

# Use of genetic algorithms and gradient based optimization techniques for calcium phosphate precipitation

L. Montastruc, C. Azzaro-Pantel, L. Pibouleau\*, S. Domenech

*Laboratoire de Génie Chimique, UMR CNRS/INP/UPS 5503, ENSIACET, 118 Route de Narbonne, 31077 Toulouse Cedex 04, France*

---

## Abstract

Phase equilibrium computations constitute an important problem for designing and optimizing crystallization processes. The Gibbs free energy is generally used as an objective function to find phase amount and composition at equilibrium. In such problems, the Gibbs free energy may be a quite complex function, with several local minima. This paper presents a contribution to handle this kind of problems by implementation of an optimization technique based on the successive use of a genetic algorithm (GA) and of a classical sequential quadratic programming (SQP) method: the GA is used to perform a preliminary search in the solution space for locating the neighborhood of the solution. Then, the SQP method is employed to refine the best solution provided by the GA. The basic operations involved in the design of the GA developed in this study (encoding with binary representation of real values, evaluation function, adaptive plan) are presented. Several test problems are first presented to demonstrate the validity of the approach. Then, calcium phosphate precipitation which is of major interest for P-recovery from wastewater, has been chosen as an illustration of the implemented algorithm.

*Keywords:* Calcium phosphate precipitation; Gibbs free energy; Genetic algorithm; SQP

---

## 1. Introduction

Phase equilibrium calculations constitute an important class of problems in chemical engineering applications and considerable literature has been devoted to numerical optimization of vapor–liquid equilibrium [1,2]. Liquid–solid equilibrium modeling is now receiving much attention due to the advent of specialty chemicals, as well as to new constraints such as environmental considerations. An accurate knowledge of such equilibria is particularly of great opportunity for designing and optimizing crystallization processes.

This paper is devoted to calcium phosphate precipitation which has been identified as a major issue. Let us recall that phosphorus can be found under various chemical forms in urban wastewater, which represents about 30–50% of the total refusal of P: insoluble or dissolved organic phosphorus, orthophosphates (until 70% sometimes) and condensed inorganic phosphates. In France, the average concentration of phosphorus in domestic wastewater is within the range

of 15–25 mg/l, which may strongly vary from day to day, even during the day. The P-discharge in aqueous natural environment leads to an excessive development of algae and, generally to a pH increase, thus corresponding to eutrophication. Consequently, the phosphorus reduction in rivers is considered as a key factor of the fight against pollution. The principal legislative tool in Europe for fighting against eutrophication is the EC Urban Waste Water Treatment Directive (271/91/EEC). This action came into force in 1991 and enabled waterbodies to be classified as sensitive areas if they display symptoms of eutrophication.

Calcium phosphate precipitation involves various parameters: calcium and phosphate ion concentrations, supersaturation, ionic strength, temperature, ion types, pH and also time. The process studied in this paper is based on calcium phosphate precipitation obtained by mixing a phosphate solution with calcium ions and a base. More precisely, it involves a fluidized bed of sand continuously fed with aqueous solutions (see Fig. 1). Calcium phosphate precipitates upon the surface of sand grains. At the same time, small particles, i.e., “fines”, leave the bed with the remaining phosphate not recovered in the reactor. A layer of fines which has agglomerated is observed at the upper zone of the fluidized bed.

---

\* Corresponding author.

*E-mail address:* luc.pibouleau@ensiacet.fr (L. Pibouleau).

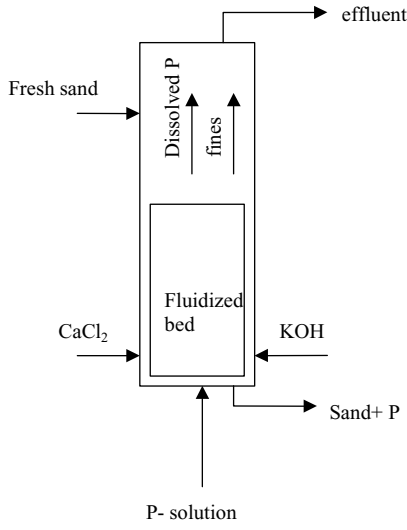


Fig. 1. Schematic representation of the pellet reactor.

Both total and dissolved concentrations of phosphorus, pH and temperature were measured at the outlet stream. The experimental results were previously presented in detail in [3].

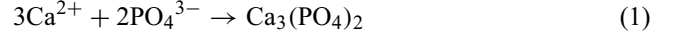
In fact, the nature of the calcium phosphate precipitate depends on the supersaturation of the various species. More precisely, the co-crystallization of amorphous calcium phosphate (ACP,  $\text{Ca}_3(\text{PO}_4)_2$ ) and dicalcium phosphate dihydrate (DCPD,  $\text{CaHPO}_4$ ) which may occur in the range of pH to be considered has been taken into account in this study [2].

This paper first presents the development of a simple thermochemical model, enough representative of calcium phosphate precipitation using a Debye–Huckel-based approach for activity coefficient modeling. A preliminary study has shown that the initialization problem of classical optimization techniques which may be used (for instance, sequential quadratic programming) is crucial for the robustness of the code. For this purpose, an optimization strategy combining a two-stage approach, i.e. a genetic algorithm for initialization and identification of the search zone followed by an SQP method to refine the solution is proposed in this paper. Several test problems have been used to demonstrate the validity of the approach. Then, the specific results obtained for calcium phosphate precipitation are discussed.

## 2. Chemical equilibrium model for calcium phosphate precipitation (system $\text{CaCl}_2\text{--H}_3\text{PO}_4\text{--KOH}$ )

To model the evolution of phosphate conversion rate as a function of pH with respect to ACP and DCPD precipitation, mass and electroneutrality conservation balances have been taken into account, as well as the supersaturation relative to each species. During this precipitation, the aqueous species considered are, on the one hand, for phosphoric acid,  $\text{H}_3\text{PO}_4$ ,  $\text{H}_2\text{PO}_4^-$ ,  $\text{HPO}_4^{2-}$ ,  $\text{PO}_4^{3-}$ , and on the other hand, the concentration of  $\text{Ca}^{2+}$  ion and the corresponding calcium salts.

