



الجمهورية الجزائرية الديمقراطية الشعبية

RÉPUBLIQUE ALGÉRIENNE DÉMOCRATIQUE ET POPULAIRE

وزارة التعليم العالي والبحث العلمي

Ministère de l'enseignement supérieur et de la recherche scientifique

Université 20 Août 1955-Skikda
Faculté des Technologie
Département de Génie civil
Laboratoire LMGHU

Référence: D012124010D

جامعة سكيكدة 20 أوت 1955
كلية التكنولوجيا
قسم الهندسة المدنية
مخبر المواد، الجيوتقني، السكن و العمران

المرجع: D012124010D

Thèse présentée en vue de l'obtention du diplôme

de Doctorat L.M.D

En Génie Civil

Option : Management De Projets

Présentée par : **DEHCHAR Chemseddine**

Thème

Amélioration et augmentation de la productivité dans la construction

Sous la direction du **Pr. BOUABAZ Mohamed**

Soutenue le : 01/06/2024

Devant le jury composé de :

Pr. HIDJEB	Mustapha	Université 20 Août 1955 Skikda	Président
Pr. BOUABAZ	Mohamed	Université 20 Août 1955 Skikda	Rapporteur
Pr. SASSI BOUDEMAGH	Souad	Université de Constantine 3	Examinatrice
Pr. MEKIMAH	Sabri	Université 20 Août 1955 Skikda	Examineur

Remerciement

Je remercie d'abord ALLAH le tout puissant qui m'a guidé et qui m'a donné la force et la volonté de réaliser ce travail.

Je remercie mon Directeur de thèse, Monsieur Bouabaz Mohamed Professeur au département de génie civil à l'université du 20 Août 1955 - Skikda, d'avoir accepté de diriger ce travail avec beaucoup de disponibilité et d'efficacité, pour m'avoir fait partager son expérience, pour les conseils et encouragements qu'il a su me prodiguer pendant toute la durée de cette thèse.

Je manifeste mes honnêtes remerciements à Monsieur HIDJEB Mustapha Professeur au département de génie civil à l'université du 20 Août 1955 – Skikda, pour l'honneur et le plaisir qu'il m'a fait en acceptant de présider ce jury.

Je remercie vivement Monsieur, MEKIMAH Sabri, Professeur au département des sciences de gestion à l'université du 20 Août 1955 – Skikda, et Madame SASSI BOUDEMAGH Souad Professeur au département de Management de Projets à l'université de Constantine 3, qui ont accepté d'examiner ce travail et être membre de jury. Je leur suis très reconnaissant d'y avoir consacré une partie de leur temps si précieux.

Mes remerciements sont adressés à l'ensemble du corps enseignant en graduation et en post-graduation qui ont contribué à ma formation.

Je présente ici mes sincères remerciements à Messieurs Gahdouni Messaoud et Ziouan Borhane, Directeurs des pôles de l'entreprise Cosider Construction, pour leur aide et leurs précieux conseils.

J'ai également apprécié la compagnie de Mr Bourzem Nabil, Mr Barka Amine et Siari Haidar qui m'ont aidé dans la collecte des données et qui m'ont transmis de judicieux conseils surtout au début de l'élaboration de la présente thèse, qu'ils trouvent ici l'expression de mes sincères remerciements.

Je tiens également à remercier mes collègues de promotion pour discussions scientifiques riches que nous avons partagées.

Enfin, les mots les plus simples étant les plus forts, j'adresse toute ma gratitude à ma famille et en particulier à mes parents, qui ont su m'insuffler le courage et la volonté de me dépasser. Très humblement, je voudrais vous dire merci pour votre amour, et vos sacrifices. Que dieu Vous garde, et vous procure santé, bonheur et longue vie.

Merci à mes sœurs et mon frère Farouk pour leurs soutiens et leurs compréhensions ont été un atout majeur pour la réalisation de ce travail.

Je ne saurais terminer ces remerciements sans citer mes proches. Merci à tous de m'avoir soutenu durant ces années.

Dédicace

Je dédie cette thèse :

A mes très chers parents

A mes très chères sœurs et mon frère

A tous les membres de ma famille et tous

mes proches

A tous mes collègues

Et à tous ceux qui m'ont encouragé et

soutenu

Abstract

The improvement of productivity in the construction sector has been a longstanding challenge, prompting significant research efforts. This study specifically targets the enhancement of productivity in construction projects. To achieve this, a survey involving 150 questionnaires with 16 factors grouped into Three groups, namely: Group 01: Management, Group 02: Worker-Related Factors, Group 03: External Factors, based on a 5-point Likert scale, was conducted to identify the factors influencing productivity. These factors, selected based on their Relative Importance Index (RII), were crucial in understanding productivity dynamics.

In the next step, a dataset comprising 101 data points from different construction sites is used to build and develop two machine learning models, namely: an Artificial Neural Network (ANN) model and a parametric approach-based multiple linear regression model, for the purpose of estimating and predicting construction productivity rates under the influence of the defined factors in the initial phase. In the case of the ANN model, 70% of the dataset was allocated for training, while the remaining 30% was reserved for testing. The model's performance evaluation used the metrics of Mean Squared Error (MSE) and Mean Absolute Percentage Error (MAPE). During the testing phase, the ANN model demonstrated remarkable accuracy, with an MSE value of 0.000906 and a MAPE value of 1.224%. In contrast, the parametric model yielded comparatively lower accuracy, with an MSE of 0.031 and a MAPE of 6.64%.

These results highlight the superiority of the Artificial Neural Network (ANN) model over the parametric approach for predicting and managing productivity rates in local construction projects. The strong predictive capabilities of the ANN model, based on its ability to capture complex data relationships, make it a reliable tool for project stakeholders and construction managers. By leveraging this developed model, based on the identified influential factors, construction projects can make informed decisions, streamline their operations, and ultimately increase their productivity.

This study significantly contributes to the ongoing efforts of the construction industry to enhance productivity. Its findings underscore the potential of advanced machine learning techniques such as ANN in addressing complex construction challenges, ultimately leading to improved project outcomes, cost efficiency, and resource management.

Keywords: Artificial Intelligence, Construction Industry, Artificial Neural Network, Labour Productivity, Parametric Approach

ملخص

تحسين الإنتاجية في قطاع البناء كان تحديًا طويل الأمد، حيث أثار جهود بحث هامة. تستهدف هذه الدراسة بشكل خاص تحسين الإنتاجية في مشاريع البناء. ولتحقيق ذلك، تم إجراء استطلاع يشمل 150 استبيانًا يتضمن 16 عاملاً تم تجميعهم في ثلاث مجموعات هي: المجموعة 01: الإدارة، المجموعة 02: عوامل متعلقة بالعمال، المجموعة 03: عوامل خارجية، باستخدام مقياس ليكرت ذو 5 نقاط، بهدف تحديد العوامل التي تؤثر على الإنتاجية. هذه العوامل، التي تم اختيارها استنادًا إلى مؤشر الأهمية النسبية، كانت ضرورية لفهم ديناميكية الإنتاجية.

في المرحلة التالية، تم استخدام مجموعة بيانات تحتوي على 101 نقطة بيانات من مواقع بناء مختلفة لبناء وتطوير نموذجي تعلم آلي، وهما: نموذج للشبكة العصبية الاصطناعية ونموذج للانحدار الخطي المتعدد للتقدير والتنبؤ بمعدلات الإنتاجية تحت تأثير العوامل المحددة في المرحلة الأولية.

في حالة نموذج الشبكة العصبية الاصطناعية، تم تخصيص 70% من مجموعة البيانات للتدريب، بينما تم حجز الـ 30% المتبقية للاختبار. تم استخدام مقاييس خطأ المربع المتوسط وخطأ المتوسط المئوي المطلق لتقييم أداء النموذج. خلال مرحلة الاختبار، أظهر نموذج الشبكة العصبية الاصطناعية دقة ملحوظة، حيث كانت قيمة خطأ المربع المتوسط تبلغ 0.000906 وقيمة خطأ المتوسط المئوي المطلق تبلغ 1.224%. بالمقابل، أظهر النموذج القائم على الانحدار دقة أقل بالمقارنة، حيث كان خطأ المربع المتوسط يبلغ 0.031 وخطأ المتوسط المئوي المطلق يبلغ 6.64%.

تسلط هذه النتائج الضوء على تفوق نموذج الشبكة العصبية الاصطناعية على النهج القائم على المعلمات للتنبؤ وإدارة معدلات الإنتاجية في مشاريع البناء المحلية. تعتمد القدرات التنبؤية القوية لنموذج الشبكة العصبية الاصطناعية على قدرته على التقاط العلاقات المعقدة بين البيانات، مما يجعله أداة موثوقة لأصحاب المشاريع ومديري البناء. من خلال الاستفادة من هذا النموذج المطور، استنادًا إلى العوامل المؤثرة المحددة، يمكن لمشاريع البناء اتخاذ قرارات مستنيرة، وتبسيط عملياتها، وزيادة إنتاجيتها في النهاية.

تسهم هذه الدراسة بشكل كبير في جهود صناعة البناء المستمرة لتحسين الإنتاجية. تسلط استنتاجاتها الضوء على إمكانيات تقنيات التعلم الآلي المتقدمة مثل الشبكة العصبية الاصطناعية في مواجهة تحديات البناء المعقدة، مما يؤدي في النهاية إلى تحسين نتائج المشروع، وكفاءة تكاليف أفضل، وإدارة أفضل للموارد.

الكلمات المفتاحية: إدارة المشاريع، إنتاجية العمل، مشاريع البناء، الذكاء الاصطناعي، الطريقة البارامترية

RESUME

L'amélioration de la productivité dans le secteur de la construction a été un défi depuis longtemps, suscitant d'importants efforts de recherche. Cette étude cible spécifiquement l'amélioration de la productivité dans les projets de construction. Pour y parvenir, une enquête impliquant 150 questionnaires avec 16 facteurs groupés en Trois groupes notamment : groupe 01 : Management, groupe 02 : Facteurs liés aux travailleurs, groupe 03 : Facteurs externes, basés sur l'échelle de Likert de 5 points, a été menée afin d'identifier les facteurs influençant la productivité. Ces facteurs, sélectionnés en fonction de leur Indice d'Importance Relative (RII), étaient essentiels pour comprendre la dynamique de la productivité.

Dans l'étape qui suit, un ensemble de données comprenant 101 points de données provenant de différents chantiers de construction est utilisé pour construire et développer deux modèles d'apprentissage automatique, à savoir : un modèle de Réseau de Neurones Artificiels (RNA) et un modèle d'approche paramétrique basé sur la régression linéaire multiple, dans le but d'estimer et de prédire les taux de la productivité sous l'influence des facteurs définis en première phase. Dans le cas du modèle de RNA, 70 % de l'ensemble des données ont été alloués à l'entraînement, tandis que les 30 % restants ont été réservés pour les tests. L'évaluation de la performance du modèle a utilisé les métriques de l'Erreur Quadratique Moyenne (MSE) et de l'Erreur Moyenne Absolue en Pourcentage (MAPE). Lors de la phase de test, le modèle de RNA a démontré une précision remarquable, avec une valeur de MSE de 0,000906 et une valeur du MAPE de 1,224%. En revanche, le modèle paramétrique a donné une précision comparativement plus faible, avec un MSE de 0,031 et un MAPE de 6,64%.

Ces résultats soulignent la supériorité du modèle de Réseau de Neurones Artificiels (RNA) par rapport à l'approche paramétrique pour prédire et gérer les taux de productivité dans les projets de construction locaux. Les capacités prédictives robustes du modèle de RNA, basées sur sa capacité à capturer des relations de données complexes, en font un outil fiable pour les parties prenantes des projets et les gestionnaires de la construction. En exploitant ce modèle développé, basé sur les facteurs influents identifiés, les projets de construction peuvent prendre des décisions éclairées, rationaliser leurs opérations et, en fin de compte, accroître leur productivité.

Cette étude contribue de manière significative aux efforts continus de l'industrie de la construction pour améliorer la productivité. Ses conclusions soulignent le potentiel des techniques avancées d'apprentissage automatique telles que le RNA pour relever les défis complexes de la construction, conduisant finalement à des résultats de projet améliorés, à une meilleure efficacité des coûts et à une meilleure gestion des ressources.

Mot clés : Management de projets, Productivité du travail, les projets de construction, Réseau de neurones artificiels, Approche paramétrique.

Table des matières

Introduction générale	i
Domaine de la recherche.....	i
Problématique de la recherche	ii
Cadre de la recherche.....	iii
Structure de la thèse	iii

Première partie Aspect théorique

Chapitre I : Productivité et domaines d'application

I.1. Introduction sur la productivité.....	1
I.1.1. Définitions littéraire	2
I.1.2. Définitions mathématique	3
I.1.2.1. Productivité totale des facteurs « TFP ».....	4
I.1.2.2. Productivité Partielle	4
I.1.2.3. Productivité de travail	5
I.2. Mesure de la productivité.....	6
I.3. Technique de mesure de la productivité.....	10
I.3.1. Observation directe.....	10
I.3.2. Etude de temps	10
I.3.3. Echantillonnage d'activité.....	11
I.3.4. Questionnaire des chefs d'équipes sur le retard	12
I.3.5. Questionnaire des travailleurs sur le retard	12
I.3.6. Méthode d'enregistrement	13

Chapitre II : Etat de l'art

II.1. Introduction	14
II.2. Revue des facteurs affectant la productivité dans la construction	14
II.3. Revue bibliographique sur l'utilisation des Réseaux de Neurones Artificiels dans la construction	23
II.4. Revue bibliographique sur l'approche paramétrique dans la construction	28
II.5. Conclusion.....	30

Chapitre III : Présentation des méthodes de modélisation

III.1. Introduction	31
III.2. Réseaux de Neurones Artificiels	31
III.2.1. Historique	31
III.2.2. Domaine d'applications des Réseaux de Neurones Artificiels.	34
III.2.3. Structure des Réseaux de Neurones Biologique.....	35
III.2.3.1. Réseaux de Neurones Biologique.....	35
III.2.3.1.1. Composition	35
III.2.3.2. Réseaux de Neurones Artificiels.....	36
III.2.3.2.1. Neurone formel	39
III.2.3.3. Comparaison entre réseau de neurones biologique et le réseau de neurones artificiels.....	39

III.2.4. Architecture des Réseaux de Neurones Artificiels.....	40
III.2.4.1. Réseaux statiques.....	40
III.2.4.2. Les réseaux dynamiques.....	41
III.2.5. Les types des Réseaux de Neurones Artificiels.....	42
III.2.5.1. Types des réseaux statiques (non-bouclés)	43
III.2.5.1.1. Réseaux complètement connectés :.....	43
III.2.5.1.2. Réseaux à couches.....	43
III.2.5.1.2.1. Perceptron Monocouche	43
III.2.5.1.2.2. Perceptron Multi Couches (PMC)	44
III.2.5.1.2.3. Les Réseau de neurones à Fonctions de Base Radiales (RBF).....	46
III.2.5.2. Types des réseaux dynamiques (bouclés).....	48
III.2.5.2.1. Les réseaux à apprentissage compétitif.....	48
III.2.5.2.1.1. Vector Quantization (quantification vectorielle)	48
III.2.5.2.1.2. Les réseaux de Kohonen SOM (Self Organizing Map)	49
III.2.5.2.1.3. Learning Vector Quantization.....	50
III.2.5.2.1.4. La théorie des résonances adaptatives (ART).....	50
III.2.5.2.2. Les réseaux récurrents.....	51
III.2.5.2.2.1. Réseaux de Hopfield	51
III.2.5.2.2.2. Réseaux Elman 1990.....	52
III.2.5.2.2.3. Réseaux Jordan 1986	53
III.2.6. Apprentissage des Réseaux de Neurones Artificiels.....	54
III.2.6.1. Types d'apprentissage	54
III.2.6.1.1. Apprentissage supervisé	54
III.2.6.1.2. Apprentissage non-supervisé.....	54
III.2.6.1.3. Apprentissage renforcé.....	55
III.2.6.1.4. Apprentissage hybride.....	55
III.2.6.2. Phénomène de surapprentissage.....	55
III.2.6.3. Phénomène de sous apprentissage	56
III.2.6.4. Algorithmes d'apprentissage.....	57
III.2.6.4.1. Algorithme de rétropropagation	57
III.2.6.5. Règles d'apprentissage.....	60
III.2.6.5.1. Correction d'erreur (la règle delta).....	60
III.2.6.5.2. Règle de Hebb	60
III.2.6.5.3. Règle d'apprentissage compétitif.....	61
III.2.6.5.4. Règle de Boltzmann	61
III.2.7. Modélisation par Réseaux de Neurones Artificiels.....	62
III.2.7.1. Préparation et Introduction des données dans le modèle	62
III.2.7.1.1. Préparation (pré traitement) des données	62
III.2.7.1.1.1. Normalisation des données	62
III.2.7.1.1.2. Standardisation des données	63
III.2.7.1.1.3. Normalisation par échelle décimale.....	63
III.2.7.1.1.4. Préparation des données qualitatif	63
III.2.7.1.2. Introduction (traitement) des données.....	63
III.2.7.2. Performance de modèle	65

III.2.8. Avantages et inconvénient des Réseaux de Neurones Artificiels	67
III.2.8.1. Avantages	67
III.2.8.2. Inconvénients	68
III.3. Introduction à la Méthode Paramétrique	68
III.3.1. Types des modèles paramétriques	68
III.3.1.1. Modèles Analytiques	69
III.3.1.2. Modèles Linéaires	69
III.3.1.2.1. Régression linéaire simple	69
III.3.1.2.2. Régression linéaire multiple (RLM)	70
III.3.1.2.3. L'estimation du modèle de régression	71
III.3.1.2.3.1. L'estimation par la méthode des moindres carrés	71
III.3.1.2.3.1.1. Propriétés des estimateurs	73
III.3.1.2.3.1.2. Critère de jugement de la qualité de l'ajustement d'un modèle	74
III.3.1.2.4. Les différents tests utilisés dans la Régression linéaire multiple	75
III.3.1.2.4.1. Test de normalité :	75
III.3.1.2.4.2. Tests d'homogénéité	77
III.3.1.2.4.3. Test de linéarité et indépendance	78
III.3.1.2.4.4. Intervalle de confiance et Tests d'hypothèses	78
III.3.1.2.4.4.1. Test de nullité d'un paramètre du modèle (Test de Student)	78
III.3.1.2.4.4.2. Test de significativité global (plusieurs coefficients) : Test de Fisher	79
III.3.1.2.4.4.3. Test de nullité de tous les paramètres du modèle:	80
III.3.1.2.5. Analyse de la variance ANOVA	80
III.3.1.2.6. Performance de modèle	81
III.3.1.2.7. Analyse des résidus	81
III.3.1.2.7.1. Normalité des résidus	82
III.3.1.2.7.2. Homoscédasticité des résidus (variance constante)	82
III.3.1.2.7.3. Test de Durbin-Watson d'autocorrélation des erreurs	82
III.3.1.2.7.4. Vérification de Multicolinéarité (par le coefficient VIF)	82
III.4. Conclusion	83

Deuxième partie : De l'aspect théorique vers la pratique

Chapitre IV : Etude de la productivité dans les projets de construction

IV.1. Introduction	84
IV.2. Mesure de la Productivité dans Les Projets De Construction en Algérie	84
IV.2.1. Résultats d'investigation.	87
IV.2.2. Analyse et discussion des résultats	92
IV.3. Identification des facteurs influençant la productivité dans les projets de construction en Algérie	93
IV.3.1. Conception du questionnaire	93
IV.3.2. Type du questionnaire	95
IV.3.3. Taille de l'échantillon (sample size)	95
IV.3.4. Les niveaux d'influence (Echelle de Likert)	96
IV.3.5. Validation du questionnaire	97
IV.3.6. Statistiques relatives aux répondants	98

IV.3.7. Analyse des résultats du sondage	100
IV.3.8. Détermination des facteurs influençant la productivité.....	104
IV.3.8.1. Relatif importance index (RII%)	104
IV.3.8.1.1. Classement des facteurs.....	105
IV.3.8.1.2. Analyse des résultats	106
IV.3.8.2. Etude de significativité des facteurs.....	107
IV.3.9. Détermination des facteurs affectant négativement la productivité de travail dans les projets de construction Algériens.....	108
IV.3.10. Collection des données.....	110
IV.4. Conclusion	110

Chapitre V : Modélisation de la productivité par les Réseaux de Neurones Artificiels

V.1. Introduction.....	111
V.2. Modélisation par Réseau de Neurones Artificiels.....	111
V.2.1. Détermination des données	112
V.2.1.1. Données d'entrée.....	112
V.2.1.2. Données de sortie	112
V.2.2. Préparation des données	113
V.2.2.1. Prétraitement des données	113
V.2.2.1.1. Prétraitement des données (quantitatives) numériques	113
V.2.2.1.2. Prétraitement des données qualitatives.....	113
V.2.2.2. Traitement des données	114
V.2.3. Type d'apprentissage.....	116
V.2.4. Algorithme d'apprentissage	117
V.2.5. Structure et architecture du RNA	117
V.2.5.1. Détermination du nombre des couches et des neurones cachés.	117
V.2.5.2. Détermination des fonctions de transfert.	119
V.2.5.3. La fonction d'entraînement (fonction d'optimisation).	120
V.2.6. Phase d'apprentissage (entraînement).....	120
V.2.6.1. Résultats de la phase d'apprentissage	121
V.2.6.1.1. Poids et biais.....	121
V.2.6.1.2. Evaluation de la performance du modèle dans la phase d'apprentissage.....	124
V.2.7. Phase de test	128
V.2.7.1. Résultats de la phase de test	129
V.2.7.2. Evaluation de la performance du modèle dans la phase test.	130
V.3. Conclusion.....	132

CHAPITRE VI : Modélisation de la productivité par la régression linéaire multiple

VI.1. Introduction.....	133
VI.2. Modélisation par la Régression Linéaire Multiple.....	133
VI.2.1. Détermination des données d'entrées et de sorties.....	134
VI.2.2. Tests d'applicabilité de la Régression Linéaire Multiple (RLM).	134
VI.2.2.1. Test de normalité des données	134

VI.2.2.2. Test d'homogénéité des données	136
VI.2.2.3. Test de l'indépendance des données	138
VI.2.2.4. Test de linéarité des données	140
VI.2.3. Développement du modèle et détermination des paramètres (coefficients) de la RLM.	140
VI.2.4. Tests de significativité des coefficients obtenus	141
VI.2.4.1. Test de significativité global (tous les paramètres du modèle) Test de Fisher ...	141
VI.2.4.2. Test de nullité d'un paramètre du modèle (Test de Student)	142
VI.2.5. Analyse des résidus	142
VI.2.5.1. Normalité des résidus.....	143
VI.2.5.2. Homoscédasticité (variance constante) et linéarité des résidus	144
VI.2.5.3. Test de Durbin-Watson d'autocorrélation des erreurs	145
VI.2.5.4. Vérification de la Multicolinéarité (par le coefficient VIF).....	146
VI.2.6. Evaluation du modèle développé	146
VI.2.6.1. Résultat obtenu dans la phase de développement	146
VI.2.6.2. Evaluation des performances de la phase de développement	149
VI.2.7. Test du modèle développé par des nouvelles données d'entrées	151
VI.2.7.1. Résultat obtenu dans la phase de test.....	151
VI.2.7.2. Evaluation des performances de la phase de test	152
VI.3. Conclusion.	154
VI.4. Discussions des résultats des deux Méthodes (modèle RNA et modèle Paramétrique).....	155
Chapitre VII : Analyse de sensibilité	
VII.1. Introduction.	158
VII.2. Différents méthode d'analyse de sensibilité	158
VII.2.1. Algorithme de Garson	158
VII.2.2. Algorithme d'Olden.....	158
VII.2.3. Dérivée partielle	158
VII.2.4. Méthode de profil	159
VII.2.5. Perturbation d'entrée	159
VII.2.6. Méthode pas à pas.....	159
VII.2.7. Deep connection weight	160
VII.3. Conclusion	162
Conclusion générale	163
Bibliographie.....	165
Annexes.....	176

Liste des figures

Figure I-1 Ferrailage dalle plane.....	7
Figure I-2 Ferrailage radier général	8
Figure I-3 ferrailage linteau voile	8
Figure I-4 Coffrage poteaux.....	9
Figure I-5 Coffrage voile	9
Figure III-1 Structure d'un neurone biologique	36
Figure III-2 Différentes types des fonctions de transferts.....	38
Figure III-3 Structure typique d'un réseau de neurone artificiel.....	38
Figure III-4 Structure d'un neurone formel.	39
Figure III-5 Réseau de neurones artificiel non bouclé.....	41
Figure III-6 Réseau de neurones artificiel bouclé	42
Figure III-7 Cartographie des types des Réseaux de Neurones Artificiels	42
Figure III-8 Structure d'un Réseau de neurone artificiel complètement connectés (Principe et al, 2010).....	43
Figure III-9 Perceptron Monocouche.....	44
Figure III-10 Réseau de Neurone Multicouches	46
Figure III-11 Représentation graphique de la fonction à base radiale de de type gaussienne	47
Figure III-12 Réseau de Neurones Artificiel à Fonction de Base Radiale (RFB).....	48
Figure III-13 Principe de fonctionnement de la carte de Kohonen	50
Figure III-14 Réseaux Hopfield	52
Figure III-15 Réseau Elman	53

Figure III-16 Réseau Jordan.....	53
Figure III-17 Phénomène de surapprentissage	56
Figure III-18 Phénomène de sous apprentissage.....	57
Figure III-19 Histogramme de fréquence.....	76
Figure III-20 Boite à moustaches (Box-plot).....	76
Figure III-21 Diagramme Q-Q et Droite de Henry	77
Figure IV-1 Résultats de la productivité de la tâche du coffrage entre groupes de même chantier et entre les différents chantiers	89
Figure IV-2 Résultats de la productivité de la tâche du ferrailage entre groupes de même chantier et entre les différents chantiers	92
Figure IV-3 Représentation graphique des statistiques des postes de travail des participants.....	99
Figure IV-4 Représentation graphique des statistiques de l'expérience professionnelles des participants.....	100
Figure IV-5 Histogramme des réponses Groupe 01 : Facteurs liés à la gestion (management)	102
Figure IV-6 Histogramme des réponses Groupe 02: Facteurs liés aux ouvriers.....	103
Figure IV-7 Histogramme des réponses Groupe 03: Facteurs externes.....	104
Figure IV-8 Résultats de RII % de chaque facteur	106
Figure V-1 Schéma simplifié représente les données d'entrées et les données de sortie dans le modèle RNA	112
Figure V-2 Architecture du Réseau de Neurones Artificiels choisi.....	118
Figure V-3 Les fonctions de transferts utilisées.....	119
Figure V-4 Droite de régression du RNA durant la phase d'apprentissage.....	127
Figure V-5 Productivité Réelle et Productivité Estimée dans la phase d'apprentissage ...	128

Figure V-6 Droite de régression du RNA durant la phase du Test	131
Figure V-7 Productivité Réelle et Productivité Estimée dans la phase du Test.....	132
Figure VI-1 Histogramme de fréquence des résidus	143
Figure VI-2 Diagramme de répartition Gaussienne des résidus	144
Figure VI-3 Distribution des résidus par rapport à la réponse estimée.....	145
Figure VI-4 Droite de régression dans la phase d'apprentissage du modèle de la RLM ..	150
Figure VI-5 Les sorties du modèle de la RLM VS les résultats réels de la productivité dans la phase d'apprentissage	151
Figure VI-6 Droite de régression dans la phase test du modèle RLM	153
Figure VI-7 Les sorties du modèle de la RLM VS les résultats réels de la productivité dans la phase de test.....	154
Figure VI-8 Résultats de la productivité des deux modèles dans la phase du test.....	156
Figure VI-9 Présentation des performances des deux modèles dans la phase du test.....	157
Figure VII-1 L'importance relatif de chaque facteur.....	161

Liste des tableaux

Tableau II-1 Résumé des résultats des travaux précédents	21
Tableau III-1 Comparaison entre le Réseau de Neurones Biologique et le Réseau de Neurones Artificiel	40
Tableau IV-1 Données de la productivité Chantier 01 W. Skikda.....	85
Tableau IV-2 Données de la productivité Chantier 02 W. Constantine.....	85
Tableau IV-3 Données de la productivité Chantier 03 (N'sigha) W. Khenchela.....	86
Tableau IV-4 Données de la productivité Chantier 04 (Baghai) W. Khenchela.....	86
Tableau IV-5 Les facteurs utilisés dans le questionnaire.....	94
Tableau IV-6 Niveau d'influence et poids de l'échelle de Likert.....	96
Tableau IV-7 Les zones des réponses selon l'échelle Likert.....	97
Tableau IV-8 Résultat d'Alpha Cronbach pour l'ensemble des facteurs.....	97
Tableau IV-9 Résultat d'Alpha Cronbach pour chaque facteur.....	98
Tableau IV-10 Statistiques relatifs aux postes de travaux des participants.....	99
Tableau IV-11 Statistique des participants selon leurs expériences professionnelles.....	100
Tableau IV-12 Résultats des réponses du Groupe 01 : (Facteurs liés à la gestion (management).....	101
Tableau IV-13 Résultats des réponses du Groupe 02 : (Facteurs liés aux ouvriers).....	102
Tableau IV-14 Résultats des réponses du Groupe 03 : (Facteurs Externe).....	103
Tableau IV-15 Résultats de RII% et classement des facteurs.....	105
Tableau IV-16 Résultats de test T de Student.....	107
Tableau IV-17 Facteurs avec une moyenne de réponse entre l'intervalle de Grande et Très Grande influence.....	108

Tableau IV-18 Liste des facteurs sélectionnés.....	109
Tableau V-1 Transformation des données qualitatives en données Numérique	113
Tableau V-2 Sous base de données de sortie d'apprentissage.....	114
Tableau V-3 Sous base de données de sortie de test.....	116
Tableau V-4 Valeurs des poids et biais entre la couche d'entrée et la première couche cachée	122
Tableau V-5 Valeurs des poids et biais entre la première couche cachée et la deuxième couche cachée.....	123
Tableau V-6 Valeurs des poids et biais entre la deuxième couche cachée et la couche de sortie.....	123
Tableau V-7 Résultats du modèle et résultats réels durant la phase d'apprentissage	124
Tableau V-8 Performances du modèle dans la phase d'apprentissage	126
Tableau V-9 Résultats du modèle et résultats réels durant la phase du Test	129
Tableau V-10 Performances du modèle dans la phase du Test.....	130
Tableau VI-1 Test d'asymétrie et aplatissement de normalité des données d'origine	135
Tableau VI-2 Test d'asymétrie et aplatissement de normalité des données transformées	136
Tableau VI-3 Test de Levene d'homogénéité des données	137
Tableau VI-4 Test de Levene d'homogénéité des données après transformation de la variable dépendante	138
Tableau VI-5 Matrice de corrélation des données	139
Tableau VI-6 Estimation des coefficients (paramètres) du modèle	141
Tableau VI-7 ANOVA du modèle développé.....	141
Tableau VI-8 Résultat du Durbin-Watson	145
Tableau VI-9 Vérification de la multicolinéarité par VIF	146

Tableau VI-10 Résultat du modèle de la régression multiple dans la phase d'apprentissage (développement).....	147
Tableau VI-11 Performance du modèle de la régression Multiple dans la phase de développement (apprentissage).....	149
Tableau VI-12 Qualité d'ajustement du modèle de la RLM dans la phase de développement	150
Tableau VI-13 Résultat du modèle de la régression multiple dans la phase de test	152
Tableau VI-14 Performance du modèle de la régression Multiple dans la phase de test..	153
Tableau VI-15 Résultats de la productivité par RNA et RLM dans la phase de test	155
Tableau VI-16 Performances des deux modèles dans la phase du test	156
Tableau VII-1 L'importance de chaque facteur sur la productivité.....	161

Liste des symboles et abréviations

RNA : Réseau de Neurones Artificiels

ANN : Artificial Neural Networks

RMC : Réseau Multicouches

PMC : Perceptron Multicouche

RBF : Réseau à Fonction de Base Radiale

QV : Quantification Vectorielle

LVQ : Learning Vector Quantization

ART : Théorie des Résonances Adaptatives

SOM : Self Organizing Map (Réseau de Kohonen)

ML : Machine learning

CLP : construction Labour Productivity

RLS : Régression Linéaire Simple

RLM : Régression Linéaire Multiple

Wij : les poids dans un modèle de réseau de neurone artificiel

SPSS: Statistical Package for the Social Sciences

β : L'estimateur

R : Coefficient de Corrélation

R² : Coefficient de Détermination

F : Test de Fisher

MSE : Mean Square Error

MAPE : Mean Absolute Percentage Error

STD : standard Deviation

AA% : Average Accuracy Percentage

Introduction

générale

Introduction générale

L'industrie de la construction joue un rôle essentiel dans le développement des pays en développement, car elle est considérée comme un facteur majeur de la croissance économique globale et contribue à la prospérité économique et sociale.

Dans ce secteur, la productivité est l'un des principaux aspects utilisés en tant qu'indice de mesure de l'efficacité de la production (Muqem et al., 2011), on peut dire que la mesure de la productivité dans la construction reste flexible en raison des efforts et des comportements humains (Jarkas, 2010; Yi & Chan, 2014).

Dans les domaines de la construction, la productivité représente la relation entre la production et les entrées correspondantes (Gerek et al., 2014). Cette relation peut être définie en considérant la quantité de travail installée en tant que sortie (la production), tandis que les entrées peuvent prendre différentes formes, (le coût de l'heure de travail ou le coût de la main-d'œuvre, ou bien le nombre des heures travaillées et le nombre des ouvriers), (H. R. Thomas et al., 1990; H. R. Thomas & Mathews, 1985; Yi & Chan, 2014). Cette définition de facteur unique n'est pas réaliste et ne donne jamais de résultats précis en raison de la méconnaissance de l'influence d'autres facteurs pouvant affecter la productivité du travail, ce qui rend l'estimation de la productivité du travail plus compliquée et approximative.

Nombreuses études ont été menées afin d'analyser cette relation en prévoyant le taux de productivité de la construction en tenant compte de l'influence de différents facteurs par l'utilisation des méthodes numériques basées sur l'Apprentissage automatique (machine Learning) telles que la logique floue, les analyses de régression, le modèle de réponse à l'action et le réseau de neurones artificiels.

Domaine de la recherche

L'idée principale de cette recherche se concentre sur l'amélioration et l'augmentation de la productivité des travailleurs, et par conséquent, des entreprises spécialisées dans la construction de bâtiments pour les projets en Algérie.

Problématique de la recherche

Pendant longtemps et jusqu'à aujourd'hui, l'amélioration et l'augmentation de la productivité demeurent les préoccupations majeures pour les gestionnaires dans le domaine de la construction à travers le monde.

L'industrie algérienne de la construction est actuellement confrontée à un problème majeur pour mesurer avec précision la productivité du travail dans le secteur de la construction. Ce problème est principalement attribué à l'utilisation continue des technologies traditionnelles, qui a considérablement entravé les progrès dans le secteur de la construction, entraînant une consommation excessive de ressources en termes de main d'œuvre, de temps et de coût. Pour améliorer l'efficacité de la construction, la gestion du temps et le contrôle des coûts, il est crucial de résoudre les problèmes de productivité.

Les études antérieures ont examiné la productivité dans un contexte général en utilisant des méthodes classiques d'observation et de calcul, qui n'ont pas pris en considération l'effet de la plupart des facteurs qui influencent cette productivité. Cela a conduit à des estimations moins précises de celle-ci.

Par conséquent, il est évident qu'il existe une lacune significative dans les recherches existantes concernant des études approfondies sur les facteurs influençant la productivité, en particulier dans le contexte des projets de construction en Algérie. De plus, il manque des méthodes intelligentes qui se concentrent spécifiquement sur la mesure et l'estimation précises des taux de productivité du travail dans le cadre de l'industrie locale de la construction.

Étant donné que notre étude se focalise sur une problématique locale présentant des caractéristiques spécifiques par rapport aux autres régions du monde, nous sommes confrontés aux problématiques suivantes :

- Quels sont les facteurs qui influent le plus sur la productivité du travail dans les projets de construction algériens ?
- Comment faire pour ne pas omettre l'influence de ces facteurs ?
- Et quelles sont les méthodes utilisées pour l'estimation de la productivité ?

Cadre de la recherche

Cette recherche a pour objet, l'estimation et la prédiction de la productivité de travail, dans les projets de construction des bâtiments en Algérie, les données sont collectées quotidiennement sur chantier de quatre projets à Khanchela, Skikda et Constantine, pour le but de développer un modèle de réseau de neurones artificiel, et un autre basé sur l'approche paramétrique, aptes à prédire avec précision la productivité de cette tâche sous l'influence de plusieurs facteurs qui sont identifiés par le biais d'un questionnaire adressé aux spécialistes dans le domaine de la construction.

Structure de la thèse

Cette recherche est organisée en sept (07) chapitres qui sont présentés comme suit :

Chapitre 1 : Productivité et domaines d'application

Dans ce chapitre, on va parler de la productivité en terme général, en illustrant les différentes définitions, les types ainsi que les techniques de calcul de cette dernière.

Chapitre 2 : Etat de l'art

Après un aperçu général de la recherche, le domaine d'application ainsi que l'objectif de la thèse, le chapitre 2 présente un résumé des recherches précédentes sur la productivité et l'utilisation des techniques mathématiques, notamment les réseaux de neurones artificiels et la méthode paramétrique dans la construction.

Chapitre 3 : Présentation générale des réseaux de neurones artificiels et de la méthode paramétrique

L'historique des RNA et ces applications ainsi que ces différents types sont décrits dans ce chapitre.

Une autre approche dite méthode paramétrique sera présentée aussi dans le chapitre 3.

Chapitre 4 : Etude de la productivité dans les projets de construction

Dans ce chapitre, une analyse de la productivité observée dans différents chantiers en Algérie a été faite en premier lieu. Ensuite, un questionnaire englobant les différents facteurs

qui peuvent affecter la productivité est réalisé afin de tirer le plus important parmi ces derniers.

Chapitre 5 : Modélisation de la productivité par les réseaux de neurones artificiels

Le développement d'un modèle de réseau de neurones artificiels est présenté dans ce chapitre, en définissant la structure et l'architecture mise en œuvre, ainsi que les différents paramètres (poids, biais) utilisés.

L'évaluation des performances du modèle développé et l'analyse des résultats obtenus sont aussi discutées

Chapitre 6 : Modélisation de la productivité par la méthode paramétrique

Ce chapitre discute la réalisation d'un modèle paramétrique basé sur la régression linéaire multiple, en présentant les différents tests utilisés, et en évaluant les performances du modèle à la fin de ce chapitre.

Chapitre 7 : Analyse de sensibilité

Les différentes méthodes d'analyse de sensibilité et l'influence de chaque facteur sur le résultat final sont discutées dans ce chapitre.

Finalement, on termine cette thèse par une conclusion générale de la recherche et les recommandations pour les prochaines recherches.

Première partie
Aspect théorique

Chapitre I
Productivité et domaines
d'application

I.1. Introduction sur la productivité

La productivité est considérée comme une mesure essentielle qui influe directement sur le bénéfice ou contrairement sur le déficit d'un projet économique. Une haute productivité permet donc aux entreprises d'achever ces projets à temps et de bien payer ces travailleurs, et le contraire est parvenu lorsque des baisses de productivité sont enregistrées. Elle revêt une grande importance dans divers secteurs et a des applications variées.

1. **Optimisation des Processus de Travail** : Mesurer la productivité du travail permet aux entreprises d'identifier les processus inefficaces et de mettre en œuvre des améliorations. Cela peut conduire à une utilisation plus efficace des ressources, à une réduction des coûts et à une augmentation globale de la performance opérationnelle.
2. **Gestion des Ressources Humaines** : La productivité du travail est un indicateur clé pour évaluer la performance individuelle et collective des employés. Elle peut être utilisée pour déterminer les primes, établir des objectifs de rendement et prendre des décisions en matière de recrutement et de formation.
3. **Évaluation de la Performance de l'Entreprise** : Les entreprises utilisent la productivité du travail comme un moyen d'évaluer leur performance globale. Cela peut aider à mesurer l'efficacité par rapport aux objectifs commerciaux, à évaluer la compétitivité et à prendre des décisions stratégiques pour améliorer la rentabilité.
4. **Allocation des Ressources** : La mesure de la productivité du travail aide les entreprises à allouer efficacement leurs ressources, que ce soit en termes de main-d'œuvre, d'équipements ou de temps. Cela peut conduire à une meilleure planification des projets et à une utilisation plus judicieuse des ressources disponibles.
5. **Prise de Décisions Stratégiques** : Les dirigeants d'entreprise utilisent la productivité du travail pour prendre des décisions stratégiques concernant l'expansion, l'investissement dans de nouvelles technologies, la diversification des activités et d'autres initiatives visant à accroître l'efficacité opérationnelle.
6. **Évaluation de l'Économie Nationale** : Au niveau macroéconomique, la productivité du travail est un indicateur essentiel de la croissance économique. Les gouvernements et les institutions économiques l'utilisent pour évaluer la compétitivité d'un pays,

mesurer les progrès économiques et formuler des politiques visant à stimuler la productivité.

7. **Amélioration de la Qualité de Vie** : Une meilleure productivité du travail peut contribuer à une amélioration de la qualité de vie des travailleurs en permettant une réduction des heures de travail, une augmentation des salaires et une meilleure répartition des gains de productivité.

La productivité est une notion ancienne, qui définit en général la relation entre les résultats et les facteurs qui contribuent à la production. Cette notion est ambiguë et elle est toujours confondue avec les notions d'efficacité, de rendement et de rentabilité... etc. Différentes types et formules de la productivité peuvent enlever l'ambiguïté. Pour cela on doit définir la productivité dans un contexte clair, afin de bien comprendre son utilisation et sa formule appropriée.

En général, la revue de la littérature montre que la productivité peut se définir en deux façons: la première est une définition littéraire qui vise à expliquer ce que signifie le terme, tandis que la deuxième est mathématique qui sert à mesurer les taux et à les interpréter. (Kisi, 2015; Tangen, 2005).

I.1.1. Définitions littéraires

Plusieurs définitions littéraires de la productivité apparaissent dans les dictionnaires latins :

Historiquement, la notion de productivité était initialement définie en relation avec la main-d'œuvre, représentant la quantité de production accomplie par chaque unité de travail, d'où le terme "productivité du travail". Dans cette acception, la responsabilité exclusive de la réalisation de la production incombait à la main-d'œuvre.

Cependant, avec l'avènement de l'industrialisation, la dynamique de production a évolué vers une collaboration entre le capital et le travail. Ainsi, l'entreprise capitaliste a vu le jour, où le travail s'associe au capital pour assurer la production, marquant ainsi une montée en importance significative du travail salarié dans ce contexte.

Larousse définit la productivité comme le rapport entre le résultat d'une activité productive (biens et services) et les facteurs de production que l'on a utilisés pour parvenir à cette production.

The Concise of Oxford a dit que la productivité représente la capacité de produire, l'état d'être productif ; et l'efficacité de l'effort productif ; surtout dans l'industrie ; production par unité d'effort. Si on rentre plus dans le détail de la définition, on trouve que la capacité de produire signifie la force de production (être productif), tandis que l'efficacité de l'effort productif est la mesure de l'utilisation des ressources, alors que la production par unité d'effort représente le taux de la production par ses facteurs dans une période de temps. (Kisi, 2015; P. O. Olomolaiye et al., 1998).

The European Association of National Productivity Centres (EANPC, 2005) définit la productivité comme la manière dont les produits et/ou services sont produits de manière efficace et efficiente, où l'efficacité, c'est fournir un produit conforme qui réponde aux exigences du client, et l'efficience, c'est l'utilisation optimale des ressources pour atteindre les objectifs escomptés.

Selon plusieurs chercheurs, la productivité dans l'industrie de la construction peut se définir soit comme la relation entre les sorties par les entrées (Rojas & Aramvareekul, 2003), dans ce cas les sorties représentent la quantité de travail réalisé et les entrées sont les facteurs suivant les quels le travail a été réalisé, d'autre chercheurs ont utilisé une définition contraire de la précédente et considéré que la productivité est le ratio entre les entrées qui représente les ressource associés, sur les sorties (Dozzi & AbouRizk, 1993), dans cette définition, faibles valeurs de la productivité obtenues indiquent une meilleure performance.

Le terme productivité peut se confondre avec le terme profitabilité du fait que ce dernier est aussi signifié la relation entre les entrées et les sorties, cette relation est monétaire, c'est-à-dire que la profitabilité prend en considération l'effet des facteurs qui ont une relation avec les prix, contrairement à la productivité qui s'appuie sur des phénomènes physique.(Kisi, 2015; Pekuri et al., 2011).

La performance est un autre terme ou définition qui peut se confondre généralement avec la productivité, mais la performance a un large concept qui englobe l'aspect économique et opérationnel, et elle vise toujours l'excellence. (Pekuri et al., 2011).

I.1.2. Définitions mathématiques

Comme on l'a vu précédemment dans les définitions littéraires, on peut conclure que la productivité est la relation entre les sorties d'une opération par ses entrées qui sont

représentées soit par un seul ou plusieurs factures. Selon le cas, on distingue deux types de productivité que l'on va les définir comme suit :

I.1.2.1. Productivité totale des facteurs « TFP »

(Hannula, 2002; Rakhra, 1991; Talhouni, 1990) ont défini la productivité totale des facteurs (en Anglais; Total Factor Productivity), comme étant le rapport entre la sortie totale d'une opération (production) sur la somme des facteurs qui lui sont associés en termes de main d'œuvre, de capital, de matériel, d'énergie....etc.

La représentation mathématique du TFP, apparaitre dans (H. R. Thomas et al., 1990), et utilisée aussi par le département de commerce de l'U.S, ainsi que plusieurs organismes gouvernementaux, elle est donnée par l'équation suivante :

$$TFP = \frac{LA\ SORTIE\ TOTALE}{MAIN\ D'OEUVRE+MATERIELS+EQUIPEMENTS+ENERGIE+CAPITAL} \quad (I-1)$$

Cette formule est une formule économique, elle est définie en terme monétaire en Dollar (selon sa forme d'origine apparue dans la revue de la littérature, mais dans notre pays, on utilise le Dinar Algérien)

$$TFP = \frac{Dollar\ de\ la\ Sortie}{Dollar\ des\ Entrées} \quad (I-2)$$

Comme on l'a dit précédemment, la productivité peut s'exprimer soit par le taux des sorties sur les entrées ou le contraire, à cet égard, (Park, 2006) a utilisé l'inverse de l'équation (I-1) et (I-2) en sa définition de la productivité totale des facteurs, comme le montre l'équation (I-3) suivante :

$$TFP = \frac{Dollar\ des\ Entrées}{Dollar\ de\ la\ Sortie} = \frac{Main\ d'oeuvre+Materials+Equipements+Energie+Capital}{sortie\ totale} \quad (I-3)$$

Selon (Mady, 2013; Park, 2006; H. R. Thomas et al., 1990), ces formules sont généralement utilisables dans les analyses économiques, mais dans le domaine de la construction, elles ne sont pas très utiles.

I.1.2.2. Productivité Partielle

Comme son nom indique, la productivité partielle c'est une partie de la productivité totale des facteurs, dont on utilise une partie des facteurs ou un seul facteur, dans ce dernier cas la productivité partielle devient productivité à facteur unique (ex : si on prend la main d'œuvre

comme entrée la productivité devient une productivité de la main d'œuvre, de même si on utilise le facteur matériel comme entrée donc nous obtenons une productivité des matériels, et ainsi de suite).

La relation mathématique de ce type de productivité peut se différer d'un projet à l'autre suivant ses besoins.

Exemple :

$$PP = \frac{\textit{Sortie}}{\textit{Main d'oeuvre+Materials+Equipements}} \quad (\text{I-4})$$

$$PP = \frac{\textit{Sortie}}{\textit{Capital}} \quad (\text{I-5})$$

On peut se sortir que la productivité est la relation entre les entrées d'une activité et ces sorties. Par ailleurs, aucun commun accord ne se produit entre les managers de la construction envers une définition précise de la productivité, notamment dans les éléments d'entrées et les éléments de sorties..(H. R. Thomas et al., 1986).

I.1.2.3. Productivité de travail

Bien que la construction soit une industrie intensive en main-d'œuvre (les ouvriers), les travailleurs constituant la ressource productive prédominante, on peut affirmer que la productivité de la main-d'œuvre devient cruciale en raison de la concentration de travailleurs nécessaire pour accomplir une tâche spécifique, ce qui peut influencer positivement ou négativement l'avancement des projets (Yi & Chan, 2014).

Bien qu'il existe de nombreuses mesures de performance, la productivité de la main-d'œuvre dans la construction est directement liée à la performance en termes de coûts et de calendrier. Si la productivité de la main-d'œuvre est médiocre, à la fois la performance en termes de coûts et de calendrier sera probablement médiocre. Si le projet est accéléré, la performance en termes de calendrier peut s'améliorer, mais la performance en termes de coûts est susceptible d'être médiocre. Si le projet est réalisé "à moindre cout", la performance en termes de coûts peut être satisfaisante, mais au détriment possible d'une performance en termes de calendrier médiocre. Ce n'est que lorsque la productivité de la main-d'œuvre est bonne qu'il y a une forte probabilité que la performance en termes de coûts et de calendrier soit satisfaisante.

A cet effet, la productivité du travail, est toute productivité en relation directe avec la main d'œuvre (ouvriers), ce qui la rend flexible parce qu'elle dépend toujours à l'effort et le comportement de l'être humain. (Jarkas, 2010).

(H. R. Thomas et al., 1990) définit la productivité du travail dans les projets de construction comme suite :

$$\mathbf{Productivité} = \frac{\mathbf{Sortie}}{\mathbf{cout\ de\ main\ d'oeuvre}} \quad \mathbf{(I-6)}$$

Ou bien :

$$\mathbf{Productivité} = \frac{\mathbf{Sortie}}{\mathbf{Nbr\ d'heures\ de\ travail}} \quad \mathbf{(I-7)}$$

L'inverse de ces équations peut être aussi utilisé

$$\mathbf{Productivité} = \frac{\mathbf{Nbr\ d'heures\ de\ travail\ (ou)\ cout\ de\ main\ d'oeuvre}}{\mathbf{Sortie}} \quad \mathbf{(I-8)}$$

Ou l'équation (I-7) et (I-6) représente le taux de la production, alors que l'équation (I-8) représente le temps unitaire (Mady, 2013).

D'après les équations précédentes, on trouve deux aspects impliqués dans la définition de la productivité de la main-d'œuvre dans la construction, le premier est lié à ce qui est produit en une période de temps finie (production ou quantité de produits réalisée) et le deuxième, ce sont les heures de travail (ou bien le cout) nécessaires sur la même période nécessaire pour produire la production.

I.2. Mesure de la productivité

Bien que de nombreuses publications traitent de la productivité dans le secteur de la construction, il n'existe pas de consensus sur la définition des activités de travail ni sur un système standard de mesure de la productivité. Les chercheurs ont conclu qu'obtenir une méthode standard pour mesurer la productivité de la main-d'œuvre dans la construction est difficile en raison de la complexité des projets et de leurs caractéristiques uniques. La singularité et la non-répétitivité des opérations dans les projets de construction rendent complexe l'établissement d'une définition standard de la productivité et de son mode de mesure.

La manière dont la productivité devrait être évaluée est profondément influencée par l'objectif pour lequel les résultats seront utilisés. La difficulté de la mesure de la productivité réside dans l'établissement de normes fiables pour définir des standards, ainsi que dans le choix de la méthode de mesure et dans la prise en compte de tous les facteurs susceptibles d'influencer la productivité.

L'un des défis rencontrés par les chercheurs réside dans le fait que la mesure de la productivité au niveau du projet est influencée par la diversité et la complexité des activités à réaliser. Ainsi, il existe des disparités dans les niveaux de taux de production en fonction des types de travaux. Pour illustrer, prenons le cas du coffrage des poteaux, où le taux de productivité est moindre par rapport au coffrage des murs en raison des particularités inhérentes à chaque type de tâche. De manière similaire, cette observation s'applique à l'activité de ferrailage et de bétonnage.

Les figures suivantes montrent la diversité de type des travaux qu'on peut les trouver dans les projets de construction.



Figure I-1 Ferrailage dalle plane



Figure I-2 Ferrailage radier général



Figure I-3 ferrailage linteau voile



Figure I-4 Coffrage poteaux.



Figure I-5 Coffrage voile

I.3. Technique de mesure de la productivité

I.3.1. Observation directe

Cette technique est basée sur les observations continues des travaux effectués par les travailleurs durant la journée, ce qui donne la possibilité d'identifier les temps productifs et non-productifs.

Il est préférable que l'observateur, se place dans un endroit où il peut voir clairement les travaux à faire et de rester loin des travailleurs pour éviter toute sorte de dérangement. (Mady, 2013; Noor, 1998).

Avantages de cette technique :

Précision des données collectées.

La durée de chaque tâche est connue

Inconvénients de cette technique :

L'observateur ne peut pas suivre une équipe ayant plus de cinq ouvriers

Si l'opération d'observation est mal gérée par l'observateur, elle peut être négativement influencée le bon déroulement des travaux

I.3.2. Etude de temps

Apparue dans les fins du 19^{ème} siècle et les débuts du 20^{ème} siècle, par Frederick W. Taylor et Frank Gilbreth, cette technique est utilisée pour améliorer la productivité, en mesurant le temps de production, elle vise le développement du temps standard et/ou la production indépendamment aux taux observés (Mady, 2013; Salem, 2006).

Cette technique consiste à déterminer la durée d'exécution de travail des différentes tâches, puis à évaluer les travailleurs par rapport aux rendements théoriques donnés par les normes, et finalement à améliorer la productivité en mettant des nouvelles normes appropriées tout en respectant les temps de relaxation et les imprévus qui peuvent être survenus.

Avantages de cette technique

Ils sont très rentables et faciles à utiliser.

Il nécessite un chronomètre et un minuteur d'intervalle capable d'enregistrer une séquence spécifiée d'événements.

Inconvénients de cette technique :

Il peut être utile uniquement si l'activité implique quelques travailleurs ou des machines.

Il est intrinsèquement difficile pour un observateur unique de couvrir avec précision des activités lorsqu'elles impliquent une période substantielle d'observation sur différents cycles

I.3.3. Echantillonnage d'activité

Dans un chantier de construction, les ouvriers passent leur temps non seulement dans le travail direct pour la réalisation des tâches, mais aussi dans d'autres tâches « non directes » telles que le transport des matériaux, l'attente, la levée des réserves.....etc. (Thomas, H. R., 1991).

L'objectif de cette technique, donc est de mesurer le pourcentage de temps que les ouvriers passent dans diverses catégories de tâches.

Vu l'impossibilité d'observer et d'enregistrer tous les détails d'une ou plusieurs tâches dans un projet de construction, cette technique est référée aux toutes mesures par des observations non continues qui représentent par la suite un échantillon de l'activité totale. (Adrian, 1987; C. H. Oglesby, H. W. Parker, 1989; Mady, 2013).

Avantages de cette technique

Une procédure simple.

Aucun équipement spécial n'est nécessaire pour mener l'étude.

Les résultats sont disponibles rapidement.

L'étude est relativement peu coûteuse.

C'est une technique utile pour étudier des activités non répétitives, non cycliques, pour lesquelles des descriptions complètes des méthodes et des fréquences ne sont pas faciles à quantifier.

Inconvénients de cette technique

La technique n'est généralement pas économique pour l'étude d'un seul travailleur ou d'une seule machine.

Elle n'est pas bien adaptée pour l'échantillonnage sur des travaux à cycle court.

Il est difficile, avec cette technique, d'obtenir des données fournissant des indicateurs suffisants sur les différences individuelles.

I.3.4. Questionnaire des chefs d'équipes sur le retard

Cette technique est basée sur la réalisation d'une enquête pour déterminer les causes de retard qui ont été produites durant la réalisation d'une ou plusieurs tâches pendant la journée.

Un questionnaire devra présenter aux chefs d'équipes pour identifier chaque retard par rapport aux temps spécifié (H. R. Thomas et al., 1990), finalement on reliant les retards et ses causes on peut résoudre les problèmes et éliminer les retards (Alfeld, 1988).

Avantages de cette technique

C'est une méthode relativement peu coûteuse pour analyser les sources de retard pendant la construction.

Cette méthode est flexible et facile à mettre en œuvre.

Inconvénients de cette technique :

Il ne mesure que les pertes dues aux retards et aux travaux de reprise et ne facilite pas la mesure d'autres paramètres utiles pour déterminer l'efficacité des activités.

I.3.5. Questionnaire des travailleurs sur le retard

Cette technique est utilisée pour identifier les retards et améliorer la productivité, en réalisant un questionnaire qui contient les éventuelles causes de retard, et le destiner aux travailleurs.

Pour effectuer ce questionnaire, l'administrateur parcourt le lieu de travail, et choisit aléatoirement des travailleurs pour répondre au questionnaire et collecter les données nécessaires.

L'administrateur répète l'opération jusqu'il obtienne une taille d'échantillon adéquate.(Liou, F., and Borcharding, 1986).

I.3.6. Méthode d'enregistrement

C'est une méthode d'enregistrement qui peut être faite par différentes manières, l'enregistrement par chronomètre est généralement utilisé, elle enregistre le temps d'une ou plusieurs tâches effectuées par l'homme ou machine, si la montre de type qui s'arrête automatiquement et redémarre à chaque fois, une erreur appréciable cumulable devra être prendre en considération, cette erreur est éliminée si la montre fonctionne en continu.

Autre méthode s'appuie sur l'utilisation d'enregistrements par photographie accélérer, c'est –à-dire, prendre des séquences photos par unité de temps soit chaque une, deux, trois et même quatre secondes, pendant de longues périodes, avec un intervalle de temps précis et prédéfini, afin de pouvoir calculer le temps écoulé par rapport aux nombres des photos enregistrées. (Parker, H. W., and Oglesby, 1972) ont trouvé que la caméra time-lapse comme un excellent moyen de collecter les informations et les données, qui sert par la suite d'améliorer la productivité du travail.

La troisième méthode de collecte des données, qui peut être utilisée est l'enregistrement des séquences vidéo par caméscope qui s'avère plus précis dans la détection des problèmes et de les résoudre par conséquence

Elle peut remplacer plusieurs observateurs car elle peut capturer et enregistrer plusieurs tâches simultanées, et elle comprend les étapes suivantes (Eldin, N. and Egger, 1990; Mady, 2013).

- Collecter le matériel vidéo pour l'enregistrement
- Planifier l'opération d'enregistrement
- Conduire des séances de brainstorming pour détecter les problèmes
- Analyser des données enregistrées
- Trouver les solutions et les recommandations

Chapitre II

Etat de l'art

II.1. Introduction

Dans le but de cadrer notre recherche et englober toutes les informations en relation avec le thème traité, la revue de la littérature est une étape très importante pour achever le but suscit , elle consiste   citer les diff rentes  tudes pr c dentes en pr sentant les m thodes utilis es et les r sultats obtenus.

Dans notre recherche, la revue de la litt rature est partag e en trois parties : dans la premi re, on va  num rer les travaux en relation avec l'identification des facteurs influen ant la productivit , la deuxi me partie consiste   pr senter les  tudes de pr diction de la productivit  par la m thode de R seau de Neurones artificiels, tandis que la troisi me et la derni re partie, les recherches de pr diction et estimation de la productivit  par l'approche param trique sont revues.

II.2. Revue bibliographique sur les facteurs affectant la productivit  dans la construction

L'identification des facteurs influen ant la productivit  est une question cruciale pour les managers des projets de la construction et les chercheurs dans le domaine, ces facteurs se diff rent d'un pays   l'autre et d'un site   un autre site (P. Olomolaiye et al., 1998).

Depuis longtemps et jusqu'aujourd'hui, de nombreuses  tudes ont  t  men es pour le but d'identifier et classer l'impact de diff rents facteurs sur la productivit , les techniques d' chantillonnage, questionnaires et de diverses analyses statistiques sont utilis es par les chercheurs pour atteindre de but voulu.

Dans ce paragraphe, nous allons  num rer les divers ouvrages qui ont trait  ce sujet.

(Enshassi et al., 2007), ont conduit un sondage destin  aux projets de construction   GAZA. Palestine, pour classifier l'effet de 45 facteurs sur la productivit  de travail dans le pays. Basant sur les recherches pr c dentes dans le domaine de la construction, les auteurs ont choisi 45 facteurs divis s en 10 groupes qui sont : main d' uvre, leadership « responsable », motivation, temps, mat riel/mat riaux, supervision, projets, s curit , qualit , et facteurs externes, et envoy  par la suite aux 83 entreprises qui font partie de « the contractors union in building specialisation within the Gaza Strip », un feedback de 91% qui repr sente 76 r ponses est obtenu.

Les répondants notaient les facteurs qui leur sont apparus influencer la productivité suivant leur importance « gradué de 1 à 5 », et les chercheurs évaluaient les résultats du sondage en définissant l'importance de chaque facteur par le calcul de L'importance relative RII de chacun.

La pénurie de matériaux de construction, et le manque d'expérience des ouvriers sont classés premier et deuxième facteur qui affectent le plus la productivité de travail dans les projets de construction à GAZA « Palestine ».

(Fagbenle Olabosipo et al., 2011), ont été visités 40 chantiers de construction pour prendre les avis des travailleurs et les managers (chefs projets) des projets, sur l'identification des facteurs qui affectent la performance de la productivité de travail dans les chantiers au Nigeria.

Un questionnaire comprenant 12 facteurs, conçu de façon à choisir l'importance de chaque facteur dans une échelle ascendante de 1 à 5 (échelle de Likert), était envoyé aux 40 chefs projets, et 100 travailleurs, et seulement 28 et 52 questionnaires sont retournés respectivement, ces réponses ont été évaluées en calculant l'Index de l'Importance Relative RII de chaque facteur.

Les travailleurs ont été accordés que les faibles salaires, les facteurs négatifs (retard d'approvisionnement de matériel/matériau, conditions de vie) et le manque de motivation sont les facteurs qui affectent le plus sur la productivité selon les travailleurs, par contre les chefs de projet ont été convenus que les facteurs qui affectent le plus la productivité sont : manque de formation des travailleurs, manque de communication entre les responsables et les travailleurs (concernant le planning, les plans...) et aussi les conditions climatiques durs.

(Jarkas & Bitar, 2012), ont conduit une étude afin d'identifier les facteurs influençant sur la productivité de travail dans les projets de construction en Kuwait, 180 entreprises ont été participées au sondage, qui a englobé 45 facteurs divisés en quatre groupes notamment : « management, technologie, ressources humaines et facteurs externes », et un nombre de 157 réponses ont été obtenues.

Le questionnaire était basé sur l'échelle de Likert pour le choix des réponses suivant leur importance, de 1 qui représente « no significatif » jusqu'à 5 « très significatif », et le RII de chaque facteur est calculé par la suite.

Les résultats ont montré que le groupe de technologie est classé le premier groupe ayant affecté le plus la productivité et le coût total du projet peut être réduit en prenant en compte la phase de design.

En 2012 (Ghoddousi & Hosseini, 2012), ont choisi 31 facteurs divisés en 07 groupes pour identifier l'influence de ces facteurs à la productivité du travail dans les entreprises de construction à l'Iran.

À cet égard, un questionnaire a été envoyé aux 112 entreprises dans cette étude, avec un retour de 93 réponses, 11 parmi elles sont mal renseignées, alors que 82 réponses ont été complètes et bien répondues. Les répondants ont évalué chaque facteur selon son importance, qui l'ont jugé suivant une échelle graduée de 1 à 5 (échelle de Likert) par rapport au temps non productif comme suit :

Le temps non productif représente moins de 20% du temps logique essentiel (effet très faible).

Le temps non productif représente moins de 20-40% du temps logique essentiel (effet faible).

Le temps non productif représente moins de 40-60% du temps logique essentiel (effet modéré).

Le temps non productif représente moins de 60-80% du temps logique essentiel (effet fort).

Le temps non productif représente moins de 80-100% du temps logique essentiel (effet très fort).

La fiabilité du questionnaire a été testée, et un Alpha Cronbach de 0.8719 a été obtenu ce qui montre que le questionnaire est bien fiable et que ses résultats peuvent prendre en considération de ces études, et afin de classer chaque facteur par ordre croissant (qui influence le plus la productivité), les auteurs ont préféré d'utiliser le coefficient de variation « CV » qui représente la variation de l'écart type par rapport à la moyenne.

