Sustainable Cities and Society 53 (2020) 101979

Contents lists available at ScienceDirect

Sustainable Cities and Society

journal homepage: www.elsevier.com/locate/scs

An integrated BIM-LEED application to automate sustainable design assessment framework at the conceptual stage of building projects

Farzad Jalaei[a,*], Farnaz Jalaei[b], Sepehr Mohammadi[c]

[a] *Construction Research Center, National Research Council Canada, 1200 Montreal Road, Ottawa, ON, K1A 0R6, Canada*
[b] *Department of Civil Engineering, University of Ottawa, Ottawa, ON, Canada*
[c] *Department of Civil Engineering, Iran University of Science and Technology (IUST), Tehran, Iran*

ARTICLE INFO

Keywords:
BIM
Sustainable cities
Green building
Certification system
LEED
Sustainable design

ABSTRACT

Construction industry has become more interested in designing and constructing environmentally friendly buildings that can provide both high performance and monetary savings. Generally, the sustainability analysis is mostly conducted at the end of the design stage, once their components and elements have already been selected. However, achieving an integrated sustainable design solution prior to construction means that the design team must manage reciprocal task interdependencies when making decisions related to the selection of the most suitable design alternative that will lead to sustainable and efficient buildings. This study describes a methodology to automate the process sustainability assessment for proposed buildings by integrating Building Information Model (BIM) and LEED certification system while providing a framework to calculate the credits that building could potentially earn at the conceptual stage. This research aims to propose an integrated methodology that links BIM with green building certification systems at the early design stage of a project's life. A plug-in is developed to calculate and predict the potential accumulated LEED credits with access to the Application Program Interface (API) of the BIM tool, energy analysis and lighting simulation tool, Google Map and their associated library. The plug-in uses K Nearest Neighbour (KNN) data mining method to estimate the missing credits, which could not be calculated directly from design specifications, to propose the whole scale innovative green building evaluation interface for building projects.

1. Introduction

A sustainable city is considered as the urban systems and the infrastructure which are planned and designed for long-term requirements. It must be ensured that the city will survive over a long period of time maintaining its integrity, normal functioning, self-reliance and also ensure quality of life for its residents while ensuring its robustness and adaptive capacity simultaneously (Ali, 2008). Nowadays, the policies of the organizations are mostly established based on sustainable development characteristics in cities, running on the principles of equity between preservation of the environment and improvement of the urban areas as well as different functions and services, to accomplish a developed urban sustainability (Ahvenniemi, Huovila, Pinto-Seppä, & Airaksinen, 2017).

Buildings use a considerable amount of energy with an effect on the built environment during construction (Chong, Lee, & Wang, 2017) and as a result, sustainable buildings, which are environmentally friendly and have energy efficient structures, have been proposed to reduce the environmental impacts of buildings, save energy and water and make contributions to the occupant's health and comfort through practices such as temperature and humidity control, indoor air quality, natural lighting, and waste management (Clevenger, 2008). The design processes of sustainable buildings are arguably more complex than conventional design approaches due to the multidisciplinary design team works that are required to address the requirements of environmentally sustainable systems. Achieving an integrated design solution prior to construction means that the design team must manage reciprocal task interdependencies and address a complex of information sharing requirements surrounding data coordination and exchange across multiple disciplines (Gandhi & Jupp, 2014).

Therefore, because of the growing demand for the sustainable development and green buildings in the past few years, green building rating systems are established to estimate the life cycle performance of the buildings. These certification programs revolutionize the way cities and communities are planned, developed and operated in order to improve the quality of life of people around the world. The programs

* Corresponding author.
E-mail address: farzad.jalaei@uottawa.ca (F. Jalaei).

https://doi.org/10.1016/j.scs.2019.101979
Received 31 May 2019; Received in revised form 16 October 2019; Accepted 18 November 2019
Available online 22 November 2019

provide a framework to plan, design, measure and manage the performance of social, economic and environmental conditions on a city-wide or community level. Leadership in Energy and Environmental Design (LEED), which is one the most well-known and commonly used green building rating systems, is developed by the U.S. Green Building Council (USGBC, 2018; Wu et al., 2017) and therefore its application in design, credit analysis, and documentation, that must be submitted in order to reach the required type of certification (Eastman, Teicholz, Sacks, & Liston, 2008), is significant in the research community (Luo & Wu, 2015).

LEED® uses an online rating system, which gives scores on choice of site, water efficiency, energy and atmosphere, materials and resources, indoor environmental quality, locations and linkages, innovation in design and regional priority. When the results are combined, successful projects are categorized as Certified, Silver, Gold or Platinum (Alwan, Greenwood, & Gledson, 2015). Among the most assuring developments in the Architecture, Engineering and Construction (AEC) industry (Eastman et al., 2008; Nizam, Zhang, & Tian, 2018), Building Information Modelling (BIM) provides a collaborative platform to help the construction project to be digitally modelled in a virtual environment accurately (Wang, Qian, Cui, & Chao, 2018).

With the help of BIM, as designers use tools to assign BIM to create 3D models of buildings where the design materials and schemes are selected from the built-in database of these tools, owners are provided with a choice to visualize the building before any physical operation takes place (Jrade & Jalaei, 2013). In addition, architects and engineers can efficiently share information related to sustainability within an integrated working platform, such as day lighting and energy consumption, and in this way, the sustainability analysis would be consistently integrated into the design process (Chong et al., 2017; Hiyama, Kato, Kubota, & Zhang, 2014).

The constant use of BIM applications in simulating sustainable design of buildings with continuous progression of LEED rating system have accelerated the USGBC to reassess process of the technology and infrastructure of administering the LEED certification. Clearly, such a reassessment was revealed in the principle of LEED automation, which is a new business paradigm of LEED certification. LEED automation operates as an application carries out three key functions for project teams and users, which provides automation of different LEED documentation processes, brings customers a unified look of their LEED projects, and standardizes LEED content and its distribution throughout multiple technology platforms persistently. There are some terms like Plug-In or Add-In used in BIM tool to indicate a module containing an algorithm that makes use of the BIM tool's Application Program Interface (API) (Jalaei & Jrade, 2014), which provides the information exchange to different programs. Autodesk Revit API has been helped the AEC industry to operate different kind of integrations and analysis with different areas by allowing the user to program with any. NET compliant language including C# which is used in this study (Carvalho, Bragança, & Mateus, 2019).

This study, as represented in Fig. 1, describes an automated methodology to assess the potential rate of sustainability for proposed buildings based on LEED certification system while providing a framework for calculating LEED points in accordance to BIM at the conceptual stage. The integrated model uses the LEED for New Construction (LEED-NC) information of the buildings' components, which are stored in connection with BIM environment. In this practice, after data collection and the input data for BIM model along with an advanced database for different materials, an innovative plug-in is developed to calculate the accumulated potential LEED credits with access to API of the BIM tool (i.e. Autodesk Revit, Google Map) and their library. In addition, a questionnaire is provided through the developed plug-in in the BIM environment. Evaluation of environmental issues automatically using Google Map and its direct connection with BIM are about the innovative approaches of this study as well. This research also proposes a methodology to estimate the potential LEED points for the criteria that are not normally measurable from the design information by using the Distance Weighted K-Nearest Neighbor (DWKNN) method by evaluating 76 LEED certified projects. The basis of this research is based on sustainability potential estimation rather than merely calculation of points. Despite the previous studies that just evaluate and identify a limited numbers of criteria to apply the automation methods, the proposed methodology would enable users to consider all LEED criteria at the same time to come up with almost fully automated process to identify the potential number of points that the new constructed building must accumulate to comply with the desired level of certification. It minimizes user input and increases the calculations efficiency.

2. Literature review

Sustainability in design of buildings has become a prominent consideration in creating sustainable cities. In this regard, decisions that are made in early design stages have a significant impact on the real environmental impacts of buildings (Basbagill, Flager, Lepech, & Fischer, 2013). While it is proven that construction process has an impact on the environment (Mah, Manrique, Yu, Al-Hussein, & Nasseri, 2011), many organizations and individuals have taken actions to the growing demand for green buildings. Cidell and Cope (2014) used linear regression to demonstrate the relationship between the presence of a municipal green building policy and the number of registered green buildings for all cities in the US with population over 50,000. They also proved that the presence of a policy leads to more sustainable buildings. BIM redefines the way of how the construction sector builds and works together, as it is a smart, innovative, interactive and responsive technology which helps the professionals to take model's data and use it efficiently to provide meaningful information for other related workflows and processes. Therefore, project participants would be able to design, analyse, sequence, and explore a project within a digital environment to identify the level and effect of changes to develop a sustainable design (Hoff, 2008; Saieg, Sotelino, Nascimento, & Caiado, 2018). Sustainable design strategies combined with BIM has the potential to produce a high performance and efficient design of proposed buildings while changing the traditional design practices. BIM-based technologies have enticed increasing research attention in sustainability rating certification and therefore, its application would help project stakeholders, including clients and contractors to calculate green building rating system scores in pursuit of LEED certification (Wong & Kuan, 2014). More specifically, a number of BIM applications are developed to address sustainability issues in the design process. Most of green BIM applications are designed for building performance analyses and simulations, such as energy performance analyses and some integrated building performance optimization (Lu, Wu, Chang, & Li, 2017).

Many countries and international organizations have established rating systems for sustainable construction. Accordingly, several methodologies have been developed to initiate the degree of accomplishment of environmental purposes, guiding the planning and design stages. The first of such systems was the Building Research Establishment Environmental Assessment Method (BREEAM) (Baldwin, Yates, Howard, & Rao, 1998). Secondly, other methodologies, such as Green Star from Australia (GBCA (Green Building Council of Australia, 2008), the Comprehensive Assessment System for Building Environmental Efficiency (CASBEE) from Japan (CASBEE, 2008), the Building and Environmental Performance Assessment Criteria (BEPAC) from Canada (Cole, Rousseau, & Theaker, 1993), Green Standard for Energy & Environmental Design (G-SEED) from Korea (Wang, Lee, Arch, & Park, 2014) and the Leadership in Energy and Environmental Design (LEED) from the United States (USGBC, 1999) were developed and are currently widely applied. Other comprehensive available tools for environmental assessment methods can be found in Ding (2008), the Whole Building Design Guide (WBDG, 2008), and the World Green Building Council (WGBC, 2008).

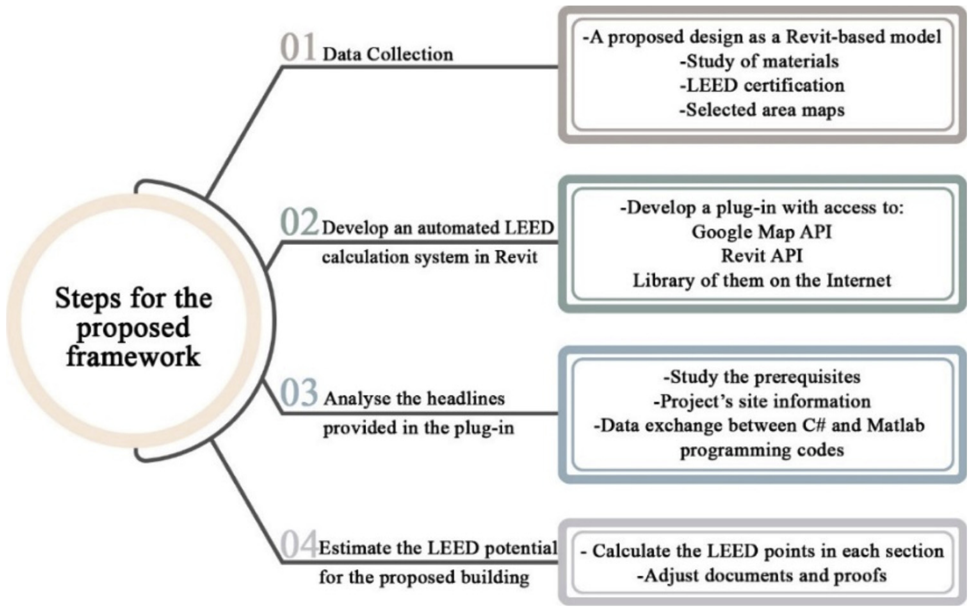

Fig. 1. Steps for the proposed framework.

Despite the extended use of existing methods and tools, LEED is a significant and commonly used rating systems and it has established a considerable credibility among the experts (Huo, Yu, & Wu, 2017; Pulselli, Simoncini, Pulselli, & Bastianoni, 2007). Moreover, since this research is concentrated on the North American market, the sustainability indicators used by the LEED rating system has a particular significance to this study. Similar to the other available rating systems, the LEED is based on credits and points (USGBC, 2005). Through each credit, the system, based on its credits, evaluates the performance of the possible building and assigns points if the requirements are fulfilled in the following areas including: Sustainable Site (SS), Water Efficiency (WE), Energy and Atmosphere (EA), Materials and Resources (MR), Indoor Environmental Quality (IEQ), Innovation and Design process (ID) and Regional Priority (RP). Each category contains a number of criteria and sub-criteria, some of them are assigned a certain number of credits and others are considered as prerequisites (CAGBC, 2011). LEED® v4 addresses 21 different market sector adaptations, including: New Construction, Core and Shell, Schools, Retail, Data Centres, Warehouses and Distribution Centres, Hospitality, Healthcare (USGBC, 2018). The number of points acquired by the project determines the level of LEED certification as follows: (LEED v4 for Building Design and Construction): 40–49 points for LEED Certified, 50–59 points for LEED Silver, 60–79 points for LEED Gold, and 80–110 points for LEED Platinum. In North America, one of the major industry drivers for addressing embodied carbon is the LEED green building rating system.

Autodesk Revit, among commercially available BIM products, is broadly used by architects, designers, engineers, and contractors (Dzambazova, Demchak, & Krygiel, 2008). One of the benefits of BIM is its potential to share a Platform API, which provides it with a chance to integrate external applications. Revit enables assigning external factors to a BIM model (Nguyen, Toroghi, & Jacobs, 2016). With the help of BIM applications, designers are provided with more integrated and visualized views of building performance in the early design phase. Considering building energy performance, despite traditional assessment methods consuming significant amount of time and effort, BIM tool has an advantage to calculate the whole building energy analysis through parameters stored on a database (Shoubi, Shoubi, Bagchi, & Barough, 2015).