The ACP precipitation equation can be written as follows:

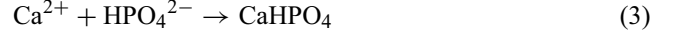


The ACP supersaturation is defined by the  $\beta$  parameter:

$$\beta_{\text{ACP}} = \frac{1}{5} \ln \left( \frac{([\text{Ca}^{2+}] \lambda_{\text{Ca}^{2+}})^3 ([\text{PO}_4^{3-}] \lambda_{\text{PO}_4^{3-}})^2}{K_{\text{SACP}}} \right) \quad (2)$$

where  $K_{\text{SACP}}$  represents the ACP solubility product.

The DCPD precipitation equation can be expressed as:



The DCPD supersaturation is defined by the following constraint:

$$\beta_{\text{DCPD}} = \frac{1}{2} \ln \left( \frac{([\text{Ca}^{2+}] \lambda_{\text{Ca}^{2+}}) ([\text{HPO}_4^{2-}] \lambda_{\text{HPO}_4^{2-}})}{K_{\text{SDCPD}}} \right) \quad (4)$$

where  $K_{\text{SDCPD}}$  represents the DCPD solubility product.

The different mass balances in the liquid phase include:

- a mass balance for calcium:

$$\begin{aligned} & [\text{Ca}^{2+}] + [\text{CaH}_2\text{PO}_4^+] + [\text{CaHPO}_4] \\ & + [\text{CaPO}_4^-] + [\text{CaOH}^+] \\ & = (\text{Ca}_{\text{Total}} - \frac{3}{2} \text{P}_{\text{total}} X_{\text{ACP}} - \text{P}_{\text{Total}} X_{\text{DCPD}}) \end{aligned} \quad (5)$$

where  $X_{\text{ACP}}$  and  $X_{\text{DCPD}}$  are the conversion ratios relative to ACP and DCPD forms, respectively, defined as:

$$X_{\text{ACP}} + X_{\text{DCPD}} = \frac{\text{P}_{\text{Total}} - \text{P}_{\text{sol}}}{\text{P}_{\text{Total}}} \quad (6)$$

- a mass balance for phosphate:

$$\begin{aligned} & [\text{H}_3\text{PO}_4] + [\text{H}_2\text{PO}_4^-] + [\text{HPO}_4^{2-}] + [\text{PO}_4^{3-}] \\ & + [\text{CaH}_2\text{PO}_4^+] + [\text{CaHPO}_4] + [\text{CaPO}_4^-] \\ & = \text{P}_{\text{Total}} (1 - X_{\text{ACP}} - X_{\text{DCPD}}) \end{aligned} \quad (7)$$

The electroneutrality requirement gives:

$$\begin{aligned} & [\text{H}_2\text{PO}_4^-] + 2[\text{HPO}_4^{2-}] + 3[\text{PO}_4^{3-}] + [\text{CaPO}_4^-] \\ & + [\text{Cl}^-] + [\text{OH}^-] = [\text{CaH}_2\text{PO}_4^+] + 2[\text{Ca}^{2+}] \\ & + [\text{CaOH}^+] + [\text{H}^+] + [\text{K}^+] \end{aligned} \quad (8)$$

The concentrations of ions and complexes are determined from chemical equilibrium relations (equilibrium constants are given for a temperature of 25 °C in molar units) (see Table 1).

The Debye–Huckel model giving the activity coefficient of each species is defined by:

$$\log_{10} \lambda = -A_{\text{DH}} z_i^2 \frac{\sqrt{I}}{1 + B_{\text{DH}} \alpha \sqrt{I}} + C_{\text{DH}} I \quad (9)$$

where

$$A_{\text{DH}} = \frac{1}{4\pi \ln 10} \left( \frac{e}{\sqrt{\epsilon k_{\text{B}} T}} \right)^3 \sqrt{\frac{\rho_0 N_{\text{A}}}{2}}$$

Table 1  
Equilibrium constants for the system Ca-PO<sub>4</sub>-H<sub>2</sub>O  $K_i = (A_i)(B_i)/(AB_i)$  [3]

$K_i$	$A_i$	$B_i$	$AB_i$	$K_i$ -value
$K_1$	H <sup>+</sup>	H <sub>2</sub> PO <sub>4</sub> <sup>-</sup>	H <sub>3</sub> PO <sub>4</sub>	$7.1285 \times 10^{-3}$
$K_2$	H <sup>+</sup>	HPO <sub>4</sub> <sup>-</sup>	H <sub>2</sub> PO <sub>4</sub> <sup>-</sup>	$6.2373 \times 10^{-8}$
$K_3$	H <sup>+</sup>	PO <sub>4</sub> <sup>-</sup>	HPO <sub>4</sub> <sup>-</sup>	$453.942 \times 10^{-15}$
$K_4$	Ca <sup>2+</sup>	H <sub>2</sub> PO <sub>4</sub> <sup>-</sup>	CaH <sub>2</sub> PO <sub>4</sub> <sup>+</sup>	$3.908 \times 10^{-2}$
$K_5$	Ca <sup>2+</sup>	HPO <sub>4</sub> <sup>-</sup>	CaHPO <sub>4</sub>	$1.8239 \times 10^{-3}$
$K_6$	Ca <sup>2+</sup>	PO <sub>4</sub> <sup>-</sup>	CaPO <sub>4</sub> <sup>-</sup>	$347.536 \times 10^{-9}$
$K_7$	Ca <sup>2+</sup>	OH <sup>-</sup>	CaOH <sup>+</sup>	$5.8884 \times 10^{-2}$
$K_w$	H <sup>+</sup>	OH <sup>-</sup>	H <sub>2</sub> O	$1.004 \times 10^{-14}$

$$B_{DH} = \sqrt{\frac{2e^2 N_A \rho_0}{\epsilon k_B T}}$$

and  $C_{DH}$  is a constant equal to 0.055 mol/l.

Note that the distance  $\alpha$  is not the same for all ions in the system (Table 2).

The function to be minimized is the Gibbs free energy  $G$  of the system expressed as a linear combination of the chemical potential of each component in each phase:

$$G = \sum_{i=1}^N \sum_{k=1}^{\pi} n_{ik} \mu_{ik} \quad (10)$$

$$\mu_{iL} = \mu_{iL}(T) + RT \ln([x_i] \lambda_{x_i}) \quad (11)$$

$$\mu_{iS} = \mu_{iS}(T) \quad (12)$$

$\mu_{iL}$  and  $\mu_{iS}$  represent the chemical potentials of the species  $i$  in the liquid and solid phases, respectively (see Table 3).

