

## Modélisation de la croissance de la tomate (*Lycopersicon esculentum*) à partir des réseaux de neurones artificiels dans la région de Daloa (Côte d'Ivoire)

### [ Modeling the growth of tomato (*Lycopersicon esculentum*) from artificial neural networks in the region of Daloa (Côte d'Ivoire) ]

*Kouame N'guessan<sup>1</sup> and Assidjo Nogbou Emmanuel<sup>2</sup>*

<sup>1</sup>Université Jean Lorougnon GUEDE (UJLoG), Laboratoire de l'Amélioration de la Production Agricole, UFR Agroforesterie, BP 150 Daloa, Côte d'Ivoire

<sup>2</sup>Institut National Polytechnique Felix Houphouët-Boigny (INP-HB), Laboratoire des Procédés Industriels de Synthèse et Environnement, BP 1313 Yamoussoukro, Côte d'Ivoire

Copyright © 2021 ISSR Journals. This is an open access article distributed under the **Creative Commons Attribution License**, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

**ABSTRACT:** The tomato is an annual herbaceous plant, of the Solanaceae family. It is cultivated for its fruits which are consumed either fresh or cooked, or processed industrially. Its growth is a complex phenomenon which involves several parameters. A study of the growth parameters carried out in the region of Daloa (Côte d'Ivoire) showed a complexity of the growth of the tomato at the level of the number of leaves, the length of the leaves, the width of the leaves, the height of the trunk and the circumference of the trunk of the tomato plant. For this purpose, mathematical models were developed to predict the growth of the tomato plant from artificial neural networks for the number of leaves, the length of the leaves, the width of the leaves, the height of the plant and the circumference of the trunk of the tomato plant. The coefficients of determination between the experimental measurements and the measurements predicted by artificial neural networks are respectively 0.9722; 0.9925; 0.997; 0.9945 and 0.9926 for plant height, number of leaves, plant circumference, leaf length and width. These results are satisfactory insofar as all the coefficients of determination ( $R^2$ ) are greater than 0.97. These coefficients close to 1 show a good interpolation between the experimental values and those predicted by the model. They indicate that the values predicted by artificial neural networks are almost more than 97% close to the experimental values. Because of this, artificial neural networks are reliable enough to predict tomato growth in leaf count, leaf length, leaf width, plant height, and trunk circumference of the tomato plant.

**KEYWORDS:** Tomato, Growth, Modeling, Neural network, Validation, Daloa.

**RESUME:** La tomate est une plante herbacée annuelle, de la famille des Solanacées. Elle est cultivée pour ses fruits qui sont consommés soit frais ou cuit, soit transformés industriellement. Sa croissance est un phénomène complexe qui fait intervenir plusieurs paramètres. Une étude des paramètres de croissance réalisée dans la région de Daloa (Côte d'Ivoire) a montré une complexité de la croissance de la tomate au niveau du nombre de feuilles, de la longueur des feuilles, de la largeur des feuilles, de la hauteur du tronc et de la circonférence du tronc de la plante de tomate. A cet effet, des modèles mathématiques ont été élaborés pour prévoir la croissance de la plante de tomate à partir des réseaux de neurones artificiels pour le nombre de feuilles, la longueur des feuilles, la largeur des feuilles, la hauteur de la plante et la circonférence du tronc de la plante de tomate. Les coefficients de détermination entre les mesures expérimentales et les mesures prédites par les réseaux de neurones artificiels sont respectivement de 0,9722; 0,9925; 0,997; 0,9945 et 0,9926 pour la hauteur de la plante, le nombre de feuille, la circonférence de la plante, la longueur et la largeur des feuilles. Ces résultats sont satisfaisants dans la mesure où tous les coefficients de détermination ( $R^2$ ) sont supérieurs à 0,97. Ces coefficients proche de 1 montrent une bonne interpolation entre les valeurs expérimentales et celles prédites par le modèle. Ils indiquent que les valeurs prédites par les

réseaux de neurones artificiels sont pratiquement à plus de 97 % proches des valeurs expérimentales. De ce fait, les réseaux de neurones artificiels sont suffisamment fiables pour prédire la croissance de la tomate au niveau du nombre de feuilles, de la longueur des feuilles, de la largeur des feuilles, de la hauteur de la plante et de la circonférence du tronc de la plante de tomate.

**MOTS-CLEFS:** Tomate, Croissance, Modélisation, Réseau de neurones, Validation, Daloa.

## 1 INTRODUCTION

Dans le secteur vivrier, les légumes jouent un rôle prépondérant dans l'alimentation quotidienne des populations. Les importations en légumes et fruits occasionnent une sortie de devise estimée à environ 10 milliards [1]. Selon l'Agence Nationale d'Appui au Développement Rural (ANADER), les quantités en produits légumiers s'élevaient en 2009 à 760 000 tonnes [2]. Parmi les légumes cultivés, c'est une culture de rente importante pour les petits exploitants et pour les agriculteurs commerciaux qui ont une exploitation moyenne. La tomate (*Lycopersicon esculentum*) est devenue un des légumes les plus importants du monde. En 2001, la production mondiale de tomate était d'environ 105 millions de tonnes de fruits frais sur une superficie évaluée à 3,9 millions d'hectares. Elle se positionne au premier rang mondial des fruits cultivés avec une production d'environ 152 millions de tonnes en 2010 [3]. En Côte d'Ivoire, les besoins en tomate étaient estimés à plus de 200 000 tonnes en 2002 [4] alors que la production nationale qui est de 52 000 tonnes/an reste largement inférieure pour couvrir cette demande [5]. Depuis quelques années, tout le territoire de la Côte d'Ivoire, subit les effets du changement climatique dont l'une des manifestations visibles est la variabilité spatio-temporelle de la pluie. A cela s'ajoute la baisse du niveau de fertilité des sols, affectant ainsi les rendements des cultures. Aujourd'hui, le constat amer du déclin de la production agricole est en grande partie justifié par ces problèmes climatiques et édaphiques. Il est donc important de trouver de nouvelles méthodes permettant la maîtrise et la prévision de ses phénomènes. Une des grandes approches de la résolution de ses problèmes est la modélisation. Ainsi, l'utilisation des outils d'aide de décisions tels que les modèles de cultures se présente comme une nécessité. Pour cela, divers modèles de culture de tomate tels que APSIM [6], Crop Syst [7], APES [8] ou modèle générique STICS (capable de simuler le fonctionnement de plusieurs cultures à partir des données sur le sol, le climat, les caractéristiques des cultures et les pratiques culturales), (Brisson *et al.* [9]; Brisson *et al.* [10]) ont été développés par la recherche afin d'étudier la productivité ainsi que l'effet des facteurs climatiques et édaphiques sur la production et décrire un grand nombre de processus écophysiologicals et agronomiques, représentatifs de la plupart des espèces annuelles cultivées [11]. Malgré ces connaissances récentes, les modèles concernant la croissance de la plante de tomate demeurent rares. C'est pourquoi la présente étude se propose de prendre part à la résolution de ce problème en modélisant plus particulièrement la croissance de la plante de tomate pour une meilleure production qui passe nécessairement par une bonne croissance.

L'objectif principal de ce travail, consiste à utiliser les réseaux de neurones artificiels pour prédire la croissance des plantes de tomate en temps réel et anticiper sur celle-ci afin de maîtriser la productivité de cette plante.

