

This article was downloaded by:

On: 27 January 2011

Access details: *Access Details: Free Access*

Publisher *Taylor & Francis*

Informa Ltd Registered in England and Wales Registered Number: 1072954 Registered office: Mortimer House, 37-41 Mortimer Street, London W1T 3JH, UK



Phosphorus, Sulfur, and Silicon and the Related Elements

Publication details, including instructions for authors and subscription information:

<http://www.informaworld.com/smpp/title~content=t713618290>

Application Des Réseaux de Neurones Avec la Régularisation Bayésienne pour la Modélisation de la Synthèse de l'hydroxyapatite Élaborée à Partir du Carbonate de Calcium et de L'acide Phosphorique

Mohamed Nohair^a; Ouafae Britel^a; Nabil Souaf^a; Driss Zakarya^a; Abdelmjid Hafid^a; Noura Mallouk^a

^a Laboratoire Catalyse, Synthèse, Chimie et Environnement, Département de Chimie, Université Hassan II-Mohammedia, Faculté des Sciences et Techniques de Mohammedia, Mohammedia, Maroc

Online publication date: 02 August 2010

To cite this Article Nohair, Mohamed , Britel, Ouafae , Souaf, Nabil , Zakarya, Driss , Hafid, Abdelmjid and Mallouk, Noura(2010) 'Application Des Réseaux de Neurones Avec la Régularisation Bayésienne pour la Modélisation de la Synthèse de l'hydroxyapatite Élaborée à Partir du Carbonate de Calcium et de L'acide Phosphorique', *Phosphorus, Sulfur, and Silicon and the Related Elements*, 185: 8, 1772 — 1781

To link to this Article: DOI: 10.1080/10426500903299877

URL: <http://dx.doi.org/10.1080/10426500903299877>

PLEASE SCROLL DOWN FOR ARTICLE

Full terms and conditions of use: <http://www.informaworld.com/terms-and-conditions-of-access.pdf>

This article may be used for research, teaching and private study purposes. Any substantial or systematic reproduction, re-distribution, re-selling, loan or sub-licensing, systematic supply or distribution in any form to anyone is expressly forbidden.

The publisher does not give any warranty express or implied or make any representation that the contents will be complete or accurate or up to date. The accuracy of any instructions, formulae and drug doses should be independently verified with primary sources. The publisher shall not be liable for any loss, actions, claims, proceedings, demand or costs or damages whatsoever or howsoever caused arising directly or indirectly in connection with or arising out of the use of this material.

APPLICATION DES RÉSEAUX DE NEURONES AVEC LA RÉGULARISATION BAYÉSIENNE POUR LA MODÉLISATION DE LA SYNTHÈSE DE L'HYDROXYAPATITE ÉLABORÉE À PARTIR DU CARBONATE DE CALCIUM ET DE L'ACIDE PHOSPHORIQUE

Mohamed Nohair, Ouafae Britel, Nabil Souaf, Driss Zakarya, Abdelmjid Hafid, and Noura Mallouk

Laboratoire Catalyse, Synthèse, Chimie et Environnement, Département de Chimie, Université Hassan II-Mohammedia, Faculté des Sciences et Techniques de Mohammedia, Mohammedia, Maroc

We have used a new, robust model mapping technique—a Bayesian-regularized neural network—to develop a quantitative relationships model for the synthesis of the phosphocalcic hydroxyapatite by precipitation from a calcium carbonate solution and a phosphoric acid solution. This model was performed by using a set of factors consisting on the pH of reactional medium, the Ca/P molar ratio of the reagents, reaction time, and the initial concentration of calcium. The results show that the method is robust and gives satisfied results. The Levenberg–Marquardt's algorithm implemented in the neural network Matlab's toolbox allowed a drastic improvement of the performance of the model. Very satisfactory results are then obtained by testing the validity by cross-validation technique.

We have also turned our interests to the explanatory capacities of our methodology to explore the relative contribution and/or the contribution profile of the input factors by using Garson weight portioning method.

Keywords Bayesian regularization; cross-validation technique; Garson weight portioning method; neural network; phosphocalcic hydroxyapatite

INTRODUCTION

Dans un récent travail réalisé par Britel et al.,¹ la synthèse de l'hydroxyapatite phosphocalcique a été optimisée en utilisant la technique des plans d'expériences.² Un modèle mathématique basé sur une équation polynomiale de second degré a été établi entre le rapport molaire Ca/P du solide préparé et les facteurs qui influent sur la précipitation, tels que le pH, le rapport molaire Ca/P des réactifs, la concentration initiale en ions calcium et la durée de la réaction. Le plan composite centré^{2,3} avec l'isovariance par rotation

Received 20 May 2009; accepted 14 August 2009.