Les résultats obtenus, ont été montrés que, l'utilisation des méthodes traditionnelles dans la construction, le manque d'expérience des chefs de projet, et manque de matériels adéquats dans le chantier sont les tops trois facteurs jugés affecter la productivité du travail dans les projets de constructions en Iran.

(Karim et al., 2013), ont été distribués un nombre total de 100 questionnaires aux entreprises de construction à la Malaisie, afin de définir les différents facteurs affectant la productivité. Seulement 37% des questionnaires ont été répondus, la statistique descriptive a été utilisée pour décrire et avoir du sens des données, et une classification des facteurs influent sur la productivité a été obtenue par le calcul de l'index moyen qui est utilisé spécifiquement pour la mesure de l'échelle ordinale « de 1 à 5 ».

L'étude a sorti que les techniques de gestion du projet étaient le facteur le plus important qui influence la productivité du travail, en d'autres parties, l'exploration et l'implantation de la technologie « nouvelle technologie » aura aidé les entreprises à augmenter la productivité des travailleurs.

(El-Gohary & Aziz, 2013), conduit un questionnaire portant sur 30 facteurs déviés en trois groupes « Humain, Industriel et Management » afin de ressortir l'effet de ces facteurs sur la productivité de travail dans les projets de construction à l'Égypte.

Le questionnaire est conçu d'une façon que les répondants cochaient la case par rapport à l'importance du facteur, basé sur une échelle de Likert de « 0 » qui représente : aucun effet jusqu'à « 5 » qui représente un très fort effet, et le classement des facteurs par la suite est obtenu suivant l'importance de chaque, qui résulte du calcul de Relatif Importance Index RII.

Le groupe management est classé premier dans cette étude. Ainsi, l'expérience et le savoir-faire des travailleurs, sont jugés comme les facteurs qui ont plus d'effet sur la productivité dans les entreprises Égyptiennes.

41 entreprises de construction dans l'Arabie saoudite ont été répondues à un sondage comprenant 32 facteurs, réalisé par (Mahamid et al., 2013), pour prendre son avis concernant l'influence des facteurs sur la productivité des projets de constructions.

Le calcul de l'Index de l'Importance Relative « RII » de chaque facteur, a montré que les trois premiers facteurs qui influent le plus sur la productivité sont : manque d'expérience des travailleurs, manque de communication et coordination entre les parties de chantier, et la mauvaise relation entre les responsables (chefs projets) et les travailleurs.

Un questionnaire a été mené à Trinidad et Tobago par (Hickson & Ellis, 2014) afin d'identifier les facteurs influençant les projets de construction. L'étude a porté sur 30 entrepreneurs membres de l'Association des Entrepreneurs. Le questionnaire comprenait 42

facteurs répartis en quatre groupes : Gestion, Technologiques, Humains/Main-d'œuvre et Externes. Des questions à choix multiples et ouvertes ont été utilisées. Les participants devaient attribuer un score à chaque facteur sur une échelle de Likert allant de 1 à 4, indiquant le niveau d'effet. Pour établir le classement de chaque facteur, les chercheurs ont calculé l'Indice d'Importance Relative (IIR%) et ont classé les facteurs en conséquence. Les trois principaux facteurs d'influence identifiés étaient : le manque de supervision de la main-d'œuvre, la planification irréaliste et les attentes concernant les performances de la main-d'œuvre, le manque de main-d'œuvre expérimentée.

Afin de comprendre l'influence des différents facteurs sur la productivité du travail dans les projets de construction à l'Egypte. Quatre groupes comprenant 27 facteurs ont fait l'objet d'un sondage réalisé par (M. Hafez, 2014), et un nombre de 55 réponses ont été reçues.

Une analyse des réponses basée sur l'identification de l'importance relative via le calcul de RII, les résultats ont montré que les retards de paiement, la compétence et l'expérience des ouvriers sont les tops trois facteurs qui affectent la productivité de travail à Egypte.

(Bekr, 2016), a été conduit aussi un sondage comportant 37 facteurs, afin d'évaluer l'influence de ces derniers sur la productivité de travail dans les projets de construction réalisés au Jordon. 150 questionnaires ont été collectés depuis les différents intervenants principaux, à savoir : maîtres d'ouvrage, les consultants « B.E.T » et les entreprises de réalisation (50 réponses par chaque catégorie), et chaque facteur est classé suivant le résultat de l'Importance Relative RII.

Comme résultat, l'auteur a signalé que, la mauvaise planification, le manque de matériaux de construction et le manque de matériel et d'équipements de construction, sont les trois principaux facteurs jugés affectant la productivité du travail dans les projets de construction au Jordan suivant la population questionnée.

Une nouvelle méthode basée sur l'évaluation de la probabilité d'occurrence, l'impact et la criticité a été utilisée par (Hwang et al., 2017), pour l'objectif de connaître l'influence de 29 facteurs sur la productivité de travail de green construction, « les projets de l'écoconstruction » réalisés à Singapour.

Pour ce faire, un questionnaire ayant 26 facteurs a été envoyé au premier lieu aux trois entreprises spécialisées dans l'écoconstruction à Singapour, pour assurer que le questionnaire

est compréhensible et enlever des éventuelles erreurs, puis il a été envoyé aux 130 spécialistes dans le domaine, et seulement 32 réponses ont été obtenues. Les répondants ont donc été invités à évaluer la probabilité et l'impact de chaque facteur, s'appuyant sur l'échelle graduée de Likert pour chacune (probabilité et impact).

Pour chaque facteur, la moyenne des probabilités et la moyenne des impacts sont donc calculées et le classement de chaque facteur est obtenu par la suite, on calcule sa criticité qui est une fonction de la probabilité fois l'impact ($C = P \cdot I$).

Les résultats de cette étude indiquent que l'expérience des travailleurs, la technologie et les changements des plans sont les trois premiers facteurs choisis par les répondants comme des facteurs influençant la productivité de travail dans les projets de l'écoconstruction à Singapour.

(Shoar & Banaitis, 2018) ont proposé une approche basée sur l'arbre des causes floues pour identifier les facteurs influençant la productivité du travail en Lituanie. Initialement, ils ont recueilli des données auprès de 15 experts qui ont évalué 27 facteurs à l'aide d'une échelle de Likert, allant de 1 à 5, pour indiquer le degré de gravité. Après cette étape, 18 facteurs présentant un indice d'importance relative (RII) supérieur à 0,7 ont été sélectionnés pour l'analyse ultérieure. Dans la phase suivante, les auteurs ont construit une structure d'arbre des causes au moyen de sessions multiples et d'entretiens avec des experts pour définir les interrelations entre les facteurs sélectionnés. Enfin, l'approche floue a été utilisée pour évaluer la contribution de chaque facteur. Les résultats de leur analyse ont indiqué que deux facteurs, à savoir "l'inflation du coût de l'exécution" et "le financement inapproprié du projet", ont été identifiés comme les facteurs les plus influents ayant un impact sur la productivité du travail en Lituanie.

Une autre étude menée par (Alaghbari et al., 2019) au YAMAN via un sondage constitué de 52 facteurs groupés en quatre groupes (08 facteurs dans le groupe humain, 17 facteurs dans le groupe management, 12 facteurs dans le groupe technique et 15 facteurs dans le groupe externe), destiné aux entreprises de construction, consiste à identifier leur influence sur la productivité de travail dans les projets de construction au YAMAN.

Un total de 150 questionnaires ont été distribués, 100 parmi eux sont envoyés aux entreprises consultantes, et les autres 50 sont envoyés aux ingénieurs et architectes qui

travaillent dans le domaine académique dans les universités publiques et privées au Yaman, et un feed-back de 103 questionnaires est obtenu.

Les réponses au questionnaire sont obtenues par rapport à l'échelle de 5 de Likert, et afin de tester la fiabilité du sondage vis-à-vis des réponses collectées, les auteurs ont calculé le coefficient d'Alpha Cronbach ce dernier donne un résultat de 0.911 qui est plus de 0.70 est enregistré, cela montre que le questionnaire est fiable et on peut compter sur ses résultats. Finalement, la recherche a ressorti que l'expérience et le savoir-faire des ouvriers sont classés comme les plus importants facteurs qui influent sur la productivité de travail dans les projets de construction au Yaman .

Une autre recherche effectuée par (Adebowale & Smallwood, 2020) visait à identifier les facteurs influençant la productivité du travail dans l'industrie de la construction en Afrique du Sud, en se basant sur les perceptions des entrepreneurs. Une enquête par questionnaire comprenant 41 facteurs a été distribuée de manière aléatoire à 96 entrepreneurs. Les réponses obtenues ont ensuite été analysées à l'aide de statistiques descriptives et inférentielles. Les facteurs ont été classés en fonction de leurs valeurs de P, et les facteurs ayant le plus d'influence significative sur la productivité du travail ont été identifiés : la bureaucratie excessive, la livraison tardive des matériaux, les mouvements industriels résultant d'activités politiques.

(Nasirzadeh et al., 2022) ont identifié les facteurs influençant la productivité du travail dans les projets de construction australiens en suivant une démarche en trois étapes. Dans la première étape, ils ont réalisé un questionnaire pour recueillir des données sur différents facteurs liés à la productivité du travail dans les projets de construction. La deuxième étape consistait à établir une boucle de rétroaction cause-effet pour identifier les liens complexes entre les 38 facteurs. Enfin, les auteurs ont utilisé la méthode du Laboratoire d'Évaluation et d'Essai de Décision (DEMATEL) pour hiérarchiser les facteurs. Les résultats de l'étude ont révélé que la taille du projet, le niveau de compétence et d'expérience, ainsi que les problèmes de communication avec les travailleurs étrangers étaient les facteurs les plus importants influençant la productivité du travail dans les projets de construction de bâtiments à plusieurs étages en Australie.

Tableau II-1 Résumé des résultats des travaux précédents

Travaux	Les Cinq Premiers facteurs				
Olomolaiye et al. 1987	Pénurie des matériaux	Retravailler	Manque matériels	Retard dans control et supervision	Absentéisme
Makulsawatudom et al. 2004	Pénurie des matériaux	Plans incomplets	Ouvriers incompetents	Manque matériels	Absentéisme
A. Enshassi et al. 2007	Pénurie des matériaux	Manque expérience des ouvriers	Manque de surveillance	Mal compréhension entre travailleurs et responsables	Changement des plans
F.Olabosipo et al. 2011 (avis des entrepreneurs)	Manque de formation	Manque de communication	Conditions climatiques	Faibles Salaires	Manque de motivation
F.Olabosipo et al. 2011 (avis des ouvriers)	Faibles Salaires	Facteurs négatifs (condition de vie, retards d'appro...)	Manque de motivation	Ouvriers incompetents	Retard d'information
Jarkas and Bitar 2012	Clarté des spécifications techniques	Changement d'ordre	Niveau de coordination entre les différents plans	Manque de contrôle des ouvriers	Proportion des Travaux sous-traités
P.Ghoddousi et al. 2012	Inutilisation de technologies modernes	Chef de projet non expérimenté	Manque matériaux /matériels	Manque d'expérience et de compétence des ouvriers	Les chefs de projet incapable de former les travailleurs
Karim, N et al. 2013	Compétence en gestion du projet	Utilisation des nouvelles technologies	Compétence et formation des ouvriers	Organisation des ouvriers	L'originalité du projet (non habituel)

El-Gohary et al. 2013	Expérience et compétence des ouvriers	Motivation	Disponibilité des matériaux	Compétence des chefs de projet	Niveau de contrôle des ouvriers
I. Mahamid al. 2013	Manque expérience	Manque de communication	Mauvaise relation entre ouvriers et responsables	Retard des paiements	Planning mal réalisé
Hafez, et al. 2014	Retard des paiements	Niveau de Compétence des ouvriers	Manque d'expérience des ouvriers	Manque de contrôle des ouvriers	Manque de motivation
Hwang et al. 2016	Manque d'expérience des ouvriers	Inutilisation des nouvelles technologies	Changement des plans	Niveau de Compétence des ouvriers	Manque de planification
Ghanim A. Bekr 2016	Mauvaise planification	Pénurie des matériaux	Manque matériels	Incompétence des ouvriers	mauvaise gestion du chantier
Aswed 2016	Age	Expérience	Nombre d'ouvriers	Conditions climatiques	Salaire des ouvriers
Alaghbari et al. 2017	Niveau de Compétence et d'expérience des ouvriers	Manque des matériaux	Chefs de projet incompétent	Manque de matériels	situation politique et sécuritaire (manifestations /problèmes politiques)

II.3. Revue bibliographique sur l'utilisation des Réseaux de Neurones Artificiels dans la construction

Dans la deuxième moitié des années 80 les réseaux de neurones artificiels ont commencé à apparaître dans le domaine de la construction en général, et utilisés pour résoudre différents types de problèmes (classifications, prédictions, estimations...).

Dans ce qui suit nous présentons une revue de l'utilisation des réseaux de neurones artificiels dans la construction particulièrement

Commençant par (Portas & AbouRizk, 1997), qui ont été utilisés, une fusion de Réseau Neurones Artificiels (RNN) et logique floue, pour estimer la productivité de coulage de béton, l'algorithme de back-propagation feed-forward a été choisi avec une fonction de transfert sigmoïdale, qui sert à transférer l'information issue de 55 facteurs d'entrée en 1ère couche, au 2ème couche (couche cachée) qui constitue 30 nœuds, la troisième couche qui est la couche de sortie est constituée de 13 nœuds flous .

Cette fusion de ces deux méthodes (RNN et logique floue) a donné de bons résultats dans l'estimation de la productivité des travaux de bétonnage dans les chantiers de construction.

(Zayed & Halpin, 2005), dans autre étude ont été réalisées deux modèles de réseaux de neurones artificiels de différents types, le premier était le feed-forward back-propagation (propagation avant de l'information et rétro propagation de l'erreur), et le deuxième s'était fully connecté (totalement connecté), afin d'estimer la productivité de la construction de pieux forés, dans chaque type les auteurs ont choisi une architecture composée de 03 couches avec sept neurones en couche d'entrée et cinq neurones en couche de sortie, tandis que différents neurones de la couche cachée sont testés dans les deux types.

Les résultats obtenus montrent que les deux modèles sont capables d'estimer avec précision, la productivité, le délai ainsi que le cout d'un projet de réalisation des pieux forés.

La productivité des travaux relatifs aux tâches principales de gros œuvre (ferrailage, coffrage et bétonnage), a été prédite avec succès par (Ezeldin & Sharara, 2006), via Réseaux de Neurones Artificiels.

Les chercheurs ont été collectés 92 données issues de six projets différents, qui l'ont été utilisées pour développer trois modèles de réseaux de neurones artificiels, chacun d'eux sert à estimer la productivité de chaque tâche (ferraillage, coffrage et bétonnage), tous les programmes ont été entraînés par 70% des données en première phase, tandis que les 30% restaient, ont été utilisées en deuxième phase pour tester la capacité des modèles à estimer les résultats, en calculant l'erreur quadratique moyenne (MSE) dans chaque phase.

L'architecture des trois modèles était presque identique, avec deux couches cachées, et utilisant le même algorithme qui était feed-forward back-propagation (propagation avant de l'information et rétropropagation de l'erreur), ainsi que les mêmes fonctions de transfert (tangente hyperbolique « tanh », sigmoïde « tansig », et linéaire « pureline ») successivement, la seule différence était dans le nombre des neurones dans les couches cachées et type de sortie (ferraillage ou coffrage ou bétonnage).

Les résultats finals ont été prouvés que les modèles développés sont capables d'estimer la productivité des trois tâches étudiées, et aider par conséquent les entreprises de mieux gérer leurs projets.

Dans l'année 2011 (Umit Dikmen & Sonmez, 2011), ont publié un article qui parle de leur étude qui a été faite dans le but d'estimer la productivité de la tâche de coffrage dans les projets de construction Turcs.

Les auteurs étaient utilisés 10 facteurs jugés influencer la productivité comme des entrées pour modéliser un réseau de neurones artificiels comportant 03 couches, avec 05 neurones en couche cachée, ont trouvé que le modèle développé fonctionne bien et donne des résultats similaires au Ministère turc des travaux publics.

(Al-Zwainy et al., 2012), ont été conduit une recherche servant à prédire la productivité du travail de pose de marbre, sous l'effet de Dix facteurs jugés influencer la productivité dans les projets de construction en Irak.

Un réseau de neurones Artificiel multicouche a été développé avec algorithme de propagation avant / rétropropagation de l'erreur, contenant 10 facteurs en couche d'entrée, une couche cachée contenant un neurone caché et un seul neurone dans la couche de sortie, ce modèle a été entraîné en première phase avec 60% des données, ensuite 25 % des données

ont été utilisées pour tester le modèle et enfin le modèle est validé par les 15% des données restantes.

La capacité du modèle développé est vérifiée en comparant les résultats obtenus dans la phase de validation par les taux réels, à cet effet, la MAPE le coefficient de corrélation R et le coefficient de détermination R² ont été calculés et des résultats satisfaisants ont été observés, ce qui veut dire que le modèle développé dans l'étude peut prédire la productivité du travail de pose de marbre avec succès.

Ils ont également constaté que l'âge des ouvriers, leur nombre et leurs expériences sont les facteurs qui ont le plus d'impact sur la productivité.

Trois ans après AL-Zwainy et al. on est donc à l'année 2015, ou (Gerek et al., 2015). Ont été développés deux types de réseaux de neurones artificiels afin d'évaluer l'influence de 16 facteurs sur la productivité de travail relative aux travaux de briquetage (maçonnerie), dans les projets de construction turcs.

Les modèles développés ayant deux algorithmes différents, le 1^{er} basé sur l'algorithme de propagation avant/ rétropropagation de l'erreur (feed-forward / back propagation) avec deux couches cachées, et le deuxième utilisait l'algorithme à fonctions de base radiales (radial basis), (que nous allons présenter dans les chapitres qui suivent) avec 11 neurones dans sa couche cachée.

75% des données collectées ont été utilisées pour entraîner les deux modèles et les 25% restaient utilisés pour les tester, et la performance de chaque modèle est évaluée en se basant sur les résultats de MSE, MAPE et le coefficient de corrélation R.

Finalement, l'étude a prouvé que les deux modèles développés sont capables d'estimer la productivité de travail relative aux travaux de briquetages (maçonneries) avec une performance élevée. Il a été trouvé aussi que le modèle avec fonctions de base radiales (radial basis) est plus performant que le modèle à algorithme de propagation avant/ rétropropagation de l'erreur (feed-forward / back propagation).

Une recherche qui ressemble à celle présentée par Gerek et al., mais qui diffère dans le nombre et les types des facteurs, où l'effet de 13 facteurs sur la productivité des maçons effectués les travaux de briquetage (maçonneries), dans les projets de construction à l'Irak a

été étudié par (Aswed, 2016), qui ont développé aussi un modèle de réseau de neurones artificiel pour le but de prédire cette productivité.

Une architecture composée de trois couches cachées avec trois neurones dans chaque couche cachée, la couche d'entrée composée de 13 neurones qui représentent 13 facteurs, et un seul neurone dans la couche de sortie qui représente la productivité estimée, la rétropropagation de l'erreur est utilisée comme algorithme avec la tangente hyperbolique « tanh » comme fonction de transfert.

Un résultat de MSE de (0.0074) et (0.0269) dans la phase d'entraînement et de test respectivement, ont été obtenus ce qui prouve que le modèle développé est capable d'estimer la productivité des travaux de briquetage avec une performance élevée.

Dans une autre étude (Golnaraghi et al., 2019), ont été développés 04 différents types de réseaux de neurones artificiels, à savoir : propagation avant/ rétropropagation de l'erreur, modèle avec fonctions de base radiales (radial basis), réseaux généraux de régression (General Regression Neural Network « GRNN ») et le quatrième, c'était un modèle hybride issu de la fusion de RNA et la logique floue (Adaptive Neuro-Fuzzy Inference System ANFIS), tous ces modèles sont développés pour le but de prédire le taux de la productivité de travail de la tâche du coffrage dans les projets de construction.

Chaque modèle était entraîné en utilisant 80% des données collectées depuis deux projets de gratte-ciels à Montréal –Canada-, et 20% des données ont été réservées pour tester les modèles, et les performances de ces derniers ont été évaluées suivant les résultats de l'erreur quadratique moyenne (MSE) et le coefficient de détermination R².

Les résultats de chaque modèle ont été comparés, et les auteurs ont conclu que les 04 modèles sont capables de prédire avec performance élevée le taux de la productivité de travail de la tâche du coffrage, ils ont trouvé aussi que le modèle développé avec propagation avant/ rétropropagation de l'erreur (feed-forward /back propagation) est le plus performant parmi les autres et donne des résultats qui s'approchent des résultats réels obtenus sur les chantiers.

(Mlybari, 2020) est affirmé que la productivité du travail dans les projets de construction en Arabie Saoudite est affectée par plusieurs facteurs et peut être quantifiée à l'aide de modèles mathématiques. À cette fin, l'auteur a développé quatre modèles différents basés sur un réseau de neurones à perceptron multicouche (MLPNN), un réseau de neurones de

régression générale (GRNN), une machine à vecteurs de support (SVM) et des arbres de régressions additives multiples (MART) pour évaluer le taux de productivité des activités de construction en béton. Les résultats des modèles développés ont été comparés. Le modèle GRNN s'est avéré être le meilleur modèle d'estimation pour la productivité de la fixation de l'acier et du coulage du béton, tandis que le modèle MART a donné de meilleurs résultats que les autres pour la productivité de l'assemblage des coffrages.

Afin d'assurer une livraison de projet efficace et efficiente, (Bokor et al., 2021) ont utilisé un réseau de neurones artificiels (ANN) pour évaluer les relations entre différents facteurs et la productivité de la pose de briques pendant la phase de pré-planification. Le modèle développé comprenait 7 facteurs d'entrée, 5 neurones cachés et une sortie. Les résultats ont montré de bonnes performances, indiquant que le modèle ANN peut être utilisé pour améliorer la livraison du projet en fournissant une estimation plus réaliste des taux de productivité. Cela conduit ensuite à une meilleure planification, estimation des coûts et allocation des ressources.

Dans une étude menée par (Juszczyk, 2020) l'auteur a utilisé cinq modèles de réseaux de neurones artificiels pour développer un modèle unique capable d'estimer les travaux de ferrailage. Il a recueilli 145 points de données sur des chantiers de construction en Pologne, qui ont ensuite été divisés en 90 % pour l'entraînement et 10 % pour le test. Le modèle proposé a montré des performances satisfaisantes, indiquant sa robustesse dans l'estimation des travaux de ferrailage.

(Sarmiento et al., 2021) ont développé un modèle de réseau de neurones artificiels (ANN) pour évaluer le taux de productivité des activités d'assemblage de tuyauterie dans les projets brésiliens. Ils ont testé 108 architectures différentes de réseau de neurones à propagation avant (FFNN), chacune avec 14 neurones d'entrée, pour sélectionner le meilleur modèle. Les résultats de la recherche ont indiqué que l'erreur quadratique moyenne (MSE) du modèle choisi atteignait la valeur de $9,67E-04$, démontrant des performances et une précision satisfaisantes.

Un modèle hybride combinant un réseau de neurones artificiels (ANN) et l'algorithme d'optimisation Grasshopper (GOA) a été développé par (Goodarzizad et al., 2021). L'objectif était d'identifier les facteurs ayant le plus grand impact sur la productivité du travail dans les projets de construction iraniens et d'améliorer la précision de la prédiction de la productivité

du travail. Parmi les 19 facteurs initialement considérés, seuls 6 facteurs ont été identifiés comme influents et utilisés pour développer le modèle ANN prédictif.

(Lawaju et al., 2021) ont utilisé des réseaux de neurones artificiels (ANN) pour prédire le taux de productivité pour le travail de maçonnerie en briques. Les auteurs ont mené une enquête par questionnaire qui incluait initialement quarante-quatre facteurs. Cependant, sur la base de l'analyse de l'indice d'importance relative (RII), seuls les 13 facteurs les mieux classés ont été considérés comme entrés dans le modèle. Le modèle développé utilisait un réseau de neurones (backpropagation-feedforward) à trois couches, avec une couche cachée composée de 30 neurones cachés. L'erreur quadratique moyenne (MSE) a été utilisée comme mesure de précision pour le modèle. Les résultats ont indiqué que le modèle développé estimait efficacement la productivité de la maçonnerie en briques en tenant compte de l'impact des 13 facteurs sélectionnés.

II.4. Revue bibliographique sur l'approche paramétrique dans la construction

Apparue dans les débuts de 1960, l'approche paramétrique a été utilisée principalement pour estimer les coûts des projets, en modélisant la relation entre la réponse du problème recherché (ex : coût de projet) et les paramètres (facteurs) qui influent sur cette réponse.

Cette relation est généralement traduite par une fonction de régression qui lie la réponse du problème avec ces paramètres, dans la plupart des études réalisées, l'approche paramétrique se réfère communément à une régression multiple.

Dans ce paragraphe, on va présenter brièvement une revue chronologique de quelques travaux déjà réalisés par l'approche paramétrique dans le domaine de la construction.

D'abord on commence par (El-Kholy, 2010), qui a été conduit une recherche sert à estimer le pourcentage de perdre ou d'augmentation de la productivité de travail dans les projets de construction en Égypte. Pour ce faire, 14 facteurs (paramètres), jugés affecter la productivité selon le résultat d'un questionnaire englobant 37 facteurs, sont utilisés par la suite pour développer un modèle paramétrique reposant sur la régression multiple, ce dernier a été construit avec les données de 15 projets, tandis que 10 projets ont été utilisés pour le tester (valider).

Les résultats d'erreur et le coefficient de corrélation R, ont montré que le modèle développé peut estimer avec succès le pourcentage de perdre ou augmentation de la productivité.

Dans une autre étude, (Al-Zwainy et al., 2013), ont été utilisés l'approche paramétrique pour estimer le taux de la productivité dans les travaux de pose des carreaux de marbre dans les projets d'Irak. Un modèle de régression multiple a été développé avec 10 facteurs (paramètres) jugés influencer cette productivité.

Un nombre total de 150 données ont été collectées, 100 parmi les, sont utilisées pour développer l'équation de régression, alors que 02 méthodes sont utilisées pour valider l'équation obtenue, la première c'était des tests statiques appliqués au coefficient de corrélation R, tandis que la deuxième c'était de tester l'équation par l'introduction de nouvelles 10 données (pour chaque facteur), la performance du modèle développé est évaluée en suite par le calcul des coefficients de corrélation R et de détermination R², ainsi que l'erreur moyenne absolue en pourcentage (MAPE), et les résultats de cette étude sont comparés par les résultats du modèle de Réseaux de neurones artificiels développé par (Al-Zwainy et al., 2012).

En fin, l'étude a ressorti qu'une prédiction proche de réalité a été enregistrée par le modèle de régression multiple et que ce dernier donne des résultats plus performants que le modèle en RNN.

Passant à l'année 2018 ou (Narayanan & Aswathi, 2018), qui ont été examinés l'influence des facteurs de constructibilité (la mesure dans laquelle la conception d'un bâtiment facilite la construction, sous réserve des exigences globales du bâtiment achevé), sur la productivité de travail, par le développement d'un modèle de régression multiple, sert à évaluer la relation entre les caractéristiques de design et la productivité de travail.

Comme résultat, les auteurs ont trouvé que le modèle développé aidait les concepteurs à bien adapter leurs plans (design) pour achever un taux productif élevé.

(Khanh et al., 2021) ont mené une étude pour examiner l'impact de l'expérience et du niveau de compétence des travailleurs sur la productivité de la construction en maçonnerie au Vietnam. Ils ont développé un modèle de régression logistique pour estimer la productivité et ont constaté qu'il n'y avait aucune différence de productivité moyenne entre les différents

groupes d'expérience des travailleurs. Cependant, ils ont observé une différence significative de productivité entre les groupes de compétences des travailleurs

Dans une autre étude (Paikun et al., 2021) ont utilisé une méthode de régression linéaire multiple pour analyser l'impact de trois facteurs sur la productivité de la main-d'œuvre dans l'installation de murs en briques légères et le crépissage de murs en briques légères en Indonésie. Les chercheurs ont obtenu une précision impressionnante de 99,43 % et 99,04 % pour les deux modèles respectifs, ce qui indique la solidité et la fiabilité des modèles développés.

Dans leur étude menée sur huit chantiers de construction à Niš, en Serbie, (Matejević-Nikolić & Živković, 2022) se sont concentrés sur l'analyse des processus de bétonnage, en particulier ceux liés aux colonnes et aux murs. Ils ont rassemblé 60 points de données pour développer une analyse de régression et des modèles de simulation visant à améliorer la précision de la prévision des taux de productivité pour les travaux en béton. Les résultats de l'étude ont indiqué que ces modèles proposés ont le potentiel d'améliorer la prise de décision et d'accroître la précision de la planification dans les projets de construction impliquant des bâtiments résidentiels à Niš, en Serbie.

II.5. Conclusion

Ce chapitre se concentre sur la revue des recherches et des études précédentes, sur le sujet de la productivité dans les projets de construction, il se constitue de trois parties, dans la première partie, les facteurs influençant la productivité sont identifiés dans différents pays et un résumé des tops 5 facteurs résultant des différentes études est présenté à la fin de cette partie.

La deuxième partie présente les différentes études sur l'estimation et la prédiction de la productivité sous plusieurs facteurs, par la méthode de Réseau de Neurones Artificiel, et différents modèles avec différentes architectures et combinaisons sont trouvés.

La troisième et dernière partie énumère les études concernant l'utilisation de l'approche paramétrique dans l'estimation de la productivité du travail dans les projets de construction.

Cette revue de la littérature nous a donné un aperçu général quant au domaine de la productivité dans les projets de construction, ce qui nous permet de bien encadrer notre sujet de recherche

Chapitre III

Présentation des méthodes de modélisation

III.1.Introduction

Une modélisation mathématique, c'est l'application des méthodes, techniques et expressions mathématiques dans le but d'interpréter une situation (problème) sous forme mathématique.

Deux méthodes de modélisation sont utilisées dans notre recherche afin d'étudier la productivité de travail dans les projets de construction Algériens.

Dans ce chapitre, on va présenter l'aspect théorique de la méthode de réseaux de Neurones Artificiels ainsi que la méthode paramétrique, en expliquant les différents types, les propriétés, ainsi que les avantages et inconvénients de chaque méthode.

III.2.Réseaux de Neurones Artificiels

Les réseaux de neurones artificiels, connus aussi sous les noms de connexionnisme ou neuromimétique, sont des modèles mathématiques qui font partie du domaine de l'intelligence artificiel, son principe est d'imiter le traitement d'information d'un cerveau biologique, ayant des capacités d'apprendre par les expériences et de s'adapter aux différentes situations, ils ont apparu de façon répandue dans la deuxième moitié du 20ème siècle et conçus principalement pour résoudre des différents problèmes (régression, classification, estimation, reconnaissance de formes...etc.).

Composés de différents types, les réseaux de neurones artificiels, touchent divers domaines, à savoir : médecine, construction, mécanique, traitement de signal...etc.

III.2.1.Historique

On commence par l'année de 1890 aux États-Unis, là où les réseaux de neurones artificiels ont fait leurs débuts, par un célèbre psychologue s'appelant W. James qui a introduit le concept de mémoire associative, ce qui lui permettra de proposer ce qui deviendra une loi de fonctionnement pour l'apprentissage des réseaux de neurones. Soixante ans plus tard, exactement en 1949, en deux scientifiques J. Mc Culloch et W. Pitts ont été créés le premier modèle mathématique à comportement binaire du réseau de neurone formel simple, capable de réaliser des formules logiques, arithmétique et complexes, durant la même année un physiologiste américain, connu sous le nom de D.Hebb qui a intéressé aux manières suivant

lesquelles les neurones se connectent, « des neurones qui s'excitent ensemble se lient entre eux » en expliquant sa théorie par l'application d'un conditionnement classique (pavlovien), en nourrissant un chien, à la même heure chaque jour, ce que lui entraîne la sécrétion de salive, à cette heure exacte même en l'absence de nourriture.

« Faisons l'hypothèse qu'une activité persistante et répétée d'une activité avec réverbération (ou trace) tend à induire un changement cellulaire persistant qui augmente sa stabilité. Quand un axone d'une cellule A est assez proche pour exciter une cellule B de manière répétée et persistante, une croissance ou des changements métaboliques prennent place dans l'une ou les deux cellules ce qui entraîne une augmentation de l'efficacité de 'A' comme cellule stimulant 'B'. » Donald Hebb, 1949

En 1957 c'est la première fois qu'un modèle du perceptron est sorti au monde, par F. Rosenblatt qu'il a fait développer le premier neuro-ordinateur et l'appliqué au domaine de la reconnaissance des formes.

Quelques années plus tard, en 1960, l'automaticien B. Widrow, a été développé un modèle qui ressemble en général à celui de Rosenblatt mais qu'il diffère dans la loi d'apprentissage. Il lui donne le nom Adaline (Adaptative Linear Element), Celle-ci est à l'origine de l'algorithme de rétropropagation de gradient.

Ensuite, les limitations du perceptron concernent spécialement l'impossibilité de traiter les problèmes non linéaires ont été présentées par Minsky et S. Papert, dans son ouvrage publié en 1969.

Les années qui suivent jusqu'à 1982 sont considérées comme des années de l'ombre, dans cette période, les recherches se poursuivaient, mais sous la couverture des autres disciplines comme le traitement adaptatif du signal, la reconnaissance de formes, la modélisation en neurobiologie. De grands noms comme : S. Grossberg, J. Hertz, B. Kosko, et autres ont été actifs durant cette vingtaine d'années.

On est au début des années 80 où la période d'ombre est arrivée à sa fin, grâce à une publication du physicien Américain J. J. Hopfield en 1982, ce dernier avait présenté sa théorie du fonctionnement et des possibilités des Réseaux De Neurones Artificiels. Contrairement aux autres auteurs qui proposent des structures et lois d'apprentissage des Réseaux de Neurones Artificiels, et étudient par la suite les propriétés de ces réseaux, J. J.

Hopfield a choisi de fixer au préalable le comportement à atteindre du réseau et réalisera par la suite une structure et la loi d'apprentissage convenable aux résultats escomptés, le modèle Hopfield est très répandu jusqu'à nos jours, et utilisé généralement dans les problèmes d'optimisation.

Le professeur finlandais T. Kohonen a inventé en 1982 le modèle de cartes auto-organisatrices (Self-Organizing Maps (SOM)), utilisé pour les problèmes à solution inconnue préalablement (apprentissage non supervisé), le principe des cartes de Kohonen est de faire une projection non linéaire des données de haute dimension sur un espace de faible dimension, et elles sont utilisées dans les cas de la classification de données. Il est à noter que Hopfield n'a pas donné des solutions aux limitations du perceptron présentées par M. Minsky, c'est pour cela que les recherches sont relancées, spécialement par l'apparition de la machine de Boltzmann en 1983, ce physicien Autrichien a fait développer un modèle capable de traiter les limitations du perceptron, mais le problème de son modèle était dans la convergence de l'algorithme qui a été très longue avec un temps de calcul énorme.

L'année 1985, a connu l'apparition de l'algorithme de rétropropagation de gradient, qui est adapté à l'apprentissage du perceptron multicouche, cet algorithme était une découverte de trois groupes de chercheurs, ce qui nous permettait par la suite de réaliser une relation non linéaire d'entrée/sortie, en la décompose en une suite d'étapes linéairement séparables, quelques années plus tard, Moody et Darken 1989 ont été réalisés l'interpolation multivariable par un modèle à fonction de base radiale « *Radial Basis Function network (RBF)* ».

En 1990, l'apprentissage par renforcement est apparu, utilisé dans l'apprentissage non supervisé, son principe consiste à récompenser les comportements souhaités et/ou à sanctionner les comportements non désirés, et c'est donc le système qui améliore ses capacités à travers les expériences successives.

Passant maintenant à l'année 2013 où les réseaux de neurones profonds (ils vont jusqu'à 10 couches, quand leurs prédécesseurs, les réseaux multicouches devaient s'arrêter à 2 ou 3 maxis) sont reconnus un grand succès en obtenant d'excellentes performances pour la classification d'images et de sons, voire même mieux que celles de l'être humain (cas classification d'images).

III.2.2. Domaine d'applications des Réseaux de Neurones Artificiels.

Dans les premières années de sa découverte, les réseaux de neurones artificiels ont été confinés dans le domaine de la psychologie et de la physiologie, et après une période de recherche et surtout avec le développement d'informatique, « l'apparition des ordinateurs », les RNA se sont alors propagés dans pas mal de disciplines vues ces performances enregistrées, et devenaient la mode scientifique du siècle. De nos jours, les réseaux de neurones artificiels sont très répandus et de nombreuses applications sont trouvées dans des domaines très variés, notant comme exemples ce qui suit :

Traitement d'image : reconnaissance de formes, de photos et de caractères, reconnaissance des signatures, des empreintes faciales et des visages... etc.

Traitement automatique des langues : traduction automatique, segmentation en mots, maisons intelligentes... etc.

Contrôle : Détection et diagnostique des pannes,etc.

Planification : Calendrier intelligent, planification des projets,etc.

Finance : Détection des fraudes, Analyse d'investissements et de fluctuations des taux de change, Couverture assurantielle et estimation des réserves, Ciblage des prospections, mesures et comparaisons des campagnes et des méthodes.

Estimation : Estimation des couts, des délais, de la productivité... etc.

Classification : Classification des espèces animales et végétales, prise de décision...etc.

Environnement : Prévisions de la qualité de l'air et de l'eau, prévisions météorologiques, détection et élimination de la pollution.....etc.

Médecine : Diagnostics médicaux, traitement des dossiers des malades « historique », robots chirurgicales...etc.

Transport : Maintenance des voies, détection des pannes, voitures autonomes...etc.

Agriculture : Irrigation intelligente, augmentation de la production, traitement des arbres...etc.

Construction : Planification des projets, estimation et prédiction des coûts des projets, amélioration de la productivité de travail, identification des caractéristiques du sol ...etc.

On peut dire que les exemples cités ci-dessus ne représentent qu'une petite partie du domaine d'application des Réseaux de neurones artificiels, alors que leurs champs d'application sont trop larges, voire illimités, en reposant sur des bases purement mathématiques qui les permettent de se pénétrer aux diverses disciplines, vu ces performances élevées qui facilitent le travail de chaque spécialiste « ingénieurs, chercheurs, managers...etc. ».

Les domaines de construction, de génie civil et du Management des projets ont aussi une grande part de l'intelligence artificielle et les réseaux de neurones, spécialement, et on va illustrer en détail dans le chapitre prochain leur utilisation dans l'amélioration et l'augmentation de la productivité du travail dans les projets de construction.

III.2.3. Structure des Réseaux de Neurones Biologique

III.2.3.1. Réseaux de Neurones Biologique

Les neurones sont des cellules nerveuses de forme variable, qui relient le corps avec le cerveau, sa mission principale est la transmission et traitement des informations « sous forme de signaux électriques » entre l'environnement extérieur et les organes du corps ou entre les différents organes au sein du même corps.

III.2.3.1.1. Composition

La composition de chaque neurone s'avère comme suite :

Le corps cellulaire : appelé aussi « SOMA », constitué un noyau, et présente de nombreuses branches liées avec les autres cellules et recevant l'information sous forme des signaux du milieu extérieur « inputs », ils sont appelés dendrites.

Axone : représente un prolongement issu du corps cellulaire vers l'extérieur sa longueur peut atteindre 1 mètre et son diamètre est de 1 à 1.5 micromètres, il sert de transmettre de l'information sortant du SOMA vers les autres cellules « output », il peut se diviser à sa fin, faisant un contact avec un grand nombre de cellules.

Synapses : trouvés à l'extrémité des axones et dendrites, sont des éléments essentiels assurant le contact avec les autres neurones.

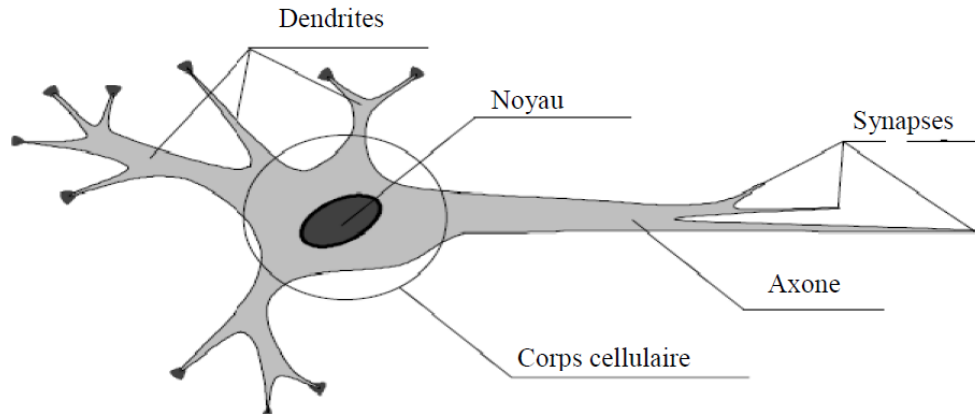


Figure III-1 Structure d'un neurone biologique

III.2.3.2. Réseaux de Neurones Artificiels

Inspirés des réseaux de neurones biologiques, les différents types des réseaux de neurones artificiels ont des structures différentes, qu'on va les présenter en détail dans les sections qui suivent, mais la suivante composition est commune entre eux :

Neurones artificiels : sont des éléments ou unités communiquant l'information avec l'environnement extérieur et intérieur « entre différentes couches ».

Couches : sont des boîtes contenant des neurones, et interconnectées entre eux, on peut distinguer trois principales couches qui sont : couche d'entrée, couche cachée et couche de sortie.

Connexions : Qui relie les neurones des différentes couches et assurent la transformation de l'information par le biais des relations mathématiques.

Le seuil d'activation (θ_i): trouver aussi sous le nom de Biais, est une valeur numérique associée à chaque neurone, qui représente la limite qui doit atteindre ou dépasser un signal à la sortie d'un neurone artificiel pour que ce dernier soit activé (prendre une valeur positive ou nulle, et prendre une valeur négative dans le cas contraire).

Poids : facteurs de pondération, associé à la connexion entre deux neurones, qui sert à mesurer l'influence d'un neurone j sur le neurone i , si le poids est positif alors la liaison entre neurone et excitatrice, par contre si le poids est négatif alors la liaison est inhibitrice.

Fonction d'entrée : un prétraitement appliqué aux entrées, et produit par la suite un réel appelé *Potentiel membranaire du neurone*, elle sert à introduire l'information provenant des entrées dans le réseau.

Cette fonction peut prendre l'une des formes suivantes :

Une combinaison booléenne des entrées.

Une combinaison linéaire des entrées : un produit scalaire entre le vecteur d'entrée et les poids associés

$$A_i(x) = b_i + \sum_j w_{ij} * x_j \quad (\text{III-1})$$

La distance entre le vecteur des entrées et le poids associés, elle est utilisée dans le type de réseau à fonction de base radiale, elle est calculée suivant les deux équations suivantes :

$$A_i(x) = b_i \sqrt{\sum_j (w_{ij} - x_j)^2} \quad (\text{III-2})$$

$$A_i(x) = b_i * \|w_{ij} - x_j\| \quad (\text{III-3})$$

Fonctions de transfert « f(x) » : fonctions appliquées au niveau des connexions, sert à introduire une non-linéarité dans le fonctionnement des neurones, différentes fonctions de transfert peut se trouver et les fonctions suivantes sont les plus utilisées :

La fonction linéaire : Figure III-2 (a)

La fonction binaire à seuil « Heaviside » : Figure III-2 (b)

La fonction linéaire à seuil : Figure III-2 (c)

La fonction sigmoïde : Figure III-2 (d)

La fonction tangente hyperbolique : Figure III-2 (e)

La fonction à base radiale : Figure III-2 (f)

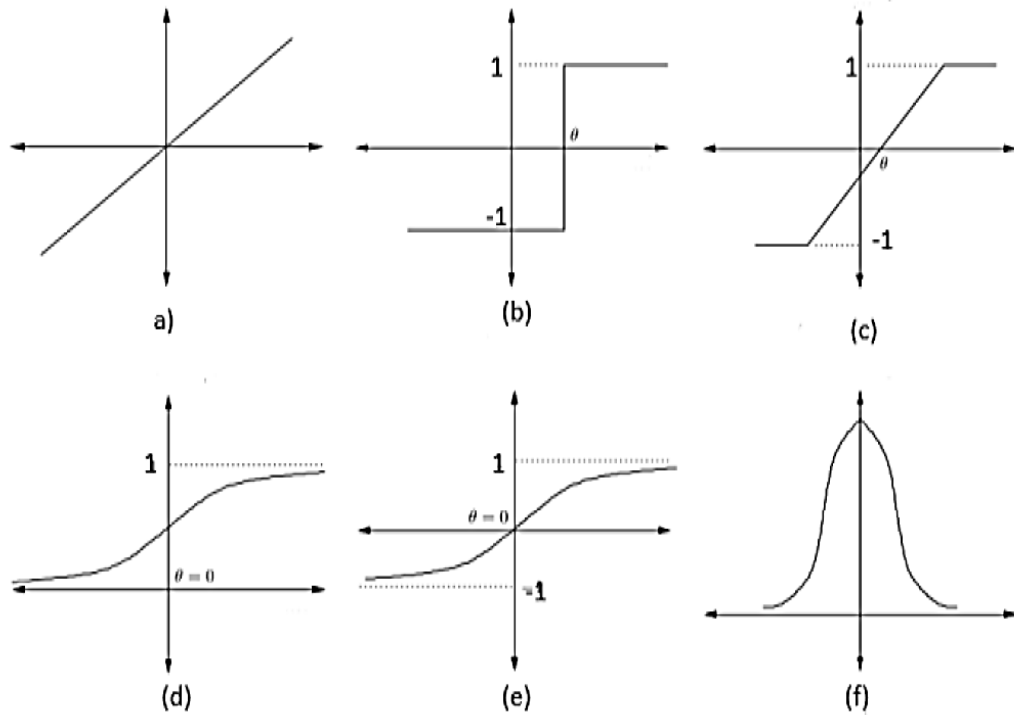


Figure III-2 Différentes types des fonctions de transferts

Fonction de Sortie $f(A(x))$: c'est l'application de la fonction de transfert à la fonction d'entrée.

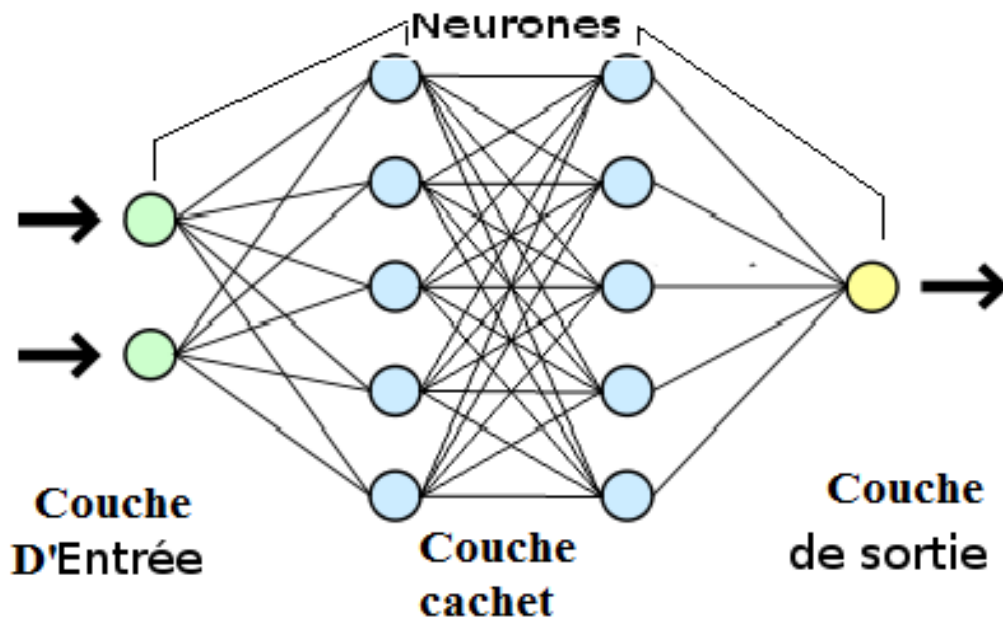


Figure III-3 Structure typique d'un réseau de neurone artificiel

III.2.3.2.1. Neurone formel

Proposé par Mc.culloche et Pitts en 1949, le neurone formel est une représentation mathématique imitant le fonctionnement du neurone biologique.

Ce neurone consiste à une unité de calcul interconnectée avec les entrées et sorties, elle reçoit un vecteur d'entrée, et effectue par la suite la sommation soit du produit scalaire des poids et les entrées soit la distance entre ces derniers (Tsopze, 2010; Viennet, 2006), le résultat de cette sommation est soustraite de la valeur du seuil, puis appliquer une fonction de transfert au résultat de cette opération, le résultat obtenu finalement et la sortie du neurone formel.

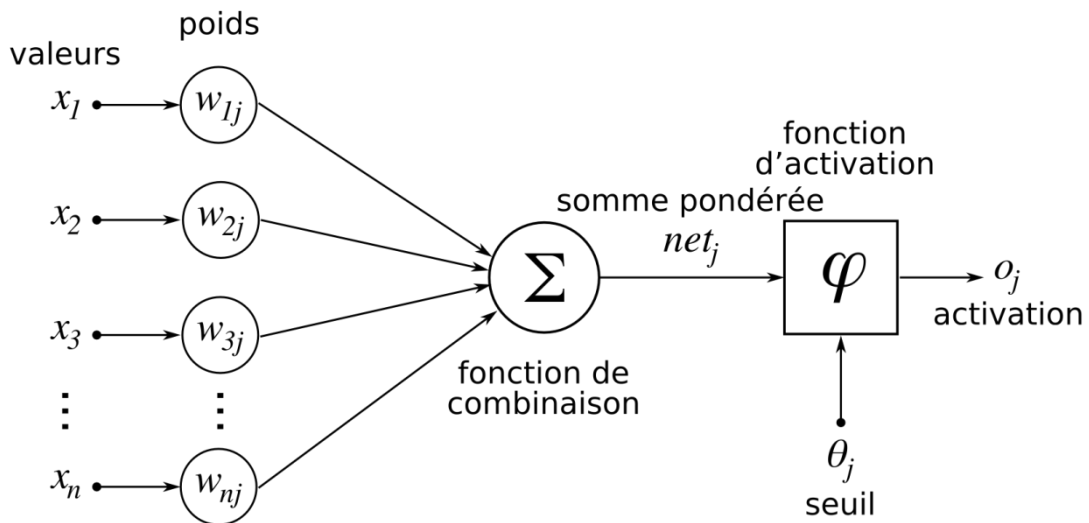


Figure III-4 Structure d'un neurone formel.

III.2.3.3. Comparaison entre réseau de neurones biologique et le réseau de neurones artificiels.

Basant sur les précédentes présentations des deux réseaux (biologiques et artificiels) on peut ressortir les différentes propriétés spéciales à chaque réseau.

Le tableau suivant montre cette comparaison

Tableau III-1 Comparaison entre le Réseau de Neurones Biologique et le Réseau de Neurones Artificiel

	Réseau Artificiel	Réseau Biologique
<i>Structure</i>	<ul style="list-style-type: none"> – Les différentes couches (entrée, cachée et de sortie) contenant les nœuds. – Fonction d’entrée, fonction du transfert et la fonction de sortie – Connexion entre neurones via des poids 	<ul style="list-style-type: none"> – Neurone – Le corps cellulaire (soma) – Axone – Synapses
<i>Propriétés</i>	<ul style="list-style-type: none"> – Déterministe – Séquences D’instructions – Tâches Répétitives – Vitesse Elevée – Programmation – Unicité Des Solutions 	<ul style="list-style-type: none"> – Parallélisme – Adaptation (Apprentissage) – Vitesse Lente – Comportement Flou – Différentes Façons D’aborder Un Problème – Différentes Solutions

III.2.4. Architecture des Réseaux de Neurones Artificiels

L’architecture du réseau de neurones artificiels, c’est l’organisation suivant laquelle un modèle peut se prendre, elle représente le nombre de couches appropriées, la manière dont ces couches sont connectées, ainsi que le type d’algorithme choisi.

On distingue deux grandes catégories d’architecture :

III.2.4.1. Réseaux statiques

Appelé aussi réseaux non bouclés, c’est un réseau avec une ou plusieurs couches, interconnectées entre elles, et le flux de l’information qui représente des fonctions mathématiques algébriques (linéaires et non linéaires) circule dans un seul sens et sans retour à partir des entrées vers la sortie du modèle passant par les couches intermédiaires « cachées » (cas de réseau multicouches).

Dans ce type de réseaux, la notion du temps n'a aucun sens fonctionnel, vu que ses entrées sont constantes, alors que les sorties seront obligatoirement constantes, tandis que le temps de calcul est considéré négligeable, ils sont couramment utilisés dans la reconnaissance des formes, la classification et la prédiction et l'estimation.(Ammar, 2007).

Les réseaux statiques « non bouclés » sont divisés en deux types principaux : **les réseaux complètement connectés et les réseaux à couches** (voir la section III.2.5.1 suivante).

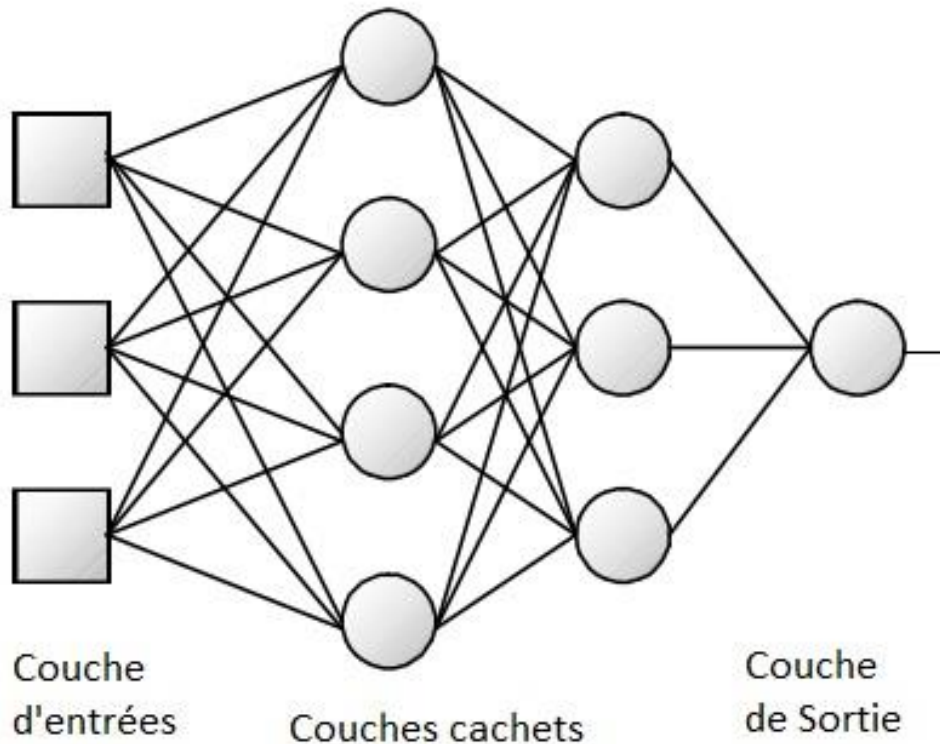


Figure III-5 Réseau de neurones artificiel non bouclé

III.2.4.2. Les réseaux dynamiques

Les réseaux dynamiques sont des réseaux bouclés, c'est-à-dire que l'information circule vers l'avant et aussi faire un cycle de retour vers une position antérieure « de départ », autrement dit l'entrée d'un neurone peut être en relation directe ou indirecte avec sa sortie, et c'est en fonction de là que la notion du temps est explicitement prise en considération.

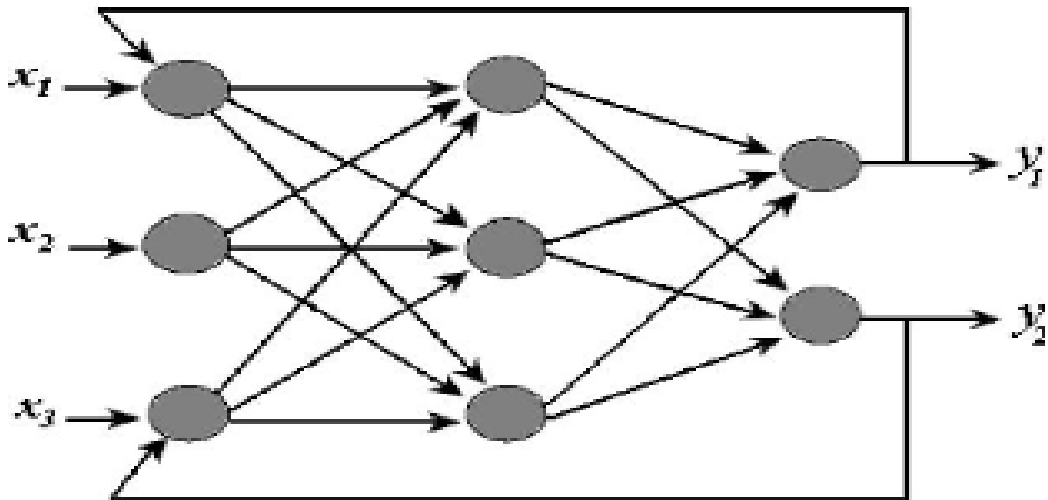


Figure III-6 Réseau de neurones artificiel bouclé

III.2.5. Les types des Réseaux de Neurones Artificiels

Comme on a présenté dans la section précédente, les réseaux de neurones artificiels sont divisés en deux catégories selon leur architecture (réseaux statiques, et réseaux dynamiques), ces derniers sont présentés de différents types, qui sont illustrés dans la figure suivante :

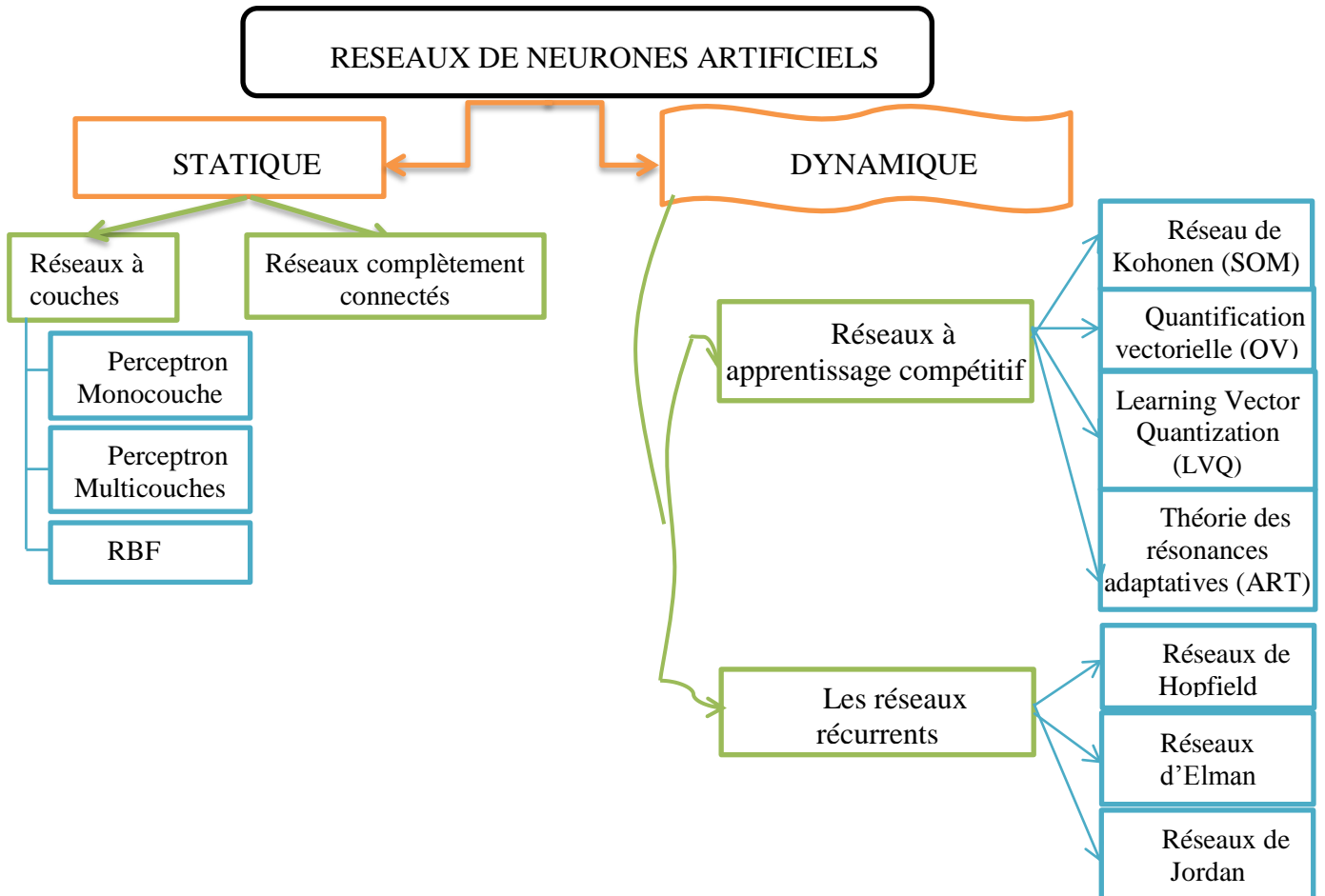


Figure III-7 Cartographie des types des Réseaux de Neurones Artificiels

III.2.5.1. Types des réseaux statiques (non-bouclés)

III.2.5.1.1. Réseaux complètement connectés :

Dans ce type les neurones sont totalement liés de façon que :

Les entrées de chaque neurone sont les entrées du système ainsi que les sorties de ceux qui le précèdent.

Les sorties de chaque neurone sont liées aux neurones qui le suivent.

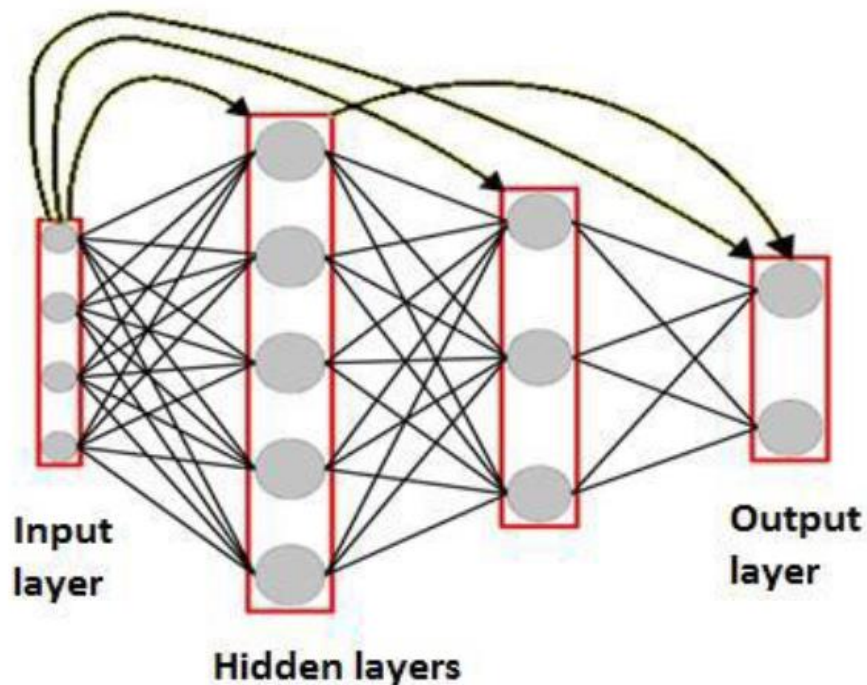


Figure III-8 Structure d'un Réseau de neurone artificiel complètement connectés
(Principe et al, 2010)

III.2.5.1.2. Réseaux à couches

III.2.5.1.2.1. Perceptron Monocouche

C'est le premier modèle de réseau de neurones artificiels, il a été développé par Rosenblatt en 1958, contrairement au neurone formel créé par McCulloch et Pitts ce modèle est capable d'apprendre de l'expérience.

