Cement and Concrete Research 109 (2018) 42–53

Contents lists available at ScienceDirect

Cement and Concrete Research

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

Computational design optimization of concrete mixtures: A review

M.A. DeRousseau, J.R. Kasprzyk, W.V. Srubar III*

Department of Civil, Environmental, and Architectural Engineering, University of Colorado Boulder, 1111 Engineering Drive, ECOT 441 UCB 428, Boulder, CO 80309-0428, USA

ARTICLE INFO

Keywords:
Cement
Concrete
Mixture proportioning
Optimization

ABSTRACT

A comprehensive review of optimization research concerning the design and proportioning of concrete mixtures is presented herein. Mixture design optimization is motivated by an ever-increasing need for designers and decision-makers to proportion concrete mixtures that satisfy multiple – oftentimes competing – performance requirements, including cost, workability, mechanical properties, durability, and environmental sustainability. In this review, we first discuss common mathematical problem formulations, decisions, objectives, and constraints pertaining to concrete mixture design optimization. Subsequently, we examine the types of models employed to approximate properties of concrete, which include a variety of linear combination, statistical, machine learning, and physics-based models that are required to optimize the proportions of a mixture. We then review and discuss computational methods used to optimize concrete mixtures in the context of surveyed literature. Finally, we highlight and discuss current trends and opportunities for advancing the field of concrete mixture design optimization in context of the current state of the art.

1. Introduction

Global consumption of ordinary portland cement (OPC) concrete, the most commonly used construction material in the world, has reached approximately 10 billion metric tons per year [1]. Its unique combination of strength, economic viability, availability of raw material resources, moldability, and durability make OPC concrete an ideal candidate for a wide variety of civil infrastructure applications. In addition, by varying the type and quantity of individual constituents in the concrete mixture (*e.g.*, cement, water, aggregate, admixtures), the fresh- and hardened-state properties of OPC concrete can be tailored to meet many different design specifications.

Concrete mixture design, also known as mixture proportioning, is the process of selecting the type and quantity of individual constituents to yield a concrete that meets specifiable characteristics for a particular application. In general, traditional approaches for proportioning concrete mixtures can be classified into two main methods: prescriptive and performance-based.

Prescriptive approaches are step-by-step design methodologies that, when followed, help the designer proportion an acceptable concrete mixture. Prescriptive proportioning methods have evolved from arbitrary 1-2-3 cement-sand-aggregate volumetric ratio methods established in the early 1900s [2] to the present-day absolute volume method (AVM) prescribed by the American Concrete Institute (ACI) [3] and Portland Cement Association [4]. Given a target compressive strength, slump (for workability), and air content (for freeze-thaw durability), the PCA methodology for designing and proportioning concrete mixtures guides the designer in selecting an appropriate water-to-cementitious materials ratio (w/c), air content, admixture dosage, and both fine and coarse aggregate content. Other prescriptive-based approaches include both the old and new Bureau of Indian Standards [5]. A primary advantage of prescriptive proportioning methods is that the mixture proportioning is directed by the method itself; the decision-maker need not make subjective design decisions. While these methods are most effective for large-volume, general construction applications, the lack of flexibility for a designer to tailor and tweak individual mixture proportions is a notable limitation of the method.

In contrast to prescriptive proportioning methods, performance-based mixture design methodologies impose no strict guidelines on the amounts and ratios of constituents. Rather, this approach allows the designer substantial leeway to meet design specifications by proportioning mixtures directly from laboratory trial batches (a trial-and-error, iterative approach) rather than the linear, non-iterative AVM. For example, if the structural design specification requires a compressive strength of 30 MPa, the designer can select any amount of cementitious material, water, and aggregate and prove, through trial-batch testing, that the mixtures sufficiently achieve the strength requirement. Fig. 1a demonstrates the process of traditional mixture design, where either prescriptive or performance-based design methods are used to decide upon mixture proportions; the output is one acceptable, but oftentimes non-optimal, design solution.

* Corresponding author.
E-mail address: wsrubar@colorado.edu (W.V. Srubar).

https://doi.org/10.1016/j.cemconres.2018.04.007
Received 10 December 2017; Received in revised form 5 April 2018; Accepted 8 April 2018
Available online 18 April 2018

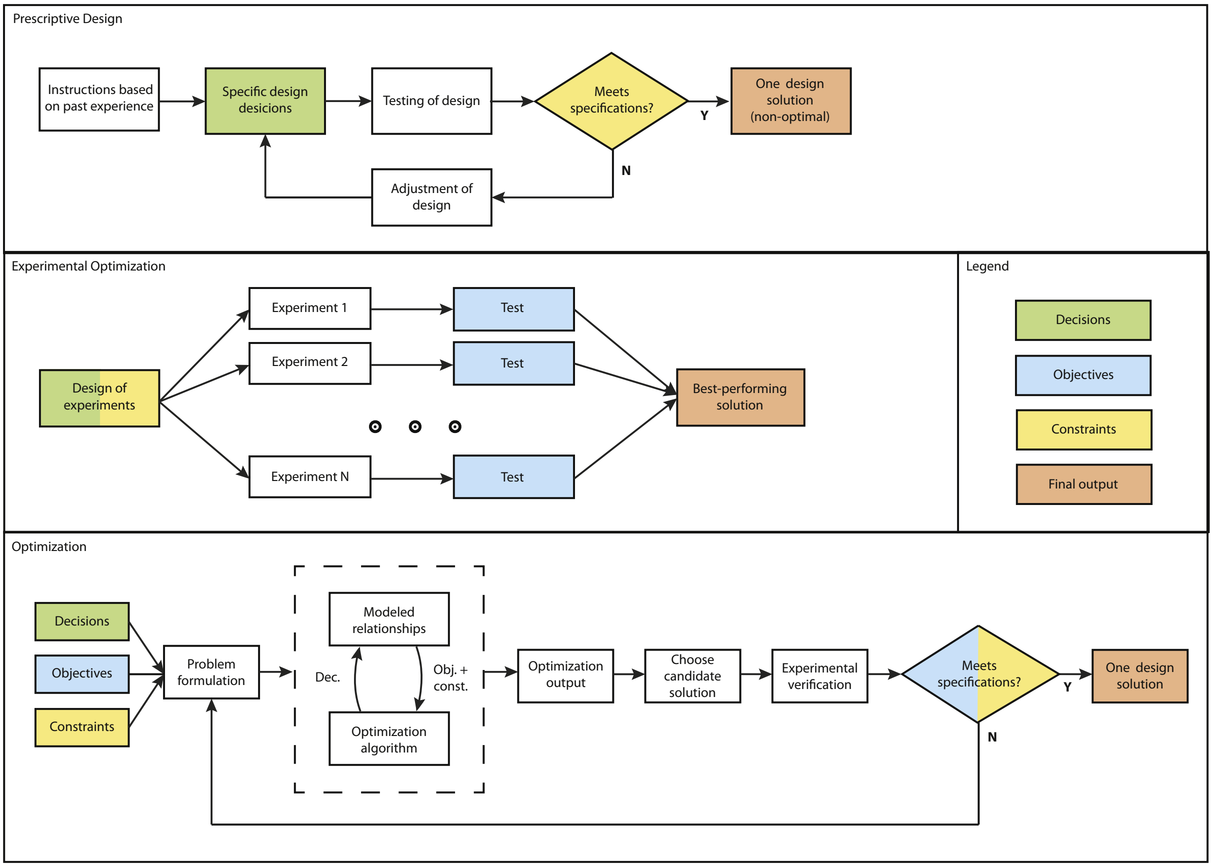

Fig. 1. (a) Traditional design; (b) experimental optimization; (c) computational optimization.

1.1. Experimental design optimization of concrete mixtures

Given the flexibility of performance-based approaches and a desire to achieve the most economical mixture design solutions that meet performance specifications, many research studies have attempted experimental optimization of concrete mixtures. The general process of experimental optimization is visualized in Fig. 1b. Soudki et al. [6], for example, aimed to experimentally maximize the compressive strength of concrete mixtures by varying the water-to-cement (w/c) ratio, coarse aggregate-to-total-aggregate ratio, total aggregate-to-cement ratio, and curing temperature. Similar studies targeted the design of concrete with experimentally maximized flexural strength [7], water absorption [8], and consistency index (a measure of workability) [8]. Despite being useful in their intent, experimental design optimization suffers from exponential increases in the required number of samples and experiments when many mixture parameters or values of those parameters are considered as variables in the optimization. As a result, detailed experimental optimization of concrete mixtures can be both time- and resource-intensive. In addition, the generalizability of the results obtained from experimental optimization is limited due to nuanced differences in concrete performance introduced by spatiotemporal environmental variability (*i.e.*, temperature, humidity) and specific constituent characteristics, such as the type and chemistry of cementitious materials and the size, shape, and texture of aggregates.

While both prescriptive and performance-based approaches yield acceptable design solutions, these methodologies do not result in truly best-performing solutions, but rather well-performing proportions of concrete mixtures. Furthermore, both approaches require a lengthy and iterative design process with only one acceptable mixture design solution. To circumvent the experimental limitations of these methodologies, a significant body of research has recently focused on formulating and validating computational design optimization approaches and tools for concrete mixture proportioning that leverage the wealth of experimental data concerning OPC concrete, advanced mathematical techniques, and the power of high-performance computing.

1.2. Computational design optimization of concrete mixtures

Computational design optimization of concrete mixtures is a mathematical—as opposed to experimental—approach to mixture proportioning. Fig. 1c illustrates that computational optimization of concrete mixtures is a process whereby an optimal design solution can be found. In computational design optimization, the decision-maker must decide upon the problem formulation, the modeled relationships, and the optimization algorithm that should be employed. The problem formulation involves defining the decision variables, objectives, and constraints of the problem. Modeling involves choosing appropriate mathematical relationships that model each objective as a function of the decision variable. An optimization algorithm is typically chosen based on its appropriateness to mathematically solve the problem.