Address correspondence to Mohamed Nohair, Laboratoire Catalyse, Synthèse, Chimie et Environnement, Département de Chimie, Université Hassan II-Mohammedia, Faculté des Sciences et Technique de Mohammedia, B.P. 146, 20650 Mohammedia, Maroc. E-mail: nohairmohamed@yahoo.fr

Tableau I Estimation des coefficients associés au modèle proposé dans la référence [1] et selon l'équation (1)

	coefficient	signification
b0	1,68	—
b1	0,03	***
b2	0,01	**
b3	0,006	*
b4	0,01	***
b12	−0,01	**
b13	0,01	*
b14	0,006	NS
b23	−0,01	**
b24	−0,003	NS
b34	0,01	**
b11	−0,001	NS
b22	−0,001	NS
b33	−0,001	NS
b44	−0,001	NS

*** significatif à un niveau 0,1%, ** significatif à un niveau 1%, * significatif à un niveau 5%.

NS: non significatif.

consistait en 31 expériences; les 16 premières expériences représentent un plan factoriel complet 2^4 expériences, auxquelles s'ajoute 7 autres au centre et 8 autres complémentaires réparties équitablement sur les quatres axes. L'ensemble des expériences est décrit dans le Tableau I, les valeurs des facteurs sont codées entre −1 et 1. Toutes les variables représentant les différents facteurs ont été normalisées et centrées pour éviter la surestimation de l'effet des variables ayant des grandeurs importantes. La relation qui en résulte par la méthode des moindres des carrées conduit à une équation de la forme suivante :

$$\hat{y} = b_0 + \sum b_i x_i + \sum \sum b_{ij} x_i x_j + \sum b_{ii} x_i^2 \quad (1)$$

D'après les valeurs des coefficients du modèle cité précédemment et comme il figure dans le Tableau II, le pH apparaît de forte contribution, suivi simultanément à un degré moindre par le rapport Ca/P des réactifs et la durée de la réaction. Pour tester la validité du modèle proposée, une seule réaction a été réalisée dans les conditions suivantes: pH = 7,75, Ca/P = 1,667, $[Ca^{2+}] = 1,25$ mol/l et Durée = 2h30 mn pour la synthèse d'un hydroxyapatite phosphocalcique dont la valeur Ca/P est égale à 1,67. Les valeurs des différents facteurs sont pré-établis par les courbes d'isoreponses. La vérification expérimentale de ce produit par la diffraction des rayons X et le spectre d'absorption infra-rouge confirme que le produit obtenu est une Hydroxyapatite stoechiométrique (Ca/P = 1,67).

Nous proposons dans ce papier un travail relatif à l'utilisation des réseaux de neurones et la régularisation Bayésienne⁴ pour la réalisation d'une relation de même type, en l'occurrence, la prédiction de l'élaboration de l'hydroxyapatite phosphocalcique par précipitation à partir d'une solution de carbonate de calcium et une solution d'acide phosphorique. Les facteurs qui régissent la précipitation, et par conséquent la composition chimique du solide préparé, sont les mêmes que ceux précédemment cités dans la modélisation par les plans d'expériences.

Récemment, les modèles linéaires largement utilisés pour établir des relations statistiques, sont supplantés par une nouvelle technique basée sur l'intelligence artificielle : Ce

Tableau II Les différentes expérimentations pour le plan composite centré étudié et les résultats d'analyse Ca/P par les réseaux de neurones

