

الجمهورية الجزائرية الديمقراطية الشعبية

RÉPUBLIQUE ALGÉRIENNE DÉMOCRATIQUE ET POPULAIRE

Ministère de l'Enseignement Supérieur et de la Recherche Scientifique



UNIVERSITÉ DE BLIDA 1

Faculté des Sciences

Département de Mathématiques

**Mémoire de Fin d'Etudes**

En vue de l'obtention de diplôme de MASTER

**Spécialité:** Recherche Opérationnelle

**Thème**

---

Les prévisions de la consommation dans le secteur des transports et la production de l'énergie électrique en l'Algérie à l'horizon 2030.

---

Présenté par : **ACHOUR Silya**

**ATTOU Fella**

Président	Boukhari Mohamed	MAA	USDB
Examineur	El moussaoui Hichem	MCB	USDB
Promoteur	M'raoui Abdelhamid	MRA	CDER
Co-promoteur	Tami Omar	MCB	USDB

Année universitaire : 2021-2022

## **Dédicaces**

Nous dédions ce modeste projet de fin d'étude :

A nos chères mamans, qui n'ont pas cessé de prier pour nous, et de nous encourager.

A nos papas qui ont contribué à la réalisation de nos rêves.

A nos frères, et sœurs, ce travail représente tous leurs soutiens, leurs aboutissements, et leurs encouragements.

A nos familles, pour leurs encouragements et leurs soutiens.

A nos chères amis, source d'espoirs et de motivation.

## Remercîments

Tout d'abords, on remercie dieu, tout puissant de nous avoir donnée le courage et la volonté d'accomplir ce travail.

Nous remercions nos chères parents, chères familles, amis pour leur soutien, et leurs encouragements, nos enseignants qui nous en aider et enseigner.

Nous tenons à remercier Mr M'raoui Abdelhamid, d'avoir accepté de nous encadrer, avec son sérieux, son aide et ses conseils précieux qui nous ont permis de rédiger ce mémoire. Nous restons et resterons très reconnaissantes.

Nous remercions Mr Elmoussaoui Hichem pour son aide et ses conseils.

Nous remercions Mr Boukhari Mohamed et Mr Tami omar d'avoir été à l'écoute tout au long de la réalisation de ce mémoire.

Nous remercions les membres de jury de nous honorer par leurs présences pour examiner notre travail.

## ملخص

نظرا لتحسن مستوى معيشي لسكان في الجزائر ومما أدى الى زيادة الطلب على الطاقة خاصة الكهرباء، أصبح من الضروري متابعة تطور استهلاك هذه الأخيرة. الهدف من دراستنا هو وضع توقعات موثوقة حول الاستهلاك المستقبلي لقطاع "النقل" خاصة المنتجات النفطية، الغاز الطبيعي وغاز البترول المسال. وكذلك قطاع الكهرباء، والحالات الخاصة للدورة المركبة والطاقة الشمسية، والتي تمثل أكبر حصة من إجمالي الاستهلاك النهائي للطاقة والكهرباء في الجزائر.

قدرت دراسة البيانات السابقة للاستهلاك خلال الفترة (1980-2019) أنه مع التطور الحالي ستستمر الكميات المستهلكة في الزيادة بشكل معتدل حتى عام 2030.

## Résumé

Vu l'amélioration de niveau de vie de la population en Algérie et ce qui a suit d'augmentation de demande d'énergie et particulièrement l'électricité, il est devenu nécessaire de suivre l'évolution de la consommation de cette dernière. L'objectif de notre étude est d'élaborer des prévisions fiables sur la consommation future du secteur « Transport » notamment les produit pétrolier, le gaz naturel et le GPL ; ainsi que le secteur de l'électricité, et les cas particuliers des cycle combiné et solaire, qui représentent la plus grande part de l'ensemble de la consommation finale de l'énergie et l'électricité en Algérie.

L'étude des données antérieures de la consommation durant la période (1980-2019) a estimé qu'avec l'évolution actuelle les quantités consommées continueront a augmenté modérément jusqu'en 2030.

## **Abstract**

Considering the improvement of the standard of living of the population in Algeria and what followed of increase of demand of energy and particularly electricity, it became necessary to follow the evolution of the consumption of the latter. The objective of our study is to elaborate reliable forecasts on the future consumption of the sector "Transport" in particular oil products, natural gas and LPG; as well as the sector of electricity, and the particular cases of the combined cycle and solar, which represent the greatest share of the whole final consumption of energy and electricity in Algeria.

The study of the previous data of consumption during the period (1980-2019) estimated that with the current evolution the consumed quantities will continue to increase moderately until 2030.

# TABLE DES MATIERES

<b>Chapitre 1</b>	<b>Généralité sur l'énergie.....</b>	<b>11</b>
1.1	<i>L'énergie renouvelable.....</i>	11
1.2	<i>L'énergie fossile.....</i>	13
1.3	<i>L'énergie en Algérie.....</i>	15
<b>Chapitre 2</b>	<b>Régression multilinéaire.....</b>	<b>18</b>
2.1	<i>Définition de la régression.....</i>	18
2.2	<i>Expression du modèle linéaire général.....</i>	19
2.3	<i>Méthode des moindres carrés.....</i>	20
2.3.1	<i>La droite des moindres carrés.....</i>	20
2.4	<i>Estimation des coefficients par la méthode des moindres carrés (MC) :.....</i>	21
2.5	<i>Propriétés.....</i>	23
2.5.1	<i>L'espérance mathématique des coefficients.....</i>	23
2.6	<i>Evaluation de la qualité du modèle.....</i>	26
2.6.1	<i>Examen graphique des résultats.....</i>	26
2.6.2	<i>Analyse de la variance.....</i>	26
2.7	<i>Analyse des résidus.....</i>	29
<b>Chapitre 3</b>	<b>Les réseaux de neurones artificiels.....</b>	<b>30</b>
3.1	<i>Biologie.....</i>	30
3.2	<i>Historique.....</i>	31
3.3	<i>Définition.....</i>	32
3.4	<i>Principe de fonctionnement.....</i>	33
3.5	<i>Fonction de perte.....</i>	35
3.5.1	<i>Régression.....</i>	35
3.5.2	<i>Classification.....</i>	36
3.6	<i>Choix de fonction de perte.....</i>	36
3.7	<i>L'optimiseur :.....</i>	37
3.7.1	<i>La descente du gradient.....</i>	37
3.7.2	<i>Adaptive moment estimation (ADAM).....</i>	38
3.8	<i>Complexité des ANN.....</i>	38
3.9	<i>Architecture des ANN.....</i>	38
3.9.1	<i>Les réseaux de neurones perceptron.....</i>	38
3.9.2	<i>Réseaux neurones FeedForward.....</i>	38
3.9.3	<i>Réseaux de neurone convolutif.....</i>	39
3.9.4	<i>Réseau neuronal récurrent.....</i>	39
<b>Chapitre 4</b>	<b>Résultats et prédiction.....</b>	<b>40</b>
4.1	<i>Analyse et interprétation.....</i>	40
4.1.1	<i>ANALYSE de la consommation par la méthode de régression.....</i>	40
4.1.2	<i>Analyse de la consommation d'électricité.....</i>	47
4.2	<i>Analyse et interprétation par la méthode des Réseaux de Neurones.....</i>	51
4.2.1	<i>Analyse de la consommation dans le secteur de transport.....</i>	51
4.2.2	<i>Analyse de la production d'électricité dans les 03 secteurs.....</i>	63
4.3	<i>Prédiction de la consommation du secteur de transport.....</i>	64
4.3.1	<i>Produit Pétrolier.....</i>	64
4.3.2	<i>Gaz Naturel.....</i>	65
4.3.3	<i>GPL.....</i>	66
4.4	<i>Prédiction de la production de l'électricité.....</i>	67
4.4.1	<i>Production Globale.....</i>	67
4.4.2	<i>Production solaire.....</i>	68
4.4.3	<i>Production des cycles combinés.....</i>	69

## LISTE DES TABLEAUX

Tableau 1 : Energie en Algérie.....	16
Tableau 2 : Fonctions de transfert $a = f(n)$ .....	34
Tableau 3. Tableau de choix de fonction de perte et de fonction d'activation.....	36

## LISTE D'ABRIVIATION

ENR : Energie Renouvelable.

$CO_2$ : Le dioxyde de carbone.

PIB : produit intérieur brut.

GPL : gaz de pétrole liquide.

MW: méga watt.

TWh: térawatt-heure.

Mtep : million de tonne équivalent pétrole.

MC : moindre carrée.

ANN : réseaux de neurones.

MSE: mean squared error.

SSE: sum of squares due to error.

## LISTE DES FIGURES

Figure 1: exemple d'une droite d'un moindres carrés .....	21
Figure 2. Schéma d'un neurone biologique. ....	31
Figure 3. Modèle d'un réseau de neurones artificiel. ....	32
Figure 4 : Fonctions d'activation : (a) du neurone « seuil », (b) du neurone « linéaire », (c) du neurone « sigmoïde ». ....	35
Figure 5 Régression linéaire pour les trois types de consommations.....	40
Figure 6 Régression exponentielle pour les trois types de consommations.....	42
Figure 7 Régression 4 <sup>ème</sup> degré pour les trois types de consommations .....	44
Figure 8 Régression linéaire pour les trois types de productions.....	47
Figure 9 Régression 4 <sup>ème</sup> degré pour les trois types de productions .....	49
Figure 10 Variation du nombre de neurones pour le secteur des transports – input année.....	51
Figure 11 Variation du nombre de neurones pour le secteur des transports – input année et PIB.....	52
Figure 12 Variation du nombre de neurones pour le secteur des transports – input année et population.....	53
Figure 13 Variation du nombre de neurones pour le secteur des transports – input année, PIB et population.....	54
Figure 14 Variation du nombre de neurones pour le gaz naturel dans les transports – input année.....	55
Figure 15 Variation du nombre de neurones pour le gaz naturel dans les transports – input année et PIB .....	56
Figure 16 Variation du nombre de neurones pour le gaz naturel dans les transports – input année et population.....	57
Figure 17 Variation du nombre de neurones pour le gaz naturel dans les transports – input année, PIB et population .....	58
Figure 18 Variation du nombre de neurones pour le GPL dans les transports – input année ..	59
Figure 19 Variation du nombre de neurones pour le GPL dans les transports – input année et PIB.....	60
Figure 20 Variation du nombre de neurones pour le GPL dans les transports – input année et population.....	61
Figure 21 Variation du nombre de neurones pour le GPL dans les transports – input année, PIB et population.....	62



Figure 22. coefficient de régression des réseaux de neurones pour la production globale d'électricité.....	63
Figure 23. coefficient de régression des réseaux de neurones pour la production des cycles combinés d'électricité.....	63
Figure 24. coefficient de régression des réseaux de neurones pour la production solaire d'électricité.....	63
Figure 25 : Prédiction de la consommation des produits pétrolier dans le secteur de transport durant les années à venir. ....	64
Figure 26 : Prédiction de la consommation du gaz naturel dans le secteur de transport durant les années à venir.....	65
Figure 27 : Prédiction de la consommation du GPL dans le secteur de transport durant les années à venir.....	66
Figure 28: Prédiction de la production d'électricité globale durant les années à venir.....	67
Figure 29 : Prédiction de la production d'électricité solaire durant les années à venir.....	68
Figure 30 : Prédiction de la production d'électricité par les cycles combines durant les années à venir.....	69

## INTRODUCTION

L'économie algérienne est fortement dépendante des revenus des hydrocarbures qui représentaient en 2020 21% du PIB, 41% des recettes fiscales et plus de 94% des exportations. La croissance que connaît le pays depuis 2001 repose sur une demande publique tirée principalement par les revenus des hydrocarbures. Or, le pétrole et le gaz sont deux ressources naturelles non renouvelables et inexploitablement indéfiniment ; la croissance infinie de leur demande est incompatible avec leur offre finie. Aujourd'hui, avec le niveau actuel des connaissances, l'approvisionnement en pétrole et en gaz de l'Algérie ne sera pas un facteur limitant majeur à l'échelle mondiale. Il peut le devenir à l'horizon qui ne peut être déterminé avec précision. Certes, cet horizon se rapproche chaque jour. Sachant que la demande de consommation de ces produits augmente de jour en jour dans notre pays.

La consommation d'électricité représentait 16 % de la consommation finale d'énergie totale en Algérie en 2020. Une centrale thermique à Cycle Combiné Gaz (CCG) produit de l'électricité grâce à la chaleur dégagée par la combustion de gaz naturel. Ce type de centrale combine deux types de turbine : une turbine à combustion et une turbine à vapeur reliées à un alternateur. Avec le même volume de combustible, ces deux turbines permettent de produire une quantité plus importante d'électricité.

Les centrales solaires convertissent l'énergie solaire en électricité, leur principal atout est le cout moins cher que les énergies classiques. Mais le principal inconvénient est leur intermittence, en effet, il y a de l'électricité lorsque le soleil est là et non quand la demande se fait ressentir.

Pour le secteur des transports, les produits pétroliers représentent la plus grande consommation. Ils totalisent les essences super, normal et sans plomb ainsi que le gasoil. Le GPL ainsi que le Gaz naturel sont aussi utilisés comme carburants pour les véhicules. Des bus au gaz naturel circulent dans la capitale et totalisent plusieurs milliers de kilomètres de trajets.

Notre étude vise à établir les prévisions de consommation dans le secteur des transports et la production de l'énergie électrique en l'Algérie à l'horizon 2022-2030.

Le mémoire se présente comme suit :

Le premier chapitre parlera en générale sur l'énergie ainsi que ces deux type l'énergie renouvelable et fossile et ensuit le contexte énergétique en Algérie.

Le chapitre deux exposera les théories des modèles utilisés, nous parlerons de la régression linéaire et ses formes et sa mathématique qui sera utilisé pour calculer, analyser et aider à la prédiction

Le chapitre trois abordera les réseaux de neurones, leur origine biologique puis leur transformation artificielle qui sera être utilisée pour calculer, analyser et aider à la prédiction.

Pour finir le chapitre quatre exposera les résultats des calculs par un programme Matlab. Les résultats seront interprétés en fonction des performances des modèles afin de permettre leur utilisation dans la prédiction des productions et consommations futurs.

# CHAPITRE 1 GENERALITE SUR L'ENERGIE

L'une des problématiques principales de notre société moderne est l'approvisionnement, la transformation et le contrôle de l'énergie. Les besoins croissants posent des problèmes de durabilité importants et l'économie des ressources énergétiques devient un problème crucial afin de maintenir de façon durable et écologique la persistance de notre monde.

L'énergie est décomposée en deux types : l'énergie renouvelable et fossile « non renouvelable ».

## 1.1 L'ENERGIE RENOUEVELABLE

Les énergies renouvelables (ENR), également appelées énergies alternatives, sont des énergies durables, utilisables dérivées de sources renouvelables naturelles illimitées telles que le soleil (énergie solaire), le vent (énergie éolienne), les rivières (énergie hydroélectrique), les sources chaudes (énergie géothermique), les marées (énergie marémotrice) et la biomasse (biocarburants)... elles permettent de produire de l'électricité, de la chaleur et du gaz.

Au début du 21<sup>e</sup> siècle, environ 80 % de l'approvisionnement énergétique mondial provenait de combustibles fossiles. La combustion de combustibles fossiles engendre beaucoup de déchets ou de émissions polluants, ce qui permet d'avoir un impact notable sur l'environnement. En outre, elle libère du dioxyde de carbone (CO<sub>2</sub>), l'un des principaux gaz à effet de serre à l'origine du réchauffement de la planète. Les stocks de combustibles fossiles sont limités.

ENR permettent de réduire les émissions de gaz et protéger le climat, diminuer la pollution de l'air et la terre ainsi avoir une bonne santé, pouvoir économiser encore plus pour le futur à venir et aussi pouvoir maîtriser à long terme la facture énergétique.

Parmi ces ENR les 05 grandes familles sont :

1. Énergie Solaire qui produit de la chaleur et de l'électricité grâce à des photovoltaïques, des thermodynamiques et des thermiques.
2. Énergie Éolienne qui produit de l'électricité, elle est terrestre et en mer.

3. Energie Biomasse qui produit de la chaleur, chauffage ainsi de l'électricité.
4. Energie Hydraulique qui produit aussi de l'électricité.
5. Energie Géothermie qui produit de la chaleur.

L'utilisation des énergies renouvelables a augmenté de 3 % en 2020, la demande de tous les autres combustibles ayant diminué. Le principal moteur a été une croissance de près de 7 % de la production d'électricité à partir de sources renouvelables. Les contrats à long terme, l'accès prioritaire au réseau et l'installation continue de nouvelles centrales ont soutenu la croissance des énergies renouvelables malgré la baisse de la demande d'électricité, les difficultés de la chaîne d'approvisionnement et les retards de construction dans de nombreuses régions du monde. En conséquence, la part des énergies renouvelables dans la production mondiale d'électricité a bondi à 29 % en 2020, contre 27 % en 2019. L'utilisation de la bioénergie dans l'industrie a augmenté de 3 %, mais a été largement compensée par un déclin des biocarburants, la baisse de la demande de pétrole ayant également réduit l'utilisation de biocarburants mélangés. La production d'électricité renouvelable en 2021 à augmenter de plus de 8 % pour atteindre 8 300 TWh, soit la plus forte croissance en glissement annuel depuis les années 1970. L'énergie solaire photovoltaïque et l'énergie éolienne ont contribué aux deux tiers de la croissance des énergies renouvelables.

Les scientifiques et les ingénieurs travaillent constamment à l'exploitation d'autres sources d'énergie renouvelables. Quatre des plus prometteuses sont l'énergie marémotrice, l'énergie des vagues et le carburant à base d'algues (ou algues), l'énergie de l'hydrogène.

Maintenant que nous disposons de moyens de plus en plus innovants et moins coûteux pour capter et conserver l'énergie éolienne et solaire, les énergies renouvelables deviennent une source d'énergie plus importante, représentant plus d'un huitième de la production américaine. L'expansion des énergies renouvelables se fait aussi bien à grande qu'à petite échelle, qu'il s'agisse de panneaux solaires installés sur les toits des maisons qui peuvent revendre de l'énergie au réseau ou de parcs éoliens offshore géants. Même des communautés rurales entières comptent sur les énergies renouvelables pour le chauffage et l'éclairage. Alors que l'utilisation des énergies renouvelables continue de croître, l'un des principaux objectifs sera de moderniser le réseau électrique, en le rendant plus intelligent, plus sûr et mieux intégré entre les régions.

L'Algérie possède de grandes réserves de pétrole et de gaz naturel et dépend largement de ces ressources pour générer des recettes d'exportation. Le secteur du gaz et du pétrole représente 45.9% du PIB algérien. Cependant, l'ENG ne compte que pour 3% de son mix énergétique, 97% de son électricité étant générée à partir d'énergies fossiles, en très grande majorité du gaz. À long terme, la reconduction du modèle national de consommation énergétique actuel peut rendre problématique l'équilibre offre-demande. De même, la production d'électricité, produite en totalité à partir des hydrocarbures se situe entre 75 à 80 TWh en 2020 et devra se situer entre 130 à 150 TWh en 2030. En ce sens et pour s'assurer de disposer d'approvisionnements énergétiques suffisants, l'intégration massive des énergies alternatives dans le mix énergétique constitue un enjeu majeur en vue de préserver les ressources fossiles, de diversifier les filières de production de l'électricité et beaucoup d'autre chose.

## **1.2 L'ENERGIE FOSSILE**

Les énergies fossiles sont des énergies produites par la combustion du charbon, du pétrole ou du gaz naturel.

Ces combustibles qui sont riches en carbone et hydrogène sont issus de la transformation de matières organiques enfouies au profond de notre terre pendant des millions d'années d'où le terme fossile; Ce sont des énergies non renouvelables puisqu'une fois utilisées elles ne peuvent être reconstituées qu'à l'échelle des temps géologiques très lointain.

On appelle hydrocarbures des composés chimiques dont les molécules sont constituées d'atomes de carbone et d'hydrogène, elles sont les principaux constituants du pétrole brut et du gaz naturel, ainsi que des produits pétroliers issus des raffineries.

Tous produits d'origine organique et extrait du sous-sol sont considérés comme des énergies non renouvelables.

Comment leurs gisements se sont-ils formés ?

Le charbon est l'enfouissement rapide de débris d'arbres et de plantes dans certains bassins sédimentaires, entraînant la formation lente et progressive du charbon, passant successivement par la tourbe, le lignite, le charbon, et enfin l'anthracite, qui est du carbone presque pur. La plupart du charbon exploitable s'est formé au Carbonifère (il y a 360 à 300

millions d'années) ou au Tertiaire plus récent. Exploité depuis des siècles, le charbon est extrait sous terre ou dans des mines à ciel ouvert. Ce fut l'un des facteurs qui contribua à la première révolution industrielle du XIXe siècle.