1.3. Scope of the review

This review fully expounds on each of the three steps of computational design optimization of concrete mixtures, namely (1) formulating, (2) modeling, and (3) solving the concrete mixture design optimization problem. Each step, explicated in a discrete section of this review, is discussed in the context of examples from the most salient and state-of-the-art literature. In addition, this review provides a critical synopsis and discussion of research and development needs that are required to advance the field of concrete mixture design optimization.

Table 1
Example decisions in concrete mixture design optimization problems.

Possible decision variable	Classification of decision variable
Cement type	Discrete
Supplementary cementitious material (SCM) type	Discrete
Coarse aggregate size	Discrete
Admixture type	Discrete
Cement content (or w/c ratio)	Continuous
SCM content	Continuous
Coarse aggregate content	Continuous
Fine aggregate content	Continuous
Admixture content	Continuous
Curing temperature	Continuous

2. Optimization problem formulation

Problem formulation, the first step of optimization, is the process of defining the *decisions*, *objectives*, and *constraints* of the optimization problem.

2.1. Decisions

Decisions, or decision levers, are independent variables of the optimization problem over which the decision-maker typically has control [9]. In other words, decisions are variables that can be altered to optimize a problem. In terms of classification, decision variables can be either discrete or continuous. Table 1 provides examples of possible decisions for concrete mixture problems. For instance, in concrete mixture optimization problems, discrete decisions could be including or not including a specific type of supplementary cementitious materials (*e.g.* fly ash) or any other mixture constituent. Continuous decisions might include the amount of a particular mixture constituent (*e.g.* mass of coarse aggregate per unit volume), a ratio of two individual components (*e.g.*, w/c ratio), or concrete curing conditions (*e.g.* ambient temperature, humidity). The multitude of decisions that could be included in a concrete mixture optimization problem are discussed in Mamlouk and Zaniewski [10].

2.2. Objectives

An *objective* is a quantification of a desired outcome for the problem that one seeks to optimize (*i.e.*, minimize or maximize). In optimization problems, objectives are modeled as functions of the decision variables (*e.g.*, the cost of a mixture is a function of the types and amounts of the mixture constituents). Furthermore, optimization problems can be formulated as single- or multi-objective, which determines appropriate solution methods. In the case of Cheng et al. [11], Meng and Valipour [12], and Ji et al. [13], single-objective problems were solved because the goal of the study was to simply optimize one objective, such as cost or compressive strength.

Multi-objective problems have been considered in more recent literature, where researchers elucidate and study tradeoffs between competing objectives. These objectives often include physical performance, cost, and environmental impacts. Unlike single-objective problems, multi-objective problems have multiple potential solutions that represent different compromises between objectives. The decision-maker is the one who must ultimately decide which of the potential solutions (and, therefore, which level of compromise among the objectives) is the most desirable for that particular problem.

2.3. Constraints

Constraints are mathematical functions of the decision variables, such that when any constraint is violated, the solution is infeasible and is discarded from the solution set. Therefore, a constraint could impose a limit on a decision variable or objective, or as a separate function of decision variable values that provides some additional restriction on an acceptable solution. As an example, a restriction on the coarse aggregate content would be a limit on a decision variable; a restriction on the workability of the hardened concrete would be classified as a constraint on an objective output. Furthermore, if both water and cement were independent decision variables, a minimum w/c ratio, a function of two decision variables, could also be defined as a constraint.

Table 2
Examples of possible constraints that could be imposed on a concrete mixture design optimization problem.

Constraint on	Type of Constraint	Imposed by
Compressive strength	Constraint on an objective	Standard (ACI 318) [3] or Designer (specification)
Cost	Constraint on an objective	Designer (specification)
Workability	Constraint on an objective	Designer (specification)
Water to cement ratio	Constraint on a decision	Standard (ACI 318) [3]
Amount of contaminants in aggregate	Constraint on a decision	Standard (ASTM C 33) [15]
Particle size distribution of aggregate	Constraint on a decision	Designer (specification)

The way constraints are imposed also warrants discussion. A constraint can be imposed by the designer or it can be imposed by a code or standard. For instance, a designer may impose a constraint on the compressive strength of the hardened concrete. In contrast, a constraint imposed by a standard might be a limit on the w/c ratio, which is often set by ACI 318. Table 2 provides more examples of possible constraints on a concrete mixture design optimization problem. Although constraints are useful for setting acceptable bounds on a problem formulation, care should be taken with their implementation because solution sets with even small violations of the constraint will be discarded and deemed unacceptable [14].

The specific decisions, objectives, and constraints of a problem are chosen such that they represent the particular goals of the concrete mixture design problem. For instance, if an engineer is designing an OPC concrete, then the mixture decision variables (*i.e.*, constituents) will only need to include the amounts of cement, water, fine aggregate, and coarse aggregate. However, if the goal is to design a more tailored or high-performing concrete (HPC), more decision variables will likely be included, such as aggregate size, superplasticizer dosage, and the amounts of fly ash, slag, or silica fume. Similarly, the objectives of the problem depend on the goals of the individual designer. For some decision-makers, cost may be the only objective; for others, assessing tradeoffs between multiple objectives might be important. For instance, a decision-maker may be required to design both a concrete mixture with minimum cost and CO_2 emissions. As will be discussed in Section 4, the nature of the decisions, objectives, and constraints determine the type of optimization problem.

3. Modeling the objectives

Models of the objectives refer to mathematical expressions that relate the decision variables to the desired objectives of a concrete mixture design problem. Models for the objectives are essential because they compute values of the objective functions that are used to optimize the problem (Fig. 2).

As was discussed in Section 2, objectives relate decision variables to the economic, mechanical, environmental, or other properties of concrete mixes; correspondingly, there are many expressions that are employed to model these objectives. These relationships include linear combination, life cycle assessment, statistical, machine learning, and physics-based models.

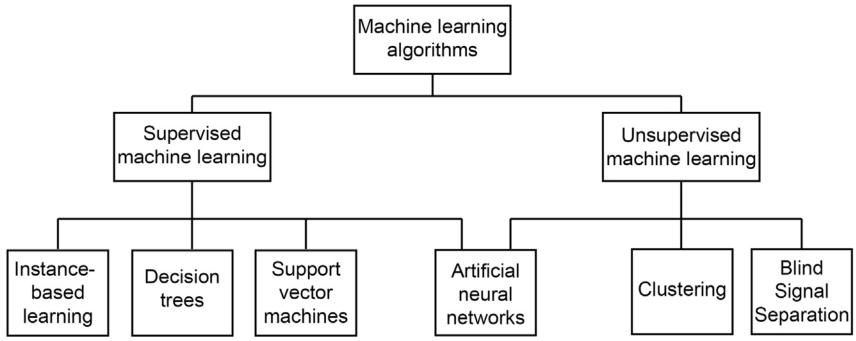

Fig. 2. Taxonomy of machine learning algorithms. Supervised machine learning techniques have been applied to concrete modeling problems. Note that artificial neural networks can be used for supervised or unsupervised machine learning.

3.1. Linear combination models

Some objectives are adequately modeled by simple linear combinations of relevant decision variables. A suitable application for these relationships is modeling the cost of a concrete mixture as a function of its individual mixture constituents. For instance, in a study conducted by the Office of Infrastructure Research and Development, the singular objective of the optimization problem was to minimize cost. A price was attributed to each mixture constituent, and the cost was modeled as a linear combination of cement, silica fume, high-range water reducing admixture (HRWRA), coarse aggregate, fine aggregate, and water [16]. Linear combination equations are the simplest and, thus, most frequently used for modeling a cost objective in concrete mixture design [17–21]. However, care should be taken when problems with different orders of magnitude are considered, because linear models may not account for economies of scale. In other words, the cost of a constituent for a small volume, for example, will not be the same for a large amount of the same constituent. Thus, appropriate cost coefficients for the size of the project being considered must be used.

3.2. Life cycle assessment models

In order to quantify environment-related objectives, life cycle assessment (LCA) models are most often used. LCA is the process of estimating environmental impacts (*e.g.*, global warming potential, water depletion, resource depletion, land transformation, eutrophication) for a particular product or process [22]. Typically, LCA software or a life cycle inventory database is used to attribute impacts to a particular material or process within the material system considered [23–25]. LCA could be considered a subset of linear combination models because the impact factors for a material or process are simply multiplied by the amount of that component in the system; however, we consider LCA models to be different from other linear combination models because of the specific and well-defined process of conducting a LCA. As an example of a LCA model, Loijos estimated the global warming potential of all major pavement types in the United States through the linear combination of factors for each stage in the pavement life cycle [26]. In this scenario, global warming potential was the objective being modeled as a linear combination of individual contributions from within the system boundary. Similar LCAs have modeled environmental impacts of specific types of concrete systems, including high strength mixtures [27], mixtures with supplementary cementitious materials (SCMs) [28,29], and mixtures with recycled aggregates [30]. Although these studies explicitly address the life cycle environmental impacts of concrete, incorporation of environmental impact objectives into concrete mixture design optimization is an emerging trend in recent literature. Only a few studies of those surveyed in this review employed objectives related to life-cycle CO_2 emissions [31,32]. As was the case for linear cost modeling, the caveat about economies of scale is also true for life cycle assessment; impact factors do not necessarily scale in a linear fashion.