Numéro	Valeurs des variables codées				(Ca/P) _{exp}	(Ca/P) _{calc} en validation croisée
	X ₁ (pH)	X ₂ (Ca/P)	X ₃ [Ca ²⁺]	X ₄ (Durée)		
1	-1	-1	-1	-1	1,643	1,643
2	1	-1	-1	-1	1,710	1,71
3	-1	1	-1	-1	1,662	1,67
4	1	1	-1	-1	1,695	1,701
5	-1	-1	1	-1	1,600	1,689
6	1	-1	1	-1	1,710	1,685
7	-1	1	1	-1	1,620	1,68
8	1	1	1	-1	1,624	1,71
9	-1	-1	-1	1	1,656	1,62
10	1	-1	-1	1	1,700	1,706
11	-1	1	-1	1	1,670	1,63
12	1	1	-1	1	1,706	1,672
13	-1	-1	1	1	1,652	1,68
14	1	-1	1	1	1,795	1,662
15	-1	1	1	1	1,610	1,624
16	1	1	1	1	1,701	1,652
17	-2	0	0	0	1,630	1,731
18	2	0	0	0	1,731	1,66
19	0	-2	0	0	1,699	1,684
20	0	2	0	0	1,622	1,695
21	0	0	-2	0	1,689	1,656
22	0	0	2	0	1,672	1,795
23	0	0	0	-2	1,660	1,699
24	0	0	0	2	1,701	1,701
25	0	0	0	0	1,684	1,685
26	0	0	0	0	1,685	1,6
27	0	0	0	0	1,680	1,7
28	0	0	0	0	1,684	1,61
29	0	0	0	0	1,685	1,662
30	0	0	0	0	1,680	1,684
31	0	0	0	0	1,680	1,68

sont les réseaux de neurones. Ils permettent de proposer un modèle de type cause à effet sans qu'il soit nécessaire de postuler a priori, un modèle de corrélation figé. Néanmoins, Ils présentent un inconvénient majeur, il réside dans le surajustement, et donc des capacités prédictives, dans certains cas, affaiblies. Ceci revient à la définition d'une architecture souple des réseaux de neurones, c-à-d le nombre de neurones et le type d'interconnexion entre elles. Il n'existe pas jusqu'à présent dans la littérature de procédure standard pour résoudre ce genre de problème, excepté la technique de l'arrêt immédiat de l'apprentissage⁴ et la régularisation Bayésienne.⁵⁻⁷ La première technique pose le problème du choix du meilleur modèle à partir de plusieurs modèles obtenus par le processus de validation. Ce processus consiste à décomposer l'ensemble initial de données en deux sous-ensembles, le premier sert à l'apprentissage et le second permet la validation et la fixation des paramètres du modèle de réseau de neurones. La seconde technique est connue par sa robustesse,⁷ puisqu'elle ne nécessite pas un ensemble de données pour la validation. Nous l'utiliserons au cours de la validation du modèle proposé.

Tableau III Analyse statistique de la régression pas-à-pas

	X_1	X_1, X_2	X_1, X_2, X_4	X_1, X_2, X_4, X_3
(R et s)	(0,68; 0,04)	(0,75; 0,03)	(0,81; 0,02)	(0,99; 0,02)

Avec: $X_1 = \text{pH}$, $X_2 = \text{Ca/P}$, $X_3 = [\text{Ca}^{2+}]$, et $X_4 = \text{Durée}$.

EVALUATION PARAMETRIQUE PAR REGRESSION LINEAIRE

Dans un premier temps, nous utilisons un modèle linéaire simple. La régression ne servirait qu'à une évaluation paramétrique des facteurs choisis, et confirmer la pertinence de leur influence sur la réaction de synthèse de l'hydroxyapatite phosphocalcique. Les différents facteurs sont indépendants, les coefficients de toutes les variables sont donc épurés des contributions fournies par les autres et représentent, donc, des effets propres, mais relatifs au modèle linéaire.

Avant tout apprentissage, toutes les variables ont été normalisées et centrées pour éviter la surestimation de l'effet des variables ayant des grandeurs importantes. Nous avons réalisé une régression pas-à-pas (stepwise regression 'forward selection'). Elle consiste en la construction des 4 modèles possibles à une variable seulement. Nous calculons les critères R et s (coefficient de corrélation et écart type) pour chacun de ces modèles et nous choisissons celui qui optimise ces critères. Nous fixons ensuite la variable déjà sélectionnée; (4 - 1 = 3) modèles sont alors construits en ajoutant une seule des variables restantes. Les critères R (coefficient de corrélation) et s (écart type) pour chacun de ces modèles sont calculés, et nous choisissons le meilleur modèle. Nous continuons le processus ci-dessus jusqu'à ce que toutes les variables soient choisies. Cette opération a permis, en commençant par la variable de forte contribution, en l'occurrence le pH de la solution, d'établir l'ordre suivant dont les paramètres statistiques est présenté dans le Tableau III.

