Scientific Journal of Latvia University of Life Sciences and Technologies
Landscape Architecture and Art, Volume 18, Number 18

DOI: 10.22616/j.landarchart.2021.18.02

Computational Analysis of Biophilic Scale Distributions of Façades in Kaunas City Centre

Marius Ivaškevičius, Huriye Armağan Doğan
Kaunas University of Technology, Lithuania

Abstract. The results of numerous studies which are performed on the concepts of Biophilic architecture demonstrate that it can influence emotional tension and health of the observers. Moreover Biophilic research exhibits that not only natural plants induce biophilic response, but also artificial, human creations with certain fractal dimensions or distributions of scales can have an impact. In that regard, the aim of this research is to describe the relation between measurable Biophilic properties of façades and the emotional tension inducing health problems measured with the count of medical emergency arrivals in the vicinity of the façades. To achieve the aim several tasks were completed, such as the development of a methodology of façade analysis, and application of it in an experiment to test the validity. The engineered features found by this research are based on statistical analysis of distributions of line lengths and distances between lines in a drawing of a façade. To test the methodology, a linear regression model with six features was trained and it achieved a 37 % confidence, measured with R^2 adjusted, predicting the number of medical emergency arrivals. Simplicity of the model allowed to make additional insights into the specificity of façade properties, and their importance to Biophilia, which establishes the scientific novelty and the significance of this research.

Keywords: Biophilia, distribution, universal scale, power law, linear regression, medical emergency

Introduction

Biophilia in its most basic definition is the love of living things and nature, which some people believe humans are born with [18]. Hence, the usage of the term is slightly different in the concept of architecture. In architecture, biophilia represents the approach to design which contains the characteristics of nature that would assist the users to feel more comfortable to be in these places. However, for the people who live in urban settings, it is not easy to feel comfortable and pleased all the time due to the social stress caused by the cities. Furthermore, according to this approach, inclusion of the properties important to Biophilia that usually are found in nature, to the artificially built environment can help to improve the well-being and the performance of the people. More widespread definition of Biophilia and Biophilic architecture is related to the usage of actual plants in the façades or the vicinity of buildings, contrary this is not the whole scope of this term. It also refers to completely artificial façades without usage of plants that have certain fractal properties of distributions of scales, therefore it is possible that the essence of Biophilia is not in the fact that observable objects are alive rather in the proportional distributions that are naturally found in live plants. Although Biofilic proprietes are shown to have an impact on people's emotional tension and improvement of health, there is no widespread practice to design façades that utilise this knowledge. Better understanding of the relation between biophilic properties of the façades and emotional stress, could enable the increase of popularity of less stressful façade design. The creation of a new, more explainable, model using the already known principles of Biophilic design, that could evaluate façades and make statistically measurable predictions, should raise awareness of this unintuitive stress in urban environments.

One of the characteristics of the environment which is proven to affect the well-being and health of people is aesthetic quality. According to Shusterman [15], evaluation of aesthetics occurs by the emotions associated with the bodily and behavioural changes that the objects evoke during the interaction. Therefore, aesthetic quality or features of an object or an artefact can catalyse the interaction. However, there is a lack of control in environmental aesthetics due to two main factors. The first factor is the insufficiency of the available tools which can help to evaluate aesthetic characteristics. The second factor is the belief that aesthetic feeling is only determined by an individual's preferences, and it is subjective. Even though many techniques and methods to create works of art that induce aesthetic feeling are invented, few of them offer means to measure the results. Methods that allow to predict aesthetic feeling are almost never used in professional practice, on the grounds that it is evaluated by the trained preferences. Despite the fact that there are established techniques in academia to train the preferences of the students, the whole system without scientifically proven control techniques is prone to human bias. This bias is powered by a very strong force - the wish to create something new. The history of creation of the artificial environment dates back to prehistoric times, however the creative process can be broken down by combinatorial analysis as a consequence although combinations are many it is a finite number. Bearing in mind all possible combinations, of which massive amounts could be easily and without doubt or loss of accuracy dismissed by the trained preferences as not associated with aesthetic feeling. Correspondingly clearly aesthetical pieces could also be identified with knowledge of art history. Furthermore, the boundary, the grey area containing doubt and uncertainty which is mostly influenced by bias and masked as fashion. This is driven by pure wish to establish new rules, new patterns, and new styles, and limited by exhaustion of original aesthetic patterns. Ability to know exact aesthetic measurement or prediction of artwork under creation, while it is in virtual form before it is realised with expensive materials and even more biased by investors wishing to have profitable investment could be the key to safe expanding of doubtless aesthetical areas of combinational patterns.

Another branch of research related to Biophilia is concentrating in image complexity measurement with fractal dimension [9], most interesting results are achieved when participants' reaction is recorded and juxtaposed with the fractal dimension of the images participants were reacting to. In the forced choice of image preference experiment 220 participants selected natural or computer generated images measuring 1.33 fractal dimension with little variance [16]. Sensitivity to fractal dimension was also measured in the study that tested 31 participants with an EEG recorder. Recordings show that the 1.32 fractal dimension causes increased activity of alpha waves and decreased activity of delta waves. Alpha waves are associated with relaxed state, and delta waves are associated with sleep states, therefore it could be summarised that images with fractal dimension 1.3 made participants more awake and relaxed [6]. But evidence is not consistent as previous studies found that image preference is correlated with fractal dimension [6], additionally Pollock paintings grew in fractal dimension constantly as author painted more [17], and study of the Pollock paintings preference showed that participants preferred painting with larger fractal dimension up to 1.8 [16]. Fractal dimension can be described as slope of the line in the log-log plot of box counts [9]. Understanding it from this standpoint enables various adaptations of the same idea in the applications where box counts are not available.

In this paper, the theory of the power-law is implemented for understanding the correlation between the features of the façades of cultural heritage artefacts and the emotional tension and health problems that can be reflected in the count of medical emergency arrivals in the vicinity of the façades. The paper is following the methodology of Salingaros and West for measuring the prosperities of the size distribution on the façades to analyse the changes and the aesthetic synergy of the adjacent buildings. The paper begins by examining the perception of the built environment both as a reflection of nature and alterations which can change the perception and can be measured by the implementation of the power-law. This is followed by giving information about the method and how it is implemented in the experiment. Furthermore, the paper discusses the results of the implementation of the power-law on the building façades in Kaunas and analyses the changes established by the interventions.