As presented in a review study by Ansah, Chen, Yang, Lu, and Lam (2019) about BIM integration with green building assessment, they conclude from the literature that the Revit API is the most commonly used approach in promoting plug-ins for Green Building Assessment Schemes (GBAS) area. They categorized databases within BIM/GBAS frameworks as augmented, external and functional databases. Within a thorough database, Materials and Resources criteria can be evaluated through BIM model. The other criteria such as energy & performance, indoor environmental quality, water efficiency and site location depend on other tools such as Web Map Service providers (i.e. Google Maps). According to Castro-Lacouture, Sefair, Flórez, and Medaglia (2009), the proper choice of materials in addition to the selection of appropriate technologies in building construction process is required for the green building movement. They proposed a study as a mixed integer optimization model integrating design and budget constraints while maximizing the number of credits fulfilled LEED rating system to help the decision making. They outlined a model as their case study in Columbia based on a modified LEED rating system for building evaluation that assimilates design, budget, and environmental requirements simultaneously to demonstrate a better set of material and their extent of use in green buildings. The designed model provides the user with a possibility to include preferred materials and design parameters through design constraints, without imposing a restrictive and costly environmental solution. While this method embracing transparent requirements along with reduced data complexity, its application must be considered as a first phase by builders towards a more prospering environmental building design.

A study conducted by Azhar, Carlton, Olsen, and Ahmad (2011) demonstrates that how planners and designers may use BIM to analyse sustainability in pursuit of LEED certification by developing a conceptual framework that illustrates the relationship between various LEED credits and the related BIM-based sustainability analysis. This framework is then validated by using data from an under-construction school building as a case study and therefore, it only reports the validation results of energy and atmosphere, water efficiency and indoor environmental quality credits. It is reported that documentation for 17 LEED credits and 2 prerequisites for a total of 38 points could be calculated directly/indirectly. The results of this study show that there is no one to one relationship between LEED certification process and BIM-based sustainability analysis because of the lack of LEED integration features in the used software and the analysis results can be used directly, semi-directly or indirectly to generate LEED documentation.

Jalaei and Jrade (2015) proposed a methodology that integrates BIM with the Canadian green building certification system (LEED©

Canada) that automatically calculates the collected points from LEED (for MR and EA criteria) and the related registration costs for green and certified materials used in designing sustainable buildings all within BIM environment. Designed plug-in link the BIM model with an external database, which maintains sustainable materials and assembly groups to be able to produce design information of the proposed building. The limitation of the proposed model is that at the design stage, the details of the construction project due to the stored information are merely for the components that are commonly used in construction projects, which might not be used globally. The database is also developed using several BIM files whereas a particular number of them had credibility. On the other hand, the proposed model is not fully automated, and the link to the database and the data model link with the environmental tool and the calculation of costs are done manually. Moreover, the evaluated LEED credits are just limited to the MR and EA credits and for the rest criteria, no strategy are proposed.

Based on the study of Nguyen et al. (2016), as the assessment of green buildings based on LEED certification is challenging and time-consuming due to its complicated process, they presented building information in a way that the data contained in LEED green building system can be extracted from the BIM model to support the assessment. Therefore, they developed a framework representing the LEED green building criteria to be implemented into a BIM platform to achieve an automated tool for the rating of a green building design. As proposed by Chen and Nguyen (2017), a framework for the integration of BIM and Web Map Service (WMS) technologies for location and transportation analysis in green building certifications due to the lack of a powerful map application in present BIM tools. Using Autodesk Revit API and Google Maps API as the development tools, it converts the integration model into the BIM-integrated plugin in Autodesk Revit. Their developed BIM-LEED framework and plugin could help to calculate the points of the two most important credits under the Sustainable Sites category, Credit 2 (denoted as SSc2) and Credit 4.1 (denoted as SSc4.1), efficiently. WMS could help users get map data, do route-planning tasks, develop documents for LEED submittal, export proof map images with all required information marked on them, and generate an Excel file containing all location-related data.

As reviewed in the recent literature and due to the changes in every LEED update for rating purposes, there are always some missing credits, which prevents achieving to the fully automated BIM- LEED integration. Since the missing values are considered as the general aspect of different research areas (Faisal & Tutz, 2017), it denotes a gap or a lack of a value for one variable in the data list. These missing credits are normally the ones that could not be measured either at the conceptual stage of the building projects or by current design and simulation tools. Since predicting missing data requires to understand the similarities between the current building project and the certified buildings that are already evaluated, it is required to identify an optimum solution to predict the missing values from available historical data. Therefore, it is common to deal with missing values in real life data sets as the values of objects' variables, which are sometimes missing under examination (Todeschini, 1990). There are a number of practices have been proposed in the literature to treat missing data (Mundfrom & Whitcomb, 1998).

Case Deletion (CD) is a common practice in most of statistical packages as well as a default method of many programs. This method involves removing all instances with missing values for at least one feature. Since relevant attributes should be conducted even with a high degree of missing values, CD is less unpredictable if it contains minimal loss of sample size while there is no structure or pattern to the missing data. This method should be applied only in practices in which data are missing completely at random, since it has been produced more biased estimates than alternative methods (Little & Rubin, 2002). Mean Substitution is a commonly practiced approach as it sustains the mean of a variable's dispersal; however, mean substitution typically distorts other attributes of a variable's dispersal (Malarvizhi & Thanamani, 2012). This method is an overestimated sample size, an underestimated variance, a negatively biased correlation, and new values' dispersal as a flawed illustration of the population values, since added values that are equal to the mean have changed the shape of the dispersal. Median Substitution method is used instead of Mean Substitution to ensure robustness as the mean is disturbed by the existence of outliers. Therefore, the median of all known values for an attribute in a class displaces the missing data for that given characteristic. In the case that the distribution of the values of a given attribute is distorted, the Median Substitution would be a recommended method.

The Nearest Neighbor rule (NN) (Cover & Hart, 1967), is one of the oldest and simplest classifiers. The basic rationale for NN is given a set of the training samples and a query, find a point that is the closest to the query, and then assign its class label to the query. KNN is an extension of NN, which a query is labelled by a majority vote of its k-nearest neighbors in the training set. According to KNN or NN, Let $T = \{(x_i, y_i)\}_{i=1}^{N}$ denote the training set, where $x_i \in \Re^m$ is training vector in the m-dimensional feature space, and y_i is the corresponding class label. Given a query x′, its unknown class y′ is assigned by two steps.

Firstly, a set of k similar labelled target neighbors for the query x′ is identified. Denote the set $T' = \{(x_i^{NN}, y_i^{NN})\}_{i=1}^{k}$, arranged in an increasing order in terms of Euclidean distance $d(x', x_i^{NN})$ between x′ and x_i^{NN}

$$d(x', x_i^{NN}) = \sqrt{(x' - x_i^{NN})^T (x' - x_i^{NN})}. \tag{1}$$

Secondly, the class label of the query is predicted by the majority voting of its nearest neighbours:

$$y' = \underset{y}{\arg\max} \sum_{(x_i^{NN}, y_i^{NN}) \in T'} \delta(y = y_i^{NN}). \tag{2}$$

Where y is a class label, y_i^{NN} is the class label for the *i*-th nearest neighbour among its *k* nearest neighbours. $\delta(y = y_i^{NN})$, the Dirac delta function takes a value of one if $y = y_i^{NN}$ and zero otherwise. Along with the simplicity, effectiveness, intuitiveness and competitive classification performance in many domains, the appeal of KNN stems from only a single integer parameter k, and the high classification performance with increasing the amount of training samples. As an improvement to KNN, Dudani introduced a distance weighted KNN rule (WKNN) with the basic idea of weighting close neighbors more heavily, according to their distances to the query (Dudani, 1976). One major challenging problem, yet to be resolved for KNN, is the selection of the neighborhood size k, which can have a significant impact on the performance of KNN-based classifiers (Wu et al., 2008). It has been found that the classification performance of KNN intrinsically results in the estimate of the conditional class probabilities from training set in a local region of data space, which contains k nearest neighbors of the query (Zavrel, 1997). The estimate is affected by the sensitivity of the selection of the neighborhood size k, because the radius of the local region is determined by the distance of the k-th nearest neighbor to the query and different k yields different conditional class probabilities. If k is very small, the local estimate tends to be very poor owing to the data sparseness and the noisy, ambiguous or mislabelled points. In order to further smooth the estimate, we can increase k and take into account a large region around the query. This study is inspired by the sensitivity issue of different choices of the neighborhood size k in KNN-based classifiers. In WKNN, the closer neighbors are weighted more heavily than the farther ones, using the distance-weighted function. The weight w_i for i-th nearest neighbor of the query x′ is defined as follow:

$$w_i' = \begin{cases} \dfrac{d(x', x_k^{NN}) - d(x', x_i^{NN})}{d(x', x_k^{NN}) - d(x', x_1^{NN})} & ,if\ d(x', x_k^{NN}) \neq (x', x_1^{NN}) \\ 1 & ,if\ d(x', x_k^{NN}) \neq (x', x_1^{NN}) \end{cases} \tag{3}$$

Then, the classification result of the query is made by the majority weighted voting:

$$y' = \underset{y}{\arg\max} \sum_{(x_i^{NN}, y_i^{NN}) \in T'} w_i' \times \delta(y = y_i^{NN}). \quad (4)$$

According to the Eq. (4), a neighbor with smaller distance is weighted more heavily than one with greater distance: the nearest neighbor gets weight of 1, the furthest neighbor a weight of 0 and the other neighbors' weights are scaled linearly to the interval in between. The key issue is the way to select a proper neighborhood size k, which largely affects the classification performance of KNN (Wu et al., 2008). As for KNN, the small training sample size could greatly affect the selection of the optimal neighborhood size k and the degradation of the classification performance of KNN is easily produced by the sensitivity of the selection of k. From a theoretical point of view, the classification performance of KNN is determined by the estimate of the conditional class probabilities of the query in a local region of the data space, which is determined by the distance of the k-th nearest neighbor to the query. Thus, the classification performance is very sensitive to the selected value of k. Furthermore, the simplest majority voting of combining the class labels for KNN could be a problem if the nearest neighbors vary widely over their distances and the closer ones more reliably indicate the class of the query object. Nevertheless, WKNN still suffers from the issue because of the existing outliers, particularly in the case of the small sample size (Fukunaga, 1990). Gou, Du, Zhang, and Xiong (2012) proposed a new distance-weighted k-nearest neighbor rule (DWKNN). In this approach, the dual distance-weight substitutes the corresponding weight for each nearest neighbor in WKNN. The dual weight is determined through multiplying the original weight from Eq. (4) by another new weight. In contrast to WKNN, the new method reduces the weight of each nearest neighbor except the first closest and the k-th nearest neighbors. It can keep from giving too much weight to the outliers by reducing the weights of other neighbors in the set of k nearest neighbors for each query and improve the classification performance. Hence, DWKNN can deal with the outliers in the local region of a data space, in order that the degree of the sensitivity of different choices of k can be degraded.

Let $\bar{T} = \{(x_i^{NN}, y_i^{NN})\}_{i=1}^{k}$ denote the set of the k-nearest neighbors to the query $\bar{x}$ arranged in an increasing order according to the distance between $\bar{x}$ and $x_i^{NN}[d(\bar{x}, x_i^{NN})]$, and $\bar{W} = \{\bar{w}1, \ldots, \bar{w}k\}$ be the set of the corresponding dual weights. DWKNN is based on WKNN to give different weights to k nearest neighbors according to their distances, with closer neighbors having greater weights. Nevertheless, different from the weights in WKNN, we assign to the i-th nearest neighbor x_i^{NN} of the query $\bar{x}$ a dual weight $\bar{w}_i$, defined by the dual distance-weighted function as below:

$$\bar{w}_i = \begin{cases} \dfrac{d(\bar{x}, x_k^{NN}) - d(\bar{x}, x_i^{NN})}{d(\bar{x}, x_k^{NN}) - d(\bar{x}, x_1^{NN})} \times \dfrac{d(\bar{x}, x_k^{NN}) + d(\bar{x}, x_1^{NN})}{d(\bar{x}, x_k^{NN}) + d(\bar{x}, x_i^{NN})} \\ 1 \end{cases}$$

$$, if\ d(\bar{x}, x_k^{NN}) \neq (\bar{x}, x_1^{NN})$$

$$, if\ d(\bar{x}, x_k^{NN}) \neq (\bar{x}, x_1^{NN}) \quad (5)$$

And then, we label the query $\bar{x}$ by the majority weighted vote of k nearest neighbors, the same as Eq. (4).

$$\bar{y} = \underset{y}{\arg\max} \sum_{(x_i^{NN}, y_i^{NN}) \in \bar{T}} \bar{w}_i \times \delta(y = y_i^{NN}). \quad (6)$$

It is important to note that the dual weight of each nearest neighbor consists of two parts: the first part is same as the weight in WKNN, the second one is a new distinct weight defined by user, and they both build on the basic idea of the distance-weighted scheme. With respect to Eq. (5), it is obvious that the dual weight $\bar{w}_i$is smaller than the weight w_i computed by Eq. (3) in WKNN, except the weights of the first and k-th nearest neighbors. As a result, the corresponding neighbor x_i^{NN} has less influence on the classification result of the query. The dual weight drops quickly from 1 at the distance of the first nearest neighbor to 0 at the distance of the furthest k-th nearest neighbor.