3.3. Statistical models

Another approach for modeling the relationship between the decision variables and a concrete mixture objective is statistical modeling. Statistical models are typically used when data that relate the decision variables and objective have been generated, but the mathematical form the relationship should take is not known. In the literature, these models are typically used to model objectives related to the mechanical properties (*e.g.*, compressive strength) of concrete because the relationships are complex, but data on these properties can be observed or collected experimentally [33].

The equation chosen for modeling a particular objective is dependent on the property being modeled and is left to the modeler's discretion according to the complexity of the study. The simplest equation to choose for property prediction is a linear equation:

$$Y = \beta_0 + \beta_i X + \varepsilon \tag{1}$$

where Y is the vector of properties being predicted, i represents the decision variables, X is the design matrix representing the values of each decision variable, β_i is a vector of multipliers representing the influence of the decision variables on each property, and β_0 is the expected value of Y when all the X variables are equal to zero. A limitation of this approach is the assumption that the relationship between the mixture constituents and concrete objectives is linear, which can lead to poor performance metrics for the model when the relationships being modeled are complex and nonlinear [34].

Equations with greater complexity, such as quadratic or polynomial models, include higher-order terms for modeling the influence of decision variables (see Eq. (2)), where the h term represents the order of the model. These models can exhibit improved predictive performance compared to purely linear equations:

$$Y = \beta_0 + \beta_1 X + \beta_2 X^2 + \ldots + \beta_h X^h + \varepsilon \tag{2}$$

We use Akalin et al. [35] as an example. In this model, the concrete mixture properties (Y_i) that are being modeled are slump flow, compressive strength, and appearance. In addition, there are six decision variables: volume fractions of cement, water, silica fume, fly ash, natural sand, crushed sand, aggregate, and chemical admixture (*i.e.*, superplasticizer). For modeling of concrete properties, second-order polynomials have been used to predict objectives, such as compressive strength and slump. These models are attractive because they add second-order terms to model the decision variables without the computational complexity of a third-order or higher model [35–45].

It is also possible to build statistical models using equations with interactive (multiplicative) terms. For instance, Simon chooses a second-order polynomial with terms representing the interactions between the decision variables to model compressive strength [46]. These models can sometimes improve predictive performance; however, the danger of adding more terms to a statistical model is that the extra terms might only be modeling random noise in the data and not actually improving predictive performance.

For statistics-based methods, comparing the predictive performance and selecting an appropriate and robust model is essential. There are several statistical performance measures that measure the goodness-of-fit and predictive ability of the model. For instance, the coefficient of determination (R^2) is a measure of the proportion of the information in the data that is explained by the model. The value of R^2 ranges from zero to one, with higher values indicating higher predictive power of the model. For example, a R^2 value of 0.8 indicates that the model is able to explain 80% of the variation of the output. R^2 is an initial indicator of goodness of fit of a model; however, R^2 should be used with other performance measures, due to the inflation of R^2 values when the number of explanatory terms in a model increases. If a model has too many explanatory terms, it produces a misleadingly high R^2 value, which is only modeling the random noise in the data.

The root-mean-square error (RMSE) is the square root of the mean-square error and indicates the average distance of a data point from the expected value provided by the model:

$$RMSE = \sqrt{\frac{\sum_{i=1}^{n} (\widehat{y}_i - y_i)^2}{n}} \tag{3}$$

where y_i is the observed value from the data, $\widehat{y}_i$ is the predicted value from the model, and n is the number of observed values. Unlike R^2, a lower RMSE indicates a better model. RMSE is also scale-dependent, meaning it does not range from zero to one, but rather can range from zero to infinity. Therefore, RMSE can be used to compare errors of different models on one dataset, but not among multiple datasets.

The mean absolute percentage error (MAPE) is a measure of prediction accuracy of a model in which the error is expressed as a percentage:

$$MAPE = \frac{100}{n} \sum_{i=1}^{n} \left| \frac{\widehat{y}_i - y_i}{\widehat{y}_i} \right| \tag{4}$$

Like RMSE, a lower MAPE indicates a better model fit. Like R^2, MAPE should be used with other performance measures because MAPE places a heavier penalty on negative errors than positive errors and cannot be used if there are values with no error, where division by zero occurs.

A critical point to consider when selecting a “best model” from the many available statistical and machine learning models is robustness. Especially for highly complex machine learning models with many parameters, there is the possibility to generate an overfitted model and conclude from the goodness-of-fit measures that the model has better predictive performance than it does in reality. To account for this problem, it is common to randomly split datasets into so-called training and testing datasets. First, a model is fit on a majority of the training dataset; then, the performance of the model is tested and reported using the reserved testing dataset. This provides greater confidence in the predictive performance of the model. A technique called k-fold cross-validation is also used to account for overfitting. Similar to the concept used in model training and testing, k-fold cross-validation partitions a dataset into k equal-sized samples. For instance in 10-fold cross-validation, the dataset is split into ten subsamples. Ten different models are fit, each with a different tenth of the original dataset being saved for model testing. The average of the statistical performance measures for the ten models is reported. Full descriptions and best practices for these methods are thoroughly discussed in Kuhn and Johnson [47]. Because of the model validation aspects of these methods, performance measures reported for cross-validated or testing datasets are more widely accepted methods for comparing models.

3.4. Machine learning models

Although statistical methods provide explicit equations relating decision variables to the objectives of concrete mixture design problems, the performance of these equations are oftentimes insufficient to describe such complex relationships. In contrast to the aforementioned traditional statistical methods, machine learning methods do not rely upon explicit equations; instead, machine learning models are learning algorithms that find patterns in a set of training data to predict future values. These techniques are more computationally expensive than statistical techniques; however, as will be shown, researchers have increasingly applied machine learning methods in concrete mixture design optimization because of their ability to account for the complexity of concrete mixtures and their properties.

The hierarchical taxonomy of machine learning methods is shown in Fig. 2. Machine learning methods, a catch-all term for algorithms used to find patterns in data, are generally classified into two main categories: unsupervised and supervised learning. In unsupervised learning, there is simply a dataset, and it is the job of the algorithm to find relationships and patterns within it; no output variables of the data guide the algorithm into learning patterns within the data. A few of the available algorithm types include clustering, neural networks, and blind signal separation techniques. Contrastingly, in supervised learning, the algorithm is trained on a dataset of both input and output variables, and the goal is to “learn” the relationship between them. In supervised learning, the output variables are known, and the algorithm can be trained by comparing predicted output values to the actual output values and by adjusting the parameters of the algorithm to improve predictive performance. Within supervised learning algorithms, there are classification-based methods (where the output variables are discrete) and regression-based methods (where the output variables are continuous). Since the output variables (*i.e.*, objectives) of a concrete mixture are nearly always continuous, regression-based algorithms are most often used in this field. Therefore, all of the examples that follow are regression-based rather than classification methods.

Although all machine learning algorithms can be used to achieve a similar goal (*i.e.*, find patterns in data and improve predictive performance), the structure of each type of algorithm and its strategies for finding those patterns can differ significantly. Each of the following machine learning methods discussed below is applicable to modeling the relationship between the decision variables and concrete objectives no matter the type of concrete or objective that is being predicted. Below, we consider some of the most commonly used supervised learning algorithms in concrete modeling, which include artificial neural networks, instance-based learning, decision-trees and support vector machines.

3.4.1. Artificial neural networks

One group of machine learning methods, artificial neural networks (ANNs), is mathematical techniques based on the idea of interconnected layers of nodes. In these types of methods, input nodes connect to output nodes through one or more layers of intermediate, or hidden, nodes. In Fig. 3, a schematic of the ANN technique is shown with an example of possible inputs and outputs to the model. The links between the nodes are weighted in order to relate importance of input variables to the output variables. In concrete mixture design using ANNs, the input layer nodes are concrete mixture decision variables and the output layer is the objective(s). For instance, one study used ANN to model the relationship between compressive strength (the output layer) and 73 decision variables of the concrete mixture (the input layer). These decision variables were grouped into four categories: material proportions, basic information, measurement variables, and temperature and humidity history of the pour [48]. The predictive performance of the model is improved by “training” the network on a dataset that changes the weights between the nodes. The intermediate layer(s) of the network do not represent real values; rather, they are used to give weight to the interactions between variables. For a comprehensive review of mathematical techniques within the ANN paradigm, see Dreiseitl and Ohno-Machado [49].

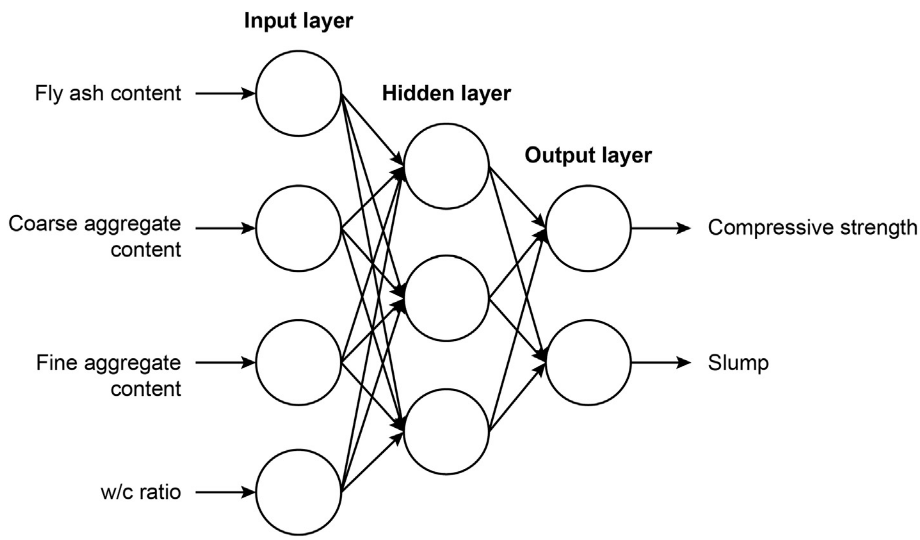

Fig. 3. Architecture of ANN. One or more hidden layers connect the output and input layer of dependent and explanatory variables, respectively.

