

GENETIC ALGORITHM AS AUTOMATED VALUATION MODEL COMPONENT IN REAL ESTATE INVESTMENT DECISIONS SYSTEM

Aneta Chmielewska

*Institute of Spatial Economy and Geography
University of Warmia and Mazury in Olsztyn, Olsztyn, Poland
e-mail: aneta.chmielewska@uwm.edu.pl, ORCID:0000-0001-5888-5874*

Jerzy Adamiczka

*Adamiczka Consulting
e-mail: jerzy@adamiczka.pl*

Michał Romanowski

*Independent researcher
e-mail: michal.romanowski87@gmail.com*

Abstract

Every real-estate related investment decision making process calls for the careful analysis of available information even though it is often carried out in conditions of uncertainty. The paper attempts to minimize the impact of the factor on the quality of real estate investment decisions through the proposal of application of tools based on the simulation of the process of natural selection and biological evolution. The aim of the study is to analyze the potential of methodology based on genetic algorithms (GA) to build automated valuation models (AVM) in uncertainty conditions and support investment decisions on the real estate market. The developed model facilitates the selection of properties adequate to the adopted assumptions, i.e. individuals best suited to the environment. The tool can be used by real estate investment advisors and potential investors on the market to predict future processes and the proper confrontation of past events with planned events. Even though genetic algorithms are tools that have already found particular application on real estate market, there are still areas that need further studies in the case of more effective uses. The obtained results allow for the possibilities and barriers of applying GA to real estate market analyses to be defined.

Key words: *genetic algorithm (GA), uncertainty, decision support system, automated valuation model (AVM), real estate market.*

JEL Classification: *C00, C10, D81, R30.*

Citation: Chmielewska, A., Adamiczka, J., & Romanowski, M. (2020). Genetic algorithm as automated valuation model component in real estate investment decisions system. *Real Estate Management and Valuation*, 28(4), 1-14.

DOI: <https://doi.org/10.1515/remav-2020-0027>

1. Introduction

Processes which are shaping the real estate market include, among others: instability of property attributes, lack of information coherence, heterogeneity in information access (Levitt & Syverson, 2008), deficiencies in cognitive abilities of real estate market entities (Burnside, et al., 2016) uncertainty of system structures (Su, et al., 2016), and the emotional approach of entities to transactions. Other factors that cause uncertainty in the real estate market include cognitive biases, such as the effect of

unnecessary information and the effect of attachment, which are the consequences of various types of heuristics and may adversely affect decision-making processes (Czechowska, 2014).

Decision-making regarding real estate is subject to great uncertainty and is very complex. Reducing the risk of making wrong decisions, and thus reducing uncertainty, can be achieved by the precise specification of goals and methods adapted to the specific circumstances of the analyzed phenomenon (Bertalanffy, 1984). Moreover, it is possible to minimize risk and uncertainty by providing relevant knowledge, processed in a specified manner and given to decision-makers. Decisions are made by analyzing information that is available for a given problem and selecting the best alternative for further action. The excess of seemingly relevant information, i.e. so-called information noise, may disrupt interpretations and develop analyses. By identifying the degree of uncertainty, we can develop methods that limit, control and account for uncertainty in the decision-making process (RAO 1994). For this reason, market analyses require solutions and methods that reflect reality, account for and, most importantly, minimize the degree of uncertainty (Ziółkowski & Niedostatkiewicz, 2019; Janowski, et al., 2018; Brzezicka, et al., 2019). The availability of information, market specificity and unpredictable and sudden changes cause all real estate investments to be subject to considerable risk and uncertainty. In the case of random events describing the image of the real estate market, tools and methods that draw inspiration from processes occurring in nature, especially those based on the concept of natural selection, find application. Genetic algorithms, which are procedures based on the basic mechanisms of biological evolution (Ahn, et al., 2012; Zavadskas, et al., 2017; Wu, et al., 2018; Kontrimas & Verikas, 2011; Chodak & Kwaśnicki, 2002) are an example of such.

The article presents the operation of the genetic algorithm as a method used to search for solutions to problems that are difficult to define in a mathematical sense. The original value of the paper constitutes a genetics algorithm application proposal in a concept of AVM as an original decision-making support system for individual (dedicated) investment consulting. AG, as a representative of evolutionary computational methods, takes into account the randomness of processes occurring in real estate market analysis.

2. Literature review

The complexity of the real estate market results, among others, from the diversity and imprecision of the attributes of space and the large and multidimensional scope of data analysis. There are many factors affecting the uncertainty of the real estate market. The valuation itself is also subject to high uncertainty, which results from the fact that land may be subject to changes in zoning and a property valuer may face the problem of determining the value of underdeveloped land affected by possible changes in land regulation (d'Amato, et al., 2019). According Renigier -Biłozor et al. (2018), these are factors such as the diversity and imprecision of spatial attributes, multidimensional scope of data to be analyzed, sensitivity of properties to environmental and economic changes and fashion, as well as heterogeneity with respect to the nature and type of individual objects. Stokey (2016) in his study gives an example of uncertainty in the context of tax policy. He assumed that uncertainty about a future tax rate creates uncertainty about the profitability of the investment. Stokey proposed a universal model of investment decisions in which uncertainty about a onetime change in tax policy induces the firm to temporarily stop investing and adopt a wait-and-see policy. Moreover, Giudice et al. (2017) proposed a procedure using numerical integration on the weights space with the Markov Chain Hybrid Monte Carlo Method, a neural network model, traditional multiple regression analysis and the Penalized Spline Semiparametric Method. This approach considers the uncertainty in parameter estimates and can extend and transform this uncertainty into predictions. As stated by Zavadskas et al. (2010), in many decisions, the consequences of the alternative courses of action cannot be predicted with certainty. Investors on the real estate market are uncertain about how returns will be generated, if a particular investment is chosen. To deal with uncertainty, subjective and imprecise data, quantitative and qualitative assessments are often required. In the normative view the decision problem can be translated into a mathematical model, future processes are stable and can be determined with an appropriate probability. In conditions where it is clearly possible to determine the functions of the target, it is possible to find an optimal solution, which however is difficult to achieve (Van Groenendaal, 2002). Current problems in the real world include complex information - accurate and objective, but also subjective and error-prone (Zavadskas, et al., 2018; Helbich, et al., 2014). In order to support decision-making processes, decision support systems are used, i.e. computer systems

