

**REPUBLIQUE ALGERIENNE DEMOCRATIQUE ET POPULAIRE
MINISTERE DE L 'ENSEIGNEMENT SUPERIEUR
ET DE LA RECHERCHE SCIENTIFIQUE**



N° d'ordre :.....
Série :.....

**Université Mohammed Seddik
BENYAHIA - Jijel
Faculté des Sciences et de la Technologie
Département : Electrotechnique**

Thèse Présentée
Pour l'obtention du Diplôme de Doctorat en Sciences
Spécialité : **Electrotechnique**

Par :

SADOU Hakim

Thème :

**Caractérisation micro-ondes pour le contrôle non-destructif :
modélisation et inversion**

Soutenue le : **15/07/2018**, Devant le Jury composé de :

Mr. MEKIDECHE Mohammed Rachid	Professeur	U.M.S.B Jijel	Président
Mr. HACIB Tarik	Professeur	U.M.S.B Jijel	Rapporteur
Mr. NAIT –Said Nasreddine	Professeur	U. Batna 2	Examineur
Mme. LOUAI Zohra Fatma	Professeur	U. Batna 2	Examineur
Mr. ALLAG Hicham	Professeur	U. M.S.B Jijel	Examineur
Mr. HASSAINE Said	MCA	U. Tiaret	Examineur

Remerciements

Tout d'abord, je remercie ALLAH Le Tout Puissant qui m'a accordé la volonté et la patience pour réaliser cette thèse "El hamdo lillah".

Je voudrais exprimer ma profonde reconnaissance à mon directeur de thèse, monsieur Tarik HACIB professeur à l'université de Jijel. Je le remercie pour tous ses précieux conseils, sa grande disponibilité, sa gentillesse mais aussi ses vastes connaissances scientifiques.

Je voudrais plus particulièrement adresser mes remerciements à messieurs Yann Le Bihan, professeur des universités à l'IUT de Cachan (Paris), Olivier Meyer, Maître de Conférences à l'UPMC-Paris 6 et Hulusi AÇIKGÖZ, HDR à l'université KTO Karatay (Konya Turquie). Premièrement, je remercie monsieur Yann Le Bihan pour ses idées, conseils, suivie et corrections d'articles. Je le remercie aussi de m'avoir accueilli pour un stage au laboratoire GeePs. Deuxièmement, je remercie monsieur Olivier Meyer comme étant le concepteur et le réalisateur de la partie expérimentale. Troisièmement, je remercie monsieur Hulusi AÇIKGÖZ pour ses idées, conseils et corrections d'articles. Je le remercie aussi de m'avoir accueilli pour un stage à l'université KTO et de m'avoir fourni les données expérimentales.

Mes remerciements vont également aux membres du jury pour la caution qu'ils ont bien voulu apporter à ce travail. J'adresse mes remerciements à :

Mohamed Rachid MEKIDECHE, professeur à l'université de Jijel. Je le remercie particulièrement d'avoir accepté de présider le jury de soutenance.

Hicham ALLAG, professeur à l'université de Jijel. Je le remercie de m'avoir honoré en acceptant d'être examinateur.

Nassreddine NAIT-SAID, professeur à l'université de Batna 2. Je le remercie de m'avoir honoré en acceptant d'être examinateur.

Fatma Zohra LOUAI, professeur à l'université de Batna 2. Je la remercie de m'avoir honoré en acceptant d'être examinatrice.

Said Hassaine, Maître de conférences à l'université de Tiaret. Je le remercie de m'avoir honoré en acceptant d'être examinateur.

Je voudrais surtout ne pas oublier mes amis et mes collègues doctorants qui ont su m'épauler aux moments difficiles de mes études supérieures et plus particulièrement pendant ces années de thèse. Qu'ils sachent à quel point je leur suis reconnaissant.

Les autres personnes qui ont toujours été présentes à mes côtés, ce sont mes parents car au delà de la thèse, c'est tout au long de mes études qu'ils m'ont encouragé, soutenu et épaulé. Si j'en suis là aujourd'hui c'est principalement grâce à eux, Merci...

Table des matières

Introduction générale	1
Chapitre I : Généralités sur les diélectriques et les différentes techniques de caractérisation micro-ondes	5
I.1. Introduction.....	5
I.2. Notions sur les micro-ondes.....	5
I.3. Permittivité complexe et pertes diélectriques	6
I.4. Phénomènes de polarisation.....	8
I.4.1. Polarisation dipolaire	8
I.4.2. Polarisation ionique	12
I.4.3. La polarisation électronique	12
I.5. Phénomène de relaxation	13
I.6. Techniques de caractérisation	14
I.6.1. Technique de mesure en cavité résonnante	14
I.6.2. Technique de mesure en espace libre	15
I.6.3. Techniques utilisant des structures de propagation	16
<i>I.6.3.1. Technique en transmission-réflexion</i>	16
<i>I.6.3.2. Ligne ou guide d'onde court-circuité</i>	17
<i>I.6.3.3. Technique de mesure avec contact</i>	18
I.7. Résolution des problèmes direct et inverse.....	19
Conclusion.....	22
Chapitre II : Modélisation électromagnétique et construction des bases de données	23
II.1. Introduction	23
II.2. Rappel sur le champ électromagnétique et les équations de Maxwell	23
II.3. Méthodes de résolution.....	25
II.3.1. Solution analytique.....	25
II.3.2. Solution numérique	26
<i>II.3.2.1. Méthode des différences finies dans le domaine temporel</i>	26

II.3.2.2. Méthode des éléments de frontières	26
II.3.2.3. Méthode des éléments finis	26
II.4. Cellule de mesure SuperMit	29
II.4.1. Modélisation analytique	31
II.4.2. Inversion itérative.....	32
II.4.3. Modélisation numérique par FEM sous Ansys	33
II.4.4. Construction des bases de données	34
II.5. Cellule de mesure SuperPol.....	36
II.5.1. Solution analytique.....	37
II.5.2. Modélisation numérique par FEM sous Ansys	37
Conclusion.....	39
Chapitre III : Modèle Neuro-Flou pour la résolution du problème inverse.....	41
III.1. Introduction	41
III.2. Réseau de Neurones Artificiels	41
III.3. Logique floue	44
III.4. Système d'inférence floue.....	48
III.5. System Neuro-Flou	50
III.6. ANFIS	52
III.7. Stratégie d'apprentissage	54
III.7.1. Passe avant	54
III.7.2. Passe arrière	56
III.8. Implémentation d'ANFIS	59
III.9. Résultats et discussions	60
III.9.1. Application sur les données de simulation.....	60
III.9.1.1. Entrées : f , G et B	60
III.9.1.2. Sélection des entrées optimales	62
III.9.2. Application sur les données expérimentales	67
Conclusion.....	71
Chapitre IV : Modèles de régression multivariés pour la résolution du problème inverse ...	72
IV.1. Introduction.....	72

IV.2. Régression multilinéaire	72
IV.2.1. Régression Non-linéaire Multiple.....	73
IV.2.2. Sélection des descripteurs optimaux.....	75
IV.2.2.1. Méthode de sélection descendante.....	75
IV.2.2.2. Méthode de sélection ascendante	75
IV.2.2.3. Méthode de sélection progressive.....	76
IV.2.3. Implémentation de la MNLR dans la sélection progressive.....	80
IV.2.4. Résultats et discussions.....	81
IV.2.4.1. Application sur les données de simulation	81
IV.2.4.2. Application sur les données expérimentales.....	86
IV.2.5. Comparaison entre MNLR et ANFIS	87
IV.3. Théorie de l'analyse en composantes principales	89
IV.4. Analyse par régression aux moindres carrés partiels	91
IV.4.1. PLSR Non-Linéaire	93
IV.4.1.1 Délinéarisation interne	93
IV.4.1.2. Délinéarisation externe.....	94
IV.4.2. Implémentation de la NLPLSR	96
IV.4.3. Résultats et discussions.....	97
IV.4.3.1. Application sur les données de la simulation	97
IV.4.3.2. Application sur les données expérimentales.....	100
IV.4.4. Comparaison entre NLPLSR et MLP	103
Conclusion.....	104
Conclusion et perspectives.....	106
Bibliographie.....	109

Liste des acronymes

ANN	Artificial Neural Network
SVM	Support Vector Machines
FEM	Finite Element Method
MNLR	Multiple Non-Linear Regression
NLPLSR	Non-Linear Partial Least Squares Regression
ANFIS	Adaptive Network based fuzzy inference system
MLR	Multiple Linear Regression
PCA	Principal Components Analysis
PLSR	Partial Least-Squares Regression
FDTD	Finite-Difference Time-Domain
PEC	Perfect Electric Conductor
PMC	Perfect Magnetic Conductor)
TEM	Transverse Électrique- Magnétique
TM _{on}	Transverses Magnétiques
TE ₁₁	Transverse Électrique
GeePs	Laboratoire de génie électrique et électronique de Paris
FIS	Fuzzy Inference System
GD	gradient descent
LSE	Least Squares Estimator
FFNN	Feed-Forward Neural network
<i>RMSE</i>	Root Mean Square Error
RBF-NN	Radial Basis Function Neural Network)
MLP	Multi Layer Perceptron
PEEK	PolyEtherEtherKetone
PSO	Particle Swarm Optimization
NIPALS	Nonlinear Iterative Partial Least Squares
SVD	Singular Value Decomposition
SIMPLS	Statistically Inspired Modification of PLS
SMART	Smooth Multiple Additive Regression Technique

Liste des symboles

μ	moment dipolaire
ξ	distance entre deux charge de signes opposés
P	polarisation
χ	susceptibilité diélectrique
ε_0	permittivité du vide
E	champ électrique
D	déplacement électrique
ε'	constante diélectrique
ε''	terme des pertes
τ_D	temps de relaxation de Debye
ε_S	constante diélectrique relative statique
ε_∞	permittivité diélectrique en très hautes fréquences
f	fréquence
ω	fréquence angulaire
τ_{CC}	temps de relaxation de Cole-Cole
τ_{CD}	temps de relaxation de Cole-Davidson
τ_{HN}	temps de relaxation de Havriliak et Negami
σ	conductivité statique du milieu
Q_i	facteur de qualité
f_i	fréquence de résonance
B	induction magnétique
H	champ magnétique
J	densité de courant
t	temps
V	fonction test
W	fonction d'interpolation
λ	fonctions barycentriques
e	circulation du champ électrique
Y_d	admittance à la discontinuité
G	partie réelle de l'admittance
B	partie imaginaire de l'admittance

Γ	coefficient de réflexion
φ	phase du coefficient de réflexion
d	longueur du guide d'ondes coaxial
μ_A	fonction d'appartenance
U	univers de discours
O_i^1	sortie de première couche d'ANFIS
O_i^2	sortie de deuxième couche d'ANFIS
O_i^3	sortie de troisième couche d'ANFIS
O_i^4	sortie de quatrième couche d'ANFIS
O_i^5	sortie de cinquième couche d'ANFIS
η	taux d'apprentissage
k	pas
f_0	fréquence de relaxation
X	matrice prédicteurs ou de descripteurs
Y	matrice de réponse ou de sortie
X^T	transposé de la matrice X
F	paramètre statistique partiel
R^2	coefficient de détermination
R	coefficient de corrélation
T	matrice des scores
P	matrice des charges pour les entrées
Q	matrices des charges pour les sorties
E	matrice des résidus pour les entrées
F	matrices des résidus pour les sorties
R	matrice des poids

Introduction générale

De nos jours, la mesure des propriétés diélectriques et magnétiques des solides et des liquides dissipatifs devient de plus en plus intéressante. Habituellement, les mesures diélectriques s'effectuent dans le cas des tissus biologiques pour les recherches sur le cancer, des matériaux de construction, des matériaux à indice négatif, du blindage électromagnétique et de la propagation de signaux sans fil.

Comme la mesure directe des propriétés physiques d'un matériau est impossible, les scientifiques se tournent vers la mesure d'autres entités qui peuvent constituer une image directe de ces propriétés. La procédure est d'exciter le matériau sous teste par une source d'énergie et d'étudier par la suite sa réponse, qui dépend sûrement de ses propriétés physiques. Il existe différentes techniques de caractérisation des matériaux : les ultrasons, les rayons X, les courants de Foucault, les rayons gamma et les micro-ondes qui sont parmi les techniques les plus utilisées. En effet, la faculté des micro-ondes à pénétrer dans les matériaux diélectriques et à être sensibles à leur permittivité fait d'eux une méthode à part entière pour ce type de milieu.

Dans le domaine de la caractérisation micro-ondes, de nombreuses techniques ont été développées pour mesurer la perméabilité magnétique et la permittivité diélectrique. Ces techniques incluent les méthodes d'espace libre, les techniques à sonde coaxiale ouverte, les cavités résonantes, les techniques à résonateur diélectrique et les techniques de ligne de transmission [Bake 04].

Un examen approfondi des travaux de recherche dans le domaine de la caractérisation micro-ondes des matériaux diélectriques révèle que la méthode de réflexion avec une ligne de transmission semble être une approche efficace et pratique, en particulier lorsqu'elle est utilisée avec une ligne coaxiale à extrémité ouverte [Zhen 91]. Cette technique est basée sur l'insertion d'un échantillon du matériau à tester au niveau de l'extrémité ouverte du guide d'onde coaxial, ensuite la permittivité du matériau sera extraite à partir de la mesure du coefficient de réflexion ou de l'admittance de la cellule de mesure.

D'après la recherche bibliographique effectuée, on a pu trouver quelques travaux sur la caractérisation micro-ondes des matériaux diélectriques en utilisant une ligne coaxiale à extrémité ouverte. Principalement, nous mentionnons ici les travaux de Belhadj et al. [Belh 86, 90, 97, 98] où ils ont introduit la méthode de raccordement modal pour résoudre analytiquement le problème direct, dans le cas d'échantillons liquides et solides à une couche, deux couches et trois couches

(Téflon , Aluminium, eau, Ethanol, ..., etc.), tandis que le problème inverse a été résolu par une procédure d'inversion itérative. Dans les références [Kaat 89, 96, 05, 10], l'auteur a étudié le changement des propriétés diélectriques de l'eau avec la fréquence et la température, examiné l'interaction des ondes électromagnétiques avec l'eau et les solutions aqueuses et fourni des recherches étendues sur les techniques de la spectrométrie diélectrique. Au cours des dernières années, Wagner et al. [Wagn 07, 11, 14] ont concentré leurs recherches sur la spectroscopie micro-ondes large bande du sol où plusieurs liquides comme le Méthanol, l'Éthanol, le Toluène, l'eau déionisée et l'eau ordinaire (Omniprésente dans le sol) ont été caractérisés comme liquides étalons. En outre, Komarov et al. [Koma 16] ont utilisé une ligne coaxiale à extrémité ouverte pour la caractérisation micro-onde de la glace de mer, tandis que dans la référence [Szyp 17] le spectre de la permittivité diélectrique complexe, l'humidité et la salinité des matériaux poreux ont été étudiés.

Souvent les chercheurs utilisent un traitement de signal pour remonter aux propriétés du matériau en question, ou parfois ils essaient d'établir un modèle semi-analytique qui donne directement les caractéristiques du matériau [Wagn 14]. Dans notre étude, la technique d'inversion directe utilisant un modèle comportemental paramétrique construit au moyen d'une base de données est adoptée. Au début, les problèmes inverses ont été résolus par des méthodes itératives qui impliquent des solutions analytiques rapides du problème direct. Cependant, les phénomènes physiques sont rarement résolus par des méthodes analytiques, c'est pourquoi les solutions numériques sont devenues nécessaires et par conséquent l'inversion itérative peut s'avérer complexe et coûteuse en temps de calcul.

Il y a quelques années, au sein du laboratoire GeePs (Laboratoire de Génie Électrique et Électronique de Paris), de nombreux travaux sur la caractérisation diélectrique utilisant la sonde coaxiale ont été réalisés, avec la nouvelle idée de l'intelligence artificielle, pour résoudre le problème inverse. Les chercheurs ont introduit divers types de réseaux de neurones artificiels (ANN, pour Artificial Neural Network) [Acik 07, 08] ainsi que diverses variantes de la méthode de machines à vecteur de support (SVM, pour Support Vector Machines) [Haci 10, 11 ; Hacib 10] comme méthodes d'inversion directe pour l'évaluation large bande de la permittivité complexe. Les résultats ont été validés par l'inversion itérative qui implique une solution analytique du problème direct. Dans ce sens, cette thèse propose d'autres outils d'inversion qui peuvent être plus efficaces, tout en résolvant le problème direct à l'aide d'une méthode numérique qui est celle des éléments finis (FEM, pour Finite Element Method).

Ce manuscrit de thèse est composé de quatre chapitres :

Le premier chapitre présente les mécanismes mis en jeu au sein d'un diélectrique excité par une micro-onde. Premièrement, on va voir les différents phénomènes de polarisation et de relaxation pour arriver à la notion de la permittivité complexe. Ensuite, les différentes techniques de caractérisation par micro-ondes vont être discutées. Enfin, la démarche adoptée pour résoudre le problème direct et inverse va être présentée, ainsi que d'autres démarches utilisées auparavant.

Le deuxième chapitre donne un aperçu sur le problème direct. Premièrement, on commence par un rappel sur les équations de Maxwell ainsi que les méthodes habituellement utilisées pour les résoudre. Ensuite, deux cellules de mesure (SuperMit et SuperPol), qui s'inscrivent dans la catégorie des techniques avec contact, sont minutieusement présentées. De plus, on décrit la solution analytique et la modélisation numérique (Logiciel éléments finis Ansys) pour chaque cellule, ainsi que la procédure de création des bases de données nécessaires pour entraîner les outils d'inversion.

Les deux chapitres restants sont consacrés, entièrement, à la résolution du problème inverse. Précisément, deux modèles de l'analyse multivariée (MNL, pour Multiple Non-Linear Regression et NLPLSR pour Non-Linear Partial Least Squares Regression) et un modèle hybride neuro-floue (ANFIS, pour Adaptive Network based fuzzy inference system) sont investis.

Dans le troisième chapitre, après des rappels sur les réseaux de neurones, la logique floue et les systèmes d'inférence floue, l'ANFIS qui est une variante des réseaux neuro-flous est décrit. Ensuite, le modèle ANFIS adopté dans notre étude (architecture, algorithme d'apprentissage, sélection des entrées optimales et mise en œuvre) est proposé. Enfin, nous terminons ce chapitre par une présentation de la procédure retenue pour l'inversion des mesures effectuées sur la cellule SuperMit.

L'application des deux modèles de l'analyse multivariée (MNL et NLPLSR) fait l'objet du quatrième chapitre. Tout d'abord, on donne un aperçu sur la régression multilinéaire (MLR pour Multiple Linear Regression) ainsi que quelques notions sur l'analyse par composantes principales (PCA, pour Principal Components Analysis) qui forment la base de la régression aux moindres carrés partiels (PLSR pour Partial Least-Squares Regression). Une étude étendue sur ces derniers est exposée dans ce chapitre. Contrairement à d'autres études qui comportent beaucoup de variables d'entrée, on cherche à créer mathématiquement de nouvelles variables pour concevoir des modèles de régression non-linéaires. Enfin, l'inversion des mesures effectuées sur les cellules SuperMit et SuperPol ainsi que des études comparatives entre les modèles développés sont présentées.

Ces travaux nous permettent de conclure en rappelant leurs apports et en indiquant les principales perspectives de recherche à développer.

Chapitre I : Généralités sur les diélectriques et les différentes techniques de caractérisation micro-ondes

I.1. Introduction

La propagation des micro-ondes dépend fortement des propriétés des matériaux. Une des propriétés les plus fondamentales d'un matériau diélectrique est sa permittivité complexe. C'est une mesure de la capacité du diélectrique à absorber et à stocker l'énergie électrique. C'est un paramètre incontournable pour l'étude des matériaux diélectriques. Des chercheurs ont tenté de retrouver théoriquement la permittivité diélectrique en proposant des modèles plus ou moins simplifiés. Parmi ces modèles, le plus connu est celui de Debye [Deby 29]. Depuis, pour s'approcher le plus possible de la réalité, d'autres modèles plus évolués comme ceux de Cole-Cole [Cole 41], Cole-Davidson [Davi 51] et Havriliak-Negami [Havr 67], ont été proposés.

Dans ce chapitre, on va étudier l'interaction des micro-ondes avec les diélectriques, les différentes techniques de caractérisation et les approches d'étude possibles.

I.2. Notions sur les micro-ondes

Le terme micro-ondes fait toujours référence aux signaux alternatifs de fréquence entre 300 MHz et 300 GHz, ce qui correspond à une longueur d'onde entre 1 m et 1 mm, respectivement. Les signaux qui ont une longueur d'onde de l'ordre du mm sont appelés ondes millimétriques. Les micro-ondes se propagent dans l'espace libre à la vitesse de la lumière mais elles perdent de vitesse en passant par des matériaux. Dans le spectre électromagnétique représenté par la figure I.1, la gamme de fréquences des micro-ondes se trouve entre les fréquences radar et l'infrarouge.

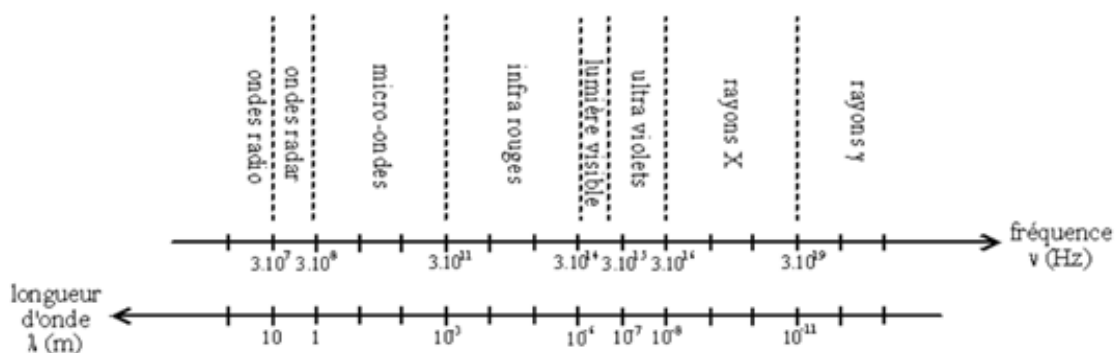


Figure I.1. Spectre électromagnétique.

Lorsque l'onde incidente passe d'un milieu 1 (Air) à un milieu 2 (Diélectrique), il y aura apparition d'une onde réfléchie dont l'amplitude est d'autant plus forte que l'écart de permittivité entre les deux milieux est élevé (Figure I.2). L'onde transmise se propage avec un affaiblissement de son amplitude dépendant du facteur de perte ϵ'' . En outre, on observe un changement de vitesse de propagation liée à la valeur de la permittivité du matériau ϵ' . Plus cette dernière est élevée plus la vitesse de l'onde dans le matériau est plus lent. Ces phénomènes d'interaction rendent les micro-ondes très adaptés pour les mesures diélectriques, contrairement aux matériaux conducteurs, où les micro-ondes ne peuvent être utilisées que pour étudier la surface extérieure (Mesures de topographie) à cause de la réflexion quasi-totale.

I.3. Permittivité complexe et pertes diélectriques

Un matériau est dit diélectrique lorsqu'il ne contient pas de charges libres dans sa structure et ne permet pas la circulation d'un courant électrique significatif sous l'influence d'un champ électrique extérieur, de ce point de vue les diélectriques sont des bons isolants électriques. Cependant, les diélectriques réels peuvent comporter des molécules polaires (Existence d'un moment dipolaire permanent) et des molécules plus ou moins facilement polarisables.

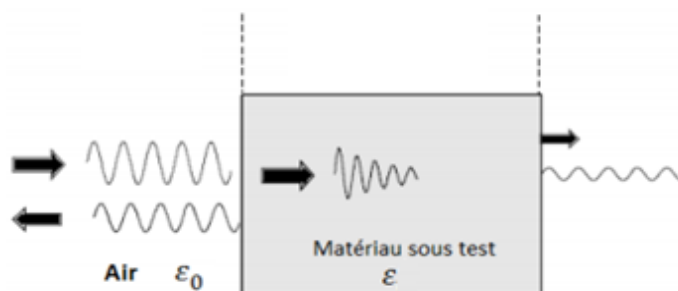


Figure I.2. Interaction micro-ondes-diélectrique.

En soumettant un matériau diélectrique à un champ électrique, plusieurs mécanismes de polarisations, s'opposant à l'action du champ, peuvent être observés et contribuent à sa permittivité globale. Le champ appliqué impose l'alignement des ions selon la direction du champ, met en rotation les molécules polaires et change le nuage électronique au niveau des atomes, et ainsi d'autres dipôles vont être créés dans le diélectrique. On définit le moment dipolaire entre deux charges de signe opposé (+ q et - q) éloignées de ξ par la relation :

$$\vec{\mu} = q \cdot \vec{\xi} \quad (\text{I.1})$$

On définit la polarisation \vec{P} induite dans le volume du matériau diélectrique comme la densité volumique de moments dipolaires. Sous l'hypothèse de linéarité et pour des matériaux isotropes, \vec{P} est directement liée au champ électrique appliqué \vec{E} par la relation :

$$\vec{P} = \varepsilon_0 \chi \vec{E} \quad (\text{I.2})$$

où ε_0 et χ sont respectivement la permittivité du vide et la susceptibilité diélectrique (Grandeur sans dimension).

Le déplacement électrique \vec{D} qui exprime la densité de flux électrique s'écrit :

$$\vec{D} = \varepsilon_0 \vec{E} + \vec{P} \quad (\text{I.3})$$

En remplaçant l'expression de \vec{P} dans \vec{D} on obtient :

$$\vec{D} = (1 + \chi) \varepsilon_0 \vec{E} \quad (\text{I.4})$$

Dans le cas général d'un milieu inhomogène et anisotrope, la permittivité $[\varepsilon]$ est un tenseur d'ordre trois. Elle représente le rapport entre le vecteur de déplacement électrique \vec{D} et le champ électrique appliqué au matériau \vec{E} : $\vec{D} = [\varepsilon] \vec{E}$. Elle se réduit à un scalaire, généralement complexe, ε dans le cas d'un milieu homogène et isotrope.

$$\varepsilon = (1 + \chi) \varepsilon_0 \quad (\text{I.5})$$

Dans la relation (I.5), $(1 + \chi)$ représente la permittivité relative ε_r du matériau.

En général les diélectriques réels sont imparfaits et présentent des pertes dus à la difficulté d'alignement que rencontrent les dipôles à cause des forces intermoléculaire et d'inertie qui assurent la cohésion de la matière et s'opposent à l'action du champ électrique extérieur, entraînant ainsi un échauffement thermique. Dans ce cas la polarisation n'arrive pas à suivre simultanément les variations du champ électrique appliqué, alors un déphasage est apparu entre ce dernier et le vecteur de déplacement électrique qui se traduit par la relation suivante :

$$\vec{D} = (\varepsilon' - j\varepsilon'') \vec{E} \quad (\text{I.6})$$

où ε' est la constante diélectrique qui représente la capacité d'un matériau à être polarisé par un champ électrique externe et ε'' est le terme d'absorption du matériau qui correspond aux pertes diélectriques dans le matériau, c'est-à-dire classiquement à sa capacité à transformer l'énergie électromagnétique en chaleur.

Un diélectrique est souvent caractérisé par sa tangente de pertes $\tan(\delta)$ qui représente les pertes dans un diélectrique et qui est définie par le rapport entre la partie imaginaire et réelle de la permittivité :

$$\tan(\delta) = \frac{\varepsilon''}{\varepsilon'} \quad (I.7)$$

avec δ est l'angle de pertes diélectriques.

I.4. Phénomènes de polarisation

Sous l'action d'un champ électrique, les atomes, les molécules et les ions se polarisent : les charges positives se déplacent dans le sens du champ et les charges négatives en sens inverse. Les barycentres des charges positives et négatives ne coïncident plus et il en résulte un dipôle induit. La polarisation totale se décompose en trois parties : dipolaire, ionique et électronique.

Dans ce qui suit, seulement les modèles qui décrivent le comportement de relaxation observé dans la gamme de fréquences qui va jusqu'à quelques gigahertz (Polarisation dipolaire) seront discutés.

I.4.1. Polarisation dipolaire

La polarisation dipolaire résulte de l'existence de moments électriques permanents, non nuls même sans l'application d'un champ électrique extérieur. On appelle souvent dipôles d'orientation les particules portant de tels moments permanents, puisque sous l'effet du champ électrique extérieur, les dipôles vont plus ou moins s'orienter suivant la direction du champ. Un milieu contenant des dipôles d'orientation est dit polaire. La figure I.3 illustre le mécanisme de la polarisation dipolaire.

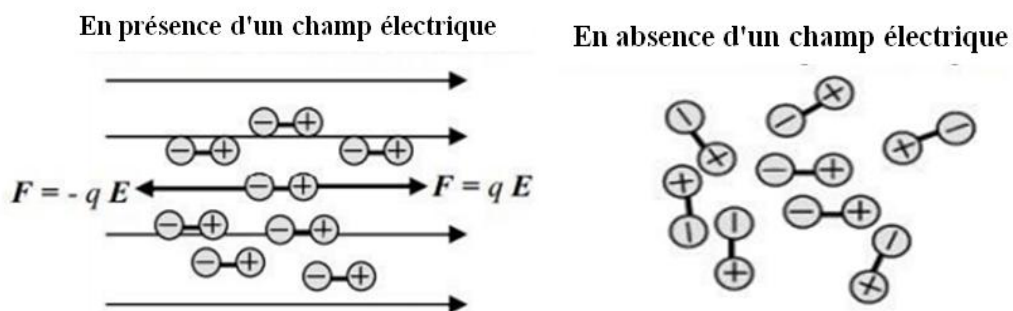


Figure I.3. Polarisation dipolaire ou polarisation d'orientation.

La formule exacte de la fonction de la réponse diélectrique a fait l'objet de nombreux travaux de recherche. Le premier model, appelé Debye [Debye 29], considère les molécules polaires non

interactives sous la forme la plus simple en négligent les processus de relaxation ionique et électronique. La relation entre la permittivité relative complexe ε_r et la fréquence f du champ électrique appliqué est de la forme suivante [Krem 03] :

$$\varepsilon_r(\omega) = \varepsilon_\infty + \frac{\Delta\varepsilon}{1 + j\omega\tau_D} \quad (\text{I.8})$$

En séparant les parties réelle et imaginaire, on obtient :

$$\varepsilon_r'(\omega) = \varepsilon_\infty + \frac{\Delta\varepsilon}{1 + \omega^2\tau_D^2} \quad \varepsilon_r''(\omega) = \frac{\omega\tau_D\Delta\varepsilon}{1 + \omega^2\tau_D^2} \quad (\text{I.9})$$

où : $\Delta\varepsilon = \varepsilon_s - \varepsilon_\infty$ est la force ou l'intensité de la relaxation diélectrique,

$\varepsilon_s = \lim_{\omega\tau_D \ll 1} \varepsilon_r'(\omega)$ est la constante diélectrique relative statique,

$\varepsilon_\infty = \lim_{\omega\tau_D \gg 1} \varepsilon_r'(\omega)$ est la permittivité diélectrique en très hautes fréquences.

Le temps de relaxation de Debye τ_D est lié à la position de pertes maximales par $\omega_p = 2\pi f_p = 1/\tau_D$ (Figure I.4). Le pic des pertes est symétrique avec une demi-largeur de 1.14 décades (À l'échèle logarithmique).

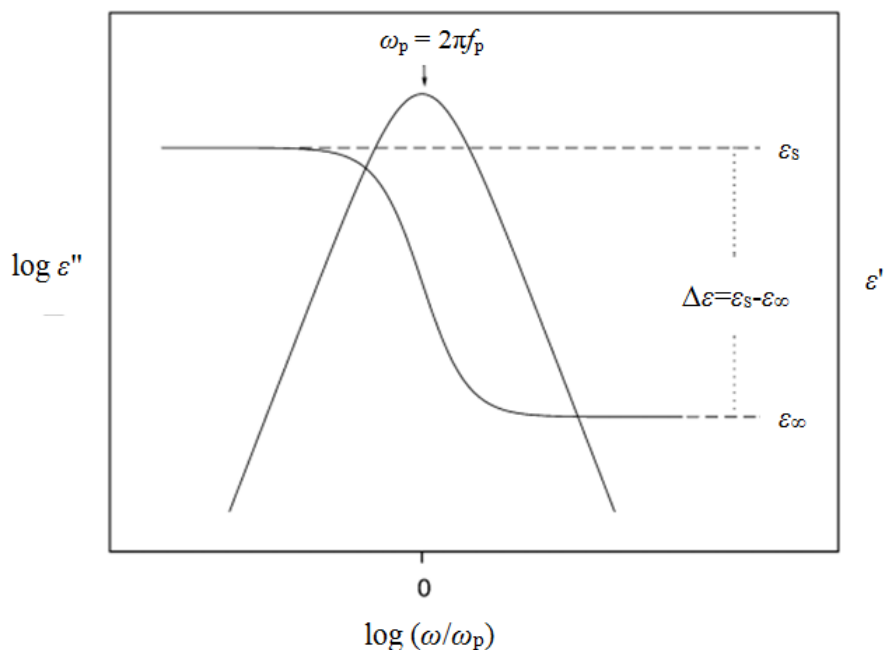


Figure I.4. Partie réel ε' et imaginaire ε'' de la fonction diélectrique complexe par rapport à la fréquence normalisée.

Le spectre diélectrique d'un matériau est souvent représenté par le diagramme de Cole-Cole (Figure I.5). C'est la représentation de $\varepsilon' - j\varepsilon''$ dans le plan complexe ($\varepsilon'' = f(\varepsilon')$) qui forme un demi-cercle de rayon $(\varepsilon_s - \varepsilon_\infty)/2$ dont le centre se situe sur l'axe des ε' en $(\varepsilon_s + \varepsilon_\infty)/2$. De gauche à droite, on part de $\varepsilon' = \varepsilon_s$ et $\varepsilon'' = 0$ (Très basse fréquence) pour arriver à $\varepsilon' = \varepsilon_\infty$ et $\varepsilon'' = 0$ (Très haute fréquence).

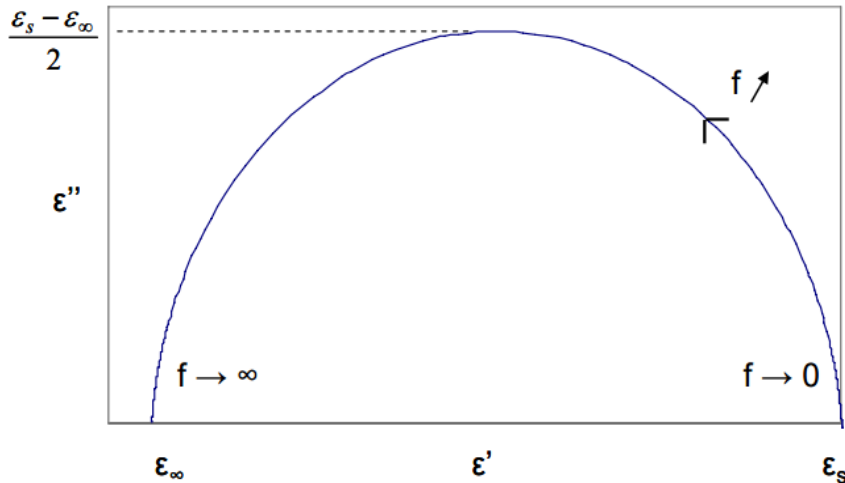


Figure I.5. Représentation en diagramme de Cole-Cole du modèle de Debye.

Dans la plupart des cas, la demi-largeur des pics de pertes mesurés est beaucoup plus large que prévu par l'équation I.8 (Jusqu'à six décades) et en outre leurs formes sont asymétriques avec une queue en hautes fréquences. C'est ce qu'on appelle un comportement de relaxation non-Debye (Ou non idéal). Dans la littérature, plusieurs fonctions de modèles empiriques, principalement des généralisations de la fonction de Debye, ont été développées et testées pour décrire des pics de perte étalés et/ou asymétriques.

Un étalement de la fonction diélectrique peut être décrit par la fonction de Cole-Cole :

$$\varepsilon_{CC}(\omega) = \varepsilon_\infty + \frac{\Delta\varepsilon}{1 + (j\omega\tau_{CC})^\beta} \quad (\text{I.10})$$

où $0 < \beta \leq 1$ conduit à un étalement symétrique de la fonction de relaxation par rapport à l'équation I.8. Pour $\beta = 1$, la fonction de Debye est obtenue. Le temps de relaxation de Cole-Cole τ_{CC} donne la position de perte maximale par $\omega_p = 2\pi f_p = 1/\tau_{CC}$.

De nombreux résultats expérimentaux, en particulier sur les liquides ou les matériaux vitreux à faible poids moléculaire, montrent que la fonction diélectrique complexe peut aussi avoir un étalement asymétrique qui peut être décrit par la fonction de Cole-Davidson :

$$\varepsilon_{CD}(\omega) = \varepsilon_\infty + \frac{\Delta\varepsilon}{(1 + j\omega\tau_{CD})^\gamma} \quad (\text{I.11})$$

Le paramètre γ ($0 < \gamma \leq 1$) décrit un étalement asymétrique de la fonction de relaxation pour les fréquences $\omega > 1/\tau_{CD}$, où τ_{CD} est le temps de relaxation de Cole-Davidson. Pour $\gamma = 1$, la fonction de Debye est à nouveau récupérée. Il est à noter que pour une fonction de modèle asymétrique comme la fonction de Cole-Davidson, le temps de relaxation caractéristique de la fonction du modèle ne coïncide pas avec le temps de relaxation lié à la position de pertes maximales. La relation entre les deux quantités dépend des paramètres de forme. Pour l'équation de Cole-Davidson (Équation I.11) on obtient :

$$\omega_p = \frac{1}{\tau_{CD}} \tan \left[\frac{\pi}{2\gamma + 2} \right] \quad (\text{I.12})$$

Un modèle plus général a été introduit par Havriliak et Negami [Havr 67] (La fonction HN) qui est en fait une combinaison des fonctions de Cole-Cole et Cole-Davidson :

$$\varepsilon_{HN}(\omega) = \varepsilon_\infty + \frac{\Delta\varepsilon}{(1 + (j\omega\tau_{HN})^\beta)^\gamma} \quad (\text{I.13})$$

Pour les paramètres de forme β et γ qui décrivent l'étalement symétrique et asymétrique de la fonction diélectrique complexe, on a : $0 < \beta \leq 1$ et $0 < \beta\gamma \leq 1$.

Les paramètres β et γ sont liés aux valeurs extrêmes de la fonction diélectrique complexe ε_s et ε_∞ de la manière suivante :

$$\begin{aligned} \varepsilon_s - \varepsilon'(\omega) &\approx \omega^m; \quad \varepsilon'' \approx \omega^m \quad \text{for } \omega \ll \frac{1}{\tau_{HN}} \quad \text{with } m = \beta \\ \varepsilon'(\omega) - \varepsilon_\infty &\approx \omega^{-n}; \quad \varepsilon'' \approx \omega^{-n} \quad \text{for } \omega \gg \frac{1}{\tau_{HN}} \quad \text{with } n = \beta\gamma \end{aligned}$$

En d'autres termes, les paramètres de forme m et n , introduits par Jonscher [Krem 03], sont les pentes de $\log(\varepsilon'')$ versus $\log(\omega)$ en basses et hautes fréquences par rapport à la position de perte maximale.

La position de pertes maximales dépend des paramètres β et γ selon :

$$\omega_p = \frac{1}{\tau_{HN}} \left[\sin \frac{\beta\pi}{2 + 2\gamma} \right]^{1/\beta} \left[\sin \frac{\beta\gamma\pi}{2 + 2\gamma} \right]^{-1/\beta} \quad (\text{I.14})$$

Les diélectriques peuvent contenir des charges libres qui proviennent des impuretés. Ces charges sont susceptibles de se déplacer sous l'action d'un champ électrique externe. Par conséquent le terme $-j\sigma/\varepsilon_0\omega$, qui représente les pertes par conduction, est souvent ajouté à l'expression de ε :

$$\varepsilon = \varepsilon' - j\varepsilon'' = \varepsilon' - j\left(\varepsilon_d + \frac{\sigma}{\varepsilon_0\omega}\right) \quad (\text{I.15})$$

avec ϵ_d est la permittivité relative due aux pertes d'origine diélectrique, σ est la conductivité statique du milieu.

Notons ici que le terme $\Delta\epsilon$ qui figure dans tous les modèles précédents exige des mesures en très basses (ϵ' tend vers ϵ_s et ϵ'' tend vers 0) et très hautes fréquences (ϵ' tend vers ϵ_∞ et ϵ'' tend vers 0). Cependant, la mesure à ces deux fréquences limites n'est pas toujours facile et nécessite un équipement coûteux.

I.4.2. Polarisation ionique

La polarisation ionique est due à la déformation des liaisons chimiques lors de l'application d'un champ électrique extérieur. Elle concerne les liaisons ioniques, fortement polaires, qui lient des atomes d'électronégativité différente. Elle se manifeste à des fréquences élevées (De l'ordre de 10^{12} Hz). Les matériaux comportant de telles liaisons sont appelés matériaux ioniques.



Figure I.6. Polarisation ionique.

I.4.3. Polarisation électronique

La polarisation électronique est due à la déformation du nuage électronique des atomes lorsqu'ils sont soumis à un champ électrique extérieur : le barycentre des charges positives s'éloigne du barycentre des charges négatives et crée un dipôle induit. Elle concerne les électrons de valence et les électrons de cœur et se manifeste donc jusqu'à des fréquences très élevées (De l'ordre de 10^{15} Hz), car les électrons, particules de très faible masse, peuvent suivre les variations très rapides du champ électrique.

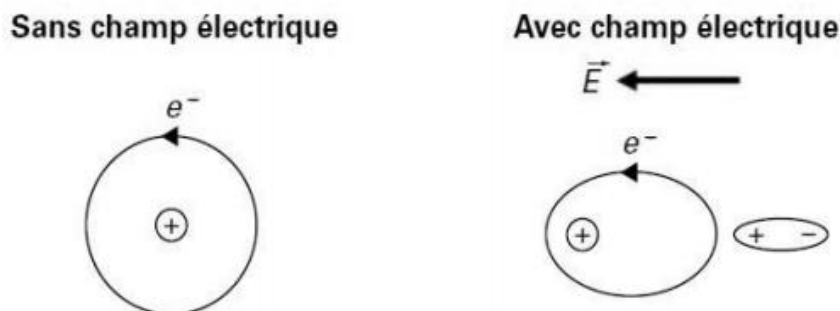


Figure I.7. Polarisation électronique.

I.5. Phénomène de relaxation

Après le phénomène de polarisation qui fait son apparition à partir des basses fréquences comme premier résultats de l'application d'un champ électrique alternatif, il vient en hautes fréquences le phénomène de relaxation. La relaxation diélectrique caractérise la difficulté que rencontrent les dipôles, liés aux différents ensembles de charges positives et négatives modélisant le diélectrique, à suivre les variations du champ électrique appliqué. En effet, le temps nécessaire pour vaincre les forces qui s'opposent à l'action du champ extérieur est appelé temps de relaxation τ .

La figure I.8 résume à la fois le phénomène de polarisation et de relaxation. En basses fréquences, les trois types de polarisation sont actifs et contribuent à une valeur élevée de ϵ' . Au fur et à mesure que la fréquence augmente, les mécanismes les plus lents disparaissent les uns après les autres pour atteindre une valeur très faible de ϵ' en très hautes fréquences. Dans les intervalles de diminution de ϵ' et l'apparition des pics sur le graphe de ϵ'' c'est la relaxation. En effet, pour des fréquences très inférieure à la relaxation (Soit dipolaire, ionique ou électronique), les dipôles ont suffisamment de temps pour suivre les variations du champ électrique, donc des pertes très faibles ($\epsilon'' \ll$) et une forte polarisation ($\epsilon' \gg$). De même si les fréquences sont très élevées, le champ varie trop vite et ainsi l'ensemble des dipôles ne peut pas suivre le renversement de la polarité ($\epsilon' \ll$). Aucune énergie n'est alors perdue dans le milieu ($\epsilon'' \ll$). Au voisinage de la fréquence qui correspond au temps de relaxation τ , un déphasage est observé entre la polarisation et le champ électrique entraînant une dissipation d'énergie sous forme de chaleur qui se traduit par un pic sur le graphe de ϵ'' .