## 2 MATERIEL ET METHODES

### 2.1 ZONE D'ÉTUDE

La phase d'expérimentation s'est déroulée dans la région du Haut Sassandra. Cette région est située au Centre-Ouest de la Côte d'Ivoire avec pour chef-lieu, la ville de Daloa (Figure 1). Le département de Daloa est situé entre 6°53'58" de latitude Nord et 6°26'32"W de longitude Ouest avec une superficie de 15 205 km<sup>2</sup> pour une population estimée à 1.430.960 habitants [12]. Le site d'expérimentation se trouve à proximité de la bibliothèque de l'Université Jean Lorougnon Guédé à environ 300 m de celle-ci. Cette zone est sujette à quatre saisons réparties comme suit: une grande saison des pluies allant d'Avril à mi-Juillet, une petite saison sèche de mi-Juillet à mi-Septembre, une petite saison des pluies de mi-Septembre à Novembre et la grande saison sèche de Décembre à Mars. Il s'agit d'une zone tropicale humide (Tableau 1) avec une végétation de forêt dense à évolution régressive due à la pratique d'une agriculture extensive et itinérante couplée à l'exploitation incontrôlée des essences forestières [13]. Le patrimoine édaphique est de type ferrallitique. Les précipitations ont connu une baisse passant de 1868,5 mm en 1968 à 1120,4 mm de pluie en moyenne en 2005. La région connaît une baisse de la pluviométrie de l'ordre de 40 % [14]. Le département est arrosé par le fleuve Sassandra et par son affluent la "Lobo" dont les ramifications, le Dé et le Gore inondent l'ensemble des localités [13].

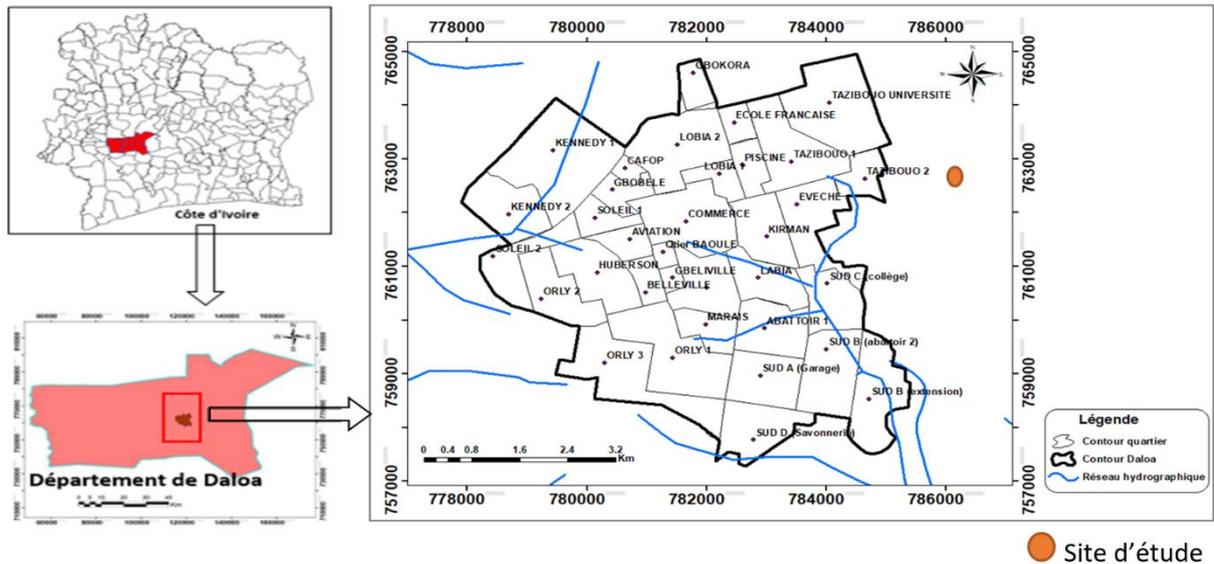


Fig. 1. Présentation de la zone d'étude

Tableau 1. Evolution des données climatiques de la région de Daloa pendant la période d'expérimentation

Mois	Température moyenne (°C)	Température maximale (°C)	Température minimale (°C)	Humidité relative (%)	Précipitations (mm)
Mai 2020	27,9	33,1	23,5	79,7	113,03
Juin 2020	26,4	30,6	22,9	84,6	107,18
Juillet 2020	25,2	28,6	22,3	86,6	105,14
Aout 2020	25	28,8	21,8	83,8	109,74

Source: [www.tutiempo](http://www.tutiempo)

## 2.2 MATÉRIEL VÉGÉTAL

Le matériel végétal est constitué de plante de tomate issue de graine de la variété UC 82.

## 2.3 MÉTHODES EXPÉRIMENTALES

### 2.3.1 DISPOSITIF EXPÉRIMENTAL

Le dispositif expérimental était un bloc de Fisher avec trois (03) répétitions. Chaque bloc était constitué de deux billons soit au total six (6) billons pour la parcelle. Chaque billon était séparé d'une allée de 1 m de large et a une dimension de 10 m x 40 cm. Les plants ont été repiqués sur les billons, espacés de 0,40 m sur la ligne et pour un total de 20 plants par billon. Le total des plants est de 120 pour une superficie de 15 m<sup>2</sup> (15 m x 15 m), soit une densité d'environ 25000 plants/hectare.

### 2.3.2 MISE EN PLACE D'UNE PLANCHE DE PÉPINIÈRE

Une planche de pépinière de 1 m de large et 3 m de long a été réalisée. Du furadan a été utilisé sur la planche pour désinfecter le sol. Le Furadan a pour matière active le Carbofuran (famille des carbamates). Il se présente sous forme de granulés et a été appliqué à la volée à la dose de 28 Kg/ha.

### 2.3.3 SEMIS EN PÉPINIÈRE

Le semis a été effectué le 30 Mai 2020 à raison de 2 graines par poquet et distants de 5 cm sur la ligne soit 20 poquets par ligne de 1 m. L'écartement entre deux lignes de semis est de 10 cm. La pépinière a été constituée de 3 lignes de semis, soit 60 poquets pour la variété. La planche a été recouverte de paille avant l'arrosage, afin d'éviter de tasser les graines.

#### **2.3.4 SUIVI ET ENTRETIEN DE LA PÉPINIÈRE**

Après la germination des graines entre le 6 et le 14 juin 2020, la paille a été enlevée. Une ombrière a été construite à 80 cm au-dessus du sol pour protéger les plants contre l'ensoleillement. Un film plastique a été posé sur l'ombrière pour éviter que les plants ne reçoivent l'eau de pluie. Des arrosages réguliers (2 fois par jour, matin et soir) ont été réalisés au début, puis réduit à un arrosage par jour à environ 10 jours après. L'ombrière a été retirée au fur et à mesure que les plants se développaient par la réduction régulière du nombre de paille de couverture. Elle a été totalement enlevée à partir du 26 juin 2020 avant le repiquage pour permettre aux plants de supporter la température ambiante avant d'être transplantés au champ.

#### **2.3.5 PRÉPARATION DU TERRAIN**

La parcelle choisie était couverte d'une jachère non améliorée. Cette parcelle a été défrichée, labourée et une fumure de fond de NPK (10-18-18), à raison de 250 kg/ha a été préalablement apportée lors de la mise en place des billons et ensuite la transplantation des plants vigoureux a été effectuée, un (01) mois après la date du semis en pépinière. Les tuteurs ont été mis en place lors du repiquage. Les observations au champ ont été effectuées et se sont achevées à la floraison.

#### **2.3.6 PLANTATION, SUIVI ET ENTRETIEN DE LA CROISSANCE DES PLANTS DE TOMATE**

Après le repiquage, un arrosage régulier (matin et soir) a été effectué selon le besoin. Le remplacement des pieds morts s'est déroulé à partir de 10 jours après le repiquage et s'est étendu sur deux semaines. Les remplacements ont été effectués après flétrissement des pieds ou après la coupe des pieds par un ravageur. Des sarclages manuels ont été effectués afin d'éviter la compétition des plantes de tomates avec les espèces sauvages. Deux traitements phytosanitaires (insecticide et fongicide) ont été effectués avec du Cigogne (40 mL dans 15 litre d'eau pour 400 m<sup>2</sup>) et du manèbe (75 g dans 15 litre d'eau). Des traitements préventifs à base de Décis (Deltaméthrine) et d'Ivory 80WP (mancozèbe) ont été effectués à titre préventif pour éviter et lutter contre la présence des ravageurs.