entering the sphere of artificial intelligence, which combine information from various sources, facilitate the organization and analysis of information, and allow easier assessment of the assumptions of specific models (Kaklauskas, et al., 2015, Zavadskas & Turskis, 2011). In the area of real estate market analysis and investment consulting, AVM is used. According to IAAO Standard on Automated Valuation Models, AVMs were defined as mathematical computer programs used by market analysts to determine the value of real estate based on location research, market conditions and real estate features based on collected information (IAAO, 2018; RICS, 2013; EMF/EAA, 2016). For example, AVMs may use regression, adaptive estimation, neural network, expert reasoning, and artificial intelligence programs. The achieved results can be used as a basis for determining the value of real estate only if the real estate appraiser is sure about the quality and significance of the input data, that is whether the output data will be credible and sufficient to perform the task. According to EVGN 11 (2017) and USPAP (2016), analysis of market evidence is possible using sophisticated IT tools such as: regression analysis (both linear and non-linear, time series analysis, geographically weighted models, simulation models (e.g. Monte Carlo simulation), neural network models, option pricing models, fuzzy logic-based models and other machine learning. The development of quantitative methods in social and economic sciences is caused by the need to standardize valuation methods and consistent approaches to the valuation of real estate (Kauko & D'amato, 2008). The International Accounting Standards/International Financial Reporting Standard (IAS/IFRS) and the International Valuation Standard (IVS) address this issue. Guidance Note n. 13 of IVS (IVSC 2005) contains a record: "The Development of Mass Appraisal Systems for property taxation should follow recognized scientific standards in statistical applications(...)". It is widely recognized today that the development of urban areas requires sustainability (Kauko & D'amato, 2008; Kauko, 2017). To achieve this, it is necessary to use of high-quality data cross-sections and a robust modelling tool. In most countries, the use of AVM is limited. In case when data are not available other kinds of AVMs with a less deterministic relation between property value and property attributes may help. It is very important to put emerging methods based on evolutionary theory (not included in the standard) in the context of normally accepted methodologies.

The specific character of information defining the real estate market and real estates is linked with the availability and imprecision of property information as well as the sudden and unpredictable changes, that often occur regarding analyses of properties and their location. The most important obstacles of gathering, processing data and elaborating the reliable analyses of the real estate markets was precisely described by Renigier-Bilozor et al. (2017). It should be clearly emphasized that the specificity of information about real estate should enforce the search for and use of methods of analysis that "understand" this specificity and allow for their inclusion, without adjusting the data to the chosen analytical method.

3. Research methodology

The elaborated procedure of a decision support system in property investment was based on two methods: genetic algorithm and Hellwig's method, both of which have been described below.

3.1. Genetic algorithm

The genetic algorithm (GA) was developed by John Holland in 1975. The goals of his research were twofold: (I) to abstract and rigorously explain the adaptive processes of natural systems, (II) to design artificial systems software that retains the important mechanisms of natural systems (Goldberg, 1989). The terminology used in genetic algorithms has been adopted from genetics as follows: population - set of individuals (search space), individual - set of chromosomes (search space, task parameter), chromosome - information set (genes) represented by a code sequence, gene - single element of the chromosome (information), and phenotype - a set of decoded values corresponding to the genotype. In the population, each individual is described by a set of chromosomes consisting of individual genes. A set of chromosomes of a single individual creates a phenotype that is subjected to the assessment of adaptation to the environment (assessment of the degree of solving a given problem). The classical genetic algorithm consists of the following steps (Rutkowski, 2009): initiation - selection of the initial chromosomal population, assessment of chromosome adaptation in the population, condition check, chromosome selection, use of genetic operators (crossover, during which the genetic material is exchanged between individuals, mutation, i.e. random gene change and selection, which

relies on the selection of individuals for further processing), creation of a new population and derivation of the "best" chromosome. The idea of evolutionary methods is to use the principle of survival of individuals in a given population, which compete with each other and cross in a random way. The evolutionary principle of survival means that individuals "fitted in a better way" have a better chance of surviving and delivering offspring. In the next generation, the number of "good" genes is greater, and the offspring are better "fitted" to the environment (Chodak & Kwaśnicki, 2002; Vandeva, 2012; Awange, et al., 2018; Andrejkova, et al., 2019). The assessment of the adaptation of a given individual to the environment is made using the fitness function. It has a significant meaning in the computational process, because, properly formulated, it allows the degree of fitting of individuals in the population to be determined and, on its basis, select individuals for the next iteration, an individual can be represented as a string of chromosomes (characters). The string is coded form of a value of a real x (individual), and $f(x)$ is the measure of the fitness of x . Finding the fittest individual means maximizing (or minimizing) the $f(x)$ function (Awange, et al., 2018).

The aim of GA is the iterative search among the generated solutions of the considered problem. This strategy leads to the determination of the fittest generation (Ławrynowicz, 2011; Wojarnik, 2015; Awange, et al., 2018; Matti & Khorsheed Al-Sulaifanie, 2018). It can be noticed that the idea of using the genetic algorithm and the decision process determine one direction - choosing the best alternative from all of the possible alternatives. In the case of large spaces and complex tasks, traditional calculation methods do not guarantee finding the optimal solution. Evolutionary calculations allow to solve inaccurate tasks with complex data, even when the objective function is not determined by an exact mathematical formula or is disturbed and imprecise (Cierpisz & Kowalik, 2000; Kotowski, 2008; Vandeva, 2012; Srinuandee, et al., 2012; Kumar, et al., 2018). Czech (2007) distinguished 4 GA features that distinguish it from traditional optimization techniques: (I) operate on code sequences, (II) operate on the whole population and not on single points, (III) search is done using the sampling method, (IV) uses random selection rules. In the case of the real estate market, deterministic models are often difficult to implement in practice due to the complexity of the processes taking place there (Pereira, 2000). GA uses only the purpose function, not its derivatives or other auxiliary information, and uses probabilistic rather than deterministic selection rules (Goldberg, 1989; Rutkowski, 2009; Lee, 2018). These features make up the "immunity of the evolutionary algorithm" giving it an advantage over analytical methods (direct and indirect), which have a limited range of applications. One of the basic characteristics of GA is the ability to find global maxima of an optimized function with properly selected task parameters. At the same time, traditional direct analytical methods find solutions to the problem only locally, because the search area is the neighborhood of the point under consideration. GAs are a type of non-deterministic method, which is why they do not always guarantee a correct solution but have universal functions that can be used to solve various problems (Lee, 2018). Genetic algorithms are tools that can be used to solve dynamic problems (variables in time) characterized by a large number of variables (discrete or continuous) and the complexity of space, ones that require fast and good results, in which the objective function can be defined as a function of several variables (Figielska, 2006; Lee, 2018). GAs found application in solving, among many others, problems such as: investment strategy design, task scheduling, financial modelling, function optimization, work scheduling, cost minimization (Winiczenko, 2008; Dubinskas & Urbšienė, 2017). Genetic algorithms have been used as tool to build a decision support system in numerous of studies (Juan, et al., 2009; Kou et al. 2012; Zavadskas et al. 2016; Lin et al. 2019). GA as elements of decision support systems also have many possible applications in the real estate field, especially in property valuation. Giudice et al. (2017) used GA in their work with the aim of identifying the effect on real estate rental values derived by geographical location of housing units. According to the authors, GA can be used to more effectively interpret different segments of local real estate markets, or even help in the prediction and interpretation of the phenomena related to the genesis of rewards of position, with particular reference to problems of transformation and investments for urban areas affected by particular projects or plans, and in order to optimize choices of goods and resources are used. For the purpose of real estate price forecasting, many authors proposed solutions based on a hybrid of the colony algorithm, grid algorithm, genetic algorithm, particle swarm optimization and support vector machines. In this case, GA was used to optimize the parameters of the Support Vector Machine simultaneously (Gu, et al., 2011; Wang, et al., 2014). Ahn et al. (2012) proposed a modified version of ridge regression - ridge regression coupled with the genetic algorithm (GA-Ridge) with the aim of forecasting real estate appraisal.