A substitution method has been applied, so that the only unknowns of the system are now the following concentrations [Ca<sup>2+</sup>], [PO<sub>4</sub><sup>3-</sup>], [H<sup>+</sup>] and the phosphate conversion rate under DCPD and/or ACP forms (see Appendix A). The system is solved for various concentrations in KOH in order to analyze pH influence on conversion. Since calcium exists in the form of calcium chloride, this concentration has been taken equal to 2[Ca<sup>2+</sup>].

To solve the system, the two cases of calcium phosphate precipitation have been dissociated, thus leading to an optimization problem with simultaneously four equations and one inequality constraint. They are summarized in Table 4.

Table 2  
Radius of hydrated ions  $\alpha$  in Å [4]

Species	Radius of the hydrated ions $\alpha$ (Å)
H <sub>2</sub> PO <sub>4</sub> <sup>-</sup>	4.0
HPO <sub>4</sub> <sup>2-</sup>	4.0
PO <sub>4</sub> <sup>3-</sup>	4.0
CaH <sub>2</sub> PO <sub>4</sub> <sup>+</sup>	5.4
CaPO <sub>4</sub> <sup>-</sup>	4.0
Ca <sup>2+</sup>	6.0
OH <sup>-</sup>	3.5
H <sup>+</sup>	9.0

Table 3  
Chemical potential values at 25 °C

Species	$\mu_i$ (kJ/mol)
H <sub>3</sub> PO <sub>4</sub>	-1142.65
H <sub>2</sub> PO <sub>4</sub> <sup>-</sup>	-1130.40
HPO <sub>4</sub> <sup>2-</sup>	-1089.26
PO <sub>4</sub> <sup>3-</sup>	-1018.80
CaH <sub>2</sub> PO <sub>4</sub> <sup>+</sup>	-1691.96
CaHPO <sub>4</sub>	-1658.30
CaPO <sub>4</sub> <sup>-</sup>	-1609.20
Ca <sup>2+</sup>	-553.54
OH <sup>-</sup>	-157.29
H <sup>+</sup>	0.0
H <sub>2</sub> O	-237.18
ACP	-1680.47
DCPD	-3844.92

Table 4  
Different cases considered in the optimization method

	Sub problem 1	Sub problem 2
ACP supersaturation	$\leq 0$	0
DCPD supersaturation	0	$\leq 0$

In the former case, DCPD (and in the latter case ACP) is the equilibrium component of the system.

The problem contains:

- 13 bounded variables
- 11 equality constraints
- One inequality constraint
- One function to minimize

### 3. System solution

#### 3.1. General principles

Several approaches have been proposed for the computation of the solutions to the phase and chemical equilibrium problem (see [1] for a review of the different contributions). A typical feature of such problems is that the generation of starting points, that are used to perform the search with conventional optimization methods (for example SQP), is a basic point to guarantee the success of the optimization procedure.

A preliminary study on the above mentioned example has shown that the use of a SPQ method only (SQP package from IMSL library) [5] is very sensitive to the choice of the initial guess and may often lead to a convergence failure. Consequently, this work is motivated by the development of a technique for the automatic generation of good starting points. For this purpose, a hybrid optimization method is proposed in this paper. A genetic algorithm (GA) is used to perform a preliminary search in the solution space and to locate the neighborhood of the solution. Then, using the best solution found with the GA as initial guess, a gradient-based

optimization method is implemented to quickly converge towards the optimum. The optimization tool (SQP) is thus used to refine the GA solution.

### 3.2. Genetic algorithm

Genetic algorithms are procedures based on the mimetics of mechanics of natural selection and genetics. Theoretically developed by Holland [6], genetic algorithms emulate the biologist evolutionary theory to solve optimization problems. They compute a set of individuals, a population, i.e., evolving through a set of biologically inspired operators constituting the reproduction scheme. In this way, new individuals are generated from parents. According to the evolutionary theory, only the most suited elements of a population can survive and generate offspring, thus transmitting their biological heredity to new generations. The heredity is enclosed in the chromosomes of individuals represented in an optimization problem by a specific numerical (often binary) code. The suitability of each element according to the optimization problem under consideration, is evaluated via a fitness value directly derived from the objective function. The evolution mechanisms are constituted by three specific reproduction procedures, i.e., selection, cross-over and mutation. The cycle of evolution is generally repeated until a predefined number of generations is reached.

#### 3.2.1. Genetic algorithm procedure

The genetic algorithm procedure can be summarized as follows:

- Generation of the initial population
  - Estimation of the fitness of the initial population
- While the total number of generations is not reached,  
Generate the offspring population
- Selection of individuals surviving according to a survival rate
  - Synthesis of offspring obtained through cross-over
  - Mutation of individuals in the entire population through a mutation rate
  - Duplication of the best individual found in the next generation via elitism

End while.

#### 3.2.2. Fitness evaluation

Like classical optimization methods, GAs also face difficulties in handling constraints. Most of GAs implementations on constrained optimization use the penalty method proposed by Goldberg [7]; the fitness function  $F(x)$  is then defined as follows:

$$F(x) = f(x) - R \left( \sum_{j=1}^{m_e} [g_j(x)]^2 + \sum_{j=m_e+1}^m [\max\{0, g_j(x)\}]^2 \right) \quad (13)$$

In this expression,  $f(x)$  represents the objective function and the two penalty terms are relative to the equality and inequality constraint violations, where  $R$  is a penalty coefficient.

In the calcium phosphate precipitation problem, the fitness function ( $F$ ) involves the objective function (Gibbs free energy) and the penalty terms ( $p$ ) denoting the equality ( $h_l(x)$ ) and inequality constraints ( $g_k(x)$ ) violations.

$$F = -G - RC \quad (14)$$

$$C = \sum_{k=1}^p [\max\{0, g_k(x)\}]^2 + \sum_{l=1}^m [h_l(x)]^2 \quad (15)$$

#### 3.2.3. Genetic coding—chromosome representation

Each problem solution, i.e., a vector consisting of continuous variables is represented by a string, denoted as a chromosome. The chromosome length depends on the number of variables of the problem considered. The variables, which are coded using a binary representation, are defined in the following way:

- the continuous variables are represented by a determined number of bits which defines the real number accuracy.
- the chromosomes contain for each variable the sign, mantissa, exponent of each real.

The binary code proposed in this paper is called the weight box. It consists in encoding each digit of real in four bits with respectively the following weights 1, 2, 3, 3 (sum of weights is equal to 9) with a precision of  $10^{-20}$ . The choice of the weight box may appear quite artificial, but it has been achieved after a sensitivity analysis on numerous numerical problems. For the sake of illustration, the real 0.185 is encoded 1000 0111 0110, the zero digit and the point being not coded (see Fig. 2).

