Article

Cost Calculation of Construction Projects Including Sustainability Factors Using the Case Based Reasoning (CBR) Method

Agnieszka Leśniak [ID] and Krzysztof Zima * [ID]

Institute of Management in Construction, Faculty of Civil Engineering, Cracow University of Technology, 31-155 Krakow, Poland; alesniak@L3.pk.edu.pl
* Correspondence: kzima@L3.pk.edu.pl; Tel.: +48-628-23-54

Abstract: The idea of sustainable development and the resulting environmentally friendly attitudes are increasingly used in construction projects. Designing in accordance with the principles of sustainable development has an impact on the costs of construction works. The authors of this paper proposed an approach to estimate the costs of sports field construction using the Case Based Reasoning method. In their analysis, they distinguished 16 factors that affect the cost of a construction project and are possible to already be described at an early stage of its preparation. The original elements of the work include: consideration of such environmental factors as the environmental impact of the building, materials used, the impact of the facility on the surroundings affecting the amount of implementation costs and development of own database containing 143 construction projects that are related to sports fields. In order to calculate the similarity of cases, different calculation formulas were applied depending on the type of data (quantitative, qualitative, uncertain, no data). The obtained results confirmed that the CBR method based on historical data and using criteria related to sustainable development may be useful in cost estimation in the initial phase of a construction project. Its application to the calculation of the costs that are related to the implementation of sports fields generates an error of 14%, which is a very good result for initial calculations. In the short run, such factors as the impact of the object and the type of materials that are used from the perspective of their influence on the environment may be decisive as far as the costs determined in the life cycle of the building are concerned, as well as the lowest costs of the building construction ensuring the appropriate quality and respect for the environment.

Keywords: Case Based Reasoning; construction; cost estimation; sports field; sustainability

1. Introduction

The implementation processes of sustainable development in the area of construction are of great importance. Construction processes have an important role in creation of built environment and their impacts have to be measured as construction contributes to air pollution, land use and contamination, usage of resources, water and materials depletion, water pollution, impacts on human health, and climate change [1]. The erection of a construction object is usually associated with the development of the biologically active surface, and subsequently, with putting a burden on the environment. The severity of these kinds of impact on the environment varies depending on the stage of the life cycle of the building. The analysis of the object vs. natural environment relation allows for distinguishing four basic stages of impact that are linked to the following processes: extraction of raw materials, production of materials, construction of the facility, operation of the facility, and its demolition [2]. As [3] proved, the results of developing sustainable architecture are based on changing the function of a

building from a linear approach to a closed circulation plan. The linear approach treats the building as the "place of processing natural resources into waste" (for instance, water is transformed into sewage, energy into heat losses, building materials into waste). The other approach, namely in the closed circulation plan, a building can change from a consumer of energy and other resources into a virtually self-sufficient unit (through energy recovery or the re-use of water or waste). Designed buildings are characterized by diversified energy demand, which depends on many factors; for example, the material and construction solutions applied, the type of object, heating system, and its efficiency. According to the Energy Performance of Buildings Directive [4], all new buildings must be nearly zero-energy buildings by the end of 2020 and all new public buildings by 2018. In general, in terms of sustainable development and sustainable construction, refurbishment of buildings is preferred to new construction, because this helps to save energy and building materials in construction phase, also reduces generation of waste and other emissions [5]. A building object should be designed, constructed, operated, and demolished in accordance with the requirements of sustainable development. The construction materials used to build objects greatly influence the impact of the building on the natural environment, which was noticed in many works, as in [6–10]. The choice of building materials with the appropriate sustainability criteria is not straightforward. What is relatively easy to determine are such objective factors as cost constraints and design considerations, yet other, often subjective, factors have an impact on the selection, which influences the achievement of sustainability goals. In [11], one can find an optimization model for sustainable materials selection, while in [12], a model using a multi-criteria analysis enabling the selection of sustainable materials. Researchers [13] noticed that due to possible harmful effects of construction equipment on the environment, their evaluation with sustainability considerations can be considered as a helpful activity to move toward the sustainability in construction. The paper [14] presents a review of the literature on the sustainable built environment, which was made on the basis of the articles that were published between 1998 and 2015. The authors believe that the welfare of the whole society depends on the sustainability of the built environment [14].

2. Literature Review

The construction objects that are built today should show adequate durability, affect the environment in a harmless way, be economical in the consumption of materials and energy, and take into account the consequences of failure from the point of view of human life and health [15]. Choosing the right construction and material solutions, as well as the technologies that are applied in the first stage of the investment process influences the costs of its implementation and subsequent operation. All of the construction projects are risky, and different techniques and tools are proposed for assessment of risk in Construction project [16,17]. In [18], authors noticed that the fluctuation of material prices is one of the risk factor leading to cost overruns problems. In construction, reliable estimation of costs is important for both the investor who finances the investment and the contractor who has to estimate the costs and achieve a satisfactory level of profit.

To determine the price of construction works, the direct costs that are connected with realization of the works, overhead costs and profit must be consider. The contractors can calculate costs of works using unit price. This method requires a lot of experience from the contractor. The unit price must include all of the mentioned cost's elements. Traditional method of cost estimating is based on calculation of costs elements of separately: the direct cost (labour, materials, and equipment), indirect cost, and profit [19]. This traditional approach is accurate but time consuming, and therefore new methods are still being sought by means of new mathematical tools that can support the effectiveness of the calculation. Studies that are worth considering include ones that are employing artificial neural networks [20–22], linear regression [23], fuzzy sets [19], and support vector machines [24]. Researchers [25] proposed a hybrid model where multivariate regression method and the artificial neural network (ANN) method have been combined to provide a cost estimate model.