Perception of the built environment and its evaluation

In the contemporary world, people who live in urban environments face several aspects of the city on a daily basis, such as its architecture, urban form, street network, and even social networks. All these different aspects establish a reflection on the perception of the built environment that can be measured by different analysing methods.

In his research Salingaros discussed the perspective of fractality and how to use the correct definition of mathematical fractals for analysing the built environment [11]. His essence laid in two principles. Salingaros named it "universal distribution" and "universal scaling". Universal scaling is the principle which governs how building façades should be subdivided. The subdivisions should be in certain proportions where smaller elements should be repeated more than big proportional to size. It is a simplified description of fractal, and the author does not hide it, in fact fractals are the main motivator for his reasoning. His idea could be simplified to: architecture should follow the same principles as nature. This is even more empathised with non-mathematical or rules of thumb which also draw examples from nature and argue how humans are evolved to perceive such a visual environment. Therefore, the fractal features of structures might have an impact on the perception of the built environment, furthermore, architecture and urban structures built in this way might encourage thriving communities.

As the size of neural avalanches rises, length also rises proportionally, this phenomenon is called universal scaling [5]. Phenomena observed in neural avalanches: power law distribution and universal scaling, correspond to Salingaros terms universal distribution and universal scaling, but the difference is that Salingaros does not argue about neural avalanches, but requirements of building façades [13]. Different scales in rat grid cells also contribute to this field of findings. Similarity of the terms is not coincidence here, phenomena described by the authors in different fields are similar. Salingaros states that façades that do not have these phenomena are not biophilic, and they cause anxiety and illness in humans. By juxtaposing Salingaros theories and finding neuroscientists it is possible to draw the conclusion that the brain is evolved to process information presented in a certain way. There are many power law distributions and universal scales in nature and fractal systems. It is interesting to point out how distributions of brain processes correspond to distributions in nature. If the straight skeleton model is a good representation of how the brain processes visual information, it could be made even better if the skeletons were analysed in the scope of power law distributions.

It is impractical to design an experiment in which one erects an experimental block or even a building just to test how design influences human behaviour. Contrary there are already cities with people and demographic statistics measured, the only thing left is to analyse the buildings. Such a de facto experiment could not bring so clearly stated results as the buildings were not built according to the hypothesis of a paper. Contrary it is possible to settle on less conclusive results for sake of practicality or even the possibility of experimentation.

As Didier Sornette states, nevertheless the uncertainty in life accompanied by potential losses and hazards, it also contains gains. The uncertainty emanates from various dynamical factors, and furthermore, it consists of a succession of choices made by limited knowledge in changing environments [1]. However, even though people have limited knowledge in situations, it is possible to make the choices by predictions which might be shaped by their prior knowledge that people gained through their ancestors. According to Ellard people might not know how to react in an environment that they are not familiar with, especially in the cities, because it is not in the codes of the city people [2]. However, it might be possible to make predictions. Predictions might not always be accurate, but if the distributions are well-defined, the deviations can be calculated. Although, not all distributions fit in a pattern, and furthermore, the process of prediction can also result with the existing uncertainty which generates problems. However, the uncertain characteristic of predictions also makes them interesting and open for scientific observations.

Predictions and logarithmic dependencies can be commonly seen in nature laws. Moreover, it can be ascertained in human senses, such as perception of light intensity, hearing of sound volume and pitch etc. All these senses have logarithmic dependencies by the power spent to produce the events, therefore, it might be possible to predict them. In statistics, these predictions can be calculated by the assistance of the power-law.

As Yaneer Bar-Yam explains, power law is a relationship between two quantities where a relative change in one quantity establishes a proportional relative change in the other quantity, independent of the initial size of those quantities [14]. Therefore, one of the quantities acts as a power for the other quantity. Moreover, the correlation of these quantities can be defined by tail graphs.

Alterations on the form and function and its effect on perception

If the original function of the structure is not convenient anymore, it is inevitable to provide a new function to the existing building, therefore, it can continue its life. The approach which facilitates giving a new function to a building is called adaptive reuse. However, adaptive reuse is not merely needed to give a new function to a structure, but it is an important strategy for the preservation of the heritage buildings. Furthermore, adaptive reuse can also be adopted for the sustainability of the cultural heritage. Cultural heritage establishes continuity in the society, furthermore it passes the cultural identity to the future generations. As Peter Bullen and Peter Love states, the most successful adaptive reuse projects are those that respect and retain a building's heritage significance as well as add a contemporary layer that provides value for the future [10]. Therefore, adaptive re-use can affect the perception of the people towards the structures, furthermore on the sustainability of the environment.

Architecture has the impact of communicating memory and identity; therefore, it has the ability to communicate values and sense of a place which fulfils the requirement of continuity for future generations. Furthermore, in a broader context, the role of cultural heritage and historical buildings become evident in sustainable development. Assigning new uses for the existing buildings which are abandoned can have an impact on the environment. The sustainable preservation of historical buildings needs to contain the blending of sustainable design, sustainable development, the process of assigning new uses for the buildings can be problematic, and it is important to have a research about the building, the impact of it on the environment, and furthermore its potentials before making the final and the most convenient decision. For achieving that, there have been various models and strategies designed by the experts.

According to the experiment performed by Imamoglu, people tend to prefer the intermediate levels of complexity when it is compared to minimum and maximum levels [7].

