

République Algérienne Démocratique et Populaire

Ministère de l'Enseignement Supérieur et de la Recherche Scientifique

ECOLE NATIONALE SUPERIEURE DES MINES ET METALLURGIE

ENSMM-Annaba-



DEPARTEMENT GENIE MINIER

Mémoire de fin d'études

MASTER

**Analyse des mesures de tomographie électrique  
ERT par inversion conventionnelle et à l'aide  
du réseau de neurones artificiel RNA**

**Présenté par :** Melle Nesrine YOUNSI

**Encadré par :** Dr Kamel OMRACI

**Membres du jury :**

Président : Dr H. CHENITI

Examinatrice: Dr S. HABES

**Septembre 2018**

# Remerciements

Tout d'abord je remercie le bon Dieu de m'avoir donné le courage et santé pour faire aboutir ce travail.

Je tiens à remercier en premier lieu mon Directeur de mémoire, Monsieur OMRACI Kamel pour ses précieux efforts et pour son infallible implication tout au long du projet, merci énormément.

J'adresse mes sincères remerciements à Monsieur Hacène CHADLI, directeur de l'ENSMM d'Annaba, Monsieur le chef de département Abdelghani MIMOUN, à tous les enseignants du département Génie Minier et à toutes les personnes qui par leurs paroles, leurs écrits, leurs conseils et leurs critiques ont guidé mes réflexions durant mes recherches d'information.

Merci à toute personne ayant contribué de près ou de loin à la réalisation de ce projet.

# Dédicace :

*A Maman et Papa ;*

*A Nadjib et Mohamed ;*

*& à Khaled.B ;*

## **Résumé :**

La tomographie de résistivité électrique ERT (Electrical Resistivity Tomography) a été largement utilisée pour l'imagerie de la distribution et de la structure de la résistivité souterraine. Au fil des ans, de nombreux algorithmes ont été développés pour résoudre la distribution de la résistivité électrique à partir des mesures ERT. Et dans le but de surmonter les difficultés d'interprétation des mesures ERT rencontrées lors de l'inversion conventionnelle, en fournissant une imagerie 2D du profil de résistivité du sous-terrain, nous allons présenter une nouvelle approche basée sur l'apprentissage supervisé pour la formation d'un réseau de neurones artificiel.

Notre travail est basé sur l'utilisation d'un RNA pour interpréter les mesures de résistivité électrique en une image 2D de la résistivité jusqu'à des profondeurs de 50 mètres en utilisant un simple levé Wenner-Schlumberger de 96 électrodes, ensuite comparer les résultats du terrain prédit avec cette méthode avec les prédictions de la méthode d'inversion classique par le programme R2.

La comparaison entre la prédiction du réseau de neurones et la prédiction du processus d'inversion itératif se fait à l'aide de deux paramètres : la fonction perte de chaque prédiction faite par RNA et R2, et la mesure de la similitude globale des valeurs entre la prédiction et le terrain réel. L'exemple étudié a démontré la capacité de reconstruire le profil de résistivité électrique en profondeur à partir d'une tomographie de résistivité électrique ERT en utilisant un réseau de neurones entraîné. Les prédictions à l'aide du réseau de neurones sont plus précises et plus rapides que les prédictions issues de la méthode d'inversion traditionnelle.

**Mots-clés :** Réseaux de neurones artificiel RNA ;Tomographie par résistivité électrique ERT ; inversion classique ; apprentissage supervisé ;Wenner-Schlumberger ; R2.

## **Abstract :**

Electrical resistivity tomography (ERT) has been widely used for imaging the distribution and structure of underground resistivity. Over the years, many algorithms have been developed to solve the distribution of electrical resistivity from ERT measurements. And in order to overcome the difficulties of interpreting the ERT measures encountered during conventional inversion, by providing a 2D imaging of the resistivity profile of the sub-surface, we will present a new approach based on supervised learning for the training of an artificial neural network.

Our work is based on the use of a RNA to interpret electrical resistivity measurements in a 2D image of resistivity up to 50 meters depths, using a single Wenner-Schlumberger survey of 96 electrodes, then compared the results predicted terrain with this method with the predictions of the classical inversion method by the R2 program.

The comparison between the prediction of the neural network and the prediction of the iterative inversion process is done using two parameters: the loss function of each prediction made by RNA and R2, and the measurement of the overall similarity of the values between

the prediction and the real terrain. The example studied demonstrated the ability to reconstruct the electrical resistivity profile in depth from an electrical resistivity tomography ERT using a trained neural network. Predictions using the neural network are more accurate and faster than predictions from the traditional inversion method.

**Keywords:** Artificial neural networks RNA, ERT electrical resistivity tomography, classical inversion, supervised learning, Wenner-Schlumberger, R2.

### ملخص :

تم استخدام التصوير المقطعي للمقاومة الكهربائية (ERT) على نطاق واسع لتصوير توزيع وبنية المقاومة تحت الأرضية. على مر السنين ، تم تطوير العديد من الخوارزميات لأيجاد خريطة توزيع المقاومة الكهربائية أنطلاقاً من قياسات ERT. ومن أجل التغلب على صعوبات المصادفة عند تفسير تطبيق الانقلاب التقليدي، وتوفير لمحة 2D عن تصوير المقاومة تحت الأرض، نقدم هنا نهج جديد يقوم على تشكيل شبكة عصبية اصطناعية.

ويستند عملنا على استخدام الشبكات العصبية الاصطناعية RNA لتفسير قياسات المقاومة الكهربائية في صورة 2D من المقاومة تصل إلى أعماق 50 متراً باستخدام 96 قطب موزعة حسب تكوين Wenner-Schlumberger ، ثم مقارنة نتائج هذه الطريقة مع التنبؤات المتحصل عليها بطريقة الانقلاب الكلاسيكي من قبل برنامج R2 .

المقارنة بين التنبؤ باستخدام الشبكة العصبية والتنبؤ عن طريق عملية انقلاب متكررة تتم باستخدام معلمتين: الأولى هي دالة الفقدان لكل من التنبؤات بين RNA و R2، الثانية تتمثل في قياس قيم التشابه الشاملة بين التضاريس المتنبئة والتضاريس الحقيقية. وأظهر المثال الذي تمت دراسته القدرة على إعادة تشكيل ملامح المقاومة الكهربائية في العمق من التصوير المقطعي للمقاومة الكهربائية باستخدام ERT باستخدام شبكة عصبية مدربة. تعد التنبؤات باستخدام الشبكة العصبية أكثر دقة وأسرع من التنبؤات المستنتجة من طريقة الانقلاب التقليدية.

**الكلمات المفتاحية:** الشبكات العصبية الاصطناعية RNA ، التصوير المقطعي للمقاومة الكهربائية ERT، الانقلاب الكلاسيكي ، Wenner-Schlumberger، R2 .

# Table des matières

Introduction générale .....	7
Chapitre I : Inversion classique des mesures de résistivité électrique ERT. ....	9
1. Introduction.....	10
2. Tomographie de résistivité électrique ERT .....	10
3. Interprétation des mesures de résistivité électrique .....	11
3.1 La technique de mesure: .....	11
3.2 Le logiciel Res2Dinv: .....	12
3.3 Théorie de l'inversion géophysique .....	13
3.4 Qualité du modèle .....	17
4. Conclusion :.....	17
Chapitre II : Les réseaux de neurones artificiels-RNA- .....	19
1. Introduction.....	20
2. Propriétés fondamentales des réseaux de neurones .....	21
2.1 Approximation universelle.....	21
2.2 Approximation parcimonieuse .....	21
3. Fonctionnement d'un réseau de neurones.....	21
4. Structure des réseaux de neurones artificiels.....	23
5. Fonction de transfert.....	23
5.1 Fonction logistique .....	24
5.2 Fonction tangente hyperbolique .....	24
5.3 Fonction ReLU .....	25
6. Types de réseaux de neurones.....	25
6.1 Réseaux non récurrents .....	25
6.2 Réseaux récurrents .....	26
6.3 Réseau NARX (Non linéaire Auto Régressive avec variables eXogènes) .....	27
7. Apprentissage d'un réseau de neurone .....	29
7.1 Apprentissage supervisé .....	29
7.2 Apprentissage non supervisé .....	30
8. Conclusion.....	30
Chapitre III : Interprétation des mesures de tomographie de résistivité électrique ERT à l'aide d'un réseau de neurones artificiel. ....	31
1. Introduction.....	32
2. Méthodologie .....	32
3. Modèle et simulations .....	32

4. Description du réseau de neurones appliqué et de la base de données utilisée pour son apprentissage .....	35
5. Résultat du RNA et comparaison avec la méthode d'inversion classique .....	37
6. Conclusion.....	39
Conclusion générale .....	41
Bibliographie.....	43

# Liste des Tableaux

<b>Tableau 1</b> : Analogie entre le neurone biologique et le neurone formel.....	23
---	----

# Liste des figures

<b>Figure 1:</b> Dispositif de mesure de tomographie électrique. L'image obtenue après l'inversion correspond à une tomographie électrique 2D (d'après Aning 2012, modifié). ..	11
<b>Figure 2:</b> Schématisation de la technique utilisée lors de la récupération des mesures à inverser avec Res2Dinv .....	11
<b>Figure 3:</b> Arrangement de blocs dans un modèle 2D (M.H. Loke, 1995-2017). .....	13
<b>Figure 4:</b> Architecture d'un RNA (Zhang, 2004). .....	22
<b>Figure 5:</b> Schéma d'un neurone artificiel (Aouedad, 2018).....	23
<b>Figure 6:</b> Tracé de la fonction logistique .....	24
<b>Figure 7:</b> Tracé de la fonction Tanh .....	25
<b>Figure 8:</b> Tracé de la fonction ReLU.....	25
<b>Figure 9:</b> Réseau de neurones non bouclé (Dreyfus, 2004). .....	26
<b>Figure 10:</b> Réseau de neurones récurrent (Touzet, 1992). .....	27
<b>Figure 11:</b> Architecture NARX, Séries-Parallèles (Topus, 2010).....	28
<b>Figure 12:</b> Architecture NARX en mode parallèle (Topus,2010). .....	28
<b>Figure 13:</b> Modèle 2D utilisé pour créer l'image synthétique de résistivité et test requis pour la mise en œuvre du RNA (Réseau de Neurones Artificiel). Figure inspirée de U.K. Sing et R.K.Siwari (2013) et modifiée par K. Omraci et N. Younsi (2018). .....	33
<b>Figure 14:</b> Un exemple représentatif du terrain utilisé pour l'entraînement et le test du réseau neuronal. ....	34
<b>Figure 15:</b> Pseudo section et schéma descriptif de la configuration Wenner-Schlumberger. $b = na$ . .....	35
<b>Figure 16:</b> Système d'apprentissage supervisé (Y. BENNANI,2014-Cours RNA-U-Paris 13) ..	36
<b>Figure 17:</b> a: Histogrammes de la fonction perte L1 des prédictions du réseau de neurones et des prédictions de la méthode d'inversion classique en utilisant R2. b: Histogrammes de distance $L_1$ entre les prédictions du RNA et les prédictions de l'inversion classique par R2 par rapport au terrain réel.. .....	38
<b>Figure 18:</b> Exemples illustrant le profil de résistivité souterraine prédit.. .....	40

# Introduction générale

Ces dernières années, l'informatique a joué un rôle clé dans les sciences de la terre. Cela est en partie dû à la nature du sujet des règles régissant de nombreux phénomènes physiques dans les sciences de la terre. Comme nos problèmes sont liés aux paramètres non linéaires de la Terre, il devient trop complexe de ne compter que sur une seule discipline et nous nous trouvons au milieu de méthodes d'analyse interdisciplinaires liées à l'explosion de l'information. Pour résoudre des problèmes complexes, nous devons nous appuyer sur une approche fondée sur la connaissance plutôt que sur les techniques mathématiques classiques. Au lieu de cela, nous devons compléter les méthodes d'analyse conventionnelles avec un certain nombre de méthodologies émergentes et de techniques informatiques telles que les systèmes experts, l'intelligence artificielle, les réseaux neuronaux, la logique floue, les algorithmes génétiques, le raisonnement probabiliste et les techniques de traitement parallèle.

La tomographie électrique (ERT *Electrical resistivity tomography*) a été largement utilisée pour l'imagerie de la distribution et de la structure de la résistivité des sols et sous-sols. Au fil des ans, de nombreux algorithmes ont été développés pour résoudre la distribution de la résistivité électrique à partir des mesures ERT.

L'application d'une campagne de tomographie électrique nécessite l'implantation d'une série d'électrodes configurées selon un dispositif adéquat, l'injection d'un courant électrique continu, et la récupération de la différence de potentiel mesuré entre les électrodes de potentiel. Étant donné que le courant circule dans un chemin inconnu qui est déterminé par le profil de conductivité du terrain, l'interprétation n'est pas simple. De plus, le nombre d'inconnues dans le problème est plus grand que le nombre d'équations; par conséquent, ce problème est mal posé et doit être résolu à l'aide de méthodes d'inversion itératives.

