

REPUBLIQUE ALGERIENNE DEMOCRATIQUE ET POPULAIRE
Ministère de l'enseignement supérieur et de la recherche scientifique

Ecole Nationale Polytechnique



Département d'hydraulique
Thèse de doctorat en sciences
Spécialité : hydraulique

**OPTIMISATION DE LA GESTION DU BARRAGE
HAMMAM DEBAGH (GUELMA)
DANS LA SEYBOUSE - ALGÉRIE**

Djamel BENGORA

Sous la direction de M. Nourreddine DECHEMI Professeur

Présenté (e) et soutenue le 07/10/ 2019

Composition du jury :

Président	M. A. BALI,	Professeur	E.N. Polytechnique
Rapporteur/ Promoteur	M. N. DECHEMI,	Professeur	E.N. Polytechnique
Examineurs	M. A. BERMAD,	Professeur	E.N. Polytechnique
	Mme. D. SOUAG GAMANE,	Professeur	U. Houari Boumediene
	M. L. HOUICHI,	Professeur	U. Batna 2
	M. L. DJEMILI,	Professeur	U. Badji Mokhtar

ENP 2019

REPUBLIQUE ALGERIENNE DEMOCRATIQUE ET POPULAIRE
Ministère de l'enseignement supérieur et de la recherche scientifique

Ecole Nationale Polytechnique



Département d'hydraulique
Thèse de doctorat en sciences
Spécialité : hydraulique

**OPTIMISATION DE LA GESTION DU BARRAGE
HAMMAM DEBAGH (GUELMA)
DANS LA SEYBOUSE - ALGÉRIE**

Djamel BENGORA

Sous la direction de M. Nourreddine DECHEMI Professeur

Présenté (e) et soutenue le 07/10/ 2019

Composition du jury :

Président	M. A. BALI,	Professeur	E.N. Polytechnique
Rapporteur/ Promoteur	M. N. DECHEMI,	Professeur	E.N. Polytechnique
Examineurs	M. A. BERMAD,	Professeur	E.N. Polytechnique
	Mme. D. SOUAG GAMANE,	Professeur	U. Houari Boumediene
	M. L. HOUICHI,	Professeur	U. Batna 2
	M. L. DJEMILI,	Professeur	U. Badji Mokhtar

ENP 2019

Dédicaces

Je dédie ce modeste travail :

A la mémoire de mes parents, de mes sœurs et de ma femme Linda que Dieu les accueille dans son vaste paradis.

A ma femme Souheila et mes enfants.

A mes amis qui sont des frères.

A tous ceux qui me sont chers.

Djamel BENGORA

Remerciements

Dans la vie, tout travail réussi nécessite en premier lieu la bénédiction de Dieu, et ensuite l'aide et le support des humains. Je tiens donc tout d'abord à remercier Dieu le tout puissant, qui m'a donné la force et la patience d'accomplir ce modeste travail.

J'exprime ma profonde gratitude et mes sincères remerciements à mon directeur de thèse, le professeur **Nourreddine DECHEMI** pour ses qualités humaines et scientifiques. Je ne pourrai jamais oublier sa gentillesse, sa générosité, son esprit de recherche et ses commentaires efficaces. Un grand merci de m'avoir donné la chance de faire une thèse malgré ma situation contraignante.

Je remercie bien vivement le professeur **Lotfi KHIARI**, de l'université Laval au Québec, Canada, qui m'a reçu au cours des trois stages effectués à l'Université Laval. Ces stages m'ont fourni l'occasion de rencontrer les professeurs **Jacques Gallichand** et **Silvio- José Gumiere** et produire un article. Je le remercie encore une fois pour son aide et ses encouragements en particulier dans le domaine de la géostatistique spatiale. Son accueil chaleureux, sa disponibilité, ses conseils précieux et sa sympathie sont fort appréciés.

Mes vifs remerciements vont également aux membres du jury pour l'intérêt qu'ils ont porté à notre recherche en acceptant d'examiner ce travail.

Le Professeur de l'école nationale polytechnique d'Alger **A. BALI**, d'avoir accepté de présider le jury. Je le remercie pour l'intérêt qu'il a manifesté pour ce travail.

Le Professeur **A. BERMAD** de l'école nationale polytechnique d'Alger, le Professeur **L. HOUICHI** de l'université Batna 2, le Professeur **L. DJEMILI** de l'université Badji Mokhtar et Madame **D. SOUAG GAMANE** Professeur à l'Université Houari Boumediene d'avoir acceptés de lire et juger mes travaux, je les remercie pour leurs critiques et précieux conseils.

Cette thèse n'aurait pu voir le jour sans les données qui m'ont été gracieusement octroyées par l'Agence Nationale des Barrages et Transfert. Je tiens à exprimer mes sincères remerciements particulièrement à Mr **Kamel BAHRI**, gestionnaire du barrage « Hammam Debagh » qui m'a facilité l'accès aux données d'exploitation.

Je tiens également à mentionner et à témoigner ma reconnaissance à tous mes amis pour l'appui moral qu'ils m'ont offerts et l'aide qui m'ont apporté, et tout particulièrement le Professeur **H. BOUZRED**, le Docteur **M.Z. BELHADJ** et le Docteur **S. HEDDAM** de l'université 20 Août 1956 de Skikda, le Docteur **Z. GUERRAICH** et les enseignants **K. BENSMAIN** et **C. BELKHIRI** de l'université Batna1.

Enfin j'aimerais remercier et adresser ma reconnaissance à tous ceux qui m'ont soutenu, de loin ou de près, durant les moments difficiles.

المخلص

تمت دراسة اشكالية استغلال وتسيير كميات المياه المخزنة في السدود لمدة قرن تقريباً، إلا أنها لا تزال تحافظ على حداتها لاعتبار ترشيد توزيع الموارد المائية ضرورة ملحة نظراً لندرة المياه والطلب المتزايد عليها وتراكم الرواسب والاحوال في السدود مما ينقص من سعة تخزينها.

وقد عجزت حتى الآن الطرق الكلاسيكية في المعالجة الصحيحة لجميع القضايا التي ينبغي أخذها بعين الاعتبار من أجل تطوير نموذج يتم بموجبه تسيير احتياطات المياه التي تتم تعبئتها بواسطة السدود على نحو أفضل.

يبدو أن الشبكات العصبية الاصطناعية كطريقة بديلة لديها القدرة على نمذجة الاستغلال الأمثل والعقلاني للسدود. إلا أنه عند تطوير النموذج الذي يتم به تسيير السدود يجب الأخذ بعين الاعتبار نقطتين منهجيتين متعلقتين بكل من زيادة وتيرة المسح الطبوغرافي وقياسات الأعماق للسدود من أجل التحيين الدوري لمنحنى " السعة مقابل الارتفاع " وكذا اختيار مدخلات النموذج لتسهيل استعماله بطريقة أسرع.

بداية قمنا بتطبيق أدوات الجيوإحصاء المكانية لتحسين حجم المسح الطبوغرافي وقياسات الأعماق لسد حمام دباغ من أجل زيادة وتيرة التنفيذ.

لقد اتبعنا نهجاً استكشافياً بمقارنة نتائج التحليل لجميع المعطيات المأخوذة من المسح الأولي لقياس الأعماق والسيناريوهات التسعة عشر المختارة عشوائياً لأجل التقليل من حجم هذا المسح. سمحت لنا هذه الدراسة بأن نوضح أن تخفيض عدد نقاط المسح من 32 164 إلى 10 000 نقطة و التي تقابل على التوالي للكثافات 42 و 13 نقطة في الهكتار له تأثير ضئيل جداً على معايير الموثوقية الجيوإحصائية.

في الواقع، أظهر هذا الحجم من 10 000 نقطة تشابه كبير مع المسح الأولي بدراسة طبيعة المعطيات وتغيراتها المكانية والتحقق من صحة أداء النموذج.

في التالي، على أساس قاعدة بيانات شهرية لفترة واحد وعشرون (21) سنة من الاستغلال التي تم جلبها من الوكالة الوطنية للسدود والتحويلات (ANBT)، اعتمدنا الطريقة الآتية:

أولاً، اخترنا أشعة المدخلات ليتم استخدامها في إنشاء نماذج التنبؤ للتزويد بالأحجام المائية الكلية (الاستهلاك الإنساني والزراعي) بعد تحديث قاعدة البيانات وذلك بالأخذ بعين الاعتبار نتائج المسح الوحيد LTB (تخفيض الحجم $0,98 \text{ hm}^3$ في السنة).

سمح لنا تطبيق تقنية اختبار جاما (Gamma Test) باختبار خمسة (05) متغيرات مهمة

($VTAL_{t-1}$, PLU_{t-1} , $EVAP_{t-1}$, APP_{t-1} et $VTAL_{t-2}$) وأربعة أشعة مدخلات من بين كل المجموعات الممكنة من تلك المتغيرات.

ثانياً، استكشفنا الفروق بين ثلاثة أنواع من النماذج العصبية الاصطناعية (MLPNN, GRNN et RBFNN) باستعمال أشعة المدخلات الأربعة المختارة. تبين من خلال البحث أن النماذج الثلاثة أدوات قوية للتنبؤ بالحجم الإجمالي المخصص للتزويد.

بناء على نتائج مرحلة التحقق من صحة أداء النماذج تبين أن نموذج GRNN1 (خمس مدخلات) هو الأفضل مع القيم الأعلى لمعامل الارتباط ($R = 0,876$) وكفاءة ناش - ساتكليف ($NSE = 0,684$) والقيم الأضعف للجذر التربيعي لمتوسط الخطأ التربيعي ($RMSE = 1,614$) ومتوسط الخطأ المطلق $MAE = 1,157$ يليه في المرتبة الثانية نموذج MLPNN1 وأخيراً RBFNN1 في المركز الثالث.

الكلمات الدالة: السد، الطمي، قياس الأعماق، النمذجة، الجيوإحصاء، اختبار جاما، الشبكات العصبية الاصطناعية.

Abstract

The problem of management of the dam is studied during nearly one century but this management preserves unquestionable news owing to the fact that the rational allowance of the water resources seems an imperative need, taking into account the scarcity of water, the quasi exponential growth of the request and the accumulation of the sediments which reduces considerably water storage capacity of the dam.

The classic methods have, until now, been unable to take care correctly all the questions which must be considered to develop a model which allows managing at best water supply mobilized by dams.

Neural networks seem to have the potential of an alternative method which can be used in the optimization of the management of dam reservoirs. However, the development of the business model must support two methodological questions concerning both the increase in the frequency of the topo-bathymetric surveys for a periodic update of the curve (capacity – height) and the selection of the input variables of the model for easy accessibility and quick use

For starters, we applied the geostatistical spatial tools to optimize the size of the topo-bathymetric survey of the Hammam Debagh dam and thus increase the frequency of realization.

We followed an exploratory approach by comparing the results of the geostatistical analysis of all the data of topo-bathymetric survey (LTB) original and 19 random scenarios of reduction in the size of this survey. This study allowed us to show that reducing the subset size from 32 164 to an optimal 10 000 points, corresponding to measurement densities of 42 to 13 points ha⁻¹, respectively, has only a slight impact on the geostatistical reliability criteria. Indeed, this size of 10 000 points resulted in a great similarity with the original LTB from the study of the normality of the variography until cross-validation.

Subsequently on a monthly database of twenty-one (21) years of operation, recovered from the ANBT we proceeded in this manner.

As a first step, we have selected the input vectors to be used for the development of the prediction models of volumes to be allocated after having updated the database taking into account the results of the unique LTB (reduction the volume of 0.98 hm³/year).

The application of the Gamma Test technique allowed us to select five (05) dominant variables ($VTAL_{t-1}$, PLU_{t-1} , $EVAP_{t-1}$, APP_{t-1} et $VTAL_{t-2}$) and four input vectors from the all combinations of these.

In a second time we explored the differences between three types of artificial neural models (MLPNN, GRNN and RBFNN) for the four input vectors retained. Indeed, it seems that all three models are powerful tools for predicting the total volume allocated.

However, it is important to note that in the validation phase, the GRNN1 model (five inputs) is considered as the best model with the Pearson Correlation Coefficient ($R = 0.876$) and the Nash-Sutcliffe Efficiency ($NSE = 0.684$), the highest and the Root Mean Squared Error ($RMSE = 1.614$) and the Mean Absolute Error ($MAE = 1.157$) the weakest; followed by the MLPNN1 model and finally the RBFNN1 in the third place.

Keywords: Dam, Siltation, Bathymetry, Modeling, Geostatistics, Gamma Test, Artificial Neural Networks (ANN).

Résumé

Le problème de gestion des barrages réservoirs est étudié depuis près d'un siècle, mais cette gestion conserve une actualité certaine du fait que l'allocation rationnelle des ressources en eau apparaît comme une nécessité impérieuse, compte tenu de la rareté de l'eau, la croissance quasi exponentielle de la demande et l'accumulation des sédiments qui réduit considérablement la capacité de stockage en eau des barrages.

Les méthodes classiques sont incapables de prendre en charge correctement l'ensemble des questions qui doivent être considérées pour élaborer un modèle qui permet de gérer au mieux les réserves d'eau mobilisées par les barrages.

Les réseaux de neurones semblent avoir le potentiel d'une méthode alternative qui peut être utilisée dans l'optimisation de la gestion des barrages réservoirs.

Cependant, l'élaboration du modèle de gestion doit prendre en charge deux questions méthodologiques. Elles concernent à la fois l'augmentation des fréquences de réalisation des levés topo-bathymétrique pour une mise à jour périodique de la courbe « capacité - hauteur » et la sélection des variables d'entrée du modèle pour une accessibilité facile et une utilisation rapide.

Pour commencer, nous avons appliqué les outils de la géostatistique spatiale afin d'optimiser la taille du levé topo-bathymétrique du barrage Hammam Debagh et augmenter ainsi les fréquences de réalisation de ces levés.

Nous avons suivi une démarche exploratoire en comparant les résultats de l'analyse géostatistique de l'ensemble des données du levé topo-bathymétrique (LTB) initial et les 19 scénarios aléatoires de réduction de la taille de ce levé. Cette étude nous a permis de démontrer que la réduction du nombre de points de mesure de 32 164 à 10 000, correspondant respectivement à des densités de 42 et de 13 points à l'hectare, a un impact très faible sur les critères géostatistiques de fiabilité. En effet, cette taille de 10 000 points a montré une grande similitude avec le LTB initial depuis l'étude de la normalité, de la variographie jusqu'à la validation croisée.

Par la suite, sur une base de données mensuelles d'une période de vingt et un (21) ans d'exploitation récupérée auprès de l'ANBT, nous avons procédé de la sorte :

Dans un premier temps, nous avons sélectionné les vecteurs d'entrée à utiliser pour l'élaboration des modèles de prédiction des volumes totaux à allouer (Alimentation en eau potable et Irrigation), et ce, après avoir procédé à la mise à jour de la base de données en tenant compte des résultats de l'unique LTB (réduction du volume de 0,98 hm³/an).

L'application de la technique du Gamma Test, nous a permis de sélectionner cinq (05) variables prépondérantes (VTAL_{t-1}, PLU_{t-1}, EVAP_{t-1}, APP_{t-1} et VTAL_{t-2}), et quatre (04) vecteurs de variables d'entrée parmi toutes les combinaisons possibles de ces dernières.

Dans un deuxième temps, nous avons exploré les différences entre trois types de modèles de neurones artificiels (MLPNN, GRNN et RBFNN) en utilisant les quatre vecteurs d'entrée retenus. En effet, il s'est avéré que les trois modèles sont des outils performants pour la prévision des volumes totaux à allouer.

Cependant, il est important de signaler qu'en phase de validation, le modèle GRNN1 (cinq variables d'entrée) est considéré comme le meilleur avec un coefficient de corrélation ($R = 0,876$) et le critère de Nash-Sutcliffe ($NSE = 0,684$) les plus élevés d'une part, et la Racine Carrée de l'Erreur Quadratique Moyenne ($RMSE = 1,614$) et l'erreur absolue moyenne ($MAE = 1,157$) les plus faibles d'autre part. Il est suivi par le modèle MLPNN1 en deuxième position et enfin le RBFNN1 en troisième place.

Mots-clés : Barrage, Envasement, Bathymétrie, Modélisation, Géostatistique, Gamma Test, Réseaux de neurones artificiels (RNA)

Table des matières

Liste des tableaux

Liste des figures

Liste des sigles

Liste des abréviations

Introduction.....	24
Chapitre 1 : Synthèse bibliographique.....	28
Chapitre 2 : Présentation de l'aire d'étude.....	34
2.1 Introduction.....	34
2.2 Localisation du site.....	34
2.3 Climat.....	36
2.4 Bassin versant Bouhamdane.....	36
2.5 Barrage Hammam Debagh.....	39
2.6 Affectation des eaux du barrage Hammam Debagh.....	42
2.6.1 Alimentation en eau potable.....	42
2.6.2 Alimentation en eau d'irrigation.....	43
2.7 Conclusion.....	45
Chapitre 3 : Collecte, critique et traitement des données.....	47
3.1 Introduction.....	47
3.2 Présentation des bases de données.....	47
3.2.1 Acquisition des données topo-bathymétriques.....	47
3.2.1.1 Système géodésique utilisé.....	47
3.2.1.2 Réseau d'appui.....	48
3.2.1.3 Levé topographique.....	49
3.2.1.4 Levé bathymétrique.....	49
3.2.1.5 Base de données recueillie.....	51
3.2.2 Acquisition des données d'exploitation.....	51
3.2.2.1 Paramètres mesurés.....	51
3.2.2.2 Paramètres estimés.....	52
3.2.2.3 Données disponibles récupérées.....	53
3.3 Critique et traitement des bases de données.....	53

3.3.1 Base de données topo-bathymétriques.....	54
3.3.2 Base de données d'exploitation.....	54
3.3.3 Traitement des données d'exploitation.....	56
3.3.3.1 Méthodologie proposée.....	56
3.3.3.2 Réajustement des volumes, surfaces et évaporations au niveau de la retenue.....	56
3.3.3.3 Etablissement du bilan hydraulique de la retenue.....	60
3.3.4 Analyse des données d'exploitation.....	64
3.3.4.1 Statistique de base des données d'exploitation traitées.....	64
3.3.4.2 Visualisation graphique de l'évolution des paramètres d'exploitation.....	65
3.4 Conclusion.....	72
CHAPITRE 4 : Optimisation du levé Topo-bathymétrique du barrage Hammam Debagh.....	74
4.1 Introduction et objectifs.....	74
4.2 Qu'est-ce que la géostatistique ?.....	74
4.3 Bref historique de la géostatistique.....	75
4.4 Domaines d'utilisation de la géostatistique.....	77
4.5 Présentation de la géostatistique.....	78
4.5.1 Définitions.....	78
4.5.1.1 Variable régionalisée.....	78
4.5.1.2 Variable aléatoire.....	79
4.5.2 Hypothèses de base.....	79
4.5.2.1 Hypothèse de stationnarité.....	79
4.5.2.2 Hypothèse intrinsèque.....	80
4.5.3 Analyse variographique.....	80
4.5.3.1 Variogramme.....	81
4.5.3.2 Propriétés théoriques du variogramme.....	81
4.5.3.3 Estimation du variogramme.....	84
4.5.3.4 Modélisation du variogramme expérimental.....	84
4.5.3.5. Calcul du variogramme expérimental et ajustement de modèles.....	90
4.5.3.6. Conclusion.....	91

4.5.4 Interpolation spatiale.....	91
4.5.4.1 Méthodes d'interpolations déterministes.....	92
4.5.4.2 Méthodes d'interpolation stochastiques (cas du Krigeage).....	96
4.5.5 Validation croisée.....	102
4.6 Application au levé topo-bathymétrique du barrage Hammam Debagh.....	103
4.6.1 Mise en œuvre de l'analyse géostatistique.....	103
4.6.1.1 Analyse exploratoire.....	103
4.6.1.2 Analyse variographique.....	105
4.6.1.3 Krigeage.....	106
4.6.1.4 Validation croisée.....	106
4.6.2 Méthodologie.....	107
4.6.3 Résultats et discussions.....	108
4.6.3.1 Analyse exploratoire.....	108
4.6.3.2 Analyse variographique.....	110
4.6.3.3 Krigeage et validation croisée.....	114
4.7 Conclusion.....	120
CHAPITRE 5 : Optimisation de la gestion des barrages par les réseaux de neurones artificiels.....	123
5.1 Introduction et objectifs.....	123
5.2 Modèle des neurones biologiques et formels.....	123
5.3 Type de Réseaux de Neurones.....	124
5.4 Perceptron multicouche (MLPNN).....	124
5.5 Réseaux de neurones à fonction de régression générale (GRNN).....	128
5.6 Réseaux de neurones à fonctions de base radiales (RBFNN).....	130
5.7 Identification des entrées pertinentes par la technique Gamma Test.....	132
5.7.1 Introduction.....	132
5.7.2 Description de la Technique Gamma Test.....	132
5.7.3 Formulation mathématique de la Technique Gamma Test.....	133
5.8 Stratégie de développement du meilleur modèle.....	135
5.8.1 Application de la Technique du Gamma Test.....	137
5.8.1.1 Sélection des variables indispensables par le Gamma Test.....	137

5.8.1.2 Classification des variables sélectionnées par le Gamma Test.....	138
5.8.1.3 Sélection des meilleurs vecteurs d'entrée.....	138
5.9 Application de trois modèles RNA pour l'estimation du volume total alloué.....	140
5.9.1 Généralités et structures des modèles.....	140
5.9.2 Résultats obtenus par les modèles RNA.....	141
5.9.2.1 Résultats obtenus par les réseaux de neurones type MLPNN.....	142
5.9.2.2 Résultats obtenus par les réseaux de neurones type GRNN.....	148
5.9.2.3 Résultats obtenus par les réseaux de neurones type RBFNN.....	153
5.9.2.4 Comparaison des résultats obtenus par les réseaux de neurones.....	158
5.10 Conclusion.....	159
Conclusion générale.....	160
Bibliographie.....	162

Liste des tableaux

Tableau 2-1 : Caractéristiques météorologiques principales (1990-2011).....	36
Tableau 2-2 : Caractéristiques du bassin versant de l'oued Bouhamdane.....	39
Tableau 2-3 : Caractéristiques du barrage Hammam Debagh.....	42
Tableau 2-4 : Superficies des secteurs et leur besoin en eau.....	43
Tableau 3-1 : Coordonnées des bornes du réseau d'appui.....	48
Tableau 3-2 : Exemples d'erreurs de volumes.....	54
Tableau 3-3 : Exemples sur l'établissement de la balance des volumes.....	55
Tableau 3-4 : Évolution des volumes d'eau selon le levé topo-bathymétrique de Hammam Debagh.....	58
Tableau 3-5 : Équations de tendance et coefficients de détermination.....	58
Tableau 3-6 : Évolution des surfaces d'eau selon le levé bathymétrique de Hammam Debagh.....	59
Tableau 3-7 : Équations de tendance et coefficients de détermination.....	60
Tableau 3-8 : Bilan hydraulique journalier du barrage Hammam Debagh.....	63
Tableau 3-9 : Statistique descriptive des données traitées (252 observations).....	64
Tableau 4-1 : Statistiques descriptives de l'ensemble de données du LTB (32164 points) et les 19 échantillons (400 à 30000 points).....	110
Tableau 5-1 : Résultats du Gamma Test sur les variables d'entrée.....	138
Tableau 5-2 : Classification des variables d'entrée par le Gamma Test.....	138
Tableau 5-3 : Résultats des combinaisons des variables d'entrée : quatre et cinq variables.....	139
Tableau 5-4 : Résultats des combinaisons des variables d'entrée : trois variables.....	139
Tableau 5-5 : Résultats des combinaisons des variables d'entrée : deux variables.....	139
Tableau 5-6 : Structures des modèles des réseaux de neurones utilisés.....	140
Tableau 5-7 : Performances des modèles MLPNN1, MLPNN2, MLPNN3 et MLPNN4.....	142
Tableau 5-8 : Performances des modèles GRNN1, GRNN2, GRNN3 et GRNN4.....	148
Tableau 5-9 : Performances des modèles RBFNN1, RBFNN2, RBFNN3 et RBFNN4.....	153
Tableau 5-10 : Comparaison des performances des modèles MLPNN1, GRNN1 et RBFNN1.....	158

Liste des figures

Figure 2-1 : Localisation du site d'étude.....	35
Figure 2-2 : Localisation et délimitation du bassin versant de l'oued Bouhamdane (Source : Carte du réseau hydro climatologique et de surveillance de la qualité des eaux du Nord d'Algérie au 1 :500 000 ème, 2005).....	38
Figure 2-3 : Coupe en travers du barrage Hammam Debagh (source : ANBT (ex ANB), 1985).....	40
Figure 2-4 : Le barrage Hammam Debagh : évacuateur de crues de type Corolle et Seuil libre (source : ANBT, 2010).....	41
Figure 2-5 : Schéma d'aménagement et organisation du périmètre Guelma-Boucheouf Source : (ONID (ex AGID) Guelma, 2011).....	44
Figure 3-1 : Schéma du réseau des bornes d'appui (Source : ANB, 2005).....	48
Figure 3-2 : Profils topo-bathymétriques (source : ANB, 2005).....	50
Figure 3-3 : Variation des volumes d'eau et de vase en fonction de la côte du niveau d'eau dans le réservoir.....	57
Figure 3-4 : Variation temporelle des volumes de vase en fonction de la côte du niveau d'eau dans le réservoir.....	58
Figure 3-5 : Variation des surfaces inondées en fonction du niveau d'eau dans le réservoir.....	59
Figure 3-6 : Variation temporelle des surfaces inondées en fonction de la côte du niveau d'eau dans le réservoir.....	60
Figure 3-7 : Schéma du bilan d'un réservoir.....	61
Figure 3-8 : Variation mensuelle des volumes stockés dans le barrage « 252 observations ».....	65
Figure 3-9 : Variation mensuelle des apports « 252 observations ».....	66
Figure 3-10 : Variation mensuelle des volumes alloués à l'AEP « 252 observations ».....	67
Figure 3-11 : Variation mensuelle des volumes alloués à l'irrigation « 252 observations ».....	68

Figure 3-12 : Variation mensuelle des volumes évaporés « 252 observations ».....	68
Figure 3-13 : Variation mensuelle de la pluie au niveau du barrage « 252 observations ».....	69
Figure 3-14 : Variation mensuelle des volumes de vidanges « 252 observations ».....	70
Figure 3-15 : Variation mensuelle des volumes de fuites « 252 observations ».....	70
Figure 3-16 : Variation mensuelle des volumes évacués à travers le barrage « 252 observations ».....	71
Figure 4-1 : Exemple de variogramme avec palier.....	84
Figure 4-2 : Modèle théorique d'un variogramme pépitique pur.....	86
Figure 4-3 : Modèle théorique d'un variogramme linéaire avec palier.....	86
Figure 4-4 : Modèle théorique d'un variogramme sphérique.....	87
Figure 4-5 : Modèle théorique d'un variogramme exponentiel.....	88
Figure 4-6 : Modèle théorique d'un variogramme gaussien.....	89
Figure 4-7 : Modèle théorique d'un variogramme linéaire sans palier.....	90
Figure 4-8 : Types d'interpolation.....	92
Figure 4-9 : Interpolation polygonale.....	93
Figure 4-10 : Interpolation par distances inverses.....	95
Figure 4-11 : Procédure de l'analyse géostatistique.....	103
Figure 4-12 : Variogrammes expérimental et théorique.....	111
Figure 4-13 : Carte variographique 2D.....	112
Figure 4-14 : Valeurs des paramètres variographiques pour les 19 échantillons de réduction selon leur taille : (a) Palier, (b) Portée, (c) Fraction structurée de la variance totale, (d) Coefficient de détermination et (e) Somme des carrés des résidus. La ligne pointillée représente les valeurs des paramètres variographiques du LTB de référence.....	113
Figure 4-15 : Validation croisée du modèle de LTB de référence (32164 points).....	115
Figure 4-16 : Evolution des paramètres de la validation croisée des 19 échantillons de réduction selon leur taille : (a) coefficient de régression, (b) intercepte, (c) coefficient de détermination, (d) erreur standard de prédiction et (e) erreur standard. La ligne pointillée représente les valeurs des paramètres de validation croisée du LTB de référence.....	117

Figure 4-17 : Evolution des volumes estimés au NNR : des 19 échantillons, La ligne pointillée représente le volume estimé au NNR avec le LTB de référence.....	118
Figure 4-18 : Courbes « volume-hauteur » du LTB de référence et de l'échantillon optimal (10 000 points).....	119
Figure 4-19 : Evolution des volumes estimés au NNR : des 12 combinaisons de taille 10 000 points. La ligne pointillée représente le volume estimé au NNR avec le LTB de référence.....	120
Figure 5-1 : Modèle du neurone formel de MCCulloch et Pitts (1948).....	124
Figure 5-2 : Architecture du perceptron multi couches (MLPNN).....	126
Figure 5-3 : Différentes fonctions d'activation : (a) Identité (Linéaire), (b) Sigmoidé et (c) Tangente hyperbolique.....	127
Figure 5-4 : Schéma détaillé d'un Réseaux de neurones à fonction de régression générale (GRNN).....	128
Figure 5-5 : Schéma détaillé d'un Réseaux de neurones à fonctions radiales de base (RBFNN).....	131
Figure 5-6 : Organigramme de développement des modèles.....	136
Figure 5-7 : Corrélation des volumes totaux alloués observés et estimés par les modèles MLPNN1, MLPNN2, MLPNN3 et MLPNN4 en phase d'apprentissage.....	144
Figure 5-8 : Evolution des volumes totaux alloués observés et estimés par les modèles MLPNN1, MLPNN2, MLPNN3 et MLPNN4 en phase d'apprentissage.....	145
Figure 5-9 : Corrélation des volumes totaux alloués observés et estimés par les modèles MLPNN1, MLPNN2, MLPNN3 et MLPNN4 en phase de validation.....	146
Figure 5-10 : Evolution des volumes totaux alloués observés et estimés par les modèles MLPNN1, MLPNN2, MLPNN3 et MLPNN4 en phase de validation.....	147
Figure 5-11 : Corrélation des volumes totaux alloués observés et estimés par les modèles GRNN1, GRNN2, GRNN3 et GRNN4 en phase d'apprentissage.....	149
Figure 5-12 : Evolution des volumes totaux alloués observés et estimés par les modèles GRNN1, GRNN2, GRNN3 et GRNN4 en phase d'apprentissage.....	150
Figure 5-13 : Corrélation des volumes totaux alloués observés et estimés par les modèles GRNN1, GRNN2, GRNN3 et GRNN4 en phase de validation.....	151

Figure 5-14 : Evolution des volumes totaux alloués observés et estimés par les modèles GRNN1, GRNN2, GRNN3 et GRNN4 en phase de validation.....	152
Figure 5-15 : Corrélation des volumes totaux alloués observés et estimés par les modèles RBFNN1, RBFNN2, RBFNN3 et RBFNN4 en phase d'apprentissage.....	154
Figure 5-16 : Evolution des volumes totaux alloués observés et estimés par les modèles RBFNN1, RBFNN2, RBFNN3 et RBFNN4 en phase d'apprentissage.....	155
Figure 5-17 : Corrélation des volumes totaux alloués observés et estimés par les modèles RBFNN1, RBFNN2, RBFNN3 et RBFNN4 en phase de validation.....	156
Figure 5-18 : Evolution des volumes totaux alloués observés et estimés par les modèles RBFNN1, RBFNN2, RBFNN3 et RBFNN4 en phase de validation.....	157

Liste des sigles

AGID	Agence Nationale de Réalisation et de Gestion des Infrastructures Hydrauliques pour l'Irrigation et le Drainage.
ANB	Agence nationale des barrages.
ANBT	Agence Nationale des Barrages et Transferts.
ANRH	Agence Nationale des Ressources Hydrauliques.
LEM	Laboratoire des Études Maritimes.
LEM - GEOID	Consortium de bureaux d'études maritimes.
ONID	Office national de l'irrigation et du drainage.
WCD	World Commission on Dams.

Liste des abréviations

ACP	Analyse en composantes principales.
AEP	Volume alloué à l'alimentation en eau potable en (hm ³).
ANFIS	Adaptive-Network-based Fuzzy Inference System.
APP :	Volume des apports globaux entrant dans le barrage en (hm ³).
APP _{t-1}	Volume des apports globaux entrant dans le barrage au temps (t-1) en (hm ³).
APP _{t-2}	Volume des apports globaux entrant dans le barrage au temps (t-2) en (hm ³).
CAO	Conception assistée par ordinateur.
DAO	Dessin assisté par ordinateur.
EVAC	Volume déversé en (hm ³).
EVAP	Volume évaporé en (hm ³).
EVAP _{t-1}	Volume évaporé au temps (t-1) en (hm ³).
EVAP _{t-2}	Volume évaporé au temps (t-2) en (hm ³).
FFBP	Feed Forward Back Propagation.
FUI	Volume des fuites en (hm ³).
FUI _{t-1}	Volume des fuites au temps (t-1) en (hm ³).
FUI _{t-2}	Volume des fuites au temps (t-2) en (hm ³).
GPS	Global Positioning System.
GRNN	General Regression Neural Networks.
GRS-80	Geodetic Reference System 1980.
GT	Gamma Test.
IRR	Volume alloué à l'irrigation en (hm ³).
ITRS	International Terrestrial Reference System.
JORADP	Journal Officiel de la République Algérienne Démocratique et Populaire.

LM	Levenberg-Marquardt.
LTB	Levé Topo-Bathymétrie.
MAE	Mean Absolute Error.
MLPNN	Multi-Layer Perceptron Neural Networks.
MLR	Multiple Linear Regression.
MNT	Modèle Numérique de Terrain.
MSE	Mean square error.
NGA	Nivellement Général d'Algérie.
NNR	Niveau Normal de la Retenue en (m).
NSE	Nash-Sutcliffe efficiency.
PHE	Niveau Plus Hautes Eaux en (m).
PHEE	Niveau Plus Hautes Eaux Exceptionnelles en (m).
PLU	Pluie mesurée au niveau du barrage en (mm).
PLU _{t-1}	Pluie mesurée au niveau du barrage au temps (t-1) en (mm).
PLU _{t-2}	Pluie mesurée au niveau du barrage au temps (t-2) en (mm).
RBFNN	Radial Basis Function Neural Networks.
RMSE	Root Mean Square Error.
RNA	Réseaux de neurones artificiels.
RSS	Residual Sum of Squares.
RTK	Real Time Kinematic.
SD	Standard Deviation.
SE	Standard Error.
SE _{pred}	Standard Error prédiction.
SIG	Système d'Information Géographique.
TGO	Trimble Geomatics Office.

TSC1	Time Stamp Counter.
UHF	Ultra hautes fréquences.
URSS	Union des républiques socialistes soviétiques.
UTM	Universal Translator Mercator.
Vfinal	Volume stocké au temps (t+1) en (hm ³).
VID	Volume de vidange en (hm ³).
Vinitial	Volume stocké au temps (t) en (hm ³).
VOL	Volume stocké dans le barrage en (hm ³).
VTAL _t	Volume total alloué à l'AEP et l'irrigation au temps (t) en (hm ³).
VTAL _{t-1}	Volume total alloué à l'AEP et l'irrigation au temps (t-1) en (hm ³).
VTAL _{t-2}	Volume total alloué à l'AEP et l'irrigation au temps (t-2) en (hm ³).
WGS-84	World Geodetic System 1984.

Introduction

Aujourd'hui plus que par le passé, la question de l'eau se pose avec acuité dans le monde. Cette inquiétude est due à la fois à la succession des années de sécheresse et à l'accroissement considérable de la demande en eau. En effet, avec la croissance démographique, l'évolution des secteurs de l'industrie et de l'agriculture, les besoins en eau se sont considérablement accrus amplifiant ainsi les problèmes de disponibilité de l'eau.

Les barrages constituent donc la solution la plus appropriée pour augmenter la disponibilité des ressources en eau par la régulation des flux d'eau entrant et sortant. Il est donc crucial de rationaliser la gestion de l'eau mobilisée par ces ouvrages dans le but de diminuer les périodes critiques hydriques à court, moyen et long terme.

Dans le domaine de la gestion des barrages réservoirs, le programme des lâchers est un des problèmes les plus complexes et délicats à résoudre. Ceci est dû essentiellement aux modifications de la structure hydrique pouvant être engendrées par la variation des demandes et des entrées d'eau. Ainsi le développement de modèles de prédiction des volumes à allouer (une composante des lâchers à programmer) s'impose d'une part par l'accroissement des besoins en eau et d'autre part par la réduction des capacités de stockage des barrages suite à leurs envasement.

Problématique et objectifs de la thèse

La problématique de cette recherche s'articule autour d'un certain nombre de questions :

- 1- Est-il possible d'augmenter la fréquence des levés topo-bathymétriques pour une estimation régulière de l'impact de l'envasement sur la disponibilité de l'eau dans les barrages et par conséquent sur la courbe « volume -hauteur » ?
- 2- Comment procédera-t-on à la mise à jour de la courbe « volume -hauteur » (pièce maitresse dans la gestion des barrages réservoirs) après chaque levé topo-bathymétrique ?

- 3- Peut-on prédire les volumes totaux à allouer par les modèles connexionnistes ?
Et peut-on optimiser le nombre de variables d'entrée de ces modèles ?

Notre travail a pour objectifs l'exploration de deux axes de recherche, à savoir :

- 1- L'optimisation des levés topo-bathymétriques, en utilisant les outils de la géostatistique afin de réduire la densité des points de levé à l'hectare et par conséquent augmenter les fréquences de réalisation de ces levés. Donc déterminer le nombre de points de levé topo-bathymétrique nécessaires et suffisants pour évaluer avec précision le volume d'eau retenu dans un barrage.
- 2- La prévision des volumes à allouer par des modèles connexionnistes tout en réduisant le nombre de variables d'entrées. Trois types de réseaux de neurones seront étudiés et comparés, à savoir :
 - ☞ Le perceptron multicouches (MLPNN).
 - ☞ Les réseaux de neurones à fonction de régression générale (GRNN).
 - ☞ Les réseaux de neurones à fonction de base radiale (RBFNN).

La sélection des variables d'entrée, en considérant leur nombre optimum, a été réalisée par la technique du Gamma Test.

Démarche et organisation de la thèse

Le premier chapitre est consacré à la synthèse bibliographique ; nous présentons notamment les travaux de recherche relatifs à l'utilisation des outils géostatistiques pour le traitement des levés topo-bathymétriques et ceux relatifs à l'emploi des techniques de l'intelligence artificielle pour l'optimisation de la gestion des barrages réservoirs.

Le second chapitre présente le Barrage Hammam Debagh objet du présent travail ainsi que son bassin versant (Oued Bouhamdane) et ses différents utilisateurs.

Le troisième chapitre est consacré à la présentation des bases de données recueillies auprès de l'exploitant du barrage, à savoir : celle du levé topo-bathymétrique et celle de l'exploitation. Les étapes de leurs acquisitions y sont détaillées. La base de données d'exploitation qui regroupe huit (08) variables observées durant 21 ans a fait l'objet

d'une critique et d'un traitement de correction. Une approche a été développée pour mettre à jour et rendre crédible l'ensemble des données d'exploitation en tenant compte de l'impact de l'envasement à travers l'unique levé topo-bathymétrique de 2004.

Le quatrième chapitre se focalise sur les grandes lignes de la théorie des variables régionalisées qui permettent de mener à bien une analyse géostatistique. Les notions de base de la géostatistique sont d'abord présentées, suivies par une mise en œuvre relatant ses différentes étapes. La fin du chapitre est dédiée à son application sur le levé topo-bathymétrique du barrage Hammam Debagh. Les résultats ont été publiés dans la revue : *International Journal of Sediment Research*.

Le cinquième chapitre présente les trois types de réseaux de neurones artificiels qui ont été utilisés pour la prévision des volumes totaux à allouer. La technique du Gamma Test est aussi exposée. Après optimisation des variables d'entrée, les trois modèles ont été testés. Nous présentons d'abord les résultats obtenus avec ces modèles de réseaux de neurones artificiels (MLPNN, GRNN et RBFNN), par la suite, nous confrontons leurs performances.

Chapitre 1

Chapitre 1 : Synthèse bibliographique

L'objectif principal des barrages réservoirs est d'assurer la disponibilité d'eau dans l'espace et dans le temps. La stratégie de gestion d'un réservoir peut être influencée par beaucoup d'aspects dépendant de ses objectifs. Outre les dissemblances dues aux divers buts, l'échelle de temps de la politique de gestion peut être implicitement modifiée. Le contrôle des crues d'un réservoir, par exemple pourrait exiger une règle de gestion à court terme (quotidienne, horaire ou même pour de plus courtes périodes), tandis qu'une politique de gestion prédéterminée à long terme (mensuelle, décadaire, hebdomadaire, etc.) pourrait être appropriée pour un réservoir assurant des objectifs comme l'approvisionnement en eau d'irrigation ou domestique.

La commission mondiale des barrages ([WCD, 2000](#)) souligne que les réservoirs d'eau ont joué un rôle important dans le développement de l'agriculture irriguée de nombreuses régions du monde. Toutefois, la croissance démographique et le développement des activités socio-économiques ont augmenté considérablement les demandes en eau et en ont rendu l'accès plus difficile. D'autre part, l'envasement des réservoirs des barrages constitue une menace pour leur durabilité, puisqu'il réduit en moyenne leur capacité de stockage d'eau d'environ 1% par année ([WCD, 2000](#) ; [Kummu & Varis, 2007](#)). Plusieurs techniques d'estimation de cette capacité de stockage sont appliquées pour assurer une gestion efficace de l'eau ([Rakhmatullaev et al., 2011](#)). L'une des méthodes les plus précises pour établir des bilans volumiques des terres érodées et d'en déduire les volumes d'eau disponibles dans les barrages est la topo-bathymétrie ([Charles, 2009](#)). Or les levés topo-bathymétriques doivent être effectués régulièrement en raison des changements rapides de la forme du fond du réservoir. Ces levés topo-bathymétriques combinent les levés topographiques dans les zones exondées et les levés bathymétriques dans les zones inondées des réservoirs.

Ces deux levés, utilisant des équipements de positionnement par GPS en mode RTK et des échosondeurs ([Barbosa et al., 2006](#)), sont indispensables pour la précision d'estimation et la gestion optimale des volumes d'eau dans les réservoirs. Cependant, les deux levés sont très élaborés, coûteux, et consomment beaucoup de temps

(Furnans & Austin, 2008), limitant ainsi leur fréquence (Merwade, 2009) surtout dans les pays à économie vulnérable. D'ailleurs la densité de points par hectare utilisée pour la réalisation des levés topo-bathymétriques est très variable de 42, 52 et 94 (points ha⁻¹) respectivement pour les trois barrages Hammam Debagh (ANB, 2005), Loch Raven et Prettyboy (Ortt et al., 2000), à 333, 375, 551 et 961 (points ha⁻¹) pour les quatre barrages Liberty (MGS, 2016), Foum El Gherza (ANBT, 2015), Triadelphia et Rocky Gorge (Ortt et al., 2007 ; MGS, 2016). La géostatistique a été utilisée pour prédire les élévations aux endroits non-mesurés et a permis de produire des modèles bathymétriques de réservoirs assez précis (Rakhmatullaev et al., 2011). Pour analyser et visualiser le fond des mers, Maleika et al. (2012) ont testé onze méthodes d'interpolation. Les meilleurs résultats ont été obtenus par l'interpolateur géostatistique de Krigeage. Cet interpolateur a permis de décrire avec précision la bathymétrie du lit d'un lac (Hera et al., 2014). D'ailleurs les outils géostatistiques ont été appliqués dans plusieurs domaines pour optimiser les schémas d'échantillonnage (Atkinson & Lloyd, 2007 ; Bechler et al., 2013 ; Wang et al., 2015). Hernandez & Emery (2009) ont utilisé l'interpolateur géostatistique pour réduire les coûts d'échantillonnage à plus de 50%. Toutefois l'optimisation de la taille des levés topo-bathymétriques a été peu abordée. Cependant, le contrôle de l'envasement des retenues présente un intérêt pratique évident pour l'exploitation des barrages. La surveillance des installations de prise d'eau et de dévasement permet de mettre à jour les courbes de remplissage des retenues qui peuvent être rendues obsolètes par des dépôts importants.

Par ailleurs la gestion des volumes d'eau retenus par les barrages constitue un souci permanent aux chercheurs qui se sont focalisé en grande partie sur l'amélioration des performances du système de mobilisation de la ressource en eau superficielle.

En effet, la prévision des volumes mensuels à allouer aux différents utilisateurs d'eau (domestique, agricole et industriel) est stratégique pour la planification et la gestion des ressources en eau ainsi que pour l'exploitation des barrages réservoirs et particulièrement pendant les épisodes de sécheresse.

Nombreuses sont les méthodes mathématiques qui ont été utilisées dans le domaine de la gestion des barrages réservoirs. Au cours des dernières décennies, les réseaux de neurones artificiels (RNA) ont été admis comme un outil alternatif pour la modélisation des variables qualitatives et quantitatives des ressources en eau (Tabari et al., 2010a ; b, 2011 ; Heddam et al., 2011 ; Marofi et al., 2011 ; Hossein zadeh Talaei et al., 2011). Quant au développement des règles d'exploitation des barrages, les réseaux de neurones artificiels (RNA) sont utilisés comme une alternative pour surmonter les difficultés liées à l'analyse par régression linéaire ou non linéaire. Ainsi, la plupart des problèmes traités avec les méthodes d'analyse par régression ont été résolus par les modèles de réseaux de neurones artificiels (RNA) qui ont prouvé qu'ils représentent une meilleure alternative (Rossi et al., 1999).

Les modèles de réseaux de neurones ont aussi été utilisés dans le domaine de la gestion des ressources en eau pour la prévision des débits d'écoulement des rivières (El-Shafie et al., 2007 ; Karunanithi et al., 1994), la prévision de chute de pluie (Ramirez et al., 2005), la prévision des crues (Chang et al., 2007) et la modélisation pluie-débit (Anctil et al., 2004 ; Kisi, 2004). Rezaeian -Zadeh et al. (2009) ont utilisé le modèle perceptron multicouches optimisé avec l'algorithme de Levenberg-Marquardt pour prédire les débits maximaux journaliers dans le bassin versant de Shirin Khosrow. Leurs résultats ont montré que la prise en compte des précipitations à l'échelle du bassin versant a augmenté l'efficacité de la prédiction. Dans une autre étude, (Rezaeian-Zadeh et al., 2010) ont utilisé la précipitation pondérée en fonction de la surface comme un vecteur d'entrée pour les réseaux de neurones.

Les réseaux de neurones artificiels ont également été utilisés pour le développement des règles d'exploitation d'un barrage réservoir (Cancelliere et al., 2002 ; Li & Huang, 2013) et la gestion de systèmes à plusieurs barrages réservoirs (Chandramouli & Raman, 2001).

Plus récemment, Kisi (2008, 2009) a proposé le modèle de neuro-ondelettes pour prédire les niveaux et les débits mensuels des lacs et a comparé les résultats à ceux des modèles de perceptron multicouches (MLP), de régression linéaire multiple et autorégressifs (MLR). Cette nouvelle technique pourrait faire la différence pour les

prévisions à court et à long terme. [Debbarma et Regulwar \(2011\)](#) ont utilisé les données historiques pour prédire les apports d'eau au niveau des barrages réservoirs par les différentes techniques de réseaux de neurones artificiels. Les résultats de la modélisation ont montré qu'une précision de prédiction raisonnable a été obtenue pour la plupart des modèles. Un modèle d'aide à la décision utilisant une combinaison de règles basée sur un système expert et les réseaux de neurones artificiels (RNA) a été développé pour l'élaboration des consignes de gestion d'un barrage réservoir dans le sud de l'Inde ([Chandramouli & Deka 2005](#)). Les auteurs ont conclu que ce modèle d'aide à la décision montre une performance plus grande que les approches basées sur la régression.

Ensemble, les réseaux de neurones artificiels (RNA) et la logique floue (connu sous le nom d'approche neuro-floue) ont été utilisés comme des outils potentiels dans le développement de modèles de contrôle intelligent ([Chang & Chang, 2001](#)). Le contrôle intelligent est une technologie qui ressemble au processus de la pensée humaine dans l'aspect décisionnel et la stratégie d'apprentissage. Les approches neuro-floues visent à combiner la capacité des ensembles flous qui servent à représenter la connaissance qui est compréhensible pour les êtres humains avec la capacité d'apprentissage des réseaux de neurones artificiels (RNA). La plus populaire méthode neuro-floue est le système de réseau adaptatif basé sur l'inférence floue (ANFIS) proposé par [Jang \(1993\)](#). L'ANFIS a été appliqué avec succès par de nombreux chercheurs à la gestion des barrages réservoirs en temps réel ([Chang & Chang, 2001](#) ; [Chang et al., 2005a, b](#)).

[Ismail et Kerem \(2012\)](#) ont utilisés deux techniques de réseau de neurones, les fonctions à bases radiales (RBF) et le Feed Forward Back Propagation (FFBP) pour prédire les trois variables aléatoires essentielles à la gestion des réservoirs, à savoir : L'apport mensuel, l'évaporation mensuelle et le volume mensuel de stockage. Les résultats de la modélisation ont montré que les deux méthodes fournissent des précisions d'estimation très satisfaisantes et que les types et le nombre des variables d'entrées ont une incidence très sensible sur la prévision résultante.

Ces expériences récentes ont montré que les réseaux de neurones artificiels pourraient constituer une alternative prometteuse pour prédire les débits de ruissèlements ([Hsu et al., 1995](#) ; [Shamseldin, 1997](#) ; [Tokar & Johnson, 1999](#) ; [Tokar & Markus, 2000](#)), générer

les apports et cartographier les règles de gestion des barrages réservoirs (Jain et al., 1999). Toutefois il convient de souligner que, comme dans tout autre modèle statistique et mathématique, les modèles RNA ont quelques inconvénients. Avoir un grand nombre de variables d'entrée est l'un des soucis les plus courants pour leur développement parce qu'ils ne sont pas conçus pour éliminer les entrées superflues. En outre, dans le cas d'un grand nombre de variables d'entrée, les variables pertinentes, redondantes et bruyantes pourraient être incluses dans l'ensemble des données, par conséquent, les variables significatives pourraient être cachées (Seasholtz & Kowalski, 1993). L'identification d'un vecteur d'entrée est donc un élément crucial pour les RNA (Rezaeian-Zadeh et al., 2013). Il y a différentes techniques pour réduire le nombre de variables d'entrée, telles que l'analyse en composantes principales (ACP) (Zhang et al., 2006 ; Zhang, 2007 ; Noori et al., 2009) et le Gamma test (GT) (Corcoran et al., 2003 ; Moghaddamnia et al., 2009). Sudheer et al. (2002) ont proposé une méthode statistique fondée sur la corrélation croisée, l'auto-corrélation et les propriétés d'auto-corrélation partielle. En outre, Kumar et al. (2005) ont déterminé des combinaisons d'entrées à partir des décalages des précipitations et du ruissellement qui ont une influence significative sur le débit prévu. Cependant il semble que la prévision du volume mensuel alloué en testant plusieurs vecteurs d'entrée et en utilisant différents modèles de RNA (MLPNN, GRNN et RBFNN) a été très peu investiguée, ainsi les études dans ce sens s'avèrent limitées. Ce qui nous a incité à inscrire ce travail dans cette optique.

Chapitre 2

Chapitre 2 : Présentation de l'aire d'étude

2.1 Introduction

Ce chapitre est dédié à la présentation de l'aire d'étude qui est le Barrage Hammam Debagh. Le bassin versant de l'Oued Bouhamdane qui est un sous bassin du bassin hydrographique la Seybouse sera également décrit. La localisation, la délimitation, les caractéristiques du bassin versant, les caractéristiques techniques du barrage ainsi que les utilisateurs de l'eau seront aussi détaillées dans ce chapitre.

