Design of experiments for screening and optimizing factors influencing extraction of bioactive compounds from saffron.

A. Douaik1

1Research Unit on Environment and Conservation of Natural Resources, Regional Center of Rabat, National Institute of Agricultural Research (INRA), PO Box 415, Rabat, MOROCCO

Abstract

The quality of saffron is evaluated by the amount of the three main secondary metabolites which are picrocrocin, safranal, and crocin that are responsible for the bitterness, the characteristic aroma, and the red yellowish color of the most expensive spice of the world, respectively. These components should be first extracted before their quantification using different analytical platforms like chromatography and spectrometry.

Currently, there are many processes for extracting these components. The efficiency of the extraction method depends on many factors such as the solvent concentration, the extraction pressure, time and temperature, the solvent ratio, etc. Consequently, the first step in the extraction process, called screening, consists of identifying the most influencing factors by checking their preliminary significance on the efficiency of the extraction process. There are many experimental designs used at this step like the full factorial design, the fractional factorial design, and the Plackett-Burman design. These designs allow a multivariate screening, taking into account any likely interactions between the different factors, and a considerable reduction in the size of the experiment. The second step in the extraction process consists in the optimization of the most influencing factors selected at the screening step. The largely used experimental designs at this second step are the Box-Behnken design, and the Box-Wilson or central composite design.

In this research work, a short presentation of the main extraction methods of the most important saffron components will be done. This will be followed by the presentation of the most common experimental designs used at the screening step. Then, the frequently used experimental designs for optimizing the extraction process will be discussed. For both kinds of experimental designs, their general principles will be reviewed and the manner to choose the most adapted among them will be stressed.

Keywords: Box-Behnken, central composite, fractional factorial, Plackett-Burman, response surface methodology.

INTRODUCTION

Saffron, the most expensive spice of the world, is valued for its three main volatile components: crocin for the color, picrocrocin for the bitterness, and safranal for the aroma. These secondary metabolites should be extracted using various extraction methods then separated and detected using different analytical methods. The extraction procedure is affected by many factors which influence the amount of volatile components. Consequently, the statistical methodology of design of experiments (DoE) should be used in order, firstly, to screen the influencing factors to keep only the most important, and, in second step, to optimize the most important factors (Myers et al, 2016). Experimental data released by the two steps of DoE need to be processed using statistical methods to derive useful information. In what follows, a general overview of the main extraction methods used for saffron will be given followed by the main analytical methods used for separating and detecting the three main secondary metabolites of saffron. Then, the main experimental designs used for screening influencing factors will be discussed. Later, experimental designs most frequently used for optimizing extraction factors will be presented. Finally, statistical methods useful for handling experimental data will be reviewed.
EXTRACTION METHODS

The most frequently used techniques for isolation of analytes from plant material, in general, are accelerated solvent extraction, microwave-assisted extraction, Soxhlet extraction, steam distillation, supercritical-fluid extraction, and ultrasonic solvent extraction (Romanik et al, 2007). For the specific case of saffron, various extraction methods have been used for extracting and quantifying secondary metabolites (Heydari and Haghayegh, 2014). These are solvent based extraction technique, steam distillation, ultrasound-assisted extraction (UAE), membrane processes, supercritical fluid extraction (SFE), solid phase extraction (SPE), etc. There are also microextraction techniques like solid phase microextraction (SPME), stir-bar sorptive extraction (SBSE), and liquid phase microextraction (LPME), etc.

The diversity of extraction techniques implies that many different factors are affecting the absorbance measurements of the secondary metabolites. Depending on the extraction technique, the main factors could be sample amount, temperature, pressure, extraction time, CO2 flow rate, filter type, particle size, stage of extract filtration, extraction solvent volume, solvent ratio, etc.