Ce modèle constitué d'une seule couche, il reçoit les informations de l'environnement extérieur sous forme des entrées, et les transmet aux cellules de décisions (neurones de

sorties) via les unités d'association qui font une somme pondérée de ces entrées et appliquent à cette somme pondérée une fonction d'activation (de transfert) de type sigmoïde.

Les sorties de ce type de réseau sont binaires (0,1) ou bipolaires (-1,1).

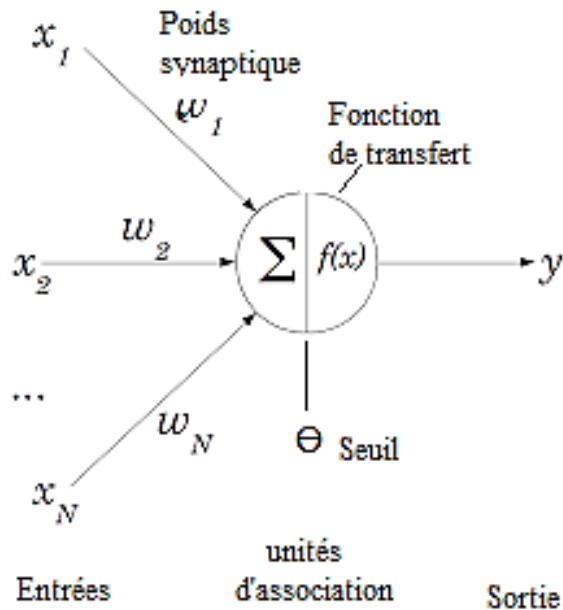


Figure III-9 Perceptron Monocouche

III.2.5.1.2.2. Perceptron Multi Couches (PMC)

Comme son nom l'indique les réseaux de neurones multicouches (Perceptron Multi Couches), constituent des couches successives dont :

- Une couche d'entrée recevant les données de l'extérieur,
- Une couche de sorties qui fournit les sorties du système
- Plusieurs couches intermédiaires appelées Couches Cachées.

On entend par couche, l'ensemble des neurones n'ayant pas de connexion entre eux, et que l'information se propage dans un seul sens entre les différentes couches (depuis la couche d'entrée vers la couche de sortie passant par les couches cachées).

Ce modèle est apparu en 1985, grâce à la découverte de l'algorithme de rétropropagation du gradient d'erreur par trois équipes indépendantes. Cet algorithme a résolu les limitations

du perceptron principalement liées aux modifications des poids pour les connexions qui ne sont pas liées avec les neurones de sortie.

Dans les modèles multicouches, une transformation paramétrique non linéaire des entrées en sorties en fonction des poids « w » et du seuil « θ » (w et θ sont les paramètres du modèle) se produira entre les différentes couches, et une minimisation du gradient d'erreur se fait au cours de l'apprentissage et se propage en arrière pour permettre au modèle d'ajuster les poids afin d'obtenir des meilleures performances (principe de l'algorithme de rétropropagation d'erreur).

Un réseau de neurones artificiel, à une seule couche cachée, avec un nombre fini de neurones et un neurone en sortie linéaire, peut approcher n'importe quelle fonction continue définie sur un domaine fini de l'espace de ses variables, avec une précision arbitraire. (Hornik, K., Stinchcombe, M., White, 1989).

On peut représenter mathématiquement la sortie du neurone i d'un modèle de Réseau de Neurones Multicouches par l'équation suivante :

$$O_i = f(\theta_i + \sum_{j=1}^n w_{ij} * x_j) \quad (\text{III-4})$$

Où :

O_i : La sortie du Neurone i

F : fonction de transfert

W_{ij} : poids associés au neurone i

X_j : les signaux entrants au neurone i provenant des sorties des neurones en couche $i-1$

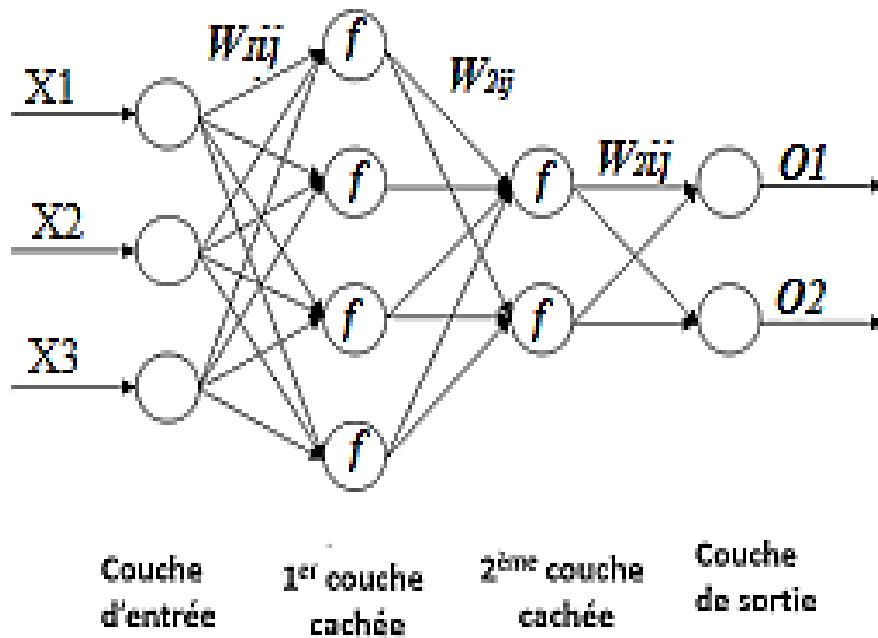


Figure III-10 Réseau de Neurone Multicouches

III.2.5.1.2.3. Les Réseau de neurones à Fonctions de Base Radiales (RBF)

Introduit la première fois en 1971 par Hardy, et en 1977, Powell a développé la théorie correspondante pour résoudre les problèmes d'interpolation multivariée, mais dans le domaine des réseaux de neurones Artificiels, le mérite revient aux Broomhead et Lowe en 1988 qui ont été les premiers à introduire les fonctions de base radiale dans ce domaine, autres contributions à la théorie et à l'application de ce type de réseaux est apparue en suite aux années 1989 et 1990 par J. Moody et C. Darken Renals, Poggio et Girosi. (Sahin, 1997).

Les réseaux à fonction de base radiale (RBF), sont des réseaux constitués de trois couches (entrées, sorties, et une seule couche cachée), la différence de ce réseau par rapport aux autres types, c'est l'activation des neurones d'entrée et de sortie dans sa couche cachée, il applique à une fonction à base radiale pour activer le vecteur d'entrée, il est à noter que divers types de fonctions peuvent être utilisées, mais celle de type gaussienne reste la plus utilisée (formule III -5).

$$a_i = \exp\left(-\frac{1}{2}\sum_{k=1}^n \frac{(x_k - c_{k,i})^2}{\sigma_{k,i}^2}\right) \quad \text{(III-5)}$$

Dans cette formule i désigne l'indice du neurone, k parcourt l'ensemble des entrées notées x_k , et c_{ki} et $\sigma_{k,i}^2$ sont des paramètres appelés respectivement centres et variances des gaussiennes, la représentation graphique de cette fonction est illustrée à la figure III-11.

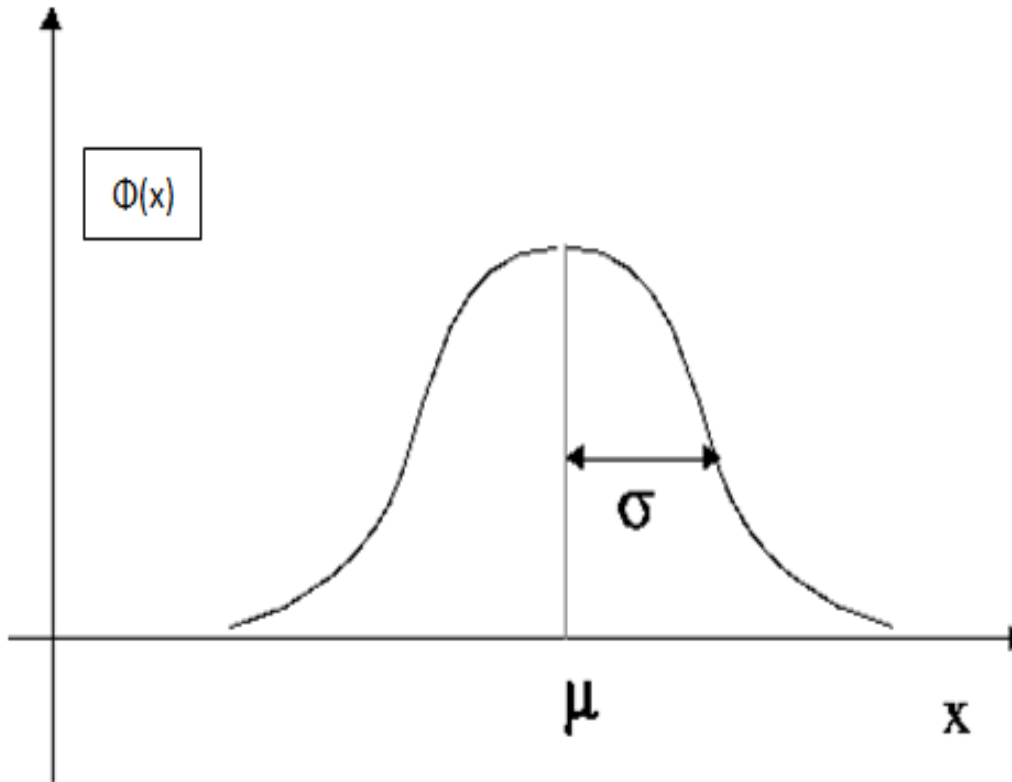


Figure III-11 Représentation graphique de la fonction à base radiale de de type gaussienne

Notant que l'activation des neurones d'entrée n'a aucune relation avec les poids, mais leur activation est relative à la distance du centre de gaussienne.

Dans l'étape qui suit, le réseau active les neurones de sorties en fonction des poids par la réalisation d'une somme pondérée des neurones cachés.

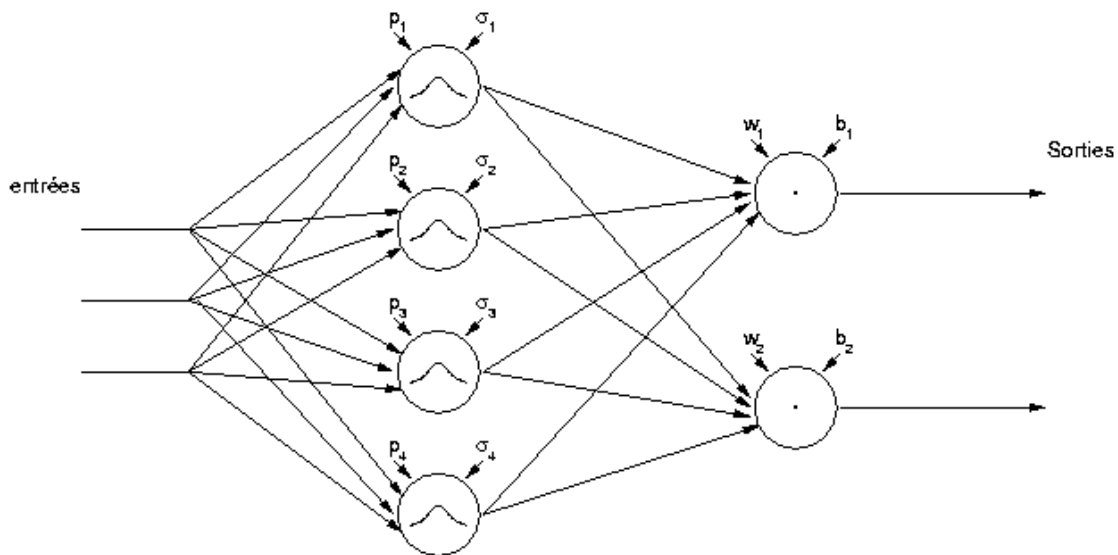


Figure III-12 Réseau de Neurones Artificiel à Fonction de Base Radiale (RFB).

III.2.5.2. Types des réseaux dynamiques (bouclés)

III.2.5.2.1. Les réseaux à apprentissage compétitif

Les réseaux à apprentissage compétitif, sont des types de modèle qui forcent une compétition des neurones afin d'activer un parmi un sous-ensemble.

D'une autre manière, leur principe consiste à grouper des vecteurs d'entrées similaires (c'est-à-dire des entrées qui sont corrélées) en sous-ensembles (LEPAGE, 2004), puis à favoriser parmi eux un seul neurone qui sera activé et transmis par la suite à la couche suivante, ce neurone appelé « **Neurone Gagnant** ».

On dispose de plusieurs méthodes d'apprentissage compétitives telles que les cartes auto-organisatrices avec la méthode d'apprentissage de Kohonen (SOM), quantification vectorielle (VQ) et Quantification vectorielle (LVQ).

III.2.5.2.1.1. Vector Quantization (quantification vectorielle)

Introduite par Grossberg en 1976, le principe de ce type de réseaux est de grouper les échantillons (les entrées du modèle) en sous-ensembles de propriétés similaires sous forme de vecteurs, contrairement à la quantification scalaire qui quantifie séparément les échantillons.(Enis Cetin & Nezih Gerek, 2006)

Chaque vecteur est appelé « *codeword* » et l'ensemble de ces *codeword* est appelé « *codebook* ».

La quantification vectorielle (QV), est une méthode utilisée dans l'apprentissage non supervisé, elle consiste à sélectionner un neurone vainqueur dont son vecteur poids est le plus proche du vecteur d'entrée parmi les autres neurones et seuls les poids du neurone vainqueur qui sont modifiés par la suite durant l'apprentissage (Decoux, 2008).

On peut la retrouver dans diverses applications comme la classification, la compression des données, la reconnaissance de formes, estimations de densité.

III.2.5.2.1.2. Les réseaux de Kohonen SOM (Self Organizing Map)

Le réseau auto-organisateur (Self Organizing Map en anglais), connu aussi sous le nom de cartes auto-adaptatives ou encore cartes (réseau) de Kohonen, par rapport à son créateur, le statisticien finlandais Teuvo Kohonen qui a développé le concept en 1982, il permet de catégoriser les données dans des groupes de propriétés similaires et chaque groupe ensuite est projeté en un nœud de la carte (RAFIK, 2008).

Les cartes Kohonen utilisent un apprentissage non supervisé et sont disposées sur une grille à une, deux ou trois dimensions (LEPAGE, 2004; TAHAR et al., 2003), représentées par des neurones chacun d'eux est lié à un vecteur référent, responsable d'une zone dans l'espace des données et associés aux vecteurs du poids, un calcul de la distance entre le vecteur d'entrée et le vecteur poids est effectué ensuite pour chaque neurone de la carte, le neurone le plus proche de l'espace d'entrées (qui a la petite distance calculée parmi l'ensemble), est appelé *le Neurone vainqueur*, ce dernier et ses voisins (le voisinage est défini par une fonction mathématique) misent à jours leurs vecteurs poids et déplacent leurs vecteurs référents vers le vecteur d'entrée, afin de minimiser la taille de voisinage et obtenir meilleures performances.

L'apprentissage s'arrête donc, lorsque les neurones ne bougent plus de façon significative. (Hugues, 1998).

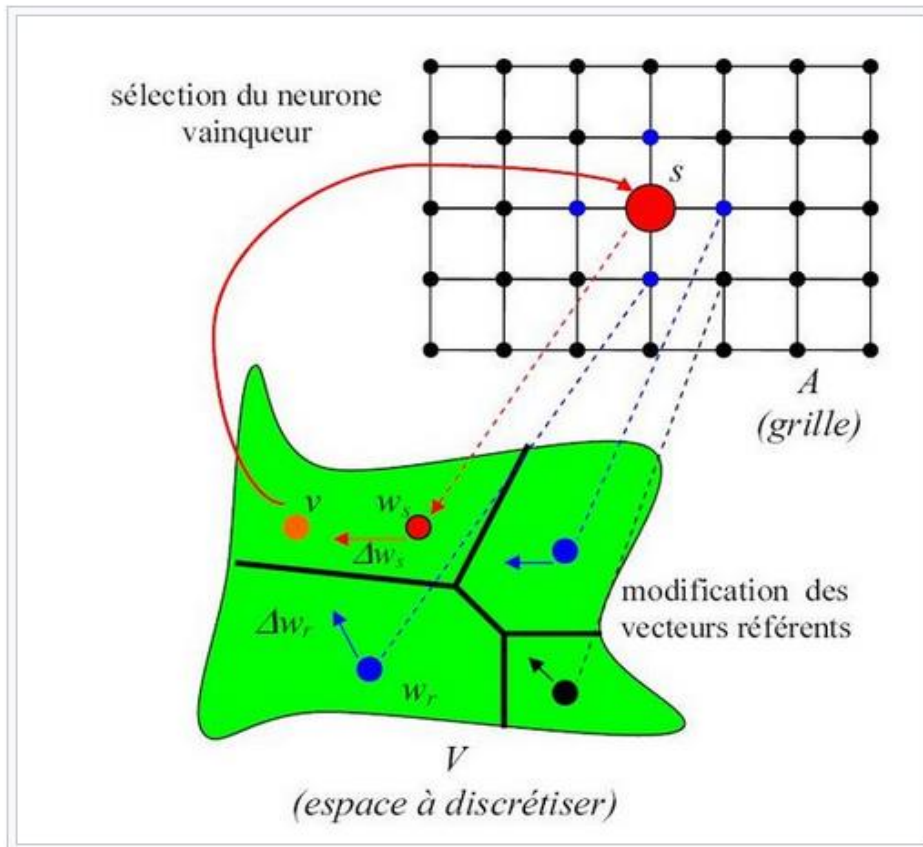


Figure III-13 Principe de fonctionnement de la carte de Kohonen

III.2.5.2.1.3. Learning Vector Quantization

Introduite par Kohonen (1988), LVQ est une méthode semblable à la quantification vectorielle, la seule différence c'est que LVQ utilise un apprentissage supervisé contrairement à la QV.

La technique d'apprentissage utilisée dans cette méthode lui permet de produire une meilleure topologie avec un nombre de neurones moins, que pour le réseau SOM. (LEPAGE, 2004).

III.2.5.2.1.4. La théorie des résonances adaptatives (ART)

ART (Adaptive Resonance Theory) C'est un réseau d'architecture évolutive introduit par Carpenter et Grossberg en 1987. Utilise un apprentissage supervisé et non supervisé, afin de résoudre des problèmes tels que la reconnaissance de formes et la prédiction.

Les réseaux ART sont capables de répondre avec succès au dilemme relatif à la « stabilité/plasticité » qui se pose dans les réseaux compétitifs, en fait un réseau plastique signifie que le réseau peut stocker de nouvelles entrées, et stable signifie qu'il peut protéger les motifs stockés d'être effacés (Manai, 2019), cependant, ART ont la capacité d'acquérir de nouvelles informations (données), sans perdre les informations qui sont enregistrées précédemment, ce qui les rend intéressants pour des entraînements continus avec les ajouts de données (Mahdi, 2011).

III.2.5.2.2. Les réseaux récurrents

Les réseaux de neurones récurrents sont des réseaux qui traitent les informations plusieurs fois formant un cycle (boucle), c'est-à-dire que les résultats sont renvoyés au sein du système au cours de l'apprentissage.

On distingue trois réseaux appartient à ce type :

III.2.5.2.2.1. Réseaux de Hopfield

Apparus en 1982 par Hopfield, ces réseaux utilisent un apprentissage non-supervisé et ils sont complètement connectés, il n'y a aucune différenciation donc entre les neurones entrés et sortis (Davallo & Naim, 1989), et que chaque neurone du réseau est responsable à la mise à jour de son propre état.

Ces réseaux sont utilisés dans la résolution de problèmes d'optimisation, ils opèrent comme une mémoire associative non-linéaire et sont capables de trouver un objet stocké en fonction de représentations partielles ou bruitées (Djarah, 2006; Mahdi, 2011).

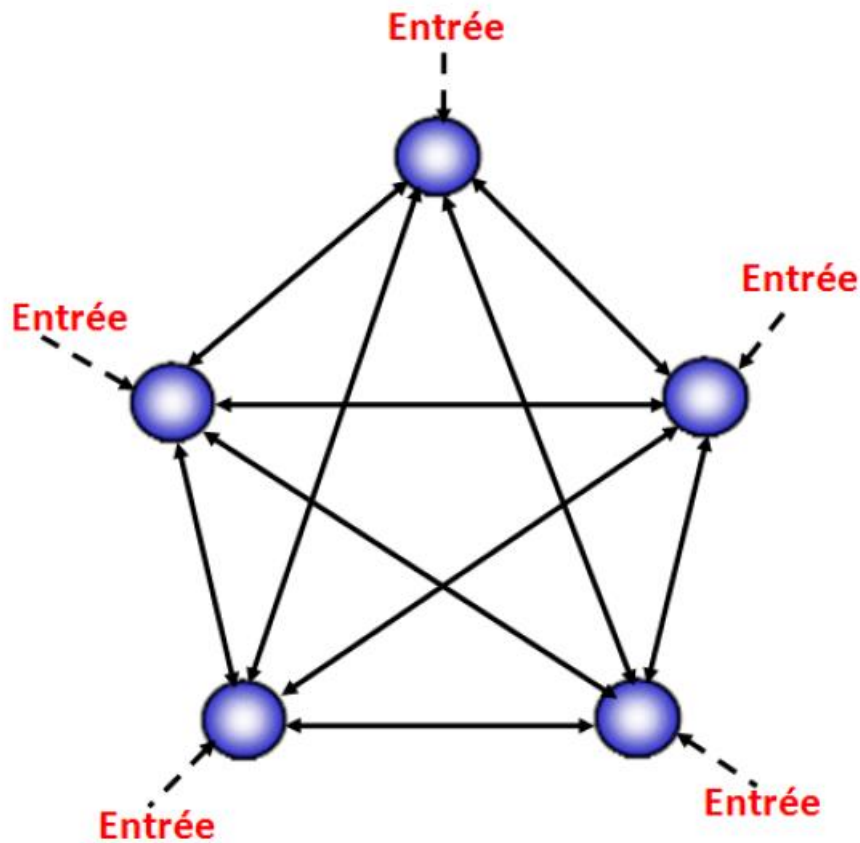


Figure III-14 Réseaux Hopfield

III.2.5.2.2.2. Réseaux Elman 1990

C'est un réseau à trois couches (entrée –cachée-sortie) et un ensemble d'unités appelées unités de contexte, les couches cachées sont connectées à la couche de sortie et aux unités de contexte qui enregistrent les sorties des neurones cachés et les propagent en arrière formant de nouvelles entrées aux neurones cachés.

Ce type est utilisé pour résoudre des problèmes de séries temporelles tels que : reconnaissance automatique de la parole, traduction automatique, reconnaissance de l'écriture manuscrite et la reconnaissance de formes.

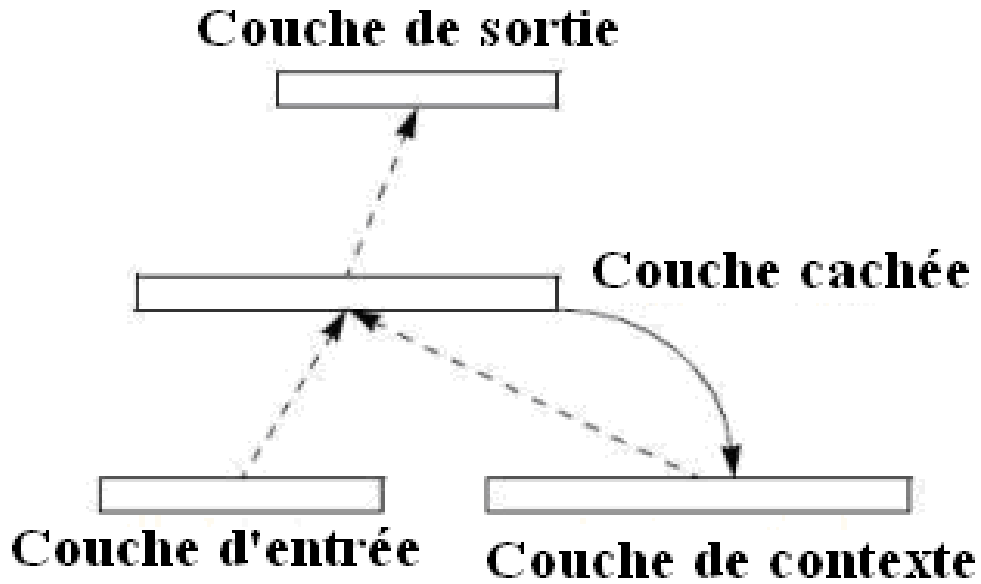


Figure III-15 Réseau Elman

III.2.5.2.2.3. Réseaux Jordan 1986

Ce type conserve le même principe du réseau Elman, la seule différence est que les unités de contexte sont alimentées par les sorties du système et non pas par les neurones cachés.

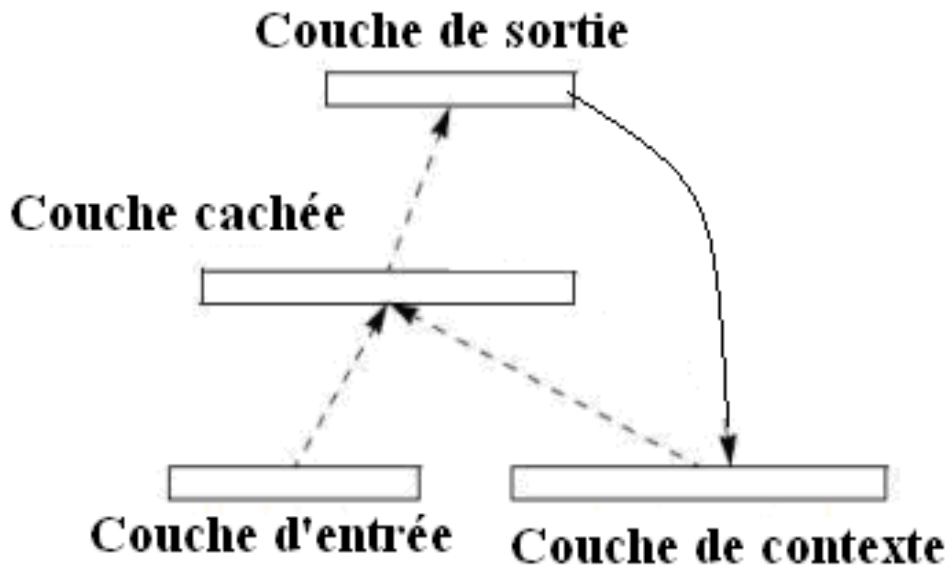


Figure III-16 Réseau Jordan

III.2.6.Apprentissage des Réseaux de Neurones Artificiels

L'apprentissage dans les réseaux de neurones artificiels, est une étape cruciale de la modélisation du réseau, il consiste principalement à modifier les paramètres du réseau par l'ajustement du poids de connexion entre les différents neurones, pour le but de rapprocher les sorties du modèle aux sorties désirées en minimisant l'erreur résultante au maximum.

Il existe plusieurs types d'apprentissage, et la manière ou bien la règle suivant laquelle l'apprentissage se produit, est appelée Algorithme d'apprentissage, qu'on peut la trouver également sous plusieurs types.

III.2.6.1. Types d'apprentissage

Différents types d'apprentissage existent, on peut citer quatre types dans ce qui suit :

III.2.6.1.1. Apprentissage supervisé

Dans ce type, les sorties du réseau sont forcées de prendre des valeurs désirées, qui sont calculées en fonction des entrées fournies et on les compare par les sorties cibles, dans ce cas, la présence d'un expert (professeur) est indispensable pour l'évaluation du réseau.

Pour atteindre les résultats désirés le réseau modifie progressivement les poids de connexion jusqu'à ce que l'erreur entre les sorties désirées du modèle et les sorties cibles sera minimisée.

III.2.6.1.2. Apprentissage non-supervisé

Contrairement à l'apprentissage supervisé, l'apprentissage non supervisé n'a pas besoin des valeurs de sorties cibles, et seules les valeurs d'entrées sont utilisées.

Dans ce type, l'apprentissage consiste à détecter des similarités dans l'ensemble des entrées, basant sur les régularités statistiques, et ajuste par la suite les valeurs des poids afin de produire des sorties proches en réponse pour des valeurs d'entrées similaires.

Ce type, reconnu aussi sous le nom d'apprentissage compétitif et d'auto-adaptateur, vu ces capacités de s'adapter automatiquement sans présence d'un expert humain.

III.2.6.1.3. Apprentissage renforcé

Ce type est similaire à l'apprentissage supervisé et certains auteurs le classent dans la catégorie des modes supervisés, cependant dans ce type les sorties cibles ne seront pas présentées au réseau, par contre un index de performance appelé aussi signal de performance est fourni au réseau.

Les réseaux de ce type sont capables donc de savoir si les sorties fournies sont correctes, mais la bonne réponse reste inconnue (Djarah, 2006).

III.2.6.1.4. Apprentissage hybride

Le nom hybride signifie une chose composée de deux éléments de différentes natures.

A cet égard ce type consiste à déterminer une partie des poids de connexion par le type d'apprentissage supervisé, et l'autre partie par le type d'apprentissage non supervisé (Djarah, 2006).

III.2.6.2. Phénomène de surapprentissage

Le surapprentissage est un phénomène apparu lorsque le modèle a trop appris les données présentées, par contre il devient incapable de reconnaître de nouvelles données.

Ce phénomène résulte d'un mauvais dimensionnement du réseau (trop de couches et d'unités cachées), ce qui lui permet d'apprendre l'échantillon d'entraînement avec un taux élevé, mais les performances du réseau en généralisation deviennent médiocres.

Dans un modèle constitué de k ensembles, où $k-1$ est utilisé pour l'entraînement du modèle et un ensemble est utilisé pour la validation du modèle, on peut parler de surapprentissage si on remarque que l'erreur dans la phase d'entraînement continue à diminuer, alors que dans la phase de validation l'erreur commence à accroître d'une manière remarquable..

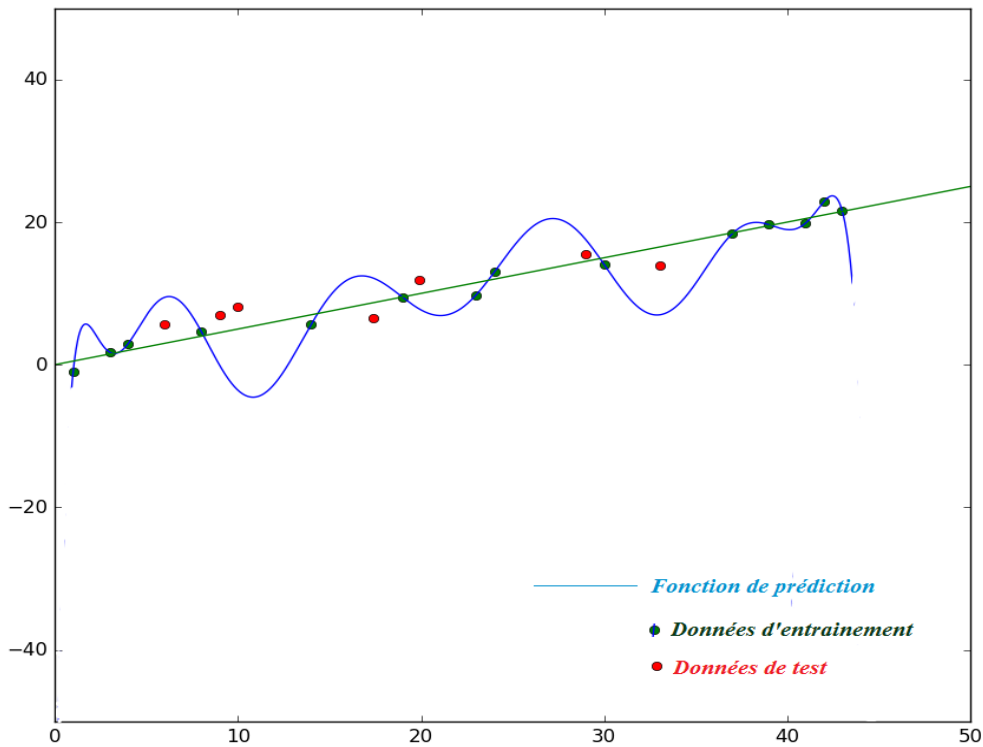


Figure III-17 Phénomène de surapprentissage

III.2.6.3. Phénomène de sous apprentissage

Le sous apprentissage apparaît lorsqu'un modèle demeure incapable de prédire les données d'entraînement, ce qui résulte en un taux d'erreur élevé dans la phase d'entraînement, et en conséquence, en phase de validation aussi, cela signifie qu'il ne pourra pas généraliser les prédictions en présence des nouvelles données.

Ce phénomène est la cause de l'utilisation de faible taille des données ou lorsqu'on essaie de construire un modèle linéaire avec des données non-linéaires.

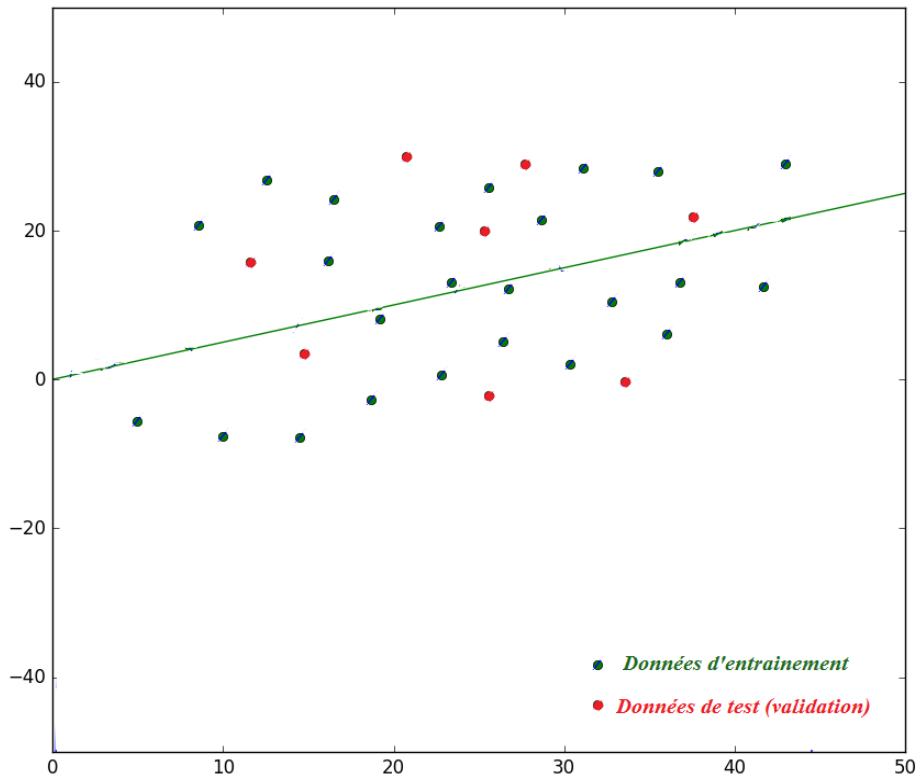


Figure III-18 Phénomène de sous apprentissage

III.2.6.4. Algorithmes d'apprentissage

L'algorithme d'apprentissage est une méthode mathématique qui permet d'ajuster les paramètres du réseau de neurones artificiels par la modification itérative « petit à petit » des poids de connexion, afin d'adapter la réponse du modèle à la réponse désirée en minimisant en conséquence l'erreur entre ces deux dernières. L'apprentissage se termine lorsque l'algorithme atteint un état stable.

On distingue plusieurs Algorithmes d'apprentissage on cite parmi lesquels :

L'algorithme de rétropropagation, l'Apprentissage de Boltzmann, Algorithme de ADALINE.

III.2.6.4.1. Algorithme de rétropropagation

Utilisé dans un apprentissage supervisé des réseaux multicouches, l'algorithme de rétropropagation est le plus répandu dans les réseaux de neurones artificiels grâce au succès marqué de cet algorithme dans certaines applications (lire un texte, détection de la fraude, lire des textes... etc.).

Il consiste de propager en première phase le vecteur d'entrée vers les sorties passant par les couches cachées, et calculer en suite l'erreur résultante entre les sorties du modèle et les sorties actuelles, dans la deuxième phase l'erreur est rétro propagée vers l'arrière de la couche de sortie vers la couche d'entrée passant par les couches cachées, pour le but de minimiser l'erreur obtenue, par ajustement des poids de connexion entre les différentes couches. Cette opération se répétera en plusieurs itérations jusqu'à l'obtention d'une erreur négligeable entre les sorties du modèle et les sorties actuelles.

Dans un réseau multicouches de « m » entrées et « n » sorties, et « L » couches cachées avec « j » unités, la valeur « I » d'un neurone caché « j » est obtenue par :

$$I_j = f\left(\sum_{i=1}^n W_{ji} * X_i + \theta_j\right) \quad (\text{III-6})$$

Et la sortie « O » d'un neurone de sortie « k » du modèle est obtenue dans la couche de sortie par :

$$O_k = g\left(\sum_{i=1}^n W_{kj} * I_j + \theta_k\right) \quad (\text{III-7})$$

Avec :

f et g : fonction de transfert dans la couche cachée et la couche de sortie respectivement.

W_{ji} et W_{kj} : les poids de connexion dans la couche cachée et la couche de sortie respectivement.

X_i : Entrée du modèle.

θ_j et θ_k : Biais dans la couche cachée et la couche de sortie.

L'erreur entre les sorties du modèle (O_k) et les sorties actuelles (Y_k) calculée comme suite :

$$E_p = \frac{1}{2} \sum_{k=1}^T (Y_k - O_k)^2 = \frac{1}{2} \sum_{k=1}^T (\delta_k)^2 \quad (\text{III-8})$$

Avec :

E : erreur (fonction de cout)

T : l'ensemble d'apprentissage

Remarque 1:

La fonction de coût présentée par l'équation III-8 est une erreur quadratique utilisée dans les problèmes de régression, ou le type d'apprentissage est supervisé et l'ensemble des données est continu, tandis que pour les problèmes de classification/apprentissage non supervisés la fonction de coût utilisée est « entropie croisée ». (Stricker, 2000).

Les poids de connexion sont ajustés par la méthode du gradient (équation III-9)

$$W_{ji}^L(\mathbf{n} + 1) = W_{ji}^L(\mathbf{n}) + \Delta W_{ji}^L(\mathbf{n}) \quad (\text{III-9})$$

$$\Delta W_{ji}^L(\mathbf{n}) = -\mu \frac{\partial E}{\partial W_{ji}^L(\mathbf{n})} \quad (\text{III-10})$$

Ou :

μ : Pas d'apprentissage représentant la vitesse de convergence, sa valeur est généralement choisie expérimentalement ($0 < \mu < 1$).

∂ : La dérivée partielle.

Pour plus de détails sur les démonstrations mathématiques, se réfère au (Decoux, 2008; Fayçal, 2010; LEPAGE, 2004; Mahdi, 2011; I. Otmani, 2011; Tsopze, 2010).

Remarque 2 :

La rétropropagation ce n'est pas seule, l'algorithme d'entraînement des réseaux de neurones artificiels, c'est uniquement la méthode de calcul du gradient, tandis que d'autres méthodes (algorithmes) d'optimisation sont utilisées pour réaliser l'apprentissage du modèle.

L'association rétropropagation-algorithme d'optimisation est la description la plus complète de l'algorithme général utilisé pour entraîner un réseau de neurones.

On distingue plusieurs algorithmes d'optimisation citant parmi les :

Algorithmes de 1^{ier} ordre :

Descente du gradient, gradient conjugué, Descente de gradient stochastique (SGD), Momentum.

Ces Algorithmes ne tiennent compte que de la première dérivée (gradient) de la fonction objective par rapport aux variables

Algorithmes de second ordre

Quasi-Newton, Newton (BFGS, Levenberg-Marquardt), Conjugate Gradient (CG), TRPO (Trust Region Policy Optimization) ...

Ces Algorithmes utilisent des informations sur la seconde dérivée (hessienne) de la fonction objective par rapport aux variables.

Les algorithmes de premier ordre ont le risque de tomber dans un minimum local, Il suffit que les gradients ou leurs dérivées soient nulles pour que le réseau se retrouve bloqué dans un minimum local, et l'utilisation des algorithmes de second ordre peut remédier à ce problème (Ammar, 2007).

D'autres algorithmes peuvent être utilisés selon le type de réseaux choisi à savoir : Adaptive Linear Element (ADALINE), MADALINE, ART1, ART2, Apprentissage de Boltzmann, Self Organizing Map (SOM).....etc.

III.2.6.5. Règles d'apprentissage

III.2.6.5.1. Correction d'erreur (la règle delta)

Utilisée dans un apprentissage supervisé, cette règle consiste à corriger la performance obtenue à la sortie du modèle par minimisation de l'erreur, entre cette sortie et la sortie désirée, et modifier en conséquence les poids de connexion entre les neurones afin de rapprocher la sortie du modèle à la sortie désirée, et obtenir une erreur la plus petite possible (tend vers zéro).

III.2.6.5.2. Règle de Hebb

Inspiré par la neurobiologie, Donald Hebb, établit un apprentissage localisé de part et d'autre d'une synapse (connexion), qui consiste à modifier les poids en fonction de l'activation d'un neurone d'entrée et d'un neurone de sortie qui sont associés à ce poids de connexion.

La forme de la règle d'apprentissage est donnée par :

$$\Delta w_j^t = \eta I_j^{t-1} O^{t-1} \quad \text{(III-11)}$$

Avec :

η : Vitesse de l'apprentissage

I_j^{t-1} : Activation du neurone d'entrée j au temps (t-1)

O_j^{t-1} : Activation du neurone de Sortie au temps (t-1)

III.2.6.5.3. Règle d'apprentissage compétitif

Comme on l'a vu précédemment, les réseaux à compétition regroupent les données corrélées (ayant les mêmes propriétés) et forcent une compétition entre les neurones du même groupe, et produisent en conséquence un neurone vainqueur et parfois un ensemble de neurones «voisins» du vainqueur.

L'apprentissage compétitif consiste à adapter seulement les poids des neurones vainqueurs et potentiellement, de ses voisins.

La règle d'apprentissage compétitif est écrite comme suit :

$$\Delta w = \begin{cases} \eta(x - w) & \text{si le neurone est vainqueur} \\ \mathbf{0} & \text{Autrement} \end{cases} \quad \text{(III-12)}$$

Avec :

η : Taux d'apprentissage $0 < \eta < 1$

X : Vecteur d'entrée

W : Vecteur poids

III.2.6.5.4. Règle de Boltzmann

Les réseaux de Boltzmann sont des réseaux récurrents symétriques totalement connectés, qui sont constitués de deux couches, la première dite visible, elle reçoit l'information de l'extérieur, et la deuxième dite cachée.

L'apprentissage consiste à modifier les poids jusqu'à l'état des neurones visibles converge à une distribution probabiliste voulue (Djarah Djalal 2006).

III.2.7. Modélisation par Réseaux de Neurones Artificiels

La modélisation par réseau de neurone artificiel, c'est la phase dont laquelle, le développeur choisit le type, la structure, l'architecture, l'algorithme d'apprentissage.... etc. appropriée au problème posé.

Cette modélisation consiste à :

- La détermination et préparation des vecteurs d'entrée et de sortie.
- La détermination du nombre des couches cachées et le nombre des neurones cachés.
- Choix des fonctions de transfert
- Réalisation de l'apprentissage
- Evaluation des résultats (Performance du modèle).

III.2.7.1. Préparation et Introduction des données dans le modèle

C'est une étape préliminaire indispensable à réaliser afin que le système du réseau de neurones artificiels puisse connaître et traite les données du problème étudié et produit en conséquence des résultats souhaitables, et sans passer pour cette étape le système peut fausser les résultats et donne des performances médiocres.

III.2.7.1.1. Préparation (pré traitement) des données

La préparation des données consiste à grouper les entrées dans un intervalle restreint, ce qui permet de système d'éviter le phénomène de saturation prématurée de la fonction d'activation de type sigmoïde. (Chamekh, 2008), ce phénomène produit des niveaux d'activation des neurones de sortie proches de 0 ou 1, et cela résulte aux très petites valeurs des poids de connexion qui sont connectés à ces derniers neurones, et qui seront alors négligeables à chaque itération suivante (Vitela & Reifman, 1997).

Différentes méthodes de préparation des données sont disponibles on distingue parmi les :

III.2.7.1.1.1. Normalisation des données

Cette méthode groupe l'ensemble des données dans un intervalle de $[-1 \ 1]$ « équation III-13 » ou $[0 \ 1]$ « équation III-14 », et ça donne au réseau la capacité de converger plus rapidement et de mieux généraliser les résultats obtenus.

$$X_n = 2 * \frac{X-MIN}{MAX-MIN} - 1 \quad \text{(III-13)}$$

$$X_n = \frac{X-MIN}{MAX-MIN} \quad \text{(III-14)}$$

III.2.7.1.1.2. Standardisation des données

La standardisation appelée aussi Z-score normalisation utilisée généralement lorsque le minimum et le maximum de l'ensemble des données est inconnu, elle consiste à transformer les valeurs en fonction de la moyenne et l'écart type, et produira par la suite une distribution de l'ensemble des données de moyenne 0 et d'écart-type 1.

$$X_n = \frac{X-\mu}{\sigma} \quad \text{(III-15)}$$

Avec :

μ : La moyenne et σ : L'écart type

III.2.7.1.1.3. Normalisation par échelle décimale.

Elle consiste de minimiser les valeurs des entrées par déplacement de la virgule décimale en les divisant par 10^i , la valeur de i dépend de la valeur absolue maximale de l'ensemble des entrées.

$$X_n = \frac{X}{10^i} \quad \text{(III-16)}$$

III.2.7.1.1.4. Préparation des données qualitatif

Le Réseau de neurones artificiels n'accepte pas les données de type qualitatif, pour cela chaque donnée de ce type doit impérativement être transformée en type quantitatif en attribuant à chaque type qualitatif un numéro différent.

III.2.7.1.2. Introduction (traitement) des données

Le développement d'un réseau de neurones artificiels repose sur une base de données relative au problème à traiter, cette dernière, après un prétraitement (voir section précédente), devra être divisée avant de l'introduire dans le réseau, en sous bases, l'une sera utilisée pour l'entraînement du modèle et la/les autres seront utilisées pour tester et valider les performances du modèle développer.

Il est à ne noter qu'elle n'existe aucune règle pour la détermination de ce partage de manière quantitative (Ammar, 2007).

Différentes méthodes de partage de la base de données pour le but de tester et valider le modèle on distingue parmi elles :

- **Méthode classique (Holdout method):**

Consiste à une division de la base de données en deux ensembles, un pour l'entraînement du modèle, et le deuxième sert à tester le modèle.

L'ensemble de données de l'entraînement sera plus grand que l'ensemble de test soit des proportions de: 80% / 20 % ou bien 75% / 35%.....etc.

- **Validation croisée :**

Dans certains cas la base de données disponible est restreinte et l'utilisation d'une partie de cette base comme un ensemble de test n'est pas pratique, le modèle ne pourra pas donc apprendre correctement, et les résultats de performances risquent de ne pas représenter de manière significative la capacité de généralisation du modèle. La validation croisée est une solution pour ce type de problème.

Elle s'agit de diviser la base de données en une série de sous-ensembles séparés d'entraînement et de validations, en plusieurs itérations, dans chacune de ces itérations les sous-ensembles changent, et les performances du modèle sont mesurées, sans jamais avoir validé le modèle sur une donnée déjà apprise, la performance du modèle alors c'est la moyenne de toutes les performances de chaque itération.

Il existe de nombreuses techniques de validation croisée, on peut distinguer parmi elle :

- **(K-folder) Validation croisée :**

Elle s'agit de diviser la base de données en k sous-ensembles séparés, l'un de ces k est utilisé pour valider le modèle et les autres k-1 sous-ensembles sont utilisés pour entraîner le modèle, ces opérations seront répétées k fois et un nombre de k performance sera obtenu à l'issue de la procédure.

La moyenne des k performances devra être calculée pour estimer la performance du modèle.

- **(Leave-one-out) Validation croisée:**

C'est un cas particulier de *k-fold* ou *k* sera égale au nombre d'échantillon (observation).

Remarque :

Dans la validation croisée il faut impérativement réserver un sous-ensemble de tests séparés aux autres sous-ensembles de validation croisée, afin de tester le modèle face aux données jamais vues par le modèle.

La base de données est donc divisée en sous-ensembles d'entraînement un autre de validation et le dernier pour le test.

III.2.7.2. Performance de modèle

La mesure de la performance est une étape très importante afin d'évaluer les capacités du modèle développé.

Les mesures de performance suivantes peuvent être utilisées pour évaluer le réseau de neurones artificiels :

- **Erreur quadratique moyenne (MSE)**

Mean Square Error en anglais, c'est une mesure de la moyenne des carrés des erreurs qui sont issues de la différence entre les résultats désirés et les sorties du modèle.

$$MSE = \frac{\sum_{i=1}^n (x_i - Y_i)^2}{n} \quad \text{(III-17)}$$

Avec :

x_i : sorties désirée

Y_i : sorties du modèle

n : nombre des observations

- **La racine de l'erreur quadratique moyenne**

Dite en anglais Root Mean Square Error, C'est la racine de MSE

$$RMSE = \sqrt{MSE} \quad \text{(III-18)}$$

- **Erreur quadratique moyenne normalisée (NMSE)**

Définit comme suit :

$$NMSE = \frac{1}{n} \sum_{i=1}^n \frac{(y_i - \hat{y}_i)^2}{\sigma_y^2} \quad (\text{III-19})$$

Où :

n est le nombre d'échantillons dans l'ensemble de données.

y_i représente les valeurs réelles ou observées.

\hat{y}_i représente les valeurs prédites par le modèle.

σ_y^2 est la variance des valeurs réelles.

- **La somme des carrés des erreurs (SSE)**

Comme son nom l'indique, c'est la somme des erreurs carrées (en anglais : Sum Square Errors).

$$SSE = \sum_{i=1}^n (X_i - Y_i)^2 \quad (\text{III-20})$$

- **Erreur absolue moyenne (MAE):**

C'est la Moyenne arithmétique des valeurs absolues des écarts calculé selon la formule suivante :

$$MAE = \frac{\sum_{i=1}^n |X_i - Y_i|}{n} \quad (\text{III-21})$$

- **Mean Absolute Percentage Error (MAPE):**

C'est l'erreur absolue moyenne en pourcentage, également appelée écart en pourcentage absolu moyen, elle est calculée selon la formule suivante :

$$MAPE = \frac{100}{n} * \sum_{i=1}^n \left| \frac{X_i - Y_i}{X_i} \right| \quad (\text{III-22})$$

- **Précision de performance (Accuracy Performance, Average Accuracy Percentage AA %):**

Définit comme (100-MAPE)%

$$AA\% = 100 - MAPE \quad (\text{III-23})$$

- **Coefficient de corrélation r :**

Dans la mesure de MSE, les résultats obtenus ne peuvent pas représenter le sens des sorties vu l'utilisation des carrées des erreurs, on peut obtenir une petite MSE mais les sorties désirées et les sorties du modèle ayant deux sens différents, le coefficient de corrélation « r » peut résoudre ce problème.

Le coefficient de corrélation « r » produit des résultats entre [-1,1], ou « 1 » représente une forte corrélation positive entre les sorties désirées et les sorties du modèle, et que ces dernières varient dans le même sens, par contre « -1 » représente une forte corrélation négative entre les sorties désirées et les sorties du modèle, et que ces dernières varient dans deux sens opposés, ainsi que le « 0 » signifie que les deux sorties ne sont pas corrélées. (Mady, 2013).

$$r = \frac{\sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum_{i=1}^n (x_i - \bar{x})^2 * \sum_{i=1}^n (y_i - \bar{y})^2}} \quad \text{(III-24)}$$

Ou :

x_i : sorties désirée

Y_i : sorties du modèle

\bar{x} et \bar{y} : La moyenne de x et la moyenne de y

III.2.8. Avantages et inconvénient des Réseaux de Neurones Artificiels

III.2.8.1. Avantages

Les RNA conviennent bien aux problèmes complexes où la relation entre les variables sont inconnues (Mady, 2013; Shahin et al., 2002).

- Ils sont capables de traiter des problèmes avec des données incomplètes ou bruitées (Mady, 2013)
- Ils ont la possibilité de s'adapter et d'améliorer leurs performances au cours d'entraînement en présence des nouvelles données.
- Ils généralisent les connaissances apprises pour produire des réponses adéquates à des situations inconnues.

- Capabilité de traiter des problèmes non-linéaires ou les résultats ne sont pas proportionnels aux données.

III.2.8.2. Inconvénients

- Les réseaux de neurones artificiels sont des boîtes noires, on ne peut pas connaître comment les résultats sont obtenus.
- Il n'existe aucune loi, normes, techniques et méthodes systématiques pour choisir l'architecture du modèle à développer et seuls les essais qui peuvent donner un modèle qui produit les résultats escomptés.

III.3. Introduction à la Méthode Paramétrique

La méthode paramétrique est une technique utilisée généralement dans les problèmes d'estimation, et elle a connu son succès dans le domaine d'estimation des coûts, elle consiste principalement à supposer au préalable la forme d'une fonction f qui relie la variable indépendante (la sortie désirée) et les variables indépendantes (les facteurs), et à définir par la suite ses paramètres, sachant que le nombre de ces derniers peut être défini à l'avance.

La régression linéaire simple ou multiple, l'analyse bayésienne et l'interpolation polynomiale, sont des exemples de la méthode paramétrique.

III.3.1. Types des modèles paramétriques

Comme on a dit précédemment la méthode paramétrique s'appuie sur une forme prédéfinie de la fonction f (dite équation centrale), et depuis son apparition différents modèles ont été développés (Halstead, 1975), IBM-FSD (Walston, C. E., & Felix, 1977) PUTNAM-SLIM (Putnam, 1978), COCOMO'81 (Boehm, 1981), COCOMOII (Boehm et al., 1995).

Ces modèles paramétriques peuvent être classifiés selon trois principaux critères :

- La forme de l'équation centrale (analytique ou linéaire)
- Le type de modèle (soit dynamique s'il contient la variable temps dans son équation centrale, ou bien statique s'il ne la contient pas).
- Modèle peut être détaillé ou global

III.3.1.1. Modèles Analytiques

Ce type de modèle utilise une fonction centrale non-linéaire, elle figure généralement dans les cas où la relation initiale entre les facteurs et la variable à exprimer est non-linéaire et elle ne pourra en aucun cas être linéaire par manque de possibilité de transformation des variables. (IDRI, 2003).

On distingue deux principaux modèles analytiques :

- Le modèle de Putnam
- Le modèle d'Halstead

Vous pouvez trouver plus de détail pour ces deux méthodes dans (Halstead, 1975) et (Putnam, 1978)

III.3.1.2. Modèles Linéaires

Contrairement aux modèles analytiques, ce type de modèle utilise une fonction centrale linéaire sous la forme suivante :

$$Y_i = \beta_0 + \beta_1 X_{i1} + \beta_2 X_{i2} + \dots + \beta_p X_{ip} + \varepsilon_i \quad \text{(III-25)}$$

Avec Y_i représente la variable à expliquer (la réponse du modèle/variable dépendante), les β_i sont les paramètres à estimer du modèle et les X_i sont les facteurs du problème, ε est l'erreur.

Ces modèles sont souvent réalisés en utilisant une étude statistique basée sur la régression linéaire simple ou multiple selon le nombre de variables X_i , ces dernières expriment avec fiabilité la variable dépendante en fonction des facteurs (variables indépendantes).

Dans ce qui suit on va présenter le principe de ces deux méthodes.

III.3.1.2.1. Régression linéaire simple

La régression linéaire simple est une méthode statistique qui sert à décrire la relation entre la variable à expliquer (variable dépendante) et une seule variable explicative (variable indépendante) (Genin, 2005), comme son nom l'indique cette relation doit être linéaire et prendre la forme suivante :

$$Y = \beta_0 + \beta X + \varepsilon \quad \text{(III-26)}$$

Ou :

Y : Variable à expliquer (variable dépendante)

X : Variable explicative (variable indépendante)

β_0 et β : Coefficients (ordonnée à l'origine et pente).

ε : erreur aléatoire.

L'estimation de la variable dépendante (Y) consiste à déterminer les valeurs des coefficients (β_0 et β) qui minimisent au maximum l'erreur, cette estimation est obtenue par l'utilisation des différentes méthodes telles que : maximum de vraisemblance, inférence bayésienne ou la méthode des moindres carrés qui est la plus utilisée. (Chouquet, 2010; A. Otmani, 2020).

III.3.1.2.2. Régression linéaire multiple (RLM)

Identiquement à la régression linéaire simple, le principe de la régression linéaire multiple est de décrire la relation entre la variable à expliquer (variable dépendante) et plusieurs variables explicatives (variables indépendantes), sous forme linéaire.

Dans RLM le résultat du modèle « variable dépendante » est expliqué par une combinaison linéaire des couples des variables indépendantes et les paramètres « coefficients » du modèle comme le montre l'équation suivante :

$$Y_i = \beta_0 + \beta_1 X_{i1} + \beta_2 X_{i2} + \dots + \beta_p X_{ip} + \varepsilon_i \quad (\text{III-27})$$

L'équation (III-27) peut s'écrire sous forme matricielle comme suit :

$$Y = X\beta + \varepsilon \quad (\text{III-28})$$

La régression linéaire multiple doit respecter les hypothèses suivantes :

- Les données de l'étude doivent être homogènes, indépendantes et suivre une loi normale
- La relation entre Y et Xi doit être linéaire
- Les termes d'erreurs sont indépendants identiquement distribués, de variance constante et, suivent une loi normale « $\mathcal{N}(0, \sigma^2)$ »
- Les paramètres inconnus β_0, \dots, β_p sont supposés constants.

III.3.1.2.3. L'estimation du modèle de régression

D'après la forme générale de la régression linéaire multiple (équation III-27), on constate que les coefficients (β_0, \dots, β_p) sont les seuls inconnus qui agissent sur l'estimation de la réponse Y, à cet effet, cette dernière ne sera obtenue qu'après l'estimation desdits coefficients.

On distingue nombreuses méthodes d'estimation (voir paragraphe III.3.1.2.1). Dans cette étude nous n'intéressons qu'à la méthode des moindres carrés et son principe sera discuté dans ce qui suit.

III.3.1.2.3.1. L'estimation par la méthode des moindres carrés.

Apparus au début du 19^{ème} siècle, cette méthode a été développée par *Carl Friedrich Gauss*, et *Adrien-Marie Legendre* indépendamment.

La méthode des moindres carrés consiste à minimiser la somme quadratique des erreurs (distance entre les valeurs réelles et les valeurs prédites), en terme mathématique, c'est l'annulation de vecteur de la dérivée partielle des erreurs en fonction de des coefficients (paramètres « β »).

L'équation de la somme des erreurs s'écrit donc :

$$f(\beta) = \sum_{i=1}^n e_i^2 = \sum_{i=1}^n (Y_i - \beta_0 - \beta_1 X_{i1} - \beta_2 X_{i2} - \dots - \beta_p X_{ip})^2$$

On a :

$$f(\beta) = \sum_{i=1}^n e_i^2 = e'e$$

Alors :

$$f(\beta) = \sum_{i=1}^n e_i^2 = e'e = (Y - X\beta)'(Y - X\beta)$$

On appliquant les propriétés des matrices on obtient :

(Soit A et B deux matrices : $(A+B)'=A'+B'$ et $(AB)'=B'A'$)

$$f(\beta) = Y'Y - Y'X\beta - \beta'X'Y - \beta'X'X\beta$$

On constate que $Y'X\beta$ et $\beta'X'Y$ est la multiplication d'un vecteur ligne Y' par un vecteur colonne $X\beta$ (ou $\beta'X'$ par Y), ce qui résulte une matrice de dimension (1.1) alors un scalaire et que :

$$\beta'X'Y = (Y'X\beta)' \text{ Ainsi que la transposée d'un scalaire est un scalaire alors : } \beta'X'Y = Y'X\beta$$

L'équation $f(\beta)$ s'écrit donc :

$$f(\beta) = \sum_{i=1}^n e_i^2 = Y'Y - 2Y'X\beta - \beta'X'X\beta \quad \text{(III-29)}$$

Calculons le vecteur des dérivées partielles $f(\beta)$ par rapport à β :

$$\frac{\partial f(\beta)}{\partial \beta} = \frac{\partial(Y'Y)}{\partial \beta} - \frac{\partial(2Y'X\beta)}{\partial \beta} - \frac{\partial(\beta'X'X\beta)}{\partial \beta}$$

Note :

A, B et X sont des matrices

1. $A'A$: est une matrice symétrique

2. $\frac{\partial(A'BX)}{\partial X} = B'A$

3. $\frac{\partial(X'AX)}{\partial X} = (A + A')X$: si A est une matrice symétrique

alors : $\frac{\partial(X'AX)}{\partial X} = 2AX$

$$\frac{\partial f(\beta)}{\partial \beta} = 0 - 2X'Y - 2X'X\beta$$

La minimisation de la somme d'erreur s'obtient en annulant sa dérivée partielle :

$$\frac{\partial f(\beta)}{\partial \beta} = 2X'Y - 2X'X\beta = 0$$

$$2X'Y = 2X'X\beta \implies X'Y = X'X\beta$$

En faisant l'hypothèse qu'il n'existe pas de colinéarité entre les variables explicatives et donc la matrice $X'X$ est inversible nous obtenons l'estimation de β :

$$\hat{\beta} = (X'X)^{-1}X'Y \quad \text{(III-30)}$$

Pour montrer que l'on a bien obtenu un minimum, on calcule la matrice hessienne des dérivées secondes

$$H = \frac{\partial^2 f(\beta)}{\partial \beta \partial \beta'} = 2X'X$$

Cette matrice est définie positive, on a donc bien un minimum

Le vecteur des prédictions des Y (appelé aussi le vecteur des valeurs ajustées) s'obtient en fonction de X et β et s'écrit donc

$$\hat{Y} = X\hat{\beta} = X(X'X)^{-1}X'Y$$

$$\hat{Y} = H_x Y \quad \text{(III-31)}$$

Avec $H_x = X(X'X)^{-1}X'$ est appelée "hat matrix"

Le vecteur des valeurs ajustées peut être interprété comme la projection de y sur le sous-espace engendré par les colonnes de la matrice X, où H_x est le projecteur.

La différence entre Y et le vecteur \hat{Y} des valeurs ajustées est le vecteur des résidus, il s'écrit comme suit :

$$e = Y - \hat{Y} = Y - X(X'X)^{-1}X'Y = (I - X(X'X)^{-1}X')Y$$

$$e = P_x Y \quad \text{(III-32)}$$

Le vecteur des valeurs ajustées peut également être interprété comme la projection de y dans le noyau de X' (ou l'orthogonal du sous-espace engendré par les colonnes de X), où P_x est le projecteur.

III.3.1.2.3.1.1. Propriétés des estimateurs

Le calcul de l'espérance mathématique des estimateurs $\hat{\beta}$ nous donne le résultat suivant :

$$E(\hat{\beta}) = \beta \quad \text{(III-33)}$$

D'après l'équation IV-9, on peut dire que l'estimateur est sans biais et la matrice de variance covariance des estimateurs sera donc :

$$E\left[(\hat{\beta} - E(\beta))(\hat{\beta} - E(\beta))'\right] = E[(\hat{\beta} - \beta)(\hat{\beta} - \beta)'] = \sigma_\varepsilon^2(X'X)^{-1} \quad \text{(III-34)}$$

Parmi les estimateurs sans biais fonctions linéaires des Y_i , il n'existe pas d'estimateurs sans biais avec une variance plus faible que celle des moindres carrés ordinaires, (Théorème de Gauss-Markov) Les estimateurs des MCO sont BLUE (best linear unbiased estimator). « pour plus de détails se conférer aux travaux (Chesneau, 2017; Rakotomalala, 2015).

La matrice de variance covariance des prédicteurs est :

$$E[(\hat{Y} - X\beta)(\hat{Y} - X\beta)'] = \sigma_\varepsilon^2 H_x \quad (\text{III-35})$$

La matrice de variance covariance des Résidus sous l'hypothèse d'hétéroscédasticité (il n'y a pas d'autocorrélation des erreurs) est :

$$E[(e - \varepsilon)(e - \varepsilon)'] = \sigma_\varepsilon^2 I \quad (\text{III-36})$$

Avec I : matrice identité.