2.2 Localisation du site

Le barrage Hammam Debagh est situé à 20 km à l'Ouest de la ville de Guelma (Algérie). Il est implanté à trois (03) Km en amont de la localité Hammam Debagh, une agglomération qui s'est développée autour des thermes de hammam Meskhoutine, dans la wilaya de Guelma. Le réservoir du barrage retient l'un des plus importants affluents de l'oued Seybouse, l'oued Bouhamdane qui s'écoule du Sud-Ouest vers le Nord-Est, le long du versant Sud-tellien, entre les massifs de djebel Taya, au Nord et du djebel Seba Mzaar, au Sud (Fig. 2-1).

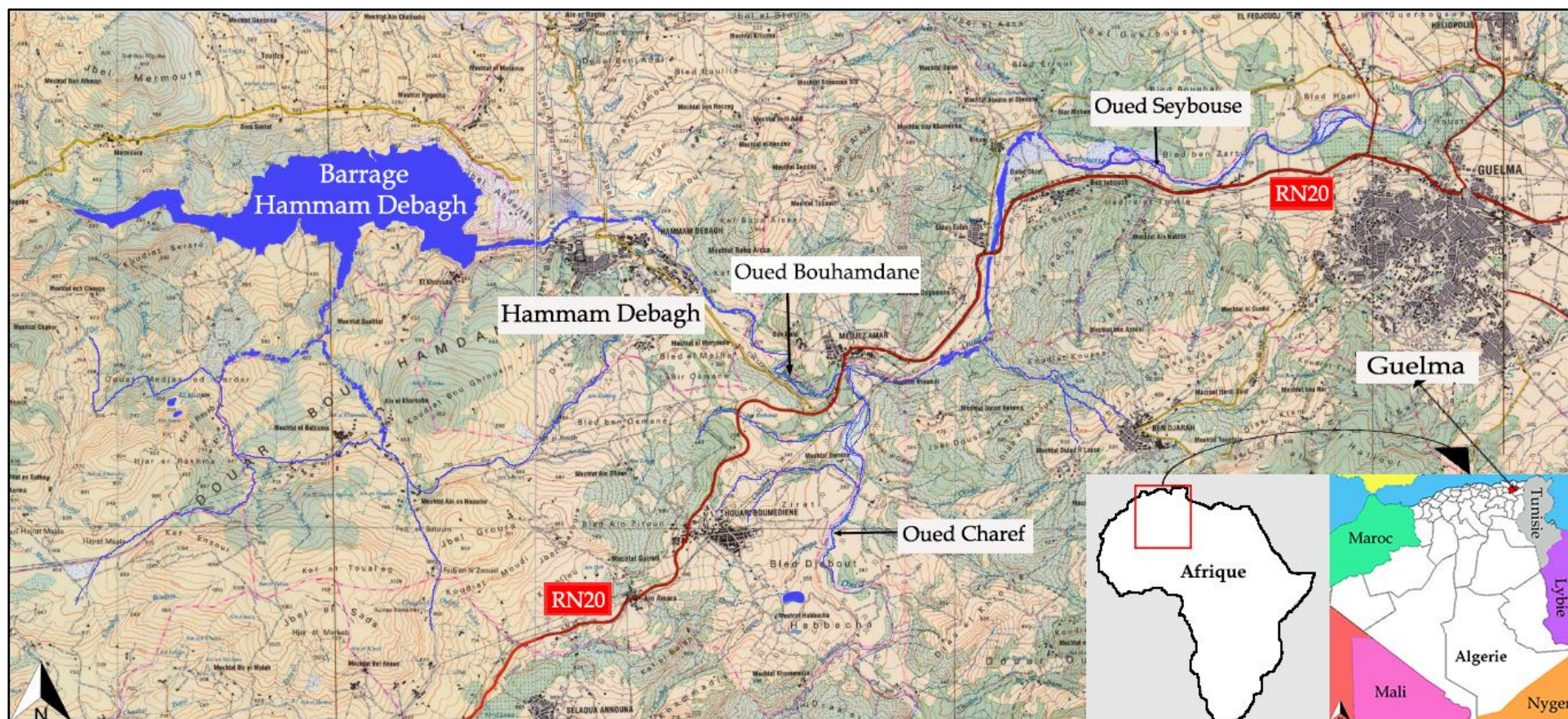


Figure 2-1 : Localisation du site d'étude

2.3 Climat

Le climat est l'une des composantes les plus importantes des bassins versants qui constitue le facteur le plus influent sur le phénomène d'écoulement à la fois annuel et extrême (crues et étiages). Le bassin versant de la Seybouse appartient au domaine des étages bioclimatiques subhumide au Nord et semi-aride dans sa partie septentrionale intégrant aussi bien la vallée de Guelma que la zone des hautes plaines.

La zone d'étude est située dans les altitudes subhumides du climat continental dont les principales caractéristiques météorologiques sont présentées au [Tableau 2-1](#). Les précipitations moyennes annuelles (21 ans) sont d'environ 605 mm. Plus de 81 % des précipitations annuelles sont observées entre novembre et mai. Sur la période juin-octobre, les précipitations observées sont faibles.

Tableau 2-1 : Caractéristiques météorologiques principales (1990-2011)

Caractéristiques	Température (°c)	Précipitation (mm)
Janvier	14,0	87,4
Février	15,1	64,7
Mars	18,2	59,9
Avril	20,8	60,1
Mai	25,9	53,0
Juin	31,1	19,7
Juillet	35,9	4,6
Août	35,4	10,1
Septembre	30,0	39,6
Octobre	25,5	38,2
Novembre	18,9	72,3
Décembre	14,7	94,9
Année	23,8	604,5

Source : (ANBT, 2011)

2.4 Bassin versant Bouhamdane

Le bassin versant de l'oued Bouhamdane (N°03) fait partie du grand bassin hydrographique la Seybouse (N° 14) dans le Nord-Est Algérien. Il occupe la partie Ouest de la wilaya de Guelma et est limité par :

1. Au Nord : le bassin hydrographique Côtiers Constantinois (N°03) ;

2. Au Sud et Sud - Est : le sous bassin versant N° 02 de l'oued Cherf du bassin hydrographique la Seybouse (N°14) ;
3. A l'Est : le bassin versant de Guelma (N°04) faisant partie du bassin hydrographique la Seybouse (N°14) ;
4. A l'Ouest et Sud - Ouest : le bassin hydrographique Kebir Rhumel (N°10).

Il présente un relief de moyenne montagne dont les altitudes sont comprises entre 270 m, à la station de Medjez Ammar 2 au niveau de la vallée et 1281 m au sommet du djebel Mzarat Sidi Chagref, avec une altitude moyenne de 785 m et un indice de pente de 0,13. Les densités de drainage permanente et temporaire sont respectivement de 0,53 km⁻¹ et 1,90 km⁻¹. La longueur du talweg principal est de 99 km avec une pente moyenne de 0,8% (ANRH, 2008).

Les oueds Zenati et Sabath, qui drainent respectivement une zone semi-aride des Hautes Plaines (pluie moyenne annuelle : 450-550 mm) et une zone sub-humide du Tell méridional (pluie moyenne annuelle 550-650 mm), forment par leur confluence l'oued Bouhamdane qui se jette dans la Seybouse à Medjaz Amar (Fig. 2-2).

Ce sous bassin versant a une lithologie très variée, elle est composée de formations superficielles (32,4%), grès, conglomérats, marno-calcaires et marnes schisteuses (40,5%), argiles et marnes (23,1%) et calcaires (4,0%). Il est relativement protégé par un couvert végétal naturel (chêne liège, maquis, reboisement) couvrant 31,7% de la surface totale dont 19,5% sont exploités comme terrains de parcours. La prédominance des surfaces occupées par les cultures céréalières, maraichères 58,1% de la surface du bassin et par l'olivier autour de 1% de la surface représente la part majoritaire du couvert végétal et témoigne de la forte activité agricole dans la région. (Louamri et al., 2013).

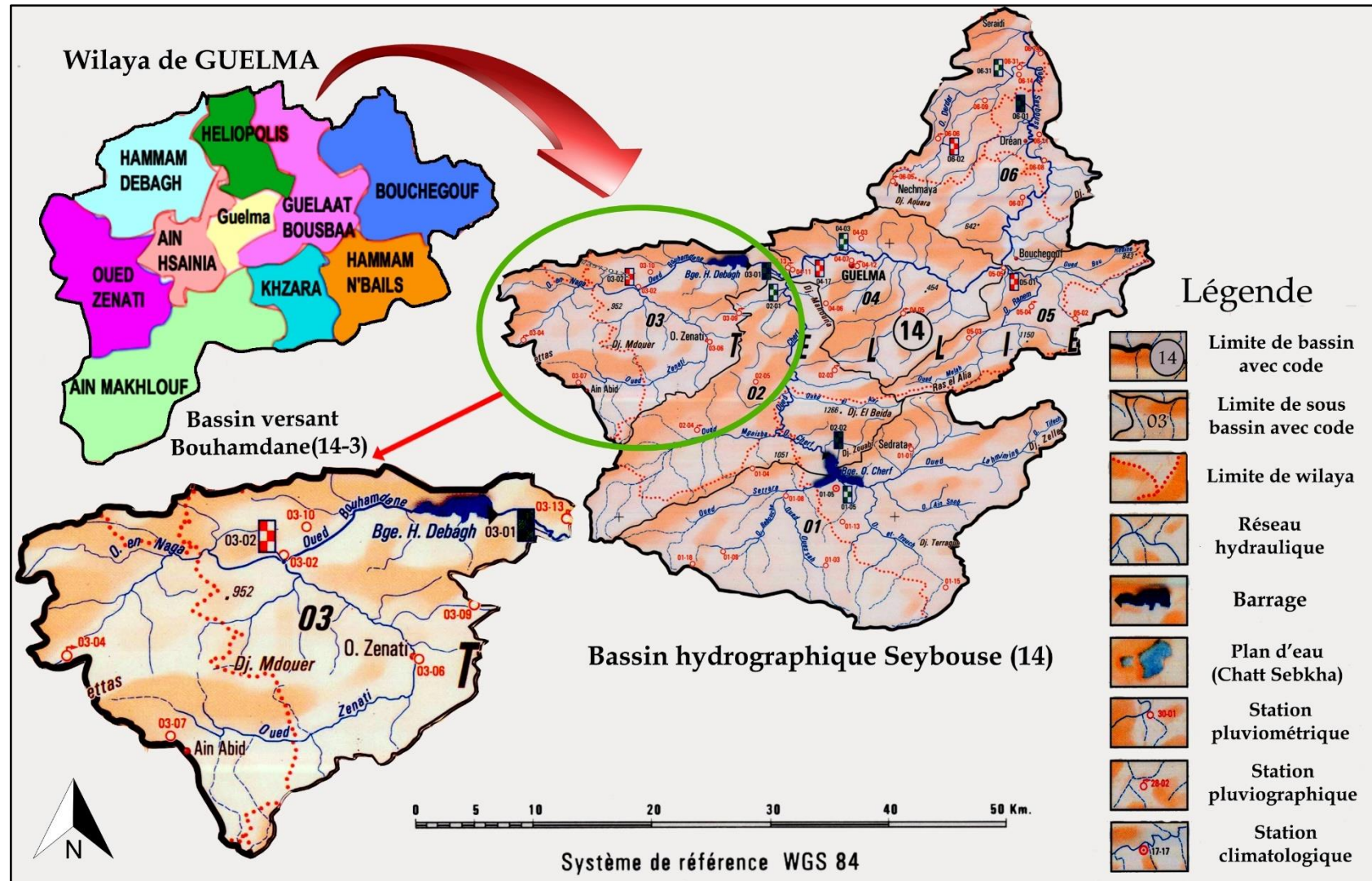


Figure 2-2 : Localisation et délimitation du bassin versant de l'oued Bouhamdane (Source : Carte du réseau hydro climatologique et de surveillance de la qualité des eaux du Nord d'Algérie au 1 :500 000 ème, 2005)

Les caractéristiques du bassin versant sont résumées dans le [Tableau 2-2](#).

Tableau 2-2 : Caractéristiques du bassin versant de l'oued Bouhamdane.

Caractéristiques du bassin versant de l'oued Bouhamdane		
Exutoire	Station hydrométrique Medjaz Amar 2 ⁽¹⁾	Barrage ⁽²⁾
Surface du bassin versant	1 105 km ²	1 070 km ²
Périmètre du bassin versant	140 Km	142 Km
Indice de compacité de Gravelius (KG)	1,18	/
Temps de concentration	15,5 heures	/
Apport moyen annuel	69 hm ³	65 hm ³
Volume régularisé	/	55 hm ³
Envasement moyen annuel calculé En fonction du levé bathymétrique de 2004	/	0,98 hm ³ /an

⁽¹⁾ (Mebarki, 2005); ⁽²⁾ (ANBT, 2010)

2.5 Barrage Hammam Debagh

Situé à 20 kilomètres à l'Ouest de la ville de Guelma, le barrage Hammam Debagh a été construit en 1980 sur l'oued Bouhamdane. Il est du type en terre zonée avec noyau central vertical formé de terre argileuse. ([Fig. 2-3](#)).

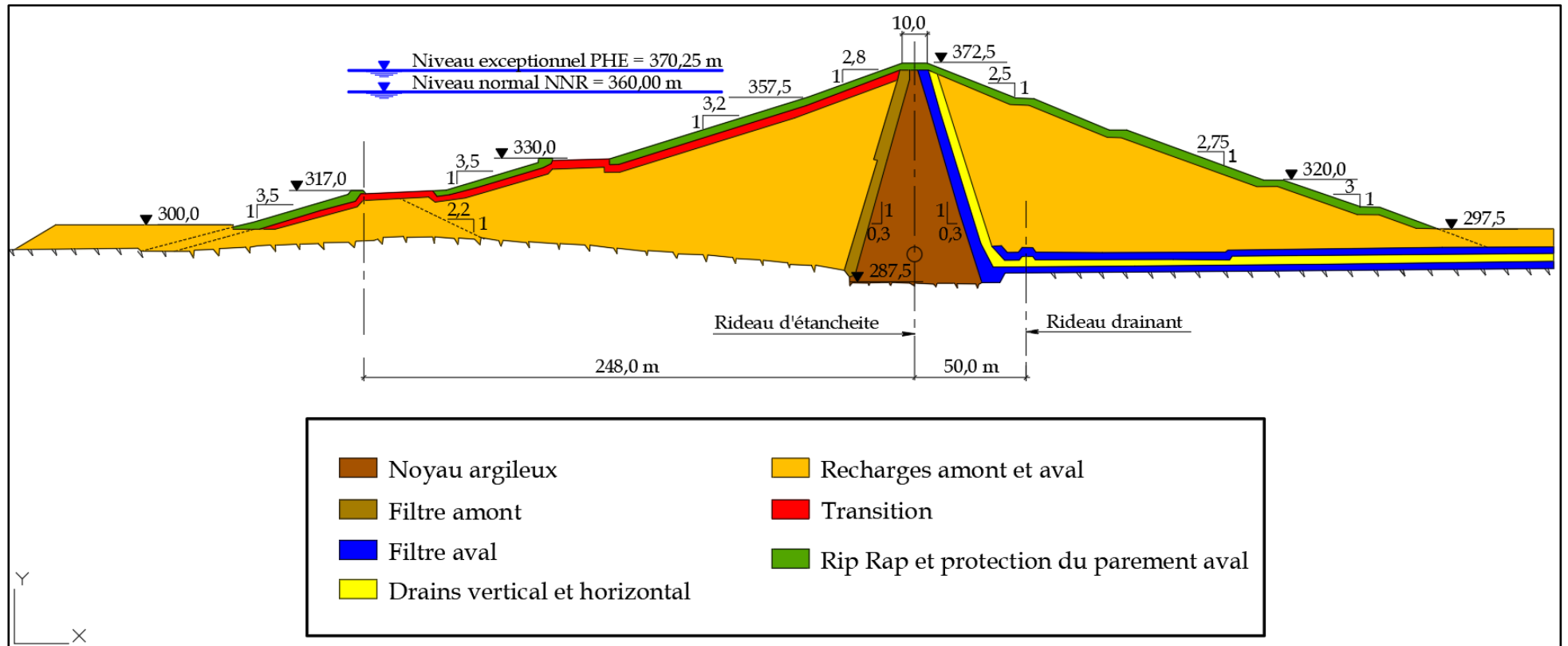


Figure 2-3 : Coupe en travers du barrage Hammam Debagh (source : ANBT (ex ANB), 1985)

Le barrage Hammam Debagh a été mis en service en décembre 1987, mais en raison des retards occasionnés dans les travaux d'équipement du périmètre situé à l'aval, l'eau mobilisée (capacité : 200 hm³) n'a pas trouvé d'utilisateur plusieurs années successives (Mebarki, 2005).

Les deux évacuateurs de crues sont du type Corolle et Seuil libre de section circulaire, en forme de tulipe (Fig. 2-4).



Figure 2-4 : Le barrage Hammam Debagh : évacuateur de crues de type Corolle et Seuil libre (source : ANBT, 2010)

Les caractéristiques du barrage sont résumées dans le [Tableau 2-3](#).

Tableau 2-3 : Caractéristiques du barrage Hammam Debagh.

Caractéristiques du barrage Hammam Debagh	
Type	Digue en terre zonée
Hauteur	95 m
Longueur en crête	430 m
Niveau Normal de la Retenue (NNR) :	360 m
Niveau des Plus Hautes Eaux (PHE)	370,25 m
Niveau des Plus Hautes Eaux Exceptionnelles (PHEE)	371,28 m
Déversoirs	Corolle et Seuil libre
Débit d'évacuation	2240 m ³ /s
Vidange de fond	218 m ³ /s
Capacité initiale	200,00 hm ³
Capacité au dernier levé topo-bathymétrique (2004)	184,35 hm ³
Envasement annuel	0,98 hm ³ /an

Source : (ANBT, 2010)

2.6 Affectation des eaux du barrage Hammam Debagh

Le but de la retenue est essentiellement de régulariser les apports en vue de satisfaire les besoins pour l'irrigation du périmètre de Guelma-Boucheouf et le renforcement de l'alimentation en eau potable de quelques villes de la wilaya de Guelma.

2.6.1 Alimentation en eau potable

Le barrage de Hammam Debagh constitue la principale ressource en eau de la wilaya de Guelma. Il alimente essentiellement le couloir Est de la vallée de Guelma constitué par les agglomérations : Hammam Debagh (18,29 l/s), Medjez Ammar (4,14 l/s), Ben Djarah (5,04 l/s), la ville de Guelma (300 l/s) et la localité de Roknia (9,77 l/s) située dans le versant Nord de la vallée de Guelma.

2.6.2 Alimentation en eau d'irrigation

Le périmètre s'étend sur 80 Km environ depuis la confluence des oueds Bouhamdane et Cherf, donnant naissance à l'oued Seybouse jusqu'au nord de la ville de Dreaan. Situé de part et d'autre de l'oued Seybouse, il dispose d'une superficie équipée de 9 940 ha et d'une superficie irrigable de 9 250 ha répartie en cinq (05) secteurs indépendants allant de 605 ha pour le plus petit secteur à 3 500 ha pour le plus grand (Tab. 2-4).

Tableau 2-4 : Superficies des secteurs et leur besoin en eau

N°	Secteur	Superficie (ha)		Besoin en Eau (hm ³)	Situation
		Equipée	Irrigable		
1	Guelma	3 500	3 255	19,40	Plaine Guelma
2	Cherf	605	565	3,40	
3	Boumahra	2 600	2 420	14,40	
4	El-Fedjouj	2 355	2 190	13,00	
5	Bouchegouf	880	820	4,90	Bouchegouf
Total		9 940	9 250	55,10	

Source : (ONID (ex AGID) Guelma, 2011)

Pour valoriser les terres des plaines de Guelma et de Bouchegouf le système d'alimentation du périmètre a été conçu de manière à utiliser au maximum la ressource en eau dont dispose la wilaya de Guelma du fait que celle-ci est sillonnée par un réseau hydrographique très important. Ainsi l'approvisionnement en eau des secteurs se fait par des lâchers d'eau à partir du barrage Hammam Debagh dans l'oued Bouhamdane et sa reprise est opérée dans chaque secteur grâce à un seuil, une station d'exhaure, un dessableur, une station de pompage et un réservoir de compensation (Fig. 2-5). Le réseau de distribution est du type ramifié et les techniques d'irrigation utilisées au sein des secteurs sont l'aspersion pour les cultures maraichères et le goutte à goutte pour l'arboriculture fruitière.

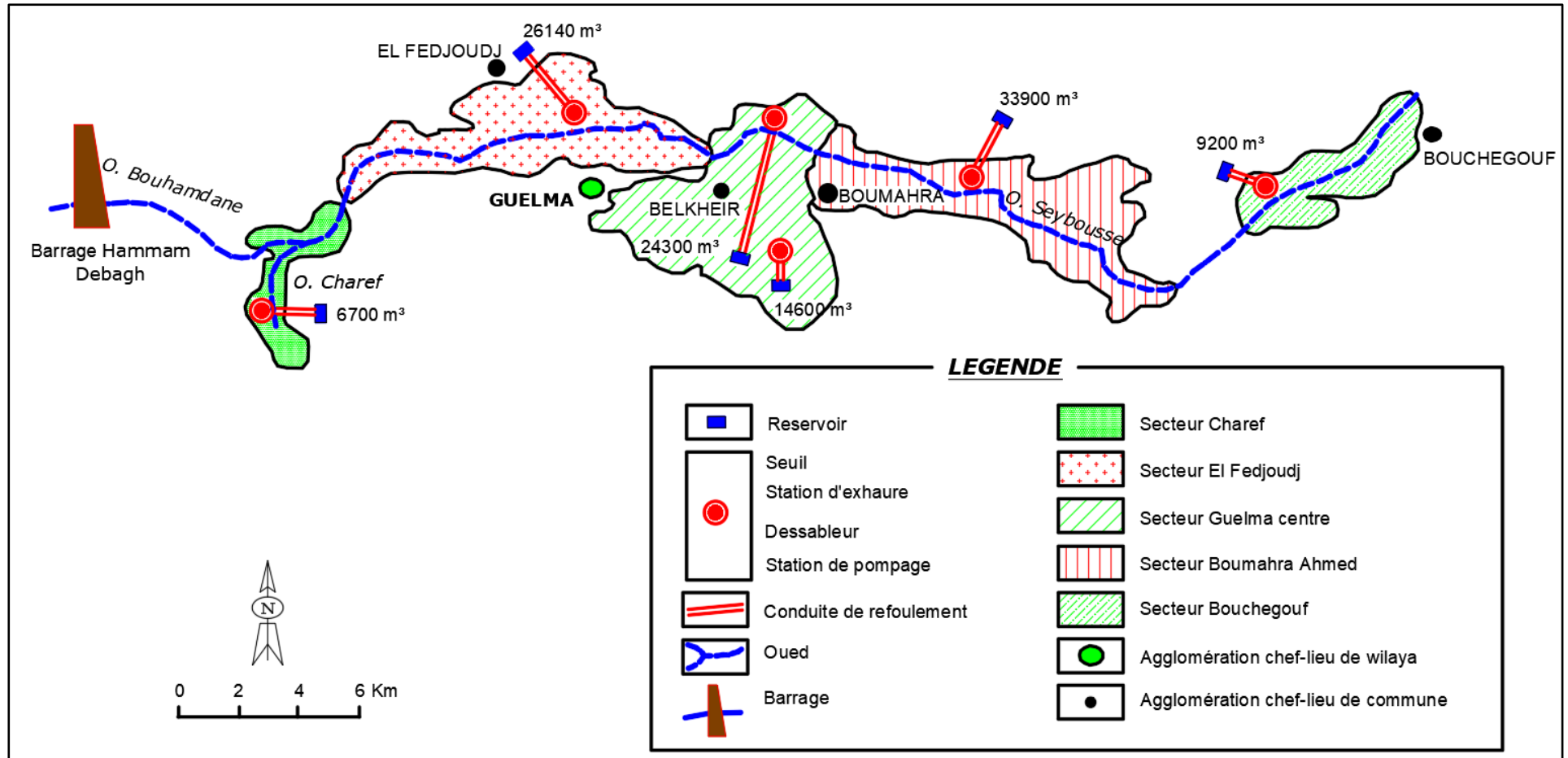


Figure 2-5 : Schéma d'aménagement et organisation du périmètre Guelma- Bouchegouf Source : (ONID (ex AGID) Guelma, 2011)

2.7 Conclusion

Le présent chapitre a en premier lieu décrit sommairement le bassin versant de l'oued Bouhamdane qui est un affluent important de la Seybouse. C'est un bassin ramassé, caractérisé essentiellement par une lithologie très variée et est relativement protégé par un couvert végétal naturel.

En deuxième lieu le Barrage Hammam Debagh a été décrit historiquement et techniquement. C'est un barrage en terre qui mobilise plus de 200 millions de m³ d'eau. Il est muni de deux évacuateurs en tulipe fonctionnant en charge. La retenue reçoit les eaux de l'oued Bouhamdane drainée par une superficie de 1070 km².

Enfin les différentes affectations des eaux mobilisées ont été identifiées. Il s'agit de l'alimentation en eau potable de quelques agglomérations de la vallée de Guelma et le périmètre d'irrigation Guelma-Bouchevouf.

Chapitre 3

Chapitre 3 : Collecte, critique et traitement des données

3.1 Introduction

Ce chapitre s'intéresse à l'acquisition, le contrôle de la fiabilité et le traitement des données utilisées pour l'optimisation de la densité des points des levés topo-bathymétriques et pour l'élaboration du modèle de prévision du volume total affecté à l'alimentation en eau potable et agricole.

Les données fournies par l'Agence Nationale des Barrages et Transferts (ANBT) constituent la source principale pour l'élaboration des travaux de cette thèse. Elles sont composées des données de l'unique levé topo-bathymétrique réalisé en 2004 et des données d'exploitation du barrage sur une période de 21 ans (1990-2011).

3.2 Présentation des bases de données

3.2.1 Acquisition des données topo-bathymétriques

Ces données sont celles acquises par le consortium de bureaux d'études LEM-GEOID lors de la campagne de levé bathymétrique du lac complétée par une topographie terrestre des berges exondées qui a été réalisée en 2004 pour actualiser les données relatives à la capacité de stockage de la retenue d'eau et pour surveiller la dynamique de l'envasement du barrage.

3.2.1.1 Système géodésique utilisé

Le levé de la retenue a été réalisé avec les paramètres géodésiques suivants ([JORADP N° 30 du 30 Avril 2003](#)) :

Système géodésique : WGS-84/ITRS ;

Ellipsoïde : GRS-80 ;

Demi grand axe : $a = 6\,378\,137$ m ;

Aplatissement : $f = 1 / 298,25722$.

Projection : UTM zone 32 Nord ;

Latitude origine : 0° N ;

Longitude Origine : 9° E ;

Facteur d'échelle à l'origine : 0,9996 ;

Constante Nord : 0 m ;

Constante Est : 500 000 m.

Systeme altimétrique : Le système altimétrique existant du barrage NGA (Nivellement Général d'Algérie pour la retenue de Hammam Debagh).

3.2.1.2 Réseau d'appui

Les coordonnées planimétriques des bornes du réseau d'appui (HD01, HD02, HD03 et HD04) et du pilier d'auscultation (HD00) sont données en projection UTM zone 32 Nord dans le référentiel géodésique WGS-84/ITRS et les élévations dans le système NGA du barrage (Tab. 3-1).

Tableau 3-1 : Coordonnées des bornes du réseau d'appui.

Nom de point	X UTM 32 N (m)	Y UTM 32 N (m)	Élévation (mNGA)
HD00	342111,26	4036664,89	374,43
HD01	341708,74	4036320,38	410,75
HD02	340348,26	4036462,95	475,68
HD03	336021,74	4038565,47	624,24
HD04	334271,53	4037987,74	405,53

Source : (ANB, 2005)

Les cinq bornes (HD00, HD01, HD02, HD03, HD04) ont été utilisées pour le positionnement de la station de référence RTK durant la réalisation des levés topographiques et bathymétriques (Fig. 3-1).

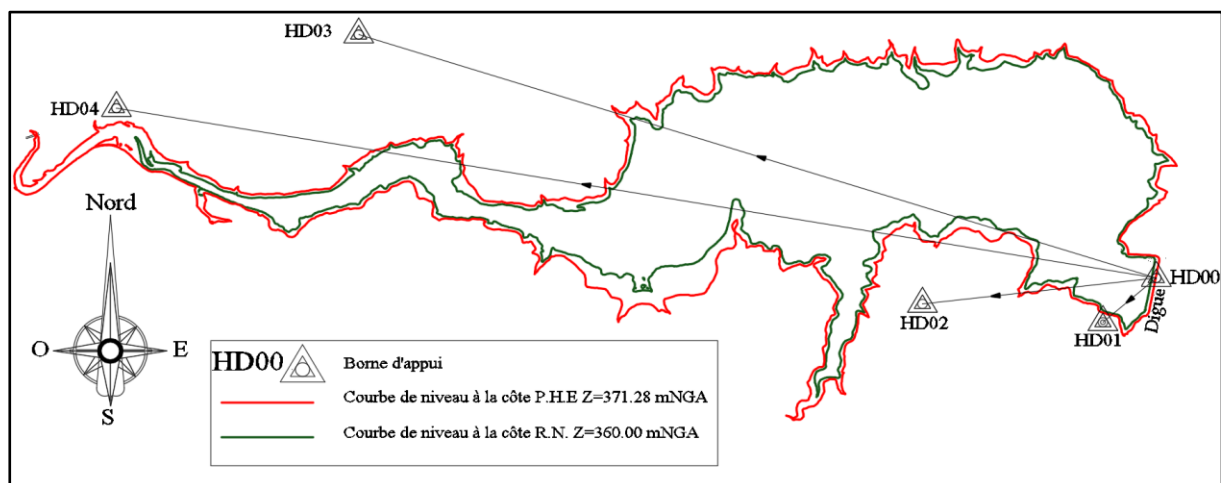


Figure 3-1 : Schéma du réseau des bornes d'appui (Source : ANB, 2005)

3.2.1.3 Levé topographique

Le levé topographique des parties émergées de la retenue a été réalisé par trois (03) récepteurs GPS Trimble 4 700, 5 700 et MS 750 en mode RTK-OTF avec contrôleurs graphiques TSC1, trois (03) radio modem Satel et antennes de transmission UHF associées et interfacées avec une station de référence composée d'un récepteur GPS Trimble 4 000 ssi avec sortie RTK, d'une radio modem Satel et d'une antenne UHF, commune avec le levé bathymétrique. Deux stratégies de levé ont été adoptées en fonction de la géomorphologie du terrain : le suivi de profils parallèles pour les zones planes ou légèrement ondulées et le levé des lignes de rupture des éléments caractéristiques de la topographie, comme les thalwegs, les oueds, les pieds et les hauts d'escarpement, la crête du barrage etc. Ainsi, sur un versant coupé de thalwegs et d'oueds, le levé suit des profils d'altitude constante. Par contre, sur un versant en « paliers » coupé par des escarpements, il est conduit en profils orientés suivant la ligne de plus grande pente. Le levé combiné de profils espacés tous les 50 m dans la zone des 1 000 m du corps du barrage puis tous les 100 m au-delà de cette limite et de lignes de ruptures caractéristiques (thalwegs, lits des oueds, hauts et bas de talus, d'escarpement, etc) dans les zones accidentées a ainsi permis une restitution optimale de la forme du terrain. Le contrôle de la qualité des données acquises a été effectué au moyen de procédures de tri automatisées mises en œuvre via le SIG (Système d'Information Géographique) en utilisant le logiciel TGO (Trimble Geomatics Office v.1.6) et la base de données GPS sous Microsoft Access, ce qui a permis de fournir une topographie avec une exactitude de ± 5 cm. Ce levé topographique, représentant une surface totale de 166,1 ha, a été réalisé en 21 jours : du 21 Février au 12 Mars 2004 ([ANB, 2005](#)).

3.2.1.4 Levé bathymétrique

Le levé bathymétrique a été réalisé à l'aide d'un appareil de contrôle de profondeur sonore Echosondeur Odom Hydrotrac, interfacé avec un récepteur GPS Trimble MS 750 en mode RTK OTF (Real Time Kinematic, résolution des ambiguïtés entières de phase On The Fly), capable de déterminer les emplacements de sondage dans le réservoir. Les profondeurs du réservoir et les coordonnées horizontales sont

enregistrées en continu dans un ordinateur portable Compaq Presario 2 100 équipé du logiciel d'acquisition et de post-traitement bathymétrique Hypack Max (Coastal Oceanographics) pour être traitées par la suite. Quelles que soient les variations de hauteur du plan d'eau dues aux conditions météorologiques ou à l'exploitation du barrage (vent, lâchers d'eau, etc.), la longueur du mât qui sépare le transducteur de l'échosondeur de l'antenne GPS ne change pas. Par conséquent, les échos du fond mesurés par le transducteur sont directement référencés à l'altitude de l'antenne GPS qui a été déterminée par rapport à la hauteur ellipsoïdale de la station de référence située à terre sur un point du réseau d'appui parfaitement connu. La calibration de l'échosondeur était régulièrement effectuée à l'aide d'une plaque métallique ad hoc immergée à la profondeur maximale. La célérité du son dans l'eau a ainsi varié entre 1 434 m/s et 1 447 m/s. La validation de l'ensemble du système du levé bathymétrique a été effectuée lors de la mise en place en affichant en temps réel sur l'écran de l'ordinateur la position et la profondeur de la plaque de calibration immergée au pied de l'embarcadère à une profondeur mesurée physiquement à l'aide d'une tige rigide graduée. Les profils bathymétriques levés avec un intervalle de 50 m dans la zone des 1 000 m du corps du barrage et un intervalle de 100 m au-delà de cette limite ont permis de reconstituer le toit des sédiments décantés dans le lac avec une précision de ± 5 cm. Ce levé bathymétrique, représentant une surface totale de 612,9 ha, a été réalisé en 13 jours du 23 Février au 06 Mars 2004 (ANB, 2005).

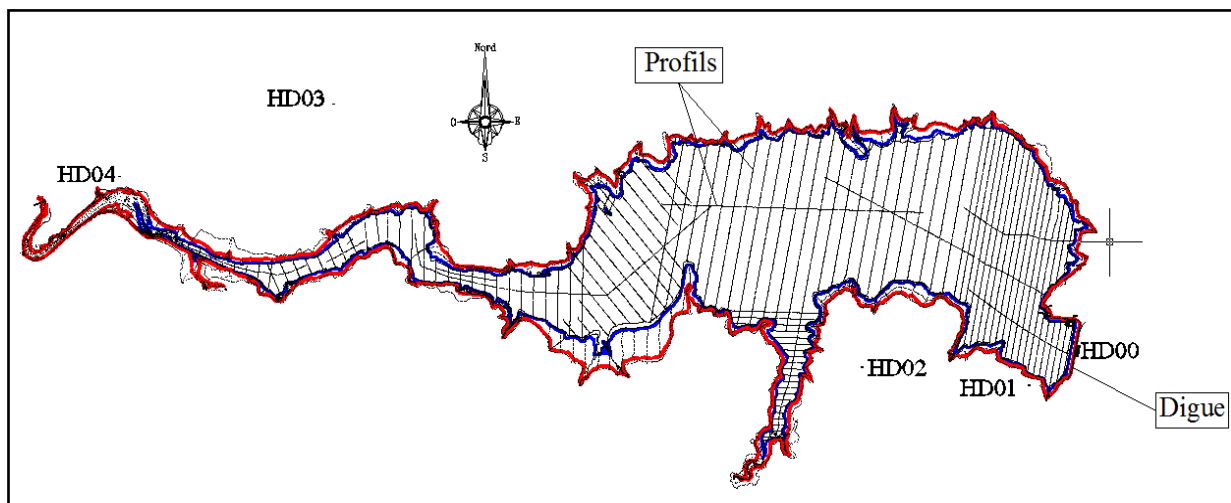


Figure 3-2 : Profils topo-bathymétriques (source : ANB, 2005)

3.2.1.5 Base de données recueillie

Les données acquises lors de la campagne de mesure topo-bathymétrique de 2004 sont 32 164 points avec trois attributs (X, Y en UTM zone 32 Nord et les élévations Z dans le système NGA) couvrant une superficie globale levée de 779 ha, répartie entre environ 21 % en levé terrestre et 79 % en levé bathymétrique.

La différence des volumes d'eau stockés à la côte normale entre l'année 1988 et la mesure bathymétrique en 2004 représente la part des sédiments érodés sur le bassin versant et transportés par les eaux de ruissellement jusqu'au barrage depuis sa mise en service. Donc la capacité du barrage, depuis sa mise en eau, a perdu à la date de mars 2004 un volume estimé à 15,65 Hm³ soit 7,83 % de sa capacité initiale (200 Hm³). Cela correspond à une perte moyenne par an de l'ordre de 978 100 m³.

3.2.2 Acquisition des données d'exploitation

Dans le but d'établir le bilan hydraulique du réservoir et de mettre en place un outil de suivi de la disponibilité en eau, une série de mesures est réalisée par l'exploitant du barrage au pas de temps journalier (tous les matins à 8 heures). Ces données sont portées régulièrement sur une fiche préétablie intitulée « *ETAT DE LA RESERVE* », puis saisies sous format numérique (Excel) pour être exploiter par la suite.

3.2.2.1 Paramètres mesurés

Les variables accessibles à l'observation sont :

1. La côte du plan d'eau : mesurée par un levé direct en utilisant un appareil topographique installé sur un point, qui sert de référence, situé sur la dalle en béton de la tour de prise d'eau à une altitude de 372,50 m vu que le barrage ne possède pas d'échelle hydrographique graduée ;
2. L'évaporation : déterminée de façon manuelle par l'observateur à partir d'un bac d'évaporation de 0,8464 m² d'ouverture ;
3. Les fuites de l'ouvrage récupérables : mesurées sur le canal principal de collecte des fuites moyennant un déversoir triangulaire ;

4. Les lâchers pour l'irrigation et la vidange : déterminées à partir de la loi d'ouverture de la vanne de vidange de fond ;
5. Les lâchers pour l'Alimentation en Eau Potable : déterminées à partir de la loi d'ouverture de la vanne de prise ;
6. Les volumes d'eau déversés : calculés en fonction de la courbe $Q_{\text{déversé}} = f(h)$.

3.2.2.2 Paramètres estimés

Les variables estimées sont :

- a. Le volume stocké dans le réservoir et la surface du plan d'eau : déduits des courbes « Hauteur - Capacité - Surface » ;
- b. L'évaporation au niveau du barrage : estimée de la manière suivante :

$$Evap\ lac\ (hm)^3 = \left[\frac{Volume\ Ajouté\ (cm^3)}{Surface\ bac\ (cm^2)} * \frac{1}{100} \right] * Surface\ d'\ eau(km^2) * 0,78 \quad (3-1)$$

La surface du bac est prise égale à 8464 cm².

- c. L'apport entrant dans le réservoir qui découle du principe de la conservation des volumes d'eau exprimés en hm³ (Albergel et al,1997) soit :

$$APP = (V_{final} - V_{initial}) + (AEP + IRR + EVAP + VID + FUI + EVAC) \quad (3-2)$$

Avec :

APP : volume des apports globaux entrant dans le barrage ;

V_{final} : volume stocké au temps (t+1)

V_{initial} : volume stocké au temps (t)

AEP : volume alloué à l'alimentation en eau potable ;

IRR : volume alloué à l'irrigation ;

EVAP : volume évaporé ;

VID : volume de vidange ;

FUI : volume des fuites ;

EVAC : volume déversé.

Signalons que les apports globaux entrant dans le barrage regroupent les volumes ruisselés, les apports souterrains et les apports pluviométriques sur la retenue.

3.2.2.3 Données disponibles récupérées

Nous disposons des données d'exploitation (côte du plan d'eau, volume stocké, apport, lâchers, pluviométrie, évaporation au niveau de la retenue du barrage, fuites de l'ouvrage, vidange et déversement) mesurées et estimées journalièrement sur 21 ans par l'Agence Nationale des Barrages et Transferts (ANBT) depuis mars 1990.

3.3 Critique et traitement des bases de données

Le seul outil dont dispose le gestionnaire pour étudier et comprendre le comportement de la retenue est l'ensemble des données observées et estimées. La qualité de ces données dépend de plusieurs facteurs d'ordre métrologique, de traitement et/ou méthodologique.

Les volumes stockés dans la retenue sont déduits de la mesure des côtes des plans d'eau en utilisant la courbe hauteur - capacité. Pour ces données, les principales sources d'erreurs sont liées à l'imprécision de la mesure des côtes des plans d'eau (par exemple, l'influence du vent à la surface des retenues). Pour une retenue de taille importante, une erreur de quelques centimètres dans la mesure de la côte peut impacter fortement l'estimation du volume stocké. La mesure du volume dépend également de la mise à jour de la courbe hauteur-capacité.

La mesure de la pluie est délicate, puisqu'elle peut être perturbée par l'influence du vent et de l'implantation du pluviomètre qui dans certains cas affecte fortement la mesure.

Les apports ne sont pas mesurés directement, mais ils sont déduits du bilan hydraulique. Ainsi, le processus d'estimation des apports est ouvert à diverses sources d'incertitudes. On peut citer par exemple des problèmes liés à la précision de la mesure des lâchers, des vidanges et des déversements.

Il est donc impératif d'effectuer un traitement et une analyse des bases de données recueillies afin de tirer le maximum d'informations et de garantir leur fiabilité.

La démarche à suivre consiste, à détecter les éventuelles anomalies des paramètres mesurés et estimés, à combler les lacunes des données recueillis et puis à réajuster les paramètres déduits des courbes (Hauteur – Capacité - surface) en tenant compte des résultats des mesures topo-bathymétriques.

3.3.1 Base de données topo-bathymétriques

Les données topo-bathymétriques recueillies, par leur mode d'acquisition et de contrôle de qualité (§ Chapitre 3.2.1), ne sont pas entachées d'erreurs. De ce fait, aucun traitement spécifique n'a été réalisé.

3.3.2 Base de données d'exploitation

La base de données qui a été mise à notre disposition est entachée par de nombreuses erreurs et anomalies : erreurs systématiques ou erreurs accidentelles (exemples : les mêmes côtes du plan d'eau correspondent à des volumes différents ou le contraire (Tab. 3-2), omission des volumes de vidange et/ou de déversement etc...).

Les deux problèmes majeurs de cette base de données sont :

1. L'incohérence des données journalières d'exploitation. En effet en utilisant les paramètres disponibles, nous n'arrivons pas à établir la balance entre les entrées, les sorties et les volumes restants dans la cuvette (Tab. 3-3).
2. Les volumes de la retenue n'ont pas été mis à jour après le levé topo - bathymétrique de 2004.

Tableau 3-2 : Exemples d'erreurs de volumes.

Date	Côte (m)	VOL (hm ³)	APP (hm ³)	AEP (hm ³)	IRR (hm ³)	EVAP (hm ³)	PLU (mm)	VID (hm ³)	FUI (hm ³)
03/01/1999	347,58	15,87	0,063	0,004	/	0,005	0,4	/	0,002
12/02/1999	353,8	160,044	1,184	0,004	/	0,001	6,0	/	0,002
01/09/1999	353,8	160,046	/	0,004	0,155	0,030	/	/	0,002
13/04/1999	357,79	184,587	/	0,004	/	0,015	/	0,069	0,002
30/04/1999	357,79	184,842	/	0,004	/	0,021	/	0,069	0,002

Avec :

Volume : le volume stocké dans le barrage ;

APP : volume des apports globaux entrant dans le barrage ;

AEP : volume alloué à l'alimentation en eau potable ;

IRR : volume alloué à l'irrigation ;

EVAP : volume évaporé ;

PLU : pluie mesurée au niveau du barrage ;

VID : volume de vidange ;

FUI : volume des fuites.

Tableau 3-3 : Exemples sur l'établissement de la balance des volumes

Date	Côte (m)	Volume (hm ³)	APP (hm ³)	AEP (hm ³)	IRR (hm ³)	EVAP (hm ³)	PLU (mm)	VID (hm ³)	FUI (hm ³)	Volume Calculé (hm ³)
31/12/1998	347,55	125,715	0,083	0,004	0	0,005	0	0	0,002	/
01/01/1999	347,56	125,767	0,063	0,004	0	0,004	0	0	0,002	125,768
02/01/1999	347,57	125,818	0,063	0,004	0	0,005	0,4	0	0,002	125,819
03/01/1999	347,58	15,870	0,063	0,004	0	0,004	0,4	0	0,002	125,871
04/01/1999	347,59	125,921	0,060	0,004	0	0,002	1,9	0	0,002	15,922
05/01/1999	347,6	125,973	0,063	0,004	0	0,004	0	0	0,002	125,974
06/01/1999	347,61	126,024	0,062	0,004	0	0,004	0	0	0,002	126,025
07/01/1999	347,61	126,024	0,011	0,004	0	0,004	0	0	0,002	126,025
08/01/1999	347,61	126,024	0,012	0,004	0	0,005	0	0	0,002	126,025
09/01/1999	347,61	126,024	0,012	0,004	0	0,005	0	0	0,002	126,025
10/01/1999	347,6	125,973	0,000	0,004	0	0,005	0	0	0,002	126,013
11/01/1999	347,59	125,921	0,000	0,004	0	0,005	0	0	0,002	125,962
12/01/1999	347,6	125,973	0,065	0,004	0	0,006	0	0	0,002	125,974
13/01/1999	347,61	126,024	0,064	0,004	0	0,006	0	0	0,002	126,025
14/01/1999	347,62	126,076	0,060	0,004	0	0,002	0	0	0,002	126,076

Nous remarquons une nette différence entre les volumes mesurés chaque jour et ceux calculés moyennant l'expression (3-2) (§ section 3.2.2.2). Donc la balance des volumes n'est pas équilibrée et de ce fait on peut conclure qu'il y a :

- Soit des erreurs de mesure où de transcription ;
- Soit des erreurs de calcul des apports ;
- Soit des oublis de saisie des volumes vidangés.

3.3.3 Traitement des données d'exploitation

L'ensemble des données d'exploitation du barrage de Hammam Debagh a subi un traitement afin de les rendre plus fiables et plus exploitables. Comme la capacité de stockage de la retenue du barrage est sujette à une réduction progressive au cours du temps du fait qu'elle est exposée à un alluvionnement plus ou moins accéléré, les volumes, les surfaces et les volumes évaporés seront réajustés en fonction de la cote du plan d'eau mesurée rapportée aux courbes Hauteur /Volume/Surface du barrage établi pour chaque année d'exploitation en tenant compte des résultats de l'unique levé topo-bathymétrique réalisé en 2004.

3.3.3.1 Méthodologie proposée

Après correction des erreurs systématiques et accidentelles, les données d'exploitation récupérées auprès de l'ANBT sont soumises à un traitement final détaillé tenant compte d'une part, de la perte de volume d'eau dans la retenue et par conséquent de la variation des courbes Hauteur/Volume/Surface dans le temps et d'autre part de l'équilibre du bilan hydraulique du réservoir.

La démarche proposée consiste donc en un traitement en deux étapes :

1. Réajustement des volumes, surfaces et évaporations au niveau de la retenue ;
2. Établissement du bilan hydrologique de la retenue.

3.3.3.2 Réajustement des volumes, surfaces et évaporations au niveau de la retenue

La gestion des barrages en exploitation nécessite l'actualisation des courbes Hauteur/Volume/Surface après chaque opération de mesures topo-bathymétriques (de préférence une période de deux années). Cependant, on signale la rareté des travaux de recherche qui traitent de la prévision de la variation du volume de la vase en fonction du niveau d'eau dans les réservoirs.

Les travaux entrepris jusqu'à nos jours traitent l'aspect quantitatif global de l'alluvionnement des réservoirs (Ben Mammou & Louati, 2007). Certaines contributions portent sur la simulation des apports en sédiments dans les barrages (Benyounes et al., 2009) ou la simulation de la sédimentation dans les retenues de barrages (Bessenasse et al., 2003). D'autres travaux s'intéressent à la prévision de l'envasement globale des barrages. Dans ce qui suit, nous avons développé une approche qui permet d'établir une relation pratique qui lie le volume d'envasement à la côte du niveau d'eau dans le lac en utilisant les données de l'unique levé topo-bathymétrique de 2004.

A- Réajustement des volumes

La différence des volumes d'eau stockés à la côte normale entre l'année 1988 et l'unique mesure bathymétrique (2004), représente la part des sédiments érodés sur le bassin versant et transportés par les eaux de ruissellement jusqu'au barrage de Hammam Debagh depuis sa mise en eau. Les résultats de ce contrôle topo-bathymétrique, en matière d'évolution des volumes d'eau et de vase en fonction de la côte du niveau d'eau dans le barrage, sont présentés par la Figure 3-3.

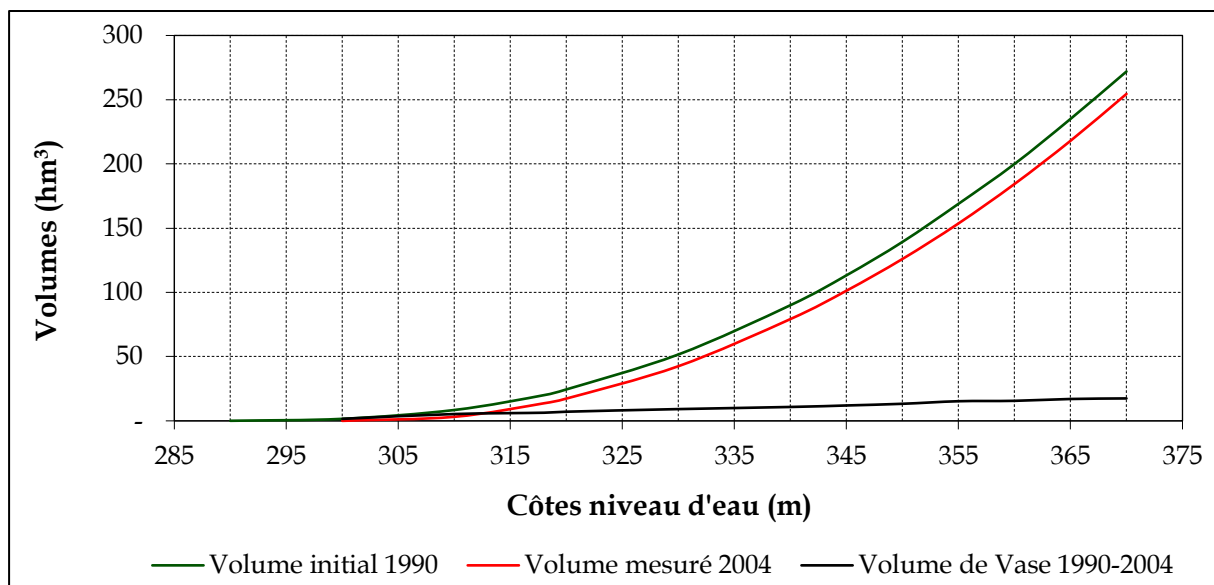


Figure 3-3 : Variation des volumes d'eau et de vase en fonction de la côte du niveau d'eau dans le réservoir.

Les résultats des mesures bathymétriques, en matière de volume total des sédiments et du taux moyen d'envasement, sont présentés dans le Tableau 3-4.

Tableau 3-4 : Évolution des volumes d’eau selon le levé topo-bathymétrique de Hammam Debagh.