Le pétrole et le gaz sont des hydrocarbures qui se forment après des millions d'années. Cela commence par l'accumulation de matière organique dans les sédiments dans des environnements pauvres en oxygène et calmes, tels que les fonds d'océans ou de lacs fermés. Elles se transforment en boues sous l'action des bactéries, et sous l'action de l'augmentation de température provoquée par l'enfouissement progressif, des "roches mères" se forment. Les hydrocarbures (pétrole et gaz) s'y forment puis commencent à migrer vers la surface pour se concentrer, si les conditions sont réunies, dans des « roches réservoirs », dans les volumes fermés (pièges) que constituent les sédiments. Ces dépôts sont plus ou moins faciles à obtenir. Ils doivent d'abord être localisés et étudiés en plusieurs étapes d'exploration. Vient ensuite le temps de l'exploitation des gisements d'hydrocarbures, un processus long et très technique qui nécessite d'énormes investissements en capital et dépend des contraintes économiques actuelles.

Le pétrole brut ne peut être utilisé tel quel Il doit être transformé en carburants et en différents autres produits nécessaires aux industries chimiques et plastiques par le biais du raffinage, un processus en plusieurs étapes.

Le pétrole a été, depuis le début du XX<sup>e</sup> siècle, le facteur déclenchant de la deuxième révolution industrielle.

Les produits pétroliers sont des dérivés utilisables du pétrole brut issu de son raffinage, on cite certain de ces produits :

- Essence à moteur.
- Carburant diesel.
- Mazouts légers.
- Carburant d'aviation
- Mazouts lourds.
- Asphalte.
- Gaz de pétrole liquide.
- Gaz de chauffage.

Le gaz de pétrole liquéfié ou GPL est un mélange d'hydrocarbures légers stocké à l'état liquide et issu du raffinage du pétrole (40 % des ressources mondiales) et plus généralement du traitement du gaz naturel (60 % des ressources mondiales).

Le GPL est composé d'environ 80% de butane (C<sub>4</sub>H<sub>10</sub>) et 20% de propane (C<sub>3</sub>H<sub>8</sub>). C'est un carburant dit propre, issu du raffinage du pétrole brut. Son utilisation présente des avantages par rapport au gazole en permettant de réduire de :

- 50% les émissions d'oxyde d'azote,
- 50% celles des monoxydes de carbone,
- 90% celles des hydrocarbures et particules.

Le gaz naturel est plus léger que l'air alors que le GPL est plus lourd. Le gaz naturel est stocké à l'état gazeux, le GPL est stocké dans le réservoir à l'état liquide. La température d'auto-inflammation est de 420°C pour le GPL (225°C pour l'essence) et 540° C pour le gaz naturel.

Sans cesse, de nouvelles technologies contribuent à modifier les conditions économiques de l'exploitation des hydrocarbures, en permettant d'accéder à des gisements nouveaux et de baisser les coûts de production.

Le paysage a ainsi été profondément modifié par la part grandissante de l'offshore (et, plus récemment, de l'offshore profond) dans l'exploitation pétrolière et gazière. Les exploitations se font par des profondeurs d'eau de plus en plus importantes, grâce à des technologies spectaculaires.

### **1.3 L'ENERGIE EN ALGERIE**

Le secteur de l'économie énergétique algérienne domine l'économie algérienne, les hydrocarbures représentant à eux seuls 60 % des recettes budgétaires et 98 % des recettes d'exportation ; En 2019, l'Algérie était le 16e producteur mondial de pétrole, le 10e producteur de gaz naturel et le 7e exportateur de gaz naturel ; Plus de 99 % de la production et de la consommation d'énergie, y compris le secteur électrique, proviennent des hydrocarbures.

Pourtant, l'Etat algérien commence à réfléchir à la résolution des problèmes écologiques en investissant dans les énergies renouvelables. Selon le Plan algérien de développement des



énergies renouvelables et de l'efficacité énergétique (PENREE) de 2012, l'Algérie vise à atteindre 22 000 MW de capacité installée d'énergie renouvelable d'ici 2030. Mais sept ans après le programme, ce qui a été réalisé est les réalisations sont modestes : le solaire n'a produit que 0,7 % de l'électricité du pays en 2017, et l'éolien 0,01 %.

	Populatio n	Consommatio n d'énergie primale	Productio n	Exportatio n nette	Consommatio n d'électricité	Emissio n du $CO_2$
Année	Millions	Mtep	Mtep	Mtep	TWh	Mt $CO_2$ éq
1990	25.19	22.9	100.11	77.34	13.69	51.16
2000	31.18	26.99	142.21	114.96	21.12	61.49
2008	34.81	37.27	162.03	123.63	32.90	88.10
2009	35.40	40.76	153.04	111.64	30.61	94.41
2010	36.04	40.09	150.51	109.41	36.58	95.77
2011	36.72	41.82	145.83	103.26	41.18	102.11
2012	37.44	45.97	143.76	97.13	46.28	110.67
2013	38.19	47.58	137.67	89.19	48.78	113.88
2014	38.3	51.77	143.22	89.66	53.05	122.93
2015	-----	54.29	142.78	88.78	57.57	-----
2016	-----	53.75	153.28	98.87	60.06	-----
2017	41.3	55.52	152.85	96.47	64.81	130.5
Variatio n de 1990- 2017	+59%	+150%	+53%	+25%	+373%	155%

**TABLEAU 1 : ENERGIE EN ALGERIE**

Les hydrocarbures représentent 60 % des recettes budgétaires et 98 % des recettes d'exportation. Le modèle rentier, la « malédiction », a dévoré capitaux et talents, poussant le pays à importer tout ce dont il avait besoin au point de réduire la fabrication à 5 % du PIB. Les importations sont passées de 9 milliards de dollars en 1990 à 60 milliards de dollars en

2016, tandis que les exportations ont diminué de moitié depuis 2014 en raison de la chute des prix du pétrole et du gaz, mais ont également commencé à décliner en raison d'un manque d'investissement dans de nouveaux prix ; étant donné que sa rentabilité au Sahara est si faible que le gaz de schiste ne pourra pas prendre le relais.

En 2015, l'Algérie a généré 98 % de ses recettes en devises et 60 % de ses recettes budgétaires à partir du gaz et du pétrole ; elle estime que ces ressources ont fondu depuis 2014 avec la chute du prix du baril. Les revenus en 2016 n'auraient dû être que de 22 milliards de dollars, contre 58 milliards de dollars en 2014. De plus, le développement du gaz de schiste aux États-Unis a réduit l'un des principaux débouchés d'exportation du pays. L'État a prévu d'économiser environ 20 milliards de dollars en 2016 en supprimant plusieurs subventions à la consommation, ce qui a permis de connaître une augmentation générale des produits de première nécessité et des prix du gaz naturel, de l'électricité et des carburants.

Le 14 novembre 2019, le Parlement algérien a abrogé la fameuse loi « 49-51 », qui limitait à 49 % la participation des sociétés étrangères dans les sociétés algériennes. Il a également adopté une loi ouvrant l'industrie des hydrocarbures aux entreprises multinationales. Sonatrach, seul opérateur public, n'a plus les moyens d'explorer et de développer de nouvelles zones, mais de nombreux appels d'offres n'ont pas trouvé de destinataires en raison d'un cadre juridique et fiscal qui décourage les étrangers. Le pétrole et le gaz fournissent les deux tiers des recettes budgétaires et 95 % des entrées de devises.

## CHAPITRE 2 REGRESSION MULTILINEAIRE

En mathématiques, la régression recouvre plusieurs méthodes d'analyse statistique permettant d'approcher une variable à partir d'autres qui lui sont corrélées. Par extension, le terme est aussi utilisé pour certaines méthodes d'ajustement de courbe. En apprentissage automatique, on distingue les problèmes de régression des problèmes de classification. Ainsi, on considère que les problèmes de prédiction d'une variable quantitative sont des problèmes de régression tandis que les problèmes de prédiction d'une variable qualitative sont des problèmes de classification. Certaines méthodes, comme la régression logistique, sont à la fois des méthodes de régression au sens où il s'agit de prédire la probabilité d'appartenir à chacune des classes et des méthodes de classification...etc.

Son objectif est d'apprendre à analyser et prédire un événement quelconque statistique.

### 2.1 DEFINITION DE LA REGRESSION

La régression est une méthode de prévision mathématique très utilisée en économie. À partir d'un ensemble de valeurs expérimentales, représentées par des points sur un graphique, on cherche à calculer la courbe qui reproduit le mieux les variations de la grandeur à étudier, c'est-à-dire celle qui passe par tous les points ou le plus proche possible. La régression est donc l'opération qui consiste à faire passer une droite ou tout autre courbe mathématique "le plus près possible" d'un certain nombre de points obtenus d'une manière expérimentale.

Une des méthodes les plus employées pour obtenir un résultat fiable est celle des "moindres carrés" qui permet de tenir compte, dans une plus grande mesure, des points éloignés de la droite moyenne, afin de ne pas fausser cette dernière en négligeant involontairement une donnée atypique. Disposant de l'équation de cette courbe, on peut alors effectuer des interpolations, pour calculer l'ordonnée de points intermédiaires, mais aussi des extrapolations, pour prévoir les valeurs que prendra la variable de test dans l'avenir.

Les tableurs permettent, pour affiner les corrélations, d'utiliser un modèle linéaire une droite d'équation  $Y = a * X + b$  — mais aussi des modèles logarithmiques ou exponentiels. Le

modèle exponentiel est utilisé pour illustrer des phénomènes dont les variations sont très rapides : l'accroissement de la variable est ici proportionnel à sa valeur.

## 2.2 EXPRESSION DU MODELE LINEAIRE GENERAL

Le modèle linéaire général est un outil fondamental pour les analyses statistiques qui consiste à considérer comme modèle de régression une fonction multilinéaire :

$$f : R^n \rightarrow R$$

### Les hypothèses suivantes :

1- Les  $\varepsilon_j$  sont des termes d'erreur, non observés, indépendants et identiquement distribués ;  
 $E(\varepsilon_j) = 0$  et  $\text{Var}(\varepsilon) = \sigma^2 I$ .

2- Les termes  $X_j$  sont supposés déterministes (facteurs contrôlés) ou bien l'erreur  $\varepsilon$  est indépendante de la distribution conjointe de  $x_1, \dots, x_p$ . On écrit dans ce dernier cas que :

$$E(Y | x_1, \dots, x_p) = a_0 + a_1 x_1 + \dots + a_p x_p \quad \text{et} \quad \text{Var}(Y | x_1, \dots, x_p) = \sigma^2$$

3- Les paramètres inconnus  $a_0, \dots, a_p$  sont supposés être constants.

4- En option, pour l'étude spécifique des lois des estimateurs, une quatrième hypothèse considère la normalité de la variable d'erreur  $\varepsilon$  ( $N(0, \sigma^2 I)$ ).

Les  $\varepsilon_i$  sont alors i.i.d. de loi  $N(0, \sigma^2)$ .

Les données sont rangées dans une matrice  $X$  ( $n \times (p+1)$ ) de terme général  $x_j^i$ , dont la première colonne contient le vecteur 1 ( $x_0^i = 1$ ), et dans un vecteur  $Y$  de terme général  $y_i$ . En notant les vecteurs  $\varepsilon = [\varepsilon_1 \dots \varepsilon_n]$  et  $A = [a_0 \dots a_p]$  Le modèle s'écrit en forme matricielle:

$$y = XA + \varepsilon.$$

## 2.3 METHODE DES MOINDRES CARRES

Une situation courante en mathématique est d'avoir à sa disposition deux ensembles de taille  $n$ ,  $\{y_1, y_2, \dots, y_n\}$  et  $\{x_1, x_2, \dots, x_n\}$ , obtenus expérimentalement ou mesurés. Le problème de la régression consiste à rechercher une relation pouvant éventuellement exister entre les  $x$  et les  $y$ , par exemple de la forme  $y = f(x)$ . Lorsque la relation recherchée est affine, c'est-à-dire de la forme  $y = ax + b$ , on parle de régression linéaire. Mais même si une telle relation est effectivement présente, les données mesurées ne vérifient pas en général cette relation exactement. Pour tenir compte dans le modèle mathématique des erreurs observées, on considère les données  $\{y_1, y_2, \dots, y_n\}$  comme autant de réalisations d'une variable aléatoire  $Y$  et parfois aussi les données  $\{x_1, x_2, \dots, x_n\}$  comme autant de réalisations d'une variable aléatoire  $X$ . On dit que la variable  $Y$  est la variable dépendante ou variable expliquée et que la variable  $X$  est la variable explicative.

### 2.3.1 LA DROITE DES MOINDRES CARRES

Les données  $\{(x_i, y_i), i = 1 \dots n\}$  peuvent être représentées par un nuage de  $n$  points dans le plan  $(x, y)$ , le diagramme de dispersion. Le centre de gravité de ce nuage peut se calculer facilement : il s'agit du point de coordonnées

$$(\bar{x}, \bar{y}) = \left( \frac{1}{n} \sum_{i=1}^n x_i ; \frac{1}{n} \sum_{i=1}^n y_i \right)$$

Rechercher une relation affine entre les variables  $X$  et  $Y$  revient à rechercher une droite qui s'ajuste le mieux possible à ce nuage de points. Parmi toutes les droites possibles, on retient celle qui jouit d'une propriété remarquable : c'est celle qui rend minimale la somme des carrés des écarts des valeurs observées  $y_i$  à la droite  $y_i = ax_i + b$ .

Si  $\varepsilon_i$  représente cet écart, appelé aussi résidu, le principe des moindres carrés ordinaire (MCO) consiste à choisir les valeurs de  $a$  et de  $b$  qui minimisent.

$$E = \sum_{i=0}^n \varepsilon^2 = \sum_{i=0}^n (y_i - (ax_i + b))^2$$

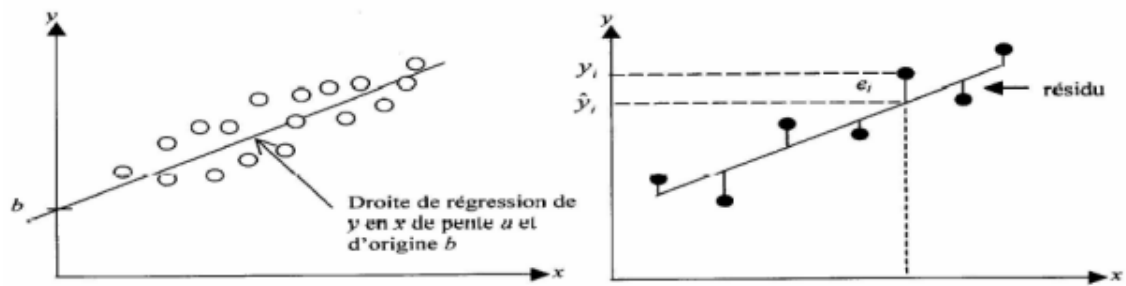


FIGURE 1: EXEMPLE D'UNE DROITE D'UN MOINDRE CARREE

Un calcul montre que ces valeurs, notées  $\tilde{a}$  et  $\tilde{b}$  sont égales à :

$$\begin{cases} \tilde{a} = \frac{\sum_{i=0}^n (x_i - \bar{x})(y_i - \bar{y})}{\sum_{i=0}^n (x_i - \bar{x})^2} \\ \tilde{b} = \bar{y} - \tilde{a}\bar{x} \end{cases}$$

On exprime souvent  $a$  au moyen de la variance de  $x$ ,  $s_x^2$  et de la covariance des variables.

## 2.4 ESTIMATION DES COEFFICIENTS PAR LA METHODE DES MOINDRES CARRES (MC) :

Lorsque on estime les inconnues  $a_1 a_2 \dots \dots a_n$  par  $\tilde{a}_1 \tilde{a}_2 \dots \dots \tilde{a}_n$  nous pouvons calculer la réponse au point  $i$  par :

$$\hat{y} = \tilde{a}_0 + \tilde{a}_1 x_{i1} + \tilde{a}_2 x_{i2} + \dots + \tilde{a}_n x_{in}$$

La valeur  $\hat{y}_i$  diffère du résultat expérimental  $y_i$  de la quantité  $e_i$  (de même que les  $\tilde{a}_i$  sont les estimateurs de  $a_i$ , les  $e_i$ , sont des estimations des  $\varepsilon_i$ ).

$$y_i = \hat{y}_i + e_i$$

Ce système devient :

$$\begin{cases} \hat{y}_1 = \tilde{a}_0 + \tilde{a}_1 x_{11} + \tilde{a}_2 x_{12} + \dots + \tilde{a}_n x_{1n} + e_1 \\ \hat{y}_2 = \tilde{a}_0 + \tilde{a}_1 x_{21} + \tilde{a}_2 x_{22} + \dots + \tilde{a}_n x_{2n} + e_2 \\ \vdots \\ \hat{y}_n = \tilde{a}_0 + \tilde{a}_1 x_{n1} + \tilde{a}_2 x_{n2} + \dots + \tilde{a}_n x_{nn} + e_n \\ i = \overline{1..n} \end{cases}$$

On cherche la valeur de  $\tilde{a}_i$  qui minimise la somme des carrés des écarts  $\sum e_i^2$ . Plus précisément, nous recherchons le modèle linéaire le plus proche de tous les points expérimentaux.

Pour simplifier l'écriture, on utilise la notation matricielle :

$$\begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix} = \begin{pmatrix} 1 & x_{11} & \dots & x_{1n} \\ 1 & x_{21} & \dots & x_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ 1 & x_{n1} & \dots & x_{np} \end{pmatrix} \begin{pmatrix} \tilde{a}_0 \\ \tilde{a}_1 \\ \vdots \\ \tilde{a}_n \end{pmatrix} + \begin{pmatrix} e_1 \\ e_2 \\ \vdots \\ e_n \end{pmatrix}$$

Le système à résoudre s'écrit :  $y = X \hat{A} + e$ , et le critère des moindres carrés ( $e^t e$ ) soit minimale on a :

$$\begin{aligned} e^t e &= (y - X \hat{A})^t (y - X \hat{A}) \\ &= (y^t - X^t \hat{A}^t) (y - X \hat{A}) \\ &= y y^t - \hat{A}^t X y - y^t X \hat{A} + \hat{A}^t X^t X \hat{A} \\ &= y y^t - 2 \hat{A}^t X y + \hat{A}^t X^t X \hat{A} \end{aligned}$$

On dérive par rapport à  $\hat{A}$  est on obtient :

$$\frac{\partial e^t e}{\partial \hat{A}} = \frac{\partial y^t y}{\partial \hat{A}} - 2 \frac{\partial \hat{A}^t X y}{\partial \hat{A}} + \frac{\partial \hat{A}^t X^t X \hat{A}}{\partial \hat{A}}$$

Avec :  $-\frac{\partial y^t y}{\partial \hat{A}} = 0$  Puisque  $y^t y$  ne dépend pas de  $\hat{A}$ .

$-\frac{\partial \hat{A}^t X^t y}{\partial \hat{A}} = X^t y$  Puisque  $\hat{A}^t X^t y$  est une forme linéaire en  $\hat{A}$ .

$$-\frac{\partial \hat{A}^t X^t X \hat{A}}{\partial \hat{A}} = 2X^t X \hat{A} \quad \text{Puisque } \hat{A}^t X^t X \hat{A} \text{ est une forme quadratique en } \hat{A}.$$

$$\text{D'où :} \quad \frac{\partial e^t e}{\partial \hat{A}} = -2X^t y + 2X^t X \hat{A}$$

$$\text{Sachant que :} \quad \frac{\partial e^t e}{\partial \hat{A}} = 0$$

$$\text{Alors on aura :} \quad -2X^t y + 2X^t X \hat{A} = 0 \Rightarrow X^t X \hat{A} = X^t y$$

Si la matrice  $(X^t X)^{-1}$  n'est pas singulière alors on aura :

Cette relation est fondamentale, et nous l'utiliserons continuellement à l'avenir. C'est valable pour tous les modèles polynomiaux, quels que soient leur degré et leur nombre de coefficients. Les coefficients ainsi obtenus sont utilisés pour écrire des modèles mathématiques. Ce modèle permet de calculer la réponse dans tout le domaine (souvent appelée la réponse prédite  $\hat{y}$ ) dans tout le domaine d'étude.

$$\hat{y} = X \hat{A}$$

Elles sont très utiles pour faire des prévisions, tracer les diagrammes d'iso-réponses, chercher des optimums ou des valeurs bien précises. C'est un précieux outil de prévision. [1]

## 2.5 PROPRIETES

Les coefficients  $\tilde{a}_i$  sont des fonctions de résultats expérimentaux, elles doivent être considérées des variables aléatoires pour pouvoir calculer l'espérance mathématique et la variance des coefficients.

### 2.5.1 L'ESPERANCE MATHEMATIQUE DES COEFFICIENTS

Soit  $\hat{A} = (X^t X)^{-1} X^t y$  son espérance est :

$$\begin{aligned} E(\hat{A}) &= E[(X^t X)^{-1} X^t y] \\ &= (X^t X)^{-1} X^t E(y) \end{aligned}$$

Les éléments de X sont fixes.



