite is the case for old buildings. In each group, buildings with higher boiler capacity and higher total heated area demonstrate higher mean and standard deviation of indoor temperature and also higher annual final energy consumption for heating. Data points that belong to buildings of the medium age group are spread out rather than showing distinctive pattern.

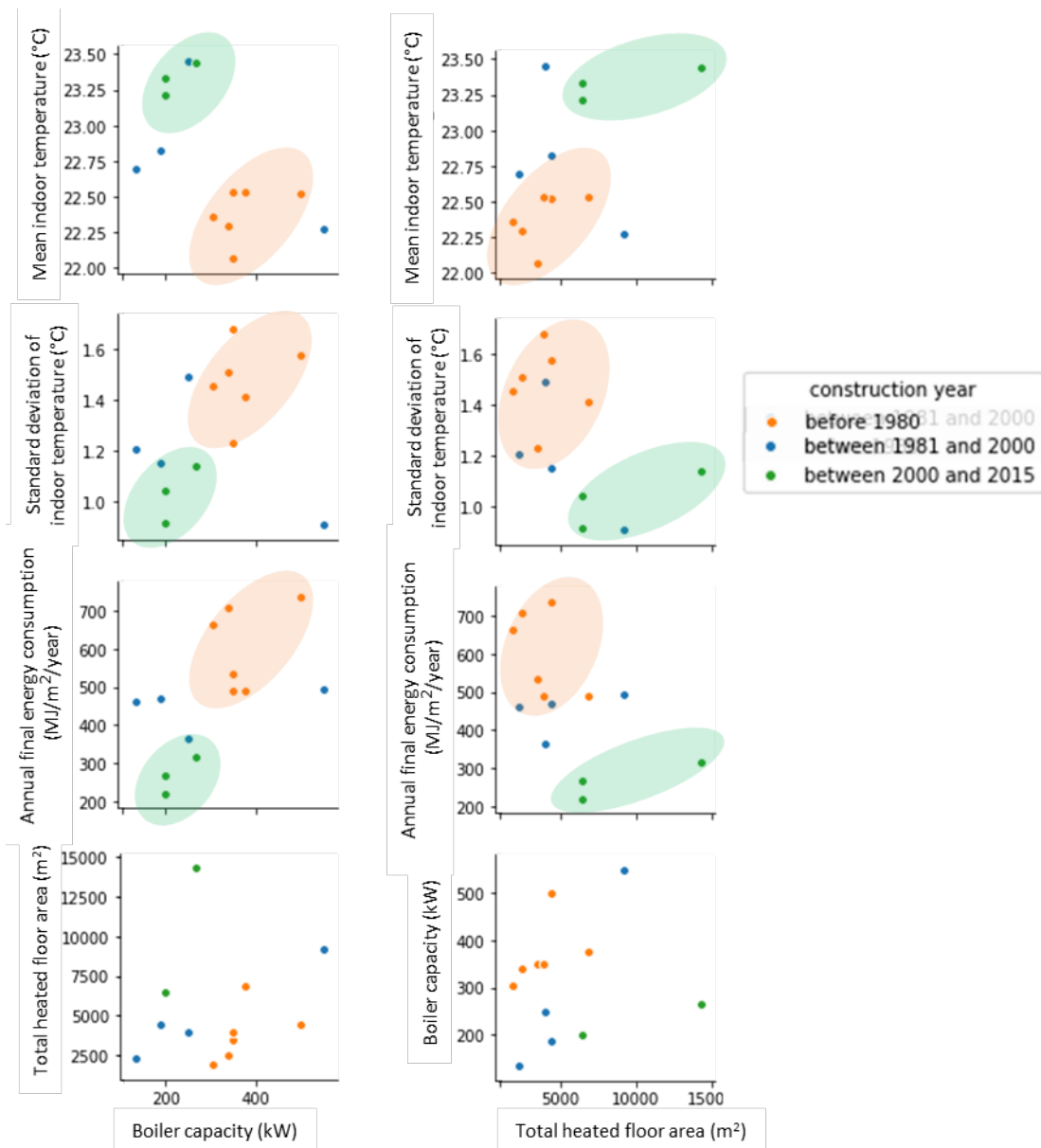


Figure 2. Scatter plot matrix representing the relationship between the level (mean) and deviation (standard deviation) of indoor temperature and the explanatory variables