For example, the calcium phosphate precipitation problem considered here contains five variables. All these variables belong to the interval  $]0, 1[$  and the required precision is  $10^{-20}$ . Each variable is represented by 80 bits ( $20 \times 4$  bits by integers) and so the chromosomes contain 400 genes (Fig. 3).

#### 3.2.4. Generation of the initial population

The initial population, synthesized through a random number generator, presents some typical features. Since, the problem contains both equality and inequality constraints.

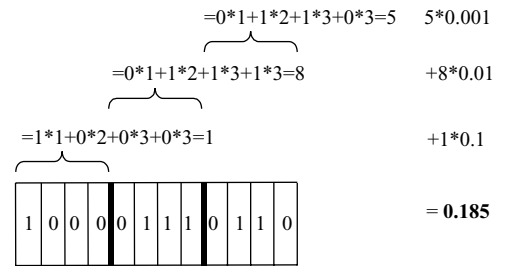


Fig. 2. Coding example with the weight box.

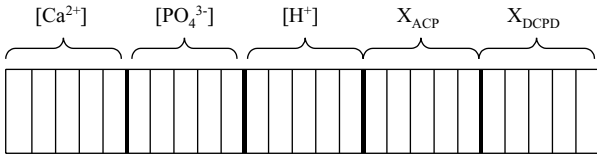


Fig. 3. Chromosome representation.

So, on the one hand, some variables are generated by a random generator, whereas the others are deduced from equality constraints by the use of a substitution procedure. On the other hand, all individuals in the initial populations must satisfy the inequality constraints with a given threshold ( $10^{-6}$ ). If it is not the case, the concerned individual is not accepted in the initial population.

### 3.2.5. Selection process

The selection process implies a fixed percentage (survival rate) of individuals that survive in the new generation. The selection is performed with the classical Goldberg biased roulette wheel [6]. In this way, the probability ( $\text{prob}_i$ ) that the individual  $i$  survives is given by:

$$\text{prob}_i = \frac{F_i}{\sum_{j=1}^{N_{\text{pop}}} F_j} \quad (14)$$

$N_{\text{pop}}$  represents the individuals number in the population. The survival rate is one of the important features of a GA implementation. If this rate is too low, the procedure may evolve too slowly, requiring an important number of generations for scanning a significant part of the search space. In turn, a too high survival rate may brew too violently the populations, destroying irremediably good solutions. So, an intermediate value of 0.6 was chosen as a default parameter.

### 3.2.6. Crossover

Once the surviving individuals are determined, the population is completed with new individuals obtained through cross-over mechanisms performed on two parents randomly paired in the whole current population. The crossover procedure adopted here is the classical one point permutation operation [7].

### 3.2.7. Mutation

Like the crossover procedure, the mutation operation is performed on the entire population with a fixed percentage for the mutation rate. The mutation technique chosen here consists in replacing a randomly selected gene by its binary complement [7]. The mutation rate constitutes another key point in a GA implementation, which creates genetic diversity in the current population. The mutation rate should remain rather low in order not to disturb too much the algorithm evolution. However, a too low rate value should be inefficient to introduce new individuals in terms on genetic codes in the population. A commonly used value of 0.1 was chosen in this study.

### 3.2.8. Elitism strategies

In the GA implemented in this work, the best individual of the current population in terms of the objective function, is systematically duplicated in the following generation. This very classic procedure is commonly called “elitism”. The use of elitism presents the advantage to guarantee the survival of the genes of the strongest individual, by favoring harmonious evolution (that is without excess of risky new development in the solutions space). This advantage yet represents a weapon with double sharp edge, because one favors the local investigation to the detriment of the global perspective, that can give premature convergence towards a local optimum.

## 4. Method validation

The method has first been validated on mathematical examples with known solutions taken from the literature (see Appendix B). Table 5 presents the numbers of the bounded variables and the number of both equality or inequality constraints.

For example 1, the fitness evolution versus generation number is shown in Fig. 4, where the optimum is represented by a dotted line. The value of the GA parameters are given in Table 6. Like any stochastic procedure to solve a given problem, the GA has to be run 60 times in this study with different initial populations.

From Tables 5 and 6, it can be observed that the survival and mutation rates do not have a significant influence on the method convergence which is confirmed by 60 runs of the GA + SQP coupling.

As in the previous case for examples 2 and 3, the survival and mutation rates do not have a major influence on the success of method. However, for a survival rate equal to 0.6 and a mutation rate equal to 0.4, the best success rate value (near to 85%) is obtained. Moreover, for example 2, the maximum generation number has no significant influence on the success rate, whereas for example 3 if the maximum generation number is decreased from 1000 to 500, the success rate decreases by 15% (Tables 7 and 8).

Compared with a classical GA implementation, some specific features have been introduced in the procedure used in this study. On the one hand, some variables are classically

Table 5  
Constraints and variables numbers for each example

	Example number		
	1	2	3
Optimization problem	Drying process	Alkylation process	Chemical complex mixture
Number of variables	2	10	10
Number of equality constraints	0	6	3
Number of inequality constraints	3	0	0

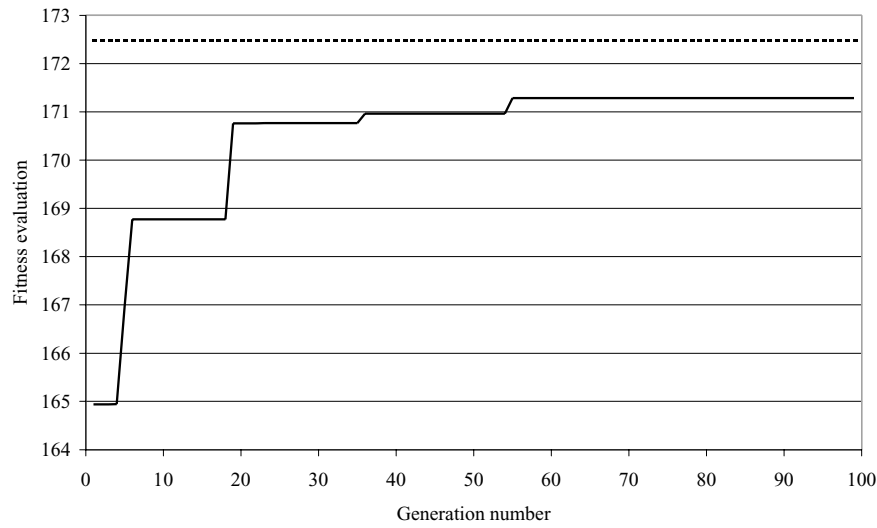


Fig. 4. Fitness evolution vs. generation number (example 1).

