CrossMark

Multivariate Normal Boundary Intersection based on rotated factor scores: A multiobjective optimization method for methyl orange treatment

Fabiano Luiz Naves [a], Taynara Incerti de Paula [b], Pedro Paulo Balestrassi [b, *],
Washington Luis Moreira Braga [b], Rapinder Singh Sawhney [c], Anderson Paulo de Paiva [b]

[a] Chemical Engineering and Statistics Department, Federal University of São João Del Rei, São João Del Rei, MG, Brazil
[b] Institute of Industrial Engineering and Management, Federal University of Itajubá, Itajubá, MG, Brazil
[c] Industrial and Systems Engineering Department, The University of Tennessee at Knoxville, Knoxville, TN, USA

A B S T R A C T

This paper presents the multiobjective optimization of methyl orange treatment with ozone using Normal Boundary Intersection and response surface models of rotated principal component factor scores for the expected value $E[f(x)]$ and prediction variance $Var[f(x)]$ of dye removal (Y_1) and chemical oxygen demand removal (Y_2). The innovation and the main contribution of this paper consists of building a 2-dimensional equispaced and convex Pareto Frontier for rotated factor scores representing the original multivariate set, reducing the number of objective functions without inverting the correlation among the original responses. Furthermore, this proposal provides a practical way to generate the narrowest possible prediction confidence intervals for a desired optima using the fuzzy membership function criterion in order to select the best compromise solution between $E[f(x)]$ and $Var[f(x)]$. To illustrate the proposal's feasibility, a central composite design for the ozonation process of methyl orange solution with three factors (x_1 = pH, x_2 = air flow and x_3 = ozone dosage) was run. The optimization results showed a maximum dye removal of 94.1% $\pm$ 4.3 with a respective chemical oxygen demand removal of 88.4% $\pm$ 5.3 obtained at $x^* = [9.5; 7.1$ l.min-1; 18.4 g h^{-1}]. However, this point have presented the largest 95% prediction confidence interval. Based on the fuzzy membership of Pareto set it was possible to select the narrowest 95% confidence intervals with maximum removal rates ($Y_1 = 90.5 \pm 2.2$ and $Y_2 = 88.3 \pm 2.7$), obtained at $x^* = [7.9, 5.6$ l min^{-1}, 18.4 g h^{-1}]. Confirmation runs and comparisons among several optimization methods were done and indicated that the results fell within the respective confidence intervals for predictions, which corroborates the good adequacy of the proposal.

© 2016 Elsevier Ltd. All rights reserved.

1. Introduction

Response surface methodology (RSM) has been extensively used in the modelling and optimization of industrial an urban wastewater treatments. Examples include the wastewater treatment of livestock (Tak et al., 2015), meat industry (Thirugnanasambandham et al., 2015), tobacco wastewater (Pi et al., 2015), leather industry (Boopathy and Sekaran, 2013), textile industry (Sheydaei et al., 2014), steel industry (Anouzla et al., 2009), petroleum refinery (Shahrezaei et al., 2012) among others. The most used RSM designs in wastewater treatment are the central composite design (CCD) (Asfaram et al., 2015a, 2015b; Li et al., 2015) and the Box-Behnken design (BBD) (Nair and Ahammed, 2015). Both designs are capable of generate nonlinear objective functions like full quadratic models (Montgomery, 2009). The central composite design (CCD) is a response surface array composed by three groups of points: a factorial design with 2^k experiments (where k is the number of controllable factors), $2k$ axial points and n center points. The

---

distance between center and axial points are generally defined as $2^{(k/4)}$. Sometimes, the axial points produce a set of unfeasible experiments, since they represent extreme high or low conditions for the controllable factors (Myers and Montgomery, 2009). Then, if the chosen levels for the controllable factors lead to impracticable experiments, Box-Behnken design is more recommended since it is a design formed by combinations of factorial and center points without outside or extreme points. These designs have been used to build objective functions related to several characteristics of wastewater treatments like color removal (Tak et al., 2015), chemical oxygen demand (COD) (Nair and Ahammed, 2015), dye removal from aqueous solutions (Asfaram et al., 2015a), biological oxygen demand (BOD) (Lu et al., 2013), total organic carbon (TOC) (Arslan-Alaton et al., 2009), turbidity (Nair and Ahammed, 2015) and decolorization efficiency (DEC) (Sheydaei et al., 2014) among others. According to the objective of treatment these responses must be maximized, like in case of DEC or color removal or minimized, like in the case of COD. COD is one of the most important characteristics in wastewater treatments like is in the tanning process of leather industry (Dixit et al., 2015) which involves azo dyes (Orange II) and textile wastewater treatment based on advanced oxidation process (Asghar et al., 2015).

Box-Behnken Design (BBD) has been applied in several works. Šereš et al. (2016), for example, applied BBD in the treatment of vegetable oil refinery wastewater using alumina ceramic membrane. In it, two concave objective functions for the microfiltration process, permeate flux and COD, were fitted. Although the stationary points were not the same, no optimization routine was employed in this case. Other examples include the use of BBD in the biosorptive decolorization process by a green type sorbent (Akar et al., 2016), the ammonium nitrogen (NH$_4$–N) removal from ammoniacal waste (Kumar and Pal, 2013), a Polyaluminium chloride-based water treatment sludge (Nair and Ahammed, 2015), a livestock wastewater treatment based on electrocoagulation process (Tak et al., 2015), a meat industry wastewater by electrochemical treatment (Thirugnanasambandham et al., 2015) and in the Ca/MG/Al coagulation process of tobacco wastewater (Pi et al., 2015).

CCD has been wider used in the wastewater treatment. Sheydaei et al. (2014), for example, optimized a photo-Fenton decolorization process of Orange 29 applied to textile wastewater using CCD; Li et al. (2015) studied the photo-Fenton decolorization of Orange II. Lu et al. (2011) used this design for Photo-Fenton pretreatment of carbofuran; Wang et al. (2014), used it in the coagulation-flocculation process of tobacco slice wastewater; Studies of Asfaram et al. (2015a, 2015b, 2015c), Asfaram et al. (2015d), Dastkhoon et al. (2015), Bagheri et al. (2016), Dil et al. (2016, 2017), Asfaram et al. (2016, 2017) used CCD for optimize the removal of dyes by adsorption processes. Torrades and Garcia Montano (2014) employed CCD in the Fenton and photo-Fenton degradation of real dye wastewater; Saeed et al. (2014) studied the palm oil mil effluent wastewater treatment by fenton using CCD; the same design was employed by Boopathy and Sekaran (2013) in the leather industry wastewater treated by electrochemical process. A face centered CCD was used by Muhamad et al. (2013) for optimization of COD, NH3–N and 2.4-DCP removal from recycled paper wastewater. It can also be cited the use of CCD in a petroleum refinery wastewater treatment by photocatalytic oxidation and mineralization using TiO2 nanoparticles (Shahrezaei et al., 2012), coagulation of highly concentrated industrial grade leather dye (Khayet et al., 2011), the electrochemical treatment of dairy industry wastewater using iron electrodes (Kushwaha et al., 2010), Fenton oxidation pretreatment of wastewater sludge (Pham et al., 2010), advanced oxidation process of Terasil Red R dye using H2O2/pyridine/Cu(II) (Lim et al., 2009), disperse azo dyes by coagulation-flocculation in the steel industrial wastewater treatment (Anouzla et al., 2009), photo-fentom-like advanced oxidation process of azo dye production wastewaters (Arslan-Alaton et al., 2009) and the advanced oxidation process by Fenton's peroxidation of olive oil mil wastewater (Ahmadi et al., 2005).

Second order surface models are generally used to build objective functions of explain the relationship between input ($\mathbf{x}$) and output (y) variables. After modelling, objective functions may be used to optimize the dependent variable (y). The simplest way to optimize a process is to find the stationary point of the objective function, taking its first partial derivative. Depending on the convexity of the second order surface model, the stationary point will be a minimum, maximum or saddle point. For example, in Pi et al. (2015) the response surface model for DEC is concave and therefore, the stationary point is a maximum. The same result may be seen in the paper of Li et al. (2015), with a saddle point for DEC in the photo-fenton decolorization of Orange II. Sheydaei et al. (2014) also modelled DEC in textile wastewater and obtained a surface model that is neither convex nor concave. The same kind of surface models are observed in the adsorption ultrasound-assisted simultaneous removal of Pb^{2+} ion and malachite green (MG) dyes in the work of Dil et al. (2017). In such cases, the stationary point is a saddle point.

The convexity of any function may be determined assessing the eigenvalues of hessian matrix (a second partial derivatives of the objective function). If the eigenvalues of the hessian are all positive then the function is convex and the stationary point is a minimum. If the eigenvalues are all negatives, function is concave and the stationary point is a maximum. If the eigenvalues are simultaneously positives and negatives then the function is neither convex nor concave and the stationary point is a saddle point (Rao, 2009). It is worth mentioning that the reduced models produced when the no significative terms are removed will imply in response surface that will be neither convex nor concave, which is very common in the wastewater treatment. Examples of such models may be seen in several works like Dastkhoon et al. (2015), Asfaram et al. (2015a).

If a convex objective function needs to be maximized, it will be necessary to add a constraint to close the solution region. In this case, the solution will fall far from the center point and the prediction variance will be the largest. Therefore, when the convexity of objective functions is not compatible with the optimization direction, it will be necessary to use a constraint function like $g(\mathbf{x}) = \mathbf{x^T x} \le \rho^2$, which represents a hypersphere outlined by the CCD design (Myers and Montgomery, 2009). According to the classical theory of DOE (Design of Experiments), the variance of prediction is a minimum in the vicinity of center points (design center) and increases in the direction of axial point (extreme points in CCD designs) (Myers and Montgomery, 2009). Therefore, every time the convexity of response surface is contrary to the optimization direction, the solution of the optimization problem will be an external point with poor predictability. This means that the (1-α)% confidence interval for the optimum will be as large as possible, and the predicted value for the optimization will be unreliable. The problem increases if more than one response is considered for optimization.

If the stationary point is a saddle point, the aforementioned constraint will be always required since the response surfaces will be neither convex nor concave. Saddle points have been extensively observed in the literature, like the response surface for COD of a photocatalytic oxidation of petroleum refinery wastewater (Shahrezaei et al., 2012), COD of dairy industry wastewater (Kushwaha et al., 2010), COD in the textile wastewater treatment by iron electrode (Lim et al., 2009) and the response surface for dye removal in the leather industry wastewater (Khayet et al., 2011).

Therefore, every time the optimization direction is not compatible with the nature of stationary point, graphical solutions like contour plots will not be useful. Even if the researcher consider the use of contour plots for optimization task, the solution tends to fall far from the center points, within of infeasible regions and with the worst possible predictability.

Besides de convexity dilemma, many works surveyed in this research have presented the use of RSM for more than one objective function which is a very common practice in engineering. Dil et al. (2017) employed CCD to the simultaneous modelling of Pb^{2+} ion and malachite green (MG) dyes removal rate by adsorption. Nair and Ahammed (2015), for example, use BBD to fit objective functions for COD and turbidity. Both functions are neither convex nor concave and presented a Pearson's correlation between the model's coefficients about r = 0.997. A high level Pearson's correlation coefficient is also found between COD (neither convex nor concave) and color removal (concave) (r = 0.998); the same convexity and correlation may be verified in the work of Thirugnanasambandham et al. (2015). Lu et al. (2013) presented the response surfaces of Carbofuran removal and BOD/DOC (concave) beside DOC removal (neither convex nor concave). These three characteristics are also correlated. Correlated response surfaces for COD (neither convex nor concave) and color removal (concave) are also found in Wang et al. (2014). Torrades and García-Montaño (2014) report two concave response surfaces for COD Fenton and Photo-fenton. In the work of Saeed et al. (2014) COD (convex) and color removal (neither convex nor concave) are also positively correlated and both function must be maximized. Boopathy and Sekaran (2013) presented convex response surfaces for COD and TKN in a leather industry wastewater. Solubilization (concave) and biodegradability (convex) are two correlated functions found in Pham et al. (2010). Arslan-Alaton et al. (2009) presented response surfaces for COD, TOC and color removal, all of them are neither convex nor concave. Ahmadi et al. (2005) built four correlated response surfaces (all of them neither concave nor convex) for COD, total polyphenols, color removal and aromaticity.

The COD reduction capacity in water is a broadly used parameter for checking the efficiency of purification treatment systems (Oguz and Keskinler, 2008). Thus many researchers have used this parameter to check the degradation of organic compounds (Zhang et al., 2007). The dye degradation has been widely studied due to high turbidity and dissolved oxygen consumption caused in receiving bodies when discarded improperly. Therefore many researchers have made efforts to determine the best conditions of degradation of these compounds in treatment plants. Ge et al. (2016) assessed the degradation of methyl orange by ozone in the presence of ferrous and persulfate ions in a rotating packed bed. Zhu et al. (2015) applied a quantitative structure-activity or property relationships (QSAR) models for degradation of organic pollutants in ozonation process under acidic condition. Jin et al. (2014) presented an alternative optimization of the decolorization of Methylene Blue and Methyl Orange dye by pulsed discharged plasma in water using response surface methodology. Li et al. (2014) promoted the degradation of methyl orange by sodium persulfate activated with zero-valent zinc. Liu et al. (2014) studied the degradation of basic and acid dyes in high-voltage pulsed discharge. Li et al. (2015) analyzed the degradation effects of UV, O_3 and UV/O_3 in the degradation of methyl orange. Mohajerani et al. (2011) employed nonlinear least-square regression to analyze the combined effect of ozonation and ultrasonolysis processes to predict the azo dye degradation. Zhang et al. (2010) analyzed the methyl orange degradation by pulsed discharge in the presence of activated carbon fibers. Tasaki et al. (2009) promoted the degradation of methyl orange using short-wavelength UV irradiation

with oxygen microbubbles. Zhang et al. (2009a) presented a procedure of degradation of C.I. Acid Orange 7 by ultrasound enhanced heterogeneous Fenton-like process. Zhang et al. (2009b) studied the organic dye removal from aqueous solution by pulsed discharge on the pinhole. Oguz and Keskinler (2008) studied the removal of color and COD from synthetic textile wastewaters using O_3, PAC, H_2O_2 and HCO_3^-. Qu et al. (2007) employed the catalytic ozonation of phenolic wastewater with activated carbon fiber in a fluid bed reactor. Grabowski et al. (2007) analyzed the breakdown of methylene blue and methyl orange by pulsed corona discharge. Zhang et al. (2007) studied the effect of granular activated carbon on degradation of methyl orange when applied in combination with high-voltage pulse discharge. Zhang et al. (2006) studied the decolorization of methyl orange by ozonation in combination with ultrasonic irradiation. Chen (2000) studied the optimal decolorization process of methyl orange by ozone.

Chen (2000) studied the degradation of a methyl orange solution with the use of ozone as an oxidant and analyzing the effects and interactions of pH, reaction temperature, stirring speed and the concentration of methyl orange solution and ozone stream using a fractional factorial 2^{5-1}. As response, the decolorization of the solution during the reaction time was analyzed. The results showed that ozone flow behaved as the most significant factor in decolorization of MO solution. Also analyzing the interactions of the effects, the interaction between temperature and ozone flow were more significant. This precisely occurs due to Henry's Law, which states that the solubility of a gas in water depends on the partial gas pressure exerted on the liquid. The proportionality constant used in this law varies with the gas and the temperature. Chen (2000) compared the decolorization of methyl orange solution concentration of 40 mg/l over time due to the air flow and ozone, obtaining a sharp difference between these two inputs. Using air there was a small amount of dye degradation in solution about 2%. However, when a bubbling ozone gas (980 ml min in pH 8.0) were used, the degradation achieved values close to 99% in 20 min of reaction. It has been found in this work also at high concentration dependence of the dye in the degradation rate, and consequently the reaction rate. The use of ozone as an oxidizer in advanced oxidation processes (AOP's) has proven effective in removing textile effluents color beyond the organic load removal quantified in terms of Chemical Oxygen Demand (COD) or Total Organic Carbon (TOC) (Poznyak et al., 2007).

Ge et al. (2016) investigated the degradation of a solution of 200 mg/L of methyl orange using ozone as an oxidant in the presence of iron ions and persulfate in a packed bed reactor. The significance of pH was checked as a function of color removal from the solution, when experiments were carried out over a pH range from 1 to 11. A 85% reduction in color was observed under a pH optimum around 4.0, which is according seen that pH values and above that may occur the precipitation of iron ions present in solution. Also evaluated was the mass transfer of ozone to methyl orange solution, conditioning this limiting step of the process.

Also in the dye treatment context, Grabowski et al. (2007) studied the degradation solution of this compound in the presence of ozone as analyzed by monitoring the pH. The solution concentration was 10 mg/L and removals of about 90% at 20 min of reaction were obtained. A cost analysis was done on the basis of corona discharge. The research point out a cost of treating a 1 m^3 of methyl orange solution in a concentration of 10 g/m^3 of approximately 0.2 €/m^3 and can reach up to € 1/m^3 in industrial scales.

Methylene blue is also a dye widely used as a standard compound for investigation of the treatment parameters that can also be used for scale up in industrial processes. The industrial environment, as waste textile industries for example, has in its

composition the mixture of these dyes, which can be explained due to its small grip in dyeing fabric. Therefore the mixture of methylene blue and methyl orange solution was studied by Jin et al. (2014), using RSM to determine the optimum factor levels used (power to generate ozone gas flow and spacing between electrodes which are responsible for generating the potential difference in the equipment). Degradations about 94.5% and 80.2% of both color was observed after 30 min of treatment. The higher degradation percentages for the methylene blue are to be expected, since a chloride ion in the center of the molecule is responsible for the establishment thereof, easily removed. However, the methyl orange consists of an azo compound, which has a triple introduction of two nitrogen atoms which is more complex to be broken. The dye mixture as well as being a reality in the textile industry should be studied more in intrinsic form, analyzing how its kinetic and thermodynamic parameters can be changed according to the concentration and physical effects related to the molecular size of each of compounds known as steric effects. Another important issue to be discussed in dye degradation is the issue related to alkalinity and acidity of these compounds. Methyl Orange azo compound is considered an acid while methylene blue is a basic compound. In this context, Liu et al. (2014) studied the mixture of these compounds, classified by him as acidic and basic. The color removal over 21 min in a reactor achieved values higher than 90%, which was considered by the author as a high amount of degradation. In addition to the color removal Liu et al. (2014) also studied the reduction of total organic carbon (TOC), which had a removal of about 35% also in 21 min of reaction. The analysis due to the interference of the pH during the reaction showed that the basic pH is more efficient in the removal of color, especially in the removal of TOC. The concentration of dissolved ozone in the solution was also studied, showing a higher percentage of dissolved ozone at pH 4, being considerably reduced at pH 10. Besides the use of ozone as the oxidant, forms of implementation and catalyst compositions in addition to the use of UV radiation has been widely used in order to reduce the treatment time, coupled to reducing the cost of the process and increasing its efficiency. Using UV, O_3 and UV/O_3, Li et al. (2015) showed the treatment of Methyl orange solution, achieving a 93% color removal over 60 min at pH 9.

In the paper of Oguz and Keskinler (2008) ozone was used together with hydrogen peroxide, activated carbon and carbonate in order to keep the residence time of ozone in solution. Color removal and COD were recorded, with reductions of 99% and 95%, respectively, over 30 min of reaction. The use of activated carbon as catalyst, together with carbonate ions present in the medium, further enhances the retention of the ozone in solution through a surface effect, increasing the mass transfer of the gas, evidencing the high COD removal in 30 min. Even using a hybrid process combining UV and ozone, Tasaki et al. (2009) used a mercury vapor lamp of 8 W low pressure along with ozone dispersed in the form of microbubbles (5.79 mm) to treat methyl orange solution. In this context the removal of COD and color were respectively around 85% and 97%. As already mentioned the importance of pH control during treatment, tests were run to detect the best pH during treatment, indicating values near neutrality, i.e., pH 6.9.

Still in the context of hybrid processes, Zhang et al. (2006) used ozonation and ultrasound, seeking a greater reduction in color and COD in a shortest time. In this context it was found the greatest color removal rate with the increase of ultrasound power, following a pseudo first order kinetics. The subject of use of ultrasound for the removal of color, can be a factor which considerably increases the cost of the process, with little difference of color removal from the combination of ultrasonic and only ozone and ozone as oxidant.

During the production of ozone at high potential differences it can be formed various other oxidants, such as oxygen free radicals, H_2O_2, and hydroxyl radicals (Zhang et al., 2009b). These products generated during the production of ozone, may also act in the treatment process. Therefore, Zhang et al. (2009b) used these generated radicals for decolorization of a methyl orange solution, which evidenced a greater color removal by increasing the power of the ozone generating equipment. However, very high powers can promote the formation of undesirable compounds, such as oxides of nitrogen class that can sequester the oxygen present in the input feedstock which is very air thus lowering the ozone production (Tang et al., 2009). Following this context, power it is determined at the optimized pattern, and that any other around the value of this point may decrease the process efficiency and increase process costs.

As it will be seen in next sections, mean and variance of each characteristic of interest may be extremely correlated even as the multiples responses of wastewater treatment, like COD, DEC and TOC, for example. Most of these characteristics must be maximized while the prediction variance should be minimized. As these responses are general positively correlated it is not possible to separate them in blocks of maximization and minimization without any multivariate statistical procedure. In general, two clusters may be formed: one with the means of wastewater treatment characteristics (DOC, DEC, TOC, among others) and other with the prediction variance equations. To separate these blocks of responses and to reduce the number of clusters, Factor Analysis (FA) may be a feasible alternative. Factor analysis is a multivariate statistical procedure capable of separate groups of objects (or responses in this case) based on their correlations, enabling the creation of a new responses in terms of independent factor scores which may be extracted from the original data set using principal component analysis (PCA) (Johnson and Wichern, 2007). Taking the factor score response surfaces for block of means and prediction variances the trade-off between the minimization of variances and maximization of wastewater treatment efficiency may be conducted using the NBI method, a multiobjective optimization algorithm capable of generate a convex and equispaced Pareto Frontiers.

The aforementioned discussion was not found elsewhere and it would be very useful to embody the trade-off aspect between the optimality and predictability of the optimal conditions established for the wastewater treatments. To accomplish this objective, this paper presents a multivariate optimization entitled NBI-FA approach. To illustrate its adequacy, the proposal is applied to a treatment of Methyl Orange (MO) using ozone.

