A simple, yet reasonable approach towards optimal mixture design

Technical Report

Mathijs Schuurmans
info@optia.be

January 13, 2025

Optia BV,
Rotspoelstraat 152, 3001
Heverlee, Belgium

Contents

1 Introduction

The challenge to reduce greenhouse emissions while maintaining excellent quality of end products has driven industrial demand for new material mixtures, which can be produced more sustainably, e.g., at lower temperatures, or using more abundant resources, like waste streams from related industries. Traditionally, developing new mixtures is a costly iterative process, involving large amounts of experimentation in the lab, which often still leads to suboptimal outcomes.

Significant efficiency gains can be made—both in experimentation time and in quality of the final solution—through the use of mathematical modeling and numerical optimization. Unfortunately however, a naive application of general-purpose statistical software or pretrained AI models typically does not provide the desired results: They output black-box solutions that are inscrutable and which hardly provide any guarantees regarding their quality. Furthermore, in order to train models with sufficient modeling capacity to represent the processes of interest (e.g., the relation between cement's composition and its mechanical properties), the amount of training data required to determine its parameters far exceeds what is expected to be available.

Instead, we present a hybrid approach towards building mathematical models that predict material properties from a given mixture composition. We consider simple model classes whose parameters can be estimated from only a small number of experiments, but impose domain-specific constraints on the model to ensure that the output is guaranteed to obey known physical laws. Given this model, we can then synthesize mixtures to optimize a given objective, while satisfying hard constraints (e.g., minimize CO_2 emissions while keeping hardened compressive strength larger than 20 MPa.)

We first formally describe the type of problem we aim to tackle, which is split into two parts. Then, we illustrate the proposed approach on a basic case study related to mixture design of concrete.

2 Problem statement

We present our methodology in a slightly generalized abstract form, emphasizing its versatility with respect to the application domain.

2.1 Optimal design

Suppose we are able to decide on a decision vector $x \in X \subseteq \mathbb{R}^{n_x}$, where X is an arbitrary compact set. In the case study of section 3, x will contain the weight fractions of mixture components in a concrete mixture. However, it could additionally involve process parameters, like the turning rate of the mixer, the temperature, ... Furthermore, we may additionally consider parameters which may affect the end result, but which are beyond our control, such as the humidity, ambient temperature, etc. Although this is straightforwardly accommodated in the presented framework, we will ignore such external factors here for ease of exposition.

Given the design x (and the environmental parameters), the process we are aiming to optimize, will produce an output, which is characterized by a vector $y \in \mathbb{R}^{n_y}$ of properties: $y = p(x)$, where p is some *a priori* unknown function. In the case study in section 3, y represents the 3-day, 7-day, and 28-day compressive strength of the concrete [MPa], as well as the CO_2-equivalent of greenhouse gasses emitted to produce a kilogram of the mixture [kg]. We will assume that an output y is admissible if it satisfies some inequality $g(y) \leq 0$, where g is a known function. These inequalities can represent design limitations dictated by a spec, or simply physical constraints that simply have to hold for the solution to make sense.

Based on our preferences, we can finally evaluate the resulting product based on an objective function $L : \mathbb{R}^{n_x+n_y} \to \mathbb{R}$. Given that the process model p is known, L allows us to compute the quality of any design x as $L(x, p(x))$. By convention, the lower this value, the better. Our goal, however, is not simply to predict the quality of a given design x, but compute the design $x^\star$ that achieves the best quality among all possible designs X. Formally, this task is expressed by the following optimization problem

$$\mathbf{P}: \quad \operatorname*{\mathbf{minimize}}_{x \in X} L(x, p(x)) \text{ s.t. } g(p(x)) \leq 0. \tag{1}$$

Of course, in order to solve (1), we require an estimate of the process model p, as well as analytical expressions for the objective and constraint functions L and g. A simple yet effective approach for estimating p is presented in section 2.2. Typically g is known directly from the task at hand. For instance, lower bounds on the compressive strength of the designed concrete mixture are usually specified by the envisioned application. Similarly, the objective L may be known in simple cases: For instance, we may simply want to minimize the CO_2 emissions, subject to the strength requirements (as we do in section 3). However, there could also be multiple objectives at play, which may even be at odds with one another. In that case, the designer is required to assign a weight (signifying the importance) to each of these desiderata, in order to obtain an overall objective which can be optimized. Sometimes it is difficult for a practitioner to manually select such weights. Fortunately, there exist simple methods to estimate the weights from direct feedback of the designer. For the sake of brevity, we do not consider this case here.