Les pertes atteignent la valeur maximale à la fréquence f_p , liée au temps de relaxation donné par la relation suivante :

$$\tau = 1/\omega_p = 1/2\pi f_p \quad (I.16)$$

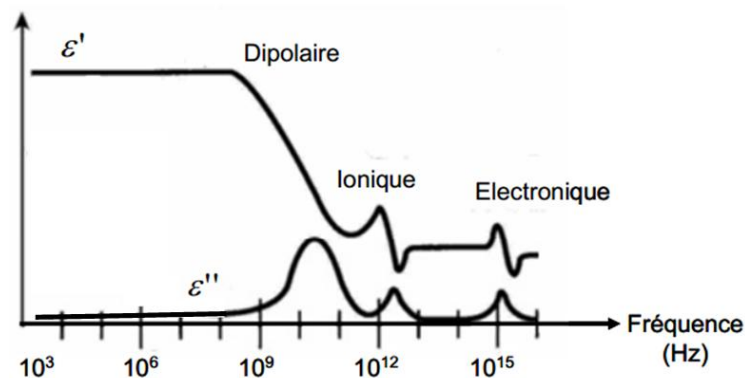


Figure I.8. Différents phénomènes de polarisation et relaxation.

I.6. Techniques de caractérisation

Plusieurs techniques de caractérisation de matériau sont décrites dans la littérature. Le choix dépend de la bande de fréquence à couvrir, les propriétés physiques du matériau et le type du matériau à caractériser (Gaz, liquide, solide ou poudre). De plus, l'anisotropie et la non homogénéité (Matériaux composites ou multicouches) jouent un rôle décisif dans ce choix.

Les techniques de caractérisation peuvent être divisées en trois catégories :

I.6.1. Technique de mesure en cavité résonnante

Une cavité résonnante est une enceinte métallique avec des parois conductrices (Souvent un bon conducteur comme le cuivre pure). Elle prend souvent des formes canoniques (Cylindrique ou rectangulaire) où la distribution des champs est très bien connue et facilement calculable (Figure I.9). Le principe de la mesure est basé sur le changement de la fréquence de résonance (f_i) et du facteur de qualité (Q_i) entre le cas d'une cavité à vide et une cavité en présence d'un échantillon à l'intérieur. Ce principe est illustré par la figure I.10.

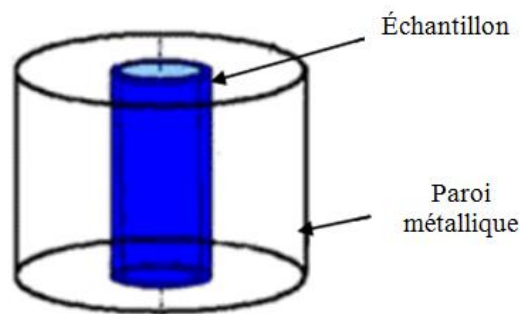


Figure I.9. Représentation schématique d'une cavité cylindrique.

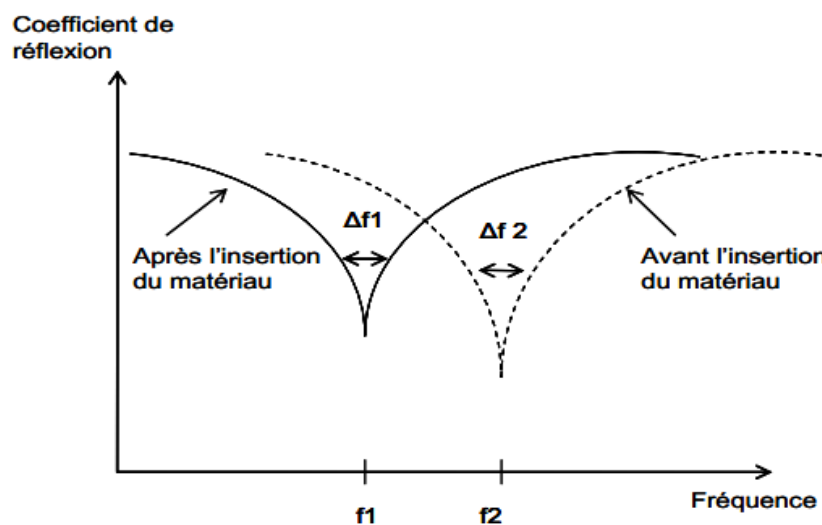


Figure I.10. Technique de perturbation dans une cavité résonnante.

$$Q_i = f_i / \Delta f_i, \quad i = 1, 2 \quad (\text{I.17})$$

Δf_i est la largeur à -3dB du pic de résonance à la fréquence f_i .

Cette technique est bien adaptée pour la caractérisation de matériaux diélectriques à faibles pertes et disponibles en faible quantité. Les matériaux gazeux, liquides et solides, y compris les poudres, ont été étudiés [Kaat 10]. Elle est très attractive comme outil pour quantifier la permittivité puisque elle est très précise [Terh 07], sans contact, non destructive et expérimentalement simple. Cependant, elle a plusieurs limitations. Si l'échantillon est trop grand ou possède une permittivité ou une tangente de perte trop élevée, la distribution du champ est significativement modifiée, ce qui rend l'approche de faibles perturbations moins précis [Orlo 14]. De plus, elle n'est malheureusement utilisable que dans une bande de fréquences très étroite (Au voisinage de la fréquence de résonance). Elle est donc inadaptée pour des mesures en large bande.

I.6.2. Technique de mesure en espace libre

Cette technique est basée sur la mesure des coefficients de transmission et réflexion qui permettent ensuite de remonter aux caractéristiques diélectriques du matériau sous teste. Ce dernier est placé au mi-chemin entre deux antennes ; émettrice et réceptrice (Mode réflexion-transmission). En mode réflexion l'antenne réceptrice est remplacée par un plan de référence qui forme un court-circuit (Figure I.11). La mesure en espace libre est bien adaptée pour les matériaux disponibles sous forme de grandes plaques [Elma 16]. Bien qu'elle est simple, large bande et applicable dans plusieurs domaines du Contrôle Non-Destructif (CND) comme la détection de défauts et la mesure d'épaisseur des couches minces, elle souffre de plusieurs inconvénients ; la surface de l'échantillon doit être large et plane pour éviter les diffractions au niveau des bords du matériau, sinon il faut avoir recours à l'utilisation d'antennes focalisantes. L'autre inconvénient majeur est les réflexions multiples entre les deux antennes via la surface de l'échantillon et celles due aux objets environnants ce qui la rend relativement moins précis derrière les deux autres techniques (La cavité résonnante et les structures de propagation).

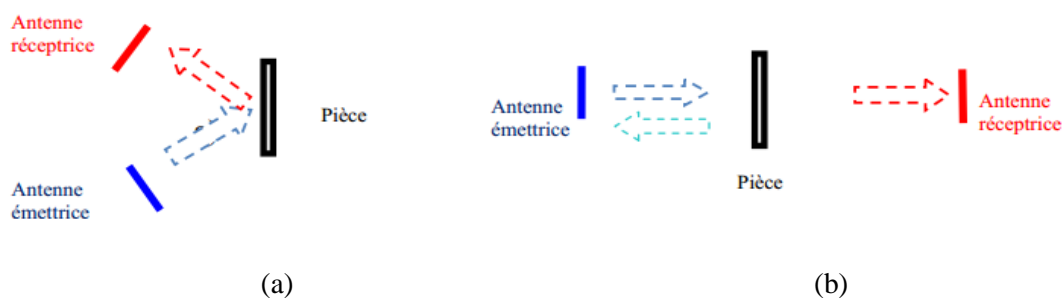


Figure I.11. Technique de mesure en espace libre :

(a) en réflexion, (b) en réflexion-transmission.

I.6.3. Techniques utilisant des structures de propagation

Ces méthodes sont moins précises que la technique de la cavité résonnante et plus précises que la mesure en espace libre [Terh 07]. Elles peuvent être divisées en trois catégories :

I.6.3.1. Technique en transmission-réflexion

Le principe de la mesure en transmission-réflexion est d'insérer l'échantillon à caractériser à l'intérieur d'un guide d'onde (Souvent rectangulaire) ou dans une ligne de transmission (Souvent coaxiale) pour former une discontinuité (Figure I.12). Ensuite l'échantillon est soumis à un champ électromagnétique incident émis par l'un des deux ports connectés au niveau des extrémités. Généralement, l'échantillon occupe soit toute la section du guide ou seulement une partie ce qui nécessite un usinage rigoureux du matériau de manière à ce qu'il puisse être parfaitement inséré dans la structure de propagation. Ceci est un inconvénient dans certaines applications où la pièce doit pouvoir être réutilisée.

A l'aide d'un analyseur de réseaux les paramètres de la matrice S sont mesurés. Théoriquement, les coefficients de réflexion S_{11} et de transmission S_{21} sont suffisants pour remonter à la permittivité et la perméabilité du matériau.

La mesure en réflexion-transmission est bien adaptée aux matériaux magnétiques solides et les diélectriques anisotropes et elle est couramment utilisée pour mesurer les échantillons à grande perte et à perte moyenne. Cette technique ne fonctionne que dans la bande de fréquence où seulement le mode fondamentale peut se propager dans la cellule de mesure ; Transverse Électrique (TE_{10}) pour le guide d'onde rectangulaire (Bande étroite) et Transverse Électrique-Magnétique (TEM) pour la ligne coaxiale (Large bande).

On peut citer ici comme exemple les travaux de Belhadj et al [Belh 90, 98] où les échantillons forment des discontinuités dans un câble coaxial. De plus, Yakovlev et al [Yako 05] ont étudié la caractérisation des diélectriques par insertion d'un échantillon à l'intérieur d'un guide d'onde

rectangulaire (Le cas d'un guide d'onde court-circuité (Voire le paragraphe suivante) a été traité aussi).

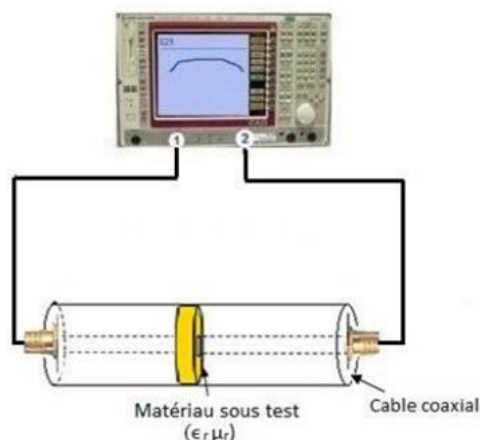


Figure I.12. Schéma de mesure en câble coaxiale.

1.6.3.2. Ligne ou guide d'onde court-circuité

Cette technique utilise un seul port d'excitation tandis que l'autre extrémité est court-circuitée (Figure I.13) ; c'est à dire un cas particulier de la technique en transmission-réflexion. Le court-circuit (CC) peut être fixe ou mobile dans la ligne. L'avantage d'un CC mobile (Figure I.14) est qu'il permet d'effectuer séparément plusieurs mesures pour une fréquence donnée et ainsi permet de placer l'échantillon soit dans une zone de champ électrique fort ; un emplacement favorable à la mesure de la permittivité diélectrique, ou dans une zone de champ magnétique fort qui est favorable à la mesure de la perméabilité magnétique. Pour la mesure simultanée de la permittivité et la perméabilité, les données de mesure pour deux différentes positions de CC sont nécessaires. Si la permittivité est le seul paramètre à rechercher, la mesure d'un seul coefficient de réflexion est suffisante. La mesure en réflexion est bien convenable à la caractérisation des matériaux diélectriques homogènes et isotropes, solides ou liquides.

On peut citer ici comme exemple l'article de Terhzaz et al [Terh 07] où un échantillon du Téflon ou du Plexiglas est inséré dans un guide d'onde rectangulaire court-circuité.

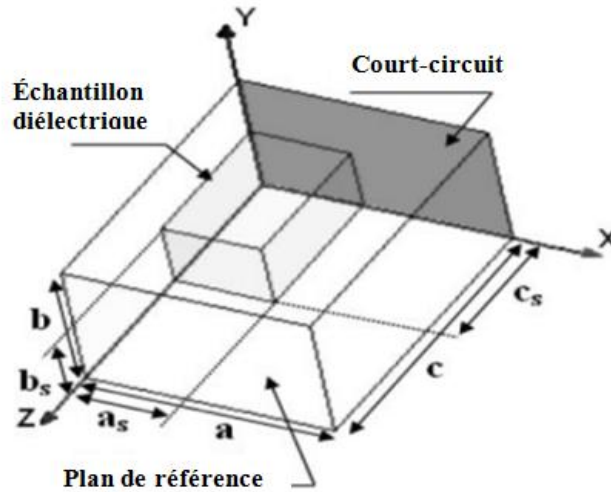


Figure I.13. Guide d'onde court-circuité chargé par un échantillon diélectrique.

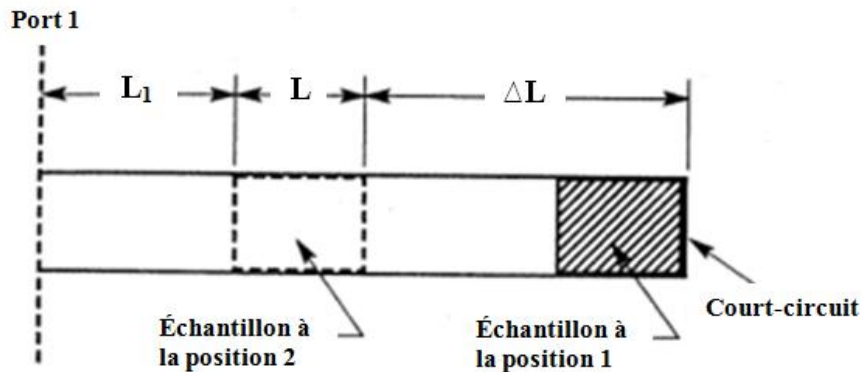


Figure I.14. Ligne de transmission court-circuitée avec deux positions de mesure.

I.6.3.3. Technique de mesure avec contact

Dans ce type de mesure, la sonde (Souvent une ligne coaxiale) est pressée contre un échantillon et la permittivité peut être déterminée à partir de la mesure du coefficient de réflexion. A l'origine, Cette technique a été développée pour la caractérisation large bande des tissus biologiques et des diélectriques liquides. De plus, elle a été adaptée pour la détection à distance, les applications agricoles, géotechniques et physiques sur le sol [Wagn 14, Fila 06].

On distingue deux catégories : caractérisation en milieu confiné et caractérisation en milieu ouvert. Pour la première catégorie, un échantillon d'une pièce doit être inséré dans une cellule fixée sur le guide d'onde (Figure I.15.b), c.-à-d. il faut sacrifier la pièce pour le teste. Par contre, la deuxième catégorie entre dans le cadre des méthodes non destructives. Généralement, on utilise une ligne coaxiale ou un guide d'onde à extrémité ouverte qui rayonne dans un demi-espace (Figure I.15.a). La difficulté majeure de la technique avec contacte réside dans la métrise de

l'interface guide d'onde / échantillon. La présence d'un gap d'air peut diminuer considérablement la sensibilité du signal de la sonde. Souvent, on peut éliminer ce problème en faisant recours à la métallisation des surfaces de contact. Comme toutes les méthodes qui utilisent des structures de propagation, seulement le mode fondamentale se propage.

Des exemples de cette technique en milieu ouverts sont publiés dans les articles de Zheng et al. [Zhen 91], Wagner et al. [Wagn 14], Misra [Misr 95] (Sonde coaxiale) et Bakhtiari et al. [Bakh 93] (Guide d'onde rectangulaire). Dans le cas de mesure en milieu confiné, on peut citer les travaux de Acikgoz et al. [Acik 07, 08 ; Acikgo 08] et Hacib et al. [Haci 10,11] (Ligne coaxiale).

Dans ce travail de thèse, la technique de mesure avec contact en milieu confiné est adoptée vue sa simplicité de mise en œuvre. La structure de propagation est un câble coaxial qui permet une caractérisation large bande. L'avantage d'une technique large bande est qu'elles permettent d'augmenter les informations sur la structure caractérisée et d'examiner les différents phénomènes de relaxation.

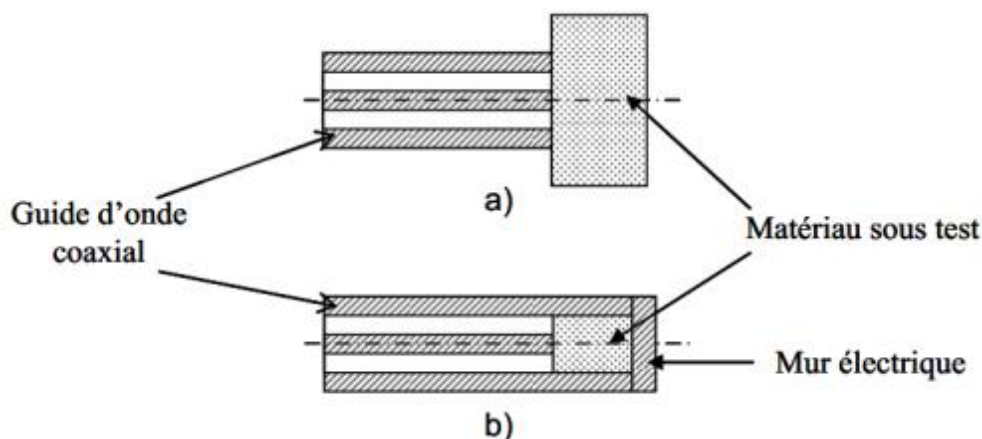


Figure I.15. Mesure en contact : a) mesure en milieu ouvert, b) mesure en milieu confiné.

I.7. Résolution des problèmes direct et inverse

Comme la mesure directe des propriétés physiques d'un matériau est impossible, les scientifiques se tournent vers la mesure d'autres entités qui peuvent constituer une image directe de ces propriétés. La technique est d'exciter le matériau sous teste par une source d'énergie et étudier ensuite la réponse du matériau qui dépend sûrement de ses propriétés physiques (Figure I.16). L'opération qui permet d'établir un model pour remonter aux paramètres de recherche à partir des observations est le problème inverse. Le problème direct est l'opération qui donne les observations en connaissant les propriétés du matériau. Les observations peuvent être calculés analytiquement ou numériquement, ou bien mesurés expérimentalement. En général la solution du

problème inverse n'est pas unique, il faudra donc inclure dans le processus d'inversion une procédure permettant d'évaluer la fiabilité de la reconstruction par rapport à la réalité. Dans le domaine de la caractérisation micro-ondes, quand on parle de paramètres de recherche, il s'agit de la permittivité diélectrique ou la perméabilité magnétique, alors que les observations sont généralement le coefficient de réflexion et de transmission.

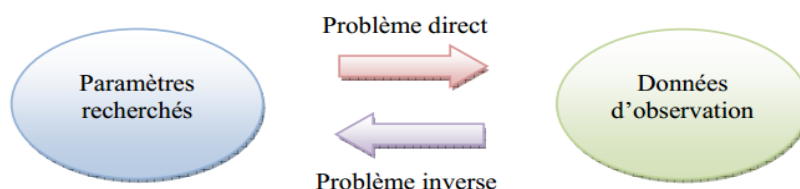


Figure I.16. Problème direct/Problème inverse.

Pour résoudre les problèmes direct et inverse dans le domaine de la caractérisation micro-ondes, beaucoup d'approches peuvent être envisagés. Avant tout, s'il existe une formule analytique d'inversion (Figure I.17), ce sera le meilleur choix sur tous les plans (Simplicité, rapidité et efficacité) mais cela est très rare dans la pratique car les phénomènes physiques sont très complexes.

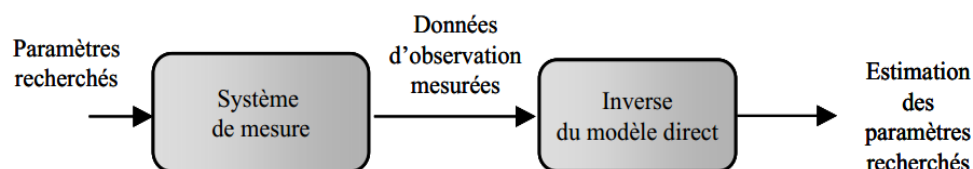


Figure I.17. Processus d'inversion directe utilisant l'inverse du modèle direct.

Une autre approche consiste à insérer une solution analytique dans une procédure itérative d'inversion (Figure I.18). En effet, pour des géométries simples et pour des fréquences relativement basses, on peut établir une formulation analytique du problème directe. La recherche de cette solution peut cependant être très longue et fastidieuse. De plus, la relation établie peut être très complexe et nécessite un coût de calcul important. L'inversion itérative cherche à minimiser l'écart entre la sortie du model direct et les données de mesure en utilisant un algorithme d'optimisation. Le processus itératif se poursuit jusqu'à ce qu'un critère d'arrêt soit satisfait. Le critère d'arrêt généralement utilisé est lié soit au nombre d'itérations maximal soit à l'erreur minimale acceptée. Un exemple de cette approche est détaillé dans [Belh 98] où le

problème direct est résolu analytiquement en utilisant la méthode de raccordement modal, tandis que la méthode du gradient est utilisée pour résoudre le problème inverse.

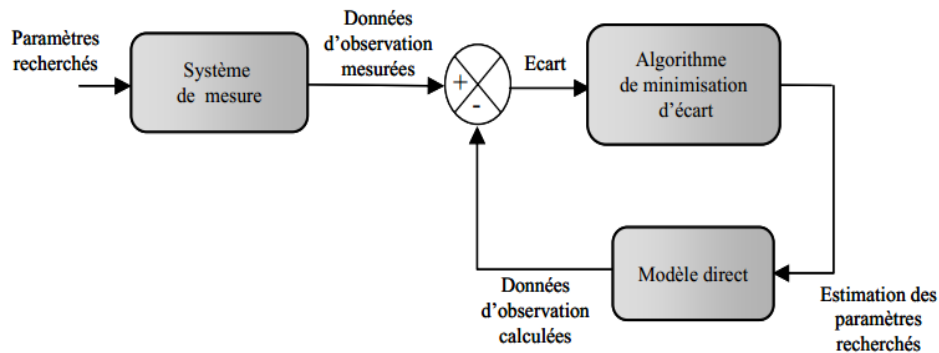


Figure I.18. Schéma du processus itératif d'inversion.

Par ailleurs, pour des géométries complexes, la recherche de la solution analytique est rarement triviale voire impossible. La solution pour contourner ce problème est le recours à une modélisation numérique. Cependant, Les méthodes numériques souvent utilisées comme les différences finies, les éléments finis (FEM pour Finite Element Method), les éléments de frontières et les différences finies dans le domaine temporel (FDTD pour Finite-Difference Time-Domain) sont coûteuses en temps de calcul. Par conséquent, l'insertion d'une solution numérique dans une procédure itérative sera très coûteuse en temps de calcul pour des mesures multifréquences sachant que la méthode numérique est sollicitée plusieurs fois pour chaque fréquence. Un exemple de cette approche est publié dans [Terh 07]. L'étude consiste à insérer un échantillon qui occupe partiellement la section d'un guide d'onde rectangulaire. Le problème direct est résolu pour le coefficient de réflexion en utilisant la FDTD, tandis que la méthode de Newton-Raphson est utilisée pour résoudre itérativement le problème inverse.

Récemment, pour des géométries complexes et sur une large bande de fréquence, les approches qui associent les méthodes numériques avec des outils de l'intelligence artificielle ont vu le jour. L'idée est d'élaborer un modèle inverse paramétrique dont on règle les paramètres internes à l'aide d'une base de données obtenue par simulation ou par des moyens expérimentaux (Figure I.19). Cette base contient des exemples connus de couples paramètres recherchés-données d'observation. Parmi les travaux réalisés utilisant cette approche pour des sondes coaxiales en contact, on peut citer les travaux d'Acikgoz et al. [Acik 07, 08] et Hacib et al. [Haci 10, 11] dans lesquels quelques variantes des réseaux de neurones artificiels (ANN pour Artificial Neural Networks) et des Machines à vecteurs de support (SVM pour Support Vector Machines) ont été associées avec la FEM. Ce travail de thèse s'inscrit dans cette ligne de recherche dans lequel on va

exposer d'autres méthodes d'inversion qui peuvent substituer aux méthodes précédentes et les surpasser en termes de rapidité, efficacité et simplicité.

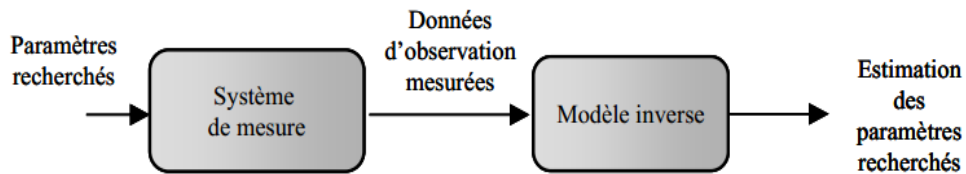


Figure I.19. Schéma du processus d'inversion directe basé sur un modèle inverse.

Conclusion

Dans ce chapitre, nous avons essayé de donner le contexte exact du travail. Il s'agit d'une méthode de caractérisation électromagnétique et plus précisément une technique par micro-ondes. Cette étude implique un type exact de matériau ; ceux sont les diélectriques avec pertes. La technique de mesure s'inscrit dans la catégorie des techniques avec contacte dans laquelle on utilise un câble coaxial qui joue le rôle d'un guide d'onde pour exciter le matériau à caractériser. C'est une méthode en réflexion seul, c.-à-d. il suffit de mesurer le coefficient de réflexion au niveau de la discontinuité guide d'onde-échantillon. Pour remonter aux propriétés du matériau, nous avons exposé l'approche permettant de résoudre les problèmes direct et inverse.

Chapitre II : Modélisation électromagnétique et construction des bases de données

II.1. Introduction

Dans cette thèse, l'approche utilisant un modèle inverse qui réalise une inversion instantanée du modèle direct est adoptée. Habituellement les méthodes inscrites dans cette approche ont besoin d'une base de données pour l'apprentissage. Contrairement à d'autres travaux où la base de données est construite expérimentalement, ce qui est une tâche très difficile puisque elle exige des mesures sur plusieurs matériaux, dans notre travail la base de données est créée par simulation numérique. En se servant de la souplesse d'un modèle numérique, il est possible d'agir à volonté sur les limites du domaine d'apprentissage imposées par la permittivité des matériaux à caractériser. De plus, la modélisation numérique permet de choisir plus facilement les exemples de la base d'apprentissage de façon à augmenter la précision de l'inversion. Dans ce travail, nous utilisons la MEF pour générer les bases nécessaires.

II.2. Rappel sur le champ électromagnétique et les équations de Maxwell

La théorie classique du champ électromagnétique s'est fixée dans les années 1860, lors que Maxwell compléta les travaux d'Ampère et de Faraday et obtint un ensemble cohérent d'équations différentielles qui régissent les variations spatio-temporelles des champs d'excitations (Champ électrique E et champ magnétique H) et les champs qui caractérisent la réponse du milieu au champ électromagnétique (Champ de déplacement D et champ de l'induction magnétique B). En introduisant la densité de courant de conduction J et la densité de charge ρ , la propagation d'une onde électromagnétique dans un domaine Ω , borné par des frontières Γ (Figure II.1) est gouvernée par les équations de Maxwell qui s'écrivent en fonction du temps t comme suit :

$$\text{Équation de Maxwell-Gauss} \quad \nabla \cdot \vec{D} = \rho \quad (\text{II.1})$$

$$\text{Équation de conservation du flux magnétique} \quad \nabla \cdot \vec{B} = 0 \quad (\text{II.2})$$

$$\text{Équation de Maxwell-Faraday} \quad \nabla \times \vec{E} = - \frac{\partial \vec{B}}{\partial t} \quad (\text{II.3})$$

$$\text{Équation de Maxwell-Ampère} \quad \nabla \times \vec{H} = \frac{\partial \vec{D}}{\partial t} + \vec{J} \quad (\text{II.4})$$

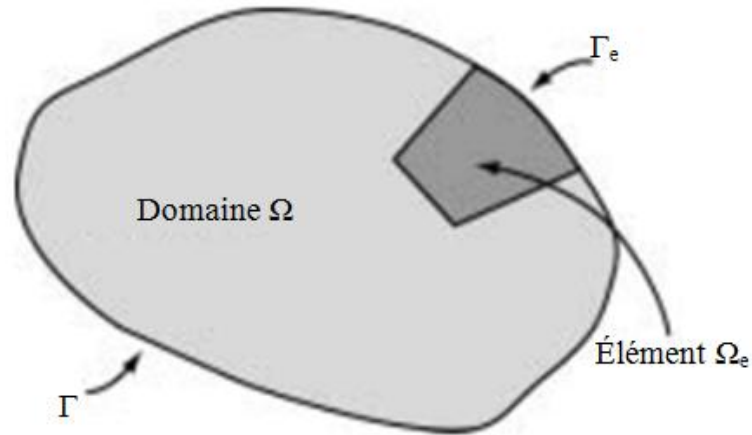


Figure II.1. Domaine de propagation de l'onde électromagnétique.

Si nous considérons des champs électriques et magnétiques variant sinusoidalement dans le temps, nous pouvons écrire, avec ω la fréquence angulaire :

$$E(t) = E e^{j\omega t} \quad (\text{II.5})$$

$$H(t) = H e^{j\omega t} \quad (\text{II.6})$$

En tenant compte des relations constitutives d'un milieu linéaire, homogène et isotrope avec pertes, on peut écrire :

$$\vec{D} = \varepsilon \vec{E} \quad (\text{II.7})$$

$$\vec{B} = \mu \vec{H} \quad (\text{II.8})$$

$$\vec{J} = \sigma \vec{E} \quad (\text{II.9})$$

avec ε la permittivité diélectrique, μ la perméabilité magnétique et σ la conductivité électrique du milieu.

Dans notre étude, les diélectriques sont parfaits (ρ et σ sont nuls) et amagnétiques ($\mu = \mu_0$). Les équations de Maxwell s'écrivent alors :

$$\nabla \cdot \varepsilon \vec{E} = 0 \quad (\text{II.10})$$

$$\nabla \cdot \mu_0 \vec{H} = 0 \quad (\text{II.11})$$

$$\nabla \times \vec{E} = -j\omega\mu_0 \vec{H} \quad (\text{II.12})$$

$$\nabla \times \vec{H} = j\omega\varepsilon \vec{E} \quad (\text{II.13})$$

À partir de (II.12), on a : $\vec{H} = \frac{1}{-j\omega\mu_0} \nabla \times \vec{E}$

En remplaçant \vec{H} dans (II.13), on obtient :

$$\nabla \times \left(\frac{1}{-j\omega\mu_0} \nabla \times \vec{E} \right) = j\omega\varepsilon \vec{E} \quad (\text{II.14})$$

De même, en suivant le même raisonnement, nous aurons :

$$\nabla \times \left(\frac{1}{j\omega\varepsilon} \nabla \times \vec{H} \right) = -j\omega\mu_0 \vec{H} \quad (\text{II.15})$$

En raisonnant en champ électrique, le problème électromagnétique peut être résolu en associant l'équation (II.14) aux conditions aux limites suivantes :

$\vec{n} \times \vec{E} = \vec{0}$: Condition aux limites de type PEC (Perfect Electric Conductor) ou mur électrique : qui impose la nullité de la composante tangentielle du champ électrique sur les parties métalliques.

$\vec{n} \times \vec{H} = \vec{0}$: Condition aux limites de type PMC (Perfect Magnetic Conductor) ou mur magnétique : nullité de la composante tangentielle du champ magnétique sur les plan de symétrie.

$\vec{n} \times \vec{E} = \vec{n} \times \vec{E}_0$: Condition aux limites de type source d'excitation. \vec{E}_0 est un champ électrique appliqué sur la surface d'excitation.

avec \vec{n} est le vecteur normal sortant de la surface sur laquelle est appliquée la condition aux limites.

II.3. Méthodes de résolution

On peut envisager de résoudre les équations de Maxwell de manière analytique ou de manière numérique :

II.3.1. Solution analytique

Dans le cas d'une géométrie simple (2D ou 3D axisymétrique), il est possible de résoudre le problème direct d'une manière analytique. Les matériaux doivent avoir des propriétés linéaires. Bien que la résolution analytique soit souvent exacte et rapide, la recherche de la solution peut être très longue et fastidieuse avec une relation finale parfois très complexe qui nécessite un coût de calcul important. La résolution analytique est très performante pour le calcul de grandeurs globales (Le coefficient de réflexion et l'admittance du capteur dans notre cas). Par contre, pour les grandeurs locales (\vec{E} , \vec{H} , ..., etc) il faut un calcul pour chaque point de la géométrie considérée afin de former une cartographie du champ, donc un temps de calcul considérable. Toutefois, les méthodes analytiques peuvent servir à valider les différentes méthodes numériques [Thol 95, Mare 04].

À titre d'exemple, on peut citer ici les travaux de Belhadj [Belh 86, 90, 97, 98].

II.3.2. Solution numérique

II.3.2.1. Méthode des différences finies dans le domaine temporel

La méthode FDTD a été développée par Yee en 1966. Appuyée par le développement de l'outil informatique, elle est présente aujourd'hui dans plusieurs domaines de l'ingénierie. La FDTD peut simuler le comportement d'une onde électromagnétique dans tout type de milieu (Diélectrique, métal, plasma, ...), sa formulation théorique est relativement simple et ne fait intervenir aucune inversion de matrice [Belk 08]. Cependant, elle s'adapte difficilement aux géométries complexes qui présentent des courbures car elle nécessite un découpage régulier du domaine d'étude. La formulation de la FDTD est basée sur la discrétisation directe des équations de Maxwell II.3 et II.4, les deux autres (II.1 et II.2) sont implicites dans l'algorithme de la FDTD.

II.3.2.2. Méthode des éléments de frontières

La méthode des éléments de frontière, basée sur la théorie des équations intégrales de frontière, avait été développée depuis les années soixante comme alternative à la FEM. Elle avait été appliquée dans différents problèmes de l'ingénierie, y compris les problèmes thermiques, élastostatiques et la mécanique des solides et fluides. Maintenant, grâce à ses avantages et la puissance croissante des calculateurs, elle gagne d'autres domaines comme l'électromagnétisme. L'attraction principale de la méthode réside dans la simplicité des données nécessaires pour la résolution car seulement la surface de la frontière doit être discrétisée (Moins de degrés de liberté) donc la dimension du problème est réduite de un. De plus, elle permet l'étude de problèmes posés dans des domaines infinis de manière plus naturelle, donc plus adaptée, par exemple, pour la modélisation de la propagation des ondes dans l'espace libre, contrairement à la FEM malgré que cette dernière a été développée pour s'adapter aux domaines infinis [Breb 87, Prem 00].

II.3.2.3. Méthode des éléments finis

La FEM est une méthode d'approximation. Elle a pour but de remplacer un problème mathématique défini sur un milieu continu (Équation II.14) par un problème discret sous forme matricielle. La première étape de la FEM est d'obtenir la formulation intégrale du problème. L'intérêt de cette formulation est de tenir compte de l'équation locale et de tout ou une partie des conditions aux limites, et ainsi obtenir une approximation par sous-régions qui donne, après assemblage, une solution sur la totalité du domaine. Il existe plusieurs approches permettant d'obtenir une formulation intégrale : la méthode des résidus pondérés, la méthode variationnelle, ... etc.

La formulation variationnelle est obtenue en faisant le produit scalaire de l'équation (II.14) avec une fonction test \vec{V} et en intégrant sur le domaine d'étude. Ce qui donne :

$$\int_{\Omega} \nabla \times \left(\frac{1}{j\omega\mu_0} \nabla \times \vec{E} \right) \cdot \vec{V} d\Omega - j\omega \int_{\Omega} \varepsilon \vec{E} \cdot \vec{V} d\Omega = 0 \quad (\text{II.16})$$

En faisant une intégration par partie et en appliquant le théorème de Stokes-Green, on obtient :

$$\int_{\Omega} \nabla \times \vec{V} \cdot \left(-\frac{1}{j\omega\mu_0} \right) \nabla \times \vec{E} d\Omega + \int_{\Gamma} \left(-\frac{1}{j\omega\mu_0} \nabla \times \vec{E} \wedge \vec{V} \right) \cdot \vec{n} d\Gamma - j\omega \int_{\Omega} \varepsilon (\vec{E} \cdot \vec{V}) d\Omega = 0 \quad (\text{II.17})$$

Le terme de bords (L'intégral sur Γ) s'annule naturellement en présence de conditions aux limites de type PMC. En présence de conditions de type PEC il s'annule en choisissant \vec{V} telle que $\vec{V} \wedge \vec{n} = \vec{0}$. Le problème s'écrit alors :

$$\int_{\Omega} \left(-\frac{1}{j\omega\mu_0} \right) \nabla \times \vec{E} \cdot \nabla \times \vec{V} d\Omega - j\omega \int_{\Omega} \varepsilon (\vec{E} \cdot \vec{V}) d\Omega = 0 \quad (\text{II.18})$$

Le maillage est la deuxième étape de la résolution. Il consiste en la division du domaine d'étude en un nombre élevé d'éléments géométriques (Mailles). La FEM donne une solution approchée du champ qui est obtenue par juxtaposition des champs locaux définis dans chaque élément. Un élément fini est une maille possédant un certain nombre de nœuds ainsi qu'une famille de champs locaux destinés à fournir une solution approchée en tout point de la maille. Le choix de la famille de champs locaux concerne à la fois les degrés de liberté dans les éléments et les fonctions d'interpolations qui définissent le champ local. Le nombre de degrés de liberté par nœud varie selon la nature du champ inconnu (Scalaire, vecteur à deux ou trois composantes) et la dimension de l'espace physique (1D, 2D ou 3D).

Par exemple, pour les éléments finis d'arêtes tétraédriques (Figure II.2) du premier ordre, la fonction d'interpolation relative à l'arête « a » est donnée par :

$$\vec{W}_a = \lambda_i \vec{\nabla} \lambda_j - \lambda_j \vec{\nabla} \lambda_i \quad (\text{II.19})$$

avec λ_i, λ_j sont les fonctions barycentriques associées aux sommets respectifs i et j de l'arête « a ».

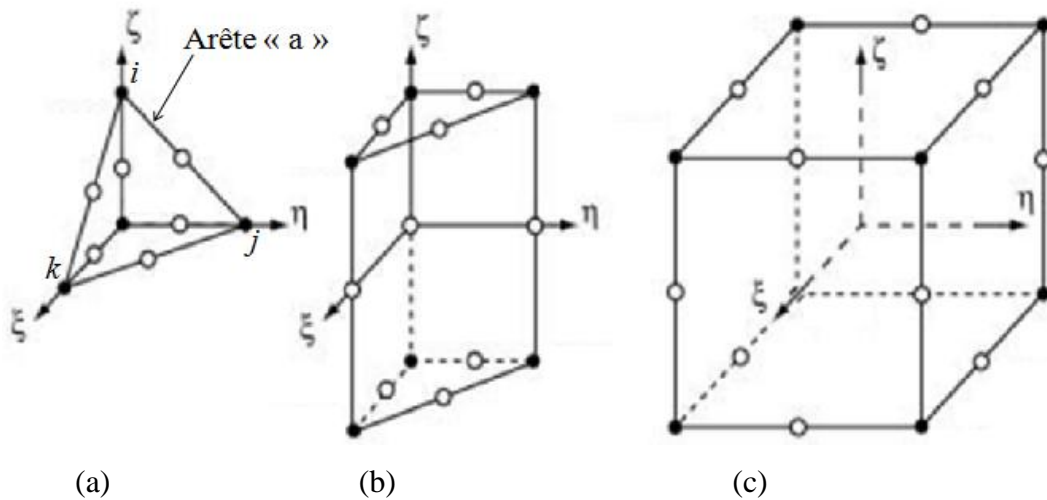


Figure II.2. Les principaux éléments de référence pour les maillages de volume, (a) tétraèdre, (b) pentaèdre et (c) hexaèdre.

Remarque : dans la figure II.2, les cercles creux représentent les nœuds supplémentaires pour les éléments quadratiques.

\vec{W}_a satisfait la relation suivante :

$$\int_{\text{arête } a'} \vec{W}_a \cdot d\vec{l} = \begin{cases} 1 & \text{si } a = a' \\ 0 & \text{si } a \neq a' \end{cases} \quad (\text{II.20})$$

Le champ électrique est ainsi écrit en termes de fonctions d'interpolation associées aux arêtes des éléments :

$$\vec{E} = \sum_{a=1}^{N_a} \vec{W}_a e_a \quad (\text{II.21})$$

où N_a est le nombre total d'arêtes du maillage et e_a est la circulation du champ électrique le long de l'arête « a ».

En substituant (II.21) dans (II.16) et en choisissant comme fonctions de test les fonctions d'interpolation $\vec{W}_{a'}, \forall a' \in [1, N_a]$ (Principe de Galerkin), on obtient le problème discret suivant :

$$\sum_{a=1}^{N_a} e_a \left[\int_{\Omega} -\frac{1}{j\omega\mu_0} \nabla \times \vec{W}_a \cdot \nabla \times \vec{W}_{a'} d\Omega - j\omega \int_{\Omega} \varepsilon (\vec{W}_a \cdot \vec{W}_{a'}) d\Omega \right] = 0 \quad (\text{II.22})$$

Finalement, en prenant en compte les termes d'excitation, on est conduit à un problème discrétisé qui se présente sous forme d'un système matriciel dont une ligne est donnée par :

$$\sum_{a=1}^{N_a} [-R_{aa'} + K^2 S_{aa'}] e_a = b_{a'} \quad (\text{II.23})$$

avec:

$$R_{aa'} = \int_{\Omega} \nabla \times \vec{W}_a \cdot \nabla \times \vec{W}_{a'} d\Omega \quad (\text{II.24})$$

$$S_{aa'} = \int_{\Omega} \varepsilon (\vec{W}_a \cdot \vec{W}_{a'}) d\Omega \quad (\text{II.25})$$

$$K^2 = \omega^2 \mu_0 \varepsilon \quad (\text{II.26})$$

$b_{a'}$ est un terme d'excitation sur l'arête a' .

II.4. Cellule de mesure SuperMit

La cellule SuperMit (Figure II.3) est une jonction entre un guide d'onde coaxial et un guide d'onde circulaire terminé par un court-circuit (Mur électrique). Le matériau sous teste occupe d'une façon homogène tout le volume du guide circulaire. Dans le cas des matériaux liquides, une fenêtre coaxiale étanche est insérée entre les deux guides. L'ensemble est relié à un analyseur de réseau qui permet d'effectuer des mesures sur le coefficient de réflexion ou l'admittance du capteur (Figure II.4).

Le protocole de mesure SuperMit a été développé par Belhadj et al [Belh 86] au laboratoire LDIM de l'université UPMC-Paris6. Cette méthode de mesure permet la caractérisation des diélectriques solides et liquides à condition qu'ils soient homogènes et isotropes. C'est une technique large bande ; la seule limitation est l'apparition des modes supérieurs en hautes fréquences.