Période	Volume initial (hm ³)	Volume restant (hm ³)	Volume de vase (hm ³)	Envasement annuel (m ³ /an)	Perte annuelle (%)
1988- 2004	200,00	184,35	15,65	978 125	0,49

Les volumes d’eau stockés réellement au pas de temps journalier ou mensuel durant l’année considérée sont déduits de la relation suivante :

$$\text{Volume réajusté (hm}^3\text{)} = \text{Volume initial (hm}^3\text{)} - \text{Volume vase (hm}^3\text{)} \quad (3-3)$$

Les volumes journaliers de la vase sont estimés par les équations des courbes de tendance traduisant la variation dans le temps de l’envasement moyen en fonction de la côte du niveau du plan d’eau observée.

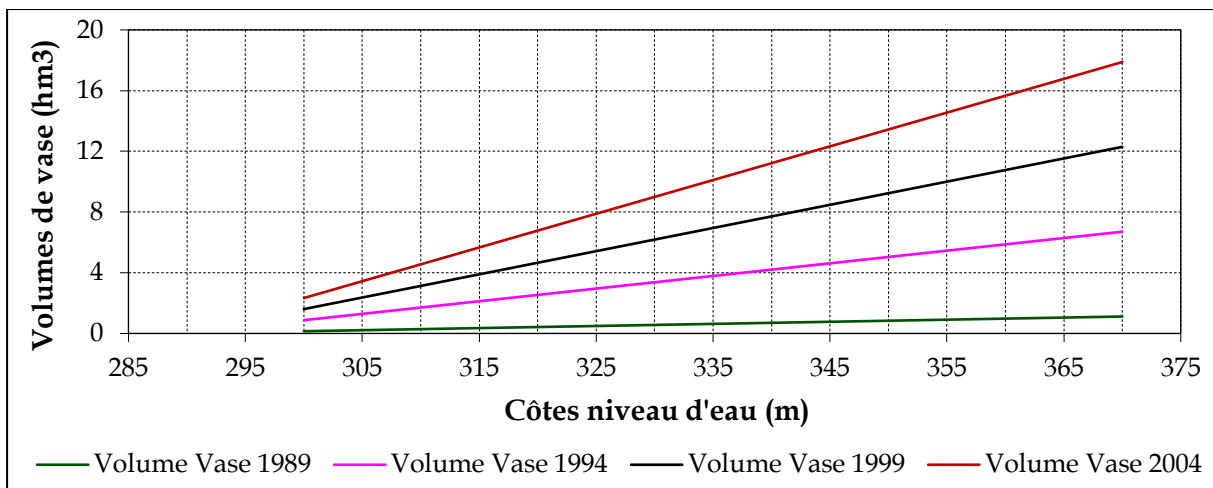


Figure 3-4 : Variation temporelle des volumes de vase en fonction de la côte du niveau d’eau dans le réservoir.

Tableau 3-5 : Équations de tendance et coefficients de détermination.

Période	Equations de tendance	Coefficients de détermination
1989	Volume de vase= 0,0139 * Côte - 4,0181	0,9915
1994	Volume de vase = 0,0833 * Côte - 24,109	0,9915
1999	Volume de vase = 0,1527 * Côte - 44,199	0,9915
2004	Volume de vase = 0,2221 * Côte - 64,290	0,9915

B- Réajustement des surfaces

L'évolution des surfaces inondées initiale et mesurée en 2004 en fonction de la cote du niveau d'eau dans le barrage est présenté par la [Figure 3-5](#).

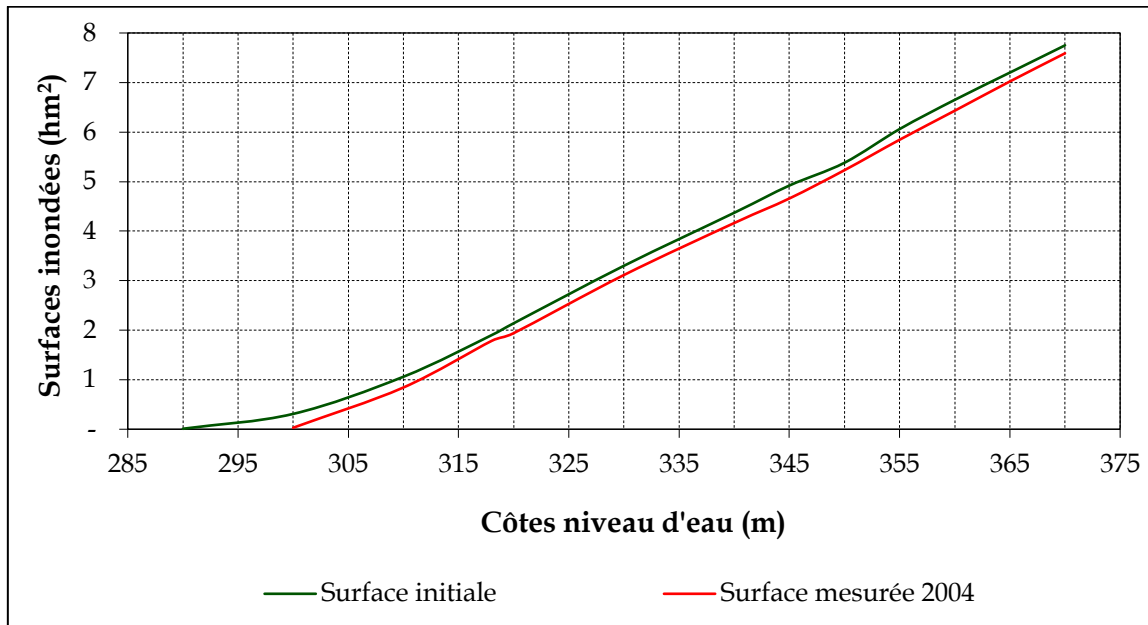


Figure 3-5 : Variation des surfaces inondées en fonction du niveau d'eau dans le réservoir.

Les résultats des mesures bathymétriques, en matière de surfaces perdues et de taux moyen de perte de surfaces, sont présentés dans le [Tableau 3-6](#).

Tableau 3-6 : Évolution des surfaces d'eau selon le levé bathymétrique de Hammam Debagh.

Période	Surface initiale (hm ²)	Surface restante (hm ²)	Surface perdue (hm ²)	Perte annuelle (m ² /an)	Perte annuelle (%)
1988- 2004	6,65	6,43	0,22	13750	15,04

Les surfaces du plan d'eau réajustées sont données directement par les équations des courbes de tendance traduisant leur évolution dans le temps en fonction des côtes du plan d'eau observées quotidiennement ou mensuellement durant l'année considérée. Nous donnons ci-après, à titre non exhaustif, les équations de tendance relatives à la variation des surfaces d'eau prévue pour les années 1989, 1994, 1999 et 2004 :

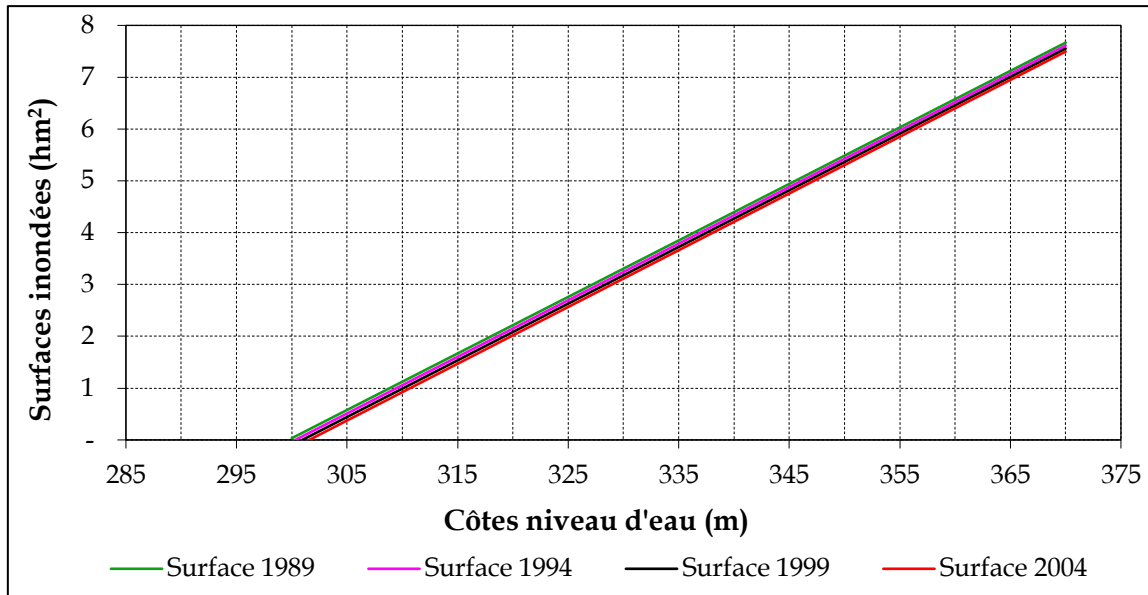


Figure 3-6 : Variation temporelle des surfaces inondées en fonction de la côte du niveau d'eau dans le réservoir.

Tableau 3-7 : Équations de tendance et coefficients de détermination.

Période	Equations de tendance	Coefficients de détermination
1989	Surface = 0,1091*Côte - 32,688	0,9982
1994	Surface = 0,1092*Côte - 32,806	0,9985
1999	Surface = 0,1094*Côte - 32,924	0,9986
2004	Surface = 0,1096*Côte - 33,042	0,9987

C- Réajustement des volumes évaporés au niveau de la retenue.

A l'issue de l'actualisation des surfaces, nous procédons au réajustement des volumes évaporés au niveau du réservoir en utilisant l'expression (3-1) (§ Section 3.2.2.2).

3.3.3.3 Etablissement du bilan hydraulique de la retenue

Le bilan des volumes d'un réservoir consiste en une balance d'eau effectuée jour par jour ou mois par mois à partir des données d'exploitation du barrage.

La Figure 3-7 schématise les états possibles de variation de la capacité du réservoir du barrage en utilisant l'expression (3-2) (§ Section 3.2.2.2) et les données d'exploitation après critique, traitement et actualisation.

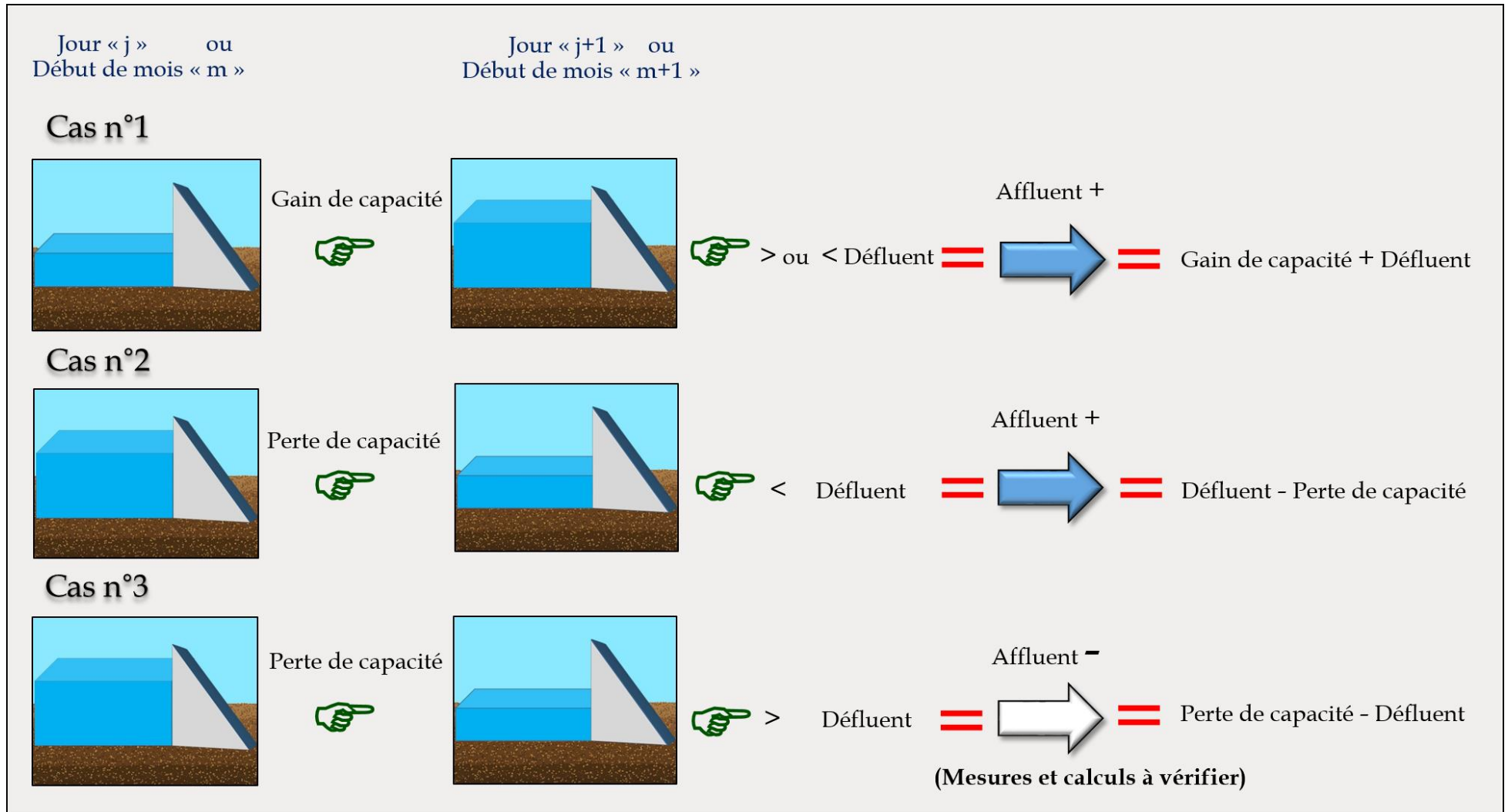


Figure 3-7 : Schéma du bilan d'un réservoir

TERMINOLOGIE :

Gain de capacité = Augmentation du volume de la retenue au début du jour « j+1 » où du mois « m+1 » ;

Perte de capacité = Diminution du volume de la retenue au début du jour « j+1 » où du mois « m+1 » ;

Défluent = Somme des sorties du barrage au jour « j+1 » où au mois « m +1 » ;
= Consommation + évaporation + fuites + vidange + déversement ;

Affluent = Somme des entrées = Apport du jour « j+1 » où du mois « m+1 »).

Le **Tableau 3-8** rend compte des éléments de ce bilan à l'échelle journalière.

Tableau 3-8 : Bilan hydraulique journalier du barrage Hammam Debagh.

Date	Côte	Volume	APP	AEP	IRR	EVAP	PLU	VID	FUI	Volume	Gain	Perte	Défluent	Affluent+	Affluent-
01	02	03	04	05	06	07	08	09	10	11	12	13	14	15	16
31/12/1998	347,55	116,843	0,083	0,004	0	0,005	0	0	0,002	/	/				
01/01/1999	347,56	116,894	0,060	0,004	0	0,004	0	0	0,002	116,894	0,050		0,010	0,060	
02/01/1999	347,57	116,943	0,060	0,004	0	0,005	0,4	0	0,002	116,943	0,049		0,011	0,060	
03/01/1999	347,58	116,994	0,060	0,004	0	0,004	0,4	0	0,002	116,994	0,050		0,010	0,060	
04/01/1999	347,59	117,043	0,057	0,004	0	0,002	1,9	0	0,002	117,043	0,049		0,008	0,057	
05/01/1999	347,6	117,093	0,060	0,004	0	0,004	0	0	0,002	117,093	0,050		0,010	0,060	
06/01/1999	347,61	117,143	0,059	0,004	0	0,004	0	0	0,002	117,143	0,049		0,010	0,059	
07/01/1999	347,61	117,143	0,010	0,004	0	0,004	0	0	0,002	117,143	0,000		0,010	0,010	
08/01/1999	347,61	117,143	0,011	0,004	0	0,005	0	0	0,002	117,143	0,000		0,011	0,011	
09/01/1999	347,61	117,143	0,011	0,004	0	0,005	0	0	0,002	117,143	0,000		0,011	0,011	
10/01/1999	347,6	117,093	0	0,004	0	0,005	0	0,038	0,002	117,093		0,049	0,049		0,038
11/01/1999	347,59	117,043	0	0,004	0	0,005	0	0,039	0,002	117,043		0,050	0,050		0,039
12/01/1999	347,6	117,093	0,062	0,004	0	0,006	0	0	0,002	117,093	0,050		0,012	0,062	
13/01/1999	347,61	117,143	0,061	0,004	0	0,006	0	0	0,002	117,143	0,049		0,012	0,061	
14/01/1999	347,62	117,193	0,058	0,004	0	0,002	0	0	0,002	117,193	0,050		0,008	0,058	

Comparées aux données du **Tableau 3-3**, nous concluons que :

1. Les volumes évaporés au niveau du réservoir n'ont pas beaucoup changé malgré le réajustement des surfaces ;
2. Les volumes stockés ont significativement changé ;
3. L'équilibre de la balance des volumes nécessite encore un réajustement des apports et des volumes vidangés (colonnes 4 et 9).

3.3.4 Analyse des données d'exploitation

L'objectif est de procéder à une analyse statistique et une visualisation graphique de l'évolution mensuelle des différents paramètres du bilan hydrique, à savoir : le volume stocké dans la retenue, l'apport entrant dans la cuvette du barrage, le volume alloué à l'alimentation en eau potable, le volume lâché pour satisfaire les besoins en eau d'irrigation, le volume évaporé, la quantité de pluie mesurée au niveau du barrage, le volume des vidanges, le volume des fuites récupérables et le volume déversé.

3.3.4.1 Statistique de base des données d'exploitation traitées

Nous avons procédé à l'analyse statistique des données retenues après correction pour la période 1990-2011. Les résultats obtenus sont résumés dans le [Tableau 3-9](#) qui porte sur les entités statistiques : la valeur minimale (Min), la valeur moyenne (Moy), la valeur maximale (Max), l'écart type (ET) et le coefficient de variation (CV).

Tableau 3-9 : Statistique descriptive des données traitées (252 observations)

Statistiques	Volume (hm ³)	APP (hm ³)	AEP (hm ³)	IRR (hm ³)	EVAP (hm ³)	PLU (mm)	VID (hm ³)	FUI (hm ³)	EVAC (hm ³)
Min	32,784	0,000	0,000	0,000	0,032	0,000	0,000	0,000	0,000
Moy	134,241	7,967	0,573	1,527	0,447	50,566	2,727	0,087	2,166
Max	205,153	130,286	1,549	9,362	1,637	278,600	39,858	0,248	82,789
ET	37,565	19,646	0,570	2,370	0,321	51,036	5,811	0,050	10,170
CV	0,280	2,466	0,996	1,552	0,717	1,009	2,130	0,575	4,696

Avec :

Volume : volume stocké dans le barrage ;

APP : volume des apports globaux ;

AEP : volume alloué à l'alimentation en eau potable ;

IRR : volume alloué à l'irrigation ;

EVAP : volume évaporé ;

PLU : pluie ;

VID : volume des vidanges ;

FUI : volume des fuites ;

EVAC : volume déversé.

3.3.4.2 Visualisation graphique de l'évolution des paramètres d'exploitation

L'évolutions des différents paramètres d'exploitation au cours du temps est visualisée par les graphiques suivants :

A- Les volumes stockés dans le barrage

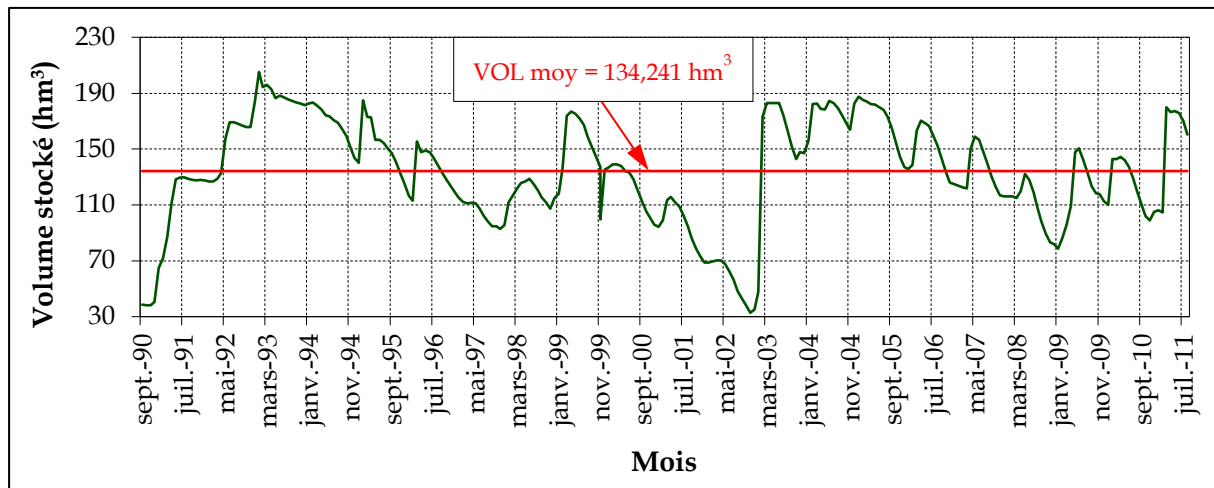


Figure 3-8 : Variation mensuelle des volumes stockés dans le barrage « 252 observations ».

La [Figure 3-8](#) montre que le marnage du réservoir est régulier. Cela est traduit par l'équilibre des nombres de mois dont les volumes stockés sont inférieurs ou supérieurs à la moyenne qui est de $134,241 \text{ hm}^3$, soit 126 mois correspondant à la moitié de la durée d'exploitation (21ans).

Dans l'ensemble, en comparant les volumes stockés pendant les deux décades d'exploitation (1990-2000 et 2001-2011), nous constatons que la première décade a enregistré le plus important volume stocké, soit $205,153 \text{ hm}^3$ dépassant ainsi le volume utile du barrage 200 hm^3 . En revanche les volumes stockés durant la deuxième décade n'ont jamais atteint le seuil de 190 hm^3 . Cela est dû, à notre avis, d'une part à l'accroissement des volumes alloués pour satisfaire les demandes en eau potable et d'irrigation et d'autre part à une gestion prudente des volumes d'eau mobilisés par le barrage.

B- Les apports entrant dans le barrage

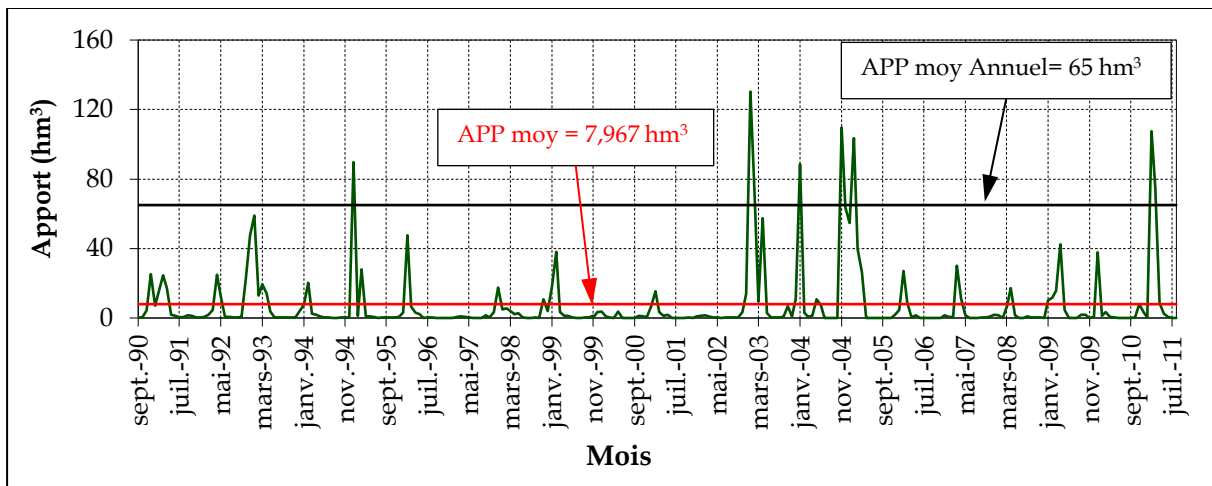


Figure 3-9 : Variation mensuelle des apports « 252 observations ».

La variation des apports mensuels durant 21 ans d'exploitation du barrage, de septembre 1990 à août 2011, est remarquablement irrégulière. Ainsi, les apports peuvent passer d'une valeur élevée à une valeur relativement faible d'un mois à l'autre, comme il est constaté entre les mois de janvier et février 1995. Les exemples de ce genre, au cours de la période considérée, sont nombreux.

La Figure 3-9 montre que 81 % de l'ensemble des mois de la période d'exploitation ont enregistré un apport inférieur à la moyenne qui est de $7,967 \text{ hm}^3$. Nous noterons aussi que seulement 3,2 % des apports dépassent l'apport moyen annuel de l'oued Bouhamdane (65 hm^3) et que 7,9 % des apports sont nuls.

Ainsi, tenant compte de la moyenne des apports mensuels, nous remarquons que le nombre de mois humides s'élève à 48, alors que le nombre de mois secs est de 204.

Globalement, la comparaison des apports des deux décades d'exploitation (1990-2000 et 2001-2011), révèle que la première décade a enregistré de faibles apports, soit 34,5% de l'apport globale reçu en 21 ans.

C- Les volumes alloués à l'alimentation en eau potable

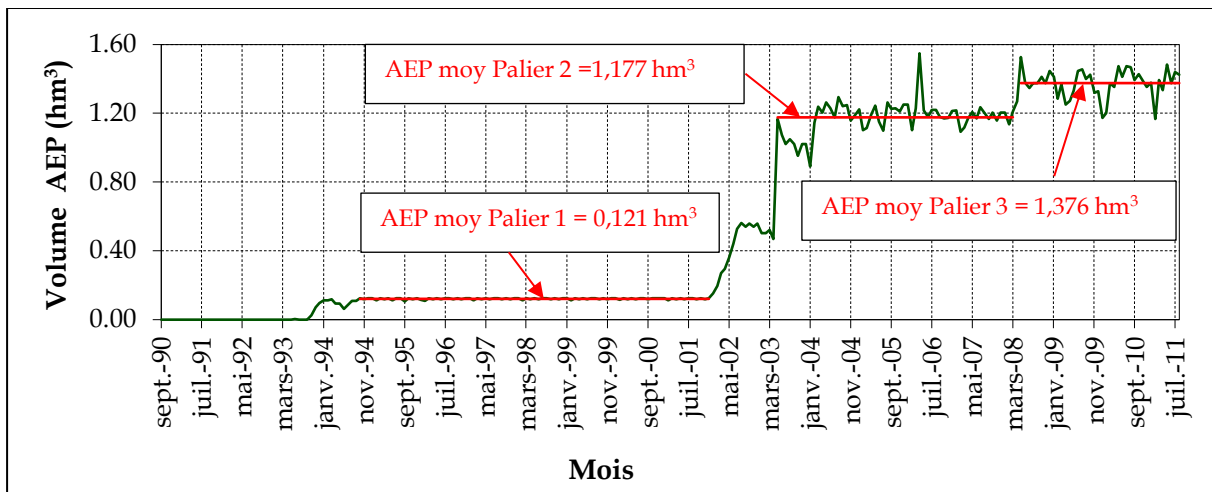


Figure 3-10 : Variation mensuelle des volumes alloués à l'AEP « 252 observations ».

De prime abord, nous remarquons que la dotation en eau potable a évolué selon trois paliers durant toute la période d'exploitation s'étalant entre 1990 et 2011 (Fig. 3-10). Cette évolution est due principalement au raccordement progressif des agglomérations au réseaux d'adduction à partir de la station de traitement des eaux. En effet, les trois premières années d'exploitation ont servi au remplissage du barrage. Une allocation progressive est pratiquée pendant une année (12 mois) pour atteindre le premier palier d'une valeur moyenne de $0,121 \text{ hm}^3$. Ce palier de dotation a été maintenu durant plus de sept ans (87 mois). Un autre système de dotation a été adopté sur une période d'une année et demi (16 mois), menant ainsi à un accroissement remarquable des dotations qui ont atteint un deuxième palier d'une valeur moyenne de $1,177 \text{ hm}^3$. Ce dernier a duré cinq années (60 mois).

Enfin, le troisième palier d'une durée de plus de trois années (40 mois) où les dotations adoptées fluctuent autour d'un seuil d'une valeur moyenne de $1,376 \text{ hm}^3$.

Nous distinguons aussi que les volumes alloués à l'alimentation en eau potable ont connu un accroissement à partir du début de la deuxième décennie d'exploitation. Cela est dû essentiellement à l'augmentation de la demande des villes concernées et au rabattement de la ressource souterraine.

D- Les volumes alloués à l'irrigation

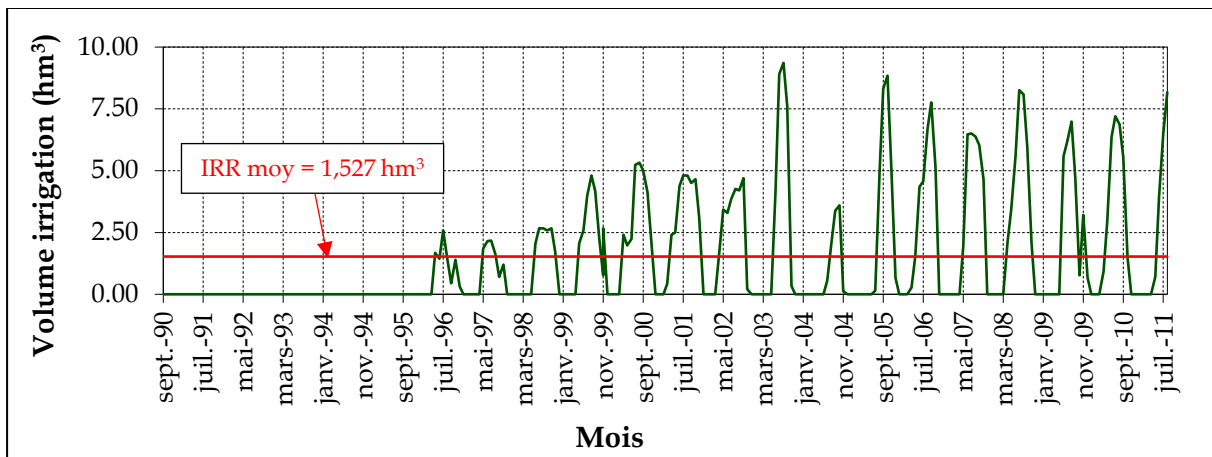


Figure 3-11 : Variation mensuelle des volumes alloués à l'irrigation « 252 observations ».

Nous constatons qu'environ les six premières années (68 mois) n'ont connu aucune allocation pour l'irrigation. Cela s'explique par le retard enregistré dans la réalisation des infrastructures du périmètre d'irrigation. Pour le reste des mois, nous constatons des fluctuations relativement importantes. En effet, la Figure 3-11 montre que 8,3 % de l'ensemble des mois de la période d'exploitation ont enregistré une dotation en eau d'irrigation inférieure à la moyenne ($1,527 \text{ hm}^3$) et que 33,3 % des allocations dépassent la dotation moyenne.

En outre, les mois de la première décennie (1990-2000) ont connu les plus faibles volumes alloués avec un maximum de $5,332 \text{ hm}^3$. En revanche, les maximums alloués ont été observés durant la deuxième décennie d'exploitation (2001-2011) avec des volumes excédant le maximum de la première décennie.

E- Les volumes évaporés

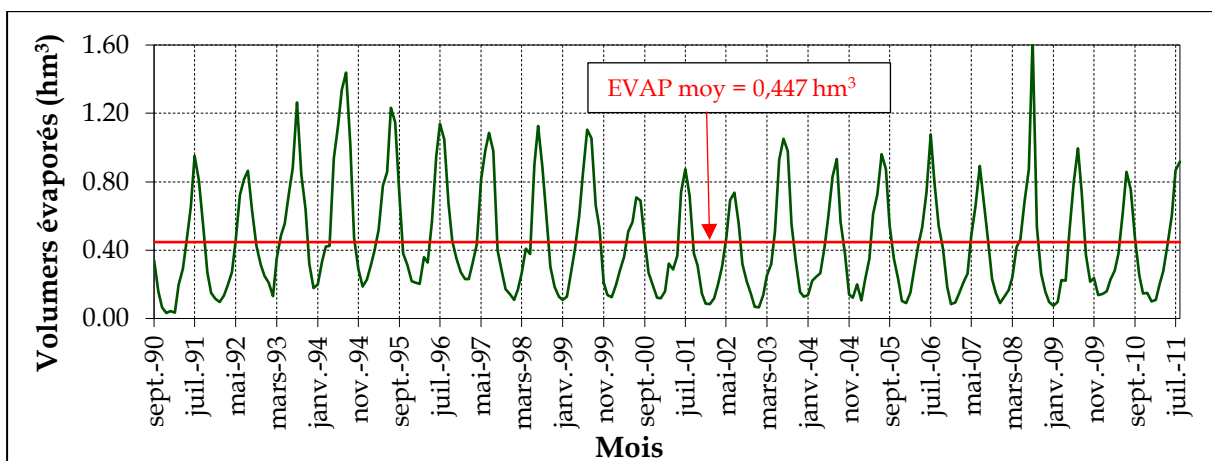


Figure 3-12 : Variation mensuelle des volumes évaporés « 252 observations ».

Nous remarquons que l'évolution des volumes évaporés est cyclique et assez régulière avec l'enregistrement de deux pics, l'un (volume max évaporé = 1,438 hm³) dans la première décade d'exploitation et l'autre (volume max évaporé = 1,637 hm³) durant la deuxième. Ceci est dû certainement à l'augmentation du plan d'eau conséquemment aux forts apports et aux températures élevées observées durant ces périodes.

F- La pluie

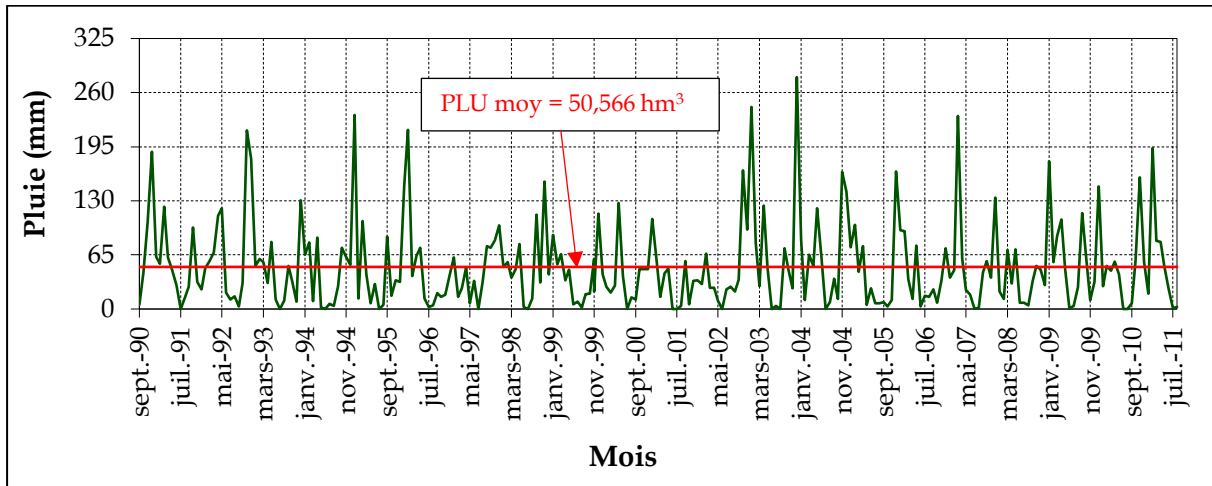


Figure 3-13 : Variation mensuelle de la pluie au niveau du barrage « 252 observations ».

La [figure 3-13](#) nous permet de constater une variation importante des précipitations mesurées à la station pluviométrique du barrage durant 21 ans, de septembre 1990 à août 2011. Ainsi, cette figure fait ressortir deux mois exceptionnels avec des valeurs remarquables relativement par rapport aux autres. L'un dans la première décade avec un total mensuel de 242,50 mm et l'autre dans la deuxième décade avec un total mensuel de 278,60 mm.

Par ailleurs par analogie à la variation des apports, l'analyse de la variation des quantités de pluie mensuelles a permis de noter qu'environ 56 % de l'ensemble des mois de la période d'exploitation ont enregistré une quantité de pluie inférieure à la moyenne (50,57 mm). Nous noterons aussi qu'environ 6 % des mois sont marqués par une pluviométrie nulle.

Enfin, la comparaison des totaux des deux décades nous permet de conclure qu'ils sont presque équilibrés (total première décade = 6342,7 mm et total deuxième décade = 6399,9 mm).

G- Les volumes des vidanges

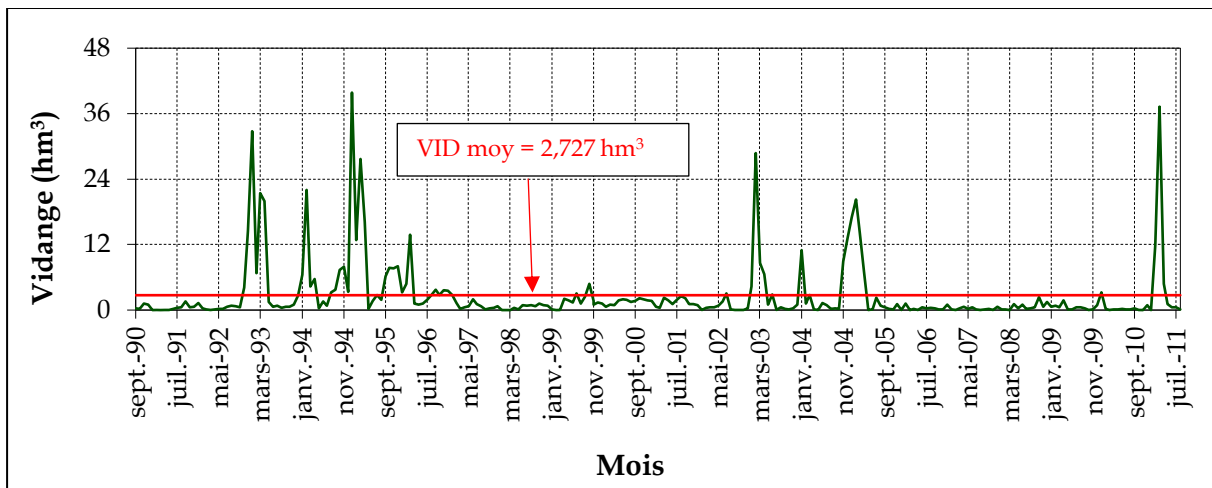


Figure 3-14 : Variation mensuelle des volumes de vidanges « 252 observations ».

En analysant la [Figure 3-14](#), nous remarquons que les volumes importants vidangés sont très épars dans le temps du fait qu'ils sont liés d'une part à la variation des apports entrant dans le barrage et d'autre part à la spécificité des consignes de gestion (programmation des opérations de vidange et état de l'envasement de la retenue).

H- Les volumes de fuites

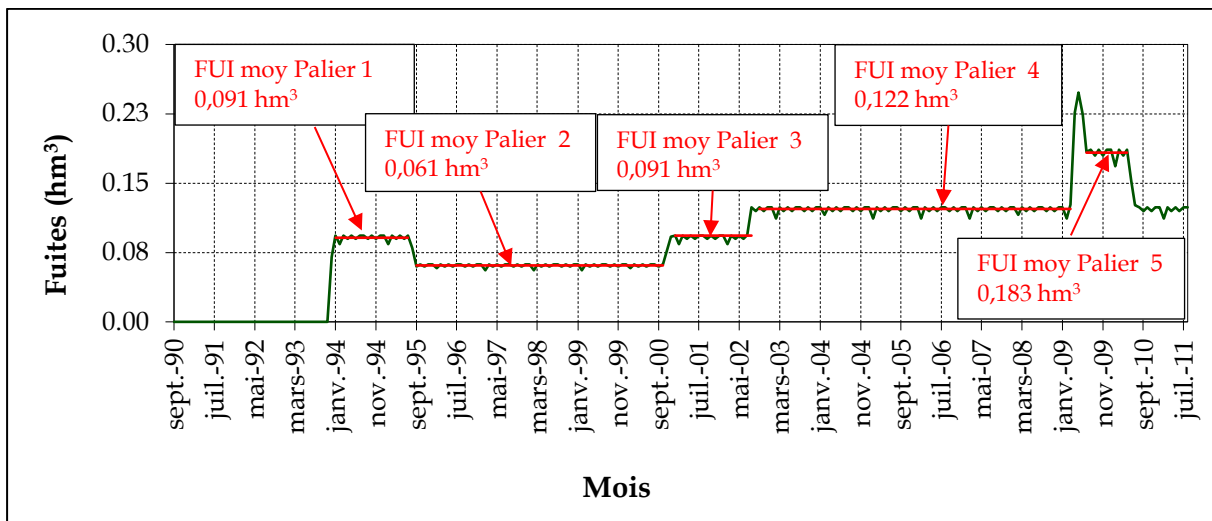


Figure 3-15 : Variation mensuelle des volumes de fuites « 252 observations ».

Nous constatons que pendant la période de mise en eau du barrage, environ trois années (38 mois), l'exploitant n'a enregistré aucune fuite. En revanche, les volumes de fuites ont évolué selon cinq paliers durant dix-huit années (18) d'exploitation s'étalant entre 1993 et 2011 ([Fig. 3-15](#)). Ces changements de palier sont dus à plusieurs paramètres à identifier et qui n'intéressent pas le présent travail.

I- Les volumes évacués

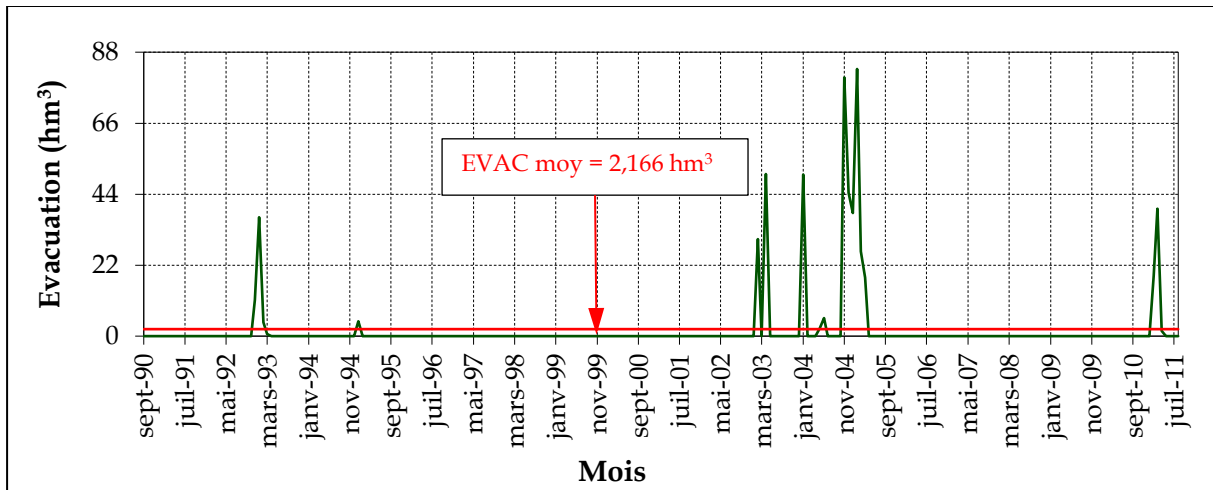


Figure 3-16 : Variation mensuelle des volumes évacués à travers le barrage « 252 observations ».

Nous constatons que pendant une période d'exploitation de 21 ans, le nombre de déversement est assez réduit. Ces déversements coïncident avec des niveaux de remplissage élevés du barrage.

3.4 Conclusion

La base de données d'exploitation du barrage de Hammam Debagh qui nous a été fourni par l'ANBT comporte des erreurs aussi bien systématiques qu'accidentelles. L'ensemble des données a subi un traitement afin de les rendre plus fiables et plus exploitables.

Les paramètres déduits de la courbe initiale (Hauteur/Volume/Surface) ont été actualisés par les résultats de l'unique levé topo-bathymétrique de 2004. Ce réajustement nous a permis d'établir la balance des volumes et d'assurer ainsi la cohérence des résultats du modèle développé dans le cadre de cette recherche.

En outre, après le traitement des données d'exploitation, nous avons procédé à une visualisation graphique de l'évolution dans le temps des différentes grandeurs du bilan hydraulique. Ainsi, toutes les figures reflètent des fluctuations remarquables des différents paramètres notamment les apports, les volumes alloués à l'irrigation, les volumes de vidange et les volumes évacués. Ces fluctuations sont moins accentuées pour le volume stocké, la pluie, le volume alloué à l'alimentation en eau potable et l'évaporation. Quant aux fuites, on constate une variation par palier quasi-constant. Ces observations sont approuvées par les coefficients de variations (CV) consignés dans le [Tableau 3-9](#).

Enfin, il nous semble plus judicieux d'avoir une vision plus critique sur l'ensemble des résultats des études précédentes ayant utilisé ces données dans leur état brut. Forcément, les résultats que nous avons obtenu à la fin de cette étude, sont différents des études antérieures. Cette différence est expliquée et justifiée par la démarche d'actualisation et de réajustement que nous avons opéré au préalable.

Chapitre 4

CHAPITRE 4 : Optimisation du levé Topo-bathymétrique du barrage Hammam Debagh

4.1 Introduction et objectifs

L'optimisation de la gestion des barrages en exploitation nécessite la mise à jour de la courbe « capacité-hauteur ». Pour cela, une estimation régulière des volumes de vase déposés dans la retenue est indispensable. Une telle opération nécessite des mesures topo-bathymétriques périodiques (de préférence une périodicité minimale) au cours de la durée d'exploitation. Or ces levés sont caractérisés par une forte densité de mesure consommant beaucoup de temps et d'argent. Dans le but d'optimiser la taille de ces levés topo-bathymétriques et augmenter leur fréquence, nous avons traité les données du premier levé topo-bathymétrique du barrage Hammam Debagh (Algérie) selon dix-neuf (19) scénarios de réduction de sa taille avec les procédures géostatistiques. Ce chapitre présente les idées directrices pouvant aider le géostatisticien à mener à bien une analyse géostatistique. Ce sont donc les méthodes et non les techniques mathématiques, qui sont exposés dans ce chapitre. Les notions de base de l'analyse géostatistique sont d'abord présentées, puis sa mise en œuvre est exposée en termes de ces différentes étapes. La fin du chapitre est consacrée à l'application sur le levé topo-bathymétrique réalisé en 2004 au niveau du barrage Hammam Debagh (Algérie).

4.2 Qu'est-ce que la géostatistique ?

La géostatistique est un ensemble de techniques statistiques pour l'analyse de données spatiales ([Journel & Huijbregts, 1978](#)).

La géostatistique propose une façon de décrire la continuité spatiale des phénomènes naturels et fournit les adaptations des techniques de régression classiques pour tirer profit de cette continuité ([Isaaks & Srivastava, 1989](#)).

La géostatistique est une discipline adaptée à la modélisation de variables ou de phénomènes structurés dans l'espace. Ces variables désignent par exemple les

propriétés des roches dans les formations géologiques. Ce type de variables présente une continuité spatiale qui se traduit par le fait que deux mesures, effectuées en deux points distants, de la variable considérée vont avoir tendance à être d'autant plus semblables que les points de mesure sont proches (Armstrong, 1998).

La géostatistique peut être considérée comme un ensemble de techniques numériques qui traitent la caractérisation des attributs spatiaux. Elle utilise des modèles essentiellement aléatoires d'une manière similaire à la façon dont l'analyse des séries chronologiques caractérise les données temporelles (Olea, 1999).

La géostatistique se définit comme l'étude des phénomènes régionalisés, c'est-à-dire des phénomènes qui s'étendent dans l'espace et y présentent une « organisation » ou une « structure ». Par « espace », nous entendons en général l'espace géographique, mais il peut aussi s'agir de l'axe temporel ou d'espaces plus abstraits (Emery, 2001).

Au sens le plus large, La géostatistique est une branche des statistiques visant à donner une description de quantités distribuées spatialement, ou encore spatio-temporellement. Elle est l'étude de la manière dont se répartissent dans l'espace les grandeurs de la variable d'intérêt telles que :

- ☞ La température, précipitation, neige, . . .
- ☞ La concentration d'un polluant, ozone, . . .
- ☞ La teneur d'un minerai d'un gisement, . . .
- ☞ L'altitude topographique, l'altitude bathymétrique,

4.3 Bref historique de la géostatistique

Historiquement, on peut dire que la géostatistique est aussi vieille que l'art des mines lui-même. Dès que les miniers se sont préoccupés de prévoir à l'avance les résultats de leurs travaux avenir, plus particulièrement, dès qu'ils ont commencé à prélever, à analyser des échantillons et à calculer des teneurs moyennes pondérées par des puissances et des zones d'influence, on peut dire que la géostatistique était née.

Cette naissance remonte à l'introduction des fonctions aléatoires pour l'étude des problèmes essentiellement géostatistiques par les écoles française et soviétique

(P. Levy, A. Kolmogorov, A. Khintchine) dès les années 1930. Les outils théoriques que nous utilisons en géostatistique linéaire étaient en place dès les années 1940 (H. Cramer, N. Wiener, S. Bochner). Dans l'espace d'une décennie, la géostatistique s'est élaborée indépendamment dans le domaine minier, dans le domaine forestier en Suède (B. Matern, 1949) et en météorologie en URSS (L.S. Gandin, 1950) (Chauvet, 2006).

D'une manière générale, on peut distinguer trois phases essentielles dans l'évolution de la géostatistique :

La première phase : est liée à l'estimation des gisements d'or exploités par les mineurs de l'Afrique du Sud. C'est dans les années 50, que les professeurs H.S. Sichel (1949-1952), D.G. Krige (1951) et le hollandais H.J. de Wijs (1952), de l'université du Witwatersrand en Afrique du Sud ont mis au point, pour les besoins des évaluations minières, les méthodes appelées plus tard géostatistiques pour pallier aux insuffisances de la statistique classique constatées dans l'étude des gisements très dispersés.

Deux aspects caractérisent cette première phase de l'évolution de la géostatistique :

- ☞ Pratique : les moyens de calculs étaient rudimentaires, mais l'abondance des publications en formules d'approximation et en courbes ou abaques constitue un véritable capital afin d'éviter aux utilisateurs de reprendre des calculs fastidieux.
- ☞ Théorique : les formalismes élaborés se placent souvent dans le cadre d'une loi de distribution donnée. Le modèle log-normal pour lequel se manifeste un engouement extraordinaire dans les années 1950 a été le plus utilisé par rapport au modèle gaussien (inadapté aux variables disséminées) (Chauvet, 2006).

La seconde phase : que l'on peut situer de 1965 à la fin des années 1970, s'est manifesté quand on a clairement réalisé l'insuffisance du calcul des probabilités classiques et la nécessité de réintroduire les caractères spatiaux des répartitions. Georges Matheron (ingénieur français du corps des mines), Matern (en foresterie) et Gandin (en météorologie) ont développés les théories et les outils utilisés précédemment par Krige ; d'où la naissance de la géostatistique linéaire stationnaire.

Parallèlement, on cherche à élargir les hypothèses de travail : c'est le développement d'une géostatistique non stationnaire en 1973 pour traiter des phénomènes présentant une tendance puis d'une géostatistique non linéaire en 1975 pour résoudre des problèmes de dépassements de seuil ou de changement de support.

En 1980 apparut la géostatistique non linéaire. C'est en 1985 que commencèrent les premières simulations (un domaine très actif en recherche aujourd'hui) (Chauvet, 2006).