This paper is organized as follows: section 2 presents the stochastic nature of the response surfaces and how the predictability of such surfaces can be determinate; section 3 presents the basic concepts of factor analysis (FA) and the respective varimax rotation method; section 4 presents the general idea of Pareto frontier and Normal Boundary Intersection (NBI) method; section 5 describes the general framework of NBI-FA method; section 6 presents a case study of Methyl Orange (MO) degradation using ozone; section 7 presents the results, comparisons among several optimization methods, confirmation runs and a technical discussion.

2. The stochastic nature of response surfaces

Response Surface Methodology (RSM) is a collection of mathematical and statistical tools used to model and analyze problems whose desired responses are influenced by many variables (Montgomery, 2009). In general, the relationship between dependent and independent variables is unknown, then a reasonable approximation for the real relationship between the response (Y) and the set of independent variables ($\mathbf{x}$) can be obtained using a higher order polynomial, like the second-order model, in some region of interest if there is significant curvature in the system.

Such model can be described as:

$$Y(\mathbf{x}) = \beta_0 + \sum_{i=1}^{k} \beta_i x_i + \sum_{i=1}^{k} \beta_{ii} x_i^2 + \sum \sum_{i<j} \beta_{ij} x_i x_j + \varepsilon \tag{1}$$

where β is the polynomial coefficient, k is the number of factors and ε is the error term.

Montgomery (2009) does not consider that a specific polynomial model approximates a real model for the whole experimental space covered for the independent variables, but for a specific region, however, the approximation is usually efficient. The Ordinary Least Squares (OLS) method is used to estimate the parameters (β). The residuals set is than obtained as:

$$L = \sum_{i=1}^{n} \varepsilon_i^2 = \sum_{i=1}^{n} \left(Y_i - \beta_0 + \sum_{i=1}^{k} \beta_i x_i + \sum_{i=1}^{k} \beta_{ii} x_i^2 + \sum \sum_{i<j} \beta_{ij} x_i x_j \right)^2 \tag{2}$$

In matrix form, $\mathbf{Y} = \beta\mathbf{X} + \varepsilon$ and $\mathbf{L} = \mathbf{y^T y} - \beta^T \mathbf{X^T y} - \mathbf{y^T X}\beta + \beta^T \mathbf{X^T X}\beta$. Taking the derivative of $\mathbf{L}$ in terms of the vector of coefficients β, it can obtain the estimates $\widehat{\beta}$ of real coefficients. This method is called Ordinary Least Squares (OLS) and can be expressed as follows:

$$\frac{\partial \mathbf{L}}{\partial \beta} = \frac{\partial \left(\mathbf{y^T y} - \beta^T \mathbf{X^T y} - \mathbf{y^T X}\beta + \beta^T \mathbf{X^T X}\beta \right)}{\partial \beta}$$

$$= \frac{\partial \left(\mathbf{y^T y} - 2\beta^T \mathbf{X^T y} + \beta^T \mathbf{X^T X}\beta \right)}{\partial \beta} = 0 \tag{3}$$

$$\frac{\partial \mathbf{L}}{\partial \beta} = -2\mathbf{X^T y} + 2\mathbf{X^T X}\widehat{\beta} = 0 \therefore \widehat{\beta} = \left(\mathbf{X^T X} \right)^{-1} \mathbf{X^T y} \tag{4}$$

Since $\widehat{\beta}$ depends upon the data, such estimated coefficients can be considered random variables with expected value $E(\beta) = \widehat{\beta}$ and $Cov(\widehat{\beta}) = \sigma^2 (\mathbf{X^T X})^{-1}$, where $\mathbf{X}$ is design matrix and σ^2 is mean square error. It can be shown that:

$$Cov\left(\widehat{\beta}\right) = E\left\{ \left[\widehat{\beta} - E\left(\widehat{\beta}\right)\right]^T \left[\widehat{\beta} - E\left(\widehat{\beta}\right)\right] \right\}$$

$$= E\left[\left(\widehat{\beta} - \beta\right)\left(\widehat{\beta} - \beta\right) \right] Cov\left(\widehat{\beta}\right)$$

$$= E\left\{ \left[\left(\mathbf{X^T X}\right)^{-1}\left(\mathbf{X^T}\varepsilon\right) \right]\left[\left(\mathbf{X^T X}\right)^{-1}\left(\mathbf{X^T}\varepsilon\right) \right] \right\}$$

$$= E\left\{ \left[\left(\mathbf{X^T X}\right)^{-1}\left(\mathbf{X^T}\varepsilon\right) \right]\left[\left(\varepsilon^T \mathbf{X}\right)\left(\mathbf{X^T X}\right)^{-1} \right] \right\} Cov\left(\widehat{\beta}\right)$$

$$= E\left\{ \left[\left(\mathbf{X^T X}\right)^{-1}\left(\mathbf{X^T X}\right) \right]\left[\left(\varepsilon^T \varepsilon\right)\left(\mathbf{X^T X}\right)^{-1} \right] \right\}$$

$$= E\left(\varepsilon^T \varepsilon\right)\left(\mathbf{X^T X}\right)^{-1} = \sigma^2\left(\mathbf{X^T X}\right)^{-1} \tag{5}$$

The covariance matrix of the estimated coefficients $[Cov(\widehat{\beta})]$ it is generally used to build the (1-α)% confidence intervals for new observations $CI_{(1-\alpha)} = \widehat{y}_0 \pm t_{\alpha/2,n-p}\sqrt{\widehat{\sigma}^2[1 + \mathbf{x}_0^T (\mathbf{X^T X})^{-1}\mathbf{x}_0]}$. It is possible to show that:

$$E\left[\widehat{Y}(\mathbf{x}) \big| \mathbf{x}_0\right] = \mathbf{x}_0^{(m)\mathbf{T}}\left[\left(\mathbf{X^T X}\right)^{-1}\left(\mathbf{X^T Y}\right) \right] \quad \text{and} \quad Var\left[\widehat{Y}(\mathbf{x}) \big| \mathbf{x}_0\right]$$

$$= \sigma^2\left[\mathbf{x}_0^{(m)\mathbf{T}}\left(\mathbf{X^T X}\right)^{-1}\mathbf{x}_0^{(m)} \right] \tag{6}$$

For $k = 2$, for example, $\mathbf{x}_0^{(m)\mathbf{T}} = \mathbf{x}_0^{(2)\mathbf{T}} = [1, x_1, x_2, x_1^2, x_2^2, x_1 x_2]$.

It is straightforward that the precision of the predictions are entirely related to the choice of a new point. In the optimization case, this means that the optimum $\mathbf{x}_0$ can present a poor predictability depending on its location on the experimental space. So, if a precise estimate of the optimum is desired, the value of $Var[\widehat{Y}(\mathbf{x})]$ must be minimized. In the algebraic form, $Var[\widehat{Y}(\mathbf{x})]$ can be written as:

$$Var\left[\widehat{Y}(\mathbf{x})\right] = \sum_{i=1}^{n} \left\{ \frac{\partial\left[\widehat{Y}(\mathbf{x})\right]}{\partial\beta_i} \right\}^2_{\widehat{\beta}_i} \sigma^2_{\widehat{\beta}_i}$$

$$+ 2\left\{ \sum_{i=1}^{n-1} \sum_{j=i+1}^{n} \left\{ \frac{\partial\left[\widehat{Y}(\mathbf{x})\right]}{\partial\beta_i} \right\}_{\widehat{\beta}_i} \left\{ \frac{\partial\left[\widehat{Y}(\mathbf{x})\right]}{\partial\beta_j} \right\}_{\widehat{\beta}_j} \times r_{\beta_i \beta_j} \right.$$

$$\left. \times \sqrt{\sigma^2_{\beta_i} \sigma^2_{\beta_j}} \right\} \tag{7}$$

Equations of mean and variance are generated from columns of a design of experiment (DOE) matrix. Since the columns in DOE arrays are independents, equation (7) will be simplified, such as:

$$Var\left[\widehat{Y}(\mathbf{x})\right] = \sum_{i=1}^{n} \left\{ \frac{\partial\left[\widehat{Y}(\mathbf{x})\right]}{\partial\beta_i} \right\}^2_{\widehat{\beta}_i} \sigma^2_{\widehat{\beta}_i} \tag{8}$$

In Eq. (8), $\sigma^2_{\widehat{\beta}_i}$ is obtained from the main diagonal of $[Cov(\widehat{\beta})]$. For a response surface with $k = 3$ factors, after solving the partial derivatives, $Var[\widehat{Y}(\mathbf{x})]$ may be written as:

$$Var\left[\widehat{Y}(\mathbf{x})\right] = \sigma^2_{\beta_0} + x_1^2\sigma^2_{\beta_1} + x_2^2\sigma^2_{\beta_2} + x_3^2\sigma^2_{\beta_3} + x_1^4\sigma^2_{\beta_{11}} + x_2^4\sigma^2_{\beta_{22}}$$

$$+ x_3^4\sigma^2_{\beta_{33}} + x_1^2 x_2^2\sigma^2_{\beta_{12}} + x_1^2 x_3^2\sigma^2_{\beta_{13}} + x_2^2 x_3^2\sigma^2_{\beta_{23}} \tag{9}$$

Since $Var[\widehat{Y}(\mathbf{x})]$ is convex, the minimum value of predicted variance will occur at the center of the experimental design, near center points. However, many times the center point will not be an adequate value for the optimum of $E[\widehat{Y}(\mathbf{x})]$. Then, to find out an optimum that meets the minimization of $Var[\widehat{Y}(\mathbf{x})]$ and the optimization (maximization/minimization) of $E[\widehat{Y}(\mathbf{x})]$ it is necessary to use a trade-off strategy based on bi-objective optimization. Several methods could be used to achieve such trade-off. In this paper, it is suggested to focus the stochastic objective function formed by $E[\widehat{Y}(\mathbf{x})]$ and $Var[\widehat{Y}(\mathbf{x})]$ using NBI. So, this problem leads to this paper's first proposition. Fig. 1 illustrates the trade-off between minimal prediction variances and maximum efficiency of wastewater treatment. As it can be seen in Fig. 1, there is a conflict between these two objective functions that may be minimized using a multiobjective optimization based on Pareto frontiers. One efficient option is the NBI method which will be presented in the next section.
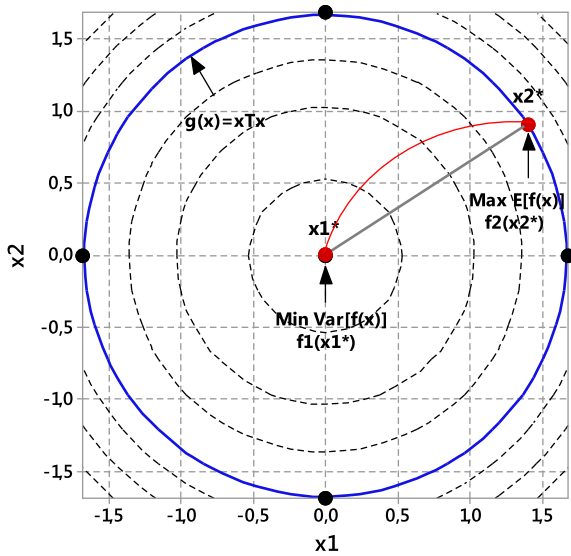
Fig. 1. Schematic illustration of NBI for E[f(x)] versus Var [f(x)].

3. Normal Boundary Intersection

Normal Boundary Intersection (NBI) (Das and Dennis, 1998) is a method developed to compensate the shortcomings attributed to the method of weighted sums as its inability to find a uniform spread of Pareto optimal solutions, even if a uniform spread of weight vectors are used. Besides the lack of uniformity among the Pareto points in the frontier, if the Pareto set is non convex, the Pareto points on the concave parts of the trade-off surface will be missed. Pareto Frontier has been used as an important element in the decision making process. Yan et al. (2016) built a Pareto frontier between energy consumption and thermal efficiency comparing the Fitness Sharing Genetic Algorithm (FSGA) and the traditional GA; Jia et al. (2006) used an Elitist Non-dominated Sorting Genetic Algorithm (NSGA) to build Pareto frontiers for environmental versus economic objective functions. Although the objectives have been properly achieved, it is observed that the Pareto frontiers are not equispaced.

The use of Pareto frontiers with response surfaces has several practical advantages since allows the researcher to establish a priori and a posteriori solutions. In other words, if the quality of Pareto solutions is related to the weights, such relationship may be modelled and also optimized, leading to the discovery of better solutions in the space mapped by the Pareto Frontier. Comparing to the use of Grey Relational Analysis (GRA) using Taguchi designs in bi-objective problems (Jozić et al., 2015) and multiobjective problems (Yan and Li, 2013) significant differences may be observed. For example, GRA is measured (and calculated) only in the design points, without modelling or optimization, and the decision making process is done only choosing the best GRA index. Besides, the importance of each response before the transformation to the index is fixed which, in turns, does not allow the exploitation of other possibilities.

Among the several multiobjective optimization methods that are capable of building Pareto frontiers, Normal Boundary Intersection (NBI) stands as one of the most promising. NBI has been recently used for multiobjective optimization of several engineering applications like in machining process with control and noise variables (Brito et al., 2014), multivariate robust parameter design (Lopes et al., 2016), dry end milling process (Costa et al., 2016), environmental and economic hydrothermal self-scheduling (Ahmadi et al., 2015a), planning of generation and transmission

expansion (Mavalizadeh et al., 2015), economic emission dispatch (Ahmadi et al., 2015b), hydrothermal scheduling (Ahmadi et al., 2015c), resource scheduling of renewable energy based on micro grids (Izadbakhsh et al., 2015), multiobjective decision making framework for electricity retailer in energy markets (Charwand et al., 2015) among others.

The original formulation of NBI, which is graphically illustrated for a bi-objective case in Fig. 2, can be mathematically written as:

$$
\begin{aligned}
&\underset{(\mathbf{x},t)}{Max\, t}\\
&S.t: \quad \overline{\boldsymbol{\Phi}}\boldsymbol{\beta} + t\hat{\mathbf{n}} = \overline{\mathbf{F}}(\mathbf{x})\\
&\qquad \mathbf{x}\in\varOmega\\
&\qquad g_j(x) \leq 0\\
&\qquad h_j(x) = 0
\end{aligned} \tag{10}
$$

In Eq. (10), t is a scalar perpendicular to the utopia line; $\boldsymbol{\Phi}$ is the payoff matrix obtained by a calculation of the individual minimum of each objective function; $\overline{\boldsymbol{\Phi}}$ is the scaled payoff, $\boldsymbol{\beta}$ is the vector of weights which represents different points along the utopia line; $\overline{\mathbf{F}}(\mathbf{x})$ is the vector of scaled objective functions and $\hat{\mathbf{n}}$ is a quasi-normal vector.

The solution that minimizes the ith objective function $f_i(x)$ will be denoted by $f_i^*(x_i^*)$. Otherwise, $f_i(x_i^*)$ is obtained when the individual optimal solution x_i^* is substituted in the objective functions. In the payoff matrix $\boldsymbol{\Phi}$ and in the scaled payoff $\overline{\boldsymbol{\Phi}}$ the ith row includes the maximum and minimum values of the functions $f_i(x)$ that represent, respectively, their upper and lower limits. This values are used to normalize the objective space, mainly by writing it in terms of different scales or units.

The individual minimum vector $f^U = [f_1^*(x_1^*), ..., f_i^*(x_i^*), ..., f_m^*(x_m^*)]^T$, is known as Utopia point. The Utopia point is a specific point, generally outside of the feasible region, that corresponds to all objectives simultaneously being at their best possible values. Otherwise, considering the vector with the worst (maximum) values of each objective function $f^N = [f_1^N, ..., f_i^N, ..., f_m^N]^T$, it is obtained Nadir point. The two anchor points connected by the Utopia line are obtained when the ith objective is minimized independently, while f_i^* represents the individual minima of the ith objective. The payoff matrices are given as:

$$
\boldsymbol{\Phi} = \begin{bmatrix}
f_1^*(x_1^*) & \cdots & f_1(x_i^*) & \cdots & f_1(x_m^*)\\
\vdots & \ddots & \vdots & & \vdots\\
f_i(x_1^*) & \cdots & f_i^*(x_i^*) & \cdots & f_i^*(x_m^*)\\
\vdots & & \vdots & \ddots & \vdots\\
f_m(x_1^*) & \cdots & f_m(x_i^*) & \cdots & f_m^*(x_m^*)
\end{bmatrix} \Rightarrow \overline{\boldsymbol{\Phi}}
$$

$$
= \begin{bmatrix}
\overline{f}_1^*(x_1^*) & \cdots & \overline{f}_1(x_i^*) & \cdots & \overline{f}_1(x_m^*)\\
\vdots & \ddots & \vdots & & \vdots\\
\overline{f}_i(x_1^*) & \cdots & \overline{f}_i^*(x_i^*) & \cdots & \overline{f}_i^*(x_m^*)\\
\vdots & & \vdots & \ddots & \vdots\\
\overline{f}_m(x_1^*) & \cdots & \overline{f}_m(x_i^*) & \cdots & \overline{f}_m^*(x_m^*)
\end{bmatrix} \tag{11}
$$

Where: $\overline{f}_i(\mathbf{x}) = \left[\dfrac{f_i(\mathbf{x}) - f_i^U}{f_i^N - f_i^U}\right] = \left[\dfrac{f_i(\mathbf{x}) - f_i^I}{f_i^{MAX} - f_i^I}\right]$

Normal Boundary Intersection (NBI) can be understood as a perpendicular line to the utopia line (or the convex hull of individual minima - CHIM) in a point that is so far from the CHIM. Normal line is defined as $\vec{\mathbf{r}}(t) = [x_0 \quad y_0 \quad z_0]^T + t \times \vec{\nabla}f[x_0 \quad y_0 \quad z_0]^T$, where t is a scalar. In other words, it is necessary just one point on the plane or surface and a specific direction vector to its construction. Considering that the t is the distance
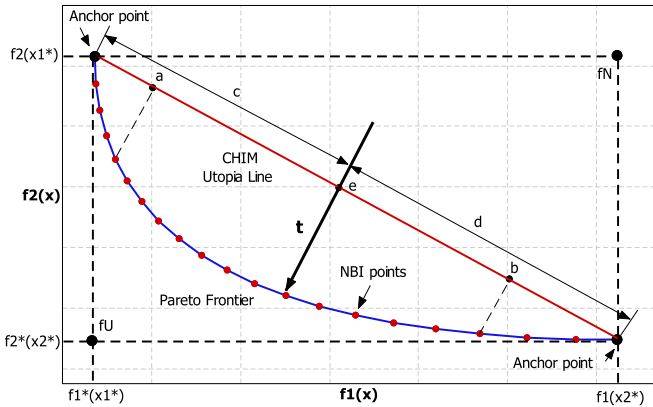
Fig. 2. Pareto Frontier for bi-objective problems obtained with NBI (Costa et al., 2016).

between $\mathbf{P_0}$ (any point in CHIM) and a point $\mathbf{P_{max}}$ (a point in the Pareto Frontier or "Boundary"), when t is a maximum, the normal line will intercept the boundary keeping the orthogonality. This condition will be repeated for any point in the CHIM represented by the vector or weights. This is essentially the nature of NBI method (Costa et al., 2016).

The classical formulation of NBI can be rewritten for the class of bi-objective problems as:

$$\begin{cases} \underset{\mathbf{x}}{Min} \quad F(\mathbf{x}) = \bar{f}_1(x) \\ St. : \quad \bar{f}_1(\mathbf{x}) - \bar{f}_2(\mathbf{x}) + 2\beta_1 - 1 = 0 \\ \quad \mathbf{x} \in \Omega \\ \quad g_j(\mathbf{x}) \leq 0 \\ \quad h_{j+1}(x) = 0 \end{cases} \tag{12}$$

The Pareto set obtained with NBI provides several optimal and feasible solutions, however, from different perspectives, some solutions may be more appropriate than others. Thereby, the choice of the best solution of the frontier may be done using some performance index, whose value can help the researcher to select the most appropriate optimal. Several indexes and algorithms have been recently proposed to this task, like TOPSIS (Technique for Order Preference by Similarity to Ideal Solution) (Ahmadi et al., 2015b), Entropy and Global Percentage Error (GPE) (Rocha et al., 2015) and Fuzzy Decision Maker (DM) (Ahmadi et al., 2015a) which will be employed in this paper. Basically, Fuzzy DM works calculating a membership function to every objective function value represented in the Pareto set, and determining the relative distance from each objective to the ideal solution. In other words, larger fuzzy memberships indicates the optima is near to the Utopia and far from the pseudo Nadir point.

For the response surfaces whose objective is the minimization, Ahmadi et al. (2015a) recommends that the individual fuzzy membership may be calculated as:

$$\mu_n^r = \begin{cases} 1 & f_n^r \leq f_n^U \\ \dfrac{f_n^{SN} - f_n^r}{f_n^{SN} - f_n^U} & f_n^U \leq f_n^r \leq f_n^{SN} \\ 0 & f_n^r \geq f_n^{SN} \end{cases} \tag{13}$$

where f_n^r is function value for a specific weigh in the Pareto frontier, f_n^U is the Utopia point and f_n^{SN} is the pseudo Nadir point.

For response surfaces that must be maximized, the individual fuzzy membership may be calculated as:

$$\mu_n^r = \begin{cases} 0 & f_n^r \leq f_n^{SN} \\ \dfrac{f_n^r - f_n^{SN}}{f_n^U - f_n^{SN}} & f_n^{SN} \leq f_n^r \leq f_n^U \\ 1 & f_n^r \leq f_n^U \end{cases} \tag{14}$$

Based on these equations, the total membership index can be stated as:

$$\mu^r = \dfrac{\sum_{n=1}^{p} w_n \mu_n^r}{\sum_{n=1}^{p} w_n} \tag{15}$$

where w_n is the relative significance of each objective function. Generally, this weight is arbitrary and may be chosen according to the researcher desire. In the present paper, since the considered objective functions are response surfaces of rotated factor scores, the values of w_n will be fixed as eigenvalues associated to each factor score.

A considerable drawback of NBI is the large number of sub-problems involved in the computation of Pareto frontier. According to Das and Dennis (1998) the number of NBI grid is equal to:

$$N_{sub} = \begin{pmatrix} n + p - 1 \\ p \end{pmatrix} \tag{16}$$

where n is the number of objective functions and p is the inverse of increment used in the Pareto grid.