En désignant un vecteur de coefficients par  $A$  et  $\varepsilon$  (erreur résiduelle) sont des vecteurs de  $N$  écarts entre résultats expérimentaux et réponses théoriques, alors :

Soit l'équation de système  $y = XA + \varepsilon$

$$E(y) = E(XA + \varepsilon)$$

$$= E(XA) + E(\varepsilon)$$

$$= XE(A)$$

$$= XA$$

Par hypothèse du MC on a  $E(\varepsilon) = 0$  et nous trouvons :

$$E(\hat{A}) = (X^t X)^{-1} X^t X A$$

$$= A$$

Sachant que  $(X^t X)^{-1} X^t X = I_n$

Le résultat trouver signifie que les distributions des  $\tilde{a}_i$  sont centrés sur les valeurs vrais  $de a_i$ . [1]

### 2.6.2 Variance des coefficients

Par définition la variance de  $\hat{A}$  est :

$$var(\hat{A}) = E[(\hat{A} - A)(\hat{A} - A)^t]$$

Remplaçons  $\hat{A}$  par  $(X^t X)^{-1} X^t y$  et  $y$  par  $XA + \varepsilon$  nous obtenons :

$$(\hat{A} - A) = (X^t X)^{-1} X^t (XA + \varepsilon) - A$$

$$= A + (X^t X)^{-1} X^t \varepsilon - A$$

$$= (X^t X)^{-1} X^t \varepsilon$$

Alors : 
$$(\hat{A} - A)^t = \varepsilon^t X (X^t X)^{-1}$$

Donc : 
$$\begin{aligned} \text{var}(\hat{A}) &= E[(X^t X)^{-1} X^t \varepsilon \varepsilon^t X (X^t X)^{-1}] \\ &= (X^t X)^{-1} X^t E(\varepsilon \varepsilon^t) X (X^t X)^{-1} \end{aligned}$$

On a : 
$$\begin{aligned} E(\varepsilon \varepsilon^t) &= E[(\varepsilon - 0)(\varepsilon - 0)^t] \\ &= \text{var}(\varepsilon) \\ &= \sigma^2 \end{aligned}$$

Ainsi on obtient : 
$$\begin{aligned} \text{var}(\hat{A}) &= (X^t X)^{-1} X^t \sigma^2 X (X^t X)^{-1} \\ &= \sigma^2 (X^t X)^{-1} X^t X (X^t X)^{-1} \\ &= \sigma^2 (X^t X)^{-1} \end{aligned}$$

En multipliant les éléments de la matrice  $(X^t X)^{-1}$  par  $\sigma^2$ , nous obtenons la matrice de variance-covariance :

$$\text{var}(\hat{A}) = \begin{pmatrix} \text{var}(\hat{a}_0) & \dots & & \dots & \text{cov}(\hat{a}_q, \hat{a}_0) \\ \vdots & & & & \vdots \\ & & \text{var}(\hat{a}_i) & \text{cov}(\hat{a}_j, \hat{a}_i) & \\ & & \text{cov}(\hat{a}_i, \hat{a}_j) & \text{var}(\hat{a}_j) & \\ \vdots & & & & \vdots \\ \text{cov}(\hat{a}_0, \hat{a}_q) & \dots & & \dots & \text{var}(\hat{a}_q) \end{pmatrix}$$

La matrice de variance-covariance ci-dessus est comme  $(X^t X)^{-1}$ , une matrice carrée symétrique. Les éléments de la diagonale sont les variances des coefficients, tandis que les termes hors diagonale correspondent aux covariances. [1]

## 2.6 EVALUATION DE LA QUALITE DU MODELE

Pour savoir si le modèle exprime bien le bon résultat, on utilise certaines techniques statistiques :

### 2.6.1 EXAMEN GRAPHIQUE DES RESULTATS

Quand cet examen est possible, il permet d'éviter de grandes fautes d'interprétation.

### 2.6.1 ANALYSE DE LA VARIANCE

L'analyse de la variance joue un rôle très particulier en statistique, elle a pour but d'étudier l'influence des variables quantitative ; on s'intéressera dans le cas où les facteurs sont fixes par l'expérimentateur.

L'analyse de la variance se résume à une comparaisant multiples de moyennes de différents échantillons.

#### 2.6.1.1 Propriétés d'y, $\hat{y}$ et e

Pour le modèle des moindres carrées on a :

Les n résultats expérimentaux  $y_i$  et les réponses calculées par le modèle  $\hat{y}_i$  ont la même moyenne générale  $\bar{y}$ .

La moyenne des écarts  $e_i$  est égale à 0.

La somme des carrées vérifient la relation suivante :

$$\sum_{i=1}^n e_i^2 = \sum_{i=1}^n y_i^2 - \sum_{i=1}^n \hat{y}_i^2$$

On démontre cette équation :

$$e^t e = y^t y - 2\hat{A}^t X^t y + \hat{A}^t X^t X \hat{A}$$

$$X^t y = X^t X \hat{A}$$

$$e^t e = y^t y - \hat{A}^t X^t X \hat{A}$$

On utilise les scalaires correspondant :  $e^t e = \sum_{i=1}^n e_i^2$

$$y^t y = \sum_{i=1}^n y_i^2$$

$$\hat{A}^t X^t X \hat{A} = \hat{y}^t \hat{y} = \sum_{i=1}^n \hat{y}_i^2$$

D'où coup les propriétés de  $y$  et  $\hat{y}$ ,  $e$  sont la base de l'analyse de variance.

### 2.6.1.2 Signification du modèle [1]

L'équation précédente devient :  $\sum_{i=1}^n y_i^2 = \sum_{i=1}^n \hat{y}_i^2 + \sum_{i=1}^n e_i^2$

On retranche  $n\bar{y}^2$  aux deux membres de l'équation :

$$\sum_{i=1}^n y_i^2 - n\bar{y}^2 = \sum_{i=1}^n \hat{y}_i^2 - n\bar{y}^2 + \sum_{i=1}^n e_i^2$$

La dispersion des  $y_i$  autour de leur moyenne, encore appelée dispersion totale, se décompose en deux part :

- la dispersion des  $\hat{y}^2$  autour de leur moyenne. Elle correspond à la dispersion expliquée par le modèle.

- La dispersion des écarts autour de leur moyenne correspond à la dispersion résiduelle. Elle est appelée somme des carrés des écarts Les degrés de liberté correspondant à chacune de ces sommes de carrés sont :

$$\left\{ \begin{array}{l} n - 1 \text{ pour } \sum_{i=1}^n y_i^2 - n\bar{y}^2 \\ p \text{ pour } \sum_{i=1}^n \hat{y}_i^2 - n\bar{y}^2 \\ n - p - 1 \text{ pour } \sum_{i=1}^n e_i^2 \end{array} \right.$$

Par définition la variance est obtenue en divisant la somme des carrés par le nombre de degrés de liberté de ses sommes. On obtient ainsi les variances :

$$\left\{ \begin{array}{l} \text{variancetotal} = \frac{\sum_{i=1}^n y_i^2 - n\bar{y}^2}{n-1} \\ \text{varianced'ajustement} = \frac{\sum_{i=1}^n \hat{y}_i^2 - n\bar{y}^2}{p} \\ \text{variancerésiduelle} = \frac{\sum_{i=1}^n e_i^2}{n-p-1} \end{array} \right.$$

### 2.6.1.3 Le coefficient de corrélation multiple $R^2$

On appelle  $R^2$  le rapport de la partie des résultats expliqués par le modèle à la totalité de ce qu'il devrait expliquer :

$$R^2 = \frac{\sum_{i=1}^n (\hat{y} - \bar{y})^2}{\sum_{i=1}^n (y - n\bar{y})^2} = 1 - \frac{\sum_{i=1}^n e_i^2}{\sum_{i=1}^n (y - n\bar{y})^2}$$

On remarque que  $R^2$  est compris entre 0 et 1. S'il est proche de 1, veut dire les réponses sont bien expliquées, s'il est proche de 0, alors les résultats ne sont pas bien expliquer.

Attention ! le rapport  $R^2$  n'est pas une qualité d'un modèle. Tout dépend du nombre de résultats et du modèle choisi. Par exemple, si l'on possède deux réponses et que l'on souhaite utiliser un modèle du premier degré, on aura toujours un  $R^2$  de 1. En effet, par deux points il passe toujours une droite. Le modèle explique bien les deux réponses. Si l'on a trois réponses le  $R^2$  ne sera probablement plus de 1. Mais si on veut faire passer une courbe du second degré par les trois points, le  $R^2$  sera à nouveau égal à 1. C'est pour cette raison que certains auteurs ont introduit le  $R_{ajuster}^2$ . [1]

### 2.6.1.4 Le coefficient $R_{ajuster}^2$

Il est donné par la relation suivante : 
$$R_{ajuster}^2 = \frac{\frac{\sum_{i=1}^n e_i^2}{n-p-1}}{\frac{\sum_{i=1}^n y_i^2 - n\bar{y}^2}{n-1}}$$

Ces deux rapports varient entre 0 et 1. Ils mesurent en quelque sorte le degré d'ajustement du modèle. Plus ils sont voisins de 1 et plus le modèle représente bien les points expérimentaux. [1]

## **2.7 ANALYSE DES RESIDUS**

On connaît déjà qu'un résidu est la différence entre les réponses mesurées et les réponses calculées par la méthode des moindres carrées on a :  $\tilde{e} = y_i - \hat{y}_i$

On peut tracer des graphes en portant les résidus en ordre et d'autres grandeurs en abscisses. Le graphe ainsi indique l'évolution des résidus en fonction de la valeur des réponses. On peut donc détecter des distributions particulières de résidus.

## CHAPITRE 3 LES RESEAUX DE NEURONES ARTIFICIELS

### 3.1 BIOLOGIE

Pour mieux comprendre le concept des réseaux de neurones artificiels, il faut, tout d'abord, comprendre le concept des neurones biologiques de l'être humain.

Un neurone est une cellule nerveuse constituant la base du système nerveux spécialisée dans le traitement des signaux électriques. En biologie, le cerveau humain contient un grand nombre de neurones fortement interconnectés constituant des réseaux de neurones.

Chaque neurone est une entité autonome au sein du cerveau. Un neurone comprend un corps cellulaire ou cellule somatique ou soma, centre de contrôle de celui-ci, qui fait la somme des informations qui lui parviennent. Il traite ensuite l'information et renvoie le résultat sous forme de signaux électriques, du corps cellulaire à l'entrée des autres neurones au travers de son axone. Les axones reliant les neurones entre eux jouent donc un rôle important dans le comportement logique de l'ensemble. Le neurone est également constitué de plusieurs branches nommées dendrites, qui sont les récepteurs principaux du neurone, par lesquelles transite l'information venue de l'extérieur vers le corps cellulaire. Les synapses du neurone quant à elles reçoivent les informations des autres neurones via l'axone et permettent donc aux neurones de communiquer entre eux.

Les synapses ont une espèce de « mémoire » qui leur permet d'ajuster leur fonctionnement. En fonction de leur activation répétée ou non entre deux neurones, les connexions synaptiques vont se modifier. Certaines synapses vont être excitatrices et activer les neurones et d'autres vont être inhibitrices et stopper le signal électrique appelé aussi influx nerveux. Les synapses inhibitrices assurent le contrôle du neurone et son bon fonctionnement. L'action inhibitrice d'une synapse peut être suspendue temporairement dans certains cas tel que l'apprentissage. Chez certaines personnes, le mécanisme responsable de l'inhibition est altéré et entraîne une excitation durable des neurones pouvant conduire à l'épilepsie.

Le traitement de l'information par chaque neurone montre que celle-ci n'est pas stockée dans les neurones mais est bien le résultat du comportement de tout le réseau interconnecté.

L'information est donc principalement dans l'architecture et dans la force des connexions neuronales.

Les neurones biologiques ont deux propriétés qui peuvent être appliquées au niveau des neurones artificiels :

L'excitabilité : c'est-à-dire la capacité de réagir sous l'effet de stimulations extérieures et de convertir celles-ci en impulsions

La conductivité : c'est-à-dire la capacité de transmettre les impulsions nerveuses.

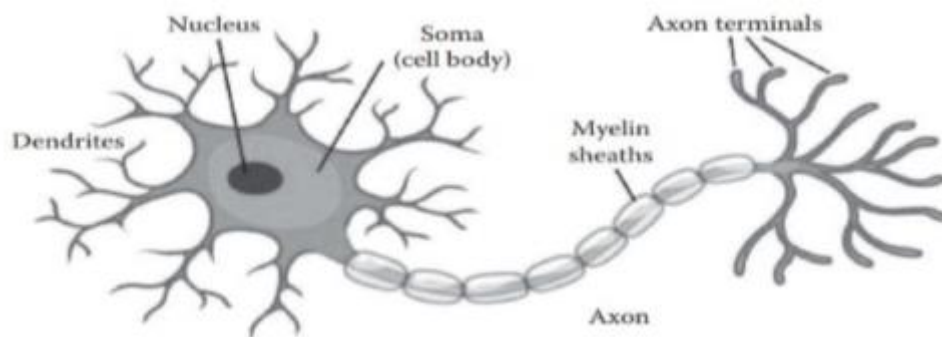


FIGURE 2. SCHEMA D'UN NEURONE BIOLOGIQUE.

### 3.2 HISTORIQUE

J. Mc Culloch et W. Pitts, sont les premiers à montrer en 1943, que des réseaux de neurones formels simples peuvent réaliser des fonctions logiques, arithmétiques et symboliques complexes (tout au moins au niveau théorique). [2]

Vers la fin des années 1940, Donald Hebb a ensuite proposé une théorie fondamentale pour l'apprentissage, en écrivant « The Organization of Behavior » mentionnant que si deux nerfs se déclenchent en même temps, la connexion entre eux est renforcée. [2]

En 1958, Rosenblatt et ses collègues ont construit un réseau et démontré ses habilités à reconnaître des formes. Un an après, B. Widrow, un automaticien, développe le modèle



ADALINE et MADALINE, le premier réseau neuronal appliqué à un problème du monde réel, utilisant un filtre adaptatif qui élimine les échos sur les lignes téléphoniques. [3]

En 1969, M. Minsky et S. Papert publient un ouvrage qui met en exergue les limitations théoriques du perceptron. Limitations alors connues, notamment concernant l'impossibilité de traiter par ce modèle des problèmes non linéaires. [3]

Aujourd'hui, les réseaux de neurones sont en voie de développement, et fournissent de nouvelles théories et algorithmes envahissant plusieurs domaines. [3]

### 3.3 DEFINITION

Un réseau neurones (ANN) est une série d'algorithmes qui cherche à reconnaître les relations sous-jacentes dans un ensemble de données par un processus qui imite le fonctionnement du cerveau humain. Il est constitué d'un ensemble de nœuds (neurones), qui sont reliés par des liens alimentés par un poids.

Les ANN jouent un rôle important dans l'apprentissage automatique (ML) et soutiennent le domaine plus large de la technologie de l'intelligence artificielle (AI). [4]

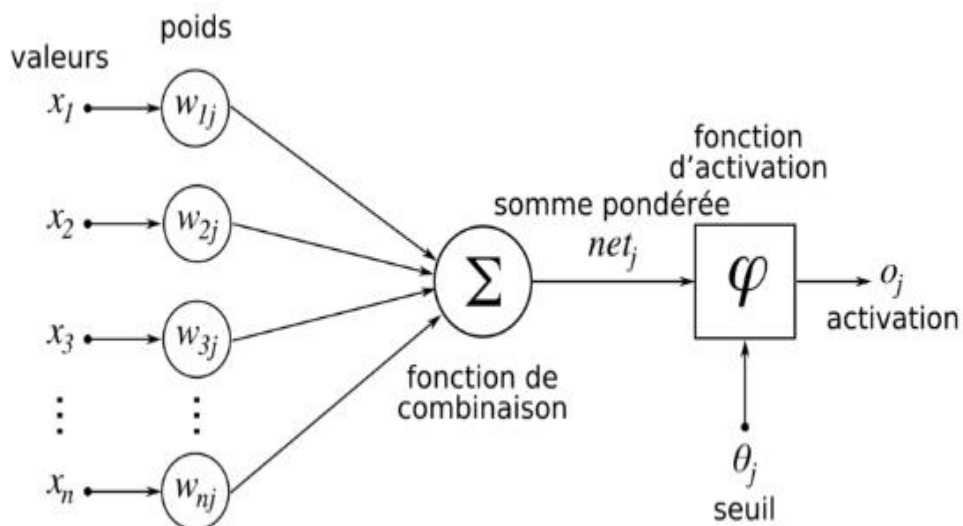


FIGURE 3. MODELE D'UN RESEAU DE NEURONES ARTIFICIEL. [5]

### 3.4 PRINCIPE DE FONCTIONNEMENT

Le principe du fonctionnement du réseau de neurones est qu'une fonction d'activation  $g$  appliqué à une somme  $S$  pour déterminer la sortie  $Y$ , la somme  $S$  est une addition de multiplication entre les signaux d'entrée  $X_i$  reçu par le neurone connecté et un poids  $W_i$

La somme pondérée s'écrit comme suit : [6]

$$S = \sum_{i=1}^n W_i * X_i \mp B$$

Avec :

$S$  : la somme pondérée

$W_i$  : poids du neurone

$X_i$ : *signal d'entrée*

$B$  : biais (ajouté pour affiner la transformation vers le signal de sortie.)

L'équation de la sortie  $Y$  :

$$Y = g\left(\sum_{i=1}^n W_i * X_i \mp B\right) = g(S)$$

Avec :

$Y$  : signal de sortie

$g$  : fonction d'activation de la somme pondérée.

### 3.5 Fonction d'activation

Différentes fonctions de transfert pouvant être utilisées comme fonction d'activation du neurone sont énumérées au tableau suivant : [7]

Fonction	Relation d'entrée/sortie
Seuil	$a = 0 \text{ si } n < 0$ $a = 1 \text{ si } n \geq 0$
Seuil symétrique	$a = -1 \text{ si } n < 0$ $a = 1 \text{ si } n \geq 0$
Linéaire	$a = n$
Linéaire saturée	$a = 0 \text{ si } n < 0$ $a = n \text{ si } 0 \leq n \leq 1$ $a = 1 \text{ si } n > 1$
Linéaire saturé symétrique	$a = -1 \text{ si } n < 0$ $a = n \text{ si } 0 \leq n \leq 1$ $a = 1 \text{ si } n > 1$
Linéaire positive	$a = 0 \text{ si } n < 0$ $a = n \text{ si } a \geq 0$
Sigmoïde	$a = \frac{1}{1 + e^{-n}}$
Tangente hyperbolique	$a = \frac{e^n - e^{-n}}{e^n + e^{-n}}$
Compétitive	$a = 1 \text{ si } n \text{ maximum}$ $a = 0 \text{ autrement}$
Sofmax	$a = \frac{e^{z_i}}{\sum_j e^{z_j}}$ <p>i= index du neurone sortie</p> <p>j=index de tous les neurones</p>

**TABLEAU 2 : FONCTIONS DE TRANSFERT  $a = f(n)$ . [7]**

Les trois les plus utilisées sont les fonctions « seuil » (en anglais « hard limite »), « linéaire » et « sigmoïde ».

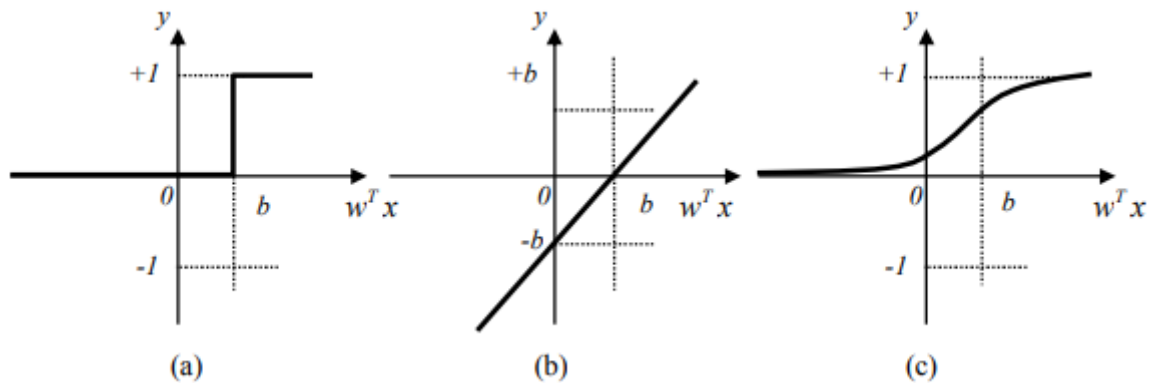


FIGURE 4 : FONCTIONS D'ACTIVATION : (A) DU NEURONE « SEUIL », (B) DU NEURONE « LINEAIRE », (C) DU NEURONE « SIGMOÏDE ».

### 3.5 FONCTION DE PERTE

L'objectif des algorithmes d'apprentissage automatique est de minimiser la fonction objective (ou de maximiser). La fonction objective est appelée « une fonction de perte » ou bien « fonction de coût ».