Ces résultats confirment ceux qui résultent de la modélisation par les plans d'expériences. Le pH est de forte influence, suivi successivement par le rapport Ca/P et la durée de la réaction. La concentration en ions Ca^{2+} contribue faiblement. Par ailleurs le modèle du plan composite réalisé a montré que les coefficients des variables relevées au degré 2 sont faibles et insignifiants statistiquement au risque de 5%. La faiblesse de la contribution des ions de calcium ne remet pas en cause sa présence en tant que facteur influent, puisque les coefficients de toutes les interactions avec ce facteur sont statistiquement significatifs dans le modèle réalisé à partir du plan composite centré.

REALISATION DU MODELE NEURONAL POUR LA MODELISATION DE LA SYNTHESE DE L'HYDROXYAPATITE

Un neurone est avant tout un opérateur mathématique. Il réalise une somme pondérée suivie d'une fonction non linéaire "f". Cette fonction doit être bornée, continue et dérivable, elle peut aussi avoir la forme d'une fonction seuil si le résultat recherché est de type booléen (soit 0 ou 1).⁷ Les fonctions les plus fréquemment utilisées sont les fonctions sigmoïdes. Par ailleurs, Il est démontré qu'un réseau de neurones ayant une couche cachée avec une combinaison d'une fonction sigmoïde et une fonction linéaire est un approximateur universel dans le sens qu'il peut approximer toute fonction multivariable non linéaire avec une précision fixée, pourvu qu'il soit soumis à un apprentissage efficace.⁷

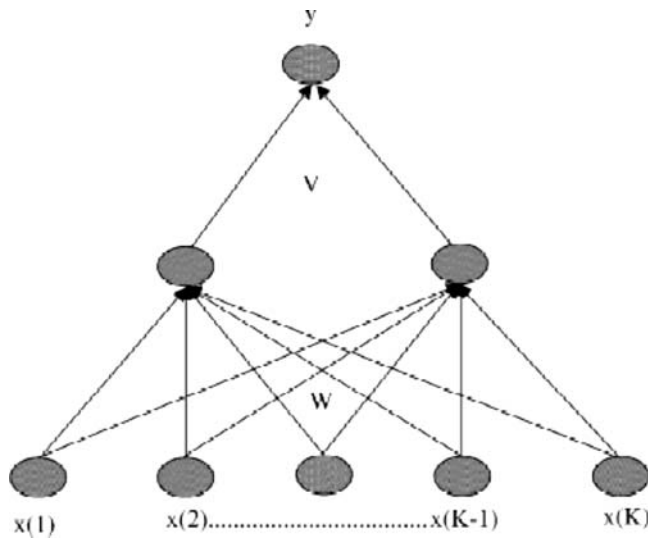


Figure 1 Architecture d'un réseau de neurones à trois couches. La matrice de poids entre les variables d'entrées et les variables de la couche cachée est représentée par W . La matrice de poids entre les variables de la couche cachée et la réponse est représentée par V .

On distingue deux grands types d'architectures de réseau de neurones en modélisation: les réseaux de neurones non bouclés et les réseaux de neurones bouclés. Dans notre cas, nous utiliserons la première variante puisqu'on ne cherche pas une relation dépendante du temps. Une seule architecture de type réseau à couches (MLP) sera utilisée (Figure 1): une couche de neurones d'entrées est connectée à une couche de neurones «cachés», qui est reliée à une couche de neurones de sortie. L'activité des neurones d'entrée code l'information qui est présentée au réseau; ce sont les variables explicatives externes citées précédemment. L'activité de chaque neurone caché est déterminée par l'activité des neurones d'entrées et les valeurs des coefficients de connexion ou poids. L'activité des neurones de sortie dépend de l'activité des neurones cachés et des coefficients de connexion reliant ces neurones à ceux de sortie. La couche de sortie contient un seul neurone, représentant le rapport Ca/P du produit final.