2.2 Model estimation

For simplicity, we will model the process p by a linear mapping of a fixed feature vector $z(x) \in \mathbb{R}^{n_z}$

$$y = p(x) = Pz(x).$$

In order to estimate the value of P, we assume that we have access to a (small) dataset of experimental measurements $D = \{(z^{(i)}, y^{(i)})\}_{i=1}^m$, and furthermore, we have domain knowledge in the form of inequalities that are required to hold, namely $h_i^\top y \leq 0$, $i = 1, \dots, n_h$. These inequalities need to hold regardless of the choice of x, or equivalently,

$$\max_{z \in Z} h_i^\top P z = \sigma_Z(P^\top h_i) \leq 0, \tag{2}$$

where σ_Z denotes the support function of a convex set $Z = \{z(x) : x \in X\}$ [1]. In many cases, e.g., when Z is a polyhedral set, (2) can be reformulated exactly as a finite-dimensional constraint on P, using duality [2]. An example is provided in section 3.

Let $E : \mathbb{R}^{n_y} \times \mathbb{R}^{n_y} \to \mathbb{R}$ be a loss function (e.g., the squared error: $E(\hat{y}, y) = \|\hat{y} - y\|_2^2$), and let R be a regularization function. Then, the parameters are obtained by solving the following optimization problem

$$\begin{aligned} \underset{P}{\textbf{minimize}}\ & R(P) + \sum_{i=1}^m E(Pz^{(i)}, y^{(i)}) \\ \textbf{subj. to}\ & \sigma_Z(P^\top h_j), \quad \forall j \in \{1, \dots, n_h\}. \end{aligned} \tag{3}$$

Note that if E is the squared norm, and $n_h = 0$, (3) is separable in the rows of P and the problem reduces to a series of classical (regularized) linear regression problems (one for each component of y). In general, (3) is slightly more complex, but still very manageable to solve. In fact, if Z is a polyhedral set, the constraints (2) can be reformulated as linear constraints, and (3) is a quadratic program, for which many efficient solvers exist, e.g., [3, 4].

3 Case study – concrete mixture design

We illustrate the described methodology by means of a concrete mixture design application. Here, the decision vector x describes the weight fractions of the mixture components, and therefore,

$$x \in \Delta = \{x \in \mathbb{R}^n \mid x \geq 0, \textstyle\sum_{i=1}^n x_i = 1\}.$$

The outputs are the compressive strengths [MPa] at 3, 7 and 28 days, respectively denoted by y_{S3}, y_{S7}, y_{S28}; the CO_2 emissions [kg] per kg of concrete, y_{env}; and the monetary cost per kg of concrete [€], y_{cost}. In summary, $y = (y_{S3}, y_{S7}, y_{S28}, y_{env}, y_{cost})$. For simplicity, let us use an affine basis, namely $z(x) = (x, 1)$.

In order to estimate the model parameters, we utilize the freely available *Concrete Compressive Strength* dataset [5] from the UC Irvine Machine Learning Repository. We use a randomly selected subset containing 70% of the datapoint of the dataset for training, and evaluate the model using the remaining 20%. Since we are modeling the (likely nonlinear) relation between the concrete mixture to its properties by an affine map, it is unlikely we can trust the model far outside the data distribution. For this reason, we set $Z = \mathbf{conv}\{z(x^{(1)}), \ldots, z(x^{(n)})\}$ equal to the convex hull of the data points in the training set.

Assuming a properly prepared mixture, let us suppose that it is known beforehand that the strength of the concrete at 3, 7 and 28 days is strictly increasing by at least some factor $\alpha \geq 0$. We can impose this knowledge into the statistical model through a constraint of the form (2): we can express the constraint that the 7-day strength is at least $\alpha \times 100\%$ larger than the 3-day strength as

$$
\begin{aligned}
& y_{S7} \geq (1+\alpha) y_{S3} \\
\iff & \max_{z \in Z} (1+\alpha) s_1^\top P z - s_2^\top P z \leq 0 \\
\iff & (1+\alpha) s_1^\top P z^{(i)} - s_2^\top P z^{(i)} \leq 0, \quad i \in \{1, \ldots, n\}
\end{aligned}
$$

where s_i is the i'th standard basis vector. Equivalently, we may impose that $((1+\alpha)P_1 - P_2) z^{(i)} \leq 0$, for all datapoints $z^{(i)} = z(x^{(i)})$, where P_j denotes the j'th row of P. This is a linear constraint on P. The same argument can be applied to impose an analogous increase between the 7-day and 28-day strength. As a final constraint ensuring sanity of the output, we impose that for all $z \in Z$, $y \geq 0$, as negative values are physically meaningless for the output quantities under consideration.

