

الجمهورية الجزائرية الديمقراطية الشعبية

République Algérienne Démocratique Et Populaire

وزارة التعليم العالي و البحث العلمي

Ministère de l'enseignement supérieur et de la recherche scientifique

Université Badji Mokhtar, Annaba

Faculté des science de l'ingénieur

Département d'électronique



جامعة باجي مختار – عنابة

كلية علوم المهندس

قسم الإلكترونيك

MEMOIRE

Présenté en vue de l'obtention du diplôme de

MAGISTER

Thème :

Identification des éléments d'une centrale électrique thermique
Approche multi-modèle

Option : Automatique Industriel

Par : Bechinia Issam

Directeur de mémoire : Mr Lafifi Mohamed Mourad

M.C U-Annaba

Jury:

Président: Mr. Abbassi Hadj Ahmed

Pr U-Annaba

Examineurs : Mr. Kherfane Hamid

M.C U-Annaba

Mr. Kermiche Salah

M.C U-Annaba

Mr. Bensaoula Salah

M.C U-Annaba

Année 2010

Dédicace

Je dédie ce travail à :

*Mes chers parents qui m'ont
beaucoup aidé,*

À mes frères et sœurs.

A toute la famille et les amis

B. Issam

بِسْمِ اللَّهِ الرَّحْمَنِ الرَّحِيمِ

وَمَا تَوْفِيقِي إِلَّا بِاللَّهِ عَلَيْهِ

تَوَكَّلْتُ وَإِلَيْهِ أُنِيبُ

يقول عماد الدين الأصفهاني □ : * إني رأيت أنه لا يكتب أحد كتاباً في يومه إلا قال في غده لو غير هذا لكان حسن ، و لو زيد هذا لكان يستحسن ، و لو قدم هذا لكان أفضل ، و لو ترك هذا لكان أجمل ، و هذا من أعظم العبر ، و هو دليل على استيلاء النقص على جملة البشر *

¹محمد بن صفي الدين الملقب بعماد الدين الأصفهاني (1125 - 20 يونيو 1201م) (519 هـ - 13 رمضان 597 هـ) مؤرخ وأديب وشاعر توفى في

Remerciement

Le plus grand Merci tout d'abord revient à « Dieu » qui, lui seul, nous a guidé dans le bon sens durant notre vie et qui nous a donné le courage, la volonté, et la force pour élaborer ce travail de recherche.

*Je remercie chaleureusement Messieurs **Lafifi Mohamed Mourad**, Mc à l'université de Badji Mokhtar, Annaba directeurs de la mémoire pour m'avoir accueilli et m'avoir encadré e et encouragé e tout au long de ces travaux.*

Par leur grandes compétences scientifique et leurs remarquables qualités humaines, ils ont su rendre chaleureuses et fructueuses ces quatre années de recherche.

*J'exprime mes profonds remerciements et tous mes respects à Monsieur **A. Hadj Ahmed** Pr au département d'électronique, pour avoir bien voulu accepter de présider ce jury.*

*Je remercie M^r **H. Kherfane**, M^r **S.Kermiche**, M^r **S.Bensaoula** M.C au département d'électronique. D'avoir accepté d'examiner ce mémoire et d'être membre du jury.*

*Je remercie M^r **A. Hichem**, MC à l'Université d'Evry Val d'Essonne, Genopole, Mr **J.RAGOT** Pr à l'Institut National Polytechnique de Lorraine, Mr **Rodolfo Orjuela** Pr à l'université de Nancy, Mr **Silvio Simani** Mc à l'université de Ferrara. Qui m'ont aidé de près ou de loin à la réalisation de ce modeste travail.*

Résumé

A notre époque, et sans électricité, la vie quotidienne serait difficilement envisageable. Il est donc nécessaire de savoir la produire de manière efficace et continue.

Pour répondre à la consommation croissante d'électricité, il a fallu inventer et construire des usines capables de produire de l'électricité en grande quantité. La turbine et l'alternateur sont les deux pièces maîtresses de ces générateurs d'électricité.

Pour bien exploiter et utiliser ces éléments d'une manière efficace et fiable, les ingénieurs font largement appel aux modèles non linéaires pour décrire leurs comportements dynamiques.

Si les modèles non linéaires sont en mesure de décrire correctement les comportements non linéaires d'un système, ils peuvent néanmoins s'avérer, en fonction de leur complexité mathématique, difficilement exploitables dans un contexte de synthèse d'une loi de commande et/ou de mise en place d'une stratégie de diagnostic du système.

Nous proposons, dans cet travail, une procédure d'identification hors ligne des systèmes non linéaires (quelques éléments d'une centrale électrique thermique) par une approche multi modèle à deux niveaux, qui alterne entre les paramètres β des fonctions de validité par une procédure non linéaire, θ étant fixé et l'optimisation des paramètres θ des modèles locaux par l'estimateur des moindres carrés éventuellement régularisés, β étant fixé .

Les multi modèles constituent une alternative intéressante et un outil puissant dans la modélisation des systèmes non linéaires. L'approche multi modèle est basée sur la décomposition du comportement dynamique du système en un nombre L de zones de fonctionnement, chaque zone étant caractérisée par un sous-modèle. En fonction de la zone où le système évolue, la sortie de chaque sous-modèle contribue plus ou moins à l'approximation du comportement global du système non linéaire. La contribution de chaque sous-modèle est définie par une fonction de pondération.

Abstract

At our time, and without electricity, the daily life would be not easily possible. It is thus necessary to know to produce it efficiently and continuous.

To answer the increasing consumption of electricity, it was necessary to invent and build factories able to produce electricity in great quantity. The turbine and the alternator are the two showpieces of these generators of electricity.

For exploiting well and using these elements in an effective and reliable way, the engineers largely call upon the nonlinear models to describe their dynamic behaviors.

If the nonlinear models are able to correctly describe the nonlinear behaviors of a system, they can nevertheless prove, according to their mathematical complexity, not easily exploitable in a context of synthesis of a law of order and/or installation of a strategy of diagnosis of the system.

We propose, in this work, a procedure of Identification off line of the nonlinear systems (some elements of a thermal power plant) by an approach on two levels, which alternates between the parameters β of the functions of validity by a nonlinear procedure, θ being fixed and the optimization of the parameters θ possibly regularized local models by the estimator of least square, β being fixed.

The model multi constitute an interesting alternative and a powerful tool in the modeling of the nonlinear systems. The approach multi model is based on the decomposition of the dynamic behavior of the system in a number L of zones of operation, each zone being characterized by a sub-model. According to the zone where the system evolves, the exit of each sub-model contributes more or less to the approximation of the total behavior of the nonlinear system. The contribution of each sub-model is defined by a weight function.