One of the methods that is proposed in the literature, which can be used in estimating costs in construction, is CBR (Case Based Reasoning) [26]. Acquiring knowledge as a result of researching the

correctness of data, assimilating or formulating new concepts that are based on examples from the past, (CBR—Case Based Reasoning) in contrast to relying on individual experiences, can accelerate the process of estimating costs in construction. CBR can be defined as systems that solve new problems by adapting the results that were used to solve old issues [27]. In [28], the authors compare the system of inference from cases to the black box. The input data describe the problem; the output data form a solution to the problem, while the memory of past cases and the box contain a reasoning mechanism. In the case-based inferencing, the basic source of knowledge is a database that is containing not rules but a set of cases from the problems encountered and resolved. New problems are solved by searching for the most similar cases and their possible adaptation. The CBR models used for cost estimates may be based on both quantitative and qualitative data [29]. In [30,31], it was noticed that, although some information is not specified, CBR models for long periods of use maintain quality and ability to solve problems and work better than other models. The use of CBR to estimate construction costs is based on searching for similar investments already completed. This provides a simple way to measure construction costs, given that, according to most studies, there are non-linear relationships between cost and factors that affect it [32–35]. An interesting example of the use of CBR in the cost estimation process can be a model using the AHP method to determine the weights of criteria, proposed by [26] or the CBR model using genetic algorithms to estimate the construction costs [36] or unit cost of residential construction projects [37]. Models are also being created that predict both construction time and cost at an early stage of a construction project [38]. Ryu et al. [39] proposed the CONPLA-CBR tool that generates master schedules at the preconstruction stage.

The paper presents the concept of supporting the estimation of construction costs in the initial investment phase based on Case Based Reasoning (the CBR method). The process of information management and the use of historical data differ from the models that have been proposed so far. Typically, global assessment of construction projects or facilities based on a few or several criteria describing the construction project in a general way, such as the usable area, cubic volume of the building, height, complexity, or location. Such criteria often do not take into account differences in the type of materials used and the details of solutions or environmental impact. In cost calculations, fuzzy logic is rarely used when there is uncertain or imprecise information for the problem of cost calculation for the implementation of construction works.

In the proposed algorithm, among the proposed 16 explanatory variables, it is suggested to take into account the features related to the environmental impact of the building, its influence on the surroundings and the parameters of construction materials in relation to the idea of sustainable development. Four different calculation formulas are also proposed for quantitative and qualitative data, for cases without data or ones with uncertain or inaccurate data. It is worth noting that there are no proposals for methods and tools supporting the cost estimation of such sports facilities as football pitches, treadmills, or skate parks.

3. Cost Calculations Method Based on Case Based Reasoning with Sustainability Criterion

The database CSDB (Cost Solution DataBase) that was developed to support index cost estimation with the CBR method includes cases containing four groups of information: information about geometry of building object—GEO, solution of problem (unit cost of object)—SoP, description of the construction—DoC, data for adaptation process—Ad. The database CSDB was defined by the following formula:

$$CSDB = U_{i=1}^{n} Case_i \{GEO_i, SoP_i, DoC_i, Ad\} \tag{1}$$

where:

$Case_i$—*i*-th case from the database,
GEO_i—graphical representation of the construction elements for the *i*-th case,
SoP_i—solution of problem,
DoC_i—description of construction,

Ad—data for adaptation process such date, localization of construction, and
n—number of old cases in the database.

GEO is information about the graphic appearance of the construction object or element in the form of quantitative data. For example, it can be information about the volume of an object, surface, or dimensions, such as length, width, and height. Solution of problem SoP contains information about a unit price of construction object referenced to, for example, surface area or volume of object. Description of construction DoC contains all the necessary information about factors (including sustainability factors) influencing to the cost of the project in the form of qualitative data.

The presented CSBD database was created in order to support the process of estimating costs in the early phase of the investment, including quantitative variables (defining the bill of quantities) and qualitative variables. The novel element of supporting the process of cost estimation is one that is involving the variables affecting sustainable development. The authors of the paper note the increasing influence of factors that are related to sustainable development and their impact on the investment price. The assessment of sustainable development consists of three basic elements: impact on the environment (consumption of natural resources, consumption of energy resources, and emissions of pollutants into the environment), comfort, and quality of life, as well as cost (the amount of funds allocated to achieve the expected value in use and reduction of the impact on the environment). In terms of costs, this means constructing a building at the lowest possible cost while obtaining adequate quality and respect for the environment, as well as the low costs of its operation.