#### **2.3.7 MESURE ET APPRÉCIATION DE LA CROISSANCE ET DU DÉVELOPPEMENT**

Les caractéristiques de croissance et de développement mesurées sont les mêmes que celles définies par Shankara *et al.* [15].

##### **2.3.7.1 PARAMÈTRES DE CROISSANCE DE LA TOMATE**

Les mesures de croissance sont réalisées en centimètre et notées sur une fiche. Les observations ont porté sur l'appareil végétatif, notamment la hauteur et la circonférence des plantes de tomate, le nombre de feuilles, la longueur et la largeur des feuilles. Elles ont démarré le 07 juillet 2020 et se sont achevées à la floraison. Les mesures des paramètres de croissance et de développement ont été faite de façon hebdomadaire. Les observations ont porté sur un effectif de 60 plants, pris de façon aléatoire sur la parcelle choisit parmi les 120 plantes de tomate de la parcelle.

##### **2.3.7.2 DIMENSIONS FOLIAIRES**

La croissance foliaire a été appréciée de façon hebdomadaire en mesurant la longueur (Lg) de la dernière feuille fonctionnelle, de la base du limbe jusqu'à l'extrémité de feuille. La largeur (lg) de la feuille a été évaluée en considérant la plus grande dimension de la partie médiane de cette même feuille. Les valeurs ont été exprimées en centimètres.

##### **2.3.7.3 DIMENSIONS DU TRONC**

La hauteur et la circonférence ont été les paramètres mesurés pour suivre la croissance du pseudotrunc. La hauteur du tronc (Hp) a été mesurée à partir du collet jusqu'au sommet de la plante. La circonférence du tronc, par contre, a été évaluée à la base de la plante au niveau du collet.

##### **2.3.7.4 PARAMÈTRES DE DÉVELOPPEMENT**

Le développement de la plante de tomate a été suivi par le comptage du nombre de feuilles fonctionnelles sur les 20 plantes prises de façon aléatoire dans chaque bloc.

### 2.3.8 DÉMARCHE DE MODÉLISATION DE LA CROISSANCE PAR LE MODÈLE EMPIRIQUE

Dans cette étude, la modélisation de la croissance de la tomate a été assurée par un modèle empirique utilisant comme outil les réseaux de neurones artificiels (RNA). La démarche de modélisation utilisée est celle de Kouame [16].

#### 2.3.8.1 MODÉLISATION DE LA CROISSANCE DE LA TOMATE PAR LE RÉSEAU DE NEURONES ARTIFICIELS

La modélisation de la croissance de la plante de tomate a été réalisée grâce à un réseau de neurones artificiels à travers un programme élaboré dans le logiciel Matlab R2010b (Math Works Inc., Massachusetts USA). Ce programme a permis de mettre en place un modèle empirique qui s'est élaboré en plusieurs étapes.

#### 2.3.8.2 DÉTERMINATION DES PARAMÈTRES D'ENTRÉES ET DES PARAMÈTRES DE SORTIES DU RÉSEAU DE NEURONES ARTIFICIELS.

Les paramètres de croissance (paramètres explicatifs) suivis ont été retenus pour expliquer et prédire la croissance de la tomate. Ils ont représenté respectivement les paramètres d'entrées à l'instant  $t$  et les paramètres de sorties à l'instant  $t+1$  du réseau de neurones artificiels. Pour faciliter l'expression mathématique du modèle, des symboles ont été attribués aux différents paramètres. Il comprend cinq neurones d'entrée correspondant aux cinq paramètres suivis expérimentalement et cinq neurones de sortie correspondant aux cinq variables de sortie à l'instant  $t+1$  (Figure 2).

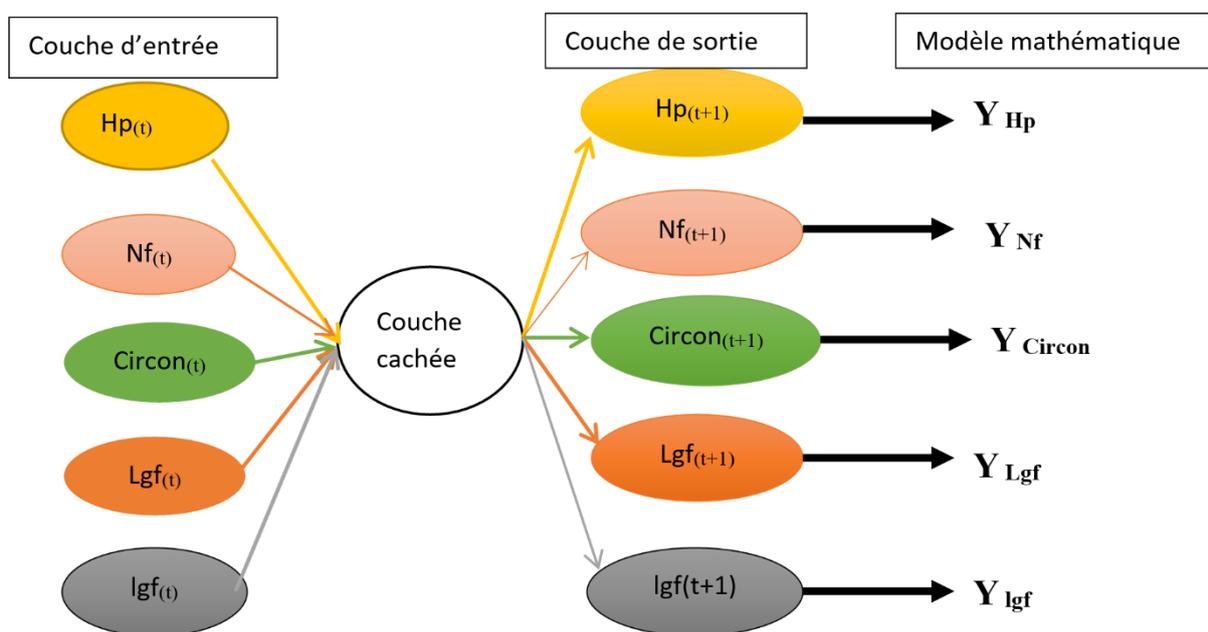


Fig. 2. Réseau de neurones bouclés avec 5 couches d'entrées, une couche cachée et 5 couches de sortie

#### 2.3.8.3 CONSTRUCTION ET PRÉPARATION DE LA BASE DE DONNÉES

La base de données a été constituée des mesures prises au champ lors de la croissance de la plante de tomate. Cette base de données implémentées dans le module réseau de neurones artificiels du logiciel Matlab R2010b (Math Works Inc., Massachusetts USA), a subi une étape de normalisation dans la gamme  $[-1; +1]$ . Cette normalisation permet de réduire les erreurs de dispersions en accélérant la convergence du réseau pendant l'apprentissage, et d'augmenter son efficacité de généralisation. Cette base de données normalisée a été par la suite scindée en trois sous-ensembles, de façon aléatoire, pour la détermination des poids de connexion du réseau. Ce partage permet d'éviter les deux pièges que peut rencontrer le réseau de neurones artificiels, à savoir, l'existence de minima locaux et le sur-apprentissage. Un premier groupe de données (50 %) a servi de base d'apprentissage. Dans le second groupe, un quart de données a été utilisé pour la validation. Le deuxième sous-groupe de données (25 % restant) a permis de tester la validité du modèle qui a été retenu.