ANN methods gained acceptance as a modeling method in the mid-1980s. The American Institute of Physics commenced their first annual meeting related to Neural Networks for Computing in 1985. The Institute of Electrical and Electronic Engineers' (IEEE's) 1st International conference on Neural Networks was held in 1987 [50]. In the field of concrete property modeling, Kasperkiewicz et al. was among the first to model compressive strength of HPC using ANNs; this study included six decision variables and obtained a R^2 of 0.757 [51]. Since then, researchers have applied ANN methods to wide-ranging problems in cement and concrete research; they have been used to model several types of concrete properties, including compressive strength, slump, filling capacity, and segregation and for many types of concrete including high performance concrete [51,52], self-consolidating concrete [53,54], ready mix concrete [55], high strength concrete [55,56], ultra-high performance concrete [57], recycled aggregate concrete [58], and structural lightweight concrete [59].

Since each type of concrete includes different mixture constituents and because each of these studies utilizes different experimental datasets, the models trained in each study are different and, thus, provide different predictive performances of the properties of the mixture. Compared to traditional statistical techniques, ANNs have the advantage of modeling the relationship between a large number of decision variables and objectives; statistical techniques can suffer from over-fitting with many decision variables and explanatory terms in the model. However, one disadvantage of ANNs is that the dataset used for training the model must be large and diverse for the model to be accurate over a wide range of decision variables and applicable to more than one type of concrete.

3.4.2. Instance-based learning

A second group of algorithms, known as instance-based learning techniques, refers to algorithms that compare a data point of input values (a "query") to some number of nearest neighbors within the dataset to determine the output variable(s) of that data point. In instance-based learning, an entire dataset is stored for retrieval and comparison to a new query. This query is classified based on existing data that it is close to, which can be determined in several ways. For instance, in metric distance minimization, the properties of the query are predicted based on the closest single data point to it [60]. The k-nearest-neighbors algorithm is similar, except that this algorithm returns an average of the properties of k nearest neighbors. Locally weighted regression is a type of instance-based learning that builds a local regression model by finding points near a query and weighting them by distance. Generally, these algorithms differ in the number of local points that are considered.

Instance-based learning concepts were first developed in the 1960s with Sebestyen (1962), Nilsson (1965), and Cover and Hart (1967) sequentially publishing studies which developed the nearest-neighbors algorithm [61–63]. And although research in predicting concrete properties has been dominated by ANN algorithms, a few studies have used instance-based learning in order to explore the utility of these less-used, computationally simpler machine learning algorithms. For instance, Ahmadi-Nedushan employed multiple k-nearest-neighbors algorithms to predict concrete compressive strength [64]. One of the conclusions of this study was that these algorithms are less computationally intensive than ANNs and do not require training on the data; however, instance-based algorithms can be sensitive to the presence of localized outliers. Although these algorithms generally do not predict as well as the other learning algorithms mentioned herein, they remain a useful technique for private companies and public entities alike, that have access to a large pre-existing dataset of mixture decisions and resulting concrete properties and that can employ a simple instance-based algorithm to quickly predict the properties of future queries.

3.4.3. Decision trees

Decision trees are a machine learning paradigm in which formal rules are obtained from patterns in the data. Like ANNs, decision trees must be trained on data in order to predict the properties of a query. For discrete problems, the algorithms are classification trees. These algorithms classify attributes of the data in order to make predictions about new queries. For continuous datasets, the algorithm is called a regression tree. In this method, the algorithm conducts a series of tests on the data in order to repeatedly partition it. The algorithm selects the partition that minimizes the squared sum of the deviations from the mean in the two partitions. Partitioning continues until user-supplied stopping rules dictate that a branch should not be split further but rather become a terminal node. This terminal node contains the predicted values of the output variable.

The first regression tree algorithm was published by Morgan et al. around the time that other machine learning algorithms were first being developed [65]. Basic tree-based methods experience trouble finding the model with best predictive performance. As a result, research in machine learning since the 1960s has focused on augmenting the concept of a simple tree-based model with additional design features. For instance, multiple additive regression trees (MART) incorporate "boosting" or fitting of a series of models, each with a lower rate of error than the previous, and combining them for better performance. Another method is bootstrap aggregating or "bagging" regression trees.

Bootstrap aggregating utilizes random resampling of the dataset (with replacement) to develop an individual regression model [66]. This procedure is repeated several times to develop multiple regression models, which are then averaged to determine the final predictive model. Another variation, called model trees, go a step further than regression trees and fit a regression model within each node of the tree. Among studies that compare the performance of machine learning algorithms for predicting concrete mixture properties, variations on regression trees often perform better than the other machine learning algorithms. For instance, Erdal demonstrated that regression tree ensembles that use bagging and boosting have superior performance to the simple decision tree model for concrete property prediction [67]. In addition, the model tree approach has been employed to predict compressive strength for a variety of concrete types including HPC [68], recycled aggregate concrete (RAC) [69,70], fiber-reinforced polymer [71], and high-volume mineral-admixture concrete [72].

3.4.4. Support vector machines

Support vector machines (SVMs) are a type of supervised learning technique that use the concept of a "margin" which is either side of a hyperplane that separates two general data classes. The goal with this technique is to maximize this margin between these two data classes, thus creating the largest possible distance between the hyperplane and the data points on either side. The optimum separating hyperplane is found, and the data points that lie on its margin are called support vector points. The solution to the problem is the linear combination of these support vector points. For a more in-depth explanation of the methods used in SVM learning, Byun provides a valuable overview [73].

The original SVM algorithm was invented in 1964 by Vapnik and Chervonenkis [74]; however, it was not until 1992 that SVMs approached their current form when Boser et al. developed a training algorithm to optimize margin classifiers [75]. SVMs were not applied to the field of concrete property modeling until much later, but have since become a widespread machine learning method for concrete performance prediction. For instance, in 2007, Gupta et al. were among the first to predict concrete compressive strength using SVM, where, for the best model, $R^2 = 0.992$ [76]. This study illustrated that SVMs are a useful modeling tool for concrete property modeling particularly when a dataset is small, because the user only needs to define two parameters. A number of other researchers have since used SVMs to predict a variety of concrete mixture properties, such as elastic modulus [77], compressive strength [68,76,78–80], and splitting tensile strength [81].

3.4.5. Comparison of machine learning models

Ideally, the body of research that has applied machine learning algorithms to predict concrete properties would lead to identifying the algorithm that has the best performance. Indeed, machine learning methods can outperform multiple linear regression. However, making definitive conclusions about the "best" algorithm remains difficult—a principle of the No Free Lunch Theorem [82], which states that even very good models perform poorly on some types of problems. Thus, the best performing algorithm is often dependent on the dataset or problem from which it is learning.

Although there is no universal best or silver-bullet algorithm, the bagging, boosting, and other calibrating improvements that have been made with regression tree algorithms have led to predictive models with much improved performance for concrete-related problems. In studies comparing multiple machine learning algorithms, regression and model tree variants often perform the best or at least very well. For instance, Chou compared the average performance of ANN, MART, SVM, multiple linear regression, and bagging regression trees to determine the best learning algorithm for predicting compressive strength of concrete mixtures [68]. All of the algorithms outperformed multiple linear regression by a large margin, and the MART performed best in R^2 and RMSE. The MAPE metric was best with the bagging regression tree. Another study by Deepa et al. compared the performance of an ANN, model tree, and multiple linear regression and concluded that the model tree performed best, closely followed by the ANN in comparison of R^2 and RMSE values [70]. Therefore, at present, regression tree algorithms are a good choice for modeling concrete properties; however, machine learning is a rapidly growing field that has demonstrated potential for expansion and improvement in property prediction of fresh- and hardened state concrete.

3.5. Physics-based models

Physics-based models are a set of mechanistic relationships between the decision variables and objectives. These models differ from those discussed thus far in that physics-based models seek to represent the underlying physics of a system and can only be used if the objective in question can be modeled mechanistically. If the mechanisms and interactions are well-understood, physics-based models are useful for accurate prediction of objectives from a set of decision variables. However, for cement and concrete, these types of models do not exist for the whole mixture design problem; rather, they exist for certain subproblems associated with concrete properties. For instance, some physics-based models predict aggregate packing density as a function of particle size distribution. These models use a particle packing model to correlate particle size distribution to aggregate packing density and to strength and modulus of elasticity, as well as creep and shrinkage. Other research correlates rheological properties of fresh concrete, such as plastic viscosity, to the explanatory variables from the concrete mixture [83–85]. These models are based on the science of rheology and fluid mechanics and are important for predicting fresh-state properties like workability and slump. Chidiac and Mahmoodzadeh [86] evaluated the predictive capabilities of these plastic viscosity models and found that the model from Sudduth et al. performed best among the physics-based models. Still others have modeled how characteristics of aggregates (*e.g.* mineralogy, material properties, and aggregate source) affect concrete performance [87]. A possible future area of concrete property modeling is so-called hybrid modeling, which crosses statistical learning models with physical models for concrete behavior. These types of models have been used in other fields of civil engineering, such as building systems [88,89].

Like statistical and machine learning models, physics-based models can also be tested on their statistical ability to describe data. Barnhouse and Srubar [90], for example, demonstrated that established physics-based models for hydraulic conductivity of macroporous concrete actually had a negative R^2 value, indicating that data is sometimes poorly described by physics-based models.