Devant cette nécessité de trouver des solutions à des problèmes inverses ou des problèmes mal posés, plusieurs méthodes sont apparues afin de remédier à ce genre de problèmes. Dans ce mémoire, une nouvelle méthode d'interprétation des mesures ERT est présentée, en utilisant l'apprentissage supervisé pour former un réseau de neurones.

Dans le premier chapitre, nous allons introduire la tomographie électrique, son intérêt ainsi que le concept de traitement de ses données avec un processus d'inversion classique qui permet d'obtenir une meilleure idée de la résistivité électrique, ce traitement est effectué d'une manière itérative qui requiert l'utilisation de quelques logiciels d'analyse, soient Res2Dinv ou R2.

Dans le deuxième chapitre, nous introduisons les réseaux de neurones artificiels (RNA) qui sont issus de la combinaison entre la neurobiologie comme idée de base, la physique comme champs d'application et les mathématiques et l'informatique en tant que moyens de réalisation. Ils traitent l'information d'une manière parallèle et distribuée, de telles propriétés

leur confèrent bien la qualité de robustesse vis-à-vis des données incomplètes, incertaines et bruitées. Chose qui motive leur utilisation dans le cadre du problème inverse.

Le troisième chapitre de ce mémoire, présente l'interprétation des mesures ERT est en utilisant l'apprentissage supervisé pour former un réseau de neurones, nous sommes en mesure d'interpréter la mesure ERT en une image 2D de la résistivité souterraine jusqu'à des profondeurs de 50 mètres en utilisant un simple levé Wenner-Schlumberger de 96 électrodes espacées d'un mètre. Le réseau de neurones est formé et testé à l'aide de données simulées et ses résultats sont supérieurs à ceux d'une méthode d'inversion classique.

Enfin, une conclusion générale viendra finaliser ce mémoire pour faire la synthèse des résultats d'analyse et de comparaison entre les deux méthodes.

# **Chapitre I : Inversion classique des mesures de résistivité électrique ERT.**

## 1. Introduction

Les méthodes géophysiques permettent de déduire les caractéristiques internes des milieux auscultés en étudiant les variations d'un champ physique mesuré, et en effectuant des mesures indirectes et non intrusives de la résistivité électrique ou de la vitesse des ondes acoustiques.

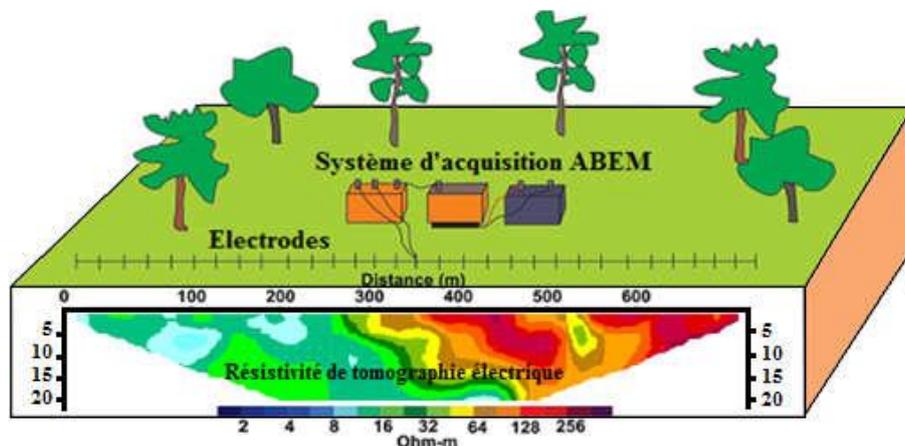
À la différence des méthodes géotechniques, les méthodes géophysiques reconstruisent les propriétés d'un milieu à partir de données obtenues de manière indirecte. Mais le problème rencontré est l'incertitude et l'imprécision de ces données mesurées lors de leur interprétation.

Donc, afin de quantifier les propriétés des matériaux en place, la majorité des méthodes géophysiques utilise les « problèmes inverses » ou présente des variations « apparentes » d'observables. La résolution du problème inverse (inversion) caractérise l'action de trouver un modèle expliquant au mieux les données mesurées.

## 2. Tomographie de résistivité électrique ERT

En prospection géophysique, la distribution des résistivités électriques en fonction de la profondeur peut être présentée en imagerie électrique 2D ou 3D, cette procédure est appelée tomographie électrique (figure 1), elle est appliquée à différentes échelles (de celle du laboratoire à celle des structures kilométriques). Elle a pour objectif la reconnaissance multidimensionnelle du milieu étudié.

La résistivité électrique est la capacité d'un milieu à s'opposer au passage d'un courant électrique. Ce paramètre varie fortement dans la nature et dépend de nombreux paramètres physiques. Les méthodes de tomographie de résistivité électrique consistent à faire circuler un courant électrique dans le terrain ou dans un massif rocheux entre deux électrodes de courant et de mesurer la différence de potentiel entre deux autres électrodes, appelées électrodes de potentiel, à proximité des électrodes de courant. Puisque l'intensité du courant est connue et que la différence de potentiel est mesurée, il est alors possible de déterminer quelle est la résistivité apparente. Cette résistivité électrique apparente dépend de la configuration des électrodes de courant et de potentiel. Cette méthode d'auscultation géophysique connaît un renouveau sous la forme de prospection et de modélisation en deux ou trois dimensions.



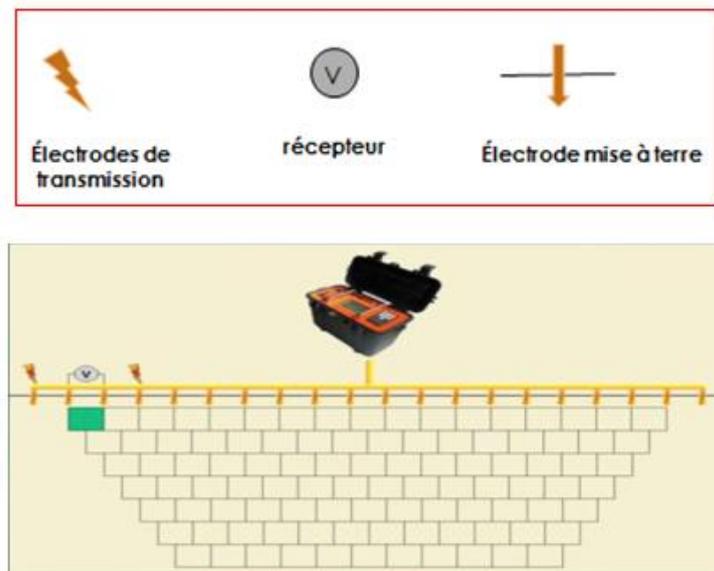
**Figure 1:** Dispositif de mesure de tomographie électrique. L'image obtenue après l'inversion correspond à une tomographie électrique 2D (d'après Aning 2012, modifié).

### 3. Interprétation des mesures de résistivité électrique

Pour procéder à l'interprétation des données mesurées de résistivité électrique, il faut d'abord construire une pseudo-section, et pour la construire, il faut installer un panneau électrique en implantant une multitude d'électrodes, injecter un courant électrique et effectuer les mesures de potentiel qui sont automatiquement sauvegardées. Ensuite établir à partir de ces résultats la carte qui présente la distribution des résistivités apparentes calculées dans le corps. L'objectif principal de cette étape est d'avoir une idée initiale pour l'interprétation quantitative, il faut donc noter que la carte obtenue ne représente pas la distribution réelle des résistivités électriques.

#### 3.1 La technique de mesure:

- Chacune des électrodes est connectée soit à l'émetteur ou au récepteur.
- le tableau commute automatiquement les électrodes
- Effectuant la mesure de  $\rho_{app}$  pour différentes combinaisons (figure 2).
- Le processus d'inversion inverse les données en section transversale 2D.



**Figure 2:** Schématisation de la technique utilisé lors de la récupération des mesures à inverser avec Res2Dinv

Étant donné que le courant circule dans un chemin inconnu qui est déterminé par le profil de conductivité du terrain, l'interprétation n'est pas simple. De plus, le nombre d'inconnues dans le problème est plus grand que le nombre d'équations; par conséquent, ce problème est mal posé et doit être résolu à l'aide de méthodes d'inversion itératives.

Dans le cas d'un milieu qui n'est pas homogène, tels que celui que l'on rencontre le plus souvent dans la nature, la pseudo-section ne peut pas être interprétée directement, d'autant plus que celle-ci dépend également de la configuration des électrodes. La pseudo-section peut être produite dans le cas d'un profil en solutionnant la loi d'Ohm et l'équation de Laplace. Ce

processus est appelé modélisation directe. Seules des interprétations qualitatives peuvent être effectuées à partir de celle-ci.

Pour obtenir une mesure quantitative de la résistivité électrique vraie, il est nécessaire de réaliser une procédure dite d'inversion. Lorsqu'un levé de résistivité électrique est réalisé, le but consiste à obtenir le modèle synthétique de résistivité électrique vraie du milieu étudié à partir de la pseudo-section de résistivité électrique apparente observée. Ce processus s'appelle l'inversion puisqu'il s'agit de l'inverse de la modélisation directe.

L'inversion est une méthode itérative automatique qui peut fournir un modèle de la résistivité électrique vraie afin de placer les structures à une profondeur adéquate. Par contre, la solution mathématique de l'inversion n'est pas unique car plusieurs modèles peuvent correspondre à la même pseudo-section. Le modèle qui correspond le mieux à la résistivité électrique apparente mesurée sur le terrain n'est donc pas nécessairement le bon (Frohlich *et al*, 1994).

L'inversion des données issues antérieurement de la pseudo-section peut s'effectuer grâce aux programmes comme **Res2Dinv** (Loke, 1997) ou **R2** (A. Binley, 2016).

A chaque itération, le profil de résistivité est modifié d'une certaine manière afin de minimiser une fonction de coût (qui inclut la différence entre les solutions directes simulées et les mesures réelles et certaines autres contraintes). Cependant, la solution finale est sensible à la fois à la conjecture initiale et au choix des contraintes et de la fonction de coût.

Le coût est évalué par calcul de la norme L2, cette norme étant communément utilisée dans la résolution du problème inverse. Ainsi pour chaque modèle de conductivité, on calcule le potentiel pour chaque injection, on prélève le potentiel en surface aux mêmes électrodes que dans le protocole utilisé pour l'acquisition sur le terrain, et on compare le potentiel modélisé au potentiel mesuré en calcul :

$$\text{Coût} = \sum [ddpmod(i) - ddpmes(i)]^2 \quad (1)$$

Où :

$ddpmod(i)$  : est la différence de potentiel modélisée pour le quadripôle  $i$ .

Et  $ddpmes$  : est la différence de potentiel mesurée de ce même quadripôle  $i$ .

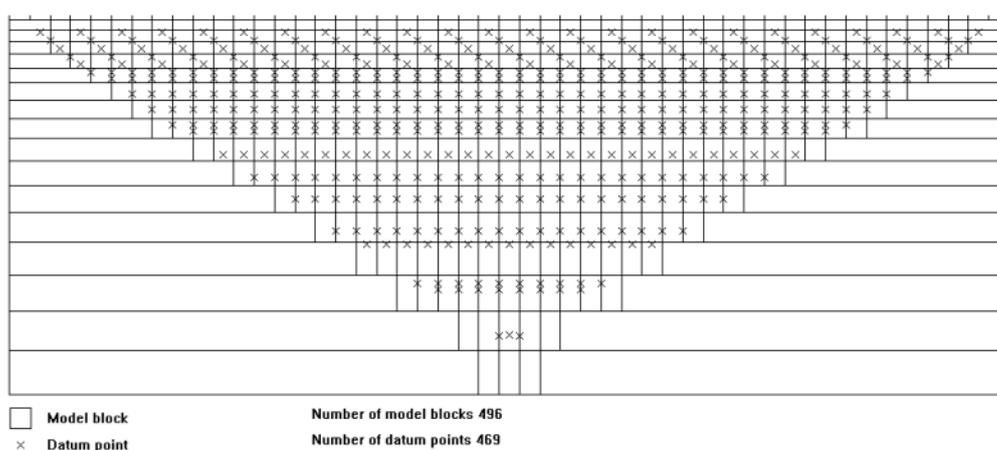
Le fait que le coût soit calculé sur la différence de potentiel et non sur la conductivité (puisque c'est l'inconnue) complique l'inversion. En effet, le modèle de conductivité pour lequel le coût calculé à partir des différences de potentiel est moindre, n'est pas forcément le meilleur modèle en termes de conductivité. On pourra également avoir plusieurs modèles de conductivité expliquant les données sans correspondre au modèle vrai.

### 3.2 Le logiciel Res2Dinv:

Le logiciel d'inversion Res2Dinv utilise un ensemble de blocs rectangulaires, leur disposition est liée en premier lieu à la distribution des points de données dans la pseudo-section, et leur tailles et agencement sont conçus automatiquement par le programme (cf.

figure 3). La profondeur de la dernière rangée de bloc est déterminée à partir de la profondeur d'investigation qui correspond aux points de données ayant le plus grand écartement d'électrodes.