3.2. Hellwig method

The heuristic method developed by Hellwig in 1968 belongs to a group of methods of multidimensional comparative analysis and is historically the first method of linear ordering proposed in the field of economics - taxonomy (Bąk, 2016). Bąk (2016), in his study, states that a multidimensional comparative analysis enables the analysis of complex phenomena, i.e. those on which the state simultaneously affects many attributes and factors. The Hellwig method is one of the tools used to select explanatory variables, where the basis for calculations is the vector of correlation coefficients between the explained variable and explanatory variables and the matrix of correlation coefficients between explanatory variables (Omiotek & Wójcik, 2014). An explanatory variable is called a synthetic variable and is hidden, while explanatory variables are alternatively diagnostic variables that can be directly observed (Omiotek & Wójcik, 2014). The assumption of Hellwig's method is to find a set of explanatory variables with the largest possible integral information capacity. For the optimal combination, variables weakly correlated with each other, but strongly correlated with the explained variable, are selected. Variables belonging to this combination should be used to build the future model. Rosienkiewicz (2012) emphasized that the advantage of this method is the special property of parameter H (information capacity index), which is a normed quantity, contained in the range $\langle 0,1 \rangle$. The closer the H index is to the unity, the more information the variables provide. Adamowicz and Janulewicz (2012) argue that the main advantage of using the Hellwig method is connected with the fact that it synthesizes factors of various nature and assigns them a synthetic aggregate measure. This measure allows a synthetic comparison of the examined units.

3.3. Data description

In order to verify the developed decision-making procedure, information from the local residential real estate market of the city of Wrocław was collected from a transaction and price register. The analysis includes data on 2578 transactions of residential apartments from the secondary market, which took place in 2018. Based on the information contained in the transaction price register, 12 variables (attributes) that each property has been described with were distinguished: x_1, x_2, \dots, x_{12} . The variables were encoded in the form of integers (x_1 - location: 1 - province Krzyki, 2 - province Fabryczna, 3 - province Psie Pole, 4 - province Śródmieście, 5 - province Stare Miasto; x_2 - type of ownership: 1 - ownership + ownership, 2 - ownership + perpetual usufruct, 3 - cooperative ownership right to the premises as follows; x_3 - number of rooms: 1,2,3,4,5,6; x_4 - number of kitchens: 0,1; x_5 - number of bathrooms: 0,1,2; x_6 - number of toilets: 0,1,2; x_7 - number of foyers: 0,1,2,3; x_8 - number of kitchen annexes: 0,1; x_9 - number of dressing-rooms: 0,1,2,4; x_{10} - number of utility rooms: 0,1,2; x_{11} - number of garages: 0,1,2,3; x_{12} - floor: -1,0,1,2,3,4,5,6,7,8,9,10,11,12,13,14,16).

The modified genetic algorithm used to build the structure of the decision support system in an automated way takes unmodified (raw) data from the database maintained in the form of a data table. The fields of property features are presented in a uniform manner for the entire database, to enable the simultaneous processing of quantitative and qualitative data.

3.4. Genetic algorithm

The problem is defined by the environment in which there is a population of individuals (real estate). The program user enters the input data: the price of the property, the number of properties to be generated and the number of processing epochs. The randomness of creating further sets of genes should take into account their level of significance and impact on the transaction price. Depending on the amount specified, the algorithm searches a database of 2578 properties and selects, on the basis of the similarity criterion (Formula 1), individual / individuals whose transaction price is close to the defined price:

$$|P - P_i| = \min \quad (1)$$

where:

P - real estate price defined by the operator,

P_i - transaction price of real estate in the database.

The absolute value of the transaction price difference and the price assumed by the user is the criterion of similarity between the individual-real estate generated by the genetic algorithm and the real estate from the transactional database. The similarity criterion is used to select, from the database,

the model real estate (M), and is used to examine the degree of adaptation to the environment of generated individuals. The absolute value of the transaction price difference and the price assumed by the user should be as close to zero as possible. The first stage of the program is to initiate the initial population (P_{in}), which is completely random and consists in selecting the desired number of individuals represented by code strings. The next step is to assess the adaptation of individuals from the initial population to the environment, which consists in calculating the fitness function value for each chromosome. The level of solving the problem by the individual is determined using the fitness function F (Formula 3), which takes into account the similarity (R) of the tested property to the model real estate (M) (Formula 2).

$$R = \frac{\sum |x_M - x_n| * W_n}{\max(1, x_n)}, \text{ where } n = (1, \dots, 12) \quad (2)$$

$$F = \frac{1}{1+R} ; \quad (3)$$

where: x_M – the attribute value of the model real estate, x_n – the attribute value of the tested property, W_n – the weight of the attribute (Hellwig), M – model real estate, L – number of attributes, $L=12$.

In order to verify the possibility of genetic algorithm application for individual (dedicated) investment consulting, the structure of decision-making system components was developed (Figure 1).