Figure 1 presents the algorithm for the cost estimation method based on the Case Based reasoning method, taking into account the above assumptions.
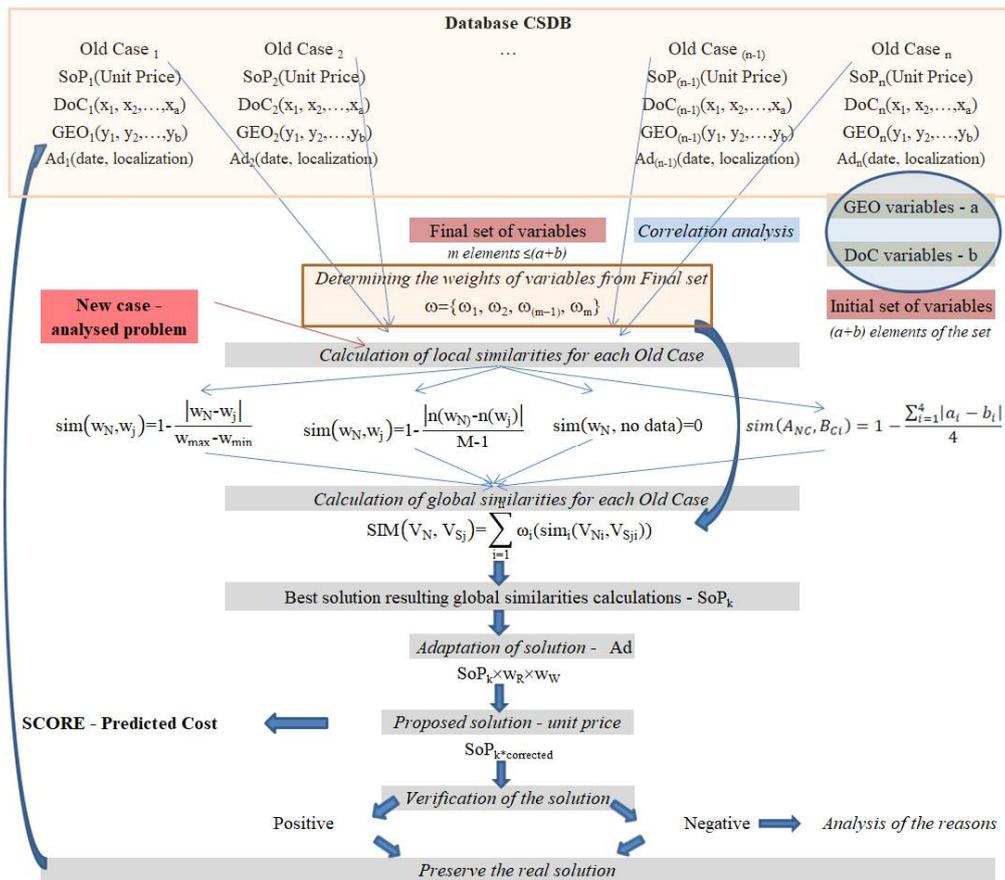

Figure 1. Algorithm used in the method proposed.

The first step in the method that is proposed is to define a preliminary set of variables, divided into GEO geometry variables, and variables describing the DoC structure. Using the case database collected, one can perform a correlation analysis in step 2 to eliminate the variables that have no impact or have a small impact on the unit price of construction works. The result of the correlation analysis is to create the final set of variables (Final set). The third step of the analysis is the assessment of the validity of variables, which can be performed by means of expert assessment; alternatively, it may result from the determined strength of correlation between the individual variables and the price.

The fourth step of the method begins the proper analysis based on the inference from cases. The new case providing a problem to analyse is compared in pairs with all the old cases from the database. The purpose of the comparison is to find the most similar cases, and, by analogy, to determine the unit price of the construction works for the new case.

The similarity of cases can be calculated using different formulas depending on the type of explanatory variables. The calculation formulas applied in the algorithm are presented below:

- For quantitative explanatory variables (as per [40]):

$$\text{sim}\left(w_N, w_j\right) = 1 - \frac{\left|w_N - w_j\right|}{w_{max} - w_{min}} \tag{2}$$

where:

w_N—value of the explanatory variable for the new case,
w_j—value of the explanatory variable for the *j*-th old case, and
w_{max}, w_{min}—minimum and maximum values for all the old cases included in the database.

- For qualitative explanatory variables:

$$\text{sim}\left(w_N, w_j\right) = 1 - \frac{\left|n(w_N) - n(w_j)\right|}{M - 1} \tag{3}$$

where:

$n(w_N)$, $n(w_j)$—place in an ordered array of values $n(w) = 1, 2, \ldots, n$,
M—number of values.

- For uncertain or inaccurate variables:

$$\text{sim}(A_{NC}, B_{Ci}) = 1 - \frac{\sum_{i=1}^{4}\left|a_i - b_i\right|}{4} \tag{4}$$

where:

$\text{sim}(A_{NC}, B_{Ci})$—similarity between fuzzy numbers,
A_{NC}—fuzzy number for the new case,
B_{Ci}—fuzzy number for the old case taken from the database, and
a_i, b_i—characteristic points for fuzzy numbers $A_{NC} = (a_1, a_2, a_3, a_4)$ and $B_{Ci} = (b_1, b_2, b_3, b_4)$.

The formula for uncertain or inaccurate variables is given as an example and concerns a situation in which the shape of the membership function is trapezoidal. It is defined for a situation in which two trapezoidal membership functions assume values $A_{NC} = (a_1, a_2, a_3, a_4)$ and $B_{Ci} = (b_1, b_2, b_3, b_4)$. The characteristic points allow for describing the limit values for the shape of the membership function that accepts the values 0 and 1. This allows for describing the fuzzy number using the four real numbers, which allows for the quick execution of actions using only these characteristic point values. When adopting other forms of membership functions, other calculation formulas should be used.

- In the absence of data for an explanatory variable for a new case, an old case or both:

$$\text{sim}(w_N, w_S) = 0 \tag{5}$$

where:

w_N—an explanatory variable for the new case and
w_S—an explanatory variable for the old case.

The formulas given above are applicable when comparing local similarity, defined as a pair comparison of similarities for subsequent explanatory variables.