La fonction de perte évalue l'écart entre les prédictions réalisées par le réseau de neurones et les réponses réelles. Plus l'écart est minime, plus le réseau est précis. [8]

Les fonctions de perte peuvent être classées en deux catégories :

#### 3.5.1 REGRESSION

- Erreur quadratique :

$$MSE = \frac{1}{n} \sum_{i=1}^n (Y_i - \hat{Y}_i)^2$$

- Huber Loss :

$$L_{\delta}(Y, \hat{Y}) = \begin{cases} \frac{1}{2} (Y - \hat{Y})^2 & \text{si } |Y - \hat{Y}| \leq \delta \\ \delta |Y - \hat{Y}| - \frac{1}{2} \delta^2 & \text{sinon} \end{cases}$$

Avec :

$\delta$  : hyper paramètre ajustée.

### 3.5.2 CLASSIFICATION

- Entropie croisée binaire :

$$J = \sum_{i=1}^n Y_n \log(\hat{Y}) + (1 - \hat{Y}) \log(1 - \hat{Y})$$

- Entropie croisée multi-classe :

$$J = - \sum_{i=1}^n Y \log \hat{Y}$$

### 3.6 CHOIX DE FONCTION DE PERTE

Le choix de la fonction de perte réside dans le problème établi, Chaque choix doit adopter une fonction d'activation adéquate pour avoir une meilleure précision. Le choix des fonctions de perte et des fonctions d'activation se résume dans le tableau suivant :

Type de problème	Fonction d'activation	Fonction de perte
Classification binaire	Sigmoïde	Entropie croisée binaire
Multi-class classification à une seule étiquette	Softmax	Entropie croisée
Multi-class classification à plusieurs étiquettes	Sigmoïde	Entropie croisée binaire
Régression sur des valeurs arbitraire	Linéaire	Erreur quadratique moyenne
Régression sur des valeurs comprise entre 0 et 1	Sigmoïde	Erreur quadratique moyenne

**TABLEAU 3. TABLEAU DE CHOIX DE FONCTION DE PERTE ET DE FONCTION D'ACTIVATION.**

### 3.7 L'OPTIMISEUR :

Les optimiseurs sont des fonctions mathématiques utilisés pour minimiser la fonction de perte et de maximiser l'efficacité de la production. Citons : [9]

#### 3.7.1 LA DESCENTE DU GRADIENT

La descente du gradient est un algorithme d'optimisation basé sur une fonction convexe a pour but de minimiser (trouver le minimum local) la fonction de perte en faisant plusieurs itérations, et se déplaçant dans la descente opposée à celle du gradient. Cette méthode nécessite une grande quantité de mémoire et ralentit le processus. L'algorithme est donné par :

$$W_{i+1} = W_i - \alpha \frac{\partial J}{\partial W_i}$$

Avec :

$W_{i+1}$  : Nouveau poids

$W_i$  : Ancien poids

$\alpha$  : Taux d'apprentissage

$J$  : Fonction de perte.

##### 3.7.1.1 La descente du gradient stochastique :

Elle consiste à minimiser la fonction dans tous les échantillons du dataset, ce qui peut être lent si le data set est large.

##### 3.7.1.2 La descente du gradient stochastique avec Momentum :

Est une méthode d'optimisation stochastique développée de la descente du gradient stochastique. Elle consiste à ajouter un moment  $\beta$  à la descente du gradient stochastique régulière.

$$W_{i+1} = \beta W_i - \alpha \frac{\partial J}{\partial W_i}$$

### **3.7.2 ADAPTIVE MOMENT ESTIMATION (ADAM)**

L'optimiseur d'ADAM est l'un des algorithmes les plus populaires et les plus utilisés. Il combine entre deux autres méthodes, la descente du gradient avec Momentum et Adadelta.

### **3.8 COMPLEXITE DES ANN**

La complexité des réseaux de neurones se mesure par le nombre de neurones dans la couche cachée. Plus un réseau de neurones comporte de neurones, plus le système sera flexible mais également complexe. Des réseaux de neurones flexibles permettent d'approcher des fonctions de n'importe quelle complexité entre les variables d'entrée et les variables cible. Par conséquent, pour modéliser un fichier de données, il est important d'utiliser des réseaux de neurones relativement flexibles, contenant suffisamment de neurones dans la couche cachée. Le choix du nombre optimal de neurones dépend en grande partie du domaine de la problématique, mais d'une manière générale, il est lié au nombre d'entrées.

### **3.9 ARCHITECTURE DES ANN**

#### **3.9.1 LES RESEAUX DE NEURONES PERCEPTRON**

Le modèle du perceptron, est l'un des modèles de neurones les plus simples et les plus anciens. Il s'agit de la plus petite unité du réseau neuronal qui effectue certains calculs pour détecter des caractéristiques ou des renseignements commerciaux dans les données d'entrée. Il accepte des entrées pondérées et applique la fonction d'activation pour obtenir la sortie comme résultat final. Il s'agit donc d'un classificateur binaire.

On distingue deux types : le perceptron monocouche et le perceptron multicouche.

#### **3.9.2 RESEAUX NEURONES FEEDFORWARD**

La forme la plus simple de réseaux neuronaux où les données d'entrée se déplacent dans une seule direction, passant par des nœuds neuronaux artificiels et sortant par des nœuds de sortie.

On peut les classer en réseaux neuronaux de type FeedForward à monocouche ou multicouches. [10]

### **3.9.3 RESEAUX DE NEURONE CONVOLUTIF**

Le réseau neuronal à convolution contient une disposition tridimensionnelle des neurones, au lieu du réseau bidimensionnel standard. La première couche est appelée couche convolutive. Chaque neurone de la couche convolutive ne traite que les informations provenant d'une petite partie du champ visuel. Les caractéristiques d'entrée sont prises en compte par lots.

### **3.9.4 RESEAU NEURONAL RECURRENT**

Conçu pour sauvegarder la sortie d'une couche, le réseau neurone récurrent est réinjecté à l'entrée pour aider à prédire le résultat de la couche. La première couche est généralement un réseau neuronal à propagation directe, suivie d'une couche de réseau neuronal récurrent où une fonction de mémoire permet de se souvenir de certaines informations obtenues au cours de l'étape temporelle précédente. La propagation directe est mise en œuvre dans ce cas. Elle stocke les informations nécessaires à son utilisation future. Si la prédiction est fautive, le taux d'apprentissage est utilisé pour effectuer de petits changements. Ainsi, il augmente progressivement vers la bonne prédiction pendant la rétro propagation. [11]



# CHAPITRE 4 RESULTATS ET PREDICTION

## 4.1 ANALYSE ET INTERPRETATION

### 4.1.1 ANALYSE DE LA CONSOMMATION PAR LA METHODE DE REGRESSION

#### 4.1.1.1 Analyse de la consommation dans le secteur de transport

##### Régression linéaire

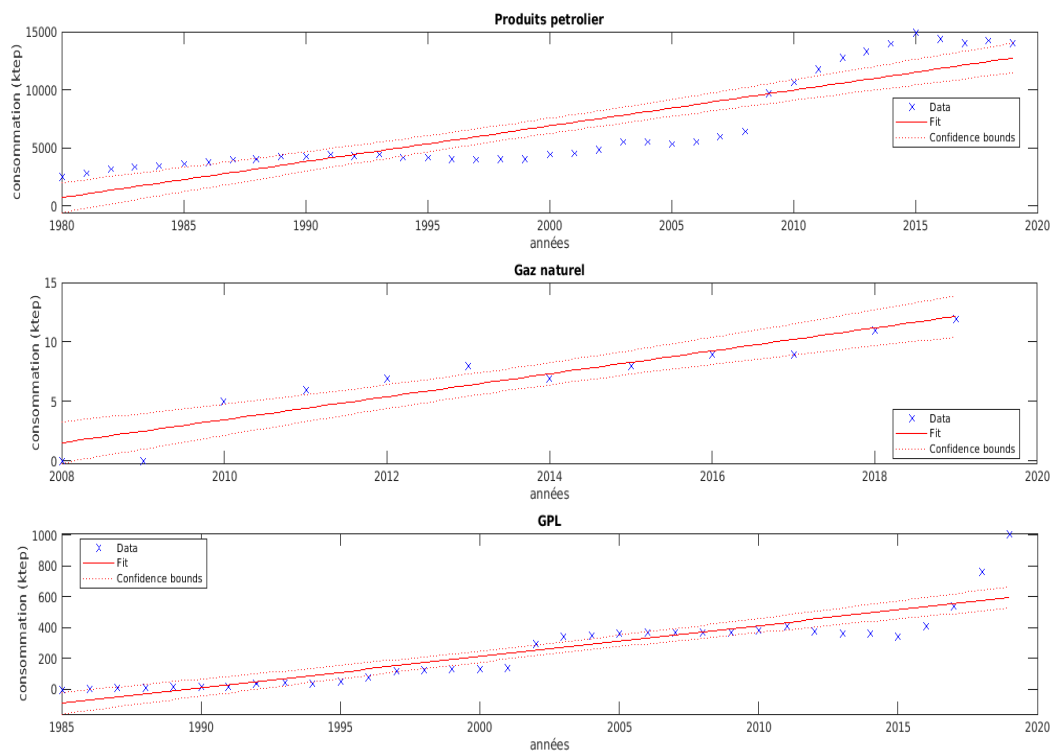


FIGURE 5 REGRESSION LINEAIRE POUR LES TROIS TYPES DE CONSOMMATIONS

Modèle mathématique linéaire pour les produits pétroliers

$$f(x) = p_1 * x + p_2$$

Coefficients:

$$p_1 = 307.7$$

$$p_2 = -6.086e^{05}$$

Goodness of fit:

$$SSE: 1.583e^{08}$$

Coefficient de corrélation: 0.761

$R_{ajuster}^2$ : 0.755

**Conclusion :**

Modèle rejetée.

Modèle mathématique linéaire pour le gaz naturel :

$$f(x) = p_1 * x + p_2$$

Coefficients :

$$p_1 = 0.2413$$

$$p_2 = -480.4$$

Goodness of fit:

SSE: 235.6

Coefficient de corrélation: 0.5684

$R_{ajuster}^2$ : 0.557

**Conclusion :**

Modèle rejetée.

Modèle mathématique linéaire pour le GPL :

$$f(x) = p_1 * x + p_2$$

Coefficients :

$$p_1 = 17.66$$

$$p_2 = -3.509e^{04}$$

Goodness of fit:

SSE:  $4.259e^{05}$

Coefficient de corrélation: 0.796

$R_{ajuster}^2$ : 0.7907

### **Interprétation**

Le graphe représente le taux de consommation du GPL selon le secteur de transport entre 1980 et 2020.

On remarque que le graphe ne fait qu'augmenter avec une vitesse acceptable vu la droite obtenue.

Donc avec le modèle linéaire on peut prédire que ce sera le cas pour les futures années.

## Régression exponentielle

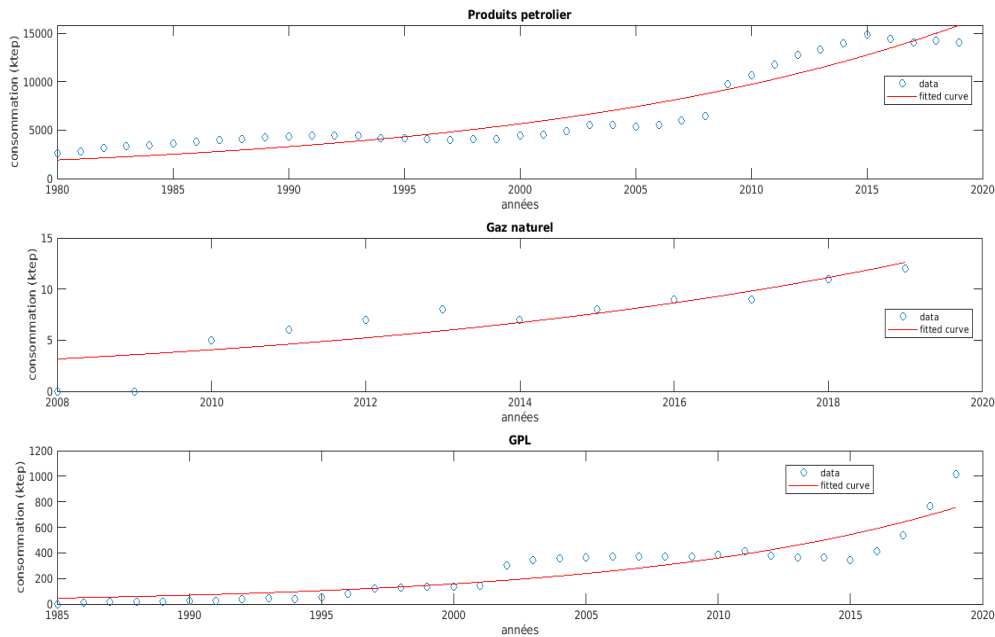


FIGURE 6 REGRESSION EXPONENTIELLE POUR LES TROIS TYPES DE CONSOMMATIONS

Modèle mathématique exponentielle pour les produits pétroliers

$$f(x) = a * e^{bx}$$

Coefficients :

$$a = 4.727e^{-44}$$

$$b = 0.0542$$

Goodness of fit:

$$SSE: 6.977e+07$$

Coefficient de corrélation: 0.8948

$$R_{ajuster}^2: 0.892$$

### Interprétation

Le graphe représente le taux de consommation des produits pétroliers selon le secteur de transport entre 1980 et 2020.

On remarque que le modèle choisi ici présent croit, ce qui signifie que la demande de consommation augmente avec un taux acceptable, cela prouve que durant les années à venir elle continuera d'augmenter puisque les produits pétrolier joue un rôle très important dans le domaine de transport en Algérie.

Le modèle exponentielle est l'un des modèles le plus adapter vu que le résidu ajuster est très proche de 1.

Modèle mathématique exponentielle gaz naturel :

$$f(x) = a * e^{bx}$$

Coefficients :

$$a = 9.353e^{-140}$$

$$b = 0.1598$$

Goodness of fit:

SSE: 55.37

Coefficient de corrélation: 0.8986

$R^2_{ajuster}$ : 0.8959

### **Interprétation**

Le graphe représente le taux de consommation du gaz naturel selon le secteur de transport entre 1980 à 2020.

On remarque que le modèle choisi ici présent croit à une vitesse lente, cela prouve que durant les années à venir il continuera d'augmenter avec une quantité acceptable.

Le modèle exponentielle est l'un des modèles le plus adapté vu que le résidu ajuster est très proche de 1.

Modèle mathématique exponentielle pour GPL :

$$f(x) = a * e^{bx}$$

Coefficients:

$$a = 1.659e^{-71}$$

$$b = 0.08401$$

Goodness of fit:

SSE: 3.097e+05

Coefficient de corrélation: 0.8517

$R^2_{ajuster}$ : 0.8478

### **Interprétation**

Le graphe représente le taux de consommation GPL selon le secteur de transport entre 1980 et 2020.

On remarque que le graphe croit, ce qui signifie que la demande de consommation augmente, et elle continuera d'augmenter tant qu'il n'y a pas un produit qui lui fera face et lui prendra sa valeur.

Le modèle exponentielle est l'un des modèles le plus adapté vu que le résidu ajuster est très proche de 1.

### Modèle 4<sup>ème</sup> degré

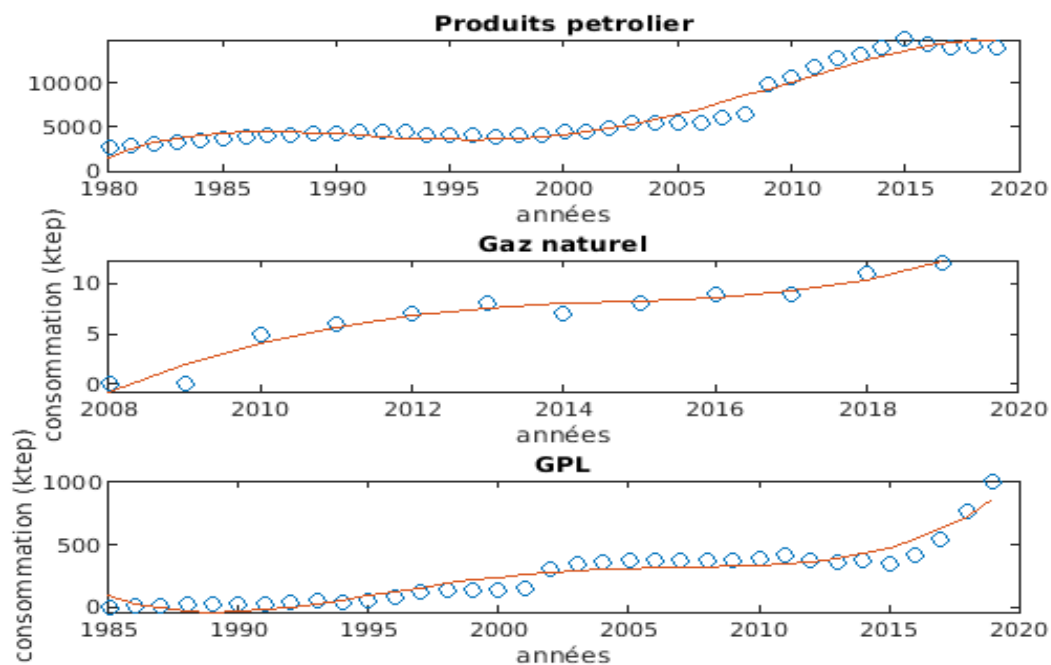


FIGURE 7 REGRESSION MODELE DE 4EME DEGRE POUR LES TROIS TYPES DE CONSOMMATIONS

Modèle mathématique 4eme degré pour produits pétrolier :

$$f(x) = p_1 * x^4 + p_2 * x^3 + p_3 * x^2 + p_4 * x + p_5$$

Coefficients :

$$p_1 = -0.06711$$

$$p_2 = 537$$

$$p_3 = -1.611e^{06}$$

$$p_4 = 2.149e^{09}$$

$$p_5 = -1.074e^{12}$$

Goodness of fit:

SSE: 2.397

Coefficient de corrélation: 0.9638

$R_{ajuster}^2$ : 0.9597

### **Interprétation**

Le graphe représente le taux de consommation des produits pétroliers selon le secteur de transport entre 1980 et 2020.

On remarque que le graphe durant la période 1980 et 1985 croit, et de 1985 à 1995 décroît, mais de 1995 à 2000, a augmenté d'une grande vitesse ce qui signifie que ça va encore augmenter durant les années à venir.

Le modèle quadratique est l'un des modèles le plus adapté vu que le résidu ajuster est très proche de 1 par rapport au modèle exponentielle.

Modèle mathématique 4eme degré pour le gaz naturel:

$$f(x) = p_1 * x^4 + p_2 * x^3 + p_3 * x^2 + p_4 * x + p_5$$

Coefficients :

$$p_1 = -1.623e^{-05}$$

$$p_2 = 0.1305$$

$$p_3 = -393.2$$

$$p_4 = 5.268e^{+05}$$

$$p_5 = -2.646e^{+08}$$

Goodness of fit:

SSE: 35.7759

Coefficient de corrélation: 0.9345

$R_{ajuster}^2$  :0.92

### **Interprétation**

Le graphe représente le taux de consommation du gaz naturel selon le secteur de transport entre 1980 et 2020.

On remarque que le graphe croit, d'où le taux de consommation augmentera durant les années à venir.

Le modèle quadratique est l'un des modèles le plus adapté vu que le résidu ajuster est très proche de 1 par rapport au modèle exponentielle.

Modèle mathématique 4eme degré pour GPL :

$$f(x) = p_1 * x^4 + p_2 * x^3 + p_3 * x^2 + p_4 * x + p_5$$

Coefficients :

$$p_1 = 0.003619$$

$$p_2 = -28.93$$

$$p_3 = 8.676e^{+04}$$

$$p_4 = -1.156e^{+08}$$

$$p_5 = 5.778e^{+10}$$

Goodness of fit:

SSE: 2.191

Coefficient de corrélation: 0.8951

$R_{ajuster}^2$ : 0.8831

### **Interprétation**

Le graphe représente le taux de consommation du GPL selon le secteur de transport entre 1980 et 2020.

On remarque que le graphe au début décroît, puis après l'année 1990, le taux de consommation augmente, d'où alors cette dernière va augmenter encore plus durant les années à venir.

Le modèle quadratique est l'un des modèles le plus adapté vu que le résidu ajuster est très proche de 1 par rapport au modèle exponentielle.

