Environmental and Climate Technologies
2022, vol. 26, no. 1, pp. 1234–1243
https://doi.org/10.2478/rtuect-2022-0093
https://content.sciendo.com

Finding the Generic Hygrothermal Properties of Historical Bricks by Supervised Agglomerative Clustering

Ritvars FREIMANIS[1*], Zigmārs ZUNDANS[2], Roberts BALINS[3], Ruta VANAGA[4], Andra BLUMBERGA[5]

[1–5]*Institute of Energy Systems and Environment, Riga Technical University, Azenes iela 12/1, Riga, Latvia*

***Abstract* – Finding the generic hygrothermal properties of historical brick for application in Heat Air and Moisture (HAM) simulation programs such as *Delphin, Wufi*, etc., is the main objective of this paper. In this paper hygrothermal properties and *Delphin* simulation results of 40 different historical brick samples from the 17th to 20th Century, were used. The clustering results of hygrothermal properties were cross-examined with the results of clustering results of *Delphin* simulation data. Six and three clusters were found to be optimal, accordingly for Hygrothermal properties and *Delphin* results data groups. After cross-examination, a total of 9 combined clusters were recognized, with two dominant clusters containing 67.5 % of all samples (30 and 37.5 %), four of the clusters had only one sample in them, and other clusters had two, three, and four samples in them. Additionally, all the resulting clusters were compared with the brick sample groups that were created based on the description of the brick: color, material type, and year of manufacturing.**

***Keywords* – Agglomerative; brick; clustering; *Delphin*; generic; historical; hygrothermal; properties; supervised**

Nomenclature		
ML	Machine Learning	Method of Data Analysis
P	Parameters cluster	Cluster group from brick parameters
R	Delphin results cluster	Cluster group from Delphin results
HAM	Heat, Air, and Moisture	Type of simulation program
CT	Count	Unit for figures
FEC	Final Energy Consumption	Energy consumption by the end user

1. INTRODUCTION

The building sector accounts for one-third of the total Final Energy consumption (FEC) globally, and up to 42 % in the European Union (EU) [1], [2]. Therefore, by implementing energy efficiency measures, such as insulation of building envelope we can decrease the FEC [3]. The building envelope can be insulated either from the outside or from the inside, and in some cases, for example, for buildings with preservable façade, internal insulation is the only option. When insulating a building for the inside, the hygrothermal properties of the building envelope change and may cause durability and performance problems [4]. As the

* Corresponding author.
E-mail address: ritvars.freimanis@rtu.lv

©2022 Author(s). This is an open access article licensed under the Creative Commons Attribution License (http://creativecommons.org/licenses/by/4.0).

topicality of energy efficiency increases, so is the topicality of internal insulation [5]–[11]. To avoid long and expensive experiments numeric simulations are widely used to find solutions for complex problems, including the prediction of the performance of refurbished/insulated historical masonry buildings [12]–[15]. However, these numerical calculations are based on rather complex functions that describe heat and moisture storage and transport within the materials [16]. Moreover, hygrothermal properties of the materials are used as input parameters for this calculation, therefore, accurate measurements of such parameters determine the accuracy and reliability of obtained results [17]. The amount of input parameters for such calculations is rather large and to obtain most of the data, specialized laboratory equipment and procedures must be used, which can be both, time-consuming and expensive [16]. Moreover, indoor and outdoor climatic conditions also have an impact on the output results.

By analysing multiple brick samples within one region (Latvia), it should be possible to obtain a cluster of samples that would generate similar output results. Afterward, general parameters for these clusters could be identified for use in numerical models. Such clusters are important to reduce the number of necessary parameters to be tested. With these clusters, it would be possible to include brick within a specific cluster based on a few properties or even brick types. Afterward, the numerical simulations could use general properties from specific clusters.

At the time of writing, no such clusters or attempts to generate one have been found in the scientific literature by the author. A similar study has been done by Hansen et.al. [18], where they tried to determine the key parameters that influence the outcome of the *Delphin* simulations the most, therefore minimizing the amount of necessary input data for reliable output data. In their study, they found that the water uptake coefficient has the most impact on the relative humidity results within the masonry and that water vapour diffusion factor and thermal conductivity showed seasonal impact on the results, and that density and open porosity had no significant impact on the results [18].