The next, fifth calculation step is to calculate the global similarity, which is the weighted sum of the local similarities of all the explanatory variables that are used in analysis and collected in the final set of variables. In order to calculate global similarity, the following formula was used:

$$\text{SIM}(V_N, V_{Sj}) = \sum_{i=1}^{n} \omega_i \left(\text{sim}_i (V_{Ni}, V_{Sji}) \right) \tag{6}$$

where:

ω_i—weight of the *i*-th explanatory variable,
$\text{SIM}(V_N, V_{Sj})$—global similarity between the old V_j and the new case V_N, and
$\text{sim}_i(V_{Ni}, V_{Sji})$—local similarity for the *i*-th explanatory variable between the old V_j and the new case V_N.

It is necessary to determine a few basic assumptions in the course of the analysis that uses the CBR method:

1. old cases with the highest global similarity $\text{SIM}(V_N, V_j)$ are selected. The minimum number of cases entering the selected set of cases is 3. This assumption is to limit the possibility of choosing an accidental solution—if three solutions are chosen, then the possible extreme solution is rejected to limit the possibility of overestimating or underestimating the price of works;
2. the minimum value of global similarity for cases included in the set of solutions must be greater than 70%;
3. the resulting value of similarities is given as a percentage and is a natural number; and,
4. cases from the set of selected cases are rejected as extreme when the difference between the selected cases is greater than 50%.

The next step is to adapt the calculated unit price for the new case. The value correction concerns a situation in which there is a difference between the duration of the construction works and the location between the new case and the selected old cases. To do the adaptation, two coefficients were applied: regional coefficient determined in the case of a difference in the location of a construction project, (introduces price correction due to the difference in the local prices of services and materials) and inflation rate (introduces a price correction due to changes in prices of services and materials over time). Corrections are made individually for each of the unit prices that were selected in the course of the analysis, resulting from the selection of the old case as the most similar and meeting the conditions that are given in the assumptions.

The unit price of the new case, resulting from several finally selected old cases after adaptation, is calculated as the weighted average of the weights of individual old cases (where the basis for the calculation of the weight is the degree of the similarity to the new case) and the unit prices of old cases. The proposed unit price is accepted as the final price of the new case and in the next step it is multiplied by the amount of works from the bill of quantities. After the settlement of the construction works, the price result obtained is verified by the actual costs. The verified result is saved in the database for its subsequent use.

4. An Example of Supporting Cost Calculation with the Use of the CBR Method, Taking into Account the Factors of Sustainable Construction

The calculation example concerns the estimation of the costs of the implementation of sports fields in the early phase of the investment, and is thus based on the investment concept. The cost analysis is performed in accordance with the algorithm provided above, which is based on the CBR method. The purpose of the exemplary cost analysis is to determine the price for the implementation of a new sports facility, which is a sports field that is based on historical data from the investments already completed.

In the first step, a preliminary set of criteria was compiled. The set of criteria was created on the basis of literature studies and the analysis of advertisements that were placed in Poland on the public procurement website. The analysis included the selection of information describing the subject of the contract.

On the basis of literature studies [22,41–43] and the analysis of announcements in public procurement, 16 variables explaining sports fields were distinguished:

1. Quantitative variables (type of information—GEO)
   intended use of the field (five types of fields)
   surface area of the field (variables range: 275–8714 m^2)
   surface area of the access paths and routes (variables range: 0–1753 m^2)
   green surface area (variables range: 0–6017 m^2)
   surface area of the ball containment netting (variables range: 0–2212 m^2)
   fence length (variables range: 0–602.5 m)
2. Qualitative variables (type of information—DoC)
   type of the material for sports surface (six types of materials)
   type of the material for access routes (five types of materials)
   type of the fence (five types of fences)
   type of sports equipment—handball (yes or no)
   type of sports equipment—volleyball (yes or no)
   type of sports equipment—basketball (yes or no)
   type of sports equipment—football (yes or no)
   type of sports equipment—tennis (yes or no)
   impact of the construction on the environment (rating 1–5)
   impact on the surroundings (rating 1–5)

The presented example includes the analysis of the costs of sports fields, both single- and multifunctional. The information contained in the database comes from the investments executed in 2014–2016 and includes 143 building projects concerning the construction of sports fields. The data comes from advertisements and building cost estimates prepared for tender proceedings for the implementation of sports fields in Poland. The data from cost estimates and the public procurement description were placed in the database that was created. The data includes all of the variables highlighted in the course of the analysis of the initial set of criteria: GEO and DoC, and the information about the SoP solution and data allowing for the later adaptation of the Ad solution (tender date and location of a construction project). The method of recording information in subsequent records in the database is described below.

Case$_n$ {GEO = (intended use of the field, surface area of the field, surface area of the access paths and routes, green surface area, surface area of the ball containment netting, fence length); DoC = (type of the material for sports surface, type of the material for access paths, type of the fence, type of sports equipment—handball, type of sports equipment—volleyball, type of sports equipment—basketball, type of sports equipment—football, type of sports equipment—tennis, impact of the construction on the environment, impact on the surroundings); SoP = (unit price of the field surface area); and, Ad = (location, date of the bid)}

The GEO information contains basic parameters concerning the size of individual construction works. On the other hand, the DoC information contain a description of the scope of works and the types of solutions used in the implementation of individual elements of the sports field and its surroundings, and the characteristics of the undertaking related to the idea of sustainable development.