Le principe des réseaux de neurones est le même que n'importe quel phénomène statistique caractérisé par une relation déterministe entre des causes et des effets. Nous présentons au réseau des exemples d'apprentissage, c'est-à-dire des jeux d'activités des neurones d'entrée ainsi que l'activité du neurone de sortie correspondant.^{8,9} L'apprentissage consiste tout simplement à calculer les coefficients de connexion entre les différentes couches de telle manière que les sorties du réseau de neurones soient, pour les exemples utilisés, aussi proches que possible des sorties désirées. Pour ce faire, on utilise des algorithmes d'optimisation: ils cherchent à minimiser une fonction de coût qui constitue une mesure de l'écart entre la réponse réelle du réseau et la réponse désirée. Cette optimisation se fait de manière itérative, en modifiant les poids en fonction du gradient de la fonction du coût; le gradient est estimé par une méthode spécifique aux réseaux de neurones, dite méthode de rétropropagation.¹⁰ Les poids sont initialisés aléatoirement avant l'apprentissage, puis modifiés itérativement jusqu'à obtention d'un compromis satisfaisant

entre la précision de l'approximation sur un ensemble de validation disjoint au précédent, et celle obtenue sur l'ensemble des données d'apprentissage.

MISE EN ŒUVRE DES MODÈLES DE RÉSEAUX DE NEURONES

Pour enseigner une tâche à un réseau, on doit ajuster les coefficients de chaque neurone et minimiser ainsi la différence entre la sortie désirée et la sortie effective. Ce procédé impose de calculer la dérivée d'une quantité nommée coût " J " par rapport aux coefficients de connexion, c'est à dire comment l'erreur varie quand en fonction de chaque coefficient de connexion. La méthode la plus utilisée pour déterminer ces dérivées est la méthode de rétropropagation.

Une fois que les activités de tous les neurones de sortie ont été déterminées par l'ensemble d'apprentissage, le réseau calcule le coût " J " défini par l'expression suivante:

$$J(w) = \sum_{k=1}^N (y - g(x, w))^2 = \sum_{k=1}^N J(w) \quad (2)$$

Où y : rapport Ca/P observé, $g(x, w)$ = rapport Ca/P calculé, N est le nombre d'expériences et w représente le poids de connexion entre les différentes couches

Pour évaluer ce gradient, il suffit donc d'évaluer le gradient du coût partiel $J(w)$ relatif à l'observation k , et de faire ensuite la somme sur tous les exemples. L'algorithme de rétropropagation consiste essentiellement en l'application répétée de la règle des dérivées composées suivante:

$$\left(\frac{\partial J^k}{\partial w_{ij}} \right) = \left(\frac{\partial J^k}{\partial v_i} \right) \left(\frac{\partial v_i}{\partial w_{ij}} \right)_k = \delta_i^k x_j^k \quad (3)$$

δ_j^k désigne la valeur du gradient du coût partiel par rapport au potentiel du neurone i pour l'exemple k , il peut être exprimé de la manière suivante:

$$\delta_j^k = \left(\frac{\partial J^k}{\partial v_i} \right) = -2g(x^k, w) \left(\frac{\partial g(x, w)}{\partial v_i} \right)_k \quad (4)$$

La même approche est appliquée aux neurones cachés, les paramètres du réseau sont modifiés par la formule suivante, à l'itération i de l'apprentissage

$$w(i) = w(i - 1) - \mu_i \nabla J(w(i - 1)) \quad (5)$$

La quantité μ_i représente le pas du gradient ou pas d'apprentissage.

Cette méthode représente des inconvénients liés par exemple à la possibilité de la faiblesse de la pente de la fonction de coût, donc le minimum peut être dû à ce facteur, plutôt au fait qu'on est au voisinage du vrai minimum de la fonction de coût. Pour remédier ce problème, d'autres techniques sont introduites en se basant sur la méthode des dérivés de second ordre, elle dérive de la méthode de Newton en développant la fonction de coût au voisinage d'un minimum w^* , parmi ces techniques on cite l'algorithme de Levenberg-Marquardt.⁷

D'autres techniques sont aussi utilisées, parmi lesquelles celle qui consiste à ajouter au terme de gradient $-\mu \nabla J$ un terme proportionnel à la variation des coefficients à l'itération précédente $\lambda[w(i-1)-w(i-2)]$, le coefficient λ a pour rôle de diriger la variation des valeurs

des poids dans la même direction que celle utilisée dans l'itération précédente, et par la suite augmenter la vitesse de convergence et éviter les oscillations de la fonction de coût. Tous ces algorithmes sont implémentés dans le toolbox de Matlab, Neural Network version 7.0.¹¹

La corrélation entre les valeurs calculées (prédites) et observées (expérimentales) est examinée au moyen d'une relation linéaire entre elles :

$$Ca/P_{\text{exp}} = a * Ca/P_{\text{pred}} + b. \quad (6)$$

Ca/P désigne le rapport entre le calcium et le phosphore du produit final.