The analysis of the results in simulation is encouraging

ملخص:

إن العيش في هذه الأيام دون كهرباء، يعتبر من الأمور التي يصعب على الإنسان أن يتخيلها. لأجل ذلك يحاول جاهدا أن يوفرها بعدة طرق و ذلك من أجل ضمان استمراريتها و فعاليتها فيقوم بإنشاء المحطات الكهربائية و إن اختلفت طرق عملها إلا أنها تشترك في الغالب في عنصرين أساسيين و هما التربين و المولد و اللذان يعملان على إنتاج هذه الطاقة بشكل مستمر و منتظم. من أجل عمل هذين العنصرين بكفاءة و فعالية يحاول المهندسون دائما الحصول على أفضل محاكات لعملهما باستعمال المعادلات الرياضية و التي تكون في الغالب غير خطية.

لكن يرجح في هذه المعادلات الغير خطية أنها تحتوي على درجة من الصعوبة و التعقيد، مما يصعب استغلالها على أحسن وجه، كما يصعب وضع خطط التحكم و مراقبة الأخطاء و التوقفات المفاجئية التي يمكن أن تحدث أثناء العطب.

نقترح في هذا العمل طريقة بحث عن أفضل تمثيل رياضي من خلال معادلات رياضية بسيطة تمثل هذه الأجزاء التي تكون المحطات الكهربائية، حيث تتلخص في: بحث من خلال مرحلتين:

- في المرحلة الأولى نقوم بالبحث عن معلّات الدالة بيتا بطريقة غير خطية و تكون معلّات الدالة بيتا ثابتة

- في المرحلة الثانية نقوم بالبحث عن معلّات الدالة بيتا بطريقة (المربع الأقل) بحيث تكون معلّات الدالة بيتا ثابتة.

إن تمثيل الأنظمة الحقيقية من خلال نماذج متعددة، تمثلها في مراحل معينة من مراحل عملها يعتبر من الطرق الناجعة، و التي تسهل و تيسر بناء استراتيجيات التحكم و تحديد الأعطاب و الخلل الذي يمكن أن يصيب النظام بصفة مفاجئة.

النتائج المتحصل عليها تعتبر مشجعة و تفتح آفاقا أخرى للتطوير.

Introduction générale

Un système est un ensemble d'objets ou de phénomènes liés entre eux et isolés artificiellement du monde extérieur. La modélisation regroupe un ensemble de techniques permettant de disposer d'une représentation mathématique du système à étudier. La modélisation théorique requiert une connaissance précise des phénomènes intervenant dans le système et une aptitude à les représenter par des équations mathématiques. Les systèmes physiques sont souvent complexes et difficilement exploitables, notamment pour la commande et le diagnostic des systèmes.

Le grand problème de l'automaticien est de représenter ces systèmes physiques avec une précision suffisante et un modèle de structure simple. Le dilemme réside alors entre la fidélité du modèle vis-à-vis du processus réel et l'adéquation de ce modèle à une forme mathématiquement exploitable.

En automatique, pour décrire le comportement d'un système, une hypothèse communément faite est la linéarité du système, car les techniques d'analyse des modèles linéaires, dits LTI, ont été largement développées dans la littérature. Cependant, l'hypothèse de linéarité n'est vérifiée que dans une plage de fonctionnement restreinte autour d'un point d'équilibre du système. Alors, les performances du modèle se dégradent dès qu'on s'en éloigne et la recherche d'un modèle plus adapté et notamment non linéaire devient nécessaire.

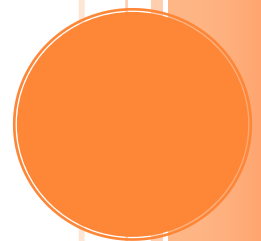
La structure mathématique qui puisse remédier à l'inconvénient cité ci-dessus, tout en gardant la simplicité mathématique des modèles LTI, est l'approche globale, dite multi-modèle ; c'est une

représentation poly topique convexe pouvant être obtenue soit directement à partir d'un modèle mathématique non linéaire par transformation directe d'un modèle affine [Morè 00] ou par linéarisation autour de différents points de fonctionnement [Murr 97], soit à partir de données sur les entrées et les sorties d'un système physique [Gass 00b].

Dans ce mémoire nous proposons, une procédure d'identification hors ligne des systèmes non linéaires (quelques éléments d'une centrale électrique thermique) par une approche multi modèle à deux niveaux, qui alterne entre les paramètres β des fonctions de validité par une procédure non linéaire, θ étant fixé et l'optimisation des paramètres θ des modèles locaux par l'estimateur des moindres carrés éventuellement régularisés, β étant fixé .

Chapitre I

La production d'électricité



Chapitre I

La production de l'électricité



Examine each question in terms of what is ethically and aesthetically right, as well as what is economically expedient. A thing is right when it tends to preserve the integrity, stability, and beauty of the biotic community. It is wrong when it tends otherwise.

-Aldo Leopold-

I-1-Introduction :

A notre époque, et sans électricité, la vie quotidienne serait difficilement envisageable. Il est donc nécessaire de savoir la produire de manière efficace et continue.

Pour répondre à la consommation croissante d'électricité, il a fallu inventer et construire des usines capables de produire de l'électricité en grande quantité. La turbine et l'alternateur sont les deux pièces maîtresses de ces générateurs d'électricité.

Dans le cas des usines thermiques, la turbine est entraînée par la vapeur produite dans les chaudières où l'on brûle les combustibles, alors que dans le cas des usines hydroélectriques, la turbine est animée par la force de l'eau. La turbine est couplée à un alternateur, un grand aimant cerclé d'une bobine, qui va produire un courant alternatif en tournant. Une fois le courant produit, il doit être amené jusque chez le consommateur.

A la sortie de la centrale, un premier transformateur, un survolteur, augmente la tension du courant à 400 ou 800000 V. Ceci permet de minimiser les pertes d'énergie pendant le transport. Près du point de livraison, un deuxième transformateur, un sousvolteur, fait l'opération inverse : il abaisse la tension du courant pour la mettre aux normes du réseau domestique. Il existe d'autres manières efficaces de produire de l'électricité : les panneaux solaires transforment la lumière du soleil en électricité et les éoliennes utilisent la force du vent. Il faut savoir qu'il existe également des usines marémotrices qui utilisent la force des marées, que la géothermie exploite les gisements d'eau

chaude stockés dans le sous-sol terrestre, tandis que les usines à biomasse utilisent les déchets comme source d'énergie.

I-2-DE L'ALTERNATEUR DE VELO VERS LES CENTRALE :

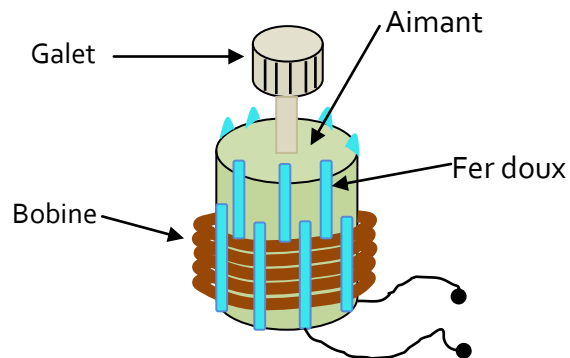


Figure I- 1 : Schéma d'une dynamo de vélo

I-2-1-Expérience N°01 :

On approche un aimant droit d'une petite aiguille aimantée mobile que l'on appelle habituellement une boussole.