The features that are related to the impact of a building object on the environment (sustainability factors) included in the initial set of variables involve in point 2. Qualitative variables (type of information—DoC) such variables:

- impact of the construction on the environment (for instance, energy demand, use of renewable energy, efficiency of energy systems);
- materials (for instance, materials with low environmental impact, materials with low risk of health hazard, recyclable materials): type of the material for sports surface and type of the material for access routes; and,
- impact on the surroundings (such as air pollution, noise, vibration, wind effects and shading of the area, the effect of the thermal island)

The determination of the correlation strength was made on the basis of the Guilford scale:

$r = 0$ no correlation,
$0 < r < 0.1$ barely perceptible correlation,
$0.1 < r < 0.3$ poor correlation,
$0.3 < r < 0.5$ average correlation,
$0.5 < r < 0.7$ high correlation,
$0.7 < r < 0.9$ very high correlation,
$0.9 < r < 1$ almost full correlation, and
$r = 1$ full correlation.

Table 1 depicts the results of the Pearson and Spearman correlation analysis. For quantitative variables, the analysis involved Pearson's correlation; while for quality variables, it used Spearman's correlation. In order to create the final set of explanatory variables, variables that showed no correlation or very low correlation were rejected in the course of the analysis. The final set of criteria with designated weights is shown in Table 2. The weights were determined based on the correlation strength of individual variables with the unit price of a construction work.

Table 1. The correlation coefficients of explanatory variables from initial set.

Variables	Correlation Coefficient:
Surface area of the fields	−0.192
Surface area of the access paths	0.232
Green surface area	0.079
Fence length	0.249
Surface area of the ball containment netting	0.056
Intended use	−0.472
Material for sport surface	−0.299
Material for access paths	−0.283
Type of sports equipment—Handball	0.244
Type of sports equipment—Basketball	0.477
Type of sports equipment—Volleyball	0.363
Type of sports equipment—Football	−0.289
Type of sports equipment—Tennis	−0.071
Fence type	−0.045
Impact of the construction on the environment	0.095
Impact on the surroundings	−0.640

Table 2. The final set of explanatory variables with their weights.

Variables	Weights ω_i
Surface area of the fields	6.2%
Surface area of the access paths	7.5%
Fence length	8.0%
Intended use	15.2%
Material for sport surface	9.7%
Material for access paths	9.1%
Type of sports equipment—Handball	7.9%
Type of sports equipment—Basketball	15.4%
Type of sports equipment—Volleyball	11.7%
Type of sports equipment—Football	9.3%

The final set of variables consists of two explanatory variables that are related to the idea of sustainable development and included in the Materials (materials for sports surfaces and materials for access). The total validity of these variables is 18.7%, thus they are quite a significant element affecting the unit price of a construction project, although not being decisive. Due to the low correlation with the unit price, the other variables that are related to sustainable construction were rejected, that is, the impact of the building on the environment and the impact on the surroundings, which is two of the four factors that were originally considered.

After creating the final set of variables and determining their validity, local similarities were calculated for the subsequent explanatory variables between the New Case and all the Old Cases from the database. Local Formulas (2)–(5) were used to calculate local similarities. An example of calculating a local similarity for a quantitative variable of an access surface, between the New Case and the Old Case 7 using the Formula (2) has been shown below:

$$\mathrm{sim}(w_N, w_j) = 1 - \frac{|458 - 295|}{5569 - 0} = 1 - 0.029 = 0.971 \cong 97\% \tag{7}$$

where:

w_N equals 458 and determines the size of the access surface for the New Case—458 m^2,
w_j equals 295, which means the access surface for Old Case 7 equal 295 m^2,
w_{max} equals 5569, which means the maximum access area for the whole set of cases is equal 5569 m^2, and
w_{min} is equal 0, which means the minimum access space for the whole set of cases is equal 0 m^2 (some orders for the implementation of sports fields did not include the implementation of access routes).

On the other hand, the local similarity for the quality variable: material for sports surfaces, between the New Case and the Old Case 7 using the formula (3) is shown below:

$$\mathrm{sim}(w_N, w_j) = 1 - \frac{|5 - 5|}{6 - 1} = 1 - 0 = 1 = 100\% \tag{8}$$

where:

$n(w_N)$ is equal 5 and determines the value given to the material which is the polyurethane from which the New Case surface is planned to be made,
$n(w_j)$ is equal 5, which means the value given to the material which is the polyurethane from which the Old Case 7 surface was made, and
M = 6, which results from the specification of six types of materials used for sports surfaces found in the database (natural grass; surface from natural dried wood chips—technologically softened along fibers with 5–50 mm fraction; artificial grass; brick flour; polyurethane surface; asphalt surface).

A list of calculated local similarities and global similarity computed in accordance with the formula (6) taking into account the weights given in Table 2 between the New Case and the selected Old Cases is shown in Table 3.

Table 3. The values of the local similarities and the global similarities for the chosen cases.

Variables	Local Similarities		
	Case 7	Case 8	Case 9
Surface area of the fields	94%	88%	94%
Surface area of the access paths	97%	87%	92%
Fence length	86%	79%	79%
Intended use	86%	0%	100%
Material for sport surface	100%	0%	100%
Material for access paths	100%	100%	0%
Type of sports equipment—Handball	100%	100%	100%
Type of sports equipment—Basketball	100%	0%	100%
Type of sports equipment—Volleyball	100%	0%	100%
Type of sports equipment—Football	100%	100%	100%
Global similarities	**98%**	**45%**	**88%**

The highest similarity of 98% occurred for three cases: Case 7—SIM($V_{\text{Test case 1}}$, $V_{\text{Case 7}}$); Case 23—SIM($V_{\text{Test case 1}}$, $V_{\text{Case 23}}$); and, Case 83—SIM($V_{\text{Test case 1}}$, $V_{\text{Case 83}}$).