La troisième phase : Il s'agit de la géostatistique de troisième génération (Matheron & Kleingeld, 1987), actuellement en plein expansion. Dans un contexte informatique de plus en plus performant, la géostatistique se développe dans les directions les plus variées. Cela se manifeste bien sûr au niveau des champs d'application, qui ne se limitent plus désormais aux ressources naturelles (mines ou pétrole), mais surtout, plus fondamentalement, les recherches s'orientent vers des directions théoriques extrêmement diverses.

On peut donc penser que cette troisième phase de la géostatistique est une étape de synthèse, dont il est encore trop tôt pour prévoir les aboutissants.

4.4 Domaines d'utilisation de la géostatistique

La géostatistique a commencé à se développer il y a plus de cinquante ans pour répondre à des problèmes posés par l'évaluation des gisements miniers et s'est étendue depuis à d'autres domaines d'activité. Au niveau international, le transfert de la géostatistique s'est effectué d'abord vers la pédologie, au début des années 1980 (Burgess & Webster 1980, 1980a), puis vers l'écologie à la fin des années 1980 (Robertson, 1987). Aujourd'hui, la géostatistique est utilisée dans de nombreux domaines tels que la foresterie, la météorologie, la sylviculture, la bathymétrie, la topographie (MNT), l'environnement, l'agriculture de précision, l'halieutique, l'épidémiologie, le génie civil et plus généralement dans n'importe quelle discipline manipulant des données localisées dans l'espace et nécessitant des modèles décrivant la dépendance spatiale entre ces données (Cressie, 1993).

Exemples d'applications de la géostatistique dans divers domaines.

- ☞ Estimation et planification des mines et des gisements pétroliers.
- ☞ Prospection géochimique et géophysique.
- ☞ Cartographie automatique (par ordinateur).
- ☞ Filtrage de signal.
- ☞ Simulations d'écoulements, prédiction et simulation de conductivités hydrauliques.
- ☞ Caractérisation de sites contaminés.
- ☞ Cartographie météorologique.
- ☞ Classification de sols.
- ☞ Estimation de la biomasse et de sa localisation en pêches.
- ☞ Estimation de la compaction du noyau imperméable d'un barrage (géotechnique).
- ☞ Répartition spatiale de la déformabilité des roches au pourtour d'une excavation.
- ☞ Charges hydrauliques et directions d'écoulement.
- ☞ Analyse et caractérisation d'images (biomédical, télédétection).
- ☞ Représentation numérique-analytique de surfaces pour la CAO-DAO.

4.5 Présentation de la géostatistique

4.5.1 Définitions

4.5.1.1 Variable régionalisée

La variable régionalisée $z(x)$ est considérée comme une réalisation d'une fonction aléatoire $Z(x)$ c'est-à-dire une famille de variables aléatoires dépendant de la localisation. Elle possède une structure dans l'espace ou dans le temps : celle-ci reprend toujours la même valeur au même endroit ou au même moment (Matheron, 1965).

En d'autres termes, sa variance est fonction de l'intervalle de temps (noté h) ou de la distance h qui sépare les observations.

4.5.1.2 Variable aléatoire

C'est une variable qui peut prendre un certain nombre de valeurs auxquelles sont associées des probabilités. En plus sa variance se maintient presque constante quel que soit l'intervalle de temps (noté h) ou la distance h qui sépare les observations.

4.5.2 Hypothèses de base

4.5.2.1 Hypothèse de stationnarité

Du point de vue mathématique, une fonction aléatoire $Z(x)$ est stationnaire de second ordre quand l'espérance mathématique existe et ne dépend pas du point x et que la covariance entre chaque paire $[Z(x+h), Z(x)]$ existe et ne dépend que de h :

1. $E[Z(x)]$ existe et prend la même valeur à tous les points x . En d'autres termes l'espérance mathématique ne dépend pas de x .

$$E(Z(x)) = \mu \text{ et } E[Z(x+h) - Z(x)] = 0 \quad (4-1)$$

2. La fonction de covariance existe, et est une fonction unique de la distance h séparant les observations spatiales ou temporelles,

$$Cov(Z(x), Z(x+h)) = C(h) \quad (4-2)$$

Ou le variogramme $\gamma(h)$ ne dépend pas de la localisation dans le temps ou dans l'espace x , il ne dépendra que de h (soit en module, soit en module et en direction).

Au niveau des modèles probabilistes, la stationnarité est une propriété sans ambiguïté : c'est l'invariance par translation de la loi spatiale du processus. Autrement exprimé, cela signifie que la loi d'un multiplet quelconque de points (de dimensions et orientation fixées) ne dépend pas de l'implantation de ce multiplet.

Cette propriété ne contient aucune hypothèse spécifique concernant les moments qui peuvent parfaitement ne pas exister. Il apparaît vite qu'un tel modèle est totalement inapplicable aux variables régionalisées.

C'est pourquoi il nous faut introduire la propriété de stationnarité d'ordre 2 (les espérances des valeurs ponctuelles et des doublets de points du processus existent et sont invariantes par translation). On se contente ainsi d'une propriété concernant au plus les lois bi-variables, mais on exige l'existence des moments d'ordre un et deux construits sur les valeurs ponctuelles (Chauvet, 2006).

Ces hypothèses ont uniquement pour but de permettre l'estimation des paramètres statistiques de notre modèle à partir des données. Elles visent essentiellement à "détacher" les deux premiers moments de localisations précises en permettant des translations des emplacements x et $x + h$. La covariance et le variogramme deviennent donc des fonctions dépendant uniquement de l'écart entre les points d'observation et non plus de leur localisation exacte dans l'espace ou dans le temps.

Cette hypothèse de stationnarité d'ordre deux ne peut pas être validée par un test statistique sur les données expérimentales, c'est plus l'expérience de l'utilisateur qui permet d'adopter ou d'écarter l'hypothèse stationnaire du second ordre. En cas d'une grande hétérogénéité du phénomène, zones très contrastées, il est préférable de les traiter séparément.

4.5.2.2 Hypothèse intrinsèque

Une fonction aléatoire $Z(x)$ est intrinsèque quand ses variations $[Z(x + h) - Z(x)]$ sont stationnaires d'ordre deux.

$$E[Z(x + h) - Z(x)] = 0 \quad (4-3)$$

$$Var[Z(x + h) - Z(x)] = E[Z(x + h) - Z(x)]^2 = 2\gamma(h) \text{ ne dépend que de } h. \quad (4-4)$$

Toute fonction stationnaire d'ordre 2 est intrinsèque, mais la réciproque n'est pas vraie. L'hypothèse intrinsèque est plus générale que celle de la stationnarité. En particulier, la covariance d'une fonction intrinsèque peut ne pas exister (elle n'existe que si le variogramme est borné), $Cov(0)$ est finie.

4.5.3 Analyse variographique

L'analyse variographique est l'inférence du variogramme à partir de données expérimentales. Elle permet de restituer des informations quant à la distribution

spatiale de la variable régionalisée, c'est-à-dire de la réalisation de la fonction aléatoire d'intérêt. L'analyse variographique constitue une étape cruciale dans une étude géostatistique : elle permet postérieurement d'estimer les valeurs inconnues de la variable régionalisée et d'assortir leur estimation d'une précision.

4.5.3.1 Variogramme

En géostatistique, l'idée fondamentale est que la nature n'est pas entièrement « imprévisible ». Deux observations proches devraient, en moyenne, se ressembler davantage que deux observations éloignées. C'est pourquoi on cherche préalablement à quantifier cette continuité. La fonction la plus utilisée pour décrire cette continuité est le variogramme (Matheron, 1970). Donc le variogramme explore la structure de la variabilité spatiale.

Considérons tout d'abord que la variable régionalisée $z(x)$ est une réalisation d'une fonction aléatoire stationnaire $Z(x)$ d'ordre deux. Soit deux points x et $x + h$ séparés d'une distance h . la différence des valeurs en ces deux points est nulle en moyenne, quelle que soit la localisation du point x et quelle que soit la longueur et l'orientation du vecteur h .

La variance de cette différence, quant à elle, est d'autant plus petite que les points sont rapprochés et ne dépend que du vecteur h . C'est uniquement aux points (x) et $(x + h)$ qu'on dispose d'information.

Le variogramme ou semi-variogramme se définit comme suit sous l'hypothèse intrinsèque :

$$\gamma(h) = \frac{1}{2} \text{Var} [Z(x + h) - Z(x)] = \frac{1}{2} E[Z(x + h) - Z(x)]^2 \quad (4-5)$$

Où x : Vecteur uni, bi ou tri-dimensionnel ;

h : Vecteur distance ou écart entre les observations.

4.5.3.2 Propriétés théoriques du variogramme

Le variogramme est une fonction paire, à valeurs positives appartenant à la classe des fonctions de type négatif conditionnel (Christakos, 1984).

$$\sum_{i=1}^k \sum_{j=1}^k \lambda_i \lambda_j \gamma(x_i - x_j) \leq 0 \quad (4-6)$$

Cette propriété, est une condition nécessaire et suffisante pour que $\gamma(h)$ soit le variogramme d'une fonction aléatoire intrinsèque. L'adjectif conditionnel se réfère au fait que l'inégalité n'est vraie que pour une classe restreinte de pondérateurs $(\lambda_i)_i$, à savoir ceux de somme totale nulle (Emery, 2001).

Parmi les propriétés du variogramme, trois d'entre elles sont fondamentales :

- ☞ Le variogramme est pair : $\gamma(h) = \gamma(-h)$.
- ☞ Il s'annule à l'origine : $\gamma(0) = 0$.
- ☞ Il est positif ou nul : $\gamma(h) \geq 0$.

En général le variogramme est croissant quand h augmente : en effet, plus les échantillons sont éloignés, moins ils sont corrélés et donc plus grand est l'écart moyen $(Z(x+h) - Z(x))^2$.

En pratique, à cause en particulier des effets de bords, le variogramme calculé est croissant jusqu'à un maximum puis globalement légèrement décroissant ou stable (Emery, 2001).

A- Isotropie

Le variogramme $\gamma(h)$ est isotrope s'il est identique dans toutes les directions, c'est-à-dire s'il ne dépend que du vecteur h de translation entre les points (x) et $(x+h)$. S'il dépend aussi de l'orientation du vecteur h , il est alors anisotrope.

B- Effet de pépite

C'est la valeur verticale non nulle que prend le variogramme $\gamma(h)$ entre l'origine (0) et la valeur de h la plus petite (Fig. 4-1). Il s'agit donc d'une variation à très courte échelle qui peut être due aux causes suivantes :

- ☞ Erreurs de localisation ;
- ☞ Existence de microstructures, inaccessibles à l'échelle des observations ;

- ☞ Existence d'erreurs de mesures (erreurs d'analyse, d'échantillonnage et précision analytique) ;
- ☞ Existence de micro-variabilité des phénomènes étudiés.

Le cas extrême est l'effet de pépité pur qui caractérise une absence totale de corrélation entre les échantillons.

C- Portée

Distance à laquelle le variogramme atteint un palier. Cette portée mesure la distance à laquelle deux observations ne sont plus liées linéairement. À cette distance la valeur du variogramme correspond à la variance de la variable aléatoire (Fig. 4-1).

D- Palier

C'est la valeur maximale que prend $\gamma(h)$ quand h devient grand et qui correspond à la dissemblance maximale. Dans ce cas la fonction aléatoire est stationnaire d'ordre deux et sa variance est a priori égale au palier. Un palier peut n'être atteint qu'asymptotiquement du fait que la portée réelle est infinie. En pratique, dans une telle situation une portée est définie par la distance à laquelle le variogramme atteint 95% de la valeur de son palier. (Fig. 4-1).

Remarque : Si le variogramme est non borné, il ne possède ni portée ni palier. La variance de la fonction aléatoire n'est pas définie pour un tel variogramme. Cette fonction aléatoire n'est donc pas stationnaire de deuxième ordre, mais seulement stationnaire intrinsèque.

L'absence de palier indique que le phénomène observé évolue dans l'ensemble du domaine.

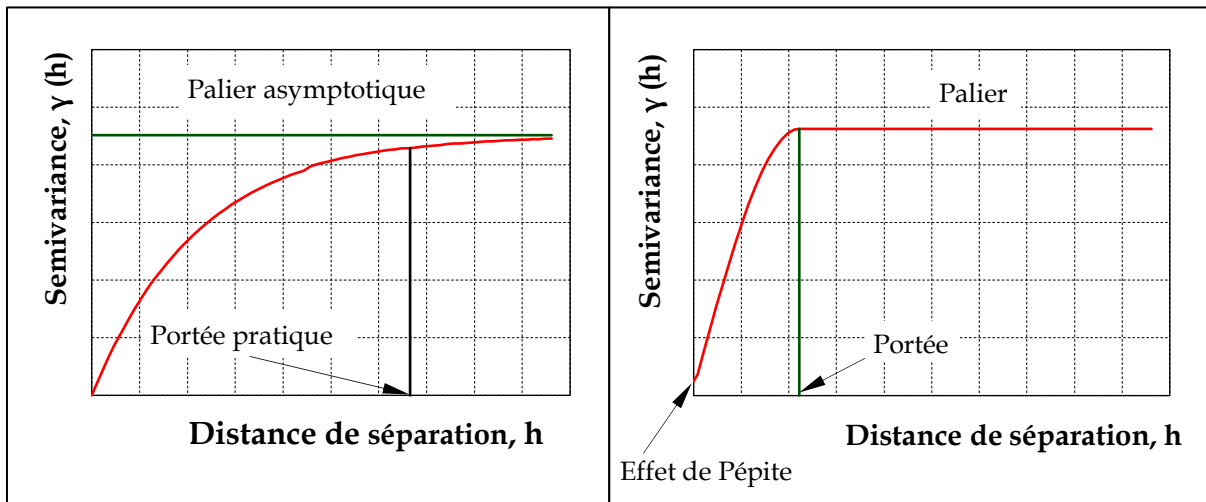


Figure 4-1 : Exemple de variogramme avec palier

4.5.3.3 Estimation du variogramme

Le variogramme est dit expérimental du fait qu'il utilise les données mesurées. Il est donné par l'estimateur suivant :

$$\gamma^*(h) = \frac{1}{2N(h)} \sum_{i=1}^{N(h)} [Z(x_i + h) - Z(x_i)]^2 \quad (4-7)$$

Où $N(h)$ nombre de paires d'observation dont l'écart est de h .

h étant une distance égale au multiple de l'intervalle de l'échantillonnage.

Arnaud et Emery (2000) conseillent de calculer le variogramme expérimental seulement pour les distances inférieures à la moitié de la distance maximale entre deux points d'observation, car la dispersion de $\gamma^*(h)$ autour de $\gamma(h)$ augmente lorsque h devient grand.

4.5.3.4 Modélisation du variogramme expérimental

La dernière phase de l'analyse variographique consiste à ajuster "au mieux" le variogramme expérimental par une fonction théorique de type négatif conditionnel. Vérifier cette propriété est une tâche très difficile, c'est pourquoi, on choisit souvent une courbe théorique obéissant à l'un ou l'autre des modèles de variogramme les plus couramment utilisés en pratique. Par ailleurs, la gamme des modèles proposés et le

formalisme développé dans la littérature géostatistique mettent en général l'utilisateur à l'abri de grosses bévues mathématiques (Chauvet, 2006).

Les deux principales caractéristiques d'un variogramme sont :

- ☞ Son comportement à l'origine qui traduit le degré de régularité de la régionalisation ;
- ☞ La présence ou l'absence d'un palier.

Les modèles variographiques les plus fréquemment utilisés sont donnés ci-après (Arnaud & Emery, 2000) :

A- Modèles avec palier ou de transition

a- Modèle pépitique pur

Modèle caractérisé par l'absence totale d'une structure spatiale. Autrement dit, il n'y a aucune auto corrélation spatiale dans les données même à des petites distances correspondant au pas d'échantillonnage (Fig. 4-2). Cet effet pépitique pur est explicable par l'un ou l'autre des artefacts suivants :

- ☞ Des erreurs de mesure ou plus généralement des incertitudes sur les mesures ;
- ☞ Des erreurs d'acquisition ;
- ☞ Des échelles non appropriées pour étudier les phénomènes spatiaux. Cet artefact est généralement rencontré dans le cas de présence d'une structure spatiale dont l'échelle est nettement inférieure à la maille d'échantillonnage. Il s'agit dans ce cas d'une micro-régionalisation.

$$\gamma(h) = \begin{cases} 0 & \text{si } h = 0 \\ C & \text{si } h > 0 \end{cases} \quad (4-8)$$

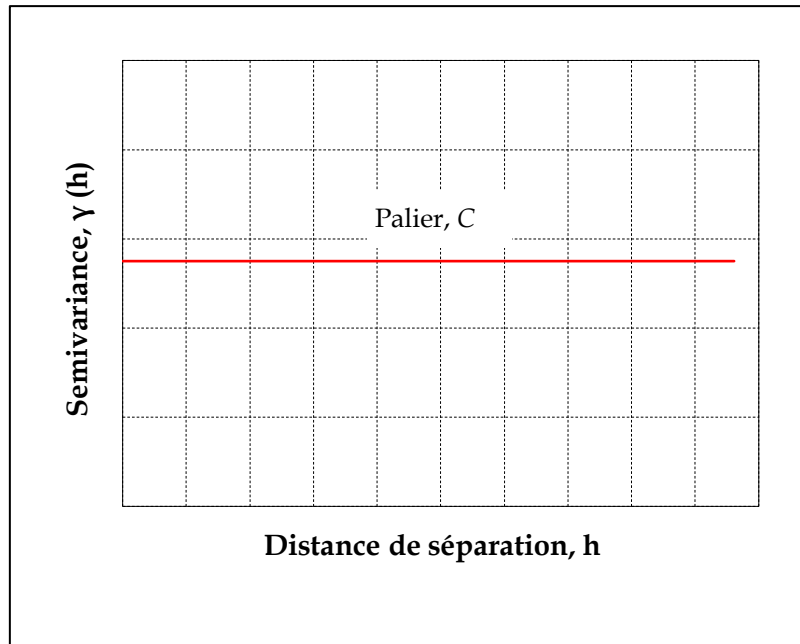


Figure 4-2 : Modèle théorique d'un variogramme pépétique pur

b- Modèle linéaire avec palier

Ce modèle montre que la variable régionalisée est moins régulière, elle est continue mais pas différentiable. C'est un modèle de palier C et de portée A_0 (Fig. 4-3) d'usage moins fréquent.

$$\gamma(h) = \begin{cases} C \left[\frac{h}{A_0} \right] & \text{si } 0 \leq h \leq A_0 \\ C & \text{si } h \geq A_0 \end{cases} \quad (4-9)$$

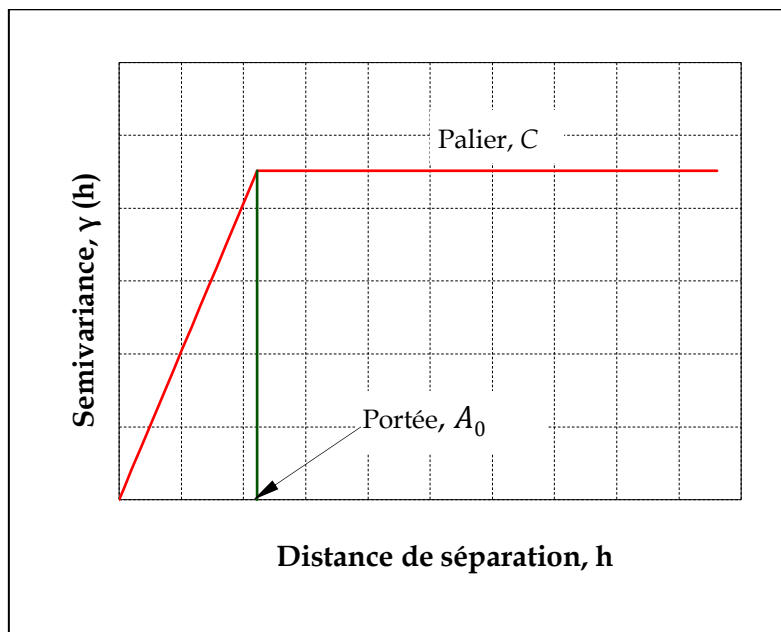


Figure 4-3 : Modèle théorique d'un variogramme linéaire avec palier

c- Modèle sphérique

C'est une fonction monotone croissante qui caractérise les phénomènes de transition dont la variabilité est bornée (Fig. 4-4). Ce modèle représente le variogramme théorique le plus utilisé pour l'ajustement variographique.

$$\gamma(h) = \begin{cases} C_0 + C \left[1.5 \frac{h}{A_0} - 0.5 \left(\frac{h}{A_0} \right)^3 \right] & \text{si } 0 < h < A_0 \\ C_0 + C & \text{si } h \geq A_0 \end{cases} \quad (4-10)$$

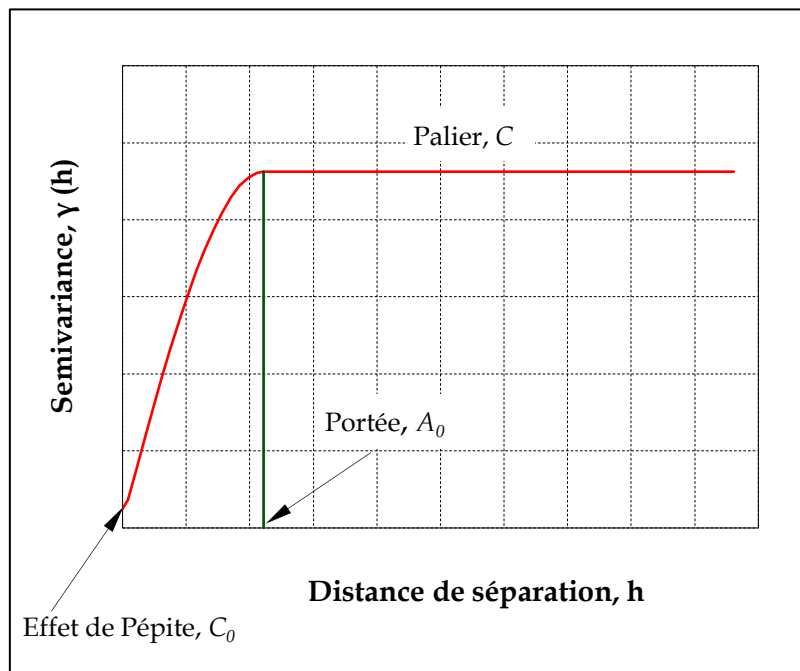


Figure 4-4 : Modèle théorique d'un variogramme sphérique

d- Modèle exponentiel.

Ce modèle est caractérisé par une croissance linéaire de la variance à l'origine et un accroissement rapide aux grandes valeurs de h sans jamais atteindre le palier (Fig. 4-5). Dans ce cas, la portée réelle est infinie mais une portée pratique égale à $3A_0$ est définie pour une variance égale à 95% de la valeur de son palier.

$$\gamma(h) = C \left[1 - \exp\left(-\frac{3h}{A_0}\right) \right] \quad \text{si } h \geq A_0 \quad (4-11)$$

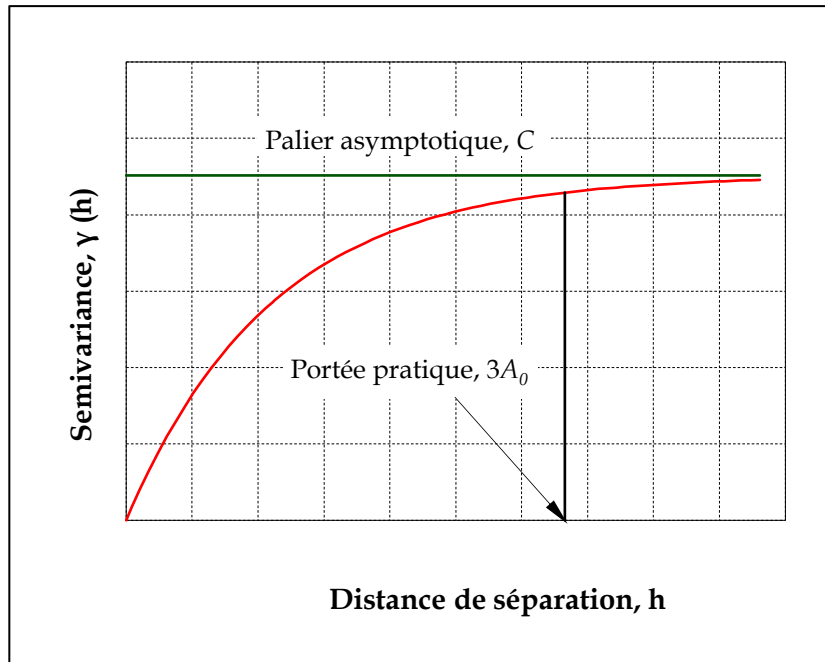


Figure 4-5 : Modèle théorique d'un variogramme exponentiel

e- Modèle gaussien

Le modèle gaussien décrit un phénomène spatial fortement régulier. Sa dérivée à l'origine (O) est nulle démontrant une très forte dépendance entre les observations séparées de distances h faibles et au fur et à mesure que la distance h augmente, la variance augmente au début d'une façon moins que proportionnelle jusqu'au point d'inflexion (point de changement de concavité de la courbe) puis d'une façon plus que proportionnelle sans jamais atteindre le palier (Fig. 4-6). Comme pour le modèle exponentiel, la portée réelle du modèle gaussien est infinie. De la même façon que le modèle exponentiel, une portée pratique est définie par une distance égale à $A_0\sqrt{3}$ pour laquelle le variogramme atteint 95% de la valeur de son palier.

$$\gamma(h) = C \left[1 - \exp\left(-3\left(\frac{h}{A_0}\right)^2\right) \right] \quad \text{si } h \geq A_0 \quad (4-12)$$

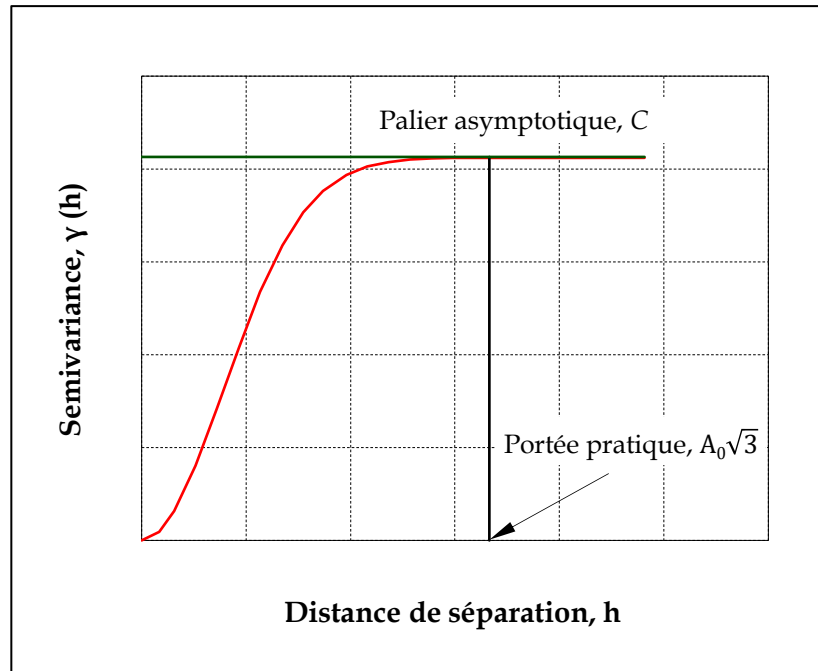


Figure 4-6 : Modèle théorique d'un variogramme gaussien

B- Modèles sans palier

a- Modèle linéaire sans palier

Le variogramme n'est pas bornée, il décrit une fonction aléatoire de variance à priori infinie, donc incompatible avec l'hypothèse de stationnarité d'ordre 2 mais compatible avec l'hypothèse de stationnarité intrinsèque. Cette allure linéaire montre que la variable régionalisée est moins régulière, elle est continue mais pas différentiable (Fig. 4-7).

$$\gamma(h) = mh \quad \text{si } h \geq A_0 \quad (4-13)$$

m : Pente de la droite à l'origine.

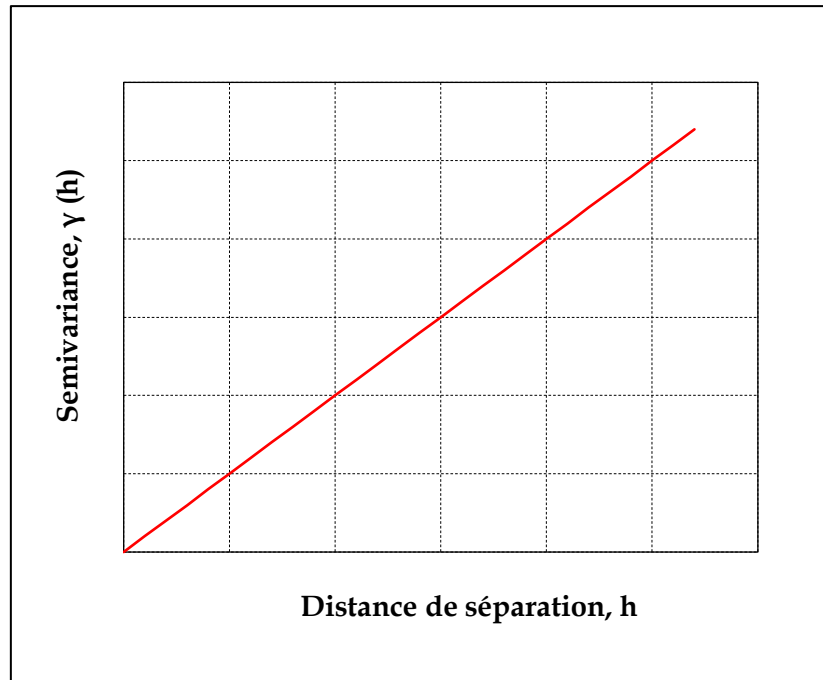


Figure 4-7 : Modèle théorique d'un variogramme linéaire sans palier

Il est important de noter que ces six modèles ne sont pas les seuls que l'on peut utiliser, en réalité, il en existe un très grand nombre.

Remarque : Il y a un lien étroit entre la nature de la variable étudiée et le type de modèle que l'on est susceptible de rencontrer. Ainsi, le modèle gaussien exprime une très grande continuité typique d'une variable comme la topographie, l'épaisseur d'une formation géologique, la charge hydraulique. Pour des variables comme les teneurs de gisement, les propriétés mécaniques des roches, les analyses géochimiques, en général les modèles sphérique et exponentiel sont beaucoup plus courants. Pour des variables discrètes, le modèle gaussien est à proscrire. Les variogrammes linéaires représentent plutôt des phénomènes qui évoluent régulièrement sans qu'une véritable structure soit décelable.

4.5.3.5. Calcul du variogramme expérimental et ajustement de modèles

Pour le calcul du variogramme expérimental et le choix du meilleur modèle à ajuster [Marcotte \(2013\)](#) recommande :

- 1- Accorder plus de poids aux points du variogramme expérimental calculés avec beaucoup de paires.
- 2- Essayer d'avoir $N(h) \geq 30$ pour chaque point expérimental du variogramme. Si ce n'est pas possible pour certaines classes, accorder moins d'importance à ces points. Si le nombre de paires est très faible (≤ 10), ne considérer plus du tout le point.
- 3- Accorder plus de poids aux premiers points du variogramme (h petit) car ce sont ces valeurs qui ont le plus d'impact dans les calculs géostatistiques.
- 4- Lorsque h dépasse environ $d_{max}/2$, ne pas tenir compte des valeurs du variogramme. d_{max} est la taille du phénomène étudié dans la direction considérée.
- 5- Chercher à obtenir des modèles les plus simples possible qui rendent bien compte des valeurs expérimentales.

4.5.3.6. Conclusion

L'analyse variographique fait appel à la fois à l'information disponible sur le phénomène étudié, à un certain savoir-faire dans le choix et l'utilisation des outils et à un sens de l'approximation. C'est l'étape essentielle de toute étude géostatistique et certainement celle qui se prête le moins à l'automatisme (Emery, 2001).

4.5.4 Interpolation spatiale

L'interpolation spatiale est une modélisation mathématique parfois utile lors de l'étude d'un phénomène naturel qui s'étend continûment sur le territoire et dont l'homme de science a, tous les temps, été confronté. Le phénomène naturel examiné est représenté par des mesures localisées sur le territoire. C'est donc un ensemble de procédures qui consistent à estimer la valeur d'une variable dans des sites non échantillonnés ou non mesurés.

Ces problèmes ont longtemps été résolus d'une façon plus ou moins objective du fait que les procédés de résolution ont été limités par les moyens techniques disponibles.

Toutefois les méthodes d'interpolation ont connu un développement opérationnel avec l'évolution des moyens de calcul puissants.

Selon les modèles mathématiques sur lesquels elles reposent, ces techniques se divisent usuellement en deux groupes :

- ☞ Méthodes déterministes ;
- ☞ Méthodes stochastiques.

Pour faire un choix d'une méthode d'interpolation locale, il est important parfois de comparer les modèles déterministes, purement mathématiques (et qui ne nécessitent aucune modélisation du phénomène étudié), à des méthodes stochastiques qui permettent d'appréhender la structure spatiale du phénomène naturel étudié comme l'interpolation par Krigeage (Fig. 4-8).

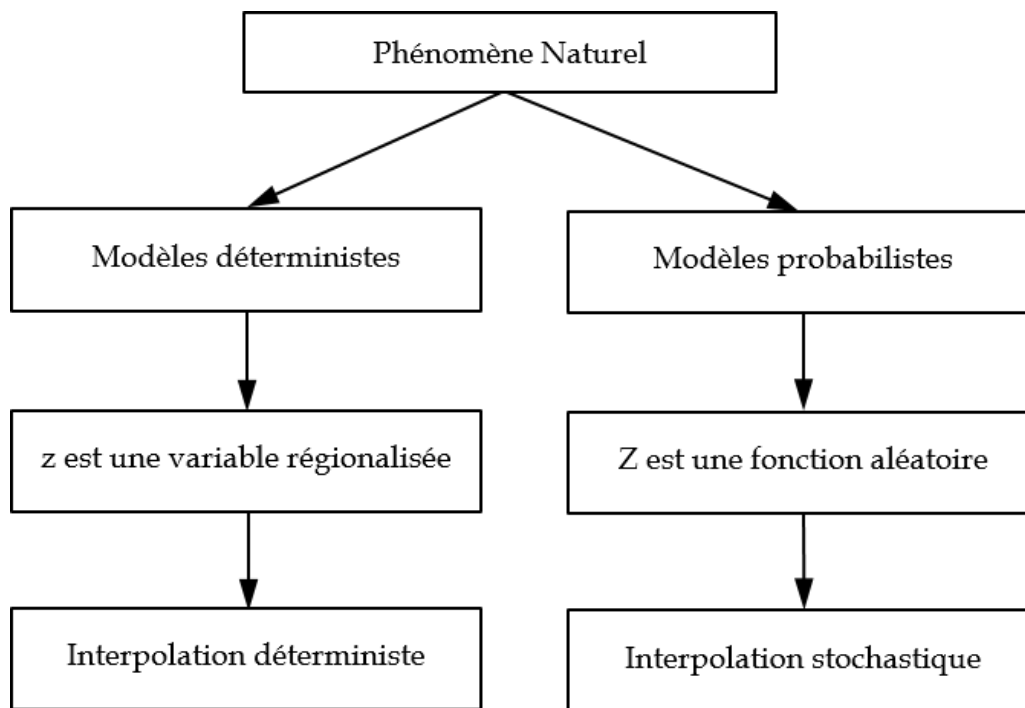


Figure 4-8 : Types d'interpolation

Le choix de la méthode dépend de la nature de la variable étudiée, du degré de précision et du temps de traitement alloué.

4.5.4.1 Méthodes d'interpolations déterministes

Ce sont des méthodes d'interpolation classiques basées sur des algorithmes purement déterministes, généralement géométriques, sans tenir compte du phénomène

physique qui nous intéresse. Elles reposent sur des hypothèses difficilement vérifiables mais qui permettent d'aboutir à des calculs simples. Les méthodes mathématiques de ce type sont nombreuses et diversement complexes, pour lesquelles chaque valeur interpolée utilise simplement les valeurs mesurées les plus proches et non l'ensemble des points.

Dans ce qui suit nous présenterons les méthodes les plus utilisées pour des fins de cartographie à savoir :

La méthode des plus proches voisins

Appelée aussi méthode de polygones de Thiessen (proposée en 1911 pour évaluer des moyennes spatiales sur un domaine à partir d'informations ponctuelles), elle accorde toute l'importance aux voisins. Une interpolation de ce genre nécessite la définition d'un secteur autour du point échantillonné pour lequel on doit prédire des valeurs et trouver les points d'échantillonnages inclus dans ce secteur.

La construction du polygone de Thiessen commence par faire relier chaque point aux points avoisinants par des segments de droite. Ensuite on coupe chacun des côtés des triangles ainsi formés par la médiatrice. Le résultat de cette opération est la création d'objets surfaces dont les poids sont rattachés respectivement à chacune de ces nouvelles zones (Fig. 4-9).

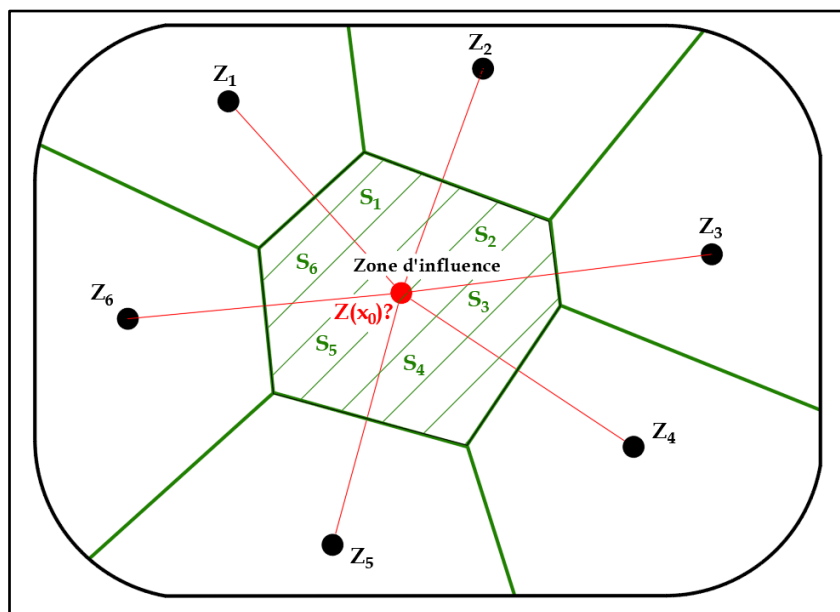


Figure 4-9 : Interpolation polygonale

Où Z_1, Z_2, Z_3, Z_4, Z_5 , et Z_6 représentent les valeurs mesurées de la variable régionalisée, objet de l'étude et $Z^*(x_0)$ la valeur à estimer.

Et S_1, S_2, S_3, S_4, S_5 , et S_6 les surfaces d'influence relative du polygone de Thiessen (S).

Chaque point a un poids λ_i , proportionnel à la surface, défini comme suit :

$$\lambda_i = \frac{S_i}{\sum_{i=1}^n S_i} \quad \text{et} \quad Z^*(x_0) = \sum_{i=1}^n \lambda_i Z_i \quad \text{avec} \quad \sum_{i=1}^n \lambda_i = 1 \quad (4-14)$$

Nous remarquons alors que les termes λ_i appelés aussi coefficients de Thiessen ne dépendent que de la répartition spatiale de la variable régionalisée par rapport au domaine d'étude. Ils ne dépendent pas de la valeur de celle-ci. Une fois la construction géométrique effectuée, on appliquera λ_i à tous les points à étudier.

Incontestablement, l'unique avantage de cette méthode réside dans sa simplicité. Préalablement, on ne dispose d'aucune information objective sur la représentativité des interpolations. Si la densité des points d'échantillonnage x_i est "très forte", la représentativité peut être correcte, dans d'autres cas, au contraire les résultats peuvent être très trompeurs.

Cette méthode géométrique, basée sur un principe de bon sens, semble donner le plus de satisfaction aux utilisateurs.

La méthode des distances inverses

Cette méthode présente l'avantage de prendre en considération plus de données du champ d'observation. Elle ne se limite pas au voisinage immédiat du point connu de la variable régionalisée. Chaque point inconnu se verra attribuer une moyenne pondérée par les distances d_i respectives que sépare ce point $Z(x_0)$ des points connus d'une grille (Z_1, Z_2, Z_3, Z_4, Z_5 et Z_6).

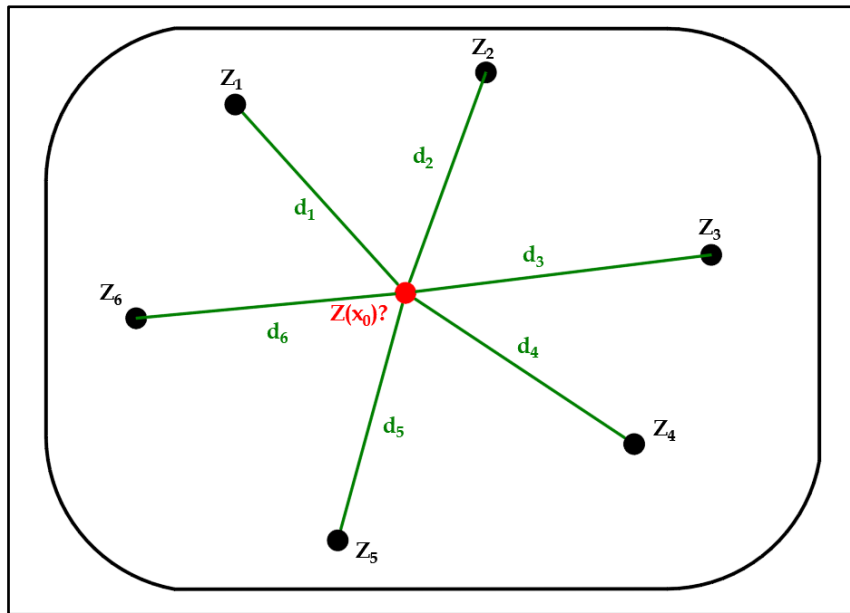


Figure 4-10 : Interpolation par distances inverses

Les facteurs de pondération (poids) sont calculés proportionnellement à l'inverse de la distance (ou à l'inverse de la distance élevée à une puissance). En conséquence un poids plus important est affecté aux sites les plus proches et un poids moindre aux sites les plus éloignés.

La division des inverses des distances d_i ($i=1,2,3,4,5$ et 6) par la somme des inverses des distances sert à créer des poids λ_i définis comme suit :

$$\lambda_i = \frac{1/d_i}{\sum_{i=1}^n 1/d_i} \quad \text{et} \quad Z^*(x_0) = \sum_{i=1}^n \lambda_i Z_i \quad \text{avec} \quad \sum_{i=1}^n \lambda_i = 1 \quad (4-15)$$

Pour que l'estimation soit non biaisée, on cherche toujours à créer des poids λ_i dont la somme doit être toujours égale à 1, de sorte que, dans le cas particulier où la variable aléatoire est toujours égale à la même valeur, l'estimation redonnera cette même valeur.

Cette méthode d'interpolation fournit une surface continue et des courbes relativement "lisses". La pondération étant positive, les valeurs interpolées sont limitées par les valeurs minimales et maximales du champ d'observation.

Cependant cette méthode présente quelques limites, à savoir :

- Elle est indifférente à la configuration géométrique des observations. Seule la distance compte.

- Elle tend à surpondérer les données groupées alors qu'elles sont redondantes.
- Elle induit parfois des changements assez importants dans l'allure des courbes iso-valeurs en raison de l'influence du point le plus proche.

4.5.4.2 Méthodes d'interpolation stochastiques (cas du Krigeage)

Le Krigeage (baptisé ainsi en l'honneur d'un des précurseurs de la géostatistique, Daniel Krige, 1951) est une méthode d'interpolation spatiale qui prévoit la valeur d'une variable en des sites non échantillonnés. Elle utilise une combinaison linéaire sans biais et à variance minimale des observations du phénomène en des sites voisins, en prenant en compte aussi bien leur configuration géométrique que la structure spatiale de la régionalisation. (Matheron, 1962).

Le fondement de base du Krigeage s'appuie sur une modélisation probabiliste du phénomène étudié. L'interpolation par Krigeage est la première méthode d'interpolation spatiale qui tient compte de la dépendance spatiale.

L'intérêt des méthodes d'interpolations géostatistiques, par rapport aux méthodes d'interpolations mathématiques déterministes, est l'estimation précise et sans biais de la grandeur recherchée, d'une part, et d'autre part, le calcul de précision de l'estimation à l'aide d'une variance de l'erreur de l'estimation (Emery, 2001).

Cette erreur est définie comme suit :

Erreur d'estimation, $e = Z^*(x_0) - Z(x_0)$

$Z^*(x_0)$ estimateur de $Z(x_0)$ au point x_0 .

Le Krigeage dépend de la forme de la tendance de $Z(x)$, on distingue :

- ☞ Le Krigeage ordinaire, si l'espérance mathématique $E(Z(x)) =$ constante m inconnue ;
- ☞ Le Krigeage simple, si l'espérance mathématique $E(Z(x)) =$ constante m connue ;
- ☞ Le Krigeage universel, si l'espérance mathématique $E(Z(x))$ est une combinaison linéaire de fonctions connues.

Dans les 3 types de Krigage, l'estimation de la valeur de Z en un point $x_0 : Z^*(x_0)$, satisfait les quatre contraintes suivantes :

- 1- Contrainte de linéarité : L'estimée doit être une combinaison linéaire des données $Z^*(x_0) = \sum_{i=1}^n \lambda_i Z_i$;
- 2- Contrainte d'autorisation : l'espérance mathématique $E[Z^*(x_0) - Z(x_0)]$ et la variance $Var[Z^*(x_0) - Z(x_0)]$ doivent obligatoirement exister. Cette contrainte n'intervient que dans le cas où la fonction aléatoire du modèle de base de $\gamma(h)$ est stationnaire intrinsèque ;
- 3- Contrainte de non-biais : L'espérance de l'erreur de prévision est nulle $E[Z^*(x_0) - Z(x_0)] = 0$;
- 4- Contrainte d'optimalité : Les poids λ_i doivent être estimés de manière à minimiser la variance de l'erreur de prévision $Var[Z^*(x_0) - Z(x_0)]$.

Or, la variance d'estimation n'est autre que la variance de l'erreur d'estimation et la valeur minimale de la variance de prédiction est appelée variance de Krigage, elle est notée σ_e^2 .

Ainsi, on obtient :

$$\sigma_e^2 = Var(Z^*(x_0) - Z(x_0)) = Var(Z^*(x_0)) + Var(Z(x_0)) - 2Cov(Z^*(x_0), Z(x_0)) \quad (4-16)$$

Si on remplace $Z^*(x_0)$ par son expression $\sum_{i=1}^n \lambda_i Z_i$ de la contrainte 1 de linéarité, la variance d'estimation devient une expression du covariogramme :

$$\sigma_e^2 = Var(Z(x_0)) + \sum_{i=1}^n \sum_{j=1}^n \lambda_i \lambda_j Cov(Z_i, Z_j) - 2 \sum_{i=1}^n \lambda_i Cov(Z_i, Z(x_0)) \quad (4-17)$$

Sachant que $C(h) = \sigma_e^2 - \gamma(h)$ et que $\sum_{i=1}^n \lambda_i = 1$, la variance d'estimation peut s'exprimer en terme du variogramme comme suit :

$$\sigma_e^2 = Var(\sigma^2 - \gamma(x_0, x_0)) + \sum_{i=1}^n \sum_{j=1}^n \lambda_i \lambda_j (\sigma^2 - \gamma(x_i - x_j)) - 2 \sum_{i=1}^n \lambda_i (\sigma^2 - \gamma(x_i, x_0)) \quad (4-18)$$

Les poids (λ_i) sont solutions du système de Krigage résultant de ces 4 contraintes. Le principe consiste donc à choisir des pondérations λ_i correspondant à une erreur d'estimation σ_e^2 minimale. Mathématiquement, la solution est obtenue par la

minimisation d'une fonction quadratique (convexe) sous contrainte d'égalité $\sum_{i=1}^n \lambda_i = 1$, en se servant de la méthode de Lagrange.

Dans le cas de stationnarité de second ordre, on note l'existence de deux types de Krigeage. La différence entre eux réside dans la connaissance de la moyenne de la variable du champ à estimer. Nous distinguons :

- ☞ **Krigeage simple** : s'applique sur une variable aléatoire de moyenne connue.
- ☞ **Krigeage ordinaire** : s'applique sur une variable aléatoire de moyenne inconnue ;

La version la plus courante est le Krigeage ordinaire, car dans les applications pratiques, il est rare que l'on connaisse avec certitude la valeur de la moyenne.

A- Krigeage simple (KS)

Dans ce type de Krigeage noté (KS), on suppose que Z est d'espérance mathématique nulle.

Parfois on connaît la moyenne m du champ à estimer ou du moins on en possède une estimation fiable. Alors sans imposer la contrainte que la somme des poids λ_i soit égale à 1, on peut former un estimateur sans biais.

$$Z^*(x_0) = \sum_{i=1}^n \lambda_i Z_i + \left[1 - \sum_{i=1}^n \lambda_i \right] m \quad (4-19)$$

En remplaçant $Z^*(x_0)$ par son expression $\sum_{i=1}^n \lambda_i Z_i$ la variance d'estimation devient une expression de la semi-variance :

$$\sigma_e^2 = Var(Z(x_0)) + \sum_{i=1}^n \sum_{j=1}^n \lambda_i \lambda_j Cov(Z_i, Z_j) - 2 \sum_{i=1}^n \lambda_i Cov(Z_i, Z(x_0)) \quad (4-20)$$

Sans poser de contrainte, en dérivant σ_e^2 par rapport à chacun des poids λ_i , on aboutit au système de Krigeage simple :

$$\sigma_{ks}^2 = Var(Z(x_0)) - \sum_{i=1}^n \lambda_i Cov(Z_i, Z(x_0)) \quad (4-21)$$

B- Krigeage ordinaire (KO)

Le krigeage simple supposant que l'espérance de la fonction aléatoire soit connue est rarement vérifiée. Cette méthode a donc été généralisée au cas où l'espérance est inconnue et constante sur tout le champ, c'est-à-dire sur le voisinage de krigeage. Il s'agit du krigeage ordinaire (Matheron, 1970), appelé aussi par certains auteurs Krigeage ponctuel. Noté *KO*, il est le plus fréquemment utilisé (Gratton, 2002). Si on suppose maintenant que la variable Z est d'espérance m inconnue, la condition de non biais s'écrit : $m \sum_{i=1}^n \lambda_i = 0$.

On forme le Lagrangien $L(\lambda)$ comme suit :

$$L(\lambda) = \sigma_e^2 + 2\mu \left[\sum_{i=1}^n \lambda_i - 1 \right] \quad (4-22)$$

Où μ est le multiplicateur de Lagrange. Le minimum est atteint lorsque toutes les dérivées partielles par rapport à chacun des poids λ_i et par rapport à μ s'annulent. Selon cette condition lagrangienne, les poids λ_i apparaissent alors comme solution du système linéaire suivant :

$$\begin{aligned} \sum_{j=1}^n \lambda_j \gamma(x_i, x_j) + \mu &= \gamma(x_i, x_0), \quad (i = 1 \dots n) \\ \sum_{j=1}^n \lambda_j &= 1 \end{aligned} \quad (4-23)$$

La variance d'estimation (σ_e^2) minimale est appelée variance de Krigeage σ_{KO}^2 .