This study proposes a methodology to automate calculation of the sustainability rate of a building based on LEED certification system. With an integrated BIM-LEED platform, owners and designers get benefit from the developed model at the conceptual design process. In the Revit environment as a BIM tool, a plug-in is developed to automatically calculate the number of LEED V4 points. The proposed plug-in is in completion of the previous endeavour of authors to enhance BIM-LEED automation towards being fully automated by using the API of Revit software and the C # coding using the Visual Studio environment (Jalaei & Jrade, 2015). In order to assess the checklist tiles of LEED certification, there is a need of an access to the API of Google Map and send information to the main source of this software. The input data is sent by the plug-in to the database of this software within the network, and would be converted to the C # code, and the output information is received. To cover more criteria and credits, it is needed to collect a handful of information through questionnaire and some others from an energy analysis tool integrated in BIM tool, user guideline in using relevant software and analyzing relevant data by the plug-in. For the missing values, an implementation of the modified DWKNN method should be conducted in the backend of the proposed plug-in tool to enable the developed tool to predict the whole scale LEED points that a concept design could potentially earn.

3. Methodology

The merit of the proposed method in this study compared to previous studies is to simulate, predict and calculate most of the credits in each section of the LEED checklist and provide a solution for each category. The credits of 7 categories from the guide list would be evaluated including: Location and Transportation (LT), Sustainable Sites (SS), Energy and Atmosphere (EA), Materials and Resources (MR), Indoor Environmental Quality (IEQ), Innovation in Design (ID) and Regional Priority (RP). The obtained points from all assessed credits have to be adjusted based on total credit points since there are categories in which a limited number of credits could have been examined or have not been reviewed in general such as SS and IEQ. For this missing data prediction, a mathematical data mining approach would be employed. The two remaining credits (SS and IEQ) could not be directly calculated from the available data and specifications in the design. The nature of evaluating and quantifying SS and IEQ is through the auditing the building site or installing devices to measure the indoor air quality of the built environment. These types of data could not be collected from the BIM model at the early design stage. In order to calculate the whole scale LEED credits for the design, it is required to estimate missing credits from the credits that directly calculated from the BIM model. As explained and justified in the literature review, the best and simplest method in machine learning in this matter is DWKNN that uses a database in which the data points are separated into several classes to predict the classification of a new sample point.

This research is explored in the following four phases and a flowchart of the proposed framework, which determines the potential LEED-based sustainability of a building in accordance with the categories considered in the guidebook is presented in Fig. 2. In this flowchart, the components of the proposed framework, the correlation between them and the stages of calculation are specified. At first, there are some prerequisites for the estimation of potential LEED points of the building, which is required to be considered. They include: 1) The study area is supposed to be more than 93 square meters, 2) the building has not to be within 38 m of protected wetlands, and 3) the building has not to be close to high-quality agricultural or endangered species land. In case that the proposed building meets these requirements, the next steps would be followed. The credits of 6 categories from the LEED v4 rating system for building design and construction have been evaluated,

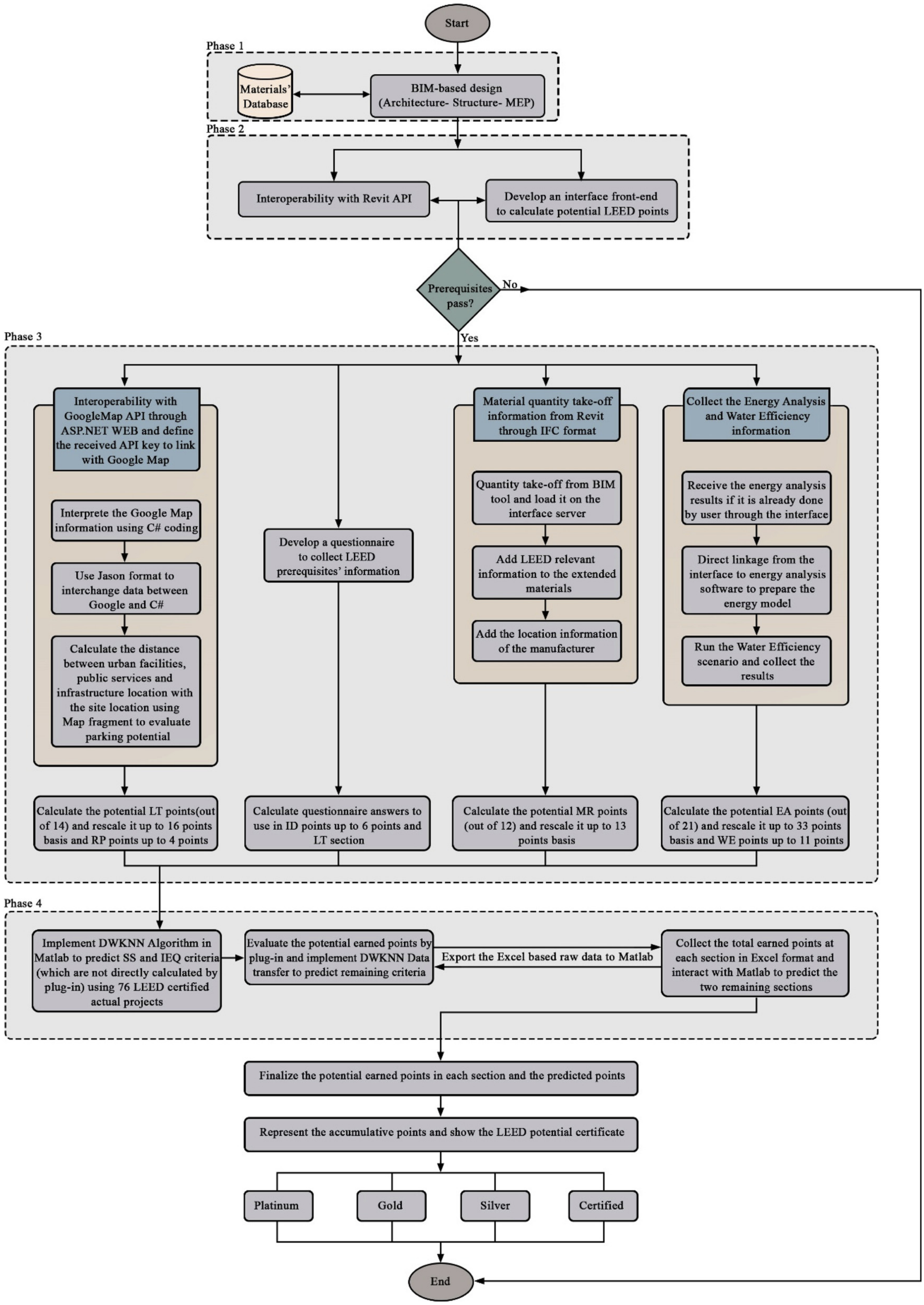

Fig. 2. Flowchart of the integration process.

which include: 14 points of LT, 11 points of WE, 22 points of EA, 12 points of MR, 6 points of ID and 4 points of RP, which would be a total of 69 points out of 110. Depending on the type of building, the questionnaire would be answered by the users. Some credits of LT and ID are assessed within answering to these questions. Considering the case study location on the map, the access to the neighborhood facilities, the number of services within particular distances, walking pathways, access to public transportation services and other mentioned items in the rating system need to be estimated. By using the API of the google map and linking it to the Revit, the mentioned data could be retrieved and added to the database. As a result, a number of credits of LT and RP would be collected. Then, the used materials in the proposed design are displayed in 5 sections (i.e. walls, windows, columns, ceilings and floors) for users in order to monitor the details of the materials, such as manufacturer's location, credits of Materials and Resources (MR1, MR2, MR3, MR4, MR5) and ID would be stored in the relevant database. Following, the state of energy and water efficiency analysis is provided and with user guide, it would be possible to access the analysed data using energy analysis tool. Following, by using the output data of each stage, reviews and relevant points would be analysed. Since for SS and IEQ, few credits could be assessed by conducting calculation methods (2 points for SS and 2 points for IEQ), authors propose using the WKNN method to calculate the credits for the missing categories from data mining on the 76 LEED certified projects.

3.1. Phase 1

Phase 1 includes BIM-based design and architectural plans. The use of a database of various types of available materials is also a prerequisite in the process and needs to be collected in this section for the verification of the green building. The developed database in this research is collected from the literature as well as the USGBC websites, published data and suppliers' web pages, which consists of information related to green components and then, implemented to the library of them stored in BIM's tool. Consequently, users would be capable of selecting a variety of sustainable components from the database with their associated specifications, potential LEED points, and manufacturer contact information (Jrade & Jalaei, 2013). Revit is considered as a BIM parameter modelling tool and BIM model contains parameters related to sustainable design, which automatically could be seized throughout the design process. The BIM model in Revit uses elements such as components of a building, including site, architectural elements (exterior/interior walls, roof, doors, etc.), structural elements (foundation, column, beam) and etc. Sustainability indicator parameters while being created and stored as project parameters could be used to establish the potential level of sustainability of a proposed building project. Furthermore, the Revit material schedules are vital in order to input materials information into the model. This information eases the process of energy analysis based on LEED certification reporting processes (Barnes & Castro-Lacouture, 2009).

3.2. Phase 2

Phase 2 develops a sustainability evaluator plug-in within the BIM environment to automatically calculate the potential points of the building. This plugin has been expanded with access to the Revit API. Moreover, for the analysis of LEED categories, it is required to access the Google Map API and send information to its library. By sending the input data through the developed plug-in to its library on the network and converting the collected information to the C# code, the output information would be received. There are some data that should be collected through the questionnaire for those non-existing inputs in the design. The building analysis begins right after the fulfilment of the prerequisites and implementation of architectural plans.

3.3. Phase 3

Phase 3 consists of material and energy analysis of building to cover 6 categories from LEED rating system. The data schema files for exporting BIM information to the energy simulation software include gbXML, IFC, ifcXML and ecoXML files, as presented in Azhar, Brown, and Farooqui (2009). To perform an energy analysis, the green building XML schema—known as "gbXML file format is required to be exported from Revit. The Green Building Studio (GBS), which is a cloud-based tool that provides users with energy analysis possibility through a gbXML file format is used in this study. This service includes high-quality graphic and fast energy performance in buildings based on building resources. Using LEED New Construction (NC), the total achievable points for energy simulation is a maximum of 22 points based on the type of building. Furthermore, it is necessary to achieve a minimum energy cost reduction of 10 % in order to get LEED certification. Building zones should be modelled according to the American Society of Heating, Refrigerating, and Air Conditioning Engineers (ASHRAE) Standard 90.1 (White et al., 2007). Each of the rooms is required to be assigned to spaces, and each space is assigned to the same air conditioning unit by using the MEP zoning function. Energy simulation shall be implemented with the proposed model as a baseline model following the guidelines of ASHRAE 90.1-2010, Appendix G. By comparing the baseline and proposed and model simulation results, the actual evaluation of the relative energy cost-savings rate shall be assessed. The ASHRAE Standard 90.1 presents the requirements for the envelope, the standard for the Heating, Ventilation and Air--Conditioning (HVAC) equipment and lighting power density for the baseline model. To create a proposed case, the first step is to enter the location information and the weather data, which transfers from BIM model automatically into GBS. Weather data is applied to both the baseline and proposed models. The next step is to create an energy model template. Through using a template, the purpose specific occupancy density, lighting load density, and the instrumental load density as well as each load specific schedule are then set. This setting would be either automatically by the software or manually by user. Furthermore, thermal properties such as infiltration, room temperature, and the envelope would be prepared. After creating a template, the gbXML that went through the verification stage is imported to GBS through the embedded plug-in in Revit for an energy simulation. This analysis enables user to cover credits of EA, WE plus daylight information.

The Material & Resources (MR) category, is also calculated based on the application of shared parameters. Currently, BIM tool (i.e. Revit) supports direct data export through Industry Foundation Classes (IFC) for the materials QTO, where the shared parameters related to the MR could be stored on the interface server. It strictly controls the export with predefined tables and data types. MR credits, including list of subcategories from MR1 to MR5, which all deal with unique material properties that contribute to sustainability. These properties are not predefined in the BIM tool platform and thus require customization in its external database. Simply put, in order to calculate the value of materials that satisfy a certain criterion, which could vary among "reused", "recycled", "regional", and "renewable", it is desirable to tag the materials in the building information model with such features. For every building component, that credit information is input by users in the interface, which is connected to BIM tool.

3.4. Phase 4

Phase 4 according to the mentioned 6 categories, it is necessary to consider the earned points based on the total credits. Thus, the calculated points have to be adjusted based on 110 points since there are two categories (i.e. SS and IEQ, which have not been assessed thoroughly or at all. Therefore, a missing value prediction method needs to be applied. Overall, among the different types of available methods for detecting the missing data, the KNN algorithm, which is described in the literature review section, is one of the most fundamental and efficiently implemented algorithms for nonlinear data interpretation (Bhattacharya, Ghosh, & Chowdhury, 2017). In addition to the several number of improved KNN algorithms which have been proposed and examined in the literature (Fan, Guo, Zheng, & Hong, 2019), the DWKNN method is used in this study since it has a desirable accuracy for the answer and the categorization of credits and expert's opinion. Thus, final points of 76 LEED certified buildings selected and reviewed from USGBC published evaluated and certified projects is considered as a statistical society as collected in the Table 1.

One important issue here is the information of the examined projects that is based on older versions of the LEED with 7 categories while the version reviewed in this study is the LEED V4, which includes 8

Table 1
Information on the obtained points of various LEED certified projects in different categories.