This study is the further research of the project ‘Improvement of building energy efficiency technologies’ (No. 03000–3.1.2-e/163) supported by the National Research Program of Latvia. In this study, we are trying to find one or more generic materials that could be used in the *Delphin* simulations, for the historical buildings in the region of Latvia.

2. METHODOLOGY

To analyse the Hygrothermal properties and *Delphin* simulation data [19], a Hierarchial clustering dendrogram, using an agglomerative clustering method, is created. *Jupyter Lab* [20] is used as an environment for this supervised Machine Learning (ML) clustering analysis. To perform the clustering *Python* Libraries from *SciKit Lea*rn are used [21]. Used *Python* libraries including *Sklearn* and *Matplotlib*. *Sklearn* is built on *NumPy* and *Scipy* and it is used for the calculations of the clustering process. *Matplotlib* is used for the visualization of the results, including the dendrogram.

After the initialization of the used libraries, the data for the clustering are imported and assigned to the dedicated variable ‘data’. Used data can be found in the *Zenodo* data repository [22]. Before the clustering, all the data entries are normalized with the *StandardScaler* function, built into the *Sklearn* library. Data normalization is necessary to avoid some data having more weight than other data on the results.

When performing clustering analysis, all the possible cluster count solutions are generated meaning, at the one end of the spectrum all the samples are in one large cluster, and on the other side of the spectrum, each sample creates an individual cluster.

The difference between one large and many small clusters is the total distance between all the samples in the cluster. If there is data only from one sample present in the cluster, the distance is zero (the smallest possible distance). And if all the samples are combined into one large cluster, the distance between all the sample data is the largest possible. This can be shown graphically with the dendrogram (see Fig. 1).

In the hierarchical clustering dendrogram, the vertical axis represents the distance, and the horizontal lines in the graph represent the clusters. You can see from Fig. 1, that the largest possible distance for this case is around 20 (the distance has no unit – because we are using normalized data with different original units). When following the dendrogram from the top down, it can be traced how this one cluster of 40 samples starts to split into smaller clusters, containing fewer samples in each cluster, and at the same time decreasing the total distance of these new clusters.

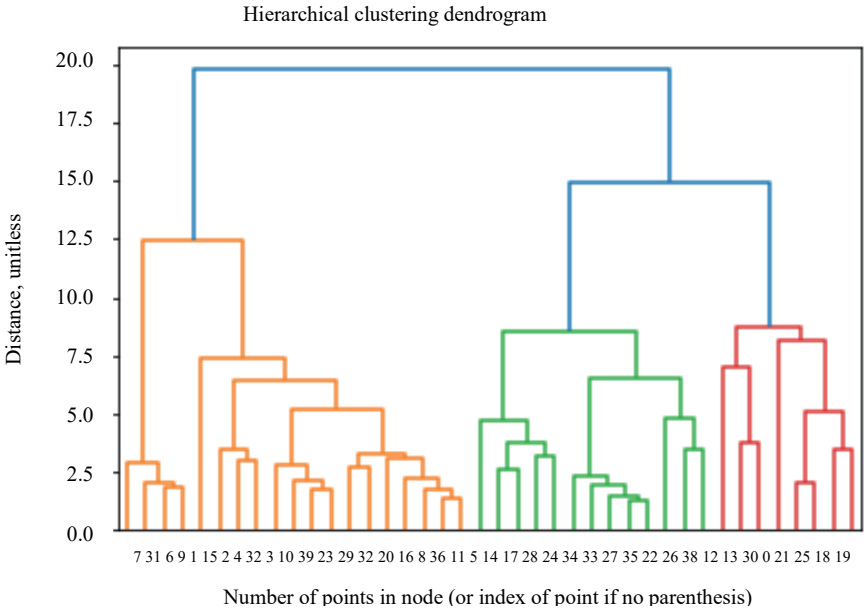


Fig. 1. Dendrogram of Hierarchical clustering.

To extract the results, all possible cluster counts (for both datasets) are generated by changing the *n_cluster* parameter in *Sklearn* and taking the array of corresponding results (in the array each number represents a cluster that the sample belongs to – the location of the number in the array correspond to the specific brick sample). Afterward, the method *model. labels_* is called to show the labels of clusters (meaning each label represents a cluster).