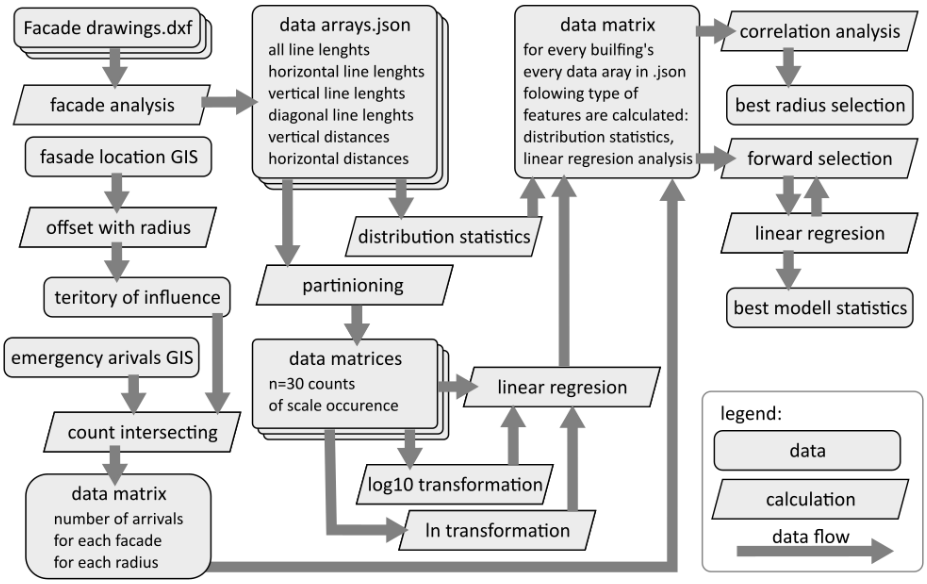

Fig. 1. Data flow overview [created by the authors]

“Organized-complexity: intricate yet coherent designs – and extends to symmetries of abstract face-like structures”. Using the methodology evaluator should choose the mark for each concept: 0–not present, 1– some, 2–abundant. Sum of all marks is a biophilic index showing how healthy influence building makes its users [12]. Although methodology is quite clear and most concepts of evaluation are backed up by experimental research, it is lacking examples of execution. However, the evaluation of the aesthetic preferences can help to understand the approach of people towards heritage buildings in the built environment. When people do appreciate the heritage buildings, they are more likely to spend an effort to protect them. Therefore, what they appreciate and what people do not is an important topic in the process of interventions and alterations for the heritage buildings, both for their life cycle but also for the sustainability. To lay the foundation for the tools that would allow interactive measurement of the façades in the perspective of biophilia, it is required to make a prototype that incorporates and tests all concepts with a small dataset.

Methods

The prototype consists of the following modular parts or modules such as, data acquisition, data preparation, calculation of engineered features, statistical analysis and modelling of medical emergency arrival counts in the vicinity of façades.

Input data of this research consists of 3 parts: façade drawing in dxf format, façade locations and medical emergency arrivals in .shp format. Façade drawings were analysed to create engineered features that will be described in more detail. Engineered features were calculated in two ways: using distribution statistics, and partitioning data with the same method that is used to make distribution plots and making regression analysis of this partitioned data. Main idea of this branch is that log-transformed partitioned data would be linear, and regression analysis would detect that. Façade locations were offseted and medical emergency arrival locations that intersect those offsets were counted and assigned to corresponding façades. Multiple repetitions of this process with varying radiuses allowed selection of the most influential radius that was selected as result of correlation analysis. Final linear regression model was made using forward selection. In the rest of this chapter every step of this process will be discussed in more detail.

There were four sources of data acquisition of the building façades. One of the sources was the architect of the building who provided the .dwg drawings of the façade from his reconstruction project (One of the blueprints of the building was achieved this way). Another one of the other sources was Kaunas municipality, which provided paper copies of several façades from their archive. Some façade drawings were also taken from Kaunas historical archive. Additionally owners of one building provided copies of the pages from the building's reconstruction project. However, most of the façades were drawn as a result of a field survey by one of the authors of this research. The preparation of the drawings followed various steps. The first step was measuring and registering information regarding the structures. In this stage, instead of preparing sketches of the buildings, photographs were used by the author. Firstly, buildings were measured at the site by the assistance of a laser-meter. This was followed by matching the measured distances and the

Fig. 2. Photo of the façade with overlaid drawing [created by the authors]

photographs in Photoshop for scaling the photographs regarding the measurements. After preparing the scaled photographs, the buildings were drawn in AutoCAD as .dwg format files.

Façade analysis pipeline was developed in a modular way allowing the inspection of results of each module as they were saved in files on the hard disk. Additionally modules provide a certain aspect of reusability.

To ensure that data is collected accurately, some steps of data preparation were performed manually. AutoCAD provides numerous objects that enable users to work more effectively. AutoCAD's native format .dxf is close sourced, despite the fact that it is reverse engineered and many details about its inner workings are known, the .dxf format was chosen which is also supported by AutoCAD. It does not support all types of objects that .dxf supports, but instead documentation is available online resulting in several open source libraries that allow interaction with this format being created. Advantage was taken of those libraries and for that purpose files were converted to .dxf format. Façade drawing files were organised to have one façade per each file. Building façade drawings are divided in separate files and saved in AutoCAD open ". DXF" format. The "explode" command was used in AutoCAD to convert most objects to lines. However, not all objects were able to be converted to lines with AutoCAD explode command. In that regard, those objects were treated as special cases in the analysis module.

In the façade analysis module, line lengths and distances between lines were collected, and they were sorted in different groups, such as: vertical, horizontal, and diagonal. There was an additional group where all line lengths were collected without filtering. In the AutoCAD package every line is stored as a separate object. The script which was used in this research accesses all objects, filters lines, extracts coordinates, measures distances, compares coordinates of start points and end points (to detect verticality and horizontality). For every .dxf file contained in directory and generates a .json file containing arrays of line coordinates (for double checking data integrity), line lengths, and distances between lines. Line lengths of vertical and horizontal lines were just absolute differences between corresponding axis coordinates of endpoints. Lengths of diagonal lines were calculated using Pythagoras theorem. Distances between vertical and horizontal were measured the same as distances but between closest lines that have intersecting projections in the perpendicular axis. Although this was a seemingly simple concept, it was somewhat difficult to achieve, therefore a separate module was developed to test it thoroughly. Data was grouped by line orientations into vertical, horizontal, and diagonal. There was no diagonal group for distances, because angles varied and there was no simple concept to measure distances between them.

Several object types**:** splines, circles and arcs do not convert into lines with usage of "explode" command in AutoCAD. However, the research aimed to use the information that these objects hold as well, therefore, the objects were converted before they were sorted out as lines. Spline objects have an array of "fit_points" which are treated as polylines and every line segment is added to corresponding sets. From the circle object coordinates of center point and radius were taken and used to calculate edges of the octagon. Edges were treated as lines and added to corresponding sets resulting in 2 horizontal lines, 2 vertical lines and 4 diagonal lines. From arc object coordinates of center point, radius, start and end angles were taken and used to calculate three lines and added to corresponding sets. Data collected in such way was not analysable with statistical methods that required to have data matrix, because lists of lengths and distances of lines in the drawing varied in length, and data matrix requires everything to be the same length, therefore data was additionally aggregated using various statistical functions that aligned data to fit into data matrix, additionally selected aggregation methods took this research closer to a theoretical background.