Project.NO	SS	WE	EA	MR	IEQ	IN	RP	Project.NO	SS	WE	EA	MR	IEQ	IN	RP
1	11	7	27	3	8	5	2	39	23	4	17	6	10	6	3
2	15	6	7	4	7	6	4	40	22	10	33	3	6	5	4
3	18	6	17	8	10	4	3	41	15	4	30	5	4	4	2
4	11	10	31	6	13	6	3	42	23	8	11	6	5	6	4
5	19	10	14	5	3	5	4	43	23	2	20	6	9	3	1
6	20	10	13	3	6	4	4	44	19	6	12	5	8	5	4
7	5	7	31	4	10	3	2	45	16	10	12	3	6	5	4
8	23	4	9	2	7	6	3	46	9	7	17	6	7	2	2
9	20	6	8	2	6	6	4	47	16	4	33	7	11	4	4
10	21	8	4	8	10	5	4	48	18	8	10	7	12	5	1
11	20	4	11	6	6	4	1	49	11	8	13	7	9	4	2
12	11	10	11	6	14	6	3	50	21	5	10	2	13	3	2
13	24	8	30	4	7	5	4	51	22	6	14	6	8	5	4
14	24	8	21	7	10	6	4	52	20	7	16	6	10	5	1
15	21	8	10	6	12	5	2	53	21	6	17	5	3	6	3
16	20	4	11	6	12	5	2	54	21	10	33	6	7	6	4
17	13	4	13	6	9	6	2	55	23	8	28	6	12	4	4
18	25	7	24	5	12	6	4	56	10	7	6	6	5	3	3
19	17	10	3	6	2	6	4	57	16	3	24	6	9	5	2
20	16	5	5	2	6	5	3	58	20	10	17	6	5	4	4
21	16	5	21	3	14	4	2	59	26	10	33	7	15	6	4
22	21	10	9	6	12	6	3	60	18	10	31	5	9	3	4
23	17	7	6	6	8	5	2	61	18	6	11	5	5	1	4
24	21	6	21	6	9	2	4	62	17	7	12	2	7	4	4
25	9	10	23	4	6	4	4	63	18	4	21	6	8	3	3
26	11	8	33	6	14	6	3	64	18	6	18	4	7	5	2
27	11	2	6	9	11	2	3	65	16	10	10	7	12	4	4
28	9	7	32	9	8	1	3	66	15	8	10	8	9	3	2
29	21	5	29	9	12	6	4	67	17	5	5	6	12	5	1
30	23	10	30	5	8	6	4	68	18	4	15	6	12	5	2
31	22	3	17	7	9	6	1	69	15	6	10	6	11	4	2
32	11	7	15	5	9	5	2	70	14	6	9	5	10	5	2
33	23	10	10	5	5	4	4	71	13	6	9	7	10	6	2
34	6	4	24	6	8	4	2	72	22	10	9	6	8	4	4
35	21	6	11	5	9	6	2	73	9	7	19	3	3	3	1
36	20	4	8	6	7	4	1	74	21	6	30	6	9	6	4
37	9	2	18	6	11	3	2	75	15	10	21	6	9	4	4
38	20	4	17	4	7	6	3	76	21	7	12	8	11	6	4

categories. Therefore, it is needed to map the differences between categories for the precise calculation. Sustainable Site with 26 points in previous versions is now divided into two categories including Sustainable Sites and Location and Transportation with 10 and 16 points, respectively. For this purpose, the information of the various credits of Sustainable Sites category would be updated on the basis of the new version and the projects' information would be mapped according to the new classification as presented in Table 2. The information of the projects and all credits of Sustainable Sites in older versions of LEED were closely studied. Since the 6 LEED credits (i.e. LT, WE, MR, EA, ID, RP) could be calculated by the proposed methodology, the data represented in the Table 2, which is the converted version of the 76 studied projects based on the LEED V4 (SS in the old version is converted to LT and SS), would be considered as a training data template in calculating the missing SS and IEQ credits that could not be calculated directly.

This method aims to obtain the values relevant to the mentioned 5 factors from the plug-in and to check the Euclidean distance between the earned points in the N-dimensional space with all available points. Then, all points must be arranged according to the Euclidean distance, which is arranged from lowest to highest. The proposed plug-in obtains the information from the user, BIM, energy analysis and simulation tools to calculate the points from LEED mentioned credits. With the request the final analysis from user, it would calculate the earned points in different categories and would generate 7 numbers for LT, WE, EA, MR, ID, RP and SS criteria. For the 6 primary points (all the points except SS), the points are divided by the maximum point that could be earned in each category (i.e. 16, 11, 33, 13, 6, 4 respectively). These six generated numbers would be saved as an output in an Excel file. By running a developed DWKNN algorithm in MATLAB, the saved spreadsheet file would be imported as an input data in MATLAB. Running the algorithm, the actual project data coming from BIM plug-in as well as the actual project input and output data that are already imported in MATLAB, it merely needs an excel file with 6 data to do calculations. The numbers of SS and IEQ for every K would be generated in MATLAB until it comes to a converging point, then the generated indices for each credit would be taken, and the corresponding output would be returned to the plug-in. The first number related to SS, which is multiplied by its maximum point (10) and the second number related to the IEQ that is multiplied by its maximum points (16). By this rational prediction, the points for SS and IEQ are obtained based on the calculated points for the design as well as the points earned by the LEED certified projects. While the points are calculated by the plug-in for all categories, it has reached to its ultimate purpose to estimate the potential LEED points for any proposed projects. After completion of required information by the user within the plug-in, energy analysis, data analysis and accumulating the earned points, it would be possible to evaluate the sustainability of the project based on LEED certification system (Certified; Silver; Gold; or Platinum levels of certification). The earned points by the project in different categories and the total earned points would be calculated and displayed by the plug-in inside the BIM environment and displayed to the user in order to determine the strengths and weaknesses of the project.

4. Model implementation and validation

In this section, the examination of developed plug-in and the items assessed within its environment, and the detailed explanation of

Table 2
Mapped version of 76 studied projects based on LEED V4 classification.

Project.NO	LT	SS	WE	EA	MR	IEQ	IN	RP	Project.NO	LT	SS	WE	EA	MR	IEQ	IN	RP
1	6	5	7	27	3	8	5	2	39	16	7	4	17	6	10	6	3
2	12	3	6	7	4	7	6	4	40	15	7	10	33	3	6	5	4
3	13	5	6	17	8	10	4	3	41	12	3	4	30	5	4	4	2
4	6	5	10	31	6	13	6	3	42	16	7	8	11	6	5	6	4
5	12	7	10	14	5	3	5	4	43	16	7	2	20	6	9	3	1
6	17	3	10	13	3	6	4	4	44	12	7	6	12	5	8	5	4
7	3	2	7	31	4	10	3	2	45	9	7	10	12	3	6	5	4
8	17	6	4	9	2	7	6	3	46	6	3	7	17	6	7	2	2
9	14	6	6	8	2	6	6	4	47	8	6	4	33	7	11	4	4
10	16	5	8	4	8	10	5	4	48	16	2	8	10	7	12	5	1
11	16	4	4	11	6	6	4	1	49	6	5	8	13	7	9	4	2
12	6	5	10	11	6	14	6	3	50	16	5	5	10	2	13	3	2
13	16	8	8	30	4	7	5	4	51	16	6	6	14	6	8	5	4
14	16	8	8	21	7	10	6	4	52	16	4	7	16	6	10	5	1
15	16	5	8	10	6	12	5	2	53	16	5	6	17	5	3	6	3
16	16	4	4	11	6	12	5	2	54	16	5	10	33	6	7	6	4
17	9	4	4	13	6	9	6	2	55	16	7	8	28	6	12	4	4
18	16	9	7	24	5	12	6	4	56	6	4	7	6	6	5	3	3
19	12	5	10	3	6	2	6	4	57	11	5	3	24	6	9	5	2
20	12	4	5	5	2	6	5	3	58	16	4	10	17	6	5	4	4
21	11	5	5	21	3	14	4	2	59	16	10	10	33	7	15	6	4
22	16	5	10	9	6	12	6	3	60	12	6	10	31	5	9	3	4
23	12	5	7	6	6	8	5	2	61	12	6	6	11	5	5	1	4
24	16	5	6	21	6	9	2	4	62	12	5	7	12	2	7	4	4
25	7	2	10	23	4	6	4	4	63	15	3	4	21	6	8	3	3
26	5	6	8	33	6	14	6	3	64	15	3	6	18	4	7	5	2
27	8	3	2	6	9	11	2	3	65	12	4	10	10	7	12	4	4
28	5	4	7	32	9	8	1	3	66	11	4	8	10	8	9	3	2
29	15	6	5	29	9	12	6	4	67	11	6	5	5	6	12	5	1
30	16	7	10	30	5	8	6	4	68	9	9	4	15	6	12	5	2
31	16	6	3	17	7	9	6	1	69	9	6	6	10	6	11	4	2
32	7	4	7	15	5	9	5	2	70	9	5	6	9	5	10	5	2
33	16	7	10	10	5	5	4	4	71	6	7	6	9	7	10	6	2
34	3	3	4	24	6	8	4	2	72	16	6	10	9	6	8	4	4
35	16	5	6	11	5	9	6	2	73	5	4	7	19	3	3	3	1
36	16	4	4	8	6	7	4	1	74	16	5	6	30	6	9	6	4
37	4	5	2	18	6	11	3	2	75	9	6	10	21	6	9	4	4
38	15	5	4	17	4	7	6	3	76	16	5	7	12	8	11	6	4

exercised categories and instances from LEED certification are investigated. To assess the input and output information and how the developed framework works, the plug-in is run to analyze a case study. The process of credits' evaluation and computation of points are presented in the form of divided categories within the plug-in's environment based on the principle of user's practice to maintain consistency of performance. The case study in this research is an office building in the Calgary's downtown west core office district, AB, Canada with a gross area of 13,685 square meters. According to the design concept, the outdoor area is to be dedicated to seasonal exhibitions and street theatres. As explained before, the architectural and MEP plans are required for lighting and energy analysis purposes. The energy analysis and requirements such as improvement level of energy consumption basis in Appendix G of the ASHRAE 90.1 would be accessible to make the plugin fully functional.

As shown in Fig. 3, once the plug-in runs, there are some prerequisite questions for the minimum required area of the project (69 m^2). Since the conditions are met, access to other sections is provided. If the second question is answered in the affirmative, another question appears asking the specialist to consider the extent of the innovation regarding the design and environmental issues of the project in order to consider the innovation points in design.

In the project location tab of the plug-in as illustrated in Fig. 4, it is mostly involved with site location and transportation issues and the city's construction and communications. The main focus of this section is on reducing traffic by using environmentally friendly vehicles, increasing access to welfare services without the use of vehicles and, if necessary, easy access to public transportation. Some significant credits of Location and Transportation and Sustainable Sites categories including sensitive land protection (option 1), surrounding density and diverse uses, access to quality transit and site assessment are analysed within this section. In this tab, the coordination of the project location is transferred from BIM model. By clicking on the location analysis button, as shown in Fig. 4, the plug-in performs its assessments by connecting to the Google Map server, analysing the requested information from the server and converting to the points of credits.

The developed plugin receives the surrounding existing density within a ¼-mile (400-meter) radius of the project, the building's main entrance within a ½-mile (800-meter) walking distance of the main entrance of existing and publicly available diverse uses (bank, hospital, post office, restaurants, library and …), entry of the project within a ¼-mile (400-meter) walking distance of existing or planned bus, streetcar, or rideshare stops, or within a ½-mile (800-meter) walking distance of existing or planned bus rapid transit stops, light or heavy rail stations, commuter rail stations, or commuter ferry terminals from Google Map servers. Moreover, for the Regional priority credit's assessment in this section, which can be earned up to four out of the six credits, there are some identified credits by the USGBC regional councils and chapters as having additional regional importance for the project's location. Therefore, it is possible to be connected to the USGBC website by clicking on the "Regional Priority credits lookup" button as it is an available database of Regional Priority credits and their geographic applicability. In this page, after choosing a preferred design system based on the proposed project and the type of LEED version, the location of the project would be analysed based on six credits and one credit is awarded for each Regional Priority credit achieved, up to a maximum of four. Then, it can be manually entered within the box provided in the plug-in for the assessment of final results.

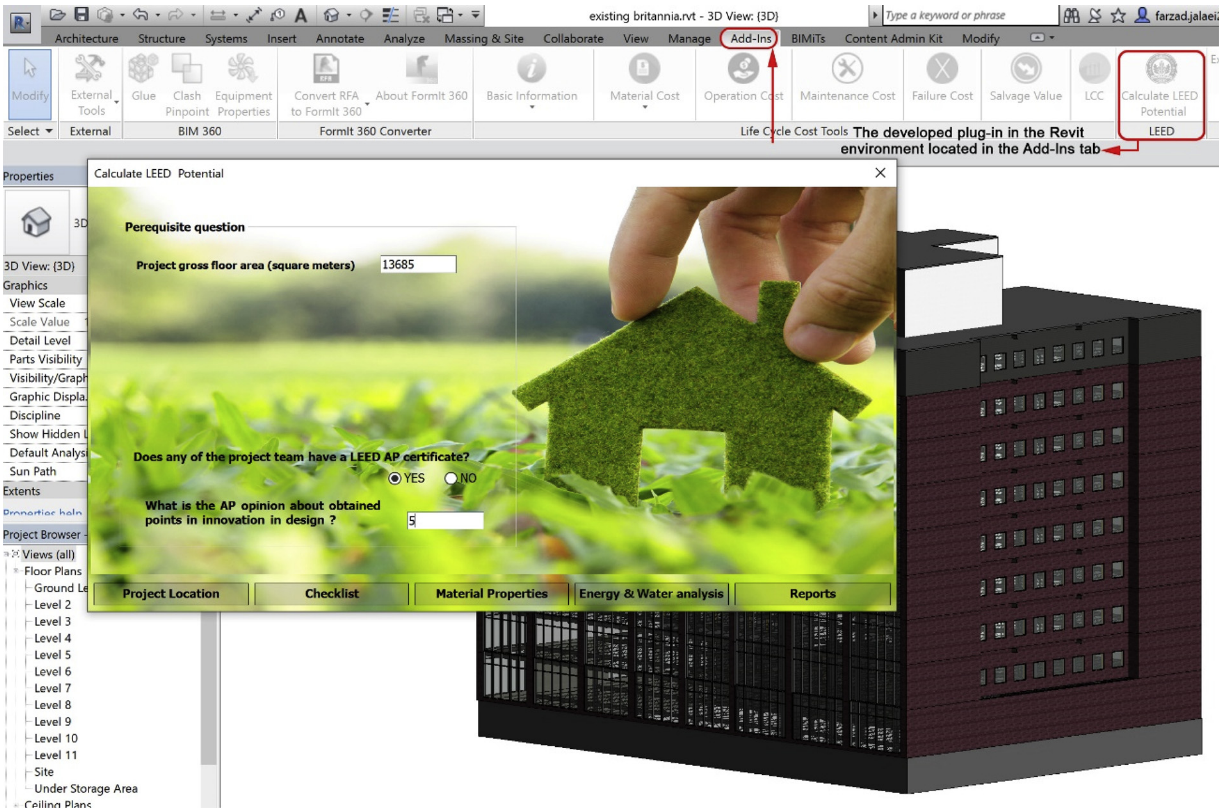

Fig. 3. Required prerequisites of the plug-in.