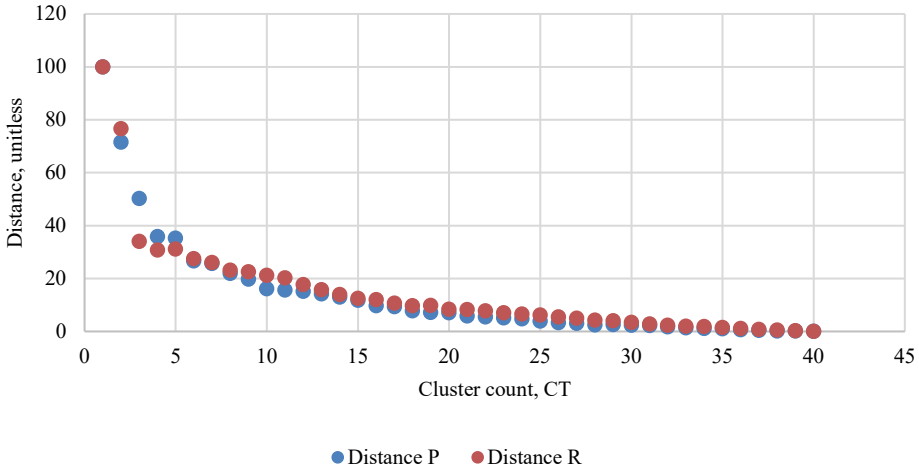


Fig. 2. Normalized Clustering results

To find the optimal combination (cluster count vs. distance) of the generated clusters, both cluster group results are normalized (see Fig. 2) and multiplied together (see Fig. 3). Afterward, for all the optimal combinations of clusters their respective average distance is compared to their cluster count (see Fig. 4).

To multiply both cluster groups from Parameters (P) and *Delphin* results (R) together, a vector of normalized distance data in descending order is created for each cluster group. These vectors are then multiplied together. All the results of the multiplied cluster results within the optimal range are shown in Fig. 3.

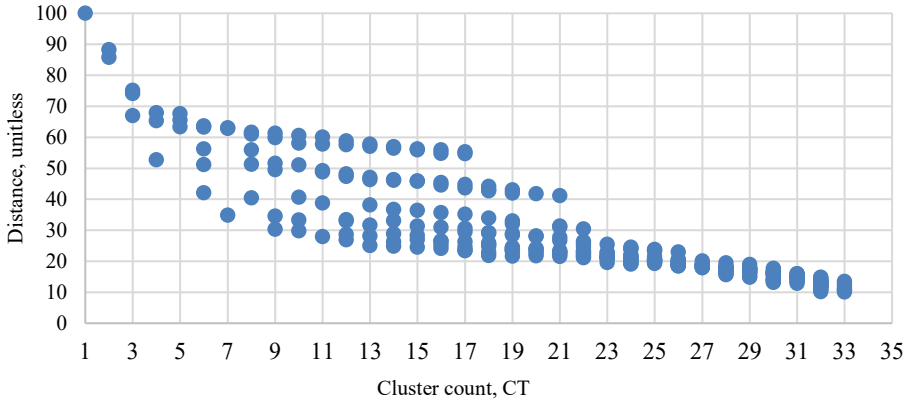


Fig. 3. Multiplied cluster results.

It can be seen in Fig. 3, that there is more than one combination of clusters for the same cluster count, and because we are looking for the most optimal combination, only the combinations with the lowest distances are used in those cases.

Afterward, the cluster count and distance of the selected combinations from Fig. 3, are normalized to the scale (0–1), with 0 being the worst case (highest cluster count; highest distance) and 1 being the best result (lowest cluster count, lowest distance), and plotted one against the other (See Fig. 4).

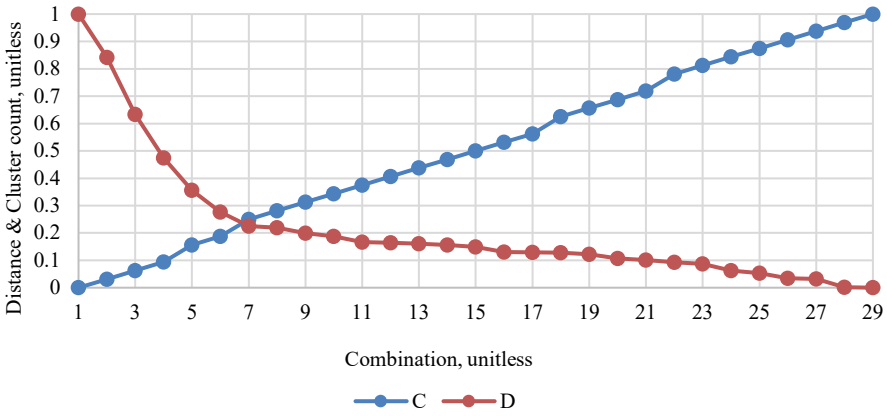


Fig. 4. Finding the optimal cluster count (C–Cluster; D–Distance).