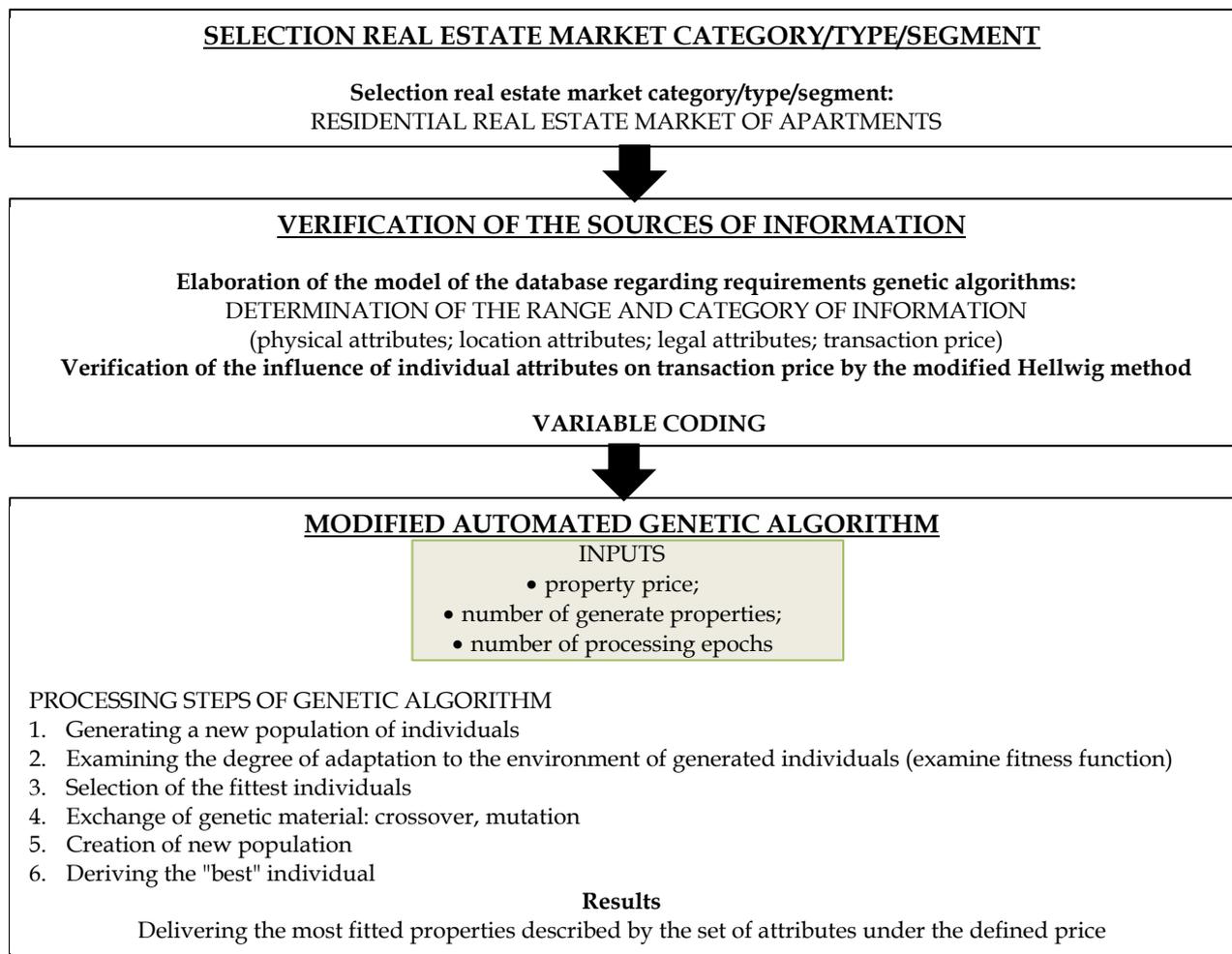


Fig. 1. Components of decision support property investment consulting system. *Source:* own study.

By means of such a defined **fitness function**, each individual is enumerated for its adaptation to the environment. After assessing the adaptation of chromosomes in the population, the retention condition is checked. In this case, the algorithm stops after a specified number of processing epochs given by the operator.

The next step is to use the **selection function**. The selection of chromosomes is based on selection

on the basis of the determined values of fitness functions of those individuals who will participate in the creation of a new generation, i.e. the next generation of individuals. The selection of individuals for the next generation takes place in accordance with the principle of natural selection - individuals with the highest value of adaptation functions have the greatest chance of participating in the creation of a new population, the so-called parental population. The population size in each generation is the same; therefore individuals selected at the selection stage occupy the places of weaker individuals in the population. Chromosomes (real estate) selected in this way are subjected to the action of genetic operators for the purpose of recombining genes.

In the described example, two genetic operators were used: **crossbreeding and mutation**. First, the subjects are subjected to a crossover operator whose probability was defined at the level of $pc = 0.5$ for each gene. A crossover operation is a multi-argument transformation that allows parents to receive progeny that will be part of the next generation. This type of crossing is based on the fact that each gene from the parent's coding sequence has a 50% chance to enter the genome of the offspring. The next step is to use a mutation operator that introduces genetic diversity and prevents the premature convergence of the algorithm. The mutation in a chromosome depends on a random change in the value of the gene (attribute). The mutation probability is defined for each gene and is equal to $pm = 0.05$. Too strong of a mutation resulted in the opposite of the intended effect, i.e. an increase in the number of so-called "bad mutations". Chromosomes obtained as a result of the action of genetic operators are part of the new population. In every generation, a new set of artificial individuals (bit sequences of generated real estates) is created, formed from a combination of fragments of the best-adapted representatives of the previous generation.

Deriving the "best" individual: The program operator obtains decoded information about the individual, i.e. the real estate described by a set of 12 attributes. From the perspective of the program user, the program generates a set of real estate attributes, the price of which corresponds to the price entered at the beginning of the program operation.

4. Results

4.1. Hellwig's method

A crucial step of GA procedure involves the feeding of the algorithm with appropriate information. In order to develop the fitness function of the GA, it is necessary to establish the weights of particular variables. The aforementioned weight can be established either on the basis of random selection, expert knowledge or calculation according to selected methods. The authors intention was to base analysis on the numerically reliable approach, where particular conclusions are drawn from data analysis. In order to determine the significance of individual attributes of real estate, the authors used a modified version of Hellwig's method based on Spearman's correlation. The heuristic method developed by Zdzisław Hellwig includes both, correlation of explanatory variables with the explained variable as well as correlations between pairs of explanatory variables (Renigier-Biłozor & Biłozor, 2016). Attributes $x_1, x_2, x_3, x_4, x_5, x_6, x_7, x_8, x_9, x_{10}, x_{11}, x_{12}$ were adopted for further analysis as explanatory variables with different levels of significance W_i , determined using the authors' modification of the Hellwig method. To determine the significance of the attributes, the first step was to create vector of r_j correlation coefficients between the explanatory variables (x_1, \dots, x_{12}) and the dependent variable (P_i), and the appoint a matrix of correlation coefficients r_{ij} between all explanatory variables. An explanatory variable (decisional) for each property is its unmodified transaction price P_i from the transaction database. The relationship between attributes (x_1, \dots, x_{12}) and the transaction price (P_i) can be presented in the form of a function (Formula 4):

$$f(x) = P_i(x_{1-i}, \dots, x_{12-i}), \text{ where } (i = 1 \dots 2578) \quad (4)$$

Individual information capacity indicators were developed initially for all variables x_i ($i = 1, \dots, 12$) according to Formula 5:

$$h_{kj} = \frac{r_j^2}{\sum_{\substack{i=1 \\ j \neq i}}^m |r_{ij}|} \quad (5)$$

where: r_j - correlation between P and x_j , r_{ij} - correlation between x_i and x_j , k - the number of combinations, $k = 1, 2, \dots, m$.