L'estimateur sans biais de σ_ε^2 s'écrit comme suit :

$$\widehat{\sigma}_\varepsilon^2 = \frac{\|Y - X\beta\|^2}{n - (p+1)} = \frac{\sum e_i^2}{n - p - 1} = \frac{SSE}{n - p - 1} \quad (\text{III-37})$$

III.3.1.2.3.1.2. Critère de jugement de la qualité de l'ajustement d'un modèle

L'évaluation de l'erreur s'écrit :

$$SST = SSR + SSE \quad (\text{III-38})$$

Avec :

$SST = \sum (Y_i - \hat{Y})^2$: Somme totale des carrés

$SSR = \sum (\hat{Y}_i - \bar{Y})^2$: La somme des carrés de la régression

$SSE = \sum e_i^2$: La somme des carrés des résidus

Et le coefficient de détermination R^2 représente le rapport entre SSR et SST

$$R^2 = \frac{SSR}{SST} \quad (\text{III-39})$$

Ce coefficient R^2 se définit alors comme la proportion de variance expliquée dans la variance totale, il donne des résultats compris entre 0 et 1, il mesure la qualité d'ajustement du modèle, plus R^2 tend vers 1 plus le modèle est meilleur. (Chesneau, 2017).

III.3.1.2.4. Les différents tests utilisés dans la Régression linéaire multiple

Avant de modéliser un problème avec la régression linéaire multiple, il faut tout d'abord vérifier les hypothèses de cette méthode (voir section III.3.1.2.2).

III.3.1.2.4.1. Test de normalité :

On peut tester la normalité des données de l'étude par différents tests soit statistiques ou graphiques qu'on va présenter quelques-uns :

➤ Tests statistiques :

Test de Kolmogorov-Smirnov et Shapiro-Wilk : Dans ces deux tests on peut dire que l'échantillon de l'étude provient d'une population normale si la signification (P- Value) est supérieure à un niveau α choisi (ex : 0.05)

Coefficients d'asymétrie et d'aplatissement :

Les données suivent une loi normale si le coefficient d'asymétrie tend vers « 0 » et le coefficient d'aplatissement vaut « 3 » (0 si normalisé), ou d'une autre manière si le z-value du coefficient d'asymétrie (et/ou coefficient d'aplatissement) devra être comprise entre [-1.96, +1.96] avec le z-value = coefficient d'asymétrie (d'aplatissement) / Standard error. (N. Thomas & Thomas, 2016).

On peut distinguer d'autres tests statistiques tel que : test de Lilliefors, test de Anderson-Darling, test de Jarque-Bera et le test D'Agostino. (Rakotomalala, 2015).

➤ Tests graphiques :

Histogramme de fréquence

C'est un outil graphique simple à réaliser, il consiste à partager l'intervalle de définition d'une variable en k intervalles de largeur égale, ensuite à tracer une série de barres dont la hauteur est proportionnelle à l'effectif associé à l'intervalle.

On compare ensuite la distribution (Histogramme) avec la courbe de densité (courbe représentant une loi normale), et on peut prononcer la normalité des données si la distribution semble symétrique.

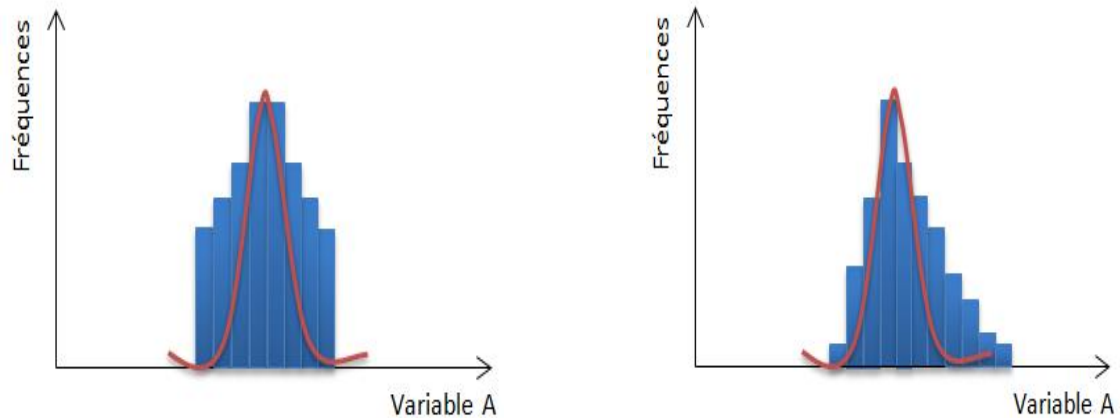


Figure III-19 Histogramme de fréquence

Boîte à moustache (Box-plot)

Box-plot en anglais, c'est un graphe permettant de représenter la distribution des données en utilisant quelques paramètres de localisation comme la médiane, le 1er et le 3eme quartile.

Les données sont normales lorsque la distribution constatée est symétrique .

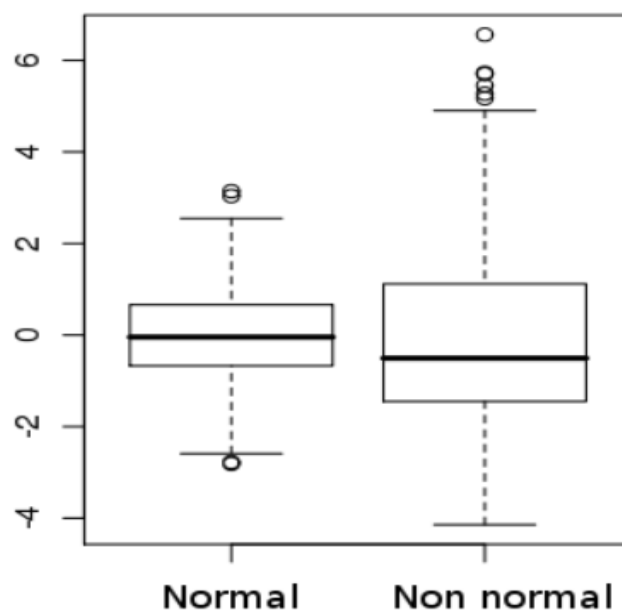


Figure III-20 Boite à moustaches (Box-plot)

Q-Q Plot et Droite de Henry

Q-Q plot (le diagramme Quantile-Quantile) est une méthode graphique reposant sur le calcul des Quantiles, elle sert à comparer l'ajustement d'une distribution des données observées à celle d'un modèle théorique (ex : normale, exponentielle...).

La droite de Henry, c'est une représentation de la loi gaussienne (loi normale), et on parle d'une distribution normale des données si on constate que les nuages des points (Quantile observé «X» ; Quantile théorique «X*») sont compatibles avec la droite de Henry et alignés sur la diagonale principale. (Rakotomalala, 2015)

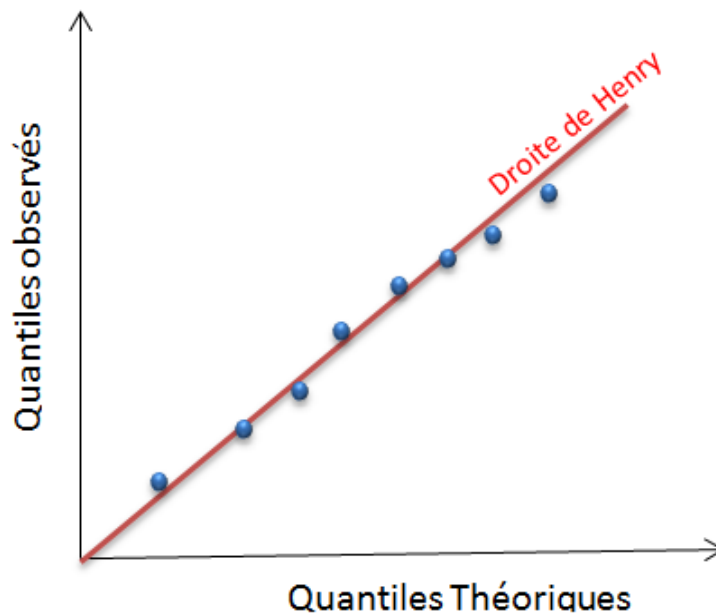


Figure III-21 Diagramme Q-Q et Droite de Henry

III.3.1.2.4.2. Tests d'homogénéité

Ces tests sont réalisés pour évaluer l'égalité des variances (homogénéité de la variance ou homoscedasticité), on distingue différents tests d'homogénéité :

Test de Breusch-Pagan : C'est un test basé sur un test du χ^2 Chi-Deux.

Test de Goldfeld et Quandt : s'appuie sur la loi de Fisher.

Test de Bartlett : vérifier que la signification de ce test soit supérieure à un niveau choisi (EX : 5%)

Test de Levene : Il utilise les écarts absolus des valeurs aux moyennes des groupes, et teste par la suite l'hypothèse nulle que les variances de population sont égales, un P-value supérieur à un niveau choisi (EX : 5%) indique l'homogénéité.

Test de Brown–Forsythe : il est identique au test de Levene la seule différence est que ce test utilise les écarts absolus aux médianes de groupes

III.3.1.2.4.3. Test de linéarité et indépendance

Le test de linéarité sert à vérifier la relation de linéarité entre les variables explicatives (les données notées X_i) et la variable à expliquer (la réponse notée Y), il est réalisable en calculant la corrélation entre X_i/Y , une forte corrélation (« r » proche de 1) indique qu'il existe une relation linéaire entre les données et la réponse.

Le test de l'indépendance sert à vérifier que les variables (ou les résidus) sont indépendantes, en calculant la corrélation « r » entre eux, une faible valeur de « r » indique l'indépendance entre les X_i (e_i).

III.3.1.2.4.4. Intervalle de confiance et Tests d'hypothèses

L'intervalle de confiance est un intervalle autour de la valeur estimée qui contient la valeur réelle avec une grande probabilité. Il doit être associé à un niveau de confiance (appelé aussi seuil de signification) noté « α », généralement représenté en pourcentage (05%, 01%, 0.1%).

Un test d'hypothèse consiste à déterminer les valeurs selon lesquelles la statistique (la variable aléatoire) peut varier, en effet il fournit une règle de décision permettant, sur la base de résultats d'échantillon, de faire un choix entre deux hypothèses statistiques (notées par la suite H_0 et H_1), tout en acceptant un risque « α » défini au préalable.

III.3.1.2.4.4.1. Test de nullité d'un paramètre du modèle (Test de Student)

Ce test est utilisé pour savoir si une variable joue un rôle explicatif dans un modèle, en testant la significativité du coefficient de la variable explicative sous l'hypothèse nulle.

$H_0 : \beta_j=0 \rightarrow$ le coefficient n'est pas significatif

$H_1 : \beta_j \neq 0 \rightarrow$ le coefficient est significatif

La statistique de *Student* sous l'hypothèse nulle s'écrit :

$$T = \frac{\widehat{\beta}_i}{\widehat{\sigma}_{\widehat{\beta}_i}} \sim \text{Student (n-p-1)} \quad (\text{III-40})$$

Ou « n » est la taille d'échantillon et « p » est le nombre des variables.

Si $|T| \geq t_{1-\alpha/2; n-p} \rightarrow$ on rejette H_0 et on conclut donc que le coefficient est significatif (significativement différent de zéro et la variable joue un rôle explicatif dans le modèle)

Si contraire $|T| \leq t_{1-\alpha/2; n-p} \rightarrow$ On accepte H_0 et on conclut donc que le coefficient n'est pas significatif.

Remarque :

$(t_{1-\alpha/2; n-p})$ représente le quantile d'ordre $\alpha/2$ d'une variable aléatoire de Student à $n-p$ degrés de liberté, qu'on peut l'extraire du tableau des valeurs du quantile.

Lorsque la taille d'échantillon est grande ($n > 30$), on peut comparer $|T|$ directement avec le seuil critique de la loi normale centrée et réduite qui est 1,96 (pour un risque de 5%) car d'après le théorème central limite, la loi de Student tend vers une loi normale lorsque T est suffisamment grand. (Hamisultane, 2002).

III.3.1.2.4.4.2. Test de significativité global (plusieurs coefficients) : Test de Fisher

Ce test est utile pour tester la significativité de plusieurs coefficients d'un modèle, ce qui nous permet de connaître l'influence de q variables explicatives (avec $q \leq p$) sur la variable à expliquer « Y ».

Les hypothèses du test de Fisher sont les suivantes :

$H_0 : \beta_1 = \beta_2 = \dots = \beta_q = 0 \rightarrow$ l'ensemble de « q » coefficients du modèle est non significatif

$H_1 : \rightarrow$ il existe au moins un coefficient non nul.

La statistique de ce test s'écrit donc :

$$F = \frac{SSR_0 - SSR_1}{SSR_1} * \frac{n-p-1}{q} \sim F(q, n-p-1) \quad (\text{III-41})$$

Ou « n » est la taille d'échantillon et « p » est le nombre des variables.

SSR_0 : Somme des carrés des résidus du modèle "réduit" sous H_0 .

SSR_1 : Somme des carrés des résidus du modèle de référence.

La règle de décision est la suivante :

Si la valeur de $F \geq F_{\text{tabulé}}$ (donnée par le tableau des quantiles de la loi Fisher pour un risque $1-\alpha$ donné) alors on rejette H_0 et on accepte en revanche H_1

Remarque :

Lorsqu'on accepte l'hypothèse H_1 qui indique qu'il y a au moins un coefficient non nul, le test ne précise pas lesquels. Il est moins précis que le test de Student (Hamisultane, 2002).

III.3.1.2.4.4.3. Test de nullité de tous les paramètres du modèle :

Ce test est utilisé pour connaître la nullité de tous les coefficients associés aux variables explicatives.

$H_0: \beta_1 = \beta_2 = \dots = \beta_p = 0 \rightarrow$ tous les coefficients du modèle associé aux variables explicatives sont nuls

Le modèle s'écrit donc sous H_0 :

$$Y_i = \beta_0 + e_i \quad \text{(III-42)}$$

La statistique de Fisher- Snédécour s'écrit :

$$F = \frac{SSL_1}{SSR_1} * \frac{n-p-1}{p} \sim F(p, n - p - 1) \quad \text{(III-43)}$$

Où SSL_1 est la somme des carrés du modèle de référence.

Si la valeur de $F > F_{\text{tabulé}}$ (donnée par le tableau des quantiles de la loi Fisher pour un risque $1-\alpha$ donné) alors on rejette H_0 .

III.3.1.2.5. Analyse de la variance ANOVA

Cette analyse consiste à vérifier si les moyennes des groupes proviennent d'une même population, elle permettant de savoir en conséquence si la variable dépendante de l'étude est en relation avec les variables explicatives (indépendantes).

Si la valeur de significativité (p-value) est inférieure au risque α on conclut que toutes les moyennes de population ne sont pas égales, et on peut accepter que les variables explicatives sont en relation avec la variable dépendante.

III.3.1.2.6. Performance de modèle

Pour évaluer la qualité d'ajustement d'un modèle de régression linéaire on peut procéder au calcul des mesures suivantes :

- *Coefficient de corrélation « r »*

Mesure la relation entre la variable dépendante et les variables explicatives, un coefficient de corrélation proche de 1 exprime une forte relation entre lesdites variables.

$$r = \frac{Cov(X,Y)}{\sigma_x \sigma_y} \quad \text{(III-44)}$$

- *Coefficient de détermination « R2 »*

Il situe entre 0 et 1 plus le résultat et proche de 1 plus le modèle et bon.

$$R^2 = \frac{SSR}{SST} \quad \text{(III-45)}$$

- *Coefficient de détermination ajusté.*

Si on ajoute des nouvelles variables à l'équation de régression le coefficient de détermination tend à augmenter aussi même si ces variables ne sont pas significatives, le recours au coefficient de détermination ajusté est la solution de ce problème

$$R^2_{ajusté} = 1 - \frac{(1-R^2)(n-1)}{(n-k-1)} \quad \text{(III-46)}$$

Ou : n est le nombre d'échantillon et k le nombre des variables explicatives.

- *Autres mesures de performance comme (MSE, MAPE, AA%, SSE, MAE...)* (voir section 2.7.2 du chapitre III).

III.3.1.2.7. Analyse des résidus

Après avoir réalisé un modèle de régression linéaire, il doit respecter les hypothèses statistiques pour nous l'acceptons comme un modèle fiable, l'analyse des résidus est faite en général pour se confirmer à ces hypothèses considérées au préalable (voir section 2.7.2 chapitre III).

III.3.1.2.7.1. Normalité des résidus

Les tests statistiques et graphiques de normalité sont les mêmes présentés dans le Chapitre III section 3.1.2.4.1

III.3.1.2.7.2. Homoscédasticité des résidus (variance constante)

Les tests présentés dans le Chapitre III section 3.1.2.4.2 sont utilisés pour vérifier l'homoscédasticité des résidus.

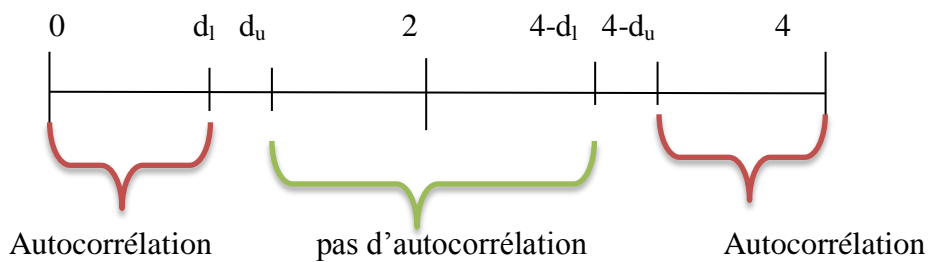
III.3.1.2.7.3. Test de Durbin-Watson d'autocorrélation des erreurs

Ce test est utilisé pour vérifier l'indépendance des résidus (les corrélations en série entre les résidus), sa formule est la suivante :

$$DW = \frac{\sum_{i=2}^n (e_i - e_{i-1})^2}{\sum_{i=1}^n e_i^2} \tag{III-47}$$

Où e_i sont les résidus du modèle général estimé.

Le Durbin-Watson (DW) varie entre 0 et 4, et un résultat autour de 2 signifie « pas d'autocorrélation » et les résidus sont indépendants.



Les valeurs d_L et d_U sont déterminées à partir de la table de Durbin et Watson en fonction de la taille de l'échantillon et du nombre de variables explicatives pour un risque fixé.

Remarque :

Un DW trouver dans les intervalles $[d_L, d_U]$ ou $[4-d_L, 4-d_U]$ signifie qu'il y a une autocorrélation des résidus.

III.3.1.2.7.4. Vérification de Multicolinéarité (par le coefficient VIF)

La colinéaires définie comme deux variables fortement corrélées entre eux, ainsi qu'une forte corrélation entre plus de deux variables indépendantes s'appelle Multicolinéarité.

La colinéarité peut exister, même lorsque les coefficients de corrélation linéaire entre les variables sont faibles. (Foucart, 2006).

Le coefficient VIF (en anglais : variance inflation factors) est la mesure utilisée pour vérifier la multicollinéarité des facteurs (variables explicatives), un résultat élevé (plus de 5 ou 10) (Chesneau, 2017; Kan Phaobunjong, B.S., 2002), indique un problème de multicollinéarité.

$$VIF = \frac{1}{1-R_j^2} \quad \text{(III-48)}$$

Où R_j^2 est le coefficient de détermination de la Régression linéaire multiple de X_j sur les autres variables.

III.4. Conclusion

Dans ce chapitre, nous avons présenté l'aspect théorique des deux méthodes de modélisation mathématique : la méthode de Réseau de Neurones Artificiels et l'approche paramétrique, les propriétés, les différents types et algorithmes des deux méthodes sont décrits en détail.

Ce chapitre est un élément de base dans notre étude, la compréhension des explications abordées dedans est essentielle pour la compréhension des chapitres prochains.

Deuxième partie

De l'aspect théorique vers la pratique

Chapitre IV

Etude de la productivité dans les projets de construction

IV.1. Introduction

Dans ce chapitre nous allons analyser la productivité dans les projets de construction en Algérie.

Premièrement on va mesurer la productivité dans différents projets et observer les résultats obtenus, puis on va déterminer les facteurs affectant cette productivité par l'interprétation des résultats du sondage effectué.

Les résultats obtenus dans ce chapitre seront utilisés dans les chapitres qui suivent.

IV.2. Mesure de la Productivité dans Les Projets De Construction en Algérie

La productivité de travail dans les projets de construction se définit comme la quantité de travail produite par un nombre des ouvriers dans une période de temps, donc les seuls facteurs qui interviennent dans cette relation sont le nombre des ouvriers et le temps, cette définition semble manquer de précision, à cet effet une investigation dans les chantiers de construction en Algérie a été faite, dans le but de confirmer les doutes quant à cette définition.

Cette investigation consiste à rassembler les données de quatre chantiers de réalisation des bâtiments de différentes régions (Wilaya) de l'Algérie et les analyser par la suite en calculant la productivité du travail de différentes tâches (coffrage, ferrailage), de deux groupes (groupe A et B) de travail dans chaque chantier.

Les tableaux suivants présentent les données collectées des quatre chantiers :

Tableau IV-1 Données de la productivité Chantier 01 W. Skikda

TACHES	Chantier N° : 01 Wilaya de Skikda					
	Groupe A			Groupe B		
	Nbr des ouvriers	Nbr des heures de travail	Quantité travaillée	Nbr des ouvriers	Nbr des heures de travail	Quantité travaillée
Coffrage	14	08	180 M ²	14	08	235.50 M ²
	10	06	161.44 M ²	09	06	161.44 M ²
	14	05	235.50 M ²	15	05	176.63 M ²
	12	04	88.31 M ²	14	04	58.87 M ²
	13	06	235.50 M ²	13	06	206.06 M ²
Ferrailage	03	08	490.95 kg	04	08	1244 kg
	03	08	402.90 kg	03	08	476.76 kg
	11	08	2535 kg	09	08	2450 kg
	02	06	273.73 kg	02	06	330 kg
	05	08	1715 kg	04	07	1707 kg

Tableau IV-2 Données de la productivité Chantier 02 W. Constantine

TACHES	Chantier N° : 02 Wilaya de Constantine					
	Groupe A			Groupe B		
	Nbr des ouvriers	Nbr des heures de travail	Quantité travaillée	Nbr des ouvriers	Nbr des heures de travail	Quantité travaillée
Coffrage	16	08	310 M ²	14	06	295 M ²
	14	08	280 M ²	14	06	280 M ²
	12	10	300 M ²	12	08	305 M ²
	13	06	240.50 M ²	13	06	210 M ²
	11	05	180 M ²	12	05	180 M ²
Ferrailage	11	08	2 205 kg	11	06	2 150 kg
	08	08	1 830 kg	08	08	1 650 kg
	10	08	2 000 kg	11	06	2 040 kg
	12	04	1 800 kg	12	08	2 103 kg
	09	07	1 850 kg	09	07	1 780 kg

Tableau IV-3 Données de la productivité Chantier 03 (N'sigha) W. Khenchela

TACHES	Chantier N° : 03 (N'sigha) Wilaya de Khenchela					
	Groupe A			Groupe B		
	Nbr des ouvriers	Nbr des heures de travail	Quantité travaillée	Nbr des ouvriers	Nbr des heures de travail	Quantité travaillée
Coffrage	11	05	140.06 M ²	10	05	140.06 M ²
	14	07	132.08 M ²	13	07	132.08 M ²
	10	08	101.60 M ²	10	07	101.60 M ²
	15	04	120 M ²	13	04	132.08 M ²
	08	06	81.28 M ²	08	04	81.28 M ²
Ferrailage	03	08	562 kg	04	08	560 kg
	09	04	1 540 kg	09	08	1 920 kg
	08	08	1 640 kg	08	08	1 500 kg
	10	07	2 430 kg	11	08	2 300 kg
	10	06	2 350 kg	10	08	2 150 kg

Tableau IV-4 Données de la productivité Chantier 04 (Baghai) W. Khenchela

TACHES	Chantier N° : 04 (Baghai) Wilaya de Khenchela					
	Groupe A			Groupe B		
	Nbr des ouvriers	Nbr des heures de travail	Quantité travaillée	Nbr des ouvriers	Nbr des heures de travail	Quantité travaillée
Coffrage	02	08	38.60 M ²	02	08	34.80 M ²
	05	08	71.20 M ²	05	08	75 M ²
	04	06	55.18 M ²	04	08	56 M ²
	08	08	100 M ²	06	08	100 M ²
	06	10	110 M ²	06	10	120 M ²
Ferrailage	11	08	2 245 kg	11	06	2 300 kg
	03	08	412 kg	03	05	400 kg
	06	08	946 kg	06	08	1 020 kg
	08	08	1 412 kg	08	07	1 501 kg
	10	08	2 015 kg	09	06	2 130 kg

IV.2.1. Résultats d'investigation.

Les résultats de la productivité des différents chantiers et groupes sont représentés dans les figures suivantes :

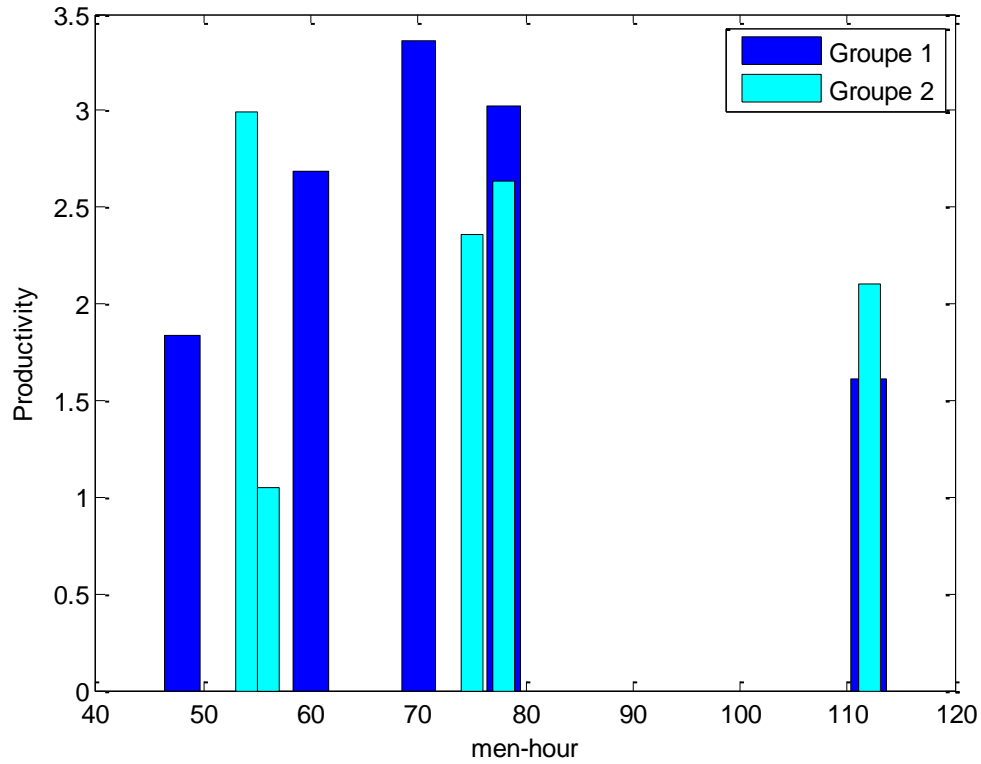


Figure IV.1.a. Chantier de Skikda

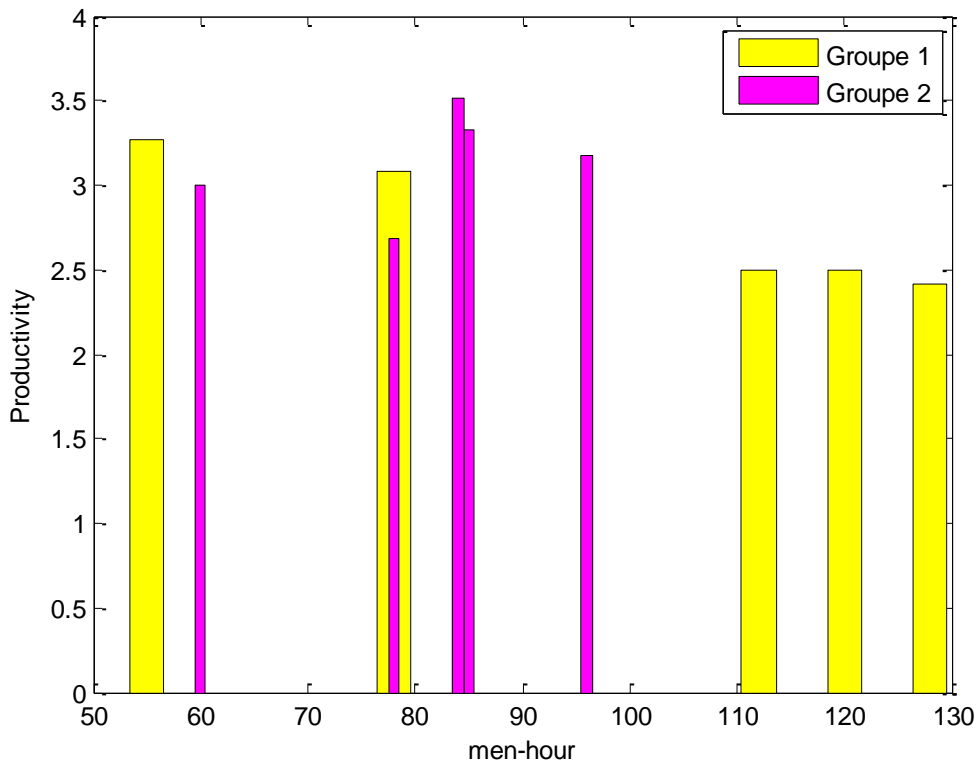


Figure IV.1. b. Chantier de Constantine

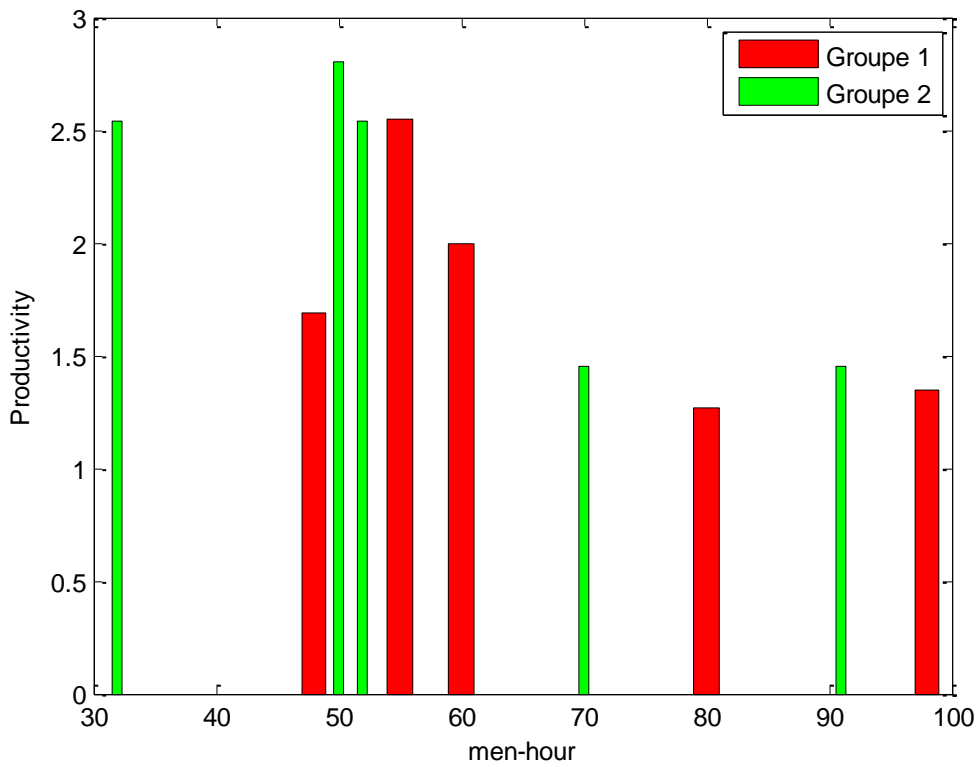


Figure IV.1. c. Chantier de N'sigha

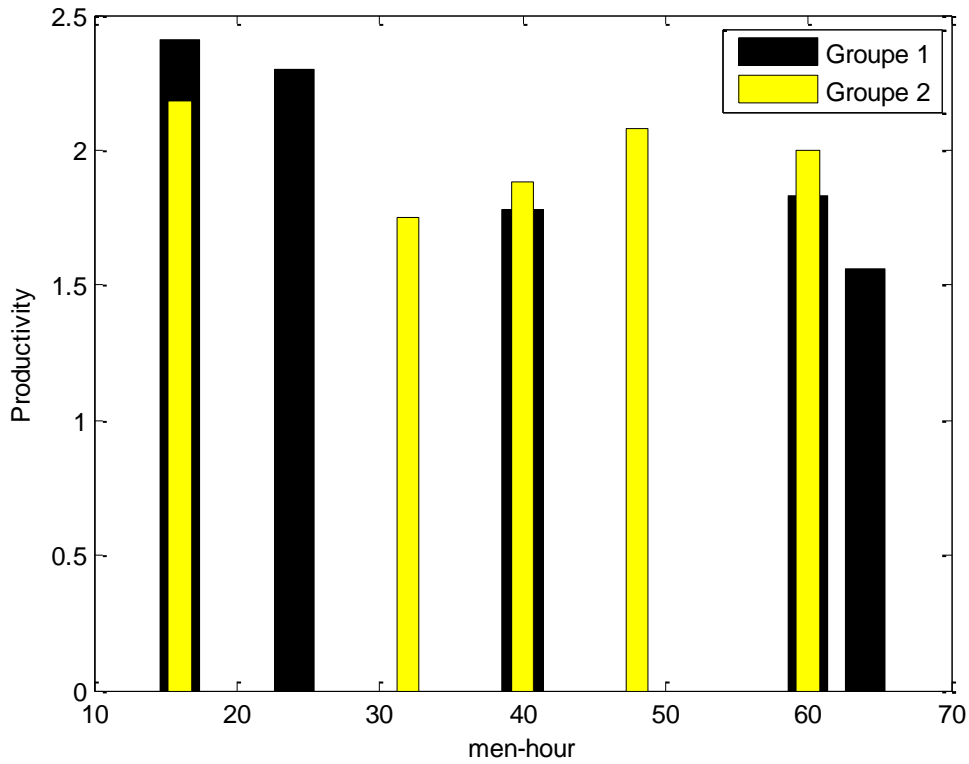


Figure IV.1. d. Chantier de Baghai

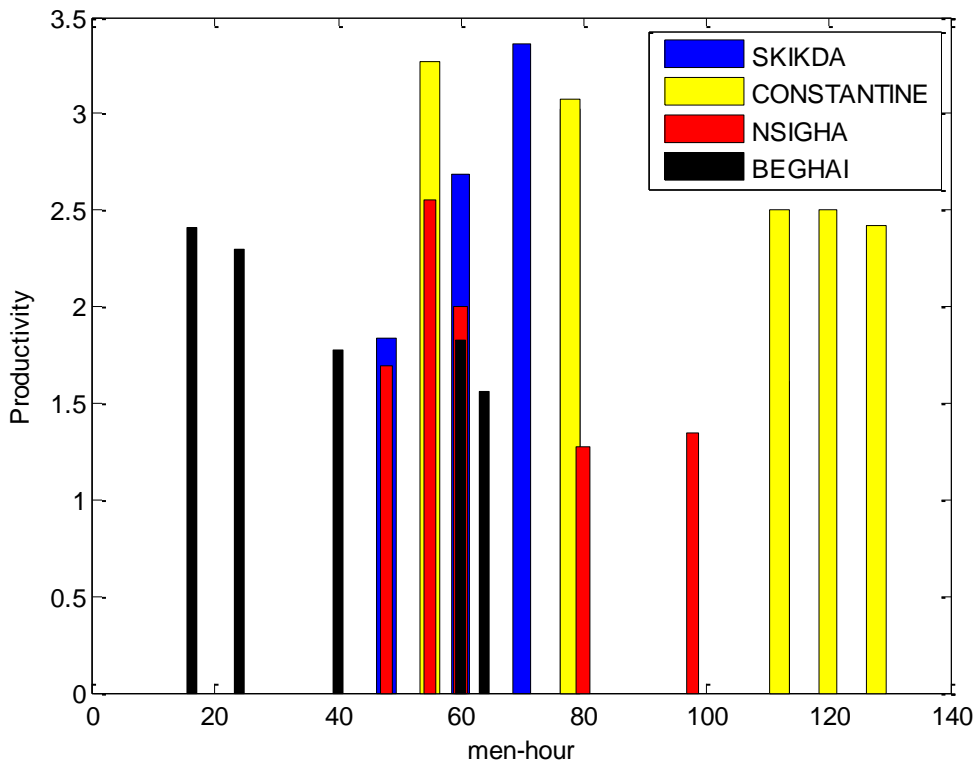
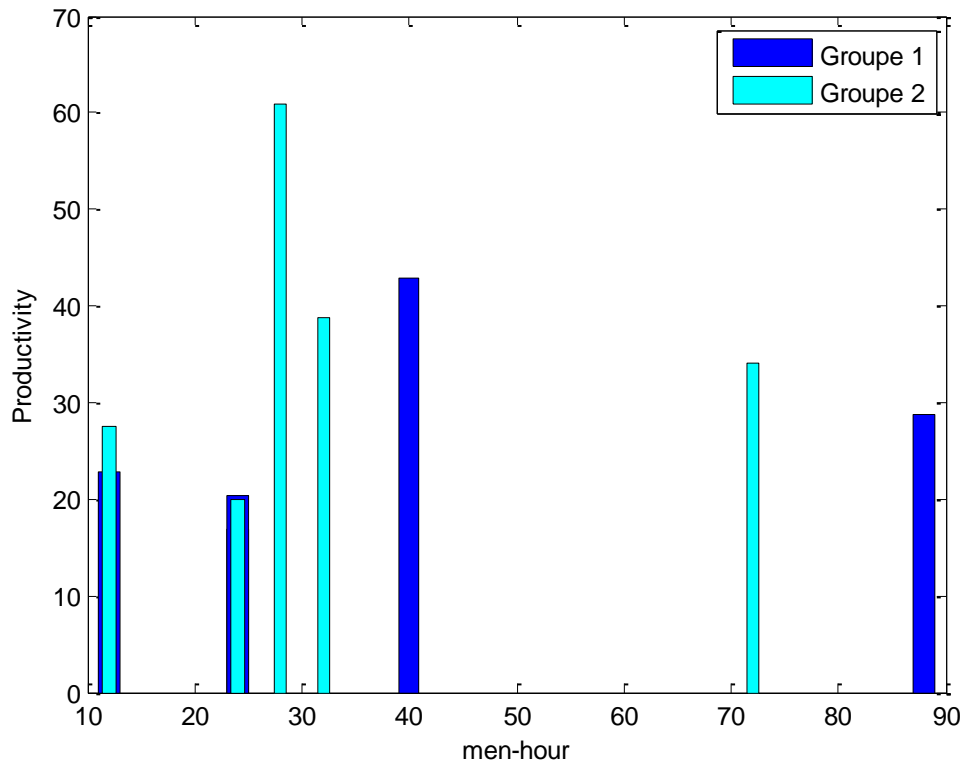


Figure IV.1. e. L'ensemble des chantiers

Figure IV-1 Résultats de la productivité de la tâche du coffrage entre groupes de même chantier et entre les différents chantiers



IV.2. a. Chantier de Skikda

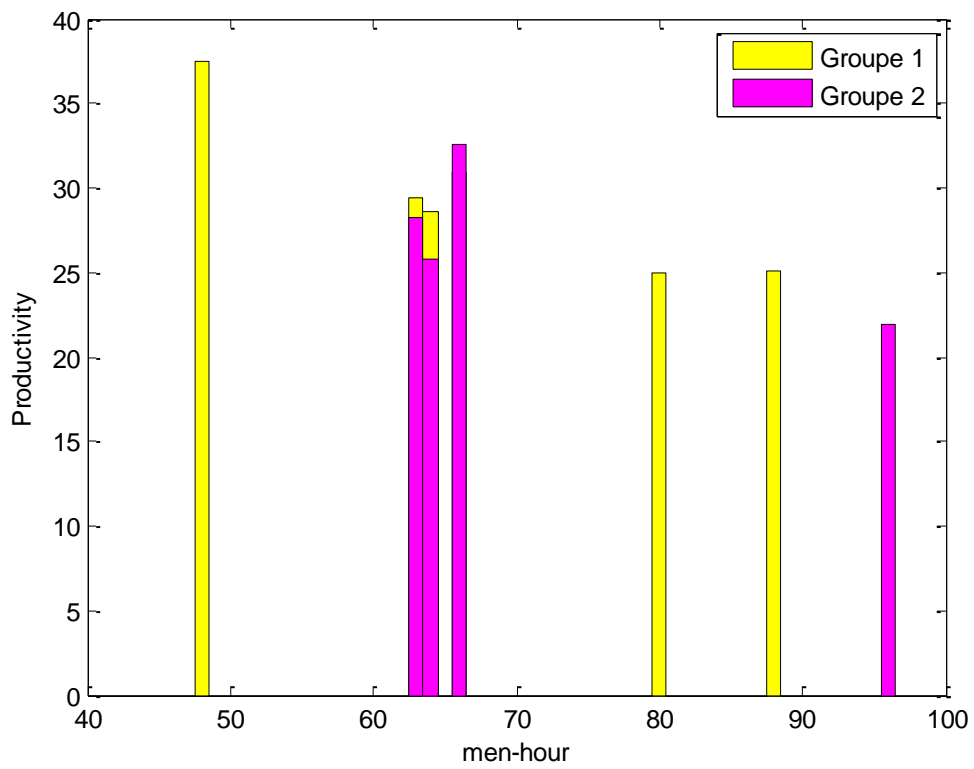


Figure IV.2. b. Chantier de Constantine

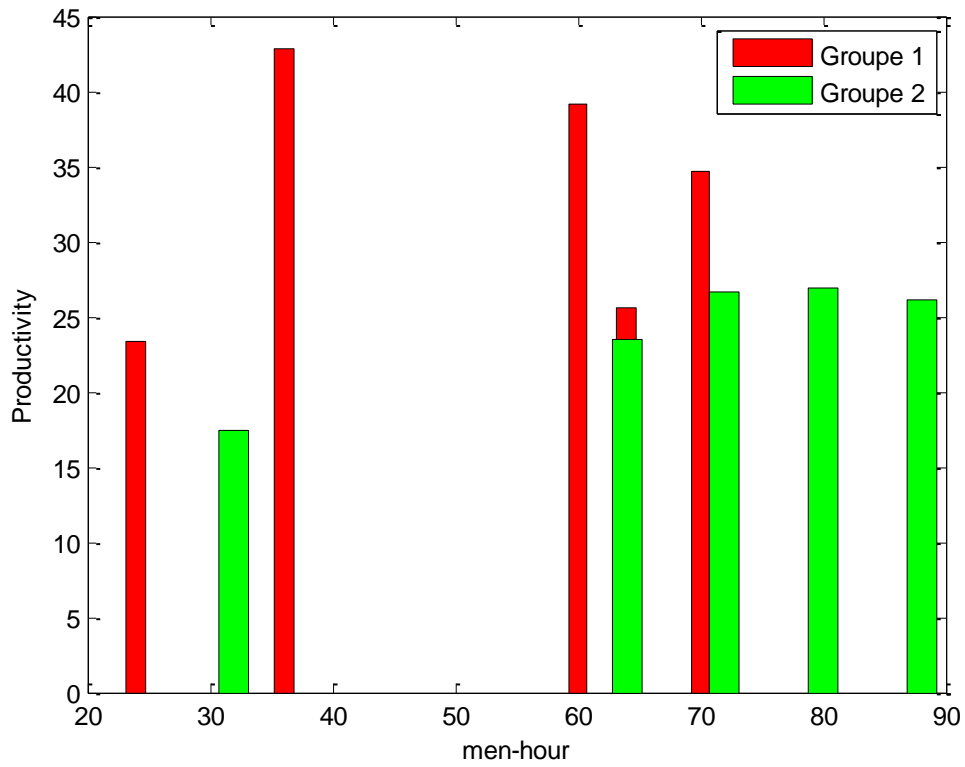


Figure IV.2.c. Chantier de N'sigha

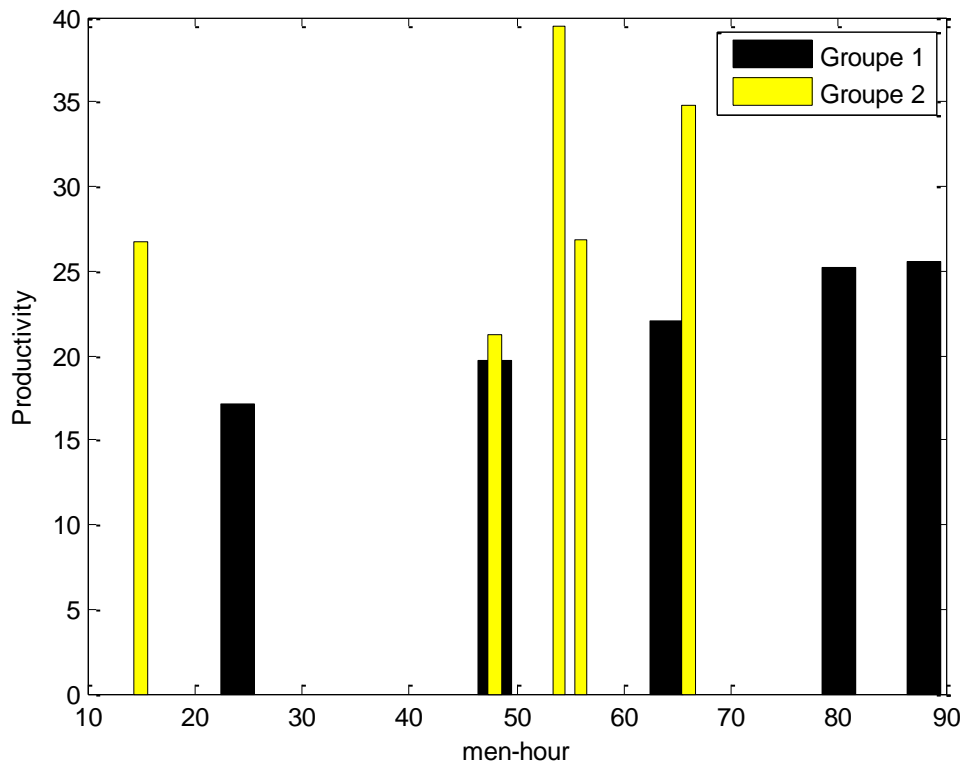
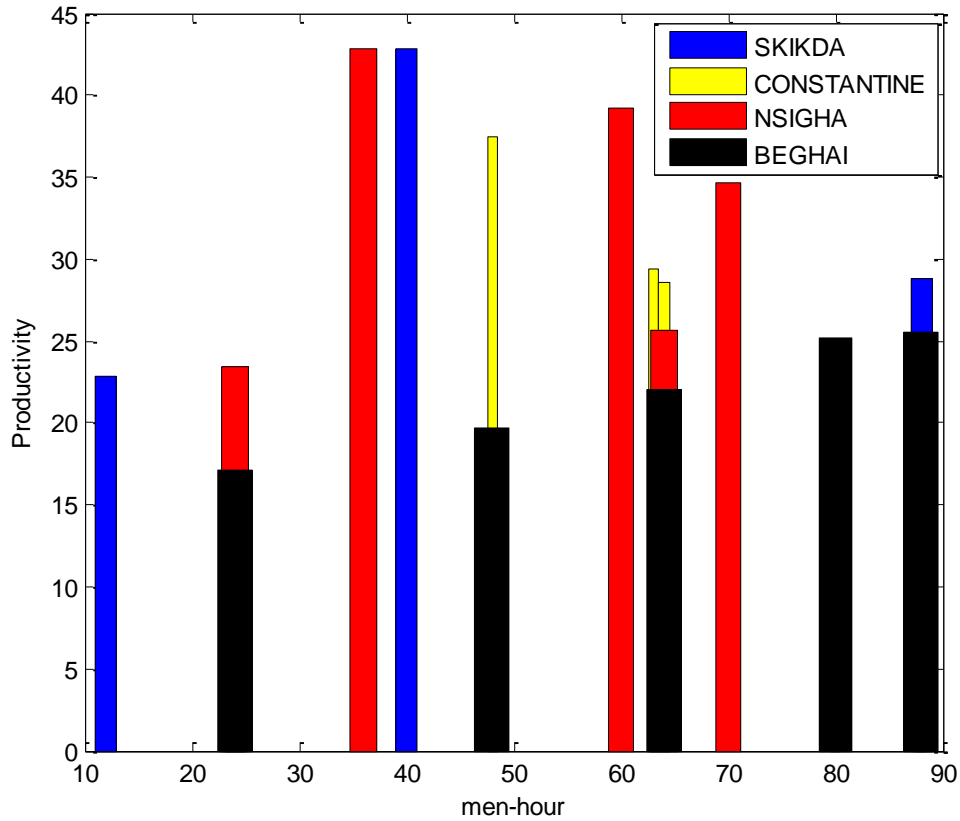


Figure IV.2.d. Chantier de Baghai



IV Figure IV.2.e. L'ensemble des chantiers

Figure IV-2 Résultats de la productivité de la tâche du ferrailage entre groupes de même chantier et entre les différents chantiers

IV.2.2. Analyse et discussion des résultats

D'après les résultats obtenus des différents chantiers et aussi dans les mêmes chantiers on constate :

La productivité de travail d'une tâche, avec une quantité identique et les mêmes nombres d'ouvriers et/ou d'heures de travail, diffère d'un chantier à l'autre.

La productivité de travail d'une tâche dans le même chantier, produite par différents groupes ayant les mêmes entrées/sorties, ne présente pas des résultats identiques.

En conséquence, les résultats obtenus montrent que la notion de la productivité de travail qui dépend seulement du nombre des ouvriers et du nombre d'heures travaillées ne semble pas précise, et des facteurs, autres que ces derniers, peuvent influencer la productivité.

À cet effet, nous optons dans ce qui suit pour l'identification des facteurs qui peuvent affecter la productivité de travail dans les projets de construction Algériens.

IV.3. Identification des facteurs influençant la productivité dans les projets de construction en Algérie

L'identification des facteurs influençant la productivité du travail, depuis des années et jusqu'à maintenant, a connu plusieurs recherches dans différents pays, mais aucun standard, normes ou manuels.....etc. n'est défini des facteurs propres à une tâche, un type de projet, ou encore une région bien déterminée (voir chapitre II). Et les chercheurs ont conclu que la productivité peut se différer d'un chantier à l'autre et d'un pays à l'autre.

À cet égard, nous choisissons dans cette étude de faire un questionnaire destiné aux spécialistes du domaine de la construction en Algérie, englobant les chefs de projets, les responsables techniques, les responsables travaux, les ingénieurs de génie civil, les architectes....etc. afin de tirer leurs impressions quant aux facteurs affectant la productivité du travail dans les projets de construction Algériens.

IV.3.1. Conception du questionnaire

Les facteurs qui influent la productivité du travail dans le domaine de la construction, issus des résultats des différentes recherches et études, précédentes sont regroupés en premier lieu dans le but de tirer les facteurs qui seront utilisés par la suite dans notre questionnaire (Tableau II.1 chapitre II).

En conséquence, 16 facteurs ont été tirés de la recherche bibliographique, groupés en Trois groupes notamment : Groupe 01 : Management, Groupe 02 : Facteurs liés aux Travailleurs, Groupe 03 : Facteurs externes, ces derniers forment le squelette de notre questionnaire.

Le tableau suivant présente les facteurs utilisés dans le questionnaire :

Tableau IV-5 Les facteurs utilisés dans le questionnaire

Questions (Factures)		
Groupe 01 : Management	Groupe 02 : Facteurs liés aux Travailleurs	Groupe 03 : Facteurs externes
Faible taille d'équipe Les faibles salaires Etat des matériels de construction (coffrage, échafaudage...) Le retard de paiement par le maître de l'ouvrage Encombrement dans l'espace de travail Manque ou faible contrôle Faible motivation des ouvriers (primes, jours de récupération...)	Les heures de travail/absentéisme Expérience des ouvriers Quantité de travail réalisée Âge des ouvriers	Conditions météorologiques (Pluie, Soleil) Humidité Température élevée Basse température Pénurie des matériaux de construction

Le questionnaire se constitue de trois parties suivantes :

Première partie : cette partie contient les informations générales et professionnelles des répondants à savoir :

Poste de travail

L'expérience professionnelle dans le domaine

Et les réponses de cette partie sont ouvertes.

Deuxième partie : elle englobe 16 facteurs partagés en Trois groupes avec cinq niveau l'influence de l'échelle de Likert

Niveau 1 : Pas du tout important

Niveau 2 : Peu important

Niveau 3 : Normal

Niveau 4 : Important

Niveau 5 : Très Important

Dans cette partie les réponses sont fermées avec une grille de choix multiples.

Troisième partie : partie ouverte aux constats et suggestions des répondants.

IV.3.2. Type du questionnaire

Le questionnaire a été conçu et rédigé à l'aide du site internet nommé : «SurveyMonkey.com», ce dernier donne la possibilité de créer le sondage et le transmettre via e-mail aux différents répondants, et sert aussi à visualiser quelques statistiques relatives aux résultats obtenus.

Le questionnaire est constitué des réponses définies, c-à-dire que les réponses sont suggérées et non ouvertes, dans la deuxième partie, tandis que la première et la troisième partie, sont constituées des réponses ouvertes qui sont dédiées aux informations professionnelles ainsi qu'aux avis et aux suggestions des répondants.

IV.3.3. Taille de l'échantillon (sample size)

Dans tous les sondages (enquêtes), la taille de l'échantillon est un facteur essentiel dans sa réussite. Il est utilisé pour déterminer l'exactitude et la fiabilité des résultats.

Dans notre recherche, les ingénieurs en génie civil et les architectes, qui travaillent dans le secteur algérien de la construction, constituaient notre population cible ; un échantillon statiquement représentatif de la population a été obtenu en utilisant une formule développée par (Alaghbari et al., 2017; Hogg & Tanis, 2009) telle qu'exprimée ci-dessous

$$n = \frac{m}{1 + \frac{(m-1)}{N}} \quad (\text{IV-1})$$

Où, n est la taille de l'échantillon de la population limitée, m est la taille de l'échantillon de la population illimitée et N est la taille de l'échantillon de la population disponible.

m est estimé selon l'équation suivante :

$$m = \frac{z^2 * p(1-p)}{\epsilon^2} \quad (\text{IV-2})$$

Dans cette équation, z représente la valeur statistique du niveau de confiance utilisé (1,96 et 1,645) pour les niveaux de confiance de 95 % et 90 %, respectivement, ϵ est l'erreur d'échantillonnage de l'estimation ponctuelle, et p est la valeur de la proportion de la population qui est étant estimée.

Dans l'équation. (IV-2) la seule inconnue est la valeur de « p », pour laquelle (Levine et al., 2002) ont proposé d'utiliser une valeur conservatrice de 0,50 pour obtenir une taille d'échantillon au moins aussi grande que nécessaire.

La valeur de m peut être obtenue qu'en utilisant un niveau de confiance de 90 %, de la taille d'échantillon illimitée de la population comme suit :

$$m = \frac{(1.645)^2 * 0.50(1-0.50)}{0.1^2} = 67.65 \approx 68 \quad (\text{IV-3})$$

Dans cette étude, un nombre total de 150 échantillons disponibles ont été sélectionnés parmi des ingénieurs civils et des architectes expérimentés, sur cette base, la taille de l'échantillon représentatif requis de la population est obtenue comme suit :

$$n = \frac{68}{1 + \frac{(68-1)}{150}} = 46.89 \approx 47 \quad (\text{IV-4})$$

Un nombre total de 150 échantillons ont été distribués par e-mail et seuls 65 échantillons de retour ont été reçus, représentant 43.33% du total des échantillons envoyés. Ce nombre de questionnaires remplis atteint et dépasse la taille d'échantillon requise (n=47).

Les répondants dans les échantillons reçus sont des ingénieurs en génie civil et des architectes, travaillant dans l'industrie de la construction en Algérie, avec 46.15 % d'entre eux ayant une expérience de 01 à 05 ans, tandis que 44,62 % des répondants ont une expérience de 05 à 10 ans, et 6,15 % ont une expérience variée entre 10 et 15 ans ainsi que les 3.08% ont 15 ou plus d'expérience.

IV.3.4. Les niveaux d'influence (Echelle de Likert)

La deuxième partie du questionnaire englobe 16 facteurs partagés en trois groupes. Ces facteurs sont évalués selon leurs importances (effet négatif sur la productivité) par les répondants, pour cela, la méthode d'évaluation de l'échelle de Likert est adoptée dans cette étude, cette méthode consiste de diviser le niveau d'influence de chaque question (réponse) en cinq niveaux et chaque niveau correspond à un poids (Tableau IV-6).

Tableau IV-6 Niveau d'influence et poids de l'échelle de Likert

Niveau d'influence	Pas du tout important	Peu important	Normal	Important	Très Important
Poids	1	2	3	4	5

La détermination des zones d'influence se fait par le calcul de la médiane, sachant que l'échelle de Likert est constituée de 05 points, et la différence entre 01 et 05 est 04, alors les pas des intervalles sont définis en divisant 04 par 05 qui donnent 0.8.

Les résultats sont montrés dans le tableau suivant :

Tableau IV-7 Les zones des réponses selon l'échelle Likert

Niveau d'influence	Pas du tout important	Peu important	Normal	Important	Très Important
Intervalle	[1 ; 1.80]	[1.81 ; 2.60]	[2.61 ; 3.40]	[3.41 ; 4.20]	[4.21 ; 5]

IV.3.5. Validation du questionnaire

Avant de transmettre le questionnaire aux répondants il est nécessaire de tester la fiabilité de ce dernier, (les réponses aux questions portant sur le même sujet devront être corrélées) (A. Otmani, 2020; Shehu & Akintoye, 2010) Alpha Cronbach, est le test utilisé dans cette étude pour mesurer la cohérence interne des questions, un résultat plus ou égal à 0.7 indique la fiabilité de notre test.

Un nombre de 10 questionnaires sont partagés aux répondants en premier lieu, afin d'analyser sa fiabilité, et les résultats obtenus sont utilisés pour calculer l'Alpha Cronbach donné par l'équation suivante :

$$\alpha = \frac{k}{k-1} \left(1 - \frac{\sum_{i=1}^k \sigma_{yi}^2}{\sigma_x^2} \right) \quad (\text{IV-5})$$

Où k est le nombre d'items, σ_{yi}^2 représente est la variance de l'item i et σ_x^2 est la variance du score total.

Logiciel SPSS est utilisé dans cette recherche pour vérifier la fiabilité du questionnaire par le calcul d'Alpha Cronbach, et les résultats sont présentés aux tableaux suivants :

Tableau IV-8 Résultat d'Alpha Cronbach pour l'ensemble des facteurs

Cronbach's Alpha	N of Items
0,833	16

Tableau IV-9 Résultat d'Alpha Cronbach pour chaque facteur

	Scale Mean if Item Deleted	Scale Variance if Item Deleted	Corrected Item-Total Correlation	Cronbach's Alpha if Item Deleted
Faible taille d'équipe	59,0000	43,556	0,184	0,839
Encombrement dans l'espace de travail	59,6000	49,156	-0,262	0,880
Manque ou faible contrôle	58,7000	41,567	0,463	0,823
Les faibles salaires	59,0000	37,556	0,793	0,801
Le retard de paiement par le maitre de l'ouvrage	59,3000	38,900	0,655	0,810
Faible motivation des ouvriers (primes, jours de récupération...)	59,4000	39,600	0,656	0,812
Etat des matériels de construction (coffrage, échafaudage...)	58,6000	43,156	0,448	0,825
Age des ouvriers	60,0000	41,556	0,485	0,822
Expérience des ouvriers	59,3000	45,122	0,140	0,836
Les heures de travail/absentéisme	59,3000	36,678	0,899	0,795
Quantité de travail réalisée	59,5000	42,056	0,348	0,829
Conditions météorologiques (Pluie, Soleil)	59,9000	38,322	0,508	0,820
Une forte ou basse humidité	59,7000	41,344	0,385	0,827
Température élevée	59,6000	40,933	0,561	0,818
Basse température	59,8000	39,067	0,607	0,813
Pénurie des matériaux de construction	58,8000	38,844	0,782	0,805

D'après les résultats obtenus, présentés aux tableaux ci-dessus (Tableau IV-8 et IV-9) on peut constater que la valeur globale d'Alpha Cronbach de l'ensemble des 16 facteurs est de 0.833, cette dernière est supérieure à 0.70.

On peut constater aussi que toutes les valeurs d'Alpha Cronbach de chaque facteur sont bornées entre 0.795 et 0.880 et qui sont supérieures à 0.7, ce qui signifie que notre questionnaire est fiable.

IV.3.6. Statistiques relatives aux répondants

Dans cette partie on va présenter les résultats obtenus, par 65 participants de notre sondage, travaillant dans le domaine de la construction en Algérie, et qui occupent différents postes de travail.

Tableau IV-10 Statistiques relatifs aux postes de travaux des participants

Poste de travail	Nombre	Pourcentage
CHEF PROJET	9	13,85%
ING GC	28	43,08%
ARCHITECTE	22	33,85%
Planificateur	6	9,23%
TOTAL	65	100%

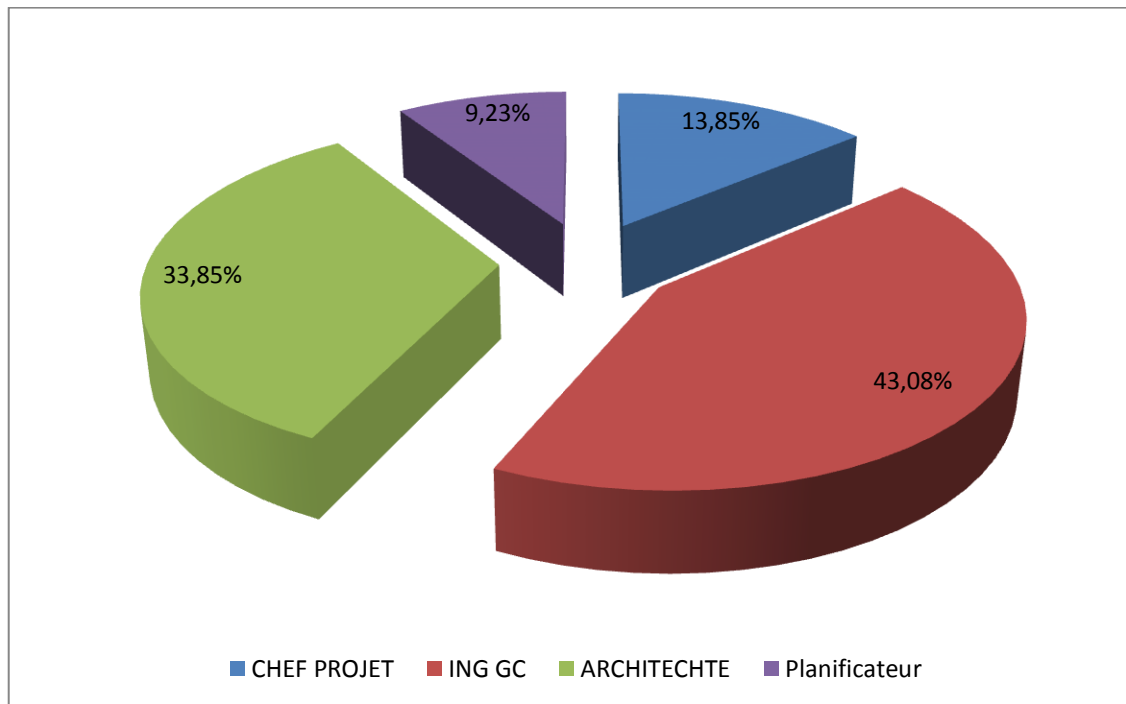


Figure IV-3 Représentation graphique des statistiques des postes de travail des participants

D'après le tableau IV-10 et la figure IV-3 on constate que les participants au sondage se constituent de : Chefs de projet, Ingénieurs en Génie Civil, Architectes et Planificateurs des Projets, ces derniers forment un nombre total de 65 participants.