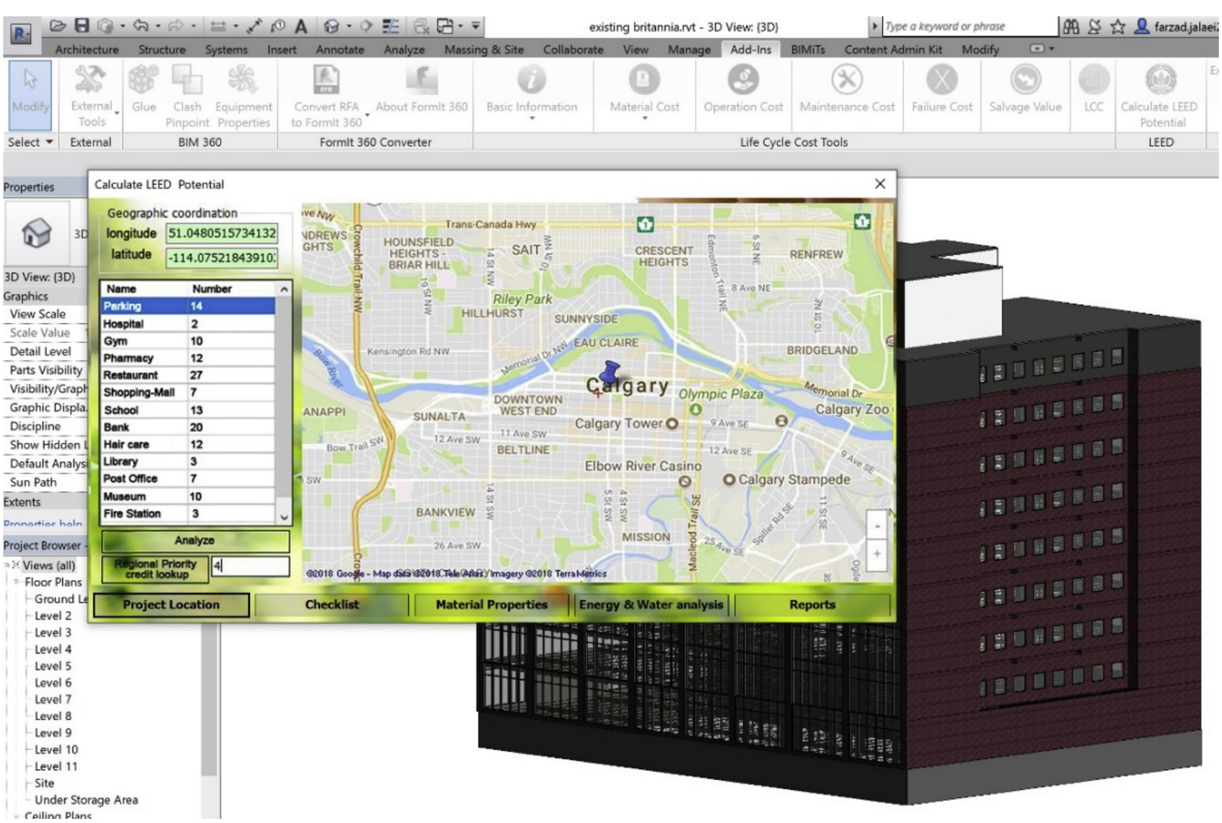

Fig. 4. Project location analysis.

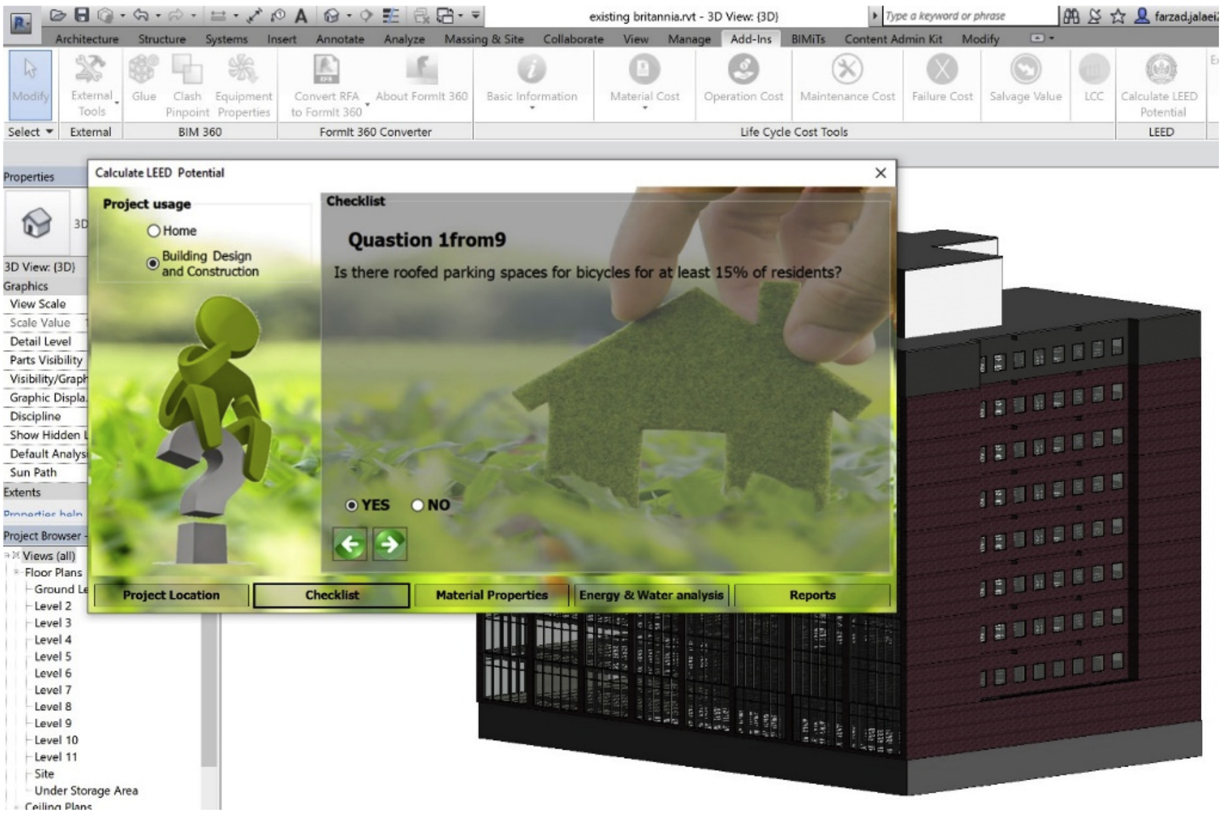

Fig. 5. Checklist questions.

As it is illustrated in Fig. 5, the provided questions are mostly concerned with the neighborhood services and to encourage residents and users to use environmentally friendly vehicles and public transportation services to improve air quality and reduce energy consumption. Based on the answered questions, some credits of LT (i.e. sensitive land protection (option 2), bicycle facilities, reduced parking footprint and green vehicles), SS (site assessment, heat island reduction and light pollution reduction), IEQ (Environmental tobacco smoke control) and Innovation are analysed. In the 10 questions that are provided in this section, the aim is to collect information about a short-term or long-term bicycle storage located within 100 feet (30 m) walking distance of any entrance of the building, roofed parking spaces for bicycles for at least 15 % of residents, a discounted parking rate of at least 20 % for green vehicles for preferred parking spaces? (the discounted rate must be publicly posted at the entrance of the parking area and permanently available to every qualifying vehicle), Electrical Vehicle Supply Equipment (EVSE), which is installed in at least 2 % of the parking spaces of the building (these spaces are identified to be used only for plug-in electric vehicles), at least 5 % of parking spaces in the building, which should be used only for green vehicles, smoking prohibited outside the building except in designated smoking areas located at least 25 feet (7.5 m) from all entries, outdoor air intakes, and operable windows, the lighting control facilities with minimum three lighting levels provided for at least 90 % of occupied spaces in the building (these facilities enable occupants to adjust the lighting to suit their individual tasks), information about the location of the building project if it is in the path of the prime farmland, unique farmland or state-wide or local farmland or if project areas on or within 50 feet (15 m) of a wetland and if the project areas is on or within 100 feet (30 m) of a water body (LEED V4, 2018).

The next tab is material properties. In this section, the main objective is to review and analyse the credits regarding the MR category and to receive information about the materials used in the project. As shown in Fig. 6, all materials used in the project are divided into 5 categories, which is directly extracted from BIM model namely doors, windows, walls, floors and columns. By double clicking on each material extracted from the BIM model, a window is opened, asking user for the specific information related to that material. In this section, there are several ways to take information from the user.

The user requires to consider the information taken from the manufacturer of these materials and in the case that the requirements of credits (MR1, MR2, MR4, MR5) are met, then the innovation points and the price of the materials need to be specified in order to calculate the relevant points. Building life-cycle impact reduction is involved with the optimization of environmental impacts of building products and reuse of materials, and it would be displayed as MR1. At the early design stage, the environmental impacts of a project can be reduced by reusing existing building resources through life-cycle assessment. For the analysis of building and material reuse section of this credit, the plug-in calculates the area of reuse materials as a percentage of the surface area including structural elements such as floors, and interior elements such as walls, doors, floors. The result area is then divided into areas of all doors, windows, columns, floors and walls in the project. Therefore, if the percentages of completed project surface areas are calculated as 25 %, 50 % and 75 %, the earned points would be as 2, 3 and 4 respectively. Building production disclosure and optimization is mostly concerned about the use of products with available life cycle information, which have environmentally, economically, and socially preferable impacts, and it would be implied as MR2. For the analysis of this credit, 1 point can be earned for the Environmental Product Declarations (EPDs) in the case that at least 20 different permanently installed products are used in the project from at least five different manufacturers based on some criteria stored in the database. The information about the products is entered by 1 point for multi-attribute optimization can be earned by using products in accordance with the stored criteria in the plug-in and it would be assessed by cost. Indeed, the plug-in calculates the price of products in all five categories in the database (doors, windows, columns, floors and walls) that have MR2

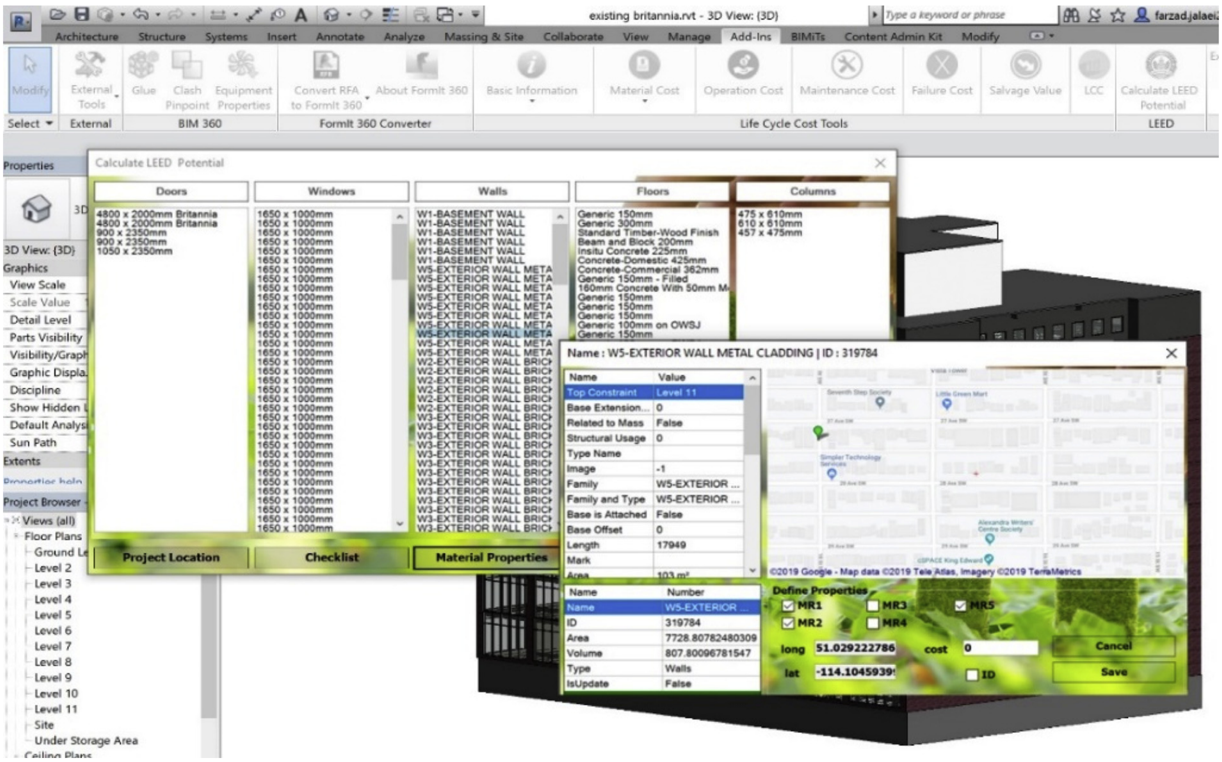

Fig. 6. Material properties and details of used materials.

credentials. After dividing the result to the cost of all products in 5 categories in the project, if the result percentage is about 50 % of the total value, this option would be fulfilled. Building production disclosure and optimization (sourcing of raw materials) is about encouraging the use of products and materials with available life cycle information, which have environmentally, economically, and socially preferable effects, and it would be presented as MR3. The assessment of this credit with maximum 2 points is possible with 1 point for raw material source and extraction reporting and 1 point for leadership extraction practices. It is possible to determine the location of the manufacturer of the materials using the map. In case that the BIM model uses the families designed by the companies, the manufacturers information including their location is already exist in the BIM family. When specifying the location of the manufacturer, the distance between the factory and the project site would be calculated and displayed to the user. By calculating the distance, the plug-in automatically analyses the MR3.