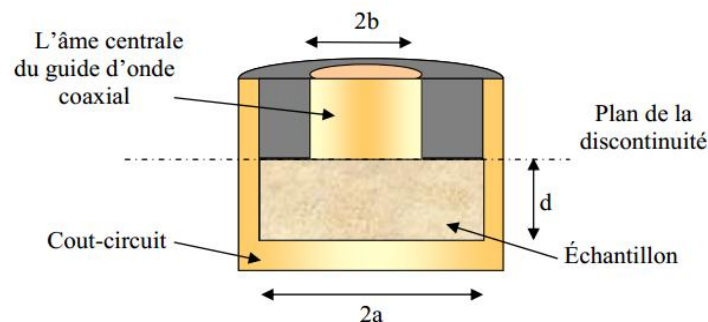


Figure II.3. Vue en coupe de la cellule de mesure SuperMit

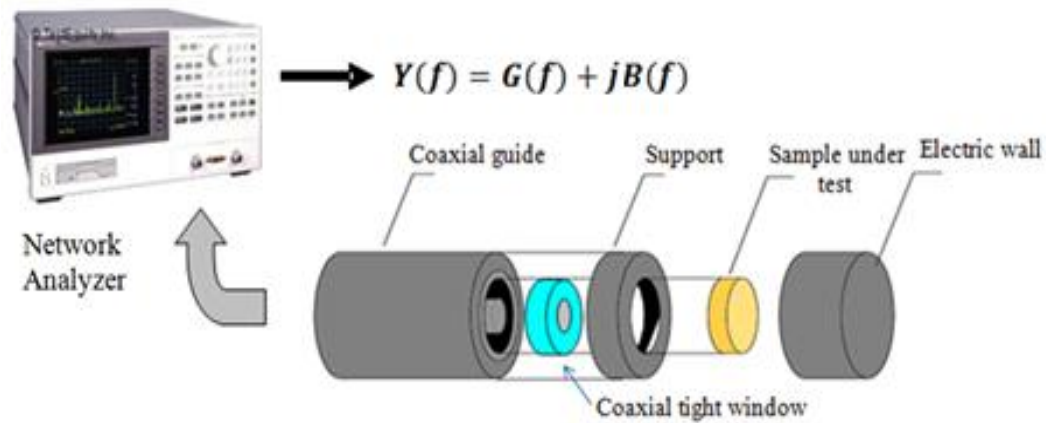


Figure II.4. Cellule SuperMit pour les liquides, connectée à un analyseur de réseau.

Le protocole requiert la propagation du mode fondamentale Transverse Électrique- Magnétique (TEM) d'une onde électromagnétique depuis l'analyseur jusqu'à la discontinuité (Plan de référence). A ce niveau, des modes d'ordre supérieurs Transverses Magnétiques (TM_{on}) seront générés et le contraste de permittivité va engendrer une onde réfléchiée. Généralement, le guide coaxiale doit être suffisamment long pour éviter la superposition des modes TM_{on} avec d'autres modes (Autres que le TEM) engendrés au début de la ligne par l'analyseur. La principale limitation de cette technique est l'apparition du mode Transverse Électrique (TE_{11}). En effet, ce dernier est évanescent au dessous de sa fréquence de coupure mais il peut se propager avec le mode fondamentale au-delà de cette fréquence. Cette fréquence de coupure est théoriquement de 19,6 GHz pour le standard APC7 ($a = 3.5$ mm, $b = 1.52$ mm) utilisé dans la cellule SuperMit.

Dans le cas des échantillons liquides, la permittivité et l'épaisseur de la fenêtre étanche sont choisis de telle sorte que seul le mode TEM existe au niveau de la discontinuité guide coaxiale-fenêtre. Afin d'éviter une réflexion élevée à l'interface guide coaxiale-fenêtre, la partie réelle de la permittivité de la fenêtre doit être aussi faible que possible. Pendant ce temps, pour éviter les pertes dans la fenêtre, la partie imaginaire doit être aussi faible. Dans notre étude, la fenêtre a une forme coaxiale pour assurer une meilleure transmission du signal. Elle est constituée du Plexiglas ($\epsilon' = 2.7$) avec une épaisseur de l'ordre du mm.

La figure II.5 montre la cellule de mesure SuperMit en 2D.

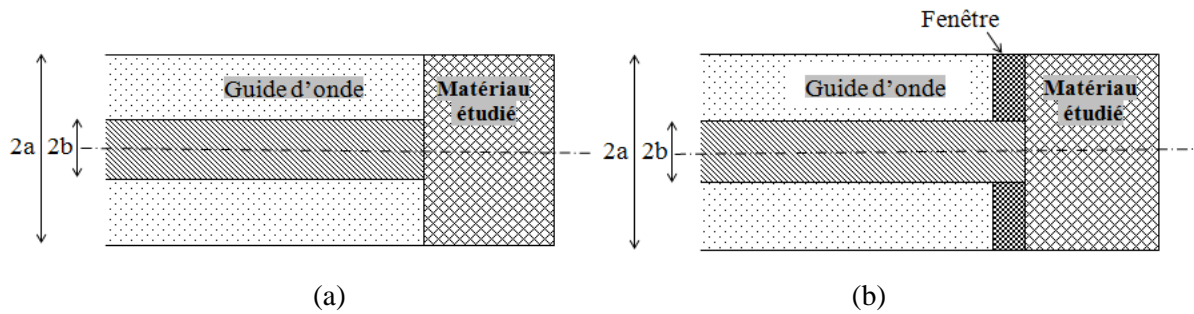


Figure II.5. Schéma 2D de la cellule SuperMit: (a) pour solides, (b) pour liquides

II.4.1. Modélisation analytique

Afin de l'insérer dans une procédure d'inversion itérative, Belhadj et al [Belh 86] ont établi une solution analytique pour le protocole SuperMit sans aucune restriction en prenant en compte tout les modes d'ordres supérieurs engendrés à l'interface entre le guide coaxial et l'échantillon étudié. En utilisant la méthode de raccordement modal, l'étude donne les éléments du schéma équivalent (Figure II.6) sur une large bande de fréquence (Du courant continu à 12 GHz). Pour un matériau aimantique, cette solution est donnée par :

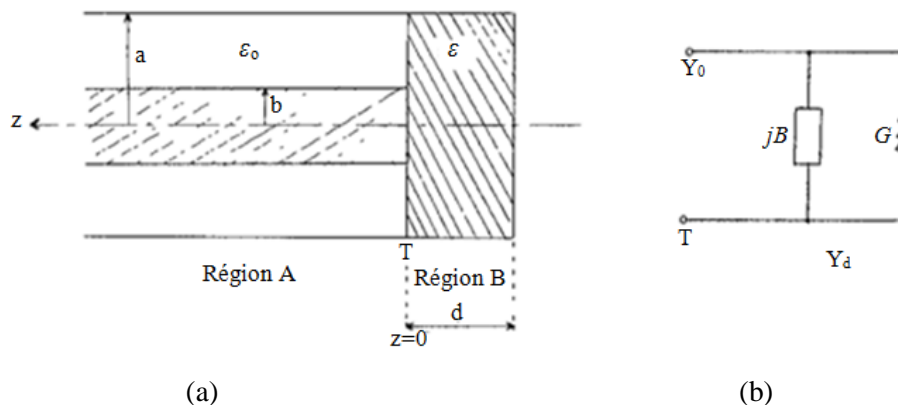


Figure II.6. Cellule SuperMit : (a) géométrie en 2D, (b) schéma équivalent.

$$Y_d = Y_0 \frac{jk_0 \varepsilon}{\ln \frac{b}{a}} \cdot \sum_{n=1}^{\infty} \frac{2 J_0^2(k_{Bn} a) \coth \left[\sqrt{k_{Bn}^2 - k_0^2} \varepsilon d \right]}{\sqrt{k_{Bn}^2 - k_0^2} \varepsilon \cdot [k_{Bn} b J_1(k_{Bn} b)]^2} \cdot \left[1 - a \sum_{m=1}^{\infty} \frac{A_m}{A_0(1+\Gamma)} \cdot \frac{Z_1(K_{Am} a)}{\left(\frac{k_{Am}}{k_{Bn}}\right)^2 - 1} \right] \quad (\text{II.27})$$

avec:

b : rayon de l'âme centrale du câble coaxial.

a : rayon du conducteur externe du câble coaxial.

d : épaisseur de l'échantillon.

ε : permittivité relative complexe du matériau.

k_0 ; constante de propagation dans l'air.

k_{Am} : constante de coupure du mode d'ordre supérieur TM_{0m} dans le guide coaxial.

k_{Bn} : constante de coupure du mode d'ordre supérieur TM_{0n} dans la cellule.

Y_0 : admittance caractéristique de la ligne coaxiale (0,02 S).

Γ : coefficient de réflexion de l'onde TEM sur le plan de la discontinuité.

A_0 : amplitude de l'onde TEM incidente.

A_m : amplitude du mode d'ordre supérieur TM_{0m} .

J_0, J_1 : fonctions de Bessel de première espèce respectivement d'ordre 0 et 1.

Z_1 : combinaison des fonctions de Bessel d'ordre 0 et 1 de première (J_0, J_1) et de seconde (N_0, N_1) espèce :

$$Z_1(k_{Am} a) = J_1(k_{Am} a) + \left[\frac{-J_0(k_{Am} a)}{N_0(k_{Am} a)} N_1(k_{Am} a) \right] \quad (\text{II.28})$$

k_{Am} et k_{Bn} représentent les solutions des deux équations suivantes :

$$J_0(k_{Bn} a) = 0 \quad (\text{II.29})$$

$$J_0(k_{Am} b) * N_0(k_{Am} a) - J_0(k_{Am} a) * N_0(k_{Am} b) = 0 \quad (\text{II.30})$$

II.4.2. Inversion itérative

Pour résoudre le problème inverse itérativement, l'expression analytique de l'admittance du capteur (Équation II.27) est insérée dans une procédure itérative. Premièrement, pour une fréquence donnée, un couple de ε' et ε'' est choisi arbitrairement. Ensuite, l'admittance correspondante Y_d est calculée puis comparée à la valeur mesurée pour avoir l'erreur à minimiser. Un algorithme d'optimisation essaye de minimiser l'erreur en ajustant les valeurs de ε' et ε'' par $d\varepsilon'$ et $d\varepsilon''$ respectivement (C'est à dire, dans l'itération suivante, ε' deviendra $\varepsilon' \pm d\varepsilon'$ et ε'' deviendra $\varepsilon'' \pm d\varepsilon''$). Le processus itératif se poursuit jusqu'à ce que la précision souhaitée soit atteinte. Finalement, le couple ($\varepsilon', \varepsilon''$) qui donne l'erreur minimale est retenu comme étant la permittivité complexe du matériau à la fréquence considérée (Et ainsi de suite pour les autres fréquences).

Cette méthode d'inversion est utilisée depuis longtemps au GeePs (Laboratoire de génie électrique et électronique de Paris (LGEP précédemment)). Elle est très rapide et efficace. Dans notre travail, les résultats de l'inversion itérative établis au GeePs seront utilisés comme outil de validation.

II.4.3. Modélisation numérique par FEM sous Ansys

En 1971, Ansys a été publié pour la première fois. C'est un logiciel de simulation numérique par éléments finis multidisciplinaire (Transfert de chaleur, écoulement de fluides, électromagnétisme, mécanique, ..., etc). Aujourd'hui, il est présent dans plusieurs domaines de l'ingénierie, y compris l'aérospatiale, l'automobile, l'électronique et le nucléaire [Moav 99]. Ansys offre la possibilité de travailler soit directement sur l'interface graphique, soit en exécutant un fichier de commandes rédigé sur un éditeur de texte.

Ansys propose un module pour chaque domaine physique avec des fonctions spécifiques. La modélisation de la propagation des ondes en hautes fréquences se déroule sur le module hyperfréquence qui dispose d'éléments de maillage hexaédriques (hf120) et tétraédriques (hf119). Dans notre cas, l'onde se propage dans des pièces circulaires, donc les éléments tétraédriques seront les plus appropriés puisque ils sont très adaptés aux géométries avec courbures.

Étant donné que le module hyperfréquence d'Ansys n'offre pas une étude en 2D, notre problème est modélisé en 3D malgré l'axisymétrie de la géométrie. Pour alléger les calculs, seulement une portion angulaire de quelques degrés (5 degrés) est modélisée. Quant au maillage, la géométrie est discrétisée en utilisant des éléments tétraédriques du second ordre. Généralement, 10 éléments par longueur d'onde au minimum sont suffisants pour obtenir une bonne précision de la simulation (Figure II.7). Concernant les conditions aux limites, des conditions de type PEC (Mur électrique) sont appliquées sur les surfaces des conducteurs internes et externes du guide coaxial ainsi que les parois externes de l'échantillon diélectrique (Support d'échantillon en matériau bon conducteur). Sur les plans de symétrie ou de coupure, un mur magnétique (PMC) est appliqué. À l'entrée du guide d'ondes coaxial, une source d'excitation est appliquée (Figure II.7).

Bien que la cellule SuperMit admette une solution analytique, la base de données nécessaire pour les outils d'inversion sera créée à partir des données issues de la modélisation par éléments finis. Le but est de réaliser une étude générale, puisque les phénomènes physiques admettent très rarement des solutions analytiques.

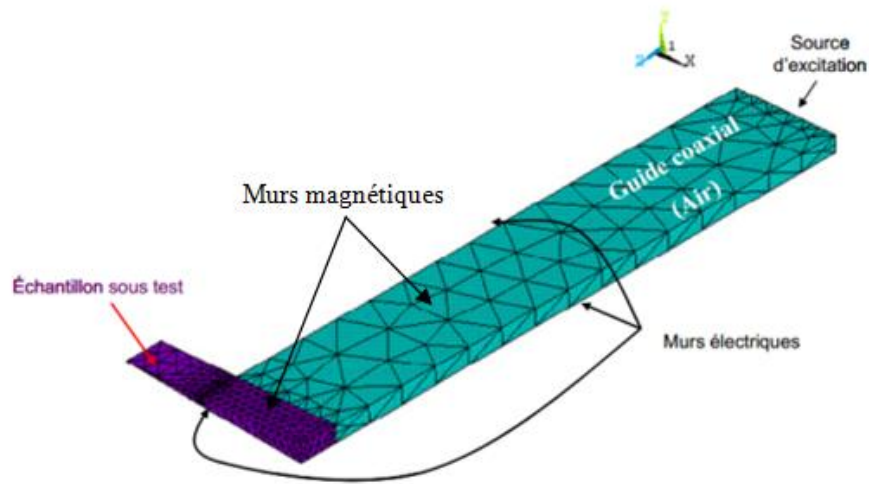


Figure II.7. Maillage et conditions aux limites pour la cellule SuperMit sans fenêtre.

L'admittance à la discontinuité est donnée par la relation suivante :

$$Y = Y_0 \frac{1 - \Gamma}{1 + \Gamma} = G + jB \quad (\text{II.31})$$

où $Y_0 = 0,02S$ est l'admittance caractéristique de la ligne coaxiale (50 Ω).

Γ est le coefficient de réflexion ramené à la discontinuité.

Cependant, la simulation donne le coefficient de réflexion à l'entrée du câble coaxial, au niveau du port d'excitation. Donc il faut le ramener à la discontinuité (Plan d'étalonnage par l'analyseur de réseau). Sachant que le module du coefficient reste le même quelque soit la position sur le guide, seule la phase φ varie d'une position à l'autre. La relation suivante permet de passer du plan d'excitation au plan d'étalonnage [Acikgo 08]:

$$\varphi_{\text{discontinuité}} = \varphi_{\text{excitation}} + 4\pi \frac{d}{\lambda} \quad (\text{II.32})$$

où d est la longueur du guide d'ondes coaxial et λ la longueur d'onde.

II.4.4. Construction des bases de données

Comme l'objectif général de cette thèse est la mise en œuvre d'une inversion directe basée sur des modèles paramétriques, donc une base de données est nécessaire. Cette base sert à faire apprendre, à l'outil d'inversion, le comportement fréquentiel de la permittivité en fonction de l'admittance. La base de données est constituée d'exemples reliant les entrées (f , ε' et ε'') et les sorties (G et B) du problème direct. La base de données est créée en résolvant numériquement le problème direct qui donne l'évolution fréquentiel de l'admittance ($Y(f) = G(f) + jB(f)$) du capteur mené d'un diélectrique dont la permittivité complexe peut prendre toutes les valeurs possibles.

Dans notre étude, on suppose connaître le matériau sous teste. Donc les limites supérieures seront les valeurs maximales (ε'_{\max} et ε''_{\max}) que peut prendre la permittivité complexe du matériau à caractériser ou d'un autre matériau dont les propriétés diélectriques sont proches sur la bande de fréquences dans laquelle on désire le caractériser (De 1 MHz à 1.8 GHz). En ce qui concerne les bornes inférieures, on a: $\varepsilon'_{\min} = 1$ pour l'air et $\varepsilon''_{\min} = 0$ pour un matériau sans pertes. Les éléments d'entrée de la base de données (ε' , ε'') sont sélectionnés de manière à être uniformément répartis sur le domaine de variation, tandis que les points de fréquence sont répartis suivant une loi logarithmique. Avant d'entrer dans la routine de génération de la base de données, toutes les N valeurs des vecteurs générés ε' , ε'' et f sont randomisées (Distribuées aléatoirement).

Trois matériaux vont être caractérisés dans ce travail en utilisant la cellule SuperMit; deux liquides (L'eau et l'Éthanol) et un solide (Le PEEK (PolyEtherEtherKetone)), alors deux base de données seront créés; une pour élaborer un modèle capable de prédire la permittivité des échantillons liquides (Cellule SuperMit avec fenêtre coaxiale), et une autre pour les échantillons solides (Cellule SuperMit sans fenêtre coaxiale). Pour les échantillons liquides, la base de données est créée en variant ε' entre 1 et 100, alors que ε'' va de 0 à 100. Dans le cas des échantillons solides, le matériau testé (PEEK) a une faible permittivité relative (Surtout pour la partie imaginaire), par conséquent une distribution plus dense des exemples est requise, donc la gamme de permittivité couverte par la base de données est réduite ($1 \leq \varepsilon' \leq 10$ et $0 \leq \varepsilon'' \leq 10$).

Les exemples de la base de données sont créés en utilisant un code de calcul développé sous l'environnement Matlab où le fichier de commandes Ansys est inséré dans une boucle et lancé en mode "batch" à chaque itération par une commande spécifique de Matlab.

Pour le problème inverse, f , G et B forment les entrées, tandis que ε' et/ou ε'' forment les sorties. Le tableau II.1 illustre le processus de génération des exemples de la base de données.

Tableau II.1. Procédure de génération de la base de données.

N° d'exemple	Entrées du problème direct			Sorties du problème direct	
	ε'	ε''	f	G	B
1	ε'_1	ε''_1	f_1	G_1	B_1
2	ε'_2	ε''_2	f_2	G_2	B_2
...
...
N	ε'_N	ε''_N	f_N	G_N	B_N
	Sorties du problème inverse			Entrées du problème inverse	

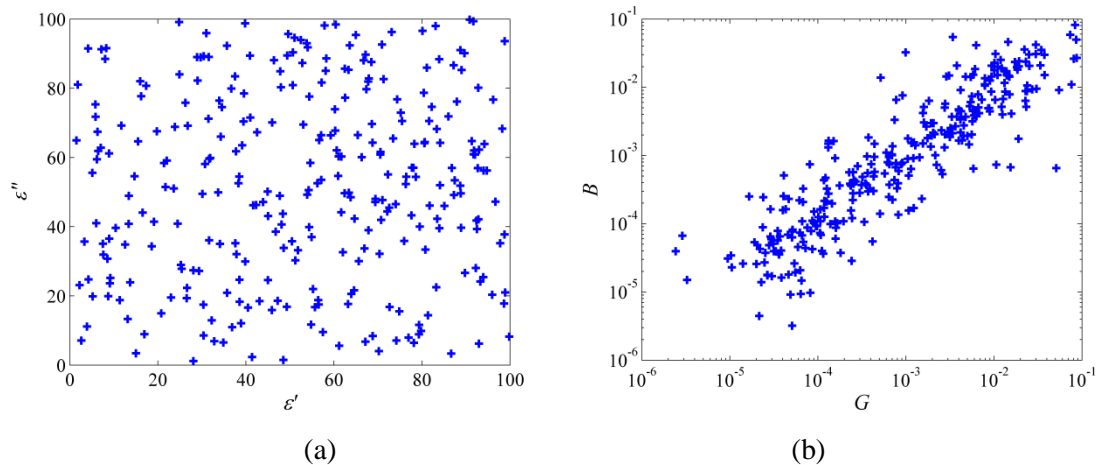


Figure II.8. Répartition des exemples de la base de données, (a) dans l'espace des paramètres recherchés, (b) dans l'espace des observations.

La figure II.8 présente un ensemble de 300 exemples de la base de données générée aléatoirement pour la cellule de mesure SuperMit. Malgré qu'elle n'offre pas une couverture idéale de l'espace des observations (Le plus important) par rapport à l'espace des paramètres recherchés, on verra plus tard (Chapitre III et IV) qu'elle donne de bons résultats avec tous les outils d'inversion.

Les variables de la base de données présentent une grande différence dans les valeurs numériques; G et B sont très faibles, f est très élevé, alors que ϵ' et ϵ'' ont des valeurs moyennes. Pour éviter l'influence de la représentation numérique et donner les mêmes poids aux différentes variables sur le modèle d'inversion, toutes les variables (ϵ' , ϵ'' , f , G et B) sont centrées sur zéro en soustrayant leurs moyennes, et normalisées en divisant leurs valeurs par leurs écarts-types.

II.5. Cellule de mesure SuperPol

La cellule de mesure SuperPol a été mise au point par Meyer [Mey96]. Elle consiste en une jonction entre un guide d'ondes coaxial et un guide circulaire. La seule différence avec la cellule SuperMit est le remplissage inhomogène du guide circulaire. Effectivement, le matériau sous teste est situé dans la continuité de l'âme centrale et maintenu par une couronne en Téflon (Figure II.9) qui a une constante diélectrique très faible ($\epsilon' = 2.1$) et des pertes négligeables ($\epsilon'' = 3.10^{-4}$). L'ensemble est relié à un analyseur de réseau. L'intérêt d'une telle configuration est qu'elle permet d'avoir un champ électrique homogène dans tout le matériau ce qui est favorable pour plusieurs applications telles que le chauffage par micro-ondes. D'ailleurs, à l'origine Meyer a développé cette cellule dans le but d'avoir un chauffage homogène en plaçant le matériau dans la zone du champ homogène et intense (Le prolongement de l'âme centrale). Cette configuration permet une

caractérisation à partir des basses fréquences jusqu'à quelques gigahertz en utilisant le standard GR900 (Rayon interne = 3.1 mm, rayon externe = 7.14 mm). Le protocole de mesure est identique à celui de la cellule SuperMit.

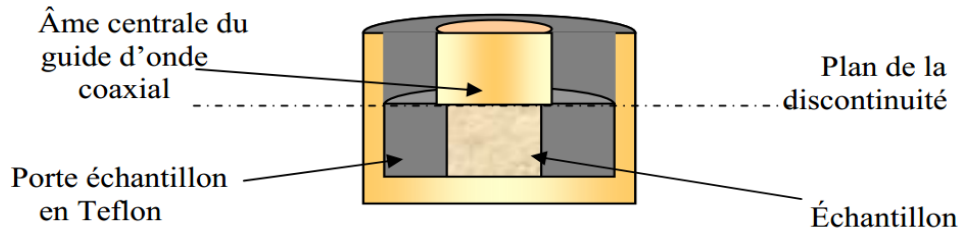


Figure II.9. Vue en coupe de la cellule de mesure SuperPol.

II.5.1. Solution analytique

La solution analytique est basée sur les travaux de Kolodziej et al. [Kolo 71] et Badot [Bado 88]. La seule différence est la présence d'une couronne d'air autour du matériau sous teste contrairement à la cellule SuperPol où le matériau est maintenu par une couronne en Teflon. Malheureusement, la fréquence maximale autorisée est loin au-dessous de la fréquence de coupure du mode fondamentale TEM : par exemple, en APC7, la limite supérieure est d'environ 10 GHz et en GR900 elle est de 2 GHz.

La procédure de l'inversion itérative est identique à celle utilisée avec la cellule SuperMit à condition de rester dans la bande autorisée. Pour aller au-delà de la fréquence limite, il faut explorer d'autres approches.

Là encore, les résultats de l'inversion itérative établis au GeePs seront utilisés comme outil de validation.

II.5.2. Modélisation numérique par FEM sous Ansys

Suivant la même démarche utilisée avec la cellule SuperMit, seule une portion angulaire (5 degrés) de la cellule SuperPol est modélisée. Le maillage tétraédrique est encore utilisé, les conditions aux limites et l'excitation restent les mêmes (Figure II.10). Là encore, la simulation par Ansys donne le coefficient de réflexion au niveau du port d'excitation, donc sa phase doit être translatée vers la discontinuité.

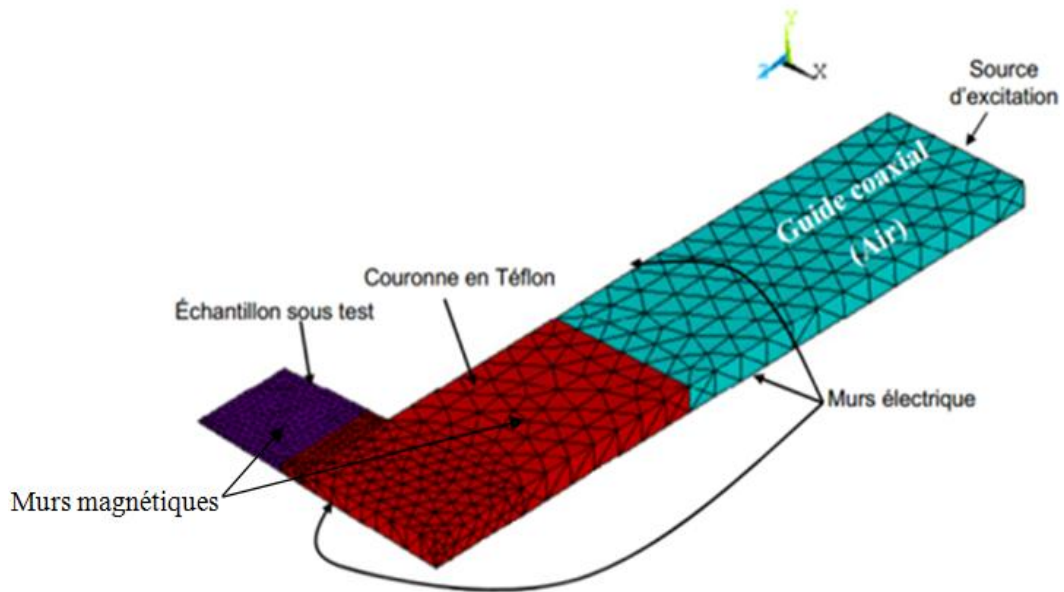


Figure II.10. Maillage par éléments finis de la cellule SuperPol avec couronne en Téflon.

La cellule SuperPol est utilisée pour caractériser l'Éthanol entre 1 MHz et 1.8 GHz, ce qui demande la création d'une autre base de données en vue d'entraîner les outils d'inversion. Contrairement à la procédure décrite dans la section II.4.4, cette base de données est créée d'une manière structurée. Supposant qu'on veut construire une base de données contenant 1000 exemples, donc chaque variable (f , ε' et ε'') va prendre 10 valeurs dans son intervalle de variation ($10^3=1000$). ε' et ε'' suivent une distribution linéaire, tandis que f suit une distribution logarithmique dans la bande de fréquences.

Les figures II.11 et II.12 présentent la répartition de 100 exemples de la base de données dans l'espace des paramètres recherchés (ε' et ε'') et l'espace des observations (G et B) pour une fréquence de 1 MHz et 1.8 GHz respectivement. Une distribution uniforme et régulière des exemples dans l'espace des observations est constatée à partir de la figure II.11.b, ce qui traduit un comportement linéaire de la sonde en basses fréquences. En hautes fréquences, un comportement non linéaire commence à se manifester avec des régions plus denses en exemples (Figure II.12.b). C'est cette non linéarité qui nous a obligés de restreindre le domaine d'apprentissage ($1 \leq \varepsilon' \leq 30$ et $0 \leq \varepsilon'' \leq 30$) pour mieux couvrir les zones de forte sensibilité (Figure II.11.a et II.12.a).

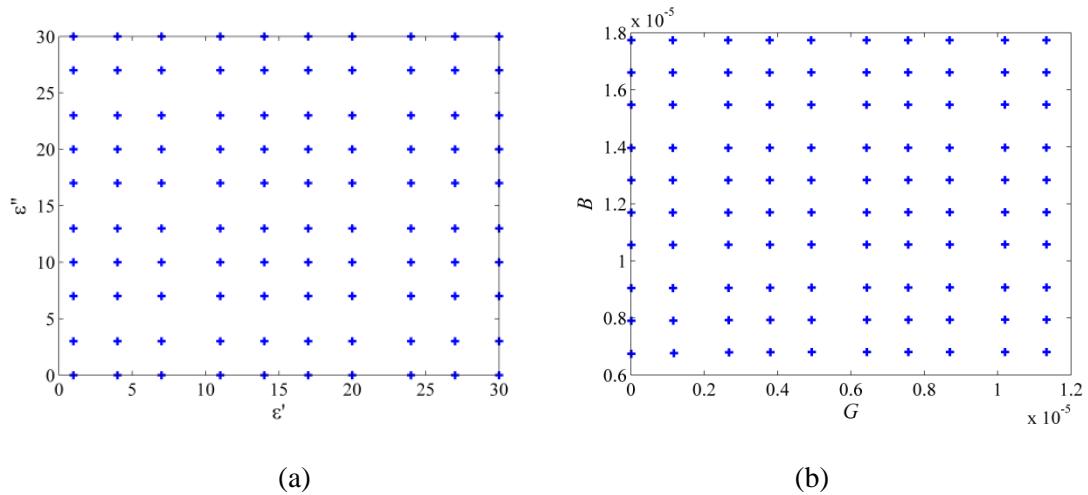


Figure II.11. Distribution de 100 exemples de la base de données structurée à 1 MHz, (a) dans l'espace des paramètres recherchés, (b) dans l'espace des observations

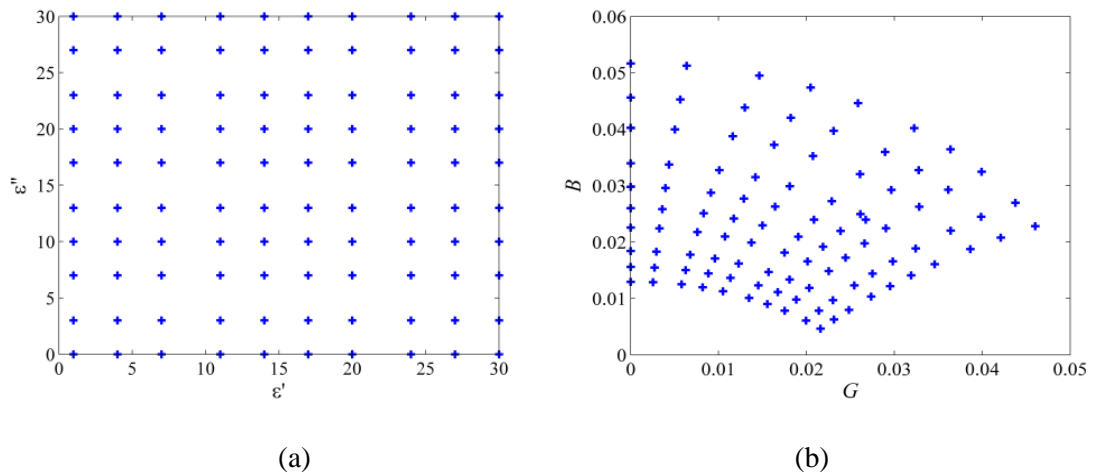


Figure II.12. Distribution de 100 exemples de la base de données structurée à 1.8 GHz, (a) dans l'espace des paramètres recherchés, (b) dans l'espace des observations

Conclusion

Dans ce chapitre, la modélisation électromagnétique d'une sonde micro-ondes (Résolution du problème direct) est présentée. Cette modélisation permet de créer les bases de données nécessaires pour réaliser une inversion directe et rapide. Deux types de résolution peuvent être envisagés : analytique et numérique. La méthode analytique de raccordement modal est brièvement décrite dans ce chapitre. Quant à la modélisation numérique, la méthode des éléments finis est largement présentée. Cependant, l'étude se déroule sous le logiciel Ansys. Bien que, les deux cellules de mesures étudiées dans cette thèse (SuperMit et SuperPol) admettent des solutions analytiques dans la bande de fréquence de fonctionnement, la procédure de création des bases de

données décrite dans ce chapitre est basée sur la simulation numérique par Ansys, et cela dans le but de réaliser une étude plus générale.

Chapitre III : Modèle Neuro-Flou pour la résolution du problème inverse

III.1. Introduction

L'idée du réseau de neurones artificiels (ANN pour Artificial Neural Networks) et de la logique floue est inspirée de l'organisation du cerveau humain et sa capacité de traiter les différents types de problèmes. Le cerveau humain est capable de s'adapter, d'apprendre et de décider même dans les cas imprécis et incertains. Pour ces raisons, beaucoup de chercheurs dans différentes disciplines se sont intéressés à comprendre son principe de fonctionnement pour pouvoir l'appliquer afin de résoudre les différents problèmes telle que l'optimisation, la classification et la reconnaissance, d'où l'apparition des ANN, la logique floue et les systèmes neuro-flous.

Dans ce chapitre, après une brève présentation des ANN, une étude approfondie de la logique floue et l'inférence floue va être livrée. Ensuite, on va aborder les systèmes neuro-flous et en particulier le réseau ANFIS (Adaptive Network based Fuzzy Inference System) qui va être utilisé pour résoudre le problème inverse.

III.2. Réseau de Neurones Artificiels

Le ANN est un modèle mathématique simplifié d'un réseau neuronal naturel. Inspiré de l'efficacité du processus cérébral, Il a été employé avec succès pour résoudre des problèmes complexes dans divers domaines d'application, y compris la classification, la reconnaissance de formes, la prédiction, l'optimisation, l'approximation de fonctions et les systèmes de contrôle. L'intérêt croissant des ANN dans ces domaines provient de la capacité d'apprentissage de ces modèles, qui manipulent des variables sans imposer de relations entre elles [Mata 11].

Un neurone formel, conçu pour imiter le rôle d'un neurone biologique (Figure III.1), est l'élément principal d'un ANN. C'est un opérateur mathématique avec des entrées (X_1, X_2, \dots, X_n) et une sortie Y , liées par une fonction de transfert f , et interconnectées par des connexions synaptiques ou des poids w . La figure III.2 illustre comment l'information est traitée par un seul neurone.

Différents modèles d'ANN ont été proposés depuis sa conception dans les années 1940, mais le Perceptron Multi Couche (MLP pour Multi Layer Perceptron) est le plus largement utilisé. Un

MLP (Figure III.3) est l'association en couches de neurones formels. La première couche reçoit les entrées et la dernière couche produit les sorties. Les couches intermédiaires n'ont aucun lien avec le monde extérieur et sont appelées couches cachées. Chaque neurone d'une couche est connecté à tous les neurones de la couche suivante. Les informations sont constamment transmises en avant d'une couche à l'autre. La connexion entrante a deux valeurs associées, une valeur d'entrée et un poids.

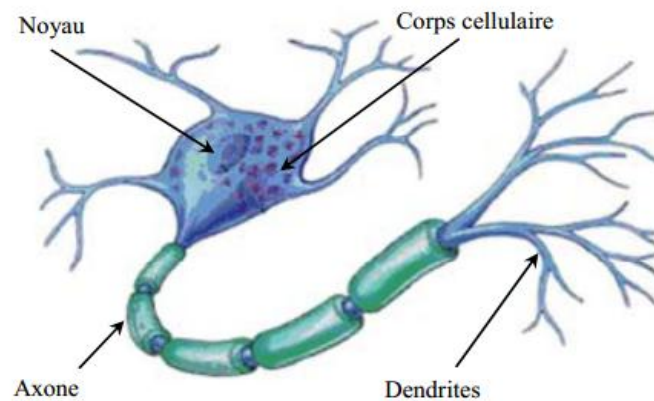


Figure III.1. Neurone biologique.

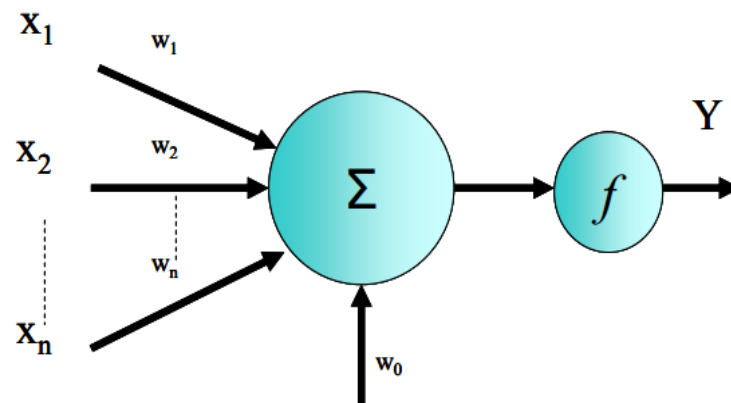


Figure III.2. Neurone formel.

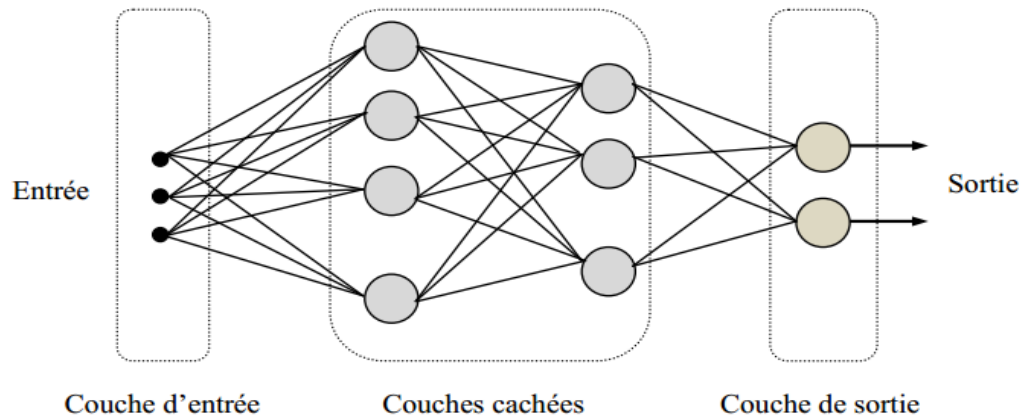


Figure III.3. Réseau MLP à deux couches cachées.

Les modèles MLP apprennent par un processus itératif, en ajustant les poids de manière à pouvoir apprendre correctement les données d'apprentissage et donc, après la phase de test, à prédire des données inconnues. La connaissance est donc stockée sous la forme d'un ensemble de poids de connexion.

L'apprentissage est le processus de modification des poids de connexion d'une manière ordonnée en utilisant une méthode d'apprentissage appropriée. Le modèle ANN utilise un mode d'apprentissage dans lequel une entrée est présentée au réseau avec une sortie, et les poids sont ajustés de sorte que le réseau tente de produire la sortie désirée. Les poids après l'opération d'entraînement contiennent des informations significatives alors qu'avant ils sont aléatoires et n'ont aucune signification.

L'algorithme de rétro-propagation, l'un des algorithmes d'apprentissage les plus célèbres pour le MLP, est une technique de descente du gradient destinée à minimiser une fonction d'erreur ou de coût, dans laquelle les poids sont ajustés avec un petit taux de variation à chaque itération. L'algorithme de minimisation le plus souvent utilisé est celui de Levenberg-Marquardt.

Dans ce travail de thèse, le réseau MLP va être utilisé seulement dans une étude comparative qui va être présentée au quatrième chapitre. Plus de détails sur ces réseaux ainsi que leurs applications pour la caractérisation micro-ondes des diélectriques peuvent être trouvés dans la référence [Acikgo 08].

III.3. Logique floue

La logique floue est la science qui s'intéresse à modéliser théoriquement les connaissances humaines pour pouvoir utiliser efficacement la capacité de raisonnement et prendre des décisions. Elle est capable de donner des résultats sous forme de recommandation pour un intervalle spécifique de l'état de sortie [Baou 14]. La logique floue est basée sur la théorie des sous-ensembles flous qui a été introduite pour la première fois en 1965 par Zadeh [Zade 65]. Le concept de sous-ensemble flou exprime l'idée d'appartenance partielle à une classe, de catégorie aux limites mal définies, de gradualité dans le passage d'une situation à une autre, dans une généralisation de la théorie classique des ensembles, admettant des situations intermédiaires entre le tout (1) et le rien (0) [Meun 07].

Un sous-ensemble flou A de l'univers de discours U est défini par une fonction d'appartenance qui associe à chaque élément x de U le degré $\mu_A(x)$, compris entre 0 et 1, avec lequel x appartient à A :

Si μ_A est la fonction d'appartenance à A , on a :

$$\forall x \in U \quad \mu_A(x) \in [0 ; 1] \quad (\text{III.1})$$

Dans le cas particulier où $\mu_A(x)$ ne prend que des valeurs égales à 0 ou 1, le sous-ensemble flou A est un sous-ensemble classique de U . Un sous-ensemble classique est donc un cas particulier des sous-ensembles flous. Dans ce cas on a :

$$\forall x \in U \quad \begin{cases} \mu_A(x) = 1 & \text{si } x \in A \\ \mu_A(x) = 0 & \text{si } x \notin A \end{cases} \quad (\text{III.2})$$

À titre d'exemple, une définition des grandes personnes, pourrait être «taille supérieure à 180 cm». Cette déclaration peut être représentée sous la forme d'un ensemble classique tel que : grand = $\{x \mid x \geq 180\}$. La figure III.4.a montre comment les ensembles classiques imposent une limite claire. Cependant, dans les ensembles flous (Figure III.4.b), cette déclaration est représentée par une fonction d'appartenance à l'ensemble flou grand $\mu(x)$ qui associe à chaque élément x , de l'univers de discours, un degré d'appartenance dans l'intervalle $[0, 1]$. Dans ce cas, on peut dire qu'une personne est grande à 95% par exemple.

Univers de discours : Gamme de taille (De 150 cm à 190 cm).

Variable linguistique : Taille.

Valeur linguistique : Grand.

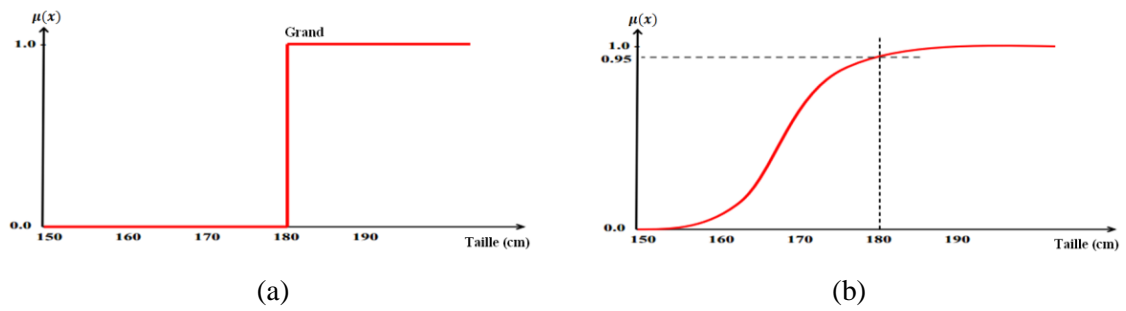


Figure III.4. Ensemble (a) classique et (b) flou d'une personne de grande taille.

La construction d'un ensemble flou dépend de deux choses : l'identification d'un univers de discours convenable et la spécification d'une fonction d'appartenance appropriée sous forme d'une courbe qui définit comment chaque point de l'espace d'entrée est projeté à un degré d'appartenance. Habituellement, la détermination des ensembles flous repose sur la connaissance des experts humains. Récemment, les méthodes intelligentes, qui peuvent dériver automatiquement les ensembles flous à l'aide des bases de données, ont été développées.

La figure III.5 donne quelques exemples de fonctions d'appartenance pouvant être utilisées pour décrire les ensembles flous.