For example, suppose that there are four objective functions and the Pareto set may be built with increments of 5%. Following Eq. (16), the number of subproblems is equal to 1771. Adopting increments of 10%, the number of subproblems decreases to 286. However, if some dimensionality reduction strategy is employed allowing the use of just two objective functions, with increments of 5%, only 21 subproblems should be solved. According to Čuček et al. (2014) if several objective function are simultaneously considered, the time spent in obtaining the entire solution space increases, making difficulty the visualization and interpretation of solution of the objective and solution spaces. Thereby, a dimensionality reduction technique may be important in order to facilitate the comprehension of solution space (Čuček et al., 2014). Among the available reduction dimensionality techniques it can be cited the researches of Costa et al. (2016), Lopes et al. (2016), Brito et al. (2014) that used Principal Component Analysis (PCA). Differently, this work proposes the use of Factor Analysis (FA) as a way to deal with several correlated objective functions.

4. Factor analysis

Factor analysis (FA) is a multivariate statistical technique capable of describing in a few underlying but unobservable factors, the covariance relationships among many variables. Supposing that the variables can be grouped by their correlations, all the variables belonging to a specific group will be highly correlated among themselves, but will have very small correlation with variables of different groups (Johnson and Wichern, 2007). This technique can be considered an extension of principal component analysis.

A factor model postulates that an observable random vector $\mathbf{X}$, with p components, mean vector $\boldsymbol{\mu}$ and variance-covariance matrix $\boldsymbol{\Sigma}$, is linearly dependent upon a few unobservable random variables $F_1, F_2, \ldots, F_m$, called common factors, and p additional sources of variation called errors (or specific factors), such as:

$$\underset{(p\times1)}{\mathbf{X}} - \underset{}{\boldsymbol{\mu}} = \underset{(p\times m)}{\mathbf{L}} \; \underset{(m\times1)}{\mathbf{F}} + \underset{(p\times1)}{\boldsymbol{\varepsilon}} \tag{17}$$

In Eq. (17), $\mathbf{L}$ is the matrix of factor loadings that can be calculated following the spectral decomposition of the variance-covariance matrix, such as:

$$\boldsymbol{\Sigma} = \lambda_1\mathbf{e}_1\mathbf{e}_1^T + \lambda_2\mathbf{e}_2\mathbf{e}_2^{\mathbf{T}} + \cdots + \lambda_p\mathbf{e}_p\mathbf{e}_p^{\mathbf{T}}$$

$$= \begin{bmatrix} \sqrt{\lambda_1}\mathbf{e}_1 & \sqrt{\lambda_2}\mathbf{e}_2 & \cdots & \sqrt{\lambda_p}\mathbf{e}_p \end{bmatrix} \begin{bmatrix} \sqrt{\lambda_1}\mathbf{e}_1^T \\ \sqrt{\lambda_2}\mathbf{e}_2^{\mathbf{T}} \\ \vdots \\ \sqrt{\lambda_p}\mathbf{e}_p^{\mathbf{T}} \end{bmatrix} = \mathbf{L}\mathbf{L}^{\mathbf{T}} \tag{18}$$

The factor loadings are then the eigenvectors of variance-covariance matrix $\boldsymbol{\Sigma}$, scaled by $\sqrt{\lambda_j}$ factor. Since $\boldsymbol{\Sigma} = E(\mathbf{X}-\boldsymbol{\mu})(\mathbf{X}-\boldsymbol{\mu})^{\mathbf{T}}$, then:

$$\begin{aligned} \boldsymbol{\Sigma} &= E\left[(\mathbf{X}-\boldsymbol{\mu})(\mathbf{X}-\boldsymbol{\mu})^{\mathbf{T}}\right] = E\left[(\mathbf{L}\mathbf{F}+\boldsymbol{\varepsilon})(\mathbf{L}\mathbf{F}+\boldsymbol{\varepsilon})^{\mathbf{T}}\right] \\ &= E\left[(\mathbf{L}\mathbf{F}+\boldsymbol{\varepsilon})\left(\mathbf{L}^{\mathbf{T}}\mathbf{F}^{\mathbf{T}}+\boldsymbol{\varepsilon}^{\mathbf{T}}\right)\right]\boldsymbol{\Sigma} \\ &= E\left[\mathbf{L}\mathbf{F}\left(\mathbf{L}^{\mathbf{T}}\mathbf{F}^{\mathbf{T}}\right) + \boldsymbol{\varepsilon}\left(\mathbf{L}^{\mathbf{T}}\mathbf{F}^{\mathbf{T}}\right) + (\mathbf{L}\mathbf{F})\boldsymbol{\varepsilon}^{\mathbf{T}} + \boldsymbol{\varepsilon}\boldsymbol{\varepsilon}^{\mathbf{T}}\right]\boldsymbol{\Sigma} \\ &= \mathbf{L}\underbrace{E\left(\mathbf{F}\mathbf{F}^{\mathbf{T}}\right)}_{\mathbf{I}}\mathbf{L}^{\mathbf{T}} + \underbrace{E\left(\boldsymbol{\varepsilon}\mathbf{F}^{\mathbf{T}}\right)}_{0}\mathbf{L}^{\mathbf{T}} + \mathbf{L}\underbrace{E\left(\mathbf{F}\boldsymbol{\varepsilon}^{\mathbf{T}}\right)}_{0} + \underbrace{E\left(\boldsymbol{\varepsilon}\boldsymbol{\varepsilon}^{\mathbf{T}}\right)}_{\Psi} = \mathbf{L}\mathbf{L}^{\mathbf{T}} + \boldsymbol{\Psi} \end{aligned} \tag{19}$$

where $\boldsymbol{\Psi}$ is a diagonal matrix formed by the specific variances ψ_i, such as $\psi_i = \sigma_i^2 - h_i^2$. The ith communality h_i^2 is the sum of squares of the loadings of the ith variable on the m common factors. Like in the theory of principal components, the original data set can also be represent by uncorrelated factors called "Factor scores" only considering $\mathbf{Z}$, the matrix of standardized values of $\mathbf{X}$, and the matrix of loadings $\mathbf{L}$, such as:

$$\mathbf{F} = \mathbf{Z}\left[\mathbf{L}\left(\mathbf{L}^{\mathbf{T}}\mathbf{L}\right)^{-1}\right] \tag{20}$$

The factor loadings can be rotated using an orthogonal transformation until a simpler structure is achieved, i.e., the factor rotation can lead to a separation of groups that is easier of interpretation. In this case, $\mathbf{L}^* = \mathbf{L}\mathbf{T}$, where $\mathbf{T}\mathbf{T}^{\mathbf{T}} = \mathbf{T}^{\mathbf{T}}\mathbf{T} = \mathbf{I}$. This transformation does not alter the communalities and specific variances. The most effective rotation procedure was suggested by Kaiser (Johnson and Wichern, 2007) and is known as *Varimax criterion*. In this approach, the rotated coefficients are scaled by the square root of the communalities, producing $\tilde{\ell}_{ij}^* = \tilde{\ell}_{ij}^*/\widehat{h}_i$. Then, the *Varimax* procedure selects the orthogonal transformation that:

$$MaxV = \frac{1}{p}\sum_{j=1}^{m}\left[\sum_{i=1}^{p}\tilde{\ell}_{ij}^{*4} - \left(\sum_{i=1}^{p}\tilde{\ell}_{ij}^{*2}\right)^2\bigg/p\right] \tag{21}$$

The aforementioned statistical and optimization methods (RSM, FA and NBI) may be combined in a new algorithm capable of solving the trade-off problem between expected value and prediction variance for optimal condition for any optimization process, like wastewater treatments. According to the number of objective functions, NBI-FA method may be structured in the following propositions:

Proposition 1. Suppose that NBI may be applied to solve the simultaneous optimization of the expected value of Y, $[\widehat{Y}(\mathbf{x})]$ and the predicted variance of $[\widehat{Y}(\mathbf{x})]$, $Var[\widehat{Y}(\mathbf{x})]$. Consider that two specific weights β_1 and β_2, such that $\beta_1 + \beta_2 = 1$ and $\beta_1 > 0$, $\beta_2 > 0$, can be arbitrarily chosen to express the degree of importance of each one of the two objective functions. So, combining them into a weighted sum, it may write:

$$\underset{\mathbf{x}\in\Omega}{Min}\, MSE = \beta_1 E\left[\widehat{Y}(\mathbf{x})\right] + \beta_2\left\{\sum_{i=1}^{n}\left\{\frac{\partial\left[\widehat{Y}(\mathbf{x})\right]}{\partial\beta_i}\right\}_{\widehat{\beta}_i}^2 \sigma_{\beta_i}^2\right\} \tag{22}$$

If these two objective function are scaled by their respective Utopia $f_{(\bullet)}^I$ and Nadir $f_{(\bullet)}^N$ points, such as $\overline{f}_1(\mathbf{x}) = \overline{f}_{(\mu)}(\mathbf{x})$ and $\overline{f}_2(\mathbf{x}) = \overline{f}_{(\sigma^2)}(\mathbf{x})$, and if an experimental space constraint $g_2(\mathbf{x}) = \mathbf{x}^{\mathbf{T}}\mathbf{x} \leq \rho^2$ is added to the problem, then a NBI formulation for solving the stochastic problem attributed to response surfaces will be written as:

$$\begin{aligned} Min\, \overline{f}_{(\mu)}(\mathbf{x}) &= \left[\frac{f_{(\mu)}(\mathbf{x}) - f_{(\mu)}^U}{f_{(\mu)}^N - f_{(\mu)}^U}\right] s.t.: \quad g_{1(\mu,\sigma^2)}(\mathbf{x}) \\ &= \left[\frac{f_{(\mu)}(\mathbf{x}) - f_{(\mu)}^U}{f_{(\mu)}^N - f_{(\mu)}^U}\right] - \left[\frac{f_{(\sigma^2)}(\mathbf{x}) - f_{(\sigma^2)}^U}{f_{(\sigma^2)}^N - f_{(\sigma^2)}^U}\right] + 2\beta_i - 1 \\ &= 0 \quad g_2(\mathbf{x}) = \mathbf{x}^{\mathbf{T}}\mathbf{x} \leq \rho^2 \quad 0 \leq w_i \leq 1 \end{aligned} \tag{23}$$

where:

$$\begin{aligned} f_\mu(\mathbf{x}) &= E\left[\widehat{Y}(\mathbf{x})\right], f_{\sigma^2}(\mathbf{x}) = Var\left[\widehat{Y}(\mathbf{x})\right], \overline{f}_{(\mu)}(x) \\ &= \frac{f_{(\mu)}(x) - f_{(\mu)}^U}{f_{(\mu)}^N - f_{(\mu)}^U} \text{ and } \overline{f}_{(\sigma^2)}(x) = \frac{f_{(\sigma^2)}(x) - f_{(\sigma^2)}^U}{f_{(\sigma^2)}^N - f_{(\sigma^2)}^U} \end{aligned}$$

Fig. 3 presents a schematic view of how NBI method can be used to solve the problem (see Fig. 4).

Suppose now that there are response surfaces representing each one of $\mathbf{p}$ correlated objective functions $\widehat{Y}_p(\mathbf{x})$, that can be deployed in two blocks of objective functions: means, $f_{p_\mu}(\mathbf{x}) = E[\widehat{Y}_p(\mathbf{x})]$ and variances, $f_{\sigma_p^2}(\mathbf{x}) = Var[\widehat{Y}_p(\mathbf{x})]$. As mentioned earlier, it can be noted that since $\mathbf{x}_0^{(m)\mathbf{T}}$ and $\mathbf{X}$ are the same in both models, $E[\widehat{Y}(\mathbf{x})]$ and $Var[\widehat{Y}(\mathbf{x})]$ will be highly correlated. Suppose it is necessary to simultaneously maximize $E[\widehat{Y}(\mathbf{x})]$ and minimize $Var[\widehat{Y}(\mathbf{x})]$ but these functions exhibits a highly positive correlation. How is it possible to build a Pareto Frontier for these objective functions if axes are dependents?

Proposition 2. If the objective functions are positively correlated with different optimization directions, than it is necessary to project a new system of axes that are independent. This task could be accomplished using Factor Analysis with scores extracted by principal component analysis (PCA) and applying a varimax rotation to obtain the independent axes with equivalent weights. The original responses will be replaced by the factor scores of the original ones. Then the new expected value and variance of the original models will be written as:
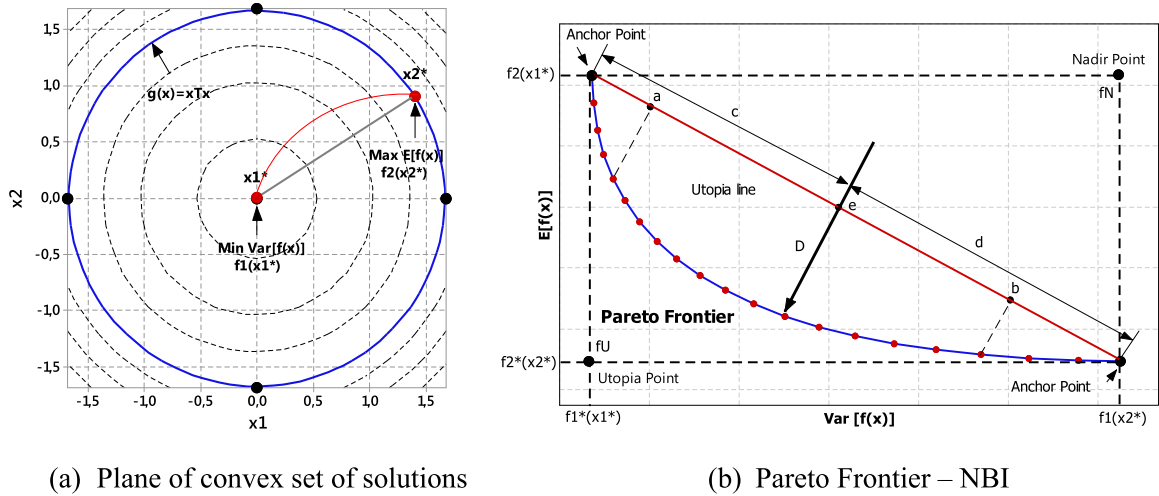
(a) Plane of convex set of solutions

(b) Pareto Frontier – NBI

Fig. 3. Schematic illustration of NBI for E[f(x)] versus Var [f(x)].
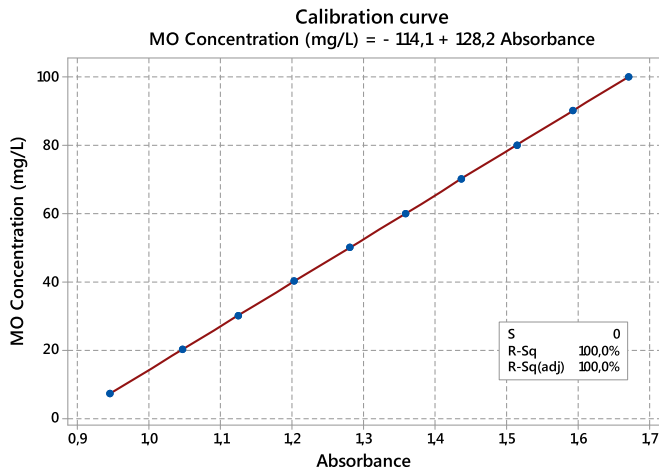
Fig. 4. Calibration curve of MO concentration versus absorbance.

$$E\left[\widehat{F}_s(\mathbf{x})\middle|\mathbf{x}_0\right] = \mathbf{x}_0^{(m)\mathbf{T}}\left[\left(\mathbf{X}^\mathbf{T}\mathbf{X}\right)^{-1}\left(\mathbf{X}^\mathbf{T}\mathbf{F}_s\right)\right]$$
$$= \mathbf{x}_0^{(m)\mathbf{T}}\left[\left(\mathbf{X}^\mathbf{T}\mathbf{X}\right)^{-1}\left(\mathbf{X}^\mathbf{T}\mathbf{Z}\left[\mathbf{L}\left(\mathbf{L}^\mathbf{T}\mathbf{L}\right)^{-1}\right]\right)\right] \quad (24)$$

where: $\mathbf{F} = \mathbf{Z}[\mathbf{L}(\mathbf{L}^\mathbf{T}\mathbf{L})^{-1}]$, $\mathbf{L}$ is the matrix of loading vectors and $\mathbf{Z}$ is the standardized value of the original responses as mentioned earlier.

So, the FA-NBI optimization approach can be written as:

$$Min\ \overline{f}_{(\mu)}(\mathbf{x}) = \left[\frac{\widehat{F}_{s(\mu)}(\mathbf{x}) - \widehat{F}_{s(\mu)}^U}{\widehat{F}_{s(\mu)}^N - \widehat{F}_{s(\mu)}^U}\right]\ s.t.:\quad \overline{g}_{1(\mu,\sigma^2)}(\mathbf{x})$$

$$= \left[\frac{\widehat{F}_{s(\mu)}(\mathbf{x}) - \widehat{F}_{s(\mu)}^U}{\widehat{F}_{s(\mu)}^N - \widehat{F}_{s(\mu)}^U}\right] - \left[\frac{\widehat{F}_{s(\sigma^2)}(\mathbf{x}) - \widehat{F}_{s(\sigma^2)}^U}{\widehat{F}_{s(\sigma^2)}^N - \widehat{F}_{s(\sigma^2)}^U}\right] + 2\beta_i - 1$$

$$= 0 \quad g_2(\mathbf{x}) = \mathbf{x}^\mathbf{T}\mathbf{x} \leq \rho^2 \quad 0 \leq w_i \leq 1$$
$$(25)$$

To compare the results of this proposal with traditional Multiobjective Optimization Problems (MOP) algorithms, it will be considered the Weighted Sums methods (Eq. (26)) and the Global Criterion Method (Eq. (27)). Such equations can be stated respectively as follows:

$$Minimize\quad F(\mathbf{x}) = w_1\left[\overline{f}_1(\mathbf{x})\right] + (1 - w_1)\left[\overline{f}_2(\mathbf{x})\right]$$
$$Subject\ to\quad \mathbf{x} \in \mathbf{\Omega} \tag{26}$$

$$Minimize\quad F(\mathbf{x}) = w_1\left[\frac{f_1(\mathbf{x}) - f_1^N(\mathbf{x})}{f_1^U(\mathbf{x}) - f_1^N(\mathbf{x})}\right]^2 + (1 - w_1)\left[\frac{f_2(\mathbf{x}) - f_2^N(\mathbf{x})}{f_2^U(\mathbf{x}) - f_2^N(\mathbf{x})}\right]^2$$

$$Subject\ to\quad \mathbf{x} \in \mathbf{\Omega}$$
$$(27)$$

To illustrate the fundamental steps of NBI-FA, next section will present a case study involving the multiobjective optimization of methyl orange degradation by ozone. A central composite design for three factors ($x_1 = $ pH, $x_2 = $ air flow and $x_3 = $ ozone dosage) will be developed to assess the expected value E[f(x)] and prediction variance Var[f(x)] of dye removal (Y_1) and chemical oxygen demand removal (Y_2). Principal component factor analysis will be considered to extract underlying data representing the original data set. Applying varimax rotation, factor scores will be computed which enables the modelling of E[$f_1(x)$], E[$f_2(x)$], Var[$f_1(x)$] and Var[$f_2(x)$], respectively.

5. Materials and methods

To illustrate the feasibility of NBI-FA approach as an alternative to solve multiobjective optimization problems, the proposal is applied to treatment of Methyl Orange (MO) using ozone. Ozonation is an environmentally sound technique because of no sludge formation have the potential to perform decolorization and degradation in a single step. According to Robinson et al. (2001), the use of ozone started in the early 1970s, and it is a very good oxidizing agent due to its high instability compared to other oxidizing agents. Some of the advantages of this treatment process are: the decomposition of residual ozone into water and oxygen, less space requirement for equipment installation, less hazardous as no stock of H_2O_2 is required for the oxidation step and ease of operation (Asghar et al., 2015). Besides the low cost and high removal rates, the main advantage of ozonation is the complete degradation of the organic compound which does not happen in conventional and biological processes because of the high

recalcitrance. Another advantage is the decomposition of ozone into oxygen during the reaction, which does not compromise other unit operations for separating secondary compounds formed during the reaction. The main drawback is that ozone can not be stored; all the needed solution must be generated and promptly used. However, there are researchers who reported that the half life of ozone in solution can reach up to 165 min, depending on the conditions mainly governed by Henry's Law.

The innovation of this study concerns the use of statistical tools and NBI for the understanding of solution treatment process containing a dye. This is a very common compound in the textile industry. Otherwise, the considerable difficulty in maintaining a sewage treatment plant at a steady state justifies the need of models that can simultaneously correct the variance and the means of flows faster and accurately while maintaining the final disposal condition at an accepted level. Setups capable of maximizing the removal rates for decolorization and COD with low values of variance favor the obtaining of a more stable process.

5.1. Reagents

The dye chosen for the study was the methyl orange (MO) which is known as a standard mono azo compound. The molecular weight of the dye was 327.33 g/mol with maximum absorption at $\lambda = 420$ nm. In this study, some important parameters (initial color, ozone dosage and pH) were varied to determine the optimum condition of the treatment process in degrading MO. It was also used sulfuric acid and sodium hydroxide for pH control solutions of 0.5 mol/L. All chemicals were used as received without further purification. The Methyl orange solution was purchased from VETEC Ltda (Rio de Janeiro, Brazil) and the sulfuric acid and sodium hydroxide were purchased from Bioquimis (Rio de Janeiro, Brazil). The ozone generator was built in the laboratory of the Federal University of São João Del Rei (MG — Brazil).

The methyl orange solution was chosen in this study because it is a standard compound for this type of analysis with an azo characteristic difficult to degrade. Otherwise, small masses of this compound in water confer high color in solution. So, discharges in an inappropriate manner can contaminate water bodies reducing sunlight penetration in the photic zone and thus reducing the generation of oxygen produced from algae present in this region.

5.2. Reactor configuration

The reactions were done in a useful volume of reactor of 2 L, made in acrylic, composed of ozone gas input and two outputs, one for collecting samples to be analyzed and the other for output of residual ozone which will then be degraded in a tank with potassium iodide solution. The reaction time of recirculation in the reactor was kept constant in 30 min. Meanwhile, the initial pH of the influent (pH 8), temperature (30 °C), dye concentration (100 mg/L) and hydraulic retention time of the reactor (30 min) were kept constant to reduce the number of factors and to simplify the experimental design. The selection of factors and the chosen levels were determined based on previous studies and initial screening experiments of various parameters and different levels. To further refine the input factors order when applying a surface response methodology, the model was evaluated in terms of curvature using a 2^3 full factorial design with six center points.