Feature engineering

Second module performs analysis of lengths and distances arrays using "statsmodels" library, using Kolmogorov-Smirnov test on a list of available distributions in the library, among which is most important for this paper power law distribution. Some distribution functions require no parameters therefore tests are straight forward, contrary others require one or multiple parameters. Functions requiring multiple parameters were discarded as there were no strong hypotheses attached to them, contrarily one parameter distribution functions were tested 10 times by scanning beta parameters from 0.0 to 2.0. The range was chosen to capture the power law distribution parameter of 1.3-1.5 discussed previously in more detail.

Although several statistical distributions tests derived from power law distribution are also related to fractal dimension there is potentially another way to test it using sets of line lengths and distances. In the essence of fractal dimension analysis, named box counting, there is a fixed size box in which details of analysed drawings are counted. There are other fractal analysis methodologies that are not using boxes, newer the less fixed size of test area is used. In the effort to stay close to this methodology, it was chosen to partition line lengths and distances between lines as it achieved when calculating histogram plots, as bare data cross plots: line counts versus line lengths were very jagged and not analysable. Data was partitioned, by calculating 30 bucket histograms for each array. 30 buckets were chosen as a rule of thumb minimum requirement to perform linear regression analysis, also to provide consistency as array sizes that vary from 33 to 35079. Also following the methodology of calculating fractal dimension as closely as possible correlation analysis is performed on all arrays. One statistic of linear regression analysis, that is of the most interest for this study, is slope coefficient of line in log plot. This is achieved by log transforming the data before linear regression analysis. Contrary this is not enough because literature analysis results call for a certain slope coefficient which is between 1.3 and 1.4 depending on source. Therefore for the scope of feature engineering the new variables were defined as absolute difference of slope coefficient and 1.35.

Sometimes significant parameters are found in unexpected fields: one study found that GIF compression algorithms effectiveness is a good predictor of beauty judgments recorded from participants of experiment [4]. Therefore a decision was made to test as many engineered features as it was possible to generate in an effective manner.

All engineered features have their own ranges of resulting statistical values, therefore without taking into account the individual characteristics of each statistic interpretation of the final regression model would be very complicated. To simplify this problem, all engineered features were normalised to min-max, hence the trade was made to sacrifice the ability to interpret the actual statistic to gain the ability to compare them with each other. Potentially it is possible to make additional compensation for the inability to interpret engineered features by visually comparing façades of extreme examples of the statistical values in the question.

Geographical data: locations of façades and arrivals of medical emergencies were aggregated using GeoPandas Python package. Façades were converted to territories of influence using the buffer method. Resulting offsets were used to count intersecting emergency arrival points using spatial join method. Although this procedure is simple and could be done manually with GIS software, scripting allowed repeatable processes that could be performed on several buffer radiuses with relatively little effort.

Resulting data matrices were joined together based on façade identification codes consisting of shortened street name, number of buildings, and façade orientation. Multiple radiuses of façade influence allowed the selection of the most influential radius using analytical methodologies which was used to make a final linear regression model that allowed the inspection of most influential façade analysis features.

Results and Discussions

Preliminary analysis allowed the identification of two buildings that were outliers in the dataset. One of them was sourced by the author of the building's reconstruction; it was outlier in the engineered features. It is suspected that underlying pipeline, or the philosophy in the process of making façade drawings or using certain functions to draw façades resulted in this extreme discrepancy. Additionally, this building's reconstruction was followed by controversies which resulted in license loss and informal policies change in the organization that is responsible for heritage conservation. Although it is not believed that engineered features of this building could predict the controversies it suffered, it is not possible to test it and this is not subject of this research, therefore it was found safe to remove it from the dataset. Another outlier was in the number of emergency arrivals. It was a large police station, although the under mechanism is not known regarding what is happening, therefore, it is assumed that it is related to organizational operations of the police station and the medical emergency services. Those data points were not filtered by excluding points labelled "patient transfer". Additionally, it could be that victims of crime walk to the station and then a policeman calls for a medical emergency. For the reasons discussed above, it was considered safe to remove this building from the dataset as there is high probability that the accumulation of emergency arrival points near it has more to do with label limitations of the emergency arrivals database or the organizational reasons. Those outliers were affecting correlation analysis in numerous variable combinations, as it was visually apparent in the scatter plots. Therefore, the following analysis results will be with the dataset that has those outliers removed.

To select the most influential radius correlation analysis was repeated for every radius and the best feature along with its correlation was selected.

In the figure nr. 3 it is possible to see that best parameters are found by scanning through distance

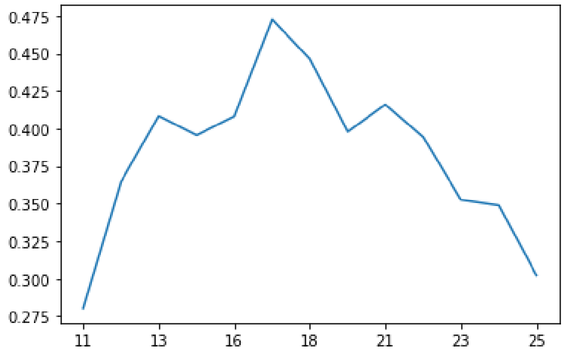

Fig. 3. Best Pearson correlation change [created by the authors]

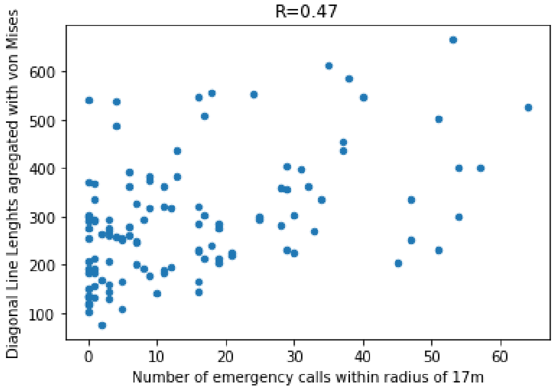

Fig. 4: Scatter plot of engineered features calculated from diagonal line lengths using the von Mises statistic versus number of emergency arrivals in 17 meter radius[created by the authors]