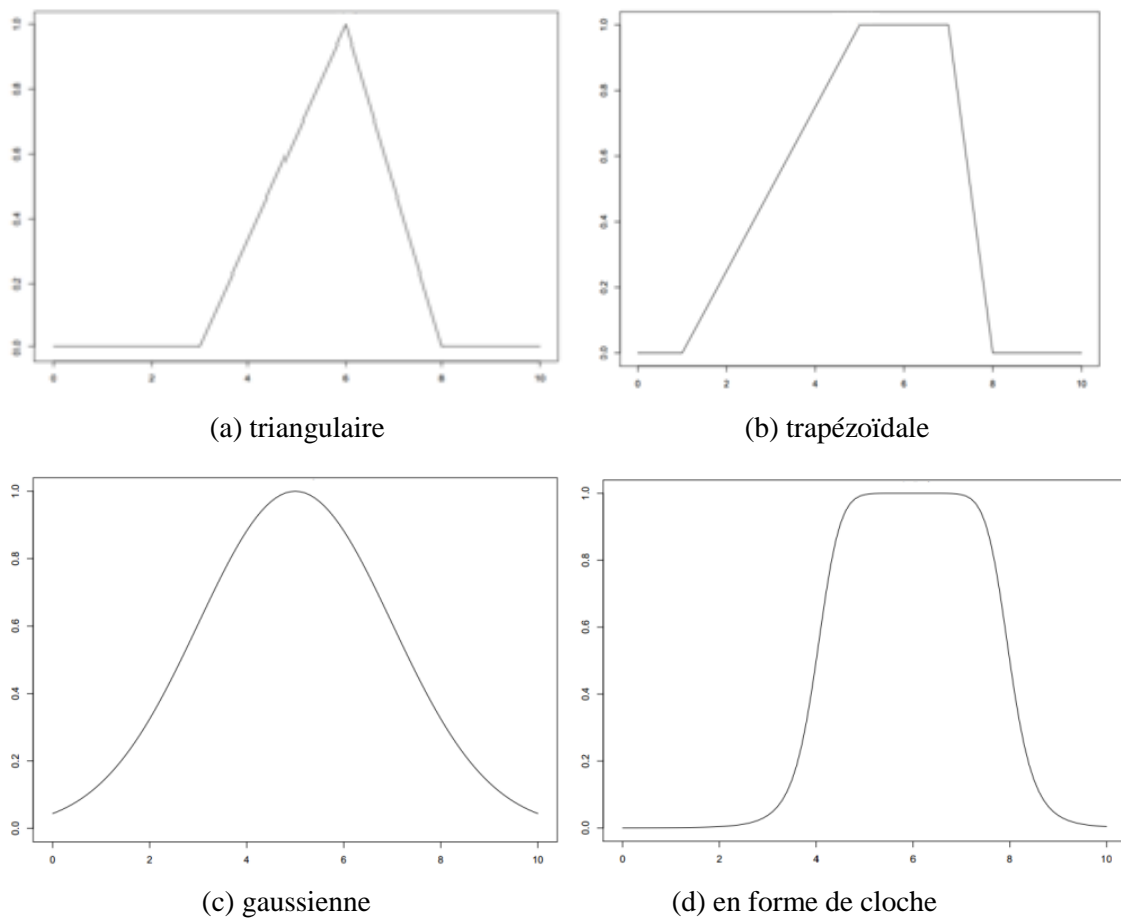


Figure III.5. Exemples de fonctions d'appartenance pour les ensembles flous.

D'une façon générale, le problème de la taille des personnes peut être traité en utilisant trois ensembles flous ; petit, moyen et grand. D'après la figure III.6, une personne de 1.78 m est à la fois membre de l'ensemble flou moyen et de l'ensemble flou grand, avec un degré d'appartenance de 0.2 et 0.8, respectivement.

Des ensembles flous sous forme d'expressions linguistiques peuvent être connectés en utilisant l'instruction SI-ALORS pour former une série de règles floues. Ces dernières s'expriment comme suit :

SI antécédent (s) ALORS conséquent (s)

Interpréter une règle SI-ALORS implique deux parties distinctes : d'abord évaluer l'antécédent, ce qui implique de fuzzifier l'entrée et ensuite appliquer ce résultat au conséquent connu sous le nom d'implication, cela s'appelle le raisonnement flou. Les règles floues peuvent avoir plusieurs antécédents, connectés par des opérateurs flous appropriés.

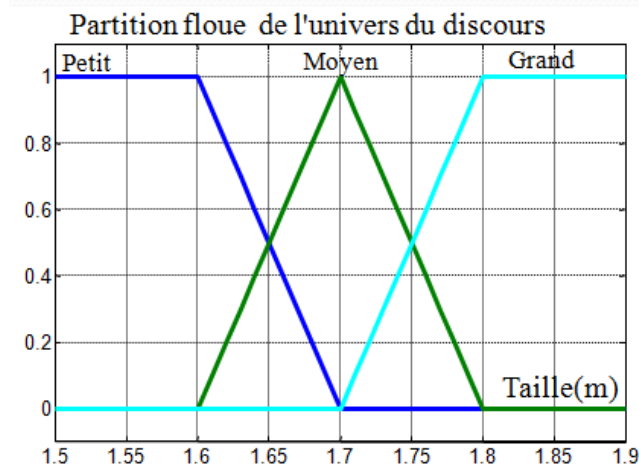


Figure III.6. Ensembles flous de personnes de tailles petite, moyenne et grande

On définit les trois opérations floues principales connues sous les noms d'intersection ou de conjonction floue (AND), d'union floue ou de disjonction (OR) et de complément flou (NOT). Les opérations de chaque opérateur peuvent être présentées comme suit :

L'union de deux ensembles flous A et B :

$$\mu_{A \cup B}(x) = \max(\mu_A(x), \mu_B(x)) \quad (\text{III.3})$$

L'intersection entre deux ensembles flous A et B :

$$\mu_{A \cap B}(x) = \min(\mu_A(x), \mu_B(x)) \quad (\text{III.4})$$

Le complément d'un ensemble flou A :

$$\mu_{\bar{A}}(x) = 1 - \mu_A(x) \quad (\text{III.5})$$

La figure III.7 montre les résultats des trois opérations avec A et B représentent les ensembles flous "petit" et "moyen" respectivement (Selon la figure III.6).

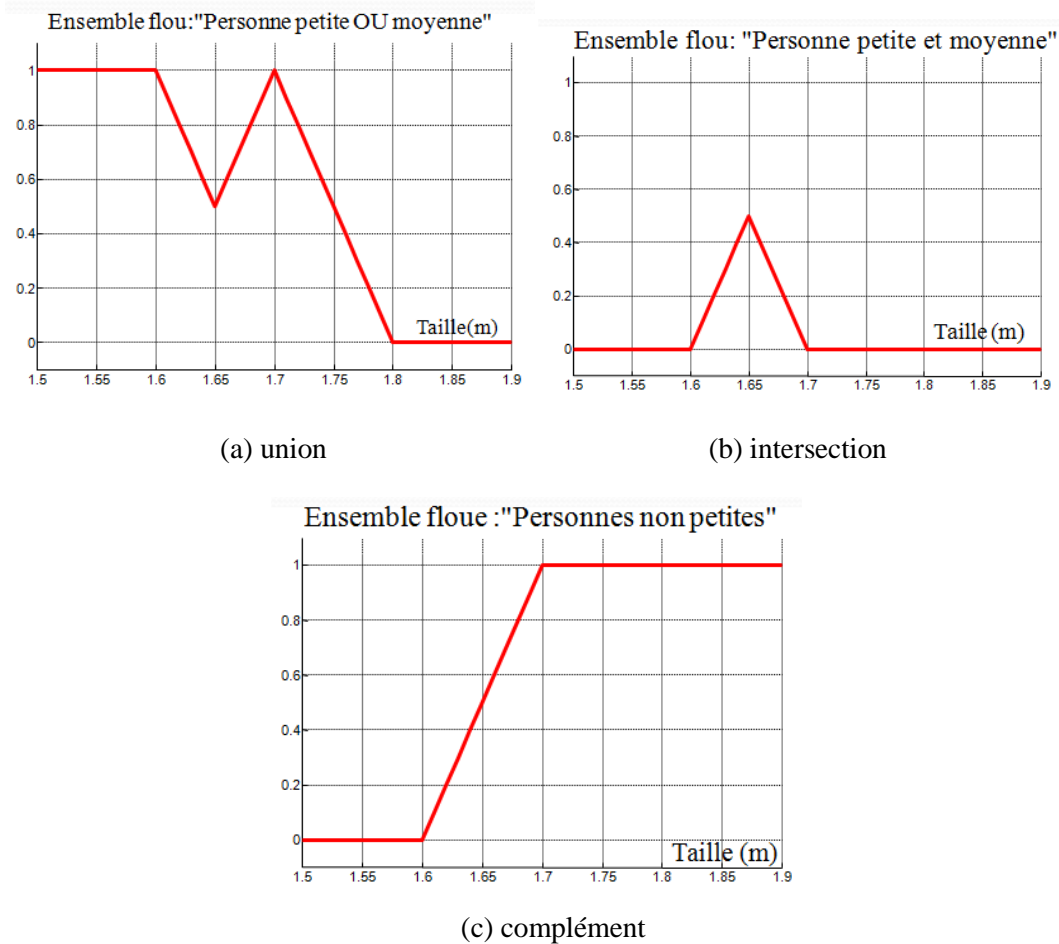


Figure III.7. Opérations sur les ensembles flous.

Une autre manière d'exprimer les opérateurs AND et OR est donnée comme suit (Opérateurs de Sugeno) :

$$\mu_{A \cap B}(x) = \mu_A(x) \times \mu_B(x) \quad (\text{III.6})$$

$$\mu_{A \cup B}(x) = \mu_A(x) + \mu_B(x) - \mu_A(x) \times \mu_B(x) \quad (\text{III.7})$$

Après avoir évalué tous les antécédents multiples qui sont liés par des opérateurs flous appropriés, toutes les parties sont calculées simultanément et résolues en une seule valeur numérique. Les conséquences peuvent également être composées de plusieurs parties qui sont ensuite agrégées en une seule sortie d'un ensemble flou. Cette sortie doit être défuzzifiée pour obtenir une valeur numérique d'une solution précise. La défuzzification fait référence à un

processus permettant d'extraire la valeur nette d'un ensemble flou en tant que valeur représentative [Jang 95]. Plusieurs méthodes de défuzzification sont suggérées dans la littérature : centre de gravité, moyenne des maximums, bissectrice de surface et de nombreuses autres méthodes.

III.4. Système d'inférence floue

Un système d'inférence floue (FIS pour Fuzzy Inference System) est un processus de cartographie d'une entrée donnée à une sortie en utilisant la théorie des ensembles flous. La figure III.8 montre un exemple de FIS avec cinq blocs fonctionnels. En général, les étapes principales effectuées dans un FIS sont les suivantes [Hamd 13] :

- Le composant de fuzzification : transforme chaque variable d'entrée classique en un degré d'appartenance basé sur les fonctions d'appartenance définies,
- Le moteur d'inférence : conduit le processus de raisonnement flou en appliquant les opérateurs flous appropriés afin d'obtenir l'ensemble flou à accumuler dans la variable de sortie,
- Le bloc de défuzzification : transforme la sortie floue en une sortie nette en appliquant une méthode de défuzzification spécifique.

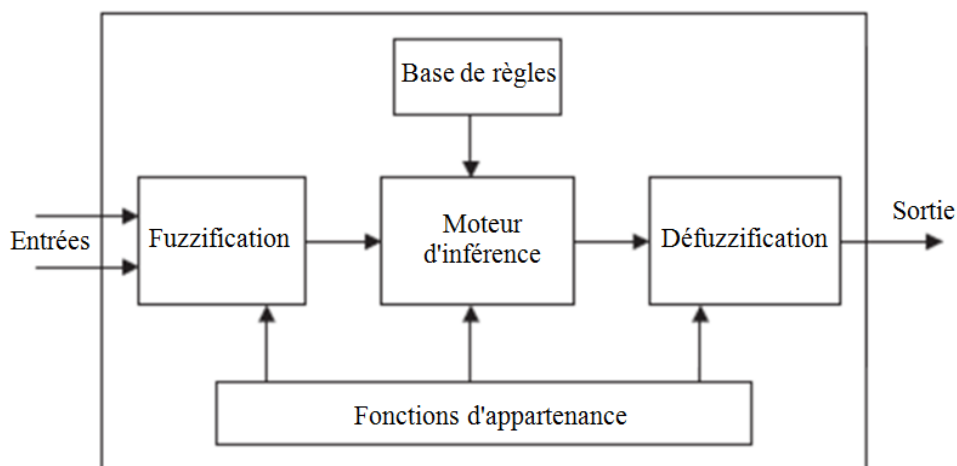


Figure III.8. Système d'Inférence Floue.

Les deux types d'inférence floue les plus couramment utilisés sont la méthode de Mamdani et la méthode de Sugeno. La différence entre ces deux méthodologies d'inférence floue est la spécification de la partie conséquente. Concernant la méthode de Mamdani [Mamd 75], les conséquents sont des ensembles flous, et la sortie finale nette est basée sur la défuzzification de la sortie floue globale en utilisant divers types de méthodes de défuzzification. En revanche, dans la méthode de Sugeno [Suge 85], les conséquents sont des nombres réels, qui peuvent être exprimés

par une fonction linéaire des entrées ou par une constante. La sortie finale est la moyenne pondérée des sorties de toutes les règles. La règle floue du style Sugeno du premier ordre a la forme suivante :

$$\text{SI } x \text{ est } A \text{ et } y \text{ est } B \text{ ALORS } z = ax + by + c \quad (\text{III.8})$$

où A et B sont des ensembles flous dans la prémisse, tandis que $ax + by + c$ est une fonction mathématique linéaire dans la conséquence. Lorsque ($a = b = 0$), on a le modèle Sugeno d'ordre zéro. La figure III.9 montre la procédure du raisonnement flou pour un modèle flou de Sugeno du premier ordre.

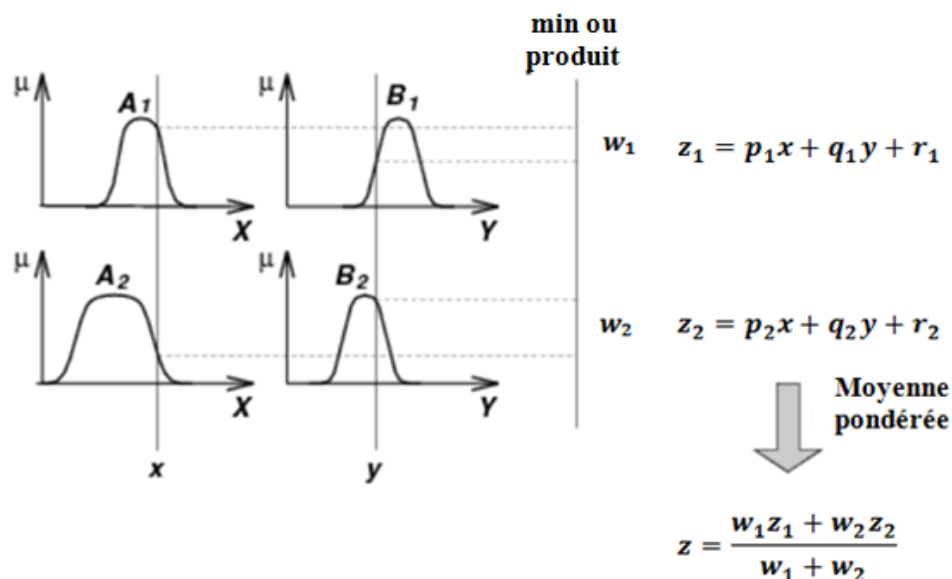


Figure III.9. Modèle flou de type Sugeno du premier ordre.

La création de règles floues est généralement assurée par un expert en se basant sur son expérience. Le partitionnement des entrées en degré d'appartenance et le choix de combinaison valide d'entrées, pour donner des valeurs typiques aux sorties, prennent généralement beaucoup de temps. Il existe plusieurs techniques qui peuvent automatiser cette tâche et réduire considérablement le temps de développement. La méthode de l'univers égalisé (Equalized universe method), la méthode de regroupement (Clustering method) et les algorithmes génétiques représentent quelques exemples de méthodes qui servent à définir les ensembles flous [Miku 05].

La méthode de l'univers égalisé, qui va être intégrée par la suite dans notre outil d'inversion, est la manière la plus simple pour former les antécédents des règles floues [Jang 93]. Les espaces d'entrées sont également partitionnés en un nombre choisi de groupes et les fonctions d'appartenance sont définies régulièrement sur la plage de variation de chaque variable d'entrée.

Pour générer les règles floues, la méthode de partitionnement de grille [Jang 95] doit être implémentée. Cette dernière fonctionne en énumérant toutes les combinaisons possibles des fonctions d'appartenance de toutes les entrées. Par exemple, un modèle flou avec deux entrées et trois fonctions d'appartenance assignées à chaque entrée donnerait $3^2 = 9$ règles floues SI-ALORS. La figure III.10 présente l'association des fonctions d'appartenance de toutes les entrées qui génère neuf règles.

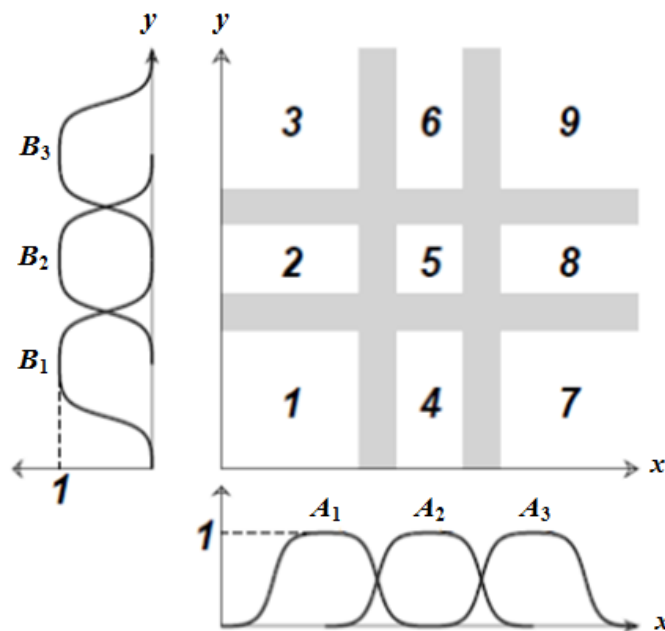


Figure III.10. Méthode de l'univers égalisé avec neuf régions floues.

La clef de succès et l'intérêt du SIF est qu'il offre la capacité de transformer des données en connaissances qui peuvent être comprises par les gens. La représentation des règles pour exprimer le comportement du système est construite à partir des connaissances d'un expert humain. Cependant, le recours à un expert pour former les règles floues (Définir les fonctions d'appartenance, les opérateurs flous et la base de connaissances) et le manque de capacité d'apprentissage sont quelques-unes des limites du SIF. L'avantage des systèmes flous, en termes de transparence, est qu'ils peuvent être combinés avec les ANN, qui fournissent une capacité d'apprentissage. La combinaison de ces techniques découle du terme de système neuro-flou.

III.5. System Neuro-Flou

Les principaux avantages des techniques floues sont l'approche naturelle de la modélisation et la bonne compréhensibilité de la description, en employant des règles linguistiques. Cependant, comme il n'y a aucune méthode formelle pour déterminer les paramètres de ces techniques (Ensembles et règles floues), l'exécution d'un système flou peut prendre beaucoup de temps. Dans

ce sens, il serait intéressant de disposer d'algorithmes permettant l'apprentissage automatique de ces paramètres [Baou 14].

L'une des méthodes qui permet de répondre à ces exigences est la théorie des réseaux de neurones qui emploie des échantillons (Données d'observation) pour l'apprentissage. Dans ce cas, la base de données remplace l'expert. La combinaison des deux techniques nous donne les systèmes neuro-flous [Raco 06].

Les règles floues codées dans le système Neuro-Flou représentent les échantillons imprécis et peuvent être vues en tant que prototypes imprécis des données d'apprentissage. Un système neuro-flou ne devrait, par contre, pas être vu comme un système expert (Flou), et il n'a rien à voir avec la logique floue dans le sens stricte du terme. On peut aussi noter que les systèmes neuro-flous peuvent être utilisés comme des approximateurs universels [Vasi 08].

Il existe plusieurs façons de combiner les ANN et les systèmes flous. En général, les systèmes neuro-flous peuvent être classés en quatre types, en fonction de la façon dont les combinaisons entre le système neuronal et le système flou sont effectuées :

- Réseau flou neuronal,
- Système neuronal/flou simultanément,
- Modèles neuro-flous coopératifs,
- Modèles neuro-flous hybrides.

Notre étude s'intéresse, particulièrement, à ce dernier type d'architecture. Dans ce cas, un système flou utilise un algorithme d'apprentissage inspiré par les réseaux de neurones pour déterminer ses paramètres grâce au traitement des observations.

Les différentes architectures des systèmes neuro-flous hybrides sont présentées dans la figure III.11. FALCON (a) et GARIC (b) interprètent la règle floue sous forme d'ensembles flous de sorties avec une structure à 5 couches. NEFCLASS (c) interprète la règle floue sous forme de composants singleton de sorties avec une structure à 3 couches et ANFIS (d) interprète la règle floue sous forme de combinaisons linéaires des variables d'entrées avec une structure à 6 couches.

Dans notre étude c'est l'architecture ANFIS qui va être adoptée. Elle constitue l'objet du reste de ce chapitre.

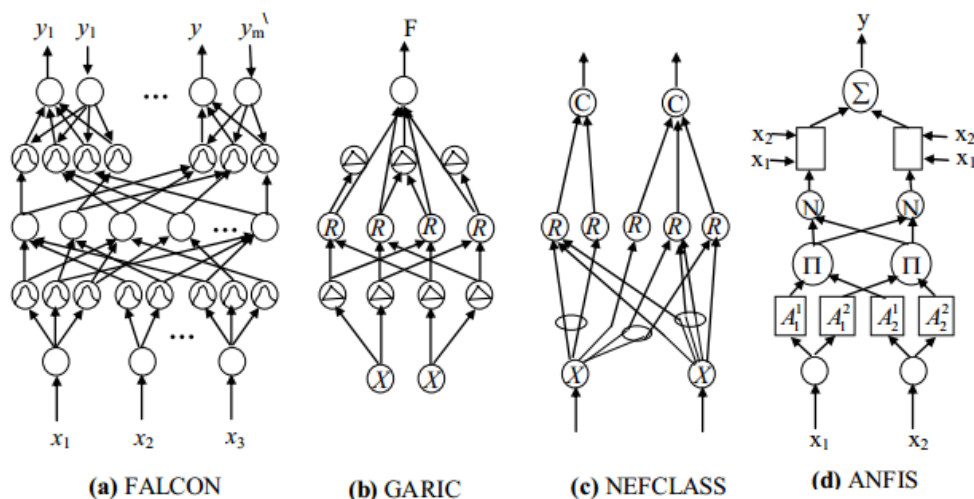


Figure III.11. Différentes architectures des systèmes neuro-flous hybrides.

III.6. ANFIS

ANFIS, introduit pour la première fois par Jang [Jang 93], est une combinaison hybride de FIS et ANN implémentée dans le cadre d'un système adaptatif pour illustrer le comportement entrée-sortie de systèmes complexes. L'idée derrière cette combinaison est de concevoir un modèle qui utilise un système flou pour représenter les connaissances d'une manière interprétable, qui a la capacité d'apprentissage dérivée des ANN et qui peut ajuster les règles linguistiques directement à partir des données afin d'améliorer les performances du système. Ce système réalise une approximation linéaire de la variable de sortie en partitionnant l'espace d'entrée en plusieurs sous-ensembles flous, chacun est régi par une règle SI-ALORS (If-Then) floue. Les paramètres des règles floues sont ajustés par un algorithme d'apprentissage hybride qui combine l'algorithme de descente du gradient (GD pour gradient descent) et l'algorithme d'estimation des moindres carrés (LSE pour Least Squares Estimator) [Jang 93, 97] en analysant tous les exemples de la base d'apprentissage. L'avantage principal d'une telle approche est qu'elle converge beaucoup plus rapidement, puisqu'elle réduit les dimensions de l'espace de recherche de la méthode de rétro-propagation utilisée dans les ANN à propagation avant (FFNN pour Feed-Forward Neural network) [Jang 97]. Ce modèle donne de très bons résultats en poursuite de trajectoire, en approximation non linéaire, en commande dynamique et en traitement du signal [Vasi 08].

ANFIS a besoin d'une base de données relativement petite (Le nombre d'exemples dans la base d'apprentissage doit être au moins 5 fois le nombre total des paramètres d'ANFIS). Ceci est un avantage très important dans les cas où les données disponibles sont limitées telles que la modélisation de la survie dans les recherches médicales ou dans les cas où des mesures expérimentales sont nécessaires pour créer les bases de données. Une base de données réduite

signifie une convergence rapide et une réduction du temps d'élaboration du modèle. Cependant, il a quelques limitations ; ANFIS ne supporte que les systèmes de type Sugeno avec les contraintes suivantes [Karr 00] :

- Systèmes de type Sugeno du premier ordre ou d'ordre zéro (Fonction d'appartenance de sortie linéaire ou constante),
- Une seule sortie, obtenue à l'aide de la méthode de défuzzification de la moyenne pondérée,
- Le poids de chaque règle est toujours un.

De plus, ANFIS souffre d'un problème connu sous le nom de "la malédiction de la dimensionnalité". Selon la procédure de partitionnement de grille (Grid partition), un système flou avec n entrées et m fonctions d'appartenance pour chaque entrée donne m^n règles. Ce problème peut être résolu en utilisant des stratégies de partitionnement plus complexes [Jang 95, Weia 07]. Le partitionnement de grille est celui choisi dans notre étude car nous n'avons que 3 entrées et 2 fonctions d'appartenance affectées à chaque entrée.

L'architecture de base d'ANFIS contient un FFNN à cinq couches, comme le montre la figure III.12. Les nœuds de chaque couche ont la même fonctionnalité. Par exemple, pour un système à deux entrées (x, y) et une sortie (z) avec deux fonctions d'appartenance affectées à chaque entrée. Les règles floues sont définies sous la forme suivante [Jang 93, Jang et Sun 93] :

$$\text{if } (x \text{ is } A_i) \text{ and } (y \text{ is } B_i) \text{ then } (f_i = p_i x + q_i y + r_i) \quad (\text{III.9})$$

où A_i et B_i sont les ensembles flous (Variables linguistiques), f_i est la sortie dans la région floue spécifiée par la règle floue i . p_i , q_i et r_i sont les paramètres conséquents qui sont ajustés pendant le processus d'entraînement [Selm 14].

Le processus d'ajustement des paramètres en utilisant un algorithme d'apprentissage en deux passes est maintenant décrit en détails.

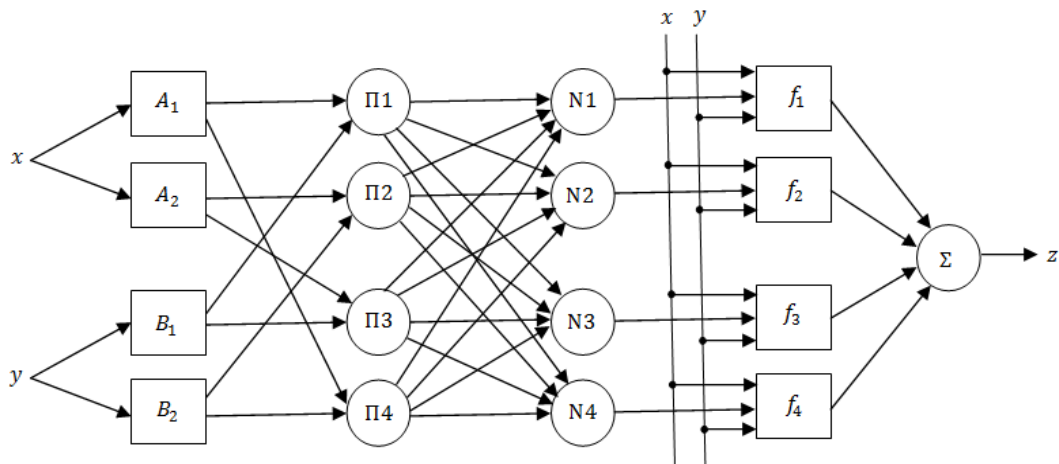


Figure III.12. Architecture d'un réseau ANFIS à quatre règles.

III.7. Stratégie d'apprentissage

III.7.1. Passe avant

Dans la passe avant, les sorties des neurones sont calculées couche par couche et les paramètres conséquents sont identifiés par LSE à la couche 4, pour obtenir la sortie unique finale à la couche 5. A ce stade, les paramètres antécédents (Fonctions d'appartenance) sont supposés être constants le long du cycle courant pour tous les exemples de la base d'entraînement [Hine 97].

L'opération de la passe avant sur chaque couche est décrite ci-dessous :

La couche 1 : nœuds de fuzzification. Tous les nœuds sont adaptatifs. Les sorties de la couche 1 représentent les degrés d'appartenance floue des entrées, qui sont données par :

$$\begin{cases} O_i^1 = \mu_{A_i}(x) & i = 1,2 \\ O_i^1 = \mu_{B_{i-2}}(y) & i = 3,4 \end{cases} \quad (\text{III.10})$$

Par exemple, si une fonction d'appartenance en forme de cloche (Bell-shaped) est utilisée, alors :

$$\mu_{A_i}(x) = \frac{1}{1 + \left\{ \left(\frac{x - c_i}{a_i} \right)^2 \right\}^{b_i}} \quad (\text{III.11})$$

où a_i , b_i et c_i sont appelés les paramètres prémisses.

La couche 2 : nœuds de règles. Les nœuds de cette couche sont des nœuds fixes, ce qui signifie qu'ils dépendent des sorties de la couche précédente. Chaque sortie de nœud représente la force de

déclenchement d'une règle (Ou le poids de la règle) en tant que produit des degrés d'appartenance.

Donc, on a :

$$O_i^2 = w_i = \mu_{A_i}(x) \mu_{B_i}(y), \quad i=1,4 \quad (\text{III.12})$$

La couche 3 : nœuds de normalisation. Tous les nœuds sont également fixes, cette couche normalise les résultats entrants de la couche précédente. Par conséquent, la sortie du nœud i est égale au rapport entre le poids de la règle i et la somme des poids de toutes les règles tel que :

$$O_i^3 = \bar{w}_i = \frac{w_i}{\sum_{i=1}^4 w_i}, \quad i=1,4 \quad (\text{III.13})$$

La couche 4 : nœuds de défuzzification. Chaque nœud calcule la contribution de chaque règle à la sortie globale. Tous les nœuds sont adaptatifs, et la fonction du nœud est le produit du poids normalisé par un polynôme du premier ordre (Ou d'ordre zéro) des entrées (x, y) .

$$O_i^4 = \bar{w}_i f_i = \bar{w}_i (p_i x + q_i y + r_i), \quad i=1,4 \quad (\text{III.14})$$

où p_i , q_i et r_i représentent les paramètres conséquents qui peuvent être calculés en utilisant LSE.

La procédure de calcul est la suivante :

Premièrement, la sortie cible (Sortie désirée) est calculée par :

$$t = \sum_i \bar{w}_i f_i = \bar{w}_1 (p_1 x + q_1 y + r_1) + \dots + \bar{w}_4 (p_4 x + q_4 y + r_4) \quad (\text{III.15})$$

Si on a P paires entrée-sortie dans la base d'entraînement, les P équations linéaires peuvent être exprimées sous la forme matricielle suivante :

$$\begin{bmatrix} t_1 \\ \cdot \\ \cdot \\ \cdot \\ t_i \\ \cdot \\ \cdot \\ \cdot \end{bmatrix}_{P \times 1} = \begin{bmatrix} \bar{w}_{1,1} x_1 & \bar{w}_{1,1} y_1 & \bar{w}_{1,1} & \cdot & \cdot & \cdot & \bar{w}_{4,1} x_1 & \bar{w}_{4,1} y_1 & \bar{w}_{4,1} \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ \bar{w}_{1,i} x_i & \bar{w}_{1,i} y_i & \bar{w}_{1,i} & \cdot & \cdot & \cdot & \bar{w}_{4,i} x_i & \bar{w}_{4,i} y_i & \bar{w}_{4,i} \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ \bar{w}_{1,P} x_P & \bar{w}_{1,P} y_P & \bar{w}_{1,P} & \cdot & \cdot & \cdot & \bar{w}_{4,P} x_P & \bar{w}_{4,P} y_P & \bar{w}_{4,P} \end{bmatrix}_{P \times 12} \times \begin{bmatrix} p_1 \\ q_1 \\ r_1 \\ \cdot \\ \cdot \\ \cdot \\ p_4 \\ q_4 \\ r_4 \end{bmatrix}_{12 \times 1} \quad (\text{III.16})$$

D'une façon générale, on peut écrire :

$$T_{P \times 1} = A_{P \times m^n(n+1)} X_{m^n(n+1) \times 1} \quad (\text{III.17})$$

où n représente le nombre de variables d'entrée, m le nombre de fonctions d'appartenance par entrée, et m^n représente le nombre maximal de règles utilisées [Pill 14].

Comme P (Nombre de lignes de A) est habituellement supérieur à $m^n(n+1)$ (Nombre de lignes de X), il s'agit d'un problème surdéterminé et il n'y a pas généralement de solution exacte (C'est à dire: $X = inv(A) \times T$). Au lieu d'une inversion directe, une technique de régression telle que LSE de X , X^* , est sollicitée pour minimiser l'erreur au carré $\|AX - T\|^2$. Avec :

$$X^* = (A^T A)^{-1} A^T T \quad (\text{III.18})$$

où A^T est le transposé de A et $(A^T A)^{-1} A^T$ est la pseudo-inverse de A [Alot 08, Sidd 13].

Ensuite, le vecteur X^* des paramètres conséquents $\{p_i, q_i, r_i\}$, calculé en utilisant la sortie désirée T , sera utilisé pour estimer la sortie de la couche 5.

La couche 5 : nœud de sortie. Il n'y a qu'un seul nœud fixe. Ce nœud calcule la sortie globale comme la somme de tous les signaux entrants. Ceci s'exprime comme suit :

$$O_i^5 = \sum_{i=1}^{m^n} \bar{w}_i f_i = \frac{\sum_{i=1}^{m^n} w_i f_i}{\sum_{i=1}^{m^n} w_i} \quad (\text{III.19})$$

D'une autre manière,

$$O_i^5 = A \times X^* \quad (\text{III.20})$$

Maintenant, on peut calculer l'erreur pour l'entrée p des données d'entraînement, à partir de l'équation ci-dessous.

$$E_p = (T_p - O_p)^2 \quad (\text{III.21})$$

où T_p est la sortie désirée et O_p la sortie d'ANFIS.

III.7.2. Passe arrière

Dans la passe arrière, les signaux d'erreur sont rétro-propagés et les paramètres antécédents sont mis à jour suivant la règle de la chaîne (Règle de dérivation en chaîne d'une fonction composée) tandis que les paramètres conséquents sont constants.

Selon la GD, le taux d'erreur $\frac{\partial E_p}{\partial O}$ pour l'exemple d'entraînement numéro p et pour chaque sortie O peut être calculé à partir de l'équation III.21 comme suit :

$$\frac{\partial E_p}{\partial O_{i,p}} = -2(T_{i,p} - O_{i,p}) \quad (\text{III.22})$$

Les taux d'erreur pour un nœud dans la position numéro i d'une couche interne peuvent être dérivés par la règle de la chaîne comme étant une combinaison linéaire des taux d'erreurs des nœuds dans la couche suivante (N nœuds). Alors, on a :

$$\frac{\partial E_p}{\partial O_{i,p}} = \sum_{m=1}^N \frac{\partial E_p}{\partial O_{m,p}} \frac{\partial O_{m,p}}{\partial O_{i,p}} \quad (\text{III.23})$$

Maintenant si α (a_i , b_i ou c_i dans notre cas) est un paramètre du réseau adaptatif en question, nous avons :

$$\frac{\partial E_p}{\partial \alpha} = \sum_{O^* \in S} \frac{\partial E_p}{\partial O^*} \frac{\partial O^*}{\partial \alpha} \quad (\text{III.24})$$

où S est l'ensemble des nœuds dont les sorties dépendent de α . La dérivée de l'erreur globale E par rapport à α est alors :

$$\frac{\partial E}{\partial \alpha} = \sum_{p=1}^P \frac{\partial E_p}{\partial \alpha} \quad (\text{III.25})$$

Cette erreur peut être exprimée selon l'architecture d'ANFIS comme suit :

$$\frac{\partial E}{\partial \alpha} = \frac{\partial E}{\partial O_5} \frac{\partial O_5}{\partial O_4} \frac{\partial O_4}{\partial O_3} \frac{\partial O_3}{\partial O_2} \frac{\partial O_2}{\partial O_1} \frac{\partial O_1}{\partial \alpha} \quad (\text{III.26})$$

La dérivée de la sortie à la couche 5 par rapport à la sortie de la couche 4 est :

$$\frac{\partial O_5}{\partial O_4} = \frac{\partial(\sum f_i \bar{w}_i)}{\partial(f_i \bar{w}_i)} = 1 \quad (\text{III.27})$$

La dérivée de la couche 4 par rapport à la couche 3 est :

$$\frac{\partial O_4}{\partial O_3} = \frac{\partial(f_i \bar{w}_i)}{\partial(\bar{w}_i)} = f_i \quad (\text{III.28})$$

où f_i est l'ensemble des paramètres conséquents obtenu à partir de (III.18).

La dérivée de la couche 3 par rapport à la couche 2 est :

$$\frac{\partial O_3}{\partial O_2} = \frac{\partial}{\partial w_i} \left(\frac{w_i}{\sum_{k=1}^n w_k} \right) = \frac{\sum_{k=1}^n w_k - w_i}{\left(\sum_{k=1}^n w_k \right)^2} \quad (\text{III.29})$$

où n est le nombre total de règles dans la deuxième couche.

La dérivée de la couche 2 par rapport à la couche 1 est :

$$\frac{\partial O_2}{\partial O_1} = \frac{\partial}{\partial A_m} \left(\prod_{A_k \in \mathcal{R}(A_m), A_k \neq A_m} A_k \right) \quad (\text{III.30})$$

où $A_k \in \mathcal{R}(A_m)$ correspond aux degrés d'appartenance qui font partie intégrante de la règle contenant l'ensemble flou A_m .

Finalement, si nous avons des fonctions d'appartenance en forme de cloche (Équation III.11), les dérivées par rapport aux prémisses seront [Wang 06] :

$$\begin{cases} \frac{\partial O_1}{\partial a} = \frac{2b}{a} O_1(1-O_1) \\ \frac{\partial O_1}{\partial b} = \begin{cases} -2 \ln \left| \frac{x-c}{a} \right| O_1(1-O_1) & \text{if } x \neq c \\ 0 & \text{if } x = c \end{cases} \\ \frac{\partial O_1}{\partial c} = \begin{cases} \frac{2b}{x-c} O_1(1-O_1) & \text{if } x \neq c \\ 0 & \text{if } x = c \end{cases} \end{cases} \quad (\text{III.31})$$

Donc, lors du passage avant de l'itération suivante, le paramètre a (De même pour b et c) va prendre cette nouvelle valeur :

$$a_{new} = a_{old} - \eta \frac{\partial E}{\partial a} \quad (\text{III.32})$$

Le taux d'apprentissage η est donné par l'expression suivante :

$$\eta = \frac{k}{\sqrt{\sum_{\alpha} \left(\frac{\partial E}{\partial \alpha} \right)^2}} \quad (\text{III.33})$$

où k est le pas (The step size), c'est à dire la longueur de chaque transition de gradient dans l'espace des paramètres. Habituellement, nous pouvons changer la valeur de k pour accélérer la convergence [Alot 08].

Puisque les paramètres antécédents et conséquents sont découplés à l'aide d'un algorithme d'apprentissage hybride, une accélération supplémentaire de l'apprentissage est possible en utilisant d'autres variantes de la méthode du gradient ou d'autres techniques d'optimisation [Sidd 13].

III.8. Implémentation d'ANFIS

Avant d'utiliser ANFIS, la détermination du type et du nombre de fonctions d'appartenance attribués à chaque entrée a un effet crucial sur les résultats d'inversion. Selon la méthode de la validation simple (Split-sample procedure) [Acik 07], la base de données déjà créée (Section II.4.4) est divisée en deux sous-ensembles différents ; 2/3 pour l'apprentissage et 1/3 pour la validation.

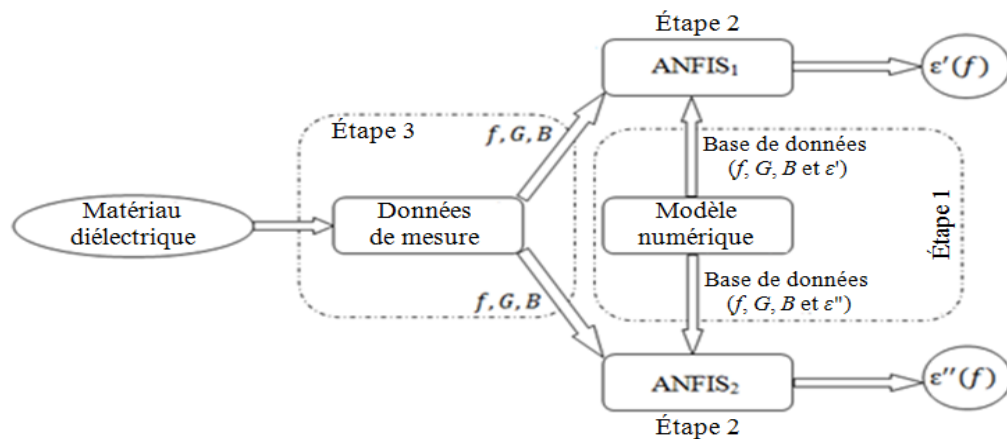
La base d'apprentissage doit respecter les conditions suivantes :

- Balayer régulièrement tout le domaine d'étude dans l'espace des sorties du problème direct. Cela est très important lorsqu'on a affaire à une fonction fortement non linéaire. Des études récentes montrent l'efficacité des bases adaptatives pour renforcer l'apprentissage [Acikg 08 ; Chel 14, 16].
- Le nombre d'exemples doit être au moins 5 fois le nombre total des paramètres (Prémises et conséquents) pour obtenir une bonne généralisation. Cependant, au-delà de cette taille, c'est seulement la pertinence des données, que la base contient, qui s'avère critique [Wang 04].

ANFIS a été entraîné et testé plusieurs fois en utilisant divers nombres de fonctions d'appartenance. À chaque fois, 2/3 exemples, de la base de données, ont été choisis d'une façon aléatoire pour former la base d'apprentissage et 1/3 exemples pour la base de validation. En analysant la précision des tests de quelques paires apprentissage-validation (Environ 10), le nombre optimal de fonctions d'appartenance est celui correspondant à l'estimation d'erreur de généralisation minimale, obtenue sur l'ensemble de validation.

ANFIS est un modèle multi-entrées avec sortie unique. Par conséquent, puisque le problème proposé manipule trois entrées et deux sorties, il est nécessaire d'utiliser deux modèles ANFIS identiques pour l'estimation des paramètres, l'un fournit ε' et l'autre fournit ε'' .

La figure III.13 montre l'organigramme qui résume les différentes étapes de l'étude.

Figure III.13. Procédure d'extraction des paramètres ε' et ε'' .

III.9. Résultats et discussions

III.9.1. Application sur les données de simulation

III.9.1.1. Entrées : f , G et B

Suivant la procédure décrite dans la section précédente, les deux modèles ANFIS optimaux élaborés sont composés, chacun, d'un réseau avec 2 fonctions d'appartenance en forme de cloche pour chaque variable d'entrée (8 règles floues SI-ALORS du type Takagi-Sugeno [Suge 85]) et une fonction d'appartenance linéaire associée à la sortie. Donc, un total de 50 paramètres linéaires et non linéaires qui nécessite une base de données qui contient au minimum 375 exemples ; 250 pour l'apprentissage et 125 pour la validation. La fonction d'appartenance en forme de cloche est choisie en raison de sa combinaison entre rapidité de convergence et moindre effort de calcul [Chik 14].