5.3. Experimental procedures

For each experiment, ozone was introduced into the reactor at different flow rates and pre-fixed pH's. The air flow rate was controlled by a flow meter, calibrated for use with air, with range between of 0–15 l/min. All experiments were carried out using 2 L samples in an ozone reactor with a height of 65 cm and an inner diameter of 16.5 cm. Dye concentration and pH were measured prior to initiation of treatment in the reactor. Ozone was produced using OM1 equipment, built and validated by the experimenter owner. The maximum flow of ozone produced by the machine using 100% of its capacity was 9.3 g O_3/h.

5.4. Analytical methods

The concentration of the dye was determined from a calibration curve as a function of absorbance as determined by a HACH spectrophotometer, and the concentration of stock solutions prepared. The pH was measured by a portable digital pH/Mv meter. The total organic load removal was determined by an organic carbon analyzer (TOC), tag Shimadzu TOC-LCSH model. The concentration of the color was measured using a Hach DR 2800 spectrophotometer. The removal efficiency of COD and color were obtained using in the following equation:

$$Y(\%) = 100 \times \left[C_i - \frac{C_f}{C_i} \right] \tag{28}$$

where C_i and C_f respectively refer to the initial and final COD and color.

The oxidation of the dye can take place via two different kinetics: a direct reaction, using ozone as the oxidant, and a parallel reaction in which there is the formation of hydroxyl radicals, which have a higher oxidation potential than the ozone only. In both cases, the pH will interfere directly in the process of decomposition of ozone influencing the mass transfer of the gas inside the reactor. In this study it can be seen that the pH played a significant role in the process: more alkaline pH's favors the indirect reaction and more acid pH's favors the forward reaction, favoring the degradation of most of the dye via ozone. The removal of COD and color were increased by increasing pH. These phenomena are attributed to the ability of O_3 to initiate hydroxyl radical at high pH levels, which has an oxidation potential ($E_0 = 2.80$) higher than O_3 ($E_0 = 2.07$) in the direct reaction at acidic condition.

6. Experimental design

In the present study, a central composite design (CCD) was run to model and optimize the experimental parameters and assess the relationships between three significant independent variables: (1) pH, (2) Air Flow an (3) Ozone dosages. The respective levels adopted in this work are shown in Table 1. The complete CCD design is presented in Table 2.

The pH, air flow and the ozone dosage were combined in a $2^k = 2^3 = 8$ factorial points, six axial points ($2k = 6$) and six center points ($cp = 6$), resulting in 20 experiments. An axial distance of 1.682 was adopted which corresponds to a spherical constraint $g(\mathbf{x}) = \mathbf{x}^T\mathbf{x} \leq 2.829$.

Table 2 also presents the responses of dye reduction (Y_1) and

Table 1
Factors and levels adopted in the study.

Factors	Units	Levels				
		−1.682	−1.000	0.000	+1.000	+1.682
pH (x_1)	—	3	5	7	9	11
Air flow (x_2)	l.min^{-1}	2	3	5	7	9
Ozone Dosage (x_3)	g.h^{-1}	8	13	18	21	34

Table 2
Central composite design.

x_1	x_2	x_3	Y_1	Y_2	Var[$Y_1(x)$]	Var[$Y_2(x)$]	F_1	F_2
−1.0000	−1.0000	−1.0000	81.02	74.69	6.6905	8.2202	0.2906	−0.8049
1.0000	−1.0000	−1.0000	85.03	84.12	6.6905	8.2202	0.4437	0.2586
−1.0000	1.0000	−1.0000	69.87	65.44	6.6905	8.2202	−0.1018	−2.5450
1.0000	1.0000	−1.0000	83.21	79.32	6.6905	8.2202	0.3732	−0.2614
−1.0000	−1.0000	1.0000	89.78	82.64	6.6905	8.2202	0.6002	0.6112
1.0000	−1.0000	1.0000	81.45	78.97	6.6905	8.2202	0.3134	−0.4564
−1.0000	1.0000	1.0000	90.02	82.69	6.6905	8.2202	0.6084	0.6380
1.0000	1.0000	1.0000	91.06	88.02	6.6905	8.2202	0.6538	1.1206
−1.6818	0.0000	0.0000	83.24	73.96	8.0785	9.9243	0.8486	−0.5216
1.6818	0.0000	0.0000	92.40	85.96	8.0785	9.9243	1.1796	1.2235
0.0000	−1.6818	0.0000	87.89	79.98	8.0785	9.9243	1.0165	0.3592
0.0000	1.6818	0.0000	87.67	80.87	8.0785	9.9243	1.0109	0.4017
0.0000	0.0000	−1.6818	67.64	67.50	8.0785	9.9243	0.3123	−2.4914
0.0000	0.0000	1.6818	86.28	83.28	8.0785	9.9243	0.9689	0.4403
0.0000	0.0000	0.0000	89.78	87.30	1.0043	1.2381	−1.3772	0.4487
0.0000	0.0000	0.0000	90.78	88.96	1.0043	1.2381	−1.3403	0.6643
0.0000	0.0000	0.0000	87.25	87.65	1.0043	1.2381	−1.4614	0.2295
0.0000	0.0000	0.0000	87.82	86.28	1.0043	1.2381	−1.4450	0.1863
0.0000	0.0000	0.0000	86.48	84.78	1.0043	1.2381	−1.4929	−0.0507
0.0000	0.0000	0.0000	88.90	89.89	1.0043	1.2381	−1.4016	0.5495

COD removal (Y_2). The variances for the predictions of dye removal and chemistry demand oxygen were calculated according to Eq. (8) and the rotated varimax factor scores (F_1 and F_2) according to Eq. (20).

7. Results, modelling, comparisons and discussion

7.1. Response surface modelling

Tables 3 and 4 present the ANOVA for the full quadratic models of dye reduction (Y_1) and COD removal (Y_2), respectively. It is possible to observe an adjusted determination coefficient (R^2 adj.) of 93.63% for $Y_1(\mathbf{x})$ and of 92.81% for $Y_2(\mathbf{x})$. The P-values presented in these tables reveal that some effects (and terms) are not statistically significant and may dropped from the model.

Fig. 5 presents the Pareto Chart for the standardized effects of $Y_1(x)$ and $Y_2(x)$ with the respective (a) T-Value statistic (blue line) and (b) Bonferroni Limit (red line) for all effects of a full quadratic model. It is also noticed that the residual from these models are normally distributed. Although some terms are not significant, we will adopt the full quadratic model for all response surfaces used in this work.

Table 3
ANOVA for full quadratic model of dye reduction (Y_1).

Source	DF	Adj SS	Adj MS	F	P-Value
Model	9	778.478	86.498	32.020	0.000
Linear	3	353.273	117.758	43.590	0.000
x_1	1	47.484	47.484	17.580	0.002
x_2	1	0.892	0.892	0.330	0.578
x_3	1	304.898	304.898	112.870	0.000
Square	3	240.509	80.170	29.680	0.000
x_1^2	1	0.830	0.830	0.310	0.592
x_2^2	1	0.930	0.930	0.340	0.570
x_3^2	1	239.840	239.840	88.790	0.000
2-Way Interaction	3	184.696	61.565	22.790	0.000
$x_1 x_2$	1	43.711	43.711	16.180	0.002
$x_1 x_3$	1	75.891	75.891	28.100	0.000
$x_2 x_3$	1	65.094	65.094	24.100	0.001
Error	10	27.012	2.701		
Lack-of-Fit	5	13.910	2.782	1.060	0.475
Pure Error	5	13.102	2.620		
Total	19	805.491			
Model	R^2:	96.65%	R^2 adj.:	93.63%	

Table 4
ANOVA for full quadratic model of COD removal (Y_2).

Source	DF	Adj SS	Adj MS	F	P-Value
Model	9	843.831	93.759	28.260	0.000
Linear	3	373.982	124.661	37.570	0.000
x_1	1	149.277	149.277	44.990	0.000
x_2	1	0.873	0.873	0.260	0.619
x_3	1	223.832	223.832	67.460	0.000
Square	3	321.656	107.219	32.310	0.000
x_1^2	1	79.927	79.927	24.090	0.001
x_2^2	1	69.158	69.158	20.840	0.001
x_3^2	1	227.223	227.223	68.480	0.000
2-Way Interaction	3	148.193	49.398	14.890	0.001
$x_1 x_2$	1	22.613	22.613	6.820	0.026
$x_1 x_3$	1	58.590	58.590	17.660	0.002
$x_2 x_3$	1	66.990	66.990	20.190	0.001
Error	10	33.179	3.318		
Lack-of-Fit	5	16.390	3.278	0.980	0.510
Pure Error	5	16.790	3.358		
Total	19				
Model	R^2:	96.22%	R^2 adj.:	92.81%	

Applying the ordinary least squares in the results shown on Table 2, it is possible to obtain the 6 s order models for dye reduction (Y_1) and COD removal (Y_2), their respective prediction variances and rotated factor scores. These models are summarized in Table 5. The full quadratic models of the six responses are all significant, with an acceptable level of R^2 adj. The response surfaces of expected value (mean) and variance of Y_1 and Y_2 are shown on Figs. 6 and 7.

Table 6 presents the correlation analysis of the six responses. It is possible to verify that dye removal (Y_1) and COD removal (Y_2) are positively correlated (r = 0.880) as well as the variances (r = 1.000). Therefore, since these responses are positively correlated with inverse optimization directions it is necessary to separate the two groups of objective functions.

Table 7 shows the results of a factor analysis based on principal components. This table also presents the rotated and unrotated results. It is possible to note that with two factors, 97.7% of the variance can be explained.

Observing the rotated factor loadings (values in bold), it is easy to conclude that factor 2 is highly correlated with expected value of dye removal (Y_1) and COD removal (Y_2), while factor 1 are correlated with the prediction variances of these responses. Moreover,
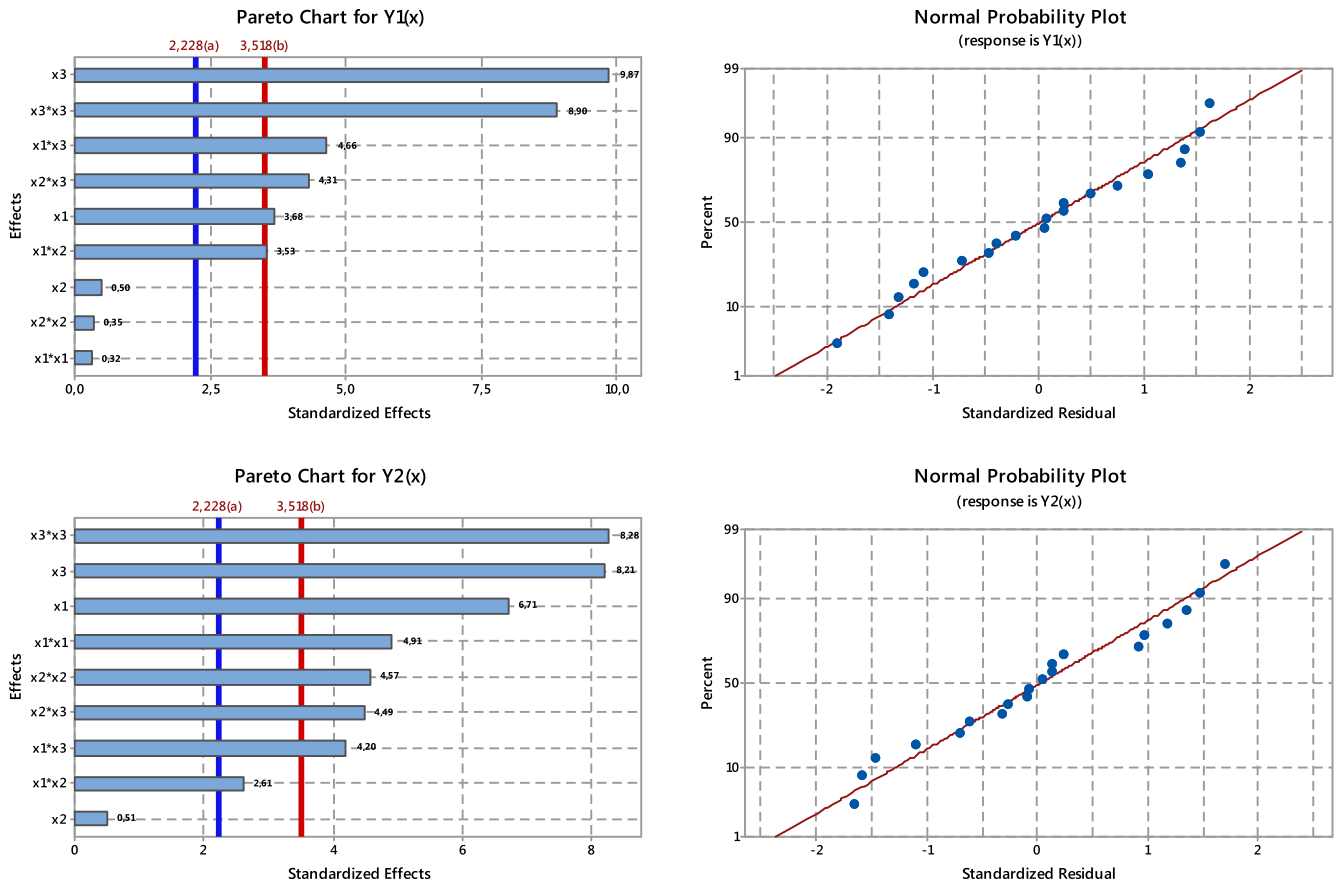
Fig. 5. Pareto Chart for standardized effects of $Y_1(x)$ and $Y_2(x)$ and residual plots.

Table 5
Full quadratic models.

Terms	$Y_1(x)$	$Y_2(x)$	Var$[Y_1(x)]$	Var$[Y_2(x)]$	$F_1(x)$	$F_2(x)$
b_0	88.502	87.430	1.0630	1.3101	−1.3990	0.3400
b_1	1.865	3.306	0.0000	0.0000	0.0691	0.4172
b_2	−0.256	−0.253	0.0000	0.0000	−0.0091	−0.0428
b_3	4.725	4.048	0.0000	0.0000	0.1665	0.7466
b_{11}	−0.240	−2.355	2.1176	2.6002	0.7272	−0.0071
b_{22}	−0.254	−2.191	2.1176	2.6002	0.7270	0.0033
b_{33}	−4.080	−3.971	2.1176	2.6002	0.5951	−0.4938
b_{12}	2.337	1.681	0.0000	0.0000	0.0820	0.3463
b_{13}	−3.080	−2.706	0.0000	0.0000	−0.1090	−0.4915
b_{23}	2.852	2.894	0.0000	0.0000	0.1010	0.4830
R^2:	96.65%	96.22%	93.91%	93.89%	93.12%	96.56%
R^2. adj.:	93.63%	92.81%	88.42%	88.39%	86.92%	93.47%

when the factors are rotated, their respective eigenvalues becomes more similar than without varimax rotation. For Pareto frontier and NBI optimization, this fact is very important, since means that the two axes of the Pareto frontier have approximately the same weights. Although the unrotated factors have the same explanation, the loadings are weaker and eigenvalues are very different. If one uses the functions of these unrotated scores to build a Pareto frontier, a bias could be introduced in the optimization analysis, forcing one block of response to present better results than the other, independently from the assigned weights. So, NBI could now be applied with the second order model for factor score 1 and 2 representing respectively the block of prediction variances and means.
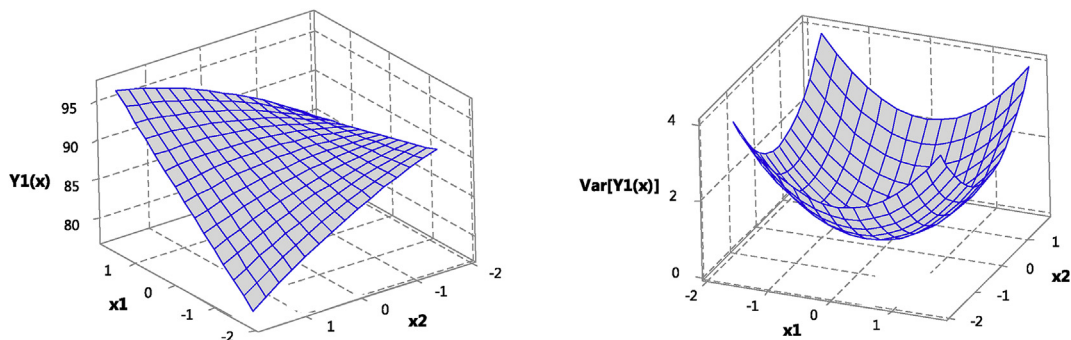
Fig. 6. Response surfaces of $Y_1(\mathbf{x})$ and Var$[Y_1(\mathbf{x})]$. Hold value $x_3 = 0$.
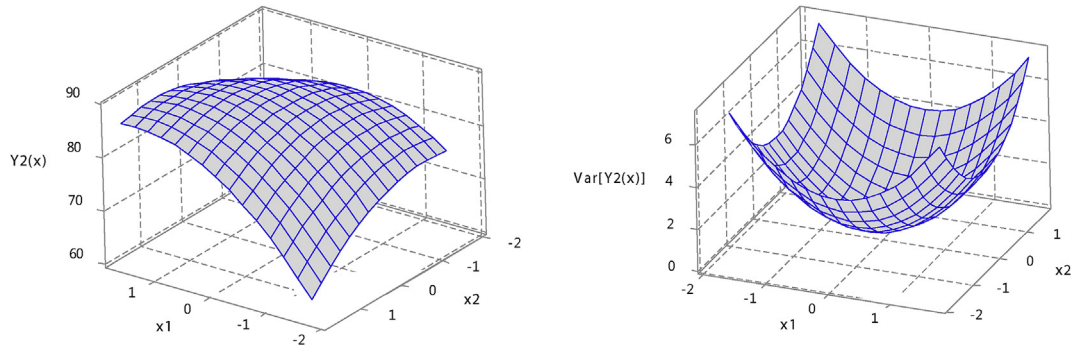
Fig. 7. Response surfaces of $Y_2(\mathbf{x})$ and $Var[Y_2(\mathbf{x})]$. Hold value $x_3 = 0$.

Table 6
Correlation matrix of original variables and rotated factor scores.

Variable	$Y_1(x)$	$Y_2(x)$	$Var[Y_1(x)]$	$Var[Y_2(x)]$	$F_1(x)rot$
$Y_2(x)$	0.881				
	0.000[a]				
$Var[Y_1(x)]$	−0.313	−0.519			
	0.179	**0.019**			
$Var[Y_2(x)]$	−0.313	−0.519	1.000		
	0.179	**0.019**	**0.000**		
$F_1(x)rot$	−0.106	−0.342	0.977	0.977	
	0.655	0.139	**0.000**	**0.000**	
$F_2(x)rot$	0.974	0.912	−0.209	−0.209	0.000
	0.000	**0.000**	0.376	0.376	1.000

[a] -Values in bold represent significant correlations (P-Value<5%).

Table 7
Factor Analysis based on principal components.

	Variable	Factor 1	Factor 2	Communality
Unrotated	$Y_1(\mathbf{x})$	−0.732	0.652	0.961
	$Y_2(\mathbf{x})$	−0.865	0.448	0.949
	$Var[Y_1(\mathbf{x})]$	0.866	0.499	0.999
	$Var[Y_2(\mathbf{x})]$	0.866	0.499	0.999
	Eigenvalues	2.7833	1.1252	3.9085
	% Variance	0.6960	0.2810	0.9770
Rotated	$Y_1(\mathbf{x})$	−0.106	**0.974**[a]	0.961
	$Y_2(\mathbf{x})$	−0.342	**0.912**	0.949
	$Var[Y_1(\mathbf{x})]$	**0.977**	−0.209	0.999
	$Var[Y_2(\mathbf{x})]$	**0.977**	−0.209	0.999
	Eigenvalues	2.0395	1.8690	3.9085
	% Variance	0.5100	0.4670	0.9770

[a] − Values in bold are the factors loadings relatives to factors 1 and 2.

7.2. NBI-FA multiobjective optimization

Using the NBI-FA method and varying iteratively the weights from 0 up to 1, it was obtained the results shown in Table 8. Figs. 8 and 9 show the Pareto frontiers of expected mean and variances for Y_1 and Y_2 as well as for $F_1(x)$ versus $F_2(x)$.

It can be verified that the Pareto frontiers are equispaced and convex. Moreover, it is possible to note that as the weights for the response surface of factor score 1 increases, the prediction variance and, consequently, the standard error of prediction of Y_1 and Y_2 decreases while the confidence interval for the predictions becomes narrower (Fig. 10).

The optimization results showed a maximum dye removal of 94.1% ± 4.3 with a respective chemical oxygen demand removal of 88.4% ± 5.3 obtained at $x^* = [9.5; 7.1 \text{ l min}^{-1}; 18.4 \text{ g h}^{-1}]$. However, this point have presented the largest 95% prediction confidence interval. Based on the fuzzy membership of Pareto set it was

possible to select the narrowest 95% confidence intervals with maximum removal rates ($Y_1 = 90.5 \pm 2.2$ and $Y_2 = 88.3 \pm 2.7$), obtained at $x^* = [7.9. 5.6 \text{ l min}^{-1}. 18.4 \text{ g h}^{-1}]$.

7.3. Comparisons among several MOP methods

To assess the efficiency of NBI-FA method, several comparisons were done. The main results were plotted in Figs. 11 and 12. Firstly, NBI, Weighted Sums (WS) and Global Criterion Method (GCM) were applied directly in $Y_1(x)$ and $Y_2(x)$. These results may be compared to the values of $Y_1(x)$ and $Y_2(x)$ obtained by the optimization of $F_1(x)$ and $F_2(x)$ using NBI (Table 8) according to Fig. 11. It is interesting that NBI outperforms WS and GCM for $Y_1(x)$ and $Y_2(x)$ when applied directly in $Y_1(x)$ and $Y_2(x)$, generating an equispaced and convex frontier, but inverts the correlation of the original response that is positive (Pearson's correlation coefficient r = +0.881), which is a serious mistake. In other words, if the correlation between the response were not removed from the original data set, the objective functions modelled by RSM will not be able to respect the original correlation.