The crossing point of Cluster and Distance lines can be considered as the optimal solution towards the smallest possible cluster count with the smallest possible distance. The resulting combination of the P and R clusters is then used to create new Combined clusters and these Combined clusters are used for further analysis, including, comparison to the clusters, based on the description of the brick samples (Year of manufacturing, colour, material type) and for the identification of the general groups/clusters of bricks.

3. RESULTS

All the possible cluster counts from *Delphin* results data, plotted against those clusters' corresponding distances, are shown in Fig. 5.

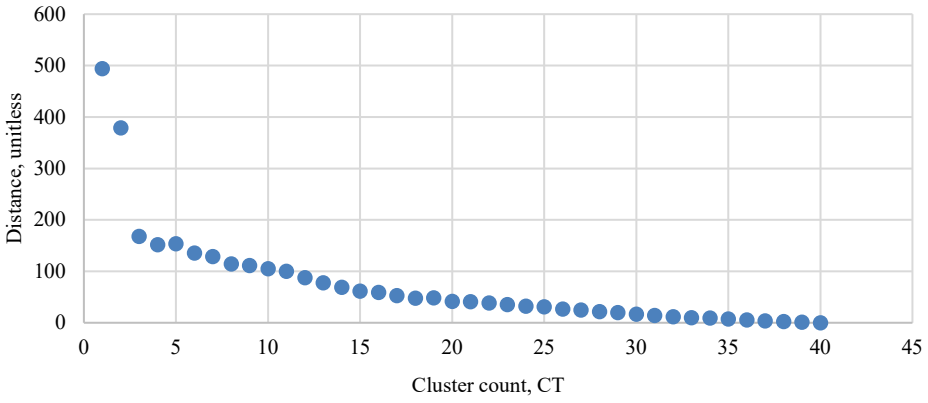


Fig. 5. Cluster combinations of *Delphin* simulations data.

It can be seen if Fig. 5, that there is a significant difference in distance between a cluster count of 2 and a cluster count of 3 when comparing to the remaining cluster combinations.

All the possible cluster counts for hygrothermal properties data, plotted against those clusters' corresponding distances, are presented in Fig. 6.

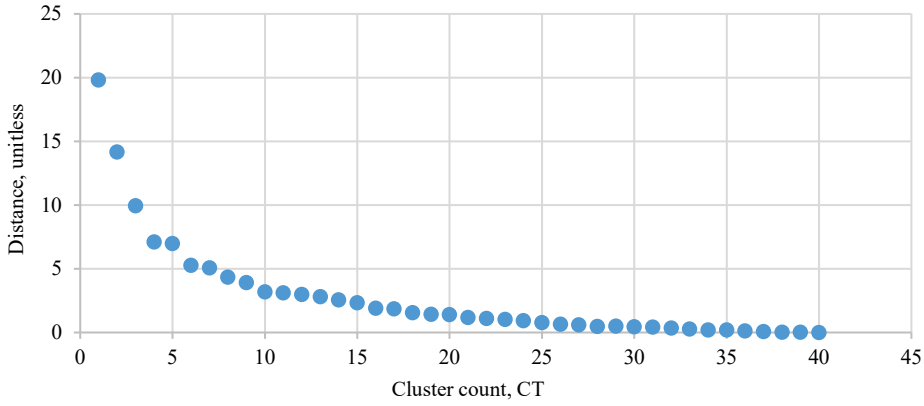


Fig. 6. Cluster combinations of hygrothermal properties data.

In Fig. 6, there is a visible exponential growth of a distance when the cluster count goes down, with the 'knee' being in a combined range of 3 to 10 clusters.