La qualité des résultats de chaque modèle ANFIS est évaluée par la mesure de la racine carrée de l'erreur quadratique moyenne ($RMSE$ pour Root Mean Square Error) et le coefficient de corrélation linéaire (R). Ce dernier est fréquemment utilisé pour évaluer la relation linéaire entre les valeurs estimées (Prédites par ANFIS) et désirées (Incluses dans la base de données). Il est compris entre 0 et 1. Plus il est proche de 1, plus le modèle est de bonne qualité. Alors que la $RMSE$ mesure l'écart entre la droite de régression des valeurs désirées et les valeurs prédites par le modèle. Plus la $RMSE$ est faible, plus le modèle est de bonne précision. Dans ce cas, la $RMSE$ et le R peuvent être définis comme suit [Haci 10] :

$$RMSE = \sqrt{\frac{1}{N} \sum_{i=1}^N (y_i - \hat{y}_i)^2} \quad (\text{III.34})$$

$$R = \frac{\sum_{i=1}^N (y_i - \bar{y}_i)(\hat{y}_i - \bar{\hat{y}}_i)}{\sqrt{\sum_{i=1}^N (y_i - \bar{y}_i)^2 \sum_{i=1}^N (\hat{y}_i - \bar{\hat{y}}_i)^2}} \quad (\text{III.35})$$

où y_i la valeur des propriétés du matériau, \hat{y}_i la valeur de la prédiction par ANFIS, \bar{y}_i et $\bar{\hat{y}}_i$ les moyennes de y_i et \hat{y}_i respectivement et N le nombre d'exemples dans la base de données.

Les tableaux III.1 et III.2 résument les performances de chaque modèle.

Tableau III.1. Performances des deux modèles ANFIS (Échantillons liquides).

	Sortie	Base d'apprentissage		Base de validation	
		<i>RMSE</i>	<i>R</i>	<i>RMSE</i>	<i>R</i>
ANFIS ₁	ε'	5.0525×10^{-1}	0.8717453	5.3545×10^{-1}	0.8427342
ANFIS ₂	ε''	1.6250×10^{-1}	0.9860151	2.3359×10^{-1}	0.9710715

Tableau III.2. Performances des deux modèles ANFIS (Échantillons solides).

	Sortie	Base d'apprentissage		Base de validation	
		<i>RMSE</i>	<i>R</i>	<i>RMSE</i>	<i>R</i>
ANFIS ₁	ε'	3.8314×10^{-1}	0.9615412	4.1928×10^{-1}	0.9095823
ANFIS ₂	ε''	2.8610×10^{-1}	0.9575002	3.4085×10^{-1}	0.9493843

A partir des tableaux III.1 et III.2, la mise en œuvre d'ANFIS en utilisant f , G et B comme entrées a conduit à des résultats jugés insatisfaisants (Voir même médiocres). Nous suggérons que cette situation est liée à l'approximation locale d'ANFIS. En effet, d'autres auteurs [Jang et Sun 93] ont démontré l'équivalence entre ANFIS et les réseaux à base de fonctions radiales (RBF-NN, pour Radial Basis Function Neural Network) (Connues comme réseaux à approximation locale) sous certaines restrictions mineures, contrairement à MLP connu comme réseau à approximation globale. La distinction entre locale et globale est faite par l'étendue de la surface d'entrée couverte par l'approximation de la fonction [Hine 97]. Effectivement, ANFIS avec f , G et B comme entrées approxime bien les paramètres ε' et ε'' si la base de données est divisée suivant la fréquence f (Par exemple la première bande de fréquence entre 1 MHz et 10 MHz, la seconde entre 10 MHz et 1.8 GHz), mais cette solution n'est pas pratique.

Maintenant, la problématique est de trouver les entrées pertinentes capables de prédire la permittivité diélectrique relative avec une bonne précision.

III.9.1.2. Sélection des entrées optimales

Afin de sélectionner les 3 entrées pertinentes pour chaque modèle, la méthode de Jang [Jang 96, Syam 15, Vasi 08] a été utilisée. Elle est basée sur l'hypothèse que le modèle ANFIS, avec la plus petite erreur d'apprentissage après une époque, a le plus grand potentiel pour atteindre la plus faible erreur après avoir terminé l'entraînement.

Par exemple, si nous avons un problème avec 10 entrées et nous voulons trouver les 3 entrées les plus influentes sur la sortie, nous devons construire $C_3^{10} = 120$ modèles ANFIS à 3 entrées (Chacun avec une combinaison différente). Ensuite, chacun sera entraîné une seule fois et celui qui a la combinaison d'entrées avec la plus petite erreur d'apprentissage sera sélectionné. Notons que l'apprentissage de tous ces 120 modèles, pour une époque, implique moins de temps de calcul que l'apprentissage d'un seul modèle ANFIS pour 120 époques, car ils sont entraînés uniquement avec la passe avant de l'algorithme d'apprentissage (LSE seulement). Donc il semble que la procédure de sélection d'entrée n'est pas vraiment coûteuse en temps de calcul.

Dans notre cas, on a seulement trois entrées (f , G et B). L'idée est de générer mathématiquement de nouvelles entrées comme combinaisons d'anciennes entrées afin d'intégrer les corrélations possibles entre eux. Chaque nouvelle entrée x s'écrit comme suit :

$$x = f^i * G^j * B^k \quad \text{avec } i = -1, \dots, 1, j = -1, \dots, 1 \text{ et } k = -1, \dots, 1 \quad (\text{III.36})$$

Ici la combinaison lorsque $i = j = k = 0$ est exclue.

Donc, on obtient 26 nouvelles entrées ($1/fGB$, $1/fG$, ..., f , ..., G , ..., B , ..., fGB , ... etc), ça veut dire $C_3^{26} = 2600$ modèles ANFIS à trois entrées.

Dans la figure III.14, l'évolution de la $RMSE$ sur la base d'apprentissage pour toutes les combinaisons possibles est montrée. Il est raisonnable que la combinaison avec la plus petite erreur d'apprentissage correspondante aux entrées B/f , B et G (La combinaison numéro 1384 avec le point rouge) soit utilisée pour prédire ε' des échantillons liquides, par exemple. Notons ici que la procédure de sélection d'entrées conduit à une combinaison différente après chaque exécution en raison de l'initialisation aléatoire d'ANFIS, mais toujours la précision requise est acquise, et souvent la convergence du modèle optimal est atteinte après quelques itérations. Les résultats concernant les échantillons liquides, présentés ci-dessous, sont obtenus en utilisant (B/f , B , G) et (G/f , GB , $1/GB$) comme entrées pour prédire ε' et ε'' respectivement, alors que (B/f , G/B , fB) et (G/f , G/B , GB) sont utilisés pour prédire ε' et ε'' , respectivement, pour les échantillons solides. Il faut noter que l'architecture optimale du réseau ANFIS obtenue précédemment (2 fonctions

d'appartenance en forme de cloche pour chaque variable d'entrée et une fonction d'appartenance linéaire associé à la sortie) est retenue pour chercher les entrées optimales.

A titre d'exemple, la figure III.15 montre l'évolution de la *RMSE* sur la base d'apprentissage et de validation du modèle destiné à prédire ε' des liquides. D'après les deux courbes, l'apprentissage commence avec une erreur très faible à la première itération pour atteindre les performances optimales à l'itération 804 avec une *RMSE* de validation minimale. Ces résultats (ainsi que tous les autres résultats) sont obtenus avec un pas d'apprentissage de 0.01, un taux de décroissance du pas d'apprentissage de 0.9 et un taux d'incrément du pas d'apprentissage de 1.1.

La figure III.16 présente la répartition finale (Itération 803) des fonctions d'appartenance sur la plage de variation de chaque entrée normalisé (*B/f*, *B* et *G*) du modèle ANFIS destiné à la prédiction de ε' des échantillons liquides.

Les tableaux III.3 et III.4 résument les performances de chaque modèle.

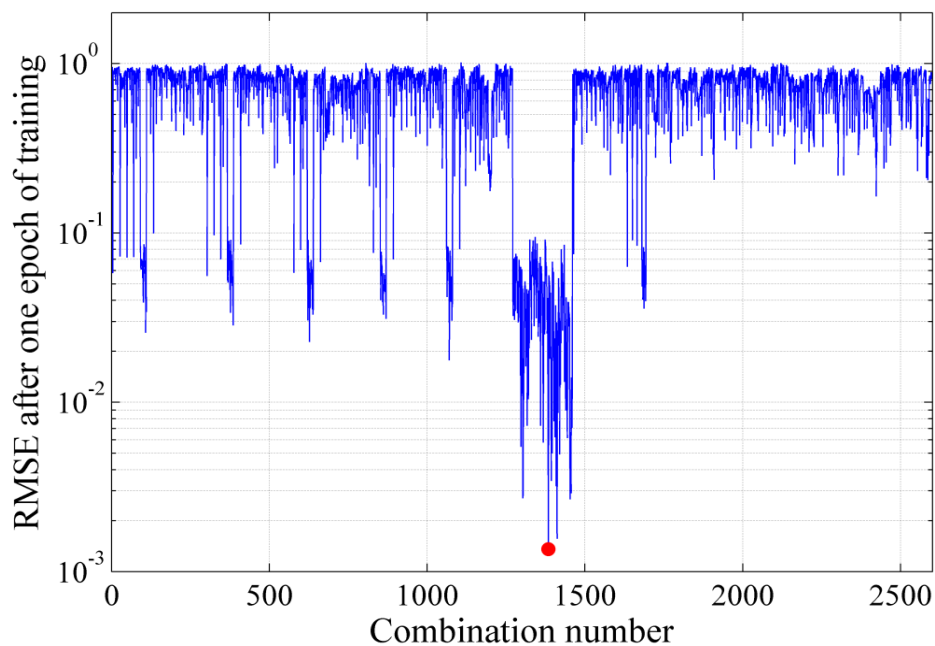


Figure III.14. Sélection d'entrées pour le modèle destiné à prédire ε' des échantillons liquides.

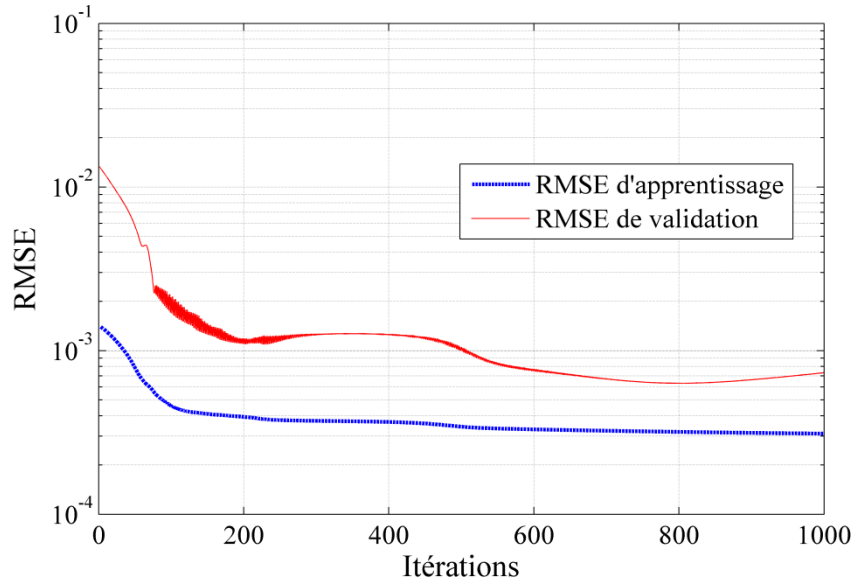


Figure III.15. Évolution de l'erreur d'apprentissage et de validation pour le modèle destiné à prédire ε' des liquides.

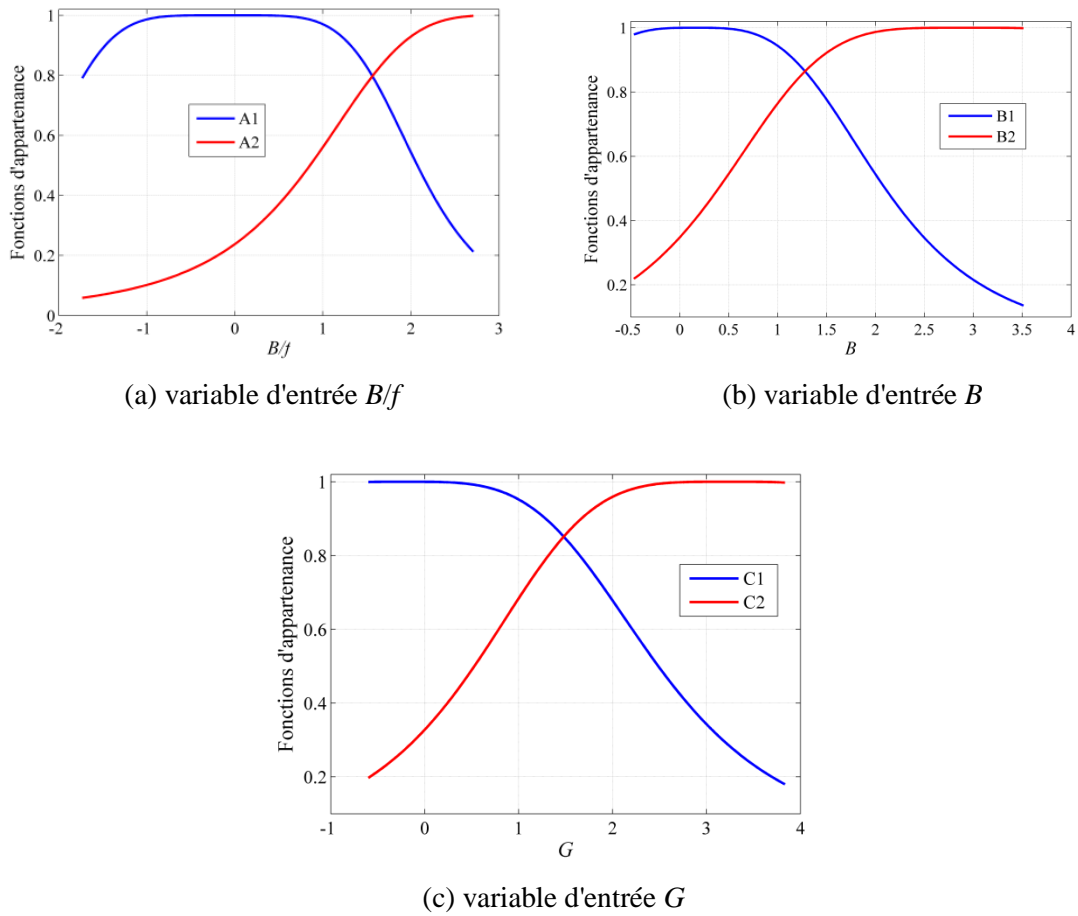


Figure III.16. Fonctions d'appartenance optimales pour le modèle destiné à prédire ε' des liquides.

Tableau III.3. Performances des deux modèles ANFIS (Échantillons liquides).

	Sortie	Base d'apprentissage		Base de validation	
		<i>RMSE</i>	<i>R</i>	<i>RMSE</i>	<i>R</i>
ANFIS ₁	ε'	3.6442×10^{-4}	0.9999999	7.4285×10^{-4}	0.9999997
ANFIS ₂	ε''	9.4046×10^{-4}	0.9999996	1.2871×10^{-3}	0.9999992

Tableau III.4. Performances des deux modèles ANFIS (Échantillons solides).

	Sortie	Base d'apprentissage		Base de validation	
		<i>RMSE</i>	<i>R</i>	<i>RMSE</i>	<i>R</i>
ANFIS ₁	ε'	1.5811×10^{-4}	0.999999989	3.6183×10^{-4}	0.999999934
ANFIS ₂	ε''	1.8838×10^{-4}	0.999999997	2.1591×10^{-4}	0.999999977

Une comparaison entre les valeurs de la permittivité prédites par ANFIS et celles contenues dans les bases d'apprentissage et de validation, pour les échantillons liquides, est montrée dans les figures III.17 et III.18, respectivement. Plus les points sont concentrés autour de la diagonale, plus la prédiction du modèle est meilleure. Les résultats obtenus montrent que la plupart des points sont distribués le long de la diagonale, ce qui explique la valeur élevée de *R*. Cela signifie que les deux modèles ANFIS ont atteint un degré d'apprentissage très élevé. De plus, dans le cas des solides, la figure III.19 montre la bonne concordance entre les permittivités prédites par ANFIS et celles contenues dans la base de validation. Comme prévu, selon les valeurs de *RMSE* et *R* présentées dans les tableaux III.3 et III.4, les résultats obtenus sur la base de validation sont un peu moins précis que ceux obtenus sur la base d'apprentissage, ce qui est logique. Cependant, le degré de généralisation est largement suffisant pour espérer de bons résultats d'inversion sur les données expérimentales.

Une fois le réseau prédit correctement les sorties contenues dans la base de validation (Données simulées), les données expérimentales peuvent être présentées aux entrées pour l'inversion.

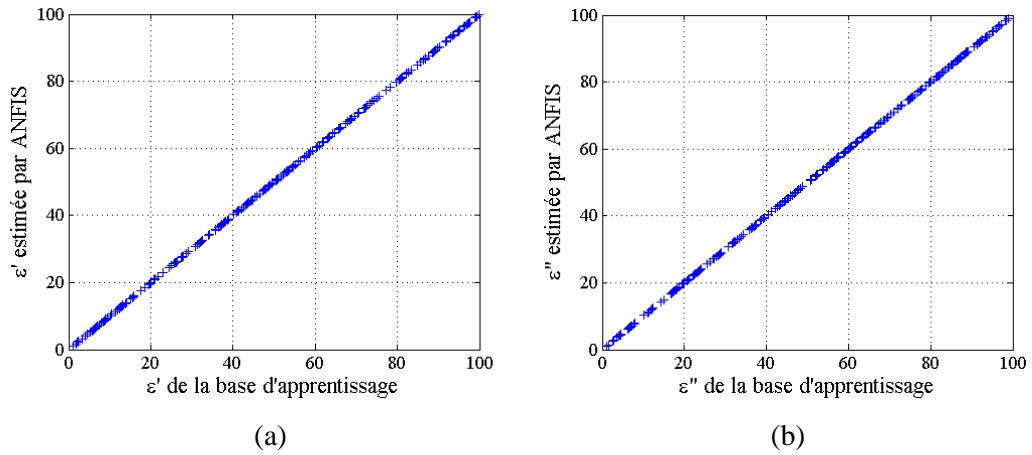


Figure III.17. Comparaison entre la permittivité fournie par ANFIS et celle contenue dans la base d'apprentissage pour les liquides, (a) partie réelle ϵ' , (b) partie imaginaire ϵ'' .

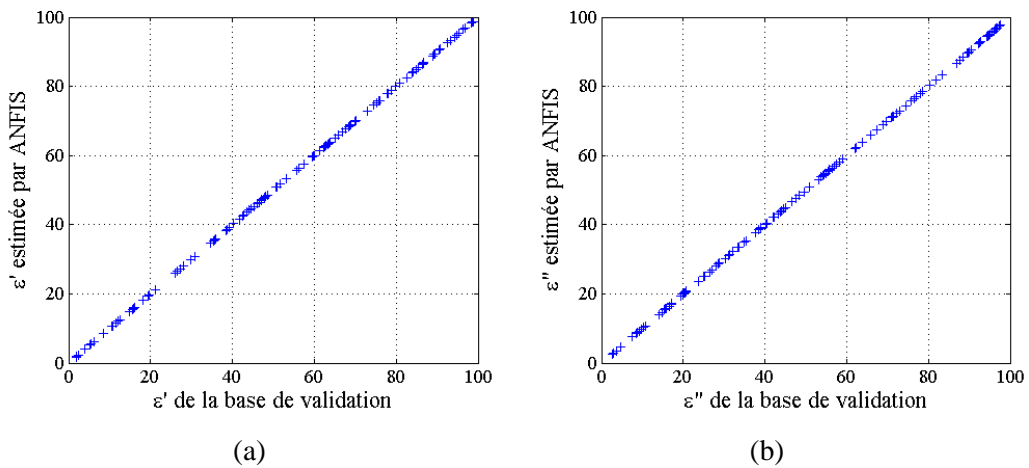


Figure III.18. Comparaison entre la permittivité fournie par ANFIS et celle contenue dans la base de validation pour les liquides, (a) partie réelle ϵ' , (b) partie imaginaire ϵ'' .

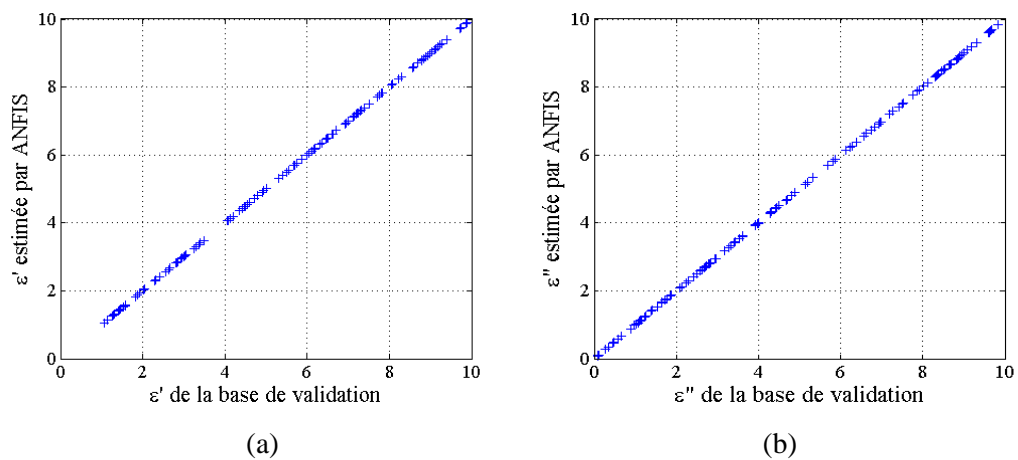


Figure III.19. Comparaison entre la permittivité fournie par ANFIS et celle contenue dans la base de validation pour les solides, (a) partie réelle ϵ' , (b) partie imaginaire ϵ'' .

III.9.2. Application sur les données expérimentales

Tout d'abord, des mesures ont été effectuées à l'aide d'un analyseur d'impédance Agilent 4291A sur un échantillon liquide dont les caractéristiques diélectriques sont connues (Figure III.20). Le guide coaxial utilisé avec la cellule SuperMit est le standard APC7 (*rayon externe* = 3.5 mm, *rayon interne* = 1.52 mm). La bande de fréquence de mesure s'étend de 1 MHz à 1.8 GHz. La fenêtre coaxiale étanche utilisée pour caractériser les liquides est en plexiglas et a une épaisseur de 0.82 mm et une constante diélectrique de 2.7. L'épaisseur de l'échantillon sous teste est de 13.8 mm. Ici, la sonde se tient verticalement de sorte que l'échantillon liquide soit placé dans la partie supérieure pour éviter l'interaction du champ électromagnétique avec les bulles susceptibles de se produire dans l'échantillon. C'est la raison pour laquelle les échantillons liquides doivent être relativement longs. Pour éviter les erreurs de mesure, une procédure d'étalonnage préalable avec une sonde ouverte et une sonde court-circuitée est effectuée sur les matériaux en question (Liquides et/ou solides).

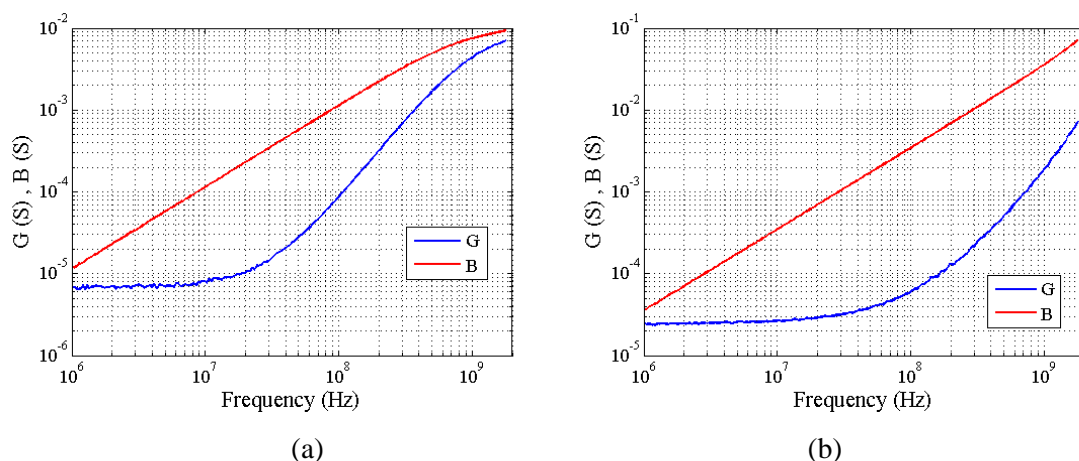


Figure III.20. Mesures expérimentales de la partie réelle (G) et imaginaire (B) de l'admittance de la cellule de mesure ; (a) l'Éthanol, (b) l'eau.

Afin d'évaluer les performances de notre modèle, les résultats d'inversion ont été comparés avec ceux du modèle de Debye qui exprime la permittivité complexe des matériaux polaires comme suit [Deby 29, Belh 98] :

$$\varepsilon = \varepsilon_{\infty} + \frac{\varepsilon_s - \varepsilon_{\infty}}{1 + j(f/f_0)} \quad (\text{III.37})$$

où ε_s et ε_{∞} sont les constantes diélectriques en basses et hautes fréquences, respectivement, f la fréquence de fonctionnement et f_0 la fréquence de relaxation.

- Dans le cas de l'Éthanol, nous avons : $\varepsilon_s = 25.07$, $\varepsilon_\infty = 4.2$ et $f_0 = 1.1$ GHz à 20 °C [Belh 97].
- Dans le cas de l'eau, nous avons : $\varepsilon_s = 80.21$, $\varepsilon_\infty = 5.7$ et $f_0 = 16.98$ GHz à 20 °C [Kaat 05].

Les évolutions des permittivités diélectriques de l'Éthanol et de l'eau, sont présentées dans les figures III.21 et III.22 respectivement. Dans ces figures, une superposition des résultats obtenus à partir de l'inversion des mesures par l'ANFIS, de ceux analytiques obtenus à partir du modèle de Debye ainsi que de ceux de l'inversion itérative des mesures, a été effectuée. D'après ces figures, on peut constater le bon accord entre l'inversion et les résultats analytiques. La légère différence est due aux conditions expérimentales telles que la température, la présence d'impuretés dans l'échantillon et l'imprécision des mesures.

Une différence significative dans ε'' est constatée. Ceci revient au fait que le modèle de Debye ne prend pas en compte les phénomènes de conduction qui apparaissent en basses fréquences. En effet, dans la formule de Cole-Cole [Cole 41, Bond 08] et les autres formules qui apparaissent après, le terme de la conductivité du courant continu $-j\sigma/2\pi f\varepsilon_0$ est ajouté, ce qui augmente fortement ε'' surtout en basses fréquences. σ est la conductivité ionique (Statique) du matériau. Cette dernière est inconnue, car elle dépend des impuretés et des conditions atmosphériques. D'ailleurs, c'est pour cette raison (Et d'autres) qu'on a contenté d'introduire un modèle simple comme celui de Debye pour la comparaison.

A environ 1,1 GHz, le graphique de la figure III.21 montre un phénomène de relaxation. En ce qui concerne la courbe de l'eau (Figure III.22), la relaxation ne peut pas être visualisée puisqu'elle apparaît en hautes fréquences. De manière générale, les résultats d'ANFIS coïncident très bien avec la procédure d'inversion itérative et ils sont en bonne concordance avec ceux de la littérature scientifique [Acik 07, Hacib 10, Kaat 05, Buck 58, Wagn 14, Greg 12]. De plus, les résultats d'inversion correspondent bien aux valeurs théoriques de la permittivité en basses fréquences. En effet, suivant la référence [Acik 07], la permittivité théorique en basses fréquences de l'eau et de l'Éthanol à 25°C est de 79 et 24, respectivement.

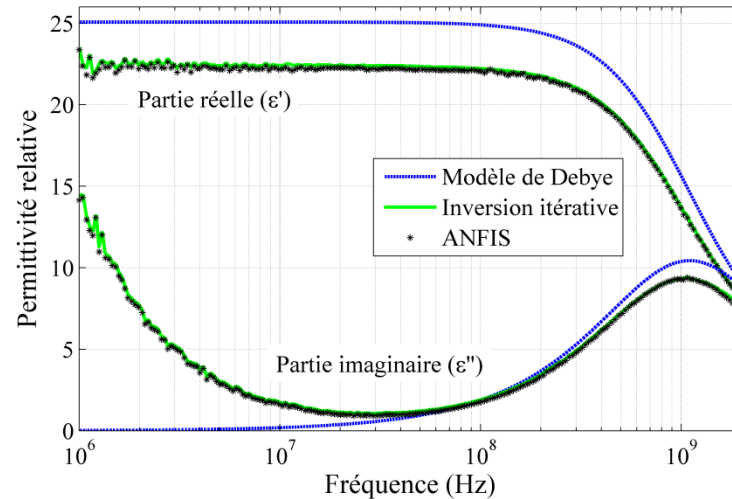


Figure III.21. Évolution de la permittivité de l'Éthanol obtenue par ANFIS, l'inversion itérative et le modèle de Debye.

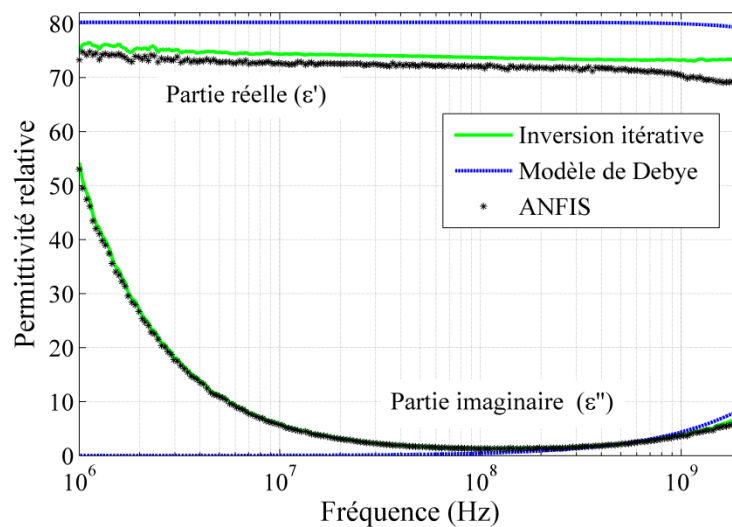


Figure III.22. Évolution de la permittivité de l'eau obtenue par ANFIS, l'inversion itérative et le modèle de Debye.

Dans un deuxième temps, des mesures ont été effectuées sur un échantillon solide ; le PEEK, d'une épaisseur de 0,7 mm (Cellule de mesure SuperMit sans fenêtre étanche). Ce matériau a une faible permittivité relative (Surtout pour la partie réelle de l'admittance G) ce qui nécessite plus de précision sur les mesures. Pour éviter les effets de contact, les échantillons sont métallisés sur des surfaces de contact avec les parties conductrices de la ligne coaxiale. La figure III.23 montre les mesures expérimentales de l'admittance complexe de la cellule de mesure avec la fréquence (f).

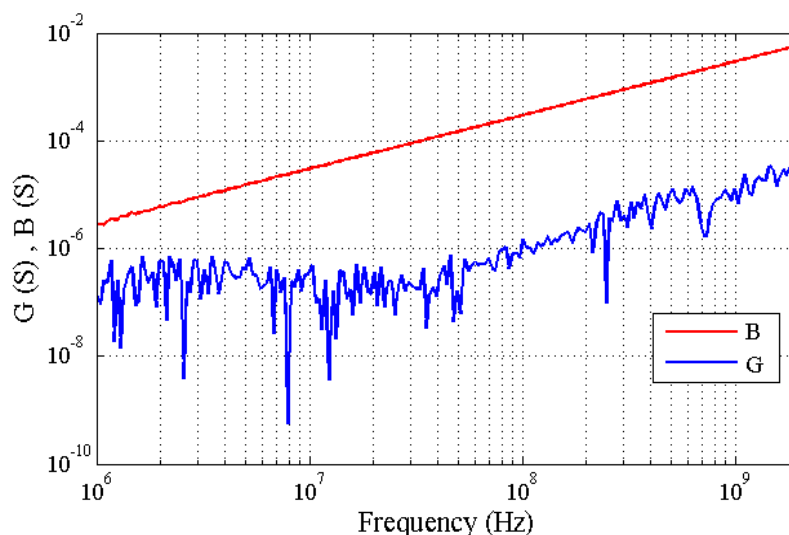


Figure III.23. Mesures expérimentales de la partie réelle (G) et imaginaire (B) de l'admittance de la cellule de mesure avec un échantillon en PEEK.

Finalement, les courbes de la partie réelle et de celle imaginaire de la permittivité relative, présentées dans la figure III.24, concordent très bien avec ceux de l'inversion itérative. En effet, d'après la référence [Good 18], le PEEK connu comme matériau sans pertes ($\varepsilon'' \approx 0$) a une constante diélectrique comprise entre 3.2 et 3.3 pour une fréquence comprise entre 50 Hz et 10 KHz. De plus, les résultats sont en bon accord avec celles des références bibliographiques [Acik 07, Raje 08]. Les valeurs dispersées et négatives de ε'' , en basses fréquences, sont dues aux valeurs très faibles de l'admittance (En particulier pour G qui est l'image de ε'') ainsi qu'à l'imprécision de l'analyseur pour les fréquences inférieures à 10 MHz.

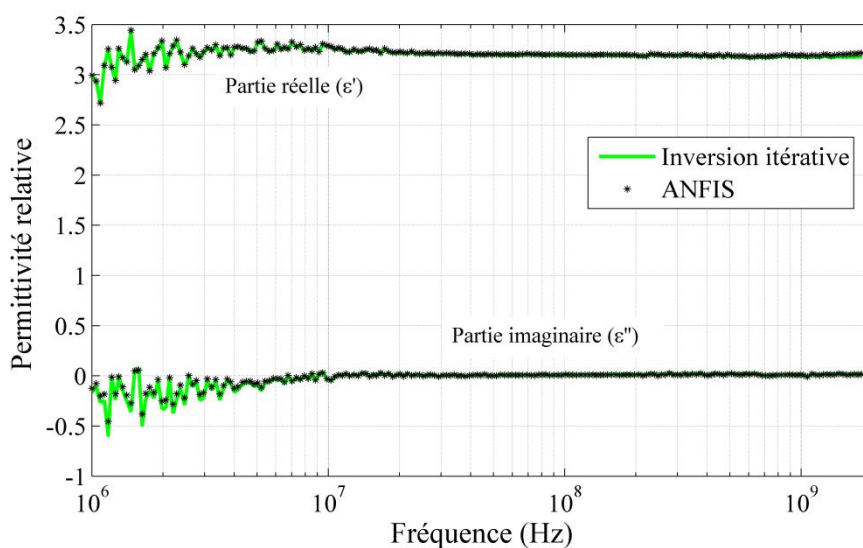


Figure III.24. Évolution de la permittivité du PEEK obtenue par ANFIS et l'inversion itérative.

Conclusion

Dans ce chapitre, on a vu la mise en œuvre d'ANFIS pour résoudre le problème inverse. La qualité supérieure de l'inversion effectuée sur des données de simulation et de mesure témoigne de l'adéquation et l'efficacité de cet outil pour notre problème. Bien qu'ANFIS avec des entrées optimales réalise l'inversion avec une très bonne précision, les mauvais résultats avec les entrées originales et la nécessité de passer par une procédure de sélection des entrées optimales constituent deux défauts de ce modèle. Cependant, le modèle développé est relativement simple, facile à mettre en œuvre, il converge rapidement et il donne, après apprentissage, une réponse instantanée. De plus, ANFIS s'entraîne à l'aide d'une base d'apprentissage relativement réduite, ce qui peut être dans certains cas très important.

Chapitre IV : Modèles de régression multivariés pour la résolution du problème inverse

IV.1. Introduction

Lorsqu'on collecte des informations quantitatives sur des individus, on les regroupe sous forme d'un tableau. Dans la statistique, on peut étudier la variable d'un point de vue uni variable, c'est à dire étudier les paramètres statistiques pour chaque variable, comme on peut s'intéresser à l'interaction entre les différentes variables, ce que l'on appelle étude multivariée. La régression multilinéaire (MLR pour Multiple Linear Regression), l'analyse des composantes principales (PCA pour Principal Components Analysis) et la régression aux moindres carrés partiels (PLSR pour Partial Least-Squares Regression) sont des méthodes d'analyses multivariées qui servent à résumer le maximum d'informations possibles en perdant le minimum pour faciliter leur étude.

Dans ce chapitre, on va exposer les aspects théoriques de la MLR et la PCA qui vont nous aider à bien comprendre la PLSR. Cette dernière ainsi que la MLR vont être utilisées, par la suite, comme outils d'inversion. De plus, deux études comparatives vont être présentées : d'une part, la MLR va être comparée au système d'inférence neuro-flou adaptative (ANFIS pour Adaptive Network based Fuzzy Inference System), d'autre part, la PLSR va être comparée au perceptron multi-couche (MLP pour Multi Layer Perceptron).

IV.2. Régression multilinéaire

La régression linéaire est la plus ancienne technique statistique de régression et elle peut être considérée comme une approche de référence pour les nouvelles techniques. Les méthodes statistiques examinent simultanément un grand nombre d'observations pour établir la relation entre le comportement observé et les causes. La MLR est une technique statistique bien connue qui sert à établir une relation quantitative entre un groupe de variables prédictors (Ou descripteurs) mis sous forme d'une matrice \mathbf{X} et une réponse univariable \mathbf{Y} en ajustant une équation linéaire aux données observées. Dans le cas d'une seule variable indépendante, on parle de la régression linéaire ordinaire. Les constantes de régression qui donnent les erreurs minimales du modèle sont déterminées par la technique des moindres carrés [Mata 11, Chan 11, Bagh 12, Shek 17].

Supposons qu'il existe P variables indépendantes et N observations $(x_1, x_2, \dots, x_i, \dots, x_P, \mathbf{Y})$ où x_i est le $i^{\text{ème}}$ variable indépendant de \mathbf{X} mis sous forme d'une colonne de N valeurs. Chaque ligne du

modèle liant les variables indépendantes x_i à la variable dépendante Y (Une observation des entrées et de la sortie) s'écrit comme suit :

$$Y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \dots + \beta_j x_j + \dots + \beta_P x_P + \epsilon \quad (\text{IV.1})$$

Le modèle obtenu pour un système de N équations peut être exprimé sous la forme matricielle suivante :

$$\begin{bmatrix} y_1 \\ y_2 \\ y_3 \\ \vdots \\ y_N \end{bmatrix} = \begin{bmatrix} 1 & x_{11} & \dots & x_{1P} \\ 1 & x_{21} & \dots & x_{2P} \\ 1 & x_{31} & \dots & x_{3P} \\ \vdots & \vdots & \ddots & \vdots \\ 1 & x_{N1} & \dots & x_{NP} \end{bmatrix} \begin{bmatrix} \beta_0 \\ \beta_1 \\ \beta_2 \\ \vdots \\ \beta_P \end{bmatrix} + \begin{bmatrix} \epsilon_1 \\ \epsilon_2 \\ \epsilon_3 \\ \vdots \\ \epsilon_N \end{bmatrix} \quad (\text{IV.2})$$

En écriture générale on a :

$$Y = X\beta + \epsilon \quad (\text{IV.3})$$

où Y est un vecteur ($N \times 1$) de la variable dépendante, X est une matrice ($N \times (P + 1)$) des P variables indépendantes, β est un vecteur ($(P + 1) \times 1$) des coefficients de régression, et ϵ est un vecteur ($N \times 1$) d'erreurs aléatoire.

Les modèles MLR sont basés sur les moindres carrés. Ces modèles sont ajustés de telle sorte que la somme des carrés des différences entre les valeurs observées et prédites (L) est minimale :

$$L = \sum_{i=1}^N \epsilon_i^2 = \epsilon^T \epsilon = (Y - X\beta)^T (Y - X\beta) \quad (\text{IV.4})$$

En notation matricielle, l'estimateur des moindres carrés de β est :

$$\hat{\beta} = (X^T X)^{-1} X^T Y \quad (\text{IV.5})$$

avec $(X^T X)^{-1} X^T$ est le pseudo inverse de X .

Le modèle approché s'exprime comme suit :

$$\hat{Y} = X \hat{\beta} \quad (\text{IV.6})$$

Finalement, le vecteur des résidus est noté :

$$\hat{\epsilon} = Y - \hat{Y} \quad (\text{IV.7})$$

IV.2.1. Régression Non-linéaire Multiple

Notre modèle MLR est sensé capable d'établir la relation qui existe entre la permittivité diélectrique (ϵ' ou ϵ'') qui représente la sortie Y et les grandeurs mesurés (f , G et B) qui forment la matrice des descripteurs X . Bien que la MLR soit une méthode linéaire, alors qu'on est certain

qu'il existe une relation non linéaire entre les entrées et la sortie. Donc, modéliser cette relation par un polynôme d'ordre un des entrées (Équation IV.1) est impossible. La conception d'un modèle de Régression Non-Linéaire Multiple (MNLR pour Multiple Nonlinear Regression) peut résoudre ce problème.

Depuis longtemps, les chercheurs, souvent en conception d'expériences (Plan d'expérience), ont proposé plusieurs méthodes pour intégrer la non-linéarité dans les modèles de régression. Dans notre étude, une démarche inspirée de la méthode du bloc X étendu [Berg 97] est adoptée. À l'origine, cette méthode consiste à étendre la matrice des descripteurs X en ajoutant des termes au carré, des produits croisés et éventuellement des termes d'ordre supérieurs des descripteurs originaux (f , G et B). Dans notre cas, les descripteurs supplémentaires sont générés d'une autre manière (Équation IV.9).

Au début, les N observations de la base de données déjà construite (Section II.5.2) sont organisées sous la forme matricielle suivante (Même chose pour ε) :

$$\mathbf{X} = \begin{bmatrix} f_1 & G_1 & B_1 \\ f_2 & G_2 & B_2 \\ \vdots & \vdots & \vdots \\ f_N & G_N & B_N \end{bmatrix}_{(N \times 3)} \quad \mathbf{Y} = \begin{bmatrix} \varepsilon'_1 \\ \varepsilon'_2 \\ \vdots \\ \varepsilon'_N \end{bmatrix}_{(N \times 1)} \quad (\text{IV.8})$$

Chaque colonne $\{x_m\}_{m=1}^P$ de la nouvelle matrice des descripteurs s'écrit comme suit :

$$x_m = f^i * G^j * B^k, \quad i = -2, \dots, 2, \quad j = 0, \dots, 2, \quad k = 0, \dots, 2 \quad (\text{IV.9})$$

Sauf la combinaison quand $i=j=k=0$.

Par conséquent, 44 nouveaux descripteurs ($P = 44$) sont créés. Par exemple, on a: $1/f^2$, B/f^2 , B^2/f^2 , ..., G^2 , B^2 , fB , ..., $1/f$, $1/B$, $1/G$, ..., B^2/G^2 , f^2G , fG^2B , ..., $f^2G^2B^2$.