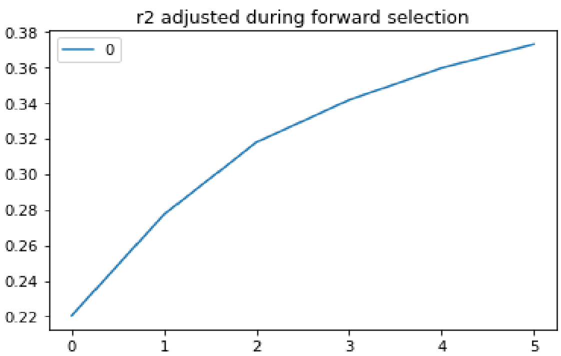

Fig. 5.Change of R^2 adjusted in the best model during the forward selection process [created by the authors]

radiuses. The features that produce best correlation values differ depending on radius. For the radius of 11 meters best correlation was calculated with all line lengths with inverse gaussian statistics. For the radius range from 12 to 24 meters the best correlation is for diagonal lengths that was calculated with von Mises statistic. From 12 to 17 meters the dominating component was uniform distribution, but for 23 and 24 meters the normal distribution component was more important. For 25 meters and above horizontal lines were dominating. Most relevant features will be discussed later. This analysis is used for best radius selection, and later analysis was concentrating on the model's radius with best correlation that is 17 meters.

Engineered features calculated from diagonal line lengths using the von Mises statistic had weak Pearson correlation with the number of emergency arrivals in the 17 meter radius (R^2=0.47). In the scatter plot (Fig. 4) a slight visual trend is visible, a linear model with only this feature would have high error values, nevertheless it is possible that other features will compensate for that noise and make the model more accurate.

Using the most influential radius which was selected using correlation it was possible to make regression analysis including more engineered features and hopefully produce a model with less error. Linear regression models are very robust in usage because they can be explained and understood, contrary to many newer machine learning approaches. Despite the existence and spectacular results of machine learning specialisation named “explainable AI”, more difficult models do not approach the level of expandability linear regression has. Although it is possible to make a regression model using all engineered features it would not be very comprehensible considering the number of features and the probability of features not contributing to models accuracy. There are multiple methodologies available for feature selection, upon trying several of them stepwise selection was selected for the final model. Despite the critique of the forward selection model it is also known that it is producing acceptable results when sample size is low and suggested models to replace stepwise selection are not suitable for such small sample size [3]. Forward selection is a type of stepwise selection model, where one starts with a small model with only one feature and tries to find the next one that increases chosen measurement most. In this case it is R^2 adjusted. To describe in more detail, the algorithm consists in two nested cycles, a deeper one which is repeated most times, cycles all features in the “available” list, makes linear regression models with them one by one and measures R^2 adjusted. Then a higher order cycle which only repeats as many times as it chooses to select features, picks the feature that produced the best model and removes it from the “available” list, also adds it to the “selected” list. Therefore, in the next iteration of deeper cycle models are made including all features in the selected list and one feature under test from the available list. This is time consuming but useful when there are many features available as in this case.

Dynamics of R^2 adjusted was measured during forward selection algorithm execution, which shows

Fig. 6. Drawing of Laisvės al. 94 façade that ranked near the top of ranked facades by the feature that was used for the second coefficient of the model [created by the authors]

Fig. 7. Drawing of Laisvės al. 88 façade that ranked near the bottom of ranked facades by the feature that was used for the second coefficient of the model [created by the authors]

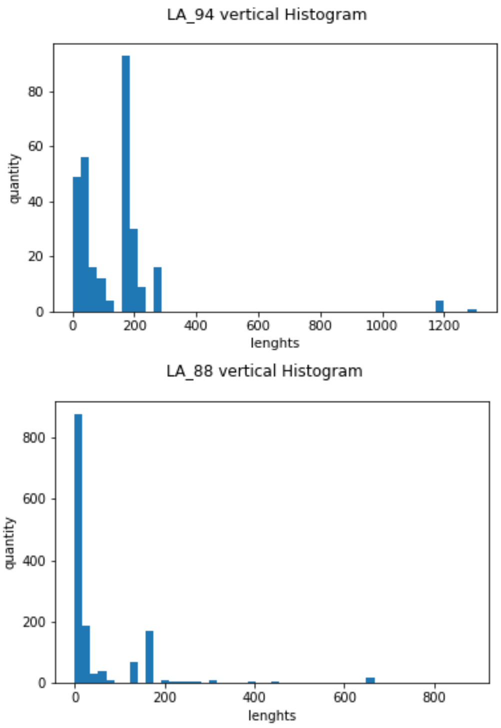

Fig. 8. Histograms façades ranked high and low in features derived from vertical line lengths [created by the authors]

that every next feature that is selected improves the model with a lesser amount. Speed of improvement decreases. The algorithm was stopped as 6 features were selected to achieve balance between accuracy and explainability of the model. Additional motivation to stop is to address the flaw of the forward selection model as it was shown that there is probability to select features that do not improve understanding of actual phenomena [3].

Selected final models accuracy measured by R^2 adjusted statistics can be interpreted in the following way: using six engineered features from the façade analysis model can explain 37 % of emergency arrivals in locations of the façades, and inductively it can explain occurrence of health issues with similar degree. Therefore, it would be possible to create façades by using its knowledge with the intention to reduce a similar amount of emergency calls by removing the cause of health issue occurrence. In the following paragraphs engineered features will be discussed to gain insights of how the knowledge gained by the model would be possible to use in improvement of façade designs that reduce stress and health issues. It would not be correct to make conclusions about the model in the classical approach because the model's endogenous variable does not have predictive purpose, or at least argument could be made that it would be more correct to treat it this way. It is possible to interpret the coefficients by forming such a conclusion that reducing this coefficient by one, emergency per certain time would be reduced by a certain number. Instead, a decision was made to treat them as unitless numbers of potential to influence human stress and health issues. Although this approach omits potential explicitness that model provides it also makes it simpler, which will help to include explanations of meaning of engineered features.

Linear regression model has an intercept which graphically represents the point on the Y axis where the line of the model intersects when all the features are equal to zero. It is negative, therefore it can be assumed that the model predicts that there will not be such a case that the façade will have all engineered features equal to zero, because it is impossible to have negative emergency arrival numbers.