Figure I- 2 : Aimants

I-2-1-1-OBSERVATIONS :

Lorsque l'aimant est approché de l'aiguille aimantée, celle-ci est déviée par rapport à sa position de repos.

I-2-1-2-CONCLUSION :

La perturbation de l'aiguille aimantée est due à la présence de l'aimant qui crée un champ magnétique.

On utilise l'aiguille aimantée pour détecter la présence d'un champ magnétique.

I-2-2-Expérience N°02 :

On déplace de diverses manières un aimant droit au voisinage d'une bobine reliée à un ampèremètre à cadran et à zéro central.

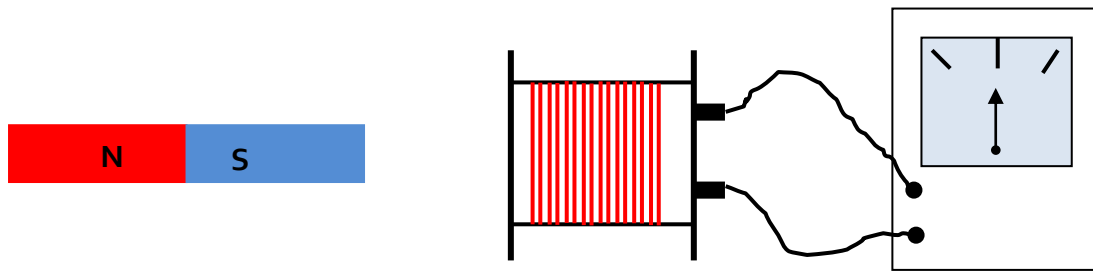


Figure I- 3 : Expérience N°02

I-2-2-1-Observations :

- Si l'aimant est immobile par rapport à la bobine, l'aiguille de l'ampèremètre ne bouge pas.
- Si l'on approche l'aimant de la bobine, l'aiguille de l'ampèremètre dévie donc un courant apparaît dans la bobine.
- Si l'on éloigne l'aimant de la bobine, l'aiguille de l'ampèremètre dévie dans l'autre sens donc le courant circulant dans la bobine change de sens.
- Plus le déplacement est effectué de manière rapide, plus la déviation de l'aiguille est grande donc plus le courant circulant dans la bobine est intense.
- Si l'aimant effectue un mouvement de va-et-vient, il circule un courant alternatif dans la bobine.

I-2-2-2-Conclusion :

Le déplacement d'un aimant au voisinage d'une bobine fait apparaître un courant électrique qui circule dans celle-ci : c'est le phénomène d'induction électromagnétique.

I-2-3-Expérience N°03 :

A l'aide d'un moteur ou d'une perceuse, on fait tourner un aimant droit devant une bobine. On visualise la tension électrique aux bornes de la bobine grâce à un oscilloscope.

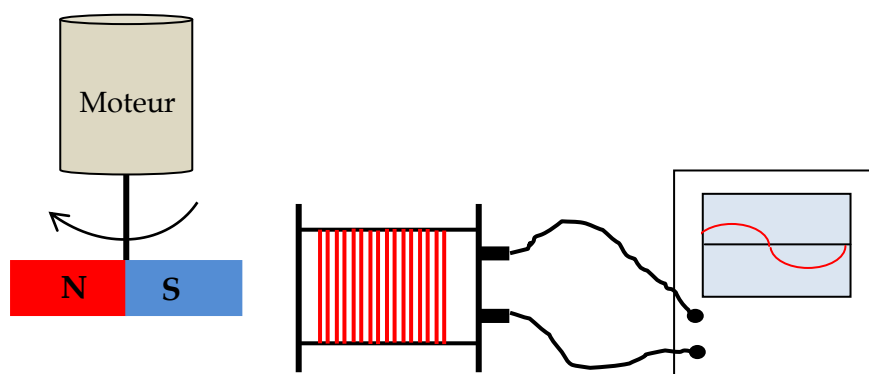


Figure I- 4 : Expérience N°03

I-2-3-1-Observations :

Quand l'aimant tourne, il apparaît une tension alternative aux bornes de la bobine (Figure I-4). Au cours de cette expérience nous avons reproduit ce qui se passe dans l'alternateur vélo.

I-2-3-2-Conclusion :

Un alternateur est un convertisseur d'énergie mécanique en énergie électrique.

I-3-LA PRODUCTION DE L'ELECTRICITE :

La production d'électricité est tout simplement une conversion, une transformation d'énergie mécanique (liée au mouvement) en énergie électrique.

La plupart du temps l'électricité est produite à partir d'une source de chaleur, en utilisant la vapeur d'eau comme colporteur d'énergie. La vapeur fait tourner des turbines qui sont couplées à des générateurs électriques. La vapeur peut être produite en utilisant la plupart des sources d'énergie. Les énergies hydrauliques et éoliennes étant des exceptions puisque c'est l'énergie de l'eau et du vent en déplacement qui produit un travail directement dans une turbine couplée à un générateur.

Les centrales nucléaires utilisent souvent un circuit primaire et secondaire de vapeur, afin d'isoler physiquement le réacteur nucléaire de la salle des générateurs et du reste des installations.

De petites installations (brûlant habituellement du gaz naturel) combinent la génération d'électricité et de chaleur (pour le chauffage domestique ou pour des processus industriels). Ces centrales électriques combinées ont le meilleur rendement, après les centrales hydroélectriques. Cette technique porte le nom de cogénération.

Des expériences sont en cours pour utiliser la géothermie pour produire de l'électricité en creusant à très grande profondeur dans des roches dures, ce qui permettent de réchauffer un fluide caloporteur alimentant en vapeur une turbine (via une pompe à chaleur quand la température est trop faible).

Tous les véhicules automobiles non électriques utilisent un petit alternateur couplé mécaniquement au moteur principal pour une génération locale d'électricité basse tension, une batterie d'accumulateur le remplace pendant l'arrêt du moteur principal.

Des unités d'appoint ou de secours, appelées groupes électrogènes permettent une fabrication d'électricité ponctuelle, ils utilisent tous un moteur à explosion pour entraîner la génératrice.

Il existe des groupes transportables pour une utilisation des outils électriques hors des lieux électrifiés.

De gros générateurs sont utilisés pour pallier une rupture de fourniture toujours possible du fournisseur d'électricité. Les hôpitaux, certains services publics et, les grandes entreprises ne pouvant supporté un arrêt brutal de leurs processus industriels possèdent des groupes électrogènes à démarrage automatique.