Dans ce cas, la nouvelle matrice des descripteurs X prend la forme suivante (Y garde la même forme) :

$$\mathbf{X} = \begin{bmatrix} (1/f^2)_1 & (B/f^2)_1 & \dots & (B)_1 & \dots & (G)_1 & \dots & (f^2G^2B)_1 & (f^2G^2B^2)_1 \\ (1/f^2)_2 & (B/f^2)_2 & \dots & (B)_2 & \dots & (G)_2 & \dots & (f^2G^2B)_2 & (f^2G^2B^2)_2 \\ \vdots & \vdots & \ddots & \vdots & \ddots & \vdots & \ddots & \vdots & \vdots \\ (1/f^2)_N & (B/f^2)_N & \dots & (B)_N & \dots & (G)_N & \dots & (f^2G^2B)_N & (f^2G^2B^2)_N \end{bmatrix}_{(N \times 44)} \quad (\text{IV.10})$$

IV.2.2. Sélection des descripteurs optimaux

Généralement, le développement d'un modèle MNLR avec le minimum de descripteurs est toujours souhaitable. Parmi les 44 nouveaux descripteurs, il peut y exister quelques uns redondants, colinéaires ou fortement corrélés qui peuvent nuire à la qualité du modèle. Un nombre réduit de descripteurs veut dire un modèle simple et avec moins de variables à mesurer. Un nombre excessif de paramètres peut entraîner le sur-apprentissage (Overfitting), [Bagh 12, Sun 16, Goh 14] c'est-à-dire, un bon apprentissage et une mauvaise généralisation. Il existe plusieurs méthodes qui permettent de choisir les descripteurs les plus pertinents ou les plus significatifs statistiquement, tels que les algorithmes génétiques [Pour 14], l'optimisation par essaims particulaires (PSO, pour Particle Swarm Optimization) [Bagh 12] et les méthodes pas à pas [Mari 06] qui ajoutent ou éliminent une par une les variables du modèle courant jusqu'à ce qu'un critère d'arrêt soit validé. Parmi ces méthodes pas à pas, les plus performantes sont très certainement la méthode ascendante (Forward) [Adhi 10], la méthode descendante (Backward) [Robi 15, Gham 17] et la méthode progressive (Stepwise) [Roy 09, Guil 11, Ghun 12 14]. Dans cette étude, la méthode progressive comme combinaison des méthodes ascendante et descendante est adoptée.

IV.2.2.1. Méthode de sélection descendante

La sélection descendante est une méthode d'élimination successive des descripteurs non significatifs. Initialement, tous les descripteurs sont utilisés pour déterminer $\hat{\beta}$ de l'équation IV.5 ainsi que la p -valeur (Cette notion sera discutée en détails dans la section IV.2.2.3) qui correspond à chaque descripteur. A partir du groupe des descripteurs non significatifs, c-à-d qui ont une p -valeur supérieure à un certain seuil, le descripteur le moins significatif (La plus grande p -valeur) est enlevé à chaque fois. Le modèle MNLR est reproduit de nouveau avec le groupe modifié. Cette procédure est répétée jusqu'à ce que tous les descripteurs soient statistiquement significatifs (p -valeur inférieure à un seuil) [robi 15].

IV.2.2.2. Méthode de sélection ascendante

Dans la méthode de sélection ascendante, les descripteurs les plus statistiquement significatifs sont ajoutés au modèle MNLR l'un après l'autre. Au début, on suppose qu'il n'y a aucun candidat dans le modèle. Après, on calcule la p -valeur pour chaque descripteur candidat supposant qu'il est dans le modèle. Ensuite, on choisit le descripteur avec la plus faible p -valeur en dessous d'un certain seuil. A chaque itération, on ajoute celui qui améliore le plus les performances du modèle (Coefficient de détermination R^2 , soit le carré du coefficient de corrélation linéaire R) jusqu'au

moment où aucun descripteur ne vérifie la condition permettant de rentrer dans le modèle. Notons ici qu'une fois une variable est ajoutée, elle ne peut pas être supprimée.

IV.2.2.3. Méthode de sélection progressive

La sélection progressive est une méthode bien connue qui sélectionne les entrées du modèle de prédiction en fonction de leur signification statistique. La sélection progressive peut être d'une grande valeur en présence d'entrées redondantes, ce qui peut entraîner une réduction de l'efficacité de la prédiction. Cette méthode combine les deux procédures précédentes, ascendante et descendante. La procédure de cette méthode exige, à chaque étape, un réexamen de toutes les variables introduites précédemment dans le modèle. Elle consiste donc, en une sélection ascendante suivie d'une élimination descendante [Mari 06 ; Roy 09 ; Ghun 12, 14].

Pour quantifier si une variable explicative (Descripteur) a un effet statistiquement significatif sur un modèle de régression, on procède généralement aux tests statistiques. Cela revient à tester si le régresseur (Coefficient) associé à la variable explicative est statistiquement différent de zéro ($\beta_j \neq 0$). Pour cela, un test d'hypothèse est défini, permettant de formuler une règle de décision pour juger si la différence, entre un modèle MNLR avec β_j et un modèle MNLR sans β_j , est significative ou non [Grig 12, Bouc 16].

a. Tests d'hypothèse

Un test d'hypothèse est un test statistique qui permet d'associer une valeur à une prise de décision, dans le but de quantifier la confiance en une hypothèse. Deux hypothèses peuvent être formulées dans un cadre statistique :

- l'hypothèse nulle \mathbb{H}_0 qui exprime une absence d'effet. Cela se traduit par le fait que le régresseur étudié ne contribue pas à expliquer une partie de la variabilité des observations.
- l'hypothèse de rejet \mathbb{H}_1 qui représente l'hypothèse alternative à \mathbb{H}_0 .

Dans le contexte de notre étude, l'hypothèse \mathbb{H}_0 correspond à « il n'y a aucun effet du régresseur β_j sur le modèle MNLR », et l'hypothèse alternative \mathbb{H}_1 correspond à « il y a un effet du régresseur β_j sur le modèle MNLR ».

b. Règle de décision

Un test permet d'accepter ou de rejeter l'hypothèse \mathbb{H}_0 . Par ailleurs, soit \mathbb{H}_0 , soit \mathbb{H}_1 est vraie. La décision engendre quatre possibilités caractérisées par des probabilités reliées aux deux types

d'erreur que l'on peut commettre (Tableau IV.1). Soient α et β les probabilités d'erreur (Ou risques) de première et deuxième espèce, respectivement (Figure IV.1) :

- α est la probabilité de choisir \mathbb{H}_1 et rejeter \mathbb{H}_0 alors que \mathbb{H}_0 est vraie. Cela correspond à détecter un Faux Positif (*FP*) avec une probabilité de α .
- $1-\alpha$ est la probabilité de choisir \mathbb{H}_0 alors que \mathbb{H}_0 est vraie. ou encore, la probabilité d'avoir un Vrai Négatif (*VN*).
- β est la probabilité de choisir \mathbb{H}_0 et rejeter \mathbb{H}_1 alors que \mathbb{H}_1 est vraie. Ce cas correspond à la probabilité β d'avoir un Faux Négatif (*FN*).
- $1-\beta$ est la probabilité de choisir \mathbb{H}_1 alors que \mathbb{H}_1 est vraie, ou encore, la probabilité d'avoir un Vrai Positif (*VP*).

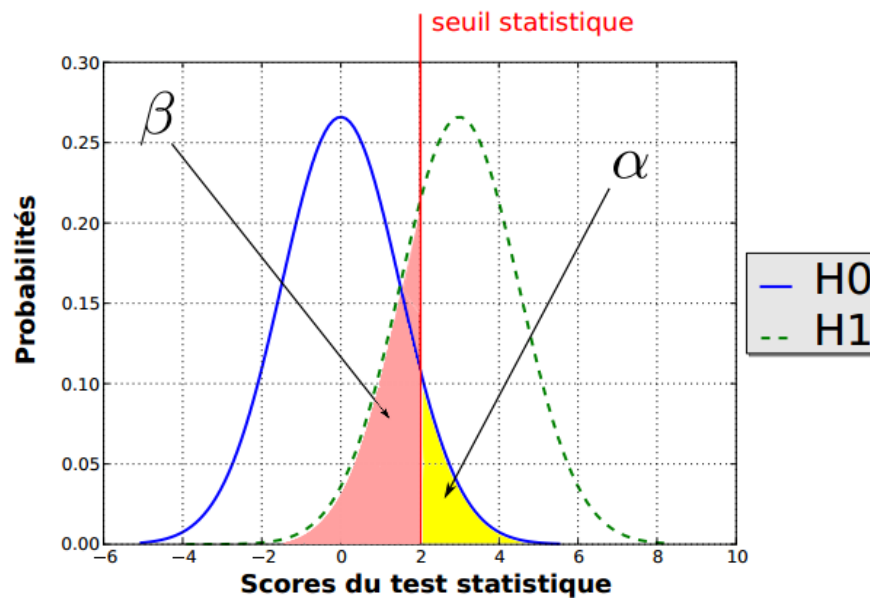


Figure IV.1. Illustration des erreurs de première (α) et deuxième (β) espèce.

Tableau IV.1. Choix possibles lors de la décision.

Décision \ Vérité	\mathbb{H}_0	\mathbb{H}_1
	\mathbb{H}_0	$1 - \alpha$ (<i>VN</i>)
\mathbb{H}_1	α (<i>FP</i>)	$1 - \beta$ (<i>VP</i>)

Il existe deux types de tests d'hypothèse : les tests paramétriques pour lesquels la loi de probabilité f_0 sous l'hypothèse \mathbb{H}_0 est paramétrée, et les tests non paramétriques pour lesquels la forme analytique de f_0 n'est pas connue. En règle générale, la distribution sous l'hypothèse \mathbb{H}_1 n'est pas connue et ne peut pas être modélisée.

c. Test de Fisher

Le test statistique de Fisher est un test paramétrique. Il permet d'évaluer si une des variables explicatives (Descripteurs) contribue de manière significative dans le modèle de régression. Pour cela, le test compare les résidus de deux modèles imbriqués : un modèle dit complet qui prend en compte toutes les variables explicatives (SSR_2) et un modèle dit restreint où la variable d'intérêt n'est pas considérée (SSR_1). Le paramètre statistique partiel F est calculé par la relation suivante :

$$F = \frac{\frac{SSR_1 - SSR_2}{P_2 - P_1}}{\frac{SSR_2}{N - P_2}} \sim F_{d_1, d_2} \quad (\text{IV.11})$$

où SSR est donné par la relation suivante :

$$SSR = \sum_{i=1}^N \epsilon_i^2 = \epsilon^T \epsilon = \|\mathbf{Y} - \mathbf{X}\hat{\beta}\|^2 \quad (\text{IV.12})$$

P_1 et P_2 représentent, respectivement, le nombre de variables explicatives des deux modèles et N le nombre d'observations. En supposant que les résidus suivent une distribution normale ($\epsilon \sim \mathcal{N}(0, \sigma^2)$) de variance σ^2 (Équation IV.13), F suit une distribution de Fisher avec $d_1 = P_2 - P_1$ et $d_2 = N - P_2$ degrés de liberté, sous l'hypothèse nulle \mathbb{H}_0 que le modèle 2 ne fournit pas un meilleur ajustement des variables explicatives aux données que le modèle 1.

$$\sigma^2 = \frac{1}{N - P} \|\mathbf{Y} - \mathbf{X}\hat{\beta}\|^2 \quad (\text{IV.13})$$

où P est le nombre de variables explicatives du modèle en question.

Deux stratégies sont possibles pour fixer le seuil de décision. La première consiste à imposer la valeur de l'erreur α et à calculer la valeur statistique (seuil) associée en fonction de la loi f_0 (Figure IV.1). Ensuite, l'hypothèse \mathbb{H}_0 est acceptée ou rejetée si le résultat du test est, respectivement, inférieur ou supérieur à ce seuil. La deuxième stratégie consiste à calculer la p -valeur associée à la valeur du score en fonction de la loi f_0 (Équation IV.14) (Figure IV.2) pour ensuite la comparer avec la valeur souhaitée de l'erreur α :

$$p\text{-valeur} = \int_{\text{score}}^{\infty} f_0(s) ds \quad (\text{IV.14})$$

Dans le cas des tests non paramétriques, la loi de probabilité sous l'hypothèse \mathbb{H}_0 n'a pas d'expression analytique, rendant impossible la modélisation de la distribution et le calcul des p -valeurs.

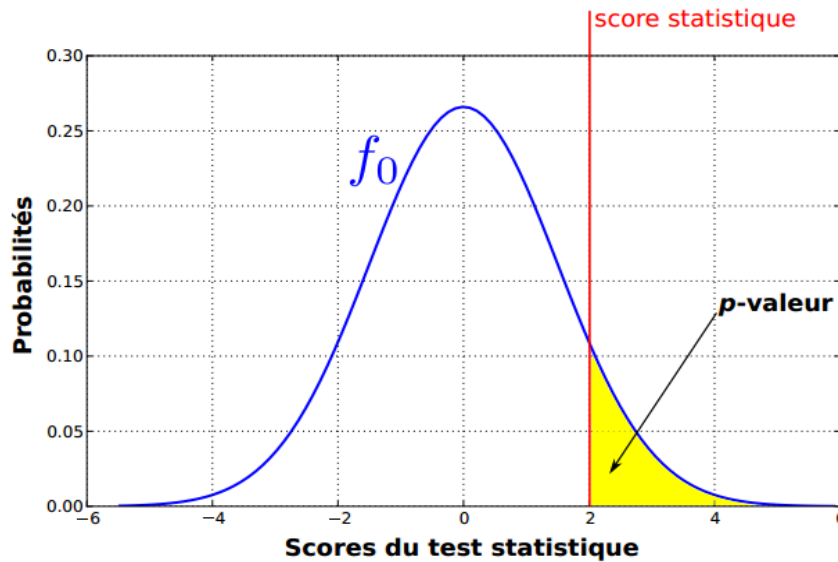


Figure IV.2. Calcul de la p -valeur.

Notons ici, que l'aire sous la courbe $f_0(s)$ de la figure IV.2 représente la distribution sous l'hypothèse \mathbb{H}_0 pour une valeur de la statistique du test.

Durant la sélection progressive, les termes sont ajoutés ou supprimés du vecteur des entrées en fonction de la p -valeur d'entrer et la p -valeur de suppression. La p -valeur (Niveau de signification) représente la probabilité de test de l'hypothèse nulle \mathbb{H}_0 avec une tolérance de α_{ajou} pour l'ajout d'un terme et une tolérance de α_{supp} pour supprimer un terme. L'hypothèse nulle exprime une absence d'effet. Cela se traduit par le fait que le descripteur étudié ne contribue pas à expliquer une partie de la variabilité des observations.

Le processus de la sélection progressive peut être résumé dans les étapes suivantes :

Étape 1 : Le processus est initié en ajoutant le descripteur avec la p -valeur la plus basse. Étant donné que cette valeur est inférieure à une valeur sélectionnée α_{ajou} , typiquement entre 0.05 et 0.1. Le processus est répété jusqu'au moment où il n'y a aucun descripteur ne puisse être additionné avec une tolérance d'entrée inférieure à α_{ajou} . L'étape 2 peut ensuite être exécutée.

Étape 2 : Le descripteur avec la plus grande p -valeur est supprimé. Étant donné qu'elle est supérieure à une valeur sélectionnée α_{supp} , typiquement entre 0.05 et 0.1. Le processus est répété pour tous les descripteurs existants, et si aucun descripteur supplémentaire ne peut être supprimé, l'étape 1 est revérifiée ; sinon, le processus est terminé.

Il convient de noter que différents modèles avec différents descripteurs d'entrée peuvent être obtenus en fonction du premier terme ajouté au modèle. Dans notre cas, il est évident qu'on choisit ce terme à partir des descripteurs originaux (f , G ou B).

IV.2.3. Implémentation de la MNLRL dans la sélection progressive

La nouvelle base de données (44 descripteurs) est divisée aléatoirement en deux sous-ensembles ; une base de calibration (2/3 de la base de données) pour réaliser la sélection progressive afin d'obtenir le modèle optimale, et une base de prédiction ou de teste (1/3 de la base de données) pour évaluer les performances du modèle obtenu.

Pour évaluer la qualité du modèle optimale, il existe plusieurs mesures statistiques selon le domaine d'application. Dans notre étude, le paramètre de Fisher F (Équation IV.11), la racine carrée de l'erreur quadratique moyenne $RMSE$ (Équation IV.15), le coefficient de détermination R^2 (Équation IV.16) et le coefficient de détermination ajusté R_a^2 (Équation IV.17) sont calculés comme mesures statistiques. Le meilleur modèle est celui qui a F , R^2 et R_a^2 les plus élevés et $RMSE$ la plus faible.

Bien que R^2 mesure, en réalité, la qualité de l'ajustement linéaire de l'équation de régression, il a l'inconvénient d'augmenter de plus en plus en ajoutant des descripteurs (Même non pertinents). Or un modèle n'est pas nécessairement meilleur parce qu'il a plus de descripteurs. Donc, R^2 n'est pas un bon indicateur pour évaluer un modèle. Le R_a^2 est une version modifiée de R^2 qui n'augmente pas nécessairement en raison du nombre de descripteurs supplémentaires [Sun 16].

$$RMSE = \sqrt{\frac{1}{N} \sum_{i=1}^N (y_i - \hat{y}_i)^2} \quad (IV.15)$$

$$R^2 = 1 - \frac{SSR}{SST} \quad (IV.16)$$

$$R_a^2 = 1 - \frac{N-1}{N-P-1} \frac{SSR}{SST} \quad (IV.17)$$

où :

$$SSR = \sum_{i=1}^N (\hat{y}_i - y_i)^2 \quad (IV.18)$$

$$SST = \sum_{i=1}^N (y_i - \bar{y}_i)^2 \quad (IV.19)$$

Avec SSR la somme carrée du résidu entre la prédiction et l'observation, SST la somme totale des carrés, y_i est la valeur des propriétés du matériau, \hat{y}_i est la valeur de la prédiction par le modèle MNLR, \bar{y}_i est la moyenne des y_i , N est le nombre d'observations et P le nombre de descripteurs.

Dans ce qui suit, nous définissons : la *RMSE* de calibration (*RMSEC*) et de prédiction (*RMSEP*), le R^2 de calibration (R^2_{cal}) et de prédiction (R^2_{pred}) et le R^2_a de calibration ($R^2_{a\text{ cal}}$) et de prédiction ($R^2_{a\text{ pred}}$).

IV.2.4. Résultats et discussions

IV.2.4.1. Application sur les données de simulation

Dans cette partie, la base de calibration (250 exemples) est utilisée pour établir les descripteurs optimaux. La procédure de sélection progressive est lancée deux fois ; la première pour estimer ε' et la deuxième pour estimer ε'' . À chaque fois, un modèle initial avec les descripteurs originaux (f , G et B) est choisi, avec $\alpha_{\text{ajou}} = 0.05$ et $\alpha_{\text{supp}} = 0.1$. Le tableau IV.2 résume les résultats de la procédure de sélection progressive de ε' et ε'' . La procédure a abouti à deux modèles avec 25 variables (plus une constante) et 22 variables (plus une constante) pour ε' et ε'' , respectivement. Concernant ε'' , on remarque que le processus a permis l'ajout de 25 variables avec une p -valeur très faible durant la première séquence. Ensuite, 6 variables à forte p -valeur ont été exclus du modèle. Dans le cas de ε' et après l'ajout de 22 variables par le processus, il n'y avait pas encore de variables qui respectent la condition de suppression.

Une étude approfondie du tableau IV.2 peut aider à fixer les idées sur les notions de l'hypothèse nulle \mathbb{H}_0 et l'hypothèse alternative \mathbb{H}_1 . À l'étape 1 de la procédure de sélection progressive de ε' , l'hypothèse \mathbb{H}_0 (Absence de l'effet de la variable f^1B sur le modèle initial) est rejetée (p -valeur = 1.11×10^{-269} inférieur au seuil α_{ajou}) en faveur de l'hypothèse \mathbb{H}_1 (Il y'a un effet de la variable f^1B sur le modèle initial). Donc, la variable f^1B est ajoutée avec un niveau de confiance de $(1 - 1.11 \times 10^{-269}) \times 100$ % ($\approx 100\%$). À l'étape 26 de la procédure de sélection progressive de ε'' , l'hypothèse \mathbb{H}_0 (Absence de l'effet de la variable G^2B sur le modèle) est retenue (p -valeur = 0.951977 supérieur au seuil α_{supp}). Donc, la variable G^2B est supprimée avec un niveau de confiance de 0.951977×100 % ($> 95\%$).

Tableau IV.2. Résultats de la procédure de sélection progressive.

La sélection progressive pour ε'			La sélection progressive pour ε''		
Étape N°	variable	p-valeur	Étape N°	variable	p-valeur
1	f^1B	$p=1.11 \times 10^{-269}$	1	f^1G	$p=7.47 \times 10^{-268}$
2	B^2	$p=1.49 \times 10^{-61}$	2	f^1G^2B	$p=1.48 \times 10^{-83}$
3	G^2	$p=8.43 \times 10^{-49}$	3	GB	$p=1.08 \times 10^{-20}$
4	G^2B	$p=2.23 \times 10^{-66}$	4	G^2B^2	$p=7.81 \times 10^{-101}$
5	fGB^2	$p=3.66 \times 10^{-36}$	5	B^2	$p=3.30 \times 10^{-14}$
6	fG^2	$p=2.77 \times 10^{-11}$	6	$f^2G^2B^2$	$p=3.05 \times 10^{-11}$
7	fB^2	$p=3.07 \times 10^{-24}$	7	f^1GB	$p=6.40 \times 10^{-20}$
8	G^2B^2	$p=2.48 \times 10^{-15}$	8	f^2G^2B	$p=7.01 \times 10^{-38}$
9	GB	$p=4.80 \times 10^{-33}$	9	f^2G^2	$p=3.59 \times 10^{-07}$
10	GB^2	$p=1.38 \times 10^{-10}$	10	fGB	$p=2.04 \times 10^{-16}$
11	fG^2B	$p=3.70 \times 10^{-08}$	11	f^2GB^2	$p=6.72 \times 10^{-30}$
12	f^2GB	$p=4.09 \times 10^{-22}$	12	fG	$p=7.55 \times 10^{-45}$
13	f^2	$p=1.61 \times 10^{-08}$	13	f^1B	$p=1.20 \times 10^{-08}$
14	f^2B^2	$p=2.01 \times 10^{-08}$	14	f^2B	$p=2.51 \times 10^{-08}$
15	f^1G	$p=1.51 \times 10^{-08}$	15	f^2G^2	$p=8.46 \times 10^{-06}$
16	f^2GB	$p=3.35 \times 10^{-21}$	16	f^2GB	$p=9.48 \times 10^{-06}$
17	f^2G^2	$p=2.88 \times 10^{-07}$	17	f^2G	$p=1.00 \times 10^{-05}$
18	f^1B^2	$p=1.42 \times 10^{-05}$	18	B^2f^2	$p=0.00156382$
19	B^2f^2	$p=0.00016631$	19	f^2GB^2	$p=0.00373547$
20	$fB,$	$p=0.00378768$	20	f^1GB^2	$p=6.03 \times 10^{-24}$
21	f^2B	$p=0.00031629$	21	G^2B	$p=8.16 \times 10^{-08}$
22	f^2G	$p=0.0400738$	22	GB^2	$p=0.00464238$
			23	f^1	$p=0.00148283$
			24	f^1G^2	$p=0.00997521$
			25	$f^2G^2B^2$	$p=0.0262598$
			26	(-) G^2B	$p=0.951977$
			27	(-) f^1GB^2	$p=0.822729$
			28	(-) f^2B	$p=0.668341$
			29	(-) B^2f^2	$p=0.556953$
			30	(-) B^2	$p=0.184997$
			31	(-) f	$p=0.334746$

Remarque : le signe (-) veut dire variable enlevée.

Tableau IV.3. Modèles optimaux issus de la sélection progressive.

Modèle MNLN pour ε'		Modèle MNLN pour ε''	
Descripteur	Coefficient de régression	Descripteur	Coefficient de régression
B^2f^2	7.7377×10^{-03}	f^2GB^2	-3.4461×10^{-02}
f^2GB	7.0147×10^{-03}	f^2G^2	3.2687×10^{-03}
f^2G^2	3.1710×10^{-03}	f^2G^2B	-3.3656×10^{-02}
f^1B	$1.0641 \times 10^{+00}$	$f^2G^2B^2$	6.0492×10^{-02}
f^1B^2	-7.1420×10^{-02}	f^1	-9.7664×10^{-04}
f^1G	-1.0508×10^{-02}	f^1B	8.1663×10^{-04}
B	8.7680×10^{-02}	f^1G	$1.0775 \times 10^{+00}$
B^2	-3.5704×10^{-01}	f^1GB	8.0506×10^{-02}
G	-1.4298×10^{-02}	f^1G^2	2.2133×10^{-02}
GB	1.4648×10^{-01}	f^1G^2B	-2.8608×10^{-02}
GB^2	-4.4042×10^{-01}	B	-4.8954×10^{-04}
G^2	2.4870×10^{-01}	G	-4.2525×10^{-02}
G^2B	-7.7774×10^{-02}	GB	-6.9074×10^{-01}
G^2B^2	1.4414×10^{-01}	GB^2	2.9697×10^{-01}
f	-2.0146×10^{-02}	G^2B^2	-5.1784×10^{-02}
fB	-1.0804×10^{-01}	fG	2.5334×10^{-01}
fB^2	2.4660×10^{-01}	fGB	-1.0371×10^{-01}
fGB^2	4.5741×10^{-01}	f^2G	-1.0559×10^{-01}
fG^2	7.2487×10^{-02}	f^2GB	3.1529×10^{-01}
fG^2B	-2.7798×10^{-01}	f^2GB^2	-2.0413×10^{-01}
f^2	3.8756×10^{-02}	f^2G^2	-9.7278×10^{-02}
f^2B	5.2235×10^{-02}	$f^2G^2B^2$	9.0133×10^{-02}
f^2B^2	-1.3012×10^{-01}		
f^2G	-1.5387×10^{-02}	Constante	2.0433×10^{-04}
f^2GB	-9.3232×10^{-02}		
Constante	2.4482×10^{-05}		
$RMSEC = 1.8084 \times 10^{-2}$ $R^2_{cal} = 0.999996$ $R^2_{a cal} = 0.999995$ $F = 2.3929 \times 10^6$ p -valeur = 0 $N = 250$ observations $P = 25$ descripteurs		$RMSEC = 1.0976 \times 10^{-2}$ $R^2_{cal} = 0.999998$ $R^2_{a cal} = 0.9999984$ $F = 7.9418 \times 10^6$ p -valeur = 0 $N = 250$ observations $P = 22$ descripteurs	

Les descripteurs optimaux du modèle MNLR ainsi que leurs coefficients sont présentés dans le tableau IV.3. L'avantage principal d'un modèle MLR est d'établir une équation simple sous forme d'un polynôme qui relie les entrées à la sortie, c'est à dire un modèle facilement interprétable. Cependant, dans notre étude, la majorité des variables sont fictives donc peu importe les descripteurs du modèle optimal. Contrairement à la plus part des études qui emploient la MLR (Dans la chimiométrie par exemple) où les descripteurs, les coefficients et les inter-corrélations sont examinés soigneusement. Néanmoins, on peut dire que les variables avec un coefficient de régression élevé (f^1B pour ε' et f^1G pour ε'') ont une grande influence sur la sortie.

Le tableau IV.3 présente également, toute en bas, quelques mesures statistiques qui caractérisent les deux modèles. On peut constater un coefficient de corrélation ajusté $R^2_{a\text{ cal}}$ élevé et une $RMSEC$ faible ce qui indique que le modèle est fiable. Une valeur élevée de F et une p -valeur du critère de Fisher inférieur à 0.05 montrent que l'équation de régression est statistiquement significative. On remarque aussi que R^2_{cal} et $R^2_{a\text{ cal}}$ sont assez proches puisque $P \ll N$.

Pour tester la capacité de généralisation, les deux modèles sont utilisés pour prédire les permittivités complexes de la base de test (125 exemples). À partir du tableau IV.4, qui résume les performances de chaque modèle en termes de $RMSE$, R^2 et R^2_a (Les résultats de la base de calibration sont également présentés), on peut remarquer que les coefficients de détermination obtenus à partir des valeurs simulées et prédites des permittivités sont très élevés ($R^2_{a\text{ cal}} > 99\%$ et $R^2_{a\text{ pred}} > 99\%$). Donc, on peut conclure que la totalité de la variance est largement exprimée par les deux modèles. De plus, une faible différence entre $R^2_{a\text{ cal}}$ et $R^2_{a\text{ pred}}$ est remarquée, ce qui signifie la robustesse du modèle [Roy 09]. A partir de la $RMSEP$, on peut constater que les deux modèles MNLR ont une capacité de généralisation suffisante pour leur permettre d'inverser les données de mesure.

Tableau IV.4. Performances des deux modèles MNLR.

	Sortie	Base de calibration			Base de test		
		$RMSEC$	R^2_{cal}	$R^2_{a\text{ cal}}$	$RMSEP$	R^2_{pred}	$R^2_{a\text{ pred}}$
MNLR ₁	ε''	1.0976×10^{-2}	0.999998	0.999998	6.3037×10^{-2}	0.99995	0.99989
MNLR ₂	ε'	1.5769×10^{-2}	0.999997	0.999996	4.1775×10^{-2}	0.999980	0.999951

Une autre manière de montrer les performances des deux modèles est de tracer les valeurs estimées en fonction des valeurs cibles. Les figures IV.3 et IV.4 montrent une comparaison entre les sorties des modèles MNLR (\hat{y}_i) et les permittivités qui appartiennent à la base de calibration et de test (y_i), respectivement. Les résultats obtenus sur la base de calibration (Figures IV.3.a et IV.3.b) montrent une bonne superposition du nuage des points (y_i, \hat{y}_i) sur la droite de régression. De même, d'après les figures IV.4.a et IV.4.b, une bonne concordance est constatée, ce qui explique bien le fait qu'on a pu obtenir une *RMSE* et un R^2 excellent sur la base de test (Tableau IV.4).

En comparant les figures ci-dessous avec les figures III.17, III.18 et III.19, présentées dans le chapitre III, on peut constater une différence dans la disposition des points sur les droites de régression. Ceci est dû à la manière dont les bases de données ont été générées (Section II.5.2). Dans les figures III.17, III.18 et III.19, les points (Les signes +) sont distribués sur toute la droite de régression. Par contre, dans les figures ci-dessous, les points ne couvrent pas toute la droite. Dans ce cas, chaque signe (+) comporte plusieurs points superposés, donc plusieurs valeurs répétées de ε' (ou de ε'').

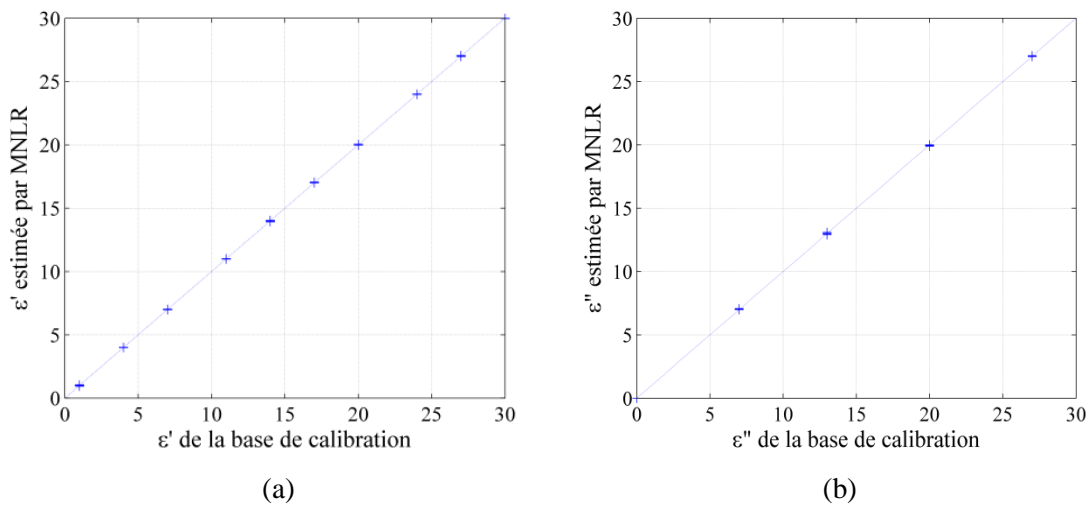


Figure IV.3. Comparaison entre les permittivités fournies par la MNLR et celles contenues dans la base de calibration, (a) partie réelle ε' , (b) partie imaginaire ε'' .

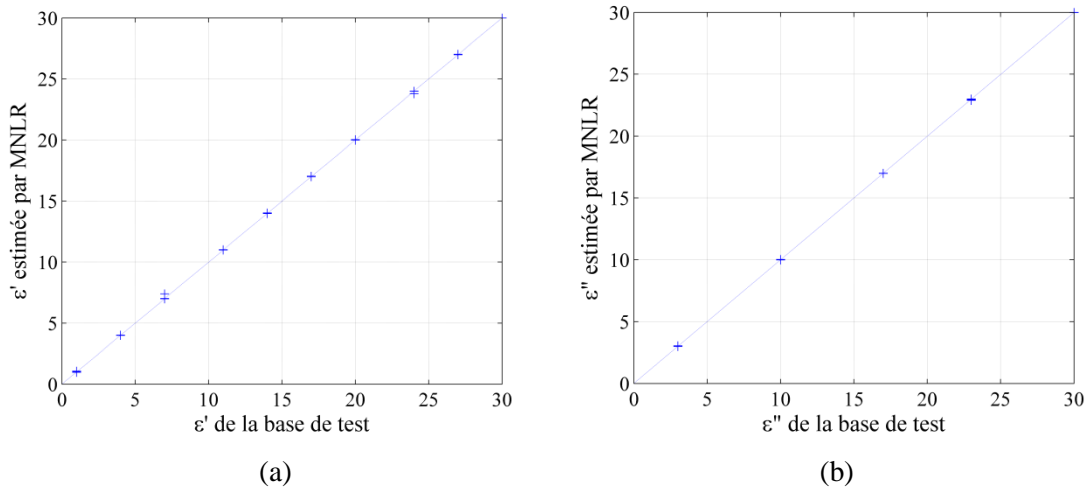


Figure IV.4. Comparaison entre les permittivités fournies par la MNLR et celles contenues dans la base de test, (a) partie réelle ϵ' , (b) partie imaginaire ϵ'' .

IV.2.4.2. Application sur les données expérimentales

Les données expérimentales peuvent être présentées au modèle MNLR pour l'inversion dès qu'il prouve son efficacité sur les données de simulation. Les mesures ont été effectuées, en utilisant un analyseur d'impédance Agilent 4291A, sur un échantillon diélectrique (l'Éthanol) dont les caractéristiques diélectriques sont connues. Le guide coaxial utilisé avec la cellule de mesure SuperPol est le standard GR900 (rayon externe = 7.14 mm, rayon interne = 3.1 mm). La bande de fréquence de mesure est comprise entre 1 MHz et 1.8 GHz. L'épaisseur de l'échantillon sous teste est de 2.6 mm. La figure IV.5 montre l'admittance complexe mesurée de la cellule de mesure.

La figure IV.6 présente les résultats de l'inversion des mesures. Une bonne concordance entre les résultats obtenus par la MNLR et ceux de l'inversion itérative a pu être constatée. De plus, les résultats obtenus sont en bon accord avec ceux présentés dans la littérature [Haci 10, 11 ; Acik 08].

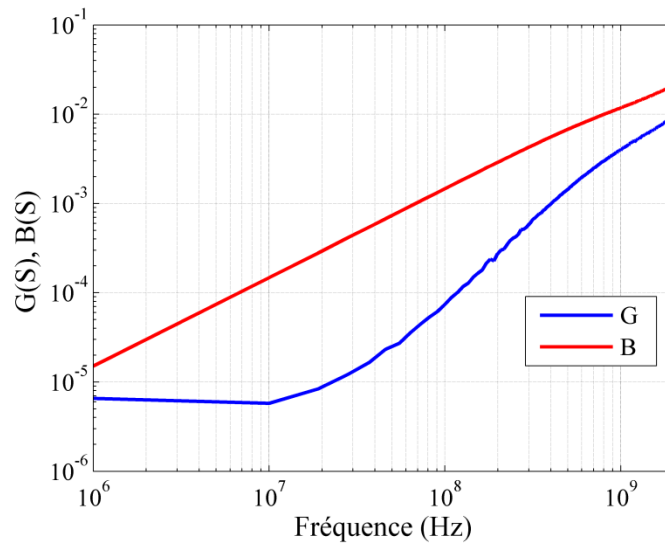


Figure IV.5. Mesures de la partie réelle (G) et imaginaire (B) de l'admittance de la cellule de mesure.

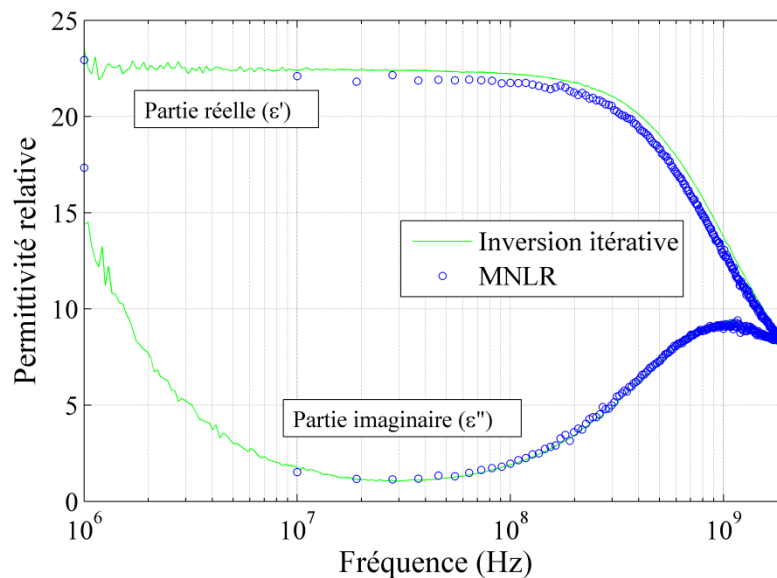


Figure IV.6. Évolution de la permittivité de l'Éthanol obtenue par la MNLR et l'inversion itérative.

IV.2.5. Comparaison entre MNLR et ANFIS

Comme la MNLR et l'ANFIS sont deux modèles d'inversion multi-entrées avec sortie unique, il est très intéressant de faire une comparaison entre eux. Ces deux modèles sont employés pour estimer ε' de l'Éthanol en utilisant les mesures effectuées sur la cellule SuperPol. La base de données générée, suivant la procédure décrite précédemment dans la section II.5.2, est utilisée pour entraîner les deux modèles. Le modèle ANFIS optimal est obtenu à partir de la procédure de

sélection d'entrées (3 entrées, 2 fonctions d'appartenance en forme de cloche associées à chaque entrée et 8 règles). Tandis que le modèle MNLR optimal est celui issu de la sélection progressive. La comparaison se fait sur la $RMSE$, R^2 et le temps de calcul CPU. Le temps de la sélection progressive des descripteurs pour la MNLR et le temps de la sélection des entrées optimales pour l'ANFIS ne sont pas comptabilisés. L'erreur de calibration calculée dans la MNLR est utilisée comme critère d'arrêt pour l'ANFIS.

Le tableau IV.5 résume les performances de chaque modèle. Pour la MNLR, les temps donnés dans ce tableau sont ceux nécessaires pour calculer les coefficients de régression et faire l'inversion des mesures. Tandis que pour l'ANFIS, ce sont les temps nécessaires pour l'apprentissage du modèle optimal et l'inversion des mesures. En plus, des résultats obtenus en utilisant la $RMSEC$ de la MNLR comme critère d'arrêt (ANFIS converge toujours après une seule itération), ce tableau présente les résultats avec l'erreur de validation minimale (ANFIS converge après 132 et 189 itérations pour des bases d'apprentissage de 250 et 2000 exemples respectivement).

D'après les résultats obtenus, on a pu constater que les deux modèles sont très efficaces et ont des performances comparables. On remarque, pour le modèle ANFIS, qu'une augmentation du nombre d'exemples dans la base d'apprentissage ne réduit pas l'erreur de calibration mais elle améliore la qualité de prédiction. En effet, d'après les résultats obtenus sur la base de test, cette augmentation contribue à renforcer la capacité de généralisation. Bien que l'ANFIS, quelque soit la taille de la base d'apprentissage, atteint l'erreur d'arrêt après une seule itération et au bout d'un temps comparable à celui de la MNLR, il nécessite plus de temps (132 itérations), avec 250 exemples, pour établir un modèle robuste ($R^2_{a\text{ cal}}$ et $R^2_{a\text{ pred}}$ élevés et très proches) comme celui de la MNLR. Concernant MNLR, on remarque qu'une augmentation du nombre d'exemples dans la base d'apprentissage n'augmente pas les performances du modèle. Pour une base d'apprentissage de 2000 exemples, bien que les deux modèles donnent rapidement de bons résultats, ils présentent une grande limite dans plusieurs études scientifiques où la création des bases de données nécessite des mesures ou la collecte de données statistiques.

Enfin, on peut conclure que les deux modèles, la MNLR avec 250 exemples et l'ANFIS avec exactement le même nombre d'exemples et après 132 itérations d'apprentissage, constituent deux outils d'inversion robustes, fiables et efficaces. Néanmoins, la MNLR est plus simple et plus rapide. Notons ici, que les deux modèles donnent, après apprentissage ou calibration, une réponse en temps réel.

Tableau IV.5. Comparaison entre MNLR et ANFIS

	Temps (s)	$RMSEC$	R^2_{cal}	Base de calibration	$RMSEP$	R^2_{pred}	Base de test
MNLR	0.011	1.5769×10^{-2}	0.999997	250	4.1775×10^{-2}	0.999980	125
	0.015	4.4087×10^{-2}	0.999977	2000	5.1468×10^{-2}	0.999969	125
ANFIS	0.016 1 itération	3.1246×10^{-3}	0.999989	250	7.5878×10^{-2}	0.994324	125
	0.698 132 itérations	1.0830×10^{-3}	0.999999	250	8.8438×10^{-3}	0.999922	125
	0.078 1 itération	1.0028×10^{-2}	0.999899	2000	1.1921×10^{-2}	0.999859	125
	7.43 189 itérations	2.0143×10^{-3}	0.999996	2000	1.5385×10^{-3}	0.999997	125

IV.3. Théorie de l'analyse en composantes principales

La théorie de l'analyse en composantes principales (PCA) est un outil statistique utilisé pour explorer des séries complexes d'observations multivariées. Il permet de résumer une grande quantité de données grâce à la reconnaissance de son contenu d'information le plus pertinent. La PCA se comporte comme une technique de filtrage-compression qui capture les principales tendances dans les données tout en révélant leur structure sous-jacente [Dubi 13].

Dans un ensemble de N « objets » (Les unités expérimentales) P attributs sont mesurés. Un nuage de points apparaîtrait lorsque les objets sont représentés dans un espace où chaque axe correspond à une variable. La PCA se concentre sur la « forme » d'un tel nuage, en essayant de capturer les directions dominantes du nuage (Les vecteurs propres). Lorsque ces données sont projetées sur un sous-espace de dimension inférieure, délimité par les vecteurs propres du nuage, une visualisation plus claire de la relation entre les objets est obtenue (Figure IV.7).

Supposent qu'on a une matrice de prédicteurs $X (N, P)$ et une matrice de réponse $Y (N, R)$, où N est le nombre d'observations, P le nombre de prédicteurs et R le nombre de sorties ($R=1$ réponse univariée, $R>1$ réponse multivariée). La PCA est une méthode non supervisée, c'est à dire elle vise simplement à décrire les données d'entrée, à savoir la matrice X en ignorant totalement la matrice Y . Elle transforme la matrice X en projetant ses éléments (Cercles bleus, figure IV.7) dans un sous-espace de dimension réduite (Cercles grises). D'une autre manière, elle cherche de nouveaux axes (PC1, PC2 et PC3) qui expriment mieux les éléments de X . Les nouveaux axes

doivent être mutuellement orthogonaux, et se croisent à la moyenne de tous les points du nuage (Cercle jaune). Cette technique exploratoire permet de constater si les exemples forment des groupes distincts, si un exemple est anormalement éloigné des autres il peut donc être considéré comme aberrant (Cercle rouge).