4. Optimization methods

The previous section focused on the development of equations and models for modeling a concrete mixture; this section discusses mathematical optimization techniques that are most commonly employed to optimize the design of concrete mixtures.

4.1. Formal definition of optimization

The formal optimization of a problem involves finding the best feasible values of the objective functions in a defined domain. Utilizing terminology of Coello Coello et al. [9], optimization is broadly defined as:

Minimize:

$$F(x) = (f_1(x),...,f_k(x)) \tag{5}$$

subject to:

$$g_i(x) = 0,\ i = \{1,..., m\} \tag{6}$$

and:

$$h_j(x) \geq 0, j = \{1,\ldots,p\} \tag{7}$$

A multi-objective problem solution minimizes the objective function in the vector $F(x)$, where x is an n-dimensional decision variable vector $x = (x_1 \ldots x_n)$ in the decision space, Ω. Single-objective optimization can be considered a special case, in which there is a single scalar objective. There exist m inequality constraints and p equality constraints that subject a solution x to a set of requirements that must be met; the solution, x, is feasible if it meets all constraints.

In the presence of objective conflicts, a single solution will not simultaneously optimize all components of the $F(x)$ objective function. Thus, the goal of multi-objective optimization is to generate a tradeoff set of solutions to the problem, using the concept of Pareto optimality. A vector of decision variables, x, is Pareto optimal if no other feasible vector can minimize some objective without causing a simultaneous increase in one or more other objectives. Furthermore, a Pareto optimal set is the set of all decision variable vectors, where the corresponding objectives of the problem cannot be improved simultaneously.

In the case of single-objective optimization, an optimal solution is a solution from the decision space that has the minimum value of the singular objective function while still maintaining acceptable values of the constraints. With single-objective problems, there are no tradeoffs or Pareto curves to consider; instead, the output is a single solution in which the objective function has been optimized.

In the field of optimization, the characteristics of the decisions, objectives, and constraints determine the type of optimization problem and what methods are best-suited to finding a solution. Because of the diversity of different decisions, objectives, and constraints available, one can argue that there are infinite possible problem formulations in concrete mixture design optimization that blend different types of objectives and mathematical solution techniques. However, in this review we focus on a few of the most important types of problem formulations in this field and the optimization techniques that are applicable to them. In the following sections, we concentrate on: linear problems, special cases of nonlinear problems, problems with machine learning objective functions, and multi-objective problems with at least one nonlinear objective.

4.2. Linear programs

An optimization problem is linear (also called a linear program (LP)) when the objective function(s) are linear and where each constraint is either a linear equality or linear inequality. One of the benefits of a linear problem formulation is that if a feasible solution to the problem exists, then it is guaranteed that one or more optimal solutions can be found. LP is a subset of mathematical programming, where exclusively linear expressions are used; furthermore, one of the first and most popular mathematical programming techniques is the simplex method. This method tests adjacent vertices of the feasible set of solutions and determines whether the objective function improves, worsens, or remains unchanged [91]. Another technique for finding the optimum value of an LP is the interior point method; here, the solution is found by iteratively moving through the interior of the feasible region [92]. One example of a LP formulation is a study in which the objective of the design problem was to minimize the rate of corrosion [84]. The rate of corrosion was assumed to be a linear function of the decision variables (cementitious materials content, water-to-cementitious-materials ratio, fine-to-total aggregate ratio, concrete cover thickness, and chloride concentration). In this particular study, the optimization problem was a LP because both the objective and constraints were assumed to be linear. However, in the field of concrete mixture design optimization, studies like this are rare, because many properties of concrete are not well-modeled by linear functions of the decision variables. Thus, a discussion of the methods for solving problems with nonlinear objectives and constraints is merited.

4.3. Special cases of nonlinear programs

An optimization problem is nonlinear when at least one of the objectives or constraints is a nonlinear function. Nonlinear programs (NLP) allow decision-makers to include objectives and constraints that are better represented by nonlinear expressions of the decision variables (*e.g.*, compressive strength). However, unlike, LPs, even if a solution to an NLP exists, there is no single technique that is guaranteed to find the solution for any kind of NLP.

Nevertheless, the literature suggests several techniques to employ nonlinear models for a problem's objectives or constraints, while still retaining the ability to use mathematical programming methods. One such special case is the use of quadratic programming; this is characterized by the use of a quadratic relationship to model the objective function of the problem while retaining all linear constraints [93]. For constrained quadratic problems, the interior point method and active set methods are appropriate solution methods.

Quadratic programming is common in the area of concrete mixture optimization because it enables researchers to incorporate some of the nonlinearities of concrete behavior while guaranteeing the optimal solution is found if it exists. In the literature, the process of optimizing concrete mixtures using quadratic objective functions in an optimization study is often called the response surface methodology (RSM); RSM involves (1) factorial experimental design, (2) fitting the best polynomial equation, and (3) optimizing that equation under the given constraints [94]. In the field of concrete mixture design optimization, researchers have conducted RSM studies with quadratic expressions to optimize compressive strength [37,40,44], slump [41], and void content [43] subject to a variety of linear constraints. In this process, it is not necessarily true that a quadratic model will be better than a linear model for a given objective function. Statistical significance tests are typically performed on each term of the model to determine which are necessary and to determine if a quadratic model is indeed better than a linear model [95]. In the aforementioned RSM studies, it was demonstrated that retaining quadratic terms was statistically significant, meaning they outperform a purely linear objective function.

Another special case of NLPs is characterized by a linear objective and one or more nonlinear constraints. In these optimization problems, a solution technique called the gradient projection method can be invoked [96]. In the realm of concrete mixture optimization, this type of problem formulation and optimization method is most often selected when the objective is to minimize cost and the problem has one or more nonlinear constraints. As was discussed in Section 3, cost is often assumed to be a linear objective because the total cost of a concrete mix is simply a linear combination of the cost of the mixture constituents. This objective can combined with nonlinear design constraints such as meeting targets for compressive strength, slump, air content, and paste drying time [38–40,42,44].

4.4. Metaheuristic optimization

An alternative optimization approach, known as metaheuristic optimization – or "guided search" – seeks to find the global optimal solution to optimization problems, although finding the truly global solution is not guaranteed. Metaheuristic optimization techniques are useful when the relationships being used to model the objectives and constraints do not lend themselves well to the aforementioned mathematical programming approaches. Two examples of this are when there are multiple nonlinear objectives and constraints and when the objectives or constraints are modeled using machine learning methods.

There are many metaheuristic optimization algorithms that have been developed to search for solutions to optimization problems; we direct the reader to Dasgupta et al. [97] and Coello Coello et al. [9], which provide discussion and analysis of the metaheuristic algorithms available. In the field of concrete mixture optimization, metaheuristics based on the concepts of evolution and musical search for perfect harmony are most common. Evolutionary algorithms (EA), for instance, are iterative algorithms where the concepts of natural selection and evolution inspire the mechanisms for generating and searching for optimal solutions [98]. EAs work by (1) generating a population of potential solutions to the problem, (2) evaluating those potential solutions based on fitness functions, (3) selecting some portion of the best solutions to go through crossover and mutation algorithms in order to introduce diversity to the population, and (4) repeating this process over many generations [9]. Harmony search (HS) algorithms are inspired by a musician's process of searching for perfect harmony during improvisation. Like EAs, HS algorithms generate a set of random possible solutions to the problem, change some values of the set of possible solutions, and reevaluate the new solutions over many cycles [99]. Although research in the area of concrete mixture optimization uses algorithms of several names, EAs and HS algorithms are all broad categories of solvers, and it is difficult to show that one of these metaheuristics performs significantly differently from the others. In fact, Weyland demonstrated that HS algorithms do not provide any novelty as a new metaheuristic beyond the new terminology [100]. Despite the lack of a significant difference between EAs and HS, we continue to use this terminology below because both are referred to separately in the literature.

4.4.1. Metaheuristic optimization for multi-objective problems

Given that concrete mixture design involves multiple competing performance criteria, multi-objective optimization is a growing area of concrete mixture optimization research. Metaheuristic algorithms are often the optimization tool of choice for multi-objective problems in concrete mixture design that include multiple nonlinear relationships. In this field of optimization, evolutionary, genetic, and harmony search algorithms dominate the literature. For instance, some studies have looked to minimize cost while maximizing physical concrete performance measures [36]. Other studies in this field seek to minimize both cost and environmental impacts [12,31,32,101]. Still others aim to optimize multiple physical performance objectives at once [102–104]. In all of these cases, evolutionary and harmony search algorithms were the methods chosen to search for the Pareto optimal set because many nonlinear objectives and constraints were selected.

4.4.2. Metaheuristic optimization for problems employing machine-learning models

Another common use of metaheuristic algorithms in the area of concrete mixture optimization is for problems which incorporate machine learning models. Since these models are not explicit expressions, mathematical programming techniques cannot be applied. Instead, the use of metaheuristics has allowed for the optimization of problems with objectives or constraints that are modeled by machine learning algorithms. For example, Cheng et al. maximized the compressive strength of HPC by employing an evolutionary algorithm when the compressive strength was modeled by a SVM [11]. As another example, Ji et al. solved for the minimum cost of reactive powder concrete when the constraints (compressive strength, splitting tensile strength, and slump) were modeled by an ANN; this was accomplished using a harmony search algorithm [13]. Using both machine learning algorithms and metaheuristic optimization algorithms in tandem can become computationally intensive. However, using machine learning to model the objectives has been shown to provide greater predictive performance of actual concrete mixture behavior compared to linear or quadratic models of properties like compressive strength or slump [105,68]. Therefore, for highly nonlinear concrete objectives and where accurate modeling is desired, using machine learning models and metaheuristic optimization can be useful tools.