I-4-LES CENTRALES THERMIQUE :

Une centrale thermique est une centrale électrique qui produit de l'électricité à partir d'une source de chaleur. Cette source peut être un combustible brûlé (tels gaz naturel, certaines huiles minérales, charbon, déchets industriels, déchets ménagers), la fission de noyaux d'uranium 235 ou de plutonium 239.

La source de chaleur chauffe un fluide (souvent de l'eau) qui passe de l'état liquide à l'état gazeux (vapeur par exemple), gaz qui entraîne une turbine accouplée à un alternateur qui transforme l'énergie cinétique de la turbine en énergie électrique.

Pour entraîner la turbine, il faut que la pression soit plus faible à sa sortie qu'à l'entrée. La baisse de pression à la sortie de la turbine est obtenue en condensant le gaz, en aval de la turbine, à l'aide d'une source froide.

Le fluide condensé est en général réutilisé comme source de vapeur et effectue un cycle thermodynamique fermé.

La cogénération consiste à produire conjointement de la chaleur pour un procédé industriel ou le chauffage urbain, afin d'améliorer le rendement global.

Les centrales thermiques se répartissent en trois grandes catégories, selon la nature de leur source de chaleur :

- ❖ Centrales nucléaires (**figure I-5**)
- ❖ Centrales à flamme (charbon, fioul ou gaz), (**figure I-6**)
- ❖ Centrales récupérant de la chaleur préexistante (solaire, géothermique...). (**figure I-7, I-8, I-9, I-10, I-11**)

I-4-1-LES CENTRALES NUCLEAIRES :

Une centrale nucléaire est un site industriel utilisant la fission de noyaux atomiques pour produire de la chaleur, dont une partie est transformée en électricité (entre 30 % et 40 % en fonction de la différence de température entre la source froide et chaude). C'est la principale mise en œuvre de l'énergie nucléaire dans le domaine civil.

Une centrale nucléaire est constituée d'un ou plusieurs réacteurs nucléaires dont la puissance électrique varie de quelques mégawatts à plus de 1 500 mégawatts pour le réacteur soviétique de grande puissance RBMK. Selon les promoteurs du futur réacteur européen EPR, il devrait atteindre une puissance record de 1 600 mégawatts.

En 2006, 442 réacteurs fonctionnent dans 31 pays différents dans le monde, soit un total de 370 gigawatts produisant environ 17 % de l'électricité mondiale. La catastrophe de Tchernobyl a conduit à plusieurs moratoires ; la baisse des prix du pétrole durant les années 1990 a renforcé cette tendance, conduisant à construire moins de nouveaux réacteurs dans le monde. Parallèlement, les centrales vieillissent : en 2006, la majorité des réacteurs avaient de 15 à 36 ans, sept ayant de 37 à 40 ans.

I-4-1-1-Description :

Une centrale nucléaire regroupe l'ensemble des installations permettant la production d'électricité sur un site donné. Elle comprend fréquemment plusieurs tranches, identiques ou non ; chaque tranche correspond à un groupe d'installations conçues pour fournir une puissance électrique donnée (par exemple 900 MWe, 1 300 MWe ou 1 450 MWe). Une tranche comprend généralement :

- ❖ le bâtiment réacteur, généralement double enceinte étanche qui contient le réacteur nucléaire, les générateurs de vapeur (trois ou quatre selon la génération), un pressuriser qui a pour fonction de maintenir l'eau du circuit primaire à l'état liquide, le groupe motopompe primaire servant à faire circuler le fluide caloporteur (eau) et une partie du circuit d'eau *secondaire* et le circuit d'eau *primaire*, dont le rôle principal est d'assurer le transfert thermique entre le cœur du réacteur et les générateurs de vapeur ;
- ❖ le bâtiment salle des machines, qui contient principalement :
 - une ligne d'arbre comprenant les différents étages de la turbine à vapeur et l'alternateur,
 - le condenseur,

- ❖ des bâtiments annexes qui contiennent notamment des installations diverses de circuits auxiliaires nécessaires au fonctionnement du réacteur nucléaire et à la maintenance, les tableaux électriques alimentant tous les auxiliaires et générateurs diesel de secours ;
- ❖ un aéroréfrigérant atmosphérique (la partie la plus visible d'une centrale nucléaire), ou simplement une station de pompage pour les tranches dont le refroidissement utilise l'eau de mer ou de rivière.

Les autres installations de la centrale électrique comprennent :

- ❖ un ou plusieurs postes électriques permettant la connexion au réseau électrique par l'intermédiaire d'une ou plusieurs lignes à haute tension, ainsi qu'une interconnexion limitée entre tranches ;
- ❖ un bâtiment administratif...

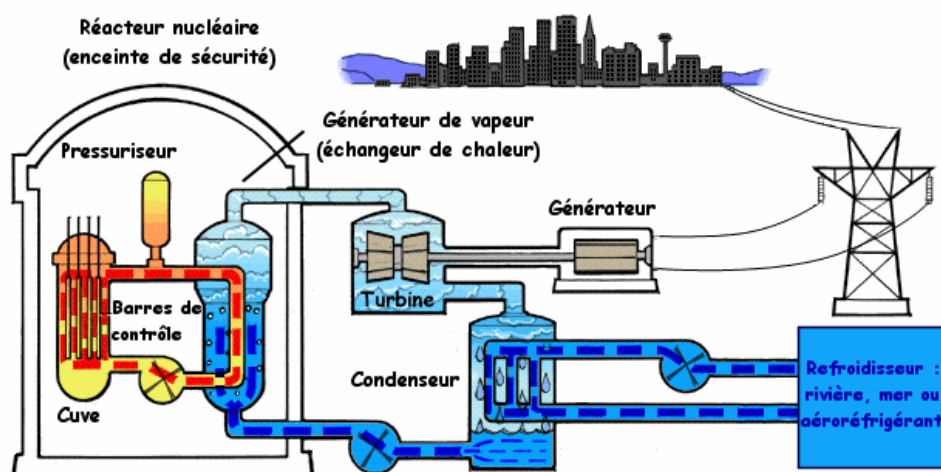


Figure I- 5 : Centrale nucléaire

I-4-1-2-FONCTIONNEMENT TECHNIQUE :

Dans une tranche nucléaire, le réacteur nucléaire est en amont d'une installation thermique qui produit de la vapeur transformée en énergie mécanique au moyen d'une turbine à vapeur ; l'alternateur utilise ensuite cette énergie mécanique pour produire de l'électricité.

La différence essentielle entre une centrale nucléaire et une centrale thermique classique est matérialisée par le remplacement d'un ensemble de chaudières consommant des combustibles fossiles par un réacteur nucléaire.

Pour récupérer de l'énergie mécanique à partir de chaleur, il est nécessaire de disposer d'une source chaude et d'une source froide.