For the selected cases in the analysis of three cases with the highest similarity rate, the adaptation of unit prices of these old cases was made due to the time difference in the calculation and the difference in the prices of services and materials in different regions of Poland. The inflation factor considering the time difference was calculated on the basis of the price index "The Sekocenbud forecasting and indexation bulletin" [44]. The regional coefficient, in turn, was calculated based on the newsletter "The Sekocenbud regional price bulletin" [45]. The adjusted unit prices are shown in Table 4.

Table 4. Adjusted unit prices of selected cases.

Case	Case 7	Case 23	Case 83
Unit price	80.25 €	84.23 €	79.34 €
Regional factor	1.028	0.960	1.028
Indexation factor	100.1%	101.3%	101.3%
Adjusted price	82.58 €	81.91 €	82.62 €

For example, the price after the adaptation of Case 7 has been calculated as 80.5 € × 1.028 × 1.001, where the regional factor is—1.028 and the indexation factor is equal to (1 + 0.01%) = 100.1%. The final price for test case 1 is an arithmetic mean of the unit prices of case 7, 23, 83, and equals:

$$\frac{\left(C_{jA}^{\text{Case 7}} + C_{jA}^{\text{Case 23}} C_{jA}^{\text{Case 83}}\right)}{3} = \frac{(82.58 + 81.91 + 82.62)}{3} = 82.37 \text{ €/m}^2 \tag{9}$$

After calculating the unit price after adaptation for the New Case equal 82.37 €/m^2 it is necessary to specify the value of the works. For this purpose, the unit price was multiplied by the number of the works representing the construction works consisting in the performance of the sports field, namely the area of the pitch.

The value of the works for the New Case is thus finally: 82.37 [€/m^2] × 1200 [m^2] = 98 844.00 €.

5. Discussion

An analysis of the error that was generated by the forecasts for the costs of sports field construction using the presented method was made. Testing was performed on selected 15 cases for which the costs

of building a construction project were known. Mean Absolute Estimate Error (MAEE) for 10 test cases was 14%. The error range for individual cases is 2–34%, except that only for three test cases it exceeded 20%. In five cases, the error was smaller than 20%. Figure 2 shows the distribution of the MAEE error in the following ranges: up to 10%, 10–20%, and above 20%.
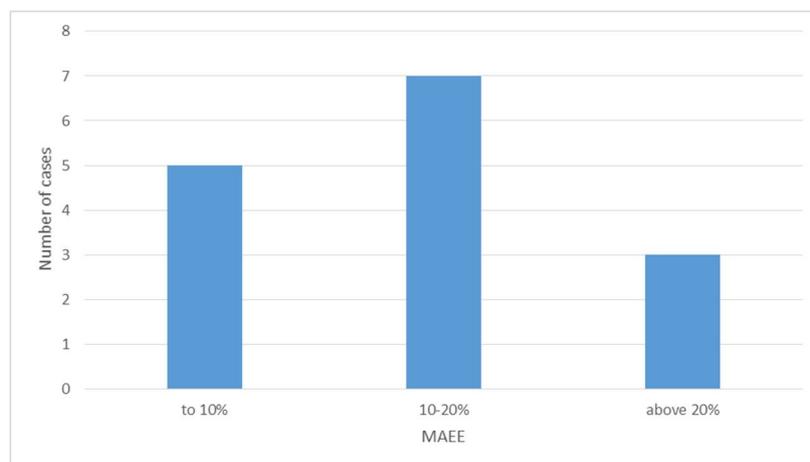

Figure 2. The distribution of the Mean Absolute Estimate Error (MAEE) error with the cost calculation broken down into percentage ranges.

In accordance with the PMI 2008 [46] guidelines, the error of 14% meets the requirements of an acceptable cost calculation error made in the early stages of construction projects in the range of −30% to +50%.

Also, the requirements of the American Association of Cost Engineers (AACE)—error size of −10% to +15%—and the requirements of the Construction Industry Institute—error size of −30% to +50% (for: [47]) are met.

It is worth noting that, in similar studies, only the use of artificial neural networks in supporting costs in such an early investment phase gives similar good results [48,49]. However, even in these cases, the maximum calculation error exceeds 50%, and in the presented method it was 34%.

It seems that both the use of the CBR method in supported cost calculation and the use of variables based on the idea of sustainable development can bring good results. According to the authors, along with the development of the idea of sustainable development of such variables affecting price, there will be increasingly more research on the impact of these factors on costs.

For the CBR method, further research must address the impact of changing calculation formulas that estimate local and global similarities on the MAEE error. The search for formulas reflecting the specificity of data used during calculations may bring even better results. In-depth research into the adaptation process of selected solutions should also be performed.

6. Conclusions

Supporting cost calculations at an early stage of a construction project is a vital problem. Despite the fact that cost-support models that are based on numerous mathematical methods have already been presented in the world literature many times, only a few generate a calculation error below 20%. It seems that only these models that are based on historical data, artificial neural networks and case based reasoning methods reach an acceptable level of error. The vast majority of methods do not take into account the variables that are based on the idea of sustainable development.

The authors believe that the proposed method using Case Based Reasoning has the potential and can be useful for practice. It generated an Mean Absolute Estimate Error of 14%. It should be noticed that only for 3 test cases the MAE Error exceeded 20%. To present performing of the model an example

of calculation the unit price for the New Case was successfully conducted. The value of the works for the New Case finally equalled 98 844.00 €.

The original element of the of cost estimation process using the proposed model is involving the factors affecting sustainable development. The authors note the increasing influence of factors that are related to sustainable development on the investment price. Environmental impact, comfort, and quality of life, in the near future can be decisive in terms of costs determined in the lifecycle of a building, taking into account the construction of a building object as the cheapest cost when obtaining adequate quality and respect for the environment, while obtaining low operating costs.