Table 6  
Values of genetic algorithm parameters (example 1)

	Example number		
	1	2	3
Total number of variables	2	10	10
Randomly generated	2	4	7
Deduced from the constraints	0	6	3
Number of randomly generated individuals to create the initial population	200	550	70000
Constraint value for the initial population	$1 \times 10^{-8}$	$1 \times 10^{-30}$	$1 \times 10^{-3}$
Survival rate	0.6	0.6	0.6
Mutation rate	0.1	0.4	0.4
Maximum number of generations	100	100	1000
Population size	100	100	100

Table 7  
Success rate for different survival rates and for a mutation rate fixed at 0.1 (example 1)

Survival rate	Success rate (%)
0.5	86.44
0.6	92.19
0.7	85.41
0.8	86.44

Table 8  
Success rate observed for different mutation rates and for a survival rate fixed at 0.6 (example 1)

Mutation rate	Success rate (%)
0.05	94.20
0.10	92.19
0.20	89.74
0.30	90.91
0.40	87.87

Table 9  
Success rates and CPU times for the three examples

Example number	Success rate (%)	CPU time (s)
1	92.19	2
2	83.33	3
3	85.00	40

generated through a random function, whereas others are deduced from the constraints. Each constraint value is determined so that the individual initial population number is not too high in particular for the problem 3. All initial population variables respect the interval limits. The method of the initial population generation allows to reduce the number of individuals generated by the random function (a decrease of 40% concerning the CPU time is observed). On the other hand, the initial populations satisfy the constraints at a given threshold. This procedure allows to reach a short CPU time (see Table 9 for example).

The best individual obtained from the GA will now serve as an initial point for the SQP method. It is important to note that the best individual constitutes the so-called scale factor required by the SQP method.

## 5. Application to calcium phosphate precipitation—results and discussion

For the calcium phosphate precipitation problem, the number of variables has been reduced by a substitution method from 13 to 5 which thus presents some advantages and drawbacks. Of course, it can be said that the calcium phosphate precipitation problem requires two different ways of handling the model equations, i.e., GA reduced set and SQP full set. For the genetic algorithm, it can be considered as an asset since the number of initial variables to be generated by the random function is reduced and the

Table 10  
Success rate vs. genetic algorithm parameters (calcium phosphate precipitation example)

Maximum number of generations	Survival rate	Mutation rate	Success rate
1000	0.4	0.4	56/60 as 93.33%
1000	0.5	0.4	52/60 as 86.66%
1000	0.6	0.4	57/60 as 95.00%
1000	0.7	0.4	54/60 as 90.00%
1000	0.8	0.4	53/60 as 88.33%
1000	0.6	0.3	52/60 as 86.66%
1000	0.6	0.2	56/60 as 93.33%
1000	0.6	0.1	50/60 as 83.33%
500	0.6	0.4	50/60 as 83.33%

different individual chromosomes are smaller than 400 bits (instead of 1040 bits when each variable is represented by 80 bits) which consequently decreases also the CPU time. Whereas for the SQP method, this reduction leads to an ill-conditioned problem, so the problem is solved with the original 13 variables. The best success rate has reached up to 95% for 60 runs.

The survival rate has not a major influence on the convergence, but the maximal success rate is obtained for a survival rate value equal to 0.6. In Table 10, it can also be seen that the mutation rate has a significant impact. In a general way, if this rate decreases, the success rate decreases too, due to a poor scanning of the search space. For the same scanning reason, the maximum generation number plays an important role on the convergence.

Some additional GA runs have been carried out with the following “worst” values for the genetic algorithm parameters: maximum generation number equal to 500, 0.6 for the survival rate, and 0.1 for the mutation rate. As expected, these values lead to the lowest success rate for the method (success rate is equal to 0 for 20 runs). This means that the

Table 11  
Gibbs free energy and penalty term evolution during the two-step solution procedure (calcium phosphate precipitation example)

	Gibbs free energy (J/mol)	Penalty term
Initialization of the GA procedure	-6236.95	$5 \times 10^{-6}$
Best solution obtained by the GA and used as an initialization for the SQP procedure	-4573.10	$1.2 \times 10^{-7}$
Best solution found by SQP	-5425.07	$10^{-32}$

search space scanning must be as large as possible to guarantee a good level of the success rate.

The numerical values obtained for the Gibbs free energy and the sum of constraints are presented in Table 11 ( $R = 10^{-6}$ ), where on the one hand, the results obtained after the GA phase and, on the other hand, those obtained at the end of the GA + SQP phase, are given. At step 2 of procedure, the SQP method is initialized with the best solution (i.e. best individual) found by the GA. It must be noted, that the best solution obtained from SQP exhibit a higher value for the Gibbs free energy higher than the initialization value used with GA, but the penalty term is smaller.

Experimental points have been used to validate the approach (see Fig. 5). They correspond to an initial phosphorus concentration of 50 mg/l, a Ca/P molar ratio equal to 3 and a temperature equal to 20 °C. It can be seen that a good agreement is observed between experimental and predicted values (see Fig. 5) [3].

## 6. Conclusions

In this paper, a hybrid optimization procedure combining a Genetic Algorithm and a SQP method has been developed and tested on a solid-liquid equilibrium

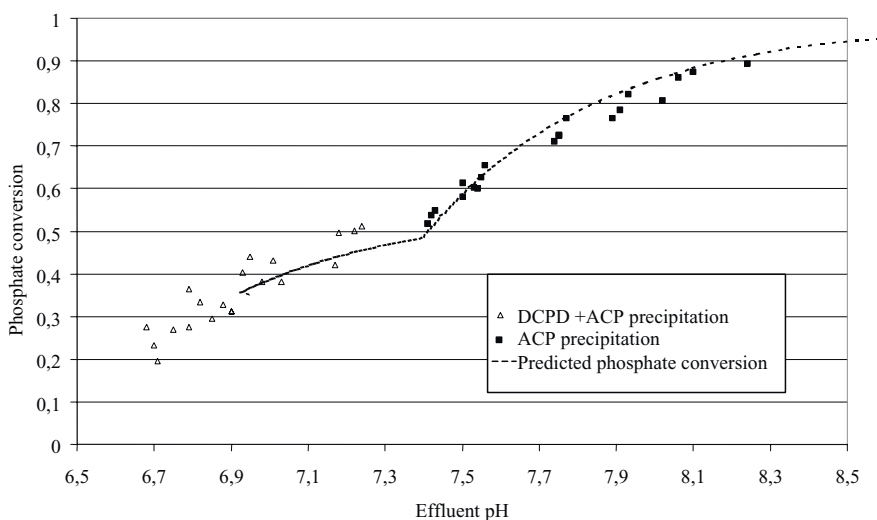


Fig. 5. Phosphate conversion vs. effluent pH.