5. Future trends in concrete mixture design optimization

This review has illustrated the wide variety of modeling and optimization techniques that have been applied in the field of concrete mixture design optimization. Decision-makers have many options for tools and approaches when optimizing problem formulations; however, there are several key areas where the field could be expanded or improved. Below, we discuss potential future opportunities for such expansions and improvements to the field of concrete mixture design optimization.

5.1. Include alternative objective functions

In the field of concrete mixture design, compressive strength and cost are often considered the two most important objectives in optimizing a mixture. Correspondingly, many optimization and modeling studies in the literature focus solely on one or both of these objectives. However, in real-world concrete design and placement, many additional important objectives and constraints must be considered including fresh state properties, such as set time and slump. Set time, for instance, is dependent on cement type, cement fineness, the w/c ratio, and the use of SCMs and admixtures [10]. Slump, an indicator of workability, is influenced by the amounts of water, air entraining agents, water reducers, superplasticizers, and the shape and gradation of aggregates [10]. Since fresh concrete must be workable and set time properly tailored for specific applications, an improved ability to accurately predict and subsequently optimize these properties with respect to cost and strength objectives would aid designers. Thus, we argue that applying some of the more advanced statistical and machine learning techniques to modeling and optimizing these properties would be a valuable contribution to the field.

Furthermore, it is now more common for environmental objectives to be thoroughly considered and quantified. Concrete is the most-used engineering material in the world and, consequently, its production energy and global warming potential – a relative measure of all greenhouse gas emissions of a product or process - are particularly high. An increase in global awareness of so-called embodied emissions of buildings [106] has prompted an industry-wide movement over the past decade to mandate quantification through lifecycle assessment (LCA) and subsequent reductions in the environmental impacts of building materials. While global warming potential and depletion of non-renewable energy resources have emerged as primary metrics reported in LCAs, ozone depletion, acidification, and eutrophication are environmental impact categories that should also be considered in the design of more environmentally sustainable concrete mixtures. As was discussed in Section 3.1, some researchers have estimated life cycle environmental impacts of certain types of concretes in an optimization context. However, only a few studies have attempted to minimize environmental objectives by changing the decision variables [31,32]; instead, most studies calculate the impacts with mixture designs set *a priori*. Therefore, incorporating environmental objectives into optimization studies will advance the field by allowing researchers to more holistically examine impacts of different decision variables on environmental as well as structural and economic objectives and constraints.

5.2. Incorporate many objectives

Another suggestion for future work in the field is to consider many objectives simultaneously. Currently, most optimization studies in concrete mixture design consider only one or two objectives. However, in real applications, a relatively large number objectives are simultaneously important. Therefore, a valuable advancement would be to develop tools that consider the tradeoffs between many different and usually competing objectives. The advantages and disadvantages of so-called *many-objective optimization* are discussed in Fleming et al. [107]. The assumed advantage of many-objective optimization is that considering more objectives adds information and helps decision-makers understand tradeoffs between objectives. However, there can also be challenges associated with considering many objectives, including the problem of data visualization, how to process and analyze the data (*i.e.*, how to choose a solution), and high computational intensity. To partially relieve the first two of these challenges, data visualization techniques, such as parallel coordinates, can be helpful for visualizing, understanding, and choosing between Pareto-optimal solutions. This particular data visualization technique was pioneered by Alfred Inselberg in the 1970s as a way to visualize high-dimension data [108]. As an example of many-objective optimization, a four-objective study could consider cost, compressive strength, CO_2e emissions, and durability of the concrete. The added value of including more objectives is that the decision-maker can analyze how these objectives interact. For instance, it is possible that including durability would quantitatively demonstrate to a decision-maker that a higher-cost solution could pay off environmentally and economically in the long-run.

Optimization with many objectives has become well-applied in other civil engineering disciplines. For instance, in the realm of building systems engineering, one study looked at the challenge building owners face to identify an optimal set of building upgrade measures for maximizing the sustainability of their buildings. This study included objectives of (1) minimizing environmental impacts (*e.g.*, CO_2 emissions, refrigerant impacts, water consumption), (2) minimizing building upgrade costs; and (3) maximizing the number of earned points from the United States Green Building Council (USGBC) Leadership in Energy and Environmental Design (LEED) rating system [109]. Similarly, in the field of water resources engineering, many-objective optimization has been applied to enhance decision-making. For instance, Smith et al. [110] propose a participatory framework for determining what objectives are important to water supply managers. The objectives included minimizing time in restriction, minimizing costs, maximizing total end-of-year storage, and maximizing the amount of time a reservoir spends above a given elevation [110]. Given the success in other civil engineering fields in considering multiple objectives, it is reasonable that a many-objective approach in the field of concrete mixture design may be valuable.

5.3. Standardize optimization terminology

Finally, we suggest that optimization terminology be standardized across the field of concrete mixture design. Since concrete behavior is considered in many realms, such as experimental, computational, and industry settings, the term *optimization* can have many meanings. On the experimental side, optimization often refers the use of regression and response-surface methodologies to look at relationships between several explanatory variables and a response variable [111]. The optimal response is found under the specific range of decision variables considered. Typically, experimental design methods, such as the Box-Behnken or full factorial designs methods, are used [112]. Thus, optimization in this context refers to the experimental design and finding a best-performing response (as represented in Fig. 1a). In other studies, optimization refers to the mathematical and computational methods used to solve the mixture design problem when relationships between the decision variables and objectives have been modeled (represented in Fig. 1b). In other cases, the term optimization is sometimes used to refer to the improvement of individual constituents of concrete mixtures. An example of this is when the aggregate grading is altered to improve packing [20,113] or lowering the cement content in order to reduce environmental impacts [114].

In order cultivate more standardized terminology, the use of modifiers would help clarify the user's meaning of the word *optimization*. For instance, specifying that researchers have conducted *experimental optimization* or *computational optimization* would clarify which techniques are used in the study. Furthermore, studies that optimize a constituent of the mixture design should be clear about the constituent system which is being optimized as well as the optimization approach (*i.e.*, experimental *vs.* computation).

6. Conclusions

Traditionally, in the field of concrete mixture design, decision-makers have cared most about achieving required targets for certain concrete properties, such as attaining a specified early-age compressive strength while maintaining a workable concrete. Recently, however, the field has beendeveloping and applying new computational tools and techniques to design concrete mixtures with tailored and fine-tuned properties. Recent work has focused on modeling and optimizing important concrete properties - such as compressive strength and slump - as a function of individual mixture constituents, necessitating models that relate consitutents to concrete performance in the fresh- and hardened-state.

This paper examined the types of problem formulations that are typical in the field of concrete mixture design optimization and discussed methods available for modeling and optimizing concrete problem formulations, as well as their applicability to different types of design problems. The methods used to model concrete objectives can involve models based on linear combination, statistics, machine learning, and physics. In the realm of optimization, mathematical programming and metaheuristic search methods are commonly used.

This review also highlighted future directions of research in this field. Increasingly, decision-makers must consider multiple objectives and satisfy many design criteria, prompting an ever-growing need for concrete mixtures that are cheaper, stronger, more workable, durable, and more environmentally sustainable. Thus, developing appropriate and accurate models for these objectives and implementing multi-objective optimization methods will contribute to progress in the field. Advancements in the area of life cycle assessment and data visualization will also contribute to designers' abilities to make informed decisions about concrete mixtures that most optimally satisfy multiple performance criteria.

Acknowledgments

This research was made possible by the Department of Civil, Environmental, and Architectural Engineering, the College of Engineering and Applied Sciences, and the Sustainable Infrastructure Materials Laboratory (SIMLab) at the University of Colorado Boulder, with support from the National Science Foundation (Award No. CMMI-1562557). This work represents the views of the authors and not necessarily those of the sponsors.

References

[1] Annual world production, Available: http://inventor.grantadesign.com/en/notes/science/material/C04%20Annual%20world%20production.htm , Accessed date: 6 September 2017(Online).
[2] D.A. Abrams, Design of Concrete Mixtures, Structural Materials Research Laboratory, Chicago, IL, 1918.
[3] A.C.I. 318, Building Code Requirements for Reinforced Concrete, American Concrete Institute, Farmington Hills, MI, 2011, p. 318.
[4] S. Kosmatka, B. Kerkhoff, W.C. Panarese, Design and Control of Concrete Mixtures, (2002).
[5] Indian Standard Concrete Mix Proportioning - Guidelines (First Revision), Bureau of Indian Standards, New Delhi, India, 2009 (IS-10262).
[6] K.A. Soudki, E.F. El-Salakawy, N.B. Elkum, Full factorial optimization of concrete mix design for hot climates, J. Mater. Civ. Eng. 13 (6) (Dec. 2001) 427–433.
[7] S. Hınıslıoğlu, O.Ü. Bayrak, Optimization of early flexural strength of pavement concrete with silica fume and fly ash by the Taguchi method, Civ. Eng. Environ. Syst. 21 (2) (Jun. 2004) 79–90.
[8] S.L. Correia, F.L. Souza, G. Dienstmann, A.M. Segadães, Assessment of the recycling potential of fresh concrete waste using a factorial design of experiments, Waste Manag. 29 (11) (Nov. 2009) 2886–2891.
[9] C.A. Coello Coello, G.B. Lamont, D.A. Van Veldhuizen, Evolutionary Algorithms for Solving Multi-objective Problems, second edition, Springer, 2007.
[10] M.S. Mamlouk, J.P. Zaniewski, Materials for Civil and Construction Engineers, 2nd ed., Pearson Education, Inc., Upper Saddle River, NJ, 2006.
[11] M.-Y. Cheng, D. Prayogo, Y.-W. Wu, Novel genetic algorithm-based evolutionary support vector machine for optimizing high-performance concrete mixture, J. Comput. Civ. Eng. 28 (4) (2014).
[12] W. Meng, M. Valipour, Optimization and performance of cost-effective ultra high