Le calcul des résistivités apparentes se fait par la modélisation utilisant la méthode des différences finies ou bien par la méthode des éléments finis. Il est adapté aux différentes configurations possibles (Wenner, Schlumberger, dipôle-dipôle, pôle-pôle etc.), en plus de toutes ces configurations, Res2Dinv peut prendre en charge des configurations non-conventionnelles avec un nombre approximativement illimité d'électrodes. On peut donc traiter des pseudo-sections avec jusqu'à 16000 positions d'électrodes et 100000 points de données en une seule fois. Le programme est adéquat aussi à des milieux aquatiques et à des forages transversaux.



**Figure 3:** Arrangement de blocs dans un modèle 2D (M.H. Loke, 1995-2017).

Le logiciel supprime automatiquement les mauvaises mesures pour qu'elles ne perturbent pas le calcul du modèle. Il fonctionne avec un système itératif, c'est à dire pour chaque modèle qu'il propose, il le compare avec la pseudo-section, jusqu'à ce qu'il trouve un pourcentage d'erreur acceptable qui soit le minimum possible et inférieur à celui introduit au préalable lors des choix des paramètres. Ceci dit que le résultat du problème inverse n'est pas unique, ce qui nécessite l'évaluation de la fiabilité du modèle obtenue.

Lors de l'inversion on peut rencontrer deux genres d'erreur, une systématique et l'autre aléatoire, la première concerne les problèmes techniques qui surgissent dans l'acquisition des données, et la deuxième concerne les paramètres incontrôlables comme la variation diurne par exemple.

### 3.3 Théorie de l'inversion géophysique

Nous cherchons, en inversion géophysique, à trouver un modèle avec une réponse similaire aux valeurs mesurées réelles. Le modèle obtenu par inversion classique n'est qu'une représentation idéalisée d'une section de la terre ayant un ensemble de paramètres (quantités physiques) que nous voulons estimer à partir des données mesurées. La réponse du modèle est

la donnée synthétique qui peut être calculée à partir de la relation mathématique définissant le modèle pour un ensemble donné de paramètres de celui ci.

Toutes les méthodes d'inversion tentent essentiellement de déterminer un modèle pour le sous-terrain dont la réponse concorde avec les données mesurées soumises à certaines restrictions. Dans la méthode utilisées par le programme Res2Dinv, les paramètres du modèle sont les valeurs de résistivité des mailles (ou blocs) du modèle, alors que les données sont les valeurs de résistivité apparente mesurées.

Le lien mathématique entre les paramètres du modèle et la réponse du modèle pour les modèles de résistivité 2D est fourni par les méthodes de différences finies ou d'éléments finis.

Dans toutes les méthodes d'optimisation, un modèle initial est modifié de manière itérative afin de réduire la différence entre la réponse du modèle et les valeurs de données observées. L'ensemble des données observées peut être écrit comme un vecteur de colonne y donné par :

$$y = col(y_1, y_2, \dots, y_m) \quad (2)$$

Tel que :

m : Nombre de mesures.

La réponse du modèle f peut être écrite aussi sous une forme similaire :

$$f = col(f_1, f_2, \dots, f_m) \quad (3)$$

Il est courant d'utiliser le logarithme des valeurs de résistivité apparente pour les données observées et la réponse du modèle, et le logarithme des valeurs du modèle comme paramètres du modèle.

Les paramètres du modèle peuvent être représentés par le vecteur suivant :

$$q = col(q_1, q_2, \dots, q_n) \quad (4)$$

Où n est le nombre des paramètres du modèle.

La différence entre les données mesurées (résistivité apparente mesurée) et la réponse du modèle (résistivité apparente calculée) est donnée par le vecteur de discordance suivant :

$$g = y - f \quad (5)$$

En optimisation avec la méthode des moindres carrés, le modèle initial est modifié de façon à avoir une différence minimisée de la somme des carrés des erreurs E entre les résistivités apparentes mesurées et calculées.

$$E = g^T g = \sum_{i=1}^n g_i^2 \quad (6)$$

Et pour réduire cette erreur, l'équation suivante de Gauss-Newton est utilisée afin de déterminer la modification des paramètres du modèle qui devrait minimiser la somme des erreurs carrés.

$$J^T J \Delta q_i = J^T g \quad (7)$$

Où :

$\Delta q$  : Le vecteur de changement de paramètre du modèle.

J : La matrice jacobienne (m×n) des dérivées partielles.

Les éléments de la matrice Jacobienne sont donnés par :

$$J_{ij} = \frac{\partial f_i}{\partial q_j} \quad (8)$$

Le changement dans la  $i^{\text{ème}}$  réponse du modèle est dû au changement dans le  $j^{\text{ème}}$  paramètre du modèle. Un nouveau modèle est obtenu après le calcul du vecteur de changement de paramètres par :

$$q_{k+1} = q_k + \Delta q_k \quad (9)$$

En pratique, l'équation simple des moindres carrés est rarement utilisée seule en inversion géophysique. Dans certains cas le produit matriciel  $J^T J$  pourrait être singulier, et donc l'équation n'aura pas de solution pour  $\Delta q$ . Il existe un autre problème: si le produit matriciel est presque singulier, cela peut se produire si un modèle initial médiocre, très différent du modèle optimum, est utilisé. Le vecteur de changement de paramètre calculé en utilisant l'équation (7) peut avoir des composantes qui sont trop grandes pour que le nouveau modèle calculé avec l'équation (9) puisse avoir des valeurs qui ne sont pas réalistes. Afin d'éviter ce problème, une méthode courante « la modification de Marquardt-Levenberg » de l'équation de Gauss-Newton est utilisée :

$$(J^T J + \lambda I) \Delta q_k = J^T g \quad (10)$$

Sachant que :

I : La matrice identité.

$\lambda$  : Facteur d'amortissement ou le Marquardt.

Cette méthode est également connue sous le nom de méthode de régression de crête. Le facteur d'amortissement contraint effectivement la plage de valeurs que les composantes du vecteur de changement de paramètre peuvent prendre. Alors que la méthode de Gauss-Newton tente de minimiser la somme des carrés du seul vecteur de discordance, la modification de la méthode de Marquardt-Levenberg réduit également une combinaison de l'amplitude du vecteur de discordance et du vecteur de changement de paramètre. Cette méthode a été utilisée avec succès dans l'inversion des données de résistivité lorsque le

modèle consiste en un petit nombre de couches. Cependant, lorsque le nombre de paramètres du modèle est grand, comme dans le modèle d'inversion 2D constitué d'un grand nombre de petites cellules ou mailles, le modèle produit par cette méthode peut avoir une distribution de résistivité erratique avec des zones de résistivité élevée ou faible. Pour surmonter ce problème, l'équation de Gauss newton des moindres carrés est encore modifiée afin de minimiser les variations spatiales dans les paramètres du modèle et exprimée comme suit :

$$(J^T J + \lambda F) \Delta q_k = J^T g - \lambda F q_k \quad (11)$$

Où :

$$F = \alpha_x C_x^T \alpha_x + \alpha_z C_z^T \alpha_z \quad (12)$$

Et :

$C_x$  et  $C_z$  sont les matrices de lissage dans les directions x et z.

$\alpha_x$  et  $\alpha_z$  sont les poids relatifs donnés aux facteurs de lissage dans les directions x et z.

Une forme commune de la matrice de lissage est la matrice de différence de premier ordre donnée par :

$$C = \begin{bmatrix} -1 & 1 & 0 & 0 & \dots & \dots & \dots & 0 \\ 0 & -1 & 1 & 0 & \dots & \dots & \dots & 0 \\ 0 & 0 & -1 & 1 & 0 & \dots & \dots & 0 \\ & & & \dots & & & & \\ & & & & \dots & & & \\ & & & & & \dots & & \\ & & & & & & \dots & \\ & & & & & & & \dots & \\ & & & & & & & & \dots & 0 \end{bmatrix}$$

Et :

$C_x$  : Filtres de rugosité horizontaux.

$C_z$  : Filtres de rugosité verticaux.

J : Matrice jacobienne de dérivées partielles.

$J^T$  : Matrice transposée de J.

$\lambda$  : Facteur d'amortissement.

q : Vecteur de changement de modèle.

g : Vecteur de données erronées.

Parmi les avantages de cette méthode est l'adaptation aux différents types de données grâce à la capacité de ce logiciel d'ajuster les différents paramètres  $\lambda$  et C. l'optimisation utilisée cherche à minimiser la différence entre les valeurs de résistivité apparente calculées et

celles mesurées tout en ajustant la résistivité des blocs modèles soumis aux contraintes de lissage employées. Cette différence est traduite par RMS (*root-mean-squared error*) ou la racine carrée de l'erreur quadratique moyenne (ou encore l'écart type). Il faut noter que le modèle avec l'erreur RMS la plus faible ne donne pas forcément le meilleur modèle du point de vu géologique, car il montre parfois des variations importantes et irréalistes dans les valeurs de résistivités. En général il est recommandé de choisir un nombre d'itérations entre 3 et 6 de façon à avoir un RMS qui ne change pas significativement, cela donne une approche prudente.

### 3.4 Qualité du modèle

#### 3.1.1. RMS (root-mean-squared error)

Le modèle avec l'erreur RMS la plus faible ne donne pas forcément le meilleur modèle. Il est recommandé de choisir un nombre d'itération entre 3 et 6 de façon à avoir un RMS qui ne change pas significativement.

#### 3.1.2. Fonction de vraisemblance :

Il faut d'abord calculer l'erreur notée  $\varepsilon$  sur les valeurs de la résistivité après inversion afin de déterminer la fonction de vraisemblance  $L_p$  selon la théorie de l'erreur de Gauss.

Où :

$$L_p = \sigma \sqrt{2\pi\varphi} \quad (13)$$

$$\varphi(\varepsilon) = \frac{1}{\sigma\sqrt{2\pi}} e^{\left(\frac{-\varepsilon^2}{2\sigma^2}\right)} \quad (14)$$

#### 3.1.3. Fonction de sensibilité :

L'avantage de cette fonction est d'avoir une vision sur l'influence des variations de résistivité du terrain sur la mesure du potentiel. Sachant que si cette fonction est importante près des électrodes, alors l'influence sur la mesure de potentiel est considérable.

## 4. Conclusion :

Dans l'inversion conventionnelle ou dite itérative, le processus est constitué d'une boucle d'itération au sein de laquelle se trouve le modèle direct, qui doit être connu, de façon exacte ou approchée. La sortie du modèle direct est comparée aux acquisitions effectuées, et l'erreur ainsi observée est proposée en entrée d'un algorithme d'optimisation qui modifie alors les estimations des caractéristiques recherchées. À chaque itération, cette modification a pour but de minimiser l'erreur obtenue. L'inversion est supposée correcte et le processus est arrêté lorsque cette erreur devient inférieure à une valeur limite déterminée comme critère d'arrêt des itérations. Si toutes les méthodes d'inversion itératives répondent à ce processus, la différence réside essentiellement dans l'algorithme d'optimisation qui effectue l'estimation des caractéristiques par modifications successives. De tels algorithmes sont relativement nombreux. Parmi les plus utilisés nous avons : la méthode des dérivées, la méthode linéaire itérative, méthode de descente, méthode de recherche systématique, méthode de recuit, ou encore par le point de vue probabiliste. Un inconvénient de cette technique est le temps de

calcul requis, en particulier si le modèle utilisé est complexe. De plus, l'inversion itérative est assez sensible à la précision du modèle direct : un modèle inexact peut aboutir à une solution non satisfaisante tandis qu'un modèle trop précis risque de rendre le processus peu robuste vis-à-vis de perturbations. Le choix de l'initialisation des caractéristiques estimées est aussi important. En effet, ces algorithmes d'optimisation recherchent des minima, et il arrive que le résultat soit un minimum local et non global. Cependant, des méthodes existent pour pallier ces problèmes, et l'inversion itérative est présentée comme une méthode très efficace et précise lorsqu'elle est correctement mise en œuvre.

# **Chapitre II : Les réseaux de neurones artificiels-RNA-**

## 1. Introduction

Depuis une vingtaine d'années, l'utilisation des réseaux de neurones artificiels (RNA) s'est étendue dans des domaines très divers de l'industrie et des services. En géophysique les RNA ont été utilisés pour plusieurs méthodes, par exemples pour détecter les premières arrivées d'ondes sismiques (Murat and Rudman, 1992; McCormack et al., 1993); classifier les différents signaux par l'inversion (Roth and Tarantola, 1994; Langer et al., 1996); transformer un problème de déconvolution sismique dans un réseau de Hopfield afin de réduire le temps de calcul (Wang and Mendal, 1992); pour caractériser la distribution de résistivité du sous-terrain par l'inversion de données magnétotelluriques (Zhang and Paulson, 1997) et électromagnétiques (Poulton et al., 1992). Ils sont particulièrement utilisés pour résoudre des problèmes de classification, de prédiction, de reconnaissance des formes, de catégorisation, de mémoire associative et d'optimisation (Drew et Monson, 2000). Par l'entraînement d'un système non linéaire de multiples variables, les RNA peuvent prédire la variable indépendante (Huang et al., 1996). Par conséquent, les RNA constituent une technique d'approximation de systèmes complexes, qui sont difficiles à modéliser par les méthodes statistiques classiques. Puisqu'il existe un grand nombre de structures des réseaux de neurones artificiels, il est nécessaire de faire un résumé de leurs caractéristiques afin de choisir un réseau de neurone approprié au développement de ce mémoire.