optimization problem. Let us recall that genetic algorithms [6] differ from most classical optimization methods since no assumption about the problem mathematical properties is required. GAs use a guided random search in which many different solutions to a problem are investigated and refined simultaneously to identify near-optimum solutions. A major interest of such methods is that they lead to reasonable solutions, even starting from poor initial guesses. A GA implementation requires the definition of some parameters, i.e., population generation mode, population size, (i.e., the number of individuals forming a population which must be sufficiently large to create genetic diversity, in order to cover as well as possible the solution space) crossover probability, mutation rate, survival, crossover and mutation mechanisms. The two-level strategy leads to an efficient search: the GA provides good starting points for the subsequent SQP method, thus favoring the local search.

Calcium phosphate precipitation which is of major interest for P-recovery from wastewater has been chosen as an illustration of the procedure. The two-stage optimization strategy presented in this study may have widespread applications for solid-liquid equilibrium calculations.

#### Appendix A. Gibbs energy minimization substitution

$x(1) = [\text{Ca}^{2+}]$ ,  $x(2) = [\text{PO}_4^{3-}]$ ,  $x(3) = [\text{H}^+]$ ,  $x(4) = X_{\text{ACP}}$ ,  $x(5) = X_{\text{DCPD}}$   
 $[\text{K}^+] = 2.92 \times 10^{-3} \text{ mol/l}$ ,  $P_{\text{Total}} = 1.6 \times 10^{-3} \text{ mol/l}$ ,  
 $\text{Ca}_{\text{Total}} = 4.8 \times 10^{-3} \text{ mol/l}$ ,  $[\text{Cl}^-] = 9.6 \times 10^{-3} \text{ mol/l}$

$$\begin{aligned} \min G = & \left( \frac{x(3)^3 x(2) \lambda_{\text{H}}^3 \lambda_{\text{PO}_4}}{K_3 K_2 K_1} \right) \left( -1142.65 \times 10^3 + 298.15 R \ln \left( \frac{x(3)^3 x(2) \lambda_{\text{H}}^3 \lambda_{\text{PO}_4}}{K_3 K_2 K_1} \right) \right) \\ & + \left( \frac{x(3)^2 x(2) \lambda_{\text{PO}_4} \lambda_{\text{H}}^2}{K_3 K_2 \lambda_{\text{H}_2\text{PO}_4}} \right) \left( -1130.40 \times 10^3 + 298.15 R \ln \left( \frac{x(3)^2 x(2) \lambda_{\text{PO}_4} \lambda_{\text{H}}^2}{K_3 K_2} \right) \right) \\ & + \left( \frac{x(3) x(2) \lambda_{\text{H}} \lambda_{\text{PO}_4}}{K_3 \lambda_{\text{HPO}_4}} \right) \left( -1189.26 \times 10^3 + 298.15 R \ln \left( \frac{x(3) x(2) \lambda_{\text{H}} \lambda_{\text{PO}_4}}{K_3} \right) \right) \\ & + x(2) (-1018.80 \times 10^3 + 298.15 R \ln(x(2) \lambda_{\text{PO}_4})) + \left( \frac{x(1) x(3)^2 \lambda_{\text{H}}^2 \lambda_{\text{PO}_4} \lambda_{\text{Ca}}}{K_4 \lambda_{\text{CaH}_2\text{PO}_4} K_3 K_2} \right) \\ & \times \left( -1691.96 \times 10^3 + 298.15 R \ln \left( \frac{x(1) x(3)^2 \lambda_{\text{H}}^2 \lambda_{\text{PO}_4} \lambda_{\text{Ca}}}{K_4 K_3 K_2} \right) \right) \\ & + \left( \frac{x(1) x(3) x(2) \lambda_{\text{H}} \lambda_{\text{PO}_4} \lambda_{\text{Ca}}}{K_3 K_5} \right) \left( -1658.30 \times 10^3 + 298.15 R \ln \left( \frac{x(1) x(3) x(2) \lambda_{\text{H}} \lambda_{\text{PO}_4} \lambda_{\text{Ca}}}{K_3 K_5} \right) \right) \\ & + \left( \frac{x(1) x(2) \lambda_{\text{Ca}} \lambda_{\text{PO}_4}}{K_6 \lambda_{\text{CaPO}_4}} \right) \left( -1609.20 \times 10^3 + 298.15 R \ln \left( \frac{x(1) x(2) \lambda_{\text{Ca}} \lambda_{\text{PO}_4}}{K_6 \lambda_{\text{CaPO}_4}} \right) \right) \\ & + x(1) (-553.54 \times 10^3 + 298.15 R \ln(x(1) \lambda_{\text{Ca}})) + \frac{K_w}{\lambda_{\text{OH}} x(3) \lambda_{\text{H}}} \left( -157.29 \times 10^3 + 298.15 R \ln \left( \frac{K_w}{x(3) \lambda_{\text{H}}} \right) \right) \\ & + x(3) (298.15 R \ln(x(3) \lambda_{\text{H}})) - 1680.47 x(4) P_{\text{Total}} - 3844.92 x(5) P_{\text{Total}} \end{aligned}$$

where

$$\begin{aligned} I = & 0.5 \left( 4x(1) + \frac{x(1)x(3)^2 \lambda_{\text{H}}^2 \lambda_{\text{PO}_4} \lambda_{\text{Ca}}}{K_4 \lambda_{\text{CaH}_2\text{PO}_4} K_3 K_2} + \frac{x(1) K_w \lambda_{\text{Ca}}}{\lambda_{\text{CaOH}} x(3) \lambda_{\text{H}} K_7} \right. \\ & + x(3) + [\text{K}^+] + \frac{x(3)^2 x(2) \lambda_{\text{PO}_4} \lambda_{\text{H}}^2}{K_3 K_2 \lambda_{\text{H}_2\text{PO}_4}} \\ & + 4 \frac{x(3) x(2) \lambda_{\text{H}} \lambda_{\text{PO}_4}}{K_3 \lambda_{\text{HPO}_4}} + 9x(2) + \frac{x(1) x(2) \lambda_{\text{Ca}} \lambda_{\text{PO}_4}}{K_6 \lambda_{\text{CaPO}_4}} \\ & \left. + \frac{K_w}{\lambda_{\text{OH}} x(3) \lambda_{\text{H}}} + [\text{Cl}^-] \right) \end{aligned}$$