Author Contributions: The individual contribution and responsibilities of the authors were as follows: K.Z. designed the research main idea and collected the data. Authors together analyzed the data and the obtained results. A.L. provided extensive advice throughout the study results and methodology. K.Z. and A.L. wrote the paper. All the authors have read and approved the final manuscript.

Conflicts of Interest: The authors declare no conflict of interest.

References and Notes

1. Zavadskas, E.K.; Vilutienė, T.; Tamošaitienė, J. Harmonization of cyclical construction processes: A systematic review. *Procedia Eng.* **2017**, *208*, 190–202. [CrossRef]
2. Golański, M. Wybór materiałów budowlanych w kontekście efektywności energetycznej i wpływu środowiskowego. *Budownictwo i Inżynieria Środowiska* **2012**, *3*, 39–53.
3. Bonenberg, W.; Kapliński, O. The Arcithect and the Paradigms of Sustainable Development: A Review of Dilemmas. *Sustainability* **2018**, *10*, 100. [CrossRef]
4. European Union (EU). *Directive 2010/31/EU of 19 May 2010 on the Energy Performance of Buildings*; European Union: Brussels, Belgium, 2010.
5. Zavadskas, E.K.; Antucheviciene, J.; Kalibatas, D.; Kalibatiene, D. Achieving Nearly Zero-Energy Buildings by applying multi-attribute assessment. *Energy Build.* **2017**, *143*, 162–172. [CrossRef]
6. Calkins, M. *Materials for Sustainable Sites: A Complete Guide to the Evaluation, Selection, and Use of Sustainable Construction Materials*; John Wiley & Sons: New York, NY, USA, 2008.
7. Thormark, C. The effect of material choice on the total energy need and recycling potential of a building. *Build. Environ.* **2006**, *41*, 1019–1026. [CrossRef]
8. Leśniak, A.; Zima, K. Comparison of traditional and ecological wall systems using the AHP method. In *International Multidisciplinary Scientific GeoConference Surveying Geology and Mining Ecology Management*; SGEM: Albena, Bulgaria, 2015; Volume 3, pp. 157–164.
9. Drozd, W.; Leśniak, A.; Zaworski, S. Construction Time of Three Wall Types Made of Locally Sourced Materials: A Comparative Study. *Adv. Mater. Sci. Eng.* **2018**, 1–8. [CrossRef]
10. Švajlenka, J.; Kozlovská, M. Houses Based on Wood as an Ecological and Sustainable Housing Alternative—Case Study. *Sustainability* **2018**, *10*, 1502. [CrossRef]
11. Florez, L.; Castro-Lacouture, D. Optimization model for sustainable materials selection using objective and subjective factors. *Mater. Des.* **2013**, *46*, 310–321. [CrossRef]
12. Akadiri, P.O.; Olomolaiye, P.O.; Chinyio, E.A. Multi-criteria evaluation model for the selection of sustainable materials for building projects. *Autom. Constr.* **2013**, *30*, 113–125. [CrossRef]
13. Ghorabaee, M.K.; Amiri, M.; Zavadskas, E.K.; Antucheviciene, J. A new hybrid fuzzy MCDM approach for evaluation of construction equipment with sustainability considerations. *Arch. Civ. Mech. Eng.* **2018**, *18*, 32–49. [CrossRef]
14. Ubarte, I.; Kaplinski, O. Review of the sustainable built environment in 1998–2015. *Eng. Struct. Technol.* **2016**, *8*, 41–51. [CrossRef]
15. Runkiewicz, L. Realizacja obiektów budowlanych zgodnie z zasadami zrównoważonego rozwoju. *Przegląd Budowlany* **2010**, *81*, 17–23.
16. Burtonshaw-Gunn, S.A. *Risk and Financial Management in Construction*; Routledge: Abington-on-Thames, UK, 2017.
17. Chatterjee, K.; Zavadskas, E.K.; Tamošaitienė, J.; Adhikary, K.; Kar, S. A hybrid MCDM technique for risk management in construction projects. *Symmetry* **2018**, *10*, 46. [CrossRef]

18. Odimabo, O.O.; Oduoza, C.; Suresh, S. Methodology for Project Risk Assessment of Building Construction Projects Using Bayesian Belief Networks. *Int. J. Constr. Eng. Manag.* **2017**, *6*, 221–234. [CrossRef]

19. Plebankiewicz, E.; Leśniak, A. Overhead costs and profit calculation by Polish contractors. *Technol. Econ. Dev. Econ.* **2013**, *19*, 141–161. [CrossRef]

20. Juszczyk, M. Application of PCA-based data compression in the ANN-supported conceptual cost estimation of residential buildings. *AIP Conf. Proc.* **2016**, *1738*, 200007. [CrossRef]

21. Leśniak, A.; Juszczyk, M. Prediction of site overhead costs with the use of artificial neural network based model. *Arch. Civ. Mech. Eng.* **2018**, *18*, 973–982. [CrossRef]

22. Juszczyk, M.; Leśniak, A.; Zima, K. ANN Based Approach for Estimation of Construction Costs of Sports Fields. *Complexity* **2018**, 7952434. [CrossRef]

23. Sonmez, R. Parametric range estimating of building costs using regression models and bootstrap. *J. Constr. Eng. Manag.* **2008**, *134*, 1011–1016. [CrossRef]