#### **2.3.8.4 ELABORATION DE LA STRUCTURE DU RÉSEAU DE NEURONES ARTIFICIELS**

Dans la mise au point du réseau de neurones artificiels, la partie la plus difficile est la détermination de sa structure qui se fait par le choix de l'architecture, le nombre de couches cachées, le nombre de neurones sur les couches cachées et les fonctions de transfert existant entre les neurones.

#### **2.3.8.5 NOMBRE DE COUCHES CACHÉES ET DE NEURONES CACHÉS**

Le nombre de neurones de la couche cachée augmente de manière exponentielle en fonction du nombre d'entrées. Fonctions continues: toutes les fonctions continues bornées sont représentables, avec une précision arbitraire, par un réseau à deux couches [17]. Ce théorème s'applique au réseau dont les neurones utilisent la sigmoïde dans la couche cachée et des neurones linéaires (sans seuil) dans la couche de sortie. Le nombre de neurones dans la couche cachée dépend de la fonction à approximer. Le nombre de couches cachées peut varier entre un et deux. Cependant, notons qu'un perceptron multicouche avec une seule couche cachée pourvue d'un nombre suffisant de neurones, peut approximer n'importe quelle fonction avec la précision souhaitée [18]. Le nombre de neurones cachés  $k$  a été varié arbitrairement de 1 à 15 dans la couche cachée jusqu'à aboutir à une architecture optimale.

#### **2.3.8.6 FONCTION DE TRANSFERT**

La fonction d'activation (ou fonction de transfert) sert à convertir le résultat de la somme pondérée des entrées d'un neurone en une valeur de sortie, cette conversion s'effectue par un calcul de l'état du neurone en introduisant une non-linéarité dans le fonctionnement du neurone [17]. Le biais «  $b$  » joue un rôle de seuil, quand le résultat de la somme pondérée dépasse ce seuil. Il y a plusieurs types de fonctions de transfert qui peuvent être utilisés dans les RNA:

- La fonction sigmoïde;
- La fonction tangente hyperbolique;
- La fonction de Heaviside.

Dans cette étude, une fonction sigmoïde de transfert de type tangentiel a été adoptée car celle-ci est une fonction non linéaire permettant d'exprimer les phénomènes non linéaires comme la croissance de la tomate.

#### **2.3.8.7 APPRENTISSAGE ET VALIDATION DU MODÈLE NEURONAL**

L'apprentissage est une étape très importante du développement d'un réseau de neurones durant laquelle le comportement du réseau est modifié itérativement jusqu'à l'obtention du comportement désiré, et ce par l'ajustement des poids (connexion ou synapse) des neurones à une source d'information bien définie [19]. Pour un RNA, l'apprentissage peut être regardé également comme étant le processus de la mise à jour des poids (connexion) au sein du réseau dans le but d'ajuster la réponse du réseau à l'expérience et aux exemples. Dans cette étude, pour arriver à de bons résultats, les calculs ont été répétés 1500 fois pour chaque  $k \in [1; 15]$ , nombre de neurones cachés. La phase de validation s'est déroulée simultanément avec la phase d'apprentissage. Ainsi, chaque fois que les poids ont été déterminés dans la phase d'apprentissage, l'erreur absolue  $\Delta y$  (valeur calculée  $y_{\text{préd}}$  - valeur expérimentale  $y_{\text{exp}}$ ) a été mesurée dans l'ensemble de la validation. Toutefois, la validation du modèle neuronal a été réalisée par le Coefficient de corrélation ( $R$ ) entre les valeurs prédites et les valeurs expérimentales, par l'Erreur Quadratique Moyenne (EQM) pour différents comportements architecturaux et par le coefficient de détermination ( $R^2$ ).

#### **2.3.9 TRAITEMENTS STATISTIQUES DES DONNÉES**

##### **2.3.9.1 COEFFICIENT DE CORRÉLATION ( $R$ )**

Le coefficient de corrélation ( $R$ ) est utilisé habituellement pour évaluer la performance des modèles mathématiques. Il est obtenu en faisant la régression linéaire entre les valeurs expérimentales et les valeurs prédites par le réseau. Lorsque la valeur absolue du coefficient de corrélation ( $R$ ) entre les valeurs expérimentales et celles prédites par le réseau tend vers zéro (0), le degré de liaison est faible. Par contre, lorsqu'elle est proche de 1, la liaison réciproque est forte [20]. Sa formulation est élaborée comme suit:

$$R = \frac{\sum_{i=1}^N (y_e - \bar{y}_e)(y_c - \bar{y}_c)}{\sqrt{\sum_{i=1}^N (y_e - \bar{y}_e)^2} \sqrt{\sum_{i=1}^N (y_c - \bar{y}_c)^2}} \quad (1)$$

$y_e$  étant les valeurs expérimentales;

$y_c$  les valeurs calculées par le réseau pour les  $i=1, \dots, N$ ;

$\bar{y}_e$  et  $\bar{y}_c$  qui sont les moyennes respectives des valeurs expérimentales et les valeurs calculées par le réseau;

$N$  qui correspond au nombre de variables.

### 2.3.9.2 ERREUR QUADRATIQUE MOYENNE (MEAN SQUARE ERROR) (EQM)

L'Erreur Quadratique Moyenne (EQM) est généralement utilisée pour mesurer la performance de l'étape d'apprentissage pendant le développement du modèle neuronal. L'expression de l'erreur quadratique moyenne (EQM) est présentée de la manière suivante [21] et [22].

$$EQM = \frac{1}{N} \sum_{i=1}^N (y_e - y_c)^2 \quad (2)$$

$y_e$  et  $y_c$  qui sont respectivement les valeurs expérimentales et celles calculées par le réseau pour les  $i = 1, \dots, N$

### 2.3.9.3 COEFFICIENT DE DÉTERMINATION ( $R^2$ )

Les valeurs prédites par le modèle et celles expérimentales ont été simulées avec le logiciel Matlab R2010b (Math Works Inc., Massachusetts USA). La performance de la qualité de la prévision a été déterminée par le coefficient de détermination ( $R^2$ ). Ce coefficient permet de mesurer la qualité de l'ajustement des estimations de l'équation de régression. Il est utilisé pour avoir une idée globale de l'ajustement des modèles développés dans cette étude. Lorsque  $R^2$  est proche de 1 dans le cas d'une régression simple, l'ajustement entre les valeurs expérimentales et prédites est meilleur [23]. L'expression du coefficient de détermination ( $R^2$ ) est donnée par l'équation (3) suivante:

$$R^2 = \frac{\sum_{i=1}^N (y_{préd} - \bar{y}_e)^2}{\sum_{i=1}^N (y_e - \bar{y}_e)^2} \quad (3)$$

$y_e$  et  $y_{préd}$  étant respectivement les valeurs expérimentales et calculées pour  $i=1, \dots, N$

$\bar{y}_e$  étant la moyenne des valeurs mesurées ou expérimentale

## 3 RESULTATS

Dans cette étude, la croissance de la tomate a été modélisée par le modèle empirique, en l'occurrence le réseau de neurones artificiels. Les étapes de la modélisation ont consisté à déterminer l'architecture des réseaux de neurones puis à valider le modèle neuronal optimisé. Les modèles établis représentent chaque paramètre de croissance étudié soit au total cinq modèles mathématiques pour cette étude. Chaque modèle peut permettre de prévoir la croissance soit au niveau foliaire (en nombre, en longueur et en largeur), soit au niveau du pseudotrunc (en épaisseur et en hauteur). La formule mathématique trouvée est spécifique pour ces paramètres.

### 3.1 ARCHITECTURE DU RÉSEAU DE NEURONE

La modélisation de la croissance de la tomate a débuté par la détermination de l'architecture neuronale qui permet de mieux ajuster les mesures expérimentales. Le nombre de neurones «  $k$  » dans la couche cachée qui optimise le réseau de

neurones artificiels a été déterminé. Ceci revient à déterminer les neurones cachés permettant un bon coefficient de corrélation (R) entre les valeurs prédites et les valeurs expérimentales. Les Erreurs Quadratiques Moyennes (EQM) ont été utilisés pour déterminer le réseau optimal.