ANALYTICAL METHODS

Whatever the extraction technique used, the next step is the separation, the detection, and the quantification of the secondary metabolites. In this way, many analytical methods are used (Alonso et al., 1996; Jalali-Heravi et al, 2009; Lage and Cantrell, 2009; Verma and Middha, 2010; Zalacain et al., 2005; Zougagh et al, 2005). They can be Ultraviolet-visible spectrometry and spectrophotometry, vibrational spectroscopy (Near-, NIR and mid-infrared, MIR), thin-layer chromatography (TLC), gas chromatography (GC), high performance liquid chromatography (HPLC), GC-Mass Spectrometry (GC-MS), LC-MS-MS, gas chromatography-flame ionisation detector (GC-FID), GC-olfactometry (GC-O), capillary electrophoresis (CE), micellar electrokinetic chromatographic (MEKC), etc.

DESIGN OF EXPERIMENTS FOR SCREENING

The screening methods allow detecting, from a set of potentially influential input variables, those which actually are in a fixed range of variation. It is a stage for identifying quickly in many of potentially influential factors (k), some factors that actually are influential in an experimental field set (f). They help to determine the "weight" of each level of each factor, and then rank them in order of importance. According to the principle of parsimony, or principle of "Pareto" or "Ockham’s razor“ (Morris, 2006; Kleinen, 2015), the actual number of active factors is low compared to the number of potential factors (k<<f). Screening experiments are referred as phase zero of a response surface study since it is the first phase of an experimental study (Myers et al, 2016). Phase 1 corresponds to the construction of a first-order linear model after screening is done. The subsequent phases involve second-order terms and the search for the optimum region. Screening is crucial since all the remaining phases rely on its results: if it is not correctly done, all subsequent experimentation may yield erroneous results.

The very simple way to construct a screening design is to change each of the factors one at a time, the other factors being kept constant in each case. This kind of design is called One-Factor-At-Time (OFAT) design (Czitrom, 1999). The OFAT approach has the major disadvantage of not giving any information about the interactions that may exist between the factors and, as such, is generally not advisable.

The most efficient screening designs use only two discrete levels for each factor corresponding to the low and high values; the levels of the factors are chosen in such a way that they span the complete factor space. There are three main experimental designs for the screening step: full factorial design, fractional factorial design, and Plackett-Burman design (Hanrahan and Lu, 2006). These designs are shortly described in what follows. These designs will be illustrated for an experiment on microwave-assisted extraction using 8 factors (Table 1): volume of extraction solvent, temperature, pressure, extraction time, solvent ratio, microwave power, plant particle size, and number of extraction vessels.

Table 1. Factors used for microwave-assisted extraction at the screening phase with their levels.

<table>
<thead>
<tr>
<th>Factors</th>
<th>Levels</th>
</tr>
</thead>
<tbody>
<tr>
<td>Name</td>
<td>Code</td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td>Volume of extraction solvent (ml)</td>
<td>$X_1$</td>
</tr>
<tr>
<td>----------------------------------</td>
<td>-------</td>
</tr>
<tr>
<td>Temperature (°C)</td>
<td>$X_2$</td>
</tr>
<tr>
<td>Extraction time (min)</td>
<td>$X_3$</td>
</tr>
<tr>
<td>Pressure (kPascal)</td>
<td>$X_4$</td>
</tr>
<tr>
<td>Solvent ratio (volume %)</td>
<td>$X_5$</td>
</tr>
<tr>
<td>Microwave power (Watt)</td>
<td>$X_6$</td>
</tr>
<tr>
<td>Plant particle size (mm)</td>
<td>$X_7$</td>
</tr>
<tr>
<td>Number of extraction vessels</td>
<td>$X_8$</td>
</tr>
</tbody>
</table>

**Full or complete factorial design (CFD)**

The full or complete factorial design (CFD) is the most commonly used one for screening experiments. It requires $2^k$ treatments, called also runs, $k$ being the number of factors. A treatment is the combination of any of the two levels of one factor with any of the two levels of the other factors. For a CFD, all the possible combinations are used in the experiment. In our example $k=8$, thus $2^8 = 256$ treatments are used.

A CFD allows estimating the linear or main effects of all the factors as well as their interactions or quadratic effects.