Why does it happen? Considering the fact that a bi-objective weighted sum is a convex combination of two objective functions, if the objective functions are positively correlated the application of complementary weights (w and 1-w) probably inverts this direction. Despite of this, the values of $Y_1(x)$ and $Y_2(x)$ obtained with NBI-FA presented a positive correlation (r = +0.774) which is very compatible with the original correlation.

Analogously, NBI, WS and GCM were applied with the response surfaces of rotated factor scores $F_1(x)$ and $F_2(x)$. Fig. 12 shows these comparison results. It can be noted that only NBI-FA is capable of generating equispaced and convex Pareto frontiers which proves the good adequacy of the method.

Comparing the Pareto Frontier of $E[Y_1(\mathbf{x})] \times Var[\widehat{Y}_1(\mathbf{x})]$, obtained with NBI applied directly in the expected value and variance response surfaces with the results obtained with NBI-FA, it is clear that the frontier plotted with rotated factor scores is longer than that drawn with the original values which points out the capability of NBI-FA to better explore the solution space. NBI-FA promoted larger values for dye removal (94.0%) when compared to mean and variance approach. In Fig. 13 it is possible to note that the Pareto solution obtained with the weight of 15% in the frontier of $E[Y_1(\mathbf{x})] \times Var[\widehat{Y}_1(\mathbf{x})]$ is equivalent to a solution obtained with 40% in NBI-FA. For larger weights, however, these differences practically disappear.

Fig. 14 presents a fuzzy decision maker comparisons among the Pareto sets obtained with three different approaches: NBI-FA, GCM-FA and WS-FA. According to this criterion, the best compromise solution occurs with NBI-FA method at the weight of 75%. This

Table 8
Pareto points and optimal values.

W	$Y_1(x)$	$Y_2(x)$	$Var[Y_1(x)]$	$Var[Y_2(x)]$	$F_1(x)$	$F_2(x)$	x_1^*	x_2^*	x_3^*	FDM
0.00	94.113	88.354	7.054	8.667	0.854	1.383	1.267	1.049	0.352	0.478
0.05	93.875	88.445	6.667	8.192	0.712	1.338	1.213	1.021	0.366	0.493
0.10	93.651	88.511	6.282	7.718	0.571	1.292	1.169	0.980	0.369	0.507
0.15	93.427	88.572	5.898	7.247	0.431	1.246	1.124	0.939	0.371	0.521
0.20	93.202	88.627	5.516	6.778	0.291	1.200	1.078	0.895	0.373	0.535
0.25	92.974	88.676	5.137	6.313	0.152	1.153	1.029	0.851	0.375	0.549
0.30	92.746	88.718	4.760	5.850	0.014	1.106	0.979	0.803	0.377	0.562
0.35	92.515	88.750	4.386	5.391	−0.124	1.058	0.926	0.754	0.378	0.574
0.40	92.283	88.773	4.016	4.936	−0.260	1.010	0.871	0.701	0.379	0.587
0.45	92.047	88.784	3.650	4.487	−0.395	0.961	0.813	0.646	0.379	0.598
0.50	91.809	88.781	3.289	4.044	−0.528	0.911	0.751	0.587	0.379	0.609
0.55	91.567	88.760	2.934	3.608	−0.659	0.860	0.684	0.522	0.377	0.619
0.60	91.320	88.716	2.588	3.182	−0.788	0.808	0.613	0.453	0.373	0.628
0.65	91.067	88.642	2.251	2.769	−0.913	0.754	0.535	0.376	0.366	0.636
0.70	90.804	88.526	1.930	2.375	−1.034	0.697	0.448	0.290	0.353	0.641
0.75	90.527	88.343	1.633	2.010	−1.147	0.636	0.351	0.194	0.329	0.643
0.80	90.220	88.057	1.375	1.693	−1.247	0.569	0.244	0.095	0.280	0.639
0.85	89.839	87.623	1.185	1.460	−1.326	0.491	0.141	0.023	0.192	0.626
0.90	89.329	87.041	1.084	1.336	−1.379	0.398	0.058	−0.002	0.080	0.602
0.95	88.684	86.342	1.065	1.313	−1.406	0.292	−0.007	0.003	−0.032	0.566
1.00	87.931	85.553	1.112	1.371	−1.414	0.175	−0.062	0.020	−0.138	0.522
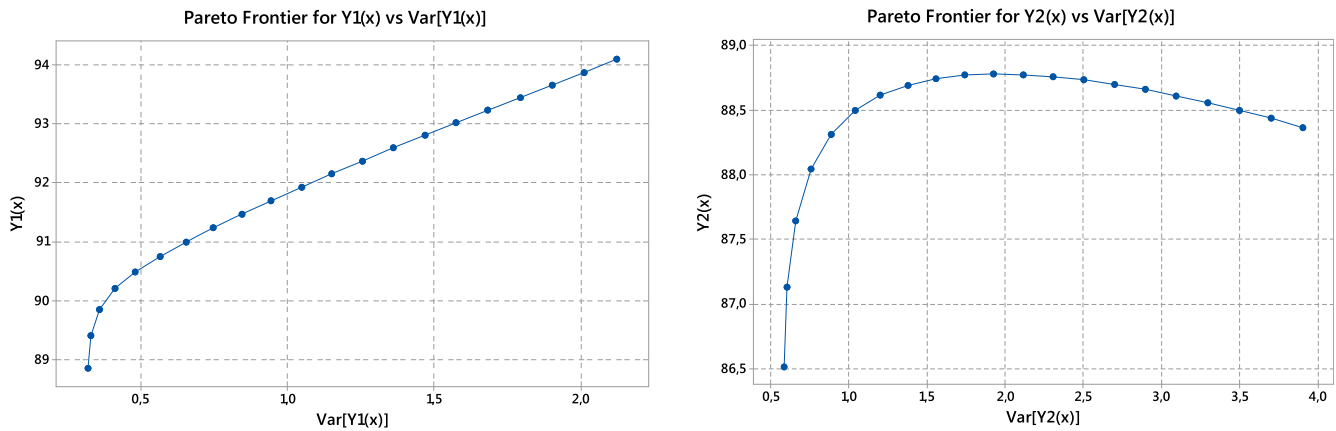
Fig. 8. Pareto frontiers for $E[Y_p(\mathbf{x})] \times Var[\widehat{Y}_p(\mathbf{x})]$ using NBI coupled with factor analysis.
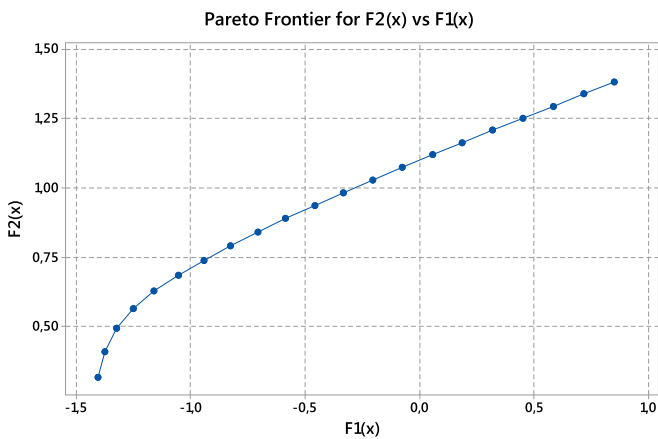
Fig. 9. Pareto frontiers for $F\{E[Y_i(\mathbf{x})]\} \times F\{Var[\widehat{Y}_i(\mathbf{x})]\}$ using NBI-FA method.

solution corresponds to the narrowest 95% confidence intervals with maximum possible removal rates for dye removal and chemical demand oxygen ($Y_1 = 90.5 \pm 2.2$ and $Y_2 = 88.3 \pm 2.7$). These optima were obtained at $x^* = [7.9. 5.6 \text{ l min}^{-1}. 18.4 \text{ g h}^{-1}]$.

The previous discussion indicates that the search of optimal conditions in a multiobjective optimization is not always a trivial task. In contrast, most of the papers devoted to optimize the wastewater treatment processes have been adopting just RSM graphical solutions, as contour or surface plots, without any optimization routine. Just a small amount of them has adopted optimization algorithms like the Desirability method (Montgomery, 2009), such as Asfaram et al. (2015d) that applied RSM and desirability for the optimization of simultaneous ultrasound-assisted ternary adsorption of dyes onto copper-doped zinc sulfide nanoparticles loaded on activated carbon, Gomes et al. (2016) that employed desirability in the optimization of tannery-dye-containing effluent treatment with leather shaving or Körbahti (2007) who applied RSM and desirability method in the optimization of electrochemical treatment of textile dye wastewater. So, for sake of comparison with traditional RSM optimization routines, it was developed a similar analysis.

Fig. 15 presents the contour plots for Y_1 (a, b and c) and Y_2 (d, e and f) for a specific hold condition [1.000; 1.000; 0.500] adopted within the most of Pareto sets used in this research (WS, GCM, NBI, NBI-FA). In a first inspection, it is possible to note that it is not so easy to find a suitable solution that meets all the objectives and constraints adopted in this work since the contour plots reveals the
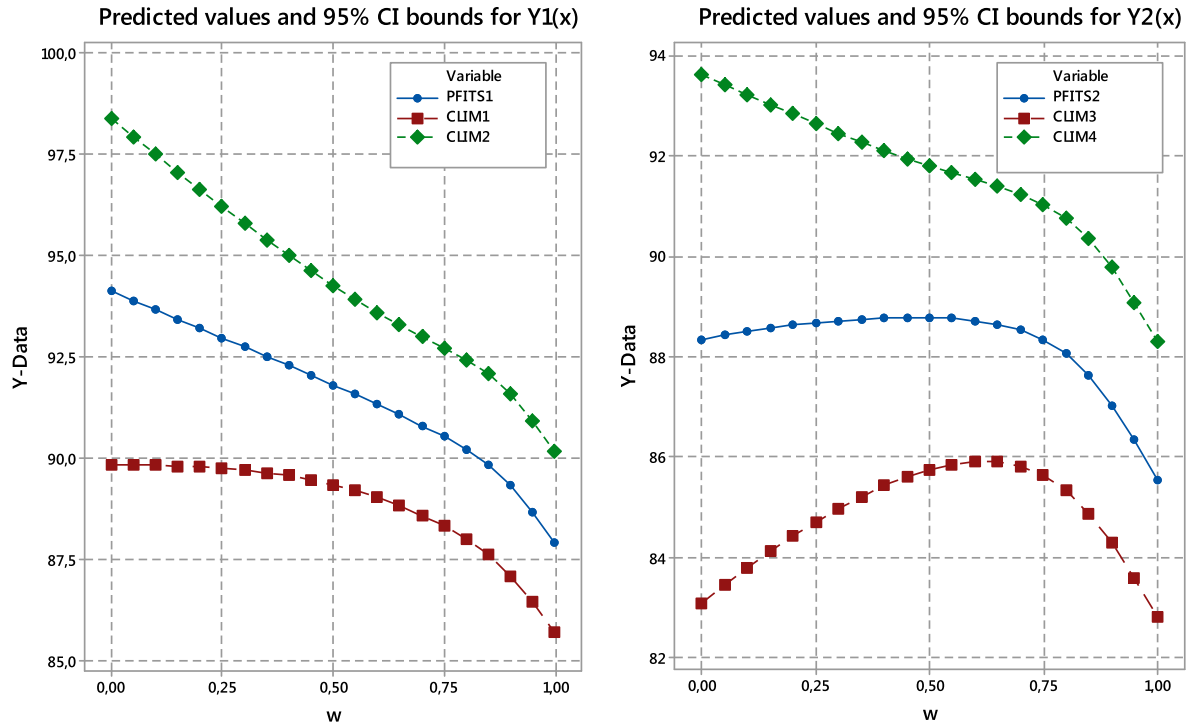
Fig. 10. Predictions and 95% C.I. bounds of Y_1 and Y_2 obtained from NBI-FA method.
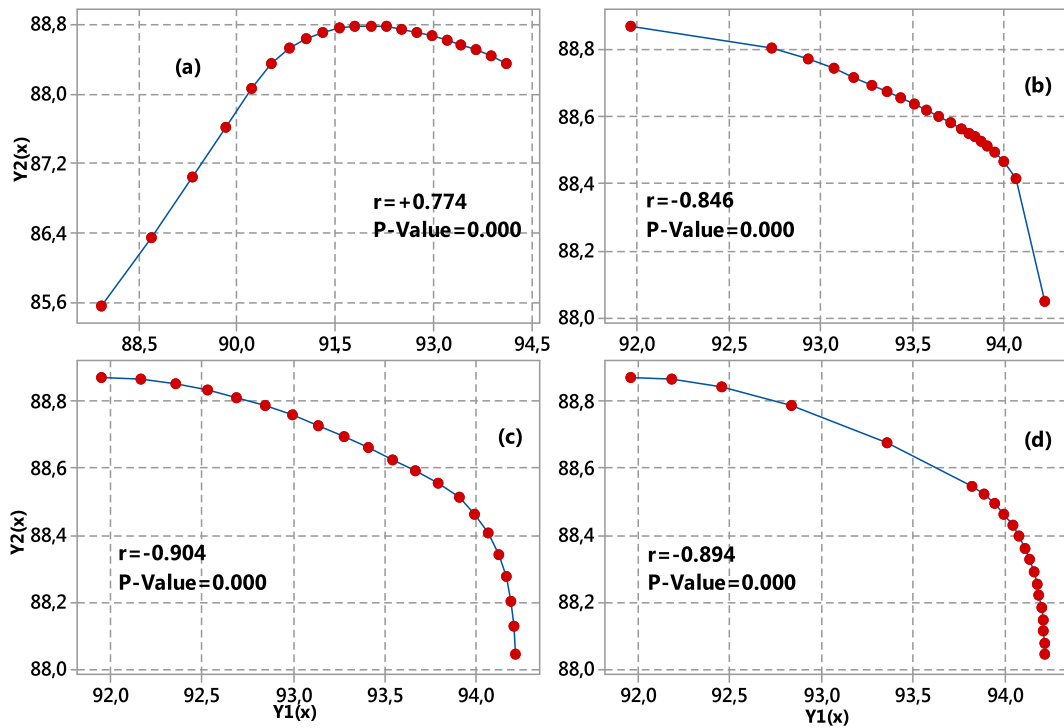

Fig. 11. Pareto Frontiers for Y_1 and Y_2 obtained with different methods.

behavior of each characteristic separately. However, even considering the entire set of objective functions and constraints, the overlaid contour plot of Fig. 16 is just able to shown a feasible region but is unable to define the best compromise solution within this space.

Otherwise, desirability method was also applied for comparison (Fig. 17). It is possible to note that this method is a feasible alternative since it is present in the most of commercial and statistical software, like Minitab 17©. Although desirability remains as a feasible alternative, the setup of its parameters is a challenge, since for each setup it will be created a new solution. Fig. 17 (a),(b), (c) and (d) show how the importance of each objective function may
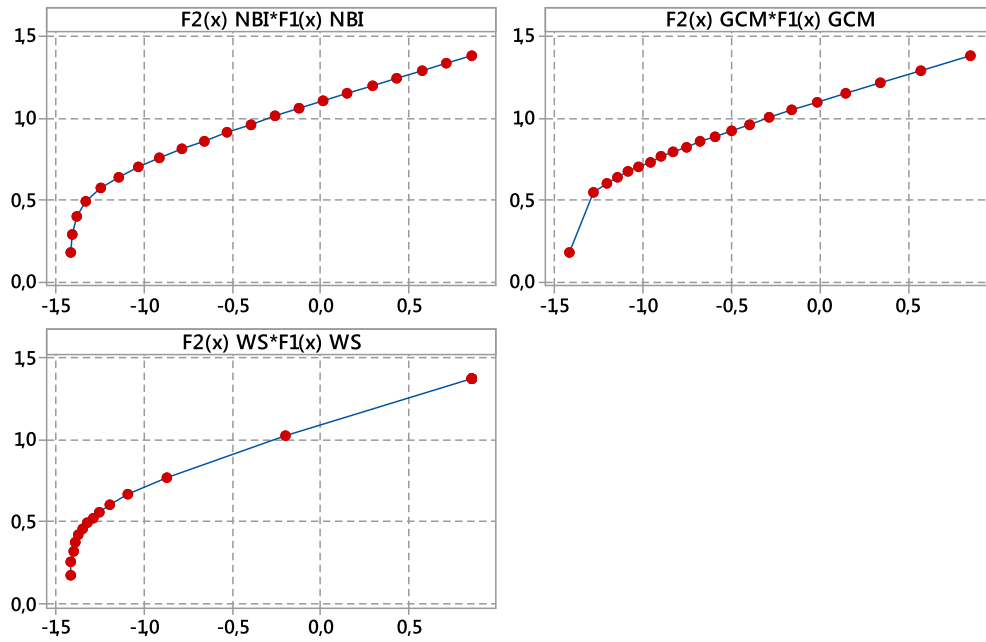
Fig. 12. Pareto Frontiers for Y_1 and Y_2 obtained with different methods.
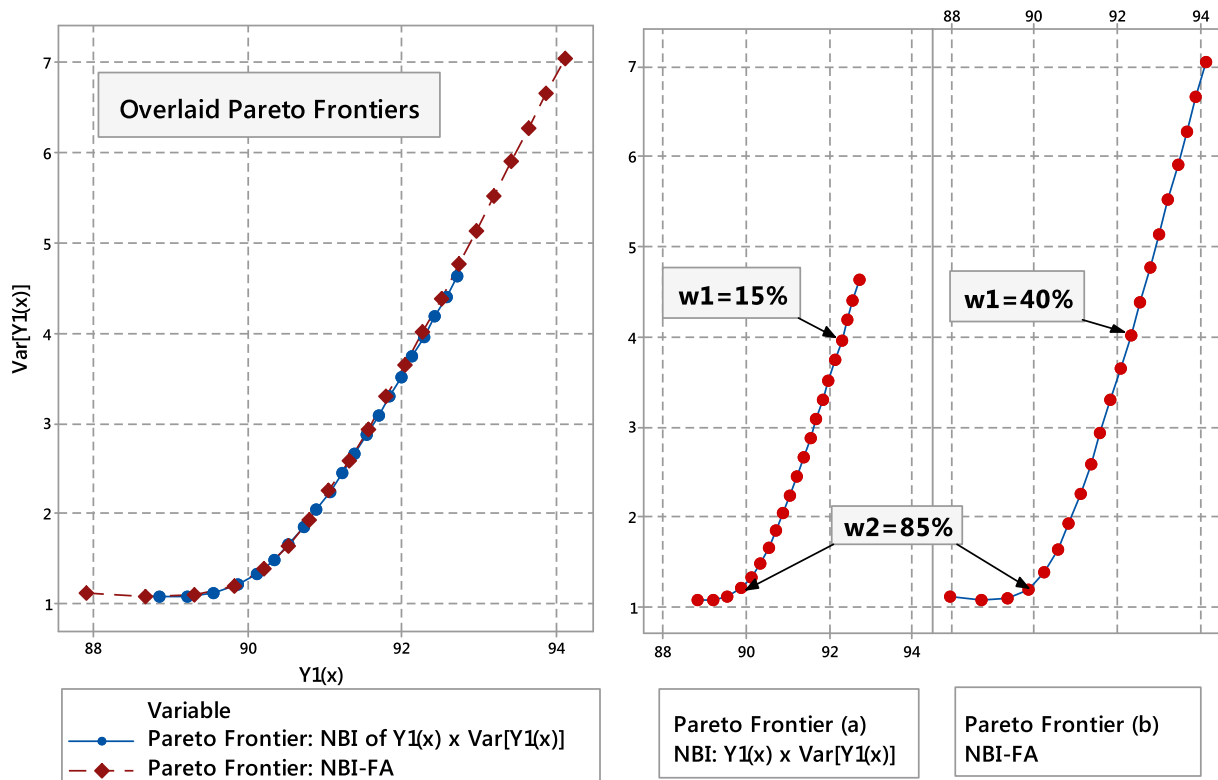
Fig. 13. Pareto Frontiers for $Y_1(x)$ and $Var[Y_1(x)]$ obtained with NBI method.

change the optimization results. So, it is necessary to run desirability several times, for different set of weights (or importance) to achieve the best compromise solution. This iteratively task is the same procedure conducted in the case of Pareto frontiers building. Therefore, even using an algorithm like desirability, a technical issue remains: how to adjust the weights to obtain the best compromise solution.