## 4.1.2 ANALYSE DE LA CONSOMMATION D'ÉLECTRICITÉ

### Régression linéaire

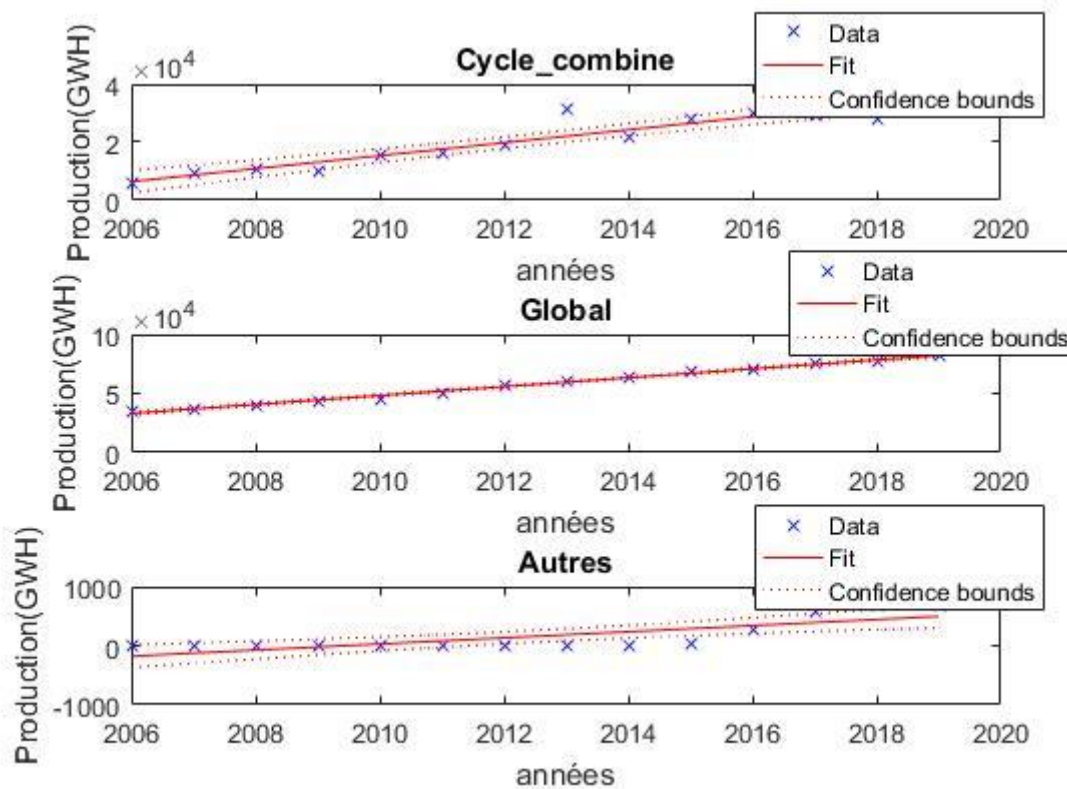


FIGURE 8 REGRESSION LINEAIRE POUR LES TROIS TYPES DE PRODUCTIONS

Modèle mathématique linéaire pour la production des cycles combinés :

$$f(x) = p_1 * x + p_2$$

Coefficients :

$$p_1 = 2232 \quad p_2 = -4.47e^{+06}$$

Goodness of fit:

$$\text{SSE: } 1.44e^{+08}$$

Coefficient de corrélation: 0.8872

$$R_{ajuster}^2: 0.8778$$

### Interprétation

Le graphe représente la production d'électricité en cycle combiné entre 2006 et 2020.

On remarque que le graphe durant ces années ne fait qu'augmenter vu la droite obtenue.



Donc avec le modèle linéaire on peut prédire que ce sera le cas pour les futures années.

Modèle mathématique linéaire pour la production globale :

$$f(x) = p_1 * x + p_2$$

Coefficients :

$$p_1 = 3.446e^{+04}$$

$$p_2 = -6.919e^{+07}$$

Goodness of fit:

$$\text{SSE: } 5.924e^{+11}$$

Coefficient de corrélation : 0.3132

$$R_{ajuster}^2: 0.256$$

**Conclusion :**

Modèle rejetée.

Modèle mathématique linéaire pour la production solaire :

$$f(x) = p_1 * x + p_2$$

Coefficients :

$$p_1 = 728.8$$

$$p_2 = -1.455e^{+06}$$

Goodness of fit:

$$\text{SSE: } 2.645e^{+07}$$

Coefficient de corrélation: 0.8204

$$R_{ajuster}^2: 0.8055$$

**Interprétation**

Le graphe représente la production d'électricité dans le secteur solaire entre 2006 et 2020.

On remarque que le graphe durant ces années ne fait qu'augmenter vu la droite obtenue.

Donc avec le modèle linéaire on peut prédire que ce sera le cas pour les futures années.

## Régression 4eme degré

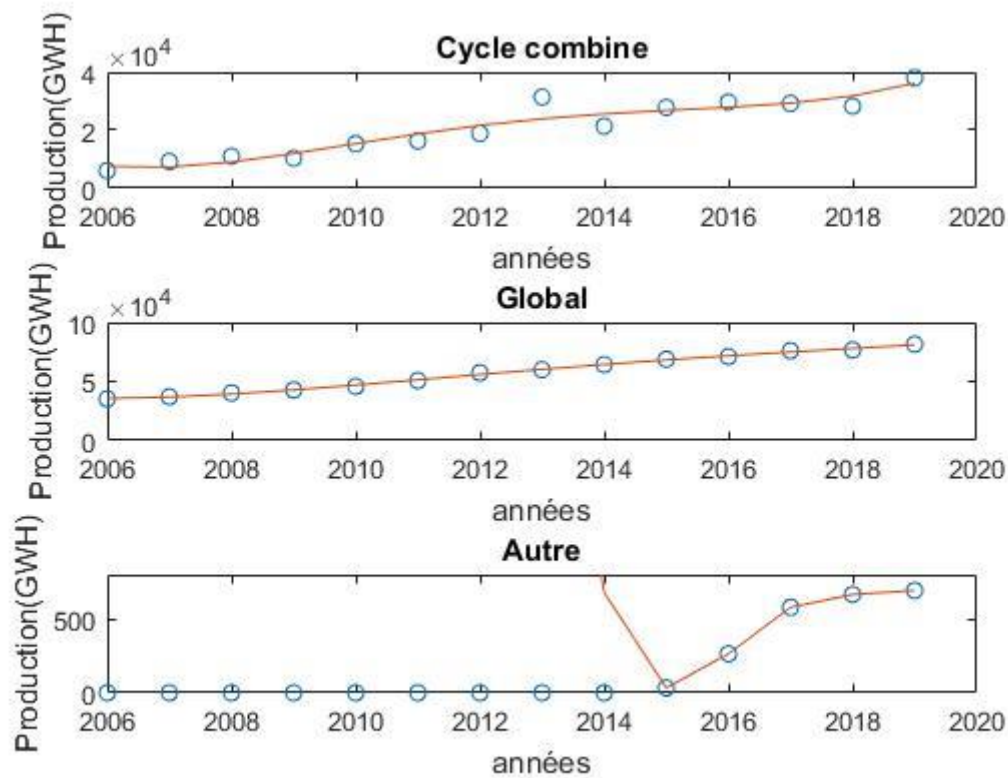


FIGURE 9 REGRESSION 4EME DEGREE POUR LES TROIS TYPES DE PRODUCTIONS.

Modèle mathématique 4eme degré cycle combiné :

$$f(x) = p_1 * x^4 + p_2 * x^3 + p_3 * x^2 + p_4 * x + p_5$$

Coefficients :

$$p_1 = 7.194$$

$$p_2 = -5.791e^{+04}$$

$$p_3 = 1.748e^{+08}$$

$$p_4 = -2.346e^{+11}$$

$$p_5 = 1.18e^{+14}$$

Goodness of fit:

$$\text{SSE: } 1.202e^{+08}$$

Coefficient de corrélation: 0.9095

$$R_{ajuster}^2: 0.8641$$

**Interprétation**

Le graphe représente la production d'électricité dans le secteur du cycle combiné entre 2006 et 2020.

On remarque que le graphe durant ces années ne fait qu'augmenter vu la droite obtenue, ce qui signifie que la demande de production sera présente d'ici les années à venir.

Modèle mathématique 4eme degré pour le global :

$$f(x) = p_1 * x^4 + p_2 * x^3 + p_3 * x^2 + p_4 * x + p_5$$

Coefficients :

$$p_1 = 197.8$$

$$p_2 = -1.591e^{+06}$$

$$p_3 = 4.802e^{+09}$$

$$p_4 = -6.439e^{+12}$$

$$p_5 = 3.238e^{+15}$$

Goodness of fit:

$$\text{SSE: } 4.671e^{+11}$$

Coefficient de corrélation: 0.4584

$$R_{ajuster}^2: 0.2177$$

**Conclusion :**

Modèle rejetée.

Modèle mathématique 4eme degré pour le solaire :

$$f(x) = p_1 * x^4 + p_2 * x^3 + p_3 * x^2 + p_4 * x + p_5$$

Coefficients :

$$p_1 = -0.2285$$

$$p_2 = 1840$$

$$p_3 = -5.558e^{+06}$$

$$p_4 = 7.459e^{+09}$$

$$p_5 = -3.754e^{+12}$$

Goodness of fit:

$$\text{SSE: } 5.193e^{+04}$$

Coefficient de corrélation: 0.9466

$$R_{ajuster}^2: 0.9229$$

## Interprétation

Le graphe représente la production d'électricité dans le secteur du global entre 2006 et 2020.

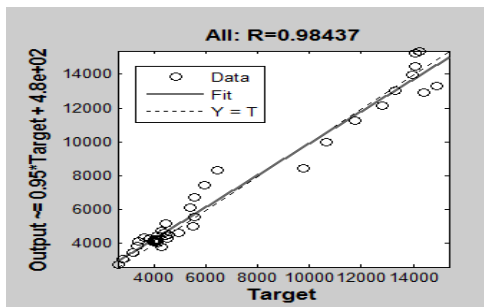
On remarque la production de l'énergie électrique de secteur solaire a commencé en 2015 et ça continuera a augmenté durant les années à venir.

## 4.2 ANALYSE ET INTERPRETATION PAR LA METHODE DES RESEAUX DE NEURONES

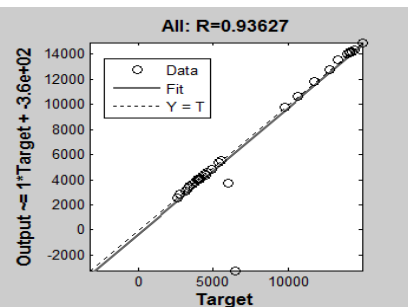
### 4.2.1 ANALYSE DE LA CONSOMMATION DANS LE SECTEUR DE TRANSPORT

La consommation des produits pétroliers pour le secteur de transport par rapport à l'année :

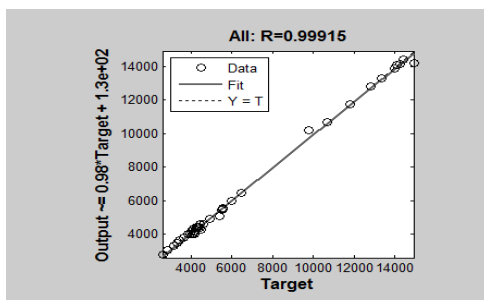
Pour N=5



Pour N=10



Pour N=15



Pour N=20

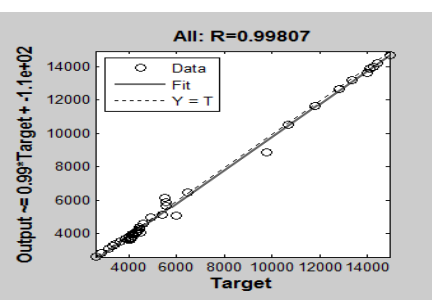
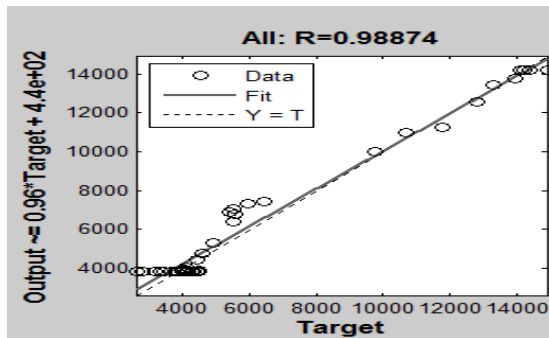


FIGURE 10 VARIATION DU NOMBRE DE NEURONES POUR LE SECTEUR DES TRANSPORTS – INPUT ANNEE

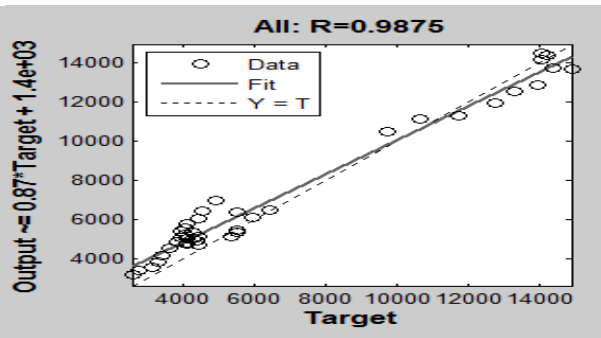
On remarque que le nombre de neurones idéal pour avoir un meilleur rendement est 15 neurones.

La consommation des produits pétroliers pour le secteur de transport par rapport à l'année et le PIB :

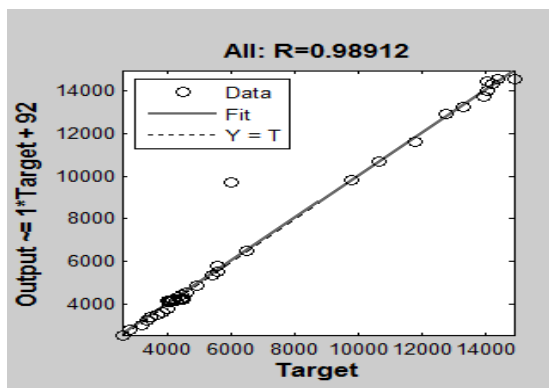
Pour N =5



Pour N =10



Pour N =15



Pour N =20

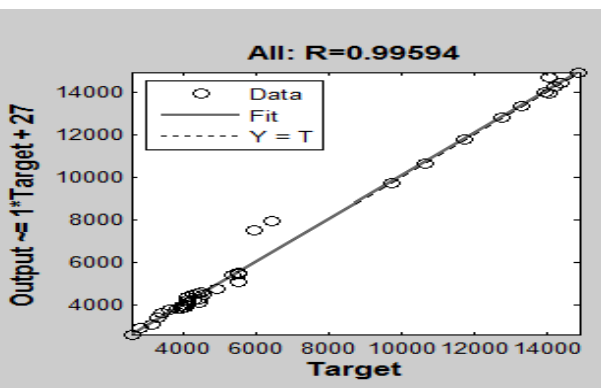
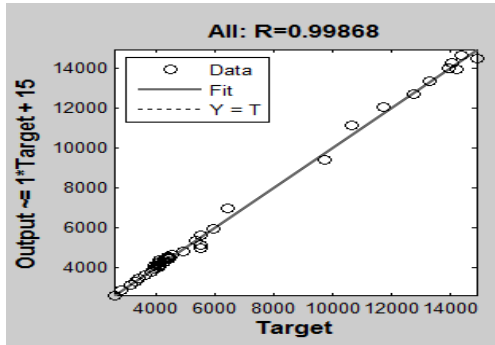


FIGURE 11 VARIATION DU NOMBRE DE NEURONES POUR LE SECTEUR DES TRANSPORTS – INPUT ANNEE ET PIB

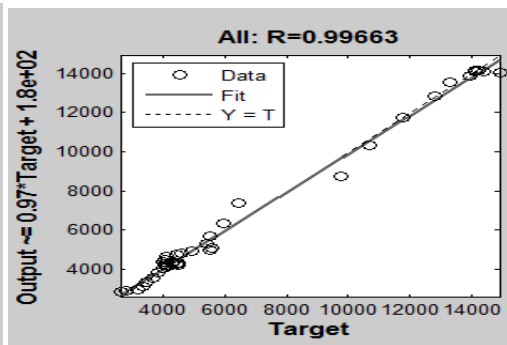
On remarque que le nombre de neurones idéal pour avoir un meilleur rendement est 20 neurones.

La consommation des produits pétroliers pour le secteur de transport par rapport à l'année et la population :

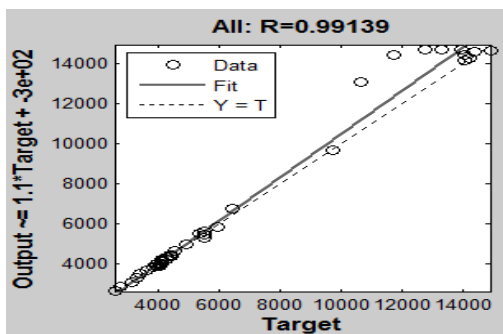
Pour N =5



Pour N =10



Pour N =15



Pour N=20

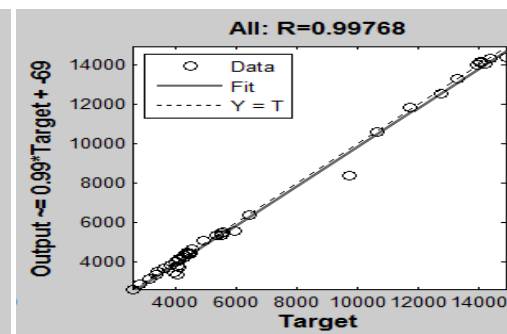
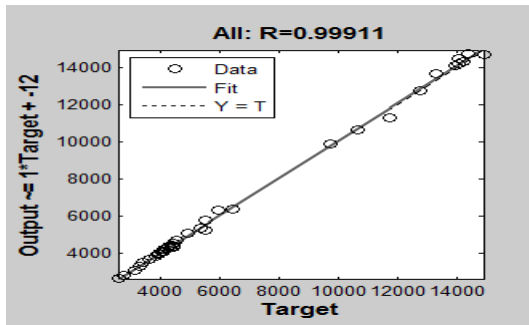


FIGURE 12 VARIATION DU NOMBRE DE NEURONES POUR LE SECTEUR DES TRANSPORTS – INPUT ANNEE ET POPULATION

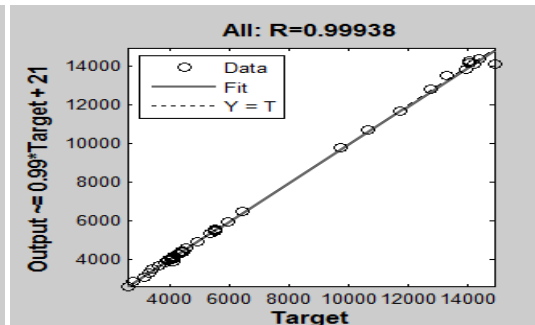
On remarque que le nombre de neurones idéal pour avoir un meilleur rendement est 5 neurones.

La consommation des produits pétroliers pour le secteur de transport par rapport à l'année et le PIB, la population :

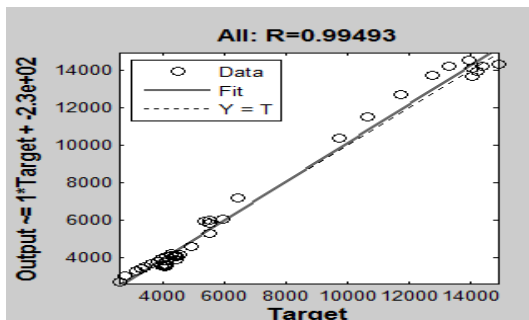
Pour N=5



Pour N =10



Pour N =15



Pour N =20

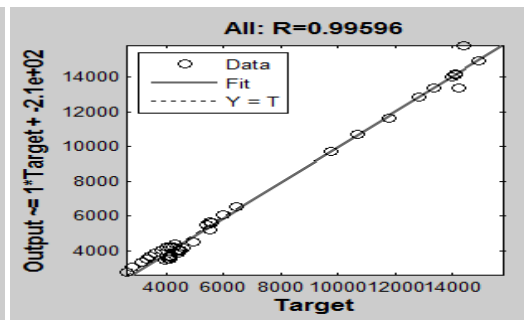
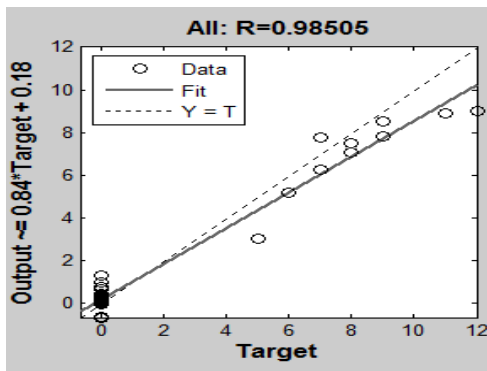


FIGURE 13 VARIATION DU NOMBRE DE NEURONES POUR LE SECTEUR DES TRANSPORTS – INPUT ANNEE, PIB ET POPULATION

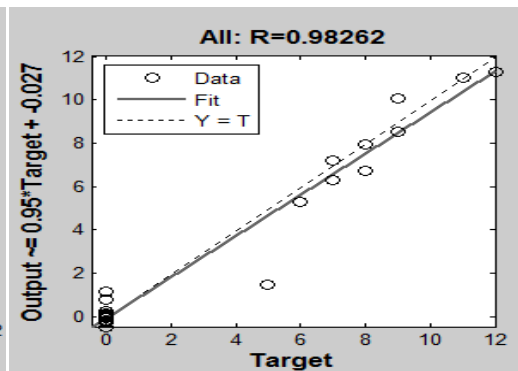
On remarque que le nombre de neurones idéal pour avoir un meilleur rendement est 10 neurones.