Le tableau 2 présente les coefficients de corrélation (R) durant les phases respectives d'apprentissage et de validation. L'analyse des résultats de ce tableau montre que les coefficients de corrélation (R) de la hauteur du pseudotrunc par exemple varient de 0,8658 à 0,9325 pour la phase d'apprentissage et de 0,7271 à 0,8784 pour la phase de validation. En considérant toutes les réponses, les coefficients de corrélation (R) évoluent dans le même ordre, c'est-à-dire de 0,4000 à 0,9325 pour toutes les phases (apprentissage et validation). Notons que les meilleures valeurs des coefficients de corrélation (R) sont obtenues au niveau de la hauteur du pseudotrunc (0,8784 à 0,9325) et les mauvaises sont observées au niveau de la circonférence du sommet et du nombre de feuilles (0,4000 à 0,4852). En observant les valeurs des coefficients de corrélation (R) pour chaque nombre de neurones cachés durant les phases d'apprentissage et de validation, il ressort que celles-ci sont nettement élevées pour certains neurones tels que 1; 2; 3; 9 et 13. La détermination du nombre de neurones le plus performant entre ces cinq neurones pré-identifiés grâce au coefficient de corrélation s'est faite par la comparaison des Erreurs Quadratiques Moyennes (EQM).

Les neurones 2 et 9 sont les neurones qui ont EQM les plus faible. De ce qui précède, une question se pose. Elle consiste à savoir, parmi les structures neuronales 2 et 9 citées, quelle est celle qui présente la meilleure performance. La détermination de cette performance entre les nombres de neurones 2 et 9 tient compte aussi du neurone le moins complexe.

En somme, en combinant tous les critères de performance du réseau c'est-à-dire le Coefficient de corrélation (R) le plus proche 1, l'Erreur Quadratique Moyenne plus faible et l'architecture neuronale la moins complexe, le meilleur compromis est le nombre de neurone 2. Ceci permet d'avoir une topologie 5-2-5. Cette topologie comporte cinq neurones dans la couche d'entrée, deux neurones dans la couche cachée et cinq neurones dans la couche de sortie. Elle a permis d'élaborer les modèles mathématiques suivants permettant de prédire la croissance de la tomate.

$$Y_{\text{Haut}} = -0,72 Y'_1 - 0,34 Y'_2 - 0,05$$

$$Y_{\text{Nf}} = -0,30 Y'_1 - 0,60 Y'_2 - 0,11$$

$$Y_{\text{Circon}} = -0,48 Y'_1 - 0,16 Y'_2 - 0,36$$

$$Y_{\text{Lgf}} = -0,63 Y'_1 - 0,25 Y'_2 - 0,02$$

$$Y_{\text{lgf}} = -0,60 Y'_1 + 0,07 Y'_2 - 0,44$$

Avec

$$Y'_1 = \text{TanH} [-1,43 \text{ Haut}_{(t+1)} - 0,54 \text{ Nf}_{(t+1)} + 0,32 \text{ Circon}_{(t+1)} - 0,34 \text{ Lgf}_{(t+1)} - 0,15 \text{ lgf}_{(t+1)} - 0,85]$$

$$Y'_2 = \text{TanH} [0,05 \text{ Haut}_{(t+1)} - 1,01 \text{ Nf}_{(t+1)} - 1,54 \text{ Circon}_{(t+1)} - 0,64 \text{ Lgf}_{(t+1)} + 0,91 \text{ lgf}_{(t+1)} + 2,52]$$

Tableau 2. Coefficients de corrélation (R) durant les phases d'apprentissage et de validation

Neurones de la couche cachée	Coefficients de corrélation (phase d'apprentissage)				
	Hp	Nf	Circon	Lgf	lgf
1	0,8658	0,6679	0,6955	0,8138	0,8195
<b>2</b>	<b>0,8676</b>	<b>0,8245</b>	<b>0,6953</b>	<b>0,8152</b>	<b>0,8230</b>
3	0,8831	0,7741	0,7110	0,8260	0,8288
4	0,8948	0,7580	0,7282	0,8377	0,8511
5	0,8933	0,8171	0,7558	0,8322	0,8456
6	0,8955	0,8002	0,7748	0,8415	0,8488
7	0,8899	0,7196	0,7178	0,8590	0,8805
8	0,9052	0,8804	0,7604	0,8501	0,8726
9	0,9011	0,8679	0,8038	0,8392	0,8551
10	0,9099	0,8321	0,7519	0,8662	0,8924
11	0,8991	0,8613	0,8034	0,8403	0,8624
12	0,9199	0,8262	0,7719	0,8521	0,8784
13	0,9116	0,8476	0,7924	0,8719	0,8874
14	0,9325	0,8469	0,7871	0,8804	0,9007
15	0,9317	0,8634	0,7853	0,8907	0,9071
	Coefficients de corrélation (phase de validation)				
1	0,8771	0,6144	0,6256	0,8294	0,8124
<b>2</b>	<b>0,8784</b>	<b>0,7146</b>	<b>0,6275</b>	<b>0,8288</b>	<b>0,8129</b>
3	0,8737	0,6741	0,6380	0,8237	0,8046
4	0,8313	0,5762	0,5749	0,8102	0,7742
5	0,8633	0,5828	0,6095	0,8461	0,8282
6	0,8436	0,6545	0,5624	0,8137	0,7724
7	0,8544	0,6408	0,5745	0,7034	0,7320
8	0,7976	0,4000	0,5401	0,7832	0,7779
9	0,8490	0,5108	0,5748	0,8212	0,8114
10	0,8087	0,7063	0,5929	0,7257	0,7244
11	0,8711	0,6844	0,6373	0,8171	0,7935
12	0,7749	0,7082	0,6019	0,7894	0,7567
13	0,8379	0,7303	0,6288	0,7582	0,7614
14	0,7303	0,4893	0,4852	0,6982	0,6890
15	0,7271	0,5459	0,5258	0,7078	0,7241

\*Les valeurs en gras correspondent aux coefficients de corrélation (R) choisis.

### 3.2 VALIDATION DU MODÈLE NEURONAL OPTIMISÉ

Après la détermination des réseaux de neurones artificiels optimisés, une autre phase est de confirmer la validité et la fiabilité des modèles. La validation des modèles neuronaux de topologie 5-2-5 a été effectuée par les coefficients de détermination ( $R^2$ ) entre les valeurs prédites par le réseau et celles expérimentales des mesures restantes de la base de données (Figure 3; 4; 5; 6 et 7). Les résultats de la validation du modèle neuronal 5-2-5 montrent que les coefficients de détermination ( $R^2$ ) varient entre 0,9659 et 0,9945. Ces valeurs qui sont proches de 1, montrant que les modèles mathématiques sont proches de la réalité entre 96,59 % et 99,45 %.