Accordingly, we solve regression problem (3), taking $R(P) = 0.001\|P\|_F^2$ and $E(\hat{y}, y) = \|\hat{y} - y\|_2^2$, the squared Euclidean norm. The resulting model is visualized in fig. 3. Most notably, the model reflects known facts that the cement-to-water is one of the most important factors in determining the strength of the mixture. Furthermore, we see that the amount of superplasticizer is predictive for the strengths at 3 and 7 days. This could indicate that the particular superplasticizer used is one that shortens the curing time of the concrete.

We evaluate the model by the mean-squared (/absolute) error (normalized by the mean-squared (/absolute) value of the output, to obtain a dimensionless error metric) with the validation set. The results are given in table 1. A visual representation of the model error is given in fig. 2. The validation points cluster around the predicted values, both at small and large values, indicating relatively little bias in the predictive model, despite its simplicity.

Next, we aim to synthesize a novel mixture, based on the predictions of this model. As an example, suppose our goal is to develop a novel concrete mixture, that minimizes

Output	compr. str. (3)	compr. str. (7)	compr. str. (28)	cost	CO_2
rMSE [%]	7.62	2.93	4.86	0.00	0.00
rMAE [%]	22.23	13.08	16.66	0.01	0.00

Table 1: Relative mean-squared error (rMSE) and relative mean-absolute error (rMAE) on the validation set.

the amount of CO_2 emissions, while achieving a 28-day compressive strength of at least $s_{\min} = 20\,\text{MPa}$. We account for a prediction error of at most 20% on the strength. Thus, we solve (1) where $L(x, y) = y_{\text{env}}$, and $g(y) = 1.2s_{\min} - y_{S28}$. The resulting mixture is shown in fig. 1. For comparison, we show the mixture in the dataset with the smallest CO_2-equivalents satisfying the strength requirement on the right. We observe that even with the 20%, safety margin on the strength requirement, the novel mixture achieves a 8,7% decrease in CO_2-equivalent. Compared to the best mixture in the dataset that achieves the desired strength *including the 20%* safety margin (not shown in the figure), the novel mixture results in a reduction of the equivalent CO_2 emissions by 17.6%.

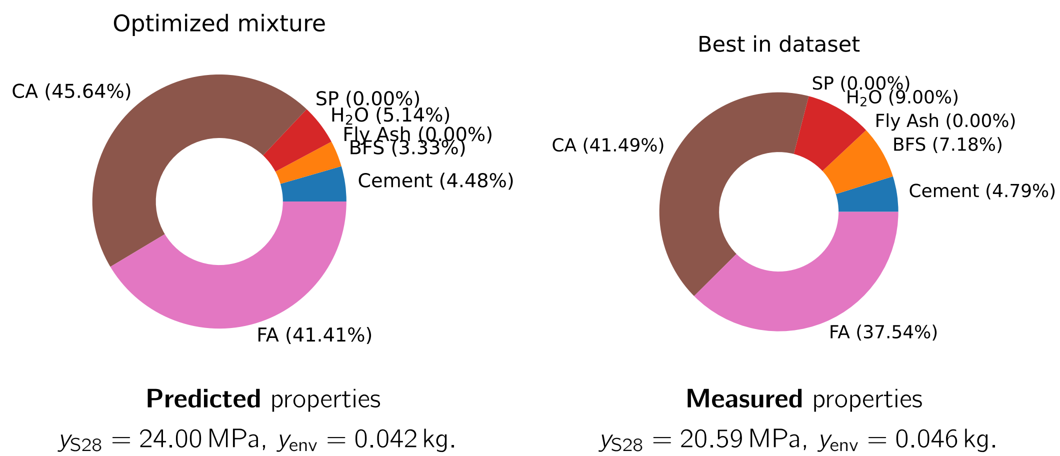

Figure 1: Comparison of the optimized mixture *(left)*, and the best mixture (according to our specifications) in the dataset *(right)*. *(CA: Coarse aggregate, FA: fine aggregate, SP: superplasticizer, BFS: Blast Furnace Slag)*

4 Conclusion

We present a simple data-driven framework for solving mixture design problems. We illustrate how certain sanity checks and other *a priori* knowledge can be easily built into simple linear models, which result in relatively accurate predictions, while simultaneously being computationally tractable for optimizing new mixtures.