Individual indicators of information capacity were used to determine the integral information capacity of explanatory variables according to Formula 6:

$$H_k = \sum_{j=1}^m h_{kj}, k = 1, 2, \dots, l \quad (6)$$

The next step involved verification of the impact of individual variables on individual information capacity indicators (H_i) and integral information capacity (H_t) by excluding the Spearman R factors corresponding to the examined explanatory variables from the calculations. Twelve variable combinations were tested for information capacity (in each combination one variable was excluded). The analysis verified the influence of variables on the value of the integral information capacity, which, for the full set of attributes, was $H_t(C_0) = 0.29283$ (Table 1).

Table 1

The value of H_t after excluding variables

Combination of variables	$H_t(C_i)$	Combination of variables	$H_t(C_i)$
$C1 = x_2, x_3, x_4, x_5, x_6, x_7, x_8, x_9, x_{10}, x_{11}, x_{12}$	0.28356	$C7 = x_1, x_2, x_3, x_4, x_5, x_6, x_8, x_9, x_{10}, x_{11}, x_{12}$	0.27222
$C2 = x_1, x_3, x_4, x_5, x_6, x_7, x_8, x_9, x_{10}, x_{11}, x_{12}$	0.24349	$C8 = x_1, x_2, x_3, x_4, x_5, x_6, x_7, x_9, x_{10}, x_{11}, x_{12}$	0.27866
$C3 = x_1, x_2, x_4, x_5, x_6, x_7, x_8, x_9, x_{10}, x_{11}, x_{12}$	0.10032	$C9 = x_1, x_2, x_3, x_4, x_5, x_6, x_7, x_8, x_{10}, x_{11}, x_{12}$	0.29896
$C4 = x_1, x_2, x_3, x_5, x_6, x_7, x_8, x_9, x_{10}, x_{11}, x_{12}$	0.20247	$C10 = x_1, x_2, x_3, x_4, x_5, x_6, x_7, x_8, x_9, x_{11}, x_{12}$	0.28589
$C5 = x_1, x_2, x_3, x_4, x_6, x_7, x_8, x_9, x_{10}, x_{11}, x_{12}$	0.21410	$C11 = x_1, x_2, x_3, x_4, x_5, x_6, x_7, x_8, x_9, x_{10}, x_{12}$	0.25122
$C6 = x_1, x_2, x_3, x_4, x_5, x_7, x_8, x_9, x_{10}, x_{11}, x_{12}$	0.30102	$C12 = x_1, x_2, x_3, x_4, x_5, x_6, x_7, x_8, x_9, x_{10}, x_{11}$	0.23211

Source: own study.

In the next step, the differences R_i between the integral information capacity of all variables $H_t(C_0)$ and the integral information capacity $H_t(C_i)$ for the twelve variable combinations were determined. The analysis showed that the variables x_1, x_2, \dots, x_{12} affect the integral information capacity in different degrees, therefore the determined values of R_i were used as the basis for determining the level of significance of individual attributes (Table 2). The weighting of the explanatory variables is determined by Formula 7:

$$W_i = \frac{[H_t(C_0) - H_t(C_i)] + C}{\sum H_t(C_0) - H_t(C_i)} * 100\% \quad (7)$$

where: $H_t(C_0)$ - integral information capacity for the complete data set, $H_t(C_i)$ - integral information capacity after exclusion of individual variables, C - constant parameter.

Table 2

The value of H_t after excluding variables

Variable	$R_i = H_t(C_0) - H_t(C_i)$	$R_i + C$	W_i [%]
x_1 (location)	0.00927	0.03927	4,32
x_2 (type of ownership)	0.04934	0.07934	8,72
x_3 (number of rooms)	0.19251	0.22251	24,45
x_4 (number of kitchens)	0.09036	0.12036	13,22
x_5 (number of bathrooms)	0.07873	0.10873	11,95
x_6 (number of toilets)	-0.00820	0.02180	2,40
x_7 (number of foyers)	0.02061	0.05061	5,56
x_8 (number of kitchen annexes)	0.01417	0.04417	4,85
x_9 (number of dressing-rooms)	-0.00614	0.02386	2,62
x_{10} (number of utility rooms)	0.00714	0.03714	4,08
x_{11} (number of garages)	0.04160	0.07160	7,87
x_{12} (floor)	0.06071	0.09071	9,97
SUM		0.91011	100

Source: own study.

As results from the conducted research (Table 3), the largest medium of information correlated with the price of the property is the number of rooms. The smallest carrier of information on the price of a property at a given configuration of variables is the number of toilets. The calculated weights W_i of the explanatory variables were used to develop the fitness function of the genetic algorithm used (see Formula 2).

4.2. Modified genetic algorithm

The developed model makes it possible to select, from the property database (the entire population of individuals), the property that best suits the assumptions made (individual best suited to the environment). The residential properties adopted for the analysis were defined as individuals (N), which formed the basis for creating a random initial population for the genetic algorithm. Individuals in the initial population and each subsequent iteration are composed of one chromosome, which consists of 12 genes, which correspond to the attribute values of individual properties (x_1, \dots, x_{12}). In the described problem, individuals have one chromosome, therefore both terms (individual, chromosome) will be used interchangeably (Formula 8).

$$N_{i(i=1\dots n)} = \{(x_1, x_2, x_3, x_4, x_5, x_6, x_7, x_8, x_9, x_{10}, x_{11}, x_{12})\} \quad (8)$$

A set of genes is the genotype of an individual. It is the basis for creating a phenotype that will be assessed for adaptation. Each chromosome encodes a set of information about the property, which is one of the possible solutions to the problem. In the first step, the user's task is to define the price of the property, the number of properties to generate corresponding to the assumed price and the number of processing epochs. The model for the comparison of real estate generated by the genetic algorithm is the model real estate (M) selected from the transaction database based on the similarity criterion (Formula 1). Table 3 presents the results of testing the operation of the genetic algorithm. For each of the two accepted property prices: PLN 100,000 (attempt 1) and PLN 300,000 (attempt 2), 7 trials (randomly adopted- taking into account the influence of the result and software computation time) were carried out. Each attempt consisted in generating one property with a random set of features for a given number of epochs (1000, 5000, 10000, 50000, 100000, 200000, 500000). The number of properties that the algorithm generates does not affect the efficiency of processing and the quality of the results obtained. This parameter only determines the number of repetitions of the entire calculation process starting with the random creation of the initial population and ending with the derivation of the best adapted individual (desired real estate).