From the cluster combinations of the combined clusters, the optimal cluster combination was found to be the 7th combination, this is the combination of the P cluster of 6 and R cluster of 3, in total 9 original clusters, with the total average normalized distance of 30.3 (see Fig. 7).

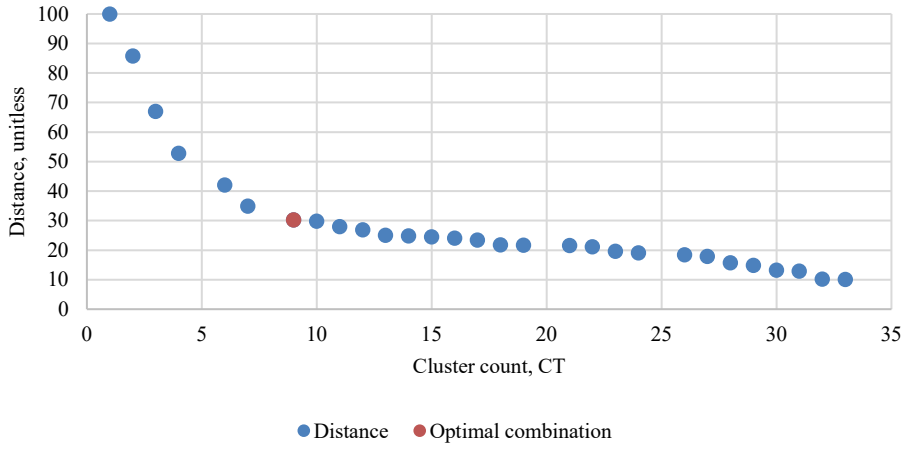


Fig. 7. Distance vs Cluster count of multiplied clusters.

From this cluster combination, 9 new clusters were created (see Table 1). If there is a zero in the box, it means that this combination of the cluster does not exist (there are any brick samples that would fit in this cluster).

TABLE 1. R AND P CLUSTER OVERLAPPING (COMBINED CLUSTERS)

		P Clusters						Total samples
		0	**1**	**2**	**3**	**4**	**5**	
R Clusters	**A**	0	0	1	0	0	0	1
	B	1	3	0	12	4	15	35
	C	2	1	0	1	0	0	4
Total samples		3	4	1	13	4	15	40

In Table 1 the distribution of samples between all the clusters is shown (in total 40 samples). Clusters 'A', 'B', and 'C' are the R clusters, with 1 sample in cluster A, 35 samples in cluster B, and 4 samples in cluster C. Clusters '0', '1', '2', '3', '4', and '5' are the P clusters with 13 samples in cluster 3, 15 samples in cluster 5, etc. After combining P and R clusters new clusters 'A1', 'B0', 'B1', 'B3, 'B4', 'B5', 'C0', 'C1', and 'C3' were created.

To cross-examine created clusters groups with the types of the used brick samples, 11 clusters or groups of the brick samples were generated based on the manufacturing year, color, and material type of the brick. For each group, a unique code was generated from the denotations of description parameters. Bricks came in three colors: Yellow Pink and Red (Y, P, R). Three types of material had been used to make these bricks: Silica, Clay, and Concrete (S, C, CO). The manufacturing year was divided into the groups of 50 years period representing the half-century with the first half being (I) and the second half (II), the century is written with Greek numbers (see Table 2).

TABLE 2. CROSS-EXAMINATION OF CLUSTERS AND BRICK TYPES

		Groups of Brick Type										
		YC 20I	YS 20II	PC 19II	RC 18II	RC 20I	RC 19II	CO 20I	RC 19I	RC 20II	RC 17II	RC 17I
Combined clusters	A2	18-4										
	B0						19-9					
	B1		19-15 19-17 20-3									
	B3				19-7	19-13 19-18 20-5 20-11	18-16 20-6 20-12	19-10	20-2		20-4	20-13
	B4	19-1 19-2 19-5								20-9		
	B5		18-5	18-6	18-7	19-6 19-11 19-12 18-15 20-7	20-1 20-10 20-15 19-4		20-14 20-18	19-16		
	C0				19-8	20-8						
	C1								19-14			
	C3								20-16			

Table 2 can be read the same way as Table 1, only the brick count inside is replaced with the specific code for each brick sample group, meaning, in group B4 there are in total 4 samples (19-1, 19-2, 19-5, 20-9), in group YC20I there is also in total 4 samples (18-4, 19-1, 19-2, 19-5), but in the combination of these two clusters B4YC20I there are 3 samples (19-1, 19-2, 19-5). All the parameters for each brick sample group can be found in the *Zenodo* data repository [22].