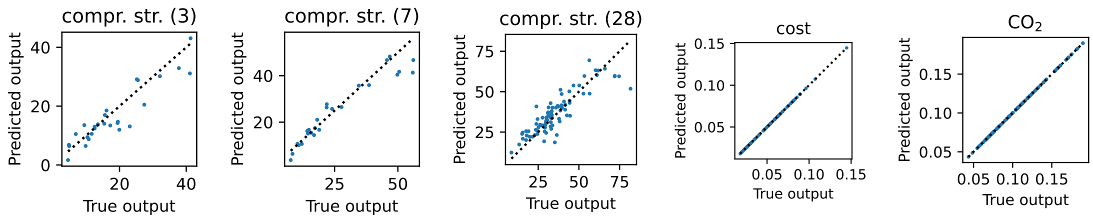

Figure 2: Prediction versus measured outputs in the validation set. For a perfect model and noiseless data, the points should lie on the diagonal drawn with a dashed line.

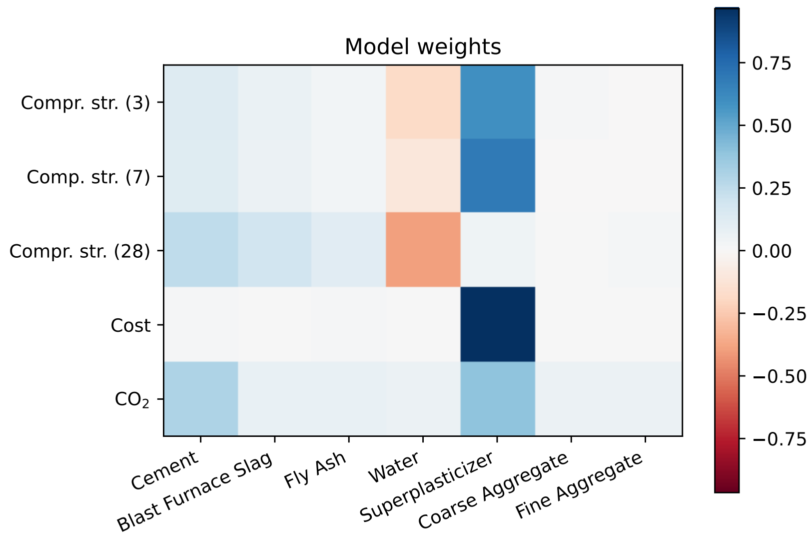

Figure 3: Weights of the model normalized row-wise.

We illustrate the methodology on a concrete mixture problem, using a publicly available dataset. We illustrate that using the method we can obtain mixtures which can be produced with around 10% less equivalent CO_2 than the best mixture in the provided dataset, while maintaining the required compressive strength.

We highlight that this case-study is a simple proof of concept, which permits many interesting extensions: using human-feedback to estimate scalarization weights for *multi-objective optimization*; integration with Design of Experiment methods to obtain the most accurate model for a given experimentation budget; more explicitly taking into account model uncertainty using robust optimization; and, of course, extensive empirical validation.

References

[1] R. T. Rockafellar, *Convex Analysis*. No. 28 in Princeton Mathematical Series, Princeton, N.J: Princeton University Press, 1970.

[2] A. Ben-Tal, L. El Ghaoui, and A. S. Nemirovskiĭ, *Robust Optimization*. Princeton Series in Applied Mathematics, Princeton: Princeton University Press, 2009.

[3] B. Hermans, A. Themelis, and P. Patrinos, "QPALM: A proximal augmented lagrangian method for nonconvex quadratic programs," *Mathematical Programming Computation*, vol. 14, pp. 497–541, Sept. 2022.

[4] B. Stellato, G. Banjac, P. Goulart, A. Bemporad, and S. Boyd, "OSQP: An operator splitting solver for quadratic programs," *Mathematical Programming Computation*, vol. 12, pp. 637–672, Dec. 2020.

[5] I.-C. Yeh, "Concrete Compressive Strength." UCI Machine Learning Repository, 1998. DOI: https://doi.org/10.24432/C5PK67.