**Fractional factorial design (FFD)**

The major limitation for CFD is that the number of treatments increases rapidly with the number of factors. Therefore, fractional factorial designs (FFD) were developed to reduce the number of treatments even for a large number of factors. They are the most widely used designs in experimental investigations. The number of treatments is only required to be a power of two. Again, for $k$ factors, each having 2 levels, the FFD has $2^{k-p}$ treatments, $p$ being the rate of reduction of the number of treatments (Box et al, 2005). An FFD is a $1/2^p$ fraction of a $2^k$ CFD (Table 2). The minimum number of treatments must be equal or greater than $k + 1$. In our example, with $k = 8$ factors, we have to use 256 treatments for a CFD. For a FFD, if $p=1$, the number of treatments will be reduced by half and the experiment will involve $2^{8-1} = 128$ treatments; if $p=2$, this number is reduced by quarter and the experiment will involve $2^{8-2} = 64$ treatments while if $p=4$, the experiment will involve $2^{8-4} = 16$ treatments. The treatments should be carefully chosen and be a representative subset of the whole set of treatments from a CFD. FFD are very efficient and generally effective. They use many fewer treatments than CFD and can estimate main effects while two-way interaction effects can or cannot be estimated unconfounded from the main effects, depending on the design which is characterized by its resolution.

Table 2. Reduction fraction and number of treatments for FFD constructed from a CFD based on 8 factors.

<table>
<thead>
<tr>
<th>$p$</th>
<th>Fraction</th>
<th>Number of treatments in a FFD</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>1</td>
<td>256 (CFD)</td>
</tr>
<tr>
<td>1</td>
<td>2</td>
<td>128</td>
</tr>
<tr>
<td>2</td>
<td>4</td>
<td>64</td>
</tr>
<tr>
<td>3</td>
<td>8</td>
<td>32</td>
</tr>
<tr>
<td>4</td>
<td>16</td>
<td>16</td>
</tr>
</tbody>
</table>
Plackett-Burman design (PBD)

Plackett and Burman (1946) developed a family of FFD where the number of treatments is only required to be a multiple of four. The Plackett-Burman designs (PBD) are identical to the usual $2^{k-p}$ regular FFD when the number of treatments is a power of two. However, the PBD are potentially useful for reducing the number of treatments for $n = 12, 20, 24, 28, 32, \ldots$ etc. They examine up to $n-1$ factors in $n$ treatments.

An example of a PBD is given in table 3 using 7 factors (all the factors from table 1 except the last one) and 8 treatments.

<table>
<thead>
<tr>
<th>Experiments</th>
<th>Factors</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$X_1$</td>
</tr>
<tr>
<td>1</td>
<td>+</td>
</tr>
<tr>
<td>2</td>
<td>+</td>
</tr>
<tr>
<td>3</td>
<td>+</td>
</tr>
<tr>
<td>4</td>
<td>-</td>
</tr>
<tr>
<td>5</td>
<td>+</td>
</tr>
<tr>
<td>6</td>
<td>-</td>
</tr>
<tr>
<td>7</td>
<td>-</td>
</tr>
<tr>
<td>8</td>
<td>-</td>
</tr>
</tbody>
</table>

PBD provides an efficient, reliable and rapid method for screening several input factors in the minimum possible experimental sets. However, PBD provides no information on, the one hand, the interactions that may exist between factors, and secondly, the magnitude of residual variations.

DESIGN OF EXPERIMENTS FOR OPTIMIZING

When the main influencing factors on a given response variable were identified during the screening phase, these factors should be optimized using some specific experimental designs that will be discussed in this section.

For $k$ quantitative factors, which can be considered as explanatory variables $x_1, x_2, \ldots, x_k$, we call response surface the surface that corresponds to equation:

$$y = f(x_1, x_2, \ldots, x_k),$$

relating the response observed at the end of the experiment $y$ and the explanatory variables $x_1, x_2, \ldots, x_k$.