- ❖ pour un réacteur de type REP (Réacteur à Eau sous Pression), la source chaude est fournie par l'eau du circuit primaire, à la température moyenne de 306 °C (286 °C en entrée et 323 °C en sortie, température de sortie variant selon la puissance de la tranche) ;
- ❖ la source froide est fournie par l'eau d'un fleuve ou de la mer, ou bien de l'air ambiant par évaporation dans des tours aéroréfrigérant.

Ainsi, une tranche nucléaire de type REP comporte trois circuits d'eau importants indépendants :

- ❖ le circuit primaire qui se situe dans une enceinte de confinement.

Il est constitué, suivant le type de tranche, de 3 ou 4 générateurs de vapeur associés respectivement à une pompe (par GV), un pressuriser assurant le maintien de la pression du circuit (155 bar) puis d'un réacteur intégrant des grappes de contrôle et le combustible. Le véhicule, en circuit fermé, de l'eau liquide qui extrait les calories du combustible pour les transporter aux générateurs de vapeur (rôle de caloporteur). L'eau du circuit primaire a aussi comme utilité la modération des neutrons (rôle de modérateur) issus de la fission nucléaire. La thermalisation des neutrons les ralentit pour interagir avec les atomes d'uranium 235 et déclencher la fission de leur noyau. Par ailleurs, l'eau procure un effet stabilisateur au réacteur : si la réaction s'emballait, la température du combustible et de l'eau augmenteraient. Cela provoquerait d'une part, une absorption des neutrons par le combustible (effet combustible) et d'autre part une modération moindre de l'eau (effet modérateur). Le cumul de ces deux effets est dit "effet puissance" : l'augmentation de ce terme provoque l'étouffement de la réaction d'elle-même, c'est un effet auto-stabilisant.

- ❖ le circuit d'eau secondaire est un circuit fermé, qui se décompose en deux parties :

- entre le condenseur et les générateurs de vapeur, l'eau reste sous forme liquide : c'est l'alimentation des générateurs de vapeur ; des pompes permettent d'élever la pression de cette eau, et des échangeurs de chaleur en élèvent la température (60 bar et 220 °C).
- cette eau se vaporise dans 3 ou 4 générateurs de vapeur (suivant le type de tranche ; 900 ou 1 300 / 1 450 MW) et les tuyauteries de vapeur alimentent successivement les étages de la turbine disposés sur une même ligne d'arbre. La vapeur acquiert une grande vitesse lors de sa détente permettant ainsi d'entraîner les roues à aubages de la turbine.

Celle-ci est composée de plusieurs étages séparés et comportant chacun de nombreuses roues de diamètre différent. D'abord, la vapeur subit une première détente dans un corps haute pression (HP ; de 55 à 11 bar), puis elle est récupérée, séchée et surchauffée pour subir une seconde détente dans les corps basse pression, (BP ; de 11 à 0,05 bar). On utilise les corps BP dans le but d'augmenter le rendement du cycle thermo-hydraulique. La sortie du dernier étage de la turbine donne directement sur le condenseur, un échangeur de chaleur dont la pression est maintenue aux environs de 50 mbar absolu (vide) par la température de l'eau du circuit de refroidissement (selon la courbe de saturation de l'eau). Des pompes à vide extraient les gaz incondensables en phase gaz du mélange (principalement l'oxygène et l'azote). L'eau condensée dans cet appareil est réutilisée pour réalimenter des générateurs de vapeur.

- ❖ Le circuit de refroidissement assure le refroidissement du *condenseur*. L'eau est refroidie par un courant d'air dans une tour aéroréfrigérante d'où une petite partie (1,5 %) de l'eau s'échappe en vapeur en forme de panache blanc. L'eau de refroidissement peut aussi être échangée directement avec un fleuve ou la mer.

L'énergie mécanique produite par la turbine sert à entraîner l'alternateur qui la convertit en énergie électrique, celle-ci étant évacuée par le réseau électrique.

Lorsque la tranche nucléaire débite de la puissance électrique sur le réseau, on dit qu'elle est "couplée" au réseau. La déconnexion intempestive de l'alternateur au réseau (ce qu'on appelle un "déclenchement"), nécessite une réduction immédiate de l'alimentation en vapeur de la turbine par des vannes de réglage disposées sur les tuyauteries de vapeur, faute de quoi sa vitesse de rotation

augmenterait jusqu'à sa destruction en raison de la force centrifuge excessive s'exerçant alors sur les aubages. Néanmoins, dans ce cas-ci, la tranche reste en service à faible puissance : la turbine est en rotation et reste prête au recouplage immédiat sur le réseau (la tranche est alors "ilotée" : elle alimente elle-même ses auxiliaires).

I-4-1-3-RENDEMENT D'UNE CENTRALE NUCLEAIRE :

Le rendement d'une centrale nucléaire est de l'ordre de 33% - auquel il faut ajouter les pertes en ligne sur le réseau Très Haute Tension. Les centrales électriques alimentées au fioul ou au charbon possèdent un rendement un peu supérieur (~40%) car elles fonctionnent avec une température de vapeur plus élevée (moins de contraintes de sécurité).

Un réacteur nucléaire électrogène ne peut être utilisé pour faire de la cogénération. Ceci reviendrait à augmenter la température de la source froide et donc diminuer la différence de température entre les sources ayant pour conséquence une baisse du rendement de production d'électricité. Dans une centrale thermique à cogénération, ce sont les gaz d'échappement qui sont utilisés pour produire de la vapeur qui sert au chauffage urbain. Par contre les eaux chaudes rejetées peuvent être captées et réutilisées, une ferme de crocodiles fonctionne de cette façon en aval de la centrale nucléaire de Cruas, en Ardèche.

I-4-2-LES CENTRALES A FLAMME:

Une centrale thermique à flamme utilise l'énergie fournie par la combustion d'un combustible (charbon, pétrole, gaz naturel, gaz issus de hauts-fourneaux). Cette combustion a lieu dans une chaudière.

La combustion dégage une grande quantité de chaleur utilisée pour chauffer de l'eau dans la chaudière (ou générateur de vapeur). On dispose alors de vapeur d'eau sous pression.

Cette vapeur sous pression fait tourner à grande vitesse une turbine qui entraîne elle-même un alternateur qui produit une tension alternative sinusoïdale. A la sortie de la turbine la vapeur est refroidie pour se transformer en eau, puis renvoyée dans la chaudière.

Le refroidissement de la vapeur issue de la turbine est confié à une réserve d'eau (cours d'eau) ou plus rarement à une tour de refroidissement analogue à celle d'une centrale nucléaire.

Une centrale thermique à flamme fournit une puissance électrique de l'ordre de quelques centaines de mégawatts (1 MW = 1 000 000 W).