La consommation du gaz naturel pour le secteur de transport par rapport à l'année :

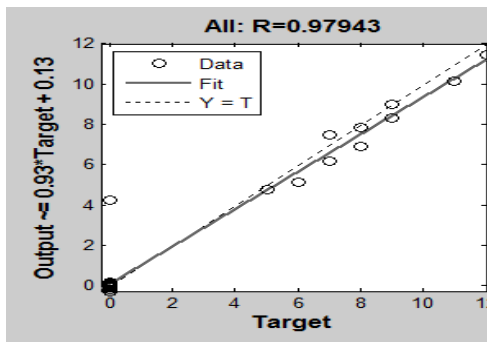
Pour N=5



Pour N=10



Pour N=15



pour N=20

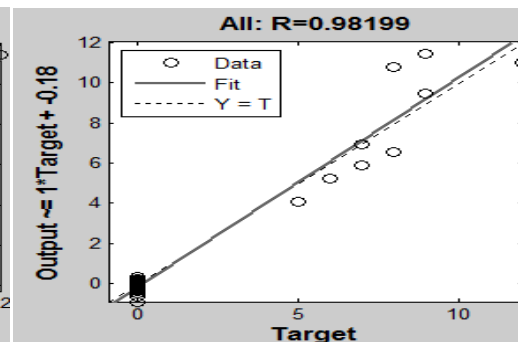


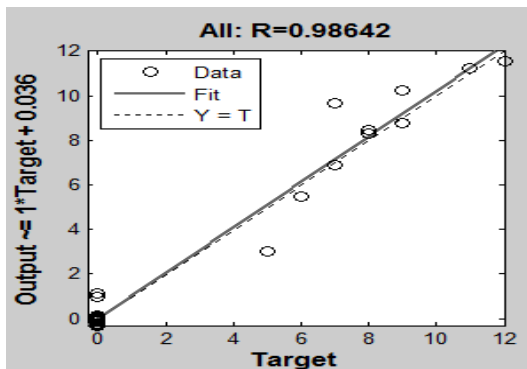
FIGURE 14 VARIATION DU NOMBRE DE NEURONES POUR LE GAZ NATUREL DANS LES TRANSPORTS – INPUT ANNEE

ON REMARQUE QUE LE NOMBRE DE NEURONES IDEAL POUR AVOIR UN MEILLEUR RENDEMENT EST 05 NEURONES.

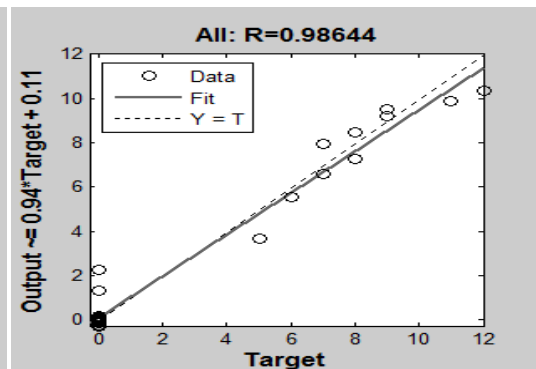


La consommation du gaz naturel pour le secteur de transport par rapport à l'année et le PIB :

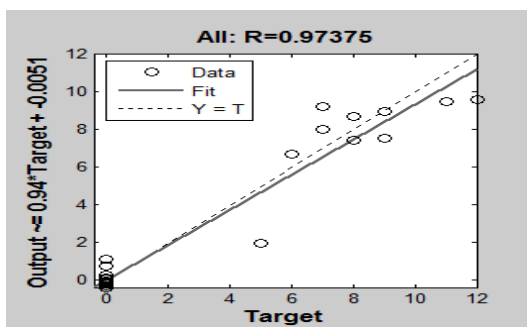
Pour N =5



Pour N=10



Pour N =15



Pour N =20

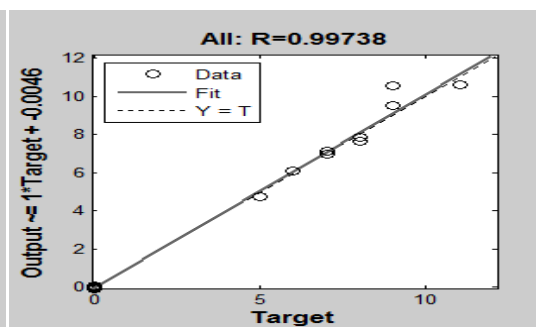
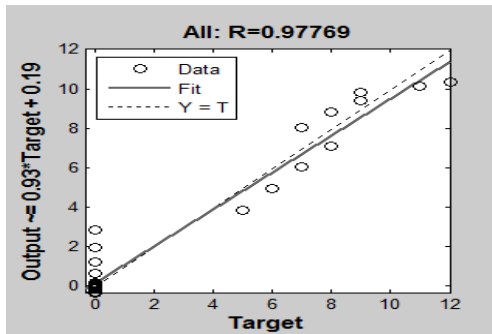


FIGURE 15 VARIATION DU NOMBRE DE NEURONES POUR LE GAZ NATUREL DANS LES TRANSPORTS – INPUT ANNEE ET PIB

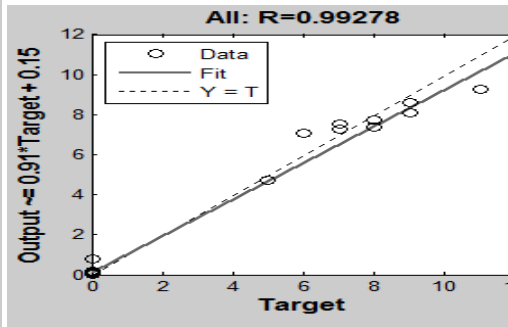
On remarque que le nombre de neurones idéal pour avoir un meilleur rendement est 20 neurones

La consommation du gaz naturel pour le secteur de transport par rapport à l'année et la population :

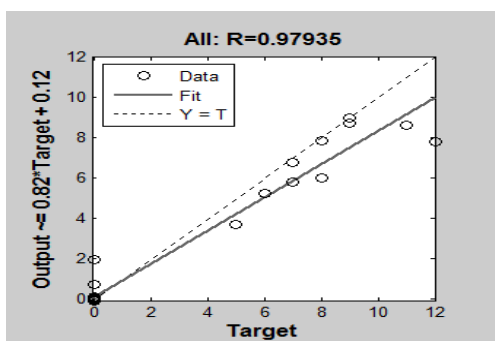
Pour N=5



Pour N =10



Pour N =15



Pour N =20

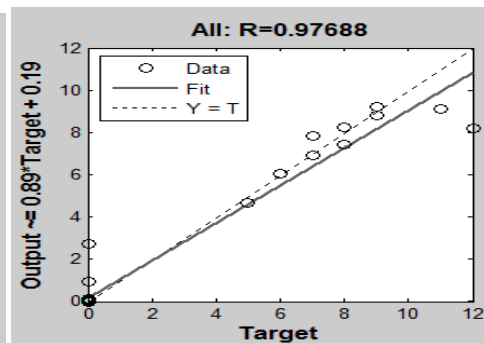
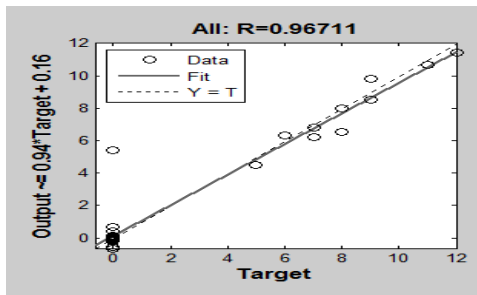


FIGURE 16 VARIATION DU NOMBRE DE NEURONES POUR LE GAZ NATUREL DANS LES TRANSPORTS – INPUT ANNEE ET POPULATION

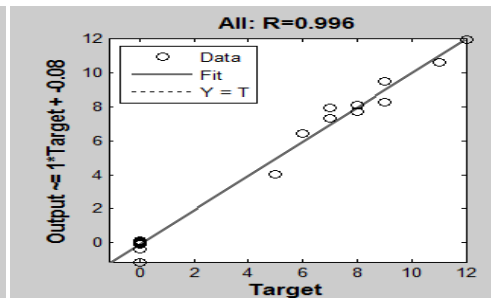
On remarque que le nombre de neurones idéal pour avoir un meilleur rendement est 10 neurones

La consommation du gaz naturel pour le secteur de transport2 par rapport à l'année et le PIB, la population :

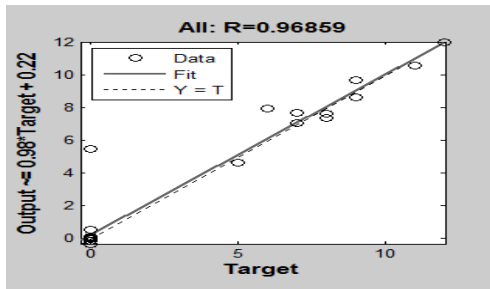
Pour N =5



Pour N =10



Pour N =15



Pour N =20

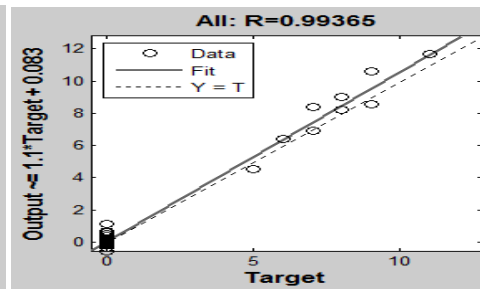
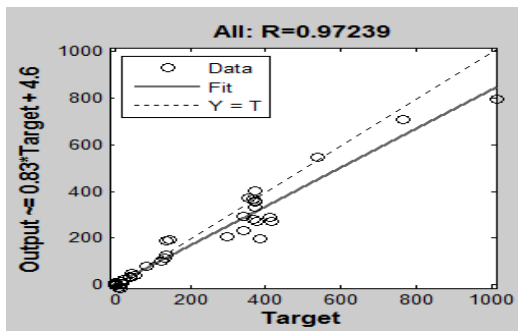


FIGURE 17 VARIATION DU NOMBRE DE NEURONES POUR LE GAZ NATUREL DANS LES TRANSPORTS – INPUT ANNEE, PIB ET POPULATION

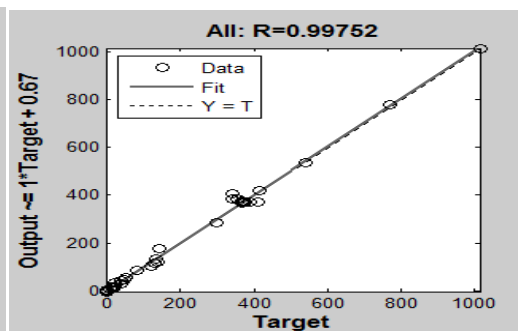
On remarque que le nombre de neurones idéal pour avoir un meilleur rendement est 10 neurones

La consommation du GPL pour le secteur de transport par rapport à l'année :

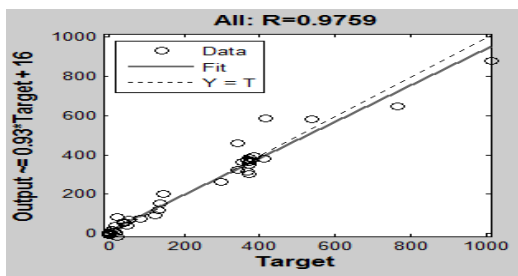
Pour N=5



Pour N=10



Pour N=15



Pour N=20

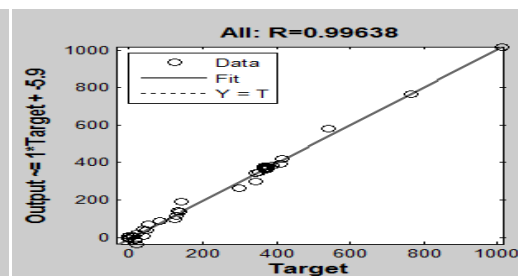
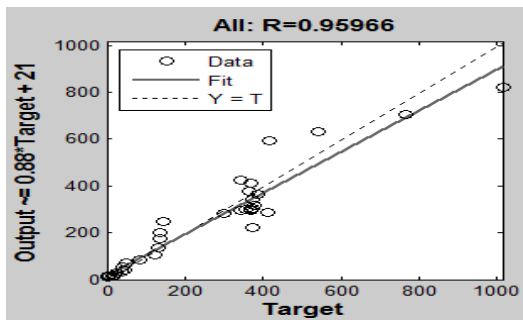


FIGURE 18 VARIATION DU NOMBRE DE NEURONES POUR LE GPL DANS LES TRANSPORTS – INPUT ANNEE

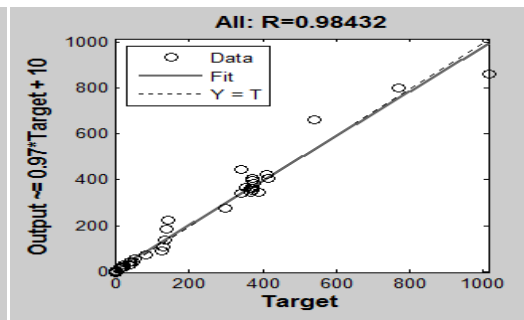
On remarque que le nombre de neurones idéal pour avoir un meilleur rendement est 10 neurones.

La consommation du GPL pour le secteur de transport par rapport à l'année et le PIB :

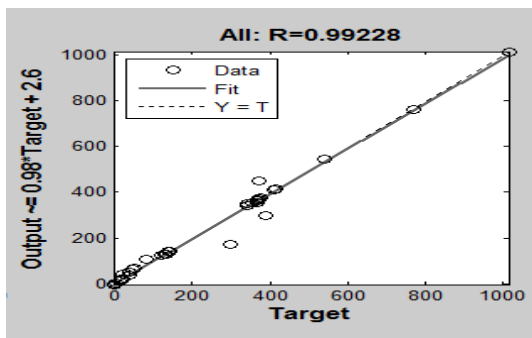
Pour N=5



Pour N =10



Pour N =15



Pour N =20

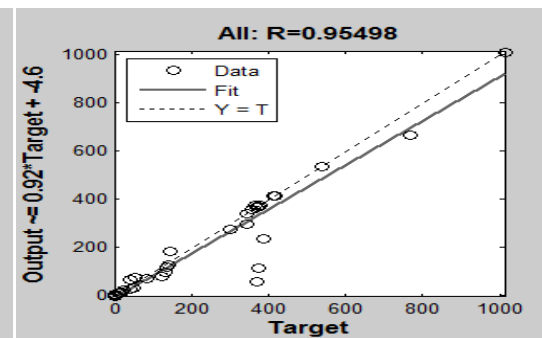
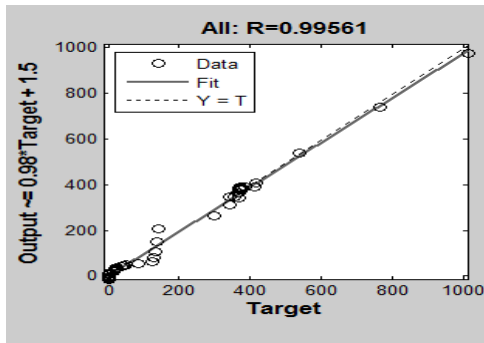


FIGURE 19 VARIATION DU NOMBRE DE NEURONES POUR LE GPL DANS LES TRANSPORTS – INPUT ANNEE ET PIB

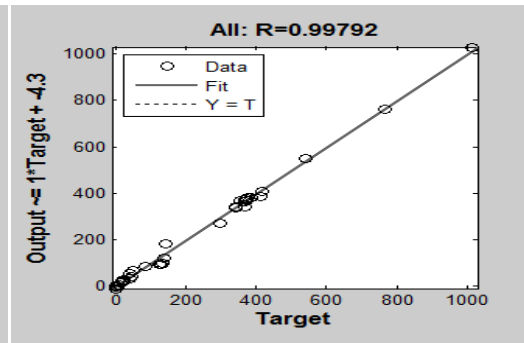
On remarque que le nombre de neurones idéal pour avoir un meilleur rendement est 15 neurones

La consommation du GPL pour le secteur de transport par rapport à l'année et la population :

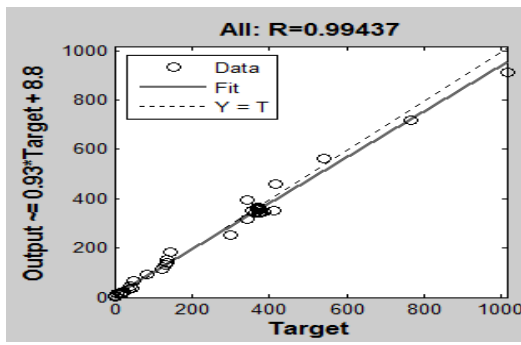
Pour N=5



Pour N=10



Pour N=15



Pour N=20

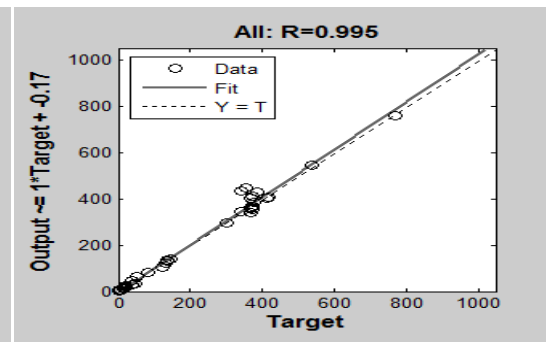


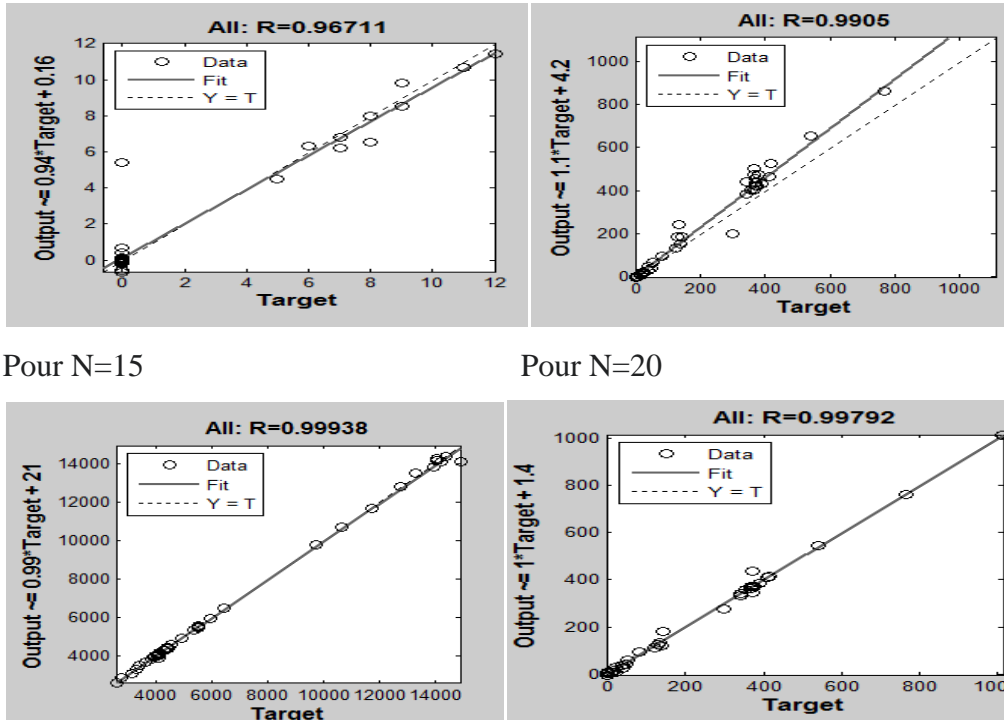
FIGURE 20 VARIATION DU NOMBRE DE NEURONES POUR LE GPL DANS LES TRANSPORTS – INPUT ANNEE ET POPULATION

On remarque que le nombre de neurones idéal pour avoir un meilleur rendement est 20 neurones

La consommation du GPL pour le secteur de transport par rapport à l'année et le PIB, la population :

Pour N=5

Pour N=10



**FIGURE 21** VARIATION DU NOMBRE DE NEURONES POUR LE GPL DANS LES TRANSPORTS – INPUT ANNEE, PIB ET POPULATION

On remarque que le nombre de neurones idéal pour avoir un meilleur rendement est 15 neurones

### Conclusion

Dans cette étude, au début on a travaillé sur une variable « l'année », puis on a ajouté une deuxième qui est « le PIB », ensuite on ajoute une troisième variables « la population ».

On remarque que le résultat très proche avec les trois variables qu'avec le résultat trouver pour une seule variable qui est l'année.

Donc l'année est le facteur le plus influent.