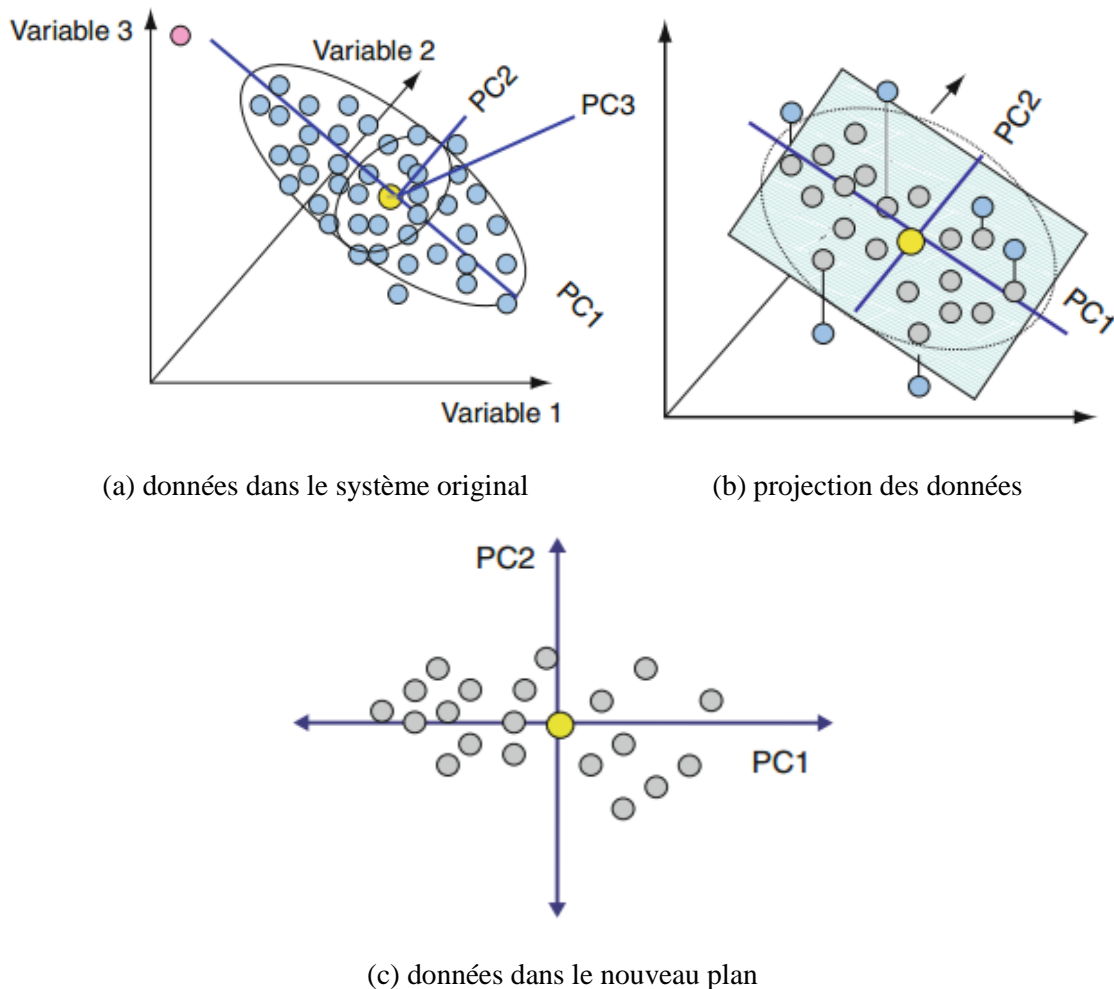


Figure IV.7. Opérations impliquées dans l'analyse par la PCA pour un système simple à trois variables.

La PCA est entièrement basée sur le calcul de la matrice des covariances (XX'). C'est une matrice carrée de dimension (P, P) et symétrique. Les valeurs situées sur la diagonale de cette matrice sont les valeurs des variances de chaque variable. Il faut ensuite calculer les valeurs et vecteurs propres de cette matrice des covariances, en prenant soin de normaliser à 1 la norme des vecteurs propres.

Le modèle PCA s'écrit comme suit :

$$X = TP' + E \quad (\text{IV.20})$$

où T la matrice des scores de dimension (N, A) , P la matrice des charges (Loadings matrix) de dimension (P, A) et E la matrice des résidus de dimension (N, P) . Dans ces conditions, A est le

nombre de composantes principales que l'on va prendre en compte pour réécrire la matrice \mathbf{X} . Notons que \mathbf{P} est la matrice des vecteurs propres de la matrice des covariances. Généralement, pour des matrices de grandes tailles, des algorithmes tels que la NIPALS (Nonlinear Iterative Partial Least Squares) et la SVD (Singular Value Decomposition) sont utilisés pour déterminer les matrices du modèle PCA (\mathbf{T} , \mathbf{P} et \mathbf{E}) [EL HA 13].

IV.4. Analyse par régression aux moindres carrés partiels

La PLSR a été initialement développée par Wold [Wold 85] et est devenue populaire dans les domaines de la chimométrie, la physiologie, le marketing, la géophysique, la supervision, ... etc. C'est une technique qui combine les principales caractéristiques de la PCA et la MLR. Le modèle PLSR sert à établir une relation entre une matrice de prédicteurs \mathbf{X} ($x_1, x_2, \dots, x_i, \dots, x_P$) et une matrice de réponse \mathbf{Y} ($y_1, y_2, \dots, y_i, \dots, y_R$), où x_i est la $i^{\text{ème}}$ variable indépendante mise sous forme d'une colonne de N observations et y_i est la $i^{\text{ème}}$ variable dépendante mise sous la même forme que x_i .

La prédiction est obtenue en extrayant de \mathbf{X} un ensemble de composantes orthogonales (Non corrélées), appelées variables latentes (Ou composantes latentes), qui ont le meilleur pouvoir prédictif en maximisant la covariance entre la projection des prédicteurs dans un sous-espace de dimension réduite et les réponses. En d'autres termes, la PLSR cherche les meilleurs axes qui, simultanément, expliquent les variations dans \mathbf{X} et prédisent \mathbf{Y} [Le Bi 14, Ruiz 13].

La puissance de la PLSR se manifeste dans le cas des problèmes à réponses multivariés (\mathbf{Y} est une matrice multidimensionnelle) avec des données bruitées, des variables indépendantes corrélées (\mathbf{X}) et un nombre limité d'observations (Même inférieur au nombre de prédicteurs) [Ruih 07, Abdi 10, Alin 09]. De plus, la PLSR effectue automatiquement une sélection de variables, facile à mettre en œuvre, statistiquement très efficace et très rapide en calculs, ce qui le rend pratique pour les applications qui exigent des bases de données larges [Boul 06].

La précision du modèle dépend fortement du nombre de composantes latentes retenues. Lorsque ce dernier est insuffisant, certaines informations utiles seront perdues (Un sous-ajustement ou under-fitting en Anglais). Contrairement, si le nombre de composantes est grand, le modèle risque d'entrer dans le phénomène de sur-ajustement (Over-fitting) et d'apporter du bruit inutile qui affecte la robustesse du modèle [Alin 09, Li 10]. Habituellement, la validation croisée est adopté pour résoudre ce problème [Wold 01, Roy 09].

Théoriquement, la PLSR suppose l'existence d'une relation linéaire entre les variables indépendantes \mathbf{X} et les variables dépendantes \mathbf{Y} . Considérons une matrice de prédicteurs \mathbf{X} ($N \times P$)

de N observations paramétrées par P paramètres (Entrées) et une variable de sortie Y ($N \times R$) supposées être à la fois (X et Y) ajustées par une combinaison linéaire d'un nombre réduit n de variables latentes.

La PLSR décompose X et Y sous la forme suivante [Le Bi 14, Boul 06] :

$$X = TP' + E \quad (IV.21)$$

$$Y = TQ' + F \quad (IV.22)$$

où T ($N \times n$) est la matrice des n vecteurs de score (n variables latentes) pour les prédicteurs et les réponses, les matrices P ($P \times n$) et Q ($R \times n$) représentent les matrices des charges (Loadings matrices) et les matrices E ($N \times P$) et F ($N \times R$) sont les matrices des résidus.

Dans cette formulation de la PLSR, les quantités X et Y sont contrôlées par la même matrice de score (Score matrix) T , tandis que chacune a sa propre matrice des vecteurs propres à savoir P pour X et Q pour Y . De même, chaque quantité a son propre résidu à savoir E pour X et F pour Y . Une telle formulation est adaptée aux problèmes de régression où X est de grande dimension (Exprimé par beaucoup de prédicteurs) mais la réponse Y est de faible dimension ($R \ll P$) ou même à variable unique ($R = 1$).

La première variable latente (Première colonne de T) $t_1 = Xr_1$ est extraite en maximisant $r_1' X' Y Y' X r_1$ avec la contrainte $\|r_1\| = 1$. Ceci conduit à calculer le vecteur propre dominant r_1 de la matrice $X' Y Y' X$. De la même manière, la deuxième variable latente (Deuxième colonne de T) $t_2 = Xr_2$ est déterminée en maximisant $r_2' X' Y Y' X r_2$ avec la double contrainte $\|r_2\| = 1$ et l'orthogonalité mutuelle des vecteurs de score $t_2' t_1 = 0$.

La procédure est répétée jusqu'à l'extraction de toutes les variables latentes, $t_k = Xr_k$ ($1 < k < n$). Ensuite, la matrice des poids R constituée de vecteurs r_k est ensuite utilisée pour établir la projection de X dans le nouvel espace de dimension réduite : $T = XR$. Après cela, la matrice des charges Q est calculée en utilisant une régression par les moindres carrés ordinaires de Y sur T : $Q = (T'T)^{-1} T'Y$.

Enfin, le modèle PLSR complet peut être établi à partir de l'équation IV.22 en supposant que l'erreur de projection est minimale, ce qui signifie que la matrice résiduelle F soit négligeable. On aura alors :

$$Y = X B_{PLS} \quad (IV.23)$$

avec $B_{PLS} = RQ'$ est la matrice des coefficients de régression PLS de $P+1$ lignes et R colonnes.

Il existe un certain nombre d'algorithmes proposés pour l'estimation pratique des coefficients du modèle PLSR (t_k , r_k , ... etc) tels que la SIMPLS (Statistically Inspired Modification of PLS) et la NIPALS. Les problèmes traités dans ce travail de thèse ont $N \gg P$ et $N \gg R$. Ceci nous amène à choisir l'algorithme SIMPLS car il assure un coût de calcul faible (Pas d'inversion de matrices) et une convergence très rapide, contrairement à l'algorithme NIPALS qui est plus couteux en temps de calcul [De Jo 93, Zamp 04, Alin 09, Huan 12].

Dans le reste du travail, le nombre maximal des variables latentes n est choisi de telle façon à avoir la même valeur que le nombre de prédicteurs P et l'inversion des mesures expérimentales est faite en utilisant un nombre optimal de variables latentes fixé par la validation croisée. Habituellement, ce nombre est limité par le critère de convergence de la procédure itérative.

IV.4.1. PLSR Non-Linéaire

Un modèle PLSR est particulièrement utile pour prédire un ensemble de variables dépendantes (Sorties) à partir d'un (Très) grand nombre de variables indépendantes (Prédicteurs) [Abdi 10]. Le but de PLS est la réduction de la dimensionnalité par la déflation itérative des matrices. En d'autres termes, remplacer les prédicteurs corrélés par certains variables latentes qui maximisent la covariance avec Y . Connaissons la nature non linéaire de la relation entre les entrées (f , G et B) et les sorties (ε' et ε'') dans la caractérisation micro-ondes, il est évident qu'un modèle linéaire comme la PLSR est incapable de remplir cette tâche avec ces trois variables d'entrée. Donc, le recours à un modèle PLSR Non-Linéaire (NLPLSR pour Non-Linear PLSR) est obligatoire.

Dans les années 1980, beaucoup de chercheurs, en particulier dans le domaine de la chimométrie, ont commencé à développer des méthodes de modélisation non linéaire basées sur des extensions de PLSR. Bien qu'elles soient adaptées au cas, de plus en plus courant, où le nombre d'observations est plus petit que le nombre de variables explicatifs, ça ne nous empêche pas d'appliquer efficacement une d'entre elles mais avec autant d'observations que de prédicteurs.

Il existe deux catégories de méthodes non linéaires basées sur la PLSR ; la délinéarisation interne et externe [Vivi 02].

IV.4.1.1 Délinéarisation interne

La délinéarisation interne rassemble un certain nombre de méthodes qui consistent à rendre non linéaire la relation entre la composante explicative t_k de la matrice des scores T et la composante associée à la réponse u_k de la matrice des scores U ($u_k = f(t_k) + \text{erreur}$). Dans ce cas,

la réponse Y a sa propre matrice des scores U ($u_1, u_2, \dots, u_k, \dots, u_n$), c'est à dire au lieu de l'équation IV.20, l'expression de Y devient $Y = UQ^T + F$. Parmi ces méthodes, on a :

PLS quadratique : c'est la première méthode de PLSR non-linéaire. Elle a été proposée par Wold [Wold 89] et elle consiste à délinéariser la relation interne par un modèle polynomial d'ordre 2 ($u = c_0 + c_1t + c_2t^2$).

PLS Non Linéaire : en 1990, Frank [Fran 90] a proposé de délinéariser la relation interne à l'aide de lisseurs, à la manière de SMART (Smooth Multiple Additive Regression Technique).

Spline PLS : cette fois Wold [Wold 92] propose, au lieu d'une fonction quadratique, d'utiliser les fonctions splines linéaires, quadratiques ou cubiques pour délinéariser la relation interne.

Il existe d'autres méthodes qui utilisent les réseaux de neurones [Qin 92, Ruih 07], les réseaux neuro-flous [Yoo 04] et la machine d'apprentissage extrême [Wang 15] pour délinéariser la relation interne.

IV.4.1.2. Délinéarisation externe

La délinéarisation externe regroupe des méthodes basées sur une transformation à priori des variables explicatives avant l'utilisation de PLSR. Parmi ces méthodes, on a :

PLS GIFI : elle consiste à adopter l'approche de Gifi [Gifi 90] pour étudier la non-linéarité présente dans les prédicteurs X qui expliquent Y . Cette approche est basée sur le codage flou des données.

PLSR Spline : elle est basée sur le recodage des prédicteurs où chaque variable x_i de la matrice X centrée et réduite est codée dans une base appartenant à l'espace des fonctions splines [Dura 01].

Régression par variables latentes implicitement non linéaire : Berglund et Wold [Berg 97] ont proposé une méthode simple et rapide pour modéliser, non-linéairement, un tableau Y par un tableau prédicteur X . Il s'agit d'étendre le tableau de prédicteurs avec les termes au carré, les termes croisés et éventuellement les termes d'ordre supérieurs. Par conséquent, cette approche ne convient donc qu'aux cas avec des variables peu nombreuses et indépendantes.

Cependant, les auteurs soulignent que cette méthode ne peut pas être appliquée lorsque le nombre de prédicteurs est important. En effet, le nombre de variables correspondants aux termes croisés ($P(P-1)/2$) devient vite très important devant le nombre des variables restantes ($2P$), qui sont les variables initiales et leurs carrés. Ceci implique que les termes croisés ont tendance à dominer les autres prédicteurs, ajoute du bruit au modèle et engendre de mauvais modèles, en

terme d'ajustement et de prédiction. Par conséquent, les auteurs proposent d'écartier les termes croisés qui seront implicitement remplacés par des produits croisés des variables latentes t_k , d'où le nom "Implicit Non-Linear Latent Variable Regression".

En PLSR linéaire, il y a une relation interne linéaire entre u_k et t_k . Comme $\mathbf{Y} = \sum_i u_i q_i'$, cela implique que le modèle de \mathbf{Y} est linéaire en composantes t_k . L'idée des auteurs ici est de faire apparaître une relation polynomiale et quadratique entre \mathbf{Y} et les variables latentes t_k en introduisant les carrés des prédicteurs dans la matrice \mathbf{X} , et ainsi des termes croisés de t_k seront implicitement inclus dans le modèle. Ceci est illustré ci-dessous avec un exemple simple.

Supposons que \mathbf{X} est une matrice de rang 2, l'équation IV.21 s'écrit alors :

$$\mathbf{X} = \mathbf{TP}' = \mathbf{t}_1 * \mathbf{p}'_1 + \mathbf{t}_2 * \mathbf{p}'_2 \quad (\mathbf{E} = 0) \quad (\text{IV.24})$$

Si nous considérons maintenant seulement le premier élément de \mathbf{X} , x_{11} , alors :

$$x_{11} = t_{11} * p_{11} + t_{21} * p_{21} \quad (\text{IV.25})$$

Le carré de x_{11} devient alors :

$$(x_{11})^2 = (t_{11} * p_{11})^2 + (t_{21} * p_{21})^2 + 2 * (t_{11} * p_{11}) * (t_{21} * p_{21}) \quad (\text{IV.26})$$

Si nous considérons l'exemple ci-dessus avec $\text{rang}(\mathbf{X}) = 2$, le bloc au carré (\mathbf{X}^2) aura un rang de 3 puisque les termes au carré et les termes croisés seront inclus, c'est-à-dire t_1^2 , t_2^2 et $t_1 t_2$. Ce sera, probablement, moins bien défini lorsque le rang de \mathbf{X} (Le nombre de composantes) est moins. Cette approche est facilement étendue aux termes cubiques, aux termes du quatrième degré, ... etc.

Les motivations de cette méthode sont multiples. Les méthodes de modélisation PLSR non-linéaires utilisant des transformations non-linéaires (Par exemple Spline PLS) sont souvent si flexibles qu'elles engendrent un sur-ajustement des données. De plus, les différents paramètres doivent être estimés par validation croisée, ainsi que le nombre de composantes principales. Cela peut être long et fastidieux dans certains cas.

Bien que cette méthode est inapplicable en pratique lorsque le nombre de prédicteurs est élevé car elle engendre une trop grosse augmentation de dimensions, un nombre raisonnable est nécessaire. Notre problème offre seulement 3 variables explicatives, donc il est évident qu'on s'attend à de mauvais résultats en utilisant un bloc de prédicteurs étendu seulement aux carrés [$\mathbf{X} \mathbf{X}^2$]. Pour pallier à ce problème, en plus des termes aux carrés, les termes d'ordre supérieurs et les produits croisés sont inclus dans notre modèle.

Étant donné qu'on possède une base de données organisée sous la forme matricielle suivante :

$$\mathbf{X} = \begin{bmatrix} f_1 & G_1 & B_1 \\ f_2 & G_2 & B_2 \\ \vdots & \vdots & \vdots \\ f_N & G_N & B_N \end{bmatrix}_{(N \times 3)} \quad \mathbf{Y} = \begin{bmatrix} \varepsilon'_1 & \varepsilon''_1 \\ \varepsilon'_2 & \varepsilon''_2 \\ \vdots & \vdots \\ \varepsilon'_N & \varepsilon''_N \end{bmatrix}_{(N \times 2)} \quad (\text{IV.27})$$

avec N le nombre d'observations (Exemples) dans la base de données.

En suivant la même démarche qu'on a développée précédemment avec la MLR (Équation IV.9), la nouvelle matrice des prédicteurs \mathbf{X} prend la forme suivante (\mathbf{Y} gardera la même forme) :

$$\mathbf{X} = \begin{bmatrix} (1/f^2)_1 & (B/f^2)_1 & \dots & (B)_1 & \dots & (G)_1 & \dots & (f^2 G^2 B)_1 & (f^2 G^2 B^2)_1 \\ (1/f^2)_2 & (B/f^2)_2 & \dots & (B)_2 & \dots & (G)_2 & \dots & (f^2 G^2 B)_2 & (f^2 G^2 B^2)_2 \\ \vdots & \vdots & \ddots & \vdots & \ddots & \vdots & \ddots & \vdots & \vdots \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \ddots & \vdots & \vdots \\ (1/f^2)_N & (B/f^2)_N & \dots & (B)_N & \dots & (G)_N & \dots & (f^2 G^2 B)_N & (f^2 G^2 B^2)_N \end{bmatrix}_{(N \times 44)} \quad (\text{IV.28})$$

IV.4.2. Implémentation de la NLPLSR

Avant la mise en œuvre de la NLPLSR, la base de données est divisée, de manière aléatoire, en deux sous-ensembles différents; un ensemble de calibration (2/3 de la base de données) utilisé pour réaliser la validation croisée et développer un modèle étalonné, et un ensemble de prédiction ou de test (1/3 de la base de données) utilisé pour évaluer les performances du modèle [Rahm 16]. L'ensemble de calibration est divisé en 10 sous-ensembles égaux. Le modèle NLPLSR est implémenté 10 fois pour différents valeurs de variables latentes (De 1 à 44) impliquant, à chaque fois, 9 sous-ensembles comme base d'apprentissage et 1 ensemble comme base de validation. Ceci jusqu'à ce que tous les sous-ensembles vont être utilisés une fois pour la validation. Par conséquent, 180 exemples sont utilisés pour l'entraînement, 20 exemples pour la validation et 100 exemples pour le test. Le nombre optimal de variables latentes est celui correspondant à la moyenne minimale (Sur les 10 sous-ensembles) de la racine de l'erreur quadratique moyenne de la validation croisée, ($CV\text{-}RMSE$) calculée sur l'ensemble de validation. Sinon, c'est celui correspond à une $CV\text{-}RMSE$ qui ne montrant pas une diminution significative.

Outre la $RMSE$ (Équation IV.29) qui indique la précision de la prédiction, les performances de notre modèle sont également mesurées par le R^2 (Équation IV.30) qui donne la proportion de variabilité de la sortie \mathbf{Y} qui est expliquée par le modèle [Asho 16].

$$RMSE = \sqrt{\frac{1}{N} \sum_{i=1}^N \|y_i - \hat{y}_i\|^2} \quad (\text{IV.29})$$

$$R^2 = \frac{\left(\sum_{i=1}^N (y_i - \bar{y}_i)(\hat{y}_i - \bar{\hat{y}}_i) \right)^2}{\sum_{i=1}^N \|y_i - \bar{y}_i\|^2 \sum_{i=1}^N \|\hat{y}_i - \bar{\hat{y}}_i\|^2} \quad (\text{IV.30})$$

où y_i est la valeur des propriétés du matériau, \hat{y}_i est la valeur de la prédiction par le modèle NLPLSR, \bar{y}_i et $\bar{\hat{y}}_i$ sont les moyennes des y_i et \hat{y}_i respectivement, et N est le nombre d'exemples dans la base de données.

IV.4.3. Résultats et discussions

IV.4.3.1. Application sur les données de la simulation

Comme il a été mentionné précédemment, le nombre des composantes latentes a un rôle crucial sur la qualité de prédiction d'un modèle PLSR. La figure IV.8 montre les évolutions de la moyenne de $CV\text{-}RMSE$, sur les 10 bases de validation, en fonction du nombre de variables latentes. Ces évolutions ont été obtenues en utilisant deux bases de données : une construite pour les échantillons liquides et l'autre pour les échantillons solides. À partir de ces courbes, on peut remarquer que pour fournir simultanément ε' et ε'' ($R = 2$) pour les échantillons solides et liquides, 26 et 29 composantes ont été retenues, respectivement. De plus, ça se voit clairement qu'au-delà de ces nombres de composantes principales, les deux modèles entrent dans le phénomène de sur-apprentissage (Over-fitting). En outre, on peut constater le haut niveau d'entraînement (Très faible $CV\text{-}RMSE$) atteint avec un nombre réduit de variables latentes.

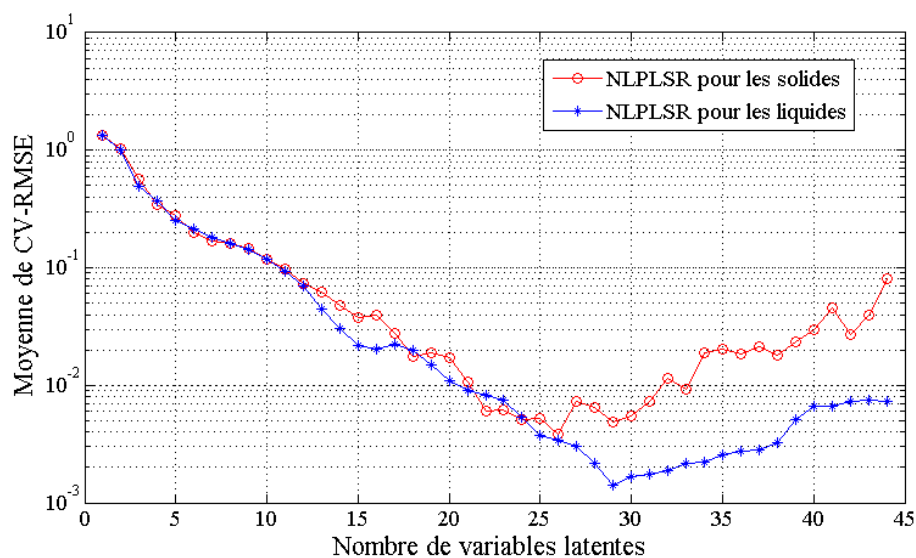


Figure IV.8. Moyenne de la $CV\text{-}RMSE$ en fonction du nombre de variables latentes.

Afin d'illustrer le choix du nombre de variables latentes, le rapport entre la variance de Y et la variance de son approximation par la PLSR (TQ') est calculé, c'est à dire $\text{Var}(TQ') / \text{Var}(Y)$. La figure IV.9 présente la variation de ce rapport en fonction du nombre de variables latentes. Cette figure comporte deux courbes similaires : une correspond aux liquides et l'autre aux solides. Ces courbes montrent bien que le maximum de la variance dans Y ($\approx 100\%$) est largement expliqué par les variables latentes retenues (29 pour les liquides et 26 pour les solides) en utilisant la procédure de la validation croisée. On remarque, clairement, que plus de 97% de la variance est expliquée par les cinq premières variables latentes. La figure IV.10 montre que la variance expliquée dans X ($\text{Var}(TP') / \text{Var}(X)$) par l'analyse PLSR a atteint son maximum ($\approx 100\%$) en utilisant le nombre optimal de variables latentes. Au-delà de ce nombre, la contribution des composantes de rang supérieur est négligeable et elles pourraient être largement affectées par le bruit lors de l'inversion des mesures. Habituellement, ce type de figures donne une idée sur le nombre de variables latentes qui peuvent être, raisonnablement, extraits par la méthode PLSR, contrairement à la figure IV.9 qui indique le nombre approprié de variables latentes requis pour l'inversion des mesures.

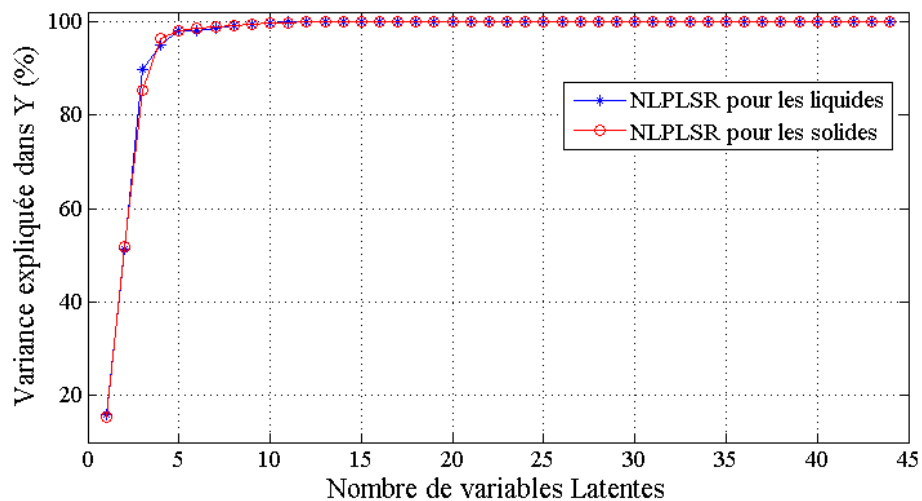
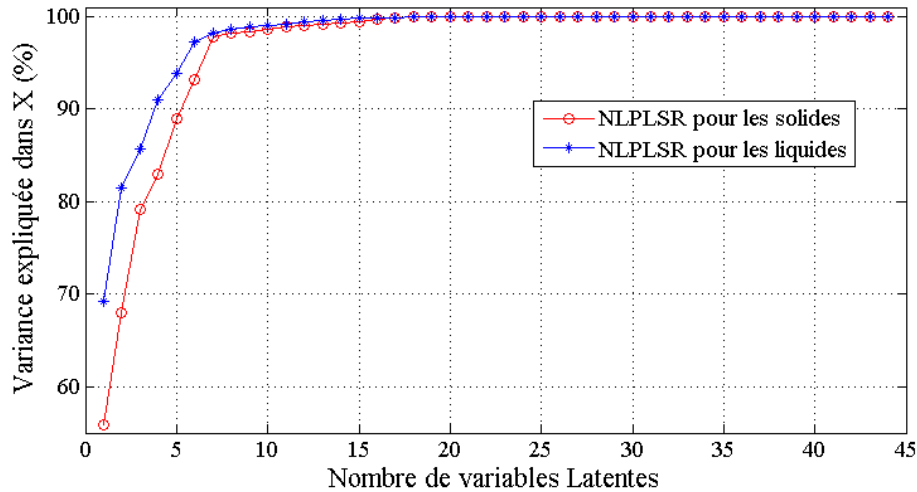


Figure IV.9. Pourcentage de la variance de Y exprimé par la PLSR.

Figure IV.10. Pourcentage de la variance de X exprimé par la PLSR.

Afin d'évaluer les performances des modèles NLPLSR, ils ont été appliqués pour prédire les valeurs de la permittivité inclus dans la base de test, pour les diélectriques liquides et solides. Dans les figures IV.11 et IV.12, les valeurs prédites de la permittivité (Sorties simultanées du modèle NLPLSR) sont tracées en fonction des valeurs souhaitées (Valeurs vraies), ceci pour les parties réelles et imaginaires. Les points tracés, dans chaque figure, forment une droite avec une pente de un, ce qui signifie qu'ils correspondent bien à une équation de forme $\hat{y} = y$ (Permittivité estimée = permittivité désirée).

D'après le tableau IV.6, qui résume les performances de chaque modèle, on peut dire que les $RMSEP$ et les R^2_{pred} calculés sur la base de test montrent bien la très bonne précision obtenue, et par conséquent la très bonne généralisation.

D'après certaines recherches, un modèle est qualifié de bon lorsque l' $RMSEC$ et l' $RMSEP$ sont proches, ce qui est le cas pour cette application [Rahm 16].

Tableau IV.6. Performances des deux modèles NLPLSR.

	Sorties	Base de calibration		Base de test	
		$RMSEC$	R^2_{cal}	$RMSEP$	R^2_{pred}
NLPLSR liquides	ε' et ε''	7.6275×10^{-3}	0.9999999	8.0919×10^{-2}	0.9999920
NLPLSR solides	ε' et ε''	5.0037×10^{-3}	0.9999966	6.0690×10^{-3}	0.9999953

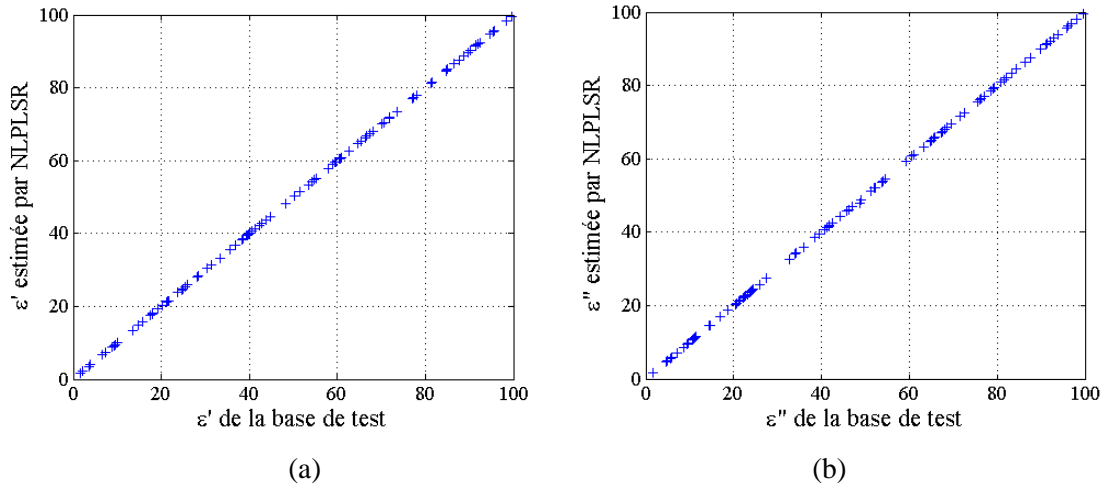


Figure IV.11. Comparaison entre les permittivités fournies par la NLPLSR et celles contenue dans la base de test dédiée aux liquides, (a) partie réelle ε' , (b) partie imaginaire ε'' .

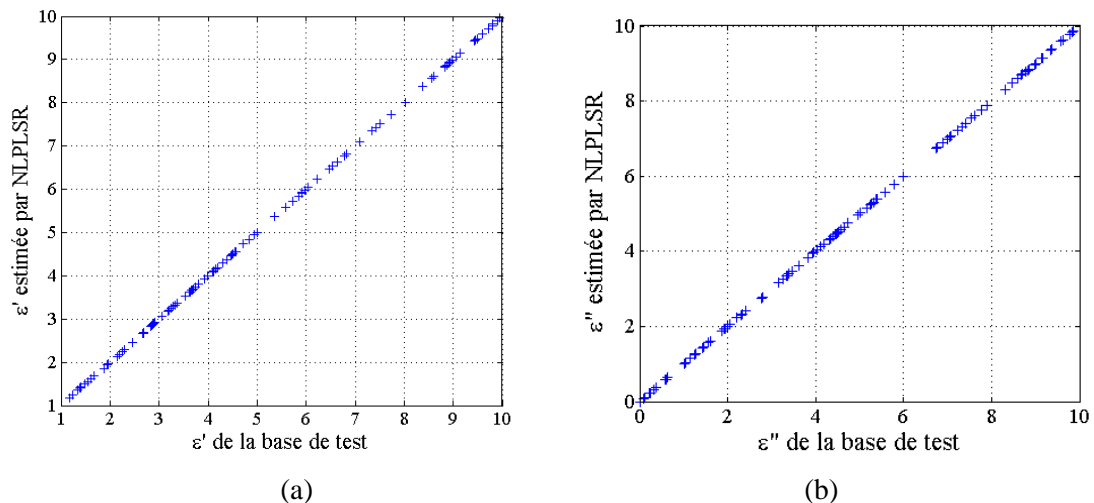


Figure IV.12. Comparaison entre les permittivités fournies par la NLPLSR et celles contenue dans la base de test dédiée aux solides, (a) partie réelle ε' , (b) partie imaginaire ε'' .

IV.4.3.2. Application sur les données expérimentales

Une fois le modèle NLPLSR prouve son efficacité sur les données de simulation, il pourra également être utilisé pour l'inversion des données expérimentales. Comme précédemment, les mesures ont été effectuées en utilisant un analyseur d'impédance Agilent 4291A sur la cellule de mesure SuperMit (Figures III.20 et III.23).

Notons ici que la matrice des prédicteurs résultant des mesures (\mathbf{X}_{mes}) a subi la même transformation que la matrice \mathbf{X} (De même que pour le cas de la MLR) des données simulées. Par conséquent, le nombre de prédicteurs est passé de 3 à 44 prédicteurs centrés et normalisés x_{mes} (Colonnes de \mathbf{X}_{mes}) :

$$x_{\text{mes}} = f_{\text{mes}}^i * G_{\text{mes}}^j * B_{\text{mes}}^k, \quad i = -2, \dots, 2, \quad j = 0, \dots, 2 \quad \text{et} \quad k = 0, \dots, 2 \quad (\text{IV.31})$$

Ici encore, la combinaison lorsque $i=j=k=0$ est exclue.

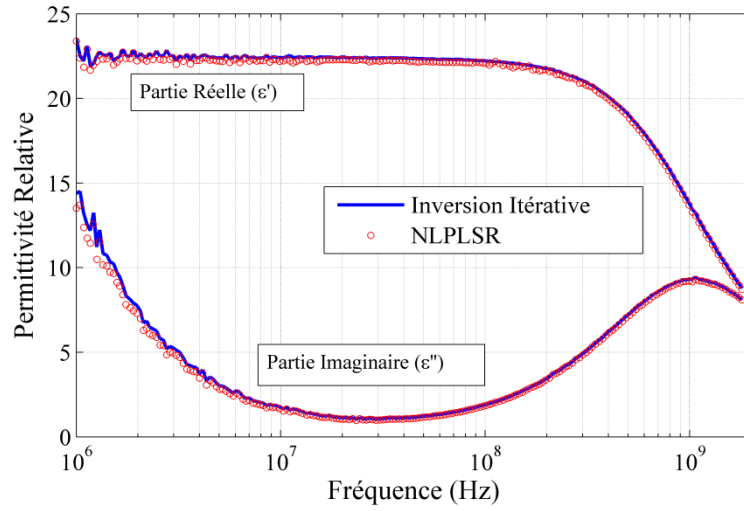
Maintenant, la nouvelle matrice des prédicteurs expérimentaux X_{mes} a la forme suivante :

$$X_{\text{mes}} = \begin{bmatrix} (1/f_{\text{mes}}^2)_1 & (B_{\text{mes}}/f_{\text{mes}}^2)_1 & \cdot & \cdot & (B_{\text{mes}})_1 & \cdot & \cdot & \cdot & (f_{\text{mes}}^2 G_{\text{mes}}^2 B_{\text{mes}}^2)_1 \\ (1/f_{\text{mes}}^2)_2 & (B_{\text{mes}}/f_{\text{mes}}^2)_2 & \cdot & \cdot & (B_{\text{mes}})_2 & \cdot & \cdot & \cdot & (f_{\text{mes}}^2 G_{\text{mes}}^2 B_{\text{mes}}^2)_2 \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ (1/f_{\text{mes}}^2)_{N_{\text{mes}}} & (B_{\text{mes}}/f_{\text{mes}}^2)_{N_{\text{mes}}} & \cdot & \cdot & (B_{\text{mes}})_{N_{\text{mes}}} & \cdot & \cdot & \cdot & (f_{\text{mes}}^2 G_{\text{mes}}^2 B_{\text{mes}}^2)_{N_{\text{mes}}} \end{bmatrix}_{(N_{\text{mes}} \times 44)} \quad (\text{IV.32})$$

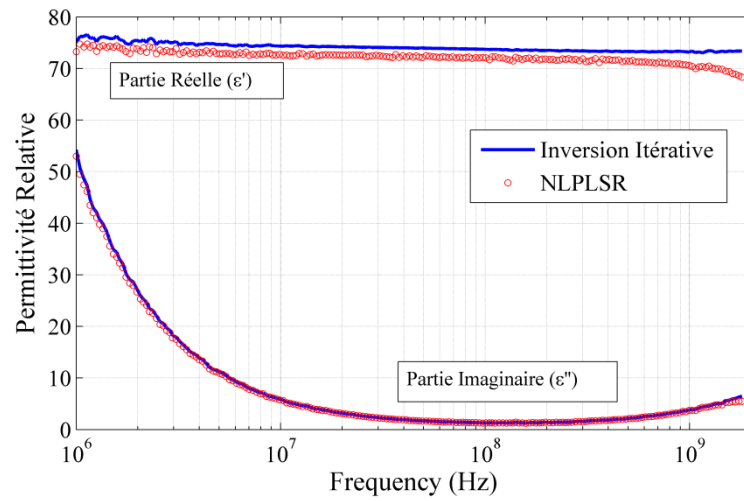
où N_{mes} est le nombre de fréquences de mesure (201).

La raison pour laquelle les indices j et k commencent par 0, contrairement à l'indice i , est d'éviter, en phase de génération (Équations IV.9 et IV.31), les valeurs élevées des prédicteurs lors de la division par G_{mes} , G_{mes}^2 , B_{mes} et B_{mes}^2 . Ce raisonnement est surtout valable pour le cas du PEEK qui est connu comme un matériau sans perte (G très faible). Par conséquent, nous aurons des matrices de prédicteurs de tailles réduites et moins de variables latentes à calculer, donc le modèle NLPLSR devient plus simple et plus rapide. Notons que tous les 44 prédicteurs sont multipliés par B_{PLS} pour obtenir l'inversion des mesures : $Y_{\text{mes}} = X_{\text{mes}} B_{\text{PLS}} + F$, quel que soit le nombre optimal de variables latentes retenues par la validation croisée.

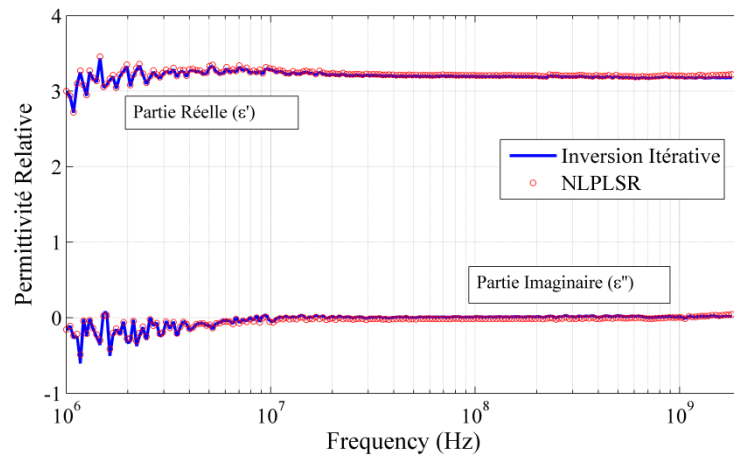
La figure IV.13 montre les résultats de l'inversion des mesures expérimentales en utilisant le modèle NLPLSR pour l'Éthanol, l'eau et le PEEK ainsi que ceux issus de l'inversion itérative. Une bonne concordance entre les résultats obtenus par les deux procédures peut être observée. De plus, ces résultats s'accordent bien avec les études précédentes [Acik 07, 08 ; Haci 10, 11 ; Kupf 05 ; Buck 58 ; Wagn 14 ; Raje 08].



(a) Éthanol



(b) Eau



(c) PEEK

Figure IV.13. Évolution de la permittivité obtenue par la NLPLSR et l'inversion itérative.

IV.4.4. Comparaison entre NLPLSR et MLP

Afin de prouver les hautes performances de la NLPLSR dans la caractérisation micro-ondes, en utilisant une ligne coaxiale à extrémité ouverte, une étude comparative avec un réseau de neurones MLP a été réalisée. Ces deux modèles sont utilisés pour prédire simultanément ε' et ε'' ($R=2$) de l'Éthanol en utilisant la base de données dédiée aux liquides (Cellule SuperMit). La comparaison est réalisée entre les modèles optimaux de la NLPLSR et du MLP obtenus en utilisant la procédure de la validation croisée. Le MLP avec une seule couche cachée comporte 3 unités d'entrée, 30 neurones dans la couche cachée avec fonction d'activation non linéaire (Sigmoid) et deux neurones en couche de sortie avec fonction d'activation linéaire. Ce réseau multicouches a été entraîné en utilisant l'algorithme de Levenberg-Marquardt [Marq 63, Haga 94]. Le modèle NLPLSR (29 variables latentes) est celui décrit à la section IV.4.3.1. Puisque l'erreur ciblée ne peut pas être imposée dans la NLPLSR, la *RMSEC* trouvée précédemment (Tableau IV.6) est utilisée comme critère d'arrêt pour le MLP.

Les résultats présentés dans le tableau IV.7 montrent une très grande différence au niveau des temps nécessaires pour élaborer chaque modèle. Concernant la taille de la base d'apprentissage utilisée par la NLPLSR, il est bien connu que la PLSR a la capacité de prédire les réponses même lorsque les observations sont moins nombreuses que les prédicteurs. Néanmoins, une base d'apprentissage plus large (Nombre important d'observations) donne généralement de meilleures performances [Asho 16]. En effet, le tableau IV.7 montre que l'augmentation du nombre d'exemples dans la base d'apprentissage (De 46 à 2000) sert à diminuer la *RMSEP* et augmenter légèrement la *RMSEC* (Les deux erreurs deviennent très proches l'une de l'autre ce qui est toujours souhaitable) pour atteindre des résultats similaires à ceux du MLP à 2000 exemples.