7.4. Power and sample size for confirmation runs

The calculation of sample size necessary to confirm the optimization results was done to accomplish with the following demands: first, to test the significant differences between a specific point in the Pareto Frontier and the mean of the confirmation runs in the same optimum using a one sample t hypothesis test; second,
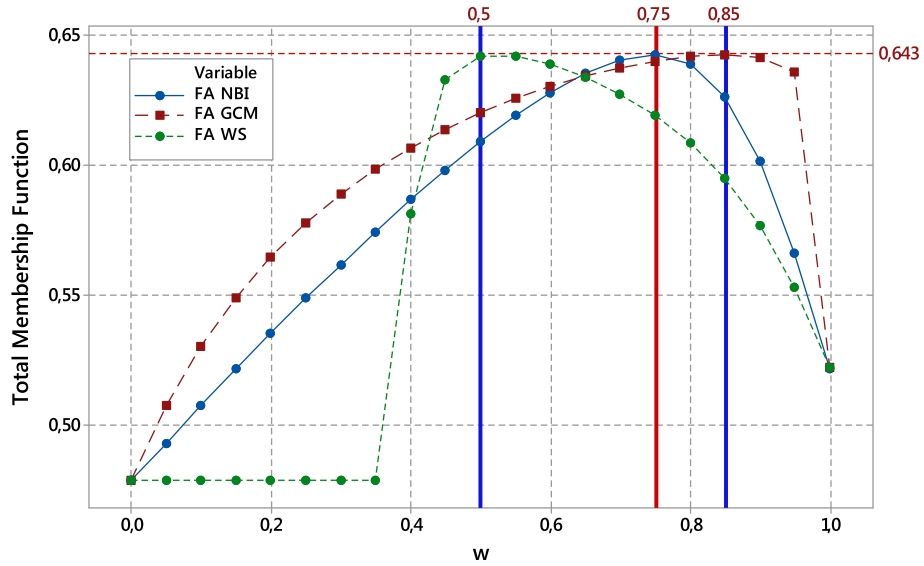
Fig. 14. Fuzzy decision maker comparisons.
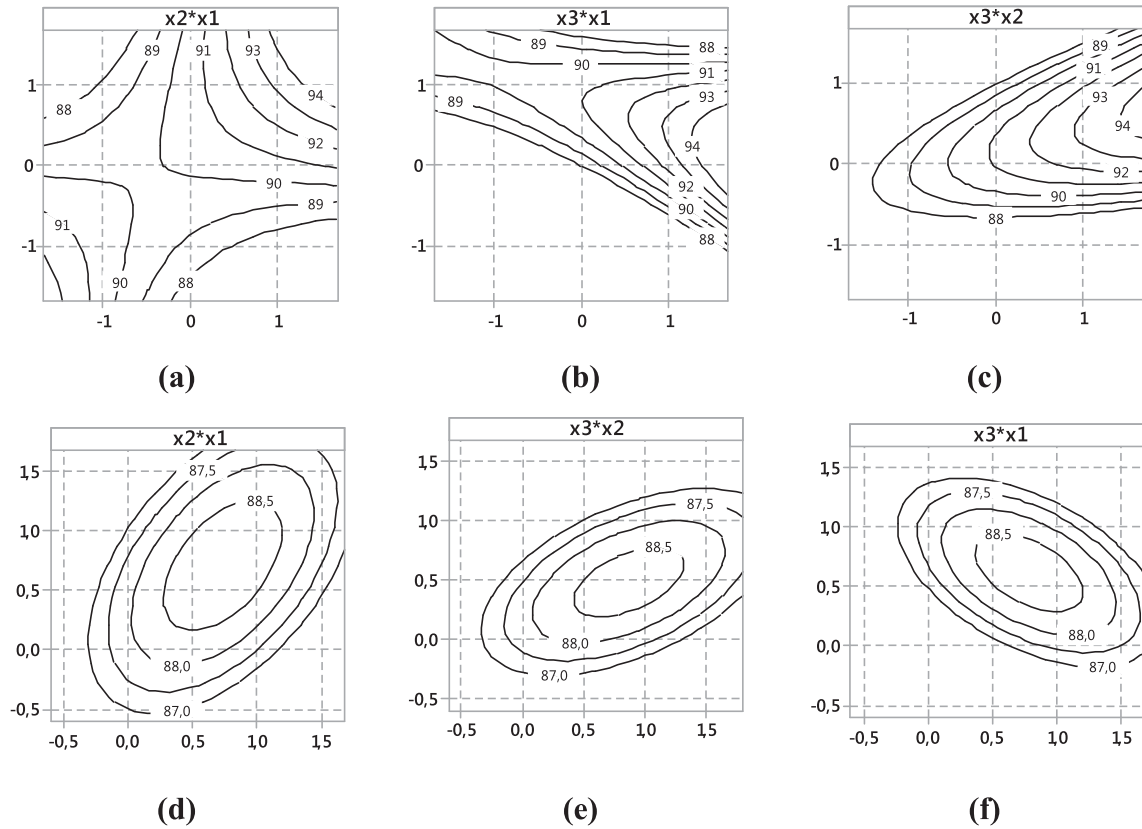
Fig. 15. Contour plots for Y_1 (a. b and c) and Y_2 (d. e and f). Hold values: [1.000; 1.000; 0.500].

to test if there is a significant difference between two far points in the frontier ($w_1 = 15\%$ and $w_2 = 85\%$) employing a 2-sample t-test. For the variances, the strategy is similar. This procedure was employed for Y_1 and Y_2, respectively. Projecting the sample size by the worst scenario, it was chosen differences between 1 up to 1.5%, with a standard deviation of 1.2%. Analogously, to test if there is a significant difference between two variances the sample size was calculated using different variance ratios (var_1/var_2), representing how many times a variance in a first condition ($w = 15\%$) is larger than in a second one ($w = 85\%$). Based on the Pareto frontier results, the larger variance ratios were 6.3, 5.0 and 4.0. In the three cases, a significance level of $\alpha = 5\%$ and a power of 80% were used. For all the tests, a sample size of $n = 20$ was a suitable choice.

Table 9 presents a summary of the main results, showing the optimum in each condition, the mean of confirmation runs for Y_1 and Y_2, the respective fitted values and upper and lower bounds for
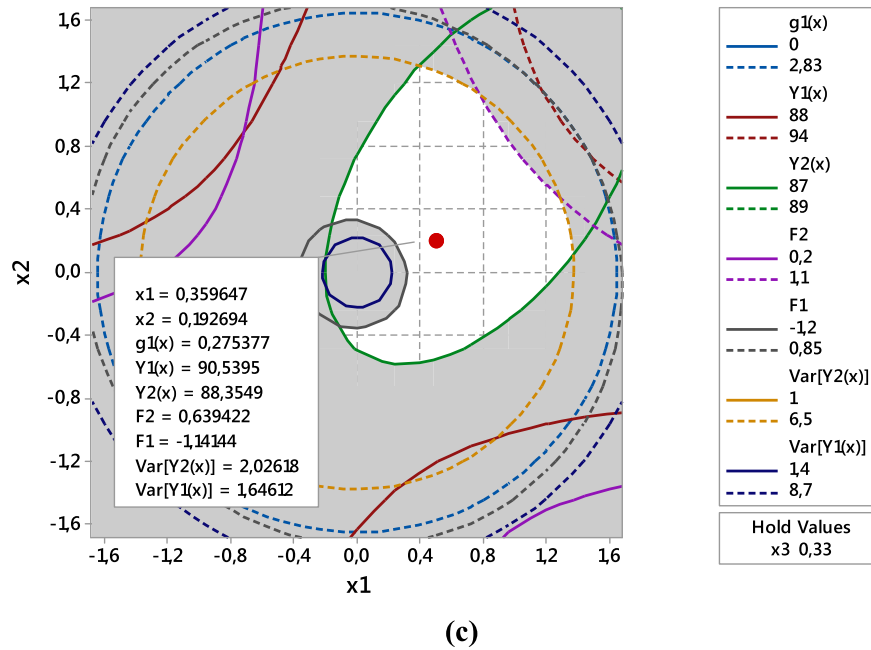
Fig. 16. Overlaid Contour plots.

95% confidence intervals. It can be noted that all sample means fall within the confidence intervals for Y_1 and Y_2, proving that the solutions are really feasible. According to one sample *t*-test results shown in Fig. 18 (b), it can be accepted the null hypothesis that the mean of confirmation runs for $Y_1(\mathbf{x})$ in the setup with weight of w = 15% is equal to the fitted value (the respective result on the Pareto frontier). The same conclusion may be drawn from the one sample *t*-test for $Y_2(\mathbf{x})$ (Fig. 19), since both p-values are lesser than α = 5%.

When comparing the difference between the expected mean of $Y_1(\mathbf{x})$ under the weights of 15 and 85% (Fig. 20), the null hypothesis is rejected with p-value = 0.000. Therefore, the Pareto solutions at these two weights are statistically different. For the predicted variances the results are similar. With p-value = 0.927 it is possible to accept the null hypothesis that the ratio between of variance of $Y_1(\mathbf{x})$ at w = 15% is at least three times larger than that obtained with w = 85%. With p-value = 0.775, the same conclusions may be drawn from two variances hypothesis test for variance of $Y_2(\mathbf{x})$ at w = 15% and w = 85% (Fig. 21).

Summarizing, it is possible to conclude that setups with lower prediction variances will also present lower means. This is in fact the essence of a trade-off between mean and variances into the solution space. Since all confirmation runs and hypothesis tests corroborated the optimization results, it is possible to affirm that NBI-FA method is a suitable method for optimization of RSM problems with correlated response data sets for mean and prediction variance of wastewater treatment process.

7.5. Discussion

Industrial wastewater treatment is a major concern in many countries nowadays. Wastewaters resulting from different processes have specific characteristics (Colla et al., 2016). As stated by Abidi et al. (2015), the wastewater produced by textile industries is colored and with high loading of inorganic salt and other chemical additives, that represent an environmental danger. In addition, the technologies used for the textile wastewater treatment require high capital and operating costs. In this context, the optimization of this

kind of process can bring very significant results. According to Ji et al. (2016), the optimization of treatment process is commonly based on the quality of the effluent, wastewater discharge standards and quality requirements for wastewater reuse. Although this study applied the NBI-FA method only for optimizing the dye removal and chemical oxygen demand removal, it could also be applied for other process characteristics and, even better, it could be applied for other processes, not only wastewater treatment.

The biological and biochemical processes that take place inside a wastewater treatment plant are strongly interrelated and involve a great number of variables, which turns the control of such plants a very complex task (Santín et al., 2016). The considerable difficulty in maintaining a sewage treatment plant at a steady state justifies the need of models that can simultaneously correct the variance and the means of flows faster and accurately while maintaining the final disposal condition at accepted levels, usually determined by national or regional policies of wastewater disposal and reuse (Corominas et al., 2013; Meneses et al., 2015; Pintilie et al., 2016). In the context of plant control, as proved in the previous sections, the NBI-FA method is capable of generating minimum prediction variances for the wastewater treatment, which was the main objective of this study. Minimum variances are very important to plan and control industrial wastewater treatments, since it promotes a more stable condition.

A more detailed analysis of the results was made so that the influence of all the three variables in the results found in the optimization could be discussed. The significative terms for color removal and COD removal models are presented in Tables 3 and 4. It can be noted that only the air flow is a non-significative parameter, although it presents significative interactions with the two other parameters tested.

For this process it was observed that any variation in the pH can directly influence its kinetic and thermodynamic. Taking into account the entire Pareto set provided in Table 6 (in uncoded values, according to Table 1), it is possible to evaluate the color removal behavior due to the change of the PH of the medium. As it can be seen in Fig. 22 (a), the scatterplot of color removal versus pH can be divided in two distinct regions: a) an exponential behavior of color
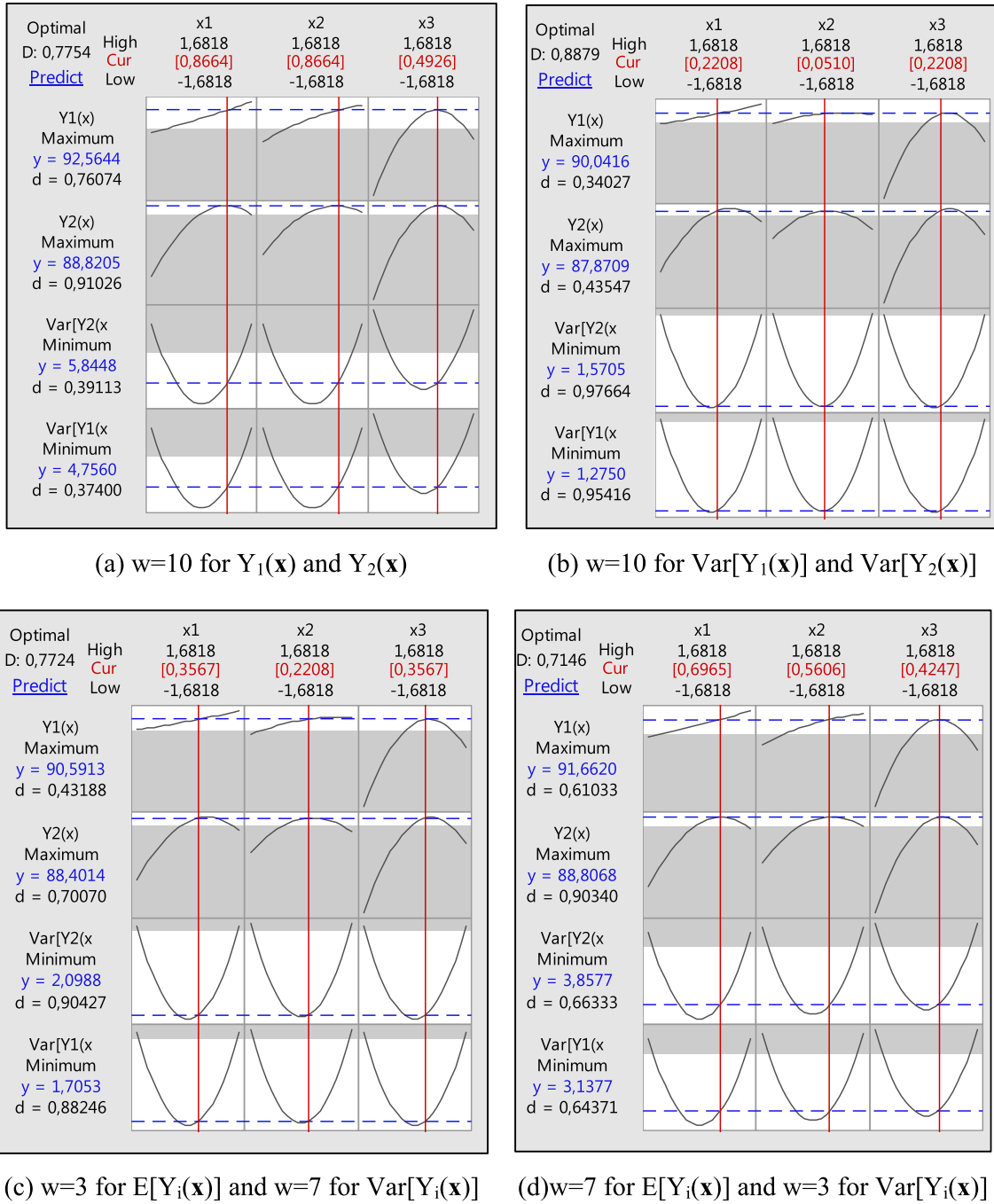
(a) w=10 for $Y_1(\mathbf{x})$ and $Y_2(\mathbf{x})$

(b) w=10 for $Var[Y_1(\mathbf{x})]$ and $Var[Y_2(\mathbf{x})]$

(c) w=3 for $E[Y_i(\mathbf{x})]$ and w=7 for $Var[Y_i(\mathbf{x})]$

(d) w=7 for $E[Y_i(\mathbf{x})]$ and w=3 for $Var[Y_i(\mathbf{x})]$

Fig. 17. Optimization using Desirability function.

Table 9
Summary of confirmation runs.

Y(x)	x1*	x2*	x3*	Y*(x)	LB	UB	Sample mean
Y1(x) 15%	1.124	0.939	0.371	93.427	89.804	97.050	92.563
Y1(x) 85%	0.141	0.023	0.192	89.839	87.615	92.063	89.760
Y2(x) 15%	1.124	0.939	0.371	88.573	84.124	93.021	88.217
Y2(x) 85%	0.141	0.023	0.192	87.623	84.893	90.353	87.153

removal from pH of 7.0−7.5 and b) an almost linear behavior for the remaining points.

Within the class of advanced oxidation processes, any increase in pH during the ozonolysis process favors the degradation of the

organic load in the solution, which comes from the methyl orange compound solubilized in distilled water. In this process, there is a significant formation of hydroxyl radicals (HO°), through an indirect reaction. However, the high molecular concentration of ozone present in solution may favor the direct reaction (Gottschalk et al., 2010).

The reaction for the degradation of methyl orange can be generalized as shown in Eq. (31).