Table 3

Generating most fitted properties to required attributes

Attempt 1: the price of the property: 100 000 PLN; the number of properties to generate: 1; number of trials: 7								
Variable	M	Number of epochs						
		1000	5000	10 000	50 000	100 000	200 000	500 000
x1	4	4	5	4	5	5	5	3
x2	2	2	2	3	2	2	2	2
x3	1	1	1	1	1	1	1	1
x4	1	1	1	1	1	0	1	1
x5	1	2	1	1	1	1	2	1
x6	0	2	0	2	0	0	2	0
x7	1	1	1	2	3	1	0	1
x8	0	0	0	1	0	1	0	1
x9	0	0	1	0	0	0	0	0
x10	0	1	1	2	0	0	1	0
x11	0	0	0	2	1	1	0	0
x12	4	4	5	5	8	4	3	12
Attempt 2: the price of the property: 300 000 PLN; the number of properties to generate: 1; number of trials: 7								
Variable	M	Number of epochs						
		1000	5000	10 000	50 000	100 000	200 000	500 000
x1	3	4	2	5	3	2	5	5
x2	2	2	2	2	2	2	2	2
x3	3	3	3	3	3	3	3	3
x4	1	0	1	1	1	1	1	1
x5	1	2	2	1	1	1	1	1
x6	0	1	1	2	2	2	1	2
x7	1	1	0	1	1	2	1	3
x8	0	0	0	1	1	0	0	0
x9	0	0	0	0	1	2	3	2
x10	0	0	1	0	0	1	1	0
x11	0	0	2	2	2	0	3	0
x12	3	4	1	3	3	7	4	8

Source: own study.

In the example, the criterion of stopping the algorithm in the form of the number of processing epochs was used. The genetic algorithm starts processing from the random creation of the initial population, and the whole process repeats itself as many times as the defined number of epochs indicates. This parameter significantly affects the rate of the algorithm - it can be noted that the GA operation has slowed down for the number of epochs above 10000. In study 1, real estate most similar to the reference real estate was generated at epochs of 1000 and 500000, and the least similar - at epochs = 10000. In study 2, the algorithm searched for the most similar property to the reference model in 10000 and 50000 iterations, and the least similar in 5000 iterations. The research shows that the smallest fluctuations in the value of features occur in the number of rooms and the largest in the number of toilets. This is due to the fact that the attributes are given either the largest or least importance of all attributes describing the real estate. Large fluctuations in relation to other variables can be noticed in the case of features whose value ranges were greater than for other attributes (location, level). GA generated such solutions, the sum of partial elements corresponding to the assumptions introduced at the beginning of the algorithm, i.e. properties with such features, which together make up the defined price.

5. Discussion and conclusions

The issue presented in the described example is an attempt to examine the potential of a methodology based on evolutionary methods to analyze the real estate market. The developed modified genetic algorithm, processing real estate transaction data, gives the opportunity to rationalize and specify the needs of potential investors in relation to market reality. The applied decision tool generates real estate with various combinations of 12 attributes, including their impact on the transaction price. Conducting research on the significance of real estate attributes based on source data and using the results of these analyses to build a modified genetic algorithm facilitates the generation of real properties adequate to the actual situation on the real estate market. Additionally, accuracy verification in random methods is a difficult issue. In the case of the genetic algorithm, among others, obtained results depend on the adopted algorithm and used parameters. Determined parameters depend on the experience of the model constructor and the aim of the research. What is more, reliability is yet another issue. The obtained result is the virtual property that can be created based on the markets reality and the investors' demands. The creation of such "individuals" (real estate) is not covered by the condition of their existence in reality. Such a solution gives the opportunity to verify the expectations of potential investors on the market, but also to assess existing real estate offers. By examining the trends in changes in the real estate market, GA can be used as a tool for modelling and forecasting future phenomena in relation to the uncertain future. In contrast to classical computational methods, genetic algorithms can be used to find single factors or their combinations that affect the phenomena occurring on the market. Knowledge of such factors in the next step can be used to improve the effectiveness of other optimization methods.

An important issue is also the calculation time and the number of processing epochs to find the optimal solution. The number of epochs defined by the program operator affects the quality of the generated results and the speed of the algorithm. The biggest disadvantage of this tool is the selection of appropriate parameters which is not trivial and does not always depend on the choice of input data. The influence of the number of iterations on the achieved results will be the subject of further research. The analysis will include the method of determining the weights of real estate attributes and their importance in the process of finding optimal solutions.

Having analyzed the literature and carried out research, it can be concluded that the application of the genetic algorithm can take place in various aspects of real estate market analysis. Genetic algorithms are excellent tools used as a component of broader methods implemented in market analysis and property valuation systems. A genetic algorithm allows risk and uncertainty to be minimized by providing relevant knowledge, in a specified manner processed and given to decision-makers. Decisions are made by analyzing information that is available for a given problem and selecting the best alternative for further action. The excess of seemingly relevant information, i.e. so-called information noise, may disrupt interpretations and the carrying out of analyses. Traditional analytical methods do not reflect the actual image of the market, which is why evolutionary methods, being one of the areas of artificial intelligence, are more and more often used. Artificial intelligence methods, as learning systems, find application in supporting decision-making processes thanks to the ability to predict future events. Awareness of the possibilities and limitations of evolutionary methods

as well as a thorough knowledge of the problem under examination allows for the proper adjustment of the algorithm's parameters, which makes it a universal and effective tool. According to the authors, it is worth undertaking efforts to popularize the use of genetic algorithms as a representation of evolutionary methods to build an AVM on the real estate market, especially due to the universality of their use and the ability to choose the right parameters for the given type of problem.

Funding: This work was supported by the National Science Centre [grant number 2019/33/B/HS4/00072].