Les réseaux de neurones artificiels (RNA) sont des modèles informatiques qui ont été inspirés des systèmes nerveux biologiques pour réaliser des tâches de reconnaissance de formes et autres traitements informatiques

Un RNA est un réseau composé de nombreuses unités de calcul simples appelées neurones ou cellules, qui sont fortement interconnectées et organisées en couches. Chaque neurone effectue une simple tâche de traitement de l'information en convertissant des entrées reçues en des sorties. Grâce aux arcs de liaison entre ces neurones, la connaissance peut être générée et stockée.

Les recherches scientifiques réalisées sur les réseaux de neurones artificiels (RNA) ont affirmé que cet outil possède une grande capacité de prédiction et s'inscrit parmi les méthodes de modélisation mathématique les plus robustes.

Leur application sur des problématiques réelles et dans différents domaines a donné une grande satisfaction et a confirmé ainsi les conclusions scientifiques. Aujourd'hui, les RNA sont utilisés pour anticiper l'évolution des indices boursiers, la consommation énergétique, la vérification d'authenticité, ... (Zhang, 2004).

Les RNA présentent plusieurs avantages qui permettent de les propulser parmi les alternatives les plus prometteuses dans la prédiction. En effet, la structure inhérente aux réseaux leur donne la capacité d'approcher n'importe quelle fonction dans un intervalle fermé. Les RNA peuvent non seulement modéliser les formes linéaires mais sont aussi très performants face à des modèles non linéaires quelconques comme cela a été prouvé dans *the Universal approximation theorem* (Hornik, 1990).

Un autre aspect très important est que les RNA sont des méthodes de type *data driven* qui n'exigent pas beaucoup d'hypothèses ou de contraintes sur le processus en question. Cette caractéristique de «l'apprentissage par les données ou par l'expérience» est hautement souhaitable dans diverses situations de prédiction où les données sont généralement plus faciles à collecter, mais sont régies par un mécanisme pas clair et pas spécifié (Danilo, 2001).

## **2. Propriétés fondamentales des réseaux de neurones**

Deux caractères fondamentaux intrinsèques aux réseaux de neurones sont abordés dans ce qui suit. Ces propriétés justifient le choix de cet outil pour notre problème.

### **2.1 Approximation universelle**

« Toute fonction bornée suffisamment régulière peut être approchée avec une précision arbitraire, dans un domaine fini de l'espace de ses variables, par un réseau de neurones comportant une couche de neurones cachés en nombre fini, possédant tous la même fonction d'activation, et un neurone de sortie linéaire »(Hornik,1990).

Cet avantage qui, certes représente un atout considérable, n'est pas une spécificité exclusive aux réseaux de neurones artificiels et plusieurs autres familles de fonctions paramétrées partagent cette propriété. L'avantage réel des RNA réside dans le caractère « parcimonieux » de l'approximation (Dreyfus, 2004).

### **2.2 Approximation parcimonieuse**

« À un même degré de précision, certains réseaux de neurones nécessitent moins de paramètres à ajuster que les autres approximateurs » (Barron, 1993). En effet, le nombre de paramètres nécessaires pour atteindre une approximation fiable demeure un critère très important dans le choix de l'approximateur. Le concepteur doit toujours opter pour l'approximateur avec le plus faible nombre de paramètres possible, dans ce cas on dit que nous cherchons l'approximation la plus parcimonieuse (Dreyfus, 2004).

- Intérêt du caractère parcimonieux:

Il est utile de préciser que l'approximateur parcimonieux en paramètres est aussi parcimonieux en données (Dreyfus, 2004).

Il sera développé ultérieurement que l'apprentissage du réseau de neurones consiste en l'ajustement des paramètres à partir des exemples. Dans le but de minimiser la fonction du « coût », le réseau doit traiter plusieurs cas avant d'arriver à une combinaison optimale des paramètres. Or, dans beaucoup de problèmes, disposer d'une grande quantité de données peut être une mission très compliquée (travailler sur des événements à fréquence hebdomadaire ou mensuelle par exemple), ce qui conforte le besoin d'un outil qui réalise une fonction avec le minimum de ressources (Dreyfus, 2004).

## **3. Fonctionnement d'un réseau de neurones**

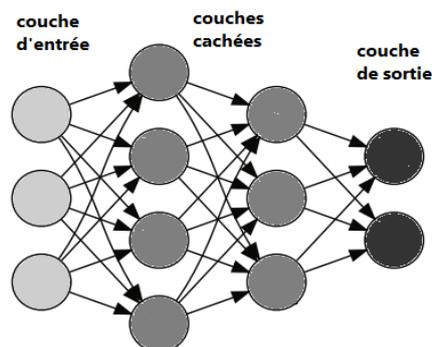
Un réseau de neurones artificiel est un ensemble de structures élémentaires appelées les « neurones ou nœuds » qui transforment un signal d'entrée  $x$  en un signal de sortie  $y$  en réalisant une opération appelée fonction de transfert (ou fonction d'activation).

Les neurones qui composent un réseau sont arrangés dans des couches. Généralement, un réseau est composé de trois couches appelées : couche d'entrée, couche interne et couche de sortie (figure 4). Il existe des modèles avec plusieurs couches internes, on parle alors d'un apprentissage profond ou *deep Learning*.

Chaque neurone du réseau est connecté avec tous les neurones qui composent la couche suivante, cette connexion est pondérée par une valeur  $\omega_{ij}$  qui représente le poids de la liaison entre le neurone  $i$  et le neurone  $j$ . L'information circule de la couche d'entrée vers la couche de sortie en passant par la couche interne (Laurne, 1993).

Chaque réseau de neurones artificiel est un modèle unique qui est caractérisé par :

- Le nombre de couches internes du réseau
- Le nombre de neurones dans chaque couche
- Les paramètres  $\omega_{ij}$  de chaque connexion
- L'entrée inhibitrice  $\theta_j$
- La fonction de transfert



**Figure 4:** Architecture d'un RNA (Zhang, 2004).

- **Neurone formel**

Comme mentionné précédemment, le neurone est l'élément chargé de transformer un signal d'entrée en un signal de sortie. Les valeurs sur lesquelles agit le neurone sont appelées les entrées du neurone, le résultat de tout le processus est appelé sortie du neurone.

Cette opération se fait en deux étapes résumées dans la figure 5 (Laurne, 1993)

1. Réaliser une somme pondérée des signaux d'entrée :

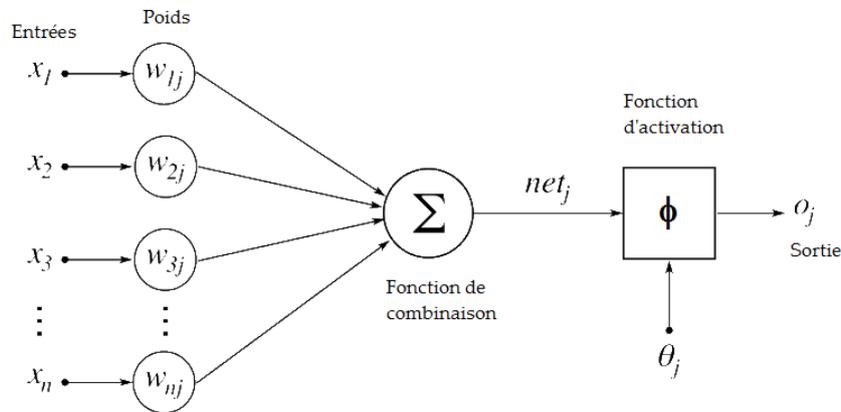
$$Z = \omega_1x_1 + \omega_2x_2 + \dots + \omega_nx_n \quad (15)$$

2. Exécuter la fonction de transfert sur  $Z$  :

$$Y = f(Z + \theta) \quad (16)$$

$f$  : fonction d'activation ou fonction de transfert

$\theta$  : biais ou seuil



**Figure 5:** Schéma d'un neurone artificiel (A.oudad, 2017).

#### 4. Structure des réseaux de neurones artificiels

McCulloch et Pitts en 1943 ont implémenté un système de réseaux neuronaux artificiels, qui est analogue aux neurones biologiques fondé sur une structure complexe (cf. tableau 1). Le système des RNA est considéré comme un arrangement d'éléments de structure identique appelés neurones interconnectés par analogie avec cellules du système nerveux humain.

Il est composé également d'une succession de couches connectées de manière à ce que chaque neurone tienne son entrée de la sortie du neurone précédant. Chaque neurone dans ce cas fonctionne indépendamment par rapport aux autres afin que l'ensemble forme un système compact. L'information est emmagasinée de façon répartie dans le réseau sous forme de coefficients synaptiques. Le neurone formel calcule régulièrement un résultat qu'il transmet ensuite aux neurones suivant, chaque calcul est associé à un poids qui définit la force de la connexion (Hagan et al. 1996).

**Tableau 1 :** Analogie entre le neurone biologique et le neurone formel

Neurone artificiel	Neurone biologique
Poids de connexion	Synapses
Signal de sortie	Axones
Signal d'entrée	Dendrite
Fonction d'activation	Soma

#### 5. Fonction de transfert

Les fonctions non-linéaires sont les plus fréquemment utilisées grâce à leurs caractéristiques de modélisation des signaux. Parmi ces fonctions, il y a les sigmoïdes (la fonction logistique et la fonction tangente hyperbolique Tanh) et plus récemment, la fonction

ReLU. Ces fonctions sont capables de généraliser ou de s'adapter à une variété de données tout en différenciant la sortie (Sharma, 2017).

Afin qu'une fonction de transfert puisse être appliquée de manière efficace, elle doit vérifier les deux critères suivants :

- Différentiable : un changement dans les x implique un changement dans y
- Fonction monotone : une fonction qui est toujours croissante ou toujours décroissante.

### 5.1 Fonction logistique

Elle a une forme en S et donne des valeurs dans l'intervalle [0,1]. Son gradient très important pour les valeurs proches de 0 fait que la moindre variation de x causera une variation significative de y, ce qui rend cette fonction très performante pour les entrées proches de 0 (Sharma, 2017). La fonction logistique, dont le graphe est donné dans la figure 6, s'écrit comme suit :

$$Y(x) = \frac{1}{1+e^{-x}} \quad (17)$$

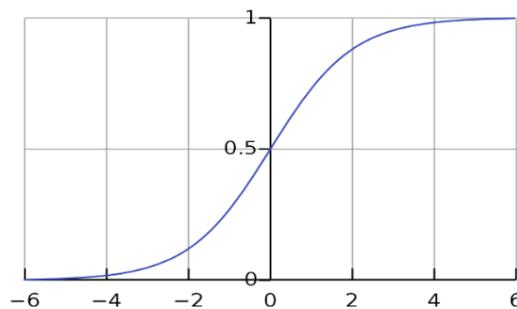
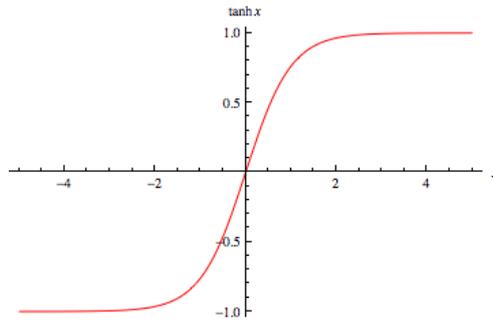


Figure 6: Tracé de la fonction logistique

### 5.2 Fonction tangente hyperbolique

Elle présente beaucoup de similitudes avec la fonction logistique. Les valeurs de sortie de cette fonction varient entre -1 et +1 (cf. figure 7), ce qui se traduit par un gradient plus important (Sharma, 2017). L'expression de la fonction tanh s'écrit comme suit :

$$\tanh(x) = \frac{2}{1+e^{-2x}} - 1 \quad (18)$$

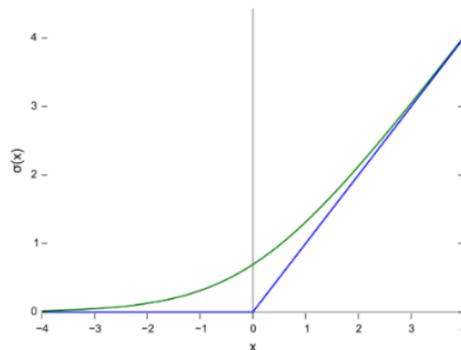


**Figure 7:** Tracé de la fonction Tanh

### 5.3 Fonction ReLU

C'est la fonction d'activation la plus récente et qui est très répandue actuellement. Elle est aussi bien utilisée dans les réseaux simples que dans les réseaux d'apprentissage profond. Il existe d'autres variantes à cette fonction qui sont LeakyReLU, PeLU, ... Elle est exprimée par (figure 8) (Sharma, 2017) :

$$F(x) = \max(0, x) = \begin{cases} 0 & \text{si } x < 0 \\ x & \text{sinon} \end{cases} \quad (19)$$



**Figure 8:** Tracé de la fonction ReLU

## 6. Types de réseaux de neurones

### 6.1 Réseaux non récurrents

Dans ce type de réseaux, l'information circule uniquement de la couche d'entrée vers la couche de sortie et le chemin inverse n'est donc pas possible comme illustré dans la figure 9. D'un point de vue graphique, le réseau non récurrent est l'équivalent d'un réseau acyclique (Dreyfus, 2004). Plusieurs types de réseaux s'inscrivent parmi les réseaux non bouclés, nous citerons principalement le FFNN (Feed Forward Neural Network) ou réseau à propagation directe, qui est le plus communément utilisé au point où si le type du réseau n'est pas précisé, c'est qu'il s'agit d'un FFNN.