La plus grande catégorie des répondants est représentée par les ingénieurs en Génie Civil par un pourcentage atteint 43.08 %, suivis par les Architectes avec 33.85 %, les chefs projets et les planificateurs sont classés 3ème et 4ème avec un pourcentage de 13.85 % et 09.23 % respectivement.

Tableau IV-11 Statistique des participants selon leurs expériences professionnelles

Expérience	Nombre	Pourcentage
[1 5 [30	46,15%
[5 10 [29	44,62%
[10 15 [4	6,15%
[15 20 [1	1,54%
20 et plus	1	1,54%
TOTAL	65	100%

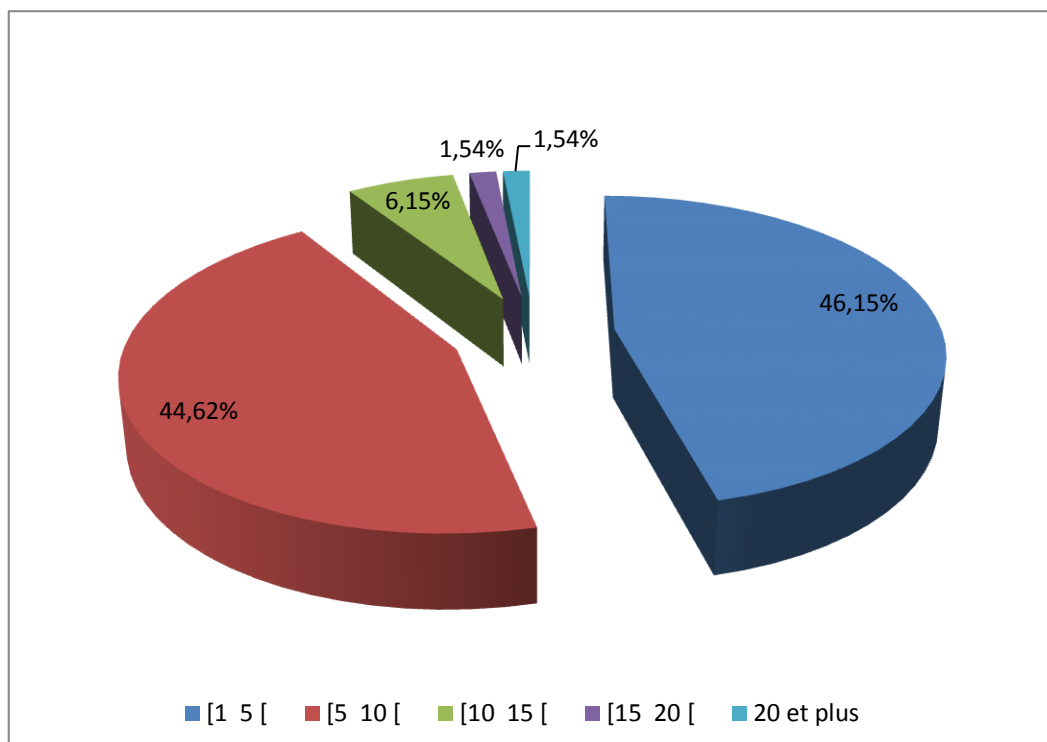


Figure IV-4 Représentation graphique des statistiques de l'expérience professionnelles des participants

IV.3.7. Analyse des résultats du sondage

Les résultats de sondage sont présentés par groupe comme suite :

Tableau IV-12 Résultats des réponses du Groupe 01 : (Facteurs liés à la gestion (management))

–		Pas du tout important	Peu important	Normal	Important	Très important	Total
Faible taille d'équipe	%	6,15%	3,08%	23,08%	32,31%	35,38%	6
	Nombre	4	2	15	21	23	5
Les faibles salaires	%	4,62%	6,15%	7,69%	23,08%	58,46%	6
	Nombre	3	4	5	15	38	5
Etat des matériels de construction (coffrage, échafaudage...)	%	3,08%	9,23%	13,85%	27,69%	46,15%	6
	Nombre	2	6	9	18	30	5
Encombrement dans l'espace de travail	%	16,92%	12,31%	18,46%	33,85%	18,46%	6
	Nombre	11	8	12	22	12	5
Manque ou faible contrôle	%	16,92%	7,69%	15,38%	38,46%	21,54%	6
	Nombre	11	5	10	25	14	5
Le retard de paiement par le maître de l'ouvrage	%	15,38%	10,77%	20,00%	27,69%	26,15%	6
	Nombre	10	7	13	18	17	5
Faible motivation des ouvriers (primes, jours de récupération...)	%	4,62%	4,62%	10,77%	36,92%	43,08%	6
	Nombre	3	3	7	24	28	5

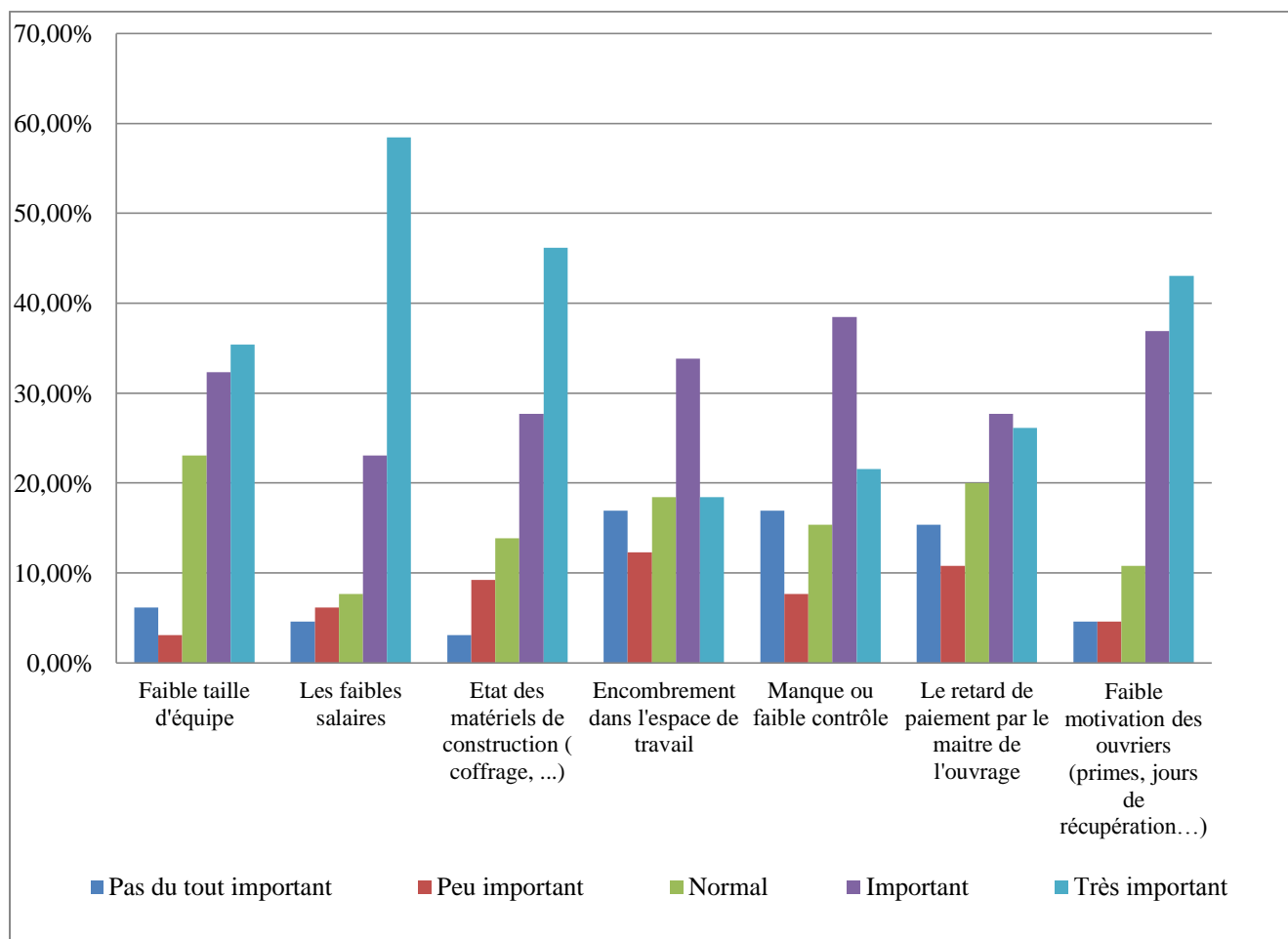


Figure IV-5 Histogramme des réponses Groupe 01 : Facteurs liés à la gestion (management)

Tableau IV-13 Résultats des réponses du Groupe 02 : (Facteurs liés aux ouvriers)

-		Pas du tout important	Peu important	Normal	Important	Très important	Total
Les heures de travail/absentéisme	%	1,54%	6,15%	21,54%	35,38%	35,38%	65
	Nombre	1	4	14	23	23	
Expérience des ouvriers	%	4,62%	6,15%	13,85%	44,62%	30,77%	65
	Nombre	3	4	9	29	20	
Quantité de travail réalisée	%	1,54%	4,62%	24,62%	40,00%	29,23%	65
	Nombre	1	3	16	26	19	
Âge des ouvriers	%	9,23%	15,38%	33,85%	30,77%	10,77%	65
	Nombre	6	10	22	20	7	

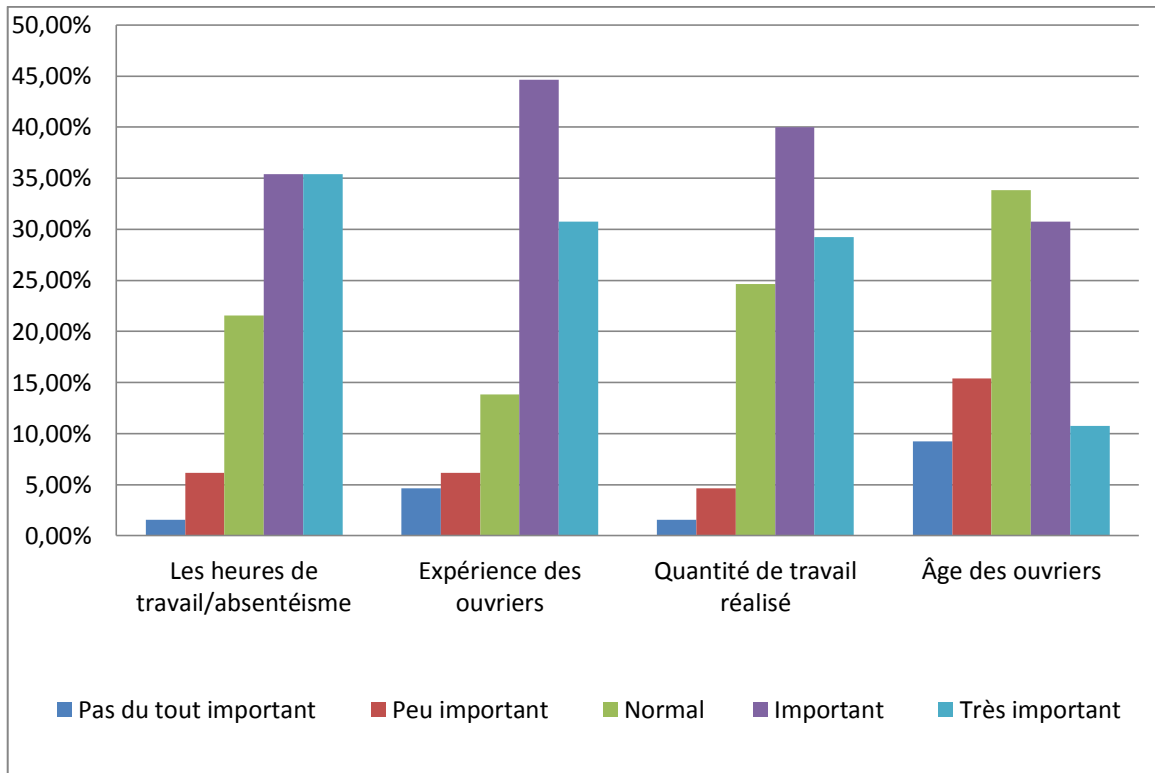


Figure IV-6 Histogramme des réponses Groupe 02 : Facteurs liés aux ouvriers

Tableau IV-14 Résultats des réponses du Groupe 03 : (Facteurs Externe)

-		Pas du tout important	Peu important	Normal	Important	Très important	Total
Conditions météorologiques (Pluie, Soleil)	%	12,31%	13,85%	32,31 %	26,15 %	15,38%	65
	Nombre	8	9	21	17	10	
Une forte ou basse humidité	%	13,85%	16,92%	26,15 %	29,23 %	13,85%	65
	Nombre	9	11	17	19	9	
Température élevée	%	7,69%	10,77%	20,00 %	33,85 %	27,69%	65
	Nombre	5	7	13	22	18	
Basse température	%	13,85%	12,31%	33,85 %	23,08 %	16,92%	65
	Nombre	9	8	22	15	11	
Pénurie des matériaux de construction	%	16,92%	13,85%	12,31 %	29,23 %	27,69%	65
	Nombre	11	9	8	19	18	

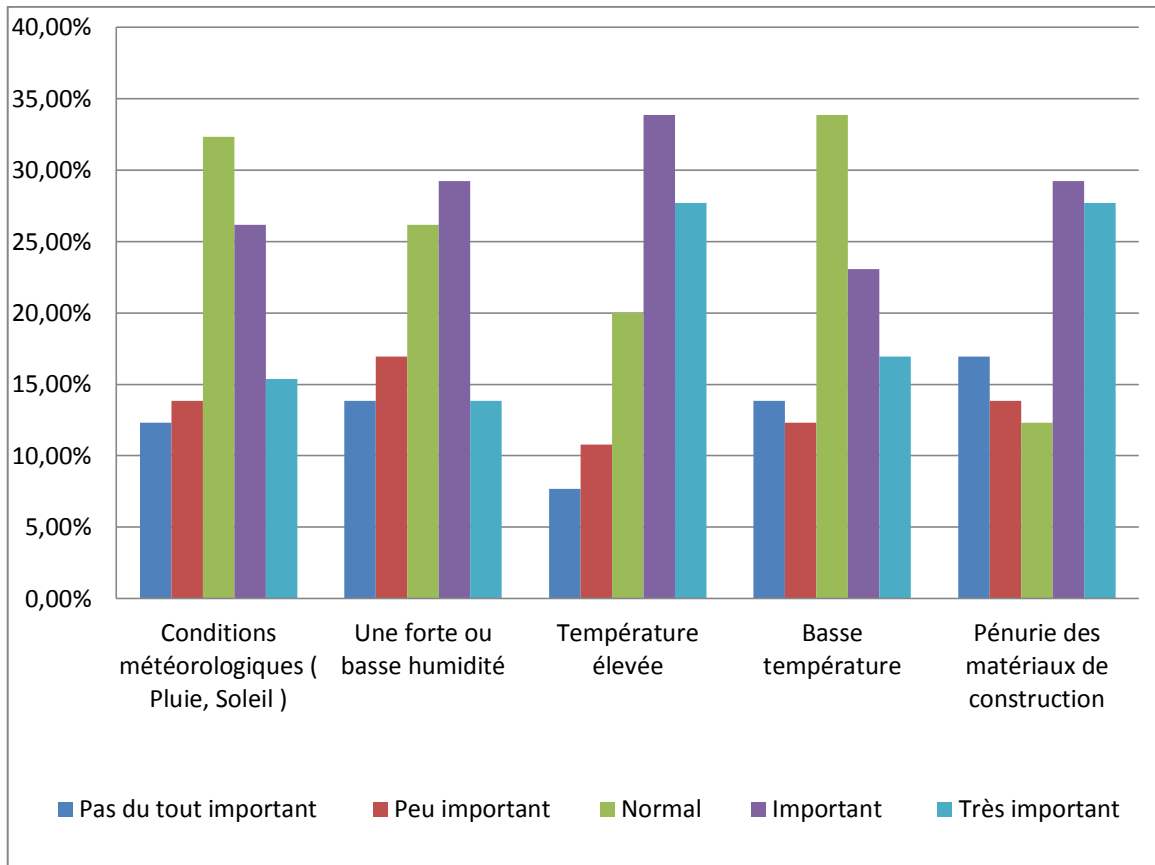


Figure IV-7 Histogramme des réponses Groupe 03 : Facteurs externes

IV.3.8. Détermination des facteurs influençant la productivité

IV.3.8.1. Relatif importance index (RII%)

Dans le but de déterminer les facteurs qui influencent le plus la productivité du travail dans les projets de construction algériens, selon les réponses des participants à notre sondage, nous avons opté pour le classement de ces facteurs selon leur indice d'importance relative RII% (en anglais : Relative Importance Index), vu que cette technique a prouvé une forte efficacité de classement, surtout dans les sondages reposant sur l'échelle de Likert (voir Chapitre II).

Le RII % est obtenu selon l'équation suivante :

$$RII (\%) = \left(\frac{\sum_{i=1}^5 n_i * x_i}{5 \sum_{i=1}^5 x_i} \right) * 100 \quad (IV-6)$$

Avec n_i est le niveau d'influence suivant l'échelle de Likert (de 1 à 5), et x_i correspond à la fréquence de chaque niveau d'influence.

IV.3.8.1.1. Classement des facteurs

D'après le calcul de l'indice d'importance relative (RII %) de chaque facteur, le classement de ces derniers est donné au tableau suivant :

Tableau IV-15 Résultats de RII% et classement des facteurs

Facteurs	RII (%)	Classement
Faible taille d'équipe	77,54	7
Les faibles salaires	84,92	1
Etat des matériels de construction	80,92	3
Encombrement dans l'espace de travail	64,92	12
Manque ou faible contrôle	68,00	9
Le retard de paiement par le maitre de l'ouvrage	67,69	10
Faible motivation des ouvriers (primes, jours de récupération...)	81,85	2
Les heures de travail/absentéisme	79,38	4
Expérience des ouvriers	78,15	5
La Quantité de travail réalisée	78,15	5
Âge des ouvriers	63,69	13
Conditions météorologiques (Pluie, Soleil)	63,69	13
Une forte ou basse humidité	62,46	16
Température élevée	72,62	8
Basse température	63,38	15
Pénurie des matériaux de construction	67,38	11

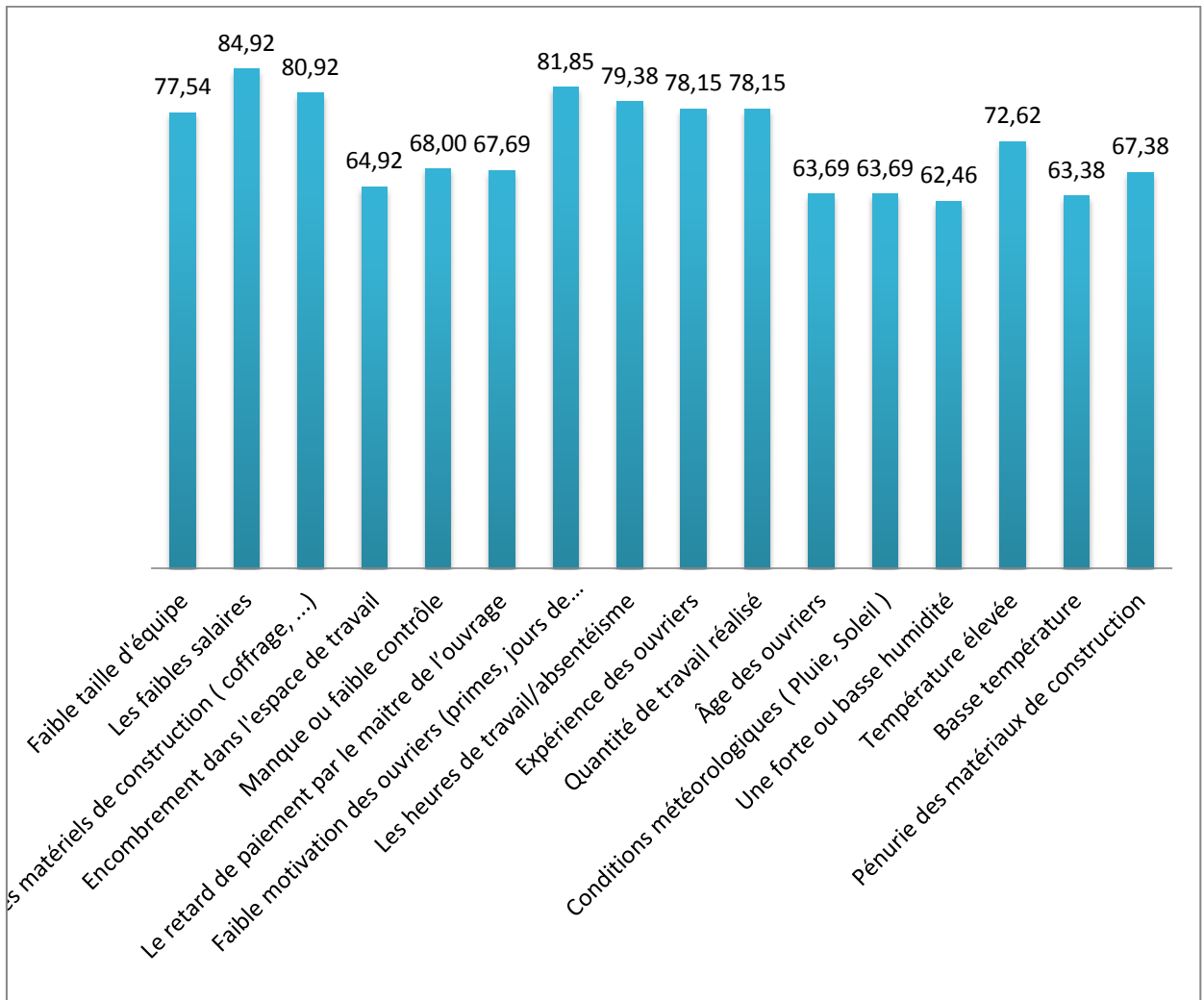


Figure IV-8 Résultats de RII % de chaque facteur

IV.3.8.1.2. Analyse des résultats

Les résultats des calculs du RII (%) montrent que trois facteurs ont un RII (%) supérieur à 80%, à savoir Les faibles salaires, Faible motivation des ouvriers, Etat des matériels de construction.

Tandis que Cinq facteurs ont un RII (%) entre 70 et 80%, à savoir, Les heures de travail/absentéisme, L'expérience des travailleurs, La Quantité de travail réalisée, Faible taille de l'équipe, Température élevée.

Les Huit derniers facteurs qui sont : Manque ou faible contrôle, Le retard de paiement par le maitre de l'ouvrage, Pénurie de matériaux de construction, Encombrement dans l'espace de travail (zone de travail), Age des ouvriers, Conditions météorologiques, Basse température ainsi que Forte ou basse humidité, ont un RII % compris entre 60% et 70%.

IV.3.8.2. Etude de significativité des facteurs

L'étude de significativité des facteurs, consiste à évaluer les résultats de la statistique de : one simple de Student (T de Student), pour un niveau de confiance (un niveau de signification) de 0.05 (05 %).

Un facteur est significatif si son T de Student est atteint ou dépasse la valeur de 1,96 (le nombre d'échantillons > 30 « chapitre III.3.1.2.4.4.1.»)

Les résultats de T de Student sont présentés dans le tableau suivant :

Tableau IV-16 Résultats de test T de Student

Facteurs	T	df	Sig. (2-tailed)	Mean Difference
Faible taille d'équipe	27,781	64	0,000	3,87692
Les faibles salaires	30,232	64	0,000	4,24615
Etat des matériels de construction (coffrage, échafaudage...)	29,021	64	0,000	4,04615
Encombrement dans l'espace de travail	19,269	64	0,000	3,24615
Manque ou faible contrôle	20,052	64	0,000	3,40000
Le retard de paiement par le maitre de l'ouvrage	19,653	64	0,000	3,38462
Faible motivation des ouvriers (primes, jours de récupération...)	30,799	64	0,000	4,09231
Les heures de travail/Absentéisme	32,529	64	0,000	3,96923
Expérience des ouvriers	29,818	64	0,000	3,90769
Quantité de travail réalisé	33,848	64	0,000	3,90769
Âge des ouvriers	22,311	64	0,000	3,24615
Conditions météorologiques (Pluie, Soleil)	20,987	64	0,000	3,18462
Une forte ou basse humidité	20,042	64	0,000	3,12308
Température élevée	24,009	64	0,000	3,63077
Basse température	20,326	64	0,000	3,16923
Pénurie des matériaux de construction	18,693	64	0,000	3,36923

Dans de tableau ci-dessus on remarque que tous les facteurs ayant un T de Student supérieur à 1.96 donc tous les facteurs sont significatifs.

IV.3.9. Détermination des facteurs affectant négativement la productivité de travail dans les projets de construction Algériens

Afin de déterminer les facteurs affectant négativement la productivité de travail dans les projets de construction Algériens, nous ne prenons en considération que les facteurs ayant une moyenne des réponses tombent dans un intervalle de Grande et très Grande influence selon les intervalles de Likert (tableau IV.7).

La moyenne des réponses se calcule par la formule suivante :

$$\text{La moyenne des réponses} = \frac{5(n5)+4(n4)+3(n3)+2(n2)+1(n1)}{N} \quad (\text{IV-7})$$

Avec n5, n4, n3, n2, n1 est le nombre des réponses par 5, 4, 3, 2, 1

N : le nombre total des répondants

Les résultats des moyennes des réponses sont présentés dans le tableau suivant :

Tableau IV-17 Facteurs avec une moyenne de réponse entre l'intervalle de Grande et Très Grande influence

Facteurs	Moyenne
Faible taille d'équipe	3,877
Les faibles salaires	4,246
Etat des matériels de construction (coffrage, échafaudage...)	4,046
Encombrement dans l'espace de travail	3,246
Manque ou faible contrôle	3,400
Le retard de paiement par le maître de l'ouvrage	3,385
Faible motivation des ouvriers (primes, jours de récupération...)	4,092
Les heures de travail/Absentéisme	3,969
Expérience des ouvriers	3,908
Quantité de travail réalisé	3,908
Âge des ouvriers	3,185
Conditions météorologiques (Pluie, Soleil)	3,185
Une forte ou basse humidité	3,123
Température élevée	3,631
Basse température	3,169
Pénurie des matériaux de construction	3,369

D'après le tableau ci-dessus on constate 08 facteurs ont une moyenne de réponse entrant dans un intervalle de 3.41 jusqu'à 5.

En conclusion, les facteurs influençant la productivité, considérés dans notre recherche sont :

- ✓ La Taille d'équipe
- ✓ Les salaires des ouvriers
- ✓ Etat des matériels de construction (coffrage, échafaudage...)
- ✓ Type de motivation des ouvriers
- ✓ Nombre des heures de travail
- ✓ Expérience des ouvriers
- ✓ Température
- ✓ Quantité de travail réalisée

Le tableau suivant résume les facteurs objets de notre recherche :

Tableau IV-18 Liste des facteurs sélectionnés

Facteurs	Description	Catégorie
La Taille d'équipe	Nombre des ouvriers qualifiés	Quantitatif
	Nombre des ouvriers non qualifiés (manœuvres)	Quantitatif
Les salaires des ouvriers	Les salaires des ouvriers qualifiés	Quantitatif
	Les salaires des ouvriers non qualifiés (manœuvres)	Quantitatif
Etat des matériels de construction (coffrage, échafaudage...)	Médiocre	Qualitatif
	Moyen	Qualitatif
	Bien	Qualitatif
Type de motivation des ouvriers	Travail à la tâche	Qualitatif
	Jours de récupération	Qualitatif
	Primes	Qualitatif
Nombre des heures de travail		Quantitatif
Expérience des ouvriers	Expérience minimale	Quantitatif
	Expérience maximale	Quantitatif
Température		Quantitatif
Quantité de travail réalisée		Quantitatif

IV.3.10. Collection des données

Après l'identification des facteurs influençant la productivité du travail, ces derniers sont observés et collectés auprès 04 chantiers différents : Skikda, Ali Mendjli W. Constantine, N'sigha W. Khenchela, et Baghai W. Khenchela.

A la fin de cette étape, une base de données comportant 101 données a été obtenue, dans le but de l'utiliser pour développer les modèles objets de notre étude.

IV.4. Conclusion

Ce chapitre présente en première phase une analyse de la productivité dans les projets de construction Algérien, ou on a traité les données depuis 04 chantiers de différentes régions du pays, puis on a conduit un sondage constitué de 16 facteurs pour les classer suivant leurs influences sur la productivité et déterminer les plus significatifs parmi eux, en utilisant les tests statistiques.

Les résultats de ce chapitre nous permettent de déterminer les facteurs nécessaires, et collecter par la suite notre base de données afin d'estimer la productivité dans les chapitres qui suivent.

Chapitre V

Modélisation de la productivité par les Réseaux de Neurones Artificiels

V.1.Introduction.

Ce chapitre consiste à la modélisation de la productivité du travail dans les projets Algériens, par la méthode du Réseau de Neurones Artificiels, afin de prédire cette productivité sous les facteurs qui l'affectent (tableau IV-18).

La structure, l'architecture, le type, et les différentes propriétés du RNA utilisé sont expliquées dans ce chapitre.

Les résultats sont obtenus en se basant sur deux phases : la première consiste à entraîner le modèle, tandis que la deuxième a pour but de tester les résultats obtenus de la première phase.

L'analyse et la discussion des résultats obtenus sont présentées à la fin de ce chapitre.

V.2.Modélisation par Réseau de Neurones Artificiels

La modélisation d'un réseau de neurones artificiel (RNA), consiste à trouver la meilleure structure et l'architecture ainsi que l'algorithme et les différentes propriétés, qui donnent des meilleurs paramètres capables de produire des résultats fiables.

Les étapes de développement du modèle de réseau de neurones artificiel sont présentées comme suite :

- Détermination des données d'entrées et sorties du modèle à développer.
- Collection des données d'entrées et sorties du modèle à développer.
- Préparation des données avant de les introduire au modèle.
- Détermination du type et algorithme d'apprentissage.
- Détermination de la structure et de l'architecture (nombre des couches et neurones cachés, choix de la fonction de transfert,...) du modèle que l'on cherche à développer.
- Réaliser l'apprentissage (l'entraînement) du modèle.
- Tester les résultats obtenus dans la phase d'apprentissage.
- Evaluation des résultats et des performances du modèle développé.

V.2.1. Détermination des données

V.2.1.1. Données d'entrée

Les données d'entrée sont des informations pertinentes, elles doivent être sélectionnées et choisies vis-à-vis de leurs influences sur les valeurs de la sortie du problème modélisé (A. Otmani, 2020), et les données qui n'ont pas une influence sur la sortie sont donc éliminées.

Dans ce contexte, les données d'entrées choisies, sont celles qui résultent de l'étude statistique du questionnaire traité dans le chapitre (IV.3) (voir Tableau IV.18).

V.2.1.2. Données de sortie

La sortie d'un modèle est le résultat obtenu par ce dernier, elle peut être connue a priori dans le cas d'un apprentissage supervisé et c'est le cas de notre étude, comme elle peut être inconnue dans le cas d'un apprentissage non supervisé.

Comme notre étude a pour objet d'estimer la productivité du travail dans les projets de construction Algériens, les données de sortie sont donc les résultats de la productivité obtenus par la formule suivante :

$$Productivité = \frac{\text{la quantité du travail effectuée}}{\text{nombre des ouvriers} * \text{nombre des heures travaillées}} \quad (V-1)$$

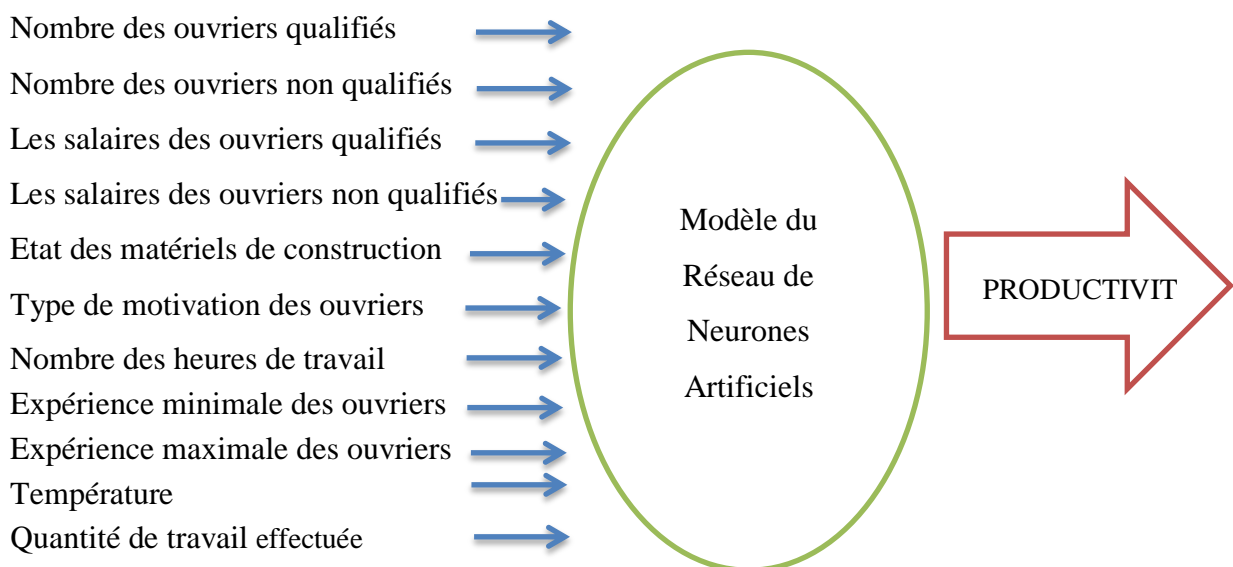


Figure V-1 Schéma simplifié représente les données d'entrées et les données de sortie dans le modèle RNA

V.2.1.3. Collection des données

Une base de données comportant 101 données a été obtenue, après observation et collection des facteurs influençant la productivité du travail (Tableau IV.18), auprès 04 chantiers situés à : Skikda ville, Ali Mendjli W. Constantine, N'sigha W. Khenchela, et Baghai W. Khenchela.

Il est à noter que les données collectées sont relatives à la tâche du coffrage des dalles pleines dans les projets suscités.

V.2.2. Préparation des données

V.2.2.1. Prétraitement des données

Dans cette étape, les données collectées ont fait l'objet d'un prétraitement avant de les introduire au modèle du Réseau de Neurones Artificiels.

V.2.2.1.1. Prétraitement des données (quantitatives) numériques

Parmi les méthodes de prétraitement présentées dans le chapitre III section 2.7.1.1, nous avons choisi la méthode de normalisation des données, afin de généraliser mieux le modèle, et notre base de données sera donc, dans un intervalle [0, 1] selon la formule suivante :

$$X_n = \frac{X-MIN}{MAX-MIN} \quad (V-2)$$

Ou X_n est la valeur normalisée et X est la valeur réelle

Min et Max sont respectivement les valeurs minimum et maximum de la valeur réelle de X .

V.2.2.1.2. Prétraitement des données qualitatives

Elle consiste en la conversion des données qualitatives en données numériques, avant de les introduire dans le modèle.

Dans notre étude, deux facteurs de type qualitatif sont présentés, ils sont convertis en forme numérique comme suite :

Tableau V-1 Transformation des données qualitatives en données Numérique

Facteur Qualitatif		Transformation Numérique
Etat des matériels de construction (coffrage, échafaudage...)	Médiocre	-0.50
	Moyen	0.00
	Bien	0.50
Type de motivation des ouvriers	Travail à la tache	0.10
	Jours de récupération	0.15
	Primes	0.20

V.2.2.2. Traitement des données

Le traitement des données consiste à partager notre base de données en plusieurs sous bases, afin de réaliser l'entraînement (l'apprentissage) et la généralisation (test) du modèle développé, et d'éviter le phénomène de sur apprentissage.

Différentes méthodes de traitement des données présentées dans le chapitre III section 2.7.1.2, dans cette étude la méthode classique dite (Holdout methode en anglais) est utilisée, elle consiste à une division de la base de données en deux ensembles, un pour l'entraînement du modèle qui sera plus grand et le deuxième sert à tester le modèle.

Une proportion de 70% / 30%, qui représente 70% de la base des données est utilisée pour la phase d'apprentissage, tandis que les 30% restants sont utilisés pour tester le modèle.

Tableau V-2 Sous base de données de sortie d'apprentissage

N° projet	Taux de la productivité réel (non normalisé)	Taux de la productivité normalisé
1	1,11	0
2	2,64	0,743378509
3	2,68	0,760042242
4	2,86	0,85030413
5	1,84	0,353849291
6	2,36	0,604115957
7	2,45	0,651742432
8	1,96	0,413448157
9	2,40	0,62742669
10	1,34	0,110109341
11	2,45	0,651776161
12	2,56	0,701884956
13	2,81	0,826046347
14	1,51	0,19193639
15	1,46	0,171612353
16	2,40	0,624284088
17	2,14	0,50010062
18	2,26	0,560121922
19	1,47	0,175133644
20	2,52	0,685819164
21	1,96	0,41347514
22	2,40	0,624284088
23	2,83	0,835069071
24	1,35	0,116893792
25	2,36	0,604083577
26	2,45	0,651762669
27	2,10	0,481445504

28	2,94	0,890077182
29	2,18	0,519356724
30	2,21	0,532564938
31	2,64	0,743378509
32	2,42	0,636564453
33	1,64	0,254585306
34	1,38	0,128506892
35	2,54	0,693937215
36	2,32	0,589027021
37	1,27	0,077104135
38	1,60	0,237113766
39	1,27	0,077104135
40	2,54	0,693937215
41	1,35	0,114869426
42	1,45	0,162170204
43	2,57	0,709115194
44	1,45	0,165223146
45	2,54	0,693937215
46	1,45	0,165223146
47	1,45	0,165223146
48	1,45	0,165223146
49	1,45	0,165223146
50	1,45	0,165223146
51	2,54	0,693937215
52	2,54	0,693937215
53	3,01	0,923601724
54	1,95	0,408726119
55	2,70	0,773516528
56	1,88	0,374852724
57	2,20	0,528613994
58	2,18	0,519356724
59	2,20	0,528613994
60	2,54	0,693937215
61	1,45	0,165223146
62	2,64	0,743378509
63	2,13	0,494949306
64	2,10	0,481532235
65	2,45	0,651727977
66	2,64	0,743378509
67	2,86	0,851125888
68	3,17	1
69	1,45	0,165223146
70	2,06	0,45992037
71	1,47	0,175093169

Tableau V-3 Sous base de données de sortie de test

N° projet	Taux de la productivité réel (non normalisé)	Taux de la productivité normalisé
1	2,63	0,736825848
2	2,40	0,624284088
3	2,44	0,647526327
4	2,34	0,597570844
5	2,72	0,780092095
6	2,54	0,693937215
7	1,45	0,165223146
8	1,54	0,207947515
9	1,45	0,165223146
10	1,45	0,165223146
11	2,26	0,560090788
12	1,31	0,09566849
13	2,74	0,790571664
14	2,04	0,452303784
15	2,02	0,441002027
16	2,54	0,693937215
17	2,45	0,65168172
18	1,58	0,226195258
19	1,58	0,226195258
20	1,84	0,353849291
21	1,45	0,165223146
22	1,45	0,165223146
23	1,46	0,171612353
24	1,48	0,176671679
25	2,40	0,624284088
26	2,18	0,519356724
27	2,18	0,519356724
28	1,45	0,165223146
29	1,45	0,165223146
30	2,20	0,527283002

V.2.3. Type d'apprentissage

L'apprentissage (dit aussi entraînement) est la phase durant laquelle le modèle ajuste les valeurs des poids et biais via des algorithmes mathématiques, afin d'obtenir de meilleures performances.

Suivant le problème que l'on cherche à modéliser, différents types d'apprentissage peuvent être utilisés (voir chapitre III section 2.6.1).

Vu que notre recherche consiste à estimer la productivité du travail dans les chantiers de construction Algériens sous différents facteurs qui l'influencent, ainsi que les valeurs de cette productivité sont connues au préalable, donc le problème étudié est un problème de régression, et le type d'apprentissage utilisé donc est l'apprentissage supervisé.

V.2.4. Algorithme d'apprentissage

L'algorithme d'apprentissage est une méthode mathématique qui permet d'ajuster les paramètres du réseau de neurones artificiels par la modification itérative « petit à petit » des poids de connexion, afin d'adapter la réponse du modèle à la réponse désirée en minimisant en conséquence l'erreur entre ces deux dernières. L'apprentissage se termine lorsque l'algorithme atteint un état stable. (Voir chapitre III section 2.6.4). Vis-à-vis le problème que l'on cherche à modéliser et le type d'apprentissage choisi (problème de régression avec un apprentissage supervisé), l'algorithme utilisé dans notre étude, donc, est la Rétropropagation de l'erreur (Backpropagation), qui sert à minimiser progressivement l'erreur obtenue entre les sorties du modèle et les sorties désirées, en modifiant les poids et les biais entre les connexions des différentes couches afin d'obtenir des résultats optimaux.

V.2.5. Structure et architecture du RNA

Dans cette étude, le modèle du Réseau de Neurones Artificiels développé à une structure Multicouches de type statique (non bouclée), la transmission de l'information sera donc de la couche d'entrée vers la couche de sortie, passant par les couches cachées en un seul sens sans retour (feedforward).

V.2.5.1. Détermination du nombre des couches et des neurones cachés.

La détermination du nombre des couches cachées ainsi que le nombre des neurones cachés dans chaque couche est une étape principale dans la modélisation de notre réseau de Neurones Artificiels, alors qu'aucune règle précise n'existe pour la détermination de ces derniers.

Le nombre des couches et neurones cachés, se détermine donc par essais de plusieurs apprentissages sur différentes architectures, et on prend en conséquence celle qui a une meilleure performance.

Dans notre cas, et après plusieurs essais d'entrainements sur différentes architectures, on a fini par choisir le modèle qui donne une meilleure performance, soit une architecture constituée de 02 couches cachées ayant 15 Neurones cachés dans la première couche cachée et 10 Neurones cachés dans la deuxième couche.

La figure suivante présente l'architecture du modèle choisi

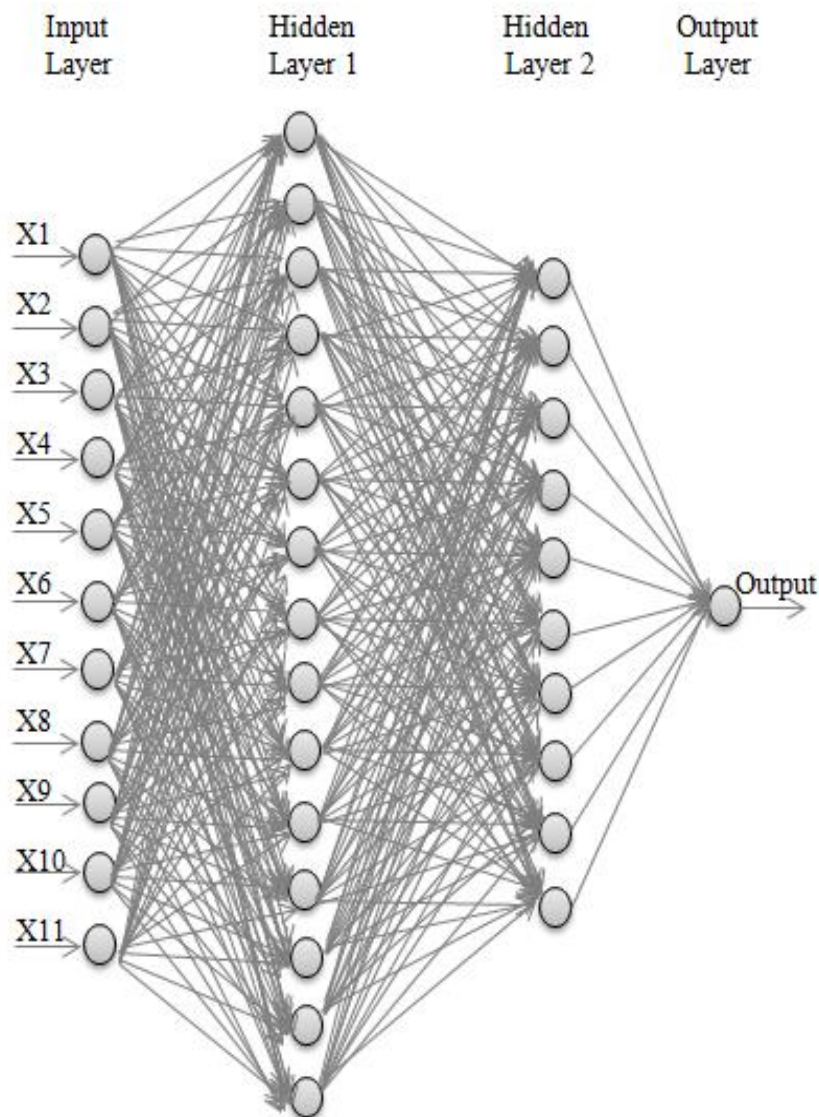


Figure V-2 Architecture du Réseau de Neurons Artificiels choisi

V.2.5.2. Détermination des fonctions de transfert.

Comme son nom l'indique, la fonction de transfert est une fonction qui permet de donner une non-linéarité à la somme pondérée des neurones. On peut trouver différentes fonctions de transfert (chapitre III 2.3.2).

Après plusieurs essais, on a abouti à choisir dans cette étude la fonction sigmoïde (Logsig) pour les connexions entre la couche d'entrée et la première couche cachée, et la fonction tangente hyperbolique (Tansig) pour les connexions entre la première couche cachée et la deuxième couche cachée, tandis que la fonction linéaire (Pureline) est utilisée pour les connexions entre la deuxième couche cachée et la couche de sortie.

$$\text{Logsig}(A) = 1/(1 + \exp(-A)) \quad (\text{V-3})$$

$$\text{Tansig}(A) = 2/(1 + \exp(-A)) - 1 \quad (\text{V-4})$$

$$F(A) = A \quad (\text{V-5})$$

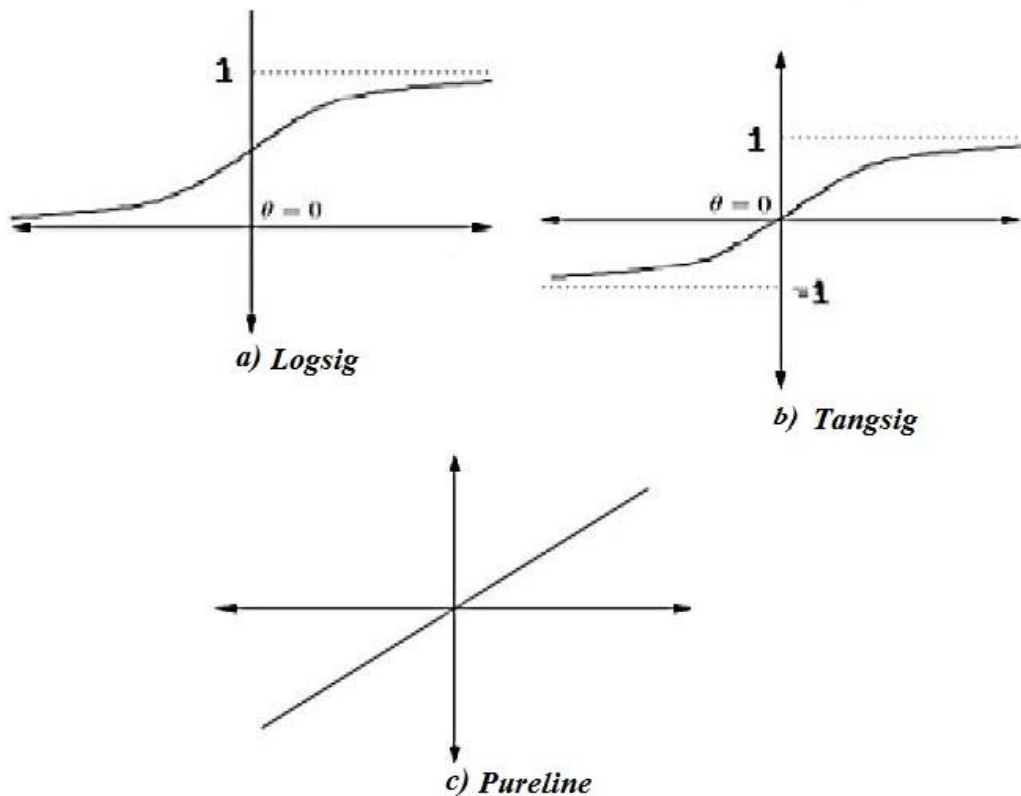


Figure V-3 Les fonctions de transferts utilisées

V.2.5.3. La fonction d'entraînement (fonction d'optimisation).

La fonction d'entraînement (apprentissage), est une fonction d'optimisation associée à un algorithme d'apprentissage (voir chapitre III section 2.6.4).

Dans notre étude, un réseau de neurones multicouches ayant une couche d'entrée, 02 couches cachées et une seule sortie, avec apprentissage supervisé et développé, l'algorithme d'apprentissage utilisé été la rétropropagation (backpropagation), et on a préféré d'utiliser une fonction d'optimisation de second ordre qui s'appelle "Bayesian regularization", cette dernière s'effectue au sein de l'algorithme de Levenberg-Marquardt (met à jour les valeurs de poids et de biais selon l'optimisation de Levenberg-Marquardt), elle consiste à ajouter un terme supplémentaire (terme de pénalité), qui représente la somme carrée des poids à l'erreur résultante, puis à propager les erreurs résultantes pour réajuster les poids et les biais, et à minimiser les erreurs pour obtenir la meilleure prédiction du résultat.

$$MSE_{br} = \gamma MSE + (1 - \gamma) MSW \quad (V-6)$$

Avec γ : performance ratio

$$\text{et } MSW = \frac{1}{n} \sum_{j=1}^n w_j^2$$

V.2.6. Phase d'apprentissage (entraînement)

Cette phase consiste à entraîner le programme développé. On le lance plusieurs fois, afin de calculer et d'ajuster les valeurs des poids et des biais dans les différentes couches en minimisant à chaque fois la valeur de l'erreur entre les résultats réels et les résultats du modèle, l'apprentissage se termine lorsqu'on obtient des résultats de l'erreur stable (mêmes résultats se produisent).

Les résultats d'apprentissage sont influencés par l'architecture du modèle développé, l'algorithme d'apprentissage, la fonction d'optimisation et les fonctions de transfert. À cet égard, notre modèle est un Réseau de Neurones Artificiels multicouches de propagation avant (feedforward) constitué de 11 Neurones en couche d'entrée, 15 et 10 neurones dans la première et la deuxième couche cachée respectivement et d'une seule sortie.

L'algorithme de rétropropagation (Backpropagation) de l'erreur avec la fonction d'optimisation du Bayesian regularization qui met à jour les valeurs de poids et de biais selon

l'optimisation de Levenberg-Marquardt est utilisé afin d'ajuster les poids et les biais et d'éviter le phénomène de surapprentissage, ainsi que les fonctions de transfert sigmoïde (Logsig), la fonction tangente hyperbolique (Tansig) et la fonction linéaire (Pureline) sont utilisées entre les différentes couches.

Il est à noter que 70 % des données collectées sont utilisées durant cette phase.

V.2.6.1. Résultats de la phase d'apprentissage

Après plusieurs entraînements, le modèle développé a appris à estimer les résultats réels de la productivité et produit des résultats proches de la réalité.

Les paramètres et les performances résultant de cette phase sont présentés comme suit :

V.2.6.1.1. Poids et biais

Les valeurs des poids et des biais des connexions entre les différentes couches résultant du modèle le plus performant dans la phase d'apprentissage sont :

Tableau V-4 Valeurs des poids et biais entre la couche d'entrée et la première couche cachée

Poids											Biais
-0,432152712	-0,274556404	-0,111894642	-0,04216896	-0,221955081	-0,126589057	0,330576925	1,216193487	-0,216393691	-0,121701369	0,006285222	-0,00828471
-0,02789528	0,064317763	0,125427113	0,18625913	0,226393055	-0,076850147	0,214992507	0,103191661	-0,087421184	-0,043111324	0,151485574	0,03260763
-0,335011257	-0,008634906	-0,218130538	-0,029332605	-0,058868926	-0,159835826	-0,023429241	0,818920266	0,130697173	0,031748736	-0,011083951	0,03568679
0,254341252	0,061353083	0,253653914	0,116848071	0,112234577	-0,098850624	0,161822208	-0,283312132	-0,13079863	-0,090948847	0,091025065	0,03182371
0,501656296	0,116387745	0,499883363	0,486787791	0,165371737	-0,266171557	0,415681787	-0,033385732	-0,152441258	0,040021063	0,227743019	0,07991949
0,272852677	0,052926649	0,265746299	0,108740842	0,104013608	-0,092208782	0,158500644	-0,322777215	-0,135979117	-0,089407197	0,084587109	0,03302243
-0,50288852	0,300426978	1,255126357	-0,45665273	-0,22038298	0,069613372	-0,268046002	-0,034884488	0,046817133	0,260906006	-0,358778184	0,45169796
-0,379449149	0,101974478	0,162185493	0,269767012	0,348198382	0,261815837	0,226749658	0,468146242	-0,266007322	0,019799291	0,069017295	0,00308396
-0,032277083	0,063514552	0,126355822	0,188401927	0,229413291	-0,076486728	0,21751217	0,108041428	-0,089900967	-0,043078045	0,15237248	0,0337335
-0,927049685	-0,07887061	-0,757415928	0,361171709	0,226315145	-0,052965486	0,355886135	-1,431159887	-0,167121403	-0,200463713	0,357227001	-1,99047124
-2,069645213	-0,12755728	-0,67088614	0,416286043	0,424489201	0,030975862	0,23458405	0,171186433	-0,302065326	-0,174862621	0,352766073	-1,03976588
-0,292915338	0,059904816	-0,155818462	0,027090704	0,016304424	-0,089620074	-0,05359517	0,589902715	0,108871621	0,060867606	0,007950099	0,03622581
-0,331580587	0,011647923	-0,18991195	0,006916872	-0,015855603	-0,104973583	-0,035721247	0,748509798	0,101579441	0,050711805	-0,007204642	0,05336218
-0,261728196	0,074627367	0,076696391	0,179034032	0,256471996	0,053855496	0,173620685	0,362221023	-0,136823021	0,027921418	0,102523425	0,03753817
-0,218210097	0,061823746	-0,057457792	0,076471623	0,087163839	-0,038708408	0,034936829	0,391591966	0,014745829	0,050025352	0,053278328	0,02703889

Tableau V-5 Valeurs des poids et biais entre la première couche cachée et la deuxième couche cachée

Poids															Biais
-0,10628	0,01180	-0,05254	0,04432	0,12514	0,04654	0,03463	-0,03095	0,01096	0,05889	0,00795	-0,02717	-0,04518	-0,01205	-0,00998	0,04453
0,03274	-0,00384	0,01581	-0,01406	-0,03936	-0,01476	-0,01148	0,00955	-0,00357	-0,01846	-0,00264	0,00790	0,01353	0,00354	0,00264	-0,01499
0,06590	-0,00755	0,03212	-0,02796	-0,07860	-0,02936	-0,02240	0,01923	-0,00702	-0,03682	-0,00503	0,01629	0,02754	0,00728	0,00568	-0,02910
-0,62987	-0,28135	-0,45800	0,02200	0,00119	0,05158	0,33766	-0,57022	-0,28675	1,20529	0,48703	-0,37127	-0,43771	-0,43820	-0,31487	0,12469
0,66818	0,18409	0,41988	-0,13535	-0,17041	-0,16439	-0,37157	0,51500	0,19008	-0,74409	-0,23878	0,31151	0,39352	0,36644	0,24598	-0,22768
0,63162	-0,28862	0,51164	-0,53940	-0,90530	-0,55863	-1,00744	0,06038	-0,28837	-0,88083	1,56774	0,39111	0,46724	-0,00491	0,17937	-0,10783
-0,09492	0,01063	-0,04674	0,03978	0,11220	0,04177	0,03128	-0,02766	0,00988	0,05269	0,00708	-0,02404	-0,04016	-0,01069	-0,00872	0,04036
-0,22381	0,02208	-0,11543	0,08828	0,24980	0,09287	0,06700	-0,06454	0,02041	0,12528	0,02302	-0,06262	-0,09997	-0,02717	-0,02566	0,08166
0,02563	-0,00302	0,01236	-0,01103	-0,03085	-0,01157	-0,00903	0,00748	-0,00281	-0,01448	-0,00209	0,00616	0,01057	0,00277	0,00205	-0,01180
-0,27954	0,02459	-0,14710	0,10667	0,29964	0,11247	0,08192	-0,08227	0,02256	0,15951	0,03472	-0,08168	-0,12793	-0,03670	-0,03532	0,09647

Tableau V-6 Valeurs des poids et biais entre la deuxième couche cachée et la couche de sortie

Poids										Biais
-0,211400123	0,066104361	0,132254197	-1,56287944	1,304569253	2,193533925	-0,189292181	-0,432547226	0,051798261	-0,530774925	-0,190839229

V.2.6.1.2. Evaluation de la performance du modèle dans la phase d'apprentissage.

La sortie du modèle qui correspond au taux de la productivité est obtenue en appliquant les poids et les biais dans l'équation empirique du réseau de neurones artificiels la suivante :

$$O_i = f(\theta_i + \sum_{j=1}^n w_{ij} * x_j) \quad (V-7)$$

Avec O_i est la sortie du modèle

X_j : représente les entrées du modèle (les facteurs)

w_{ij} et θ_i : sont les Poids est Biais des connexions

f est la fonction du transfert

Cette équation est appliquée au niveau des couches cachées afin d'obtenir les neurones cachés et au niveau de la sortie pour obtenir les résultats du taux de la productivité.

Les résultats de la sortie du modèle (taux de la productivité) de la phase d'apprentissage sont montrés dans le tableau suivant :

Tableau V-7 Résultats du modèle et résultats réels durant la phase d'apprentissage

N° Ordre	Sorties du modèle normalisée [0;1]	Résultats réelles normalisées [0;1]	Retour à l'origine Sorties du modèle	Retour à l'origine Résultats réelles
1	-0,030570009	0	1,05	1,11
2	0,743378507	0,743378509	2,64	2,64
3	0,760042238	0,760042242	2,68	2,68
4	0,850304122	0,85030413	2,86	2,86
5	0,346352048	0,353849291	1,82	1,84
6	0,604115957	0,604115957	2,36	2,36
7	0,651742422	0,651742432	2,45	2,45
8	0,41344816	0,413448157	1,96	1,96
9	0,631085311	0,62742669	2,41	2,40
10	0,11010934	0,110109341	1,34	1,34
11	0,651776154	0,651776161	2,45	2,45
12	0,701884962	0,701884956	2,56	2,56
13	0,826046367	0,826046347	2,81	2,81
14	0,191936375	0,19193639	1,51	1,51
15	0,171612341	0,171612353	1,46	1,46
16	0,624284094	0,624284088	2,40	2,40
17	0,500100625	0,50010062	2,14	2,14

18	0,56012192	0,560121922	2,26	2,26
19	0,175133645	0,175133644	1,47	1,47
20	0,689302121	0,685819164	2,53	2,52
21	0,41347513	0,41347514	1,96	1,96
22	0,624284085	0,624284088	2,40	2,40
23	0,835069079	0,835069071	2,83	2,83
24	0,116893782	0,116893792	1,35	1,35
25	0,604083578	0,604083577	2,36	2,36
26	0,651762674	0,651762669	2,45	2,45
27	0,481445492	0,481445504	2,10	2,10
28	0,890077183	0,890077182	2,94	2,94
29	0,519356721	0,519356724	2,18	2,18
30	0,532564935	0,532564938	2,21	2,21
31	0,743378507	0,743378509	2,64	2,64
32	0,636564451	0,636564453	2,42	2,42
33	0,2545853	0,254585306	1,64	1,64
34	0,1370813	0,128506892	1,39	1,38
35	0,693937223	0,693937215	2,54	2,54
36	0,58902702	0,589027021	2,32	2,32
37	0,077104128	0,077104135	1,27	1,27
38	0,237113761	0,237113766	1,60	1,60
39	0,077104121	0,077104135	1,27	1,27
40	0,693937213	0,693937215	2,54	2,54
41	0,114869416	0,114869426	1,35	1,35
42	0,153045712	0,162170204	1,43	1,45
43	0,709115197	0,709115194	2,57	2,57
44	0,165223142	0,165223146	1,45	1,45
45	0,693937224	0,693937215	2,54	2,54
46	0,16522314	0,165223146	1,45	1,45
47	0,165223146	0,165223146	1,45	1,45
48	0,165223148	0,165223146	1,45	1,45
49	0,165223143	0,165223146	1,45	1,45
50	0,165223149	0,165223146	1,45	1,45
51	0,693937218	0,693937215	2,54	2,54
52	0,693937217	0,693937215	2,54	2,54
53	0,906882081	0,923601724	2,98	3,01
54	0,408726124	0,408726119	1,95	1,95
55	0,773516528	0,773516528	2,70	2,70
56	0,374852717	0,374852724	1,88	1,88
57	0,515561625	0,528613994	2,17	2,20
58	0,51935673	0,519356724	2,18	2,18

59	0,516862787	0,528613994	2,18	2,20
60	0,668748646	0,693937215	2,49	2,54
61	0,165223152	0,165223146	1,45	1,45
62	0,743378504	0,743378509	2,64	2,64
63	0,494949298	0,494949306	2,13	2,13
64	0,481532237	0,481532235	2,10	2,10
65	0,65172797	0,651727977	2,45	2,45
66	0,743378504	0,743378509	2,64	2,64
67	0,8511259	0,851125888	2,86	2,86
68	1,000000004	1	3,17	3,17
69	0,165223142	0,165223146	1,45	1,45
70	0,420204874	0,45992037	1,98	2,06
71	0,175093163	0,175093169	1,47	1,47

Durant la phase d'apprentissage, des performances élevées sont enregistrées et les résultats des sorties obtenues représentent avec fiabilité les résultats réels.

Les performances du modèle sont évaluées par le calcul de l'erreur entre les résultats obtenus et celles réelles du taux de la productivité du travail et ils sont donnés par le tableau suivant :

Tableau V-8 Performances du modèle dans la phase d'apprentissage

MSE (données normalisées)	MSE (données réelles)	MAPE % (données normalisées)	MAPE % (données réelles)	AA % (données normalisées)	AA % (données réelles)
0,0000559	0,000237	0.489	0.269	99.511	99.731

On constate d'après le tableau V-8 que notre modèle interprète très bien les données présentées, vu les performances enregistrées, avec une précision d'estimation atteinte 99.731 %, et une moyenne des erreurs carrées égale à 2.37 E-04, tandis que l'erreur moyenne absolue en pourcentage MAPE est de 0.269.

Autre que le calcul de l'erreur, la régression linéaire est aussi une mesure qui sert à évaluer la performance du modèle développé, la droite de régression présentée dans la figure V-4, indique une bonne relation entre les entrées et sorties, ainsi qu'une valeur du coefficient de corrélation proche 1 (R=0.999) est enregistrée.