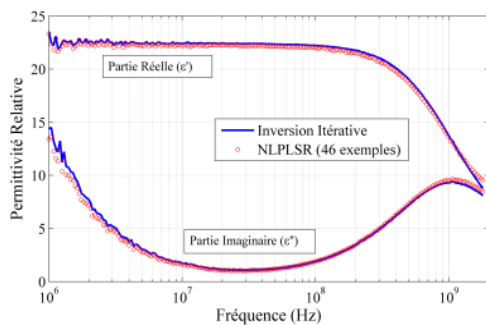
À partir de la figure IV.14 qui montre l'évolution de la permittivité de l'Éthanol avec la fréquence, on peut constater que le un dixième (200 exemples) de la base de données utilisée pour entraîner le modèle MLP optimal est largement suffisante pour la NLPLSR, alors qu'un modèle MLP avec 200 exemples est totalement inutile.

En résumé, un modèle NLPLSR très rapide avec 200 exemples et un modèle MLP très lent avec 2000 exemples sont très satisfaisants avec une variance exprimée de plus de 99% ($R^2_{\text{pred}} \approx 1$). Une base de données réduite est un avantage très important dans les cas où les données disponibles sont limitées telles que la modélisation de la survie dans les recherches médicales ou dans les cas où des mesures expérimentales sont nécessaires pour générer les bases de données.

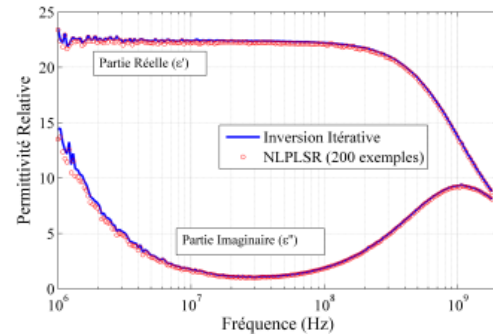
Le tableau IV.7 résume les performances de chaque modèle.

Tableau IV.7. Performances des deux modèles NLPLSR et MLP.

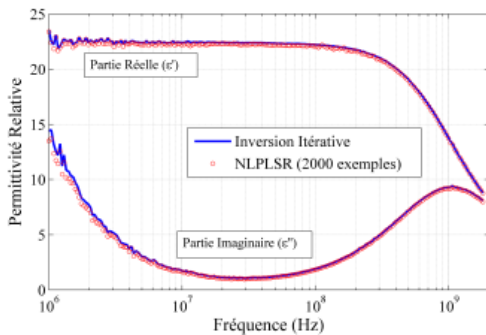
	Temps (s)	RMSEC	R^2_{cal}	Base de calibration	RMSEP	R^2_{pred}	Base de test
NLPLSR	0.019	5.212×10^{-3}	0.9999999	46	4.132×10^{-1}	0.9997870	1000
	0.022	1.235×10^{-2}	0.9999998	200	5.196×10^{-2}	0.9999966	1000
	0.060	2.038×10^{-2}	0.9999994	2000	2.509×10^{-2}	0.9999992	1000
MLP	451	4.270×10^{-2}	0.9999977	2000	4.944×10^{-2}	0.9999970	1000
	4.59	7.2297	0.93674	200	17.833	0.71027	1000



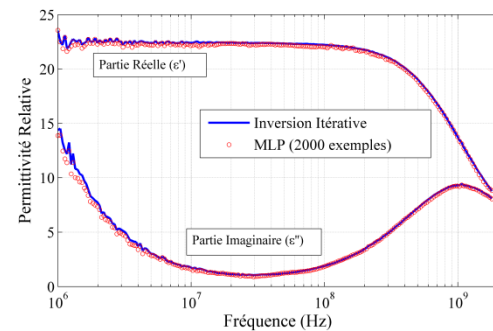
(a) NLPLSR avec 46 exemples



(b) NLPLSR avec 200 exemples



(c) NLPLSR avec 2000 exemples



(d) MLP avec 2000 exemples

Figure IV.14. Permittivité de l'Éthanol.

Conclusion

Dans ce chapitre, la MNLR et la NLPLSR, qui sont deux méthodes de régression non linéaire, sont investies pour résoudre le problème inverse. Généralement, la MLR et la PLSR sont utilisées comme outils d'inversion linéaires pour résoudre des problèmes avec un nombre important de variables explicatives. Dans cette étude, nous étions dans l'obligation de concevoir des modèles de

régression non-linéaires, en créant mathématiquement de nouveaux prédicteurs, afin d'intégrer une non linéarité entre les entrées et les sorties. Les résultats d'inversion sur les mesures confirment l'adéquation et l'efficacité de ces méthodes dans la caractérisation micro-ondes des matériaux.

D'après les études comparatives réalisées, il est très clair que les deux modèles de régression non-linéaires (MNLR et NLPLSR) surpassent le MLP et l'ANFIS en rapidité de calculs et avec une précision très satisfaisante. Cependant, l'ANFIS est le meilleur en terme de précision tout en utilisant des bases de tailles relativement réduites (Comparables à celles de la MNLR et de la NLPLSR) par rapport à celles que nécessite le MLP. Ce dernier présente, avec la NLPLSR, l'avantage de traiter des problèmes multi-sorties.

Conclusion et perspectives

Sur une large gamme de fréquence allant du quasi-statique jusqu'aux micro-ondes, les ondes électromagnétiques permettent d'effectuer des contrôles non destructifs (Milieu industriel) et non invasifs (Milieu médical). En effet, c'est à partir de la réponse d'un échantillon sous teste, suite à son excitation par une onde électromagnétique, que peuvent être mesurés les propriétés physiques au sein des matériaux industriels ou inhomogénéités dans des tissus biologiques.

L'objectif de cette thèse était de développer des modèles d'inversion directe (Modèles comportementaux paramétriques) construit au moyen de bases de données obtenues par simulations numériques. Ces modèles ont été utilisés pour la caractérisation micro-ondes des matériaux diélectriques en utilisant des protocoles de caractérisation large bande en structure cylindrique (Cellule coaxiale). Le problème direct a été résolu en utilisant la méthode des éléments finis (FEM, pour Finite Element Method). Cette dernière est considérée comme un outil efficace pour résoudre numériquement des problèmes électromagnétiques. Deux bases de données ont été générées : une aléatoire dans le cas d'une cellule SuperMit, et une autre régulière dans le cas d'une cellule SuperPol.

Premièrement, le modèle ANFIS (Adaptive Network based Fuzzy Inference System) est adopté. Ce dernier est un modèle hybride neuro-flous qui combine les concepts des réseaux de neurones artificiels (ANN, pour Artificial Neural Network) et ceux de la logique floue. Il s'agit d'un système d'inférence flou (FIS, pour Fuzzy Inference System) mis en forme d'un réseau adaptatif pour acquérir les capacités d'apprendre inspirées des ANN à l'aide d'une base de données. En effet, au cours du troisième chapitre, ANFIS a été investi pour résoudre le problème inverse impliquant des données expérimentales obtenues en utilisant la cellule SuperMit. Trois matériaux ont été caractérisés : deux liquides (L'eau et l'Ethanol) et un solide (Le PEEK). Les résultats d'inversion donnent l'impression qu'il s'agit d'un modèle très efficace et très convenable pour notre problème. Effectivement, l'ANFIS converge rapidement et donne, après apprentissage, une estimation de la permittivité diélectrique complexe des matériaux d'une manière instantanée (En temps réel). De plus, l'apprentissage se fait à l'aide d'une base de données de taille relativement réduite (Environ 250 exemples), contrairement à d'autres outils d'inversion (ANN et SVM pour Support Vector Machines) utilisés auparavant dans des travaux similaires (Environ 2500 exemples).

Cependant, avec les entrées originales (f , G et B), le modèle ANFIS a rencontré de sérieux problèmes pour qu'il arrive à une meilleure capacité de généralisation. Pour cela, nous étions obligés de chercher les entrées optimales du modèle à partir d'une liste des entrées obtenues par combinaison des entrées originales. Le passage par une procédure de sélection a permis à l'ANFIS d'atteindre un niveau très élevé de prédiction après une courte phase d'entraînement. Comme la plus part des modèles paramétriques, l'ANFIS nécessite un bon choix du type et du nombre de fonctions d'appartenance attribué à chaque entrée ainsi qu'un prétraitement des données.

D'autres méthodes, totalement différentes des réseaux neuro-flous, ont été également investies. Il s'agit de deux modèles de l'analyse multi-variée, souvent utilisés dans le domaine de la chimométrie : la MLR (Multiple Linear Regression) et la PLSR (Partial Least-Squares Regression). Ce sont des modèles statistiques qui analysent un ensemble d'observations pour ajuster un modèle qui relie les entrées et les sorties du phénomène physique observé. À l'origine, la MLR et la PLSR sont deux modèles linéaires, or une relation non-linéaire est certainement présente entre les entrées et les sorties de notre problème, donc deux modèles de régression non-linéaire ; la MNLR (Multiple Non-Linear Regression) et la NLPLSR (Non-Linear PLSR) ont été développés en se basant sur la technique du bloc des prédicteurs étendu. Cette technique de délinéarisation externe consiste à ajouter mathématiquement les termes au carré, les produits croisés et éventuellement les termes d'ordre supérieurs des prédicteurs originaux (f , G et B).

Dans la première partie du quatrième chapitre, le modèle MNLR a été utilisé pour résoudre le problème inverse impliquant la cellule de mesure SuperPol. Les mesures ont été effectuées sur un échantillon liquide (Éthanol). Avant l'implantation de la MNLR, la sélection progressive a été sollicitée pour choisir les variables pertinentes (Entrées pertinentes). Cette technique ajoute et supprime successivement les variables en se basant sur des mesures statistiques pour garder à la fin que les variables les plus statistiquement significatives. Les résultats obtenus avec les prédicteurs optimaux ont démontré l'efficacité et la robustesse de ce modèle qui peut être considéré comme un outil simple, rapide et facilement interprétable. De plus, la MNLR offre une prédiction de bonne qualité avec une base d'apprentissage de taille relativement réduite (250 exemples).

Dans la deuxième partie du quatrième chapitre, la NLPLSR a été introduit pour résoudre le problème inverse impliquant à nouveau les mesures effectuées sur la cellule SuperMit. Comme c'était le cas avec la MNLR, la matrice des prédicteurs a été étendue pour contenir d'autres combinaisons des entrées originales. Avant l'implémentation de ce modèle, la procédure de validation croisée est utilisée pour choisir le nombre optimal de composantes principales qui

offrent au modèle les meilleures performances. A partir des résultats de simulation ainsi que ceux expérimentaux, on a pu constater les excellentes performances que présente le modèle multi-sortie utilisé en termes de rapidité et de précision.

Si on veut apporter un jugement sur les modèles exposés dans cette thèse, on préfère sûrement les modèles d'analyse multi-variée en raison de leur simplicité, rapidité et qualité de prédiction. Parmi ces modèles de régression non-linéaires, la NLPLSR est, sans aucun doute, mieux que la MNLR car elle est capable de prédire simultanément plusieurs sorties. Cependant, pour des problèmes à sortie unique, l'ANFIS offre la meilleure précision de prédiction.

Au cours de cette thèse, on souligne régulièrement le fait d'avoir développé des modèles à réponse instantanée avec des bases de données issues de la simulation de tailles réduites. Effectivement, nous avons profité de la souplesse des modèles numériques pour générer facilement, une fois pour toute, des bases de données qui peuvent avoir différentes tailles et bornes de paramètres. Ainsi, on a évité de passer par une génération expérimentale d'exemples qui sera sûrement fastidieuse et pénible. L'importance des bases réduites se manifeste dans le cas des problèmes qui n'admettent pas de modèles mathématiques ou les études qui nécessitent la collecte d'observations sur les êtres vivantes ou les phénomènes physiques. Tous les modèles développés dans cette thèse effectuent, après apprentissage ou calibration, une inversion en temps réel. Donc, en pratique, ils seront très convenables à être intégré dans des appareils de mesure portables, contrairement à l'inversion itérative qui demande un peu de temps car le problème direct est résolu plusieurs fois pour chaque point de fréquence.

Les perspectives portent, notamment, sur l'application des modèles exposés dans ce manuscrit à des situations plus complexes et réalistes. Ainsi leur extension à des gammes de permittivités et de fréquences élevées est à étudier. Les matériaux à prendre en considération sont notamment les liquides biologiques, les matériaux composites et les tissus humains. Cependant, il faut s'attendre à affronter les problèmes de la non linéarité et l'apparition de résonances. Donc, le développement d'une stratégie de génération adaptative des exemples des bases de données sera intéressant.

Bibliographie

- [Abdi 10] Abdi H. "Partial least squares regression and projection on latent structure regression (PLS Regression)". Wiley Interdiscip. Rev. Comput. Stat. Vol. 2, pp. 97–106, 2010.
- [Acik 07] Acikgoz H, Le Bihan Y, Meyer O, Pichon L. "Neural networks for broad-band evaluation of complex permittivity using a coaxial discontinuity". The European Physical Journal Applied Physics. Vol. 39 No. 2, pp. 197–201, 2007.
- [Acik 08] Acikgoz H, Le Bihan Y, Meyer O, Pichon L. "Microwave characterization of dielectric materials using Bayesian neural networks". Progress in Electromagnetics Research C. Vol. 3, pp. 169–82, 2008.
- [Acikg 08] Acikgoz H, Santandréa L, Bihan Y, Gyimothy S, Pavo J, Meyer O, Pichon L. "Generation and use of optimised databases in microwave characterisation". IET Science Measurement and Technology, Institution of Engineering and Technology. Vol. 2, No 6, pp. 467–473, 2008.
- [Acikgo 08] Acikgoz H. "Technique d'inversion associant la modélisation numérique et les réseaux de neurones pour la caractérisation microondes de matériaux". Thèse de doctorat, université de Pierre et Marie Curie, Paris 2008.
- [Adhi 10] Adhikari N, Maiti M K, Jha T. "Predictive comparative QSAR modelling of (phenylpiperazinyl-alkyl) oxindoles as selective 5-HT_{1A} antagonists by stepwise regression, PCRA, FA-MLR and PLS techniques". European Journal of Medicinal Chemistry. Vol. 45, pp. 1119–1127, 2010.
- [Alin 09] Alin A. "Comparison of PLS algorithms when number of objects is much larger than number of variables". Stat Papers. Vol. 50, pp. 711–720. 2009.
- [Alot 08] Alotaibi F D, Abdennour A, Ali A A. "A robust prediction model using ANFIS based on recent TETRA outdoor RF measurements conducted in Riyadh city – Saudi Arabia. AEU – ". International Journal of Electronics and Communications. Vol. 62, pp. 674–682. 2008.
- [Asho 16] Ashourloo D, Aghighi H, Matkan A A, Mobasheri M R and Rad A M. "An Investigation Into Machine Learning Regression Techniques for the Leaf Rust Disease Detection Using Hyperspectral Measurement". IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing.; Vol. 9, pp. 4344–4351, 2016.
- [Bagh 12] Bagheri M, Rajabi M, Mirbagheri M, Amin M. "BPSO-MLR and ANFIS based modeling of lower flammability limit". Journal of Loss Prevention in the Process Industries Vol. 25, pp. 373–382. 2012.

- [Bake 04] Baker-Jarvis J, Janezic M D, Riddle B, Johnk R T, Kabos P, Holloway C L, Geyer R G, Grosvenor C A. "Measuring the permittivity and permeability of lossy materials: Solids, liquids, metals, building materials, and negative-index materials". Technical report, National Institute of Standards and Technology - NIST TN1536, 2004.
- [Baou 14] Baouche R. "Prédiction des Paramètres Physiques des Couches Pétrolifères par Analyse des Réseaux de Neurones et Analyse Faciologique". These de doctorat, Université M'hamed Bougara-Boumerdes Algérie. 2014.
- [Belh 86] Belhadj-Tahar N-E, Fourier-Lamer A. "Broad-band analysis of a coaxial discontinuity used for dielectric measurements", IEEE Transactions on Microwave Theory and Techniques, Vol. 34 No. 3, pp. 346-350. 1986.
- [Belh 90] Belhadj-Tahar N-E, Fourier-Lamer A, De Chanterac, H. "Broad-band simultaneous measurement of complex permittivity and permeability using a coaxial discontinuity", IEEE Transactions on Microwave Theory and Techniques, Vol. 38, No. 1, pp. 1–7, 1990.
- [Belh 97] Belhadj-Tahar N-E, Meyer O, Fourier-Lamer A. "Broad-band microwave characterization of bilayered materials using a coaxial discontinuity with applications for thin conductive films for microelectronics and material in air-tight cell", IEEE Transactions on Microwave Theory and Techniques, Vol. 45 No. 2, pp. 260–267, 1997.
- [Belh 98] Belhadj-Tahar N-E, Dubrunfaut O, Fourier-Lamer A. "Broad-band microwave characterization of a tri-layer structure using a coaxial discontinuity with applications for magnetic liquids and films", IEEE Transactions on Microwave Theory and Techniques. Vol. 46 No. 12, pp. 2109–2116, 1998.
- [Belk 08] Belkhir A. "Extention de la modélisation par FDTD en nano-optique", Thèse de doctorat, Université Mouloud Mammeri de Tizi ousou. 2008.
- [Berg 97] Berglund A, Wold S. "INLR, Implicit Non-Linear Latent Variable Regression". Journal of Chemometrics. Vol. 11, pp. 141–156, 1997.
- [Bond 08] Bondet de la Bernardie E, Dubrunfaut O, Badot J C, Fourier-Lamer A, Villard E, David P Y, Jannier B, Grosjean N, Lance M. "Low (10 MHz – 800 MHz) and high (40 GHz) frequencies probes applied to petroleum multiphase flow characterization", Measurement Science and Technology. Vol.19, pp. 055602–055610. 2008.
- [Bouc 16] Bouchon A. "Modèles de régression multivariés pour la comparaison de populations en IRM de diffusion". Thèse de doctorat, , université de Strasbourg, 2016.
- [Boul 06] Boulesteix A L, Strimmer K. "Partial least squares: a versatile tool for the analysis of high-dimensional genomic data". Briefings in Bioinformatics. Vol. 8, pp. 32–44, 2006.

- [Buck 58] Buckley F, Maryott A. "Tables of dielectric dispersion data for pure liquids and dilute solutions". Washington, D. C. : U.S. Dept. of Commerce, National Bureau of Standards. (1958),
- [Breb 87] Brebbia C A, Magureanu R. "The boundary element method for electromagnetic problems". *Journal of engineering analysis*. Vol. 04, No. 04, pp. 178-185, 1987.
- [Bado 88] Badot J C. "Relaxation diélectriques [1kHz-10GHz] dans les conducteurs protoniques et les matériaux hydratés", Thèse de doctorat, université de Pierre et Marie Curie, Paris, 1988.
- [Bakh 93] Bakhtiari S, Ganchev S I, Zoughi R. "Open-Ended Rectangular Waveguide for Nondestructive Thickness Measurement and Variation Detection of Lossy Dielectric Slabs Backed by a Conducting Plate". *IEEE Transactions on Instrumentation and Measurement*. Vol. 42, No. 1, pp. 19–24, 1993.
- [Chan 11] Chang C-W, Yu W-C, Chen W-J, Chang R-F, Kao W-S. "A study on the enzymatic hydrolysis of steam exploded napiergrass with alkaline treatment using artificial neural networks and regression analysis". *Journal of the Taiwan Institute of Chemical Engineers*. Vol. 42, pp. 889–894, 2011.
- [Chel 14] Chelabi M, Hacib T, Belli Z, Mekideche M, Le Bihan Y, "Optimal database combining with Multi Output Support Vector Machine for Eddy Current Testing", Ninth International Conference on Ecological Vehicles and Renewable Energies (EVER) , MC, 25 March 2014, pp. 5 p.
- [Chel 16] Chelabi M, Hacib T, Le Bihan Y, Ikhlef N, Boughedda H, Mekideche M R. "Eddy current characterization of small cracks using least square support vector machine". *J. Phys. D: Appl. Phys.* Vol. 49, pp. 155303 (7pp). 2016.
- [Chik 14] Chikh A, Chandra A. "Adaptive neuro-fuzzy based solar cell model". *IET Renewable Power Generation*. Vol. 8, pp. 679–686, 2014.
- [Cole 41] Cole K S, Cole R H. "Dispersion and absorption in dielectrics. I. Alternating current characteristics". *Journal of Chemical Physics*. Vol. 9, pp. 341–351, 1941.
- [Davi 51] Davidson D W, Cole R H. "Dielectric Relaxation in Glycerol, Propylene Glycol, and n-Propanol". *The Journal of Chemical Physics*. Vol. 19, No. 12, pp. 1484–1490, 1951.
- [Deby 29] Debye P. "Polar molecules". *The chemical catalogue*. New York, 1929.
- [De Jo 93] De Jong S. "SIMPLS: An alternative approach to partial least squares regression". *Chemometrics and Intelligent Laboratory Systems*. Vol. 18, pp. 251–263, 1993.
- [Dubi 13] Dubitzky W, Wolkenhauer O, Cho K H, Yokota H. "Encyclopedia of Systems Biology". Springer, 2013.

- [Dura 01] Durand J F. "Local polynomial additive regression through PLS and spline: PLSS". *Chemometrics and Intelligent Laboratory Systems*. Vol. 58, pp. 235-246, 2001.
- [EL HA 13] El Haddad J. "chimie appliquée à la spectroscopie de plasma induit par laser (LIBS) et à la spectroscopie terahertz", Thèse de Doctorat, Université de Bordeaux 1, 2013.
- [Elma 16] Elmajid H, Terhzaz J, Ammor H, Chaïbi M, Sánchez A M. "Determination of the Complex Permittivity of Each Layer for a Bi-layer Dielectric Material Using Transmission (ABCD) Matrix in Ku-Band Frequency". *Journal of Engineering Science and Technology Review*. Vol. 9, No. 1, pp. 61-65, 2016.
- [Fila 06] Filali B, Rhazi J E, Ballivy G. "Mesure des propriétés diélectriques du béton par une large sonde coaxiale à terminaison ouverte". *Can. J. Phys.* Vol. 84, pp. 365–379, 2006.
- [Fran 90] Frank I E. "A nonlinear PLS model". *Chemometrics and Intelligent Laboratory Systems*. Vol. 8, pp. 109–119, 1990.
- [Gham 17] Ghamali M, Chtita S, Ousaa A, Elidrissi B, Bouachrine M, Lakhlifi T. "QSAR analysis of the toxicity of phenols and thiophenols using MLR and ANN". *Journal of Taibah University for Science*. Vol. 11, pp. 1–10, 2017.
- [Ghun 12] Ghunem R A, Assaleh K, El-Hag A H. "Artificial Neural Networks with Stepwise Regression for Predicting Transformer Oil Furan Content". *IEEE Transactions on Dielectrics and Electrical Insulation* Vol. 19, No. 2, pp. 414–420, 2012.
- [Ghun 14] Ghunem R A, El-Shatshat R, Ozgonenel O. "A Novel Selection Algorithm of a Wavelet-Based Transformer Differential Current Features". *IEEE Transactions on Power Delivery* . Vol. 29, No. 3, 2014.
- [Gifi 90] Gifi A. "Nonlinear multivariate analysis". Wiley, New York. 1990.
- [Goh 14] Goh A T C, Zhang W G. "An improvement to MLR model for predicting liquefaction-induced lateral spread using multivariate adaptive regression splines". *Engineering Geology*. Vol. 170, pp. 1–10, 2014.
- [Good 18] <http://www.goodfellow.com/F/Polyetherethercetone.HTML>, accédé en Mai 2018.
- [Greg 12] Gregory A P, Clarke R N. "Tables of the Complex Permittivity of Dielectric Reference Liquids at Frequencies up to 5 GHz". National Physical Laboratory. 2012.
- [Grig 12] Grigis A. "Approches statistiques pour la détection de changements en IRM de diffusion : application au suivi longitudinal de pathologies neuro-dégénératives". Thèse de doctorat, université de Strasbourg, 2012.

- [Haci 10] Hacib T, Le Bihan Y, Mekideche M R, Acikgoz H, Meyer O, Pichon L. "Microwave Characterization Using Least-Square Support Vector Machines", IEEE Transactions on Magnetics, Vol. 46 No. 8, pp. 2811–2814, 2010.
- [Hacib 10] Hacib T, Acikgoz H, Le Bihan Y, Mekideche M R, Meyer O, Pichon L. "Support vector machines for measuring dielectric properties of materials", COMPEL: The International Journal for Computation and Mathematics in Electrical and Electronic Engineering, Vol. 29 No. 4, pp. 1081–1089, 2010.
- [Haci 11] Hacib T, Le Bihan Y, Smail M K, Mekideche M R, Meyer O, Pichon L. "Microwave Characterization Using Ridge Polynomial Neural Networks and Least-Square Support Vector Machines", IEEE Transactions on Magnetics, Vol. 47 No. 5, pp. 990–993, 2011.
- [Haga 94] Hagan M T, Menhaj M. "Training feed-forward networks with the Marquardt algorithm". IEEE Transactions on Neural Networks. Vol. 5, pp. 989–993, 1994.
- [Hamd 13] Hamdan H. "An Exploration of the Adaptive Neuro-Fuzzy Inference System (ANFIS) in Modelling Survival". These de doctorat, université de Nottingham, Nottingham, Royaume-Uni. 2013.
- [Havr 67] Havriliak S, Negami S. "A complex plane representation of dielectric and mechanical relaxation processes in some polymers", Polymer, Vol. 8 No. 4, pp. 161–210, 1967.
- [Hine 97] Hines J W. "Fuzzy and neural approaches in engineering: MATLAB Supplement". John Wiley and Sons: New York, NY; 1997.
- [Huan 12] Huang D-Y, Sun W. "A Comparison of SVM and Asymmetric SIMPLS in Emotion Recognition from Naturalistic Dialogues". Circuits and Systems (ISCAS), 2012 IEEE International Symposium on. Seoul, South Korea.
- [Jang 93] Jang J-S R. "ANFIS: Adaptive-Neuro-Based Fuzzy Inference System", IEEE Transactions on Systems, Man, and Cybernetics. Vol. 23 No. 3, pp. 665–685, 1993.
- [Jang 97] Jang J-S R, Sun C T, Mizutani E. "Neuro-Fuzzy and Soft Computing: A Computational Approach to Learning and Machine Intelligence". Prentice Hall: Upper Saddle River, NJ 07458, 1997.
- [Jang 95] Jang J-S R, Sun C T. "Neuro-fuzzy modelling and control". Proceedings of the IEEE. Vol. 83, pp. 378–406, 1995.
- [Jang 93] Jang J-S R. "ANFIS: Adaptive-Neuro-Based Fuzzy Inference System", IEEE Transactions on Systems, Man, and Cybernetics. Vol. 23 No. 3, pp. 665–685, 1993.

- [Jang et Sun 93] Jang J-S R, Sun C T. "Functional Equivalence Between Radial Basis Function Networks and Fuzzy Inference Systems", *IEEE Transactions on Neural Networks*, Vol. 4 No. 1. pp. 156–159. 1993.
- [Jang 96] Jang J-S R. "Input Selection for ANFIS Learning", in *Fuzzy Systems proceedings of the Fifth IEEE International Conference in New Orleans, Louisiana, USA, 11-11 Sept. 1996*, pp. 1493–1499, 1996.
- [Kaat 89] Kaatze U. "Complex Permittivity of Water as a Function of Frequency and Temperature", *Journal of Chemical and Engineering Data*, Vol. 34 No. 4, pp. 371–374, 1989.
- [Kaat 96] Kaatze U. "Microwave dielectric properties of water", in: Kraszewski, A. (Ed.), "Microwave aquametry. Electromagnetic wave interactions with water-containing materials", *IEEE Press, New York*, pp. 37–53, (1996).
- [Kaat 05] Kaatze U. "Electromagnetic Wave Interactions with Water and Aqueous Solutions", in Kupfer K (Ed.), "Electromagnetic Aquametry. Electromagnetic Wave Interaction with Water and Moist Substances", *Springer-Verlag, Berlin: Heidelberg*, pp. 15–37, 2005.
- [Kaat 10] Kaatze U. "Techniques for measuring the microwave dielectric properties of materials", *Metrologia*, Vol. 47 pp. S91–S113, 2010.
- [Karr 00] Karray F. "Soft computing techniques for intelligent machines", C. de Silva, Editor, *Intelligent Machines*. CRC Press. Florida; pp. 337–340, 2000.
- [Kolo 71] Kolodziej H, Sobczyk L. "Investigation of dielectric properties of potassium hydrogen maleate", *Acta Physica Polonia*, A39, p. 59, 1971.
- [Koma 16] Komarov S A, Komarov A S, Barber D G, Lemes M J L, Rysgaard S. "Open-Ended Coaxial Probe Technique for Dielectric Spectroscopy of Artificially Grown Sea Ice", *IEEE Transactions on Geoscience and Remote Sensing*, Vol. 54 No. 8 , pp. 4941–4951, 2016.
- [Krem 03] Kremer F. Schönhals A. "Broadband Dielectric Spectroscopy", *Springer-Verlag Berlin Heidelberg GmbH*, 2003.
- [Kupf 05] Kupfer K. "Electromagnetic Aquametry, Electromagnetic Wave Interaction with Water and Moist Substances". *Springer-Verlag, Berlin: Heidelberg*, 2005.
- [Le Bi 14] Le Bihan Y, Pavo J, Marchand C. "Partial least square regression: an analysis tool for quantitative non-destructive testing". *Eur. Phys. J. A ppl. P hys.* Vol. 67, pp. 30901(7pp), 2014.
- [Li 10] Li Y, Zhencai Z, Jianyi L, Hailong T. "Modeling on the Drum Loading Performance of the Helix Shearer Based on Partial Least Square Regression". *Advanced Technology of Design and Manufacture (ATDM 2010)*, *International Conference on*. 23-25 Nov. 2010, pp 134-137, Beijing, China.

- [Mamd 75] Mamdani E H, Assilian S. "An experiment in linguistic synthesis with a fuzzy logic controller". *International Journal of Man-Machine Studies*. Vol. 7, pp. 1–13, 1975.
- [Marq 63] Marquardt D W. "An algorithm for least-squares estimation of nonlinear parameters". *Journal of the Society for Industrial and Applied Mathematics*. Vol. 11, pp. 431–441, 1963.
- [Mare 04] Marescot L. "Modélisation direct et inverse prospection électrique sur des structures 3D complexes par la méthode des éléments finis". Thèse de Doctorat, Ecole Doctorale de Nantes, Juin 2004.
- [Mari 06] Maria S. "Modélisation parcimonieuse : application à la sélection de variables et aux données STAP". Thèse de doctorat, Université de Rennes 1, 2006.
- [Mata 11] Mata J. "Interpretation of concrete dam behaviour with artificial neural network and multiple linear regression models". *Journal of Engineering Structures*. Vol. 33, pp. 903–910, 2011.
- [Meun 07] Meunier B B. "La logique floue, Que sais-je". Presses Universitaires de France, 4ème édition. 2007.
- [Meye 96] Meyer O. "Instrumentation pour un contrôle de processus de réticulation sous microondes par caractérisation diélectrique large bande. Contribution à la compréhension du couplage du champ électromagnétique aux dipôles moléculaire", thèse de doctorat, université de Pierre et marie Curie, Paris, 1996.
- [Miku 05] Mikut R, Jakel J, Groll L. "Interpretability issues in data-based learning of fuzzy systems". *Fuzzy Sets and Systems*. Vol. 150, pp. 179 – 197. 2005.
- [Misr 95] Misra D. "On the Measurement of the Complex Permittivity of Materials by an Open-Ended Coaxial Probe". *Ieee Microwave and Guided Wave Letters*. Vol. 5, No. 5, 1995.
- [Moav 99] Moavini S. "Finite Element Analysis, Theory and Application with ANSYS", Prentice Hall, Upper Saddle River, New Jersey 07458, 1999.
- [Orlo 14] Orloff N D, Obrzut J, Long C J, Lam T, Kabos P, Novotny D R, Booth J C, Liddle J A. "Dielectric Characterization by Microwave Cavity Perturbation Corrected for Nonuniform Fields". *IEEE transactions on microwave theory and techniques* , Vol. 62, No. 9, 2014.
- [Pill 14] Pillai G N, Pushpak J, Nisha M G. "Extreme learning ANFIS for control applications". *IEEE Symposium on Computational Intelligence in Control and Automation (CICA)*; 2014 Dec 9–12; Orlando, FL.

- [Pour 14] Pourbasheer E, Aalizadeh R, Ganjali M R, Norouzi P, Shadmanesh J. "QSAR study of ACK1 inhibitors by genetic algorithm–multiple linear regression (GA–MLR)". *Journal of Saudi Chemical Society*. Vol. 18, pp. 681–688, 2014.
- [Prem 00] Premat E. "Prise en compte d'effets météorologiques dans une méthode d'éléments finis de frontière". Thèse de doctorat de l'institut national des sciences appliquées de Lyon, 2000.
- [Qin 92] Qin S J, McAvoy T J. "Nonlinear PLS modeling using neural networks". *Computers Chem. Engng*, Vol. 16, No. 4, pp. 379-391, 1992.
- [Raco 06] Racoceanu D. "Contribution à la surveillance des Systèmes de Production en utilisant les Techniques de l'Intelligence Artificielle". Habilitation à diriger des recherches, Université de FRANCHE-COMTÉ de Besançon, 2006.
- [Rahm 16] Rahman A, Kondo N, Ogawa Y, Suzuki T, Kanamori K. "Determination of K value for fish flesh with ultraviolet-visible spectroscopy and interval partial least squares (iPLS) regression method". *Biosystems Engineering*. Vol. 141, pp. 12–18, 2016.
- [Raje 08] Rajesh S, Murali K P, Priyadarsini V, Potty S N, Ratheesh R, Mohanan P. "Microwave Dielectric Properties of Rutile Filled PEEK Composites", *Polymer-Plastics Technology and Engineering*, Vol. 47 No. 3, pp. 242–246, 2008.
- [Robi 15] Robinson N, Guan C, Vinod A P. "Adaptive estimation of hand movement trajectory in an EEG based brain–computer interface system". *J. Neural Eng.* Vol. 12, pp. 066019 (12pp), 2015.
- [Roy 09] Roy K, Roy P P. "Comparative chemometric modeling of cytochrome 3A4 inhibitory activity of structurally diverse compounds using stepwise MLR, FA-MLR, PL S, GFA , G/PLS and ANN techniques". *European Journal of Medicinal Chemistry*. Vol. 44, pp. 2913–2922, 2009.
- [Ruih 07] Ruihua Li, Meng G, Gao N, Xie H. "Combined use of partial least-squares regression and neural network for residual life estimation of large generator stator insulation". *Measurement Science and Technology*. Vol. 18, pp. 2074–2082, 2007.
- [Ruiz 13] Ruiz M, Mujica L E, Berjaga X, Rodellar J. "Partial least square/projection to latent structures (PLS) regression to estimate impact localization in structures". *Smart Materials and Structures*. Vol. 22, pp. 025028 (11pp), 2013.
- [Selm 14] Selmani A, Seddik H, Ben braiek E. "Adaptive median filter based on ANFIS for impulse noise suppression". *IEEE 23rd International Symposium on Industrial Electronics (ISIE)*; 2014 June 1-4; Istanbul.

[Shek 17] Shekofteh H, Ramazani F, Shirani H. "Optimal feature selection for predicting soil CEC: Comparing the hybrid of ant colony organization algorithm and adaptive network-based fuzzy system with multiple linear regression". *Geoderma*. Vol. 298, pp. 27–34, 2017.

[Sidd 13] Siddique N, Adeli H. "Computational intelligence: synergies of fuzzy logic, neural networks and evolutionary computing", John Wiley and Sons, Ltd, Publication; 2013.

[Suge 85] Sugeno M. "Industrial applications of fuzzy control". Elsevier Science Pub. Co, 1985.

[Sun 16] Sun S, Bezemer R, Long X, Muehlsteff J, Aarts R M. "Systolic blood pressure estimation using PPG and ECG during physical exercise". *Physiol. Meas.* Vol. 37, pp. 2154–2169, 2016.

[Syam 15] Syamsiah A B, Rosma M D, Ajab B A, Wan Hasamudin W H. "Role of Input Selection Prediction of Physical Properties of Degradable Composites Using ANFIS", *International Journal of Materials, Mechanics and Manufacturing*, Vol. 3 No. 3, pp. 213–217, 2015.

[Szyp 17] Szyplowska A, Kafarski M, Wilczek A, Lewandowski A, Skierucha W. "Salinity index determination of porous materials using open-ended probes", *Measurement Science and Technology*, Vol. 28, pp. 014006–014012, 2017.

[Terh 07] Terhzaz J, Ammor H, Assir A, Mamouni A. "Application of the FDTD method to determine complex permittivity of dielectric materials at microwave frequencies using a rectangular waveguide". *Microwave and Optical Technology Letters*, Vol. 49, No. 8, 2007.

[Thol 95] Thollon F. "Conception et optimisation de capteurs à courants de Foucault pour la détection de défauts profonds dans des matériaux amagnétiques". Thèse de doctorat de l'école centrale de Lyon, janvier 1995.

[Vasi 08] Vasile O E. "Contribution au pronostique de défaillances par réseau neuro-flou: maîtrise de l'erreur de prédiction". Thèse de doctorat, université de FRANCHE-COMTÉ de Besançon, 2008.

[Wagn 07] Wagner N, Trinks E, Kupfer K. "Determination of the spatial TDR-sensor characteristics in strong dispersive subsoil using 3D-FEM frequency domain simulations in combination with microwave dielectric spectroscopy", *Measurement Science and Technology*, Vol. 18, pp. 1137–1146, 2007.

[Wagn 11] Wagner N, Emmerich K, Bonitz F, Kupfer K. "Experimental Investigations on the Frequency- and Temperature-Dependent Dielectric Material Properties of Soil", *IEEE Transactions on Geoscience and Remote Sensing*, Vol. 49 No. 7, pp. 2518–2530, 2011.

[Wagn 14] Wagner N, Schwing M, Scheuermann A. "Numerical 3-D FEM and Experimental Analysis of the Open-Ended Coaxial Line Technique for Microwave Dielectric Spectroscopy on

- Soil". *IEEE Transactions on Geoscience and Remote Sensing*, Vol. 52 No. 2, pp. 880–893, 2014.
- [Wang 04] Wang W Q, Goldnaraghi M F, Ismail F. "Prognosis of machine health condition using neuro-fuzzy systems". *Mechanical Systems and Signal Processing*. Vol. 18, pp. 813–83. 2004.
- [Wang 06] Wang Z, Palade V, Xu Y. "Neuro-fuzzy ensemble approach for microarray cancer gene expression data analysis". *International Symposium on Evolving Fuzzy Systems*; 2006 Sept; Ambleside, Cumbria UK.
- [Wang 15] Wang C, Hu J, Wen C. "A nonlinear PLS modeling method based on extreme learning machine". *Proceedings of the 34th Chinese Control Conference July 28-30, 2015, Hangzhou, China*.
- [Weia 07] Weia M, Baia B, Sungb A H, Liu Q, Wang J, Cather M E. "Predicting injection profiles using ANFIS". *Information Sciences*. Vol. 177, pp. 4445–4461, 2007.
- [Wold 85] Wold H. "in *Encyclopedia of Statistical Sciences*", edited by S. Kotz, N.L. Johnson (Wiley, New York, 1985), pp. 581–591.
- [Wold 89] Wold S, Wold N K, Skagerberg B. "Nonlinear PLS Modeling". *Chemometrics and Intelligent Laboratory Systems*, Vol. 7, pp. 53-65, 1989.
- [Wold 92] Wold S. "Nonlinear partial least squares modelling II. Spline inner relation". *Chemometrics and Intelligent Laboratory Systems*. Vol. 14, pp. 71-84, 1992.
- [Wold 01] Wold S, Sjostrom M, Eriksson L. "PLS-regression: a basic tool of chemometrics". *Chemometrics and Intelligent Laboratory Systems*. Vol. 58, pp. 109–130, 2001.
- [Yako 05] Yakovlev V V, Murphy E K, Eves E E, "Neural networks for FDTD-backed permittivity reconstruction". *COMPEL: The International Journal for Computation and Mathematics in Electrical and Electronic Engineering*. Vol. 24 No. 1, pp. 291-304, 2005.
- [Yoo 04] Yoo C K, Bang Y H, Lee I B, Vanrolleghem P A, Rosén C. "Application of Fuzzy Partial Least Squares (FPLS) Modeling Nonlinear Biological Processes". *Korean J. Chem. Eng.* Vol. 21, No. 6, pp. 1087–1097, 2004.
- [Zade 65] Zadeh L A. "Fuzzy Sets", *Information and Control*. Vol. 8, pp. 338–353. 1965.
- [Zamp 04] Zamprogna E, Barolo M, Seborg D E. "Estimating product composition profiles in batch distillation via partial least squares regression". *Control Engineering Practice*. Vol. 12, pp. 917–929, 2004.
- [Zhen 91] Zheng H, Smith C E. "Permittivity measurements using a short open-ended coaxial line probe", *IEEE Microwave and Guided Wave Letters*, Vol. 1 No. 11, pp. 337–339, 1991.

Résumé : Le principe de la caractérisation micro-ondes des matériaux diélectriques à l'aide d'une ligne coaxiale consiste à relier les propriétés diélectriques de l'échantillon sous teste aux mesures de l'admittance de la sonde. Le but de ce travail est de développer des outils d'inversion alternatives, capables de prédire l'évolution de la permittivité complexe sur une large bande de fréquence. Le problème direct est résolu en utilisant la méthode des éléments finis. Tandis que, le problème inverse est résolu en utilisant un système neuro-flou adaptatif, la régression non-linéaire multiple et la régression non-linéaire par les moindres carrés partiels, qui nécessitent la création d'une base de données pour leur apprentissage. L'inversion des mesures donne, après l'apprentissage, en temps réel la permittivité complexe des échantillons solides et liquides avec une très bonne précision.

Mots-clés : Caractérisation micro-ondes, Ligne coaxiale, Permittivité diélectrique complexe, Méthode des éléments finis, ANFIS, Procédure de sélection des entrées, MNLR, NLPLSR.

Abstract: The principle of microwave characterization of dielectric materials using a coaxial line is to link the dielectric properties of the sample under test to the measurements of the probe admittance. The purpose of this work is to develop alternative inversion tools, able to predict the evolution of the complex permittivity on a broad band frequency. The direct problem is solved using finite element method. Whereas, the inverse problem is solved using an adaptive neuro-fuzzy system, Multiple Non-Linear Regression and non-linear partial least square regression which need the creation of a database for their learning. Inversion of measurements gives, after training, in real time the complex permittivity of solid and liquid samples with a very good accuracy.

Keywords: Microwave characterization, coaxial line, Complex dielectric permittivity, Finite Element Method, ANFIS, Input selection procedure, MNLR, NLPLSR.

ملخص : مبدأ التوصيف بالميكروويف للمواد العازلة باستخدام خط متحد المحور هو ربط خصائص العزل للعينة تحت الاختبار إلى قياسات قابلية المجس. الغرض من هذا العمل هو تطوير أدوات عكس بديلة قادرة على التنبؤ بتطور السماحية المعقدة على تردد واسع النطاق. يتم حل المشكلة المباشرة باستخدام طريقة العناصر المنتهية. بينما يتم حل المشكلة العكسية باستخدام نظام استدلال عصبي ضبابي متكيف ، الانحدار غير الخطي المتعدد وانحدار المربعات الصغرى الجزئية والتي تحتاج إلى إنشاء قاعدة بيانات للتعلم. إن عكس القياسات يعطي ، بعد التدريب ، في الوقت الحقيقي السماحية المعقدة للعينات الصلبة والسائلة بدقة جيدة للغاية.

الكلمات المفتاحية : توصيف بالميكروويف ، خط متحد المحور ، السماحية العازلة المركبة ، طريقة العناصر المنتهية ، ANFIS ، إجراء اختيار المدخلات ، MNLR ، NLPLSR .