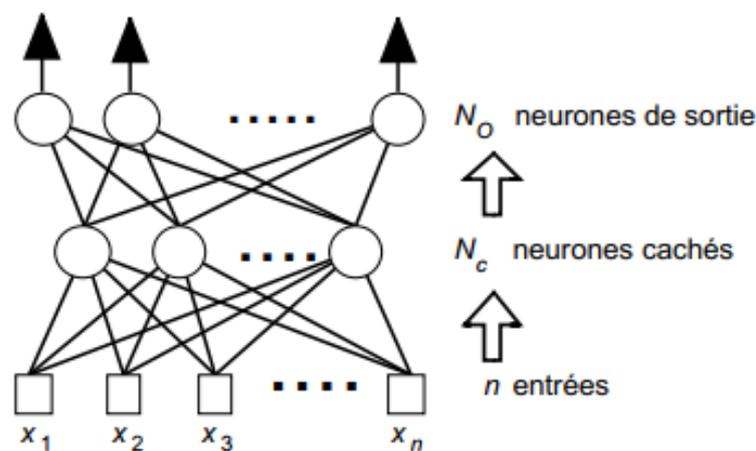
Dans cette architecture, les neurones d'entrée correspondent aux variables indépendantes  $x$  qui sont nécessaires pour prévoir la variable  $y$ .

Le FFNN est donc utilisé lorsque la variable de sortie est expliquée par plusieurs variables d'entrée appelées variables explicatives. Les neurones dans la couche cachée sont la clé pour détecter la relation entre les variables d'entrée et la variable de sortie. Avec des fonctions de transfert non linéaires, les neurones cachés peuvent traiter des informations complexes reçues des neurones d'entrée, puis envoyer ces informations traitées à la couche de sortie pour un traitement supplémentaire avant de générer des prédictions.

Il convient d'ajouter que les réseaux non bouclés sont aussi appelés réseaux statiques du fait que le temps ne joue aucun rôle et que les fonctions sont exécutées instantanément.

- **Cas particulier : Le Perceptron Multicouches**

Le perceptron multicouches est un type spécial parmi les réseaux de neurones non récurrents dont la caractéristique est que la fonction d'activation est un sigmoïde pour tous les neurones excepté les neurones d'entrée qui exécutent une fonction d'activation linéaire. Cette structure est très largement utilisée et on la trouve parfois avec la notation MLP (Multi-Layer Perceptron).



**Figure 9:** Réseau de neurones non bouclé (Dreyfus, 2004).

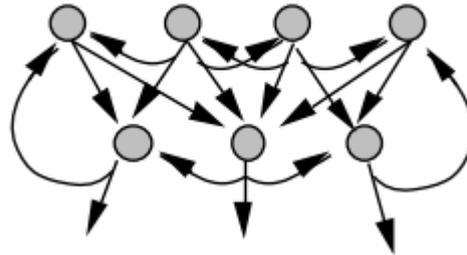
## 6.2 Réseaux récurrents

Dans la partie précédente, il a été expliqué que les réseaux de neurones non récurrents sont soumis à la contrainte d'absence de boucle et par conséquent l'information ne peut circuler que dans un seul sens.

Si nous détenons cette condition, et que nous laissons apparaître au moins un cycle à l'intérieur du réseau, nous obtenons alors un réseau récurrent où l'information peut revenir en arrière pour alimenter les neurones.

Contrairement aux réseaux *feed forward* qui appréhendent les vecteurs d'entrée et de sortie d'une manière indépendante, les RNN (Recurrent neural network) peuvent faire intégrer l'historique de l'ensemble des entrées précédentes pour chaque sortie, ce sont donc des systèmes dynamiques.

Les réseaux récurrents (cf. figure 10) permettent de traiter des données séquentielles. En effet, à l'instant  $t$  ils calculent leur sortie en fonction de l'état aux temps précédents. Ainsi, ils font évoluer un état interne qui fait office de mémoire à court terme et qui permet de prendre en compte les dépendances temporelles que manifestent les sorties (Tristan, 2016).



**Figure 10:** Réseau de neurones récurrent (Touzet, 1992).

La récurrence à l'intérieur du réseau peut être locale ou globale. Une récurrence locale est, soit une connexion entre deux neurones appartenant à la couche interne (ou entre deux couches internes), soit une boucle à l'intérieur d'un même neurone.

Une récurrence globale est, quant à elle, réalisée par une liaison partant de la couche de sortie vers la couche d'entrée (Danilo, 2001).

### 6.3 Réseau NARX (Non linéaire Auto Régressif avec variables exogènes)

#### 6.3.1 Définition

NARX est un réseau de neurones dynamique, constitué d'un Perceptron Multicouches (MLP) où la future valeur de la variable à prédire  $y(t)$  est en fonction des  $n_y$  précédentes valeurs de sortie et des  $n_u$  précédentes valeurs d'une autre variable appelée variable exogène (Dreyfus, 2004). L'équation qui modélise le système s'écrit comme suit :

$$Y(t) = f(y(t-1), y(t-2), \dots, y(t-n_y), u(t-1), u(t-2), \dots, u(t-n_u)) \quad (20)$$

Où

- $u_t$  et  $y_t$  sont respectivement l'entrée et la sortie du modèle à l'instant  $t$ .
- $n_u$  et  $n_y$  sont respectivement l'ordre du retard de l'entrée et ordre du retard de la sortie.

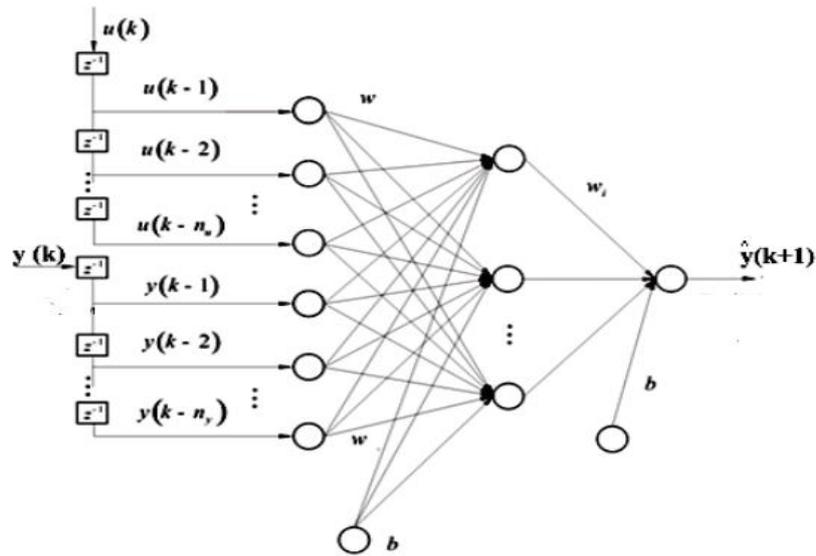
#### 6.3.2 Architecture

Le modèle dynamique estimé est construit en ajustant le poids des connexions  $\omega_{ji}$  et les biais  $b$ , via la fonction d'apprentissage. Le réseau NARX peut être représenté en adoptant l'une des deux architectures suivantes (Topus, 2010) :

a) Mode série parallèle:

Le réseau est alimenté par les vraies valeurs de sorties (figure 11) et le système est défini ainsi :

$$\hat{y}(k+1) = f(y(k), y(k-1), \dots, y(k-n_y), u(k+1), u(k), \dots, u(k-n_u)) \quad (21)$$

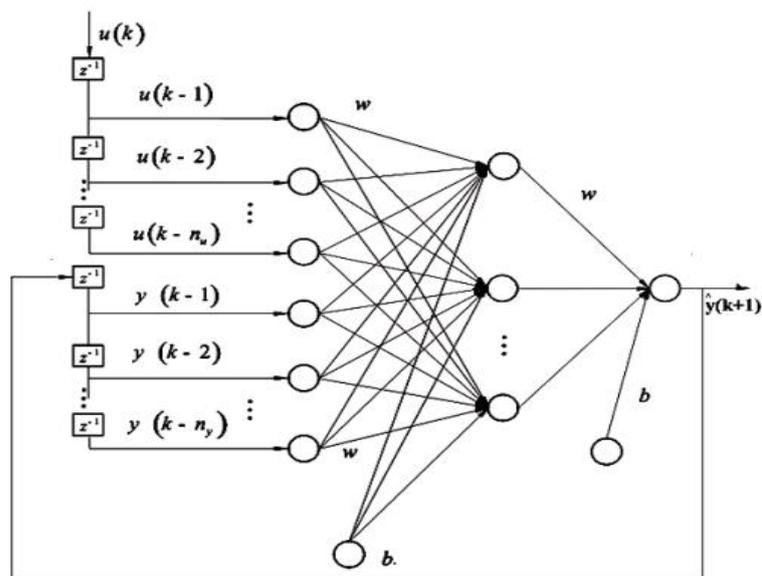


**Figure 11** : Architecture NARX, Séries-Parallèles (Topus, 2010).

b) Mode parallèle:

Dans ce cas, les sorties estimées sont ré-envoyées et incluses dans l'entrée de la structure (figure 12) qui est définie comme suit :

$$\hat{y}(k+1) = f(\hat{y}(k), \hat{y}(k-1), \dots, \hat{y}(k-n_y), u(k), u(k-1), \dots, u(k-n_u)) \quad (22)$$



**Figure 12**: Architecture NARX en mode parallèle (Topus,2010).

Alors que le réseau de type NARX-SP est utilisé en phase d'apprentissage, le réseau de type NARX -P est utilisé dans la phase de test.

## 7. Apprentissage d'un réseau de neurone

L'apprentissage est une étape très importante du développement d'un réseau de neurones durant laquelle le comportement du réseau est modifié itérativement jusqu'à l'obtention du comportement désiré, et ce par l'ajustement des poids (connexion ou synapse) des neurones à une source d'information bien définie (Hebb, 1949; Grossberg, 1982; Rumelhart et al. 1986). L'apprentissage consiste également à extraire des conformités parmi les données utilisées pour l'entraînement du réseau mais l'objectif essentiel de l'apprentissage est la résolution du problème par la prévision, la classification... etc. (Weiss et Kulikowski, 1991). Pour un RNA, l'apprentissage peut être regardé également comme étant le processus de la mise à jour des poids (connexion) au sein du réseau dans le but d'ajuster la réponse du réseau à l'expérience et aux exemples. Il existe plusieurs types de règles d'apprentissage qui peuvent être rassemblées en deux catégories qui sont l'apprentissage supervisé et l'apprentissage non supervisé.

### 7.1 Apprentissage supervisé

Les techniques de l'apprentissage supervisé sont caractérisées par la présence de données d'entrée et de données de sortie pour la construction du modèle et chaque individu se présente sous la forme du couple (input-output). L'algorithme fait associer la variable de sortie de chaque individu avec sa ou ses valeurs d'entrée et fait entraîner le modèle à réaliser des prédictions fidèles sur les données futures (Murphy, 2012).

L'échantillon  $\mathcal{D}$  de taille  $N$  impliqué dans l'apprentissage du modèle se présente comme suit :

$$\mathcal{D} = \{(X_i, Y_i)\}_{i=1}^N$$

Tel que :

$X_i = (x_{i1}, x_{i2}, \dots, x_{im})$  : vecteur d'entrée de l'individu  $i$  et  $m$  le nombre de variables

$Y_i$  = variable à prédire.

Après une étape d'entraînement, le modèle doit être testé avec un second échantillon de données où, dans cette étape, il reçoit seulement les variables d'entrée et fait des prédictions sur les sorties. Ces prédictions sont ensuite comparées avec les résultats réels pour mesurer la performance du modèle. Plus les prédictions du modèle sont proches des valeurs réelles plus le système est performant.

Ces techniques sont appelées ainsi car les données de sortie permettent à l'algorithme de comparer et de se corriger au fur et à mesure. En d'autres termes, elles vont « guider » l'algorithme vers les résultats qui doivent en ressortir.

Lorsque la valeur de sortie est une variable continue on parle d'une *prédiction*. Par exemple, anticiper la température dans une période de l'année est un problème de prédiction.

Par contre, si l'output est une variable discrète alors il s'agit d'une *classification* (exemple : classifier les clients dans deux catégories Fiables/Non-Fiables ou classifier les cellules cancéreuses en Bénigne/Maline) (Youcef, 2014).