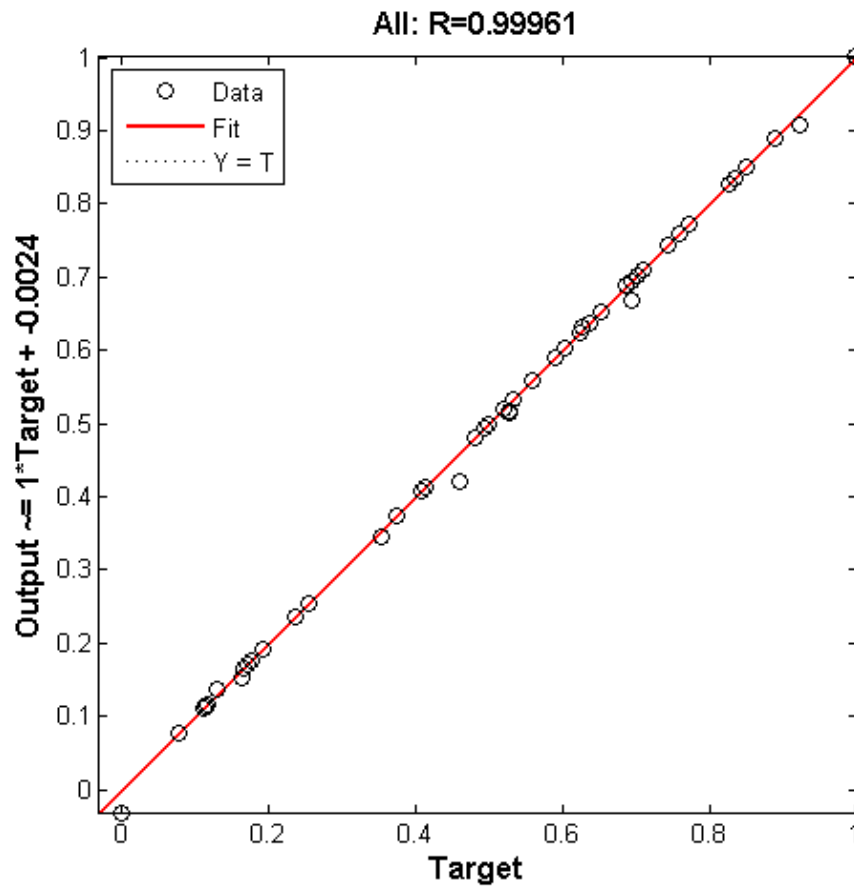


Figure V-4 Droite de régression du RNA durant la phase d'apprentissage

On peut remarquer depuis la figure suivante (figure V-5) que les résultats obtenus par le modèle du réseau de neurones artificiels développé sont presque superposés aux résultats réels de la productivité, chose qui prouve une très bonne estimation de la productivité du travail dans les projets de construction Algériens par le modèle développé dans la phase d'apprentissage.

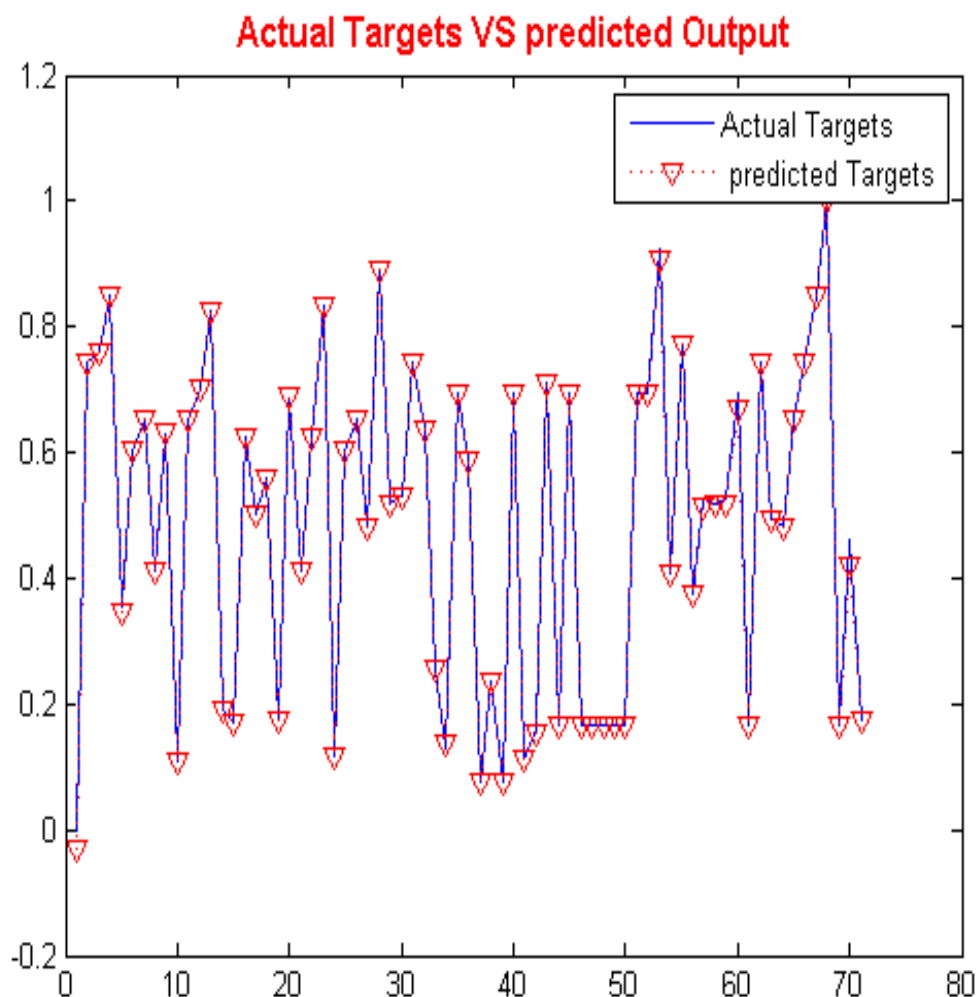


Figure V-5 Productivité Réelle et Productivité Estimée dans la phase d'apprentissage

V.2.7. Phase de test

La phase test sert à généraliser les résultats obtenus dans la phase d'apprentissage, en présentant au modèle développé des nouvelles données d'entrée (jamais vues dans la phase d'apprentissage) et en les appliquant par la suite aux valeurs des poids et des biais des connexions entre les différentes couches obtenues dans la phase d'apprentissage, et évaluant en conséquence les résultats des sorties obtenues.

Il est à noter que 30% de la base de données sont utilisés dans cette phase

V.2.7.1. Résultats de la phase de test

Les résultats de la sortie du modèle (taux de la productivité) de la phase de test sont montrés dans le tableau suivant :

Tableau V-9 Résultats du modèle et résultats réels durant la phase du Test

N° Ordre	Sorties du modèle normalisée [0;1]	Résultats réelles normalisées [0;1]	Retour à l'origine Sorties du modèle	Retour à l'origine Résultats réelles
1	0,764528771	0,736825848	2,69	2,63
2	0,600592941	0,624284088	2,35	2,40
3	0,649165423	0,647526327	2,45	2,44
4	0,594215841	0,597570844	2,33	2,34
5	0,795156607	0,780092095	2,75	2,72
6	0,705097088	0,693937215	2,56	2,54
7	0,201992477	0,165223146	1,53	1,45
8	0,199624578	0,207947515	1,52	1,54
9	0,186635504	0,165223146	1,50	1,45
10	0,140095002	0,165223146	1,40	1,45
11	0,558502679	0,560090788	2,26	2,26
12	0,102427468	0,09566849	1,32	1,31
13	0,769649251	0,790571664	2,70	2,74
14	0,429250978	0,452303784	2,00	2,04
15	0,439537551	0,441002027	2,02	2,02
16	0,683980917	0,693937215	2,52	2,54
17	0,655302414	0,65168172	2,46	2,45
18	0,228259422	0,226195258	1,58	1,58
19	0,230679224	0,226195258	1,59	1,58
20	0,345659034	0,353849291	1,82	1,84
21	0,181309782	0,165223146	1,48	1,45
22	0,180277198	0,165223146	1,48	1,45
23	0,174886595	0,171612353	1,47	1,46
24	0,177805741	0,176671679	1,48	1,48
25	0,621665334	0,624284088	2,39	2,40
26	0,519859408	0,519356724	2,18	2,18
27	0,519402607	0,519356724	2,18	2,18
28	0,150848265	0,165223146	1,42	1,45
29	0,165731452	0,165223146	1,45	1,45
30	0,543851125	0,527283002	2,23	2,20

On constate que les résultats issus de cette phase sont proche des résultats réels

V.2.7.2. Evaluation de la performance du modèle dans la phase test.

Comme dans la phase d'apprentissage, les performances du modèle sont évaluées par le calcul de l'erreur entre les résultats obtenus dans la phase de test et celle réelle du taux de la productivité du travail.

Les résultats sont illustrés dans le tableau suivant :

Tableau V-10 Performances du modèle dans la phase du Test

MSE (données normalisées)	MSE (données réelles)	MAPE % (données normalisées)	MAPE % (données réelles)	AA % (données normalisées)	AA % (données réelles)
0.000213	0.000906	4.101	1.224	95.899	98.776

Le tableau V-10 montre des performances élevées résultent dans la phase de test, ce qui indique que le modèle développé généralise très bien les nouvelles données.

Les résultats de la régression linéaire de la phase du test ont montré aussi une bonne relation entrées-sortie, avec un coefficient de corrélation proche de 1 ($R=0.998$), la figure V-6 illustre graphiquement cette relation.

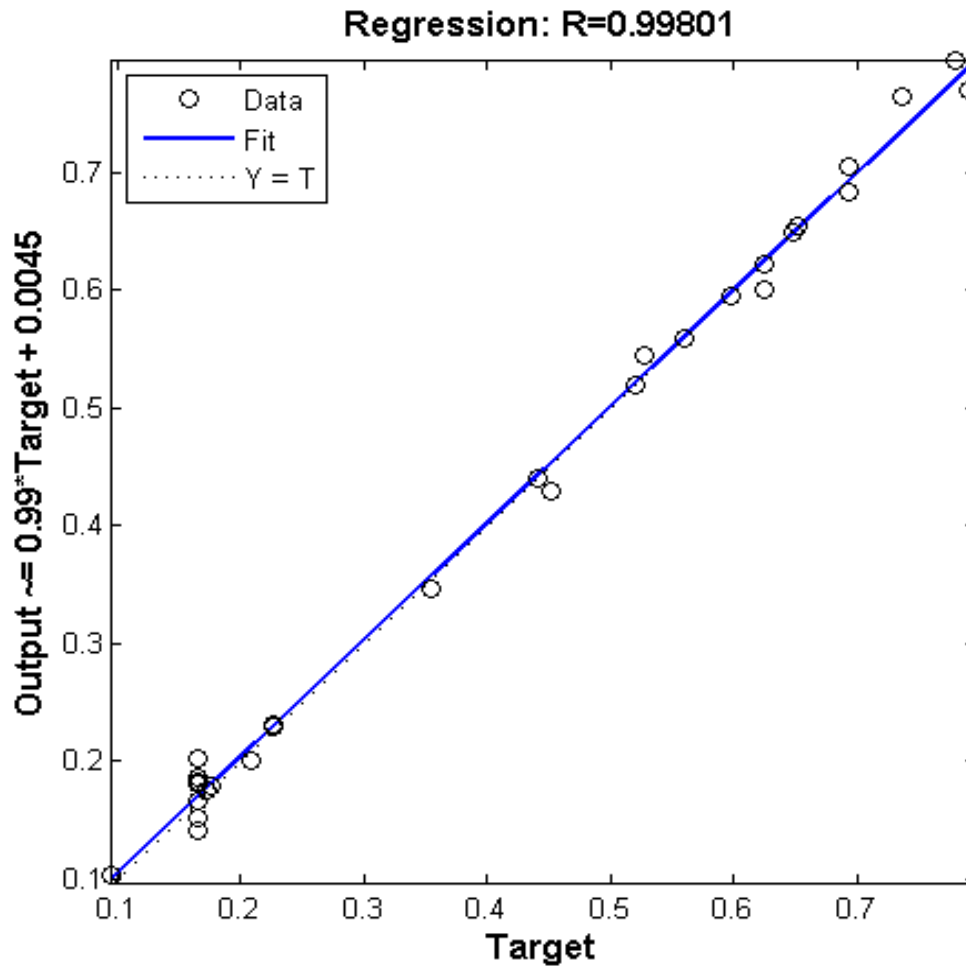


Figure V-6 Droite de régression du RNA durant la phase du Test

La comparaison entre les résultats réels et les sorties résultant du modèle développé du Réseau de Neurones Artificiels est montrée graphiquement comme suite :

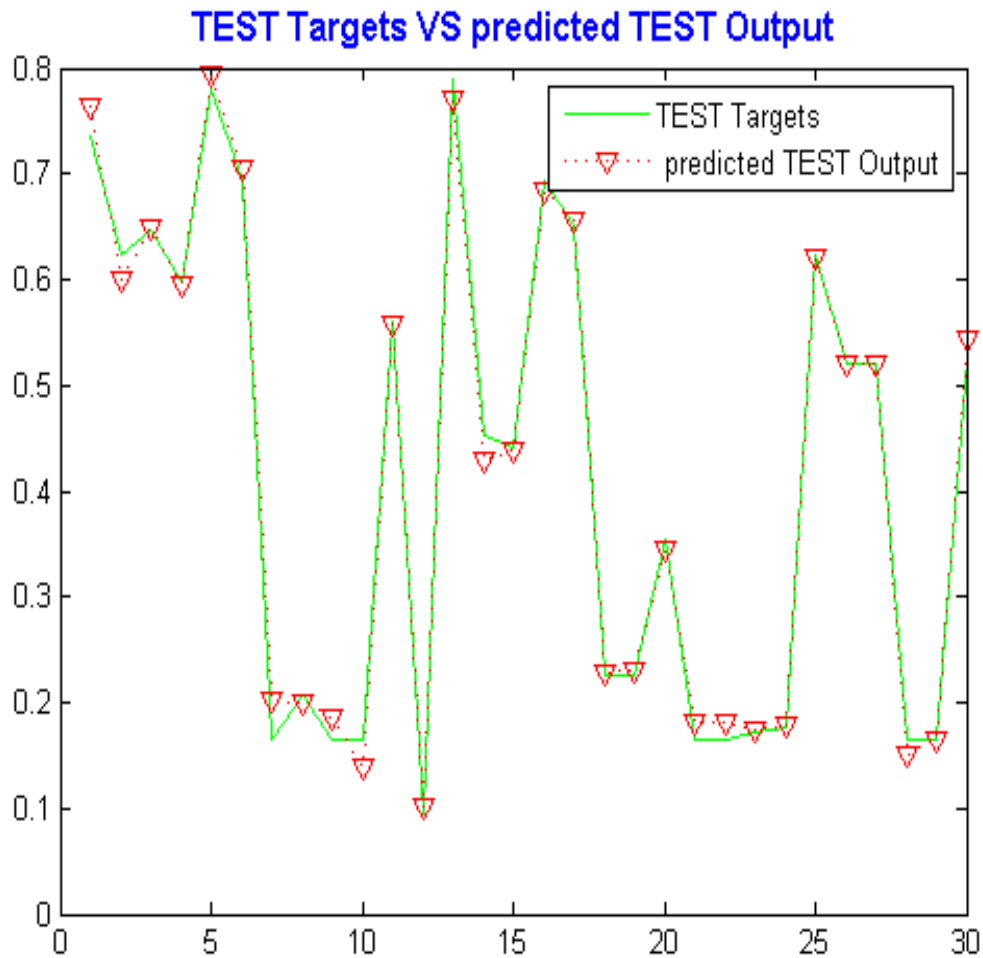


Figure V-7 Productivité Réelle et Productivité Estimée dans la phase du Test

On constate que les sorties du modèle sont superposées aux taux de la productivité réelle, chose qui prouve que le modèle développé représente avec grande confiance le taux de la productivité du travail dans les projets de construction Algériens.

V.3. Conclusion.

Dans ce chapitre, on a développé un modèle mathématique, basé sur la théorie de l'intelligence artificielle, par la méthode dite : Réseau de Neurons Artificiel (RNA), avec une architecture composée de 11 facteurs comme entrées, de deux couches cachées et d'une seule sortie, le modèle développé a montré sa puissance quant à l'estimation du taux de la productivité du travail dans les projets de construction algériens, et d'après les résultats obtenus, on peut dire qu'on peut compter sur notre modèle pour la résolution de ce type de problème.

CHAPITRE VI

Modélisation de la productivité par la régression linéaire multiple

VI.1. Introduction

La Régression Linéaire Multiple (RLM), est une méthode mathématique paramétrique, considérée comme la méthode la plus répandue de l'approche paramétrique, utilisée pour modéliser des problèmes mathématiques linéaires.

Cette méthode est utilisée dans ce chapitre afin d'estimer la productivité de travail dans les projets de construction Algériens, les tests utilisés ainsi que les étapes de développement du modèle sont présentés aussi.

VI.2. Modélisation par la Régression Linéaire Multiple.

La modélisation par la méthode de la Régression Linéaire Multiple (RLM), consiste à trouver la meilleure relation linéaire entre les données d'entrée (les facteurs) et la sortie, et à produire en conséquence une estimation de la productivité du travail proche de la réalité.

Les coefficients standardisés β_i , sont les paramètres à déterminer afin de satisfaire cette relation.

Les étapes de développement, ainsi que les conditions d'applicabilité et les différents tests utilisés du modèle de la Régression Linéaire Multiple sont présentés comme suite :

- Détermination des données d'entrées et de sorties
- Tester l'applicabilité du RLM (tester la normalité, l'homogénéité, la linéarité et l'indépendance des données).
- Développement du modèle et détermination des paramètres (coefficients standardisés) de la RLM
- Tester la significativité des coefficients obtenus (signification globale, et signification de chaque coefficient)
- Analyser les résidus (la normalité, l'homogénéité, la linéarité et l'indépendance des résidus)
- Evaluer le modèle développé
- Tester les performances de modèle développé par des nouvelles données d'entrées

VI.2.1. Détermination des données d'entrées et de sorties

Identiquement aux données entrées-sorties utilisées dans le développement du modèle de réseau de neurones artificiels dans le chapitre précédent, les données d'entrées sont celles déterminées par l'étude statistique du questionnaire traité dans le chapitre (IV.3) (voir Tableau IV.18, Figure V-1), et la sortie représente le taux de la productivité dans les projets de construction Algérien (se référer au chapitre V.2.1.2).

Dans la terminologie de la RLM les données d'entrée sont appelées variables explicatives et la sortie est appelée variable à expliquer ou la réponse.

VI.2.2. Tests d'applicabilité de la Régression Linéaire Multiple (RLM).

Avant d'entamer la modélisation par la méthode du RLM les conditions (hypothèses) suivantes doivent être vérifiées :

- Les données de l'étude doivent être homogènes
- Les données de l'étude doivent suivre une loi normale
- Les données de l'étude doivent être indépendantes
- La relation entre Y et X_i doit être linéaire

VI.2.2.1. Test de normalité des données

Comme nous avons vu dans le chapitre III section 3.1.2.4.1, la normalité des variables explicatives, peut se tester soit statistiquement, soit graphiquement.

Dans ce qui suit, nous allons tester la normalité de l'ensemble des données par le test d'asymétrie et d'aplatissement.

Tableau VI-1 Test d'asymétrie et aplatissement de normalité des données d'origine

Facteurs	N	Asymétrie		Aplatissement	
	Statistique	Statistique	Erreur. Std	Statistique	Erreur. Std
Nombre des heures de travail	70	0,137	0,287	-0,969	0,566
nombre des manœuvres	70	0,422	0,287	0,972	0,566
Nombre des ouvriers	70	-0,314	0,287	-0,251	0,566
Expérience minimale	70	1,896	0,287	3,496	0,566
Expérience maximale	70	1,784	0,287	4,749	0,566
Etat des matériels de construction	70	0,063	0,287	-0,083	0,566
Type de Motivation	70	0,693	0,287	-0,579	0,566
Surface de plancher	70	1,428	0,287	3,995	0,566
Température	70	-0,303	0,287	-0,718	0,566
Salaire des manœuvres	70	0,422	0,287	-1,429	0,566
Salaire des ouvriers	70	1,979	0,287	8,280	0,566

Depuis le tableau VI-1, on constate que quelques données telles que : Expérience minimale, Expérience maximale, Type de motivation, Quantité de travail réalisée ainsi que Salaire des ouvriers ont une z-value > 1.97 ce qui signifie qu'elles ne suivent pas une distribution normale et des transformations de ces données doivent être faites afin de satisfaire l'hypothèse de normalité.

Les transformations des données sont faites comme suite :

- Expérience Minimale transformée = $1 / (\sqrt{\text{Expérience Minimale}})$
- Expérience Maximale transformée = $\text{Ln}(\text{Expérience Maximale})$
- Type de motivation transformé = $\text{EXP}(\text{Ln}(\text{Type de motivation})/2.5)$
- Surface du plancher transformée = $\text{Ln}(\text{Surface du plancher})$
- Salaire d'ouvrier transformé = $\sin(\text{Salaire d'ouvrier} / 1.25)$
- Salaire Manœuvre transformé = $\cos(\text{Salaire Manœuvre} / 5)$

Après ces transformations le test d'asymétrie et d'aplatissement se donne les résultats suivants :

Tableau VI-2 Test d'asymétrie et aplatissement de normalité des données transformées

Facteurs	N	Asymétrie		Aplatissement	
	Statistique	Statistique	Erreur. Std	Statistique	Erreur. Std
Nombre des heures de travail	70	0,137	0,287	-0,969	0,566
Nombre des manœuvres	70	0,422	0,287	0,972	0,566
Nombre des ouvriers	70	-0,314	0,287	-0,251	0,566
Expérience minimale transformée	70	-0,083	0,287	-0,785	0,566
Expérience maximale transformée	70	0,497	0,287	-0,907	0,566
Etat des matériels de construction	70	0,063	0,287	-0,083	0,566
Type de Motivation transformé	70	0,560	0,287	-1,108	0,566
Quantité de travail réalisé transformée	70	0,011	0,287	0,188	0,566
Température	70	-0,303	0,287	-0,718	0,566
Salaire des manœuvres transformé	70	-0,371	0,287	-0,791	0,566
Salaire des ouvriers transformé	70	-0,143	0,287	-1,087	0,566

D'après le tableau VI-2, on constate que le z-value = coefficient d'asymétrie / Erreur standard, de tous les facteurs sont comprises entre [-1.96 ; +1.96], ce qui implique que les facteurs de l'étude suivent une loi normale.

VI.2.2.2. Test d'homogénéité des données

Plusieurs tests statistiques peuvent être utilisés afin de statuer sur homogénéité des données (variables explicatives), (se réfère au chapitre III.3.1.2.4.2, pour plus de détails). Dans cette étude le Test de Levene est utilisé et les résultats sont montrés dans le tableau suivant :

Tableau VI-3 Test de Levene d'homogénéité des données

Test d'homogénéité des variances				
Facteurs	Statistique de Levene	ddl1	ddl2	Sig.
Nombre des heures de travail	3,061	5	64	0,015
Nombre des manœuvres	1,198	2	66	0,31
Nombre des ouvriers	2,863	7	60	0,012
Expérience minimale transformée	2,696	7	62	0,017
Expérience maximale transformée	2,625	11	54	0,009
Etat des matériels de construction	0,136	2	67	0,873
Type de Motivation transformé	3,645	2	67	0,031
Quantité de travail réalisé transformée	2,768	15	41	0,005
Température	2,405	15	51	0,010
Salaire des manœuvres transformé	2,217	3	65	0,094
Salaire ouvrier transformé	7,392	1	67	0,008

Depuis le tableau VI-3 on constate que la P-value de la statistique de Levene pour la plupart des facteurs (08 facteurs) est inférieure à 0.05, ce qui indique que ces variables explicatives ne sont pas homogènes, tandis que trois facteurs sont homogènes avec une P-value supérieure à 0.05.

Pour remédier à ce problème d'homogénéité on a choisi de transformer la variable dépendante (la productivité réalisée)

La productivité réalisée transformée= $\arcsin(\text{productivité réalisée} / \text{productivité réalisée}_{\max})$.

Ou productivité réalisée $_{\max} = 3.17$, alors La productivité réalisée transformée prend la forme suivante :

La productivité réalisée transformée= $\arcsin(\text{productivité réalisée} / 3.17)$.

Après cette transformation le test d'homogénéité des données d'entrées devient comme suit :

Tableau VI-4 Test de Levene d'homogénéité des données après transformation de la variable dépendante

Test d'homogénéité des variances				
Facteurs	Statistique de Levene	ddl1	ddl2	Sig.
Nombre des heures de travail	2,530	5	64	0,038
Nombre des manœuvres	0,510	2	66	0,603
Nombre des ouvriers	1,371	7	60	0,234
Expérience minimale transformée	1,889	7	62	0,087
Expérience maximale transformée	1,491	11	54	0,162
Etat des matériels de construction	0,238	2	67	0,789
Type de Motivation transformé	0,977	2	67	0,382
Quantité de travail réalisé Transformée	1,716	15	41	0,086
Température	2,373	15	51	0,011
Salaire des manœuvres transformé	0,562	3	65	0,642
Salaire ouvrier transformé	0,853	4	64	0,497

On constate depuis le tableau VI-4, une amélioration de l'homogénéité des données après la transformation de la variable dépendante, et on remarque que la plupart des variables explicatives ont un P-value supérieure à 0.05, et seulement 02 variables ont une P-value inférieure à 0.05.

Depuis les résultats du tableau VI-4, on considère que notre jeu de données est homogène.

VI.2.2.3. Test de l'indépendance des données

Le test de l'indépendance sert à vérifier que les variables sont indépendantes, en calculant la corrélation « r » entre eux, une faible valeur de « r » indique l'indépendance entre les Xi.

Tableau VI-5 Matrice de corrélation des données

	Nombre des heures de travail	Nombre des manœuvres	Nombre des ouvriers	Etat des matériels de construction	Température	Exp min transformée	Exp max transformée	Surface plancher transformée	Motivation transformée	Salaire ouvrier transformé	Salaire Mo transformé
Nombre des heures de travail	1	0,143	-0,012	0,131	-0,274	-0,103	0,159	0,482	0,687	-0,097	0,113
Nombre des manœuvres		1	0,135	0,215	0,049	-0,168	0,045	0,198	0,065	0,265	-0,212
Nombre des ouvriers			1	0,312	-0,183	0,020	-0,140	0,556	0,012	-0,111	-0,162
Etat des matériels de construction				1	-0,030	-0,381	0,367	0,341	0,207	0,044	-0,151
Température					1	-0,303	0,377	-0,442	0,013	0,326	-0,016
Exp min transformée						1	-0,802	0,011	-0,407	-0,191	0,004
Exp max transformée							1	-0,069	0,537	0,145	0,139
Quantité de travail réalisé transformée								1	0,115	-0,321	-0,098
Motivation transformée									1	-0,005	0,092
Salaire ouvrier transformé										1	-0,080
Salaire Mo transformé											1

D'après le tableau VI-5, on remarque des faibles corrélations entre la plupart des variables explicatives de l'étude, des valeurs inférieures à 0.7 sont enregistrées, et seulement une valeur de $r = -0.802$ qui est légèrement supérieure à 0.7 apparaitre entre l'expérience Minimale et l'expérience Maximale.

On peut dire en conséquence que les variables explicatives sont indépendantes.

VI.2.2.4. Test de linéarité des données

Dans ce qui suit on suppose que les variables explicatives et la variable indépendante, sont reliées entre eux par une relation linéaire de la forme :

$$Y_i = \beta_0 + \beta_1 X_{i1} + \beta_2 X_{i2} + \dots + \beta_p X_{ip} + \varepsilon_i \quad (\text{VI-1})$$

Y_i : Variable à expliquer (variable dépendante)

X_{ip} : Variables explicatives (variable indépendante)

β_0 et β_p : Coefficients (paramètres du modèle).

ε_i : erreur aléatoire.

VI.2.3. Développement du modèle et détermination des paramètres (coefficients) de la RLM.

Après la vérification des hypothèses (conditions d'applicabilité) de la Régression Linéaire Multiple RLM, on va passer maintenant au développement de notre modèle, en introduisant 70% des données collectées.

Le développement du modèle de la RLM, et l'analyse statistique des résultats obtenus sont réalisés à l'aide du logiciel du calcul statistique (SPSS)

Les résultats d'estimation des coefficients non standardisés sont montrés dans le tableau suivant

Tableau VI-6 Estimation des coefficients (paramètres) du modèle

Modele	Coefficients Non-standardisés		Coefficients Standardisés	t	Sig.
	B	Erreur. Std	Beta		
(Constant)	3,225	0,329		9,799	0,000
Nombre des heures de travail	0,145	0,013	0,956	10,827	0,000
Nombre des manœuvres	0,053	0,018	0,135	2,963	0,004
Nombre des ouvriers	0,080	0,008	0,654	10,648	0,000
Expérience minimale transformée	0,083	0,066	0,088	1,248	0,217
Expérience maximale transformée	0,024	0,037	0,058	0,656	0,514
Etat des matériels de construction	-0,002	0,022	-0,004	-0,079	0,937
Type de Motivation transformé	-0,126	0,547	-0,019	-0,231	0,818
Quantité de travail réalisé transformée	-0,869	0,052	-1,344	-16,603	0,000
Température	0,003	0,003	0,047	0,913	0,365
Salaire des manœuvres transformé	-0,009	0,021	-0,020	-0,445	0,658
Salaire ouvrier transformé	-0,010	0,017	-0,027	-0,568	0,572

VI.2.4. Tests de significativité des coefficients obtenus

VI.2.4.1. Test de significativité globale (tous les paramètres du modèle) Test de Fisher

Le test de significativité globale sert à tester la fiabilité du modèle développé, en testant la nullité de tous les paramètres du modèle, un résultat de F-value de Fisher calculé, supérieur à la F-value de Fisher présentée dans le tableau de Fisher avec un risque de 5 % (0.05), indique que le modèle développé est fiable.

Le tableau suivant présente les résultats d'ANOVA du modèle développé :

Tableau VI-7 ANOVA du modèle développé

Modèle	Somme des carrés	ddl	Moyenne des carrés	F	Sig.
Régression	3,835	11	0,349	50,868	0,000
1 Résidu	0,398	58	0,007		
Total	4,233	69			

On constate depuis le tableau VI-7, que la valeur de Fisher calculée, égale à 50,868 qui est supérieure à la valeur de Fisher tabulée F_{tab} pour 11/58 degrés de liberté qui est égale à 1.979 à un niveau de signification de 0.05. Ce qui prouve que le modèle de l'étude est accepté à un niveau de signification de 5 %.

VI.2.4.2. Test de nullité d'un paramètre du modèle (Test de Student)

Depuis VI-6 on constate que 04 facteurs sur 11 ont un t-value supérieur à 1.96 donc ils ont un impact significatif sur le modèle, et les autres 07 facteurs qui ont un t-value inférieur à 1.96 sont des facteurs non significatifs qu'on peut les écarter du modèle, néanmoins, dans cette étude, on va prendre en considération tous les 11 facteurs.

D'autre constat qu'on peut relever de ce tableau, c'est que la somme des valeurs de t-student, donne une valeur positive (+9.31), ce qui prouve qu'il y a une relation significative entre la productivité et les 11 facteurs d'étude (Diab et al., 2017; A. Otmani, 2020).

Depuis ce qui précède, la relation entre les variables explicatives (les facteurs) et la réponse Y (la productivité), peut exprimer par une équation linéaire estimative comme suit :

La productivité réalisée transformée =

$$\begin{aligned}
 &+3.225 + 0.145(\text{Nombre des heures de travail}) + \\
 &0.053(\text{Nombre des manœuvres}) + 0.080(\text{Nombre des ouvriers}) + \\
 &0.083(\text{Expérience minimale transformée}) + \\
 &0.024(\text{Expérience maximale transformée}) - 0.002(\text{Etat de coffrage}) - \\
 &0.126(\text{Type de Motivation transformé}) - \\
 &0.869(\text{Surface de plancher transformée}) + 0.003(\text{Température}) - \\
 &0.009(\text{Salaire des manœuvres transformé}) - 0.010(\text{Salaire ouvrier transformé})
 \end{aligned}
 \tag{VI-2}$$

VI.2.5. Analyse des résidus

Afin de statuer sur la fiabilité de l'équation VI-2, dans la phase de développement les résidus de cette dernière doivent vérifier les hypothèses (conditions) suivantes :

VI.2.5.1. Normalité des résidus

Contrairement à l'analyse des données, et pour le but de diversifier les tests utilisés dans cette étude, nous allons tester la normalité et l'homogénéité (homoscédasticité) des résidus par des tests graphiques.

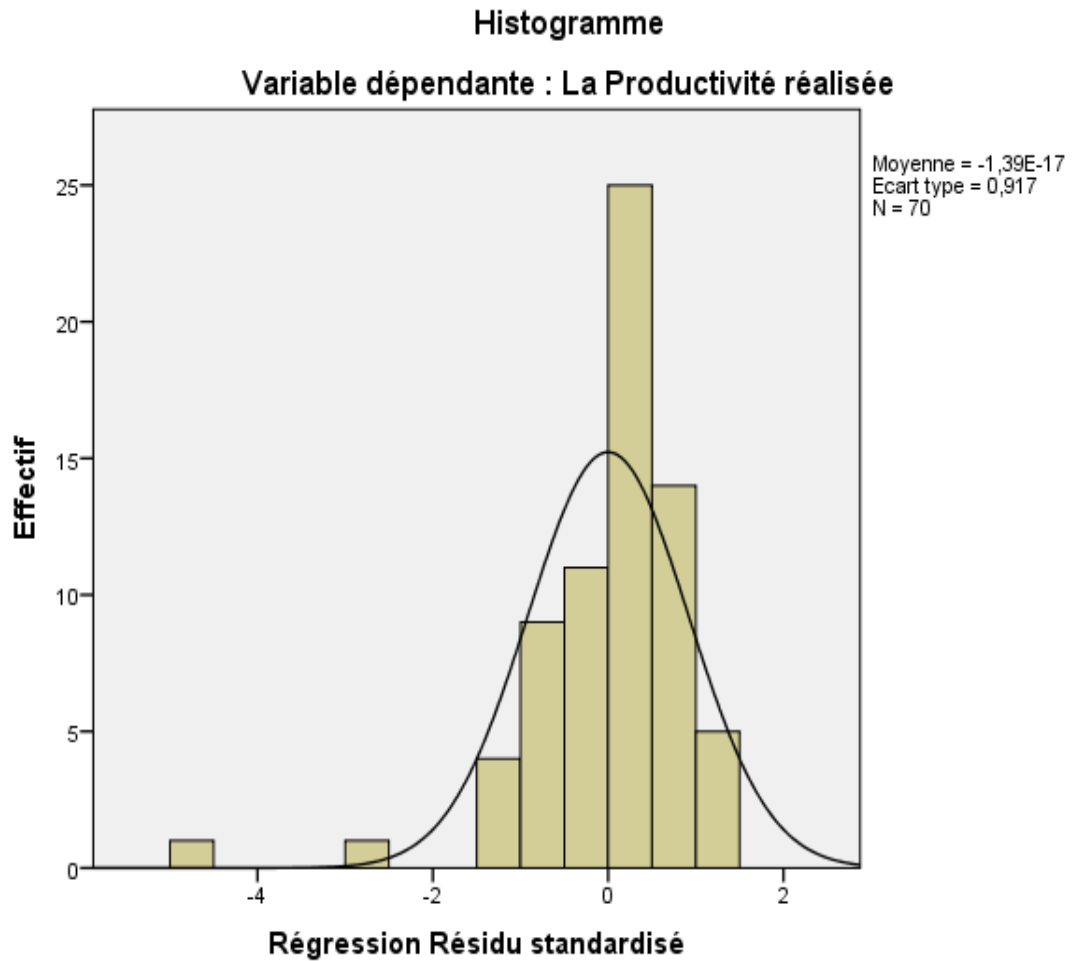


Figure VI-1 Histogramme de fréquence des résidus

Diagramme gaussien P-P de régression de Résidu standardisé

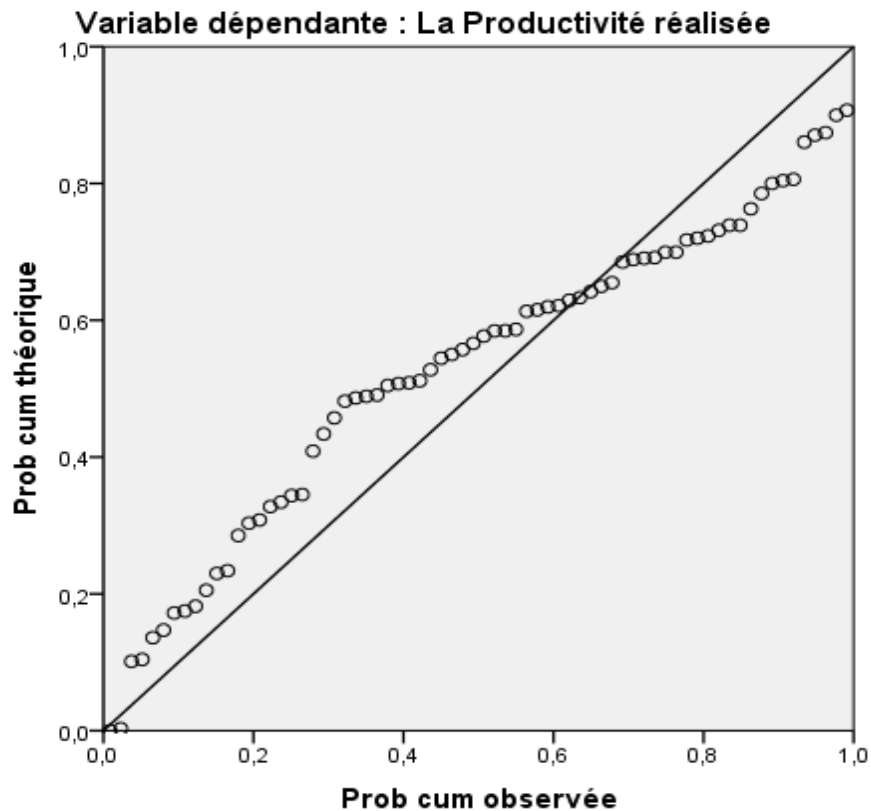


Figure VI-2 Diagramme de répartition Gaussienne des résidus

D'après la figure VI-1 on constate que les résidus suivent une distribution gaussienne, ainsi que la figure VI-2 illustre que ces derniers suivent une ligne droite et presque superposés à la droite des quantiles, en conséquence, on peut dire que les résidus suivent une loi normale.

VI.2.5.2. Homoscédasticité (variance constante) et linéarité des résidus

Le test graphique du nuage de points des sorties du modèle par rapport aux résidus est utilisé afin de statuer sur l'homogénéité et la linéarité des résidus comme le montre la figure suivante :

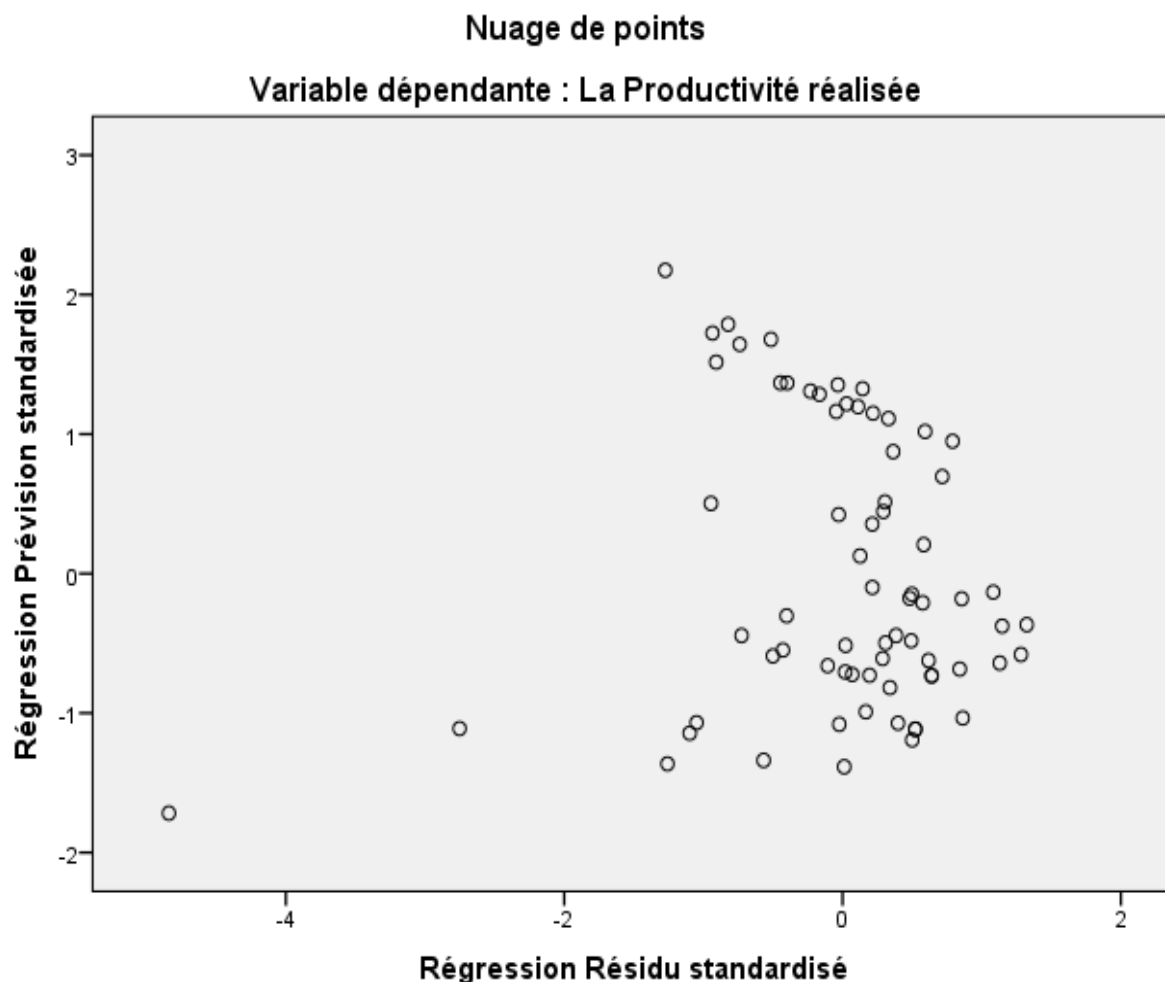


Figure VI-3 Distribution des résidus par rapport à la réponse estimée

De la figure VI-3, on peut voir que le nuage des résidus présente une forme conique située entre $[-2 ; +2]$, et qu'ils sont distribués autour de 0, ce qui prouve que les résidus du modèle sont linéaires et de variance constante.

VI.2.5.3. Test de Durbin-Watson d'autocorrélation des erreurs

Ce test est utilisé pour vérifier l'indépendance des résidus (les corrélations en série entre les résidus).

Tableau VI-8 Résultat du Durbin-Watson

Model	Durbin-Watson
1	1.942

Un résultat proche de 02 du Durbin-Watson est observé depuis le tableau VI-8, chose qui prouve que les résidus sont indépendants.

VI.2.5.4. Vérification de la Multicolinéarité (par le coefficient VIF)

Pour la vérification de la multicolinéarité des résidus, le coefficient VIF est utilisé et résultats sont présentés dans le tableau suivant :

Tableau VI-9 Vérification de la multicolinéarité par VIF

Model	Statistiques de colinéarité	
	Tolérance	VIF
Nombre des heures de travail	0,208	4,814
Nombre des manœuvres	0,782	1,279
Nombre des ouvriers	0,429	2,332
Expérience minimale transformée	0,323	3,099
Expérience maximale transformée	0,210	4,757
Etat des matériels de construction	0,646	1,549
Type de Motivation transformé	0,230	4,346
Quantité de travail réalisée transformée	0,247	4,045
Température	0,610	1,639
Salaire des manœuvres transformé	0,840	1,190
Salaire ouvrier transformé	0,713	1,402

Des faibles valeurs de VIF inférieures à 5 sont obtenues, indique le manque de la multicolinéarité des résidus du modèle développé.

VI.2.6. Evaluation du modèle développé

VI.2.6.1. Résultat obtenu dans la phase de développement

Après tous les tests réalisés sur les données et les résidus, nous avons fini de proposer l'équation VI-2, comme équation d'estimation de la productivité du travail dans les projets de construction Algérien sous l'influence des 11 facteurs.

Les résultats de la productivité issus de l'utilisation de cette équation, dans la phase d'apprentissage (développement), sont présentés dans le tableau suivant :

Tableau VI-10 Résultat du modèle de la régression multiple dans la phase d'apprentissage (développement)

N° Ordre	Sorties du modèle transformées	Résultats réelles transformée	Retour à l'origine Sorties du modèle	Retour à l'origine Résultats réelles
1	1,334	1,213	0,743	1,111
2	0,554	0,586	2,695	2,642
3	0,536	0,566	2,725	2,676
4	0,550	0,445	2,703	2,862
5	0,937	0,952	1,878	1,840
6	0,703	0,733	2,418	2,355
7	0,645	0,686	2,533	2,453
8	0,917	0,903	1,929	1,963
9	0,672	0,710	2,480	2,403
10	1,223	1,135	1,080	1,338
11	0,674	0,686	2,476	2,453
12	0,571	0,633	2,668	2,556
13	0,489	0,480	2,798	2,812
14	1,092	1,076	1,460	1,506
15	1,141	1,091	1,322	1,465
16	0,655	0,714	2,514	2,397
17	0,730	0,829	2,363	2,141
18	0,679	0,775	2,466	2,264
19	1,176	1,088	1,220	1,472
20	0,645	0,650	2,532	2,523
21	0,869	0,903	2,046	1,963
22	0,702	0,714	2,420	2,397
23	0,566	0,467	2,676	2,831
24	1,131	1,130	1,350	1,352
25	0,714	0,733	2,396	2,355
26	0,696	0,686	2,433	2,453
27	0,937	0,846	1,877	2,103
28	0,497	0,380	2,786	2,944
29	0,785	0,812	2,243	2,181

30	0,797	0,800	2,216	2,208
31	0,554	0,586	2,695	2,642
32	0,743	0,701	2,335	2,422
33	0,985	1,029	1,754	1,635
34	1,140	1,122	1,322	1,376
35	0,653	0,641	2,517	2,540
36	0,668	0,748	2,488	2,324
37	1,241	1,159	1,026	1,270
38	1,027	1,042	1,639	1,599
39	1,218	1,159	1,095	1,270
40	0,665	0,641	2,495	2,540
41	1,208	1,132	1,124	1,348
42	1,103	1,097	1,431	1,445
43	0,679	0,625	2,468	2,571
44	1,123	1,095	1,372	1,451
45	0,626	0,641	2,568	2,540
46	1,091	1,095	1,464	1,451
47	1,057	1,095	1,559	1,451
48	1,127	1,095	1,362	1,451
49	1,106	1,095	1,420	1,451
50	1,142	1,095	1,318	1,451
51	0,692	0,641	2,442	2,540
52	0,650	0,641	2,524	2,540
53	0,556	0,316	2,693	3,013
54	0,899	0,907	1,972	1,953
55	0,564	0,549	2,680	2,704
56	0,923	0,935	1,914	1,883
57	0,770	0,804	2,276	2,200
58	0,730	0,812	2,362	2,181
59	0,779	0,804	2,256	2,200
60	0,716	0,641	2,391	2,540
61	1,046	1,095	1,589	1,451
62	0,586	0,586	2,642	2,642

63	0,779	0,834	2,257	2,130
64	0,849	0,846	2,095	2,103
65	0,646	0,686	2,531	2,453
66	0,566	0,586	2,675	2,642
67	0,498	0,443	2,784	2,864
68	0,412	0,000	2,904	3,170
69	1,083	1,095	1,485	1,451
70	0,790	0,864	2,232	2,058

VI.2.6.2. Evaluation des performances de la phase de développement

Afin d'évaluer la performance de l'équation d'estimation durant la phase de développement, les mesures de MSE, MAPE et AA% sont calculées et leurs résultats sont présentés dans le tableau ci-dessous :

Tableau VI-11 Performance du modèle de la régression Multiple dans la phase de développement (apprentissage)

MSE (données transformées)	MSE (données Réelles)	MAPE % (données transformées)	MAPE % (données Réelles)	AA % (données transformées)	AA % (données Réelles)
0,0058	0,016	6,41	5.21	93.58	94.79

D'après le tableau VI-11, on constate de fortes performances d'estimation enregistrées dans cette phase, avec une précision attendue 94.79 %, l'erreur quadratique moyenne, et le MAPE sont aussi donnés de très bons résultats, chose qui prouve la puissance d'estimation de l'équation développée.

Autre Les mesures de la performance, discutées en haut, la qualité d'ajustement dans la phase de développement est aussi utilisée, et les résultats sont montrés dans le tableau et la figure suivants :

Tableau VI-12 Qualité d'ajustement du modèle de la RLM dans la phase de développement

Récapitulatif de modèle					
Modèle	R	R-deux	R-deux ajusté	Erreur standard de l'estimation	Durbin-Watson
1	0,952	0,906	0,888	0,082792	1,942

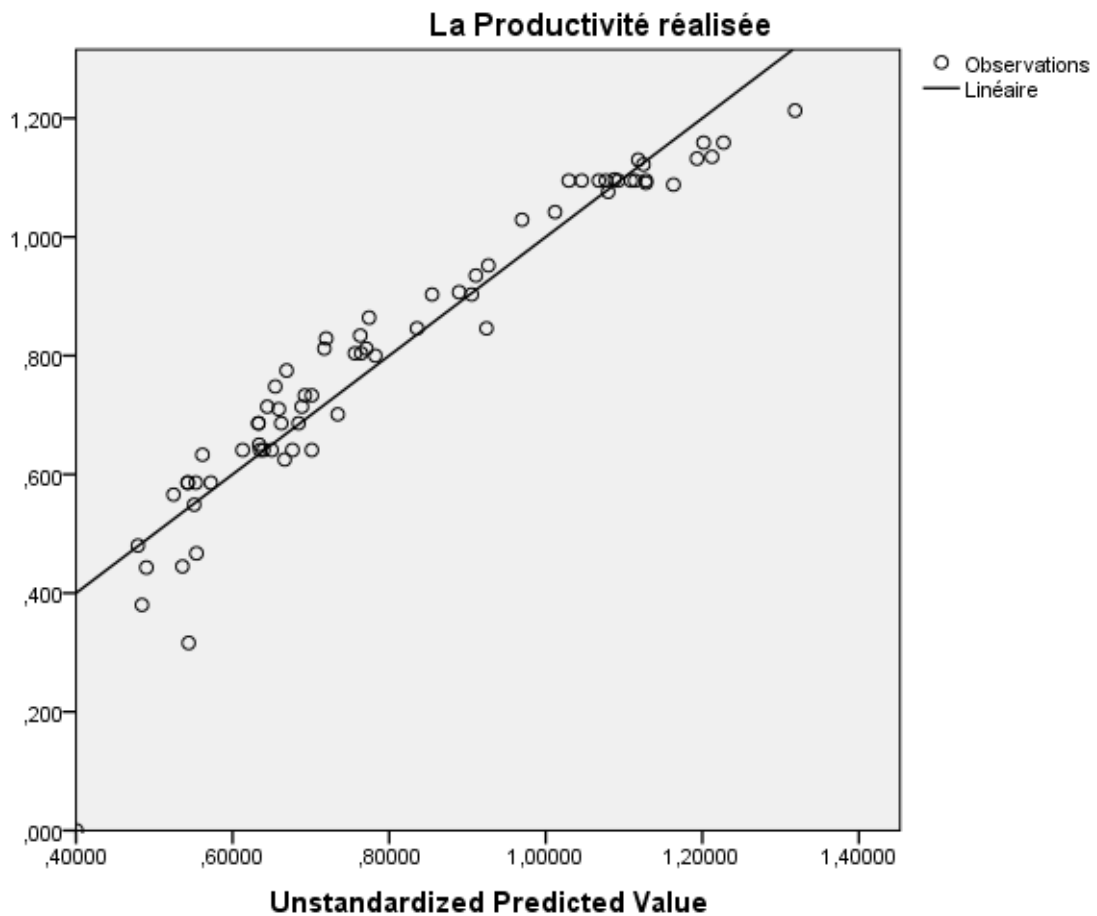


Figure VI-4 Droite de régression dans la phase d'apprentissage du modèle de la RLM

La comparaison entre les résultats réels et les sorties résultant du modèle développé de la Régression Linéaire Multiple est montrée graphiquement comme suite :

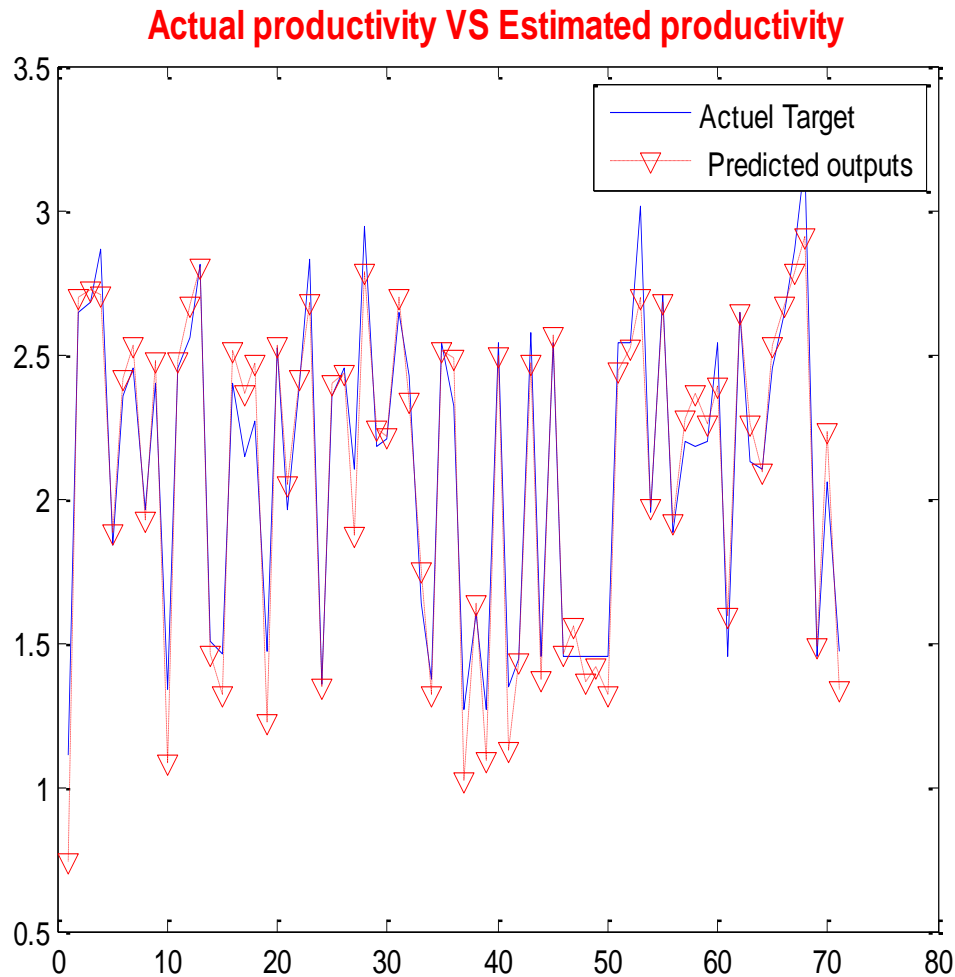


Figure VI-5 Les sorties du modèle de la RLM VS les résultats réels de la productivité dans la phase d'apprentissage

Depuis la figure VI-5 on remarque que les sorties du modèle sont superposées aux résultats réels de la productivité, ce qui indique que le modèle développé estime bien la productivité dans la phase d'apprentissage.

VI.2.7. Test du modèle développé par des nouvelles données d'entrées

VI.2.7.1. Résultat obtenu dans la phase de test

Cette phase consiste à tester la capacité d'estimation du modèle vis-à-vis des nouvelles données, en introduisant au modèle développé les 30% des données collectées qui sont réservées à l'avance, et observer les résultats obtenus par la suite.

Le tableau suivant présente les résultats d'estimation obtenus dans la phase du test :

Tableau VI-13 Résultat du modèle de la régression multiple dans la phase de test

N° Ordre	Sorties du modèle transformées	Résultats réelles transformés	Retour à l'origine Sorties du modèle	Retour à l'origine Résultats réelles
1	0,604	0,593	2,610	2,628
2	0,648	0,714	2,527	2,397
3	0,640	0,690	2,543	2,444
4	0,732	0,740	2,357	2,342
5	0,531	0,541	2,733	2,717
6	0,665	0,641	2,495	2,540
7	1,184	1,095	1,195	1,451
8	1,050	1,064	1,578	1,539
9	1,125	1,095	1,367	1,451
10	1,065	1,095	1,535	1,451
11	0,747	0,775	2,325	2,264
12	1,393	1,145	0,560	1,308
13	0,517	0,528	2,756	2,739
14	0,873	0,871	2,037	2,043
15	0,797	0,880	2,214	2,019
16	0,624	0,641	2,572	2,540
17	0,662	0,686	2,500	2,453
18	1,115	1,050	1,396	1,577
19	1,121	1,050	1,379	1,577
20	0,927	0,952	1,904	1,840
21	1,096	1,095	1,450	1,451
22	1,141	1,095	1,321	1,451
23	1,168	1,091	1,242	1,465
24	1,078	1,087	1,499	1,475
25	0,686	0,714	2,454	2,397
26	0,736	0,812	2,349	2,181
27	0,752	0,812	2,315	2,181
28	1,087	1,095	1,474	1,451
29	1,079	1,095	1,496	1,451
30	0,743	0,805	2,334	2,197

VI.2.7.2. Evaluation des performances de la phase de test

Les mêmes mesures des erreurs utilisées dans la phase de développement, sont aussi utilisées dans la phase du test afin d'évaluer les performances d'estimation du modèle développé dans cette phase.

Tableau VI-14 Performance du modèle de la régression Multiple dans la phase de test

MSE (données transformées)	MSE (données Réelles)	MAPE % (données transformées)	MAPE % (données Réelles)	AA % (données transformées)	AA % (données Réelles)
0.004	0.031	4.69	6.64	95.30	93.36

Les résultats présentés dans le tableau ci-dessus montrent des performances élevées quant à l'estimation de la productivité du travail dans les projets de construction Algériens, durant la phase du test avec une précision allant jusqu'à 93.36%, ainsi que des valeurs de MSE= 0.031 et MAPE=6.64, sont enregistrées.

On a enregistré aussi dans cette phase une bonne qualité d'ajustement avec un coefficient de corrélation qui atteint 0.963 (Figure VI-6)

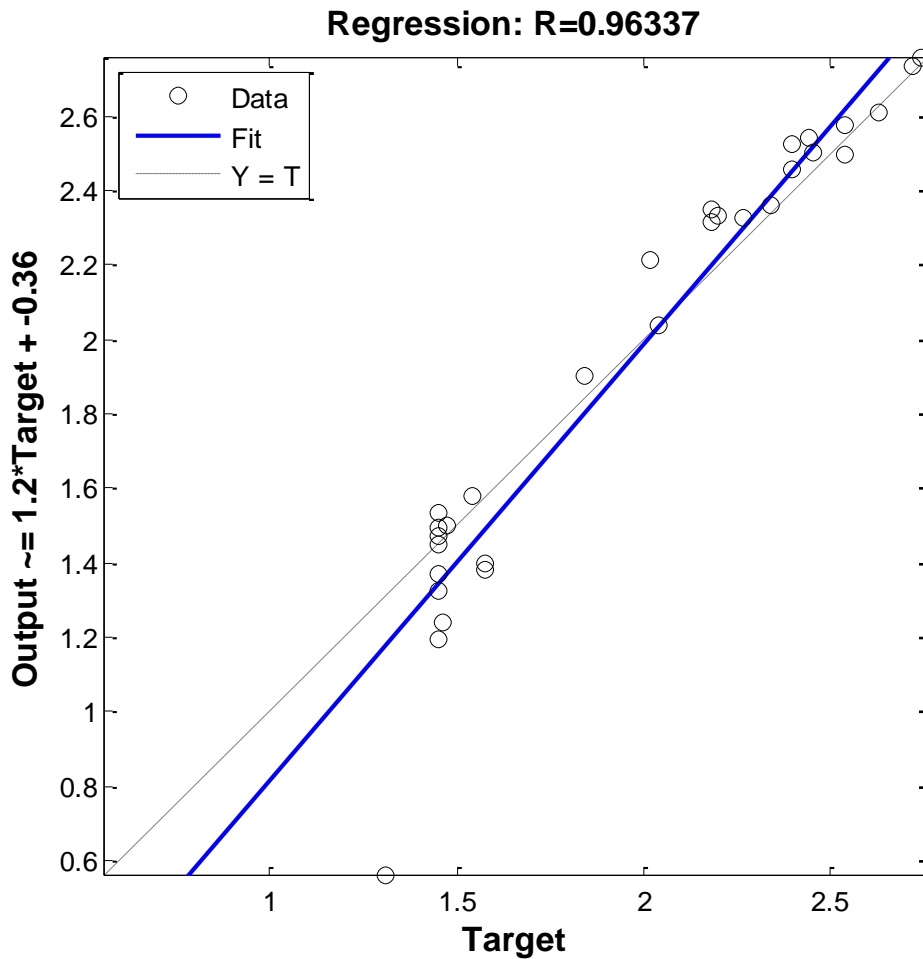


Figure VI-6 Droite de régression dans la phase test du modèle RLM

La comparaison entre les sorties du modèle et les résultats réels de la productivité dans la phase de test ne montre aucune différence remarquable entre les deux, ce qui prouve la force d'estimation du modèle développé (figure VI-7).

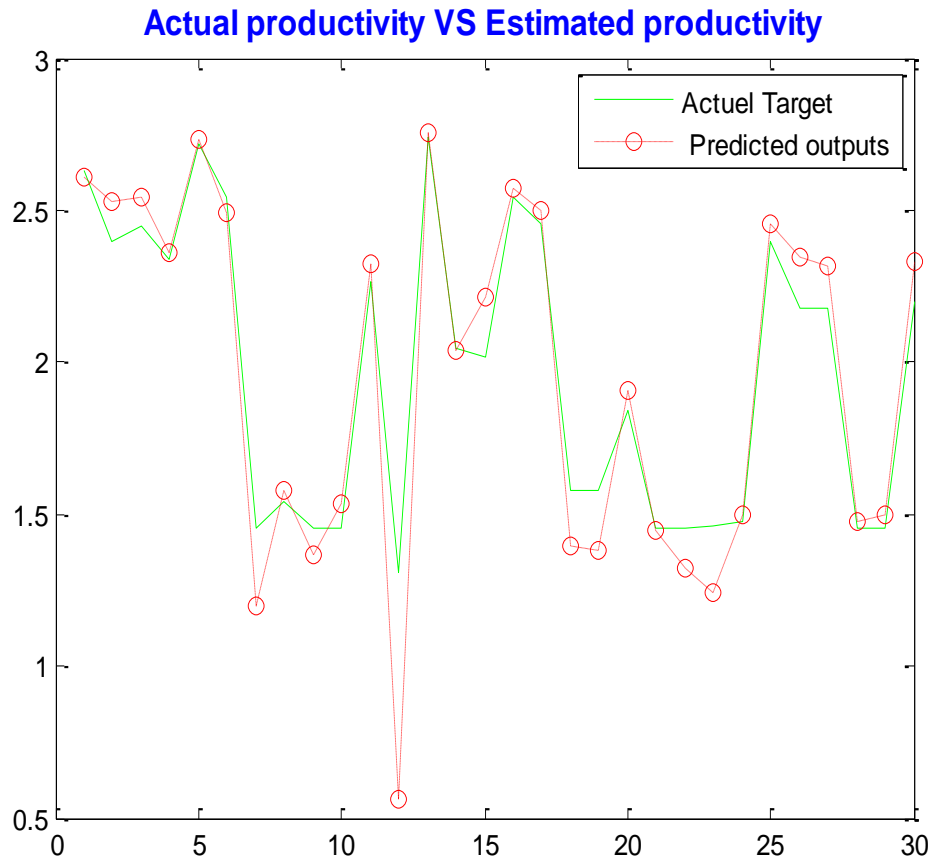


Figure VI-7 Les sorties du modèle de la RLM VS les résultats réels de la productivité dans la phase de test

VI.3. Conclusion.

Un modèle mathématique paramétrique basé sur la méthode de la Régression Linéaire Multiple est développé dans ce chapitre pour le but de prédire la productivité du travail dans les projets de construction Algérien.

Les paramètres du modèle sont développés et testés via le logiciel de calcul statistique (SPSS), et les résultats obtenus ont montré une forte capacité de prédiction allant jusqu'à 93.36 % dans la phase de test.

VI.4. Discussions des résultats des deux Méthodes (modèle RNA et modèle Paramétrique)

Dans le but d'améliorer la productivité du travail dans les projets de construction en Algérie, deux modèles basés sur la méthode des Réseaux de Neurones Artificiels et la méthode Paramétrique par la Régression Linéaire Multiple ont été développés pour estimer le taux de cette productivité sous l'effet de 11 facteurs qui l'influent.