Response surfaces can be either:

- plans or hyperplans in case of equations of first degree (for $k=2$ factors):
  $$y = b_0 + b_1x_1 + b_2x_2$$

- quadratic surfaces in case of equation of second degree (for $k=2$ factors):
  $$y = b_0 + b_1x_1 + b_2x_2 + b_3x_1^2 + b_4x_2^2 + b_5x_1x_2$$

Response surface methodology (RSM) is an effective means of optimizing the process in situations where multiple variables may influence the output (Myers et al, 2016). RSM explores the relationships between several explanatory or independent variables and one or more response or dependent variables. The main idea of RSM is to use a sequence of designed experiments to obtain an optimal response, and the experiments will be more easily arranged and interpreted using this efficient design.

One of the main advantages of the RSM is that it considers the interactive effects of the factors in addition to their main effects. Different models are fitted to the experimental data. First-order model is useful for the linear behavior of the data. A second order or quadratic model should be applied if the dataset presents curvature.

There are mainly two experimental designs that are frequently used for optimizing an analytical process: the Box-Behnken design and the Box-Wilson or Central Composite design. These two designs will be illustrated on the extraction factors from Table 1 assuming that three factors were selected during the screening phase: Volume of extraction solvent ($X_1$), temperature ($X_2$), and extraction time ($X_3$).

Central composite design (CCD)
The central composite design (CCD) of Box and Wilson (1951) has the principle of associating each time an OFAT design (pluses) to a $2^k$ full (CFD) or $2^{k-p}$ fractional (FFD) factorial design (dots) in addition to a central point (star). Figure 1 gives a graphical representation in the case of three factors, for values of the variables ($X_1$, $X_2$, and $X_3$) coded in -1 and +1 regarding the factorial part of the design.

![Figure 1. Illustration of a CCD with 3 factors, each having 5 levels.](image)

In general, for $k$ factors, the CCD consists of a central point at the origin, $2k$ radial points located at the same distance $\Delta x$ of the origin, and $2^k$ factorial points located at a distance $\sqrt{k}$ from the origin. Thus, in total, there are $1 + 2k + 2^k$ experimental points. Generally, the distance $\Delta x$ is set to $4\sqrt{2k}$.

For our case study with $k = 3$, there is one central point, $2*3 = 6$ radial points located at the same distance of 1.682, and $2^3 = 8$ factorial points located at the distance of 1.732. There are, in total, 15 experimental points.

Frequently, the central point is replicated a given number of times with the aim of estimating the residual variation whereas the radial and factorial points are not replicated.

Compared to a $3^k$ CFD, CCD has the advantage of requiring less experimental points. Four our case study, a one replication of a CFD requires $3^3 = 27$ experimental points compared to only 15 experimental points for a CCD. One limitation of the CCD is that it requires always 5 different levels for each factor ($-\Delta x$, -1, 0, +1, and $+\Delta x$) instead of 3 levels in case of CFD (-1, 0, and +1).

**Box-Behnken design (BBD)**

Box and Behnken (1960) developed a design that remedies the disadvantage of a CCD returning to three levels for each of the factors. Figure 2 presents the case of three factors.

![Figure 2. Example of a BBD with 3 factors, each having 3 levels.](image)

In our case study with $k = 3$, in addition to the central point (star), the BBD consists of 12 points (dots), which are located in the middle of the 12 edges of the cube delimiting the experimental domain. These 12 points actually form three $2^2$ factorial sets, relating to the three
pairs of factors, each time at the intermediate level of the factor that does not intervene in the considered 2-factor factorial set. In total, there are 13 experimental points. All external points constitute a polyhedron with 14 faces (8 triangular faces and 6 square faces), sometimes called a cuboctahedron. Moreover, as in the case of a CCD, the central point is generally replicated.

CONCLUSIONS

Basically, there are four main steps in the analytical process for extraction of metabolites from saffron: data collection, data preprocessing, data modeling, and model validation before its use for practical applications.

Data collection needs to be done in a rigorous way using the principles of the statistical methodology of the design of experiments. There are 2 phases in the extraction process: screening to identify the most influencing factors and optimization to optimize the selected factors at screening. During each phase, there are some appropriate experimental designs.

The experimental design and the statistical analysis of its data are crucial for the optimization of the extraction process of metabolites from saffron.

Literature cited


Box GEP and Behnken DW. (1960). Some new three level designs for the study of quantitative variables. Technometrics, 2: 455 – 475.