Les algorithmes qui sont utilisés pour résoudre les problèmes de classification incluent notamment les machines à vecteurs de support (SVM), la classification naïve bayésienne, la méthode des k plus proches voisins, ... (Mathworks, 2018).

## 7.2 Apprentissage non supervisé

Pour l'apprentissage non-supervisé, seules les données d'entrée sont disponibles pour la construction du modèle, le but étant de trouver des informations cachées et de faire émerger des similitudes complexes à l'intérieur de ces données. *Le clustering*, qui consiste en la séparation des individus dans des groupes homogènes, est la principale technique de l'apprentissage non-supervisé (Youcef, 2014).

L'avantage de cette méthode qui est connue comme un outil d'exploration, est qu'elle peut apporter de nouvelles informations dont l'utilisateur ignorait l'existence. Un exemple pratique dans la segmentation des clients, l'algorithme peut renvoyer des groupes en fonction de certains paramètres qu'un humain n'aurait jamais pris en compte à cause des préjugés pré établis et à cause d'une inertie psychologique.

Contrairement à la première technique où le modèle peut comparer les outputs prédits avec les outputs réels, il est moins évident de mesurer la performance d'un modèle de clustering car les valeurs de sortie correctes ne sont pas connues au préalable. La performance dépend donc de si des informations intéressantes dans les données ont été révélées par les clusters ou non (Youcef, 2014).

## 8. Conclusion

Dans sa forme générale, un réseau de neurones est une machine destinée à reproduire la manière dont le cerveau exécute une tâche. En statistique, il est employé pour traiter des problèmes difficilement abordables par des méthodes linéaires ou par des équations algébriques. Un réseau de neurones est constitué d'un ensemble de processeurs élémentaires, les neurones, qui sont connectés les uns aux autres et qui sont capables d'échanger des informations au moyen des connexions qui les relient. Les connexions sont directionnelles et à chacune d'elles est associé un nombre réel appelé poids de la connexion, ou poids synaptique, qui représente une mesure du lien. Pour déterminer ces poids, on a recours à des méthodes dites d'apprentissage.

Les réseaux de neurones artificiels RNA font partie des systèmes adaptatifs non linéaires, qui sont constitués d'un grand nombre d'unités élémentaires. Ils traitent l'information d'une manière parallèle et distribuée. De telles propriétés leur confèrent bien la qualité de robustesse vis-à-vis des données incomplètes, incertaines et bruitées. Chose qui motive leur utilisation dans le cadre du problème inverse en géophysique. En vu de justifier l'applicabilité des RNA, une étude comparative va être réalisée. Pour ce faire, les réseaux de neurones vont être utilisés comme outils d'inversion de données issues d'une tomographie électrique ERT (electrical resistivity tomography) et comparés à la méthode classique.

# **Chapitre III : Interprétation des mesures de tomographie de résistivité électrique ERT à l'aide d'un réseau de neurones artificiel.**

## 1. Introduction

Ces dernières années, il y a eu une avancée majeure dans la formation et la mise en œuvre de réseaux neuronaux artificiels (RNA) dans de nombreux domaines différents. De nos jours, les réseaux neuronaux (profonds) sont à la pointe de la technologie dans de nombreuses tâches de classification et de prédiction. RNA est maintenant capable de traduire des phrases, de générer de nouvelles images et même de définir un protocole de communication. L'idée d'utiliser un RNA pour interpréter les propriétés géophysiques sous la surface n'est pas nouvelle et a été suggérée il y a presque trois décennies. Au cours des dernières années, RNA a été utilisé pour analyser les mesures gravimétriques, l'inversion de la forme d'ondes sismiques, les données de radar pénétrant dans le terrain, etc. Un travail préalable a été réalisé par S. Upendra, K. Singh, et R.K. Tiwari qui a démontré la capacité d'interpréter les mesures géophysiques par des réseaux neuronaux artificiels. L'hypothèse de la solution était un profil de terrain en couches et la solution déterminait les profondeurs et les résistivités de 5 à 6 couches.

Donc, afin de surmonter les difficultés d'interprétation des mesures ERT rencontrées lors de l'inversion conventionnelle, en fournissant une imagerie 2D du profil de résistivité du sous-terrain, nous allons présenter ici une autre approche différente basée sur l'apprentissage supervisé pour la formation d'un réseau de neurones artificiel (RNA) effectuée par les chercheurs (Naeh, Peleg, Furman, & Mannor, 2018).

## 2. Méthodologie

(Naeh, Peleg, Furman, & Mannor, 2018) ont réalisé un panneau électrique dans une formation linéaire, avec un espacement  $d$  entre les électrodes, en injectant un courant continu allant jusqu'à quelques Ampères entre les électrodes injectrices, et mesuré la différence de potentiel entre les électrodes de potentiel. Cette procédure est effectuée en utilisant toutes les électrodes pertinentes pour la campagne d'investigation.

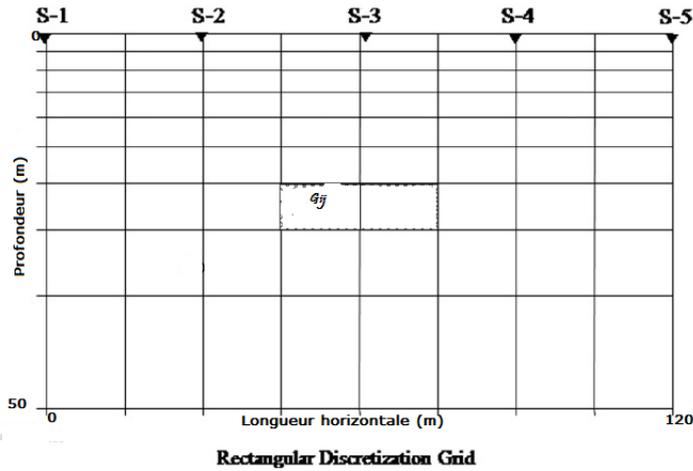
Dans un premier temps, une campagne réalisée par S. Upendra, K. Singh et R.K. Tiwari a démontré la capacité d'interpréter les mesures géophysiques par les réseaux de neurones artificiels, sachant que l'hypothèse portée sur un profil de terrain constitué de plusieurs couches, la solution a déterminé les profondeurs et les résistivités de 5 à 6 couches. Dans le cas présent le résultat est une image 2D complète qui ne nécessite pas de définir, à priori, un modèle de souterrain (comme dans la méthode d'inversion itérative). Une autre différence majeure réside dans les échelles, où dans la précédente étude, l'enquête avait des centaines de mètres de séparation d'électrodes et les profondeurs d'interprétation étaient de 1 km, au-delà de la résolution. Cependant, l'objectif visé dans ce travail en ERT est d'avoir une haute résolution de  $1m^2$  pour un terrain peu profond.

## 3. Modèle et simulations

Comme la structure du terrain rencontrée dans la nature est souvent hétérogène et, afin d'augmenter la validité de la méthode présentée, les chercheurs ont simulé un terrain non homogène en utilisant un simple dispositif Wenner-Schlumberger de 96 électrodes avec un

espacement de 1m et en supposant des valeurs de résistivité aléatoires avec des corrélations spatiales.

La structure interne du terrain simulé en 2D, est dimensionnée ainsi : 120 m de longueur pour 50 m de profondeur, on peut la présenter comme une matrice  $G \in R^{120 \times 50}$ . Où chaque élément  $G_{i,j}$ , représente la résistivité électrique d'une section transversale de  $1m^2$  en position horizontale  $i$  et en profondeur  $j$  (représente la résistivité simulée d'une maille ou bloc du modèle de terrain).



**Figure 13:** Modèle 2D utilisé pour créer l'image synthétique de résistivité et test requis pour la mise en œuvre du RNA (Réseau de Neurones Artificiel). Figure inspirée de U.K. Sing et R.K.Siwari (2013) et modifiée par K. Omraci et N. Younsi (2018).

Le profil de résistivité du terrain est un gradient linéaire perturbé par  $N=100$  valeurs aléatoires gaussiennes simulé à l'aide de l'équation 23:

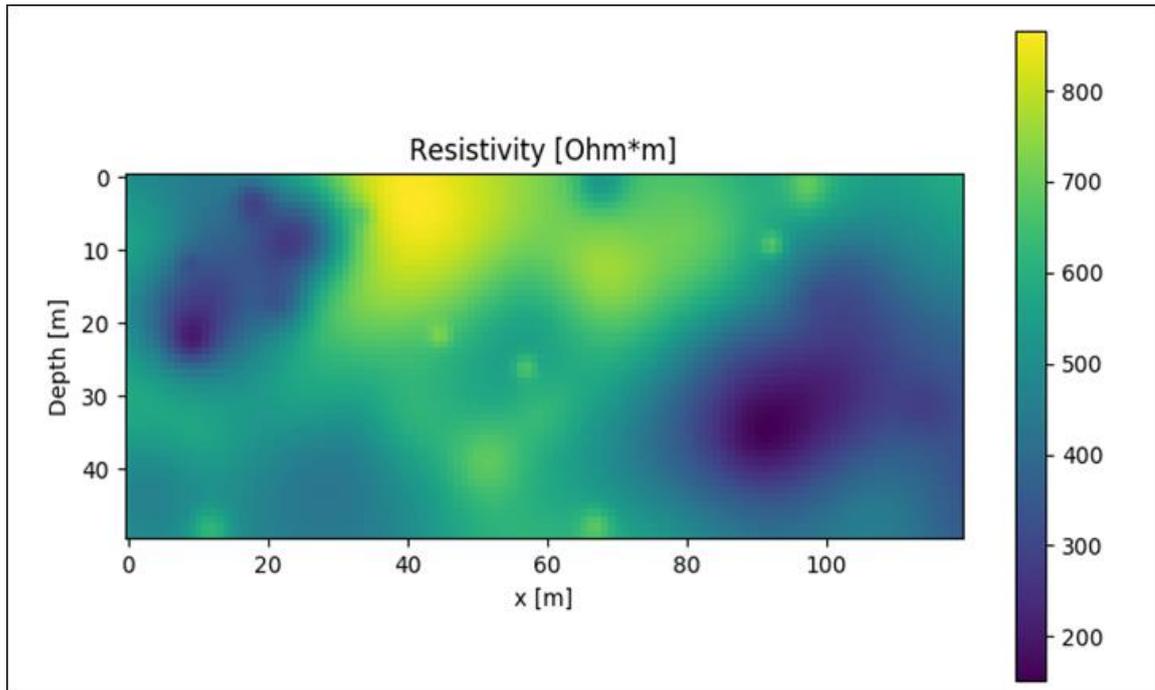
$$\rho = 600 - 0.2z + \sum_{n=1}^N A_n N(\mu_n, \sigma_n) \quad (23)$$

Où :

$Z$  : est mesurée en mètres,

$N(\mu, \sigma)$  : suit une loi gaussienne 2D centrée en  $\mu$  avec un écart type  $\sigma$ .

Les valeurs de  $A_n$  et de  $\sigma_n$  sont tirées des distributions uniformes  $U(-100, 100)$  et  $U(3, 30)$ , respectivement. Le centre des gaussiennes (valeurs moyennes),  $\mu_n$ , représente des variables aléatoires uniformément réparties dans le sous-terrain. La figure 14 illustre un terrain simulé typique.



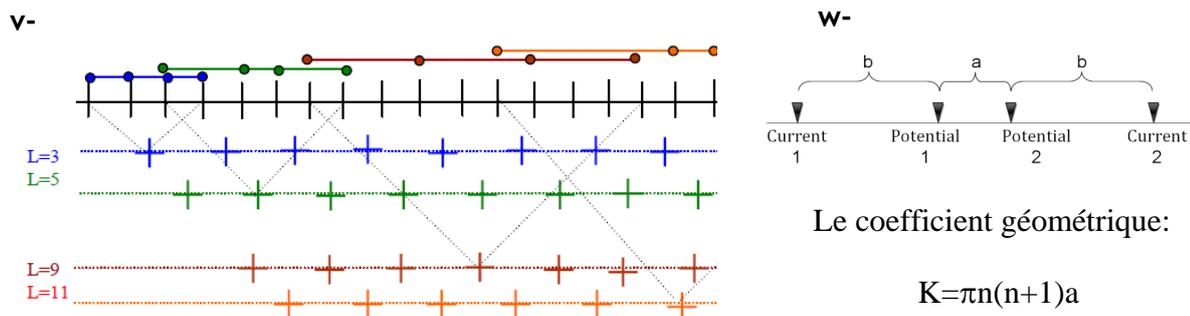
**Figure 14:** Un exemple représentatif du terrain utilisé pour l'entraînement et le test du réseau de neurones.

Il existe plusieurs configurations d'électrodes bien connus et couramment utilisés. Les principales différences entre les dispositifs sont : la profondeur d'investigation, la sensibilité de la matrice aux changements verticaux et horizontaux de la résistivité électrique, ainsi que la couverture horizontale des données et la force du signal.

Étant donné que les modèles de terrain utilisés dans les simulations ont des variations des valeurs de résistivité dans les directions horizontales et verticales, la configuration appliquée sur terrain était celle de Wenner-Schlumberger (WS) choisie pour sa sensibilité aux variations verticales et horizontales.