#### 4.2.2 ANALYSE DE LA PRODUCTION D'ELECTRICITE DANS LES 03 SECTEURS

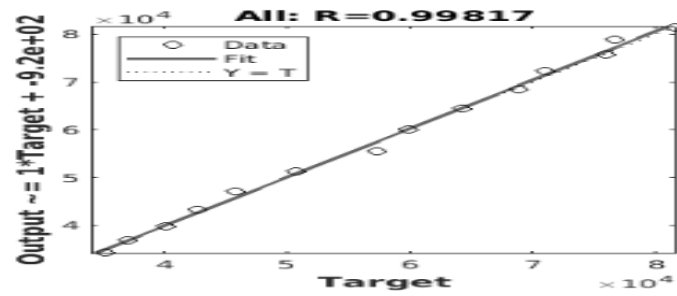


FIGURE 22. COEFFICIENT DE REGRESSION DES RESEAUX DE NEURONES POUR LA PRODUCTION GLOBALE D'ELECTRICITE

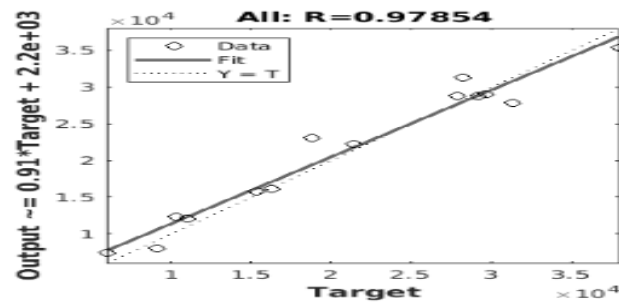


FIGURE 23. COEFFICIENT DE REGRESSION DES RESEAUX DE NEURONES POUR LA PRODUCTION DES CYCLES COMBINES D'ELECTRICITE

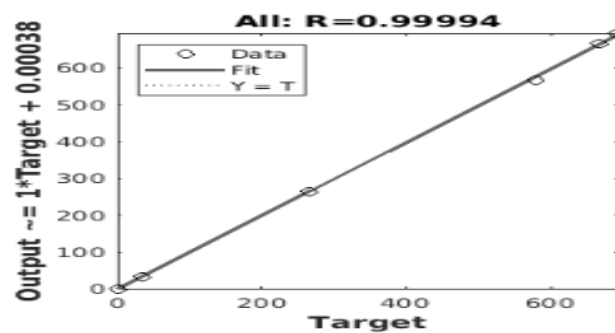


FIGURE 24. COEFFICIENT DE REGRESSION DES RESEAUX DE NEURONES POUR LA PRODUCTION SOLAIRE D'ELECTRICITE



On remarque que le coefficient de corrélation est très proche de 01, ce qui prouve que le modèle est validé.

### 4.3 PREDICTION DE LA CONSOMMATION DU SECTEUR DE TRANSPORT

#### 4.3.1 PRODUIT PETROLIER

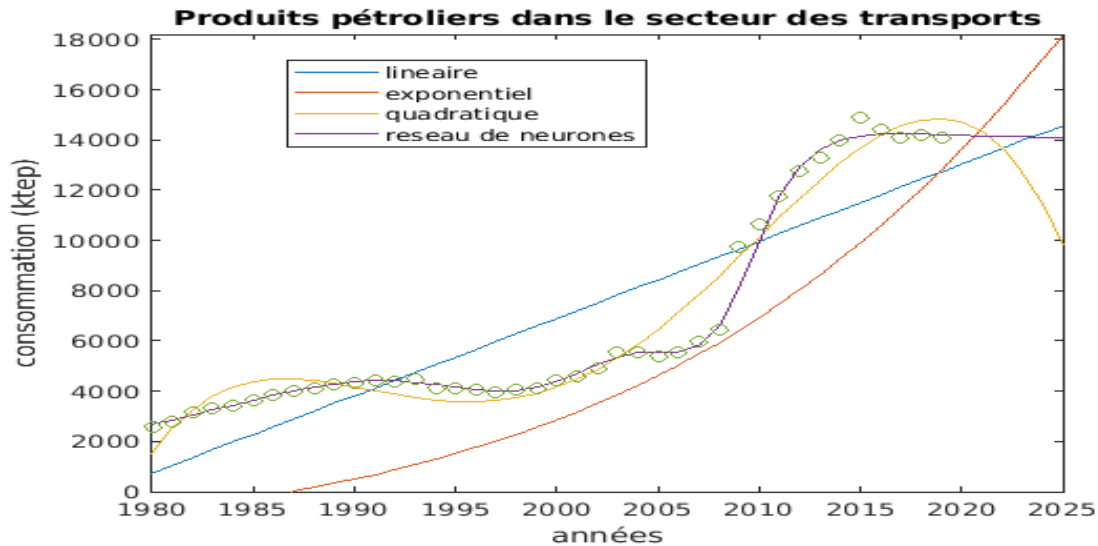


FIGURE 25 : PREDICTION DE LA CONSOMMATION DES PRODUITS PETROLIER DANS LE SECTEUR DE TRANSPORT DURANT LES ANNEES A VENIR.

#### Interprétation

Le graphe représente la prédiction de la consommation des produits pétrolier dans le secteur de transport durant les années à venir précisément jusqu'à 2030. Pour prédire la consommation nous avons utilisé plusieurs méthodes de calculs qui sont la méthode de régression « linéaire, exponentiel et 4eme degré » et la méthode des réseaux de neurones.

Nous pouvons remarquer nos données prends la même direction que les réseaux de neurones

Ce qui laisse à comprendre dans ce cas, la meilleure méthode pour prédire est la méthode des réseaux de neurones.

Durant les années 1980 à 2020, la demande de consommation des produits pétroliers n'arrête pas de croitre en suivant l'évolution de la courbe des réseaux de neurones, on peut conclure que d'ici 2030, ça continuera à augmenter, ce qui signifie que les produits pétroliers seront à la demande de plusieurs secteurs y compris le transport.

### 4.3.2 GAZ NATUREL

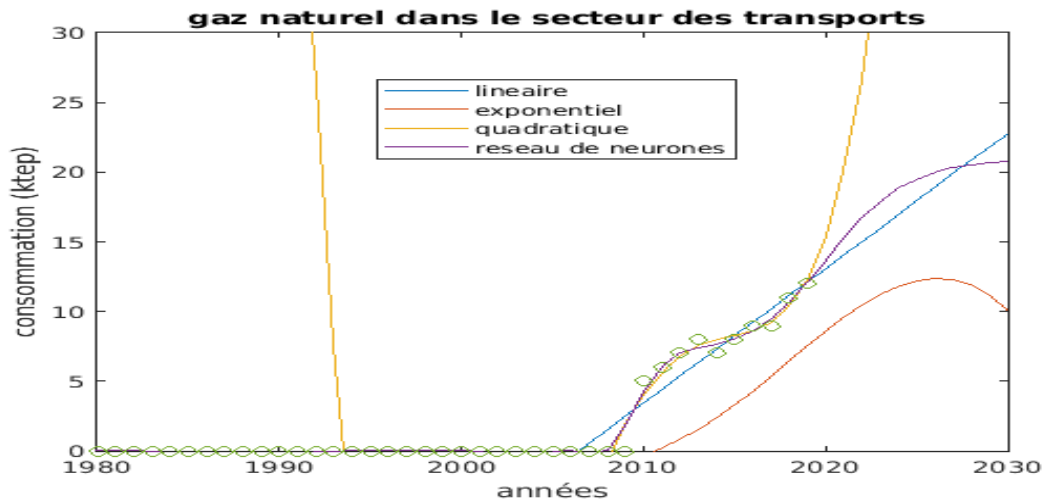


FIGURE 26 : PREDICTION DE LA CONSOMMATION DU GAZ NATUREL DANS LE SECTEUR DE TRANSPORT DURANT LES ANNEES A VENIR

#### Interprétation

Le graphe représente la prédiction de la consommation du gaz naturel dans le secteur de transport durant les années à venir précisément jusqu'à 2030. Pour prédire la consommation nous avons utilisé plusieurs méthodes de calculs qui sont la méthode de régression « linéaire, exponentiel et 4<sup>ème</sup> degrés » et la méthode des réseaux de neurones.

On remarque que les données prennent la même direction que les réseaux de neurones. D'où alors, cette dernière est la meilleure méthode pour prédire.

La consommation du gaz naturel à commencer en 2010, avec de petite quantité qui ne dépasse pas les 15tep, et ça continuera à augmenter jusqu'à 2030.

### 4.3.3 GPL

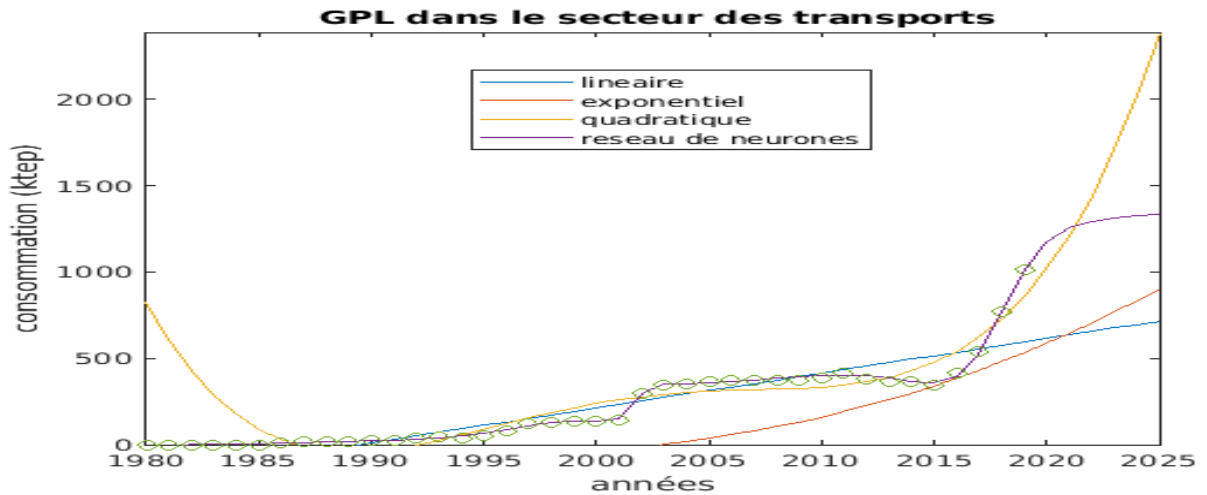


FIGURE 27 : PREDICTION DE LA CONSOMMATION DU GPL DANS LE SECTEUR DE TRANSPORT DURANT LES ANNEES A VENIR

#### Interprétation

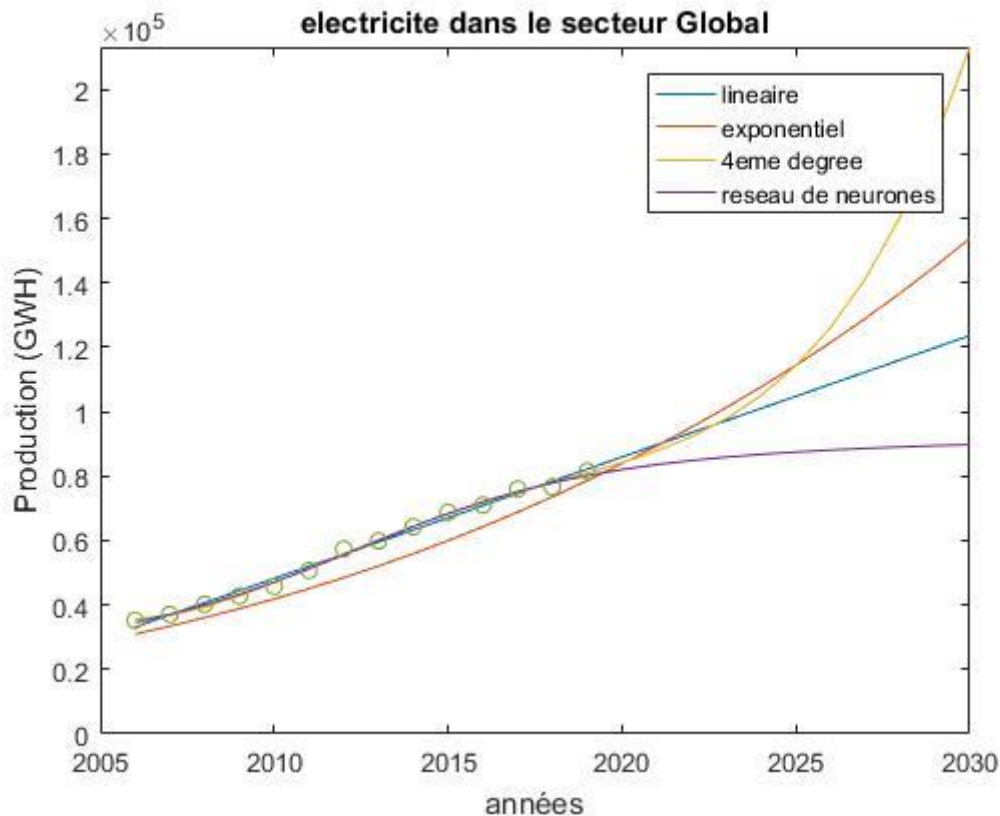
Le graphe représente la prédiction de la consommation du GPL dans le secteur de transport durant les années à venir précisément jusqu'à 2030, dans ce graphe pour prédire la consommation nous avons utilisé plusieurs méthodes de calculs qui sont la méthode de régression « linéaire, exponentiel et 4eme degré » et la méthode des réseaux de neurones.

On remarque que les données prennent la même direction que les réseaux de neurones, ce qui laisse à comprendre dans ce cas-là la meilleure méthode pour prédire et la méthode des réseaux de neurones.

C'est durant l'année 1985 que le GPL a fait apparition dans notre société, surtout dans le secteur de transport. En visualisant le graphe, on remarque la courbe croit surtout de 2015 jusqu'aujourd'hui, et prend l'échelon de la méthode des réseaux de neurones. On peut très facilement comprendre qu'elle continuera même jusqu'à 2030.

## 4.4 PREDICTION DE LA PRODUCTION DE L'ELECTRICITE

### 4.4.1 PRODUCTION GLOBALE



**FIGURE 28: PREDICTION DE LA PRODUCTION D'ELECTRICITE GLOBALE DURANT LES ANNEES A VENIR**

#### Interprétation

Le graphe représente la prédiction de la consommation d'électricité dans le secteur global durant les années à venir précisément jusqu'à 2030, dans ce graphe pour prédire la consommation nous avons utilisé plusieurs méthodes de calculs qui sont la méthode de régression « linéaire, exponentiel et 4eme degrés » et la méthode des réseaux de neurones.

On remarque que les données prennent la même direction que les réseaux de neurones, ce qui laisse à comprendre dans ce cas-là la meilleure méthode pour prédire et la méthode des réseaux de neurones.

C'est qu'en 2006 que le secteur global a commencé à produire de l'électricité jusqu'aujourd'hui, et on remarque qu'elle va augmenter d'ici 2030

#### 4.4.2 PRODUCTION SOLAIRE

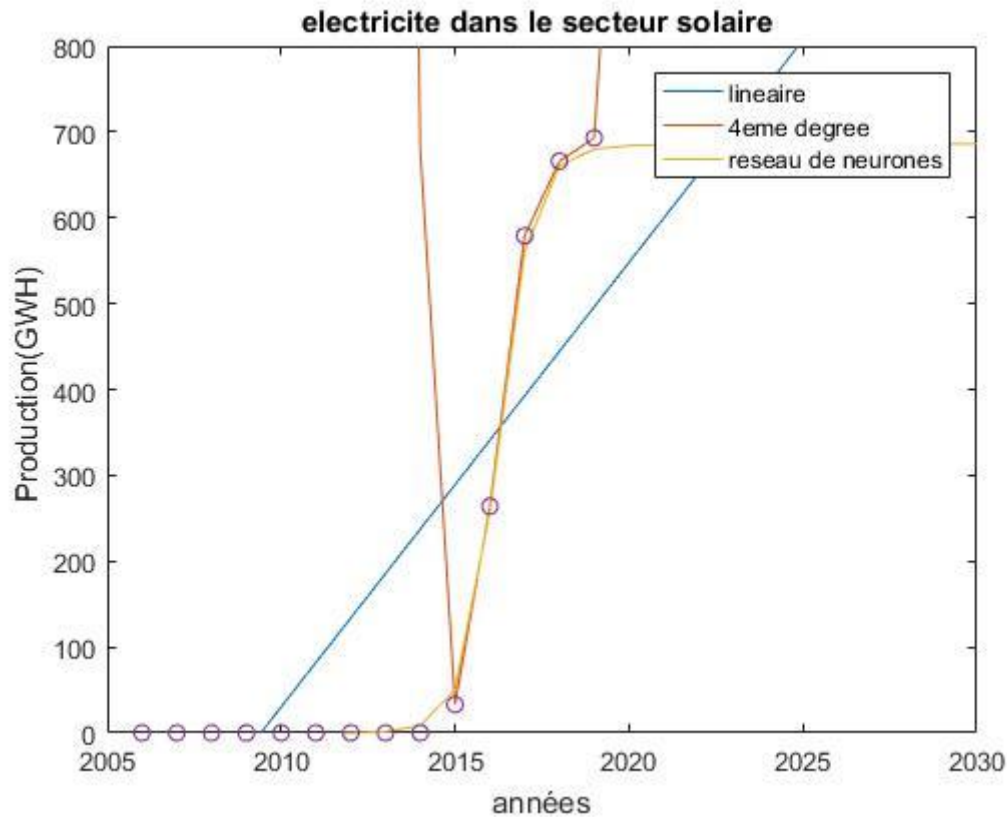


FIGURE 29 : PREDICTION DE LA PRODUCTION D'ÉLECTRICITÉ SOLAIRE DURANT LES ANNÉES À VENIR

#### Interprétation

Le graphe représente la prédiction de la consommation d'électricité dans le secteur solaire durant les années à venir précisément jusqu'à 2030, dans ce graphe pour prédire la consommation nous avons utilisé plusieurs méthodes de calculs qui sont la méthode de régression « linéaire, exponentiel et 4eme degrés » et la méthode des réseaux de neurones.

On remarque que les données prennent la même direction que les réseaux de neurones, et la régression 4eme degrés, ce qui laisse à comprendre dans ce cas-là la meilleure méthode pour prédire est l'une des deux mais le meilleur reste les réseaux de neurones.

C'est qu'après 2015 que le secteur solaire a commencé à s'intéresser à la production de l'électricité, avec de très petite quantité. Vu l'évolution de la courbe, ça va continuer à augmenter modérément durant les futures années.

### 4.4.3 PRODUCTION DES CYCLES COMBINES

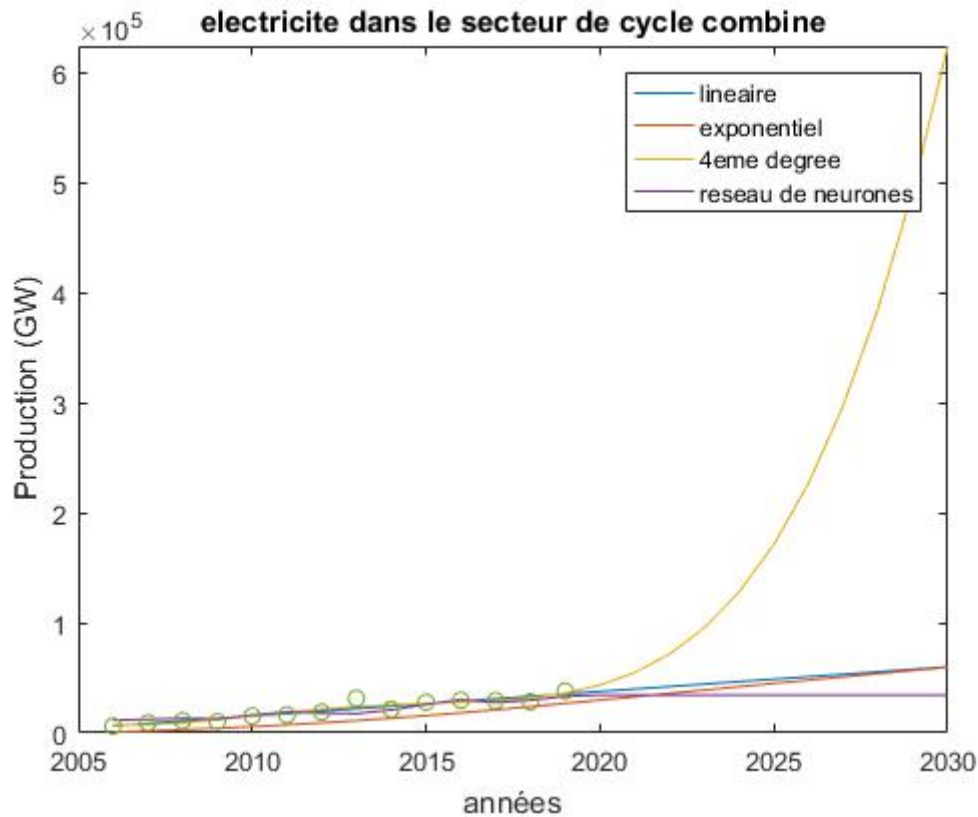


FIGURE 30 : PREDICTION DE LA PRODUCTION D'ÉLECTRICITÉ PAR LES CYCLES COMBINES DURANT LES ANNÉES À VENIR

#### Interprétation

Le graphe représente la prédiction de la consommation d'électricité dans le secteur du cycle combiné, durant les années à venir précisément jusqu'à 2030. Pour prédire la consommation nous avons utilisé plusieurs méthodes de calculs qui sont la méthode de régression « linéaire, exponentiel et 4eme degré » et la méthode des réseaux de neurones.