The strongest parameter was assigned to a feature that is calculated by testing diagonal lengths with von Mises (beta=0.05) statistic. This statistic is the result of a test that tests for von Mises distribution also known as circular normal distribution. As it is parametric distribution tests, as parameter beta is approaching zero, the test approaches the state where it tests for uniform distribution. Considering that the engineered feature was generated using it with beta=0.05 the distribution under test is a partial distribution

Fig. 9. Drawing of Kęstučio g. 15 façade that ranked near the top of ranked facades by the feature that was used for the third coefficient of the model [created by the authors]

Fig. 10. Drawing of Laisvės al. 53 façade that ranked near the bottom of ranked facades by the feature that was used for the third coefficient of the model [created by the authors]

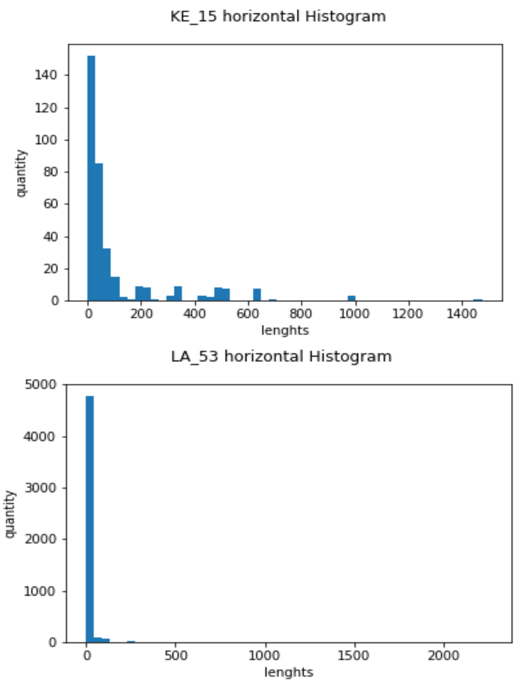

Fig. 11. Histograms façades ranked high and low in features derived from vertical line lengths [created by the authors]

between uniform and circular-normal distribution, being much closely similar to uniform. Although circular-normal distribution was an important component for winning the forward selection process it is minuscule and adds additional complexity in interpretation, therefore the interpretation of this component will be dismissed. To explain equal distribution white noise is often used as an example, where all the frequencies of sound are on the same volume. White colour is where the electromagnetic waves approach our eyes in the same levels on all visible frequency bands. Uniform distribution is when there is a similar count of data values in all data ranges. Looking from probability perspective: the probability of encountering any value in the sample is the same. To interpret this feature it could be stated that uniform distribution of diagonal line lengths in the façade is inducing stress and can trigger health problems. Regression resulted in coefficient 43 which will not be interpreted classically but remembered as benchmark value for comparison with other features in the future. Unfortunately it was impossible to gain intuition visually inspecting façades raked by this engineered feature.

The Second strongest parameter was assigned to the feature that is calculated by normalising vertical lines with base 10 logarithm and testing them with F-statistic after the regression analysis of distribution curve, which is testing if the regression model has predictive capability. Larger value means better predictions and more linearity caused by logarithmic transformation of data. It follows the assumptions made in literature analysis analogously as the first feature. F-statistic is calculated including the slope coefficients of the model, therefore slope of the linear model is important for this feature. Regression resulted in a coefficient 35 which compared to this model's strongest coefficient that is 43. It has 19% less influence on average. Visually inspecting façades in the extremes of the list ranked by this engineered feature we can identify some key properties. As this coefficient is positive higher ranked façades are believed to be inducing more stress to the spectators. Façade Laisvės al. 88 (Fig. 7) has decorations and more variation in the frames of windows and ranked low for this feature, opposed by façade of Laisvės al. 94 (Fig. 6) which has monotonic modern style and ranked high for this feature. The same conclusion can be drawn visually inspecting histograms where in a high ranking histogram we can see a second strong peak which is even stronger than the first peak. This is caused by a large amount of middle length lines. This histogram differs greatly from power-law. By comparison histogram of façade that ranked low in this feature is more compliant to power law with few exceptions. Those exceptions could be the reason why the actual best engineered feature is not a statistical indicator of power-law distribution test.

The third strongest parameter was assigned to a feature that is calculated by normalising horizontal line lengths with natural logarithm and testing them with Jarque–Bera test [8]. Jarque–Bera test is test of normality in the sample distribution, it produces its main statistic which to interpret it sample size must be taken into account, but algorithm selected not its main output but p-value which is probability that the null hypothesis is rejected, therefore sample size does not affect it. Larger result Jarque–Bera test suggests that sample has more normality, therefore sample before log-normalisation was lognormal and not distributed according to power law. As this number increases the distribution line moves from (hypothetically) power-law distribution to lognormal distribution. This could be caused by lack of small details confirming the literature analysis results. Additionally, it provides insight into most common distribution nonconformity with power-law, which is so common that it is a better predictor than test for power-law distribution itself. Regression resulted in coefficient 16 which can be compared to models' strongest coefficient that is 43. It has 63% less influence on average compared to the most influential feature and 55% less influential than the second feature. Visually inspecting façades in the extremes of the list ranked by this engineered feature we can identify some key properties. Façade Laisvės al. 53 (Fig. 10) has decorations and more variation in the windows horizontally and it ranked low for this feature, opposed by façade of Kęstučio g. 15 (Fig. 9) which has no decorations and few types repeated windows and ranked high for this feature. The same conclusion can be drawn visually inspecting histograms where in a high ranking histogram we can see multiple small peaks in the distribution contradictory to low ranking distribution, which has exceptionally high count of small detail and no visually detectable peaks after. High peak in short horizontal lines must have been caused by the decorated handrail in the balcony.

The fourth strongest coefficient was assigned to a feature that is calculated by testing horizontal distances with von Mises (beta=0.05) statistic. This feature was discussed before, therefore to conclude having in mind that a beta of 0.05 makes this statistic largely a test for uniform distribution it can be concluded that when distribution of horizontal distances between vertical lines approaches uniform distribution, the façade is more likely to induce stress and can trigger health problems. Visually inspecting façades in the extremes of the list ranked by this engineered feature we can identify some key properties. As this coefficient is negative lower ranked facades are believed to be inducing more stress to the spectators. Façade Mickevičiaus g. 45 (Fig. 13) has decorations and more variation in the distances between windows and decorations horizontally and it ranked low for this feature, opposed by façade of Kęstučio g. 3a (Fig. 12) which has no decorations and uniformly distributed windows and ranked high for this feature. The same conclusion can be drawn visually inspecting histograms where in a low ranking histogram we can see a second semi strong peak in the distribution contradictory to low ranking distribution, which has one high peak and little noise after.