Elle est donnée par la relation :

$$\sigma_{KO}^2 = \sum_{i=1}^n \lambda_i \gamma(x_0, x_i) - \gamma(x_0, x_0) - \mu \quad (4-24)$$

Remarque : Cette variance de Krigeage ne dépend pas des valeurs observées. Elle ne dépend que du variogramme et de la configuration des points servant à l'estimation par rapport au point (ou bloc) à estimer (Marcotte, 2013).

Le problème s'exprime finalement sous la forme d'un système de $n+1$ équations linéaires à $n+1$ inconnues résolues sous la forme matricielle : $Ax = b$

$$A = \begin{bmatrix} \gamma(x_1, x_1) & \gamma(x_1, x_2) & \cdots & \gamma(x_1, x_n) & 1 \\ \gamma(x_2, x_1) & \gamma(x_2, x_2) & \cdots & \gamma(x_2, x_n) & 1 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ \gamma(x_n, x_1) & \gamma(x_n, x_2) & \cdots & \gamma(x_n, x_n) & 1 \\ 1 & 1 & \cdots & 1 & 0 \end{bmatrix} \quad x = \begin{bmatrix} \lambda_1 \\ \lambda_2 \\ \vdots \\ \lambda_n \\ \mu \end{bmatrix} \quad b = \begin{bmatrix} \gamma(x_1, x_0) \\ \gamma(x_2, x_0) \\ \vdots \\ \gamma(x_n, x_0) \\ 1 \end{bmatrix} \quad (4-25)$$

Les $\gamma(x_i, x_j)$ sont les valeurs du variogramme qui correspondent à la distance h_{ij} , entre les points x_i et x_j qui ont été déjà calculés, tandis que les $\gamma(x_i, x_0)$ sont calculés à l'aide de la fonction analytique qui a été ajustée aux points du variogramme expérimental. Pour respecter la condition de non biais de la solution unique du système, on vérifie toujours que la somme des poids λ_i est égale à 1. Cette dernière contrainte introduit un degré de liberté supplémentaire dans le problème. Ce dernier est introduit en ajoutant le multiplicateur de Lagrange μ , dans le but de minimiser l'erreur d'estimation.

Remarques :

- 1- La variance de krigeage simple est toujours inférieure à la variance de krigeage ordinaire ($\sigma_{KS}^2 \leq \sigma_{KO}^2$) car on n'a pas besoin d'imposer de contrainte sur les poids λ_i .
- 2- Le krigeage ordinaire (KO) : seule l'hypothèse intrinsèque est requise. La stationnarité d'ordre 2 n'est pas nécessaire.
- 3- Le krigeage simple (KS) : la stationnarité est nécessaire. Ainsi, il n'est pas possible d'effectuer un krigeage simple si le variogramme ne présente pas de palier.
- 4- En pratique le Krigeage ordinaire (KO) et le Krigeage simple (KS) sont similaires à courte distance.
- 5- En général le Krigeage ordinaire (KO) est préférable au Krigeage simple (KS)

C- Propriétés du Krigeage

Les principales propriétés et caractéristiques associées au krigeage sont :

- 1- Contrairement aux autres méthodes d'interpolation mathématiques, le Krigeage tient compte d'une continuité du phénomène étudié (variogramme omnidirectionnel et directionnel, effet de pépité et anisotropie).
- 2- Le Krigeage est un interpolateur linéaire, sans biais conditionnel, il utilise un estimateur plus lisse que la valeur à estimer. En d'autres termes, les estimations par Krigeage sont moins variables que les valeurs réelles de la variable étudiée.
- 3- Comme pour les autres méthodes d'interpolation mathématiques, le Krigeage accorde aux points les plus proches les poids les plus importants. Cette propriété s'appelle effet d'écran. Or celui-ci va dépendre de la configuration et du modèle du variogramme utilisé, de sa portée et de son effet de pépité. Plus l'effet de pépité est important, moins il y a d'effet d'écran.
- 4- Le Krigeage est un interpolateur exact, puisqu'il retourne la valeur de l'échantillon au point échantillonné. C'est-à-dire si l'on estime un point mesuré, on retrouve la valeur mesurée.
- 5- La plus importante propriété du Krigeage est qu'il présente une variance minimale d'estimation par construction.

D- Voisinage utilisé pour le krigeage

Le voisinage de Krigeage est le champ spatial qui contient le point (x_0) non mesuré et qu'on veut estimer par interpolation et les points (x_i) mesurés qu'on utilise pour l'estimation de $Z(x_0)$. Pour aboutir à une estimation précise il faut :

- 1- Utiliser un voisinage glissant ou l'interpolation ne tient compte que des points mesurés ayant un lien de voisinage avec celui à estimer.
- 2- Utiliser un nombre de points suffisant (>10 ; peut atteindre jusqu'à 50-100).
- 3- Définir un rayon de voisinage qui tient compte de la nature de stationnarité et de l'échelle de définition du modèle variographique. S'il y a anisotropie géométrique, le cercle est substitué par une ellipse parallèle à la direction de meilleure continuité.

La taille du cercle et du nombre de voisinage doit être testé par validation croisée pour le choix de la meilleure combinaison de voisinage de Krigeage.

4.5.5 Validation croisée

La validation croisée est une pratique pour vérifier et valider tous les paramètres choisis pour décrire la continuité spatiale, à savoir : le modèle théorique des variogrammes omnidirectionnels et directionnels et les caractéristiques isotropiques de la structure spatiale. Par conséquent elle doit être réalisée avant d'entreprendre le Krigeage.

La validation croisée permet donc de valider le choix manuel des paramètres variographiques (le pas de calcul, l'étendu de calcul, la tolérance sur les angles, la tolérance sur le pas de calcul, l'axe d'anisotropie, le facteur d'anisotropie et les paramètres d'ajustement du modèle théorique tel que l'effet de pépite, la portée et le seuil).

Son principe est le suivant :

- 1- Éliminer à tour de rôle chaque observation et l'estimer à l'aide de ses voisins ;
- 2- Comparer les vraies valeurs aux valeurs estimées par un modèle de régression pour vérifier si le modèle théorique choisi se comporte comme prévu et si le voisinage utilisé est adéquat ;
- 3- Une bonne pratique de Krigeage se traduit généralement par une droite de régression passant par l'origine en formant un angle de 45° avec les deux axes des abscisses et des ordonnées positifs (se rapprochant de la première bissectrice) et par un coefficient de corrélation élevé.

Un modèle et un voisinage adéquats devraient fournir :

- 1- Une erreur de prédiction faible ;
- 2- Des coefficients de corrélation et de détermination proches de 1 ;
- 3- L'ordonnée à l'origine et le coefficient d'ajustement de la droite de régression proches de zéro et de l'unité respectivement.

4.6 Application au levé topo-bathymétrique du barrage Hammam Debagh

4.6.1 Mise en œuvre de l'analyse géostatistique

Tel que l'indique (Isaaks & Srivastava, 1989), l'analyse géostatistique a été exécutée en quatre étapes successives (Fig. 4-11) : l'analyse exploratoire, l'analyse variographique, la validation croisée et le Krigeage qui sont développées dans les paragraphes suivants.

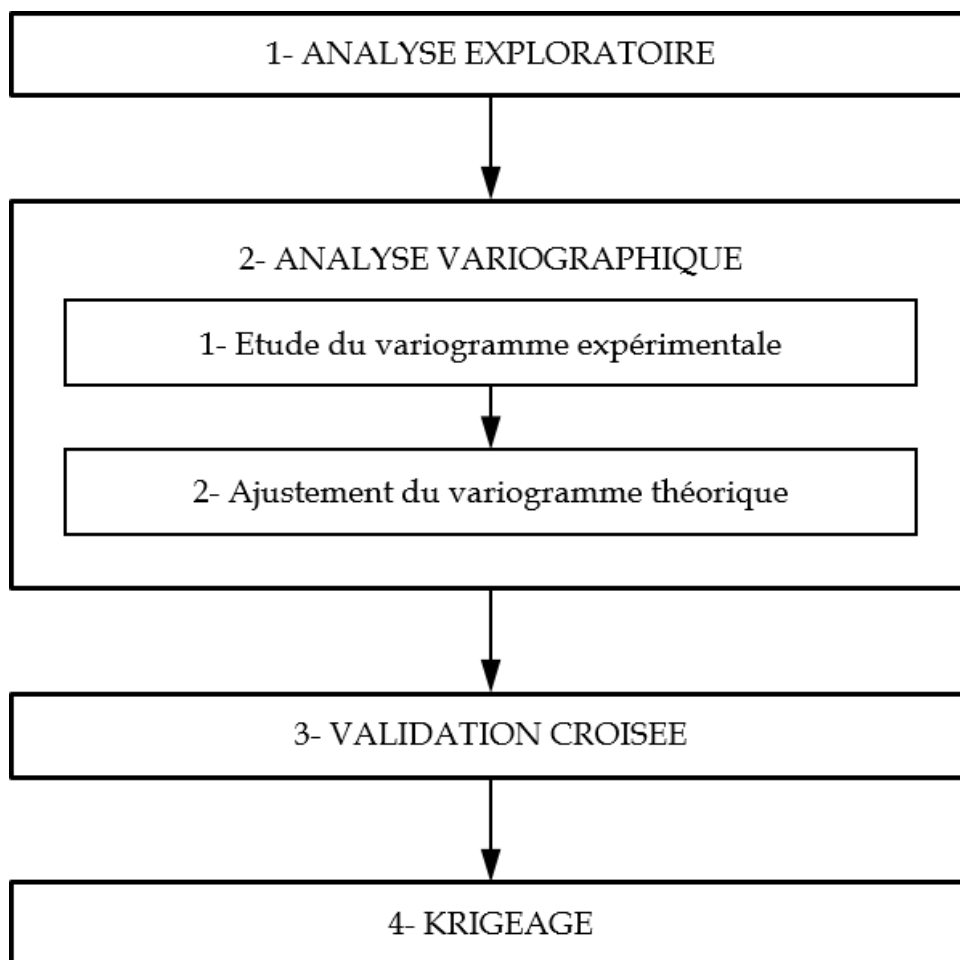


Figure 4-11 : Procédure de l'analyse géostatistique

4.6.1.1 Analyse exploratoire

La première étape pour toute analyse géostatistique repose sur l'analyse des données par les statistiques descriptives. Cette analyse permet de fournir l'information pertinente sur la nature de la distribution statistique des données dans l'espace et d'appréhender leur degré d'homogénéité et de normalité.

Li et Heap (2008) rapportent que la normalité ou la distribution des données d'une façon normale autour de leurs statistiques centrales peut influencer certaines méthodes d'interpolations spatiales. Les données spatiales sont normales si les moyennes, les médianes et les modes sont très voisins rendant ainsi l'analyse géostatistique faisable. Dans le cas contraire, on a souvent recours à une transformation mathématique de la variable d'intérêt en une forme proche de la distribution gaussienne. Cette transformation devrait, en principe, réduire la dissymétrie (mesurée par le coefficient d'asymétrie « skewness ») et le degré d'aplatissement (mesuré par le kurtosis).

Si le coefficient d'asymétrie « skewness » est positif, la dissymétrie est à droite et les trois statistiques de tendance centrales sont réparties ainsi : Mode < médiane < moyenne.

Si le coefficient d'asymétrie « skewness » est négatif, la dissymétrie est à gauche avec moyenne < médiane < mode.

Pour interpréter la valeur absolue du coefficient d'asymétrie « skewness », Bulmer (1979) a donné cette règle générale :

$0 < | \text{coefficient d'asymétrie} | < 0.5$: assez symétrique ;

$0.5 < | \text{coefficient d'asymétrie} | < 1$: dissymétrie moyenne ;

$| \text{coefficient d'asymétrie} | > 1$: forte dissymétrie.

Si le degré d'aplatissement « Kurtosis » est positif, la courbe de distribution est leptokurtique, c'est-à-dire plus pointue que la courbe normale ou gaussienne.

Si le degré d'aplatissement « Kurtosis » est négatif, la courbe est platykurtique, c'est-à-dire plus aplatie et plus arrondie que la courbe normale.

En règle générale : $0 < | \text{degré d'aplatissement} | < 0.5$: assez mésokurtique ou normale ;

$0.5 < \text{degré d'aplatissement} < 1$: courbe moyennement sur-gaussienne ;

Degré d'aplatissement > 1 : courbe fortement sur-gaussienne ;

$-1 < \text{degré d'aplatissement} < -0.5$: courbe moyennement sous-gaussienne ;

Degré d'aplatissement < -1 : courbe fortement sous-gaussienne.

Comme le Krigeage nécessite des données normalement distribuées (Sepaskhah et al., 2005), l'asymétrie et l'aplatissement, en particulier, seront examinés.

4.6.1.2 Analyse variographique

Pour explorer la structure de la régionalisation, notamment ses propriétés de stationnarité et de régularité afin de mettre en évidence la présence d'une structure spatiale (Robinson & Metternicht, 2006), les résultats sont exprimés sous la forme d'un semi-variogramme expérimental formé d'un ensemble de points de semi-variance $\gamma(h)$ de la variable d'intérêt Z en fonction de la distance h qui sépare les observations.

En pratique, le semi-variogramme expérimental est ajusté par des fonctions mathématiques en lui assimilant un semi-variogramme théorique de forme sphérique, exponentielle, linéaire ou gaussienne selon son allure et son ajustement. Comme le montre la Figure 4-12, quatre paramètres caractérisent un semi-variogramme théorique (Smith et al., 1993 ; Srivastava, 2013) :

- L'effet de pépite (C_0) qui décrit le comportement du variogramme à l'origine. Un C_0 différent de zéro est due aux erreurs de localisation ou de mesure.
- La portée (A_0) qui représente la distance d'autocorrélation de la variable d'intérêt.
- Le palier (ou seuil) ($C_0 + C$) qui correspond à la valeur autour de laquelle se stabilise le variogramme, il s'agit de la variance totale de la variable d'intérêt.

D'autre part trois indices statistiques de la qualité d'ajustement des modèles théoriques de variographie sont estimés, il s'agit :

- Du RSS (la somme des carrés des résidus ou l'erreur quadratique globale entre les données mesurées et les valeurs prédites par le modèle) ;
- De R^2 (coefficient de détermination pour cet ajustement) ;
- Du rapport $C/(C_0 + C)$ qui donne la proportion de la variance structurée (C) par rapport à la variance totale ($C_0 + C$) (Arnaud & Emery, 2000).

Le modèle théorique est d'autant plus précis, structuré et représentatif d'une autocorrélation spatiale que sa valeur de RSS s'approche de zéro, son R^2 et son rapport $C/(C_0 + C)$ s'approchent de l'unité.

4.6.1.3 Krigeage

L'interpolation stochastique par Krigeage est effectuée pour estimer les valeurs en des points non observés par une combinaison linéaire sans biais des valeurs mesurés. Par construction mathématique, le Krigeage minimise la variance sur l'erreur d'estimation. Le Krigeage ordinaire est la technique utilisée.

Le Krigeage tient compte aussi de la taille du champ à estimer, de la position des points entre eux et de la continuité du phénomène mesuré.

4.6.1.4 Validation croisée

La validation croisée (VC) est utilisée pour évaluer la qualité du modèle de semi-variogramme choisi et générer un ensemble de points mesurés et interpolés qui sont utilisés dans une régression linéaire sous la forme $y = ax + b$. Cinq paramètres sont déterminés à partir de la validation croisée :

- La pente de la droite de régression (a) ;
- L'ordonnée à l'origine (b) ;
- Le coefficient de détermination (R^2) ;
- L'erreur standard (SE) du coefficient de régression ;
- L'erreur de prédiction ($SE_{préd}$) définie par :

$$SE_{préd} = SD \times (1 - R^2)^{0.5} \quad (4-22)$$

Où SD est l'écart-type des données mesurées.

Marcotte (2013) distingue trois situations :

- 1- $SE_{préd}$ est proche de 1. La variance de Krigeage prédit correctement la précision des estimations.
- 2- $SE_{préd}$ est inférieure à 1. Le modèle variographique ajusté décrit un phénomène plus variable que le phénomène réel. La variance de Krigeage surestime la variance des erreurs.
- 3- $SE_{préd}$ est supérieure à 1. Le modèle variographique ajusté décrit un phénomène moins variable que le phénomène réel. La variance de Krigeage sous-estime la variance des erreurs.

Selon Robertson (2000) l'interpolation est d'autant plus précise et de bonne qualité que la pente (a) de sa droite de régression, son erreur de prédiction ($SE_{préd}$) et son coefficient de détermination (R^2) s'approchent de l'unité et que son intercepte (b) ainsi que son erreur standard du coefficient de régression (SE) s'approchent de zéro.

4.6.2 Méthodologie

Sur l'ensemble des 32164 valeurs originales du LTB, nous avons effectué 19 combinaisons aléatoires pour réduire la taille du levé. A cet effet, nous avons utilisé une macro de choix aléatoire développée sous le progiciel Microsoft Excel 2003. En conservant une bonne couverture de la retenue du barrage, ces choix aléatoires étaient de 400 jusqu'à 30000 points. Nous considérons donc un LTB de référence et 19 scénarios sur lesquels nous avons réalisé des analyses géostatistiques complètes en utilisant le logiciel Gamma Design Software, GS+ version 5.1. Une analyse exploratoire suivie d'une analyse variographique et une validation croisée ont été réalisées pour le LTB de référence et les 19 échantillons.

Une première optimisation des paramètres du semi-variogramme expérimental de l'ensemble des données du LTB de référence (32164 points) a été réalisée en commençant par l'histogramme de fréquence, puis la variographie en faisant varier le pas de calcul et l'étendue de calcul (active lag-distance). Le modèle théorique d'ajustement a été choisi selon le coefficient de détermination R^2 le plus élevé et la somme des carrés des écarts (RSS) la plus faible (Cressie, 1993). Pour valider ces choix, nous avons appliqué le modèle de la validation croisée en utilisant un Krigeage ordinaire avec un voisinage fixe de 64 points (choisi après plusieurs essais) et un rayon égal à la portée. L'histogramme des fréquences, le modèle variographique (γ) ainsi sélectionné et sa validation croisée (VC) seront notés : $\gamma_{réf}$ et $VC_{réf}$. Une deuxième analyse géostatistique a porté sur les 19 autres scénarios et les résultats obtenus ont été comparés à ceux de l'analyse géostatistique du LTB de référence.

Pour comparer les mêmes données spatiales à différentes tailles, Hartkamp et al. (1999) recommandent de maintenir le même pas de calcul (lag), la même étendue de calcul (active lag-distance) et d'ajuster le même modèle théorique (γ). La comparaison de

tous les paramètres de la géostatistique a permis d'en déduire la taille optimale d'un LTB. Cette taille a été validée par la constance d'estimation du volume d'eau au niveau normale de la retenue (NNR) et au niveau des plus hautes eaux (PHE). A cet effet, les modèles numériques de terrain (MNT) réalisés par Krigeage dans le logiciel GS+ ont été transférés vers le logiciel Global Mapper V17.0 (Blue Marble Geographics, 2015) dans lequel le volume a été calculé à différents niveaux du barrage, depuis le fond du réservoir jusqu'au niveau des plus hautes eaux (PHE), pour obtenir la courbe capacité-hauteur du réservoir. Cette courbe est d'une grande importance, car elle est utilisée pour gérer les lâchers du barrage.

Pour s'assurer que le choix aléatoire n'affecte pas cette optimisation, nous avons refait le tirage aléatoire 12 fois pour cette même taille optimale et nous avons vérifié la constance d'estimation de ces mêmes volumes au NNR et au PHE.

4.6.3 Résultats et discussions

4.6.3.1 Analyse exploratoire

Dans le [Tableau 4-1](#) sont consignées les statistiques descriptives pour les prédictions spatiales des altitudes topo-bathymétriques au niveau de la cuvette du barrage. La distribution spatiale est relativement uniforme aussi bien pour l'ensemble de données du LTB original que pour les 19 échantillons analysés, avec un coefficient de variation (*CV*) d'environ 6,3%. Cette variation d'altitudes au niveau du réservoir de « Hammam Debagh » est proche de celle de « Rocky Gorge » (6,4%) ; (MGS, 2016), mais différente de celle de « Triadelphia » (4,3%) (MGS, 2016) et de celle de « Foum El Gharza » (2,1%) (ANBT, 2015).

L'amplitude du coefficient de variation (*CV*) est proportionnelle à la déclivité du réservoir. Par conséquent, le coefficient de variation (*CV*) du réservoir du barrage de Foum El Gharza est très faible, ce qui correspond à un barrage typique du Sahara. Plus le coefficient de variation (*CV*) est faible, meilleures sont les performances des méthodes d'interpolation spatiale.

Les moyennes, médianes et modes ne varient pas beaucoup d'un échantillon à l'autre. Cependant, pour l'ensemble des données du LTB et tous les échantillons, nous

observons que : moyenne < médiane < mode. L'écart entre la moyenne et le mode de l'ensemble des données du LTB est de 11,63 m et celui entre la moyenne et la médiane est de 7,65 m. Ces écarts impliquent une légère dissymétrie négative.

Pour les échantillons, la même dissymétrie négative est observée pour 2000 points ou plus. En outre, l'asymétrie de ces échantillons est entièrement négative avec une valeur absolue proche de celle de l'ensemble des données du LTB (0,40). Cette valeur est inférieure à 0,5, ce qui correspond à la limite établie par [Bulmer \(1979\)](#) pour une faible asymétrie. De même, le kurtosis normalisé des échantillons de 2000 points ou plus est négatif (platykurtique), montrant un léger aplatissement du même ordre de grandeur que celui de l'ensemble des données du LTB (-0,41). Par conséquent, ces statistiques descriptives des 19 échantillons ne varient pas beaucoup de celles de l'ensemble des données du LTB. Les échantillons de plus de 2 000 points semblent suffisants pour garantir les mêmes paramètres de la distribution normale du LTB. De plus, la distribution de l'ensemble des données du LTB est proche de celle de Gauss, puisque l'asymétrie (Skewness) et l'aplatissement (kurtosis normalisé de Fisher) sont proches de zéro. Ceci est une indication que le krigeage fonctionnera convenablement puisque la normalité des données est requise.

Tableau 4-1 : Statistiques descriptives de l'ensemble de données du LTB (32164 points) et les 19 échantillons (400 à 30000 points).

Ech,	Taille	Moyenne	Mode	Médiane	E,type	Variance	Min	Max	Skew ¹	Kurt ²	CV
		(m)	(m)	(m)	(m)		(m)	(m)			%
1	400	348,16	359,8	356,13	22,04	485,94	300,32	415,24	-0,45	-0,44	6,13
2	1000	349,00	356,57	356,20	21,90	479,71	298,53	428,28	-0,41	-0,22	6,28
3	1500	347,23	359,66	355,34	21,66	469,23	298,74	414,20	-0,17	-0,60	6,24
4	2000	347,74	359,66	355,37	22,05	486,20	298,80	424,54	-0,33	-0,43	6,34
5	4000	348,03	359,62	355,78	22,49	505,87	298,63	418,29	-0,39	-0,42	6,46
6	6000	348,13	359,54	355,73	22,38	500,84	297,97	429,34	-0,40	-0,43	6,43
7	8000	348,02	359,65	355,58	22,21	493,31	297,50	425,54	-0,38	-0,43	6,38
8	10000	348,06	359,68	355,73	22,16	491,19	297,50	429,34	-0,41	-0,44	6,37
9	12000	348,11	359,71	355,71	21,98	482,97	297,50	422,71	-0,43	-0,48	6,31
10	14000	348,09	359,71	355,71	22,01	484,64	297,50	432,02	-0,41	-0,43	6,32
11	15000	347,89	359,73	355,59	22,27	496,01	297,60	432,15	-0,38	-0,44	6,40
12	16000	347,83	359,73	355,65	22,24	494,68	297,60	432,02	-0,41	-0,45	6,39
13	18000	348,16	359,71	355,68	22,03	485,43	297,50	432,15	-0,39	-0,37	6,33
14	20000	348,09	359,68	355,79	22,06	486,77	297,50	432,15	-0,40	-0,39	6,34
15	22000	348,10	359,71	355,79	22,04	485,65	297,60	429,34	-0,42	-0,42	6,33
16	24000	348,13	359,71	355,78	22,16	491,08	297,50	432,15	-0,40	-0,40	6,37
17	26000	348,10	359,71	355,72	22,08	487,74	297,50	432,15	-0,41	-0,40	6,34
18	28000	348,09	359,71	355,74	22,12	489,40	297,50	432,15	-0,40	-0,41	6,35
19	30000	348,14	359,71	355,77	22,10	488,57	297,50	432,15	-0,40	-0,40	6,35
LTB	32164	348,08	359,71	355,73	22,13	489,57	297,50	432,15	-0,40	-0,41	6,33

CV : coefficient de variation ; ¹Skewness : Asymétrie ; ²Kurtosis normalisé : Aplatissement selon Fischer ;

LTB : Levé Topo- Bathymétrique

4.6.3.2 Analyse variographique

- Variogramme de référence ($\gamma_{réf}$)

Le variogramme de référence ($\gamma_{réf}$) est illustré dans la Figure 4-12. Le modèle sphérique donne la meilleure description du variogramme expérimental avec un R^2 élevé de 0,981 et atteint un seuil (stationnarité du second ordre) à une semi-variance de 462,50 m² avec un petit effet de pépite de 11 m². Sur la base des valeurs du modèle sphérique représentées sur la Figure 4-12, le rapport $C/(C_0 + C)$ est de 97,6%, ce qui indique une forte corrélation spatiale jusqu'à une portée (A_0) de 732 m, au-delà de laquelle il n'y a pas d'influence entre les points du LTB. La portée fournit des informations précieuses sur le voisinage dans lequel les points de mesure influencent

l'estimation des élévations dans des endroits non mesurés (Cambardella et al., 1994). La valeur de A_0 est similaire à celle d'un barrage en Ouzbékistan, qui varie entre 500 et 1000 m selon l'orientation du calcul des semi-variances (Rakhmatullaev et al., 2011). Les semi-variances expérimentales oscillent entre 439,7 m² et 488,6 m², mais sont proches de la variance totale de l'ensemble des données du LTB dont la valeur est de 489,57 m². Le seuil du modèle sphérique est très proche de la variance totale. L'effet de pépité est très petit (11 m²) et inférieur au seuil, indiquant ainsi une petite erreur de mesure du LTB. La pépité est obtenue en extrapolant le modèle de semi-variogramme sphérique à une étendue de calcul nulle.

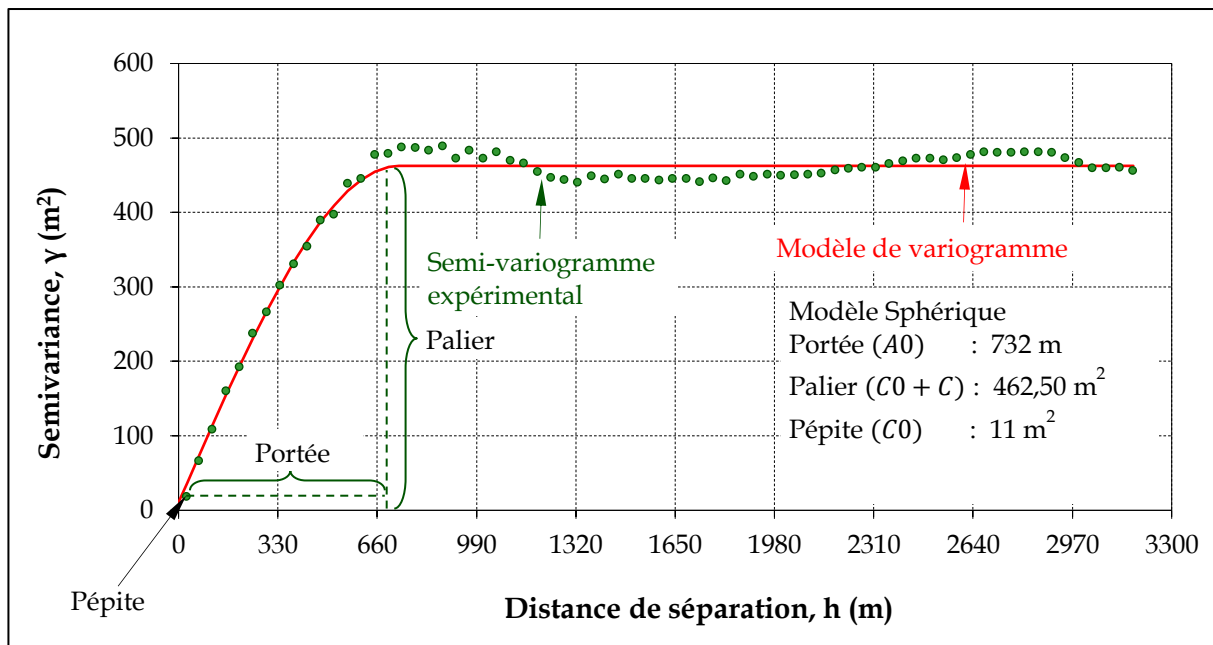


Figure 4-12 : Variogrammes expérimental et théorique

L'anisotropie observée par Rakhmatullaev et al. (2011) n'apparaît pas dans la présente étude. En effet, le diagramme variographique de l'ensemble des données du LTB (Figure 4-13) montre des lignes concentriques d'iso-semi-variance, indiquant que la structure spatiale des élévations au niveau du réservoir du barrage de « Hammam Debagh » est isotrope.

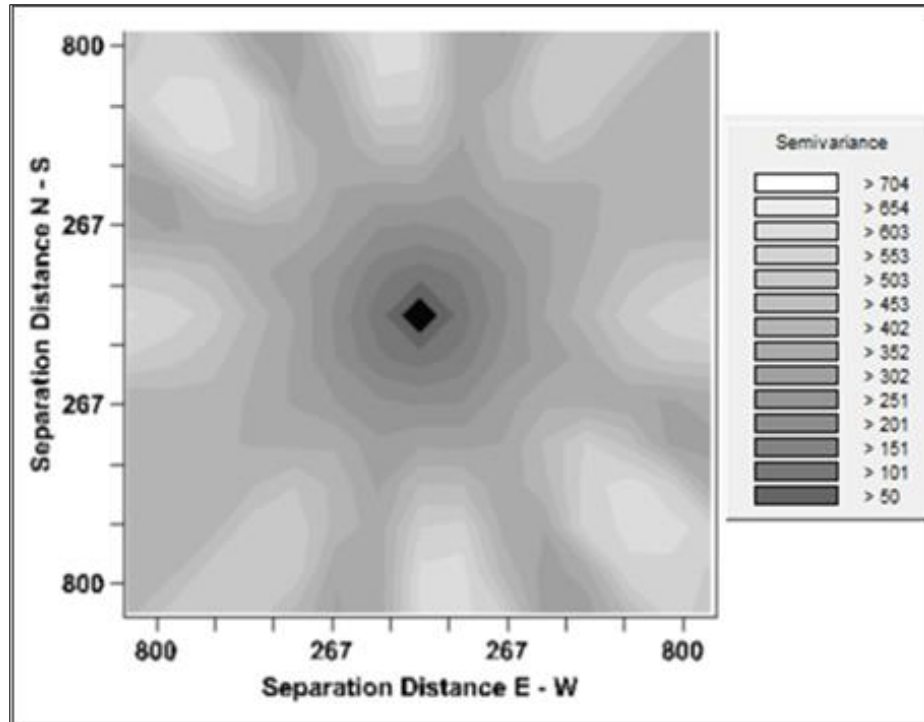


Figure 4-13 : Carte variographique 2D.

Variogrammes des échantillons ($\gamma_{éch}$)

Les variogrammes expérimentaux des 19 sous-ensembles ont été ajustés par le modèle sphérique, en gardant $C_0 = 11m^2$, pour obtenir $\gamma_{éch}$. Les paramètres des variogrammes de tous les échantillons et de $\gamma_{réf}$ sont présentés sur la Figure 4-14. Cette figure montre que les trois critères de qualité d'ajustement des semi-variogrammes : RSS , R^2 et le rapport $C/(C_0 + C)$ atteignent des valeurs quasi constantes proches de celles de $\gamma_{réf}$ lorsque le nombre de points de mesure est au-delà 8 000 points. Cela correspond à une densité de mesure de 11 points ha^{-1} . Avec cette densité, les seuils ($C_0 + C$, Figure 4-14a) et les portées (A_0 , Figure 4-14b) oscillent légèrement autour de ceux de $\gamma_{réf}$.

Li et Heap (2008) ont supposé l'existence d'une densité critique au-dessus de laquelle il n'y a pas d'amélioration de la précision de l'estimation. Les résultats trouvés sont en accord avec les résultats de Li et Heap (2008).

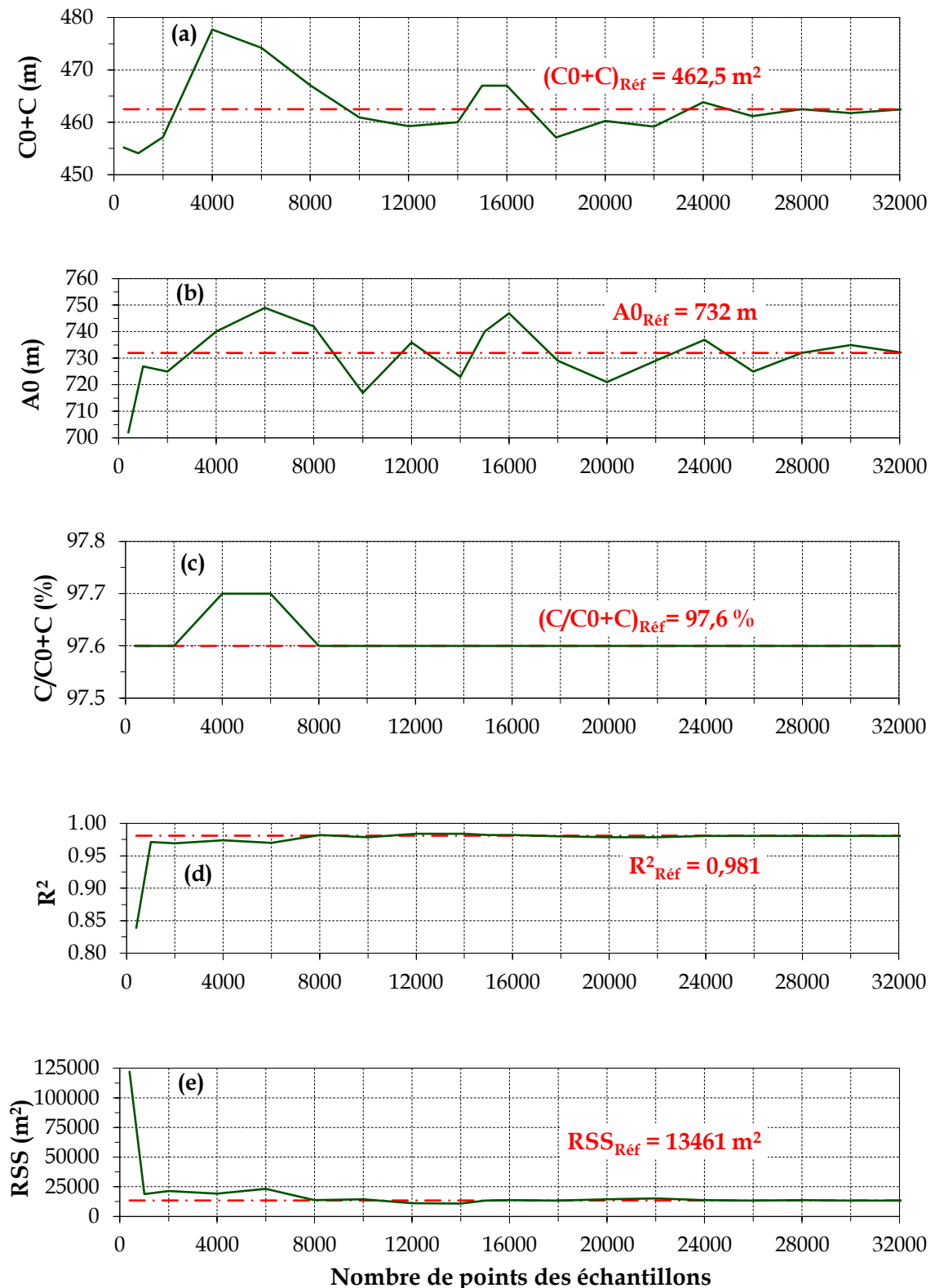


Figure 4-14 : Valeurs des paramètres variographiques pour les 19 échantillons de réduction selon leur taille : (a) Palier, (b) Portée, (c) Fraction structurée de la variance totale, (d) Coefficient de détermination et (e) Somme des carrés des résidus. La ligne pointillée représente les valeurs des paramètres variographiques du LTB de référence.

4.6.3.3 Krigeage et validation croisée

A. Krigeage et validation croisée de référence $(VC)_{réf}$

Comme le $\gamma_{réf}$ repose sur un levé de grande taille, il a permis de décrire correctement la structure spatiale de l'élévation Z . En effet, la validation croisée de l'ensemble de ces données du LTB de référence représentée par une droite de régression linéaire (valeurs mesurées / valeurs estimées) a une pente très proche de l'unité (1,004) et une ordonnée à l'origine proche de zéro (-1,581 m) (Fig. 4-15). Cette droite est presque confondue avec la première bissectrice ($Y = X$) où les valeurs prédites de l'élévation (Z^*) sont égales à celles mesurées (Z).

De ce modèle nous pouvons déduire le degré du lien relationnel (Coefficient de détermination R^2) et les statistiques des erreurs entre Z et Z^* . Ainsi, l'estimation par Krigeage de l'altitude Z^* en tout point de la retenue du barrage est d'autant plus robuste que son R^2 tend vers 1. Puisque le R^2 de la $(VC)_{réf}$ est de 0,998, donc l'estimation par Krigeage en se servant du $\gamma_{réf}$ est considéré très robuste.

La $(VC)_{réf}$ donne une valeur SE de 0,000 (Fig. 4-15), témoignant de l'absence de biais dans l'estimation de l'élévation Z . Ce biais de valeur nulle pourrait être dû à la densité de mesure élevée de l'ensemble de données du LTB (Li & Heap, 2008).

En utilisant le Krigeage comme méthode d'interpolation, Bello-Pineda & Hernández Stefanoni (2007) ont réduit l'erreur d'estimation de la profondeur d'eau dans une zone inondée de 20 à 30% par rapport à d'autres interpolateurs, comme la méthode de la distance inverse. Par conséquent, la combinaison du Krigeage avec les mesures à haute densité peut pratiquement réduire le biais à zéro. Une analyse de la validation croisée par Curtarelli et al. (2015) a montré que l'utilisation du Krigeage ordinaire sur les données du LTB entraîne des erreurs d'estimation réduites et un R^2 élevé.

Quant à la $SE_{préd}$, dite aussi la variance des erreurs standardisées, la $(VC)_{réf}$ prend une valeur proche de 1 (0,875), ce qui prouve, selon Marcotte (2013), que cette variance de Krigeage prédit correctement l'erreur d'estimation.

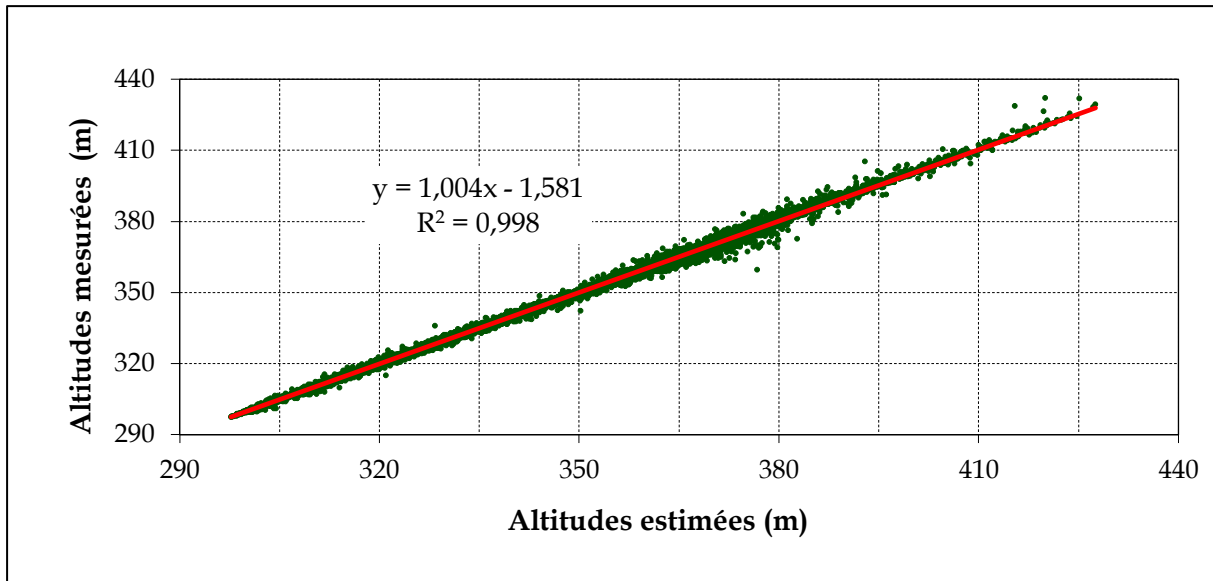


Figure 4-15 : Validation croisée du modèle de LTB de référence (32164 points)

B. Krigeage et validation croisée des échantillons (VC)éch

En prenant comme référence l'ensemble des données (32164 points) du LTB, des valeurs presque maximales sont obtenues pour les critères de fiabilité, à savoir : non-biais ($SE \approx 0$), robustesse ($R^2 \approx 1$) et de précision dans l'estimation de l'erreur de prédiction ($SE_{préd} \approx 1$). Réduire le nombre de points peut ne pas donner de si bons résultats. Par conséquent, il est important d'analyser la sensibilité de ces trois critères de fiabilité pour chacun des 19 échantillons.

Comme le montre la [Figure 4-16c](#), le R^2 s'écarte de l'unité lorsque des échantillons de moins de 10 000 points sont utilisés. De plus, $SE_{préd}$ diminue considérablement ([Fig. 4-16d](#)), ce qui indique que la variance de Krigeage sous-estime de plus en plus la variance d'erreur lorsque la taille des échantillons diminue. Un manque de biais est observé, car SE reste proche de zéro jusqu'à environ 6 000 points ([Fig. 4-16e](#)).

Le passage de 32 164 à 10 000 points a donc un impact négligeable sur les trois paramètres de fiabilité (R^2 , $SE_{préd}$ et SE) et les deux paramètres de régression (a et b) ([Fig. 4-16](#)). Cependant, plus la taille des échantillons, descend à moins de 10 000 points, $SE_{préd}$ devient plus grand et prend des valeurs supérieures à 1. Le variogramme du modèle décrit une estimation Z^* moins variable que la valeur Z observée, c'est-à-dire que la variance de Krigeage sous-estime l'erreur de variance. Ces observations concordent avec celles de [Puenta & Bras \(1986\)](#) qui ont montré que le Krigeage peut

entraîner une sous-estimation ou une surestimation importante de l'erreur de prédiction lorsque la taille d'un ensemble de données diminue.

Cette étape de l'analyse géostatistique montre que les $(VC)_{éch}$ convergent vers $(VC)_{réf}$ à partir d'une taille de levé topo-bathymétrique supérieure ou égale à 10 000 points de mesure. Pour valider le choix de la taille optimale des levés topo-bathymétriques, il est important de compléter cette analyse géostatistique par l'établissement des courbes « volume-hauteur » pour chaque échantillon et la comparaison des volumes d'eau estimés à la côte du niveau normal de la retenue (NNR).

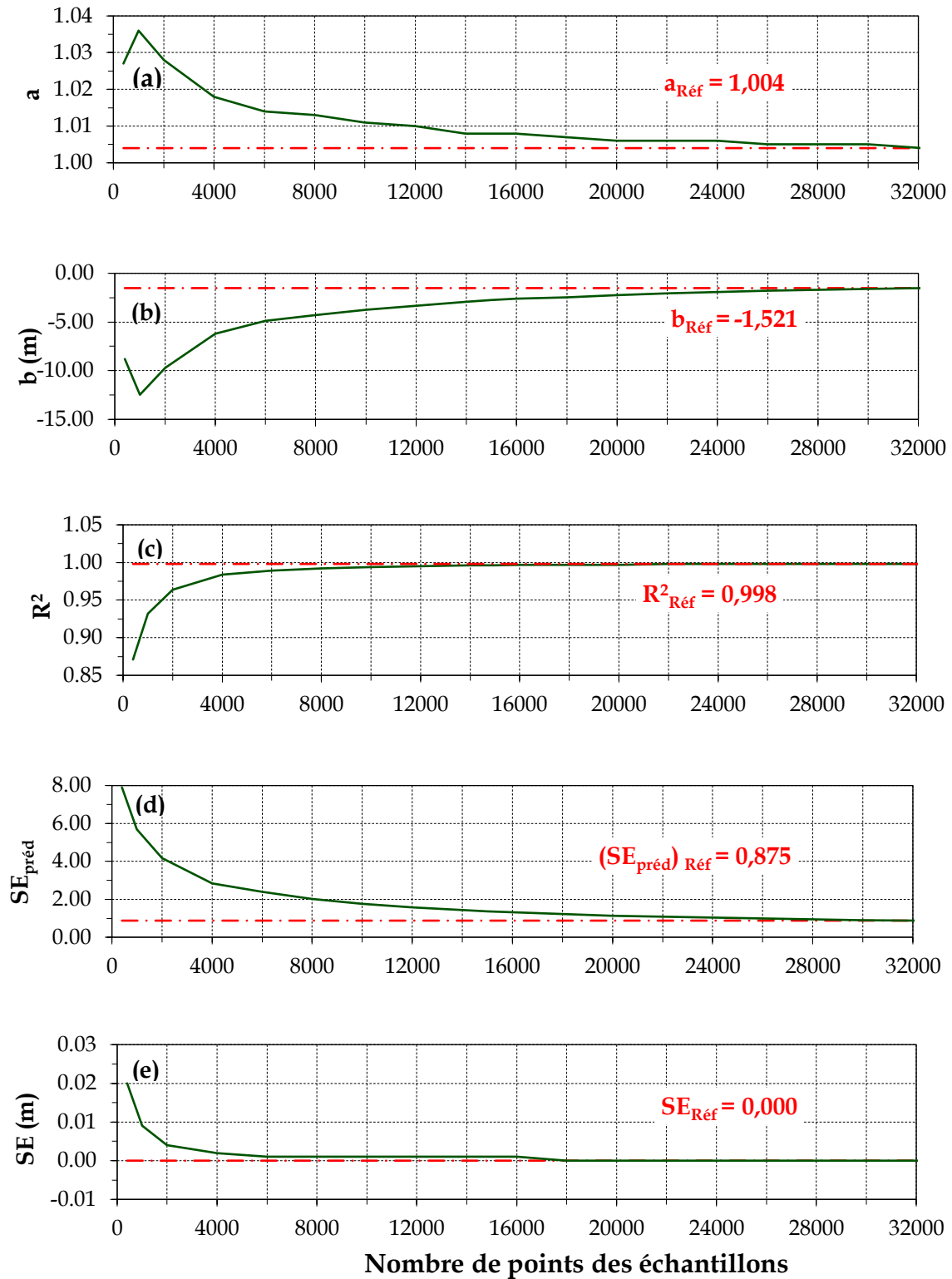


Figure 4-16 : Evolution des paramètres de la validation croisée des 19 échantillons de réduction selon leur taille : (a) coefficient de régression, (b) intercepte, (c) coefficient de détermination, (d) erreur standard de prédiction et (e) erreur standard. La ligne pointillée représente les valeurs des paramètres de validation croisée du LTB de référence.

C. Sensibilité de la courbe volume-hauteur

La Figure 4-17 montre que l'analyse de la sensibilité des volumes en fonction de la taille réduite du levé topo-bathymétrique corrobore les résultats des études géostatistiques. En effet, les tailles de 8 000 et de 10 000 points et plus, ayant montré respectivement une stabilité des paramètres variographiques et une stabilité des critères de la validation croisée, ont aussi montré un palier constant dans l'estimation des volumes d'eau dans la retenue du barrage à la côte NNR.

La réduction qui correspond au maximum de précision des paramètres géostatistiques sans affecter l'estimation du volume d'eau au NNR avoisine les 10 000 points. D'ailleurs, ce palier de stationnarité du volume commence à 6 000 points et il est presque confondu au volume de référence ($179,14 \text{ hm}^3$) quand nous considérons toutes les données du LTB de référence (32 164 points). Ce n'est qu'en dessous de 6 000 points, que nous remarquons une sous-estimation de ce volume d'autant plus importante que la taille du levé diminue. Cette sous-estimation du volume d'eau se traduit par une surestimation du volume des sédiments retenus par le barrage ; Ce qui est probablement dû à la forme concave de n'importe quel réservoir. Ainsi, La diminution du nombre des points du levé surestime le volume des sédiments.

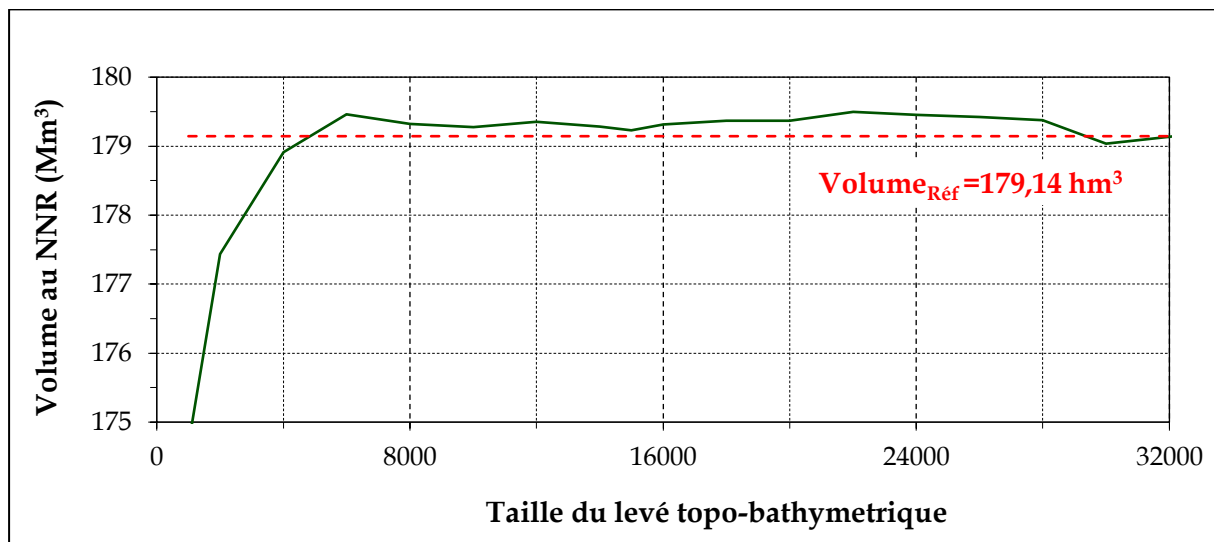


Figure 4-17 : Evolution des volumes estimés au NNR : des 19 échantillons, La ligne pointillée représente le volume estimé au NNR avec le LTB de référence.

La surestimation du volume des sédiments a un impact sur la gestion du réservoir qui est principalement basée sur la courbe « volume-hauteur ».

Cette courbe (Fig. 4-18) définit le volume d'eau dans le réservoir pour les hauteurs allant du fond (297,5 m) au niveau des plus hautes eaux du barrage (PHE ; 370,4 m).

La relation linéaire entre les volumes obtenus avec 10 000 points et ceux de la courbe de référence (32 164 points) a montré que ces deux estimations ne sont pas significativement différentes ($R^2=0,999997$, pente=1,001). La réduction du nombre de points de 32 164 à 10 000 n'a pas affecté significativement la courbe « volume-hauteur » et n'affectera pas donc la bonne gestion de l'eau mobilisée par le barrage.