performance concrete, Mater. Struct. 50 (29) (2017).
[13] T. Ji, Y. Yang, M. Fu, H.-C. Wu, Optimum design of reactive powder concrete mixture proportion based on artificial neural and harmony search algorithm, ACI Mater. J. 114 (2017).
[14] T. Clarkin, W. Raseman, J. Kasprzyk, and J. Herman, "Diagnostic assessment of constraints in multiobjective evolutionary algorithms for water resources," J. Water Resour. Plan. Manag., https://doi.org/10.1061/(ASCE)WR.1943-5452.0000940.
[15] ASTM C 33: Standard Specification for Concrete Aggregates, American Society for Testing and Materials (ASTM), West Conshohocken, PA, ASTM C 33, (2016).
[16] M.J. Simon, Concrete Mixture Optimization Using Statistical Methods: Final Report, FHWA Office of Infrastructure Research and Development, McLean, VA, 2003 (FHWA-RD-03-060).
[17] S. Yehia, C.Y. Tuan, D. Ferdon, B. Chen, Conductive concrete overlay for bridge deck deicing: mixture proportioning, optimization, and properties, ACI Struct. J. 97 (Mar. 2000) 172–181.
[18] I.-C. Yeh, Computer-aided design for optimum concrete mixtures, Cem. Concr. Compos. 29 (3) (Mar. 2007) 193–202.
[19] G. Shakhmenko, J. Birsh, Concrete mix design and optimization, Int. PhD Symp. Civ. Eng. vol. 2, 1998.
[20] F. de Larrard, T. Sedran, Optimization of ultra-high-performance concerete by the use of a packing model, Cem. Concr. Res. 24 (6) (1994) 997–1009.
[21] S. Ahmad, Optimum concrete mixture design using locally available ingredients, Arab. J. Sci. Eng. 32 (1B) (Apr. 2007) 27–33 (Springer Sci. Bus. Media BV).
[22] J. Guinée, Handbook on life cycle assessment — operational guide to the ISO standards, Int. J. Life Cycle Assess. 6 (5) (Sep. 2001) 255.
[23] "U.S. Life Cycle Inventory Database," National Renewable Energy Laboratory, Available: https://www.lcacommons.gov/nrel/search, (2012) (Online).
[24] "The Ecoinvent Database," Ecoinvent, Available: http://www.ecoinvent.org/database/database.html, (2016) (Online).
[25] "GaBi Databases: GaBi Software," thinkstep GaBi, Available: http://www.gabi-software.com/databases/gabi-databases/, (2016) , Accessed date: 3 October 2017(Online).
[26] A. Loijos, Life Cycle Assessment of Concrete Pavements: Impacts and Opportunities, Massachusetts Institute of Technology, Cambridge, Massachusetts, 2011.
[27] S. Tae, C. Baek, S. Shin, Life cycle CO2 evaluation on reinforced concrete structures with high-strength concrete, Environ. Impact Assess. Rev. 31 (3) (Apr. 2011) 253–260.
[28] D.N. Huntzinger, T.D. Eatmon, A life-cycle assessment of Portland cement manufacturing: comparing the traditional process with alternative technologies, J. Clean. Prod. 17 (7) (May 2009) 668–675.
[29] K.-H. Yang, Y.-B. Jung, M.-S. Cho, S.-H. Tae, Effect of supplementary cementitious materials on reduction of CO2 emissions from concrete, J. Clean. Prod. 103 (Supplement C) (Sep. 2015) 774–783.
[30] S. Marinković, V. Radonjanin, M. Malešev, I. Ignjatović, Comparative environmental assessment of natural and recycled aggregate concrete, Waste Manag. 30 (11) (Nov. 2010) 2255–2264.
[31] T. Kim, S. Tae, Assessment of the CO2 emission and cost reduction performance of a low-carbon-emission concrete mix design using and optimal mix design system, Renew. Sust. Energ. Rev. 25 (2013) 729–741.
[32] W. Park, Genetic-algorithm-based mix proportion design method for recycled aggregate concrete, Trans. Can. Soc. Mech. Eng. 37 (3) (2013) 345–354.
[33] D.A. Freedman, Statistical Models: Theory and Practice, 7th ed., Cambridge University Press, New York, NY, 2009.
[34] E. Ozbay, M. Gesoglu, E. Guneyisi, Transport properties based multi-objective mix proportioning optimization of high performance concrete, Mater. Struct. (2011) 139–154.
[35] O. Akalin, K.U. Akay, B. Sennaroglu, Self-consolidating high-strength concrete optimization by mixture design method, ACI Mater. J. 107 (4) (2010).
[36] A. Khan, J. Do, D. Kim, Cost effective optimal mix proportioning of high strength self-compacting concrete using response surface modeling, Comput. Concr. 17 (5) (2016) 629–638.
[37] M. Muthukumar, D. Mohan, Optimization of mechanical properties of polymer concrete and mix design recommendation based on design of experiments, J. Appl. Polym. Sci. 94 (3) (2004) 1107–1116.
[38] R. Aggarwal, M. Kumar, R. Sharma, M.K. Sharma, Reliability-based design optimization of concrete mix proportions using generalized ridge regression model, Int. J. Sci. Eng. 8 (1) (2015).
[39] S. Ahmad, S.A. Alghamdi, A statistical approach to optimizing concrete mixture design, Sci. World J. 2014 (2014).
[40] M. Berry, B. Kappes, L. Kappes, Optimization of concrete mixtures containing reclaimed asphalt pavement, Mater. J. 112 (6) (2015) 723–733.
[41] B.E. Jimma, P.R. Rangaraju, Chemical admixtures dose optimization in pervious concrete paste selection - a statistical approach, Constr. Build. Mater. 101 (1) (2015) 1047–1058.
[42] S. Kostic, N. Vasovic, B. Marinkovic, Robust optimization of concrete strength estimation using response surface methodology and Monte Carlo simulation, Eng. Optim. 49 (5) (2016) 864–877.
[43] M. Muthukumar, D. Mohan, M. Rajendran, Optimization of mix proportions of mineral aggregates using Box Behnken design of experiments, Cem. Concr. Compos. 25 (7) (2003) 751–758.
[44] B. Simsek, Y. Tansel, E. Simsek, A.B. Guvenc, Development of a graphical user interface for determining the optimal mixture parameters of normal weight concretes: a response surface methodology based quadratic programming approach, Chemom. Intell. Lab. Syst. 136 (2014) 1–9.
[45] S. Bhanja, Investigations on the compressive strength of silica fume concrete using statistical methods, Cem. Concr. Res. 32 (9) (2002) 1391–1394.
[46] M. Simon, Concrete Mixture Optimization Using Statistical Methods: Final Report, Federal Highway Administration Office of Infrastructure Research and Devleopment, McLean, VA, 2004 (FHWA-RD-03-060).
[47] M. Kuhn, K. Johnson, Applied Predictive Modeling, Springer Science & Business Media, New York, NY, 2013.
[48] S.-C. Lee, Prediction of concrete strength using artificial neural networks, Eng. Struct. 25 (7) (Jun. 2003) 849–857.
[49] S. Dreiseitl, L. Ohno-Machado, Logistic regression and artifical neural network classification models: a methodology review, J. Biomed. Inform. 35 (5–6) (2002) 352–359.
[50] O. Omidvar, Progress in Neural Networks, vol. 2, Ablex Publishing Corporation, 1994.
[51] J. Kasperkiewicz, J. Racz, A. Dubrawski, HPC strength prediction using artificial neural network, J. Comput. Civ. Eng. 9 (4) (1995) 279–284.
[52] I. Yeh, Modeling of strength of high-performance concrete using artificial neural networks, Cem. Concr. Res. 28 (12) (1998) 1797–1808.
[53] M. Nehdi, H. El Chabib, M. El Naggar, Predicting performance of self-compacting concrete mixtures using artificial neural networks, Mater. J. 98 (5) (2001) 394–401.
[54] R. Siddique, P. Aggarwal, Y. Aggarwal, Prediction of compressive strength of self-compacting concrete containing bottom ash using artificial neural networks, Adv. Eng. Softw. 42 (10) (2011) 780–786.
[55] W.P.S. Dias, S.P. Pooliyadda, Neural networks for predicting properties of concretes with admixtures, Constr. Build. Mater. 15 (7) (2001) 371–379.
[56] A. Oztas, M. Pala, E. Ozbay, E. Kanca, N. Caglar, M.A. Bhatti, Predicting the compressive strength and slump of high strength concrete using neural network, Constr. Build. Mater. 20 (9) (2006) 769–775.
[57] E. Ghafari, M. Bandarabadi, H. Costa, E. Julio, Prediction of fresh and hardened state properties of UHPC: comparative study of statistical mixture design and an artificial neural network model, J. Mater. Civ. Eng. 27 (11) (2015).
[58] I.B. Topcu, Prediction of properties of waste AAC aggregate concrete using artificial neural network, Comput. Mater. Sci. 41 (1) (2007) 117–125.
[59] M. Alshihri, A. Azmy, M. El-Bisy, Neural networks for predicting compressive strength of structural light weight concrete, Constr. Build. Mater. 23 (6) (2009) 2214–2219.
[60] O. Fuentes, R.K. Gulati, Instance-based machine learning methods for the prediction of the stellar atmospheric parameters, Astron. Data Anal. Softw. Syst. 216 (2000) 611–614.
[61] G. Sebestyen, Decision-making Processes in Pattern Recognition, Macmillan, New York, NY, 1962.
[62] N. Nilsson, Learning Machines: Foundations of Trainable Pattern-classifying Systems, McGraw-Hill, 1965.
[63] T. Cover, P. Hart, Nearest neighbor pattern classification, IEEE Trans. Inf. Theory 13 (1) (1967) 21–27.
[64] B. Ahmadi-Nedushan, An optimized instance based learning algorithm for estimation of compressive strength of concrete, Eng. Appl. Artif. Intell. 25 (5) (2012) 1073–1081.
[65] W.-Y. Loh, Fifty years of classification and regression trees, Int. Stat. Rev. 82 (3) (2014) 329–348.
[66] L. Breiman, Bagging predictors, Mach. Learn. 24 (2) (1996) 123–140.
[67] H.I. Erdal, Two-level and hybrid ensembles of decision trees for high performance concrete compressive strength prediction, Eng. Appl. Artif. Intell. 26 (7) (2013) 1689–1697.
[68] J.-S. Chou, C.-K. Chiu, M. Farfoura, I. Al-Taharwa, Optimizing the prediction accuracy of concrete compressive strength based on a comparison of data-mining techniques, J. Comput. Civ. Eng. 25 (3) (2011) 242–253.
[69] N. Deshpande, S. Londhe, S. Kulkarni, Modeling compressive strength of recycled aggregate concete by Artificial Neural Network, Model Tree, and Non-linear Regression, Int. J. Sustain. Built Environ. 3 (2) (2014) 187–198.
[70] C. Deepa, K. Sathiya Kumari, V. Pream Sudha, Prediction of the compressive strength of high performance concrete mix using tree based modeling, Int. J. Comput. Appl. 6 (5) (2010) 18–24.
[71] I. Mansouri, T. Ozbakkaloglu, O. Kisi, Predicting behavior of FRP-confined concrete using neuro fuzzy, neural network, multivariate adaptive regression splines and M5 model tree techniques, Mater. Struct. 49 (10) (2016) 4319–4334.
[72] Y. Ayaz, A.F. Kocamaz, M.B. Karakoc, Modeling of compressive strength and UPV of high-volume mineral-admixtured concrete using rule-based M5 rule and tree-model M5P classifiers, Constr. Build. Mater. 94 (2015) 235–240.
[73] H. Byun, S.-W. Lee, "Applications of Support Vector Machines for Pattern Recognition: A Survey," Presented at the Pattern Recognition With Support Vector Machines, Niagara Falls, Canada, (2002), pp. 213–236.
[74] V. Vapnik, Chervonenkis, A note on one class of perceptrons, Autom. Remote. Control. 25 (1964).
[75] B.E. Boser, I.M. Guyon, V.N. Vapnik, A training algorithm for optimal margin classifiers, (1992).
[76] S.M. Gupta, Support vector machines based modelling of concrete strength, Int. J. Intell. Technol. 3 (1) (2007) 12–18.
[77] K. Yan, C. Shi, Prediction of elastic modulus of normal and high strength concrete by support vector machine, Constr. Build. Mater. 24 (8) (2010) 1479–1485.
[78] M.-Y. Cheng, J.-S. Chou, High-performance concrete compressive strength prediction using time-weighted evolutionary fuzzy support vector machines inference model, Autom. Constr. 28 (4) (2012) 106–115.
[79] R. Siddique, Modeling properties of self-compacting concrete: support vector machiens approach, Comput. Concr. 5 (5) (2008) 461–473.