Les paramètres statistiques sont calculés à partir des relations ci-dessous :

$$R^2 = 1 - \frac{\sum (Ca/P_{\text{obs}} - Ca/P_{\text{calc}})^2}{\sum (Ca/P_{\text{obs}} - Ca/P_{\text{MEAN}})^2} \quad s^2 = \frac{\sum (Ca/P_{\text{obs}} - Ca/P_{\text{calc}})^2}{n} \quad (7)$$

R et s représentent respectivement le coefficient de corrélation et l'écart type de l'erreur.

RESULTATS ET DISCUSSIONS

L'architecture des réseaux de neurones utilisée dans ce travail consiste en une série de trois couches de neurones (Figure 1). La première couche représente les valeurs des facteurs pour chaque expérience, la dernière couche représente la valeur du rapport Ca/P du produit synthétisé. Les paramètres du réseau (les valeurs des poids) sont initialisés à l'aide d'une fonction qui fournit des valeurs aléatoires. L'apprentissage des données, c-à-d la minimisation de la fonction de coût est réalisée à l'aide de l'algorithme de Levenberg-Marquardt.⁷

Dans le but de réaliser une architecture optimale du réseau de neurones, le nombre de neurones de la couche d'entrée et cachée doit être calculé. La régression linéaire par la technique de pas-à-pas (stepwise method) suppose que les différents facteurs contribuent séparément et indépendamment, ceci correspond à un modèle linéaire idéal. En analyse statistique par les réseaux de neurones, les possibles interactions entre les différents facteurs doivent être prise en considération. De ce fait, nous avons construit intuitivement tous les modèles de 1 jusqu'à 4 variables, à partir des quatre facteurs du modèle initial. Le nombre de modèles est élevé, il est égal $\sum_{i=1}^4 C_4^i = 14$. L'architecture optimale est déterminée en choisissant la combinaison de facteurs qui produit les meilleurs paramètres statistiques (R et s). Cette procédure est réalisée par un programme développé sur la plateforme de Matlab; et elle nous fournit une combinaison semblable à celle du modèle linéaire, avec le même ordre de contribution des différents facteurs, néanmoins ils sont tous statistiquement significatives. Par ailleurs, nous avons réalisé plusieurs architectures en faisant varier le nombre de neurones de la couche cachée, nous avons remarqué qu'il n'y a pas de limite à ses capacités de prédiction, surtout lorsque le nombre de neurones dans la couche cachée augmente, mais nous avons constaté que pour un nombre de neurones dans la couche cachée supérieur à 3, le gain n'est pas trop significatif. Les paramètres statistiques qui résultent de la relation linéaire entre le rapport Ca/P observé et calculé sont performants, avec un coefficient de détermination égale à 95% et une erreur type égale à 0,01. Par ailleurs, l'algorithme de Levenberg-Marquardt nécessite à peine une centaine d'itérations pour une convergence totale de la fonction de coût, par rapport à la méthode classique (simple gradient) avec plus d'un millier d'itérations.

Notons que ces paramètres résultent d'un apprentissage sur la totalité de la base de données. Malheureusement, nous ne disposons d'aucun moyen de vérifier la nature de la fonction à approximer par les réseaux de neurones. Ce résultat peut résulter d'un apprentissage aléatoire sans aucune régularisation, d'autant plus que les valeurs des rapports Ca/P sont parfois très distants (exemple: essai 14, Ca/P = 1,795 et l'essai 18, Ca/P = 1,731). Pour pallier à cet inconvénient, nous avons réalisé une partition de l'ensemble des données en deux ensembles. Le premier sert à l'apprentissage du modèle, et le deuxième pour en fixer les paramètres par une validation. Nous effectuons un arrêt de l'apprentissage lorsque l'erreur sur l'ensemble de la validation (le second ensemble) augmente. Les paramètres statistiques régressent, mais néanmoins restent très satisfaisants ($R^2 = 94$ et $s = 0,015$). Cette méthode peut conduire à plusieurs modèles, et ce autant que le nombre de partitions possibles de la base de données initiale.