Building production disclosure and optimization (material ingredients) is involved with the use of products and materials with available life-cycle information and it would be introduced as MR4. The first item of this credit as material ingredient reporting with 1 possible point can be earned by using at least 20 different permanently installed products from at least five different manufacturers on the basis of some criteria stored in the BIM database. The information regarding these materials is either entered by the user or is available in the BIM library and if it meets the requirements, the point would be achieved. As for the second item (material ingredient optimization) and the third (product manufacturer supply chain optimization) each with 1 possible point, they can be earned by using products that document their material ingredient optimization based on their cost, and using building products according to some criteria, in turn. The plug-in calculates the price of selected products based on required criteria in all five categories in the database and by dividing the result to the cost of all products in 5 categories, if the result percentage is about at least 25 % of the total value, this option would be accomplished. It is noticeable that the maximum gained points of this credit are 2 points. Construction and demolition waste management is about reducing construction and demolition waste disposed of in landfills and incineration facilities by recovering, reusing, and recycling materials and it would be shown as MR5. The main purpose is about recycling non-hazardous construction and demolition materials. This credit with maximum 2 possible points is analysed within 2 options. Diversion and reduction of total waste material. Diverted materials must have at least three material streams and 50 % of the total construction and demolition material to get 1 point, or they must contain at least four material streams and at least 75 % of the total construction and demolition material in order to get 2 points. The plug-in calculates the volume of materials in all five categories in the BIM database as well as the number of material streams. After dividing the result to the volume of all materials in 5 categories, if the result percentage is about at least 50 % or 75 % of the total, this option would be achieved by 1 point or 2 points, respectively. As for the second option with 2 points, it is essential to not generate more than 2.5 pounds of construction waste per square foot (12.2 kg of waste per square meter) of the building's floor area. This option would be calculated by the plug-in as well.

In the energy analysis tab, assessment of credits is considered in four sections as renewable energy, daylight, water and energy efficiency. As illustrated in Fig. 7, there is a question as the following context: Is there any renewable energy system installed in the building to reduce the use of fossil fuels? If the answer to this question would be No, then the following items of this section would not be displayed and subsequently, the points for this section are not considered, but if the user's answer to this question is positive, there would be a following question and the user needs to answer the question with regard to the functionality of the installed system: what percentage of energy usage is provided by this system? With the user's response to this question, the plug-in calculates the points for this section based on Table 3 and adds points to the analysis of energy.

The other question in this tab is related to optimize energy performance of the building. As explained earlier, in order to calculate the percentage of performance efficiency in this section, the energy consumption of the building should be simulated by software for 8760 h of

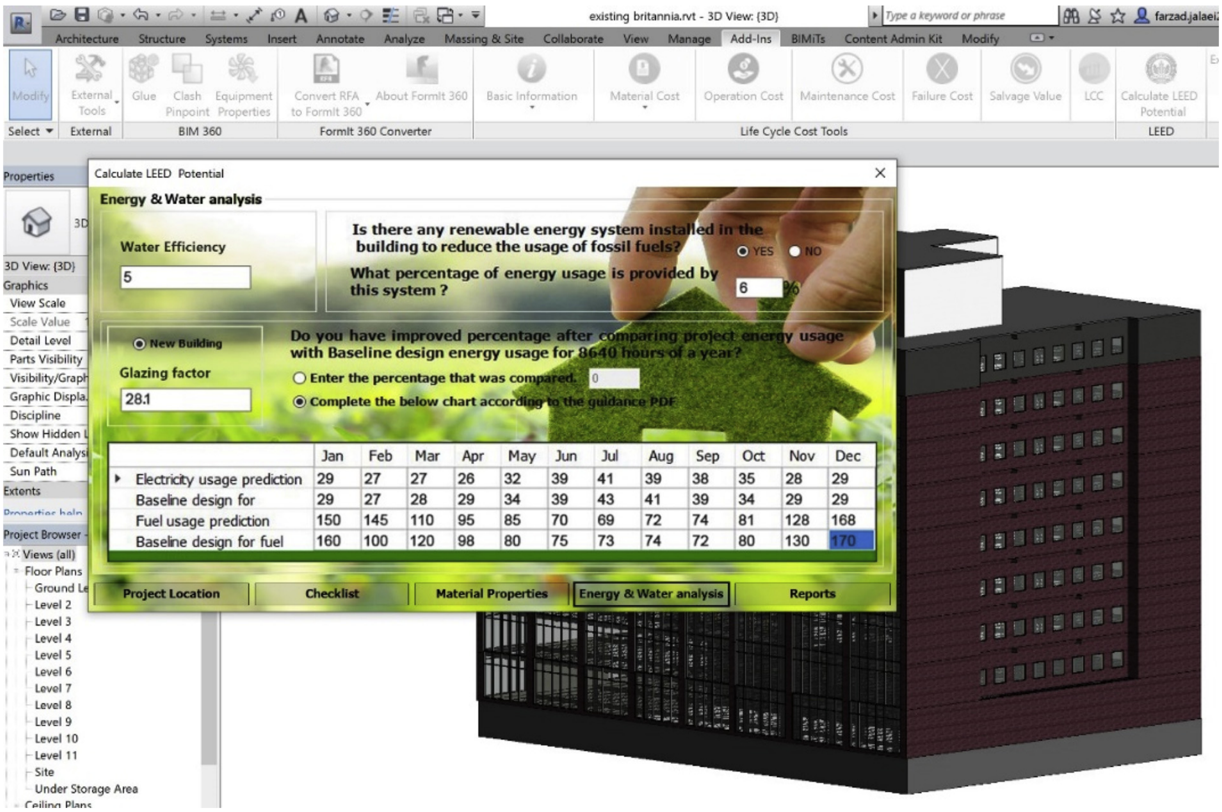

Fig. 7. Energy analysis.

Table 3
Points for percentage improvement in energy performance (LEED v4).

Percentage of renewable energy	1 %	3 %	5 %	10 %
Points	1	–	2	3

a year and be compared to the baseline energy consumption level according to the appendix G of ASHRAE 90.1. In the first place, the user is asked, if possible, to calculate the performance efficiency percentage using energy analysis software and to enter the result in its section, and if this information is not available, then the user needs to enter the results on a monthly basis in the table shown in Fig. 7. Here, instead of doing an hourly basis comparison, a monthly basis comparison is used for the convenience, which is, after all, less precise, but it has an acceptable accuracy according to the experts' views. Another instance in the plug-in's guide is about how to calculate the glazing factor percentage and by its calculation, the energy analysis tab would be completed. As for the completion of energy setting in the analysed tab of Revit and by running the energy simulation function, the gbXML file format of the BIM model would be transferred and the user would be directed to Green Building Studio web page. After detailed analysis of the project, it is necessary to perform the three different steps to calculate the glazing factor, the amount of electricity consumption and monthly fuel consumption as shown in Fig. 8, and the information for water efficiency by creating the default water and wastewater systems and identify the proposed system. GBS would be able to analyse the proposed scenario and evaluate the WE point as shown in Fig. 9.

For the case study project, GBS provides a default of 248 alternative runs besides the base run. There is a possibility to add or remove alternative runs in Design Alternative tab feature as well. The alternative designs give an overall picture on how sensitive the parameters are that affect energy usage. By modifying a broad range of design parameters, user would be able to perform efficient proposed design alongside with all the replacement feasible to make the project design become more sustainable when it is operated. Alternative designs work with parameter values that are orientation, wall construction, roof construction, window glass, skylight glass, infiltration, equipment and lighting. The alternative runs chosen from the default section are representative. Building orientation is also an important parameter that deals with the geometry, building location, size and position of the openings. Applying any changes in direction of the building could make a huge or slight impact on the results. The wall, roof construction and glazing property had extreme impacts on the energy performance. The difference came from the heat loss, solar energy income and natural lighting provided to the building. In the Run Charts tab as the second step, the energy consumption analysis result is divided into total energy consumption, electricity and fuel consumption. Furthermore, each section's consumption rate is available in two formats: energy and cost. As presented in Fig. 8, after adjusting the required energy consumption units, total energy, fuel and electricity consumption per month could be received and entered in the provided chart in the developed plug-in or even the improved percentage could be entered manually by user.

In the baseline case, the energy use intensity, which is a measure of the combined electricity and fuel used by the project, per area (square meter) per year is 1044.3 MJ/m^2/year and in the proposed case it is 740.8 MJ/m^2/year. From the entire energy consumption, the fossil fuel consumption is estimated to 62 %, nuclear 28 %, hydroelectric 5 % and the renewable energy is estimated to be 6 %. If this is converted to the annual energy cost, in the baseline case, the energy cost is $33,124, and in the proposed case it is $30,154 annually. If this was converted into a percentage, the energy saving rate would be 20.34 %. This is equivalent to eight points of 18 possible maximum points in item Energy and Atmosphere (EA) credit 1 based on the LEED NC. The glazing factor percentage is acquired from LEED Daylight tab, and then it should be entered in its identified section in the plug-in. 28.1 % for the glazing is estimated for the proposed design and it collects no LEED points for this project since LEED requires your project achieve a minimum glazing

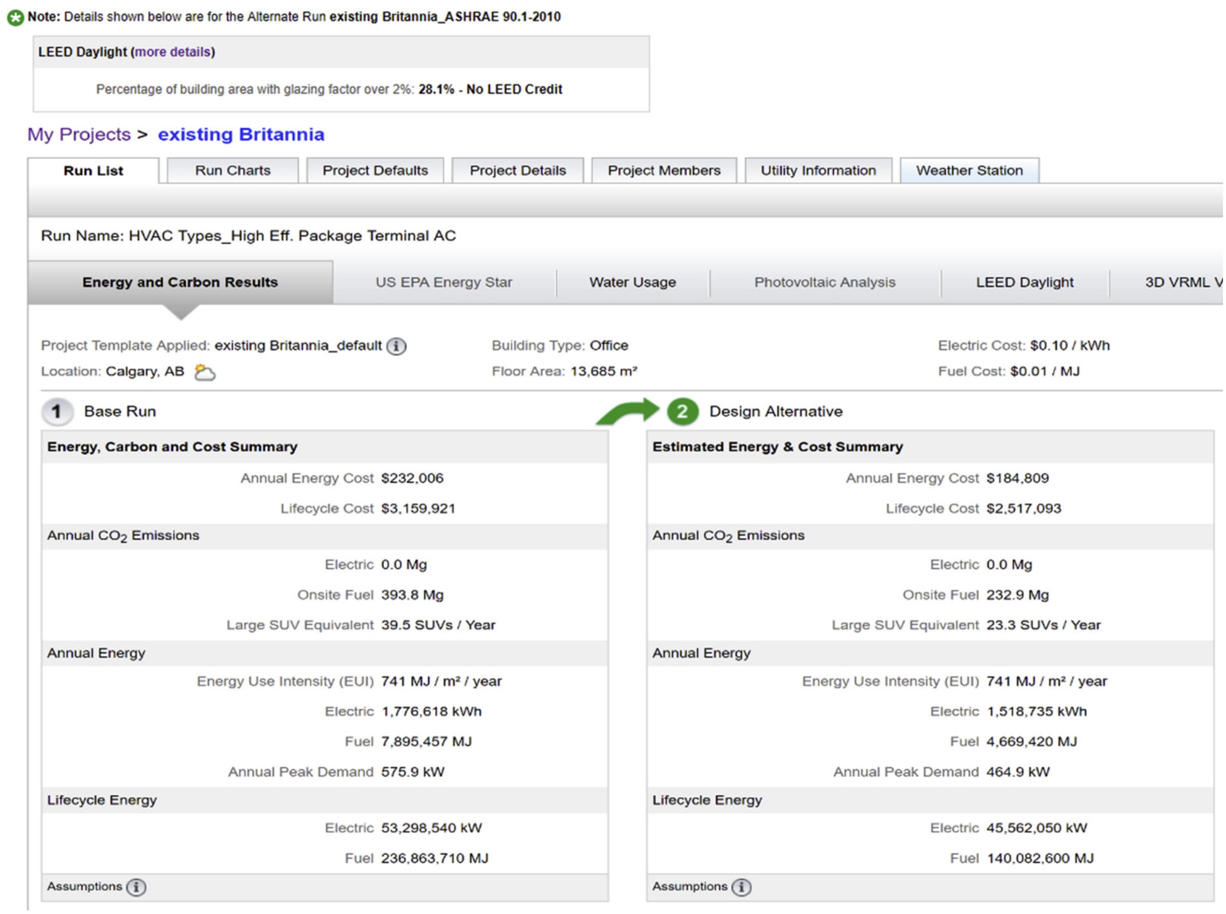

Fig. 8. Different sections for energy analysis and glazing factor.

Summary of LEED® Water Efficiency (WE) Credit Requirements

WE Credit	Description	LEED® Points	Requirement*	Your Estimated Reduction	Applicable Measures***
1.1	Water Efficient Landscaping - Reduce by 50%	1	50% reduction in potable water usage for irrigation	85%	EIRR, GREYW, NATV, NPUW, RAIN
1.2	Water Efficient Landscaping - No Potable Water Use or No Irrigation	2	100% reduction in potable water usage for irrigation	85%	EIRR, GREYW, NATV, NPUW, RAIN
2	Innovative Wastewater Technologies	1**	50% reduction in water to sewer	57%	GREYW, NPUW, RAIN, WCF
3.1	Water Use Reduction - 20%	1	20% reduction in indoor water use	54%	GREYW, RAIN, WCF
3.2	Water Use Reduction - 30%	2**	30% reduction in indoor water use	54%	GREYW, RAIN, WCF
	Your Potential LEED® WE Points:	5			LEED® NC Rating Scale Certified 26-32 points Silver 33-38 points Gold 39-51 points Platinum 52-69 points

Back

Fig. 9. Water Efficiency (WE) credit requirements for the case study.

factor of 2 % in a minimum of 75 % of all regularly occupied areas. The WE analysis would be done by GBS through using its water usage estimator function. Assuming the numbers of toilets, urinals, sinks, showers, cloth washers, dish washers and cooling towers for indoor water factors as well as irrigated areas, timed sprinklers, pool and other fixtures as an outdoor factor, the potential LEED points would be calculated by the tool. Fig. 9 shows the results for WE credit for the case building.