6. References

- Adamowicz, M. & Janulewicz, P. (2012). Wykorzystanie metod wielowymiarowych w określeniu pozycji konkurencyjnej gminy na przykładzie województwa lubelskiego [The use of multi dimensional methods in defining the competitive position of the community on the example Lubelskie voivodeship]. *Metody ilościowe w badaniach ekonomicznych*, 13(1), 17 – 28.
- Ahn, J. J., Byun, H. W., Oh, K. J., & Kim, T. Y. (2012). Using ridge regression with genetic algorithm to enhance real estate appraisal forecasting. *Expert Systems with Applications*, 39(9), 8369–8379. <https://doi.org/10.1016/j.eswa.2012.01.183>
- Andrejkova, G., Marčišínová, K. & Kudela, K. (2019). Genetic algorithms in the prediction of geomagnetic storms.
- Awange, J., Palancz, B., Lewis, R., & Volgyesi, L. (2018). Genetic algorithms. *Mathematical Geosciences*. Springer. <https://doi.org/10.1007/978-3-319-67371-4>
- Bąk, A. (2016). Porządkowanie liniowe obiektów metodą hellwiga i topsis – analiza porównawcza [Linear ordering of objects using hellwig and topsis methods a comparative analysis]. *Prace Naukowe Uniwersytetu Ekonomicznego we Wrocławiu*, 426, 22 – 31. <https://doi.org/10.15611/pn.2016.426.02>
- Bertalanffy, L. (1984). *Ogólna teoria systemów* [General system theory]. PWN.
- Brzezicka, J. (2016). Znaczenie heurystyki zakotwiczenia i dostosowania w procesie wartościotwórczym na rynku nieruchomości [Significance of anchoring and adjustment heuristic in the process of value creation on the real estate market]. *Acta Scientiarum Polonorum. Administratio Locorum*, 15(1), 31–44. <https://doi.org/10.31648/aspal.480>
- Brzezicka, J., Łaszek, J., Olszewski, K., & Waszczuk, J. (2019). Analysis of the filtering process and the ripple effect on the primary and secondary housing market in Warsaw, Poland. *Land Use Policy*, 88, 104098. Advance online publication. <https://doi.org/10.1016/j.landusepol.2019.104098>
- Burnside, C., Eichenbaum, M., & Rebelo, S. (2016). Understanding Booms and Busts in Housing Markets. *Journal of Political Economy*, 124(4), 1088–1147. <https://doi.org/10.3386/w16734>
- Chodak, G., & Kwaśnicki, W. (2002). Zastosowanie algorytmów genetycznych w prognozowaniu popytu (Application of Genetic Algorithms in Demand). *Gospodarka Materialowa i Logistyka*, 4, 2–7.
- Cierpisz, S., & Kowalik, S. (2000). Zastosowanie algorytmu genetycznego do optymalizacji układu technologicznego produkcji mieszanki węgla [Application of genetic algorithm to optimize the process of coal blend production]. *Mechanizacja i Automatyzacja Górnictwa*, 12, 5–11.
- Czech, P. (2007). Wykorzystanie algorytmów genetycznych do doboru wejść klasyfikatora uszkodzeń zębów kół przekładni opartego na sieci neuronowej PNN oraz krótkoczasowej transformacji Fouriera [The use of genetic algorithms in the task of choosing inputs for PNN neural network classifier of faults of gears – tooth which used inputs from STFT analysis]. *Problemy eksploatacji*, 3, 51–70.
- Czechowska, K. (2014). Wybrane uwarunkowania podejmowania decyzji inwestycyjnych na rynku nieruchomości – ujęcie behawioralne [Selected Determinants of Investment Decision on the Real Estate Markets – Behavioral Approach]. *Studia i Prace Wydziału Nauk Ekonomicznych i Zarządzania Uniwersytetu Szczecińskiego*, 36(1), 13–25.
- d'Amato, M., Żróbek, S., Renigier-Biłozor, M., Walacik, M., & Mercadante, G. (2019). Valuing the effect of the change of zoning on underdeveloped land using fuzzy real option approach. *Land Use Policy*, 86, 365–374. <https://doi.org/10.1016/j.landusepol.2019.04.042>
- Del Giudice, V., De Paola, P., & Forte, F. (2017). Using Genetic Algorithms for Real Estate Appraisals. *Buildings*, 7(2), 31. <https://doi.org/10.3390/buildings7020031>

- Del Giudice, V., De Paola, P., Forte, F., & Manganeli, B. (2017). Real Estate Appraisals with Bayesian Approach and Markov Chain Hybrid Monte Carlo Method: An Application to a Central Urban Area of Naples. *Sustainability*, 9(11), 2138. <https://doi.org/10.3390/su9112138>
- Dubinskas, P., & Urbšienė, L. (2017). Investment portfolio optimization by applying a genetic algorithm – based approach. *Ekonomika (Nis)*, 96(2), 66–78. <https://doi.org/10.15388/Ekon.2017.2.10998>
- European Mortgage Federation and European AVM Alliance. (2016). EMF/EAA joint paper on the use of automated valuation models in Europe.
- European Valuation Standards. (2017). EVGN11 The Valuer's Use of Statistical Tools. TEGoVA.
- Figielska, E. (2006). Algorytmy ewolucyjne i ich zastosowania. *Zeszyty Naukowe Warszawskiej Wyższej Szkoły Informatyki*, 1, 81–92.
- Goldberg, D. E. (1989). Genetic algorithms in search optimization and machine learning. Addison – Wesley Longman Publishing.
- Gu J., Zhu, M., & Jiang, L. (2011). Housing price forecasting based on genetic algorithm and support vector machine. *Expert Systems with Applications*, 38(4), 3383–3386. <https://doi.org/10.1016/j.eswa.2010.08.123>
- Helbich, M., Brunauer, W., Vaz, E., & Nijkamp, P. (2014). Spatial heterogeneity in hedonic house price models: The case of Austria. *Urban Studies (Edinburgh, Scotland)*, 51(2), 390–411. <https://doi.org/10.1177/0042098013492234>
- IAAO. (2018). Standard on Automated Valuation Models, Approved September 2003, Revised July 2018.
- International Valuation Standards. (2005). International Valuation Standards (7th ed.).
- Janowski, A., Bobkowska, K., & Szulwic, J. (2018). 3D modelling of cylindrical-shaped objects from lidar data-an assessment based on theoretical modelling and experimental data. *Metrology and Measurement Systems*, 25(1), 47–56. <https://doi.org/10.24425/118156>
- Juan, Y., Kim, J. H., Roper, K., & Castro-Lacouture, D. (2009). GA-based decision support system for housing condition assessment and refurbishment strategies. *Automation in Construction*, 18(4), 394–401. <https://doi.org/10.1016/j.autcon.2008.10.006>
- Kaklauskas, A., Zavadskas, E. K., Bardauskienė, D., & Dargis, R. (2015). Sustainable Development of Real Estate: monograph. Vilnius Technika. <https://doi.org/10.3846/2336-M>
- Kauko, T. (2017). Pricing and Sustainability of Urban Real Estate. Routledge.
- Kauko, T., & d'Amato, M. (2008). Mass Appraisal Methods, An interpersonal perspective for property valuers. Wiley-Blackwell., <https://doi.org/10.1002/9781444301021>
- Kontrimas, V., & Verikas, A. (2011). The mass appraisal of the real estate by computational intelligence. *Applied Soft Computing*, 11(1), 443–448. <https://doi.org/10.1016/j.asoc.2009.12.003>
- Kou, G., Lu, Y., Peng, Y., & Shi, Y. (2012). Evaluation of classification algorithms using MCDM and rank correlation. *International Journal of Information Technology & Decision Making*, 11(01), 197–225. Advance online publication. <https://doi.org/10.1142/S0219622012500095>
- Kotowski, S. (2008). Analiza algorytmów genetycznych jako układów dynamicznych. *PhD thesis*, Instytut Podstawowych Problemów Techniki Polskiej Akademii Nauk, Warszawa, Poland.
- Kumar, S., Jain, S., & Sharma, H. (2018). Genetic Algorithms. In N. Nhu Gia (Ed.), Anand, N., Dac – Nhung, L (pp. 27–52). Advances in Swarm Intelligence for Optimizing Problems in Computer Science.
- Lee, I. (2018). Modern Genetic Algorithms. *Korea Research Institute of Standards and Science*, 8 – 11. <https://doi.org/10.3938/PhiT.27.002>
- Levitt, S. D., & Syverson, Ch. (2008). Market Distortions When Agents Are Better Informed: The Value of Information in Real Estate Transactions. *The Review of Economics and Statistics*, 90(4), 599–611. <https://doi.org/10.1162/rest.90.4.599>
- Lin, Ch., Lee, I., & Wu, M. (2019). Merits of using chromosome representations and shadow chromosomes in genetic algorithms for solving scheduling problems. *Robotics and Computer-integrated Manufacturing*, 58, 196–207. <https://doi.org/10.1016/j.rcim.2019.01.005>
- Ławrynowicz, A. (2011). Genetic algorithms for solving scheduling problems in manufacturing systems. *The Journal of Warsaw University of Technology*, 3(2), 7–26. <https://doi.org/10.2478/v10238-012-0039-2>