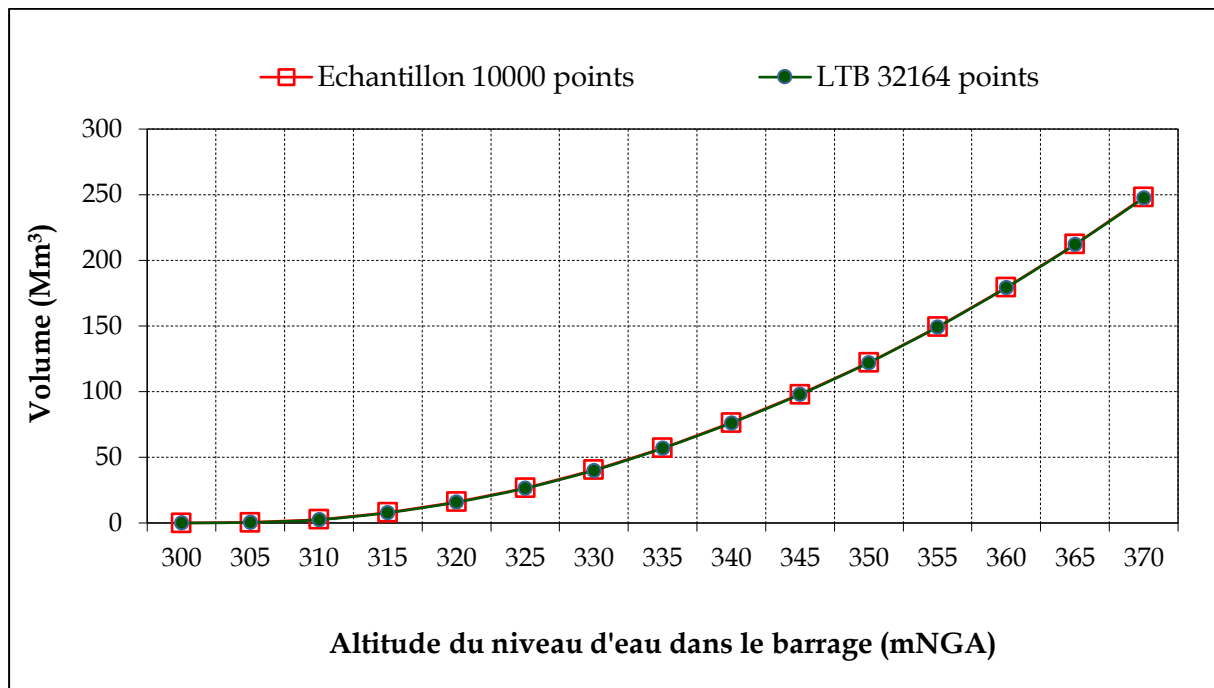


Figure 4-18 : Courbes « volume-hauteur » du LTB de référence et de l'échantillon optimal (10 000 points).

Pour éliminer l'effet du hasard, puisque les estimations précédentes sont basées sur une seule extraction aléatoire de l'échantillon de 10 000 points, l'étude a été répétée pour 11 autres extractions randomisées (soit au total douze échantillons de 10 000 points). Les variations du volume au NNR estimées pour ces derniers sont représentées sur la Figure 4-19. La moyenne des 12 valeurs du volume au NNR (179,16 hm³) n'est pas significativement différente du volume de référence (179,14 hm³).

On peut donc conclure que 10 000 points suffisent pour estimer le volume de vase dans le réservoir du barrage de Hammam Debagh.

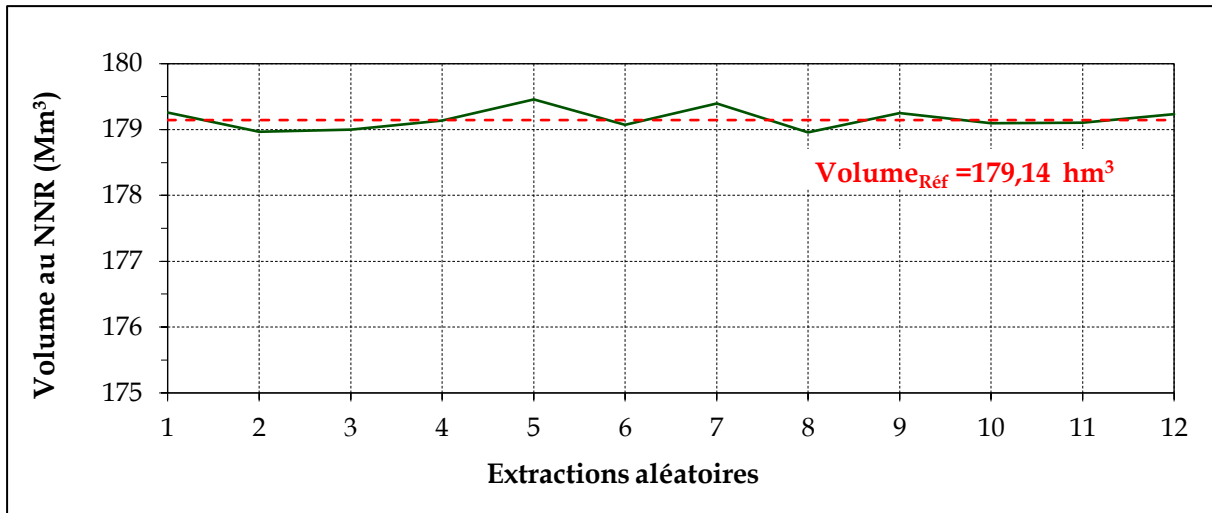


Figure 4-19 : Evolution des volumes estimés au NNR : des 12 combinaisons de taille 10 000 points. La ligne pointillée représente le volume estimé au NNR avec le LTB de référence.

4.7 Conclusion

Après avoir présenté la théorie de la géostatistique et les méthodes d'interpolation déterministes et stochastiques, une étude géostatistique a été réalisée sur le seul levé topo-bathymétrique du barrage « Hammam Debagh » en vue d'en optimiser le nombre de points.

Dix paramètres géostatistiques ont été utilisés pour étudier la ressemblance du LTB du barrage « Hammam Debagh » en Algérie (32 164 points au total) avec 19 échantillons de réduction de la taille dont le nombre de points varie entre 400 et 30 000 points extraits aléatoirement. Les comparaisons de ces paramètres ont montré que la réduction de la taille du LTB de 32 164 points à un optimum de 10 000 points, correspondant respectivement à des densités de mesure de 42 à 13 points ha⁻¹, n'a qu'un faible impact sur les critères de fiabilité géostatistiques.

Toutes les étapes de l'étude géostatistique (analyse exploratoire, analyse variographique et Krigeage/validation croisée) ont montré qu'une taille de mesure de 10 000 points fournit des critères de performance et de fiabilité similaires à ceux obtenus avec les 32 164 points mesurés. La tendance centrale et les caractéristiques de dispersion sont similaires pour ces deux tailles. Cette réduction de taille a maintenu

les trois critères variographiques (RSS , R^2 , le rapport $C/(C0 + C)$) ainsi que le palier et la portée proches de ceux obtenus avec la totalité des points de mesure (32 164 points).

À partir de l'analyse de validation croisée, on peut également conclure que le non-biais (SE), la robustesse (R^2) et l'estimation de l'erreur de prédiction ($SE_{préd}$) sont très similaires.

Cette procédure géostatistique d'optimisation de la taille a été validée en analysant 12 combinaisons de 10 000 points. Les volumes estimés ont été comparés à celui fourni par l'ensemble des données (32 164 points) du LTB.

Cette optimisation du nombre de points permet donc de réduire le coût et les délais de réalisation des levés topo-bathymétriques et d'accroître ainsi leurs fréquences. Cette augmentation de la fréquence des levés est de nature à mieux suivre l'évolution de la sédimentation dans la cuvette du barrage assurant ainsi une meilleure gestion de la ressource mobilisée.

Chapitre 5

CHAPITRE 5 : Optimisation de la gestion des barrages par les réseaux de neurones artificiels

5.1 Introduction et objectifs

Depuis qu'ils ont été proposés pour la première fois par [MCCulloch & Pitts \(1948\)](#), les réseaux de neurones artificiels sont devenus des modèles célèbres et largement utilisés dans pratiquement tous les domaines de la recherche scientifique. Cet engouement vient suite à leur capacité de résoudre des problèmes non linéaires. Ce chapitre vise à tester la possibilité de mettre en place un modèle d'intelligence artificielle pour l'optimisation de la gestion des barrages réservoirs. Pour ce faire, nous avons tenté de vérifier l'efficacité des réseaux de neurones formels pour la prévision des volumes mensuels à allouer. A cet effet, nous avons optimisé le nombre de variables d'entrée, ensuite nous avons sélectionné les meilleurs vecteurs par combinaison de celles-ci. Enfin, douze (12) modèles de réseaux de neurones ont été développés afin d'atteindre cet objectif.

5.2 Modèle des neurones biologiques et formels

Le cerveau humain se compose d'un très grand nombre de neurones biologiques spécialisés dans la transmission de l'information sous forme de signale. Les neurones biologiques se composent de : (i) soma, (ii) les dendrites et (iii) l'axone. Les neurones donc collectent l'information, assurent l'intégration de cette information et envoient l'information vers d'autres neurones. A partir de cette configuration complexe le premier modèle de neurone formel a été proposé. Le premier modèle de neurone formel, proposé par [MCCulloch & Pitts \(1948\)](#), effectue la sommation pondérée des entrées et le résultat est transformé, par la suite, par le biais d'une fonction d'activation ou de transfert vers d'autres neurones. Le tout forme un réseau dense et interconnecté. La pondération des entrées est assurée par des poids (appelés poids synaptiques).

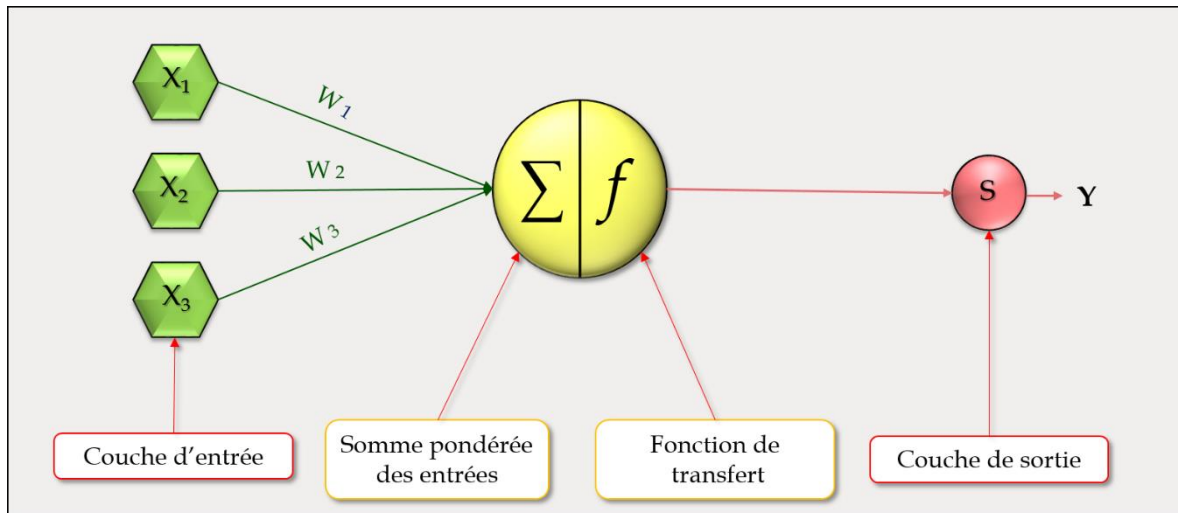


Figure 5-1 : Modèle du neurone formel de MCCulloch et Pitts (1948)

5.3 Type de Réseaux de Neurones

Un neurone réalise une fonction non linéaire, paramétrée, de ses variables d'entrée. L'intérêt des neurones réside dans les propriétés qui résultent de leur association en réseaux, c'est-à-dire, de la composition des fonctions non linéaires réalisées par chacun des neurones. On distingue deux types de réseaux de neurones : les réseaux non bouclés et les réseaux bouclés (Dreyfus et al., 2004). Comme leur nom l'indique, les réseaux de neurones non bouclés possèdent la propriété d'avoir le flux d'information qui circule des entrées vers les sorties, sans retour en arrière. Pour ce qui est des réseaux bouclés, lorsqu'on se déplace dans le réseau en suivant le sens des connexions, il est possible de trouver au moins un chemin qui revient à son point de départ (Dreyfus et al., 2004).

5.4 Perceptron multicouche (MLPNN)

Le perceptron multicouche (MLPNN) est le type de réseaux de neurones les plus utilisés et les plus reportés dans la littérature (Rumelhart et al., 1986a, b). Ce type de réseaux de neurones est largement utilisé dans le domaine des ressources en eau, l'hydrologie et l'environnement. Le modèle MLPNN se compose de trois types de couches :

- Une couche d'entrée qui contient un nombre de neurone égal au nombre de variables explicatives ;

- Une ou plusieurs couches cachées ;
- Une couche de sortie ayant un seul neurone qui correspond à la variable à expliquer.

Les paramètres du modèle MLPNN sont les poids et les biais. Ces derniers sont optimisés par un algorithme d'apprentissage généralement la rétropropagation du gradient en utilisant la partie apprentissage de la base de données.

L'algorithme d'apprentissage est répété plusieurs fois pour déterminer le meilleur modèle qui possède la plus petite erreur entre les valeurs observées et les valeurs calculées.

Le modèle utilisé dans le cadre de la présente thèse possède une fonction d'activation sigmoïde pour les neurones de la couche cachée et une fonction d'activation linéaire pour les neurones de la couche de sortie. La structure est présentée dans la [Figure 5-2](#).

Il a été démontré qu'un perceptron multicouche (MLPNN) ayant une seule couche cachée à fonction d'activation sigmoïde est un approximateur universel ([Hornik et al., 1990](#) ; [Hornik, 1991](#)). Du point de vue mathématique, un MLPNN peut être présenté comme suit :

$$I_j = \sum_{i=1}^m x_i w_{ij} + \delta_j \quad (i = 1, 2, \dots, n) \quad (5-1)$$

$$E_j = f_1(I_j) \quad (j = 1, 2, \dots, m) \quad (5-2)$$

$$f_1(x) = \frac{1}{1 + e^{-x}} \quad (5-3)$$

$$O = \sum_{j=1}^m E_j w_{ik} + \delta_0 \quad (k = 1, 2, \dots, p) \quad (5-4)$$

$$Y = f_2(O) \quad (5-5)$$

Avec :

x_i : l'entrée i ;

I_j : Etat du neurone j de la couche cachée (égale à la somme pondérée de toutes les entrées x_i plus un biais) ;

W_{ij} : Poids de la connexion entre les neurones i et j (entre la couche d'entrée et la couche cachée) ;

W_{jk} : Poids de la connexion entre les neurones j et k (entre la couche cachée et la couche de sortie) ;

f_j : Fonction d'activation du neurone j appartenant à la couche cachée (généralement la sigmoïde) ;

δ_j : Le biais associé au neurone j de la couche cachée ;

δ_k : Le biais associé au neurone k de la couche de sortie ;

m : Nombre de neurone dans la couche cachée ;

O_k : Etat du neurone k de la couche de sortie.

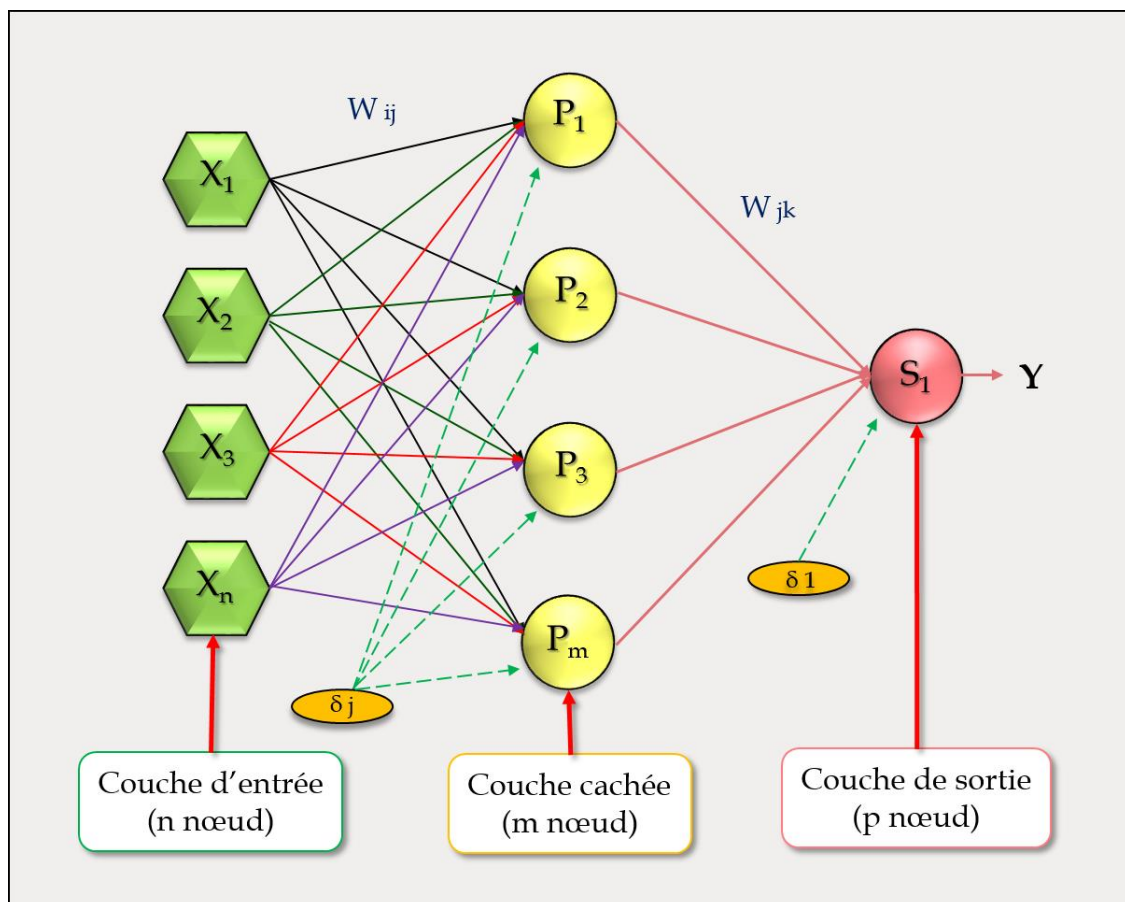


Figure 5-2 : Architecture du perceptron multi couches (MLPNN).

La fonction f (Fig. 5-1), est la fonction d'activation du neurone, il existe plusieurs types de fonctions d'activation (Fig. 5-3), parmi lesquelles :

a- La fonction identité :

$$f(A) = A \quad (5-6)$$

b- La fonction sigmoïde :

$$f(A) = \frac{1}{1 + e^{-A}} \quad (5-7)$$

c- La fonction Tangente hyperbolique :

$$f(A) = \frac{e^A - e^{-A}}{e^A + e^{-A}} \quad (5-8)$$

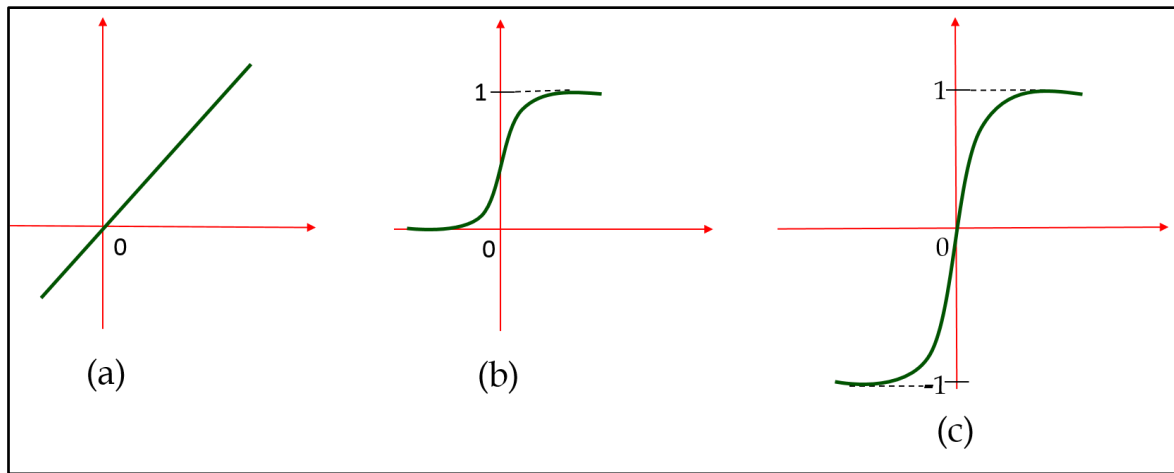


Figure 5-3 : Différentes fonctions d'activation : (a) Identité (Linéaire), (b) Sigmoïde et (c) Tangente hyperbolique.

Le problème de l'apprentissage se résume à minimiser une fonction coût, cette minimisation s'effectue à l'aide de méthodes de gradient descendant. L'algorithme le plus connu c'est l'algorithme de rétropropagation du gradient (Rumelhart et al., 1986a ; 1986b). La fonction coût la plus fréquemment utilisée est la moyenne de l'erreur quadratique (MSE), cette dernière mesure l'écart entre les sorties effectives du réseau de neurones (s) et les sorties désirées ou cibles (d).

$$E_p = \frac{1}{2} \sum_{k=1}^P (d_k - s_k)^2 \quad (\text{Erreur sur l'exemple } p) \quad (5-9)$$

$$E = \sum_{p=1}^P E_p \quad (\text{Erreur totale sur la base d'apprentissage}) \quad (5-10)$$

5.5 Réseaux de neurones à fonction de régression générale (GRNN)

Les réseaux de neurones à fonction de régression générale (GRNN) ont été introduits par Donald F. Specht (Specht, 1991). Ce sont des modèles de réseaux de neurones probabilistes, sans apprentissage, largement utilisés pour l'approximation de fonction entre un vecteur d'entrée $[X_i]$ et un vecteur de sortie désiré $[Y_i]$. Comme montré sur la Figure 5-4, le modèle GRNN se compose de quatre couches (Specht, 1991) :

- La couche d'entrée (input layer) ;
- La couche de modèle (Pattern layer) ;
- La couche de sommation (summation layer) ;
- La couche de sortie (output layer).

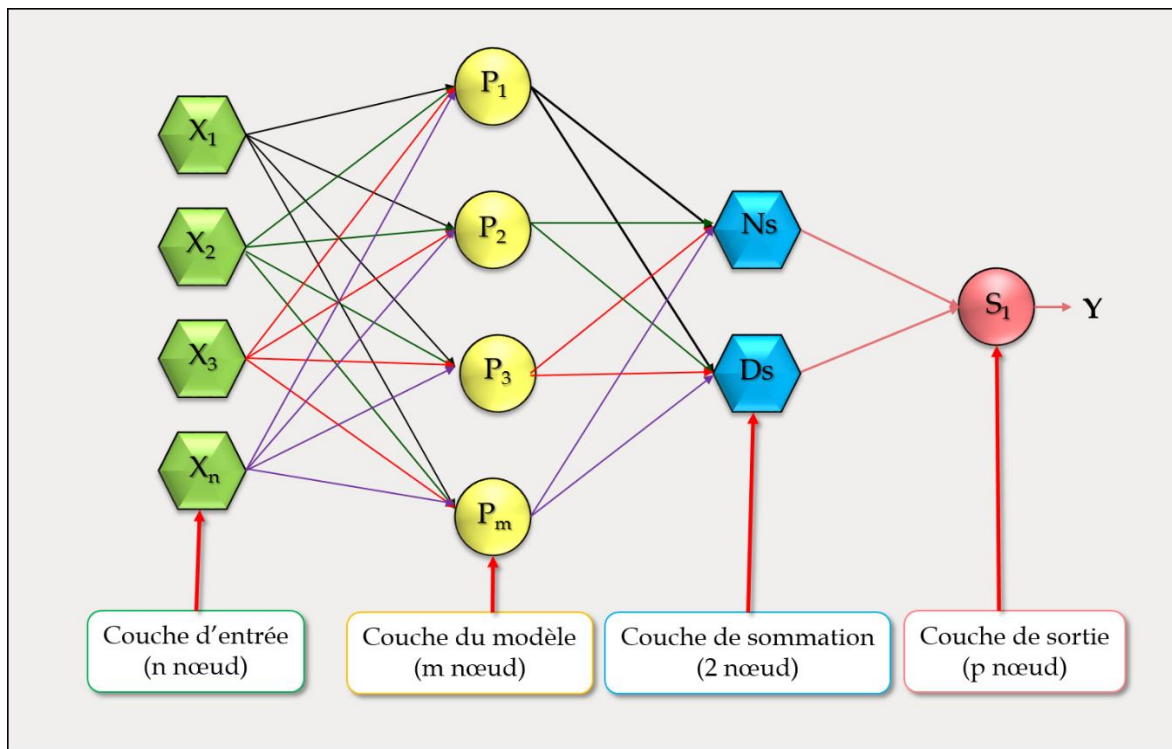


Figure 5-4 : Schéma détaillé d'un Réseau de neurones à fonction de régression générale (GRNN)

La couche d'entrée comprend un nombre de neurone égal au nombre de variables explicatives (X_i), alors que la couche de sortie comprend un seul neurone appelé neurone de décision (Y_i).

La couche de sommation comprend deux neurones très bien distingués : le numérateur (N_s , en anglais numerator) et le deuxième type de neurones appelé : dénominateur (D_s , en anglais denominator). La sortie Y est estimée par le rapport entre le numérateur et le dénominateur ($Y = N_s/D_s$), (équation 5-17). La couche de modèle se compose d'un nombre de neurones égale au nombre de vecteur (individus ou enregistrement) composant la partie d'apprentissage (d'entraînement) (Cigizoglu 2006 ; Cigizoglu & Alp 2006 ; Specht, 1991).

La formulation mathématique du modèle GRNN qui conduit à l'estimation de Y en fonction de X_i est la suivante (plus de détail peut être trouvé dans (Specht, 1991)) :

$$E[y/X] = \frac{\int_{-\infty}^{\infty} yf(\mathbf{X}, y)dy}{\int_{-\infty}^{\infty} f(\mathbf{X}, y)dy} \quad (5-11)$$

$E[y/X]$: la moyenne conditionnelle de y sur \mathbf{X} ;

$f[\mathbf{X}/y]$: la fonction de densité de probabilité.

La fonction de densité de probabilité, $f[\mathbf{X}/y]$, est estimée à partir des observations d'échantillon de x et de y . L'estimation de la fonction de probabilité $f'[\mathbf{X}/Y]$ (Eq. (5-12)) est définie pour les observations des couples, X_i et Y_i des variables aléatoires x et y :

$$f'(\mathbf{X}, Y) = \frac{1}{(2\pi)^{\frac{p+1}{2}} \sigma^{p+1}} \cdot \frac{1}{n} \sum_{i=1}^n f_X f_Y \quad (5-12)$$

Où

$$f_X = \exp \left[-\frac{(\mathbf{X}-X_i)^T(\mathbf{X}-X_i)}{2\sigma^2} \right] \quad (5-13)$$

Et

$$f_Y = \exp \left[-\frac{(Y-Y_i)^2}{2\sigma^2} \right] \quad (5-14)$$

n : Nombre des observations ;

p : Dimension du vecteur X ;

σ : Paramètre (appelé en anglais *Spread*).

Une interprétation physique de la fonction de densité de probabilité, $f'[\mathbf{X}/Y]$ est qu'elle attribue à chaque échantillon X_i et Y_i la probabilité du paramètre σ . La probabilité estimée est la somme des deux.

$$\hat{Y}(X) = \frac{\sum_{i=1}^n \exp \left[-\frac{(\mathbf{X} - \mathbf{X}_i)^T (\mathbf{X} - \mathbf{X}_i)}{2\sigma^2} \right] \int_{-\infty}^{\infty} y \exp \left[-\frac{(y - Y_i)^2}{2\sigma^2} \right] dy}{\sum_{i=1}^n \exp \left[-\frac{(\mathbf{X} - \mathbf{X}_i)^T (\mathbf{X} - \mathbf{X}_i)}{2\sigma^2} \right] \int_{-\infty}^{\infty} \exp \left[-\frac{(y - Y_i)^2}{2\sigma^2} \right] dy} \quad (5-15)$$

Où, la fonction scalaire D_i^2 est donnée par :

$$D_i^2 = (\mathbf{X} - \mathbf{X}_i)^T (\mathbf{X} - \mathbf{X}_i) \quad (5-16)$$

$$\hat{Y}(\mathbf{X}) = \frac{N_s}{D_s} = \frac{\sum_{i=1}^n Y_i \exp \left(-\frac{D_i^2}{2\sigma^2} \right)}{\sum_{i=1}^n \exp \left(-\frac{D_i^2}{2\sigma^2} \right)} \quad (5-17)$$

$$N_s = \sum_{i=1}^n Y_i \exp \left(-\frac{D_i^2}{2\sigma^2} \right) \quad (5-18)$$

$$D_s = \sum_{i=1}^n \exp \left(-\frac{D_i^2}{2\sigma^2} \right) \quad (5-19)$$

Une valeur optimale pour le paramètre σ est déterminée par des méthodes itératives. Elle doit être plus grande que 0 et peut généralement s'étendre de 0.01 à 3.

5.6 Réseaux de neurones à fonctions de base radiales

(RBFNN)

Les réseaux de neurones à fonctions de base radiales (RBFNN) ont été proposés par Moody et Darken en 1989 (Moody & Darken, 1989). Ce sont des réseaux de neurones à propagation en avant utilisant la rétropropagation comme algorithme de minimisation des erreurs et l'optimisation des paramètres. Ce sont des modèles à architecture simple et se composent de trois couches (Fig. 5-5) :

- Une couche d'entrée constituée d'un nombre de neurones égale au nombre de variables explicatives ;
- Une couche cachée de neurones RBF avec des fonctions d'activation gaussienne $\varphi_j(x)$ (Eq. (5-20)) dont les paramètres (μ_j et σ_j) sont à déterminer lors de l'apprentissage ;

- Une couche de sortie linéaire.

Les neurones de la couche cachée ont chacun un centre (μ_j) ; Où (μ_j) est un vecteur dont la dimension est égale au nombre d'entrées et un coefficient d'étalement strictement positif (σ_j) ; Où (σ_j) est l'écart type de la fonction d'activation. Les réseaux RBFNN utilisent une approche par noyaux.

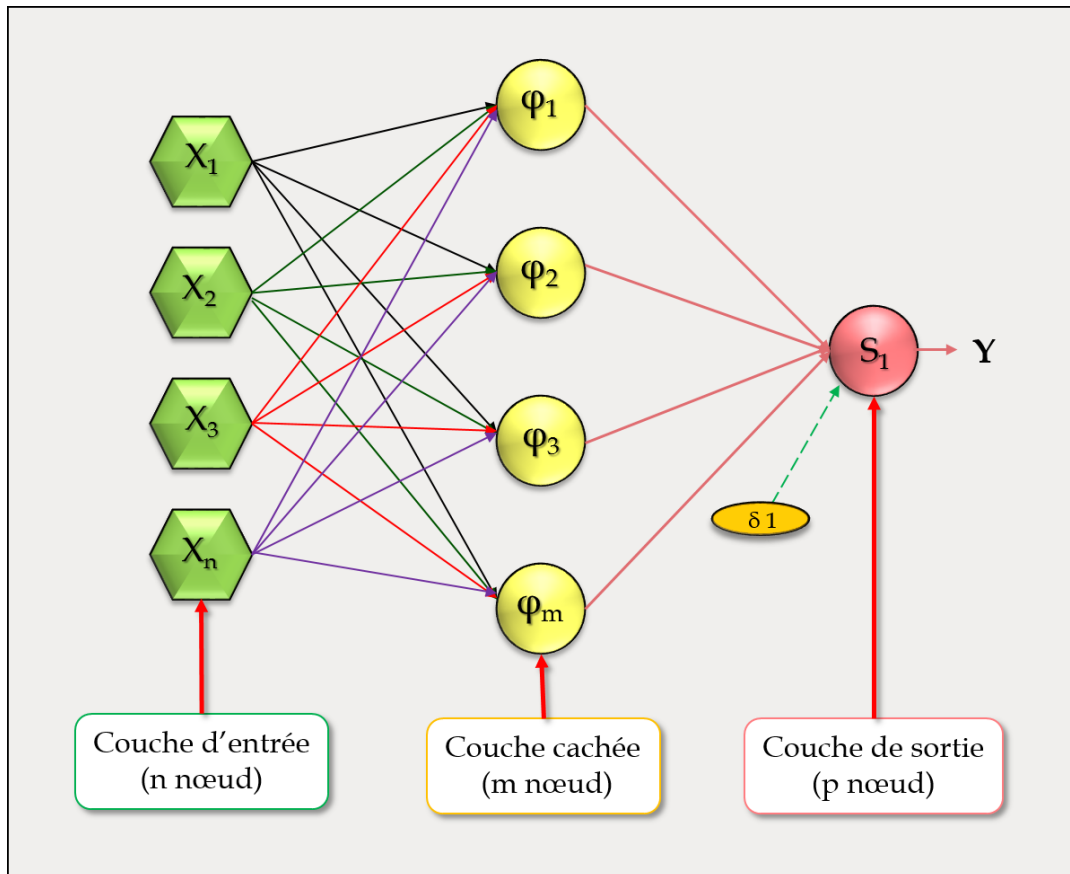


Figure 5-5 : Schéma détaillé d'un Réseau de neurones à fonctions radiales de base (RBFNN)

$$\varphi_j(x) = f(x, \mu_j, \sigma_j) = \exp \left[- \frac{\|x - \mu_j\|^2}{2\sigma_j^2} \right] \quad (j = 1, 2, \dots, m) \quad (5-20)$$

Les sorties réalisent une combinaison linéaire des fonctions radiales :

$$y_k = \sum_{j=1}^m w_{jk} f(x, \mu_j, \sigma_j) + \delta_0 \quad k = (1, 2, \dots, p) \quad (5-21)$$

Où

$\varphi_j()$: La fonction de base radiales de la couche de l'unité cachée j ;

x : Le vecteur d'entrée ;

w_{jk} : Le poids de la connexion entre l'unité cachée (φ_j) et la sortie y_k ;

σ_j : L'écart type de la fonction d'activation gaussienne ;

μ_i : Le vecteur des coordonnées du centre de la fonction d'activation gaussienne ;

δ_k : Le biais associé au neurone k de la couche de sortie.

5.7 Identification des entrées pertinentes par la technique

Gamma Test

5.7.1 Introduction

Ces variables décrivent l'état du système. Dans le cas des réseaux de neurones qui sont des systèmes non autonomes, un sous-ensemble des variables descriptives est constitué par les variables d'entrée, variables dont la valeur est déterminée extérieurement au modèle.

Une connaissance préalable du modèle peut contribuer à accroître ses performances. L'une des plus difficiles étapes dans l'élaboration d'un modèle est le choix des variables d'entrée les plus prépondérantes et les plus efficaces de la fonction non linéaire inconnue.

5.7.2 Description de la Technique Gamma Test

Le Gamma test (GT) est un outil d'analyse qui estime l'erreur quadratique moyenne minimale (MSE) au cours de la modélisation des données implicites en utilisant n'importe quelle modélisation continue non linéaire. Il permet d'examiner la relation entrée/sortie d'un ensemble de données numériques. Il aide à déterminer la taille optimale de la base des données ainsi que la meilleure combinaison des variables d'entrée pour cibler une sortie particulière. Le Gamma test a d'abord été introduit par (Stefánsson et al., 1997) et plus tard renforcé et discuté en détail par de nombreux autres chercheurs (Evans & Jones 2002 ; Evans et al., 2002). Il calcule essentiellement la partie variance de la sortie qui ne peut pas être représentée par un modèle lisse, basé sur les entrées, même si ce modèle est inconnu. L'avantage principal de cet outil est sa

vitesse d'exécution particulièrement dans le cas de l'utilisation de grandes bases de données. Il permet donc de traiter une base de données à plusieurs variables avec des milliers d'observations en un temps record et de sélectionner celles les plus pertinentes (Jones, 2001). En outre, parmi les raisons qui font que le Gamma test soit si utile, c'est qu'il peut immédiatement et directement indiquer à partir des données fournies si on dispose de suffisamment de données pour former un modèle non-linéaire lisse, comme il peut indiquer le niveau de performance de ce modèle (Durrant, 2001 ; Evans & Jones 2002).

5.7.3 Formulation mathématique de la Technique Gamma Test

L'idée de base est tout à fait distincte des tentatives antérieures avec l'analyse non linéaire. Supposons que nous avons un ensemble d'observations, $\{(x_i, y_i), 1 \leq i \leq M\}$, où les vecteurs d'entrée $x_i \in R^m$ sont des vecteurs confinés à un ensemble limité clos $C \in R^m$, et sans perte de généralité les sorties correspondantes $y_i \in R$ sont des scalaires. Les vecteurs x contiennent prédictivement des facteurs utiles influençant la sortie y . La seule hypothèse faite est que la relation sous-jacente du système est $y = f(x_1 \dots x_m) + r$, où f est une fonction régulière et r est une variable aléatoire qui représente le bruit. Sans perte de généralité, on peut supposer que la moyenne de la distribution de r est égale à zéro (puisque tout biais constant peut être inclus dans la fonction inconnue f et que la variance du bruit $Var(r)$ est bornée). Le modèle possible a un domaine qui est maintenant limité à la classe des fonctions régulières qui ont des dérivées partielles premières bornées.

La technique du Gamma Test est basée sur le principe de $N[i, k]$ qui sont les $k^{ièmes}$ ($1 \leq k \leq p$) plus proche voisin de $x_{N[i, k]}$ ($1 \leq k \leq p$) pour chaque vecteur x_i ($1 \leq i \leq M$). Plus précisément, le Gamma Test est obtenu à partir de la fonction Delta des vecteurs d'entrée :

$$\delta_M(k) = \frac{1}{M} \sum_{i=1}^M |x_{N[i, k]} - x_i|^2 \quad (1 \leq k \leq p) \quad (5-22)$$

Où $|\dots|$ est la distance euclidienne et par conséquent la fonction Gamma correspond à la valeur de la sortie (output) ou (la variable à expliquer) :

$$\gamma_M(k) = \frac{1}{2M} \sum_{i=1}^M |y_{N[i,k]} - y_i|^2 \quad (1 \leq k \leq p) \quad (5-23)$$

Où $y_{N[i,k]}$ est la valeur correspondante du $k^{ième}$ ($1 \leq k \leq p$) plus proche voisin du vecteur x_i ($1 \leq k \leq p$) dans l'équation (5-22).

Pour calculer (Γ), une droite de régression des moindres carrées est construite pour les p points ($\delta_M(k), \gamma_M(k)$) :

$$\boldsymbol{\gamma} = \mathbf{A}\boldsymbol{\delta} + \boldsymbol{\Gamma} \quad (5-24)$$

L'intersection avec l'axe vertical ($\delta = 0$) est la valeur Γ , comme on peut montrer $\gamma_M(k) \rightarrow Var(r)$ sachant qu'en probabilité $\gamma_M(k) \rightarrow 0$. Le calcul du gradient de la droite de régression peut également fournir des informations utiles sur la complexité du système d'étude. Une justification mathématique formelle de la méthode peut être trouvée dans [Evans & Jones \(2002\)](#). La sortie graphique de cette droite de régression (Eq. (5-24)) fournit une information très utile. Premièrement, il est visible que l'intersection verticale Γ de l'axe y (ou Gamma) fournit une estimation de la meilleure MSE possible en utilisant une technique de modélisation pour les fonctions inconnues régulières de variables continues ([Evans & Jones, 2002](#)). Deuxièmement, le gradient indique la complexité du modèle (un gradient plus fort indique un modèle de plus grande complexité). En pratique, le Gamma test peut être obtenu à travers l'implémentation du logiciel WinGamma™ ([Durrant, 2001](#)). Le Gamma test est un modèle conceptuel, ces résultats n'ont aucune relation avec la technique utilisée pour exprimer la fonction f . Ces résultats peuvent être standardisés en considérant le terme dit V_{ratio} qui est défini comme suit :

$$V_{ratio} = \frac{\Gamma}{\sigma^2(y)} \quad (5-25)$$

$\sigma^2(y)$ est la variance de y (la réponse du modèle ou la variable à expliquer) qui fournit la capacité de jugement formé indépendamment de l'intervalle de variation de y . Une valeur de V_{ratio} proche de zéro ($V_{ratio} = 0$) indique un très grand degré de prédiction pour la valeur à expliquer y . En fait :

$$1 - |V_{ratio}| = 1 - \frac{|\Gamma|}{Var(y)} \quad (5-26)$$

est pratiquement analogue au conventionnel R^2 (coefficient de détermination) qui exprime la capacité des données (x_i, y_i) à construire un modèle linéaire. Par contre, le V_{ratio} ici exprime la capacité des données à construire un modèle non linéaire.

5.8 Stratégie de développement du meilleur modèle

L'organigramme (Fig. 5-6) illustre la démarche entreprise pour développer le modèle le plus performant en termes de prévision des volumes totaux alloués (Alimentation en eau potable et Irrigation). La première étape concerne le traitement et la mise à jour de la base des données. La deuxième étape consiste à choisir les variables à utiliser pour la construction du modèle à base des réseaux de neurones (RNA) en détectant les liens entre les variables moyennant la corrélation et en déterminant le pas de temps ($t-i ; 1 \leq i \leq 4 ; i$ est un entier désignant le nombre de mois) favorable. La troisième étape porte sur l'application du Gamma test pour la sélection des variables indispensables, leur classification et l'extraction des meilleurs vecteurs d'entrée du modèle neuronal par combinaison de celles-ci. La quatrième étape vise le développement des modèles RNA (MLPNN, GRNN et RBFNN) en utilisant les vecteurs déterminés dans la troisième étape par le Gamma test. Enfin en utilisant les critères de performance, on adopte le modèle adéquat.

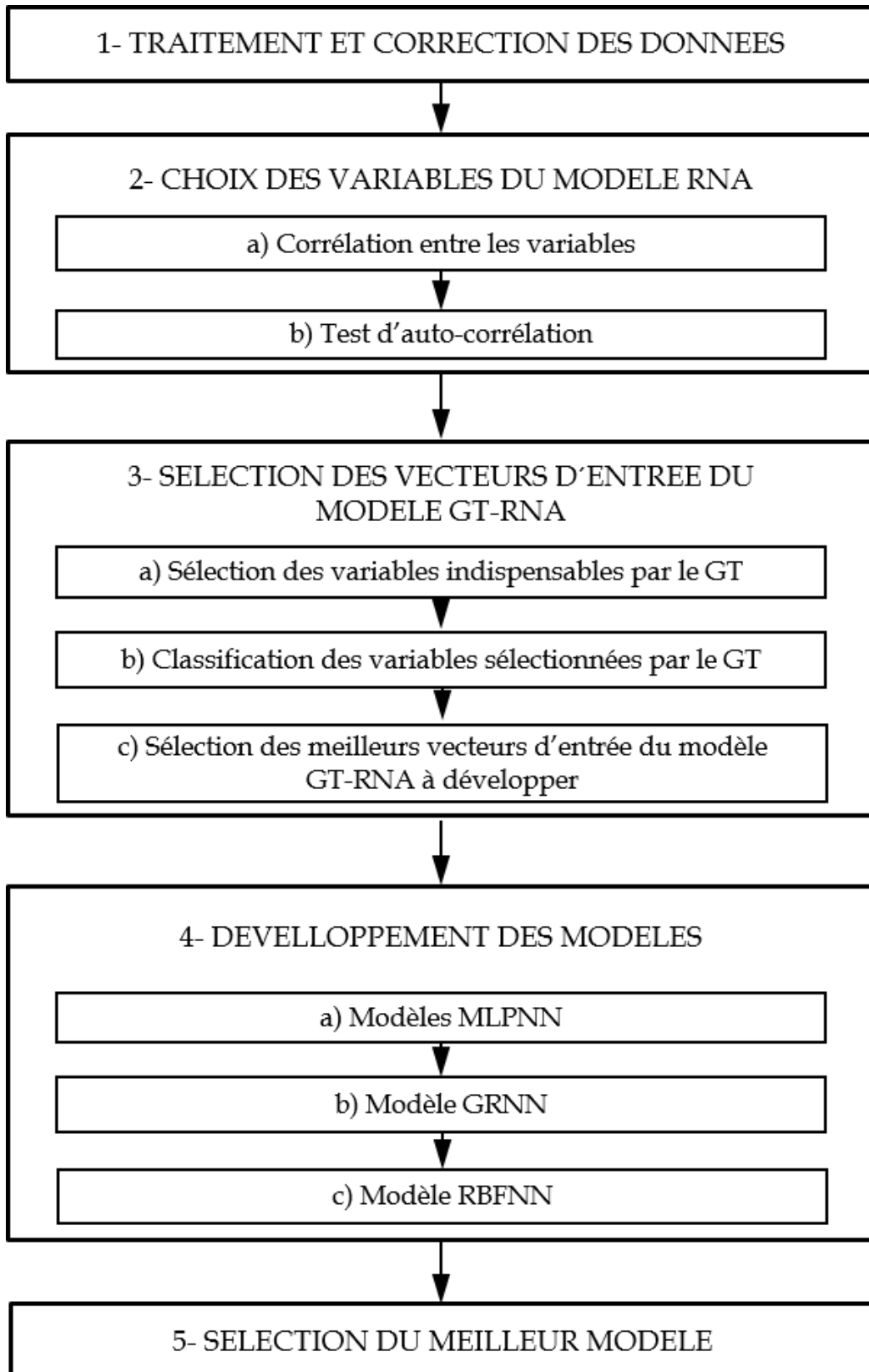


Figure 5-6 : Organigramme de développement des modèles.

5.8.1 Application de la Technique du Gamma Test

Après correction des erreurs systématiques et accidentelles ainsi que la mise à jour des courbes (Hauteur-Volume-Surface) une corrélation est établie entre les paramètres d'exploitation de l'eau mobilisée par le barrage. Les variables suivantes sont liées au volume total alloué : Apport ($R = -0.26$), pluie ($R = -0.42$), évaporation ($R = 0.62$) et fuites ($R = 0.29$). Le test d'auto-corrélation a défini un ordre de pas de temps de $t-2$, ainsi un vecteur d'entrée de dix (10) variables (APP_{t-2} , APP_{t-1} , PLU_{t-2} , PLU_{t-1} , FUI_{t-2} , FUI_{t-1} , $EVAP_{t-2}$, $EVAP_{t-1}$, $VTAL_{t-2}$, $VTAL_{t-1}$) est retenu pour l'élaboration du modèle neuronal de prévision du volume total alloué au temps t . La sélection d'une combinaison appropriée des variables d'entrée est l'étape la plus importante du processus de construction des modèles d'intelligence artificielle. Dans ce qui suit, les paramètres qui ont un impact significatif sur les résultats du modèle sont identifiés. La sélection se fait en 3 étapes :

5.8.1.1 Sélection des variables indispensables par le Gamma Test

Pour identifier les variables les plus importantes, il faut d'abord que la valeur de Gamma doit être calculée pour une combinaison de toutes ces variables (10 variables d'entrée). Dans l'étape suivante, une des variables est omise et la valeur de Gamma est recalculée pour une combinaison des autres variables (09 variables). Ensuite, la variable omise à l'étape précédente est récupérée et une autre variable est omise de la combinaison initiale (10 variables) et la valeur de Gamma est ensuite recalculée pour la nouvelle combinaison à neuf (09) variables. Ce processus se poursuit pour toutes les variables, une par une, et dans chaque étape la valeur Gamma est calculée. Dans ce procédé, l'omission de variables importantes est liée à l'augmentation de la valeur de Gamma par rapport à la combinaison originale. Les résultats pour des combinaisons différentes sont indiqués dans le [Tableau 5-1](#). Ce tableau indique que $VTAL_{t-1}$ est la variable la plus importante. Les autres variables importantes sont PLU_{t-1} , $EVAP_{t-1}$, APP_{t-1} et $VTAL_{t-2}$ respectivement.

Tableau 5-1 : Résultats du Gamma Test sur les variables d'entrée

Variables d'entrée	Valeurs de Gamma
Toutes les entrées	0.055000
Toutes les entrées-APP _{t-2}	0.051000
Toutes les entrées-APP _{t-1}	0.059000
Toutes les entrées-PLU _{t-2}	0.053000
Toutes les entrées-PLU _{t-1}	0.062490
Toutes les entrées-FUI _{t-2}	0.054700
Toutes les entrées-FUI _{t-1}	0.052950
Toutes les entrées-EVAP _{t-2}	0.051681
Toutes les entrées-EVAP _{t-1}	0.059823
Toutes les entrées-VTAL _{t-2}	0.058489
Toutes les entrées-VTAL _{t-1}	0.077689

5.8.1.2 Classification des variables sélectionnées par le Gamma Test

Les variables importantes ainsi trouvées sont classées par ordre d'influence affectant le modèle comme l'indique le [Tableau 5-2](#).

Tableau 5-2 : Classification des variables d'entrée par le Gamma Test

Variables d'entrée	Valeurs de Gamma
Toutes les entrées	0.038648
Toute les entrées-VTAL _{t-1}	0.086543
Toute les entrées-VTAL _{t-2}	0.049302
Toute les entrées- PLU _{t-1}	0.048587
Toute les entrées- EVAP _{t-1}	0.042556
Toute les entrées- APP _{t-1}	0.039651

Ce classement intervient dans les combinaisons des données d'entrée prédéterminées pour en sortir les meilleurs vecteurs d'entrée.

5.8.1.3 Sélection des meilleurs vecteurs d'entrée

Différentes combinaisons des variables d'entrée sont examinées pour évaluer leur influence sur la modélisation du VTAL_t. Les combinaisons des variables significatives à partir desquelles les meilleures peuvent être déterminées en observant les valeurs de : Gamma (Γ), Gradient (A), Standard Error (SE) et V_{ratio} sont résumées dans les [Tableaux 5-3, 5-4 et 5-5](#).