After all the required information are collected by the plug-in, it would be possible to run comparison with the LEED terms in order to identify the missing credits. By normalizing the credit points of the 76 LEED certified projects and entering the calculated credits of the case study in MATLAB, the tool evaluates the indices for SS and IEQ based on different K factor up until reaching to a converging number for each credit. The convergence is witnessed in K = 8, which gives us the 0.3 for SS and 0.5625 for IEQ, as shown in Fig. 10. By multiplying these numbrs to the maximum capacity SS and IEQ (i.e. 10 and 16), the 3 points for SS as well as the 9 points for IEQ are taken from MATLAB and it is reflected in the final report given by the plug-in for the case project. In the reports tab, by pressing the "calculate LEED potential" button, a table (Fig. 11) is displayed, which shows the potential accumulated credits in different categories as well as the total earned points. The detailed explanation related to the assessment of credits of each category and the obtained points related to the case study of this research is provided in Table 4. According to the earned points from the studied project in this research, it is managed to obtain 57 points out of 94 assessed points. In a linear transition to the 110 points scale, the final earned points could be reached to 66 points. At this point, this project

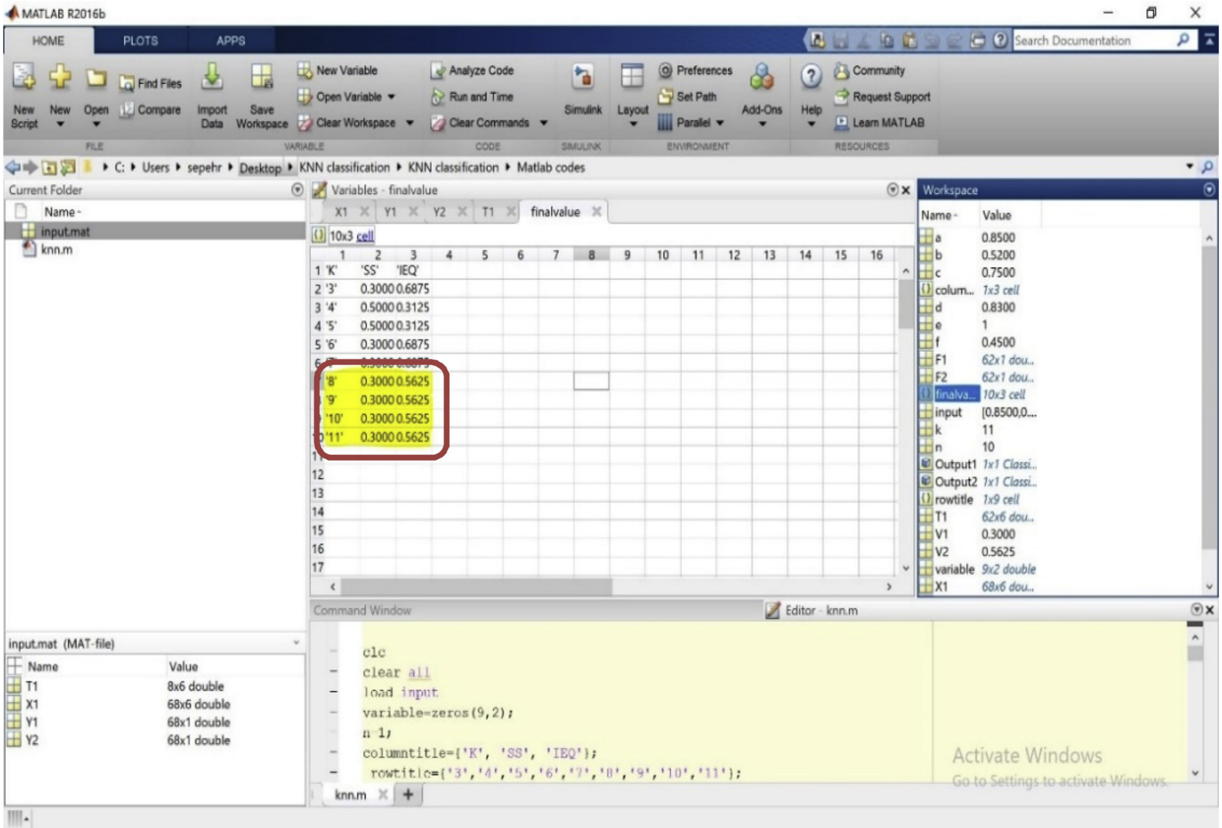

Fig. 10. The snapshot on MATLAB to calculate the missing indices (SS and IEQ) based on the calculated points and the 76 studied certified projects.

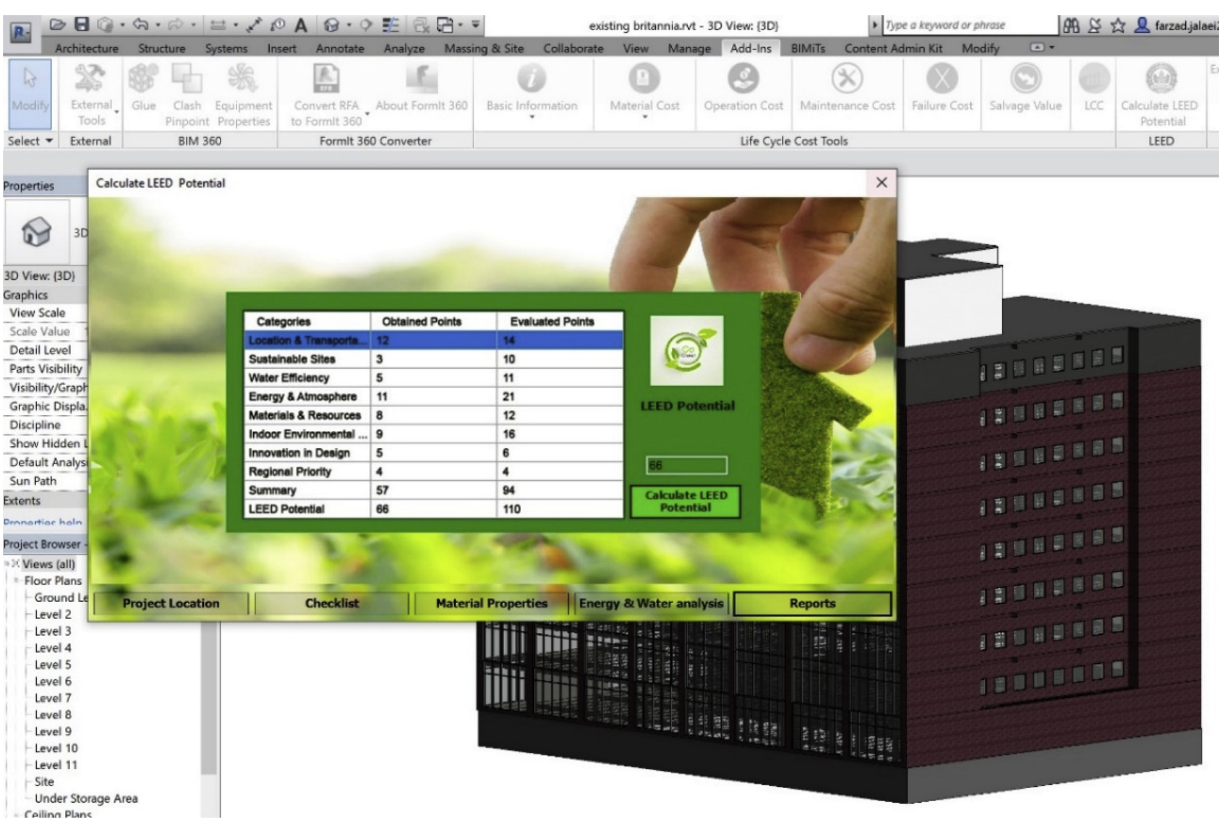

Fig. 11. The calculated LEED report from the plug-in that the project can potentially earn.

Table 4
Detailed information on the studied project.

Categories	Credits	Options	Analysis method	Possible evaluated points	Possible earned points
Location and Transporation (LT)	Sensitive land protection	Option 1	Project location	1	–
		Option 2	Checklist questions		
	Surrounding density and diverse uses	Option 1	Project location	5	5
		Option 2			
	Access to quality transit	–	Project location	5	5
	Bicycle facilities	–	Checklist questions	1	1
	Reduced parking footprint	–	Checklist questions	1	1
	Green vehicles	Option 1	Checklist questions	1	–
		Option 2	Checklist questions		
	Total			14/16	12/14
Sustainable Sites (SS)	Site assessment	Project location & checklist questions		1	1
	Heat island reduction	Option 2	Checklist questions	1	1
	Total			2/10	2/2
	KNN algorithm implementation			–	3/10
Water Efficiency (WE)	Outdoor Water Use Reduction	–	GBS Water Efficiency (WE) Analysis	–	2
	Indoor Water Use Reduction	–	GBS Water Efficiency (WE) Analysis	3	6
	Cooling Tower Water Use	–	GBS Water Efficiency (WE) Analysis	2	2
	Water Metering	–	GBS Water Efficiency (WE) Analysis	–	1
	Total			11/11	5/11
Energy and Atmosphere (EA)	Minimum energy performance	–	Energy analysis	prerequisite	prerequisite
	Optimize energy performance	–	Energy analysis	18	9
	Renewable energy production	–	Energy analysis	3	2
	Total			21/33	11/21
Materials and Resources (MR)	Construction and demolition waste management planning	–	Material properties	prerequisite	prerequisite
	Building life-cycle impact reduction	Option 3	Material properties	4	2
	Building production disclosure and optimization-environmental product declarations	Option 1 Option 2	Material properties	2	1
	Building production disclosure and optimization-sourcing of raw materials	Option 1 Option 2	Material properties	2	1
	Building production disclosure and optimization-material ingredients	Option 1 Option 2 Option 3	Material properties	2	2
	Construction and demolition waste management	Option 1 Option 2	Material properties	2	2
	Total			12/13	8/12
Indoor Environmental Quality (EQ)	Environmental tobacco smoke control	–	Checklist questions	prerequisite	prerequisite
	Daylight	Option 2	Energy analysis	–	–
	Total			-/16	-/2
	KNN algorithm implementation			–	9/16
Innovation (IN)	Innovation	–	Checklist questions & material properties	5	4
	LEED accredited professional		Checklist questions	1	1
	Total			6/6	5/6
Regional Priority (RP)	Regional priority	–	Project location	4	4
	Total			4/4	4/4
Total possible calculated points				94 /110	57/94

determines the LEED certification level to pursue Gold.

5. Conclusion

This study evaluates the implementation of BIM by representing a computer model that automates the process to identify the required number of points based on selected LEED certification categories, accumulates the total selected points as well as suggests the qualified certification. The developed model in green building practice reduces one of the biggest barriers for going green by eliminating the documentation process. It is the first model that deals with almost the whole LEED v4 categories, associated certification levels, saves users' time, and reduces users' effort. Also, it is a simple point selection method and easy to use. There is no need for users to document and record all points, just the expected points that they may achieve. Furthermore, the developed model will improve the documentation process. The model concentrates more on predicting and calculating the other LEED points that could not normally be calculated from the design by using data mining method. Overall, this research demonstrated that BIM and LEED integration was feasible with considerable constraints. In this project, 6 credits out of 8 credits (i.e. LT, WE, EA, MR, IEQ, IN, RP) are calculated by the proposed plug-in through collecting answers from BIM model, google map information for the project location and orientation, the embedded checklist questions in the plug-in to collect answers from user and the energy, lighting and water efficiency analysis results coming from GBS. Users are able to either evaluate each energy model alternative that is automatically generates by GBS and pick the optimum option as a proposed design or they could even apply their preferred changes in the design (i.e. conduct changes on project orientation, building materials and elements or the mechanical and electrical and HVAC system) and view the results of their customized design alternative. The potential LEED points for WE as well as the glazing factor that could be used for LEED daylight at IEQ would be calculated at this point. This information would be transferred by user to the plug-in manually either by entering energy performance improvement percentage or monthly energy consumption numbers as well as the numbers for WE accumulated points and glazing factor. MR would be directly calculated from BIM QTO. LT would be calculated by connecting the developed plug-in to the google map API, which shows the city facilities in the defined distance of the project location as well as collecting the answers from user from the checklist. Innovation in design points are coming from the checklist as well. RP credit points are also come from the project location information.

Although there is a potential to calculate partial points related to the SS and IEQ credits, but since the aim of the proposed plug-in is to predict and estimate the maximum LEED points that a building project could potentially earn at the conceptual design stage, there is a need to propose a solution to rescale the estimated points to suggest the most complete estimation. The KNN method is an innovative way, which is used in the backend of the plug-in to estimate the SS and IEQ points from the studied LEED certified projects that is collected from USGBC webpage. We could almost calculate 94 points using KNN as well as BIM-based solutions for the design at the conceptual stage. The plug-in would also suggest the rescaled number out of 110 by applying a linear rescaling factor as well to suggest the maximum points the project could potentially earn (i.e. for the case study, 57 out of 94 is calculated that would be over 66 in the scale of 110 points and it is suggested to register the complete design for LEED Gold certificate). The constraint in this integrated model could also be on the number of LEED certified projects that are used in this study. As much as the data for more projects become available, the proposed model could come up with more accurate results. Another limitation is in automation of energy and water efficiency results between GBS and Revit since the CSV exported file format of GBS does not include the glazing factor, monthly energy consumption and the WE earned LEED points. Thus, the transfer of information from the GBS to the plugin should be manually done by user. To achieve well balanced building performance, BIM development needs reinforced efforts at the full spectrum of issues in building design and construction.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

References

Ahvenniemi, H., Huovila, A., Pinto-Seppä, I., & Airaksinen, M. (2017). What are the differences between sustainable and smart cities? *Cities, 60*, 234–245. https://doi.org/10.1016/j.cities.2016.09.009.