- Matti, M. S., & Khorsheed Al-Sulaifanie, A. (2018). Wavelet Denoising Based on Genetic Algorithm. In 2018 International Conference on Advanced Science and Engineering, Duhok, Iraq (pp. 75 – 80). <https://doi.org/10.1109/ICOASE.2018.8548814>
- Omiotek, Z., & Wójcik, W. (2014). Zastosowanie metody Hellwiga do redukcji wymiaru przestrzeni cech obrazów USG tarczycy [The use of Hellwig's method for dimension reduction in feature space of thyroid ultrasound images]. *Informatyka, Automatyka. Pomiar w Gospodarce i Ochronie Środowiska*, 4(3), 14–17. <https://doi.org/10.5604/20830157.1121333>
- Pereira, R. (2000). Genetic Algorithm Optimization for Finance and Investments. MPRA Paper 8610. University Library of Munich.
- Rao, R. C. (1994). Statystyka i prawda [Statistics and truth]. Wydawnictwo Naukowe PWN.
- Renigier-Biłozor, M., & Biłozor, A. (2016). Informatio Capacity Database in the Rating Model on the Basis of Polish and Italian Real Estate Markets. *Real Estate Management and Valuation*, 24(3), 40–51. <https://doi.org/10.1515/remav-2016-0020>
- Renigier-Biłozor, M., Biłozor, A., & d'Amato, M. (2018). Residential market ratings using fuzzy logic decision-making procedures. *Economic Research Journal*, 31(1), 1758–1787. <https://doi.org/10.1080/1331677X.2018.1484785>
- Renigier – Biłozor, M., Biłozor, A. & Wiśniewski, R. (. (2017). Rating engineering of real estate markets as the condition of urban areas assessment. *Land Use Policy*, 61, 511–525. <https://doi.org/10.1016/j.landusepol.2016.11.040>
- Rosienkiewicz, M. (2012). Porównanie metod Akaike i Hellwiga w zakresie efektywności konstrukcji modelu regresyjnego [Efficiency comparison of Akaike and Hellwig methods in constructing regression model]. *Wiadomości statystyczne*, 10, 27 – 43.
- Rutkowski, L. (2009). Metody i techniki sztucznej inteligencji. Wydawnictwo Naukowe PWN.
- Stokey, N. L. (2016). Wait-and-See: Investment Options under Policy Uncertainty. *Review of Economic Dynamics*, 21, 246–265. <https://doi.org/10.1016/j.red.2015.06.001>
- Su, D., Xin, L., Lobonț, O., & Yanping, Z. (2016). Economic Policy Uncertainty and Housing Returns in Germany. *Journal of Economics and Business*, 34(1), 43–61.
- TEGoVA The European Group of Valuers Associations. (2016). European Valuation Standards: Automated Valuation Models (AVMs).
- USPAP Uniform Standards of Professional Appraisal Practice 2016 – 2017. (2016).
- Van Groenendaal, W. J. H. (2003). Group decision support for public policy planning. *Information & Management*, 40(5), 371–380. [https://doi.org/10.1016/S0378-7206\(02\)00044-7](https://doi.org/10.1016/S0378-7206(02)00044-7)
- Vandeva, E. (2012). MultiObjective Genetic Modified Algorithm (MOGMA). *Cybernetics and Information Technologies*, 12(2), 23–33. <https://doi.org/10.2478/cait-2012-0010>
- Wang, X., Wen, J., Zhang, Y., & Wang, Y. (2014). Real estate price forecasting based on SVM optimized by PSO. *Optik (Stuttgart)*, 125(3), 1439–1443. <https://doi.org/10.1016/j.ijleo.2013.09.017>
- Winiczenko, R. (2008). Algorytmy genetyczne i ich zastosowania [Genetic algorithms and their applications]. *Postępy Techniki Przetwórstwa Spożywczego*, 1, 107–110.
- Wojarnik, G. (2015). Metody ewolucyjne w analizie zmian kursu akcji spółek giełdowych [Evolutionary methods for the analysis of the changes in price of company stock exchange]. *Zeszyty Naukowe Uniwersytetu Szczecińskiego. Studia Informatica Pomerania.*, 36, 39–50. <https://doi.org/10.18276/si.2015.36-03>
- Wu, Q., Wu, P., Zhou, L. G., Chen, H. Y., & Guan, X. J. (2018). Some new Hamacher aggregation operators under single-valued neutrosophic 2-tuple linguistic environment and their applications to multi-attribute group decision making. *Computers & Industrial Engineering*, 116, 144–162. <https://doi.org/10.1016/j.cie.2017.12.024>
- Zavadskas, E. K., Antucheviciene, J., & Chatterjee, P. (2018). Multiple – Criteria Decision – Making (MCDM) Techniques for Business Processes Information Management. *Informations*, 10(1), 4. Advance online publication. <https://doi.org/10.3390/info10010004>
- Zavadskas, E. K., Antucheviciene, J., Turskis, Z., & Adeli, H. (2016). Hybrid multiple-criteria decision-making methods: A review of applications in engineering. *Scientia Iranica*, 23(1), 1–20. <https://doi.org/10.24200/sci.2016.2093>
- Zavadskas, E., Bausys, R., Kaklauskas, A., Ubarte, I., Kuzminske, A., & Gudiene, N. (2017). Sustainable market valuation of buildings by the single-valued neutrosophic MAMVA method. *Applied Soft Computing*, 57(C), 74–87. <https://doi.org/10.1016/j.asoc.2017.03.040>

- Zavadskas, E., Kaklauskas, A., Turskis, Z., & Tamošaitienė, J. (2008). Selection of the effective dwelling house walls by applying attributes values determined at intervals. *Journal of Civil Engineering and Management*, 14(2), 85–93. <https://doi.org/10.3846/1392-3730.2008.14.3>
- Zavadskas, E., & Turskis, Z. (2011). Daugiatiksliai sprendimų priėmimo metodai ekonomikoje: Apžvalga [Multiple criteria decision making (mcdm) methods in economics: an overview]. *Technological and Economic Development of Economy*, 17(2), 397–427. <https://doi.org/10.3846/20294913.2011.593291>
- Ziółkowski, P., & Niedostatkiewicz, M. (2019). Machine Learning Techniques in Concrete Mix Design. *Materials (Basel)*, 12(8), 1256. <https://doi.org/10.3390/ma12081256> PMID:30999557