4. DISCUSSION

Clusters generated from the *Delphin* results are creating more concise results with only three clusters, and the majority (87.5 %) of the samples being in one cluster. There are twice as many clusters (6) generated from the brick parameters with two dominant clusters containing 70 % of all the samples (37.5 and 32.5 %). After combining the cluster these three dominant clusters (B, 3, and 5) created two new dominant clusters (B3 and B5). In total 67.5 % of all samples were in these two clusters, 30 % in B3 and 37.5 % in B5.

One sample (18-4) has been recognized to be the only sample in both cluster groups. When comparing to the other brick samples based on their description, there are 3 more bricks with the same description (Yellow clay brick, manufactured in the 1st part of the 20th century). Three other bricks belong to group B4 with only one other brick being in this group, and that brick is red clay brick, manufactured in the 2nd part of the 20th century.

The majority of the samples in the largest combined cluster group of B5 are Red Clay bricks (86.7 %), with most of them being manufactured in the period from the 2nd part of the 19th century (30.8 %) to 1st part of the 20th century (38.5 %). In the second largest cluster group B3, almost all the bricks are red clay bricks, except for one brick that is made of concrete, and in this group, the majority of bricks has been manufactured in the period from the 2nd part of the 19th century (27.3 %) to 1st part of 20th century (36.4 %).

But it should be noted that the majority of the brick samples are red clay (75 %) and 60 % are manufactured in the period from the 2nd part of the 19th century (22.5 %) to the 1st part of the 20th century (37.5 %).

Looking at the *clusters from Delphin* results, where 35 of 40 bricks are in a single cluster, it could indicate, that different parameter clusters are generating similar results, in this case, parameter clusters 5 and 4 have 100 % overlap with results cluster B, followed by parameter cluster 3 with 92.3 % overlap, and cluster 1 with 25 % overlap. But it should be noted that the low number of clusters, in this case, 3, from the Delphin results data is due to the high distance of the first two cluster combination with 1 and 2 clusters, 494.17 and 378.56 accordingly, followed by a distance of 168 in combination with 3 clusters. If the first two cluster combinations are excluded from calculations of optimal cluster combination, then the optimal cluster count vs. distance would be the combination with 15 clusters and a distance of 37.5.

As the input data for P clusters are based on the numerical simulation output, these results are highly influenced by the chosen numerical calculation program (DELPHIN) and the performed simulation input data (material properties, chosen geometry, climatic conditions, boundary conditions, etc.), therefore, interpretation of the results is limited to the specific climatic condition and set of brick samples used in this study.

5. CONCLUSION

The distribution of the brick samples in the clusters, is similar to the distribution of the brick sample types, based on their description, therefore overall conclusions regarding the genericization of historical bricks cannot be made. However, regarding the red clay bricks, manufactured from the end of the 19th century till the beginning of the 20th century, there is a strong distinction between two main clusters (B3 and B5) with only two of the samples being in other clusters (B0 and C0).

The aim of the study is reached only partially, in the study 3 clusters for numerical simulation output were found, but no general brick parameters have been yet identified for those clusters, moreover, no methodology for assigning the brick sample to one of the clusters has been found.

Further research with an equal amount of brick samples with different descriptions should be carried out in a future study, moreover, different climatic conditions and geometry should be considered.

ACKNOWLEDGMENT

European Social Fund within Project No 8.2.2.0/20/I/008 'Strengthening of Ph.D. students and academic personnel of Riga Technical University and BA School of Business and Finance in the strategic fields of specialization' of the Specific Objective 8.2.2 'To Strengthen Academic Staff of Higher Education Institutions in Strategic Specialization Areas' of the Operational Programme 'Growth and Employment'.