24. Kim, G.H.; Shin, J.M.; Kim, S.; Shin, Y. Comparison of school building construction costs estimation methods using regression analysis, neural network, and support vector machine. *J. Build. Constr. Plan. Res.* **2013**, *1*, 1–7. [CrossRef]

25. Yazdani-Chamzini, A.; Zavadskas, E.K.; Antucheviciene, J.; Bausys, R. A Model for Shovel Capital Cost Estimation, Using a Hybrid Model of Multivariate Regression and Neural Networks. *Symmetry* **2017**, *9*, 298. [CrossRef]

26. An, S.-H.; Kim, G.; Kang, K. A Case-based reasoning cost estimating model using experience by analytic hierarchy process. *Build. Environ.* **2007**, *42*, 2573–2579. [CrossRef]

27. Riesbeck, C.K.; Schank, R.C. *Inside Case-Based Reasoning*; Lawrence Erlbaum Associates: Hillsdale, NJ, USA, 1989.

28. Pal Sankar, K.; Shiu Simon, C.K. *Foundations of Soft Case-Based Reasoning*; John Wiley & Sons, Inc.: Hoboken, NJ, USA, 2004.

29. Mendes, E.; Mosley, N.; Counsell, S. The application of casebased reasoning to early web project cost estimation. In *Proceedings—IEEE Computer Society's International Computer Software and Applications Conference*; Institute of Electrical and Electronics Engineers Computer Society: Oxford, UK, 2002.

30. Kim, G.-H.; An, S.-H.; Kang, K.-I. Comparison of construction cost estimating models based on regression analysis, neural networks, and case-based reasoning. *Build. Environ.* **2004**, *39*, 1235–1242. [CrossRef]

31. Duverlie, P.; Castelain, J.M. Cost estimation during design step: Parametric method versus case based reasoning method. *Int. J. Adv. Manuf. Technol.* **1999**, *15*, 895–906. [CrossRef]

32. Chou, J.-S.; Peng, M.; Persad, K.; O'Connor, J. Quantity-based approach to preliminary cost estimates for highway projects. *Transp. Res. Rec.* **2006**, *1946*, 22–30. [CrossRef]

33. Emsley, M.W.; Lowe, D.J.; Duff, A.R.; Harding, A.; Hickson, A. Data modelling and the application of a neural network approach to the prediction of total construction costs. *Constr. Manag. Econ.* **2002**, *20*, 465. [CrossRef]

34. Lowe, D.J.; Emsley, M.W.; Harding, A. Predicting construction cost using multiple regression techniques. *J. Constr. Eng. Manag.* **2006**, *132*, 750–758. [CrossRef]

35. Marir, F.; Wang, F.; Ouazzane, K. A case-based expert system for estimating the cost of refurbishing construction buildings. In Proceedings of the 4th International Conference on Enterprise Information Systems, Ciudad Real, Spain, 3–6 April 2002; Volume 1, pp. 391–398.

36. Ji, S.-H.; Park, M.; Lee, H.-S. Cost estimation model for building projects using case-based reasoning. *Can. J. Civ. Eng.* **2011**, *38*, 570–581. [CrossRef]

37. Dogan, S.Z.; Arditi, D.; Günaydın, H.M. Determining attribute weights in a CBR model for early cost prediction of structural system. *J. Constr. Eng. Manag.* **2006**, *132*, 1092–1098. [CrossRef]

38. Koo, C.; Hong, T.; Hyun, C.; Koo, K.A. CBR-based hybrid model for predicting a construction duration and cost based on project characteristics in multi-family housing projects. *Can. J. Civ. Eng.* **2010**, *37*, 739–752. [CrossRef]

39. Ryu, H.G.; Lee, H.S.; Park, M. Construction planning method using case based reasoning (CONPLA-CBR). *J. Comput. Civ. Eng.* **2007**, *21*, 410–422. [CrossRef]

40. Traczyk, W. *Inżynieria Wiedzy/Knowledge Engineering*; Akademicka Oficyna Wydawnicza EXIT: Warszawa, Poland, 2010.

41. Zima, K. Cost Estimating of Football Pitches Construction in the Concept Phase Using Case Based Reasoning. In Proceedings of the International Conference on Economics and Management Engineering (ICEME 2017), Wuhan, China, 24–26 March 2017; DEStech Transactions on Economics, Business and Management; 2017; pp. 24–26.

42. Polish Green Building Council Homepage. Available online: https://plgbc.org.pl/od-czego-zaczac/ (accessed on 3 March 2018).

43. Public Procurement Control Department Homepage. Available online: https://www.uzp.gov.pl/ (accessed on 6 March 2018).

44. Sekocenbud. Forecasting and Indexation Bulletin—ZWW, 2014–2017; PROMOCJA Sp. z o.o.: Warszawa, Poland.

45. Sekocenbud. Regional Price Bulletin—BCR, 2014–2017; PROMOCJA Sp. z o.o.: Warszawa, Poland.

46. PMI. *A Guide to the Project Management Body of Knowledge*; Project Management Institute: Newtown Square, PA, USA, 2008.

47. Kim, S.; Shim, J.H. Combining case-based reasoning with genetic algorithm optimization for preliminary cost estimation in construction industry. *Can. J. Civ. Eng.* **2014**, *41*, 65–73. [CrossRef]

48. Gunaydin, H.; Dogan, S. A neural network approach for early cost estimation of structural systems of buildings. *Int. J. Proj. Manag.* **2004**, *22*, 595–602. [CrossRef]

49. Juszczyk, M. Application of committees of neural networks for conceptual cost estimation of residential buildings. *AIP Conf. Proc.* **2015**, *1648*, 600008. [CrossRef]