Ali, M. M. (2008). The role of tall buildings in sustainable cities. *WIT Transactions on Ecology and the Environment, 117*, 345–354. https://doi.org/10.2495/SC080331.

Alwan, Z., Greenwood, D., & Gledson, B. (2015). Rapid LEED evaluation performed with BIM based sustainability analysis on a virtual construction project. *Construction Innovation, 15*(2), 134–150. https://doi.org/10.1108/CI-01-2014-0002.

Ansah, M. K., Chen, X., Yang, H., Lu, L., & Lam, P. T. I. (2019). A review and outlook for integrated BIM application in green building assessment. *Sustainable Cities and Society, 48*(July), https://doi.org/10.1016/j.scs.2019.101576.

Azhar, S., Brown, J., & Farooqui, R. (2009). BIM-based sustainability analysis: An evaluation of building performance analysis software. *Proceedings of the 45th ASC annual conference,* 1–4.

Azhar, S., Carlton, W. A., Olsen, D., & Ahmad, I. (2011). Building information modeling for sustainable design and LEED® rating analysis. *Automation in Construction, 20*(2), 217–224. https://doi.org/10.1016/j.autcon.2010.09.019.

Baldwin, R., Yates, A., Howard, N., & Rao, S. (1998). *BREEAM (Building Research Establishment Environmental Assessment Method) 98 for offices*.

Barnes, S., & Castro-Lacouture, D. (2009). BIM-enabled integrated optimization tool for leed decisions. *Proceedings of the 2009 ASCE international workshop on computing in civil engineering, Vol. 346*, 258–268. https://doi.org/10.1061/41052(346)26.

Basbagill, J., Flager, F., Lepech, M., & Fischer, M. (2013). Application of life-cycle assessment to early stage building design for reduced embodied environmental impacts. *Building and Environment, 60*, 81–92. https://doi.org/10.1016/j.buildenv.2012.11.009.

Bhattacharya, G., Ghosh, K., & Chowdhury, A. S. (2017). Granger causality driven AHP for feature weighted kNN. *Pattern Recognition, 66*, 425–436. https://doi.org/10.1016/j.patcog.2017.01.018.

Canada Green Building Council, CAGBC (2011). *Why is LEED?* Available at: https://www.cagbc.org/CAGBC/LEED/Why_LEED/CAGBC/Programs/LEED/_LEED.aspx?hkey=5d7f0f3e-0dc3-4ede-b768-021835c8ff92/ (Accessed 10 February 2018).

Carvalho, J. P., Bragança, L., & Mateus, R. (2019). Optimising building sustainability assessment using BIM. *Automation in Construction, 102*, 170–182. https://doi.org/10.1016/j.autcon.2019.02.021.

CASBEE (Comprehensive Assessment System for Building Environmental Efficiency) (2008). *CASBEE website.* Available at: http://www.ibec.or.jp/CASBEE/english/ (Accessed 2 May 2018).

Castro-Lacouture, D., Sefair, J. A., Flórez, L., & Medaglia, A. L. (2009). Optimization model for the selection of materials using the leed green building rating system. *Building a Sustainable Future Proceedings of the 2009 Construction Research Congress, 44*(6), 608–617. https://doi.org/10.1061/41020(339)62.

Chen, P. H., & Nguyen, T. C. (2017). Integrating web map service and building information modeling for location and transportation analysis in green building certification process. *Automation in Construction, 77*, 52–66. https://doi.org/10.1016/j.autcon.2017.01.014.

Chong, H. Y., Lee, C. Y., & Wang, X. (2017). A mixed review of the adoption of Building Information Modelling (BIM) for sustainability. *Journal of Cleaner Production, 142*, 4114–4126. https://doi.org/10.1016/j.jclepro.2016.09.222.

Cidell, J., & Cope, M. A. (2014). Factors explaining the adoption and impact of LEED-based green building policies at the municipal level. *Journal of Environmental Planning and Management, 57*(12), 1763–1781. https://doi.org/10.1080/09640568.2013.835714.

Clevenger, C. (2008). *Leadership in energy and environmental design (LEED). PE, RA, CEE 115/215.* Available at: https://seedconsortium.pbworks.com/w/file/fetch/44727508/LEED.pdf/ (Accessed 12 April 2018).

Cole, R. J., Rousseau, D., & Theaker, G. T. (1993). *Building environmental performance assessment criteria (BEPAC) foundation.* Vancouver, Canada.

Cover, T. M., & Hart, P. E. (1967). Nearest neighbor pattern classification. *IEEE Transactions on Information Theory, 13*(1), 21–27. https://doi.org/10.1109/TIT.1967.1053964.

Ding, G. K. C. (2008). Sustainable construction—The role of environmental assessment tools. *Journal of Environmental Management, 86*(3), 451–464. https://doi.org/10.1016/j.jenvman.2006.12.025.

Dudani, S. A. (1976). The distance-weighted k-nearest neighbor rule. *IEEE Transactions on Systems, Man, and Cybernetics, 6*, 325–327.

Dzambazova, T., Demchak, G., & Krygiel, E. (2008). *Mastering revit architecture 2009.* Indianapolis: Wiley Publishing.

Eastman, C., Teicholz, P., Sacks, R., & Liston, K. (2008). *BIM handbook: A guide to building information modeling for owners, managers, designers, engineers and contractors.* NY: John Wiley and Sons.

Faisal, S., & Tutz, G. (2017). *Nearest neighbor imputation for categorical data by weighting of attributes.* Retrieved fromhttp://arxiv.org/abs/1710.01011.

Fan, G. F., Guo, Y. H., Zheng, J. M., & Hong, W. C. (2019). Application of the weighted k-nearest neighbor algorithm for short-term load forecasting. *Energies, 12*(5), https://doi.org/10.3390/en12050916.

Fukunaga, K. (1990). Statistical pattern statistical pattern recognition. *Pattern Recognition, 22*(7), 833–834. https://doi.org/10.1016/0098-3004(96)00017-9.

Gandhi, S., & Jupp, J. (2014). BIM and Australian green star building certification. *Computing in civil and building engineering - proceedings of the 2014 international conference on computing in civil and building engineering,* 275–282. https://doi.org/10.1061/9780784413616.035.

GBCA (Green Building Council of Australia). Green star (2008). GBCA website. Available at: https://new.gbca.org.au/green-star/. (Accessed 2 May 2018).

Gou, J., Du, L., Zhang, Y., & Xiong, T. (2012). A new distance-weighted k-nearest neighbor classifier. *Journal of Information and Computational Science, 9*(6), 1429–1436.

Hiyama, K., Kato, S., Kubota, M., & Zhang, J. (2014). A new method for reusing building information models of past projects to optimize the default configuration for performance simulations. *Energy and Buildings, 73*, 83–91. https://doi.org/10.1016/j.enbuild.2014.01.025.

Hoff, J. L. (2008). Life cycle assessment and the LEED green building rating systems. *Proceedings of the RCI 23rd international convention* Available at: rci-online.org/wp-content/uploads/2008-cts-proceedings-hoff.pdf/ (Accessed 26 April 2018).

Huo, X., Yu, A. T. W., & Wu, Z. (2017). A comparative analysis of site planning and design among green building rating tools. *Journal of Cleaner Production, 147*, 352–359. https://doi.org/10.1016/j.jclepro.2017.01.099 PII: S0959-6526(17)30113-0.

Jalaei, F., & Jrade, A. (2014). An automated BIM model to conceptually design, analyze, simulate, and assess sustainable building projects. *Journal of Construction Engineering, 2014*, 1–21. https://doi.org/10.1155/2014/672896.

Jalaei, F., & Jrade, A. (2015). Integrating building information modeling (BIM) and LEED system at the conceptual design stage of sustainable buildings. *Sustainable Cities and Society, 18*, 95–107. https://doi.org/10.1016/j.scs.2015.06.007.

Jrade, A., & Jalaei, F. (2013). Integrating building information modelling with sustainability to design building projects at the conceptual stage. *Building Simulation, 6*(4), 429–444. https://doi.org/10.1007/s12273-013-0120-0.

LEED v4 for Building Design and Construction (2018). *LEED v4 for building design and construction.* Available at: https://www.usgbc.org/resources/leed-v4-building-design-and-construction-current-version (Accessed 27 July 2018).

Little, R. J., & Rubin, D. B. (2002). *Statistical analysis with missing data* (second edition). New York: John Wiley and Sons.

Lu, Y., Wu, Z., Chang, R., & Li, Y. (2017). Building Information Modeling (BIM) for green buildings: A critical review and future directions. *Automation in Construction, 83*, 134–148. https://doi.org/10.1016/j.autcon.2017.08.024.

Luo, Y., & Wu, W. (2015). Sustainable design with BIM facilitation in project-based learning. *Procedia Engineering, 118*, 819–826. https://doi.org/10.1016/j.proeng.2015.08.519.
Mah, D., Manrique, J. D., Yu, H., Al-Hussein, M., & Nasseri, R. (2011). House construction CODN2/DN footprint quantification: A BIM approach. *Construction Innovation, 11*(2), 161–178. https://doi.org/10.1108/14714171111124149.
Malarvizhi, R., & Thanamani, A. S. (2012). K-nearest neighbor in missing data imputation. *International Journal of Engineering Research and Development, 5*(1), 05–07.
Mundfrom, D., & Whitcomb, A. (1998). *Imputing missing values: The effect on the accuracy of classification. Multiple linear regression viewpoints*, *Vol. 25*. Retrieved from http://mlrv.ua.edu/1998/VOL25_N1_A4.pdf.
Nguyen, T. H., Toroghi, S. H., & Jacobs, F. (2016). Automated green building rating system for building designs. *Journal of Architectural Engineering, 22*(4), https://doi.org/10.1061/(ASCE)AE.1943-5568.0000168.
Nizam, R. S., Zhang, C., & Tian, L. (2018). A BIM based tool for assessing embodied energy for buildings. *Energy and Buildings, 170*, 1–14. https://doi.org/10.1016/j.enbuild.2018.03.067.
Pulselli, R. M., Simoncini, E., Pulselli, F. M., & Bastianoni, S. (2007). Emergy analysis of building manufacturing, maintenance and use: Em-building indices to evaluate housing sustainability. *Energy and Buildings, 39*(5), 620–628. https://doi.org/10.1016/j.enbuild.2006.10.004.
Saieg, P., Sotelino, E. D., Nascimento, D., & Caiado, R. G. G. (2018). Interactions of building information modeling, lean and sustainability on the architectural, engineering and construction industry: A systematic review. *Journal of Cleaner Production, 174*, 788–806. https://doi.org/10.1016/j.jclepro.2017.11.030.
Shoubi, M. V., Shoubi, M. V., Bagchi, A., & Barough, A. S. (2015). Reducing the operational energy demand in buildings using building information modeling tools and sustainability approaches. *Ain Shams Engineering Journal, 6*(1), 41–55. https://doi.org/10.1016/j.asej.2014.09.006.
Todeschini, R. (1990). Weighted k-nearest neighbour method for the calculation of missing values. *Chemometrics and Intelligent Laboratory Systems, 9*(2), 201–205. https://doi.org/10.1016/0169-7439(90)80098-Q.
USGBC (United States Green Building Council) (1999). *LEED – Leadership in energy and environmental design: Green building rating system, version 1.0.* US Green Building Council.
USGBC (United States Green Building Council) (2005). *LEED for new construction & major renovations (Reference guide version 2.2). Washington, D.C.* Available at: http://www.usgbc.org/ (Accessed 5 March 2018).
USGBC (United States Green Building Council) (2018). *LEED v4 for building design and construction.* Available at: http://www.usgbc.org/ (Accessed 19 May 2018).
Wang, J., Lee, K., Arch, M., & Park, I. (2014). *G-SEED: The revised Korean green building certification system* (December)16–18.
Wang, L., Qian, B., Cui, X., & Chao, L. (2018). *Application of BIM technology in interference check in construction work.* https://doi.org/10.2991/iceep-18.2018.20.
WBDG (Whole Building Design Guide) (2008). *WBDG website.* Available at: http://www.wbdg.org/ (Accessed 2 May 2018).
White, J. W., Calm, J. M., Schwedler, M., Marriot, C. E., Anis, W. A. Y., Arbore, A., ... Baxter, V. D. (2007). *ASHRAE standard 90.1-2007.* Atlanta, GA, USA: American Society of Heating. Refrigerating and Air-Conditioning Engineers, Inc.
Wong, J. K. W., & Kuan, K. L. (2014). Implementing "BEAM Plus" for BIM-based sustainability analysis. *Automation in Construction, 44*, 163–175. https://doi.org/10.1016/j.autcon.2014.04.003.
Wu, P., Song, Y., Shou, W., Chi, H., Chong, H. Y., & Sutrisna, M. (2017). A comprehensive analysis of the credits obtained by LEED 2009 certified green buildings. *Renewable and Sustainable Energy Reviews, 68*, 370–379. https://doi.org/10.1016/j.rser.2016.10.007.
Wu, X., Kumar, V., Ross, Q. J., Ghosh, J., Yang, Q., Motoda, H., ... Steinberg, D. (2008). Top 10 algorithms in data mining. *Knowledge and Information Systems, 14*(1), 1–37. https://doi.org/10.1007/s10115-007-0114-2.
Zavrel, J. (1997). An empirical re-examination of weighted voting for {k-{\sc nn}}. *Proceedings of the 7th Belgian-Dutch conference on machine learning,* 139–148. Retrieved from http://citeseerx.ist.psu.edu/viewdoc/summary?doi=10.1.1.36.573.