Le dispositif Wenner-Schlumberger (figure 15v) est caractérisé par les 3 paramètres suivants :

- 1)  $a$  : la distance entre les électrodes de potentiel.
- 2)  $b = na$  ( $n \in \mathbb{Z}$ ) : elle représente la distance entre l'électrode de potentiel et l'électrode injectrice la plus proche en terme de multiplication comme schématisé sur la figure 15w :
- 3) la position de l'électrode injectrice de courant se situant à gauche (courant 1 sur la figure 15w).



**Figure 15** : Pseudo section et schéma descriptif de la configuration Wenner-Schlumberger.  $b = na$ .

Afin de maintenir un nombre raisonnable de mesures,  $n$  a été choisi dans :  $\{1, 2\}$  et  $a$  dans  $\{1, 2, \dots, 33\}$  et pour chaque valeur de la paire  $(n, a)$ , nous avons utilisé toutes les positions possibles pour l'électrode 1 du courant.

En conséquence, ce protocole contient 1488 mesures de la résistivité électrique apparente. La simulation directe des mesures ERT a été effectuée à l'aide du logiciel **R2** (A. Binley, 2016) et le terrain simulé a été étendu loin pour minimiser les effets de limite (de bords) sur les résultats.

**NB** : R2 est programme qui offre une solution directe ou inverse pour le flux de courant 2D ou 3D dans un maillage quadrilatéral ou triangulaire. La solution inverse est basée sur une fonction objectif régularisée combinée à des moindres carrés pondérés. Les électrodes peuvent être placées à la surface du terrain ou dans des trous de forage.

#### 4. Description du réseau de neurones appliqué et de la base de données utilisée pour son apprentissage

Lors de l'interprétation des mesures issues d'une tomographie de résistivité électrique ERT (comme dans tout problème d'inversion), la question principale qui se pose est : quel est le support (ou le milieu) qui a provoqué les mesures acquises.

Ainsi, s'il existe assez d'exemples de milieux de terrain et des mesures de résistivité électrique, il est probable d'obtenir un aperçu de la relation inverse entre les données mesurées et le support qui les a provoqué. Cette approche pour l'entraînement d'un réseau de neurones est appelée « apprentissage supervisé ».

**NB**: Apprentissage supervisé = est une tâche d'apprentissage automatique consistant à apprendre une fonction de prédiction à partir d'exemples annotés.

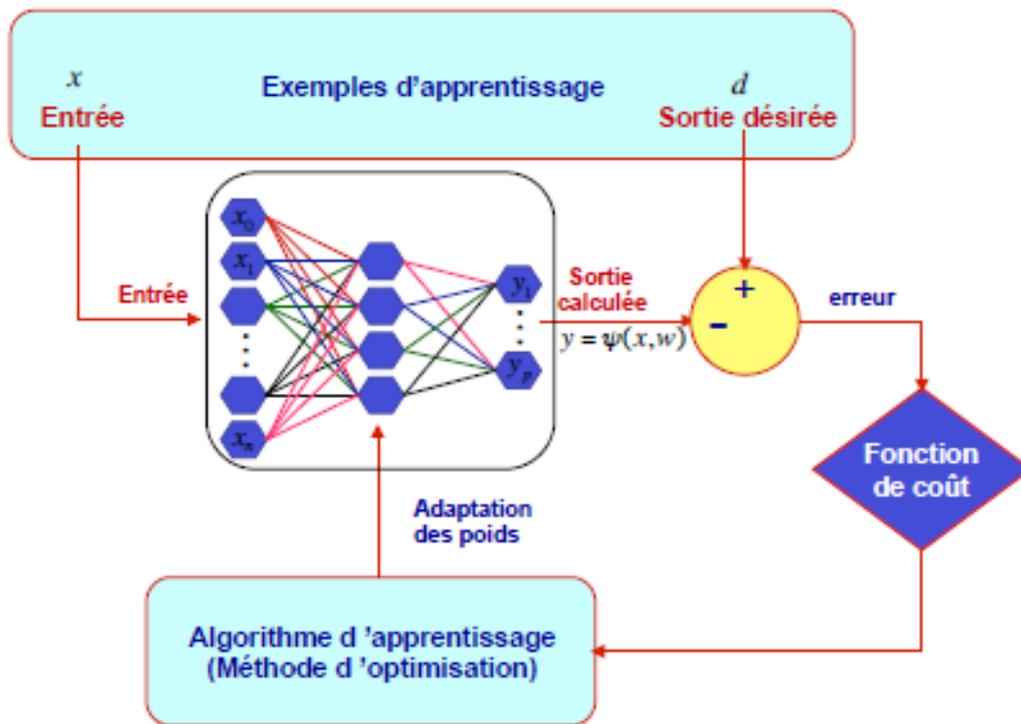
Les chercheurs (Naeh, Peleg, Furman, & Mannor, 2018) ont simulé une base de données massive de 15000 exemples de terrains simulés aléatoirement avec un profil de résistivité de terrain précédemment défini par l'équation (23). Sachant que la modélisation des 1488 mesures de résistivité électrique ont été réalisées pour chaque exemple, ce qui résulte une base de données constituée de paires  $(G_k, E_k)$ , où  $G_k$  représente la matrice du terrain, et  $E_k$  représente le vecteur de mesures de résistivité résultant.

Le jeu de donnée a été ensuite divisé aléatoirement en un ensemble d'apprentissage composé de 90% des exemples (training set), et les 1500 exemples restants ont été utilisés pour tester l'entraînement du réseau de neurones à former.

La donnée d'entrée du réseau de neurones est le vecteur  $E$  qui représente les mesures de résistivité électrique ERT constitué de 1488 mesures, et la valeur de sortie est une matrice de terrain de dimensions  $120 \times 50$ .

Pour avoir la solution la plus simple possible, une seule couche cachée a été utilisée composée de 250 neurones (cf. figure 16) avec une fonction d'activation ReLU (définie dans le second chapitre au point 5.3) :

$$\text{ReLU}(x) = \begin{cases} 0 & \text{si } x < 0 \\ x & \text{sinon} \end{cases} \quad (24)$$



**Figure 16:** Système d'apprentissage supervisé (Y. BENNANI, Cours RNA-U-Paris 13, 2014)

Le modèle global est donné par :

$$\tilde{\mathbf{g}} = \mathbf{b}^{(1)} + W_{6000 \times 250}^{(1)} \text{ReLU}(\mathbf{b}^{(0)} + W_{250 \times 1488}^{(0)} \mathbf{E}) \quad (25)$$

Où les lettres en caractères **gras** représentent les vecteurs.

- $W^{(0)}, W^{(1)}$  sont les matrices poids des couches correspondantes.
- $\mathbf{b}^{(0)}, \mathbf{b}^{(1)}$  sont les vecteurs biais correspondants.

- Le vecteur de sortie  $\tilde{g}$  est remodelé en une matrice  $\tilde{G}^{120 \times 50}$  qui est la base prédite par le modèle de réseau de neurones à former.

Pour l'apprentissage du réseau de neurones, une fonction perte (fonction coût) doit être définie. Dans de nombreuses applications, la distance Euclidienne entre le terrain prédit  $\tilde{G}$  et le terrain réel  $G$  est la plus utilisée. Cependant, dans ce scénario, il s'est avéré qu'une fonction de perte L1 se traduit par une meilleure performance et un temps de convergence plus court.

$$L_1 = \|\tilde{G} - G\|_1 = \sum_{i,j} \text{abs}(\tilde{G}_{i,j} - G_{i,j}) \quad (26)$$

Les poids et les biais ont été initialisés pour être distribués normalement avec un écart type de  $10^{-6}$  et pour l'apprentissage du RNA l'optimiseur appelé « Adam » a été utilisé. Chaque groupe comportait 50 exemples de paires (G, E), le terrain et les mesures issues de l'ERT, et le réseau de neurones a été entraîné pour 1000 époques.

## 5. Résultat du RNA et comparaison avec la méthode d'inversion classique

La manière habituelle d'examiner et d'analyser les résultats d'un processus d'inversion consiste à établir une comparaison entre les mesures ERT (E) obtenues à partir de la solution directe du terrain réel et les mesures d'ERT ( $\tilde{E}$ ) du terrain prédit et à calculer une fonction de perte  $L(E, \tilde{E})$ . Lorsque la fonction de perte  $L(E, \tilde{E})$  est suffisamment petite, le processus d'inversion s'arrête. Cependant, cela ne garantit pas que le terrain prédit  $\tilde{G}$  soit similaire au terrain réel  $G$ . Dans le scénario de mesures réelles sur le terrain, c'est le meilleur que l'on puisse faire avec un processus d'inversion. Cependant, en utilisant des données simulées, il est possible de faire une comparaison entre le véritable objectif du processus d'inversion - le terrain prédit  $\tilde{G}$  et le terrain réel  $G$ .

Afin de comparer les résultats du réseau de neurones à la méthode d'inversion classique le programme R2 a été utilisé pour prédire les valeurs de résistivité du terrain pour les 1500 exemples de test. Les paramètres utilisés pour le processus d'inversion étaient :

- Résistivité de base (background resistivity) égale à 600  $\Omega\text{m}$ .
- Tolérance égale à 0.1,
- Nombre d'itération maximal égal à 10.

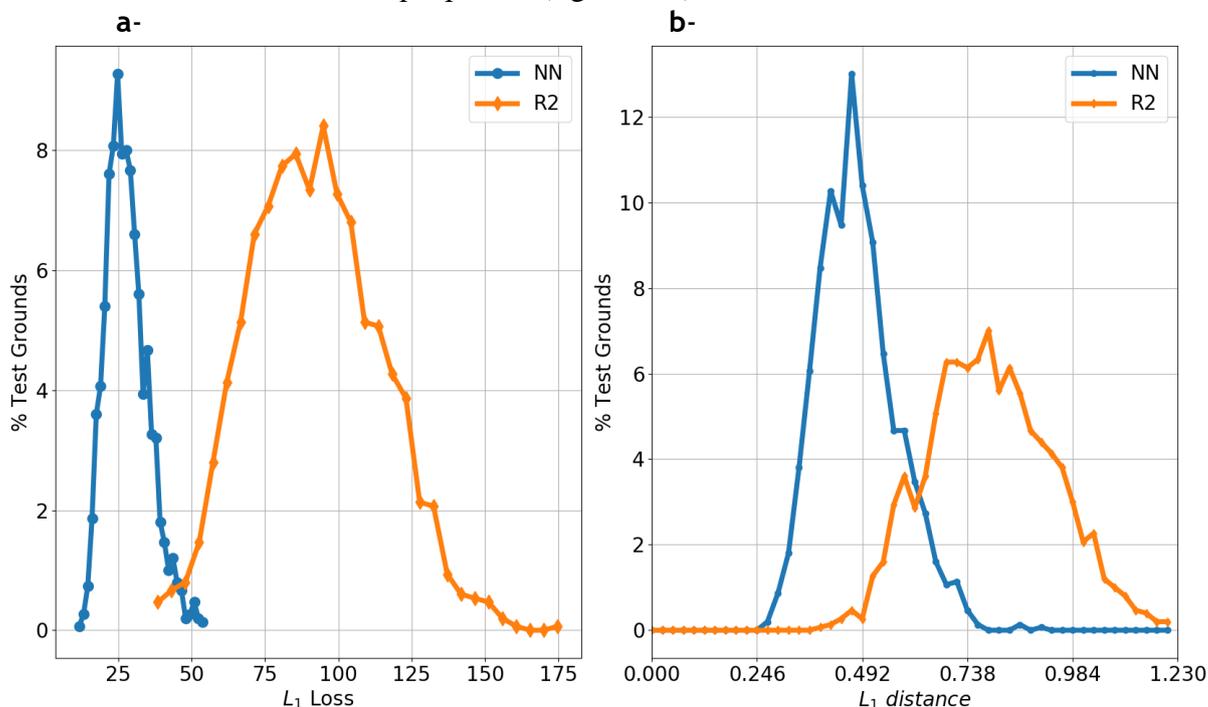
La sortie recherchée de l'inversion a la même taille et la même résolution que celle des prédictions du réseau de neurones. Et le terrain réel (de longueur de 120 m et de profondeur de 50 m) avec une résolution de 1 m<sup>2</sup>.

La comparaison entre la prédiction du réseau de neurones et la prédiction du processus d'inversion itérative se fait à l'aide de deux paramètres:

- Le premier est la fonction perte L1 (équation 26) de chaque prédiction faite par RNA et R2.

- Le second consiste à avoir une mesure de la similitude globale des valeurs entre la prédiction et le terrain réel. Cela peut être fait en utilisant la même distance  $L_1$  entre l'histogramme du terrain prédit et celui du réel. De cette façon, les caractéristiques légèrement décalées spatialement ne causeront pas de grandes difficultés.