Afin de choisir le meilleur modèle, les résultats des deux modèles dans la phase de test sont comparés par les tableaux et figures comme suite

Tableau VI-15 Résultats de la productivité par RNA et RLM dans la phase de test

N° Ordre	Résultats réels de la productivité	Sorties du modèle de la RLM	Sorties du modèle de des RNA
1	2,628	2,610	2,63
2	2,397	2,527	2,40
3	2,444	2,543	2,44
4	2,342	2,357	2,34
5	2,717	2,733	2,72
6	2,540	2,495	2,54
7	1,451	1,195	1,45
8	1,539	1,578	1,54
9	1,451	1,367	1,45
10	1,451	1,535	1,45
11	2,264	2,325	2,26
12	1,308	0,560	1,31
13	2,739	2,756	2,74
14	2,043	2,037	2,04
15	2,019	2,214	2,02
16	2,540	2,572	2,54
17	2,453	2,500	2,45
18	1,577	1,396	1,58
19	1,577	1,379	1,58
20	1,840	1,904	1,84
21	1,451	1,450	1,45
22	1,451	1,321	1,45
23	1,465	1,242	1,46
24	1,475	1,499	1,48
25	2,397	2,454	2,40

26	2,181	2,349	2,18
27	2,181	2,315	2,18
28	1,451	1,474	1,45
29	1,451	1,496	1,45
30	2,197	2,334	2,20

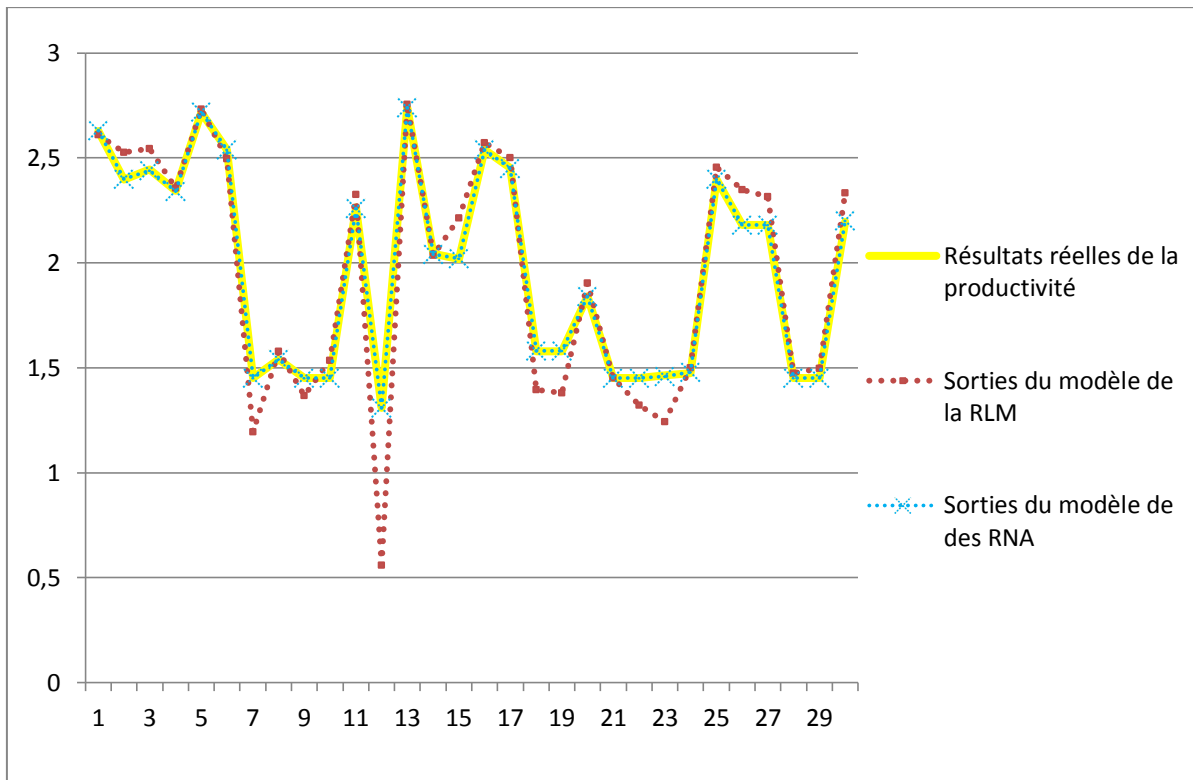


Figure VI-8 Résultats de la productivité des deux modèles dans la phase du test

Tableau VI-16 Performances des deux modèles dans la phase du test

Performances	Modèle du RLM	Modèle de la RNA
MSE	0.031	0.000906
MAPE	6.64	1.224
AA%	93.36	98.776

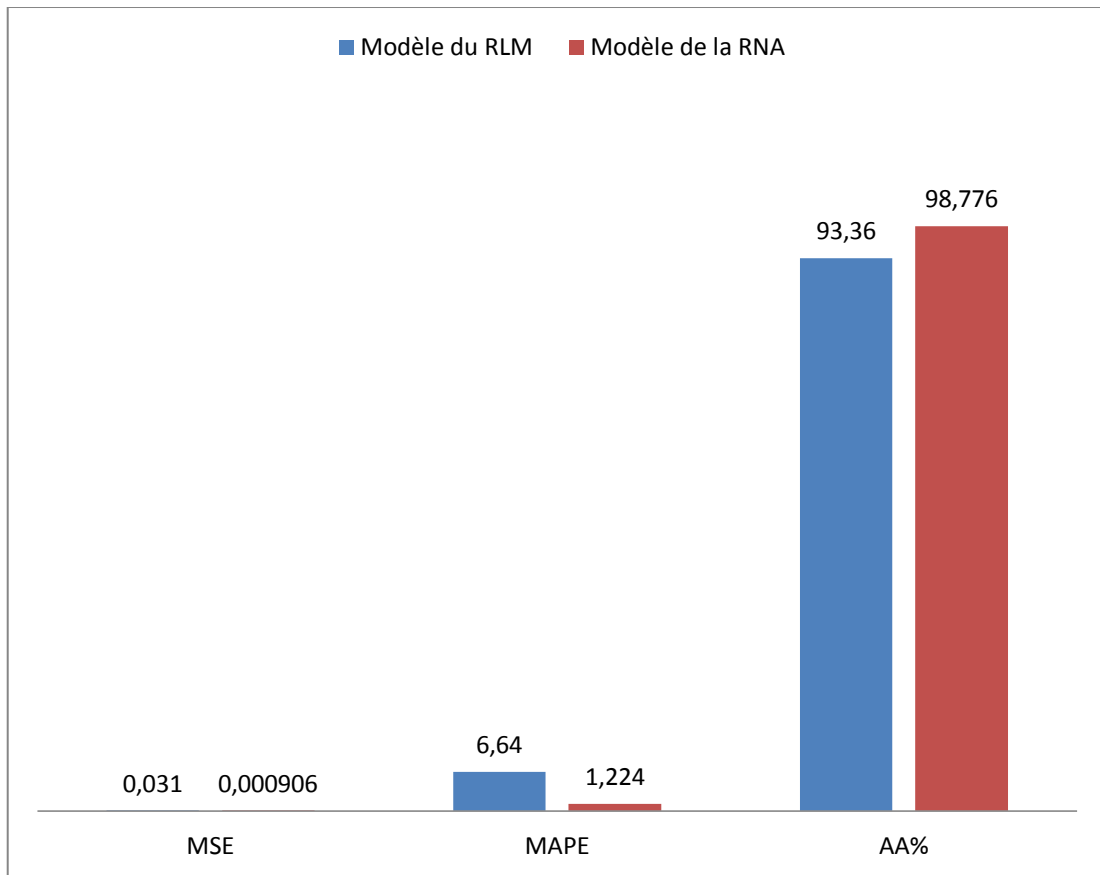


Figure VI-9 Présentation des performances des deux modèles dans la phase du test

On remarque depuis les tableaux et figures précédents que les deux modèles ont prouvé leurs forces quant à l'estimation et à la prédiction du taux de la productivité du travail dans les projets de constructions Algériens, et donnent des résultats très proches de la réalité avec une précision élevée. Néanmoins, on constate une légère différence dans les performances des deux modèles, et que le modèle de Réseau Neurones Artificiels présente des performances mieux que celles produites par le modèle du RLM.

Chapitre VII

Analyse de sensibilité

VII.1. Introduction.

L'analyse de sensibilité est une étape principale dans l'interprétation des données ; il donne une description générale de l'influence des variables en déterminant leur impact sur les résultats finaux. Plusieurs algorithmes d'analyse de sensibilité ont été fondés, qui servent à classer chaque entrée en donnant son importance relative sur la sortie.

Dans ce chapitre, on va présenter les différentes méthodes de l'analyse de sensibilité, et puis on va analyser et évaluer l'effet des 11 facteurs (introduits dans notre modèle) sur la productivité du travail.

VII.2. Différentes méthodes d'analyse de sensibilité

VII.2.1. Algorithme de Garson

proposé d'abord par (Garson, 1991), puis amélioré par (Goh, 1995), il consiste à partitionner les poids de connexion des couches d'entrée, cachée et de sortie en éléments liés à chaque variable d'entrée, en utilisant les valeurs absolues des matrices de poids, dans l'ordre pour définir l'importance relative (RIx) de chaque entrée en pourcentage, plus le (RIx est élevé), plus l'entrée correspondante a un fort impact sur la sortie, cet algorithme n'est appliqué qu'en ANN avec une couche cachée et une variable de sortie (Bhattacharjee & Tollner, 2016). Pour une description plus détaillée, voir (Ghanizadeh et al., 2020; Goh, 1995).

VII.2.2. Algorithme d'Olden

Cet algorithme basé sur le calcul du produit brut des poids d'entrée-neurones cachés-sortie, en tenant compte du signe de chaque poids, puis additionne le produit brut résultant, notez que le signe des poids de connexion représente une augmentation ou une diminution de la valeur de sortie, cet algorithme peut être utilisé dans des réseaux de neurones à sorties multiples (Bhattacharjee & Tollner, 2016; Olden & Jackson, 2002), ont donné plus de détails.

VII.2.3. Dérivée partielle

Dans cette méthode, la dérivée partielle de la sortie par rapport à l'entrée est calculée afin d'obtenir l'importance relative de chaque variable d'entrée.

Cette approche donne deux résultats, le premier relatif au profil de sortie par rapport à chaque entrée, et le second présente l'importance relative de chaque variable d'entrée (Gevrey et al., 2003).

VII.2.4. Méthode de profil

Appelée également méthode Lek, par rapport à son premier proposant (Lek, S., Belaud, A., Baran, P., Dimopoulos, I., & Delacoste, 1996), elle repose sur la division de chaque variable d'entrée en une plage égale appelée échelle entre son minimum et son maximum.

La variable étudiée est ensuite distribuée selon les points d'échelle choisis en fixant les variables d'entrée restantes successivement à son premier quartile minimum, sa médiane, son troisième quartile et son maximum ; résultant en cinq valeurs de sortie pour chaque point d'échelle et sa médiane qui seront prises en compte.

(Gevrey et al., 2003; Zhang et al., 2018), illustrent l'application de la méthode des profils

VII.2.5. Perturbation d'entrée

cette approche vise à évaluer le changement d'erreur quadratique moyenne du modèle en ajoutant un terme de bruit à l'entrée correspondante tout en gardant les autres entrées inchangées, le bruit prend une plage de ratio croissant (par exemple : [1 % 50 %] augmente par pas de 10 %) (Bai et al., 2011), et ajouté à l'entrée selon la relation ci-dessous :

$$\text{Perturbation} = X_i * \delta ;$$

$$X_i = X_i + \text{Perturbation} ;$$

Où δ est le ratio (ex : 10%).

L'importance relative de chaque entrée peut alors être obtenue en observant le changement de MSE, l'entrée avec le plus de changement, est celle qui a un fort impact sur la sortie, Voir (Bai et al., 2011; Gevrey et al., 2003).

VII.2.6. Méthode pas à pas

Considérée comme l'une des méthodes les plus importantes dans l'analyse de sensibilité (Shojaeefard et al., 2013), elle consiste à ajouter ou à exclure une entrée étape par étape et à observer le changement sur MSE, l'importance relative de la variable est classée en fonction

des plus petits changements de MSE, dans l'approche vers l'avant ou aux changements les plus importants d'un MSE, dans l'approche vers l'arrière, voir (Gevrey et al., 2003) pour plus de descriptions

VII.2.7. Deep connection weight

Comme nous l'avons vu précédemment, l'approche du poids de connexion ne peut être utilisée que dans le réseau avec une couche cachée, pour cela (O'Donoghue et al., 2017), ont proposé un algorithme de poids de connexion profond, qui sert à généraliser le poids de connexion traditionnel aux réseaux de neurones artificiels multicouches cachés. L'algorithme consiste à calculer une matrice de score qui représente le résultat du produit scalaire cumulé de toutes les matrices de poids (Equation VII-1).

$$\boldsymbol{\varphi} = \prod_{k=1}^L \mathbf{W}_k \quad (\text{VII-1})$$

Où $\boldsymbol{\varphi}$ représente la matrice de score ayant une dimension (n * m), correspondant respectivement au nombre d'entrées et de sorties \mathbf{W}_k correspond à la matrice de poids de la $k^{\text{ième}}$ couche et L est le nombre de couches. L'importance relative de chaque variable d'entrée peut être obtenue en divisant chaque élément de la matrice des scores dans sa valeur absolue par la somme absolue des éléments de chaque colonne correspondante (Martins et al., 2020).

L'équation VII-2 montre la représentation mathématique

$$RIx_{ij} = \frac{\varphi_{ij}}{\sum_{i=1}^n \varphi_{ij}} \quad (\text{VII-2})$$

RIx est l'importance relative

φ_{ij} correspond à la $i^{\text{ème}}$ ligne et à la $j^{\text{ième}}$ colonne de l'élément de la matrice des scores.

Dans cette étude, nous nous concentrons sur l'algorithme de poids de connexion profonde pour calculer l'importance relative afin d'évaluer l'influence de chaque entrée sur la sortie, les résultats montrent que la Quantité de travail réalisée est le facteur qui a le plus d'impact sur la sortie avec le « Importance relative » de 57,71 %, suivi par le Nombre des heures de travail ayant un RIX de 16,60 %, le Nombre des ouvriers a été classé le troisième facteur influent avec 15,94 %, les quatrième et cinquième influents étaient le Nombre des manœuvres et le Salaire ouvrier ayant 5,50 % et 0,88 % successivement.

Les résultats sont illustrés dans le tableau et la figure suivants :

Tableau VII-1 L'importance de chaque facteur sur la productivité

Facteurs	RIx %
Nombre des heures de travail	16.60
Nombre des manœuvres	5.50
Nombre des ouvriers	15.94
Expérience minimale	0.74
Expérience maximale	0.16
Etat de matériel de construction	0.83
Type de Motivation	0.87
Quantité de travail réalisée	57.71
Température	0.49
Salaire des manœuvres	0.23
Salaire des ouvriers	0.88

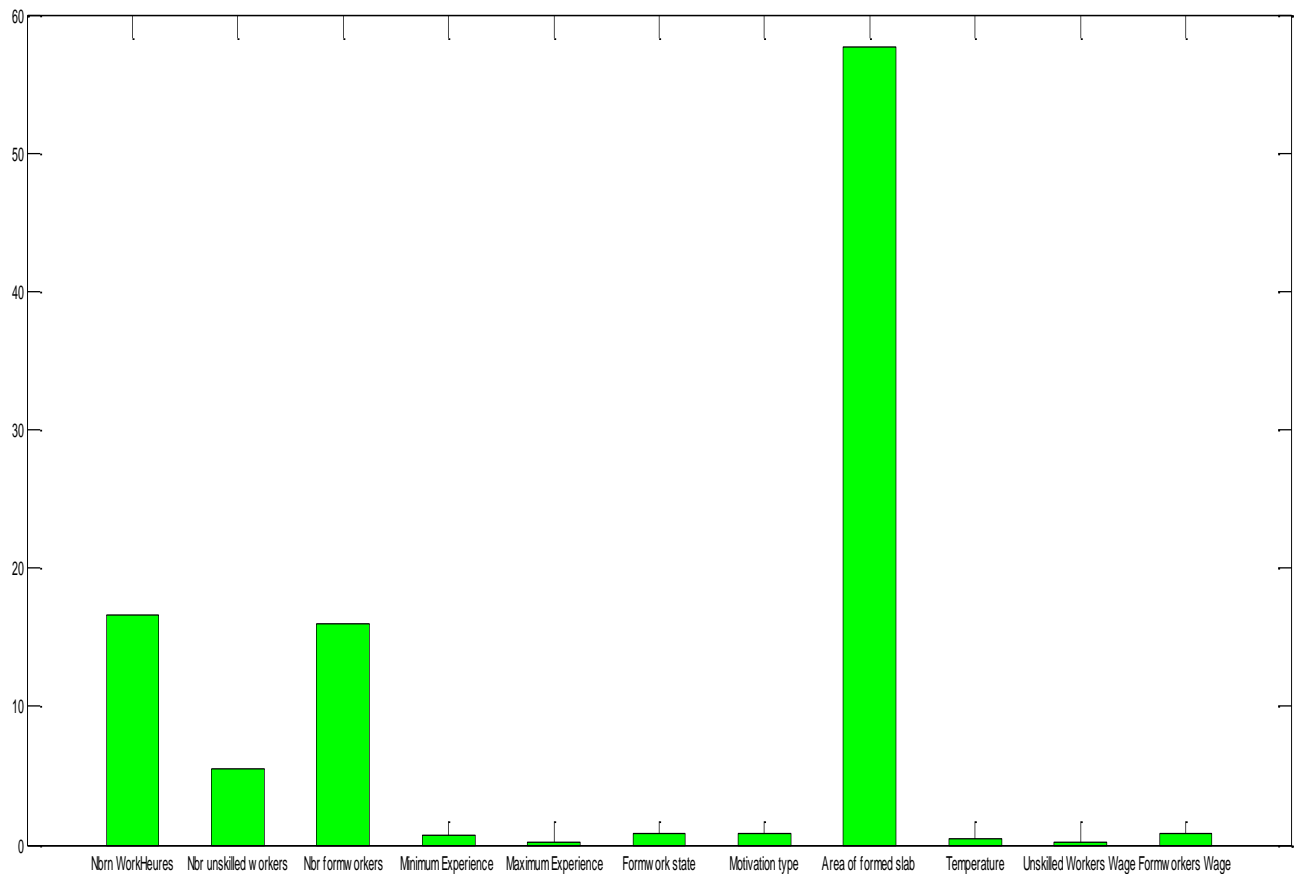


Figure VII-1 L'importance relatif de chaque facteur

VII.3. Conclusion

Ce chapitre définit différentes méthodes de l'analyse de sensibilité, et une de ces dernières « Deep connection weight » est utilisée afin d'évaluer l'effet des 11 facteurs (introduits dans notre modèle) sur la productivité du travail.

Les résultats nous donnent un aperçu sur la participation de chaque facteur sur le résultat de la productivité estimée.

Conclusion générale

Conclusion générale

Cette thèse a pour but d'améliorer et d'augmenter la productivité du travail dans les projets de construction Algérien, et afin d'atteindre l'objectif voulu, une introduction au domaine de la productivité consacrée aux différentes définitions et techniques de mesure de cette dernière, est présentée en premier lieu, puis on a revu dans le chapitre qui suit, les différents travaux précédents dans l'identification des facteurs influençant la productivité, ainsi que les différentes méthodes d'estimation et de prédiction de la productivité pour cadrer notre recherche et englober toutes les informations en relation avec le thème traité.

L'aspect théorique des méthodes utilisées dans cette thèse a été abordé dans le chapitre 03, en présentant les différents types, les Algorithmes utilisés et le mode d'application ...etc. des deux méthodes, afin de nous permettre de mieux comprendre et gérer l'outil utilisé. Une analyse des résultats de la productivité classique des différentes tâches (coffrage et ferrailage) dans quatre différents projets algérien a été faite dans le chapitre 04, et les résultats de cette analyse montrent différents taux de productivité d'une même tâche dans le même chantier et dans les différents chantiers, sur cette base un questionnaire a été conçu, distribué et analysé, afin d'identifier les facteurs qui influencent la productivité dans les projets de construction Algérien, et à la fin de ce chapitre un nombre de 11 facteurs jugés affecter la productivité sont choisis et qui vont être utilisés comme entrées dans les chapitres qui suivent.

Dans les chapitres 05 et 06, on a développé deux modèles par la méthode des Réseaux de Neurones Artificiels et la méthode paramétrique par la Régression Linéaire Multiple successivement, afin d'estimer le taux de la productivité du travail, les deux modèles ont montré de bonnes performances dans la phase d'apprentissage ainsi que dans la phase du test, et des résultats proches de la productivité réelle sont enregistrés, en conséquence les modèles développés peuvent être utilisés avec fiabilité dans la prédiction du taux de la productivité du travail dans les projets de construction Algérien, sous les 11 facteurs résultant du chapitre 04.

Dans la dernière étape de cette thèse, on a fait une étude de sensibilité, en utilisant les poids obtenus dans le modèle du RNA développé, dans le but de classer les facteurs qui affectent le plus la productivité, ce qui nous permettra de comprendre mieux l'influence de chaque facteur sur cette dernière.

- **Recommandations**

Notre étude est basée sur une base de données collectée des différents chantiers en Algérie dans la région Nord-Est et la région des Aurès, une totalité de 04 chantiers et 04 villes, avec un nombre de 101 données.

- La diversité dans les régions et chantiers ainsi que l'augmentation du nombre de la base des données peuvent améliorer les résultats obtenus.
- La transformation de l'intelligence artificielle vers l'apprentissage approfondi (deep learning) est recommandée.
- Les méthodes d'estimation utilisées dans cette étude peuvent être utilisées dans d'autres domaines de la gestion et management des projets tels que la gestion des délais, gestion des couts.

Bibliographie

- Adebowale, O., & Smallwood, J. (2020). Contractors' perceptions of factors affecting construction labour productivity in South Africa. *Journal of Contemporary Management*, 17(2), 326–346.
- Adrian, J. (1987). *Construction productivity improvement* (T. Netherlands (ed.); Livre). Elsevier Science Publishing.
- Al-Zwainy, F. M. S., Abdulmajeed, M. H., & Aljumaily, H. S. M. (2013). Using Multivariable Linear Regression Technique for Modeling Productivity Construction in Iraq. *Open Journal of Civil Engineering*, 03(03), 127–135.
<https://doi.org/10.4236/ojce.2013.33015>
- Al-Zwainy, F. M. S., Rasheed, H. A., & Ibraheem, H. F. (2012). Development of the construction productivity estimation model using artificial neural network for finishing works for floors with marble. *ARPJ Journal of Engineering and Applied Sciences*, 7(6), 714–722.
- Alaghbari, W., Al-Sakkaf, A. A., & Sultan, B. (2017). Factors affecting construction labour productivity in Yemen. *International Journal of Construction Management*, 19(1), 79–91.
<https://doi.org/10.1080/15623599.2017.1382091>
- Alaghbari, W., Al-Sakkaf, A. A., & Sultan, B. (2019). Factors affecting construction labour productivity in Yemen. *International Journal of Construction Management*, 19(1), 79–91.
<https://doi.org/10.1080/15623599.2017.1382091>
- Alfeld, L. E. (1988). *Construction productivity: on-site measurement and management*. Livre; McGraw-Hill.
- Ammar, M. Y. (2007). *Mise en oeuvre de réseaux de neurones pour la modélisation de cinétiques réactionnelles en vue de la transposition batch/continu*, Thèse de doctorat; Ecole Nationale d'Ingénieurs de Sfax, Tunisie. <http://ethesis.inp-toulouse.fr/archive/00000607/01/ammam.pdf>
- Aswed, G. K. (2016). Productivity Estimation Model for Bricklayer in Construction Projects Using Neural Network. *Al-Qadisiyah Journal For Engineering Sciences*, 9(2), 183–199.

- Bai, R., Zhang, D., & Jia, H. (2011). Factor sensitivity analysis with neural network simulation based on perturbation system. *Journal of Computers*, 6(7), 1402–1407. <https://doi.org/10.4304/jcp.6.7.1402-1407>
- Bekr, G. A. (2016). Study of Significant Factors Affecting Labor Productivity at Construction Sites in Jordan: Site Survey. *Journal of Engineering Technology*, 4(1), 92–97. <https://doi.org/10.5176/2251-3701>
- Bhattacharjee, N. V., & Tollner, E. W. (2016). Improving management of windrow composting systems by modeling runoff water quality dynamics using recurrent neural network. *Ecological Modelling*, 339, 68–76. <https://doi.org/10.1016/j.ecolmodel.2016.08.011>
- Boehm, B. W. (1981). *Software Engineering Economics* (E. Cliffs (ed.)). Livre; Prentice-Hall.
- Boehm, B. W., Clark, B., Horowitz, E., Westland, C., Madachy, R., & Selby., R. (1995). Cost Models for Future Software Life Cycle Processes: COCOMO II. *Software Process and Product Measurement*.
- Bokor, O., Florez-Perez, L., Pesce, G., & Gerami Seresht, N. (2021). Using Artificial Neural Networks to Model Bricklaying Productivity. *Proceedings of the 2021 European Conference on Computing in Construction*, 2(1997), 52–58. <https://doi.org/10.35490/ec3.2021.155>
- Bouabaz, M. & Hamami, M. (2008). A cost estimation model for repair bridges based on artificial neural network. *American Journal of Applied Sciences*, 5(4), 334 -339
- C. H. Oglesby, H. W. Parker, G. A. H. (1989). *Productivity improvement in construction*. Livre; McGraw-Hill,.
- Chamekh, A. (2008). *Optimisation des procédés de mise en forme par les réseaux de neurones artificiels*. Thèse de doctorat; Université d'Angers,.
- Chesneau, C. (2017). *Modèles de régression*. Mémoire de Master; Université de Caen.
- Chouquet, C. (2010). *Modeles lineaires*. Livre; Université Paul Sabatier.
- Davalo, E., & Naim, P. (1989). *Des réseaux de neurones* (Eyrolles). Livre.
- Decoux, B. (2008). *Neurocomputing - Réseaux de Neurones*. Cours.
- Diab, M. F., Varma, A., & Panthi, K. (2017). Modeling the Construction Risk Ratings to

- Estimate the Contingency in Highway Projects. *Journal of Construction Engineering and Management*, 143(8), 04017041. [https://doi.org/10.1061/\(asce\)co.1943-7862.0001334](https://doi.org/10.1061/(asce)co.1943-7862.0001334)
- Djarah, D. (2006). *Application des Réseaux de Neurones Pour la Gestion d ' un Système de Perception Pour un Robot Mobile d' intérieur*. Mémoire de magistère; Université de Batna.
- Dozzi, S. P., & AbouRizk, S. M. (1993). *Productivity in construction* (Livre; Institute for research in construction (ed.)). National Research Council.
- El-Gohary, K. M., & Aziz, R. F. (2013). Factors influencing construction labor productivity in Egypt. *Journal of Management in Engineering*, 30(1), 1–9. [https://doi.org/10.1061/\(ASCE\)ME.1943-5479.0000168](https://doi.org/10.1061/(ASCE)ME.1943-5479.0000168)
- El-Kholy, A. M. (2010). Productivity performance estimation in Egyptian construction projects based on regression analysis. *Mansoura Engineering Journal (MEJ)*, 35(01), 44–54.
- Eldin, N. and Egger, S. (1990). Productivity improvement tool: Camcorders. *Journal of Construction Engineering and Management, ASCE*, 116(1), 100–111.
- Enis Cetin, A., & Nezh Gerek, O. (2006). Vector Quantization. *Wiley Encyclopedia of Biomedical Engineering*, April. <https://doi.org/10.1002/9780471740360.ebs1254>
- Enshassi, A., Mohamed, S., Mustafa, Z. A., & Mayer, P. E. (2007). Factors affecting labour productivity in building projects in the Gaza strip. *Journal of Civil Engineering and Management*, 13(4), 245–254. <https://doi.org/10.1080/13923730.2007.9636444>
- Ezeldin, A. S., & Sharara, L. M. (2006). Neural networks for estimating the productivity of concreting activities. *Journal of Construction Engineering and Management*, 132(6), 650–656. [https://doi.org/10.1061/\(ASCE\)0733-9364\(2006\)132:6\(650\)](https://doi.org/10.1061/(ASCE)0733-9364(2006)132:6(650))
- Fagbenle Olabosipo, I., Ogunde Ayodeji, O., & Owolabi James, D. (2011). Factors Affecting the Performance of Labour in Nigerian Construction Sites. *Mediterranean Journal of Social Sciences*, 2(2), 251–257.
- Fayçal, M. (2010). *Modélisation par Réseaux de Neurones Artificiels (RNA) et commande Prédictive non linéaire d'une station de production d'eau froide Option*. Mémoire de magistère; Université Badji Mokhtar Annaba.
- Foucart. (2006). Colinéarité et Régression linéaire. *Mathématiques et Sciences Humaines*, 173, 5–25.

- Garson, D. G. (1991). Interpreting neural network connection weights. *Artificial Intelligence. Expert* 6, 6(4), 47–51.
- Genin, M. (2005). Régression Linaire multiple. *Journal de Travaux Public.*
- Gerek, I. H., Erdis, E., Mistikoglu, G., & Usmen, M. (2015). Modelling masonry crew productivity using two artificial neural network techniques. *Journal of Civil Engineering and Management*, 21(2), 231–238. <https://doi.org/10.3846/13923730.2013.802741>
- Gerek, I. H., Erdis, E., Mistikoglu, G., & Usmen, M. A. (2014). Evaluation of plastering crew performance in building projects using data envelopment analysis. *Technological and Economic Development of Economy*, 22(6), 926–940. <https://doi.org/10.3846/20294913.2014.909903>
- Gevrey, M., Dimopoulos, I., & Lek, S. (2003). Review and comparison of methods to study the contribution of variables in artificial neural network models. *Ecological Modelling*, 160(3), 249–264. [https://doi.org/10.1016/S0304-3800\(02\)00257-0](https://doi.org/10.1016/S0304-3800(02)00257-0)
- Ghanizadeh, A. R., Heidarabadizadeh, N., & Jalali, F. (2020). Artificial neural network back-calculation of flexible pavements with sensitivity analysis using Garson's and connection weights algorithms. *Innovative Infrastructure Solutions*, 5(2). <https://doi.org/10.1007/s41062-020-00312-z>
- Ghoddousi, P., & Hosseini, M. R. (2012). A survey of the factors affecting the productivity of construction projects in Iran. *Technological and Economic Development of Economy*, 18(1), 99–116. <https://doi.org/10.3846/20294913.2012.661203>
- Goh, A. T. C. (1995). Back-propagation neural networks for modeling complex systems. *Artificial Intelligence in Engineering*, 9(3), 143–151. [https://doi.org/10.1016/0954-1810\(94\)00011-S](https://doi.org/10.1016/0954-1810(94)00011-S)
- Golnaraghi, S., Zangenehmadar, Z., Moselhi, O., & Alkass, S. (2019). Application of Artificial Neural Network (s) in Predicting Formwork Labour Productivity. *Advances in Civil Engineering*, 2019, 1–11. <https://doi.org/https://doi.org/10.1155/2019/5972620>
- Goodarzizad, P., Mohammadi Golafshani, E., & Arashpour, M. (2021). Predicting the construction labour productivity using artificial neural network and grasshopper optimisation algorithm. *International Journal of Construction Management*, 0(0), 1–28.

<https://doi.org/10.1080/15623599.2021.1927363>

Halstead, M. H. (1975). Toward a theoretical basis for estimating programming effort. *Proceedings of the 1975 Annual Conference, ACM 1975*, 222–224.

<https://doi.org/10.1145/800181.810326>

Hamisultane, H. (2002). *Econometrie*. Mémoire de Licence; France.

Hannula, M. (2002). Total productivity measurement based on partial productivity ratios. *International Journal of Production Economic*, 78, 57–67.

Hickson, B. G., & Ellis, L. A. (2014). Factors affecting Construction Labour Productivity in Trinidad and Tobago. *The Journal of the Association of Professional Engineers of Trinidad and Tobago*, 42(1), 4–11.

Hogg, R. V., & Tanis, E. a. (2009). *Instructor's Solutions Manual Probability and Statistical Inference Eighth Edition* (Livre). Prentice Hall, Upper Saddle River, NJ.

Hornik, K., Stinchcombe, M., White, H. (1989). Multilayer feedforward networks are universal approximators. *Neural Networks* 2, 359–366.

Hugues, B. (1998). *Réseau de Kohonen pour la détection des contours d'objets dans une image à niveaux de gris* (Vol. 17997, Issue 383). Projet d'application présenté à l'école de technologie supérieure; université du Québec.

Hwang, B.-G., Zhu, L., & Ming, J. T. T. (2017). Factors Affecting Productivity in Green Building Construction Projects: The Case of Singapore. *Journal of Management in Engineering*, 33(3), 04016052. [https://doi.org/10.1061/\(asce\)me.1943-5479.0000499](https://doi.org/10.1061/(asce)me.1943-5479.0000499)

IDRI, A. (2003). *Un modèle intelligent d'estimation des coûts de développement de logiciels*. Thèse de doctorat ; Université Du Québec À Montréal.

Jarkas, A. M. (2010). Critical investigation into the applicability of the learning curve theory to rebar fixing labor productivity. *Journal of Construction Engineering and Management*, 136(12), 1279–1288. [https://doi.org/10.1061/\(ASCE\)CO.1943-7862.0000236](https://doi.org/10.1061/(ASCE)CO.1943-7862.0000236)

Jarkas, A. M., & Bitar, C. G. (2012). Factors affecting construction labor productivity in Kuwait. *Journal of Construction Engineering and Management*, 138(7), 811–820. [https://doi.org/10.1061/\(ASCE\)CO.1943-7862.0000501](https://doi.org/10.1061/(ASCE)CO.1943-7862.0000501)

Juszczyk, M. (2020). Analysis of labour efficiency supported by the ensembles of neural networks on the example of steel reinforcement works. *Archives of Civil Engineering*, 66(1), 97–111. <https://doi.org/10.24425/ace.2020.131777>

Kan Phaobunjong, B.S., M. S. (2002). *Parametric Cost Estimating Model for Conceptual Cost Estimating of Building Construction Projects*. Thèse de doctorat; University of Texas at Austin.

Karim, N. A., Hassan, S., Yunus, J., & Hashim, M. (2013). Factors Influence Labour Productivity and the Impacts on Construction Industry. *Caspian Journal of Applied Sciences Research*, 2(August), 349–354.

Khanh, H. D., Kim, S. Y., Khoa, N. Van, & Tu, N. T. (2021). The relationship between workers' experience and productivity: a case study of brick masonry construction. *International Journal of Construction Management*, 0(0), 1–10. <https://doi.org/10.1080/15623599.2021.1899593>

Kisi, K. P. (2015). *Estimation of Optimal Productivity in Labor- Intensive Construction Operations*. Thèse de doctorat; University of Nebraska-Lincoln. <http://digitalcommons.unl.edu/constructiondiss%5Cnhttp://digitalcommons.unl.edu/constructiondiss/19>

Lawaju, N., Parajuli, N., & Shrestha, S. K. (2021). Analysis of Labor Productivity of Brick Masonry Work in Building Construction in Kathmandu Valley. *Journal of Advanced College of Engineering and Management*, 6, 159–175. <https://doi.org/10.3126/jacem.v6i0.38356>

Lek, S., Belaud, A., Baran, P., Dimopoulos, I., & Delacoste, M. (1996). Role of some environmental variables in trout abundance models using neural networks. *Aquatic Living Resources*, 9(1), 23–29.

LEPAGE, R. (2004). *Reconnaissance d'algues toxiques par vision artificielle et réseau de neurones*. Mémoire maître sciences appliquées; Université Du Québec A Rimouski.

Levine, D. M., Sincich, T., & Stephan, D. (2002). *Practical statistics by example using Microsoft Excel and Minitab* (U. S. River (ed.); 2nd ed.). Livre; Prentice Hall Professional Technical Reference.

Liou, F., and Borcharding, J. D. (1986). Work sampling can predict unit rate productivity. *J. Constr. Engrg. and Mgmt*, 112(1), 90–103.

- M. Hafez, S. (2014). Critical Factors Affecting Construction Labor Productivity in Egypt. *American Journal of Civil Engineering*, 2(2), 35. <https://doi.org/10.11648/j.ajce.20140202.14>
- Mady, M. (2013). *Prediction Model of Construction Labor Production Rates in Gaza Strip using Artificial Neural Networks* (Vol. 53, Issue 9). Mémoire de master; The Islamic University – Gaza Deanery.
- Mahamid, I., Al-Ghonamy, A., & Aichouni, M. (2013). Major Factors Influencing Employee Productivity in the KSA Public Construction Projects. *International Journal of Civil & Environmental Engineering IJCEE*, 14(1), 1–5.
- Mahdi, E. (2011). *Réseaux de neurones artificiels appliqués à la méthode électromagnétique transitoire infiniTEM*. Mémoire de maîtrise en Ingénierie; Université du Québec à Chicoutimi.
- Manai, M. (2019). *A New Approach for Topic Detection using Adaptive Neural Networks*. 2017–2018. <http://arxiv.org/abs/1903.03775>
- Martins, T. D., Annichino-Bizzacchi, J. M., Romano, A. V. C., & Maciel Filho, R. (2020). Artificial neural networks for prediction of recurrent venous thromboembolism. *International Journal of Medical Informatics*, 141, 104221. <https://doi.org/10.1016/j.ijmedinf.2020.104221>
- Matejević-Nikolić, B., & Živković, L. (2022). Comparative Representation of Two Models for Predicting the Productivity of Column and Wall Concreting Process. *Buildings*, 12(11). <https://doi.org/10.3390/buildings12111809>
- Mlybari, E. A. (2020). Application of soft computing techniques to predict construction labour productivity in Saudi Arabia. *International Journal of GEOMATE*, 19(71), 203–210. <https://doi.org/10.21660/2020.71.31349>
- Muqem, S., Idrus, A., Khamidi, M. F., Bin Ahmad, J., & Bin Zakaria, S. (2011). Prediction Modeling of Construction Labor Production Rates using Artificial Neural Network Sana. *Electronic Journal of Information Technology in Construction*, 16(January 2015), 713–726.
- Narayanan, K. M. A., & Aswathi, M. P. (2018). Labour Productivity Model for Structural Elements By Varying Buildability Factors using Multiple Regression Analysis. *International Journal of Advance Research, Ideas and Innovations in Technology*, 4(2), 139–146.

- Nasirzadeh, F., Rostamnezhad, M., Carmichael, D. G., Khosravi, A., & Aisbett, B. (2022). Labour productivity in Australian building construction projects: a roadmap for improvement. *International Journal of Construction Management*, 22(11), 2079–2088. <https://doi.org/10.1080/15623599.2020.1765286>
- Noor, I. (1998). Measuring construction labor productivity by daily visits. *AACE International Transactions*, PR16.
- O'Donoghue, J., Roantree, M., & McCarren, A. (2017). Detecting feature interactions in agricultural trade data using a deep neural network. *Lecture Notes in Computer Science (Including Subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, 10440 LNCS, 449–458. https://doi.org/10.1007/978-3-319-64283-3_33
- Olden, J. D., & Jackson, D. A. (2002). Illuminating the “black box”: A randomization approach for understanding variable contributions in artificial neural networks. *Ecological Modelling*, 154(1–2), 135–150. [https://doi.org/10.1016/S0304-3800\(02\)00064-9](https://doi.org/10.1016/S0304-3800(02)00064-9)
- Olomolaiye, P., Jayawardane, A., & Harris, F. (1998). *Construction productivity management*. Longman.
- Olomolaiye, P. O., Jayawardane, A. K. W., & Harris, F. C. (1998). *Construction Productivity Management* (E. Harlow (ed.)). Livre; Addison Wesley Longman.
- Otmani, A. (2020). *Prédire et contrôler les coûts d'exploitations et maintenances des constructions*. Thèse de Doctorat; Université 20 Aout-1955-Skikda.
- Otmani, I. (2011). L'analyse discriminante et le perceptron multicouche. In *Réforme, Humanisme, Renaissance*. Mémoire de magistère; Université Mentouri Constantine.
- Paikun, P., Fatimah, C., Nugroho, N. S., & Kusmana, D. (2021). Labor Productivity Rating Model for Light Brick Wall Installation in Residential Projects. *ASTONJADRO*., 10(2), 260–270. <http://dx.doi.org/10.32832/astonjadro>.
- Park, H. S. (2006). Conceptual framework of construction productivity estimation. *Journal of Construction Engineering and Management*, 131(7), 772–778.
- Parker, H. W., and Oglesby, C. H. (1972). *Methods improvement for construction managers*. Livre; McGraw-Hill.

Pekuri, A., Haapasalo, H., & Herrala, M. (2011). Productivity and performance management—managerial practices in the construction industry. *International Journal of Performance Measurement*, 1, 39–58.

Portas, J., & AbouRizk, S. (1997). Neural network model for estimating construction productivity. *Journal of Construction Engineering and Management*, 123(4), 399–410. [https://doi.org/10.1061/\(ASCE\)0733-9364\(1997\)123:4\(399\)](https://doi.org/10.1061/(ASCE)0733-9364(1997)123:4(399))

Putnam, L. H. (1978). A General Empirical Solution to the Macro Software Sizing and Estimation Problem. *Transactions on Software Engineering*, 4(4).

RAFIK, M. (2008). *Diagnostic industriel par neuro-flou -application a un système de production*. Mémoire de magistère; Université El-hadj Lakhdar –BATNA-.

Rakhra, A. S. (1991). Construction productivity: Concept, measurement and trends, organization and management in construction. *Proc. of the 4th Yugoslavian Symp. on Construction Management*, 487–497.

Rakotomalala, R. (2015). *Pratique de la Régression Linéaire Multiple Diagnostic et sélection de variables*. Université Lumière Lyon 2.

Rojas, E. M., & Aramvareekul, P. (2003). *Labor Productivity Drivers and Opportunities in the Construction Industry*. April, 78–82.

Sahin, F. (1997). *A radial basis function approach to a color image classification problem in a real time industrial application*. Thèse de doctorat; Université Virginia Tech.

Salem, D. (2006). *Predicting productivity rates of pouring concrete in Egypt using artificial neural networks model*. Mémoire de mastère; Arab Academy for Science, Technology and Maritime Transport.

Sarmiento, M., Sc, M., Luiz, M., Ferreira, R., & Sc, D. (2021). Sizing an Artificial Neural Network for Modelling Piping Assembly Labor Productivity Under Different Conditions. *International Journal of Engineering Research and Applications*, 11(2), 16–22. <https://doi.org/10.9790/9622-1102051622>

Shahin, M. A., Jaksa, M. B., & Maier, H. . (2002). Artificial Neural Network–based settlement prediction formula for shallow foundations on granular soils. *Ustralian*

Geomechanics, 36(1), 45–52.

Shehu, Z., & Akintoye, A. (2010). Major challenges to the successful implementation and practice of programme management in the construction environment: A critical analysis.

International Journal of Project Management, 28(1), 26–39.

<https://doi.org/10.1016/j.ijproman.2009.02.004>

Shoar, S., & Banaitis, A. (2018). Application of fuzzy fault tree analysis to identify factors influencing construction labor productivity: A high-rise building case study. *Journal of Civil Engineering and Management*, 25(1), 41–52. <https://doi.org/10.3846/jcem.2019.7785>

Shojaeefard, M. H., Akbari, M., Tahani, M., & Farhani, F. (2013). Sensitivity analysis of the artificial neural network outputs in friction stir lap joining of aluminum to brass. *Advances in Materials Science and Engineering*, 2013. <https://doi.org/10.1155/2013/574914>

Stricker, M. (2000). *Réseaux de neurones pour le traitement automatique du langage: conception et réalisation de filtres d'informations*. Thèse de doctorat; ESPCI ParisTECH.

TAHAR, I., HABIB, M., & ISABATEN, H. (2003). *Approche De Classification Par Réseaux De Kohonen Pour Établir Des Cartes D ' Occupation Du Sol De La Région D ' Oran (Algérie) the Kohonen ' S Network Classification Approach Used for Land Cover Mapping : Case of Oran Area (Algeria)*. 3(October), 361–372.

Talhouni, B. T. (1990). *Measurement and analysis of construction labour productivity*. Thèse de doctorat; Dundee, UK.

Tangen, S. (2005). Demystifying productivity and performance. *International Journal of Productivity and Performance Management*, 54(1), 34–46.

Thomas, H. R., and S. (1991). Factors affecting masonry productivity. *Journal of Construction Engineering and Management*, 117(4), 626–644.

Thomas, H. R., Maloney, W. F., Horner, R. M. W., Smith, G. R., Handa, V. K., & Sanders, S. R. (1990). Modeling construction labor productivity. *Journal of Construction Engineering and Management*, Vol. 116, No. 4, December, 1990. ©ASCE, ISSN 0733-9364/90/0004-0705, 116(4), 705–726. [https://doi.org/https://doi.org/10.1061/\(ASCE\)0733-9364\(1990\)116:4\(705\)](https://doi.org/https://doi.org/10.1061/(ASCE)0733-9364(1990)116:4(705))

Thomas, H. R., & Mathews, C. T. (1985). An analysis of the methods of measuring

construction productivity. *Construction Industry Institute, Austin, Tex.*

Thomas, H. R., Mathews, C. T., & Ward, J. G. (1986). Learning curve models of construction productivity. *J. Constr. Eng. Manage*, 112(2), 245–258.

Thomas, N., & Thomas, V. A. (2016). Regression Modelling for Predicting of construction cost and Duration. *Applied Mechanics and Materials*, 857, 195–199.

Tsopze, N. (2010). *Treillis de Galois et réseaux de neurones: une approche constructive d'architecture des réseaux de neurones*. Thèse de doctorat; Université d'Artois et de l'Université de Yaoundé I. <https://www.theses.fr/2010ARTO0407>

Umit Dikmen, S., & Sonmez, M. (2011). An artificial neural networks model for the estimation of formwork labour. *Journal of Civil Engineering and Management*, 17(3), 340–347. <https://doi.org/10.3846/13923730.2011.594154>

Viennet, E. (2006). *Réseaux à fonctions de base radiales* (Lavoisier (ed.)). Livre; HAL.

Vitela, J. E., & Reifman, J. (1997). Premature Saturation in Backpropagation Networks: Mechanism and Necessary Conditions. *Neural Networks*, 10(4), 721–735.

Walston, C. E., & Felix, C. P. (1977). A method of programming measurement and estimation. *IBM Systems Journal*, 16(1), 54-73.

Yi, W., & Chan, A. P. C. (2014). Critical review of labor productivity research in construction journals. *Journal of Management in Engineering*, 30(2), 214–225. [https://doi.org/10.1061/\(ASCE\)ME.1943-5479.0000194](https://doi.org/10.1061/(ASCE)ME.1943-5479.0000194)

Zayed, T. M., & Halpin, W. D. (2005). Pile Construction Productivity Assessment. *Journal of Construction Engineering and Management*, 131(6), 168–175. [https://doi.org/10.1061/\(ASCE\)0733-9364\(2005\)131:6\(705\)](https://doi.org/10.1061/(ASCE)0733-9364(2005)131:6(705))

Zhang, Z., Beck, M. W., Winkler, D. A., Huang, B., Sibanda, W., & Goyal, H. (2018). Opening the black box of neural networks: methods for interpreting neural network models in clinical applications. *Annals of Translational Medicine*, 6(11), 216–216. <https://doi.org/10.21037/atm.2018.05.32>

Annexes

Annexe A

Questionnaire

Annexes A

QUESTIONNAIRE SUR LES FACTEURS PROVOQUANT LA BAISSE DE LA PRODUCTIVITÉ DANS LES PROJETS DE CONSTRUCTION EN ALGÉRIE.

J'ai l'honneur de bien vouloir vous adresser ce questionnaire pour que vos réponses me permettent d'identifier Les facteurs provoquant la baisse de la productivité en Algérie, ce sondage faisant l'objet d'une étude de doctorat.

1. Informations personnelles

Votre Poste de travail

Expérience Professionnelle

2. Facteurs liés à la gestion (management)

	Pas du tout important	Peu important	Normal	Important	Très important
Faible taille d'équipage	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Encombrement dans l'espace de travaille	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Manque ou faible contrôle	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Les faibles salaires	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Le retard de paiement	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Faible motivation des ouvriers (primes, jours de récupération...)	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Etat des matériels de construction (coffrage, ...)	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

3. Facteurs liés aux ouvriers

Annexes A

	Pas du tout important	Peu important	Normal	Important	Très important
Âge des ouvriers	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Expérience des ouvriers	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Heure de travaux/ Absentéisme	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
La Quantité de travail réalisé	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

4. Facteurs Externe

	Pas du tout important	Peu important	Normal	Important	Très important
Conditions météorologiques (Pluie, Soleil)	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Une forte ou basse humidité	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Température élevée	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Basse température	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Pénurie des matériaux de construction	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

5. Autres facteurs peuvent provoquer la baisse de productivité (précisez l'importance) :

Annexe B

Base de données

Annexes B

Tableau B- 1 Sous base de données d'entrées non normalisées (réelles) pour l'apprentissage

Projet	Nbr Heures	Nbr Manœuvres	Nbr ouvriers qualifiés	Min Expérience	Max Expérience	Etat des matériels de construction	Type de motivation	Quantité de travail réalisé	Température	Salaire des ouvriers qualifiés	Salaire des manœuvres
1	8,00	3	5	4	13	2	0,15	71,12	25	20 640,00	24 310,00
2	6,00	2	11	1	3	2	0,1	206,06	17	19 390,00	22 160,00
3	7,00	2	9	1	3	2	0,1	206,06	16	19 390,00	22 160,00
4	8,00	2	7	0,7	3,4	2	0,1	206,06	19	19 390,00	22 160,00
5	4,00	2	10	0,7	3,4	2	0,1	88,31	16	19 390,00	22 160,00
6	5,00	2	13	1	3	2	0,1	176,63	17	19 390,00	22 160,00
7	8,00	2	10	1	3	2	0,1	235,5	16	19 390,00	22 160,00
8	4,00	2	13	1	3	2	0,1	117,75	19	19 390,00	22 160,00
9	7,00	2	12	1	3	2	0,1	235,5	20	19 390,00	22 160,00
10	4,00	1	10	1	3	2	0,1	58,87	18	19 390,00	22 160,00
11	8,00	2	7	0,7	3,4	2	0,1	176,63	16	19 390,00	22 160,00
12	5,00	2	9	2	10	2	0,1	140,6	13	21 290,00	24 310,00
13	5,00	1	9	3	9	2	0,15	140,6	15	20 640,00	23 580,00
14	10,0	4	10	2	5	3	0,15	210,9	14	20 640,00	24 310,00
15	8,00	2	13	1	4	2	0,15	175,75	18	21 290,00	23 580,00
16	8,00	1	10	2	5	3	0,15	210,9	12	20 010,00	22 860,00
17	5,00	2	9	1	3	2	0,1	117,75	15	19 390,00	22 160,00
18	6,00	2	11	1	3	2	0,1	176,63	14	19 390,00	22 160,00
19	4,00	2	13	1	3	2	0,1	88,31	22	19 390,00	22 160,00
20	5,00	2	12	0,7	3,4	2	0,1	176,63	18	19 390,00	22 160,00
21	6,00	2	13	1	3	2	0,1	176,63	24	19 390,00	22 160,00
22	8,00	1	10	1	5	2	0,15	210,90	19	20010,00	22860,00
23	4,00	2	11	1	3	2	0,1	147,19	20	19 390,00	22 160,00
24	8,00	3	10	1	2	2	0,15	140,6	18	21290,00	23580,00
25	5,00	1	9	0,7	3,4	2	0,1	117,75	21	19 390,00	22 160,00
26	4,00	3	12	1	3	2	0,1	147,19	19	19 390,00	22 160,00
27	4,00	1	6	0,7	3,4	2	0,1	58,87	23	19 390,00	22 160,00

Annexes B

28	5,00	2	10	0,7	3,4	2	0,1	176,63	22	19 390,00	22 160,00
29	6,00	1	8	0,7	3,4	2	0,1	117,75	23	19 390,00	22 160,00
30	5,00	1	7	1	3	2	0,1	88,31	25	19 390,00	22 160,00
31	6,00	2	11	1	3	2	0,1	206,06	17	19 390,00	22 160,00
32	8,00	2	14	2	15	3	0,2	310	23	21 000,00	29 000,00
33	6,00	2	10	1	3	2	0,1	117,75	25	19 390,00	22 160,00
34	8,00	3	9	2	10	2	0,15	132,08	25	21 290,00	23 580,00
35	4,00	2	8	5	26	3	0,1	101,6	26	20 640,00	24 310,00
36	5,00	2	8	2	8	2	0,1	116,2	23	20 640,00	24 310,00
37	8,00	1	9	2	6	2	0,15	101,6	22	21 290,00	23 580,00
38	8,00	2	7	4	9	1	0,15	115,16	23	20 010,00	22 860,00
39	8,00	3	10	1	5	2	0,15	132,08	26	20 640,00	24 310,00
40	4,00	2	8	2	7	2	0,1	101,6	28	21 290,00	23 580,00
41	7,00	2	12	1	10	2	0,15	132,08	24	20 010,00	22 860,00
42	7,00	1	9	2	9	2	0,15	101,16	24	20 640,00	24 310,00
43	4,00	3	13	2	7	3	0,1	164,56	22	21 290,00	23 580,00
44	7,00	1	12	1	6	3	0,15	132,08	24	20 640,00	24 310,00
45	4,00	3	10	7	13	3	0,1	132,08	25	21 290,00	23 580,00
46	7,00	2	12	5	11	3	0,15	142,24	23	21 290,00	23 580,00
47	7,00	2	12	3	8	3	0,15	142,24	15	20 640,00	24 310,00
48	7,00	2	7	0,5	5	1	0,15	91,44	17	20 010,00	22 860,00
49	7,00	1	9	1	6	1	0,15	101,6	21	20 640,00	24 310,00
50	7,00	2	5	2	6	2	0,15	71,12	23	21 290,00	23 580,00
51	4,00	1	10	3	11	3	0,1	111,76	28	20 010,00	22 860,00
52	4,00	2	9	2	9	3	0,1	111,76	27	20 640,00	24 310,00
53	10,0	2	12	5,00	11,00	3	0,15	421,8	12	20 010,00	24 000,00
54	8,00	1	8	0,50	4,00	2	0,15	140,6	10	20 010,00	24 000,00
55	8,00	2	11	3,00	11,00	3	0,15	281,2	17	20 010,00	24 000,00
56	8,00	1	13	2,00	9,00	2	0,15	210,9	15	20 010,00	24 000,00
57	8,00	2	11	2,00	9,00	3	0,15	228,76	20	21 290,00	24 000,00
58	6,00	2	10	2,00	9,00	3	0,15	157	21	21 290,00	24 000,00
59	8,00	2	11	2,00	9,00	3	0,15	228,76	23	21 290,00	24 000,00

Annexes B

60	4,00	1	7	7	12	1	0,1	81,28	28	19 390,00	22 160,00
61	7,00	3	9	5	11	3	0,2	121,92	27	19 390,00	22 160,00
62	6,00	2	11	0,7	3,4	2	0,1	206,06	21	19 390,00	22 160,00
63	6,00	2	9	0,5	4	2	0,1	140,6	26	20 640,00	24 310,00
64	4,00	2	12	1	3	2	0,1	117,75	23	19 390,00	22 160,00
65	6,00	2	12	1	3	2	0,1	206,06	21	19 390,00	22 160,00
66	6,00	2	11	1	3	2	0,1	206,06	21	19 390,00	22 160,00
67	5,00	2	9	0,7	3,4	3	0,1	157,5	11	21 290,00	24 000,00
68	5,00	2	11	1	3	2	0,1	206,06	18	19 390,00	22 160,00
69	7,00	2	10	3	18	3	0,15	121,92	26	20 640,00	24 310,00
70	5,00	2	9	3	12	1	0,15	113,2	26	20 010,00	22 860,00
71	4,00	2	8	1	3	2	0,1	58,87	24	19 390,00	22 160,00

Annexes B

Tableau B- 2 Sous base de données de sortie non normalisées (réelles) pour test

Projet	Nbr Heures	Nbr manœuvres	Nbr ouvriers qualifiés	Min Expérience	Max Expérience	Etat des matériels de construction	Type de motivation	Quantité de travail réalisé	Température	Salaire des ouvriers qualifiés	Salaire des manœuvres
1	8,00	3	11	0,7	3,4	2	0,1	294,37	16	19 390,00	22 160,00
2	8,00	1	10	2	5	2	0,15	210,90	15	20 640,00	23 580,00
3	6,00	2	10	0,7	3,4	2	0,1	176,00	20	19 390,00	22 160,00
4	8,00	1	10	1	3	2	0,1	206,06	24	19 390,00	22 160,00
5	5,00	2	11	1	3	2	0,1	176,63	13	19 390,00	22 160,00
6	4,00	3	10	0,5	4	3	0,1	132,08	23	20 640,00	24 310,00
7	7,00	1	7	1	9	1	0,15	81,28	27	19 390,00	22 160,00
8	6,00	1	10	3	11	3	0,1	101,6	29	20 640,00	24 310,00
9	7,00	2	8	0,5	5	3	0,2	101,6	24	20 010,00	22860,00
10	7,00	1	9	5	8	3	0,2	101,6	24	20 640,00	24 310,00
11	8,00	2	11	0,7	3,4	2	0,1	235,5	17	19 390,00	22 160,00
12	3,00	2	13	0,7	3,4	2	0,1	58,87	19	19 390,00	22 160,00
13	7,00	2	9	6	14	3	0,1	210,9	22	21 290,00	23 580,00
14	8,00	2	12	0,7	3,4	3	0,1	228,76	19	21 290,00	24 000,00
15	6,00	2	11	0,7	3,4	3	0,1	157,5	9	21 290,00	24 000,00
16	4,00	3	8	1	6	2	0,1	111,76	22	20 640,00	24 310,00
17	5,00	2	10	0,7	3,4	2	0,1	147,18	24	19 390,00	22 160,00
18	4,00	2	12	0,7	3,4	2	0,1	88,31	22	19 390,00	22 160,00
19	4,00	2	12	0,7	3,4	2	0,1	88,31	24	19 390,00	22 160,00
20	4,00	2	10	1	3	2	0,1	88,31	19	19 390,00	22 160,00
21	7,00	2	10	2	9	2	0,2	121,92	28	20 010,00	22 860,00
22	7,00	3	11	0,5	4	3	0,15	142,24	29	19 390,00	22 160,00
23	8,00	2	13	1	9	3	0,15	175,75	19	20 010,00	24 000,00
24	8,00	2	10	7	13	2	0,15	141,6	20	20 010,00	24 000,00
25	8,00	1	10	3	11	2	0,15	210,9	18	20 010,00	24 000,00
26	6,00	2	10	2	9	3	0,15	157	23	21 290,00	24 000,00
27	6,00	2	7	0,7	3,4	2	0,1	117,75	21	19 390,00	22 160,00
28	7,00	2	8	5	12	1	0,15	101,6	29	21 290,00	23 580,00
29	7,00	2	10	6	14	3	0,15	121,92	27	21 290,00	23 580,00
30	4,00	4	12	2	5	3	0,1	140,60	14	20 640,00	24 310,00

Annexes B

Tableau B- 3 Sous base de données d'entrées normalisées pour apprentissage

Projet	Nbr Heures	Nbr manœuvres	Nbr ouvriers qualifiés	Min Expérience	Max Expérience	Etat des matériels de construction	Type de motivation	Quantité de travail réalisé	Température	Salaire des ouvriers qualifiés	Salaire des manœuvres
1	0,714	0,667	0,000	0,538	0,458	0,500	0,500	0,034	0,800	0,658	0,314
2	0,429	0,333	0,667	0,077	0,042	0,500	0,000	0,406	0,400	0,000	0,000
3	0,571	0,333	0,444	0,077	0,042	0,500	0,000	0,406	0,350	0,000	0,000
4	0,714	0,333	0,222	0,031	0,058	0,500	0,000	0,406	0,500	0,000	0,000
5	0,143	0,333	0,556	0,031	0,058	0,500	0,000	0,081	0,350	0,000	0,000
6	0,286	0,333	0,889	0,077	0,042	0,500	0,000	0,324	0,400	0,000	0,000
7	0,714	0,333	0,556	0,077	0,042	0,500	0,000	0,487	0,350	0,000	0,000
8	0,143	0,333	0,889	0,077	0,042	0,500	0,000	0,162	0,500	0,000	0,000
9	0,571	0,333	0,778	0,077	0,042	0,500	0,000	0,487	0,550	0,000	0,000
10	0,143	0,000	0,556	0,077	0,042	0,500	0,000	0,000	0,450	0,000	0,000
11	0,714	0,333	0,222	0,031	0,058	0,500	0,000	0,324	0,350	0,000	0,000
12	0,286	0,333	0,444	0,231	0,333	0,500	0,000	0,225	0,200	1,000	0,314
13	0,286	0,000	0,444	0,385	0,292	0,500	0,500	0,225	0,300	0,658	0,208
14	1,000	1,000	0,556	0,231	0,125	1,000	0,500	0,419	0,250	0,658	0,314
15	0,714	0,333	0,889	0,077	0,083	0,500	0,500	0,322	0,450	1,000	0,208
16	0,714	0,000	0,556	0,231	0,125	1,000	0,500	0,419	0,150	0,326	0,102
17	0,286	0,333	0,444	0,077	0,042	0,500	0,000	0,162	0,300	0,000	0,000
18	0,429	0,333	0,667	0,077	0,042	0,500	0,000	0,324	0,250	0,000	0,000
19	0,143	0,333	0,889	0,077	0,042	0,500	0,000	0,081	0,650	0,000	0,000
20	0,286	0,333	0,778	0,031	0,058	0,500	0,000	0,324	0,450	0,000	0,000
21	0,429	0,333	0,889	0,077	0,042	0,500	0,000	0,324	0,750	0,000	0,000
22	0,714	0,000	0,556	0,077	0,125	0,500	0,500	0,419	0,500	0,326	0,102
23	0,143	0,333	0,667	0,077	0,042	0,500	0,000	0,243	0,550	0,000	0,000
24	0,714	0,667	0,556	0,077	0,000	0,500	0,500	0,225	0,450	1,000	0,208
25	0,286	0,000	0,444	0,031	0,058	0,500	0,000	0,162	0,600	0,000	0,000
26	0,143	0,667	0,778	0,077	0,042	0,500	0,000	0,243	0,500	0,000	0,000
27	0,143	0,000	0,111	0,031	0,058	0,500	0,000	0,000	0,700	0,000	0,000
28	0,286	0,333	0,556	0,031	0,058	0,500	0,000	0,324	0,650	0,000	0,000
29	0,429	0,000	0,333	0,031	0,058	0,500	0,000	0,162	0,700	0,000	0,000
30	0,286	0,000	0,222	0,077	0,042	0,500	0,000	0,081	0,800	0,000	0,000
31	0,429	0,333	0,667	0,077	0,042	0,500	0,000	0,406	0,400	0,000	0,000
32	0,714	0,333	1,000	0,231	0,542	1,000	1,000	0,692	0,700	0,847	1,000
33	0,429	0,333	0,556	0,077	0,042	0,500	0,000	0,162	0,800	0,000	0,000