D'autres part, nous avons analysé la contribution relative de chaque facteur en utilisant la méthode de Garson (the Garson weight partitioning method).^{12,13} Les valeurs de poids de connexion sont exprimées en valeur absolue pour épurer chaque valeur des interactions des autres facteurs sur la réponse de sortie.

$$C\% = \frac{\sum_{j=1}^{n_H} \left\{ \frac{i_{vj}}{\sum_{k=1}^{n_v} i_{kj}} o_j \right\}}{\sum_{i=1}^{n_v} \left[\sum_{j=1}^{n_H} \left\{ \frac{i_{vj}}{\sum_{k=1}^{n_v} i_{kj}} o_j \right\} \right]} \quad (8)$$

Avec C représentant la contribution en pourcentage.

Dans cette équation, n_v est le nombre de neurones dans la couche d'entrée, n_H le nombre de neurones dans la couche cachée, o_j représente la valeur de poids de connexion entre la couche de sortie et la couche cachée, et i_{kj} représente la valeur de poids de connexion entre la couche d'entrée et la couche cachée. La valeur de cette équation exprime, sous forme de pourcentage, la contribution de chaque facteur indépendamment des autres. Les résultats sont exposés dans le Tableau IV.

Le pH est de forte contribution, suivi à moindre degré par les autres facteurs. La contribution de la concentration des ions Ca^{2+} n'est pas à négliger, à l'opposé des conclusions déduites du modèle basé sur le plan composite centré.

Notre objectif est de choisir un modèle capable d'apprendre les données qu'on lui propose, mais aussi qui offre une bonne généralisation à des fins de prédiction, en d'autre terme éviter le surajustement lors d'un test. Pour ce faire, une des techniques utilisées consiste à décomposer l'ensemble de données en trois sous ensembles, le premier sert à l'apprentissage du modèle, le deuxième pour ajuster le modèle et éviter le surajustement en effectuant un arrêt prématuré quand son erreur augmente lors du calibrage. Le dernier sous ensemble sert à tester la performance du modèle. Cette technique est lourde à réaliser et pose la problématique du choix du deuxième sous ensemble pour le calibrage du modèle. Dans notre cas, nous utilisons la méthode basée sur la technique de régularisation Bayésienne par modération des poids, elle consiste à pénaliser les valeurs élevées des poids en modifiant

Tableau IV Analyse des contributions des différents facteurs

Variable	pH	(Ca/P) _i	D	[Ca ²⁺]
Contribution%	51,14	21,61	15,73	11,50

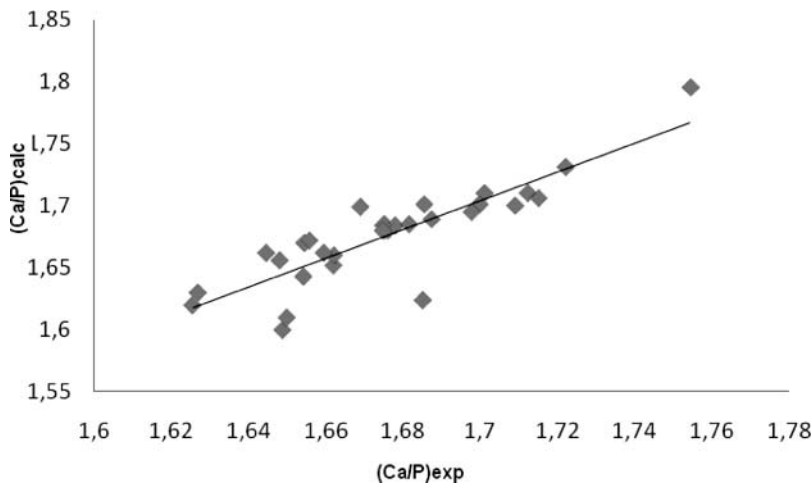


Figure 2 Relation entre rapports Ca/P observés et Ca/P calculés au moyen de la validation croisée.

la fonction de coût. Cette technique force les paramètres (les poids) à ne pas prendre des valeurs élevées, et par conséquent à éviter le surajustement.

Les résultats de la décomposition effectuée auparavant restent arbitraires puisqu'ils sont limités à un ensemble de test fixe. Donc, nous avons testé le pouvoir prédictif de notre modèle par la technique de la validation croisée (leave-20%-out cross-validation method). Dans cette procédure, nous retirons 20% de l'échantillon de base qui servira de test et le reste à l'apprentissage. Cette opération est répétée cinq fois jusqu'au test de la totalité des données disponibles. Pour toutes les décompositions, le choix des valeurs de poids initiales n'a presque pas d'effet, et on observe une convergence maximale de la fonction de coût après une centaine d'itérations en moyenne.