The fifth potential was assigned to a feature that is calculated by testing horizontal distances and testing them with F-statistics after the regression analysis of the distribution curve. F-statistics was discussed before in the previous paragraph about the second strongest coefficient, but the difference here is that data was not log-transformed. Therefore if the distribution of the distances of horizontal distances are likely to be modelled with a linear model, then the façade is more likely to induce stress and can trigger health problems. Façades ranked similarly but in reverse compared to facades discussed in the previous paragraph. It is expected as the source of distribution is the same but coefent in the model of previous feature was negative, and for this feature is positive.

The sixth feature by absolute potential is calculated by normalising diagonal lengths with natural logarithm and testing them with Jarque–Bera test. Jarque–Bera test was discussed before in analysis of the third feature. Interesting thing about this feature coefficient is that it is negative, therefore it means logo-normal distribution in the diagonal line lengths is reducing stress and probability to trigger health problems. As with previous feature that were sourced from diagonal lines it was impossible to gain visual insights from top and bottom ranked façade comparisons.

Variety of sources of feature calculations that were selected by forward selection indicate that all filters of lines were important for model improvement. Specialisation of the engineered feature by choice of certain statistical function or method allowed to capture most relevant information about the feature. It is indicated by the fact that for the next feature a different type of information source was chosen. For example, the strongest coefficient was for diagonal line lengths and another feature calculated from diagonal line length source was not chosen twice before other sources were depleted.

Buffers with radiuses are a poor visibility model. Although it definitely could be improved it is not as bad as it could be perceived at the first glance. It is true that it could include medical emergency points from which the façade is not visible, although from manual inspection it is possible to state that it is

Fig. 12. Drawing of Kęstučio 3a façade that ranked near the bottom of ranked facades by the feature that was used for the third coefficient of the model [created by the authors]

Fig. 13. Drawing of Mickevičiaus 45 façade that ranked near the top of ranked facades by the feature that was used for the third coefficient of the model [created by the authors]

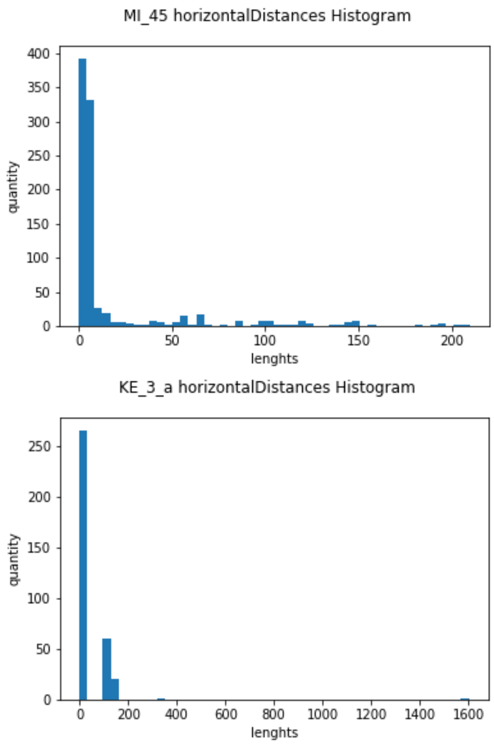

Fig. 14. Histograms façades ranked high and low in features derived from horizontal distances between lines [created by the authors]

rarely the case. This is mainly caused by the physical possibility of arrival location. For other cases it can be explained that the façade was perceived just before the emergency.

Conclusions

Final model can explain 37 % of emergency arrivals just from façade analysis methodology developed during this research. This number alone should be enough evidence that façade design should not be taken lightly. It should be used as motivation to develop further analysis methods and tools that could be used by architects to design façades with stress reduction in mind.

It is interesting that almost all possible strategy variants to calculate façade measurements (except all line lengths and vertical distances), played a significant role in the modelling. Strongest being lengths of diagonal lines, and lengths of vertical and horizontal lines following. Distances between the lines were less important than lengths of the lines, however, lengths of the vertical lines were found to be more important than lengths of horizontal lines. Contrary, horizontal distances were more important than vertical distances, as vertical distances were not selected with the forward selection model, but horizontal distances were selected twice. It still makes vertical lines more important because horizontal distances are measured between vertical lines.

The findings regarding the impact of the lines can be an essential element to use on the design process of the façades of new buildings. However, it would not only affect the new designs, but it can help to make the decisions regarding the alteration of already existing buildings as well. Even though it might be thought that the mass gives the effect of being vertical or horizontal in a design of a structure, actually the small elements have an impact as well, as it was demonstrated in this research. Even the divisions of the window frames and the transition which is gained by ornamentation or any other element can assist the perception of the façade. Assignation of different statistical tests and functions in a final model indicated that different orientations are treated differently, not only by the model but intuitively also by the human mind.

This approach of not looking at adjacency in the perceived façade is somewhat similar to the “bag of words” methodology used in computational linguistics. It is a useful concept that is used in several methodologies in this field to achieve good results. As the computational linguistic models improve when adjacency is included, this model could be also improved. Contrary to such a model much more data is required and it itself would be significantly complex and less interpretable.

Min max normalisation of engineered features made this model specialised to this particular dataset. Any façade that pushes statistics beyond the range of current data will be scaled beyond zero and one limits of the scaled data. This poses no problem because the model is linear. But it raises a question about the meaning of those features, which is now acquired by comparing the façades with extreme values. This is possible continuation of the research, to procedurally generate parametric façades, for the purpose of finding extreme values that still work as façades.

The difficulty of acquiring intuition that is required to use the findings of this research in practice is untested. It should be possible to make a diagnostic tool that automatically shows values of important features, additionally uses predictive capability of this model to help gain this intuition. Such a tool could be integrated as a plugin in a selected CAD or BIM program and provide diagnostic values for the user by analysing the façade under design.