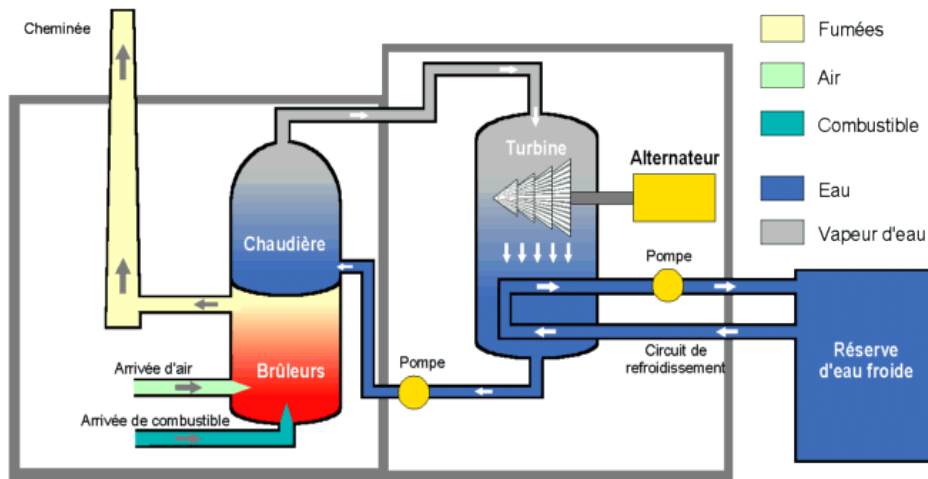


Figure I- 6: Centrale à flamme

I-4-3-CENTRALES RECUPERANT DE LA CHALEUR PREEXISTANTE (SOLAIRE, GEOTHERMIQUE...):

I-4-3-1-CENTRALE SOLAIRE THERMODYNAMIQUE :

Le soleil envoie chaque année 1.080.000.000 térawatt heures d'énergie sur la terre l'équivalent de 60.000 fois la consommation mondiale annuelle d'électricité. L'énergie solaire possède ainsi le plus grand potentiel parmi toutes les énergies renouvelables.

Une centrale solaire thermique (ou centrale solaire thermodynamique¹ ou encore heliothermodynamique) est une centrale qui concentre les rayons du soleil à l'aide de miroirs afin de chauffer un fluide caloporteur qui permet en général de produire de l'électricité. C'est par exemple la centrale solaire de Vignola (Corse-du-Sud) et ce sera le cas de la Centrale solaire Thémis (Pyrénées-Orientales) qui après avoir été reprise en 2006 pour de nouveaux projets qui seront opérationnels en 2009-2010.

➤ TYPE :

On distingue :

- ❖ centrale à tour.
- ❖ centrale constituée de capteurs cylindro paraboliques.
- ❖ centrale constituée de capteurs paraboliques.

¹ La thermodynamique, qui étudie les échanges de chaleur, démontre qu'une source chaude et une source froide sont absolument nécessaires pour échanger un travail avec une machine thermique.



Figure I- 7: Centrale Solaire de THERMIS



Figure I- 8: Centrale constituée de capteurs cylindro paraboliques



Figure I- 9: Centrale solaire parabolique (projet DISH-STIRLING à Font-Romeu-Odeillo)



Figure I- 10: Centrale solaire à tour (Solar Tower en Californie)

I-4-3-2-CENTRALE GEOTHERMIQUE:

La géothermie, du grec Géo (la terre) et thermie (la chaleur), est la science qui étudie les phénomènes thermiques internes du globe terrestre et la technique qui vise à l'exploiter. Par extension, la géothermie désigne aussi l'énergie géothermique issue de l'énergie de la Terre qui est convertie en chaleur.

On distingue trois types de géothermie :

- ❖ la géothermie peu profonde à basse température ;
- ❖ la géothermie profonde à haute température ;
- ❖ la géothermie très profonde à très haute température.

Ces trois types de géothermie prélèvent la chaleur contenue dans le sol.

L'énergie géothermique est exploitée dans des réseaux de chauffage et d'eau chaude depuis des milliers d'années en Chine, dans la Rome antique et dans le bassin méditerranéen.

L'augmentation des prix de l'énergie et le besoin d'émettre moins de gaz à effet de serre la rendent plus attrayante.



Figure I- 11: Centrale géothermique de Nesjavellir en Islande

❖ INSTALLATIONS DANS LE MONDE:

Région du monde	MWe
Asie	3 220
Amérique du Nord	2 971
Union européenne	883
Océanie	441
Amériques centrale et du Sud	416
Autres pays d'Europe	297
Afrique	128
Total mondial	8 536

Tableau I- 1: Capacité géothermique installée (2002)

I-5-LES TURBINES :

I-5-1-GENERALITES :

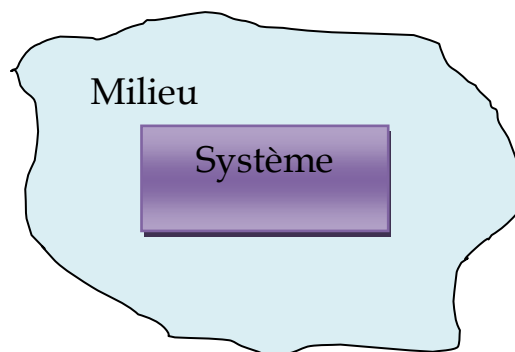
La thermodynamique est la science qui étudie et décrit le comportement de la matière ou des systèmes, en fonction des notions de température T , d'énergie (chaleur Q , travail W ...) et d'entropie S . La thermodynamique :

- ❖ étudie l'évolution ou les transformations de la matière ou des systèmes en considérant les variations d'état du système, lors d'échanges d'énergie entre le milieu extérieur et le système.
- ❖ repose sur 2 notions de base, l'énergie interne (U) et l'entropie (S) qui satisfont aux deux principes suivants, qui stipulent que :
 - l'énergie se conserve (premier principe de conservation de l'énergie).
 - l'entropie ne peut qu'augmenter (deuxième principe d'évolution).

L'objet de la thermodynamique est d'étudier le fonctionnement et le bilan d'énergie des machines thermiques et aussi les échanges ou transferts de chaleur dans un système ou entre deux systèmes.

- ❖ dans les machines thermiques on assiste à une conversion d'énergie d'une forme en une autre (chaleur -> travail ou inversement).
- ❖ dans les échanges de chaleur, il y a transfert de chaleur par suite d'une différence de température dans le système ou entre deux systèmes.

Le système est défini comme une partie de matière (de masse donnée) délimitée par rapport au milieu extérieur. Le milieu extérieur est le reste de l'espace entourant le système.



Le système peut être **ouvert** ou **fermé**, **adiabatique** c.à.d. isolé thermiquement ($Q = 0$) ou **rigide** c.à.d. indéformable ($W = 0$).