- Nombre de neurones dans la couche cachée: 3 neurones
- Nombre d'itération: 100 en moyenne

La relation entre le rapport Ca/P observée et calculée conduit à une parfaite corrélation, avec des paramètres statistiques très satisfaisants ($R = 0,92$ et $s = 0,014$) (Figure 2 et Tableau I). Rappelons que ce test porte sur la totalité des données. D'autre part, la nature de la décomposition de l'ensemble des données permet d'évaluer la performance du modèle. Lorsque les ensembles du test sont composés à partir de données prélevées un par un à intervalles égales (saut de quatre expériences), après avoir réalisé un classement par ordre croissant des données, les paramètres statistiques sont nettement améliorés ($R = 0,95$, $s = 0,012$), ceci est dû essentiellement au fait que le rapport Ca/P de l'ensemble d'apprentissage représente une parfaite empreinte des données dans l'ensemble du test.

CONCLUSION

Les réseaux de neurones avec la régularisation bayésienne fournissent un modèle robuste capable de prédire la totalité des essais des rapports Ca/P de l'hydroxyapatite phosphocalcique avec une grande performance, un coefficient de corrélation élevé avec une erreur type très faible pour la validation croisée et le test de l'ensemble des données. Il est essentiellement indépendant de l'architecture des réseaux de neurones. Il n'est nul besoin

d'effectuer une partition de l'ensemble des données pour la validation de l'apprentissage et par la suite la fixation des paramètres des réseaux de neurones. Par ailleurs, nous avons évalué au moyen de la méthode de Garson les contributions propres de tous les facteurs, épurés de toute interaction avec les autres facteurs. Les contributions de tous les facteurs sont significatives avec un ordre bien établi; le pH est de forte influence, suivi successivement par le rapport Ca/P, la durée de la réaction et la concentration des ions Ca^{2+} .

La modélisation par les plans d'expérience, dont une part des résultats est citée dans ce papier, permet d'obtenir un modèle simple et efficace dans la mesure où la relation fixe a priori, un modèle de corrélation figé. Pour ce faire, un polynôme de degré égal à 2 a été proposé. Elle n'est pas remise en cause, puisque sa validation a été vérifiée pour un rapport Ca/P = 1,67 au moyen des courbes isoréponses. Néanmoins les coefficients des variables d'ordre 2 sont non significatifs. Nous pensons qu'un plan factoriel simple aurait suffi pour obtenir le même résultat, et que seule la modélisation par les réseaux de neurones est capable d'approximer l'écart par rapport au modèle réel.

REFERENCES

1. O. Britel, M. Hamad, B. Sallek, H. Chaair, K. Digua, and H. Ouadadess, *Phosphorus, Sulfur, and Silicon*, **181**, 325 (2006).
2. J. Goupy, *Pratiquer les plans d'expériences* (Dunod, Paris, 2005).
3. SAS Institute, JMP Version 3.2, Introduction Guide, User's Guide and Statistics and Graphics Guide, USA (1995).
4. M. H. Hassoum, Ed., *Fundamentals of Artificial Neural Networks* (MIT Press, Cambridge, MA, 1995).
5. D. J. C. Mackay, *Comput. Neural Syst.*, **6**, 469 (1995).
6. D. J. C. Mackay, *Neural Comput.*, **4**, 415 (1992).
7. G. J. Dreyfus, M. Martinez, M. Samuelides, F. Gordon, S. Badran, S. Thiria, and L. Hérault, Eds., *Réseaux de neurones—Méthodologie et applications* (Eyrolles, Paris, 2002).
8. M. Nohair and D. Zakarya, *J. Mol. Model.*, **9**, 365 (2003).
9. M. Nohair, D. Zakarya, and A. Berrada, *J. Chem. Inf. Comput. Sci.*, **42**, 586 (2003).
10. S. Rudolph, *Neurocomputing*, **16**, 1 (1997).
11. MATLAB 7.0, The MathWorks, Inc. (Natick, MA, 2002).
12. G. D. Garson, *AI Expert*, **6**, 47 (1991).
13. I. A. Basheer, Y. M. Najaar, and M. N. Hajmer, *Environ. Technol.*, **17**, 795 (1996).