ACP supersaturation constraints:

$$\ln \left( \frac{(x(1) \lambda_{\text{Ca}})^3 (x(2) \lambda_{\text{PO}_4})^2}{K_{\text{SACP}}} \right) \geq 0$$

DCPD supersaturation constraints:

$$\ln \left( \frac{(x(1) \lambda_{\text{Ca}}) (x(3) x(2) \lambda_{\text{H}} \lambda_{\text{PO}_4} / K_3 \lambda_{\text{HPO}_4})}{K_{\text{SDCPD}}} \right) \geq 0$$



Electroneutrality constraints:

$$2x(1) + \frac{x(1)x(3)^2\lambda_H^2\lambda_{PO_4}\lambda_{Ca}}{K_4\lambda_{CaH_2PO_4}K_3K_2} + \frac{x(1)K_w\lambda_{Ca}}{\lambda_{CaOH}x(3)\lambda_HK_7} + x(3) + [K^+] - \left( \frac{x(3)^2x(2)\lambda_{PO_4}\lambda_H^2}{K_3K_2\lambda_{H_2PO_4}} + 2\frac{x(3)x(2)\lambda_H\lambda_{PO_4}}{K_3\lambda_{HPO_4}} \right) + 3x(2) + \frac{x(1)x(2)\lambda_{Ca}\lambda_{PO_4}}{K_6\lambda_{CaPO_4}} + \frac{K_w}{\lambda_{OH}x(3)\lambda_H} + [Cl^-] = 0$$

Calcium balance constraints:

$$x(1) + \frac{x(1)x(3)^2\lambda_H^2\lambda_{PO_4}\lambda_{Ca}}{K_4\lambda_{CaH_2PO_4}K_3K_2} + \frac{x(1)K_w\lambda_{Ca}}{\lambda_{CaOH}x(3)\lambda_HK_7} + \frac{x(1)x(2)\lambda_{Ca}\lambda_{PO_4}}{K_6\lambda_{CaPO_4}} + \frac{x(1)x(3)x(2)\lambda_H\lambda_{PO_4}\lambda_{Ca}}{K_3K_5} - \left( Ca_{Total} - \left( \frac{3}{2}x(4) + x(5) \right) P_{Total} \right) = 0$$

Phosphate balance constraints:

$$x(2) + \frac{x(3)^3x(2)\lambda_H^3\lambda_{PO_4}}{K_3K_2K_1} + \frac{x(1)x(3)^2\lambda_H^2\lambda_{PO_4}\lambda_{Ca}}{K_4\lambda_{CaH_2PO_4}K_3K_2} + \frac{x(3)^2x(2)\lambda_{PO_4}\lambda_H^2}{K_3K_2\lambda_{H_2PO_4}} + \frac{x(3)x(2)\lambda_H\lambda_{PO_4}}{K_3\lambda_{HPO_4}} + \frac{x(1)x(2)\lambda_{Ca}\lambda_{PO_4}}{K_6\lambda_{CaPO_4}} + \frac{x(1)x(3)x(2)\lambda_H\lambda_{PO_4}\lambda_{Ca}}{K_3K_5} - (P_{Total}(1 - x(4) - x(5))) = 0$$

## Appendix B. Mathematical examples

### B.1. Example 1: optimization of drying process for a through-circulation dryer

The problem proposed by Chung [8] consists in finding the air flow rate  $x_1$  and the bed thickness  $x_2$  which maximize the drying production rate:

$$P = 0.033x_1 \left[ \frac{0.036}{1 - \exp(-107.9x_2/x_1^{0.41})} + 0.095 - \frac{9.27 \times 10^{-4}x_1^{0.41}}{x_2} \ln \left( \frac{1 - \exp(-5.39x_2/x_1^{0.41})}{1 - \exp(-107.9x_2/x_1^{0.41})} \right) \right]^{-1} \quad (B.15)$$

subject to the constraints

$$0.2 - 4.62 \times 10^{-10}x_1^{2.85} - 1.055 \times 10^{-4}x_1 \geq 0 \quad (B.16)$$

$$\frac{4}{12} - 8.20 \times 10^{-7}x_1^{1.85}x_2 - \frac{2.25}{12} \geq 0 \quad (B.17)$$

$$2 - 109.6 \frac{x_2}{x_1^{0.41}} \left[ \frac{0.036}{1 - \exp(-107.9x_2/x_1^{0.41})} + 0.095 - \frac{9.27 \times 10^{-4}x_1^{0.41}}{x_2} \ln \left( \frac{1 - \exp(-5.39x_2/x_1^{0.41})}{1 - \exp(-107.9x_2/x_1^{0.41})} \right) \right] \geq 0 \quad (B.18)$$

The optimum given in the literature is  $(x_1, x_2; f) = (975.831, 0.5244; 172.487)$ .

### B.2. Example 2: alkylation process optimization

Alkylation is a commonly used process in upgrading gasoline. The problem under consideration consists in determining the best operating conditions for the alkylation process described by Paynes [9] and optimized by Sauer et al. [10] who transformed the nonlinear problems into a series of linear programming problems. Since Bracken and McCormick [11] formulated this problem as a classical nonlinear programming model, but failed to obtain a correct solution, the same notations as in [11] are used, but the problem is solved by using the proposed GA/SQP procedure.