Tableau 5-3 : Résultats des combinaisons des variables d'entrée : quatre et cinq variables

Désignation	V ₁ V ₂ PEA	V ₂ PEA	V ₁ PEA	V ₁ V ₂ EA	V ₁ V ₂ PA	V ₁ V ₂ PE
Gamma	0.03865	0.03966	0.04256	0.04859	0.04930	0.08654
A	0.10816	0.14760	0.20422	0.13358	0.12366	0.09970
SE	0.00762	0.00688	0.00861	0.00757	0.00693	0.00791
V _{ratio}	0.15459	0.15860	0.17022	0.19435	0.19721	0.34617

V₁ : VTAL_{t-1}; V₂ : VTAL_{t-2}; P : PLU_{t-1}; E : EVAP_{t-1}; A : APP_{t-1}

Tableau 5-4 : Résultats des combinaisons des variables d'entrée : trois variables

Désignation	PEA	V ₁ EA	V ₁ V ₂ A	V ₁ V ₂ P	V ₂ EA	V ₁ PA	V ₁ PE
Gamma	0.04822	0.05193	0.06117	0.12102	0.04205	0.08578	0.13858
A	0.28867	0.26645	0.00895	0.04840	0.25834	0.05387	0.08340
SE	0.01018	0.00741	0.11151	0.01672	0.00536	0.01193	0.01947
V _{ratio}	0.19288	0.20773	0.24469	0.48407	0.16819	0.34313	0.55432

V₁ : VTAL_{t-1}; V₂ : VTAL_{t-2}; P : PLU_{t-1}; E : EVAP_{t-1}; A : APP_{t-1}

Tableau 5-5 : Résultats des combinaisons des variables d'entrée : deux variables

Désignation	V ₁ V ₂	V ₁ P	V ₁ E	V ₁ A	V ₂ P	V ₂ E	V ₂ A	PE	EA
Gamma	0.13086	0.1545	0.1900	0.0946	0.1223	0.0857	0.0625	0.1456	0.0471
A	-0.0951	0.1540	0.1968	0.0489	0.1113	0.3786	0.0967	0.0046	0.8505
SE	0.00959	0.01541	0.01095	0.01823	0.01606	0.00636	0.01037	0.01694	0.00578
V _{Ratio}	0.52345	0.61818	0.76007	0.37844	0.48948	0.34308	0.25038	0.58260	0.18862

V₁ : VTAL_{t-1}; V₂ : VTAL_{t-2}; P : PLU_{t-1}; E : EVAP_{t-1}; A : APP_{t-1}

En plus de la combinaison faisant intervenir toutes les variables classées, les autres combinaisons de vecteurs de deux, trois et quatre variables correspondant aux valeurs minimales de Gamma et SE sont adoptées pour l'élaboration du modèle. Il s'avère que les valeurs de A des combinaisons retenues ne sont pas minimales ce qui reflète un modèle qui consomme plus de temps dans l'exécution. À la lumière des résultats obtenus par la technique Gamma test, nous avons retenu quatre combinaisons de variables d'entrée qui vont servir pour la construction des modèles de réseaux de neurones artificielles à savoir : les MLPNN, GRNN et RBFNN. Les combinaisons choisies sont : cinq entrées (VTAL_{t-1}, VTAL_{t-2}, PLU_{t-1}, EVAP_{t-1}, APP_{t-1}), quatre entrée (VTAL_{t-2}, PLU_{t-1}, EVAP_{t-1}, APP_{t-1}), trois entrée (VTAL_{t-2}, EVAP_{t-1}, APP_{t-1}) et deux entrée (EVAP_{t-1}, APP_{t-1}).

5.9 Application de trois modèles RNA pour l'estimation du volume total alloué

5.9.1 Généralités et structures des modèles

Nous présentons dans ce chapitre les résultats obtenus par les trois types de modèles de réseaux de neurones définis précédemment sous les abréviations (MLPNN, GRNN, et RBFNN). Ces modèles sont implémentés sous le logiciel Matlab 2009Rb et leurs structures sont telles qu'indiquées dans le [Tableau 5-6](#).

Tableau 5-6 : Structures des modèles des réseaux de neurones utilisés

Modèles			Variables d'entrées
MLPNN1	GRNN1	RBFNN1	VTAL _{t-1} , VTAL _{t-2} , PLU _{t-1} , EVAP _{t-1} , APP _{t-1}
MLPNN2	GRNN2	RBFNN2	VTAL _{t-2} , PLU _{t-1} , EVAP _{t-1} , APP _{t-1}
MLPNN3	GRNN3	RBFNN3	VTAL _{t-2} , APP _{t-1} , EVAP _{t-1}
MLPNN4	GRNN4	RBFNN4	EVAP _{t-1} , APP _{t-1}

La base de données est scindée en deux parties : une partie pour l'apprentissage des modèles (Training) et une autre pour la validation des modèles (Validation).

Pour le présent cas d'étude, nous avons réparti la base de données comme suit : 70% (125 données) pour la partie apprentissage et 30% (53 données) pour la partie validation. Les performances des modèles développés ont été évaluées pour comparer leur efficacité, leur précision et leur capacité de reproduire la réalité physique. Nous avons sélectionné quatre critères de mesure de l'erreur numérique : la Racine Carrée de l'Erreur Quadratique Moyenne (*RMSE* : Root Mean Squared Error), le Coefficient de Corrélation (Pearson Corrélation Coefficient : *R*), l'erreur absolue moyenne (*MAE* : Mean Absolute Error) et le critère de Nash-Sutcliffe (*NSE*). Ces statistiques seront calculées sur l'ensemble des données, à savoir l'apprentissage et la validation.

$$R = \frac{\frac{1}{N} \sum_{i=1}^N (O_i - O_m)(P_i - P_m)}{\sqrt{\frac{1}{N} \sum_{i=1}^N (O_i - O_m)^2} \sqrt{\frac{1}{N} \sum_{i=1}^N (P_i - P_m)^2}} \quad (5-27)$$

$$MAE = \frac{1}{N} \sum_{i=1}^N |O_i - P_i| \quad (5-28)$$

$$RMSE = \sqrt{\frac{1}{N} \sum_{i=1}^N (O_i - P_i)^2} \quad (5-29)$$

$$NSE = 1 - \frac{\sum_{i=1}^N (O_i - P_i)^2}{\sum_{i=1}^N (O_i - O_m)^2} \quad (5-30)$$

Avec :

O_i : Valeur observée ;

P_i : Valeur calculée par le modèle ;

N : Nombre total des observations (nombre d'individus) ;

O_m : Moyenne de valeurs observées ;

P_m : Moyenne de valeurs calculées.

Nous avons normalisé les données en utilisant la formule suivante :

$$x_{n_{i,k}} = \frac{x_{i,k} - m_k}{\sigma_k} \quad (5-31)$$

Avec :

$x_{n_{i,k}}$: Valeur normalisée de la variable k pour l'individu i ;

m_k : Moyenne de la variable k ;

σ_k : Ecart type de la variable k .

5.9.2 Résultats obtenus par les modèles RNA

Nous avons utilisé des réseaux MLPNN avec des fonctions d'activation sigmoïdes pour les neurones de la couche cachée et des fonctions d'activation linéaires pour le neurone de la couche de sortie du réseau, avec un algorithme d'apprentissage type Levenberg-Marquardt (LM).

Pour les réseaux de neurones GRNN, seul le paramètre : « *Spread* » est fixé. Généralement le « *Spread* » varie entre 0,1 et 3. Pour notre cas nous avons testé plusieurs « *Spread* » et les meilleurs résultats sont obtenus différemment selon le nombre d'entrée du modèle.

Pour les réseaux RBFNN, le paramètre « *Spread* » est fixé entre 0,1 et 3 selon le modèle. Comme tous les modèles à base d'intelligences artificielles, MLPNN, GRNN et RBFNN doivent subir une phase d'apprentissage (calage) pour obtenir la meilleure structure et aussi pour fixer les paramètres optimaux de chaque modèle à savoir :

le nombre de neurones dans la couche cachée pour les MLPNN et le *Spread* pour les GRNN et RBFNN.

5.9.2.1 Résultats obtenus par les réseaux de neurones type MLPNN

Le [Tableau 5-7](#), présente les résultats obtenus en différentes phases, à savoir : l'apprentissage et la validation.

Tableau 5-7 : Performances des modèles MLPNN1, MLPNN2, MLPNN3 et MLPNN4

Modèles	Apprentissage				Validation			
	<i>NSE</i>	<i>R</i>	<i>RMSE</i>	<i>MAE</i>	<i>NSE</i>	<i>R</i>	<i>RMSE</i>	<i>MAE</i>
MLPNN1	0.920	0.959	0.756	0.548	0.661	0.880	1.670	1.308
MLPNN2	0.887	0.942	0.895	0.608	0.187	0.653	2.587	1.990
MLPNN3	0.814	0.902	1.152	0.818	0.060	0.625	2.782	1.997
MLPNN4	0.671	0.819	1.529	1.110	-0.155	0.313	3.084	2.148

Comme présenté dans le [Tableau 5-7](#), en phase d'apprentissage les quatre modèles MLPNN ont des performances très variables entre l'excellent pour le modèle MLPNN1 à l'acceptable pour le modèle MLPNN4. En effet, les *RMSE* et *MAE* varient respectivement de 0,756 à 1,529 et de 0,548 à 1,110. Toujours en phase d'apprentissage les *NSE* et *R* varient de 0,671 à 0,920 et de 0,819 à 0,959 respectivement.

La [Figure 5-7](#) illustre la corrélation entre les volumes totaux alloués estimés par les quatre modèles, en phase d'apprentissage, avec ceux observés. La position du nuage par rapport à la première bissectrice confirme les critères statistiques de performance. Aussi, la [Figure 5-8](#) montre que les volumes totaux alloués estimés par les quatre modèles, en phase d'apprentissage, et ceux observés suivent la même tendance en général. Toutefois nous remarquons que les écarts augmentent en réduisant le nombre de variables d'entrée (MLPNN2, MLPNN3 et MLPNN4).

En phase de validation, et à la lumière des résultats obtenus en phase d'apprentissage, les meilleurs résultats sont obtenus par le modèle MLPNN1 avec des coefficients *R* et *NSE* égaux à 0,880 et 0,661 respectivement. Les [Figures 5-9 et 5-10](#) visualisent, respectivement et en phase validation, la corrélation des volumes totaux alloués estimés par les quatre modèles avec ceux observés, ainsi que leurs évolutions mensuelles.

La dispersion du nuage autour de la droite de régression est nettement remarquable pour les modèles MLPNN2, MLPNN3 et MLPNN4. Cela est confirmé par des importants écarts illustrés par la [Figure 5-10](#).

En accord avec les résultats obtenus en phase d'apprentissage, le meilleur modèle s'avère le MLPNN1.

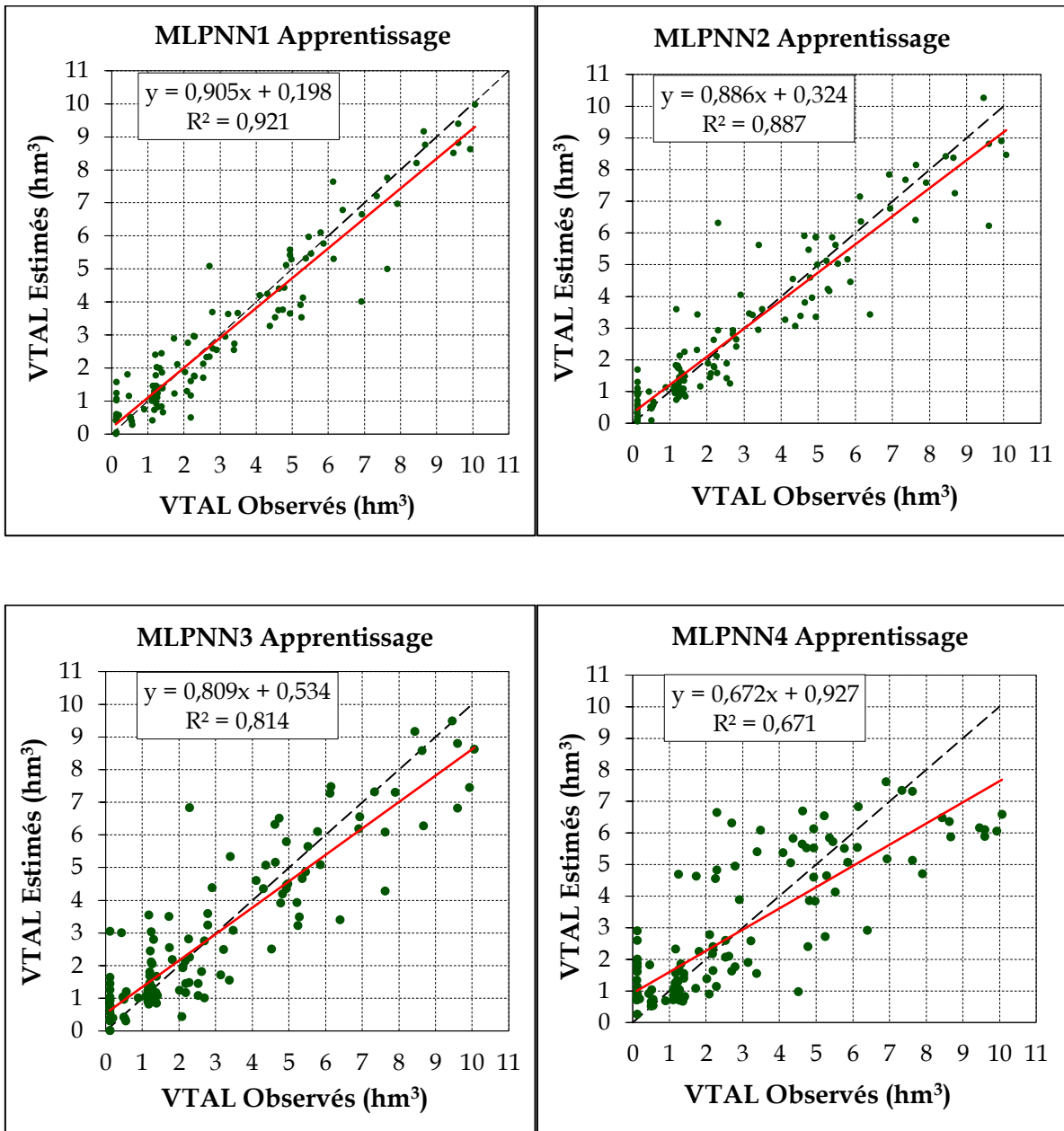


Figure 5-7 : Corrélation des volumes totaux alloués observés et estimés par les modèles MLPNN1, MLPNN2, MLPNN3 et MLPNN4 en phase d'apprentissage

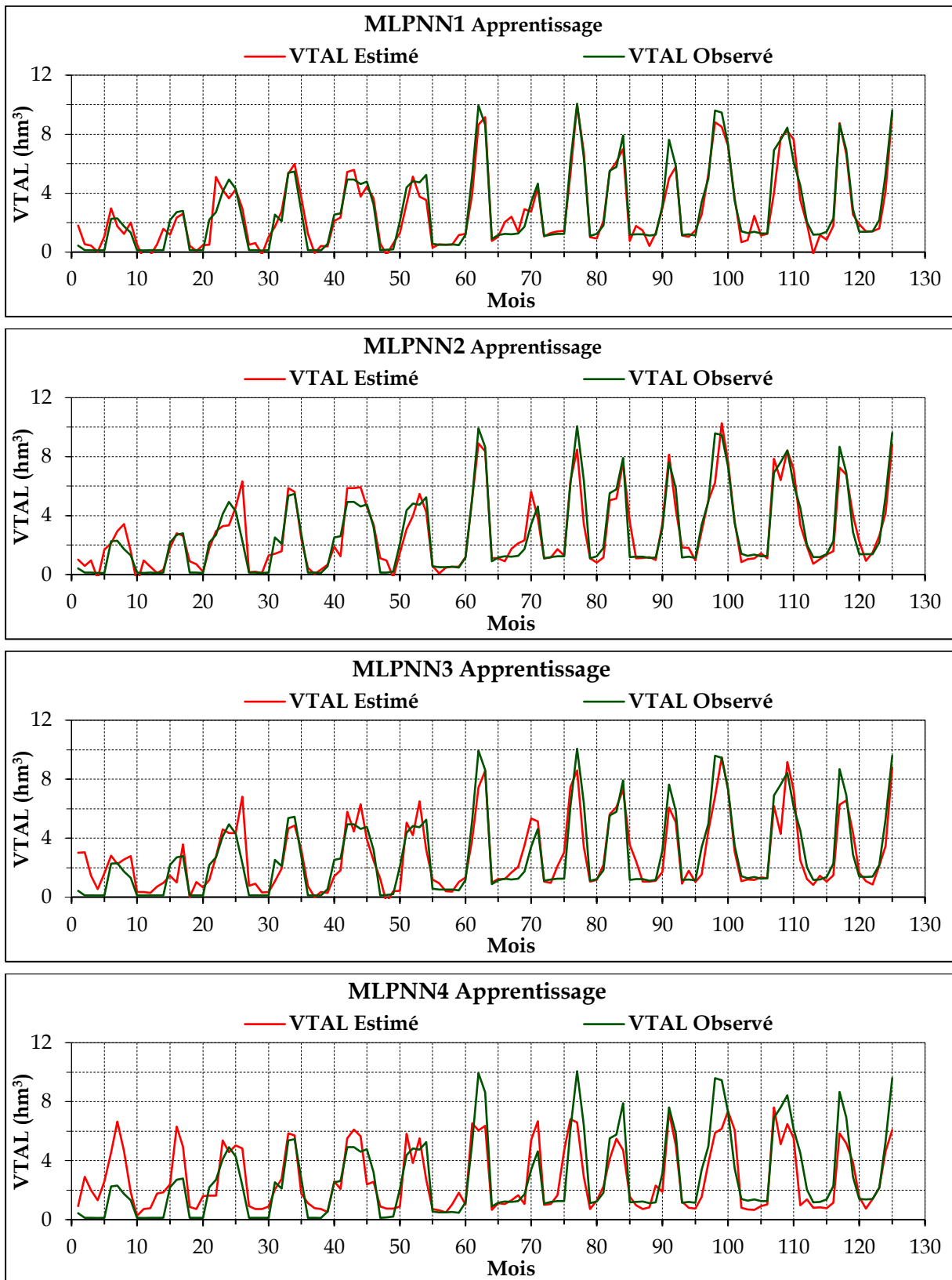


Figure 5-8 : Evolution des volumes totaux alloués observés et estimés par les modèles MLPNN1, MLPNN2, MLPNN3 et MLPNN4 en phase d'apprentissage

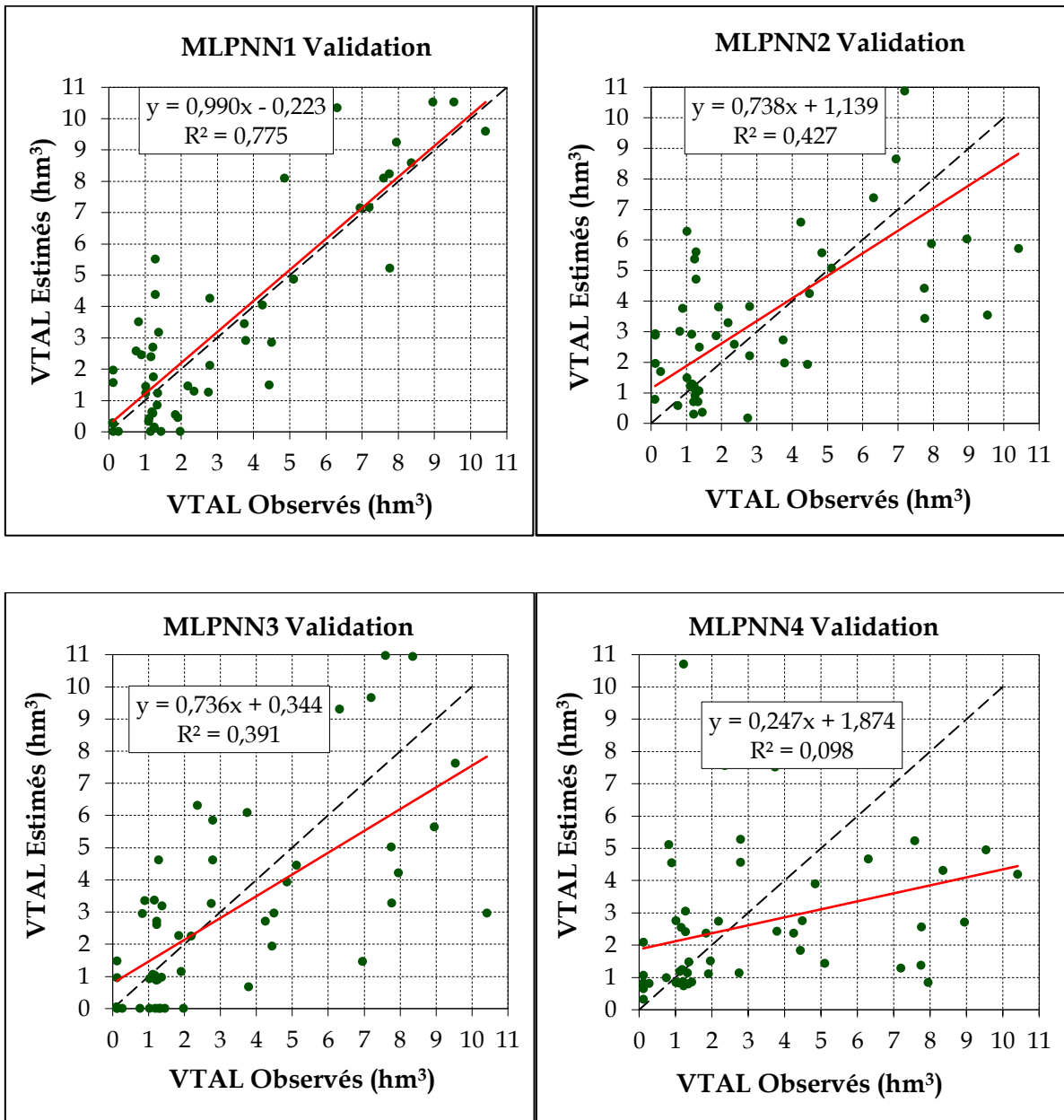


Figure 5-9 : Corrélation des volumes totaux alloués observés et estimés par les modèles MLPNN1, MLPNN2, MLPNN3 et MLPNN4 en phase de validation

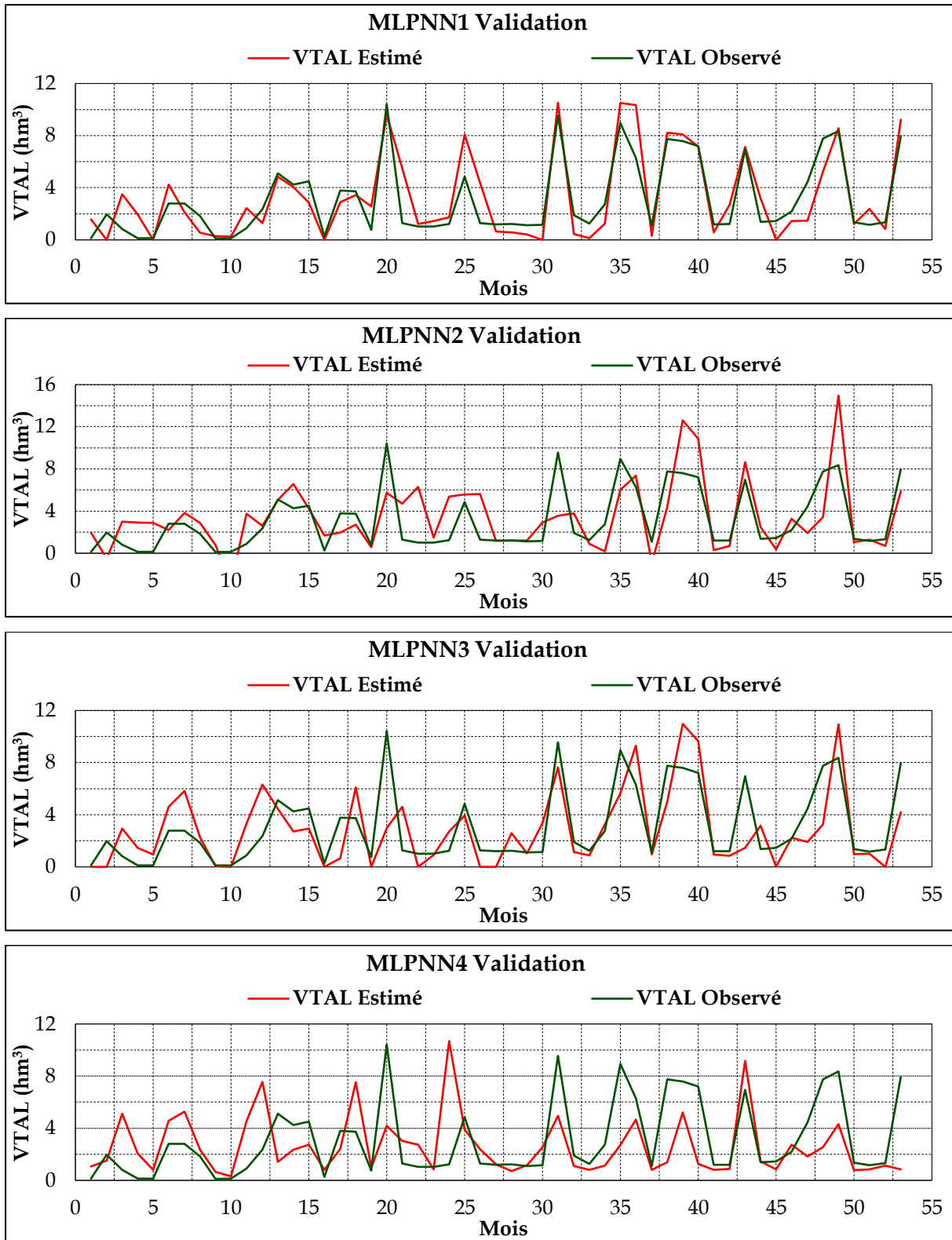


Figure 5-10 : Evolution des volumes totaux alloués observés et estimés par les modèles MLPNN1, MLPNN2, MLPNN3 et MLPNN4 en phase de validation

5.9.2.2 Résultats obtenus par les réseaux de neurones type GRNN

Les résultats obtenus par les modèles GRNN en différentes phases, l'apprentissage et la validation, sont reportés dans le [Tableau 5-8](#).

Tableau 5-8 : Performances des modèles GRNN1, GRNN2, GRNN3 et GRNN4

Modèles	Apprentissage				Validation			
	<i>NSE</i>	<i>R</i>	<i>RMSE</i>	<i>MAE</i>	<i>NSE</i>	<i>R</i>	<i>RMSE</i>	<i>MAE</i>
GRNN1	0.996	0.998	0.162	0.050	0.684	0.876	1.614	1.157
GRNN2	0.990	0.995	0.271	0.098	0.341	0.703	2.330	1.742
GRNN3	0.956	0.978	0.559	0.276	0.442	0.716	2.143	1.582
GRNN4	0.662	0.815	1.550	1.061	0.370	0.630	2.277	1.617

Le [Tableau 5-8](#) présente une comparaison entre les performances obtenues par les quatre modèles GRNN testés. Les [Figures \(5-11, 5-12 et 5-13, 5-14\)](#) présentent la corrélations et l'évolution des volumes totaux alloués observés et estimés par les quatre modèles GRNN1, GRNN2, GRNN3 et GRNN4 en phase d'apprentissage et en phase de validation, respectivement. Quant à la phase apprentissage les valeurs des quatre indices statistiques *RMSE*, *MAE*, *R* et *NSE* varient entre (0,162 et 1,550), (0,050 et 1,061), (0,815 et 0,998) et (0,662 et 0,996) respectivement. Ces valeurs nous permettent de constater qu'en retirant à chaque fois une variable d'entrée les performances du modèle se dégradent. Cela est confirmé par l'éloignement du nuage de la première bissectrice ($Y = X$) dans le cas des modèles a trois entrées (GRNN3) et a deux entrées (GRNN4) ([Fig. 5-11](#)) d'une part et d'autre part, par l'écart enregistré entre les courbes d'évolution des volumes totaux alloués observés et celles des volumes estimés par les modèles GRNN3 et GRNN4 ([Fig. 5-12](#)).

Aussi on peut facilement remarquer qu'en phase de validation les valeurs des quatre indices statistiques *RMSE*, *MAE*, *R* et *NSE* varient entre (1,614 et 2,277), (1,157 et 1,617), (0,630 à 0,876) et (0,370 à 0,684) respectivement. Les meilleurs résultats sont obtenus avec le modèle GRNN1, suivis par le modèle GRNN 3 et GRNN 2 et enfin les plus faibles performances sont obtenues avec le modèle GRNN 4 ([Fig.5-13 et 5-14](#)).

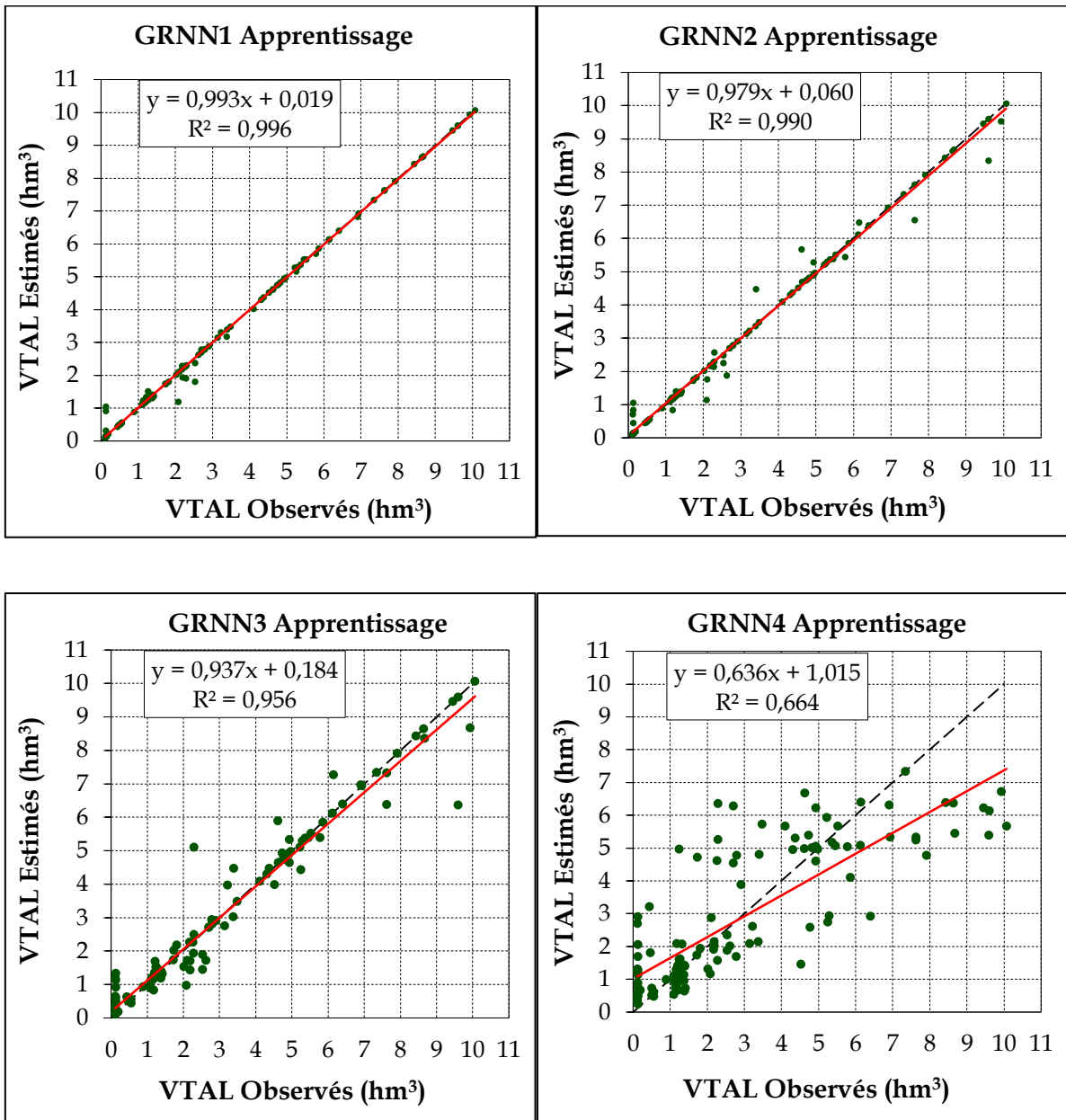


Figure 5-11 : Corrélation des volumes totaux alloués observés et estimés par les modèles GRNN1, GRNN2, GRNN3 et GRNN4 en phase d'apprentissage

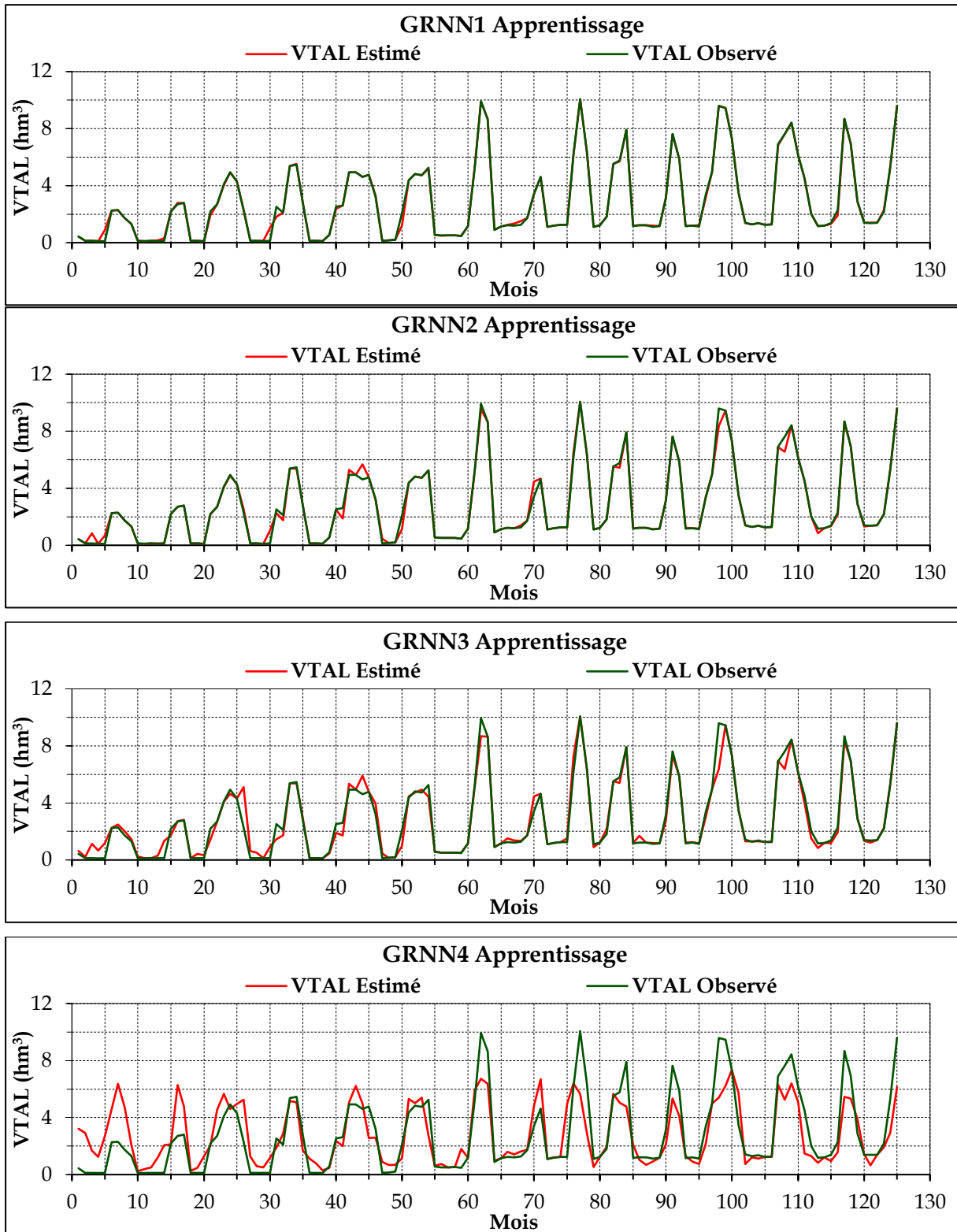


Figure 5-12 : Evolution des volumes totaux alloués observés et estimés par les modèles GRNN1, GRNN2, GRNN3 et GRNN4 en phase d'apprentissage

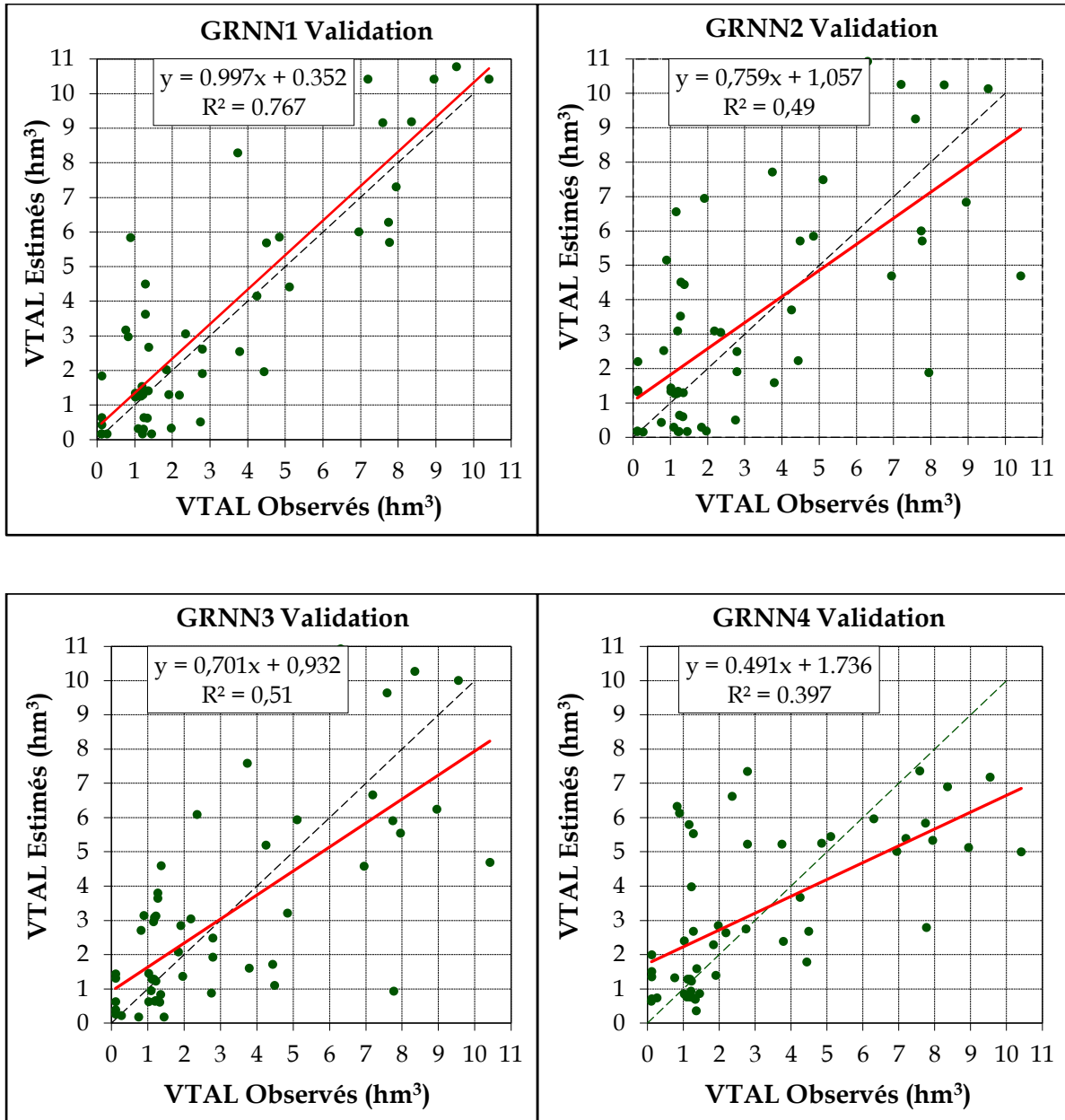


Figure 5-13 : Corrélation des volumes totaux alloués observés et estimés par les modèles GRNN1, GRNN2, GRNN3 et GRNN4 en phase de validation

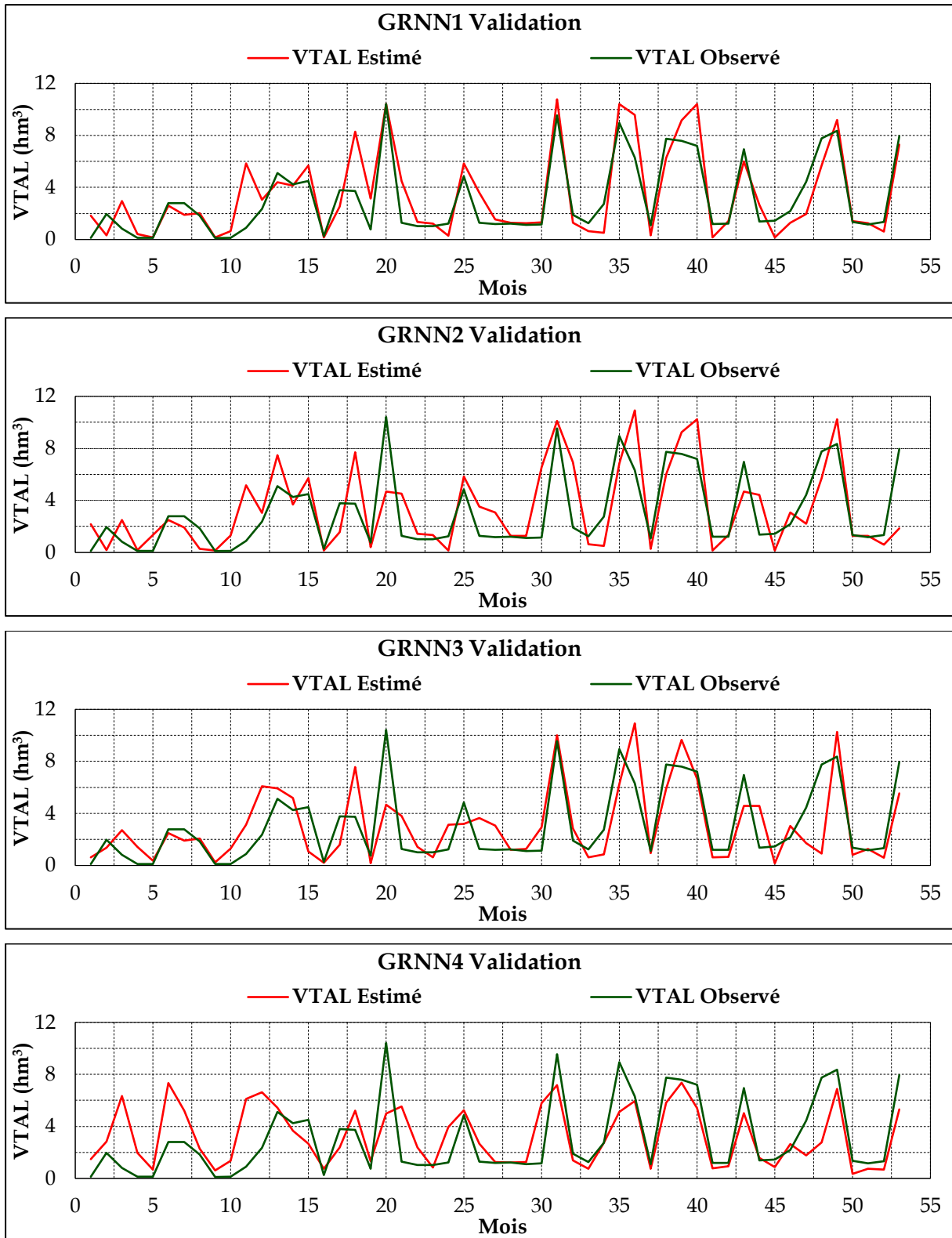


Figure 5-14 : Evolution des volumes totaux alloués observés et estimés par les modèles GRNN1, GRNN2, GRNN3 et GRNN4 en phase de validation

5.9.2.3 Résultats obtenus par les réseaux de neurones type RBFNN

Le [Tableau 5-9](#), présente les résultats obtenus en différentes phases, à savoir : l'apprentissage et la validation.

Tableau 05-9 : Performances des modèles RBFNN1, RBFNN2, RBFNN3 et RBFNN4

Modèles	Apprentissage				Validation			
	<i>NSE</i>	<i>R</i>	<i>RMSE</i>	<i>MAE</i>	<i>NSE</i>	<i>R</i>	<i>RMSE</i>	<i>MAE</i>
RBFNN1	0.990	0.995	0.265	0.164	0.508	0.816	2.012	1.645
RBFNN2	0.991	0.996	0.248	0.113	0.218	0.524	2.539	2.015
RBFNN3	0.990	0.995	0.266	0.138	0.071	0.304	2.766	2.200
RBFNN4	0.997	0.998	0.156	0.053	0.026	0.187	2.832	2.311

Nous discernons clairement à partir du [Tableau 5-9](#) que les coefficients de corrélation des quatre modèles présentent une légère variation en phase d'apprentissage (0,990 à 0,997) et une variation remarquable en phase de validation (0,816 à 0,187). Les meilleurs résultats en phase d'apprentissage sont obtenus avec le modèle RBFNN4 ([Fig. 5-15](#)). Cela est affirmé par la superposition de la courbe d'évolution des volumes totaux alloués observés sur celle des volumes estimés par le modèle RBFNN4 ([Fig. 5-16](#)).

En phase de validation le meilleur coefficient de corrélation est obtenu par le modèle RBFNN1 ($R=0,816$) et le plus faible par le modèle RBFNN4 ($R=0,187$) ([Fig. 5-17](#)). Conformément au [Tableau 5-9](#), le meilleur critère de Nash-Sutcliffe (*NSE*) est obtenu avec le modèle RBFNN1, suivis par les modèles RBFNN2 et RBFNN3. Enfin le plus faible indice est obtenu avec le modèle RBFNN4.

Concernant les indices statistiques en relation avec l'erreur de prédiction (*MAE* et *RMSE*), la plus faible Racine Carrée de l'Erreur Quadratique Moyenne ($RMSE=2,012$) est obtenue par le modèle RBFNN1, de même que pour la plus faible erreur absolue moyenne ($MAE=1,645$).

Cette analyse des critères de performance des modèles en phase validation nous permet de dire que les modèles RBFNN2, RBFNN3 et RBFNN4 ne peuvent pas prévoir les volumes totaux alloués avec satisfaction.