La figure 17 illustre les histogrammes de la fonction perte  $L_1$  pour les prédictions faites par le RNA et R2. Il est très clair que, R2 a une très grande plage d'erreurs, allant de 50 à 150, le RNA admet des erreurs autour de 50 dans le pire des cas. Comme la fonction perte  $L_1$  est très sensible à la valeur de chaque pixel de  $1m^2$  dans le terrain, la perte  $L_1$  de l'histogramme est une comparaison des distributions. En comparant les valeurs de  $L_1$  pour les deux histogrammes par rapport au terrain réel (Figure 17a), nous trouvons que le RNA a une distribution plus petite des valeurs de perte  $L_1$  et que les valeurs sont plus proches de 0 que celles résultants de l'inversion classique par R2 (figure 17b).



**Figure 17** : **a-**: Histogrammes de la fonction perte  $L_1$  des prédictions du réseau de neurones et des prédictions de la méthode d'inversion classique en utilisant R2. Dans presque tous les cas de test, la fonction perte  $L_1$  des prédictions du RNA est inférieure à celle de la méthode d'inversion conventionnelle. **b** : Histogrammes de distance  $L_1$  entre les prédictions du RNA et les prédictions de l'inversion classique par R2 par rapport au terrain réel. Dans la plupart des cas, La distance  $L_1$  des histogrammes de la prédiction faite par RNA, est inférieure à la distance  $L_1$  des histogrammes de la méthode d'inversion par R2.

NB : NN représente Neural Network c'est-à-dire réseau de neurones artificiel RNA, et  $L_1$  Loss est la fonction perte  $L_1$ .

Une illustration de la reconstruction souterraine est présentée dans la figure 18. Dans chaque colonne de 3 images, celle du milieu est le terrain réel, celle supérieure est la prévision du réseau de neurones RNA et celle du bas est la prévision de l'inversion classique utilisant

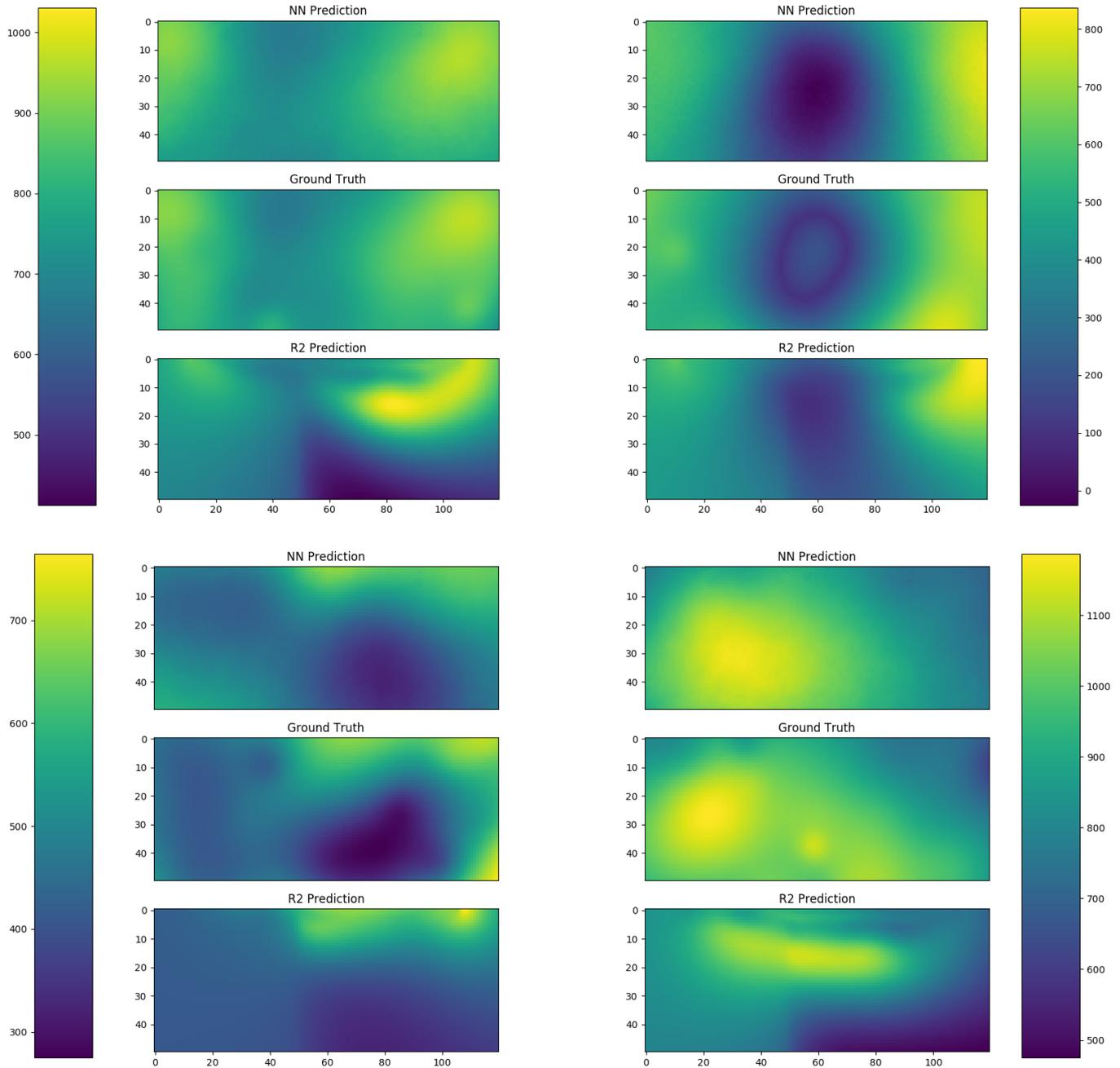
R2. Il est clair que dans tous les exemples, le RNA est plus performant que la méthode d'inversion itérative classique avec R2.

## 6. Conclusion

Les réseaux de neurones appliquent entre autres, le principe de l'approximation fonctionnelle. Si un réseau de neurones artificiels est capable d'apprendre un problème, ce dernier doit être défini comme une fonction contenant un ensemble de variables d'entrée et de sortie supportées par des exemples qu'illustrent la façon dont cette fonction devrait effectuer l'évaluation.

L'exemple étudié a démontré la capacité de reconstruire le profil de résistivité électrique en profondeur à partir d'une tomographie de résistivité électrique ERT en utilisant un réseau de neurones entraîné. Comme le montre la figure 17, les prédictions à l'aide du réseau de neurones sont plus précises que les prédictions issues de la méthode d'inversion traditionnelle. De plus, le réseau de neurones est capable de prédire la distribution de la résistivité électrique même à une profondeur de 40 à 50 mètres, qui est bien au delà de la profondeur maximale de la méthode d'inversion classique. Bien que la cartographie à ces profondeurs ne soit pas parfaite, les principales caractéristiques sont visibles et reconnaissables et, dans de nombreux cas, une légère déformation ou un léger décalage est dans la tolérance des solutions acceptables. Une autre caractéristique du réseau de neurones est que le processus de prédiction est évidemment plus rapide que les algorithmes d'inversion conventionnels et que les résultats sont obtenus en quelques millisecondes en ayant des mesures de tomographie ERT du terrain.

Chapitre 3 : Interprétation des mesures de tomographie de résistivité électrique ERT à l'aide d'un réseau de neurones artificiel.



**Figure 18** : Exemples illustrant le profil de résistivité souterraine prédit. Au sein de chaque triplet d'images, l'image du milieu est la distribution vraie des résistivités, la première est la prédiction utilisant le réseau de neurones RNA et la dernière est la prédiction de l'inversion classique effectuée par R2.

# Conclusion générale

L'appréhension de la structure interne de la Terre à partir de données collectées en surface demeure sans doute le souci constant de la plupart des géophysiciens. C'est dans ce contexte typique que s'est initialement développée la théorie du problème inverse, visant à interpréter, voire caractériser au mieux les problèmes géophysiques ainsi posés. Au cours de ces dernières années, plusieurs techniques ont été élaborées pour la résolution du problème inverse. La plupart de ces techniques cherchent à traiter son caractère mal posé qui pose en lui-même de sérieux problèmes de stabilité et de convergence.

Ce travail avait pour objectif, l'analyse d'une approche différente d'interprétation de données de tomographie électrique (ERT *Electrical Resistivity Tomography*), basée sur l'apprentissage supervisé pour la formation d'un réseau de neurones artificiel (RNA) afin de surmonter les difficultés d'interprétation des mesures ERT rencontrées lors de l'inversion conventionnelle, en fournissant une imagerie 2D du profil de résistivité du souterrain.

La tomographie électrique est une technique non-destructive, où le terrain peut être investigué sans le dégrader. Cette méthode permet d'obtenir une image de distribution des résistivités électriques en faisant circuler un courant électrique entre deux électrodes de courant et de mesurer la différence de potentiel entre deux autres électrodes, appelées électrodes de potentiel, à proximité des électrodes de courant. Puisque l'intensité du courant est connue et que la différence de potentiel est mesurée, il est alors possible de déterminer quelle est la résistivité apparente. Les mesures collectées permettent la construction d'une pseudo-section sous forme d'une carte qui présente la distribution des résistivités apparentes mesurées dans le milieu. Étant donné que le courant circule dans un chemin inconnu qui est déterminé par le profil de conductivité du terrain, l'interprétation n'est pas simple. De plus, le nombre d'inconnues dans le problème est plus grand que le nombre d'équations; par conséquent.

Ce problème est mal posé et doit être résolu à l'aide de méthodes d'inversion itératives, la manière habituelle d'examiner et d'analyser les résultats d'un processus d'inversion qui consiste à établir une comparaison entre les mesures ERT obtenues à partir de la solution directe du terrain réel et les mesures d'ERT du terrain prédit et à calculer une fonction de perte. Lorsque la fonction de perte est suffisamment petite, le processus d'inversion s'arrête. Cependant, cela ne garantit pas que le terrain prédit soit similaire au terrain réel. Dans le scénario de mesures réelles sur le terrain, c'est le meilleur que l'on puisse faire avec un processus d'inversion.

Cependant une comparaison entre les résultats du réseau de neurones à la méthode d'inversion classique a montré que les prédictions à l'aide du réseau de neurones sont plus précises que les prédictions issues de la méthode d'inversion traditionnelle. De plus, le réseau de neurones est capable de prédire la distribution de la résistivité électrique même à une

profondeur de 40 à 50 mètres, qui est bien au delà de la profondeur maximale de la méthode d'inversion classique. De plus le processus de prédiction du réseau de neurones est évidemment plus rapide que les algorithmes d'inversion conventionnels où les résultats sont obtenus en quelques millisecondes en ayant des mesures de tomographie ERT du terrain.

# Bibliographie

- Bennani, Y. (2014). *Cours RNA, Université de Paris 13* .  
[https://epat2014.sciencesconf.org/conference/epat2014/pages/EPAT\\_2014\\_ANN.pdf](https://epat2014.sciencesconf.org/conference/epat2014/pages/EPAT_2014_ANN.pdf)
- Danilo, P., Mandic, & Jonathan, A. (2001). *Chambers "Recurrent Neural Network for Prediction."*
- Fargier, Y. (2011). *éveloppement de l'imagerie de résistivité électrique pour la reconnaissance et la surveillance des ouvrages hydrauliques en terre. Science de la terre Ecole Centrale de Nantes. Nantes, France.*
- Hagan, M., Demuth, H., & Beale, M. (1996). *Neural Network design. PWS Publishing Company, Boston, Massachusetts.*
- Hornik, K. (1990). *Approximation Capabilities of FFNN.* Vienna.
- Laurne, F. (1993). *Fundamental of Neural Network.*
- Loke, M. (2017, Aout). *Geoelectrical Imaging 2-D and 3-D.* Consulté le Mars 1, 2018, sur GEOTOMO SOFTWARE / [www.geotomosoft.com](http://www.geotomosoft.com).
- Loke, M. (2002). *Tutorial : 2-D and 3-D electrical imaging surveys.*
- Marescot, L. (2004). *modélisation directe et inverse en prospection électrique sur les structures 3D complexe par la méthode des éléments finis. Génie Civil. université de Nantes et de Lausanne. Nantes.*
- Murphy, K. (2012). *Machine Learning a probabilistic perspective.*
- Naeh, I., Peleg, Y., Furman, A., & Mannor, S. (2018, Mai 11). *Interpreting Electrical-Resistivity Tomography measurements using Neural Network.* Haifa, Israel.
- Oudad, A. (2017, Mai 2). Récupéré sur [http://aoudad.perso.centrale-marseille.fr/TPE/ii\\_maths.php](http://aoudad.perso.centrale-marseille.fr/TPE/ii_maths.php).
- Sharma, S. (2017, Septembre 6). Récupéré sur <https://towardsdatascience.com/activation-functions-neural-networks-1cbd9f8d91d6>.
- Topus, V. (2014). *Hourly trafic flow prediction. Marmara University.*
- Tristan, S. (2016). *Réseaux de neurones récurrents, Université de Lyon 1? Département informatique.*
- Youcef, M. (2014). *Micro RNA Biology and Computational Analytics.* Hatfield : Humana Press.
- Zhang, G. (2004). *Neural Networks for Businnes Forecasting.* Georgia: Idea Publishing.