$$A + O_3 \rightarrow CO_2 + H_2O \tag{31}$$

where the term A represents the methyl orange solution. In kinetic terms, Eq. (31) can be written in differential form as:
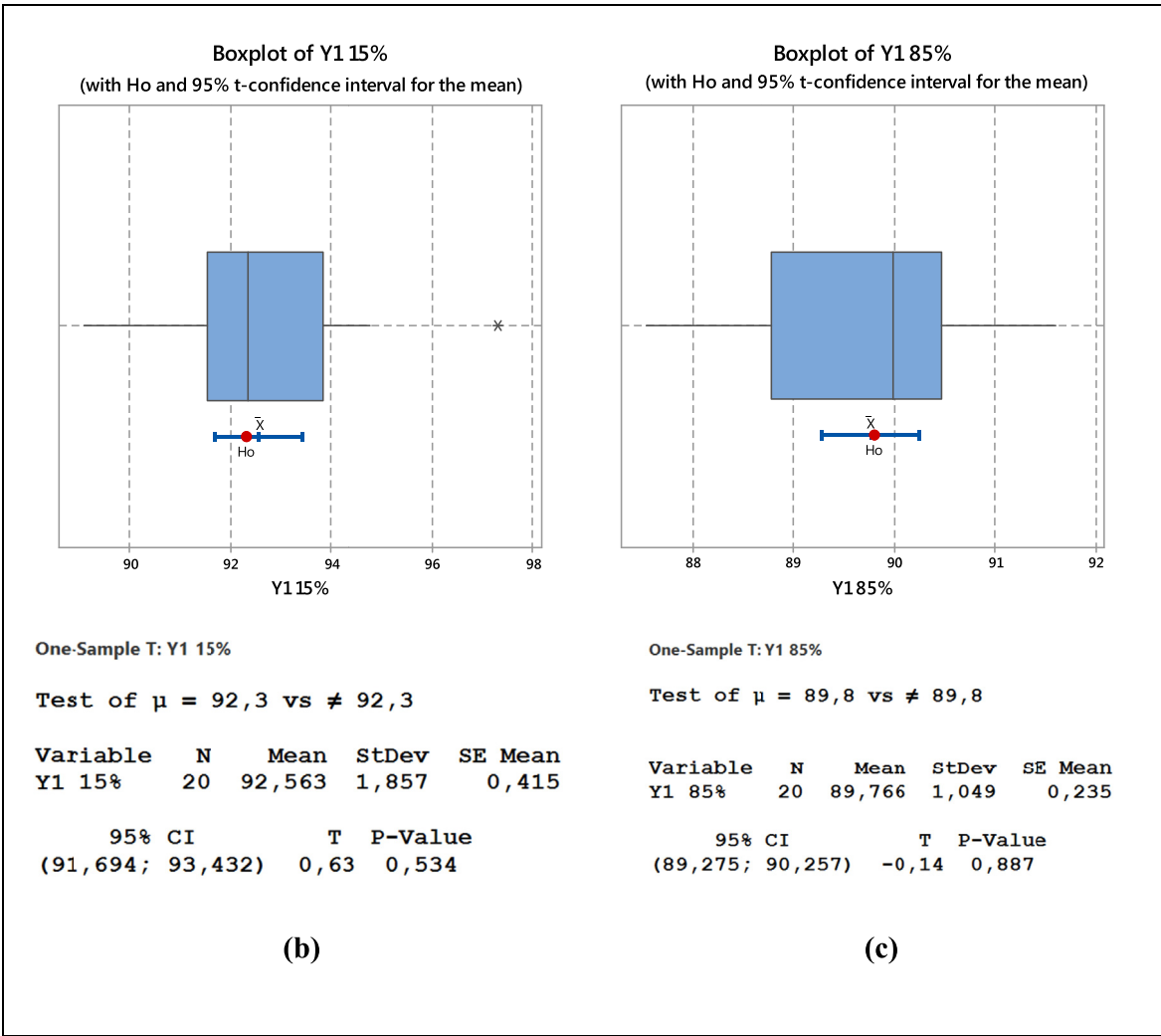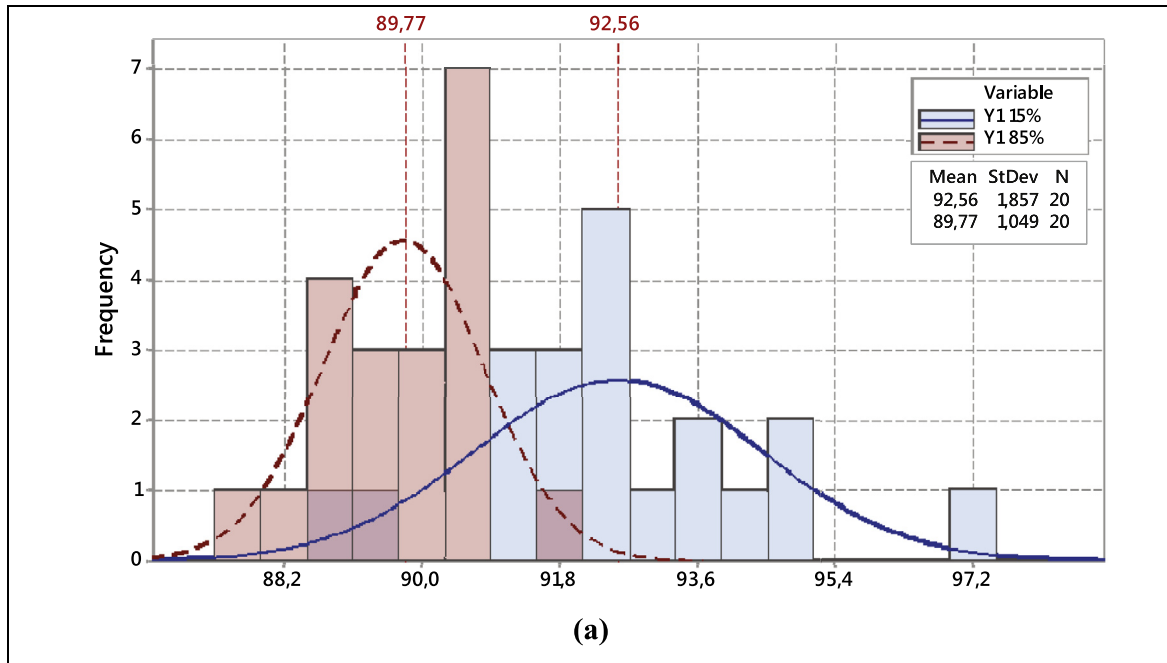
Fig. 18. Y1:Confirmation runs (a). One-sample t for w = 15% (b) and w = 85% (c).
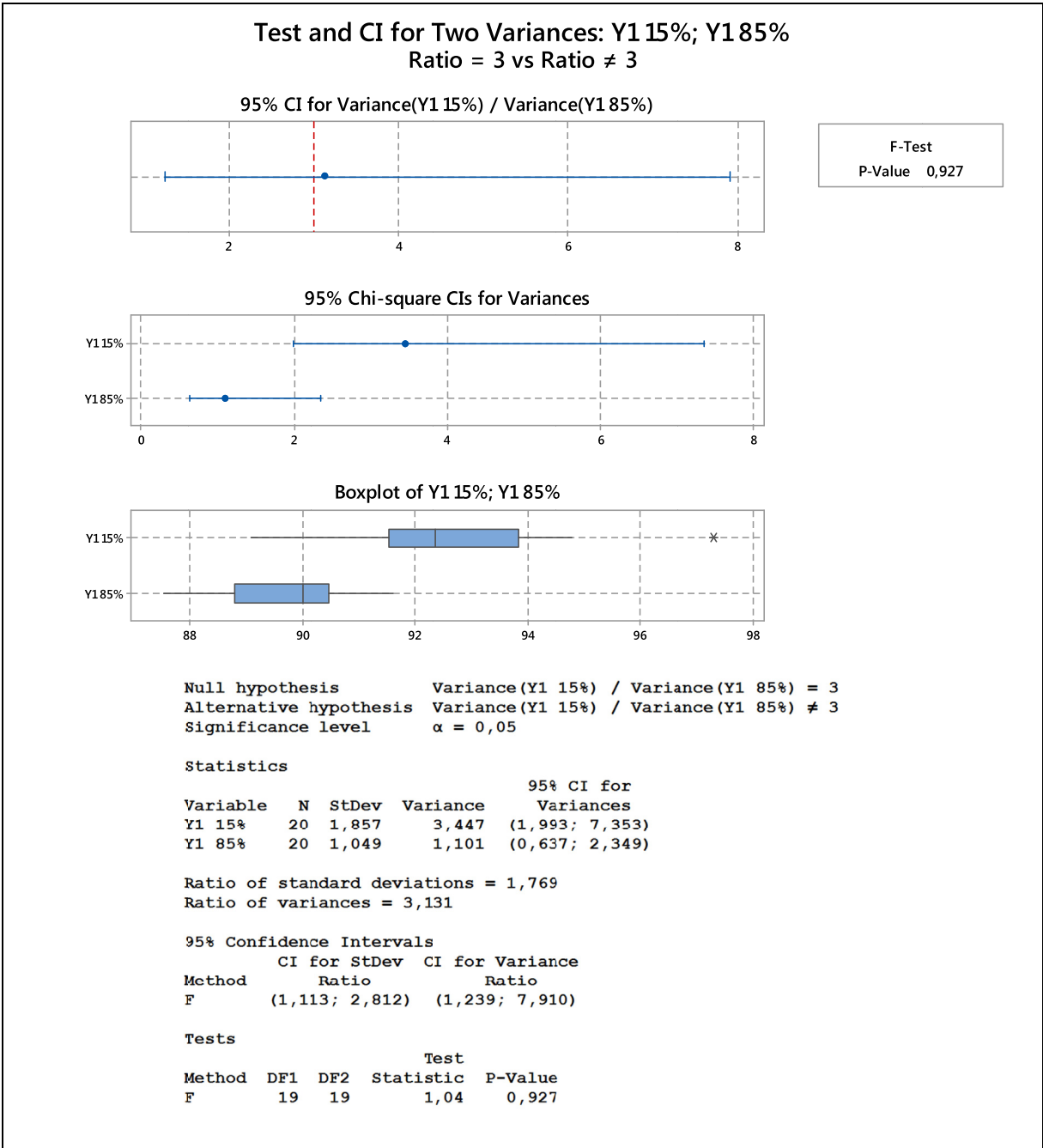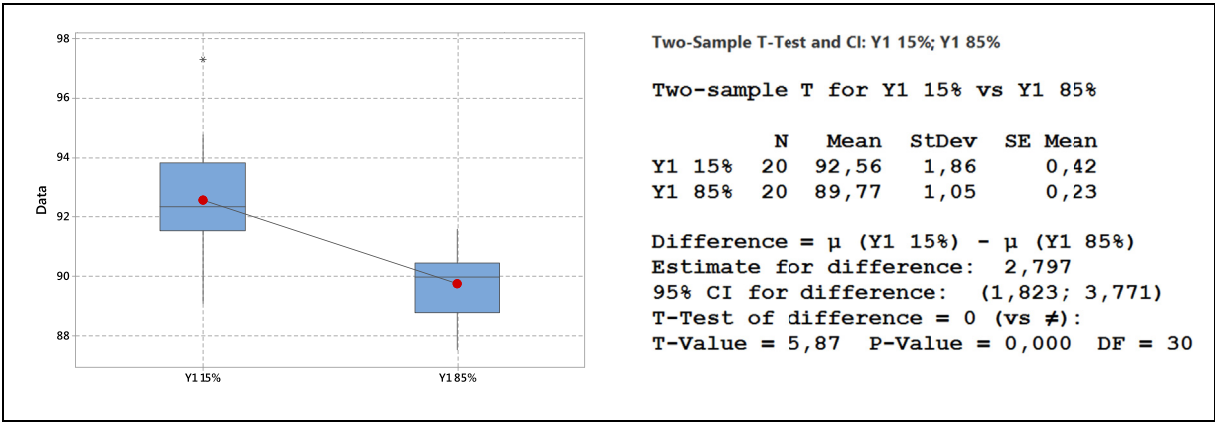
Fig. 19. Y1: 2-Sample t and 2-variances hypothesis tests for w = 15% and w = 85%.
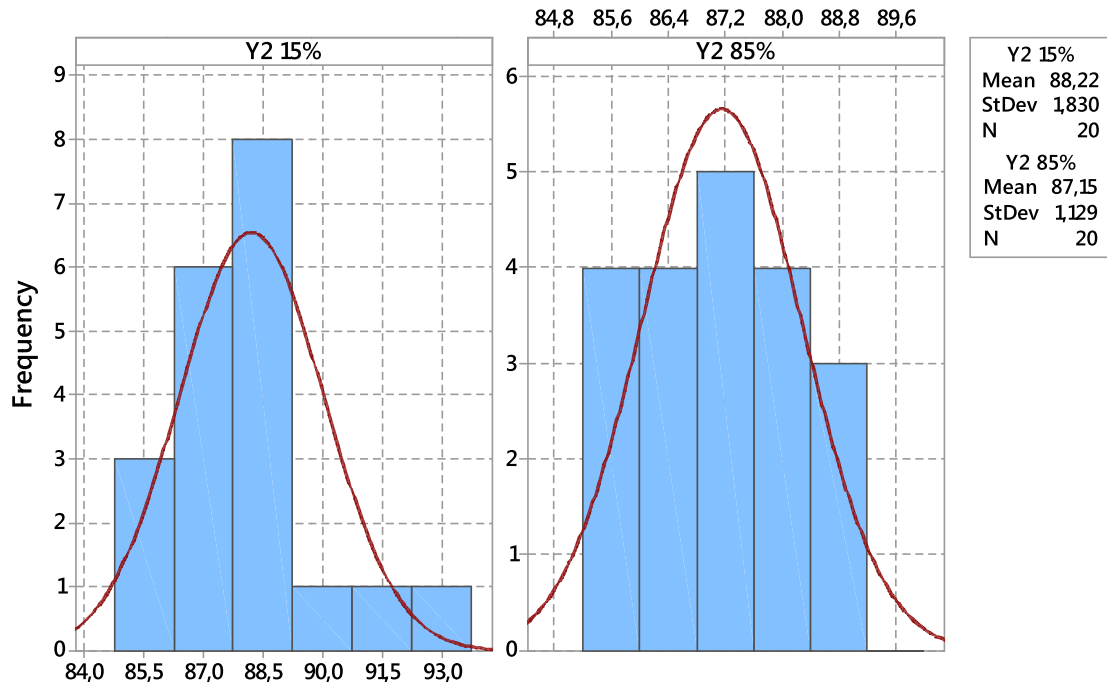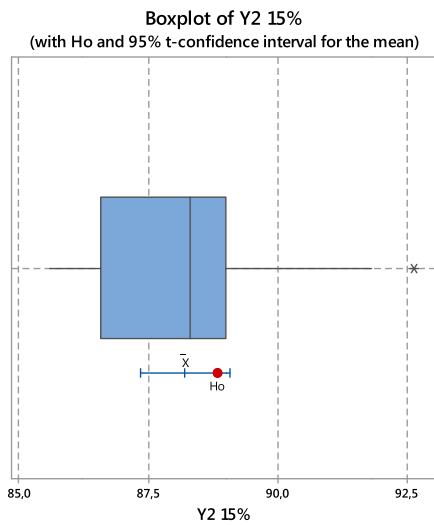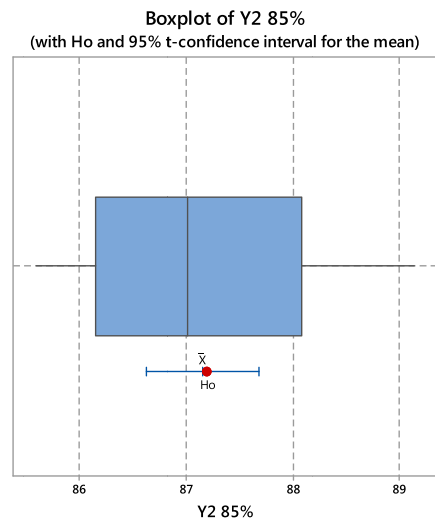
Fig. 20. Y2: Confirmation runs (a). One-sample t for w = 15% (b) and w = 85% (c).

Fig. 21. Y2: 2-Sample t and 2-variances hypothesis tests for w = 15% and w = 85%.
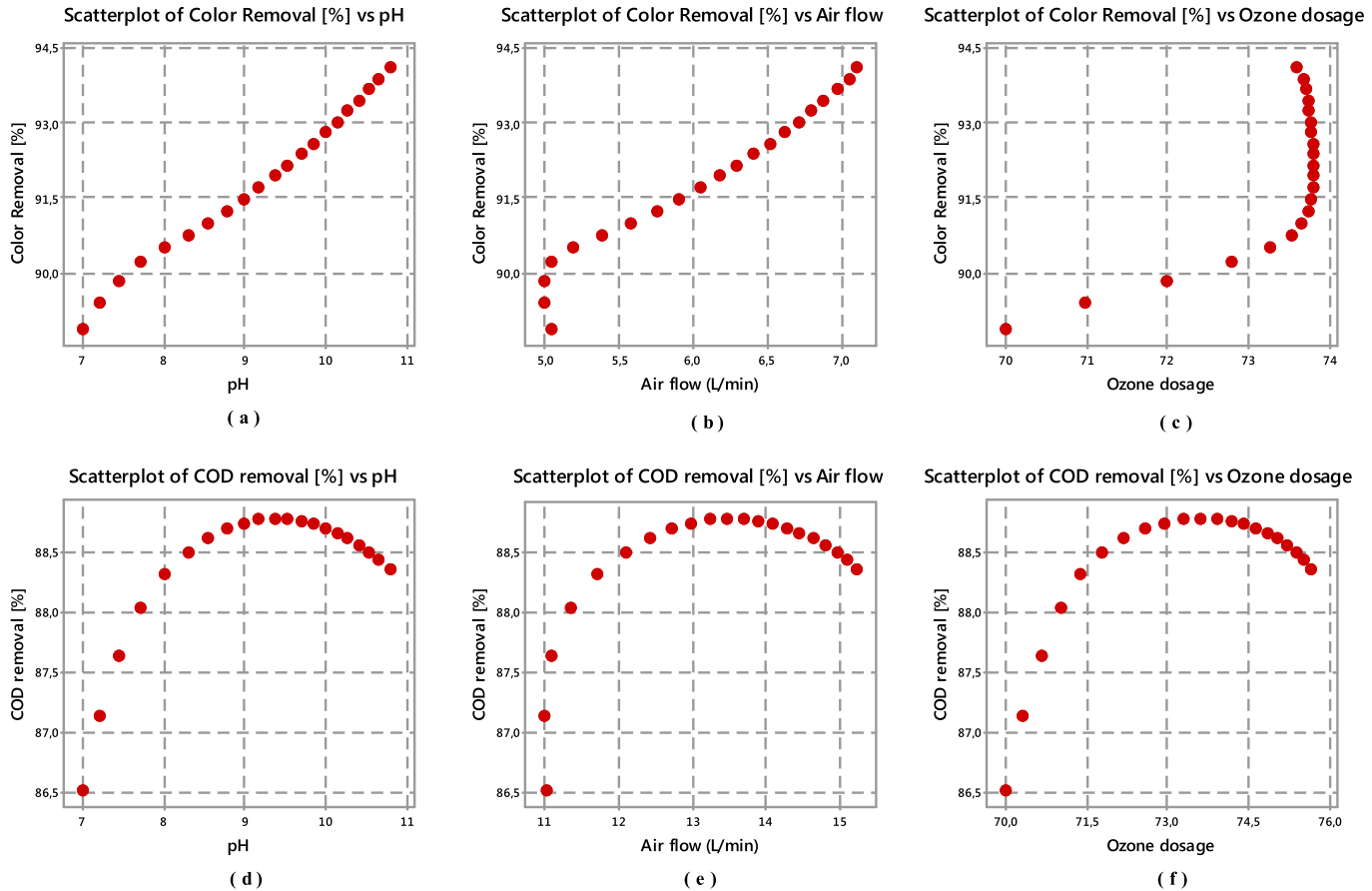
Fig. 22. Influence of pH, air flow and ozone dosage on color and COD removal.

$$\frac{dCA}{dt} = kCA[O_3] \tag{32}$$

Consequently, for elevated ozone concentrations, the percentage of color removal is reduced mainly due to the ozone consumption (direct reaction). Therefore, the reaction presents second-order kinetic, which means that the concentration of methyl orange solution varies exponentially and explains the slight bend in the interval of pH ranging from 7.0 to 7.4.

On the other hand, for a more alkaline pH, the concentration of O_3 in solution decreases, since it is converted into hydroxyl radicals. In this context, the ozone concentration in solution remains practically constant. Thus, it is possible to rewrite Eq. (32). In terms of a rate constant composed of the product of the rate constant k and the ozone concentration in the medium, as in Eq. (33), which characterizes a kinetically pseudo first order reaction.

$$\frac{dCA}{dt} = k'CA \quad \forall\ k \cdot [O_3] = k' \tag{33}$$

This approach can only be done since the rate constant k varies depending on the reaction temperature, which is characterized as constant throughout the experiments. The pseudo first order reaction explains the linear behavior observed in the second scenario of Fig. 22 (a), where, at higher pH, the color removal is linearly depending on the pH variation.

When the effect of the airflow in the color removal is analyzed, there is an effect of gas-to-liquid mass transfer, which can be described by Fick's law:

$$J = -D_A\frac{\partial C_A}{\partial x} \tag{34}$$

where J is the diffusion in the liquid in [mass].[time]$^{-1}$.[surface]$^{-1}$, D is the diffusion coefficient [surface]. [Time] $^{-1}$ and $\frac{\partial C_A}{\partial x}$ is the concentration variation along a column ∂x height. The ozone mass transfer is directly related to the reactor height in addition to the form of diffusion of this gas in the reaction medium. When the gas bubbles are considered macro bubbles in an open system at atmospheric pressure, there is a decrease of gas diffusion in the solution due to the decrease of the density of bubbles within the reactor, which can be considerably improved when the bubbles diameters are increased by mechanical, physical or chemical systems. In this study, bubbles were produced by physical process, consisting of a ceramic sphere, that generates a mixture of macro and micro bubbles and depends on the gas flow rate used. The non-significance of the airflow in this work can be explained due to the small height of the reactor column is not sufficient to compare the efficiency of mass transfer in color reduction.

It is also possible to evaluate the color removal behavior due to the ozone dosage. The ozone dosage depends on the power of the ozone generator and. according to Tang et al. (2009), variations in this power can affect the production of hydroxyl radicals and, consequently, the color removal. In order to understand the relationship between the power of the ozone generator and the ozone dosage, a calibration curve was built by applying an iodometric method in accordance with Standard Methods for the Examination of Water and Wastewater (American Public Health Association et al., 2012). This method consists in determining the
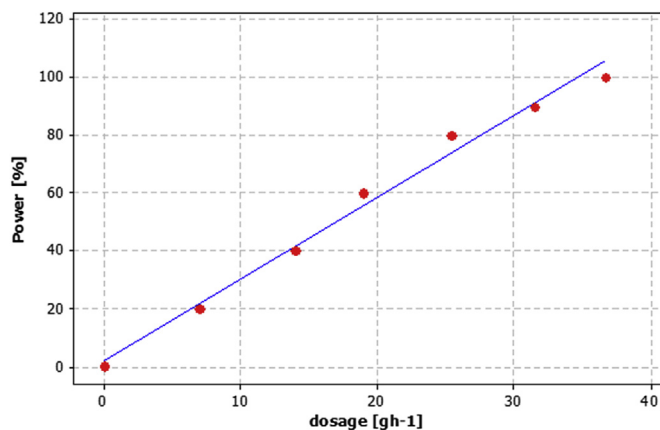
Fig. 23. Calibration curve for the ozone generator.

concentration of ozone gas by titration of a solution of potassium iodide in acidic medium of a sodium thiosulfate solution. The ozone generator power was controlled via a potentiometer with a percentage scale from 0 to 100%, corresponding to 0 and 220 W, respectively, that were determined based on the ozone dosages produced by the iodometric method. With the data collected, it was possible to build a calibration curve as shown in Fig. 23.

When a slight increase in the power of ozone generator occurs, the production of hydroxyl radicals increases as the ozone production is intensified. However, when the power is about 60 W, the production of hydroxyl radicals starts to decrease and, when the power is close to 127 W, the production of hydroxyl radical is practically undetectable (Tang et al., 2009). The Scatterplot in Fig. 22 (c) shows the behavior of color removal as a function of the ozone dosage, where it is also possible to notice two different stages. In the first part of the curve (with ozone dosage from 17.0 to 18.4), there is a gradual increase in the power of the generator and, consequently, the variation of color removal is due to the small but efficient ozone generation, explaining the large intervals in the graphic. This spacing decreases with the power increasing and it is possible to notice that values marginally above the optimal (18.4 g h^{-1}) does not necessarily causes a greater color removal.

The same analysis done for color removal can be performed for the COD removal using the data from the Pareto front (Table 6). The COD removal is chemically related to the removal of color, since the COD is only removed at an azo compound (in this case, methyl orange) if there is the splitting of the azo bond, which leads to color removal. Therefore, Fig. 22 (d) shows the relationship between pH and COD removal, where it can be seen a tendency of higher COD removal for pH ranging from 8.0 to 9.0. According to the optimization results, for w = 75%, this value approaches to 8.0. At this value of pH, as previously discussed, the indirect reaction is favored, leading to the formation of hydroxyl radicals that promotes the degradation of COD.

The relationship of O$_3$ dosage with COD removal has also a similar behavior, since its diffusion is directly linked to pH variations in the reaction medium, which can be seen in Fig. 22 (f). In this context the optimal ozone dosage can be detected around of 18.4 g h^{-1}, which corresponds to approximately 70% of the generator power, that is a range value also referenced in Tang et al. (2009).

7.6. Advantages and drawbacks of NBI-FA method

The theoretical results and the confirmation runs showed that NBI-FA is a feasible proposal to treat large scale multiobjective problems presenting some degree of correlation among several responses and promoting the reduction in prediction error of the optima. As could be seen in the literature review, a large deal of cases employing response surface modelling of wastewater treatments presented such characteristics. Analyzing the numerical results, it is possible to state the following advantages and drawbacks of NBI-FA method:

a) Advantages:
- NBI-FA method is capable of reducing the number of objective functions in a multiobjective optimization problem using the correlation or variance-covariance matrix of original responses;
- The use of rotated factor scores allows to obtain a Pareto frontier with uncorrelated axes;
- The method is able to build the frontier without inverting the original correlation among the original responses;
- Allows to obtain equispaced frontiers and solutions;
- Promotes a better exploitation of solution space;
- When treating with E[f(x)] and Var[f(x)], NBI-FA is able to find optimal solutions with the narrowest confidence intervals for the predictions.

b) Drawbacks
- Requires a larger number of orthogonal transformations, represented by the principal components and factor scores;
- The method only reduces the number of subproblems if the method is only able to find solutions if there are significant correlations between objective functions;
- Sometimes the sense of optimization of the components or factor scores will not be compatible with the direction of optimization of the original objective functions. In this case, some further transformation may be necessary, like signal-to-noise ratios or multivariate mean square error (Lopes et al., 2016; Costa et al., 2016).

8. Conclusions

Response surface methodology has been broadly used in the modelling and optimization of wastewater treatments. However, since there is no guarantee that convexity of the response surface models will be compatible with the desired direction of optimization, in many times the optimal setup defined for the wastewater treatment will not have a good predictability. So, taking this fact into consideration, this paper presented a bi-objective formulation for NBI extending its application to those problems related to expected values and prediction variances of response surfaces used to model and optimize the performance indexes for wastewater treatment. Using rotated factor scores of the original means and variances, the NBI-FA approach was capable of generate a wide set of feasible solutions. The main conclusions of this study may be summarized as follows:

- The rotated factor scores represented correctly the groups of means and variances in response surfaces problems;
- The varimax rotation produced factor scores with more similar weights (eigenvalues) than the unrotated one, which is important when the researcher desires to build a Pareto Frontier. Similar eigenvalues may describe independent and equally important axes for the Pareto Frontier;
- The NBI-FA method was capable of generate narrow confidence intervals for the prediction with satisfactory values for the wastewater treatment;
- Considering that the Pareto Frontier of four objective functions (two means and two variances) would request the perform of 1771 subproblems in the traditional NBI method, the same

problem was solved solving only 21 subproblems with increments of 5% with NBI-FA method;

- The maximum dye removal of Methyl Orange achieved when the response was optimized individually was 94.1% ± 4.3 and it was obtained at an optima point equals to $x^* = [9.8; 6.8\ l\ min^{-1}; 17.8\ g\ h^{-1}]$;
- The maximum oxygen demand removal obtained when this response was optimized individually was 88.4% ± 3.3 and this removal rate was obtained at an optima point equals to $x^* = [8.4; 6.5\ l.min-1; 19.26\ g\ h^{-1}]$.
- This individual optimum for these two response have presented the largest 95% prediction confidence intervals;
- Using the individual maximization of the response surface of the second rotated factor score (which represents both color and COD removal), the maximum dye removal of Methyl Orange achieved was 94.1% ± 4.3 and the maximum oxygen demand removal obtained when this response was optimized individually was 88.4% ± 5.3, obtained at an optima point equals to $x^* = [9.5; 7.1\ l\ min^{-1}; 18.4\ g\ h^{-1}]$;
- Based on the fuzzy membership of Pareto set, the narrowest 95% confidence intervals for dye removal and COD removal were, respectively, obtained by the convex and equispaced Pareto Frontier as $Y_1 = 90.5 \pm 2.2$ and $Y_2 = 88.3 \pm 2.7$ and it was produced by optimum $[7.9.\ 5.6\ l\ min^{-1}.\ 18.4\ g\ h^{-1}]$;
- These values are compatible with the process kinetic and the respective thermodynamic;
- Fuzzy membership criterion was important to select the best compromise Solution in the Pareto Set;
- Compared to weighted sums and global criterion method, it was possible to verify that only NBI was able to generate convex and equispaced Pareto frontiers;
- Besides, the traditional multiobjective methods did not respect the original variance-covariance structure, inverting the correlation of the original responses.
- Only NBI-FA method was capable of to generate equispaced Pareto frontiers with the same signal to the correlation between color and COD removals;
- Confirmation runs corroborated the good adequacy of present approach. Revealing very suitable values for removal rates with minimum prediction variances. These results are very important to plan and control industrial wastewater treatments, because promotes a more stable and optimal condition;
- Although these results are relative to the wastewater treatment, NBI-FA method could be extended for others processes that uses RSM with correlated responses, also presenting conflict between the optimum and the prediction variance of the several responses of interest.