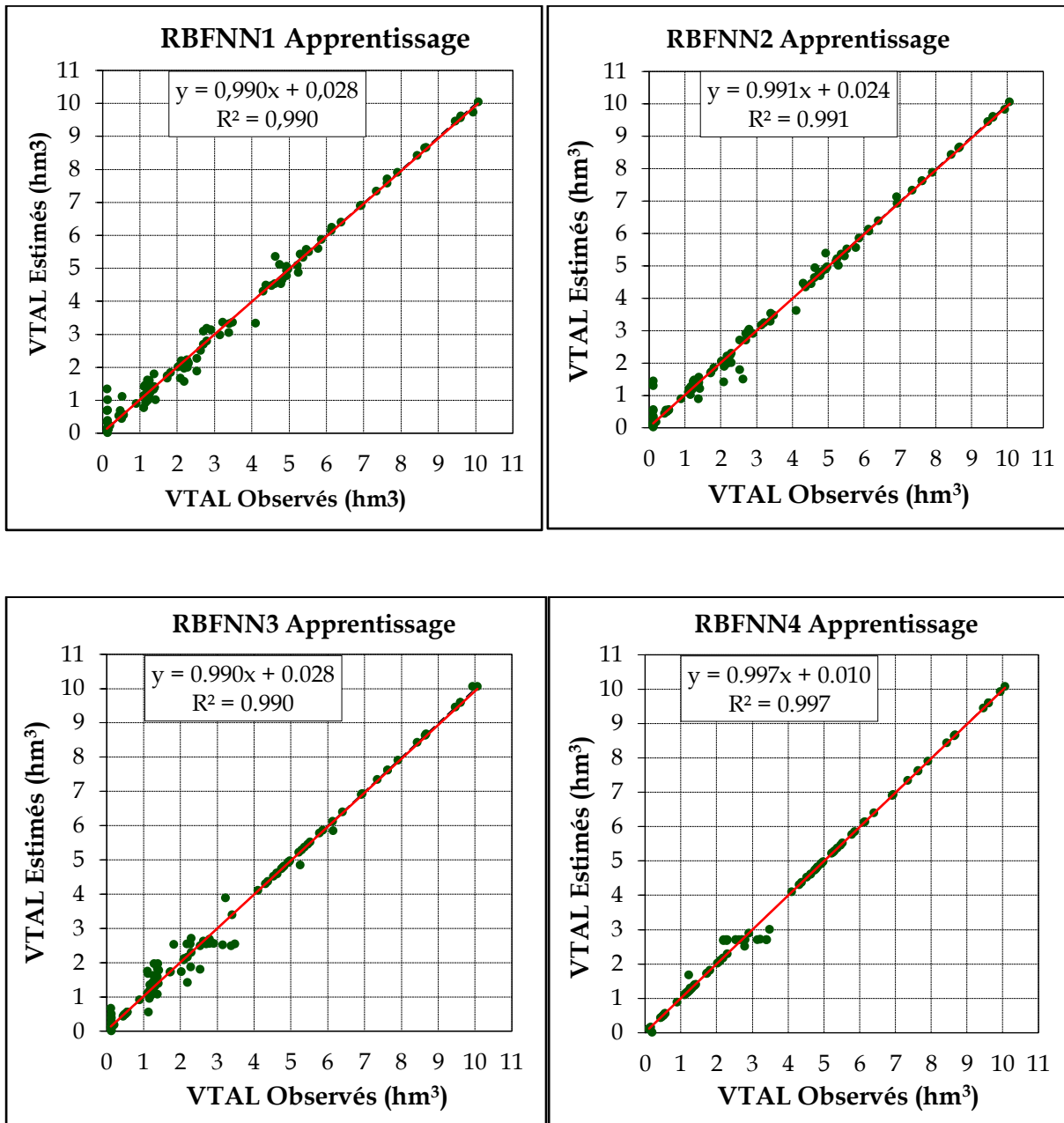


Figure 5-15 : Corrélation des volumes totaux alloués observés et estimés par les modèles RBFFNN1, RBFFNN2, RBFFNN3 et RBFFNN4 en phase d'apprentissage

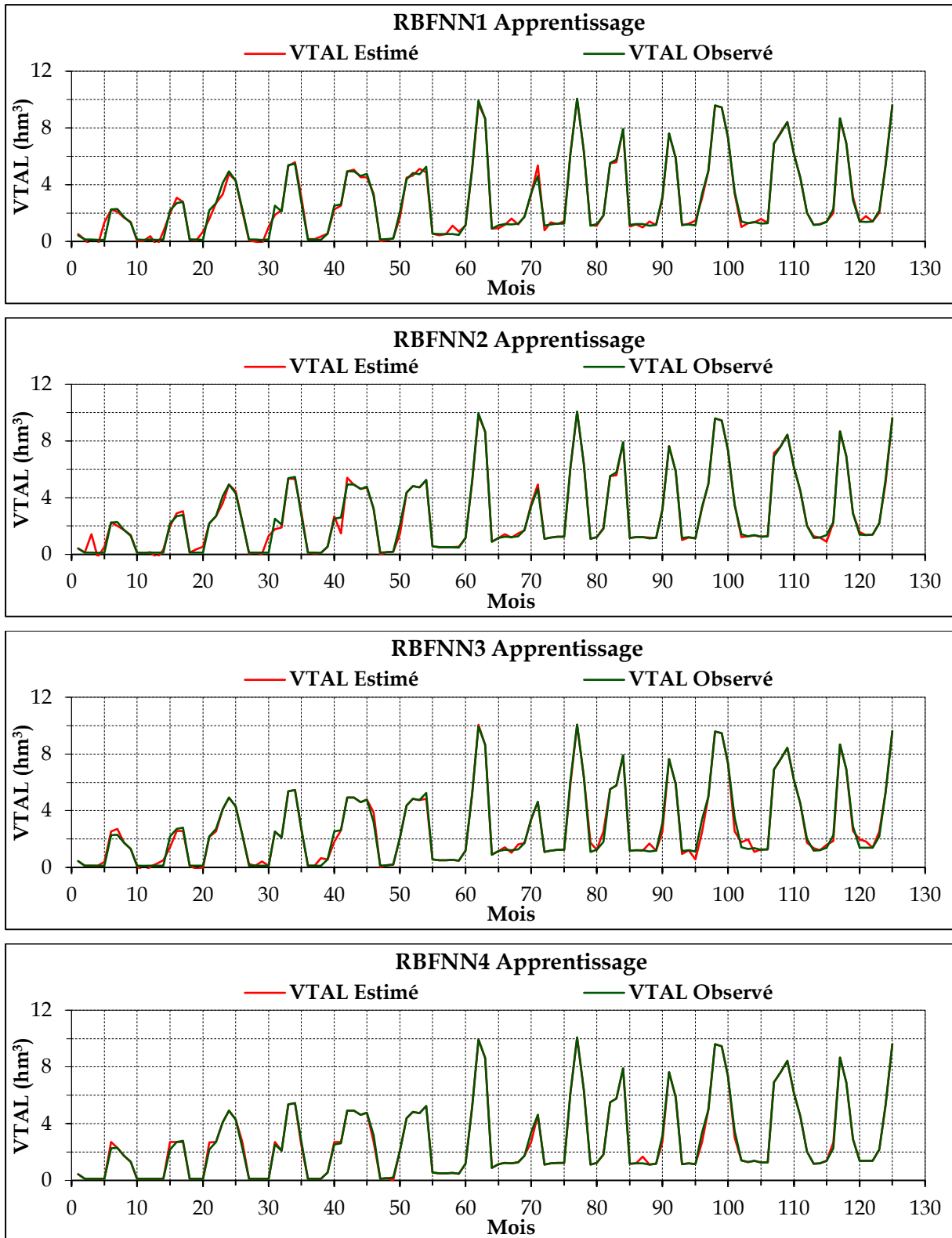


Figure 5-16 : Evolution des volumes totaux alloués observés et estimés par les modèles RBFNN1, RBFNN2, RBFNN3 et RBFNN4 en phase d'apprentissage

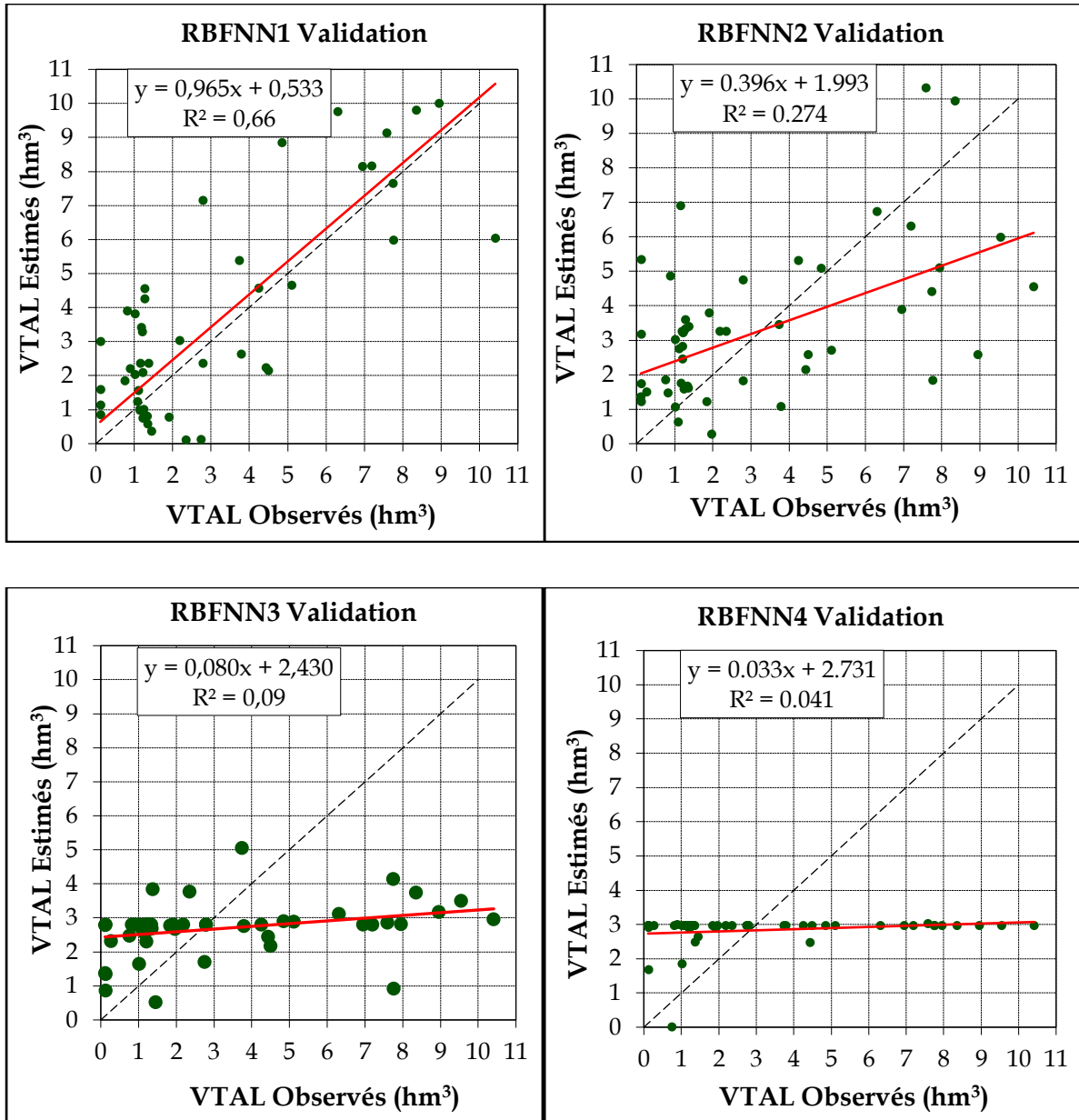


Figure 5-17 : Corrélation des volumes totaux alloués observés et estimés par les modèles RBFNN1, RBFNN2, RBFNN3 et RBFNN4 en phase de validation

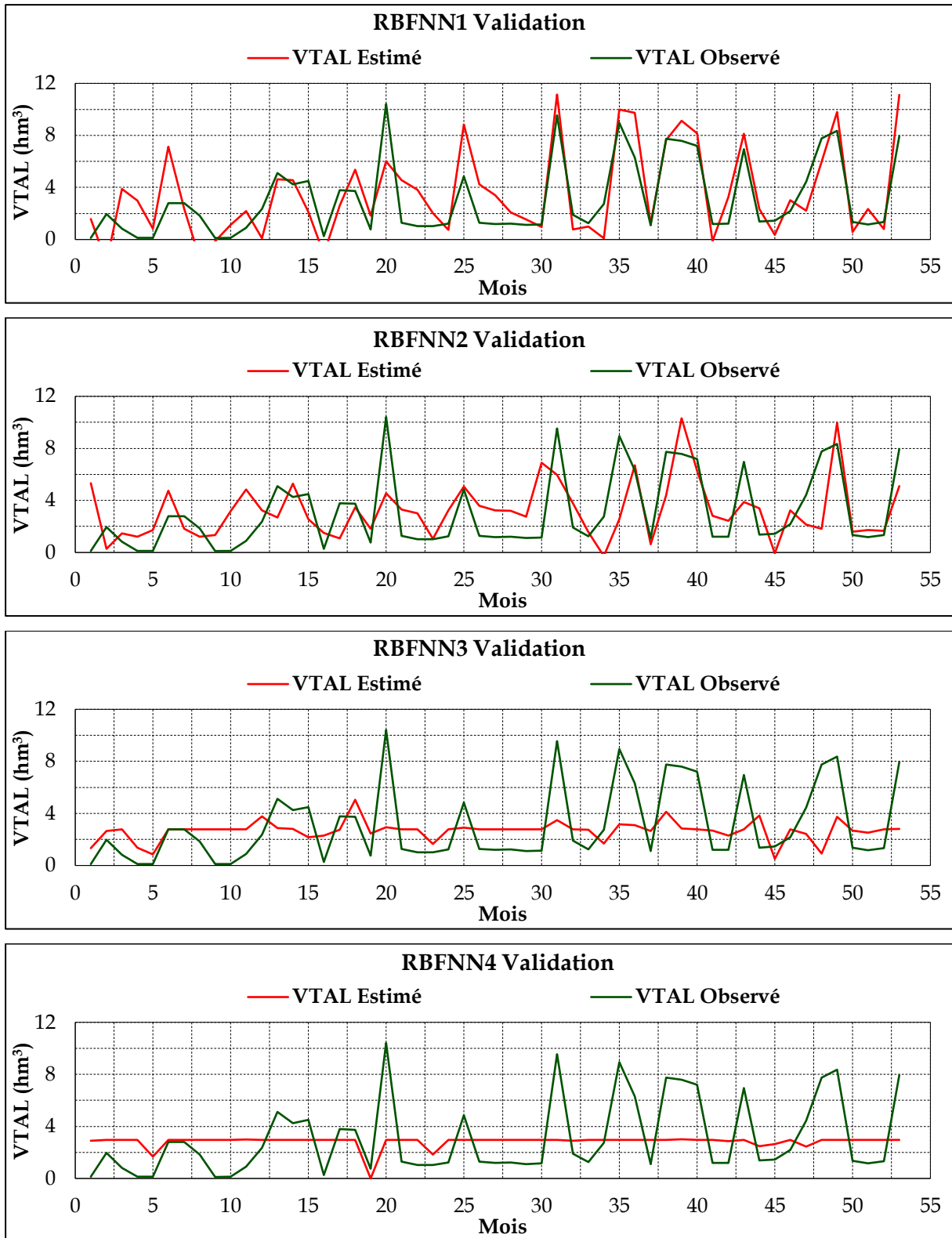


Figure 5-18 : Evolution des volumes totaux alloués observés et estimés par les modèles RBFNN1, RBFNN2, RBFNN3 et RBFNN4 en phase de validation

5.9.2.4 Comparaison des résultats obtenus par les réseaux de neurones

Le [Tableau 5-10](#) résume les performances obtenues par les trois réseaux de neurones retenus à savoir : MLPNN1, GRNN1 et RBFNN1 en phase d'apprentissage et en phase de validation.

Tableau 5-10 : Comparaison des performances des modèles MLPNN1, GRNN1 et RBFNN1

Modèles	Apprentissage				Validation			
	<i>NSE</i>	<i>R</i>	<i>RMSE</i>	<i>MAE</i>	<i>NSE</i>	<i>R</i>	<i>RMSE</i>	<i>MAE</i>
MLPNN1	0.920	0.959	0.756	0.548	0.661	0.869	1.670	1.308
GRNN1	0.996	0.998	0.162	0.050	0.684	0.876	1.614	1.157
RBFNN1	0.990	0.995	0.265	0.164	0.508	0.816	2.012	1.645

Tout d'abord nous constatons que l'utilisation de deux variables endogènes ($VTAL_{t-2}$ et $VTAL_{t-1}$) a permis d'améliorer les performances des modèles développés. En outre, en comparant les trois modèles MLPNN1, GRNN1 et RBFNN1 nous remarquons qu'ils possèdent de meilleures performances en phase d'apprentissage qu'en phase de validation avec un léger avantage du modèle GRNN1 puisqu'il est le plus compétitif avec ($R=0,876$ et $NSE=0,684$) les plus élevées et ($RMSE=1,614$ et $MAE=1,157$) les plus faibles.

Enfin, nous pouvons conclure que les trois modèles de réseaux de neurones retenus peuvent être un outil efficace pour la prévision des volumes totaux à allouer mensuellement en utilisant les cinq variables explicatives pris en considération à savoir : le volume total alloué au temp (t-1), le volume total alloué au temp (t-2), la pluie au temp (t-1), l'évaporation au temp (t-1), et l'apport au temp (t-1) ($VTAL_{t-1}$, $VTAL_{t-2}$, PLU_{t-1} , $EVAP_{t-1}$, APP_{t-1}).

5.10 Conclusion

A la lumière des résultats obtenus par les trois modèles à base des réseaux de neurones à savoir : MLPNN, GRNN et RBFNN nous pouvons conclure ce qui suit :

- Les trois modèles peuvent être considérés comme un outil performant pour la prévision des volumes totaux alloués.
- Les résultats numériques obtenus, démontrent que les performances des modèles dépendent du nombre des variables utilisées comment entrées (inputs) et le meilleur modèle est toujours obtenu en incluant les cinq variables sélectionnées par la technique du Gamma Test.
- Les résultats obtenus en phase de validation montrent que le modèle GRNN1 est le meilleur avec ($R=0,876$ et $NSE=0,684$) les plus élevées et ($RMSE=1,614$ et $MAE=1,157$) les plus faibles, suivi par le modèle MLPNN1 et enfin le modèle RBFNN1 (Tab. 5-10).
- Les résultats obtenus sont concluants et la technique des réseaux de neurones semble être une solution alternative pour la prévision des volumes totaux à allouer.

Conclusion générale

L'idée de développer un modèle de prédiction du volume total à allouer à l'alimentation en eau potable et à l'irrigation à partir des barrages réservoirs a été motivée par l'indisponibilité de l'eau, particulièrement dans les zones à forte demande des régions à vocation agricole. Ce manque d'eau est dû essentiellement à la réduction des volumes stockés causée par l'envasement et l'irrégularité des précipitations.

La gestion de ces volumes limités s'appuie essentiellement sur la courbe « capacité-hauteur ». En effet, cette courbe doit être mise à jour périodiquement par la réalisation de levés topo-bathymétriques qui non seulement sont onéreux et laborieux mais aussi consomment beaucoup de temps.

Dans cet optique, une méthodologie susceptible d'apporter une contribution à l'optimisation de la taille des levés topo-bathymétriques afin d'augmenter les fréquences de réalisation et par conséquent celles d'actualisation de la courbe « capacité-hauteur » des barrages est établie. Par la suite un modèle de prédiction du volume total à allouer est développé sur la base de la technique des réseaux de neurones artificiels.

Le barrage de Hammam Debagh est pris dans le cadre de cette thèse comme un cas d'application de la méthodologie citée ci-dessus.

Tenant compte des résultats de l'unique levé topo-bathymétrique réalisé en 2004 (taux d'envasement annuel de $0,53 \text{ hm}^3$) nous avons procédé à la mise à jour de la base des données mensuelles sur une période de vingt et un (21) années d'exploitation. La démarche préconisée consiste en un réajustement mensuel des volumes, des surfaces et des évaporations et puis l'établissement du bilan hydraulique de la retenue en utilisant le principe de la conservation des volumes d'eau.

Dans le but de réduire la taille des levés topo-bathymétriques afin d'augmenter leurs fréquences de réalisation une analyse géostatistique a été appliquée sur l'ensemble des 32164 valeurs originales du levé topo-bathymétrique et les dix-neuf (19) combinaisons aléatoires choisies. Les quatre étapes de l'analyse géostatistique (analyse exploratoire, analyse variographique et Krigeage/validation croisée) ont montré que la réduction

de la densité de mesure de 42 à 13 points ha⁻¹, (32 164 points à un optimum de 10 000 points) n'a qu'un faible impact sur les critères de fiabilité géostatistiques, tant en calage qu'en validation.

Cette procédure géostatistique d'optimisation de la taille a été validée par la comparaison des courbes « capacité-hauteur » de 12 combinaisons de la taille optimale retenue (10 000 points).

L'utilisation des réseaux de neurones artificiels pour la gestion des barrages est en effet bien adaptée, spécialement pour la prévision du volume total à allouer, grâce à leur robustesse, leur capacité à traiter des données de différentes natures et surtout leurs capacités d'apprentissage. Dans un premier temps, nous avons présenté les trois familles de réseaux de neurones les plus fréquemment utilisées, les perceptrons multicouches (MLPNN), les réseaux dites à fonction de régression générale (GRNN) et les réseaux à fonction radiale de base (RBFNN).

Dans une seconde étape pour élaborer un modèle facile à utiliser comportant un nombre optimal de variables d'entrée, ces dernières ont été réduites à 50% (de 10 à 5) en appliquant la technique du Gamma Test. Les variables prépondérantes retenues sont $VTAL_{t-1}$, PLU_{t-1} , $EVAP_{t-1}$, APP_{t-1} et $VTAL_{t-2}$ respectivement. Les différentes combinaisons de ces variables significatives nous ont permis à l'aide de la technique du Gamma Test de sélectionner quatre vecteurs d'entrée : le vecteur à cinq entrées ($VTAL_{t-1}$, $VTAL_{t-2}$, PLU_{t-1} , $EVAP_{t-1}$, APP_{t-1}), le vecteur à quatre entrées ($VTAL_{t-2}$, PLU_{t-1} , $EVAP_{t-1}$, APP_{t-1}), le vecteur à trois entrées ($VTAL_{t-2}$, $EVAP_{t-1}$, APP_{t-1}) et le vecteur à deux entrées ($EVAP_{t-1}$, APP_{t-1}). Bien entendu, le vecteur de la couche sortie est le volume total alloué au temps t ($VTAL_t$). Les trois modèles (MLPNN, GRNN et RBFNN) sont des outils performants pour la prévision du volume total alloué.

Cependant il est important de signaler qu'en phase de validation, le modèle GRNN1 (cinq variables d'entrée) est considéré comme le meilleur avec ($R=0,876$ et $NSE=0,684$) les plus élevés et ($RMSE=1,614$ et $MAE=1,157$) les plus faibles, suivi par le modèle MLPNN1 et enfin le RBFNN1.

Bibliographie

AGENCE NATIONALE DES BARRAGES (ANB). Levés bathymétriques des barrages en exploitation. Rapport technique: Barrage Hammam Debagh, Guelma, Algérie, Mars, 2005. 29 p.

AGENCE NATIONALE DES BARRAGES ET TRANSFERTS (ANBT). Levé bathymétrique avant dragage. Rapport technique: Barrage Foum el Gherza, Biskra, Algérie, Mai, 2015. 20 p.

AGENCE NATIONALE DES BARRAGES ET TRANSFERTS (ANBT). Fiches techniques des barrages en exploitation: Barrages de l'Est, Alger, 2010. 21 p. Format PPT.

AGENCE NATIONALE DES RESSOURCES HYDRAULIQUES. Etude générale des crues du Nord de l'Algérie : Modélisation des débits de crue. Rapport technique, Juillet, 2008. 73 p.

ALBERGEL, J., GUEGUEN, N., PEPIN, Y. Comment faire le suivi hydrologique d'un petit barrage. Note technique. Tunis : ORSTOM, 1997. 30 p.

ANCTIL, F., PERRIN, C., ANDREASSIAN, V. Impact of the length of observed records on the performance of ANN and of conceptual parsimonious rainfall-runoff forecasting models. *Environmental Modelling and Software*, April, 2004, Vol. 19, n° 4, p. 357-368. DOI : [10.1016/S1364-8152\(03\)00135-X](https://doi.org/10.1016/S1364-8152(03)00135-X)

ARMSTRONG, M. *Basic linear geostatistics*. Berlin : Springer Verlag, 1998. 153 p. ISBN 3-540-61845-7.

ARNAUD, M. and EMERY, X. Estimation et interpolation spatiale : Méthodes déterministes et méthodes géostatistiques. Paris : Hermès, 2000. 216 p.

ATKINSON, P.M. And LLOYD, C.D. Non-stationary variogram models for geostatistical sampling optimisation: An empirical investigation using elevation data. *Computers & Geosciences*, October, 2007, Vol. 33, n°10, p. 1285-1300.

DOI : [10.1016/j.cageo.2007.05.011](https://doi.org/10.1016/j.cageo.2007.05.011).

BARBOSA, C.C.F., NOVO, E.M.L.M., MELACK, J.M., FREITAS, R.M., FILHO, W.P. Metodologia de análise da dinâmica de área e volumen inundável: o exemplo da várzea do lago grande de curuai. *Revista Brasileira de Cartografia*, Mars, 2006, Vol. 58, p. 200-210.

BECHLER, A., ROMARY, T., JEANNÉE, N., DESNOYERS, Y. Geostatistical sampling optimization of contaminated facilities. *Stochastic Environmental Research and Risk Assessment*, December, 2013, Vol. 27, n° 8, p. 1967-1974.

DOI : [10.1007/s00477-013-0731-0](https://doi.org/10.1007/s00477-013-0731-0).

BELLO-PINEDA, J. And HERNÁNDEZ-STEFANONI, J.L. Comparing the performance of two spatial interpolation methods for creating a digital bathymetric model of the Yucatan submerged platform. *Pan-American Journal of Aquatic Sciences*, September, 2007, Vol. 2, n° 3, p. 247-254.

BEN MAMMOU, A. et LOUATI, M.H. Évolution temporelle de l'envasement des retenues de barrages de Tunisie. *Revue des sciences de l'eau*, Janvier, 2007, Vol. 20, n° 02, p. 201-210. DOI : [10.7202/015813ar](https://doi.org/10.7202/015813ar).

BENYOUNES, A., ABDELAZIZ, M., PEPIN, Y., ABERKAN, M., ALBERJEL, J. Simulation du bilan hydro-sédimentaire d'un barrage collinaire en zone marneuse méditerranéenne (Cas du barrage Saboun, Maroc). *Revue des sciences de l'eau*, Janvier, 2009, Vol. 22, n° 4, p. 487-504. DOI : [10.7202/038327ar](https://doi.org/10.7202/038327ar).

BESSENASSE, M., KETTAB, A., PAQUIER, A., GALEAS, G., RAMEZ, P. Simulation numérique de la sédimentation dans les retenues de barrages : cas de la retenue de Zardezas, Algérie. *Revue des sciences de l'eau*, 2003, Vol. 16, n° 1, p. 103-122. DOI : [10.7202/705500ar](https://doi.org/10.7202/705500ar).

BLUE MARBLE GEOGRAPHICS. Global Mapper V 17.0.2. Hallowell, Maine, U.S.A, 2015.

BULMER, M. G. *Principles of statistics*. New York : Dover publications, 1979. 252 p. ISBN: 0486637603 / 9780486637600.

BURGESS, T.M. And WEBSTER, R. Optimal Interpolation and isarithmic Mapping of Soil Properties. I- The Semi-Variogram and Punctual Kriging. *European Journal of Soil Sciences*, June, 1980, Vol. 31, n° 2, p.315-331. DOI: [10.1111/j.1365-2389.1980.tb02084.x](https://doi.org/10.1111/j.1365-2389.1980.tb02084.x)

BURGESS, T.M. And WEBSTER, R. Optimal Interpolation and Isarithmic Mapping of Soil Properties. II- Block Kriging. *European Journal of Soil Sciences*, June, 1980a, Vol. 31, n° 2, p.333-341. DOI : [10.1111/j.1365-2389.1980.tb02085.x](https://doi.org/10.1111/j.1365-2389.1980.tb02085.x)

CAMBARDELLA, C.A., MOORMAN, T.B., NOVAK, J.M., PARKIN, T.B., KARLEN, D.L., TURCO, R.F., KONOPKA, A.E. Field-scale variability of soil properties in central Iowa soils. *Soil Science Society of America Journal Abstract*, Sept, 1994. Vol. 58 n° 5, p. 1501-1511. DOI : [10.2136/sssaj1994.03615995005800050033x](https://doi.org/10.2136/sssaj1994.03615995005800050033x).

CANCELLIERE, A., GIULIANO, G., ANCARANI, A., ROSSI, G. A neural networks approach for deriving irrigation reservoir operating rules. *Water Resources Management*, February, 2002, Vol. 16, n° 1, p. 71-88. DOI : [10.1023/A:1015563820136](https://doi.org/10.1023/A:1015563820136).

CHANDRAMOULI, V. And DEKA, P. Neural network based decision support model for optimal reservoir operation. *Water Resources Management*, August, 2005, Vol. 19, n° 4, p. 447-464. DOI : [10.1007/s11269-005-3276-2](https://doi.org/10.1007/s11269-005-3276-2).

CHANDRAMOULI, V. And RAMAN, H. Multireservoir modeling with dynamic programming and neural networks. *Water Resources Planning and Management ASCE*, April, 2001, Vol.127, n° 2, p. 89-98. DOI:[10.1061/\(ASCE\)0733-9496\(2001\)127:2\(89\)](https://doi.org/10.1061/(ASCE)0733-9496(2001)127:2(89)).

CHANG, F.J., CHEN, L., CHANG, L.C. Optimizing the reservoir operating rule curves by genetic algorithms. *Hydrological Processes*, March, 2005a, Vol. 19, n° 11, p. 2277-2289. DOI: [10.1002/hyp.5674](https://doi.org/10.1002/hyp.5674).

CHANG, F.J., CHIANG, Y.M., CHANG, L.C. Multi-step-ahead neural networks for flood forecasting. *Hydrological Sciences Journal*, February, 2007, Vol. 52, n°1, p. 114-130. DOI : [10.1623/hysj.52.1.114](https://doi.org/10.1623/hysj.52.1.114).

CHANG, L.C. And CHANG, F.J. Intelligent control for modelling of real-time reservoir operation. *Hydrological Processes*, June, 2001, Vol. 15, n° 9, p. 1621-1634. DOI : [10.1002/hyp.226](https://doi.org/10.1002/hyp.226).

CHANG, Y.T., CHANG, L.C, CHANG, F.J. Intelligent control for modelling of real-time reservoir operation. part II: artificial neural network with operating rule curves. *Hydrological Processes*, April, 2005b, Vol. 19, n° 7, p. 1431-1444.

DOI : [10.1002/hyp.5582](https://doi.org/10.1002/hyp.5582).

CHARLES, J.A. The engineering behavior of fill materials and its influence on the performance of embankment dams. *Dams and Reservoirs*, March, 2009, Vol. 19, n° 1, p. 21-33. DOI:[10.1680/dare.2009.19.1.21](https://doi.org/10.1680/dare.2009.19.1.21).

CHAUVET, P. *Aide-mémoire de la géostatistique linéaire*. Paris : Presses de Ecole des Mines de Paris, 2006. 286 p.

CHRISTAKOS, G. On the problem of permissible covariance and variogram models. *Water Resources Research*, February, 1984, Vol. 20, n° 2, p. 251-265.

DOI : [10.1029/WR020i002p00251](https://doi.org/10.1029/WR020i002p00251).

CIGIZOGLU, H. K., And ALP, M. Generalized regression neural network in modelling river sediment yield. *Advances in Engineering Software*, February, 2006, Vol. 37, n° 2, p. 63-68. DOI : [10.1016/j.advengsoft.2005.05.002](https://doi.org/10.1016/j.advengsoft.2005.05.002).

CIGIZOGLU, H.K. Generalized regression neural network in monthly flow forecasting. *Civil Engineering and Environmental Systems*, August, 2006, Vol. 22, n° 2, p. 71-84. DOI : [10.1080/10286600500126256](https://doi.org/10.1080/10286600500126256).

CORCORAN, J., WILSON, I., WARE, J. Predicting the geo-temporal variation of crime and disorder. *International Journal of Forecasting*, 2003, Vol. 19, p. 623-634.

DOI: [10.1016/S0169-2070\(03\)00095-5](https://doi.org/10.1016/S0169-2070(03)00095-5).

CRESSIE, N.A.C. *Statistics for Spatial Data*. Revised reprint of the 1991 edition. New York : John Wiley & Sons, Inc, 1993.900 p. ISBN : 9780471002550.

DOI : [10.1002/9781119115151](https://doi.org/10.1002/9781119115151).

- CURTARELLI, M., Leão, J., Ogashawara, I., Lorenzzetti, J., Stech, J. Assessment of Spatial Interpolation Methods to Map the Bathymetry of an Amazonian Hydroelectric Reservoir to Aid in Decision Making for Water Management. *ISPRS International Journal of Geo-Information*, February, 2015, Vol. 4, n° 1, p. 220-235. DOI : [10.3390/ijgi4010220](https://doi.org/10.3390/ijgi4010220).
- DEBBARMA, S., REGULWAR, D.G. Inflow Prediction by Different Neural Network Architectures: A Case Study Debbarma Sentu. *International Journal of Earth Sciences and Engineering*, October, 2011, Vol. 04, n° 06 SPL, p. 225-230.
- DREYFUS, G., MARTINEZ, J.M., SAMUELIDES, M., GORDON, M.B., BADRAN, F., THIRIA, S., HERAULT, L. *Réseaux de neurones : Méthodologie et applications*. Paris : Eyrolles, 2004. 408 p. ISBN : 2-212-11464-8
- DURRANT, P.J. Win GammaTM : A Non-linear Data Analysis and Modeling Tool with Applications to Flood Prediction. 261 p. PhD thesis, Department of Computer Science : Wales, Cardiff University : 2001.
- EL-SHAFIE, A., TAHA, M.R., NOURELDIN, A. A neuro-fuzzy model for inflow forecasting of the Nile river at Aswan high dam. *Water Resources Management*, March, 2007, Vol. 21, n° 3, p. 533-556. DOI : [10.1007/s11269-006-9027-1](https://doi.org/10.1007/s11269-006-9027-1).
- EMERY, X. *Géostatistique Linéaire*. Paris : Ecole des Mines de Paris, 2001, 405 p.
- EVANS, D. and JONES, A. J. A proof of the gamma test. *Proceedings of Royal Society A: Mathematical Physical and Engineering Sciences*, November 2002, Vol. 458, n° 2027, p. 2759-2799. DOI : [10.1098/rspa.2002.1010](https://doi.org/10.1098/rspa.2002.1010)
- EVANS, D., JONES, A. J., SCHMIDT, W. M. Asymptotic moments of near neighbour distance distributions. *Proceedings of Royal Society A: Mathematical Physical and Engineering Sciences*, 2002, Vol. 458, n° 2028, p. 2839-2829. ISSN 1364-5021.
- FURNANS, J., and AUSTIN, B. Hydrographic survey methods for determining reservoir volume. *Environmental Modelling Software*, February, 2008, Vol. 23, n° 2, p.139-146. DOI : [10.1016/j.envsoft.2007.05.011](https://doi.org/10.1016/j.envsoft.2007.05.011).

GRATTON, Y. Le Krigeage : La méthode optimale d'interpolation spatiale [en ligne]. 2002. [Consulté le 13 décembre 2016] Disponible sur : www-ia.lip6.fr/~hnguyen/coursan/projets/krigeage/krigeage_juillet2002.pdf

HARTKAMP, A.D., DE BEURS, K., STEIN, A., WHITE, J.W. Interpolation Techniques for Climate Variables. *NRG-GIS Series 99-01*. Mexico, D.F. : CIMMYT, January, 1999.

HEDDAM, S., BERMAD, A., DECHEMI, N. Applications of Radial-Basis Function and Generalized Regression Neural Networks for Modeling of Coagulant Dosage in a Drinking Water-Treatment Plant: Comparative Study. *Journal of Environmental Engineering*, December, 2011, Vol. 137, n° 12, p. 1209-1214.

DOI : [10.1061/\(ASCE\)EE.1943-7870.0000435](https://doi.org/10.1061/(ASCE)EE.1943-7870.0000435).

HERA, Á., LOPEZ-PAMO, E., SANTOFIMIA, E., GALLEGO, G., MORALES, R., DURAN-VALSERO, J.J., MURILLO-DIAZ, J.M. A case study of geometric modelling via 3-d point interpolation for the bathymetry of the Rabasa lakes (Alicante, Spain). In : Pardo-Igúzquiza, E., Guardiola-Albert, C., Heredia, J., Moreno-Merino, L., Duran, J., Vargas-Guzman, J., (Eds.), *Mathematics of planet earth, Lecture Notes in Earth System Sciences*. Berlin, Heidelberg : Springer Verlag, 2014, p. 503-506. DOI : [10.1007/978-3-642-32408-6_110](https://doi.org/10.1007/978-3-642-32408-6_110).

HERNANDEZ, J. And EMERY, X. A geostatistical approach to optimize sampling designs for local forest inventories. *Canadian Journal of Forest Research*, August, 2009, Vol. 39, n° 8, p. 1465-1474. DOI : [10.1139/X09-048](https://doi.org/10.1139/X09-048).

HORNIK, K. Approximation capabilities of multilayer feedforward networks. *Neural Networks*, 1991, Vol. 4, n° 2, p. 251-257. DOI : [10.1016/0893-6080\(91\)90009-T](https://doi.org/10.1016/0893-6080(91)90009-T)

HORNIK, K., STINCHCOMBE, M., WHITE, H. Universal approximation of an unknown mapping and its derivatives using multilayer feedforward networks. *Neural Networks*, 1990, Vol. 3, n° 5, p. 551-560.

DOI: [10.1016/0893-6080\(90\)90005-6](https://doi.org/10.1016/0893-6080(90)90005-6)

HOSSEIN ZADEH, T.P., HEYDARI, M., FATHI, P., MAROFI, S., TABARI, H. Numerical model and computational intelligence approaches for estimating flow through rockfill dam. *Journal of Hydrologic Engineering ASCE*, May, 2011, Vol. 17, n° 4, p. 528-536. DOI : [10.1061/\(ASCE\)HE.1943-5584.0000446](https://doi.org/10.1061/(ASCE)HE.1943-5584.0000446).

HSU, K.L., GUPTA, H.V., SOROOSHIAN, S. Artificial neural network modeling of rainfall-runoff process. *Water Resources Research*, October, 1995, Vol. 31, n° 10, p. 2517–2530. DOI : [10.1029/95WR01955](https://doi.org/10.1029/95WR01955).

ISAAKS, E.H. And SRIVASTAVA, R.M. *An introduction to applied geostatistics*. New York : Oxford university press, 1989. 561 p. ISBN 0-19-505012-6, ISBN 0-19-505013-4.

ISMAIL, K. And KEREM C. Reservoir Management Using Artificial Neural Networks. Istanbul : 14th Regional Directorate of DSI (State Hydraulic Works), 2012.

JAIN, S.K., DAS, A., SRIVASTAVA, D.K. Application of ANN for reservoir inflow prediction and operation. *Journal of Water Resources Planning and Management ASCE*, September, 1999, Vol. 125, n° 5, p. 263–271.

DOI : [10.1061/\(ASCE\)0733-9496\(1999\)125:5\(263\)](https://doi.org/10.1061/(ASCE)0733-9496(1999)125:5(263)).

JANG, J.S.R. ANFIS: adaptive-network-based fuzzy inference system. *IEEE Transaction on Systems, Man, and Cybernetics*, May/jun, 1993, Vol. 23, n° 3, p. 665–685.

DOI: [10.1109/21.256541](https://doi.org/10.1109/21.256541).

JONES, A.J. The Win Gamma™ User Guide. Department of Computer Science: Wales, Cardiff University, 2001. 77 p.

JOURNEL, A.G. And HUIJBREGTS, Ch.J. *Mining geostatistics*. London & New York: Academic Press, 1978. 600 p. ISBN 0123910501.

KARUNANITHI, N., WILLIAM, J.G., DARRELL, W., KEN, B. Neural networks for river flow prediction. *Journal of Computing in Civil Engineering*, 1994, Vol. 8, n° 2, p. 201–220. DOI : [10.1061/\(ASCE\)0887-3801\(1994\)8:2\(201\)](https://doi.org/10.1061/(ASCE)0887-3801(1994)8:2(201)).

KISI, Ö. Neural network and wavelet conjunction model for modeling monthly level fluctuations in Turkey. *Hydrological Processes*, July, 2009, Vol. 23, n° 14, p. 2081–2092. DOI : [10.1002/hyp.7340](https://doi.org/10.1002/hyp.7340).

KISI, Ö. River flow modeling using artificial neural networks. *Journal of Hydrologic Engineering ASCE*, January, 2004, Vol. 9, n° 1, p.60–63.

DOI : [10.1061/\(ASCE\)1084-0699\(2004\)9:1\(60\)](https://doi.org/10.1061/(ASCE)1084-0699(2004)9:1(60)).

KISI, Ö. Stream flow forecasting using neuro-wavelet technique. *Hydrological Processes*, September, 2008, Vol. 22, n° 20, p. 4142–4152.

DOI ; [10.1002/hyp.7014](https://doi.org/10.1002/hyp.7014).

KUMAR, A.R.S., SUDHEER, K.P., JAIN, S.K., AGARWAL, P.K. Rainfall-runoff modeling using artificial neural networks: comparison of network types. *Hydrological Processes*, April, 2005, Vol. 19 n° 6, p. 1277–1291.

DOI: [10.1002/hyp.5581](https://doi.org/10.1002/hyp.5581).

KUMMU, M. And VARIS, O. Sediment-related impacts due to upstream reservoir trapping, the Lower Mekong River. *Geomorphology*, March, 2007, Vol. 85, n° 3-4, p. 275- 293. DOI : [10.1016/j.geomorph.2006.03.024](https://doi.org/10.1016/j.geomorph.2006.03.024).

LI, J. and HEAP, A.D. *A Review of Spatial Interpolation Methods for Environmental Scientists*. Canberra: Geoscience Australia, Record 2008/23. 137 p.

LI, Y.P., and HUANG, G.H. Risk analysis and management for water resources systems. *Stochastic Environmental Research and Risk Assessment*, March, 2013, Vol. 27, n° 3, p. 593–597. DOI : [10.1007/s00477-012-0625-6](https://doi.org/10.1007/s00477-012-0625-6).

LOUAMRI, A., MEBARKI, A., LAIGNEL, B. Variabilité interannuelle et intra-annuelle des transports solides de l'Oued Bouhamdane, à l'amont du barrage Hammam Debagh (Algérie orientale). *Hydrological Sciences Journal*, October, 2013, Vol. 58, n°7, p. 1–14. DOI: [10.1080/02626667.2013.824089](https://doi.org/10.1080/02626667.2013.824089).

MALEIKA, W., PALCZYNSKI, M., FREJLICHOWSKI, D. Interpolation methods and the accuracy of bathymetric seabed models based on multibeam echosounder data. In: Pan JS., Chen SM., Nguyen N.T. (eds), *Intelligent Information and Database Systems. ACIIDS 2012*. Berlin, Heidelberg : Springer, Lecture Notes in Computer Science, 2012, Vol. 7198, p. 466–475. DOI : [10.1007/978-3-642-28493-9_49](https://doi.org/10.1007/978-3-642-28493-9_49).

MARCOTTE, D. Géologie et géostatistique minières : géostatistique manière [En ligne]. Montréal: École Polytechnique, cours GLQ3401, 2013. Disponible sur: <<http://www.groupe.polymtl.ca/geo/marcotte/glq3401geo.html>>. Consulté le 18/12/2017)

MAROFI, S., TABARI, H., ABYANEH, H.Z. Predicting spatial distribution of snow water equivalent using multivariate non-linear regression and computational intelligence methods. *Water Resources Management*, March, 2011, Vol. 25, n° 5, p. 1417-1435. DOI 10.1007/s11269-010-9751-4.

MARYLAND GEOLOGICAL SURVEY. Reservoir bathymetry data. Disponible sur : <http://www.mgs.md.gov/publications/data_pages/reservoir_bathymetry.html>. (Consulté le 12 /12/ 2016).

MATHERON, G. and KLEINGELD, W.J. The evolution of geostatistics. In: *Proceedings of the Twentieth International Symposium on the Application of Computers and Mathematics in the Mineral Industries*. Vol. 3: Geostatistics. Johannesburg: SAIMM, 1987, p. 9-12.

MATHERON, G. *La théorie des variables régionalisées et ses applications*. Les cahiers du centre de morphologie mathématique de Fontainebleau, fascicule 5. Paris: Ecole des Mines de Paris, 1970. p. 9-162.

MATHERON, G. *Les variables régionalisées et leur estimation*. Paris : Masson, 1965. 306p.

MATHERON, G. *Traité de géostatistique appliquée*. Mémoires du Bureau de Recherches Géologiques et Minières, Tome I, Vol 14. Paris : Editions Technip, 1962. 333p.

MCCULLOCH, W. S. and PITTS, W. A logical calculus of the ideas immanent in Nervous activity. *Bulletin of Mathematical Biophysics*, 1943, Vol. 5, p. 113-115.

MEBARKI, A. *Hydrologie des bassins de l'Est Algérien : ressources en eau, aménagement et environnement*. Thèse de doctorat d'Etat : Géographie et Aménagement du Territoire, option : Hydrologie : Constantine, Université Mentouri : 2005, 360 p.

MERWADE, V. Effect of spatial trends on interpolation of river bathymetry. *Journal of Hydrology*, June, 2009, Vol. 371, n° 1-4, p. 169-181. DOI10.1016/j.jhydrol.2009.03.026.

MINISTERE DE LA DÉFENSE NATIONALE. *Arrêté du 24 Dhou El Hidja 1423 correspondant au 25 février 2003 fixant les systèmes de référence des coordonnées géographiques, planimétriques et altimétriques relatifs au territoire national; [en ligne]* Journal officiel, n° 30 du 28 Safar 1424 correspondant au 30 Avril 2003. Disponible sur: <https://www.joradp.dz/FTP/jo-francais/2003/F2003030.pdf>. (Consulté le 05/01/2017).

MOGHADDAMNIA, A., GHAFARI-GOUSHEH, M., PIRI, J., AMIN, S., HAN, D. Evaporation estimation using artificial neural networks and adaptive neuro-fuzzy inference system techniques. *Advances in Water Resources*, January, 2009, Vol. 32, n° 1, p. 88-97. DOI : [10.1016/j.advwatres.2008.10.005](https://doi.org/10.1016/j.advwatres.2008.10.005).

MOODY, J. and DARKEN, C.J. Fast learning in networks of locally-tuned processing units. *Neural Computation*, June, 1989, Vol. 1, n° 2, p. 281-294. DOI : [10.1162/neco.1989.1.2.281](https://doi.org/10.1162/neco.1989.1.2.281)

NOORI, R., KARBASSI, A.R., SABAHI, M.S. Evaluation of PCA and Gamma test techniques on ANN operation for weekly solid waste prediction. *Journal of Environmental Management*, November, 2009, Vol. 91, n° 3, p. 767-71. DOI : [10.1016/j.jenvman.2009.10.007](https://doi.org/10.1016/j.jenvman.2009.10.007).

OLEA, R.A. *Geostatistics for Engineers and Earth Scientists*. Boston : Springer , 1999, 303 p. ISBN 978-1-4613-7271-4. DOI: [10.1007/978-1-4615-5001-3](https://doi.org/10.1007/978-1-4615-5001-3).

ORTT, R.A., J.R., VANRYSWICK, S., WELLS, D. Bathymetry and sediment accumulation of triadelphia and rocky gorge reservoirs. File Report No. 07-03. Baltimore : MGS publications, 2007, 19 p.

ORTT, R.A.Jr., KERHIN, R.T., WELLS, D., CORNWELL, J. Bathymetric survey and sedimentation analysis of Loch Raven and Prettyboy reservoirs. File Report No. 99-4. Baltimore : MGS publications, 2000, 56 p.

PUENTE, C.E. And BRAS, R.L. Disjunctive kriging, universal kriging, or no kriging: Small sample results with simulated fields. *Mathematical Geology*, April, 1986, Vol. 18, n° 3, p. 287-305. DOI:[10.1007/BF00898033](https://doi.org/10.1007/BF00898033).

RAKHMATULLAEV, S., MARACHE, A., HUNEAU, F., LE COUSTUMER, P., MASHARIF, B., MOTELICA-HEINO, M. Geostatistical approach for the assessment of the water reservoir capacity in arid regions: a case study of the Akdarya reservoir, Uzbekistan. *Environmental Earth Sciences*, June, 2011, Vol. 63, n° 3, p. 447–460. DOI:10.1007/s12665-010-0711-3.

RAMIREZ, M.C.V., VELHO, H.F.C., FERREIRA, N.J. Artificial neural network technique for rainfall forecasting applied to the São Paulo region. *Journal of Hydrology*, January, 2005, Vol. 301, n°1–4, p. 146–162.

DOI: 10.1016/j.jhydrol.2004.06.028.

REZAEIAN-ZADEH, M., ABGHARI, H., VAN DE GIESEN, N., NIKIAN, A., NIKNIA, N. Maximum daily discharge volume prediction using multi layer perceptron network. *Geophysical Research Abstracts*, EGU General Assembly, 2009, n° 11, p. 3652–3662.

REZAEIAN-ZADEH, M., AMIN, S., KHALILI, D., SINGH, V.P. Daily outflow prediction by multi layer perceptron with logistic sigmoid and tangent sigmoid activation functions. *Water Resources Management*, September, 2010, Vol. 24, n° 11, p. 2673–2688. DOI :10.1007/s11269-009-9573-4.

REZAEIAN-ZADEH, M., TABARI, H., ABGHARI, H. Prediction of monthly discharge volume by different artificial neural network algorithms in semi-arid regions. *Arabian Journal of Geosciences*, July, 2013, Vol. 6, n° 7, p. 2529–2537. DOI:10.1007/s12517-011-0517-y.

ROBERTSON, G. P. Geostatistics in ecology: Interpolating with known variance. *Ecology*, June, 1987, Vol. 68, n° 3, p. 744–748. DOI : 10.2307/1938482

ROBERTSON, G.P. GeoStatistics for the Environmental Sciences. GS+ User's guide Version 5. Gamma Design Software : Plainwell, Michigan : 2000. 152 p.

ROBINSON, T.P., And METTERNICHT, G. Testing the performance of spatial interpolation techniques for mapping soil properties. *Computers and Electronics in Agriculture*, February, 2006, Vol. 50, n° 2, p. 97–108.

DOI : 10.1016/j.compag.2005.07.003.

- ROSSI, G., ANCARANI, A., CANCELLIERE, A. A comparison between neural network Approach and multiple regression procedures for deriving reservoir operation policies. In: Singh VP et al (eds.), *Proceedings of water resources planning and management*. Fort Collins, Water Resources Publications, 1999, p. 148–161
- RUMELHART, D.E., HINTON, G.E., WILLIAMS R.J. Learning internal representations by back-propagating errors. *Nature*, October, 1986b, Vol. 323, n° 6088, p. 533-536. DOI: [10.1038/323533a0](https://doi.org/10.1038/323533a0).
- RUMELHART, D.E., HINTON, G.E., WILLIAMS, R.J. Learning Internal Representations by Error Propagation. In: Rumelhart, D.E., McClelland, J.L. and the PDP Research Group, Eds. *Parallel Distributed Processing: Explorations in the Microstructure of Cognition*. Cambridge : MIT Press, 1986a, Vol. 1, p. 318-362. ISBN:0-262-68053-X
- SEASHOLTZ, M.B. And KOWALSKI, B. The parsimony principle applied to multivariate calibration. *Analytica Chimica Acta*, May, 1993, Vol. 277, n° 2, p. 165-177. DOI: [10.1016/0003-2670\(93\)80430-S](https://doi.org/10.1016/0003-2670(93)80430-S).
- SEPASKHAH, A.R., AHMADI, S.H., SHAHBAZI, R. Geostatistical analysis of sorptivity for a soil under tilled and no-tilled conditions. *Journal of Soil and Tillage Research*, September, 2005, Vol. 83, n° 2, p. 237–245. DOI : [10.1016/j.still.2004.07.019](https://doi.org/10.1016/j.still.2004.07.019)
- SHAMSELDIN, A.Y. Application of neural network technique to rainfall-runoff modeling. *Journal of Hydrology*, December, 1997, Vol. 199, n° 3–4, p. 272–294. DOI : [10.1016/S0022-1694\(96\)03330-6](https://doi.org/10.1016/S0022-1694(96)03330-6).
- SMITH, J.L., HALVORSON, J.J., PAPENDICK, R.I. Using multiple-variable indicator kriging for evaluating soil quality. *Soil Science Society of America Journal*, May, 1993, Vol. 57, n° 3, p. 743–749. DOI:[10.2136/sssaj1993.03615995005700030020x](https://doi.org/10.2136/sssaj1993.03615995005700030020x).
- SPECHT, D.F. A general regression neural network. *IEEE Transactions on Neural Networks*, Nov, 1991, Vol. 2, n° 6, p. 568-576. DOI : [10.1109/72.97934](https://doi.org/10.1109/72.97934).

SRIVASTAVA, R.M. Geostatistics : A toolkit for data analysis, spatial prediction and risk management in the coal industry. *International Journal of Coal Geology*, June, 2013, Vol. 112, p.2–13. DOI : [10.1016/j.coal.2013.01.011](https://doi.org/10.1016/j.coal.2013.01.011).

STEFÁNSSON, A., KONČAR, N., JONES, A.J. A note on the Gamma Test. *Neural Computing & Applications*, September, 1997, Vol. 5, n° 3, p. 131-133
DOI : [10.1007/BF01413858](https://doi.org/10.1007/BF01413858).

SUDHEER, K.P., GOSAIN, A.K., RAMASASTRI, K.S. A data driven algorithm for constructing artificial neural network rainfall-runoff models. *Hydrological Processes*, March, 2002, Vol. 16, n°6, p.1325–1330. DOI : [10.1002/hyp.554](https://doi.org/10.1002/hyp.554)

TABARI, H., MAROFI, S., ABYANEH, H.Z., SHARIFI, M.R. Comparison of artificial neural network and combined models in estimating spatial distribution of snow depth and snow water equivalent in Samsami basin of Irna. *Neural Computing and Applications*, June, 2010b, Vol. 19, n° 4, p. 625–635. DOI : [10.1007/s00521-009-0320-9](https://doi.org/10.1007/s00521-009-0320-9).

TABARI, H., MAROFI, S., SABZIPARVAR, A.A. Estimation of daily pan evaporation using artificial neural network and multivariate non-linear regression. *Irrigation Science*, July, 2010a, Vol. 28, n° 5, p. 399–406. DOI [10.1007/s00271-009-0201-0](https://doi.org/10.1007/s00271-009-0201-0).

TABARI, H., SABZIPARVAR, A.A., AHMADI, M. Comparison of artificial neural network and multivariate linear regression methods for estimation of daily soil temperature in an arid region. *Meteorology and Atmospheric Physics*, January, 2011, Vol. 110, n° 3-4, p. 135–142. DOI : [10.1007/s00703-010-0110-z](https://doi.org/10.1007/s00703-010-0110-z).

TOKAR, A.S. And JOHNSON, P.A. Rainfall-runoff modeling using artificial neural networks. *Journal of Hydrologic Engineering ASCE*, July, 1999, Vol. 4, n° 3, p. 232–239. DOI : [10.1061/\(ASCE\)1084-0699\(1999\)4:3\(232\)](https://doi.org/10.1061/(ASCE)1084-0699(1999)4:3(232)).

TOKAR, A.S. And MARKUS, M. Precipitation-runoff modeling using artificial neural networks and conceptual models. *Journal of Hydrologic Engineering ASCE*, April, 2000, Vol. 5, n° 2, p. 156–161. DOI : [10.1061/\(ASCE\)1084-0699\(2000\)5:2\(156\)](https://doi.org/10.1061/(ASCE)1084-0699(2000)5:2(156)).

WANG, J., YANG, R., BAI, Z. Spatial variability and sampling optimization of soil organic carbon and total nitrogen for Minesoils of the Loess Plateau using geostatistics. *Ecological Engineering*, May, 2015, Vol. 82, n° C, p. 159-164.

DOI : [10.1016/j.ecoleng.2015.04.103](https://doi.org/10.1016/j.ecoleng.2015.04.103).

WORLD COMMISSION ON DAMS. Dams and development: A new framework for decision-making. Report of WCD. London: Earthscan Publications, 2000, 404 p. Disponible sur: https://www.internationalrivers.org/sites/default/files/attached-files/world_commission_on_dams_final_report.pdf.

ZHANG, Y. Artificial neural networks based on principal component analysis input selection for clinical pattern recognition analysis. *Talanta*, August, 2007, Vol. 73, n° 1, p. 68-75. DOI: [10.1016/j.talanta.2007.02.030](https://doi.org/10.1016/j.talanta.2007.02.030).

ZHANG, Y., LI, H., HOU, A., HAVEL, J. Artificial neural networks based on principal component analysis input selection for quantification in overlapped capillary electrophoresis peaks. *Chemometrics and Intelligent Laboratory Systems*, May, 2006, Vol. 82, n° (1-2), p. 165-175. DOI:[10.1016/j.chemolab.2005.08.012](https://doi.org/10.1016/j.chemolab.2005.08.012).