[80] B.G. Aiyer, D. Kim, N. Karingattikkal, P. Samui, P.R. Rao, Prediction of compressive stregnth of self-compacting concrete using least square support vector machine and relevance vector machine, KSCE J. Civ. Eng. 18 (6) (2014) 1753–1758.
[81] K. Yan, H. Xu, G. Shen, P. Liu, Prediction of splitting tensile strength from cylinder compressive strength of concrete by support vector machine, Adv. Mater. Sci. Eng. 2013 (2013).
[82] D. Wolpert, No free lunch theorem for optimization, IEEE Trans. Evol. Comput. 1 (1997) 467–482.
[83] J. Murata, H. Kikukawa, Viscosity equation for fresh concrete, ACI Mater. J. 89 (3) (1992) 230–237.
[84] T. Roshavelov, Prediction of fresh concrete flow behavior based on analytical model for mixture proportioning, Cem. Concr. Res. 35 (5) (2005) 831–835.
[85] C. Hu, The rheology of fresh high-performance concrete, Cem. Concr. Res. 26 (2) (1996) 28.
[86] S.E. Chidiac, F. Mahmoodzadeh, Plastic viscosity of fresh concrete - a critical review of predictions methods, Cem. Concr. Compos. 31 (8) (2009) 535–544.
[87] D.P. Bentz, J. Arnold, M.J. Boisclair, S.Z. Jones, Influence of aggregate characteristics on concrete performance, National Institute of Standards and Technology, NIST Technical Note 1963, May 2017.
[88] S.A. Vaghefi, M.A. Jafari, J. Zhu, J. Brouwer, Y. Lu, A hybrid physics-based and data driven approach to optimal control of building cooling/heating systems, IEEE Trans. Autom. Sci. Eng. 13 (2) (Apr. 2016) 600–610.
[89] J. Granderson, G. Lin, Building energy information systems: synthesis of costs, savings, and best-practice uses, Energ. Effic. 9 (6) (Dec. 2016) 1369–1384.
[90] P. Barnhouse, W.V. Srubar III, Material characterization and hydraulic conductivity modeling of macroporous recyled-aggregate pervious concrete, Constr. Build. Mater. 110 (2016) 89–97.
[91] S.P. Bradley, A. Hax, Magnanti, Applied Mathematical Programming, Addison-Wesley Publishing Company, 1977.
[92] U. Diwekar, Introduction to Applied Optimization, 22 Springer, US, 2008.
[93] P.A. Jensen, J.F. Bard, Nonlinear Programming Methods S2 Quadratic Programming, University of Texas, Austin, 2012.
[94] A. Khuri, Mukhopadhya Siuli, Response surface methodology, WIREs Comput. Stat. 2 (March/April) (2010).
[95] Concrete Mixture Optimization Using Statistical Methods, Federal Highway Administration, Washington, D.C., 2003 (FHWA-RD-03-060).
[96] J. Rosen, The gradient projection method for nonlinear programming. Part II. Nonlinear constraints, J. Soc. Ind. Appl. Math. 9 (4) (Dec. 1961) 514–532.
[97] P. Dasgupta, P.P. Chakrabarti, S.C. DeSarkar, Multiobjective Heuristic Search, Springer, 1999.
[98] A. Petrowski, S. Ben-Hamida, Evolutionary Algorithms, vol. 9, John Wiley & Sons, Inc., Hoboken, New Jersey, 2017.
[99] X.Z. Gao, V. Govindasamy, H. Xu, X. Wang, K. Zenger, Harmony search method: theory and applications, Comput. Intell. Neurosci. 2015 (2015).
[100] D. Weyland, A critical analysis of the harmony search algorithm - how not to solve sudoku, Oper. Res. Perspect. 2 (2015) 97–105.
[101] N. Bouzoubaa, B. Fournier, Optimization of fly ash content in concrete, Cem. Concr. Res. 33 (2003) 1029–1037.
[102] J.-H. Lee, Y.-S. Yoon, J.-H. Kim, A new heuristic algorithm for mix design of high-performance concrete, KSCE J. Civ. Eng. 16 (6) (2012) 974–979.
[103] C.-H. Lim, Y.-S. Yoon, J.-H. Kim, Genetic algorithm in mix proportioning of high-performance concrete, Cem. Concr. Res. 34 (3) (2003) 409–420.
[104] A. Baykasoglu, A. Oztas, E. Ozbay, Prediction and multi-objective optimization of high-strength concrete parameters via soft computing approaches, Expert Syst. Appl. 36 (3) (2009) 6145–6155.
[105] J.-S. Chou, C.-F. Tsai, A.-D. Pham, Y.-H. Lu, Machine learning in concrete strength simulations: multi-nation data analytics, Constr. Build. Mater. 73 (Dec. 2014) 771–780.
[106] G.P. Hammond, C.I. Jones, Embodied energy and carbon in construction materials, Proc. Inst. Civ. Eng. Energy 161 (2) (May 2008) 87–98.
[107] P. Fleming, R. Purshouse, R. Lygoe, Many-objective optimization: an engineering design perspective, Lect. Notes Comput. Sci 3410 (2005).
[108] A. Inselberg, Parallel Coordinates: Visual Multidimensional Geometry and Its Applications, (2009) (Springer).
[109] Abdallah Moatassem, El-Rayes Khaled, Multiobjective optimization model for maximizing sustainability of existing buildings, J. Manag. Eng. 32 (4) (Jul. 2016) 04016003.
[110] R. Smith, J. Kasprzyk, L. Dilling, Participatory Framework for Assessment and Improvement of Tools (ParFAIT): increasing the impact and relevance of water management decision support research, Environ. Model. Softw. 95 (Supplement C) (Sep. 2017) 432–446.
[111] R.H. Myers, D.C. Montgomery, C.M. Anderson-Cook, Response Surface Methodology: Process and Product Optimization Using Designed Experiments, 3rd edition, Wiley, Hoboken, N.J., 2009.
[112] M. Cavazzuti, Optimization Methods: From Theory to Design Scientific and Technological Aspects in Mechanics, Springer Science & Business Media, 2012.
[113] A. Amirjanov, K. Sobolev, Optimization of a computer simulation model for packing of concrete aggregates, Part. Sci. Technol. 26 (4) (2008).
[114] E. Yurdakul, Optimizing Concrete Mixtures With Minimum Cement Content for Performance and Sustainability, Iowa State University, Ames, Iowa, 2010.