REFERENCES

[1] Freimanis R., *et al.* In-Situ Moisture Assessment in External Walls of Historic Building using Non-Destructive Methods. *Environmental and Climate Technologies* 2019:23(1):122–134. https://doi.org/10.2478/rtuect-2019-0009

[2] Zhou X., Carmeliet J., Derome D. Influence of envelope properties on interior insulation solutions for masonry walls. *Building and Environment* 2018:135:246–256. https://doi.org/10.1016/j.buildenv.2018.02.047

[3] Zhou X., Derome D., Carmeliet J. Analysis of moisture risk in internally insulated masonry walls. *Building and Environment* 2022:212:108734. https://doi.org/10.1016/j.buildenv.2021.108734

[4] Straube J., Schumacher C. Interior Insulation Retrofits of Load-Bearing Masonry Walls in Cold Climates. *Journal of Green Building* 2007:2(2):42–50.

[5] Haupl P., Jurk K., Petzold H. Inside thermal insulation for historical facades. Research in Building Physics. Boca Raton: CRC Press, 2003:463–469.

[6] Vereecken E., Roels S. Capillary active interior insulation: do the advantages offset potential disadvantages? *Materials and Structures* 2015:48:3009–3021. https://doi.org/10.1617/s11527-014-0373-9

[7] Johansson P., *et al.* Retrofitting a brick wall using vacuum insulation panels: measured hygrothermal effect on the existing structure. *Proceedings of the 10th Nordic Symposium on Building Physics* 2014:1269–1276.

[8] Kehl D., *et al.* Wooden beam ends in masonry with interior insulation–A literature review and simulation on causes and assessment of decay. *CESBP Vienna* 2013:299–304.

[9] Kohta U., Straube J., Van Straaten R. Field monitoring and simulation of a historic mass masonry building retrofitted with interior insulation. *Proceedings of the 12th International Conference on Thermal Performance of the Exterior Envelopes of Whole Buildings* 2013.

[10] Anker N., *et al.* Use of sensitivity analysis to evaluate hygrothermal conditions in solid brick walls with interior insulation. *Proceedings of the 5th International Building Physics Conference (IBPC): The Role of Building Physics in Resolving Carbon Reduction Challenge and Promoting Human Health in Buildings* 2012:377–384.

[11] Zhao J., *et al.* Evaluation of capillary-active mineral insulation systems for interior retrofit solution. *Building and Environment* 2017:115:215–227. https://doi.org/10.1016/j.buildenv.2017.01.004

[12] Sun Y., Haghighat F., Fung B. C. M. A review of the the-state-of-the-art in data-driven approaches for building energy prediction. *Energy & Buildings* 2020:221:110022. https://doi.org/10.1016/j.enbuild.2020.110022

[13] Ramirez R., *et al.* Simulation of moisture transport in fired-clay brick masonry structures accounting for interfacial phenomena. *Building and Environment* 2023:228:109898. https://doi.org/10.1016/j.buildenv.2022.109838

[14] Wasik M., Lapka P. Analysis of seasonal energy consumption during drying of highly saturated moist masonry walls in polish climatic conditions. *Energy* 2022:240:122694. https://doi.org/10.1016/j.energy.2021.122694

[15] Zhou X., Carmeliet J., Derome D. Assessment of moisture risk of wooden beam embedded in internally insulated masonry walls with 2D and 3D models. *Building and Environment* 2021:193:107460. https://doi.org/10.1016/j.buildenv.2020.107460
[16] Kunzel H. M. Simultaneous Heat and Moisture Transport in Building Components. Stuttgart: Fraunhofer IRB, 1995.
[17] Freudenberg P., Ruisinger U., Stocker E., Calibration of Hygrothermal Simulations by the Help of a Generic Optimization Tool. *Energy Procedia* 2017:132:405–410. https://doi.org/10.1016/j.egypro.2017.09.645
[18] Hansen T., *et al.* Material characterization models and test methods for historic building materials. *Energy Procedia* 2017:132:315–320. https://doi.org/10.1016/j.egypro.2017.09.738
[19] Delphin Application [Online]. [Accessed 16.03.2022]. Available: https://www.bauklimatik-dresden.de/delphin/
[20] Jupyter homepage [Online]. [Accessed 16.03.2022]. Available: https://jupyter.org/
[21] Pedregosa F., *et al.* Scikit-learn: Machine Learning in Python. *Journal of Machine Learning Research* 2011:12:2825–2830.
[22] Freimanis R., *et al.* Hygrothermal Properties of historic bricks from various sites of Latvia. *Zenodo* 2021. https://doi.org/10.5281/zenodo.5656966