Model potentially could be improved by the addition of omitted statistical functions. There are 99 statistical functions for calculating distribution statistics in the SciPy library. They were not all used for feature engineering trying to avoid feature explosion caused by two dimensional scan of the parameters. Continuing the same strategy as it was used in this study to scan parameter 10 fold, two dimensional scan would produce 100 features, utilising all 25 double parameter functions would result in additional 2500 features.

Another possible improvement could be made by utilizing parametric log-transformation before the linear regression model of partitioned line length and distances.

Most top and bottom ranked façades by features used in the final model were identifiable visually from the histograms of corresponding feature sources, namely lengths or distances by similar properties like: presence and size of multiple peaks in the histogram. Contrary functions that produce best aggregations of those distributions differ. Therefore it should be possible to find or make new statistical functions that would be specialised to diagnose scale distributions in façades.

Acknowledgments

We are grateful to Public Enterprise: Medical Emergency Station of Kaunas city (*Viešoji Įstaiga Kauno Miesto Greitosios Medicinos Pagalbos Stotis*), for providing anonymised data of geolocated emergency arrivals. We are grateful to Joint-stock company JAS for providing architectural drawings of the façade they designed in DWG format. We are grateful to the Urban and Architectural department of Kaunas Municipality for providing photocopies of some historical façades from their archive. We are grateful to Joint-stock company “Spaineta” for providing photocopies of the building in their disposition of historical façades from their archive.

References

1. **Sonette, D.** Critical Phenomena in Natural Sciences: Chaos, Fractals, Self-organization and Disorder: Concepts and Tools, Berlin: Springer Publishing, 2006, p. 1–7.
2. **Ellard, C.** Places of the Heart: The Psychogeography of Everyday Life. New York: Bellevue Literary Press, 2015, p. 32–35.
3. **Flom, P., Cassell, D.** Stopping Stepwise: Why Stepwise and Similar Selection Methods Are Bad, and What You Should Use, Presented at the NorthEast SAS Users Group Inc 20th Annual 8 Conference, 2007.
4. **Forsythe, A., Nadal, M., Sheehy,N.,** et al. Predicting beauty: fractal dimension and visual complexity. British Journal of Psychology, 2011, No.102, p.49–70.
5. **Friedman, N., Ito, S., Brinkman, B. A. W.,** et al. Universal Critical Dynamics in High Resolution Neuronal Avalanche Data. Journal of Physical Review Letters, 2012, No.108(20), p. 1–5.
6. **Hagerhall, C. M., Purcell. T., Taylor, R.** Fractal Dimension of Landscape Silhouette Outlines as a Predictor of Landscape Preference. Journal of Environmental Psychology, 2004, No.24, p. 247–255.
7. **Imamoglu, Ç.** Complexity, Liking and Familiarity: Architecture and Non-Architecture Turkish Students’ Assessments of Traditional and Modern House Facades. Journal of Environmental Psychology, 2000, No.20, p. 5–16.
8. **Jarque, C. M., Bera, A.K.** Efficient Tests for Normality, Homoscedasticity and Serial Independence of Regression Residuals. Journal of Economics Letters, 1980, No.6, p. 255–259.
9. **Mandelbrot, B.** How Long Is the Coast of Britain? Statistical Self-Similarity and Fractional Dimension. Journal of Science, 1967, No.156, p. 636–638.
10. **Plevoets, B., Van Cleempoel, K.** Adaptive Reuse as a Strategy Towards Conservation of Cultural Heritage: A Literature Review, Southampton: WIT Press, 2011, p.155–164.
11. **Salingaros, N.** Connecting the Fractal City. **In:** Principles of Urban Structure. Amsterdam: Techne Press, 2005, p.15-17.
12. **Salingaros, N.** The Biophilic Healing Index Predicts Effects of the Built Environment on Our Wellbeing. Journal of Biourbanism, 2019, No.8, p. 13–34.
13. **Salingaros, N., West, B. J.** A Universal Rule for the Distribution of Sizes. Journal of Environment and Planning B: Urban Analytics and City Science,1999, No. 26, p.909–923.
14. **Schroeder, M.** Fractals, Chaos, Power Laws. AIChE Journal, 1991, No.38, p. 159–160.

15. **Shusterman, R.** Affective Cognition: From Pragmatism to Somaesthetics. Journal of Intellectica, 2013, No.60, p. 49–68.
16. **Spehar, B., Clifford, C. W. G., Newell, B. R.,** et al. Universal Aesthetic of Fractals. Journal of Computers & Graphics, 2003, No. 27, p. 813–820.
17. **Taylor, R. P., Micolich, A.P., Jonas, D.** Fractal Analysis of Pollock's Drip Paintings. Journal of Nature, 1999, No.399, p.422–422.
18. Cambridge Advanced Learner's Dictionary with CD-ROM. 4 edition. Cambridge, Cambridge University Press, 2013.

AUTHORS:
Marius Ivaškevičius, PhD Candidate Kaunas University of technology. E–mail: marius.ivaskevicius@ktu.lt
Huriye Armağan Doğan, PhD of History and Theory of Art and Architecture, Junior Researcher at the Institute of Architecture and Construction, in the Research Centre of Architecture and Urbanism of Kaunas University of Technology; Tunelio g. 60, Kaunas LT–44405, Lithuania. E–mail: huriye.dogan@ktu.lt

Kopsavilkums. Daudzu pētījumu rezultāti, kas veikti par biofilās arhitektūras jēdzieniem, parāda, ka tā var ietekmēt novērotāju emocionālo spriedzi un veselību. Turklāt biofīlie pētījumi liecina, ka ne tikai dabiski augi izraisa biofīlu reakciju, bet var ietekmēt arī mākslīgi cilvēka radīti darbi ar noteiktiem fraktāļu izmēriem vai mērogu sadalījumu. Pētījuma mērķis ir aprakstīt saikni starp izmērāmām fasāžu biofīlām īpašībām un emocionālo spriedzi, kas izraisa veselības problēmas, veicot mērījumus fasāžu tuvumā. Mērķa sasniegšanai tika izpildīti vairāki uzdevumi, piemēram, fasādes analīzes metodikas izstrāde un tās pielietošana eksperimentā. Šajā pētījumā atrastās inženierzinātņu iezīmes ir balstītas uz statistisku analīzi par līniju garumu sadalījumu un attālumiem starp līnijām fasādes zīmējumā.