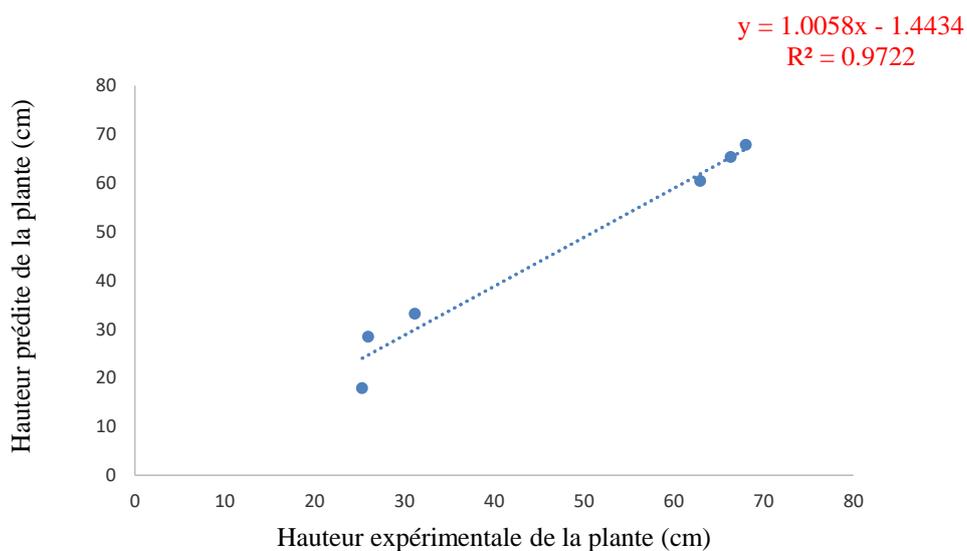


Fig. 3. Courbe de régression entre les valeurs expérimentales et les valeurs prédites par le réseau 5-2-5 pour la hauteur de la tomate

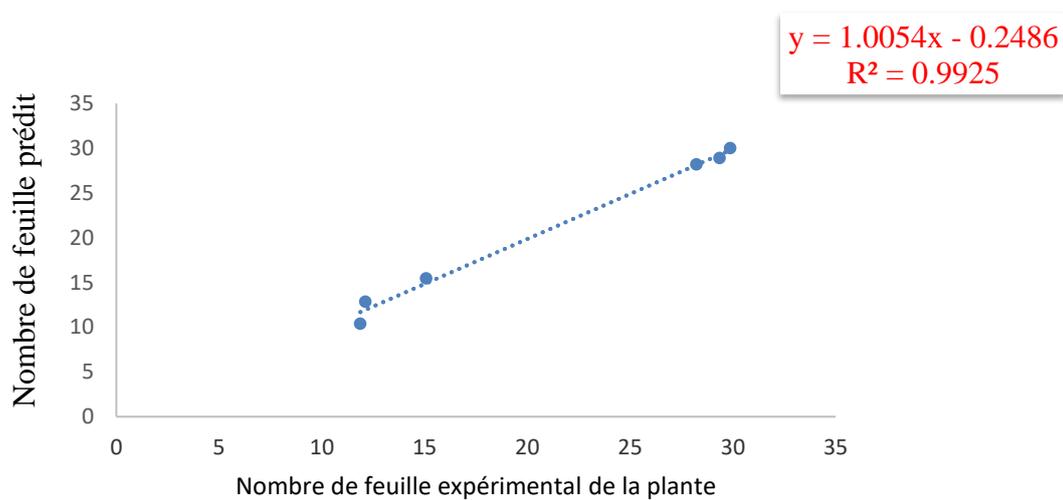


Fig. 4. Courbe de régression entre les valeurs prédites par le réseau 5-2-5 et expérimentales pour le nombre de feuilles de la tomate

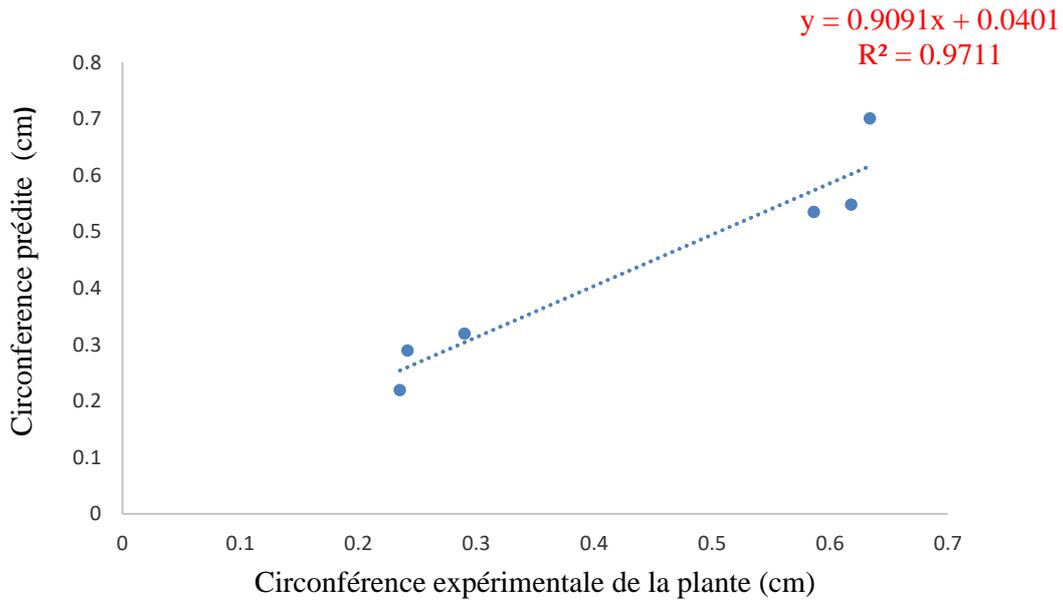


Fig. 5. Courbe de régression entre les valeurs prédites par le réseau 5-2-5 et expérimentales de la circonférence de la tomate

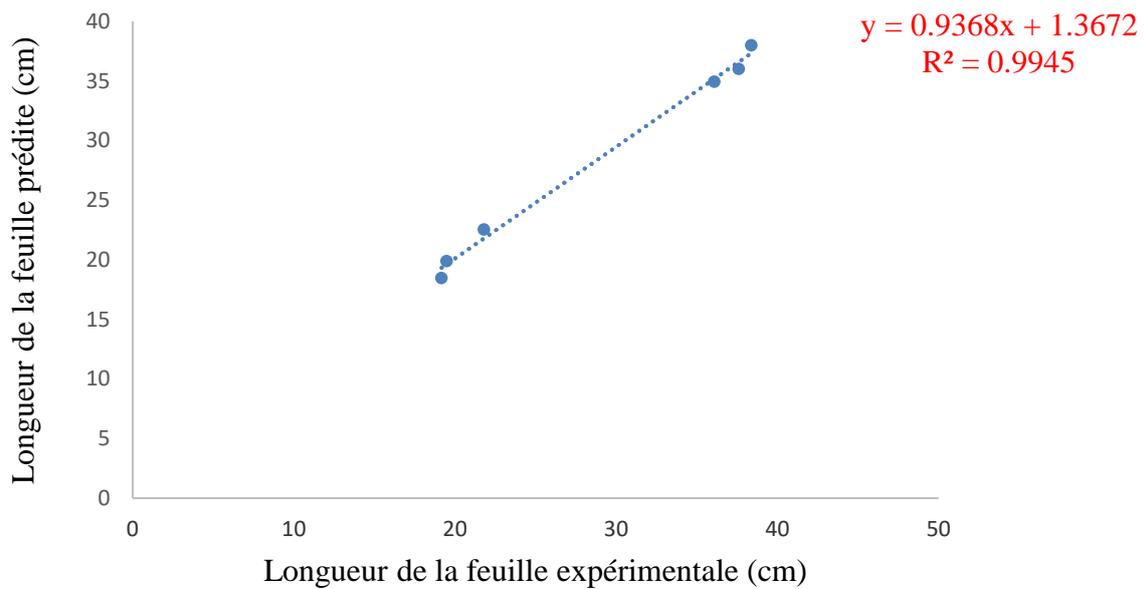


Fig. 6. Courbe de régression entre les valeurs prédites par le réseau 5-2-5 et expérimentales de la longueur des feuilles de la tomate