Système	Echange Matière	Echange Energie
Isolé	non	non
Fermé	non	oui
Ouvert	oui	oui

Tableau I- 2: Echange de masse et d'énergie entre le système et le milieu extérieur

Une turbine est un dispositif rotatif destiné à utiliser la force d'un fluide (eau, vapeur, air, gaz de combustion), dont le couple est transmis au moyen d'un arbre.

L'énergie du fluide, caractérisée par sa vitesse et son enthalpie, est partiellement convertie en énergie mécanique pour entraîner un alternateur, une pompe ou tout autre récepteur mécanique rotatif.

Schéma de principe d'une turbine

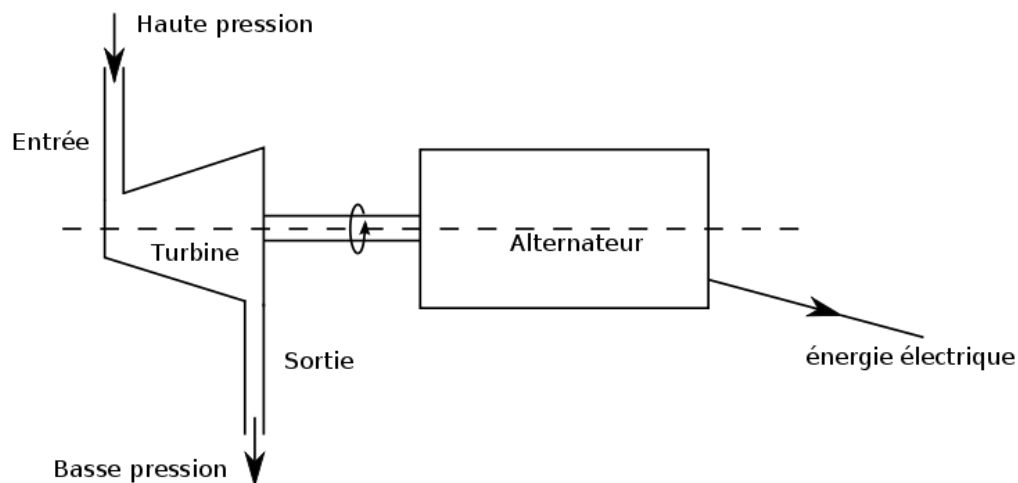


Figure I- 12:Schéma de principe d'une turbine

I-5-2-LES TURBINE A VAPEUR :

La turbine à vapeur (**figure I-12**) est un moteur thermique à combustion externe, fonctionnant selon le cycle thermodynamique dit de Clausius-Rankine (**figure I-14**). Ce cycle se distingue par le changement d'état affectant le fluide moteur qui est en général de la vapeur d'eau.

Ce cycle comprend au moins les étapes suivantes :

- ❖ L'eau liquide est mise en pression par une pompe et envoyée vers la chaudière,
- ❖ L'eau est chauffée, vaporisée et surchauffée,
- ❖ La vapeur se détend et refroidit dans la turbine en fournissant de l'énergie mécanique,
- ❖ La vapeur détendue est condensée au contact de la source froide sous vide partiel.

Le principe est donc le même que celui de la machine à vapeur à pistons. La turbine en constitue une évolution exploitant les principaux avantages des turbomachines à savoir :

- ❖ Puissance massique et puissance volumique élevée,
- ❖ Rendement amélioré par la multiplication des étages de détente.

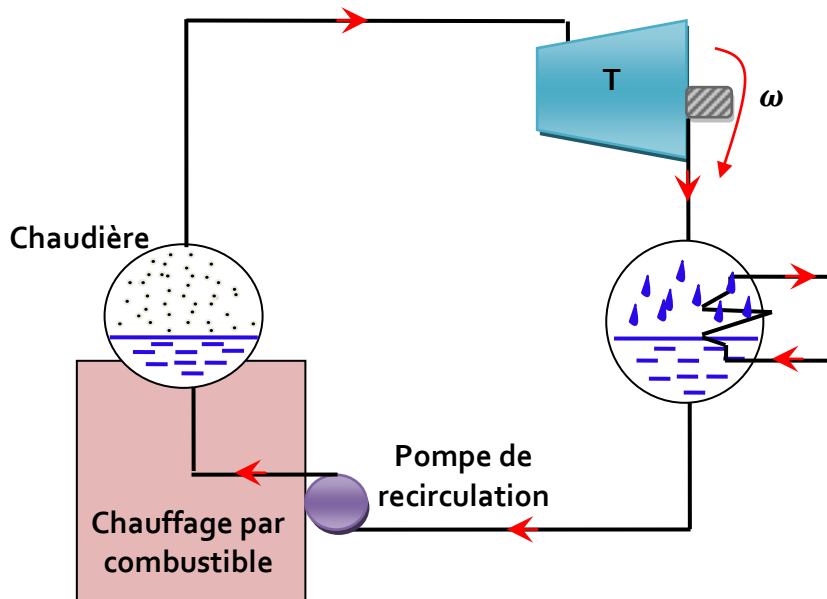


Figure I- 14: Cycle de Rankine

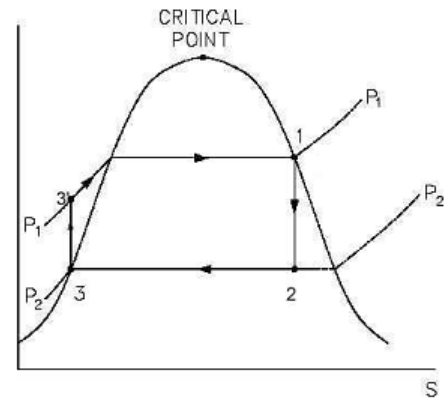


Figure I- 13: Diagramme (T, S) du cycle

Le point 1 est sur la courbe de saturation. Le cycle est effectué entre deux isobares HP et BP :

1-2	détente adiabatique (réversible ou non), de HP à BP, dans la turbine
2-3	condensation isobare et isotherme du mélange jusqu'à eau liquide
3-3'	compression de l'eau dans la pompe, passage de BP à HP
3'-1	chauffage isobare et évaporation isobare et isotherme dans le générateur de vapeur (G.V.)

Hypothèses usuelles : les points 3 et 3' sont confondus (mais à des pressions différentes).

I-5-2-1-LE CYCLE DE HIRN :

Le point 1 n'est plus sur la courbe de saturation → vapeur surchauffé.