On remarque que les données prennent la même direction que les réseaux de neurones, ce qui laisse à comprendre dans ce cas-là la meilleure méthode pour prédire et la méthode des réseaux de neurones.

C'est qu'en 2006 que le secteur du cycle combiné a commencé à produire de l'électricité. On remarque d'après la courbe, qu'elle va augmenter probablement d'ici 2030.

## CONCLUSION

Nous avons élaboré un programme Matlab pour le calcul des régressions linéaires, 4eme degré, exponentielles ainsi que les réseaux de neurones artificiels. Les données d'entrée sont les consommations de produits pétroliers, gaz naturel et GPL dans le secteur des transports. Pour le cas de l'électricité, nous nous sommes intéressés à la production électrique globale, et les cas particuliers des cycles combinés ainsi que l'énergie solaire (Il faut savoir que la construction des centrales classiques a été abandonné pour ne laisser que des programmes de construction de centrales à cycles combinées et solaire).

Les analyses des régressions montrent que globalement les réseaux de neurones ont les meilleures performances, suivie du modèle de 4eme degré puis le modèle exponentiel et en dernier le modèle linéaire.

En ce qui concerne la prédiction, la tendance la plus modéré est présenté par les réseaux de neurones. Le modèle 4eme degré prédit une baisse de la consommation dans certains cas, le modèle exponentiel nous paraît trop surestimer la consommation future.

Une analyse plus approfondie permettra d'améliorer les prédictions des réseaux de neurones en utilisant plusieurs couches, les performances doivent aller jusqu'en 2050.

## Référence

- [1] **Elmoussaoui H.** thèse de magister (2004) USDB « Contribution à la méthodologie de la recherche expérimentale »
- [2] « Neural network history »  
<https://cs.stanford.edu/people/eroberts/courses/soco/projects/neural-networks/History/history1.html>
- [3] **Marc Parizeau** « RESEAUX DE NEURONES », université LAVAL, Automne 2004  
<http://seborga1.free.fr/RNF.pdf>
- [4] **Youcef djerir** (2017), « Les Réseaux de Neurones Artificiels »
- [5] **Zayegh A. Al Bassam N.** « Neural network Principles and applications ». in digital systems, V, Asadpour, Ed. IntechOpen, 2018.  
<https://www.intechopen.com/chapters/63906>
- [6] Djamel Belhaouci, « Démystifier le machine learning, Partie 2 : les réseaux de neurones artificiels. », juri'Predis le blog.  
<https://www.juripredis.com/fr/blog/id-19-demystifier-le-machine-learning-partie-2-les-reseaux-de-neurones-artificiels>
- [7] « Réseaux de neurons formels »
- [8] **Namrata Kapoor** « Loss function- when to use», medium, Nov. 20, 2020.  
<https://towardsdatascience.com/loss-functions-when-to-use-which-one-718ebad36e0>
- [9] **Musstafa** (2021) « optimizers in deep learning », Medium, mars 28, 2021.  
<https://medium.com/mlearning-ai/optimizers-in-deep-learning-7bf81fed78a0>
- [10] **Kaadoud. I.** « Architecture des réseaux de neurones : réseaux de neurones artificiels classique » (2/3)



<https://scilogs.fr/intelligence-mecanique/architecture-des-reseaux-de-neurones-reseaux-de-neurones-artificiels-classiques-2-3/>

[11] G. L. Team, GreatLearning Blog: Free Resources What Matter to shape your career! avr. 29, 2020 « Type of neural network and definition of neural network»

<https://www.mygreatlearning.com/blog/types-of-neural-networks/>

## ANNEXE

### Consommation énergétique entre 1980-2019

#### Code pour le secteur de transport

clc;

clear;

% les donnees s arretent en 2019, apres 2019 c de la projection

year = (1980:1:2019);

```
population =[19221659 19824297 20452901 21101875 21763578 22431507
23102386 23774287 24443472 25106192 25758872 26400468 27028330
27635517 28213777 28757788 29266415 29742980 30192750 30623406
31042238 31451513 31855110 32264159 32692153 33149720 33641007
34166976 34730604 35333882 35977451 36661438 37383899 38140135
38923688 39728020 40551398 41389174 42228415 43053054];
```

```
pib =[42346380831 44348672668 45207088716 48801369800 53698278906 57937868670
63692238160 66746396416 59089067187 55634414465 62048562947 45715614560
48003078389 49945599429 42543178042 41764315330 46941582519 48177612042
48187747529 48640653469 54790392746 54744712815 56760355865 67863828413
85332581189 103198223709 117027307541 134977082624 171000699877
137211035770 161207270185 200013052199 209058991952 209755003251
213810024944 165979279263 160034163871 170097014589 174910878623
171767403748];
```

% donnees concernant les produits petroliers selon chaque secteur

% consommation des produits petrolier selon les 3 secteurs

```
Industrie_ET_BTP1 = [227 258 308 340 380 454 526 589 610 514 502 548 527 527 464 519
502 481 526 558 650 689 809 879 931 976 1036 1154 1451 1659 864 949 620 592 690 739
775 710 519 808 ];
```

```
Transports1 = [2591 2806 3149 3332 3421 3629 3838 3982 4108 4267 4304 4451 4397
4474 4141 4141 4053 3957 4053 4113 4438 4570 4927 5524 5553 5375 5533 5967 6460
9764 10660 11772 12790 13314 13970 14912 14392 14095 14222 14096];
```

```
Menages_et_autres1= [1167 1134 1225 1378 1678 1862 1738 1560 1489 1451 1410 1543
1487 1486 1309 1463 1413 1355 1482 1573 1830 1941 2280 2476 2622 2750 2917 3251
3605 896 747 827 588 886 329 324 360 533 777 1249 ];
```



```
close all;
subplot(2,2,1);
m=plot(year, Transports1);
title('Produits petroliers')
subplot(2,2,2);
plot(year, Transports2);
title('Gaz naturel')
subplot(2,2,3);
plot(year, Transports3);
title('GPL')
```

```
% regression lineaire
```

```
md1=fitlm(year,Transports1);
md2=fitlm(year(29:length(year)), Transports2(29:length(Transports2)));
md3=fitlm(year(6:length(year)), Transports3(6:length(Transports3)));
```

```
close all;
subplot(2,2,1);
plot(md1)
title('Produits petroliers')
subplot(2,2,2);
plot(md2)
title('Gaz naturel')
subplot(2,2,3);
plot(md3)
title('GPL')
```

```
% regression exponentielle
```

```
x=year';
y=Transports1';
y1=Transports2';
y2=Transports3';
```

```

y3=Transports4';
me1=fit(x,y,'exp1');
me2=fit(x(29:length(year)),y1(29:length(Transports2)),'exp1');
me3=fit(x(6:length(year)), y2(6:length(Transports2)),'exp1');
me4=fit(x,y3,'exp1');

close all;
subplot(2,2,1);
plot(me1,x,y,'o')
title('Produits petroliers')
subplot(2,2,2);
plot(me2,x(29:length(year)),y1(29:length(Transports2)),'o')
title('Gaz naturel')
subplot(2,2,3);
plot(me3,x(6:length(year)), y2(6:length(Transports2)),'o')
title('GPL')

% regression 4eme degres
mq1=polyfit(x,y,4);
mq2=polyfit(x(29:length(year)),y1(29:length(Transports2)),4);
mq3=polyfit(x(6:length(year)), y2(6:length(Transports2)),4);

close all;
subplot(2,2,1);
plot(x,y,'o',x,polyval(mq1,x),'-')
title('Produitspetrolier')
subplot(2,2,2);
plot(x(29:length(year)),y1(29:length(Transports2)),'o',x(29:length(year)),polyval(mq2,x(29:le
ngth(year))),'-')
title('Gaz naturel')
subplot(2,2,3);
plot(x(6:length(year)),
y2(6:length(Transports2)),'o',x(6:length(year)),polyval(mq3,x(6:length(year))),'-')
title('GPL')

```

```

%NFT
% year - input data.
% Produits pétroliers
% 15 neurones
tx1(1,1:40)=year(1:40);
ty1=Transports1;
net1 = fitnet([50],[20],[10],'trainbr');
% Create a Fitting Network
net1 = configure(net1,tx1,ty1);
net1.trainParam.epochs=1000;
net1.trainParam.max_fail = 6;
net1.trainParam.goal=0.01;
%net1.trainParam.lr = 0.001;
%net1.trainParam.min_grad = 1e-8;
%net1.trainParam.mu_inc = 10;
%net1.trainParam.mu_max = 1e10;
% Setup Division of Data for Training, Validation, Testing
[trainInd,valInd,testInd] = dividerand(1,0.7,0.15,0.15)
% Train the Network
[net1,tr]=train(net1,tx1,ty1);
% Test the Network
y = net1(tx1);
perf = perform(net1,y,ty1);
perf = mse(net1,y,ty1,'regularization',0.01);

% Gaz naturel
% Transports4 - target data.
% 5 neurones
tx2(1,1:40)=year(1:40);
ty2=Transports2;
net2 = fitnet([50],[20],[10],'trainbr');
net2 = configure(net2,tx2,ty2);
net2.trainParam.epochs=1000;
net2.trainParam.max_fail = 6;

```

```

net2.trainParam.goal=0.01;
[trainInd,valInd,testInd] = dividerand(1,0.7,0.15,0.15)
[net2,tr]=train(net2,tx2,ty2);
y = net2(tx2);
perf = perform(net2,y,ty2);
perf = mse(net2,y,ty2,'regularization',0.01);

```

```
% GPI
```

```
% Transports3 - target data.
```

```
% 10 neurones
```

```

tx3(1,1:40)=transpose([year(1:40)']);
ty3=Transports3;
net3 = fitnet([50],[20],[10],'trainbr');
net3 = configure(net3,tx3,ty3);
net3.trainParam.epochs=1000;
net3.trainParam.max_fail = 6;
net3.trainParam.goal=0.01;
[trainInd,valInd,testInd] = dividerand(1,0.7,0.15,0.15)
[net3,tr]=train(net3,tx3,ty3);
y = net3(tx3);
perf = perform(net3,y,ty3);
perf = mse(net3,y,ty3,'regularization',0.01);

```

```
% en un graphe
```

```
hold on
```

```
plotfit(net1,tx1,ty1);
```

```
title('Produitspetrolier');
```

```
plotfit(net2,tx2,ty2);
```

```
title('gaz naturel');
```

```
plotfit(net3,tx3,ty3);
```

```
title('GPL');
```

```
hold off
```

```
hold on
```

```
plot(md1)
```

```

hold on
plot(me1)
hold on
plot(x,polyval(mq1,x),'-')
title('Produits petroliers')
hold on
plot(md2)
hold on
plot(me2,x(29:length(year)),y1(29:length(Transports2)));
hold on
plot(mq2,x(29:length(year)),y1(29:length(Transports2)),'o',x(29:length(year)),polyval(mq3,x(
29:length(year))),'-');
title('gaz naturel')
hold on
plot(md3)
hold on
plot(me3,x(6:length(year)), y2(6:length(Transports3)))
hold on
plot(x(6:length(year)),
y2(6:length(Transports3)),'o',x(6:length(year)),polyval(mq3,x(6:length(year))),'-')
title('GPL')

% prediction produits pétroliers dans le secteur des transports
ypred1 = predict(md1,[1980:2025])
ypred11 = predint(me1,[1980:2025])
ypred111 = polyval(mq1,[1980:2025])
genFunction(net1,'ANNTransport1');
pANNTransport1 = ANNTransport1([1980:2025]);
plot([1980:2050],ypred1,[1980:2050],ypred11,[1980:2050],ypred111,[1980:2050],pANNTran
sport1,year,Transports1,'o')
ylim([0 1e5])
legend('lineaire','exponentiel','4eme degrés','reseau de neurones')
title('Produits pétroliers dans le secteur des transports')

```



```

plot([1980:2025],ypred1(1:46),[1980:2025],ypred11(1:46),[1980:2025],ypred111(1:46),[1980
:2025],pANNTransport1(1:46),year,Transports1,'o')
ylim([0 inf])
legend('lineaire','exponentiel','4eme degres','reseau de neurones')
title('Produits pétroliers dans le secteur des transports')

```

```

% prediction gaz naturel dans le secteur des transports

```

```

ypred2 = predict(md2,[1980:2025])
ypred22 = predint(me2,[1980:2025])
ypred222 = polyval(mq2,[1980:2025])
genFunction(net2,'ANNTransport2');
pANNTransport2 = ANNTransport2([1980:2050]);
plot([1980:2025],ypred2(1:46),[1980:2025],ypred22(1:46),[1980:2025],ypred222(1:46),[1980
:2025],pANNTransport2(1:46),year,Transports2,'o')
ylim([0 inf])
legend('lineaire','exponentiel','4eme degres','reseau de neurones')
title(' gaz naturel dans le secteur des transports')

```

```

% prediction GPL dans le secteur des transports

```

```

ypred3 = predict(md3,[1980:2025])
ypred33 = predint(me3,[1980:2025])
ypred333 = polyval(mq3,[1980:2025])
genFunction(net3,'ANNTransport3');
pANNTransport3 = ANNTransport3([1980:2025]);
plot([1980:2025],ypred3(1:46),[1980:2025],ypred33(1:46),[1980:2025],ypred333(1:46),[1980
:2025],pANNTransport3(1:46),year,Transports3,'o')
ylim([0 inf])
legend('lineaire','exponentiel','4eme degres','reseau de neurones')
title(' GPL dans le secteur des transports')

```

## 1. Production de l'électricité dans les 03 secteurs

```

clc;
clear;
year = (2006:1:2019);

```

```

population =[19221659  19824297  20452901  21101875  21763578  22431507
23102386  23774287  24443472  25106192  25758872  26400468  27028330
27635517  28213777  28757788  29266415  29742980  30192750  30623406
31042238  31451513  31855110  32264159  32692153  33149720  33641007
34166976  34730604  35333882  35977451  36661438  37383899  38140135
38923688  39728020  40551398  41389174  42228415  43053054];

```

```

pib =[42346380831 44348672668 45207088716 48801369800 53698278906 57937868670
63692238160 66746396416 59089067187 55634414465 62048562947 45715614560
48003078389 49945599429 42543178042 41764315330 46941582519 48177612042
48187747529 48640653469 54790392746 54744712815 56760355865 67863828413
85332581189 103198223709 117027307541 134977082624 171000699877
137211035770 161207270185 200013052199 209058991952 209755003251
213810024944 165979279263 160034163871 170097014589 174910878623
171767403748];

```

```

Global = [35227 37061 40237 42770 45734 50722 57397 59890 64241 68798 70997 76017
76663 81526 ];

```

```

Cycle_combine = [6042 9177 11017 10318 15314 16320 18855 31410 21392 27859 29664
29204 28244 38045 ];

```

```

Solaire = [0 0 0 0 0 0 0 0 33 264 579 666 693 ];

```

```

GN = { 'years'; 'Global'; 'Cycle combine'; 'Solaire' };

```

```

close all;

```

```

subplot(3,1,1);

```

```

plot(year, Cycle_combine, 'g');

```

```

title(' Cycle combiné')

```

```

xlabel('années')

```

```

ylabel('Production (GWH)')

```

```

subplot(3,1,2);

```

```

plot(year, Global, 'b');

```

```

title('Global');

```

```

subplot(3,1,3);

```

```

plot(year, Autres, 'y');

```

```

title('Solaire');
xlabel('années')
ylabel(' Production (GWH)')

% regression lineaire
md2 = fitlm(year, Cycle_combine);
md3 = fitlm (year, Global);
md4 = fitlm(year, Autres);
close all
subplot(3,1,1)
plot(md2)
title('Cycle combine')
xlabel('années')
ylabel(' Production (GWH)')
subplot(3,1,2)
plot(md3)
title('Global')
xlabel('années')
ylabel(' Production (GWH)')
subplot(3,1,3)
plot(md4)
title('Solaire')

% Regression au 4eme degree
mq2=polyfit(x,y1,4);
mq3=polyfit(x,y2,4);
mq4=polyfit(x(10:length(year)), y3(10:length(Autres)),4);
close all;
subplot(3,1,1);
plot(x,y1,'o',x,polyval(mq2,x),'-')
title('Cycle combine')
xlabel('années')
ylabel(' Production (GWH)')
subplot(3,1,2);

```

```

plot(x, y2,'o',x,polyval(mq3,x),'-')
title('Global')
xlabel('années')
ylabel(' Production (GWH)')
subplot(3,1,3);
plot(x, y3,'o',x,polyval(mq4,x),'-');
ylim([0 800])
title('Solaire')
xlabel('années')
ylabel(' Production (GWH)')

% reseaux de neurones
% year - input data.
% Transport
% 50 - 20 - 10 neurones
tx2(1,1:14)=year(1:14);
ty2=Cycle_combine;
net2 = fitnet([50,20,10], 'trainbr');
% Create a Fitting Network
net2= configure(net2,tx2,ty2);
net2.trainParam.epochs=1000;
net2.trainParam.max_fail = 100;
net2.trainParam.goal=0.01;
[trainInd,valInd,testInd] = dividerand(1,0.7,0.15,0.15)
% Train the Network
[net2,tr]=train(net2,tx2,ty2);
% Test the Network
y = net2(tx2)
perf = perform(net2,y,ty2)
perf = mse(net2,y,ty2,'regularization',0.01)

% 50 - 20 - 10 neurones
tx3(1,1:14)=year(1:14);
ty3=Global;

```

```

net3 = fitnet([50,20,10], 'trainbr');
net3 = configure(net2,tx2,ty2);
net3.trainParam.epochs=1000;
net3.trainParam.max_fail = 100;
net3.trainParam.goal=0.01;
[trainInd,valInd,testInd] = dividerand(1,0.7,0.15,0.15)
[net3,tr]=train(net3,tx3,ty3);
y = net3(tx3);
perf = perform(net3,y3,ty3);
perf = mse(net3,y3,ty3, 'regularization',0.01);

```

**% 50 - 20 - 10 neurones**

```

tx4(1,1:14)=year;
ty4=Autres;
net4 = fitnet([50,20,10], 'trainbr');
net4 = configure(net4,tx4,ty4);
net4.trainParam.epochs=1000;
net4.trainParam.max_fail = 100;
net4.trainParam.goal=0.01;
[trainInd,valInd,testInd] = dividerand(1,0.7,0.15,0.15)
[net4,tr]=train(net4,tx4,ty4);
y = net4(tx4);
perf = perform(net4,y,ty4);
perf = mse(net4,y,ty4, 'regularization',0.01);

```

**% prediction electricite dans cycle combine**

```

ypred2 = predict(md2,[2006:2030])
ypred22=predint(me2,[2006:2030])
ypred222= polyval(mq2,[2006:2030])
genFunction(net2, 'ANNTransport2');
pANNTransport2 = ANNTransport2([2006:2030]);
plot([2006:2030],ypred2(1:25),[2006:2030],ypred22(1:25),[2006:2030],ypred222(1:25),[2006
:2030],pANNTransport2(1:25),year,Cycle_combine, 'o')
ylim([0 inf])

```

```

legend('lineaire','exponentiel','4eme degres','reseau de neurones')
title('electricite dans le secteur de cycle combine')
xlabel('années')
ylabel(' Production (GWH)')
% predictionelectricite dans le secteur global
ypred3= predict(md3,[2006:2030])
ypred33 = predint(me3,[2006:2030])
ypred333 = polyval(mq3,[2006:2030])
genFunction(net3,'ANNTransport3');
pANNTransport3 = ANNTransport3([2006:2030]);
plot([2006:2030],ypred3(1:25),[2006:2030],ypred33(1:25),[2006:2030],ypred333(1:25),[2006
:2030],pANNTransport4(1:25),year,Global,'o')
ylim([0 inf])
legend('lineaire','exponentiel','4eme degres','reseau de neurones')
title('electricite dans le secteur Global')
xlabel('années')
ylabel(' Production (GWH)')

% prediction electricite dans le secteur autres
ypred4= predict(md4,[2006:2030])
ypred444 = polyval(mq4,[2006:2030])
genFunction(net4,'ANNTransport4');
pANNTransport4 = ANNTransport4([2006:2030]);
plot([2006:2030],ypred4(1:25),[2006:2030],ypred444(1:25),[2006:2030],pANNTransport4(1:
25),year,Autres,'o')
ylim([0 800])
legend('lineaire','4eme degres','reseau de neurones')
title('electricite dans le secteur solaire ')
xlabel('années')
ylabel(' Production (GWH)')

```