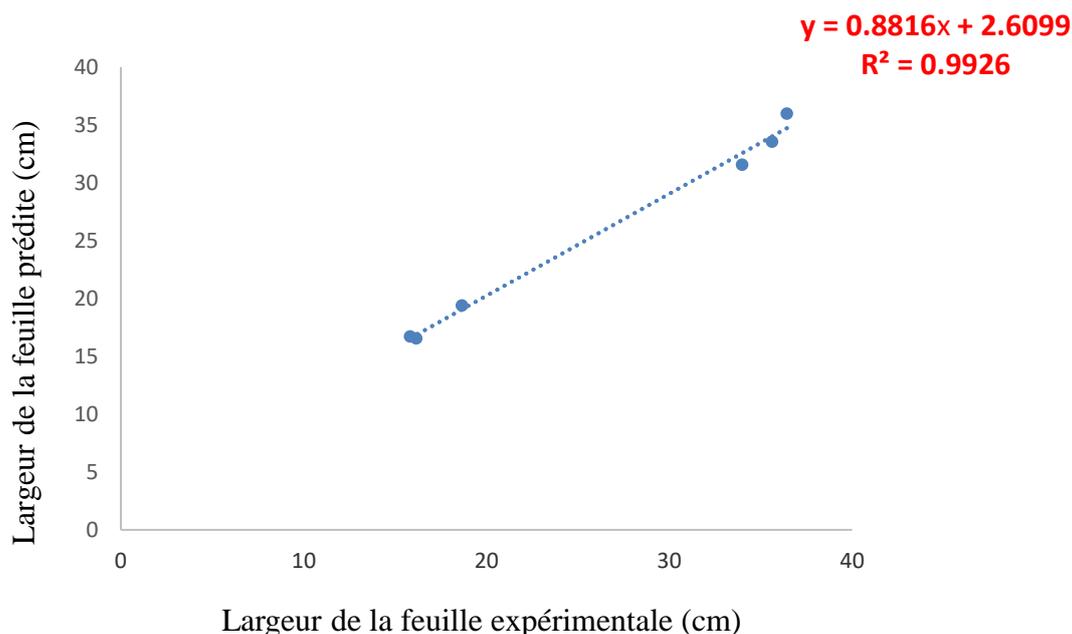


Fig. 7. Courbe de régression entre les valeurs prédites par le réseau 5-2-5 et expérimentales de la largeur des feuilles de la tomate

#### 4 DISCUSSION

Les résultats obtenus sont très variables et dépendent directement des paramètres du réseau. Les simulations réalisées afin de déterminer le nombre de neurones nécessaire pour une prédiction satisfaisante des paramètres de la croissance de la tomate ont montré que l'erreur quadratique moyenne peut être optimisée en choisissant la bonne configuration du réseau. Les simulations précédentes montrent que l'erreur diminue de manière significative à chaque fois qu'on augmente le nombre de neurones cachés. Ensuite, cette erreur est de plus en plus petite, ce qui permet au réseau d'approcher les résultats recherchés. Cependant, une fois le nombre de neurones cachés est fixé à 15 neurones, l'erreur se stabilise même si on augmente d'avantage le nombre de neurones, cela se répercute sur le réseau par une prédiction assez stable à partir de 15 neurones cachés. A cet effet, le meilleur réseau de neurones artificiels est un compromis entre les coefficients de corrélation (R) obtenus pendant la phase d'apprentissage et ceux obtenus pendant la phase de la validation [24] et [25]. En considérant toutes les réponses, y compris celle de l'EQM, l'architecture neuronale 5-2-5 (c'est-à-dire cinq neurones d'entrée, deux neurones dans la couche cachée et cinq neurones de sortie) a été retenue comme meilleure topologie. Cette stabilité de l'erreur montre que les poids sont assez bien entraînés et que cela ne sert à rien de les entraîner davantage. La modélisation réalisée afin de déterminer le nombre de couche cachée nécessaire pour une prédiction satisfaisante des paramètres de la tomate ont montré que l'EQM diminue en général de manière significative au début des premières itérations, puis cette erreur commence à fluctuer jusqu'à ce qu'elle devienne assez stable et optimale [26]. Le réseau a commencé par une petite architecture, puis il s'est développé jusqu'à atteindre une taille qui donne un bon modèle de prévision. Cette méthodologie a permis d'obtenir une variable de neurones  $k$  dans la couche cachée qui est entre 1 et 15 [27]. Ce qui a permis de faire le choix du réseau de neurones artificiels le plus performant permettant de mieux approximer la croissance de la tomate à travers les cinq paramètres de croissance. Cela montre clairement que la résolution de notre problème nécessite plus d'une couche cachée c'est-à-dire au moins 2. L'étude de l'erreur quadratique moyenne en fonction du nombre d'itérations révèle que ce dernier joue un rôle très important dans la stabilité du réseau, cependant le nombre d'itérations peut être fixé lorsque l'erreur est petite et les sorties sont satisfaisantes. Une fois qu'on a une bonne prédiction, cela ne sert à rien d'utiliser un nombre d'itérations plus grand. Car selon Thompson et al [20] et Feinberg [23], lorsque le coefficient de détermination ( $R^2$ ) est proche de 1, l'ajustement entre les valeurs expérimentales et celles prédites par un modèle mathématique est meilleur. La validation de cette modélisation s'est faite par une comparaison entre les courbes issues du modèle par réseaux de neurones et celles obtenues par les valeurs expérimentales [28]. Les coefficients de détermination ( $R^2$ ) enregistrés par le réseau de neurones artificiels de topologie 5-2-5, sont supérieurs à 0,90 et sont très proches de 1. Cela indique une bonne interpolation entre les valeurs expérimentales et celles prédites. Les résultats obtenus de la prévision et de la simulation de la croissance de la tomate

en temps réel, sont similaires à ceux de Kouame et al [29; 30] dans le cas de la modélisation de la croissance du bananier plantain par le réseau de neurones artificiels de topologie 7-2-7. Ces auteurs ont montré que les valeurs prédites par le réseau de neurones artificiels de topologie 7-2-7 sont très proches des données expérimentales dans la mesure où les valeurs des coefficients de détermination ( $R^2$ ) sont nettement supérieures à 0,97. Ce résultat confirme l'aptitude des réseaux de neurones artificiels, de façon générale, à prédire par approximations dans le domaine de la modélisation des phénomènes dynamiques tels que la croissance de la tomate. Ce dernier a la capacité de bien représenter les relations complexes non-linéaires et convient pour l'approximation de la tomate. En effet, selon Chouai et ses collaborateurs [24] et Mohamed [31], le modèle développé par les réseaux de neurones artificiels est un modèle « type boîte noire » qui constitue la forme la plus primitive de modèle mathématique. Il est réalisé uniquement à partir des mesures expérimentales ou d'observations. Il peut avoir une valeur prédictive, dans un certain domaine de validité, mais il n'a aucune valeur explicative. De ce fait, il permet au mieux d'approcher la croissance de la tomate à partir des paramètres de croissance étudiés.

## 5 CONCLUSION

Cette étude a pour but d'utiliser les réseaux de neurones artificiels pour prédire la croissance des plants de tomate en temps réel et anticiper sur celle-ci afin de maîtriser la productivité de cette plante.

La modélisation a été réalisée à partir du modèle empirique utilisant le réseau de neurones artificiels à partir des mesures expérimentales faites au champ. Elle a été réalisée à partir de cinq paramètres de croissance de la plante de tomate. Il s'agit du nombre de feuilles, de la longueur et de la largeur des feuilles, de la hauteur et de la circonférence tronc. Ces mesures ont pu révéler la complexité de la croissance de la tomate qui a été capitalisée par des modèles mathématiques.

Les modèles utilisant comme outil le réseau de neurones artificiels ont permis de prédire avec succès la croissance des plantes de tomate à partir des variables opératoires. Ces modèles ont donné une adéquation étroite entre les valeurs des variables opératoires expérimentales et celles calculées par le réseau. Ils ont permis d'approcher de façon satisfaisante la croissance de la tomate.