Acknowledgement

References

Abidi, N., Errais, E., Duplay, J., Berez, A., Jrad, A., Schaefer, G., Ghazi, M., Semhi, K., Trabelsi-Ayadi, M., 2015. Treatment of dye-containing effluent by natural clay. J. Clean. Prod. 86, 432–440. http://dx.doi.org/10.1016/j.jclepro.2014.08.043.

Ahmadi, M., Vahabzadeh, F., Bonakdarpour, B., Mofarrah, E., Mehranian, M., 2005. Application of the central composite design and response surface methodology to the advanced treatment of olive oil processing wastewater using Fenton's peroxidation. J. Hazard. Mater. 123, 33–50.

Ahmadi, A., Kaymanesh, A., Siano, P., Janghorbani, M., Nezhad, A.E., Sarno, D., 2015a. Evaluating the effectiveness of normal boundary intersection method for short-term environmental/economic hydrothermal self-scheduling. Electr. Power Syst. Res. 123, 192–204.

Ahmadi, A., Masouleh, M.S., Janghorbani, M., Manjili, N.Y.G., Sharaf, A.M., 2015b. Short term multi-objective hydrothermal scheduling. Electr. Power Syst. Res. 121, 357–367.

Ahmadi, A., Moghimi, H., Nezhad, A.E., Agelidis, V.G., Sharaf, A.M., 2015c. Multi-objective economic emission dispatch considering combined heat and power by normal boundary intersection method. Electr. Power Syst. Res. 129, 32–43.

Akar, T., Turkyilmaz, S., Celik, S., Akar, S.T., 2016. Treatment design and characteristics of a biosorptive decolorization process by a green type sorbent. J. Clean. Prod. 112, 4844–4853.

American Public Health Association, American Water Works Association, Water Environment Federation, 2012. Standard Methods for the Examination of Water and Wastewater.

Anouzla, A., Abrouki, Y., Souabi, S., Safi, M., Rhbal, H., 2009. Color and COD removal of disperse dye solution by a novel coagulant: application of statistical design for the optimization and regression analysis. J. Hazard. Mater. 166, 1302–1306.

Arslan-Alaton, I., TUreli, G., Olmez-Hanci, T., 2009. Treatment of azo dye production wastewaters using Photo-Fenton-like advanced oxidation processes: optimization by response surface methodology. J. Photochem. Photobiol. A Chem. 202, 142–153.

Asfaram, A., Ghaedi, M., Agarwal, S., Tyagi, I., Gupta, V.K., 2015a. Removal of basic dye Auramine-O by ZnS: Cu nanoparticles loaded on activated carbon: optimization of parameters using response surface methodology with central composite design. RSC Adv. 5, 18438–18450. http://dx.doi.org/10.1039/C4RA15637D.

Asfaram, A., Ghaedi, M., Goudarzi, A., Rajabi, M., 2015b. Response surface methodology approach for optimization of simultaneous dye and metal ion ultrasound-assisted adsorption onto Mn doped Fe 3 O 4 -NPs loaded on AC: kinetic and isothermal studies. Dalt. Trans. 44, 14707–14723. http://dx.doi.org/10.1039/C5DT01504A.

Asfaram, A., Ghaedi, M., Hajati, S., Goudarzi, A., 2015c. Ternary dye adsorption onto MnO 2 nanoparticle-loaded activated carbon: derivative spectrophotometry and modeling. RSC Adv. 5, 72300–72320. http://dx.doi.org/10.1039/C5RA10815B.

Asfaram, A., Ghaedi, M., Hajati, S., Goudarzi, A., Bazrafshan, A.A., 2015d. Simultaneous ultrasound-assisted ternary adsorption of dyes onto copper-doped zinc sulfide nanoparticles loaded on activated carbon: optimization by response surface methodology. Spectrochim. Acta - Part A Mol. Biomol. Spectrosc. 145, 203–212. http://dx.doi.org/10.1016/j.saa.2015.03.006.

Asfaram, a., Ghaedi, M., Azqhandi, M.H.A., Goudarzi, a., Dastkhoon, M., 2016. Statistical experimental design, least squares-support vector machine (LS-SVM) and artificial neural network (ANN) methods for modeling the facilitated adsorption of methylene blue dye. RSC Adv. 6, 40502–40516. http://dx.doi.org/10.1039/C6RA01874B.

Asfaram, A., Ghaedi, M., Hajati, S., Goudarzi, A., Dil, E.A., 2017. Screening and optimization of highly effective ultrasound-assisted simultaneous adsorption of cationic dyes onto Mn-doped Fe3O4-nanoparticle-loaded activated carbon. Ultrason. Sonochem. 34, 1–12. http://dx.doi.org/10.1016/j.ultsonch.2016.05.011.

Asghar, A., Raman, A.A.A., Daud, W.M.A.W., 2015. Advanced oxidation processes for in-situ production of hydrogen peroxide/hydroxyl radical for textile wastewater treatment: a review. J. Clean. Prod. 87, 826–838.

Bagheri, A.R., Ghaedi, M., Asfaram, A., Jannesar, R., Goudarzi, A., 2016. Design and construction of nanoscale material for ultrasonic assisted adsorption of dyes: application of derivative spectrophotometry and experimental design methodology. Ultrason. Sonochem. http://dx.doi.org/10.1016/j.ultsonch.2016.09.008.

Boopathy, R., Sekaran, G., 2013. Electrochemical treatment of evaporated residue of soak liquor generated from leather industry. J. Hazard. Mater. 260, 286–295.

Brito, T.G., Paiva, A.P., Ferreira, J.R., Gomes, J.H.F., Balestrassi, P.P., 2014. A normal boundary intersection approach to multiresponse robust optimization of the surface roughness in end milling process with combined arrays. Prec. Eng. 38.

Charwand, M., Ahmadi, A., Heidari, A.R., Nezhad, A.E., 2015. Benders decomposition and normal boundary intersection method for multiobjective decision making framework for an electricity retailer in energy markets. IEEE Syst. J. 9, 1475–1484.

Chen, L.C., 2000. Effects of factors and interacted factors on the optimal decolorization process of methyl orange by ozone. Water Res. 34, 974–982.

Colla, V., Branca, T.A., Rosito, F., Lucca, C., Vivas, B.P., Delmiro, V.M., 2016. Sustainable Reverse Osmosis application for wastewater treatment in the steel industry. J. Clean. Prod. 130, 103–115. http://dx.doi.org/10.1016/j.jclepro.2015.09.025.

Corominas, L., Acuña, V., Ginebreda, A., Poch, M., 2013. Integration of freshwater environmental policies and wastewater treatment plant management. Sci. Total Environ. 445–446, 185–191. http://dx.doi.org/10.1016/j.scitotenv.2012.12.055.

Costa, D.M.D., Brito, T.G., Paiva, A.P., Leme, R.C., Balestrassi, P.P., 2016. A normal boundary intersection with multivariate mean square error approach for dry end milling process optimization of the AISI 1045 steel. J. Clean. Pr. http://dx.doi.org/10.1016/j.jclepro.2016.01.062 od.

Čuček, L., Martín, M., Grossmann, I.E., Kravanja, Z., 2014. Multi-period synthesis of optimally integrated biomass and bioenergy supply network. Comp. Chem. Eng. 66, 57–70.

Das, I., Dennis, J.E., 1998. Normal boundary intersection: a new method for generating the Pareto surface in nonlinear multicriteria optimization problems. SIAM J. Optim. 8, 631–657.

Dastkhoon, M., Ghaedi, M., Asfaram, A., Goudarzi, A., Langroodi, S.M., Tyagi, I., Agarwal, S., Gupta, V.K., 2015. Ultrasound assisted adsorption of malachite green dye onto ZnS: Cu-NP-AC: equilibrium isotherms and kinetic studies - response surface optimization. Sep. Purif. Technol. 156, 780–788. http://dx.doi.org/10.1016/j.seppur.2015.11.001.

Dil, E.A., Ghaedi, M., Ghaedi, A.M., Asfaram, A., Goudarzi, A., Hajati, S., Soylak, M.,

Agarwal, S., Gupta, V.K., 2016. Modeling of quaternary dyes adsorption onto ZnO-NR-AC artificial neural network: analysis by derivative spectrophotometry. J. Ind. Eng. Chem. 34, 186—197. http://dx.doi.org/10.1016/j.jiec.2015.11.010.

Dil, E.A., Ghaedi, M., Asfaram, A., Hajati, S., Mehrabi, F., Goudarzi, A., 2017. Preparation of nanomaterials for the ultrasound-enhanced removal of Pb2+ ions and malachite green dye: chemometric optimization and modeling. Ultrason. Sonochem. 34, 677—691. http://dx.doi.org/10.1016/j.ultsonch.2016.07.001.

Dixit, S., Yadav, A., Dwivedi, P.D., Das, M., 2015. Toxic hazards of leather industry and technologies to combat threat: a review. J. Clean. Prod. 87, 39—49.

Ge, D., Zeng, Z., Arowo, M., Zou, H., Chen, J., Shao, L., 2016. Degradation of methyl orange by ozone in the presence of ferrous and persulfate ions in a rotating packed bed. Chemosphere 146, 413—418.

Gomes, C.S., Piccin, J.S., Gutterres, M., 2016. Optimizing adsorption parameters in tannery-dye-containing effluent treatment with leather shaving waste. Proc. Saf. Environ. Protec. 99, 98—106.

Gottschalk, C., Libra, J.A., Saupe, A., 2010. Ozonation of Water and Waste Water: a Practical Guide to Understanding Ozone and its Application, second ed. John Wiley & Sons, Inc., New York.

Grabowski, L.R., Van Veldhuizen, E.M., Pemen, A.J.M., Rutgers, W.R., 2007. Breakdown of methylene blue and methyl orange by pulsed corona discharge. Plasma Sources Sci. Technol. 16, 226—232.

Izadbakhsh, M., Gandomkar, M., Rezvani, A., Ahmadi, A., 2015. Short-term resource scheduling of a renewable energy based micro grid. Energy 75, 598—660.

Ji, Q., Tabassum, S., Hena, S., Silva, C.G., Yu, G., Zhang, Z., 2016. A review on the coal gasification wastewater treatment technologies: past, present and future outlook. J. Clean. Prod. 126, 38—55. http://dx.doi.org/10.1016/j.jclepro.2016.02.147.

Jia, X., Zhang, T., Wang, F., Han, F., 2006. Multi-objective modeling and optimization for cleaner production processes. J. Clean. Prod. 14, 146—151.

Jin, Y., Wu, Y., Cao, J., Wu, Y., 2014. Optimizing decolorization of Methylene Blue and Methyl Orange dye by pulsed discharged plasma in water using response surface methodology. J. Taiwan Inst. Chem. Eng. 45, 589—595.

Jozić, S., Bajić, D., Celent, L., 2015. Application of compressed cold air cooling: achieving multiple performance characteristics in end milling process. J. Clean. Prod. 100, 325—332.

Johnson, R.A., Wichern, D., 2007. Applied Multivariate Statistical Analysis, sixth ed. Prentice-Hall, New Jersey.

Khayet, M., Zahrim, A.Y., Hilal, N., 2011. Modelling and optimization of coagulation of highly concentrated industrial grade leather dye by response surface methodology. Chem. Eng. J. 167, 77—83.

Körbahti, B.K., 2007. Response surface optimization of electrochemical treatment of textile dye wastewater. J. Hazard. Mater. 145, 277—286.

Kumar, R., Pal, P., 2013. Turning hazardous waste into value-added products: production and characterization of struvite from ammoniacal waste with new approaches. J. Clean. Prod. 43, 59—70.

Kushwaha, J.P., Srivastava, V.C., Mall, I.D., 2010. Organics removal from dairy wastewater by electrochemical treatment and residue disposal. Sep. Purif. Technol. 76, 198—205.

Li, H., Guo, J., Yang, L., Lan, Y., 2014. Degradation of methyl orange by sodium persulfate activated with zero-valent zinc. Sep. Purif. Technol. 132, 168—173.

Li, H., Li, Y., Xiang, L., Huang, Q., Qiu, J., Zhang, H., Sivaiah, M.V., Baron, F., Barrault, J., Petit, S., Valange, S., 2015. Heterogeneous photo-Fenton decolorization of Orange II over Al-pillared Fe-smectite: response surface approach, degradation pathway, and toxicity evaluation. J. Hazard. Mater. 287, 32—41.

Lim, C.L., Morad, N., Teng, T.T., Ismail, N., 2009. Treatment of Terasil red R dye wastewater using H2O2/pyridine/Cu(II) system. J. Hazard. Mater. 168, 383—389.

Liu, Y., Fu, J., Deng, S., Zhang, X., Shen, F., Yan, G., Peng, H., Zhang, Y., 2014. Degradation of basic and acid dyes in high-voltage pulsed discharge. J. Taiwan Inst. Chem. Eng. 45, 2480—2487.

Lopes, L.G.D., Brito, T.G., Paiva, A.P., Peruchi, R.S., Balestrassi, P.P., 2016. Robust parameter optimization based on multivariate normal boundary intersection. Comp. Ind. Eng. 93, 55—66.

Lu, L.A., Ma, Y.S., Kumar, M., Lin, J.G., 2011. Photo-Fenton pretreatment of carbofuran — analyses via experimental design. detoxification and biodegradability enhancement. Sep. Purif. Technol. 81, 325—331.

Lu, X.F., Ma, H.R., Zhang, Q., Du, K., 2013. Degradation of methyl orange by UV, O3 and UV/O3 systems: analysis of the degradation effects and mineralization mechanism. Res. Chem. Intermed. 39, 4189—4203.

Mavalizadeh, H., Ahmadi, A., Heidari, A., 2015. Probabilistic multi-objective generation and transmission expansion planning problem using normal boundary intersection. Gener. Transm. Distrib. IET 9, 560—570.

Meneses, M., Concepción, H., Vrecko, D., Vilanova, R., 2015. Life Cycle Assessment as an environmental evaluation tool for control strategies in wastewater treatment plants. J. Clean. Prod. 107, 653—661. http://dx.doi.org/10.1016/j.jclepro.2015.05.057.

Montgomery, D.C., 2009. Design and Analysis of Experiments, seventh ed. John Wiley & Sons, Inc., New York.

Mohajerani, M., Mehrvar, M., Ein-Mozaffari, F., 2011. Correlation and prediction of azo dye degradation by nonlinear least-square regression in combined ozonation and ultrasonolysis processes. Water. Qual. Res. J. Can. 46, 250—258.

Muhamad, M.H., Abdullah, S.R.S., Mohamad, A.B., Rahman, R.A., Kadhum, A.A.H., 2013. Application of response surface methodology (RSM) for optimisation of COD, NH3—N and 2.4-DCP removal from recycled paper wastewater in a pilot-scale granular activated carbon sequencing batch biofilm reactor (GAC-SBBR). J. Environ. Manag. 121, 179—190.

Myers, R.H., Montgomery, D.C., 2009. Response Surface Methodology: Process and Product Optimization Using Designed Experiments, second ed. John Wiley & Sons, New York.

Nair, A.T., Ahammed, M.M., 2015. The reuse of water treatment sludge as a coagulant for post-treatment of UASB reactor treating urban wastewater. J. Clean. Prod. 96, 272—281.

Oguz, E., Keskinler, B., 2008. Removal of colour and COD from synthetic textile wastewaters using O3, PAC, H2O2 and HCO3-. J. Hazard. Mater. 151, 753—760.

Pham, T.T.H., Brar, S.K., Tyagi, R.D., Surampalli, R.Y., 2010. Optimization of Fenton oxidation pre-treatment for B. thuringiensis — based production of value added products from wastewater sludge. J. Env. Manag. 91, 1657—1664.

Pi, K., Xia, M., Yang, X., Wu, P., Chen, S., Yang, M., Gerson, A.R., 2015. Optimization of COD decrease from tobacco wastewater by Ca/Mg/Al coagulant using RSM. J. Water Proc. Eng. 5, 166—171.

Pintilie, L., Torres, C.M., Teodosiu, C., Castells, F., 2016. Urban wastewater reclamation for industrial reuse: an LCA case study. J. Clean. Prod. 139, 1—14. http://dx.doi.org/10.1016/j.jclepro.2016.07.209.

Poznyak, T., Colindres, P., Chairez, I., 2007. Treatment of textile industrial dyes by simple ozonation with water recirculation. J. Mex. Chem. Soc. 51, 81—86.

Qu, X., Zheng, J., Zhang, Y., 2007. Catalytic ozonation of phenolic wastewater with activated carbon fiber in a fluid bed reactor. J. Colloid. Interface Sci. 309, 429—434.

Rao, S.S., 2009. Engineering Optimization — Theory and Practice, fourth ed. John Wiley & Sons, Inc., New York.

Robinson, T., McMullan, G., Marchant, R., Nigam, P., 2001. Remediation of dyes in textile effluent: a critical review on current treatment technologies with a proposed alternative. Bioresour. Technol. 77, 247—255. http://dx.doi.org/10.1016/S0960-8524(00)00080-8.

Rocha, L.C.S., Paiva, A.P., Balestrassi, P.P., Severino, G., Rotela Jr., P., 2015. Entropy-based weighting for multiobjective optimization: an application on vertical turning. Math. Probl. Eng. http://dx.doi.org/10.1155/2015/608325.

Saeed, M.O., Azizli, K., Isa, M.H., Bashir, M.J.K., 2014. 2014. Application of CCD in RSM to obtain optimize treatment of POME using Fenton oxidation process. J. Water Process Eng. http://dx.doi.org/10.1016/j.jwpe.2014.11.001.

Santín, I., Pedret, C., Vilanova, R., Meneses, M., 2016. Advanced decision control system for effluent violations removal in wastewater treatment plants. Cont. Eng. Pract. 49, 60—75. http://dx.doi.org/10.1016/j.conengprac.2016.01.005.

Šereš, Z., Maravić, N., Takači, A., Nikolić, I., Šoronja-Simović, D., Jokić, A., Hodur, C., 2016. Treatment of vegetable oil refinery wastewater using alumina ceramic membrane: optimization using response surface methodology. J. Clean. Prod. 112, 3132—3137.

Shahrezaei, F., Mansouri, Y., Zinatizadeh, A.A.L., Akhbari, A., 2012. Process modeling and kinetic evaluation of petroleum refinery wastewater treatment in a photocatalytic reactor using TiO2 nanoparticles. Powder Technol. 221, 203—212.

Sheydaei, M., Aber, S., Khataee, A., 2014. Preparation of a novel γ-FeOOH-GAC nano composite for decolorization of textile wastewater by photo Fenton-like process in a continuous reactor. J. Mol. Cat. A Chem. 392, 229—234.

Tak, B.Y., Tak, B.S., Kim, Y.J., Park, Y.J., Yoon, Y.H., Min, G.H., 2015. Optimization of color and COD removal from livestock wastewater by electrocoagulation process: application of Box—Behnken design (BBD). J. Ind. Eng. Chem. 28, 307—315.

Tang, Q., Jiang, W., Zhang, Y., Wei, W., Lim, T.M., 2009. Degradation of azo dye acid red 88 by gas phase dielectric barrier discharges. Plasma Chem. Plasma Process 29, 291—305.

Tasaki, T., Wada, T., Fujimoto, K., Kai, S., Oke, K., Oshima, T., Baba, T., Kukazaki, M., 2009. Degradation of methyl orange using short-wavelength UV irradiation with oxygen microbubbles. J. Hazard. Mater 162, 1103—1110.

Thirugnanasambandham, K., Sivakumar, V., Maran, J.P., 2015. Response surface modelling and optimization of treatment of meat industry wastewater using electrochemical treatment method. J. Taiwan Inst. Chem. Eng. 46, 160—167.

Torrades, F., García-Montaño, J., 2014. Using central composite experimental design to optimize the degradation of real dye wastewater by Fenton and photo-Fenton reactions. Dyes Pigments 100, 184—189.

Wang, Y., Chen, K., Mo, L., Li, J., Xu, J., 2014. Optimization of coagulation—-flocculation process for papermaking-reconstituted tobacco slice wastewater treatment using response surface methodology. J. Ind. Eng. Chem. 20, 391—396.

Yan, J., Li, L., 2013. Multi-objective optimization of milling parameters e the trade-offs between energy, production rate and cutting quality. J. Clean. Prod. 52, 462—471.

Yan, W., Zhang, H., Jiang, Z.G., Hon, K.K.B., 2016. Multi-objective optimization of arc welding parameters: the trade-offs between energy and thermal efficiency. J. Clean. Prod. http://dx.doi.org/10.1016/j.jclepro.2016.03.171.

Zhang, H., Duan, L., Zhang, D., 2006. Decolorization of methyl orange by ozonation in combination with ultrasonic irradiation. J. Hazard. Mater. 138, 53—59.

Zhang, H., Fu, H., Zhang, D., 2009a. Degradation of C.I. Acid Orange 7 by ultrasound enhanced heterogeneous Fenton-like process. J. Hazard. Mater 172, 654—660.

Zhang, L., Sun, B., Zhu, X., 2009b. Organic dye removal from aqueous solution by pulsed discharge on the pinhole. J. Electrostat 67, 62—66.

Zhang, Y., Sun, B., Deng, S., Wang, Y., Peng, H., Li, Y., Zhang, X., 2010. Methyl orange degradation by pulsed discharge in the presence of activated carbon fibers. Chem. Eng. J. 159, 47—52.

Zhang, Y., Zheng, J., Qu, X., Chen, H., 2007. Effect of granular activated carbon on degradation of methyl orange when applied in combination with high-voltage pulse discharge. J. Colloid Interface Sci. 316, 523—530.

Zhu, H., Guo, W., Shen, Z., Tang, Q., Ji, W., Jia, L., 2015. QSAR models for degradation of organic pollutants in ozonation process under acidic condition. Chemosphere 119, 65—71.