By using the data from [10], the problem can be formulated as follows:

Maximize:

$$P = 0.063x_4x_7 - 5.04x_1 - 0.035x_2 - 10x_3 - 3.36x_5 \quad (B.19)$$

subject to the inequality constraints:

$x_1$ = olefin feed (barrels/day)	$0.010 \leq x_1 \leq 2000$
$x_2$ = isobutane recycle (thousands of pounds/day)	$0.010 \leq x_2 \leq 16000$
$x_3$ = acid addition rate (thousand of pounds/day)	$0.010 \leq x_3 \leq 120$
$x_4$ = alkylate yield (barrels/day)	$0.010 \leq x_4 \leq 5000$
$x_5$ = isobutane makeup (barrels/day)	$0.010 \leq x_5 \leq 2000$
$x_6$ = acid strength (weight percent)	$85 \leq x_6 \leq 93$
$x_7$ = motor octane number	$90 \leq x_7 \leq 95$
$x_8$ = external isobutane-to-olefin ratio	$3 \leq x_8 \leq 12$
$x_9$ = acid dilution factor	$1.2 \leq x_9 \leq 4$
$x_{10}$ = F-4 performance number	$145 \leq x_{10} \leq 162$

and the equality constraints:

$$x_4 = x_1(1.12 + 0.13167x_8 - 0.00667x_8^2) \quad (B.20)$$

$$x_5 = 1.22x_4 - x_1 \quad (B.21)$$

$$x_2 = x_1x_8 - x_5 \quad (B.22)$$

$$x_6 = 89 + \frac{x_7 - (86.35 + 1.098x_8 - 0.038x_8)}{0.325} \quad (B.23)$$

$$x_{10} = 133 + 3x_7 \quad (B.24)$$

$$x_9 = \frac{0.001x_4x_6x_9}{98 - x_6} \quad (\text{B.25})$$

The optimum given in the literature is  $(x_1, x_2, x_3, x_4, x_5, x_6, x_7, x_8, x_9, x_{10}; f) = (1728.371, 16000.000, 98.136, 3056.042, 2000.00, 90.618, 94.189, 10.414, 2.616, 149.569; 1113.231)$ .

### B.3. Example 3: chemical equilibrium in complex mixture

The chemical equilibrium problem formulated and solved in [12,13] by means of piecewise linear approximation followed by linear programming is now considered. It is known that the free energy reaches a minimum for a system at the chemical equilibrium. So, to determine the relative amounts of species in an equilibrium situation, the problem of free energy minimization can be solved. By considering the N, H, O system [12] involving 10 species, the problem consists in minimizing:

$$f(x) = \sum_{i=1}^{10} x_i \left( w_i + \ln P + \ln \frac{x_i}{\sum_{i=1}^{10} x_i} \right) \quad (\text{B.26})$$

with  $P = 750$  and  $w_i$  ( $i = 1, \dots, 10$ ) =  $-10.021, -21.096, -37.986, -9.846, -28.653, -18.918, -28.032, -14.640, -30.594, -26.111$ .

The constraints are:

$$x_1 + 2x_2 + 2x_3 + x_6 + x_{10} = 2 \quad (\text{B.27})$$

$$x_4 + 2x_5 + x_6 + x_7 = 1 \quad (\text{B.28})$$

$$x_3 + x_7 + x_8 + 2x_9 + x_{10} = 1 \quad (\text{B.29})$$

The optimum given in the literature is  $(x_1, x_2, x_3, x_4, x_5, x_6, x_7, x_8, x_9, x_{10}; f) = (7.006 \times 10^{-3}, 6.808 \times 10^{-2}, 0.9072, 3.609 \times 10^{-4}, 0.4908, 4.723 \times 10^{-4}, 1.757 \times 10^{-2}, 2.904 \times 10^{-3}, 1.518 \times 10^{-2}, 4.195 \times 10^{-2}; 43.494)$ .

## Appendix C. Nomenclature

$e$  electronic charge ( $= 1.602177 \times 10^{-19}$  C)  
 $k_B$  Boltzmann constant ( $= 1.380658 \times 10^{-23}$  J/K)

$N_A$  Avogadro number ( $= 6.022136 \times 10^{23}$  mol<sup>-1</sup>)  
 $T$  temperature (K)  
 $z_i$  charge number of ion  
 $\varepsilon$  solvent dielectric constant ( $\varepsilon = \varepsilon_r \times \varepsilon_0$ )  
 $\varepsilon_0$  vacuum permittivity ( $= 8.854187 \times 10^{-12}$  F/m)  
 $\varepsilon_r$  relative solvent dielectric constant  
 $\lambda$  ion activity coefficient  
 $\rho_0$  solvent density (kg/m<sup>3</sup>)  
 $I$  solution ionic strength (mol/l)

## References

- [1] C.M. McDonald, C.A. Floudas, Global optimization for the phase and chemical equilibrium problem: application to the NRTL equation, *Comput. Chem. Eng.* 19 (11) (1995) 1111–1139.
- [2] M.M. Seckler, Calcium phosphate precipitation in a fluidized bed, Ph.D. Thesis, Technische Universiteit Delft, 1994.
- [3] L. Montastruc, C. Azzaro-Pantel, B. Biscans, M. Cabassud, S. Domenech, L. Pibouleau, A general framework for pellet reactor modeling: application to P-recovery, *Trans. Inst. Chem. Eng. Part A: Chem. Eng. Res. Design* 81 (9) (2003) 1271–1278.
- [4] L. Montastruc, C. Azzaro-Pantel, B. Biscans, M. Cabassud, S. Domenech, A thermochemical approach for calcium phosphate precipitation modeling in a pellet reactor, *Chem. Eng. J.* 94 (1) (2003) 41–50.
- [5] K. Schittkowski, C.L. Monma (Ed.), NLPQL: a FORTRAN subroutine solving constrained non-linear programming problems, *Ann. Operation Research* 5 (1986) 485–500.
- [6] J. Holland, *Adaptation in Natural and Artificial Systems*, MIT Press, Cambridge, MA, 1975.
- [7] D.E. Goldberg, *Genetic Algorithms in Search, Optimization and Machine Learning*, Addison-Wesley, Massachusetts, 1989.
- [8] S.F. Chung, Mathematical model and optimization of drying process for a through-circulation dryer, *Can. J. Chem. Eng.* 50 (1972) 657–662.
- [9] R.E. Paynes, Alkylation What you should know about this process, *Pet. Refiner* 37 (1958) 316–325.
- [10] R.N. Sauer, A.R. Colville, C.W. Burwick, Computer points in the way to more profits, *Hydrocarbon Process. Pet. Refiner* 43 (1964) 84–90.
- [11] J. Bracken, G.P. McCormick, *Selected Applications of Nonlinear Programming*, Wiley, New York, 1968, Chapter 8.4, pp. 89–90.
- [12] W.B. White, S.M. Johnson, G. Dantzig, Chemical equilibrium in complex mixture, *J. Chem. Phys.* 28 (1958) 751–755.
- [13] G. Dantzig, S. Johnson, W. White, A linear programming approach to the chemical equilibrium problem, *Manage. Sci.* 5 (1958) 38–43.