## REFERENCES

- [1] "Côte d'Ivoire, appui à la mise en œuvre du NEPAD-PDDAA Profil de projet d'investissement bancable. Appui à la production de la banane plantain et aux productions maraîchères en zone de forêt". TCP/IVC/2903 (I) (NEPAD Ref 05/25 F) Vol IV de VI. p. 41, 2005.
- [2] B. Guede, "Le Partenaire. Bulletin de liaison de l'Anader, N°14 Août 2009, 4 p.
- [3] FAO L'actualité agricole en Méditerranée". Ed. CIHEAM, p. 33, 2012.
- [4] J.C. N'zi, C. Kouame, S.P.N. Assanvo, L. Fondio, A.H. Djidji & Sangare, "Evolution des populations de Bemisia tabaci Genn. Selon les variétés de tomates (*Solanum lycopersicum* L.) au centre de la Côte d'Ivoire". Sciences et Nature 7 (1): 31-40, 2010. <https://doi.org/10.4314/scinat.v7i1.59918>.
- [5] S. Soro, "Caractérisation agronomique et culture sous voile de variétés améliorées de tomate (*Lycopersicum esculentum* Mill.) en Côte d'Ivoire". Thèse de Doctorat d'Etat. Université d'Abobo-Adjamé, Abidjan p. 198, 2009.
- [6] B. Keating & P. Grace, "Using biophysical simulation models to assess the performance of agriculture systems in risky environments. In: CIMMYT – Risk Management for maize Farmers in Drought56 prone areas of southern Africa; Proceedings of a workshop held at Kadoma Ranch, Zimbabwe", 1 - 3 October 1997. Mexico, D. F.: 67 – 82 pp, 2003.
- [7] C.O. Stöckle, M. Donatelli & R. Nelson, "CropSyst, a cropping systems simulation model". European journal of agronomy, 18 (3-4): 289-307, 2003.
- [8] I. Van, M.K. Ewert, F. Heckeley, T. Wery, J. Alkan Olsson, J. Andersen, E. Bezlepkina, I. Brouwer, F. Donatelli, M. Flichman, G. Olsson, L. Rizzoli, A.E. van der Wal, T. Wien & J.E. Wolf, "Integrated assessment of agricultural systems - A component-based framework for the European Union (SEAMLESS)". Agricultural Systems 96: 150-165, 2008.
- [9] N. Brisson, C. Gary, E. Justes, R. Roche, B. Mary, D. Ripoche, D. Zimmer, G. Sierra, P. Bertuzzi, P. Burger, F. Bussière, Y.M. Cabidoche, P. Cellier, P. Debaeke, J.P. Gaudillère, C. Hénault, F. Maraux, B. Seguin & H. Sinoquet, "An overview of the crop model STICS". European Journal of Agronomy. 18: 309-322, 2003 [https://doi.org/10.1016/S1161-0301\(02\)00110-7](https://doi.org/10.1016/S1161-0301(02)00110-7).
- [10] N. Brisson, M. Launay, B. Mary & N. Beaudoin, "Conceptual basis, formalizations and parametrization of the STICS crop model ions". Edit QUAE p. 252, 2009.
- [11] C. Garcia, I. Aauri, N. Brisson & B. Segnin, "Estimation de l'impact du changement climatique sur les résultats agronomiques de la vigne avec le modèle STICS. Actes Mondiat". ITV ed. France, 12-15, 2004.
- [12] RGPH, "Recensement Général de la Population et de l'Habitat. Rapport d'exécution et présentation des résultats", p. 49, 2014.

- [13] A. Sangaré, E. Koffi, F. Akamou & C. Fall, "État des ressources phylogénétiques pour l'alimentation et l'agriculture: second rapport national. Ministère de l'agriculture, République de Côte d'Ivoire", p. 16, 2009.
- [14] R. Ligban, L.D. Gone, B. Kamagaté, M.B. Saley & J. Biemi, "Processus hydrogéochimiques et origine des sources naturelles dans le degré carré de Daloa Centre ouest de la Côte d'Ivoire". International Journal of Biological and Chemical Sciences, 3 (1): 38-47, 2009.
- [15] N. Shankara, L. Van, De Jeudi & M. Mardja, "La culture de la tomate production, transformation et commercialisation". p. 20, 2005.
- [16] N. Kouame, "Modélisation de la croissance du bananier plantain à partir des réseaux de neurones artificiels; cas d'Azaguié et Yamoussoukro", Thèse de Doctorat Unique, Université Félix Houphouët-Boigny, Côte d'Ivoire, p. 252, 2014.
- [17] G. Cybenko, "Approximation by superpositions of a sigmoidal function: Math. Control, Signals and Systems". 2, 303-314, 1989. <https://doi.org/10.1007/BF02551274>.
- [18] K. Hornik, "Approximation capabilities of multilayer feed forward networks". Neur. Net. comp, 4: 251-257, 1991.
- [19] D.E. Rumelhart, "Parallel Distributed Processing, Cambridge": MIT Press: The Organization of Behavior, New York: Wiley. 1: 318-362, 1986.
- [20] M. Thompson, J. Walton & S.J. Wood, "Statistical appraisal of interference effects in determination of trace elements". Ana. Sci. Chem, 104: 299-312, 1979.
- [21] M.T. Hagan, H.B. Demuth & M.H. Beale, "Neural network design". PWS, Boston, MA, p. 259, 1996.
- [22] J.C. Principe, N.R. Euliano & W.C. Lefebvre, "Neural and Adaptive systems: Fundamentals through Simulation". Inc.Stil, Irwin, p.563, 2000.
- [23] M. Feinberg, "La validation des méthodes d'analyse; une approche chimométrique de l'assurance qualité au laboratoire". Ed. Mas, Paris. p 56-57, 1996.
- [24] A. Chouai, M. Cabassud, V. Le lanm, C. Gourdon & G. Casamatta, "Use of neural networks for liquid-liquid extraction column modelling: an experimental study". Chem. Eng. And Proces. 39: 171-180, 2000.
- [25] S. Chevret, "Modèles mathématiques utilisés en médecine en France". Edition Masson (Sci. Dir.). 2 (16): 240-244, 2007. <https://doi.org/10.1016/j.reaurg.2007.03.004>.
- [26] Y. Le Cun, J.S. Denker & S.A. Solla, "Optimal brain damage: proceeding of the neural information processing systems". Sci. comp.contr, 2: 598-605, 1990.
- [27] G. Dreyfus, J.M. Martinez, M. Samuelides, B. Gordonm, F. Badran, S. Thiria & Herault, "Réseaux de Neurones Artificiels". New York, USA: Gr. Eyrol, p. 374, 2002.
- [28] W. Benaissa, "Développement d'une méthodologie pour la conduite en sécurité d'un réacteur continu intensifié", Thèse de Doctorat, I.N.P. Toulouse. p. 203, 2006.
- [29] N. Kouamé, N.E. Assidjo, A.E. Dick & A.P. Anno, "Plantain tree growth (Musa sp., AAB cultivar horn 1) modeling using the artificial neural networks method". European Scientific Journal, 33 (9): 272-285, 2013. <https://doi.org/10.19044/esj.2013.v9n33p%25p>.
- [30] N. Kouamé, N.E. Assidjo & A.E. Dick, "Illustration de la croissance du bananier plantain (Musa sp., AAB, cultivar corne 1) dans les localités de Yamoussoukro et d'Azaguié (Côte d'Ivoire) à l'aide des modèles mathématiques". Agronomie Africaine, 29 (2): 135-147, 2017.
- [31] Y.A. Mohamed, "Mise en œuvre de réseaux de neurones pour la modélisation de cinétiques réactionnelles en vue de la transposition batch/continu". Thèse de Doctorat Unique. Institut National Polytechnique, Toulouse, France, p. 156, 2007.