4. Conclusion

As the oldest (before 1980) and newest buildings (between 2000 and 2015) indicate, there is a linkage between high standard deviation and high mean of indoor temperature on the one hand with large thermal energy consumption on the other (Figure 1 and 2). Based on the regularization method, boiler capacity and total heated area are identified to have the highest explanatory power on the variability in standard deviation of indoor temperature and mean indoor temperature. To maximize thermal energy savings, old buildings (before 1980) with high boiler capacity and large heated area should therefore be prioritized for hydraulic balancing. This study is conducted with an intermediate sample size. However, by increasing the sample size, we aim to achieve a higher level of statistical significance. This would then enable the development of a statistical method that allows to reliably pinpoint buildings for which hydraulic balancing should be conducted.

References

- [1] M. Santamouris, “Energy Consumption in the European Built Environment The Role of Cooling,” 2015.
- [2] European Commission, “Energy performance of buildings directive,” 2018. [Online]. Available: <https://ec.europa.eu/energy/en/topics/energy-efficiency/energy-performance-of-buildings/energy-performance-buildings-directive>.
- [3] L. Mazzarella, “Energy retrofit of historic and existing buildings. The legislative and regulatory point of view,” *Energy Build.*, vol. 95, pp. 23–31, May 2015.
- [4] A. Ferrante, E. Cattani, A. Fotopoulou, G. Semprini, and E. Agliardi, “Assistant Buildings’ addition to Retrofit, Adopt, Cure And Develop the Actual Buildings up to zeRo energy, Activating a market for deep renovation,” 2016.
- [5] B. Lenz, J. Schreiber, and T. Stark, *Sustainable Building Services: Principles - Systems - Concepts*. 2011.
- [6] IMI hydronic engineering, “The aim of balancing is to apply adjustment of flow rates to achieve design requirements.” [Online]. Available: <https://www.imi-hydronic.com/sites/EN/international/knowledge-centre/HydronicTheory/Pages/Balancing-Protocol.aspx>. [Accessed: 15-May-2019].
- [7] R. Ruch, P. Ludwig, and T. Maurer, “Balancing Hydronic Systems in Multifamily Buildings,” 2014.
- [8] C. Ahern and B. Norton, “Energy savings across EU domestic building stock by optimizing hydraulic distribution in domestic space heating systems,” *Energy Build.*, vol. 91, pp. 199–209, 2015.
- [9] T. Cholewa, I. Balen, and A. Siuta-Olcha, “On the influence of local and zonal hydraulic balancing of heating system on energy savings in existing buildings – Long term experimental research,” *Energy Build.*, 2018.
- [10] T. Cholewa, A. Siuta-Olcha, and C. A. Balaras, “Actual energy savings from the use of thermostatic radiator valves in residential buildings – Long term field evaluation,” *Energy Build.*, vol. 151, pp. 487–493, Sep. 2017.
- [11] G. P. Henze and A. G. Floss, “Evaluation of temperature degradation in hydraulic flow networks,” *Energy Build.*, 2011.
- [12] A. Ghahramani, K. Zhang, K. Dutta, Z. Yang, and B. Becerik-Gerber, “Energy savings from temperature setpoints and deadband: Quantifying the influence of building and system properties on savings,” 2016.
- [13] A. Ashfaq and A. Ianakiev, “Investigation of hydraulic imbalance for converting existing boiler based buildings to low temperature district heating,” 2018.
- [14] T. Hastie, R. T. Martin, W. Hastie, • Tibshirani, and • Wainwright, *Statistical Learning with Sparsity -The Lasso and Generalizations*. CRC Press Taylor & Francis Group, 2016.
- [15] S. L. Kukreja, J. Löfberg, and M. J. Brenner, “A LEAST ABSOLUTE SHRINKAGE AND SELECTION OPERATOR (LASSO) FOR NONLINEAR SYSTEM IDENTIFICATION,” *IFAC Proc. Vol.*, vol. 39, no. 1, pp. 814–819, 2006.
- [16] PennState, “12.4 - Detecting Multicollinearity Using Variance Inflation Factors | STAT 501.” [Online]. Available: <https://newonlinecourses.science.psu.edu/stat501/node/347/>. [Accessed: 14-May-2019].
- [17] G. Huebner, D. Shipworth, I. Hamilton, Z. Chalabi, and T. Oreszczyn, “Understanding electricity consumption: A comparative contribution of building factors, socio-demographics, appliances, behaviours and attitudes,” 2016.