Annexes B

34	0,714	0,667	0,444	0,231	0,333	0,500	0,500	0,202	0,800	1,000	0,208
35	0,143	0,333	0,333	0,692	1,000	1,000	0,000	0,118	0,850	0,658	0,314
36	0,286	0,333	0,333	0,231	0,250	0,500	0,000	0,158	0,700	0,658	0,314
37	0,714	0,000	0,444	0,231	0,167	0,500	0,500	0,118	0,650	1,000	0,208
38	0,714	0,333	0,222	0,538	0,292	0,000	0,500	0,155	0,700	0,326	0,102
39	0,714	0,667	0,556	0,077	0,125	0,500	0,500	0,202	0,850	0,658	0,314
40	0,143	0,333	0,333	0,231	0,208	0,500	0,000	0,118	0,950	1,000	0,208
41	0,571	0,333	0,778	0,077	0,333	0,500	0,500	0,202	0,750	0,326	0,102
42	0,571	0,000	0,444	0,231	0,292	0,500	0,500	0,117	0,750	0,658	0,314
43	0,143	0,667	0,889	0,231	0,208	1,000	0,000	0,291	0,650	1,000	0,208
44	0,571	0,000	0,778	0,077	0,167	1,000	0,500	0,202	0,750	0,658	0,314
45	0,143	0,667	0,556	1,000	0,458	1,000	0,000	0,202	0,800	1,000	0,208
46	0,571	0,333	0,778	0,692	0,375	1,000	0,500	0,230	0,700	1,000	0,208
47	0,571	0,333	0,778	0,385	0,250	1,000	0,500	0,230	0,300	0,658	0,314
48	0,571	0,333	0,222	0,000	0,125	0,000	0,500	0,090	0,400	0,326	0,102
49	0,571	0,000	0,444	0,077	0,167	0,000	0,500	0,118	0,600	0,658	0,314
50	0,571	0,333	0,000	0,231	0,167	0,500	0,500	0,034	0,700	1,000	0,208
51	0,143	0,000	0,556	0,385	0,375	1,000	0,000	0,146	0,950	0,326	0,102
52	0,143	0,333	0,444	0,231	0,292	1,000	0,000	0,146	0,900	0,658	0,314
53	1,000	0,333	0,778	0,692	0,375	1,000	0,500	1,000	0,150	0,326	0,269
54	0,714	0,000	0,333	0,000	0,083	0,500	0,500	0,225	0,050	0,326	0,269
55	0,714	0,333	0,667	0,385	0,375	1,000	0,500	0,613	0,400	0,326	0,269
56	0,714	0,000	0,889	0,231	0,292	0,500	0,500	0,419	0,300	0,326	0,269
57	0,714	0,333	0,667	0,231	0,292	1,000	0,500	0,468	0,550	1,000	0,269
58	0,429	0,333	0,556	0,231	0,292	1,000	0,500	0,270	0,600	1,000	0,269
59	0,714	0,333	0,667	0,231	0,292	1,000	0,500	0,468	0,700	1,000	0,269
60	0,143	0,000	0,222	1,000	0,417	0,000	0,000	0,062	0,950	0,000	0,000
61	0,571	0,667	0,444	0,692	0,375	1,000	1,000	0,174	0,900	0,000	0,000
62	0,429	0,333	0,667	0,031	0,058	0,500	0,000	0,406	0,600	0,000	0,000
63	0,429	0,333	0,444	0,000	0,083	0,500	0,000	0,225	0,850	0,658	0,314
64	0,143	0,333	0,778	0,077	0,042	0,500	0,000	0,162	0,700	0,000	0,000
65	0,429	0,333	0,778	0,077	0,042	0,500	0,000	0,406	0,600	0,000	0,000
66	0,429	0,333	0,667	0,077	0,042	0,500	0,000	0,406	0,600	0,000	0,000
67	0,286	0,333	0,444	0,031	0,058	1,000	0,000	0,272	0,100	1,000	0,269
68	0,286	0,333	0,667	0,077	0,042	0,500	0,000	0,406	0,450	0,000	0,000
69	0,571	0,333	0,556	0,385	0,667	1,000	0,500	0,174	0,850	0,658	0,314
70	0,286	0,333	0,444	0,385	0,417	0,000	0,500	0,150	0,850	0,326	0,102
71	0,143	0,333	0,333	0,077	0,042	0,500	0,000	0,000	0,750	0,000	0,000

Annexes B

Tableau B- 4 Sous base de données de sortie normalisées pour test

Projet	Nbr Heures	Nbr manœuvres	Nbr ouvriers qualifiés	Min Expérience	Max Expérience	Etat des matériels de construction	Type de motivation	Quantité de travail réalisé	Température	Salaire des ouvriers qualifiés	Salaire des manœuvres
1	0,714	0,667	0,667	0,031	0,058	0,500	0,000	0,649	0,350	0,000	0,000
2	0,714	0,000	0,556	0,231	0,125	0,500	0,500	0,419	0,300	0,658	0,208
3	0,429	0,333	0,556	0,031	0,058	0,500	0,000	0,323	0,550	0,000	0,000
4	0,714	0,000	0,556	0,077	0,042	0,500	0,000	0,406	0,750	0,000	0,000
5	0,286	0,333	0,667	0,077	0,042	0,500	0,000	0,324	0,200	0,000	0,000
6	0,143	0,667	0,556	0,000	0,083	1,000	0,000	0,202	0,700	0,658	0,314
7	0,571	0,000	0,222	0,077	0,292	0,000	0,500	0,062	0,900	0,000	0,000
8	0,429	0,000	0,556	0,385	0,375	1,000	0,000	0,118	1,000	0,658	0,314
9	0,571	0,333	0,333	0,000	0,125	1,000	1,000	0,118	0,750	0,326	0,102
10	0,571	0,000	0,444	0,692	0,250	1,000	1,000	0,118	0,750	0,658	0,314
11	0,714	0,333	0,667	0,031	0,058	0,500	0,000	0,487	0,400	0,000	0,000
12	0,000	0,333	0,889	0,031	0,058	0,500	0,000	0,000	0,500	0,000	0,000
13	0,571	0,333	0,444	0,846	0,500	1,000	0,000	0,419	0,650	1,000	0,208
14	0,714	0,333	0,778	0,031	0,058	1,000	0,000	0,468	0,500	1,000	0,269
15	0,429	0,333	0,667	0,031	0,058	1,000	0,000	0,272	0,000	1,000	0,269
16	0,143	0,667	0,333	0,077	0,167	0,500	0,000	0,146	0,650	0,658	0,314
17	0,286	0,333	0,556	0,031	0,058	0,500	0,000	0,243	0,750	0,000	0,000
18	0,143	0,333	0,778	0,031	0,058	0,500	0,000	0,081	0,650	0,000	0,000
19	0,143	0,333	0,778	0,031	0,058	0,500	0,000	0,081	0,750	0,000	0,000
20	0,143	0,333	0,556	0,077	0,042	0,500	0,000	0,081	0,500	0,000	0,000
21	0,571	0,333	0,556	0,231	0,292	0,500	1,000	0,174	0,950	0,326	0,102
22	0,571	0,667	0,667	0,000	0,083	1,000	0,500	0,230	1,000	0,000	0,000
23	0,714	0,333	0,889	0,077	0,292	1,000	0,500	0,322	0,500	0,326	0,269
24	0,714	0,333	0,556	1,000	0,458	0,500	0,500	0,228	0,550	0,326	0,269
25	0,714	0,000	0,556	0,385	0,375	0,500	0,500	0,419	0,450	0,326	0,269
26	0,429	0,333	0,556	0,231	0,292	1,000	0,500	0,270	0,700	1,000	0,269
27	0,429	0,333	0,222	0,031	0,058	0,500	0,000	0,162	0,600	0,000	0,000
28	0,571	0,333	0,333	0,692	0,417	0,000	0,500	0,118	1,000	1,000	0,208
29	0,571	0,333	0,556	0,846	0,500	1,000	0,500	0,174	0,900	1,000	0,208
30	0,143	1,000	0,778	0,231	0,125	1,000	0,000	0,225	0,250	0,658	0,314

Annexe C

Tableau des quantiles d'ordre 0.95

(Table de Fisher)

Annexes C

Tableau C- 1 Quantiles d'ordre 0.95 de la loi de Fisher (table de Fisher) Degrés de liberté du numérateur sur la première ligne Degrés de liberté du dénominateur sur la colonne de gauche

	1	2	3	4	5	6	7	8	9	10
1	161.4	199.5	215.7	224.6	230.2	234.0	236.8	238.9	240.5	241.9
2	18.51	19.00	19.16	19.25	19.30	19.33	19.35	19.37	19.38	19.40
3	10.13	9.552	9.277	9.117	9.013	8.941	8.887	8.845	8.812	8.786
4	7.709	6.944	6.591	6.388	6.256	6.163	6.094	6.041	5.999	5.964
5	6.608	5.786	5.409	5.192	5.050	4.950	4.876	4.818	4.772	4.735
6	5.987	5.143	4.757	4.534	4.387	4.284	4.207	4.147	4.099	4.060
7	5.591	4.737	4.347	4.120	3.972	3.866	3.787	3.726	3.677	3.637
8	5.318	4.459	4.066	3.838	3.687	3.581	3.500	3.438	3.388	3.347
9	5.117	4.256	3.863	3.633	3.482	3.374	3.293	3.230	3.179	3.137
10	4.965	4.103	3.708	3.478	3.326	3.217	3.135	3.072	3.020	2.978
11	4.844	3.982	3.587	3.357	3.204	3.095	3.012	2.948	2.896	2.854
12	4.747	3.885	3.490	3.259	3.106	2.996	2.913	2.849	2.796	2.753
13	4.667	3.806	3.411	3.179	3.025	2.915	2.832	2.767	2.714	2.671
14	4.600	3.739	3.344	3.112	2.958	2.848	2.764	2.699	2.646	2.602
15	4.543	3.682	3.287	3.056	2.901	2.790	2.707	2.641	2.588	2.544
16	4.494	3.634	3.239	3.007	2.852	2.741	2.657	2.591	2.538	2.494
17	4.451	3.592	3.197	2.965	2.810	2.699	2.614	2.548	2.494	2.450
18	4.414	3.555	3.160	2.928	2.773	2.661	2.577	2.510	2.456	2.412
19	4.381	3.522	3.127	2.895	2.740	2.628	2.544	2.477	2.423	2.378
20	4.351	3.493	3.098	2.866	2.711	2.599	2.514	2.447	2.393	2.348
21	4.325	3.467	3.072	2.840	2.685	2.573	2.488	2.420	2.366	2.321
22	4.301	3.443	3.049	2.817	2.661	2.549	2.464	2.397	2.342	2.297
23	4.279	3.422	3.028	2.796	2.640	2.528	2.442	2.375	2.320	2.275
24	4.260	3.403	3.009	2.776	2.621	2.508	2.423	2.355	2.300	2.255
25	4.242	3.385	2.991	2.759	2.603	2.490	2.405	2.337	2.282	2.236
26	4.225	3.369	2.975	2.743	2.587	2.474	2.388	2.321	2.265	2.220
27	4.210	3.354	2.960	2.728	2.572	2.459	2.373	2.305	2.250	2.204
28	4.196	3.340	2.947	2.714	2.558	2.445	2.359	2.291	2.236	2.190
29	4.183	3.328	2.934	2.701	2.545	2.432	2.346	2.278	2.223	2.177
30	4.171	3.316	2.922	2.690	2.534	2.421	2.334	2.266	2.211	2.165
40	4.085	3.232	2.839	2.606	2.449	2.336	2.249	2.180	2.124	2.077
50	4.034	3.183	2.790	2.557	2.400	2.286	2.199	2.130	2.073	2.026
60	4.001	3.150	2.758	2.525	2.368	2.254	2.167	2.097	2.040	1.993
70	3.978	3.128	2.736	2.503	2.346	2.231	2.143	2.074	2.017	1.969
80	3.960	3.111	2.719	2.486	2.329	2.214	2.126	2.056	1.999	1.951
90	3.947	3.098	2.706	2.473	2.316	2.201	2.113	2.043	1.986	1.938
100	3.936	3.087	2.696	2.463	2.305	2.191	2.103	2.032	1.975	1.927
150	3.904	3.056	2.665	2.432	2.274	2.160	2.071	2.001	1.943	1.894
200	3.888	3.041	2.650	2.417	2.259	2.144	2.056	1.985	1.927	1.878
400	3.865	3.018	2.627	2.394	2.237	2.121	2.032	1.962	1.903	1.854

Annexes C

	11	12	13	14	15	16	17	18	19	20
1	243.0	243.9	244.7	245.4	245.9	246.5	246.9	247.3	247.7	248.0
2	19.40	19.41	19.42	19.42	19.43	19.43	19.44	19.44	19.44	19.45
3	8.763	8.745	8.729	8.715	8.703	8.692	8.683	8.675	8.667	8.660
4	5.936	5.912	5.891	5.873	5.858	5.844	5.832	5.821	5.811	5.803
5	4.704	4.678	4.655	4.636	4.619	4.604	4.590	4.579	4.568	4.558
6	4.027	4.000	3.976	3.956	3.938	3.922	3.908	3.896	3.884	3.874
7	3.603	3.575	3.550	3.529	3.511	3.494	3.480	3.467	3.455	3.445
8	3.313	3.284	3.259	3.237	3.218	3.202	3.187	3.173	3.161	3.150
9	3.102	3.073	3.048	3.025	3.006	2.989	2.974	2.960	2.948	2.936
10	2.943	2.913	2.887	2.865	2.845	2.828	2.812	2.798	2.785	2.774
11	2.818	2.788	2.761	2.739	2.719	2.701	2.685	2.671	2.658	2.646
12	2.717	2.687	2.660	2.637	2.617	2.599	2.583	2.568	2.555	2.544
13	2.635	2.604	2.577	2.554	2.533	2.515	2.499	2.484	2.471	2.459
14	2.565	2.534	2.507	2.484	2.463	2.445	2.428	2.413	2.400	2.388
15	2.507	2.475	2.448	2.424	2.403	2.385	2.368	2.353	2.340	2.328
16	2.456	2.425	2.397	2.373	2.352	2.333	2.317	2.302	2.288	2.276
17	2.413	2.381	2.353	2.329	2.308	2.289	2.272	2.257	2.243	2.230
18	2.374	2.342	2.314	2.290	2.269	2.250	2.233	2.217	2.203	2.191
19	2.340	2.308	2.280	2.256	2.234	2.215	2.198	2.182	2.168	2.155
20	2.310	2.278	2.250	2.225	2.203	2.184	2.167	2.151	2.137	2.124
21	2.283	2.250	2.222	2.197	2.176	2.156	2.139	2.123	2.109	2.096
22	2.259	2.226	2.198	2.173	2.151	2.131	2.114	2.098	2.084	2.071
23	2.236	2.204	2.175	2.150	2.128	2.109	2.091	2.075	2.061	2.048
24	2.216	2.183	2.155	2.130	2.108	2.088	2.070	2.054	2.040	2.027
25	2.198	2.165	2.136	2.111	2.089	2.069	2.051	2.035	2.021	2.007
26	2.181	2.148	2.119	2.094	2.072	2.052	2.034	2.018	2.003	1.990
27	2.166	2.132	2.103	2.078	2.056	2.036	2.018	2.002	1.987	1.974
28	2.151	2.118	2.089	2.064	2.041	2.021	2.003	1.987	1.972	1.959
29	2.138	2.104	2.075	2.050	2.027	2.007	1.989	1.973	1.958	1.945
30	2.126	2.092	2.063	2.037	2.015	1.995	1.976	1.960	1.945	1.932
40	2.038	2.003	1.974	1.948	1.924	1.904	1.885	1.868	1.853	1.839
50	1.986	1.952	1.921	1.895	1.871	1.850	1.831	1.814	1.798	1.784
60	1.952	1.917	1.887	1.860	1.836	1.815	1.796	1.778	1.763	1.748
70	1.928	1.893	1.863	1.836	1.812	1.790	1.771	1.753	1.737	1.722
80	1.910	1.875	1.845	1.817	1.793	1.772	1.752	1.734	1.718	1.703
90	1.897	1.861	1.830	1.803	1.779	1.757	1.737	1.720	1.703	1.688
100	1.886	1.850	1.819	1.792	1.768	1.746	1.726	1.708	1.691	1.676
150	1.853	1.817	1.786	1.758	1.734	1.711	1.691	1.673	1.656	1.641
200	1.837	1.801	1.769	1.742	1.717	1.694	1.674	1.656	1.639	1.623
400	1.813	1.776	1.745	1.717	1.691	1.669	1.648	1.630	1.613	1.597

F-value de 11 / 58 degré de liberté :

$$F_{11/50} = 1.986 \text{ et } F_{11/60} = 1.952 ; \text{ par interpolation linéaire on a } \frac{60-50}{1.986-1.952} = \frac{58-50}{F_{11/58}-1.952}$$

On obtient donc $F_{11/58} = 1.979$

Annexe D

Publication

Improvement of Productivity in Buildings Construction

Chemseddine Dehchar^{1*}, Khaled Boudjellal¹, Mohamed Bouabaz¹

¹Faculty of Technology, Department of Civil Engineering, LMGHU Laboratory, University of 20 August 1955, Skikda, Algeria

*e-mail: ch.dehchar@univ-skikda.dz / dehcharchemseddine@hotmail.com

Abstract

Improving productivity in construction projects has long been a major concern, and much research has been carried out to try to ameliorate construction productivity. To this end, this study aims to improve and increase the productivity rate of flat slab formwork used in residential construction projects. A survey consisting of 150 questionnaires was undertaken to identify the factors that influence on the productivity. Based on the relative Importance Index (RII), data on eleven factors deemed to affect productivity were selected. A collection of 100 data points from various sites were utilized to develop two models. Firstly, an Artificial Neural Network (ANN) model was employed, and secondly, a parametric approach was investigated. The data were divided into two sets, with 70% of the data used for training and the remaining 30% used for testing. The models' performance was evaluated using the Mean Squared Error (MSE) and Mean Absolute Percentage Error (MAPE) values. In the test phase, the artificial neural network model yielded an MSE value of $2.6610e^{-4}$ and a MAPE value of 4.9227, whereas the parametric model produced an MSE of 0.040 and a MAPE of 9.525. It was found that the artificial neural network model provided reliable prediction accuracy compared to the parametric model. However, the artificial neural network approach can be selected as a robust model in predicting and controlling the productivity rate in local construction projects by using the developed model based on the identified factors.

Keywords: Artificial Intelligence, Construction Industry, Artificial Neural Network, Labour Productivity, Parametric Approach

1 Introduction

In the construction industry, productivity can be defined as the quantity of production work per corresponding input [1]. This relationship can take different forms, defining labour productivity when using work-hour as an input or cost productivity when cost input is taken into account [2, 3]; however, this simple definition of factors does not seem realistic and cannot give robust and accurate results, as omitting the influence of other factors can affect productivity. Various research studies have been conducted over many years to estimate, predict and improve the productivity rate in the construction industry, taking into account the effect of different influencing factors.

The review of the literature in the area of construction and specifically productivity improvement can be divided into two main parts, the first focuses on the identification of factors affecting productivity, and the second deals with applications of artificial intelligent.

The pace of labour productivity is a critical factor that directly influences the outcome of projects, ultimately determining their success or failure [4, 5]. According to [6], factors affecting labour productivity may vary from country to country and from site to site, and may also vary within a site. [7] Stated that, labour productivity is influenced by various factors, including external conditions, site conditions, and workers' characteristics. The combined influence of these factors determines the overall level of efficiency and effectiveness attained in construction tasks. [8] Invited 180 Kuwaiti companies to participate in a sample survey that included 45 factors grouped into four groups, to identify the factors most influencing Kuwaiti projects, the result of the statistical analysis determined that the total project cost could be reduced by focusing on the design phase. The influence of several factors on labour productivity in Malaysian construction sector was examined in [9], the authors found that project management skills were the factors that had the greatest impact on productivity, and also stated that the implementation of new technologies in construction sector has positive effects on labour productivity. Another study by [10] highlights that the management group ranks first in three categories, which includes 30 factors studied for their effects on labour productivity in Egyptian construction firms, the authors also found that workers experience and skills were the most important factor. A questionnaire survey conducted in Trinidad and Tobago by [11], to identify the influencing factors on construction projects, were investigated on 30 contractors who are members of the Contractors Association. The survey included 42 factors categorized into four groups: Management, Technological, Human/Labour, and External. Both close-ended and open-ended queries were used, and the respondents were asked to score each factor based on a Likert Scale, ranging from 1 to 4, indicating the effect level. To determine the rank of each factor, the researchers calculated the Relative Importance Index (RII%) and ranked the factors accordingly. The top three influencing factors identified were: the lack of labour supervision, unrealistic scheduling and expectation of labour performance, shortage of experienced labour. In [12], 52 factors, were used to investigate their influence on productivity in Yemen, by calculating the relative importance index, the authors found that workers experience and skills were the most significant factors. [13] Proposed a fuzzy fault tree-based approach to identify factors influencing labour productivity in Lithuania. Initially, they gathered data from 15 experts who evaluated 27 factors using a Likert scale, ranging from 1 to 5, to indicate the degree of severity. After this step, 18 factors with a Relative Importance Index (RII) greater than 0.7 were selected for the subsequent analysis. In the next phase, the authors constructed a fault tree structure through multiple sessions and expert interviews to define the interrelations among the selected factors. Finally, the fuzzy approach was employed to evaluate the contribution of each factor. The results of their analysis indicated that two factors, namely "inflation in the cost of execution" and "improper project financing," were identified as the most influencing factors impacting labour productivity in Lithuania. Another study conducted by [14] aimed to identify the factors affecting labour productivity in the South Africa construction industry based on contractors' perceptions. A questionnaire survey consisting of 41 factors was randomly distributed to 96 contractors. The obtained responses were then analysed using descriptive and inferential statistics. The factors were ranked according to their P-values, and the most significant influencing factors on labour productivity were found to be: excessive bureaucracy, late delivery of materials, industrial action resulting from political activities. Authors in [15] identified factors affecting labour productivity in Australian construction projects by following a three-step approach. In the first step, they conducted a questionnaire to gather data on various factors related to labour productivity in construction projects. The second step involved drawing

a cause-and-effect feedback loop to identify the complicated interrelated links between the 38 factors. Lastly, the authors used the Decision-Making Trial and Evaluation Laboratory (DEMATEL) method to prioritize the factors. The results of the study revealed that project size, level of skill and experience, and communication problems with foreign workers were the most significant factors affecting labour productivity in Australian multi-storey building construction projects. In another research conducted by [16], authors examined the influence of weather conditions on labour productivity in the United Kingdom. The study involved collecting data over a span of 5 years, from September 2017 to August 2022. To gather information, the researchers administered 28 online questionnaire surveys to relevant participants. Based on their findings, the authors proposed implementing modular techniques in construction, wherein certain elements are produced in factory settings with controlled environments. This approach is expected to mitigate the adverse effects of weather on construction projects, consequently reducing overall costs.

In the second half of the 1980s, Artificial Neural Networks were developed and used to solve various construction problems such as cost estimation, duration and productivity. Numerous models were produced [17–19], and are still in progress to this day. [20] Examined the influence of ten factors, on Turkish formworkers' productivity. A three-layer feed-forward back propagation neural network was used, using these factors as input variables, a hidden layer containing five hidden neurons and an output neuron to estimate the labourers' productivity rate, compared to the estimate of the Turkish Ministry of Public Works and Settlement, the simulation results indicate that the model developed accurately predicted the total man-hours required. In another approach, [21], also formulated a multi-layer ANN with the backpropagation algorithm using ten influencing factors, as input variables to assess the labour productivity rate for finishing work on marble floors. The architecture of the model used was a three-layer neural network with one hidden node in the hidden layer and used the hyperbolic tangent sigmoid function as a transfer function, a degree of accuracy of 90.9% was recorded, which proved that Network, the Backpropagation Neural Network, the Radial Base Function Neural Network and the Adaptive Neuro-Fuzzy Inference System, and compared their results to choose the best model the developed model can reliably predict the productivity rate, it was also found that the most influential factors on productivity were age, number of workers and experience. Two ANNs models with different architectures were developed by [22], using three categories of influencing factors, the capabilities of a Feed-Forward neural network and a Radial-Based NN to predict masonry crew productivity were compared; the results showed that the RBNN technique predicted productivity better than the FFNN. Authors in [23] took thirteen factors that were found to have a negative effect on labour productivity and considered them as inputs for developing an ANN model to estimate bricklayer productivity, a gradient descent momentum backpropagation was used as a training algorithm for the model which has three hidden layers and a hyperbolic tangent sigmoid transfer function; accurate results were obtained using this architecture, in addition, the results of the sensitivity analyses led to the conclusion that the wall thickness was the factor that most influenced the labour productivity rate. [24] Developed four ANN models based on different techniques, including the General Regression Neural for predicting the productivity rate of formwork labours, the performance results show that the Backpropagation NN is the best model among the four techniques developed in the study. Authors in [25] stated that labour productivity in Saudi construction projects is affected by a number of factors and can be quantified using mathematical models. For this purpose, the author developed four different models based on multilayer perceptron neural network

(MLPNN), general regression neural network (GRNN), support vector machine (SVM), and multiple additive regression trees (MART) to assess the productivity rate of concrete construction activities and compared the results of the developed models. The GRNN model was found to be the best estimator model for labour productivity of steel fixing and concrete pouring, while the MART model provided better results than the others for formwork assembly productivity. In order to ensure effective and efficient project delivery, [26] utilized an artificial neural network (ANN) to assess the relationships between various factors and bricklaying productivity during the pre-planning phase. The developed model included 7 input factors, 5 hidden neurons, and one output. The results showed good performance, indicating that the ANN model can be utilized to enhance project delivery by providing a more realistic estimation of productivity rates. This, in turn, leads to better planning, cost estimation, and resource allocation. In a study conducted by [27], authors used sets of five Artificial Neural Network models to develop a single model capable of estimating steel reinforcement works. They collected 145 data points from construction sites in Poland, which were then divided into 90% for training and 10% for testing purposes. The proposed model demonstrated satisfactory performance, indicating its robustness in estimating steel reinforcement works. [28] Developed an Artificial Neural Network (ANN) model to assess the productivity rate of piping assembly activities in Brazilian projects. They tested 108 different Feedforward Neural Network (FFNN) architectures, each with 14 input neurons, to select the best model. The research results indicated that the Mean Square Error (MSE) of the chosen model reached the value of 9.67E-04, demonstrating satisfactory performance and accuracy. A hybrid model combining artificial neural network (ANN) and Grasshopper Optimisation Algorithm (GOA) was developed by [29]. The objective was to identify the factors with the greatest impact on labour productivity in Iranian construction projects and improve the accuracy of labour productivity prediction. Out of the 19 factors initially considered, only 6 factors were identified as influential and utilized to develop the predictive ANN model. [30] Employed Artificial Neural Networks (ANN) to predict the productivity rate for brick masonry work. The authors conducted a questionnaire survey that initially included forty-four factors. However, based on the Relative Importance Index (RII) analysis, only the top 13 ranking factors were considered as inputs for the model. The developed model utilized a three-layer backpropagation-feedforward neural network with one hidden layer consisting of 30 hidden neurons. The Mean Squared Error (MSE) was used as a measure of accuracy for the model. The results indicated that the developed model effectively estimated brick masonry productivity by considering the impact of the selected 13 factors. Parametric model was developed in the early 1960s, it has been used to find the relationship between the response and one or more factors, usually by applying regression analysis to the parameters [31]. [32] Developed a parametric equation based on multiple linear regression (MLR) analysis to model the labour productivity of marble floor finishing works in Iraq. 100 data including 10 influencing factors were collected and used to predict productivity, good performance, and a high correlation was obtained, indicating a good relationship between the response and the factors, and the model can be reliably used to estimate productivity. In another study conducted by [33] to estimate the percentage loss or increase in construction labour productivity, 14 influencing factors were defined and used to develop a regression model, the author argued that the proposed model can accurately predict construction labour productivity. [34] Examined the influence of various buildability factors on the labour productivity of structural elements, by developing a MLR model to assess the relationship between design characteristics and labour productivity. The results show that the model can provide information

for designers to adapt their designs to achieve the best labour productivity performance. Authors in [35] conducted a study to examine the impact of worker experience and skill level on masonry construction productivity in Vietnam. They developed a logistic regression model to estimate productivity and found that there was no difference in average productivity between different worker experience groups. However, they did observe a significant difference in productivity between worker skill groups. [36] Employed a multiple linear regression method to analyse the impact of three factors on labour productivity in lightweight brick wall installation and lightweight brick wall plastering in Indonesia. The researchers achieved an impressive accuracy of 99.43% and 99.04% for the two respective models, indicating the robustness and reliability of the developed models. In their study conducted at eight construction sites in Niš, Serbia, [37] Focused on analysing concreting processes, particularly those related to columns and walls. They gathered 60 data points to develop regression analysis and Simulation models aimed at improving the accuracy of forecasting productivity rates for concrete works. The findings of the study indicated that these proposed models have the potential to enhance decision-making and improve the accuracy of planning in construction projects involving residential buildings in Niš, Serbia.

The Algerian construction industry is currently facing a major problem in accurately measuring construction labor productivity. This issue is mainly attributed to the continued utilisation of traditional technologies, which has significantly hindered progress in the construction sector. The challenges are particularly evident in the execution of flat slab formwork activities, resulting in excessive resource consumption in terms of manpower, time, and cost. To enhance construction efficiency, time management, and cost control, it is crucial to address the productivity issues related to flat slab formwork. By optimizing the flat slab formwork productivity in Algerian building projects, it becomes possible to achieve faster construction cycles, lower labor costs, and improve overall project performance. This, in turn, leads to more efficient resource utilization and successful project outcomes.

While previous studies have examined construction productivity in a general context, however, there appears to be a noticeable gap in the existing research regarding comprehensive investigations into the factors influencing productivity, particularly in Algerian construction projects. Additionally, there is a lack of intelligent methods that specifically concentrate on accurately measuring and estimating labor productivity rates for flat slab formwork within the context of the local construction industry. Since our study focuses on a local problem with specific characteristics compared to other world regions, we are faced with the following issues:

- What are the key factors that influence productivity in Algerian construction projects?
- How do the identified factors influence the productivity rate of flat slab formwork in Algerian construction projects?
- How can we estimate this productivity?

This paper aims to address two primary objectives: identifying the key factors that affect productivity in Algerian construction projects and developing intelligent models capable of estimating the labour productivity rate of flat slab formwork. By conducting research on this topic, the study intends to fill the existing knowledge gap and contribute valuable insights to both academia and industry practitioners. The findings of this research will provide valuable guidance for productivity improvement and decision-making in the management of construction projects in Algeria.

2 Methodology

This study aims to improve the labour productivity rate of flat slab formwork in Algerian building projects. To achieve the objective mentioned, the methodology presented in this section was adopted.

2.1 Questionnaire Survey Analysis

Identifying factors that negatively affect labour productivity in construction projects has been a major concern for managers [32, 38]. After careful analysis of the top influential factors cited by researchers [10, 12], [21-23], [39-43], the factors that appeared repeatedly were identified and extracted. Moreover, by considering on-site observations of construction projects, a questionnaire survey was designed including 16 influencing factors classified into three main groups as shown in Table 1. The survey consisted of closed-ended questions that were sent by email to local expert, who were asked to rate the impact of each factor on labour productivity using a five-point Likert scale, ranging from 1 (indicating the least impact) to 5 (indicating the greatest impact). Respondents were specifically asked to indicate the extent to which a specific factor affected labour productivity according to their own perception. The designed questionnaire survey was divided into three sections:

- 1 The first shows the respondents' personal information (position and professional experience)
- 2 The second section consists of three groups with 16 influencing factors, the respondents rank the importance of each of them based on a five-point Likert scale, where 1 indicates the least impact while 5 represents the greatest impact.
- 3 The last section represents a space dedicated to the respondents' opinions.

Table 1: Questionnaire survey factors

Questions (Factors)		
Group 01: Management	Group 02: Factors related to workers	Group 03: External factors
Small crew size		Weather condition
Low wages		Humidity
Formwork condition	Work-hour/absenteeism	High temperature
Delay in payment of the company by the project owner	Worker's experience	Low temperature
Congestion in the workspace	Quantity of installed work	Shortage of building materials
Poor control	Worker's age	
Low workers motivation		

2.1.1 Sample Size

For a more diverse and precise analysis, we specifically targeted project managers, technical managers, site managers, civil engineers, and architects working in the public and private sectors to gather their perceptions and viewpoints on the various factors that influence labour productivity in the Algerian construction industry. A statically representative sample of the population was obtained using a formulae developed by [12, 44], as expressed bellow:

$$n = \frac{m}{1 + \frac{(m-1)}{N}} \quad (1)$$

where, n is the sample size of the limited population, m is the sample size of the unlimited population, and N is the sample size of the available population. m is estimated as follows:

$$m = \frac{z^2 \times p(1-p)}{\varepsilon^2} \quad (2)$$

z represents the statistic value for the confidence level used ($z = 1.645$ for 90% confidence level), ε is the sampling error of the point estimate, and p is the value of the population proportion that is being estimated, for which [45], proposed using a conservative value of $p = 0.50$ to obtain a sample size at least as large as needed. The value of m can be obtained than:

$$m = \frac{(1.645)^2 \times 0.50(1-0.50)}{0.1^2} = 67.65 \approx 68 \quad (3)$$

In this study, a total number of 150 available samples were selected, based on this; the required representative sample size of the population is obtained as follows:

$$n = \frac{68}{1 + \frac{(68-1)}{150}} \approx 47 \quad (4)$$

A total of 150 samples were distributed by email and only 56 feedbacks were received, this number of completed questionnaires met and exceeded the required sample size. The respondents in the samples are project managers, technical managers, site managers, civil engineers, and architects, working in Algerian construction industry. 48.22% of them having an experience of 1 to 4 years, while 44.64% have an experience of 05 to 09 years, and the remaining 7.14% have 10 or more years of experience.

2.1.2 Defining the Influencing Factors

The Relative Importance Index (RII) technique was used to define the most influential factors; this index was calculated according to equation (6).

$$RII(\%) = \left(\frac{\sum_{i=1}^5 n_i \times x_i}{5 \sum_{i=1}^5 x_i} \right) \times 100 \quad (6)$$

where, n_i indicates the Likert scale from 1 to 5, and x_i denotes the frequency of each n_i . The results of the RII show that three factors had an RII greater than 80%, whereas nine factors have an RII between 70 and 80%, while the RII of the other four factors is in the range of 60% and 70%. Table 2 summarizes the results of RII and the rank of each factor.

Table 2: RII (%) Results and rank of each factor

Factors	RII (%)	Rank
Small crew size	78.21	07
Low wages	89.64	01
Formwork condition	84.29	03
Congestion in the workspace.	71.79	12

Poor control	74.64	09
Delay in payment of the company by the project owner	74.29	10
Low workers motivation	86.07	02
Work-hour /absenteeism	78.93	04
Worker's experience	78.57	05
Quantity of installed work	78.57	05
Worker's age	68.57	13
Weather condition	68.57	13
Humidity	65.00	16
High temperature	75.36	08
Low temperature	67.14	15
Shortage of building materials	73.93	11

In this study, factors with an RII (%) greater than 75% were taken into account. On this basis, the most relevant factors deemed to affect the productivity rate are: workers' wages, workers' motivation, formwork condition, number of working hours, workers' experience, crew size, quantity of installed work, and temperature. Table 3 summarises all the factors considered.

Table 3: The selected influential factors

Factors	Categories	Classification
Number of working hours	Quantitative	Numerical
Number of formworkers	Quantitative	Numerical
Number of unskilled workers	Quantitative	Numerical
Minimum workers' experience	Quantitative	Numerical
Maximum workers' experience	Quantitative	Numerical
Formwork condition	Qualitative	Mediocre Average Good
Type of workers' motivation	Qualitative	Extra money Recuperation days Piecework
Temperature	Quantitative	Numerical
Wage of formworkers	Quantitative	Numerical
Wage of unskilled workers	Quantitative	Numerical
Surface area of the slab formed	Quantitative	Numerical

2.2 Data Collection

The data for this study were collected daily, based on in site observations on the flat slab formwork task, a number of 100 data samples, comprising 11 quantitative and qualitative influencing factors, were compiled from construction projects in different regions of Algeria, the qualitative factors used in this study were transformed into numerical values for use in the models (Table 4). As indicated in the previous sections, productivity can be calculated in different ways (total or partial factor productivity) [46], the type of labour productivity, which is of interest in this study, is defined as the ratio of output to worker-hour (Equation 7), noting that the output presents the surface area of the slab that was carried out in square metres.

$$productivity\ rate = \frac{output}{man \times work\ hour} \quad (7)$$

Table: 4 Transformation of qualitative factors into numerical form

Qualitative factors		Numerical transformation
Formwork state	Mediocre	-0.50
	Average	0.00
	Good	0.50
Type of motivation	Extra money	0.10
	Recuperation days	0.15
	Piecework	0.20

2.3 Artificial Neural Network Model Development

2.3.1 Data Preprocessing

The preprocessing phase consists of preparing the data before introducing them into the model, which gives the network the capacity to converge more quickly and to better generalise the results obtained. In this phase the input and target data were normalised to be within the range of [0 to 1] both for the training and the test phase using equation (8).

$$X_n = \frac{X_i - X_{i_{min}}}{X_{i_{max}} - X_{i_{min}}} \quad (8)$$

where: X_n : the value of the variable (x_i) after normalization. And X_i : the value of the variable before normalization. X_i min; X_i max are the minimum and maximum value of (X_i).

2.3.2 Data Processing

Is the process by which the collected data is organised and introduces into the ANN model, to obtain better generalisation of the results, for this purpose, the designed set was divided into training and test subsets, the training subset is used to initialise the weights and biases of the network. According to [47], each network starts training from arbitrary initial weights and biases, while the test subset is used to verify the network design, testing the efficiency of the model on the new data. During the modelling phase, a phenomenon can occur when errors begin to increase in front of the test data after providing very small values in the training sets, this phenomenon is known as over fitting, in this case, the model adjusted its parameters well and stored all the training courses, but when new data was introduced into the network, the model failed to adapt and generalise the new situation. Different methods have been used to avoid this phenomenon; in this study the Bayesian regularisation technique was used.

Simulation software was utilized to create, train and test the networks, with random division of the designed set into 70% training, and 30% into test sets.

2.3.3 Model Architecture and Adapted Algorithm

The model architecture represents the number of appropriate layers, and the way these layers are connected, as well as the type of algorithm chosen. In this study, several architectures were tested to obtain better accuracy, the model with three layers having 11 input neurons, two hidden layers having 15 and 10 hidden neurons in the first and second hidden layers respectively, and

one output neuron as presented in Figure 1. Backpropagation was used as a learning algorithm with Bayesian regulation.

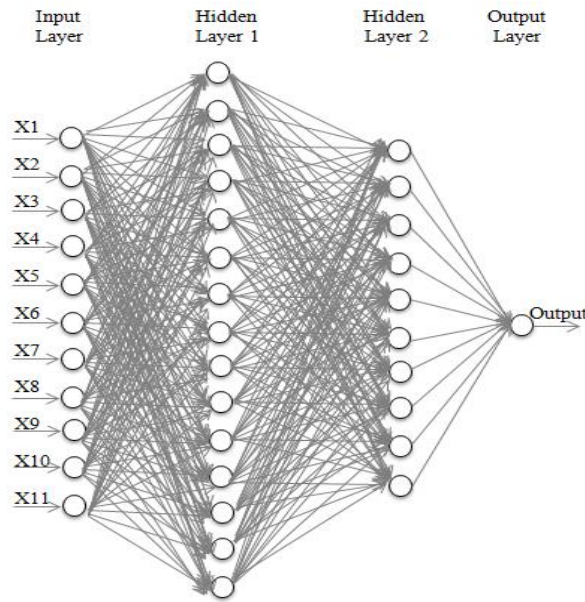


Figure 1: The architecture of the ANN model developed

2.3.1 Transfer Function

The transfer function is a monotonic, continuous, increasing and differentiable function applied to the sum of the weighted inputs of the model to produce their resulting outputs. This function consists of two sub-functions used by the neurons in the different layers. Firstly, the neuron in each layer receives a linear combination of weighted inputs, determining the sum of the information collected, using the activation function $A(x)$ which is given as follows:

$$A_i(x) = \sum_{j=1}^N (W_{ij} \times X_j) - \theta_i \quad (9)$$

where: X_j : the inputs, and W_{ij} : the weights and θ_i is the threshold (bias).

In the next step, the neurons use an $O(A)$ output function, to keep the output values within a specified range, operating it on the activation function which has a scalar format and returns the scalar output neurons. Thus, the transfer function $O(A(x))$ is the combination of the two previous functions (activation and output function), which is used to produce the final output neurons [48]. Different transfer functions are available in the Artificial Neural Network model used to pass information from one layer to the other [47], the most commonly used are log-sigmoid, tan-sigmoid, and linear (purelin). Each function of the latter has its own range that generates output results within it. The logsig produces results between 0 and 1 while tansig returns outputs in the range of -1 and 1, however, the specific range of purelin is: $[-\infty + \infty]$.

The transfer functions adopted in this study for the implementation of the ANN model were the log-sigmoid transfer function (logsig) and the hyperbolic tangent sigmoid (tansig) which were used in the first and second hidden layers, respectively, while the pureline linear transfer function was used in the output layer equation (10), equation (11) and equation (12).

$$\text{Log sig}(A) = \frac{1}{(1 + \exp(-A))} \quad (10)$$

$$\text{Tansig}(A) = \frac{2}{(1 + \exp(-2A))} - 1 \quad (11)$$

$$F(A) = A \quad (12)$$

2.3.2 Training and Testing Phase

Network training is the first process carried out by the model to select its parameters, it consists of adjusting the values of weights and biases until a minimum error value is reached and the best performance is obtained. The Bayesian regularization training algorithm was used in this study to obtain better results. This algorithm consists of adding an additional term (penalty term), which represents the sum of the squares weights to the equation error, and then propagating the resulting errors back to readjust the weights and biases, and minimizing the errors to obtain the best prediction of the output.

Out of the 100 datasets collected, a number of 70 datasets were used for training in the first phase, while the remaining 30 datasets were retained to test the model in order to generalize the model to the new data, and to obtain better performance in predicting the results.

The Mean Square Error (MSE) and the Mean Absolute Percentage Error (MAPE) were used in this study to calculate the performance of the model, as follows:

$$MSE = \frac{\sum_{i=1}^n (x_i - E_i)^2}{n} \quad (13)$$

$$MAPE = \frac{100}{n} \times \sum_{i=1}^n \left| \frac{x_i - E_i}{x_i} \right| \quad (14)$$

where, n is the number of the data, E is the model outputs, and x_i is the actual value.

2.4 Parametric Analysis

Parametric analysis is a method that links a variable to be estimated (labour productivity) and the different factors that influence it through a mathematical relationship. The technique most commonly used to develop the parametric equation is Multiple Linear Regression (MLR).

2.4.1 Development of Multiple Linear Regression Model

At this stage, 70% of the total data collected, were used to develop the model using Statistical Product and Solutions Services (SPSS) software. The correlation test was carried out to check the relationship between the response and the explanatory variables. The results of the correlation coefficient (R) and determination coefficient (R²) presented in Table 5 show a strong correlation between the response (flat slab formwork productivity rate) and the

explanatory variables, indicating that they have a good relationship. A significant value for the model was found to be less than 0.05, meaning that the regression line fits the data well.

Table 5: Model Results Summary

Coefficient of correlation (R)	Coefficient of determination (R ²)	Significance of the model
0.958	0.917	0.000

Table 6 summarises the results of MLR coefficients and the resulting equation for estimating the productivity rate of flat slab formwork in construction projects takes the form below:

$$\begin{aligned}
 Y = & 3,239 - 0,304X_1 - 0,088X_2 - 0,139X_3 - 0,021X_4 \\
 & + 0,007X_5 - 0,015X_6 - 1,356X_7 + 0,011X_8 \\
 & - 0,004X_9 + 0,000103 X_{10} - 0,000053 X_{11}
 \end{aligned}
 \tag{15}$$

Table 6: Coefficients of Multiple Linear Regression

Model	Unstandardized Coefficients	
	β	Std. Error
(Constant)	3.239	0.716
X1	-0.304	0.024
X2	-0.088	0.035
X3	-0.139	0.014
X4	-0.021	0.024
X5	0.007	0.010
X6	-0.015	0.047
X7	-1.356	1.302
X8	0.011	0.001
X9	-0.004	0.006
X10	0.000103	0.000046
X11	-0.000053	0.000039

2.4.2 Validation of the Developed Model

Model validation is the stage during which the new data was used on the explanatory variables (factors) to test the ability of the parametric equation developed in equation (15), to estimate and predict the productivity rate of the flat slab formwork. In this respect, a set of 30 data was applied to the developed parametric equation, and the results of the predicted productivity rate were compared with the actual productivity rate. The MSE and MAPE measures were calculated to verify the performance of the model with the new data.

3 Results and Discussion

This paper aims to develop a machine learning model capable of predicting and estimating the labour productivity rate for flat slab formwork assembly. In this respect, an Artificial Neural Network was first trained using 70% of the data sets and tested using 30% of the data sets. Various ANN architectures were tested with different combinations in the number of hidden

layers and hidden neurons as well as in the learning algorithm, and in the different ANN parameters, to choose the best model that gives the best performance, the final structure adopted in this study was a three-layer feed-forward back-propagation neural network having two hidden layers with 15 hidden neurons and 10 hidden neurons in the first and second hidden layers respectively. The Bayesian regularization technique was adopted as the learning algorithm. The modelling phase consists of training the network first; in this step, the model shows its ability to predict the results of the labour productivity rate with lower values of the MSE and the MAPE, as shown in Figure 2. Testing the network is the next step carried out by the model to examine its ability to handle unseen data. Good results from MSE and MAPE were obtained at this stage as well, showing the model's ability to predict the labour productivity rates of flat slab formwork, as shown in Figure 3. In line with the results from the MSE and MAPE, it can be said that the model developed performs well and that no significant difference could be observed between the predicted and actual labour productivity rates. Table 7 summarises the overall results of the measures. The regression analysis in the developed ANN model was carried out to examine the relationship between actual targets and estimated output, a correlation coefficient of 0.9994 was obtained indicating a good linear correlation between forecasted and actual labour productivity (Figure 4).

In the second part of the study, a parametric analysis of the collected data was carried out by developing a Multiple Linear Regression model. The results of the correlation coefficient and the determination coefficient rather than the significant value of the developed equation show a strong correlation between the response (labour productivity rate) and the explanatory variables (factors). A good result from MSE and MAPE was obtained, indicating that the parametric equation developed has the capacity to estimate and predict the labour productivity rates of flat slab formwork (Figure 2). A set of 30 data was applied to the developed parametric equation to validate it and the results of the predicted labour productivity rates were compared to the actual labour productivity rates. Based on the MSE and MAPE results obtained during the validation phase, it can be stated that the parametric equation developed works well and can be used reliably to predict the labour productivity rates of flat slab formwork (Figure 3). Table 8 lists the results of the performance of the parametric model.

Finally, from the results of the two methods above, it can be seen that both the ANN model and the parametric model can be reliably used to estimate and predict the labour productivity rate. However, significant differences between the performances of the two models developed were observed, so we can conclude that the ANN model gives better results and predicts the labour productivity rates of flat slab formwork better than the parametric model.

Table 7: Results of the Artificial Neural Network performance

Phase	Performance	MSE	MAPE %
Training phase		8.0796e-05	0.82832
Test phase		2.6610e-04	4.9227

Table 8: Results of the parametric equation performance

Phase	Performance	MSE	MAPE %
Development phase		0.048	9.395
Validation phase		0.040	9.525

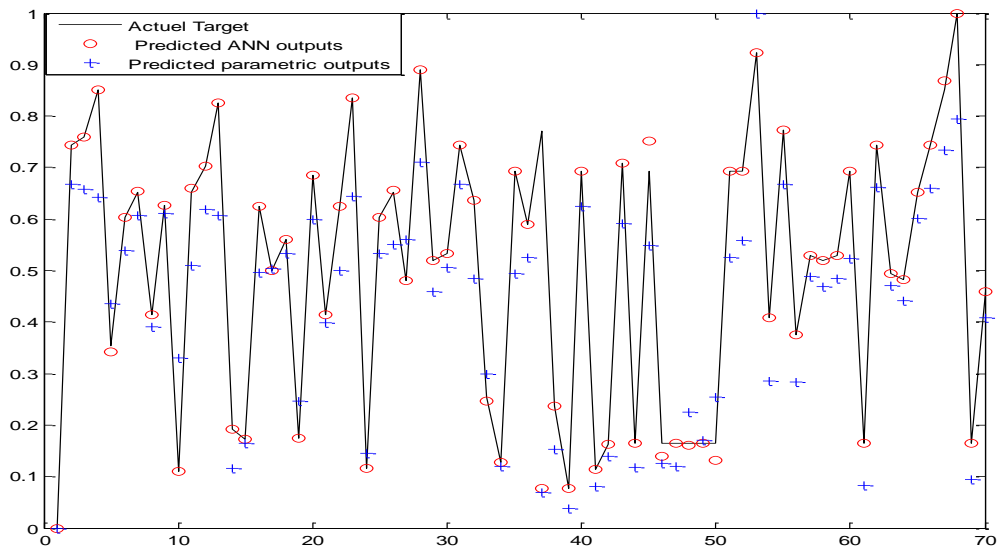


Figure 2: Predicted Neural Network and parametric outputs versus actual targets in the training phase

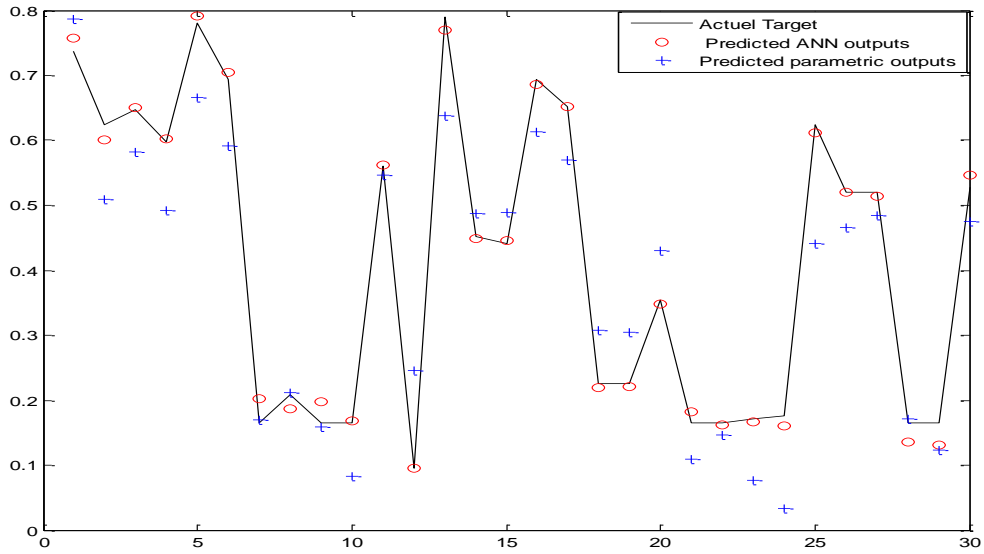


Figure 3: Predicted Neural Network and parametric outputs versus actual targets in the test phase

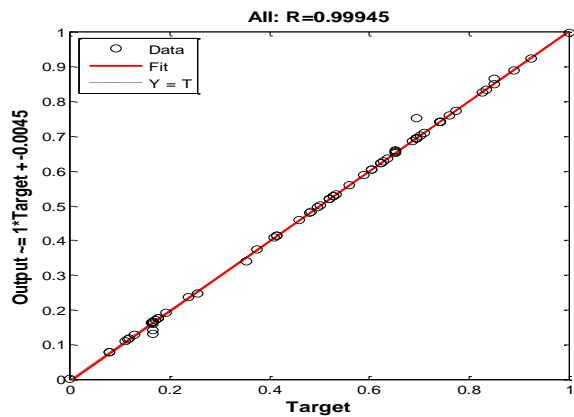


Figure 4: The regression analysis and correlation coefficient R in ANN model

Additionally, we investigated samples of previous studies conducted by many researchers on formwork productivity modelling. Results were reported regarding the performance of each model. For instance, [49] obtained a MAPE of 14.55% and 17.69% for the ANN and MLR models, respectively, for column formwork productivity. In another study by [25], the MART model was found to be the best among the four developed models, with an MSE of 0.0937. Furthermore, [24] developed four models for slabs, walls, and columns formwork productivity, it was found that Backpropagation Neural Network was the best model which achieved an MSE of 0.018.

Table 9 describes a comparison study conducted on existing models versus the developed models.

Table 9: Results comparison among existing models

Existing Research	The used model	Performances results in the test phase
Developed research study	Artificial Neural Network	MSE = 2.6610e-04 MAPE = 4.9227 %
	Multiple Linear Regression	MSE = 0.040 MAPE = 9.525 %
S. Nassar and A. Khaleel. (2019)	Artificial Neural Network	MAPE = 14.55 %
	Multiple Linear Regression	MAPE = 17.69 %
E. A. Mlybari. (2020)	MART	MSE = 0.0937
Golnaraghi et al. (2019)	Backpropagation Artificial Neural Network	MSE = 0.018

By comparing the outcomes of the mentioned models with models developed in this study, it can be observed that the latter exhibits significantly higher performance accuracy. The achieved MSE and MAPE values suggest that our ANN and MLR models are better in predicting formwork productivity and providing more accurate results. This confirms the reliability and efficiency of the methods used in this study, which leading to a significant advancement in the field of formwork productivity modelling.

4 Conclusion

This study presented a development of a multilayer Artificial Neural Network and a parametric equation based on MLR models aimed at predicting and increasing the labour productivity rate of the flat slab formwork in Algerian construction projects. A survey of 150 questionnaires was distributed to define the most influential factors. The total number of data collected on the 11 influencing factors defined in this study was 100, which were then randomly divided into two main subsets. An Artificial Neural Network model was first developed using 70% of the data set, to train the network to update weights and biases, and the remaining 30% was used to test

the network for better generalisation on the new data. The model architecture used was a three-layer feed-forward ANN with a backpropagation algorithm and Bayesian regularisation as a learning function. A parametric equation was developed in the second part of this study, also using the same 70% of the dataset to establish a mathematical equation based on the multiple linear regression method. The remaining 30% of the dataset was used to validate the equation against the new data.

The results of this study clearly show that the ANN model developed is more efficient and gives more accurate prediction results than the parametric equation. However, it can be noticed that both developed models have a good relationship between the predicted results and the actual targets. Moreover, a high performance was recorded, indicating that both models behave well with the unseen data and produce results reasonably close to the actual targets. The results of this research study allow concluding that the ANN model and the parametric equation developed can be reliably used to predict the measurement of the productivity rate with the incorporation of influencing factors deemed to affect the labour productivity of the flat slab formwork task.

References

- [1] Gerek, I. H., Erdis, E., Mistikoglu, G. & Usmen, M. A. (2014). Evaluation of plastering crew performance in building projects using data envelopment analysis. *Technological and Economic Development of Economy*, 22(6), 926–940.
- [2] Thomas, H. R., Maloney, W. F., Horner, R. M. W., Smith, G. R., Handa, V. K. & Sanders, S. R. (1990). Modeling construction labor productivity. *Journal of Construction Engineering and Management*, 116(4), 705–726.
- [3] Yi, W. & Chan, A. P. C. (2014). Critical review of labor productivity research in construction journals. *Journal of Management in Engineering*, 30(2), 214–225.
- [4] Boughaba, A. & Bouabaz, M. (2020). Identification and risk management related to construction projects. *Advances in Computational Design*, 5(4), 445–465.
- [5] Boulkenafet, N., Boudjellal, K. & Bouabaz, M. (2023). Optimization of productivity in the rehabilitation of building linked to BIM. *Advances in Computational Design*, 8(2), 179–190.
- [6] Olomolaiye, P. O., Jayawardane, A. K. W. & Harris, F. C. (1998). *Construction Productivity Management*. England: Addison Wesley Longman.
- [7] Florez-Perez, L. Song, Z. & Cortissoz, J. C. (2022). Predicting construction productivity with machine learning approaches. Proceedings of the International Symposium on Automation and Robotics in Construction. 2022-July, (107–114).
- [8] Jarkas, A. M. & Bitar, C. G. (2012). Factors affecting construction labor productivity in Kuwait. *Journal of Construction Engineering and Management*. 138(7), 811–820.
- [9] Karim, N. A., Hassan, S., Yunus, J. & Hashim, M. (2013). Factors Influence Labour Productivity and the Impacts on Construction Industry. *Caspian Journal of Applied Sciences Research*, 2(August), 349–354.
- [10] El-Gohary, K. M. & Aziz, R. F. (2013). Factors influencing construction labor productivity in Egypt. *Journal of Management in Engineering*, 30(1), 1–9.
- [11] Hickson, B. G. & Ellis, L. A. (2014). Factors affecting Construction Labour Productivity in Trinidad and Tobago. *The Journal of the Association of Professional Engineers of Trinidad and Tobago*, 42(1), 4–11.
- [12] Alaghbari, W., Al-Sakkaf, A. A. & Sultan, B., (2019). Factors affecting construction labour productivity in Yemen. *International Journal of Construction Management*, 19(1), 79–91.
- [13] Shoar, S. & Banaitis, A. (2018). Application of fuzzy fault tree analysis to identify factors influencing construction labor productivity: A high-rise building case study. *Journal of Civil*

- Engineering and Management*, 25(1), 41–52.
- [14] Adebowale, O. & Smallwood, J. (2020). Contractors' perceptions of factors affecting construction labour productivity in South Africa. *Journal of Contemporary Management*, 17(2), 326–346.
- [15] Nasirzadeh, F., Rostamnezhad, M., Carmichael, D. G., Khosravi, A. & Aisbett, B. (2022). Labour productivity in Australian building construction projects: a roadmap for improvement. *International Journal of Construction Management*, 22(11), 2079–2088.
- [16] Roberts, D., Park, K., Ankrah, N. & Kim, K. P. (2023). The effect on cost when utilising modular construction techniques-The impact of weather on labour productivity. *International Journal of Sustainable Building Technology and Urban Development*, 14(1), 77–95.
- [17] Bouabaz, M. & Hamami, M. (2008). A cost estimation model for repair bridges based on artificial neural network. *American Journal of Applied Sciences*, 5(4), 334–339.
- [18] Otmani, A., Bouabaz, M. & Al-Hajj, A. (2020). Predicting Maintenance and Rehabilitation Cost for Buildings Based on Artificial Neural Network and Fuzzy Logic. *International Journal of Computational Intelligence and Applications*, 19(1), 1–13.
- [19] Peško, I. *et al.* (2017). Estimation of costs and durations of construction of urban roads using ANN and SVM. *Complexity*. 2450370.
- [20] Umit Dikmen, S. & Sonmez, M. (2011). An artificial neural networks model for the estimation of formwork labour. *Journal of Civil Engineering and Management*, 17(3), 340–347.
- [21] Al-Zwainy, F. M. S., Rasheed, H. A. & Ibraheem, H. F. (2012). Development of the construction productivity estimation model using artificial neural network for finishing works for floors with marble. *ARN Journal of Engineering and Applied Sciences*, 7(6), 714–722.
- [22] Gerek, I. H., Erdis, E., Mistikoglu, G. & Usmen, M. (2015). Modelling masonry crew productivity using two artificial neural network techniques. *Journal of Civil Engineering and Management*, 21(2), 231–238.
- [23] Aswed, G. K. (2016). Productivity Estimation Model for Bracklayer in Construction Projects Using Neural Network. *Al-Qadisiyah Journal For Engineering Sciences*, 9(2), 183–199.
- [24] Golnaraghi, S., Zangenehmadar, Z., Moselhi, O. & Alkass, S. (2019). Application of Artificial Neural Network (s) in Predicting Formwork Labour Productivity. *Advances in Civil Engineering*, 2019, 1–11.
- [25] Mlybari, E. A. (2020). Application of soft computing techniques to predict construction labour productivity in Saudi Arabia. *International Journal of GEOMATE*, 19(71), 203–210.
- [26] Bokor, O., Florez-Perez, L., Pesce, G. & Gerami Seresht, N. (2021). Using Artificial Neural Networks to Model Bricklaying Productivity. *Proceedings of the 2021 European Conference on Computing in Construction*, 2021, July (52–58).
- [27] Juszczak, M. (2020). Analysis of labour efficiency supported by the ensembles of neural networks on the example of steel reinforcement works. *Archives of Civil Engineering*, 66(1), 97–111.
- [28] Sarmiento, M., Sc, M., Luiz, M., Ferreira, R. & Sc, D. (2021). Sizing an Artificial Neural Network for Modelling Piping Assembly Labor Productivity Under Different Conditions. *International Journal of Engineering Research and Applications*, 11(2), 16–22.
- [29] Goodarizad, PP., Mohammadi Golafshani, E. & Arashpour, M. (2021). Predicting the construction labour productivity using artificial neural network and grasshopper optimisation algorithm. *International Journal of Construction Management*, 1–28.
- [30] Lawaju, N., Parajuli, N. & Shrestha, S. K. (2021). Analysis of Labor Productivity of Brick Masonry Work in Building Construction in Kathmandu Valley. *Journal of Advanced College of Engineering and Management*, 6, 159–175.
- [31] Phaobunjong, K. (2002). *Parametric Cost Estimating Model for Conceptual Cost Estimating of Building Construction Projects*. University of Texas at Austin.
- [32] Al-Zwainy, F., Abed - Alla Eiada, A. & Abed - Almajed Khaleel, T. (2016). Application

- Intelligent Predicting Technologies in Construction Productivity. *American Journal of Engineering and Technology Management*, 1(3), 39–48.
- [33] El-Kholy, A. M. (2010). Productivity performance estimation in Egyptian construction projects based on regression analysis. *Mansoura Engineering Journal (MEJ)*, 35(01), 44–54.
- [34] Narayanan, K. M. A. & Aswathi, M. P. (2018). Labour Productivity Model for Structural Elements By Varying Buildability Factors using Multiple Regression Analysis. *International Journal of Advance Research, Ideas and Innovations in Technology*, 4(2), 139–146.
- [35] Khanh, H. D., Kim, S. Y., Van Khoa, N. & Tu, N. T. (2021). The relationship between workers' experience and productivity: a case study of brick masonry construction. *International Journal of Construction Management*, 1–10.
- [36] Paikun, P., Fatimah, C., Nugroho, N. S. & Kusmana, D. (2021). Labor Productivity Rating Model for Light Brick Wall Installation in Residential Projects. *ASTONJADRO*, 10(2), 260–270.
- [37] Matejević-Nikolić, B. & Živković, L. (2022). Comparative Representation of Two Models for Predicting the Productivity of Column and Wall Concreting Process. *Buildings*, 12(11), 1809.
- [38] Al-Zwainy, F. M. S., Abdulmajeed, M. H. & Aljumaily, H. S. M. (2013). Using Multivariable Linear Regression Technique for Modeling Productivity Construction in Iraq. *Open Journal of Civil Engineering*, 03(03), 127–135.
- [39] Enshassi, A., Mohamed, S., Mustafa, Z. A. & Mayer, P. E. (2007). Factors affecting labour productivity in building projects in the Gaza strip. *Journal of Civil Engineering and Management*. 13(4), 245–254.
- [40] Fagbenle Olabosipo, I., Ogunde Ayodeji, O. & Owolabi James, D. (2011). Factors Affecting the Performance of Labour in Nigerian Construction Sites. *Mediterranean Journal of Social Sciences*, 2(2), 251–257.
- [41] El-Gohary, K. M., Aziz, R. F. & Abdel-Khalek, H. A. (2017). Engineering Approach Using ANN to Improve and Predict Construction Labor Productivity under Different Influences. *Journal of Construction Engineering and Management*, 143(8), 1–10.
- [42] Hafez, S. M. (2014). Critical Factors Affecting Construction Labor Productivity in Egypt. *American Journal of Civil Engineering*, 2(2), 35.
- [43] Kaming, P. F., Olomolaiye, P. O., Holt, G. D. & Harris, F. C. (1997). Factors influencing craftsmen's productivity in Indonesia. *International Journal of Project Management*, 15(1), 21–30.
- [44] Hogg, R. V. & Tanis, E. a. (2009). *Instructor 's Solutions Manual Probability and Statistical Inference Eighth Edition*. Upper Saddle River, NJ(USA): Prentice Hall.
- [45] Sincich, T., Stephan, D. & Levine, D. M. (2002). *Practical statistics by example using Microsoft Excel and Minitab*. Prentice Hall Professional Technical Reference.
- [46] Gopal, T. G. S. R. & Murali, K. (2015). A Critical Review on Factors Influencing Labour Productivity in Construction. *Journal of Mechanical and Civil Engineering (IOSR-JMCE)*, 12(5), 47–51.
- [47] Heravi, G. & Eslamdoost, E. (2015). Applying Artificial Neural Networks for Measuring and Predicting Construction-Labor Productivity. *American Society of Civil Engineers*, 141(10), 04015032.
- [48] Duch, W. & Jankowski, N. (1999). Survey of neural transfer functions," *Neural Computing Surveys*, 2, 163–212.
- [49] Nassar, Y. S. & Khaleel, T. A. (2019). Building Model to Predict Labour Productivity Using Multiple Linear Regression Technique for " Formwork Concrete Columns ". *Kufa Journal of Engineering*, 10(2), 12–26.