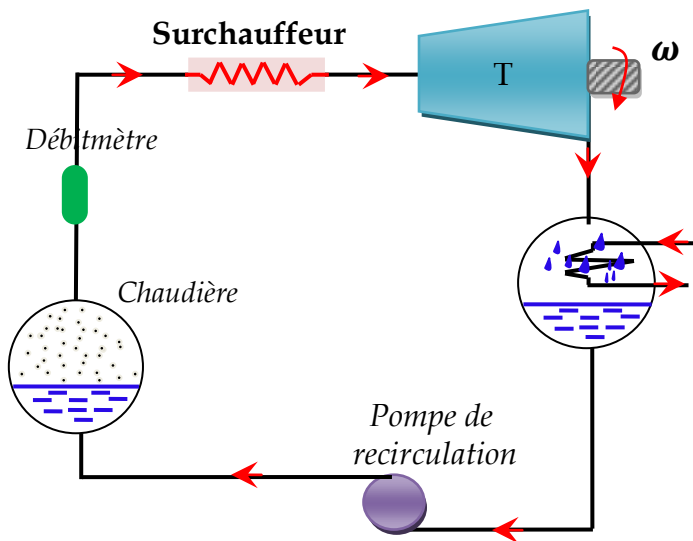


Figure I- 15: Cycle de Hirn

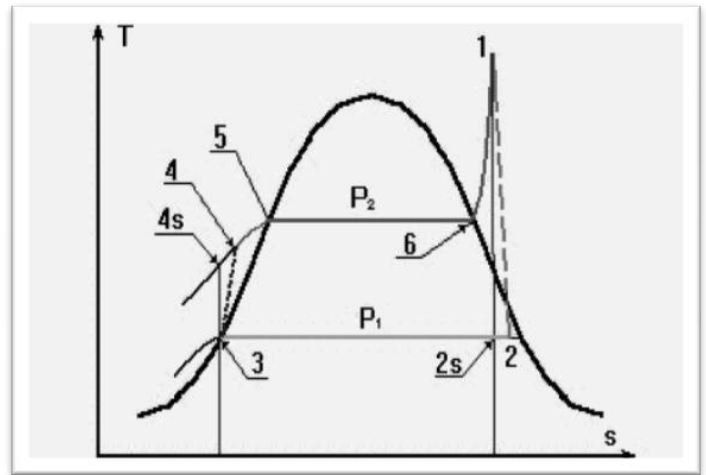


Figure I- 16: Diagramme (T, S) du cycle

	détente adiabatique irréversible, de HP à BP, dans la turbine
2-3	condensation isobare et isotherme du mélange jusqu'à eau liquide
3-4	compression de l'eau dans la pompe, passage de BP à HP
4-5-6	chauffage isobare et évaporation isobare et isotherme dans le G.V.
6-1	chauffage isobare de la vapeur

I-5-2-2-CYCLE A RESURCHAUFFE :

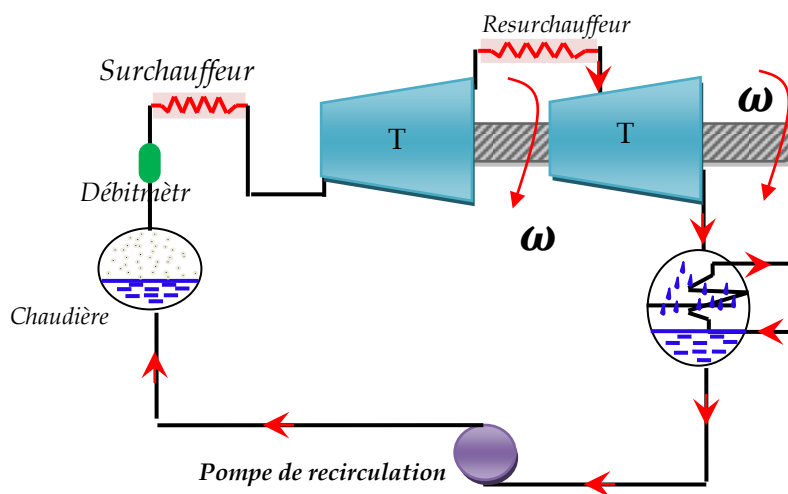


Figure I- 17: Cycle à Resurchauffé

1-2	1 ^{ère} détente adiabatique de HP à MP, dans la turbine HP
2-3	Chauffage isobare de la vapeur
3-4	2 ^{ème} détente adiabatique de MP à BP dans la turbine BP
4-5	Condensation isobare et isotherme du mélange jusqu'à l'eau liquide
5-6	Compression de l'eau dans la pompe, passage de BP à HP, W_{pompe} négligé
6-7	Chauffage isobare et évaporation isobare et isotherme dans le GV
7-1	Chauffage isobare de la vapeur

I-5-2-3-CYCLE A SOUTIRAGE :

Tous les cycles de centrale sont à soutirage. Il peut y a voir 6 à 8 soutirages effectués.

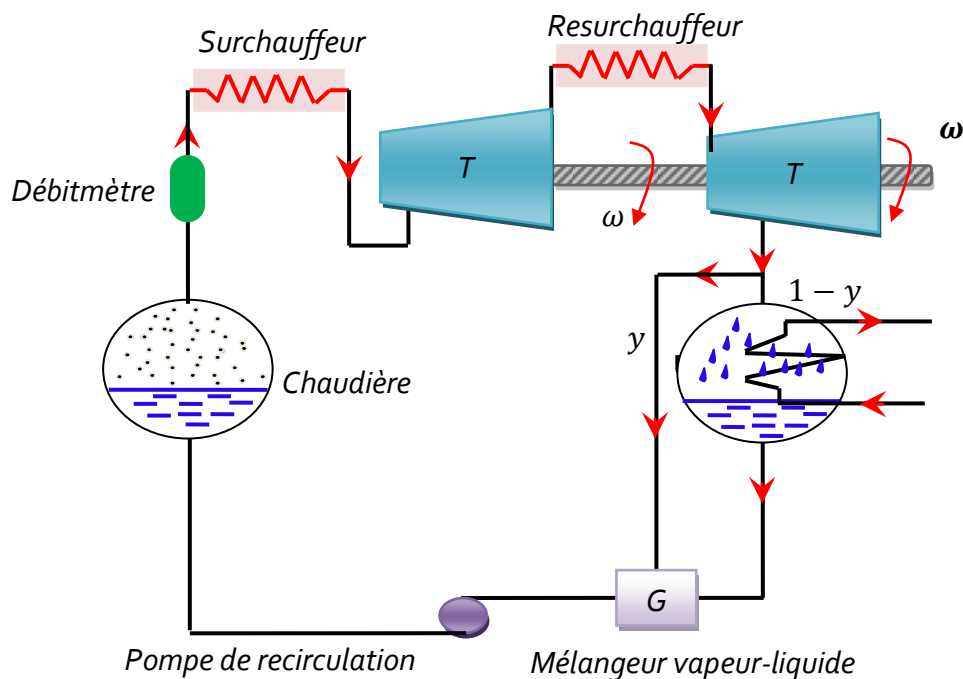


Figure I- 18: Cycle à soutirage

I-5-2-4-MODELISATION USUEL D'UN CYCLE A SOUTIRAGE :

Quantité initiale de fluide(1 + x). Soutirage d'une fraction.

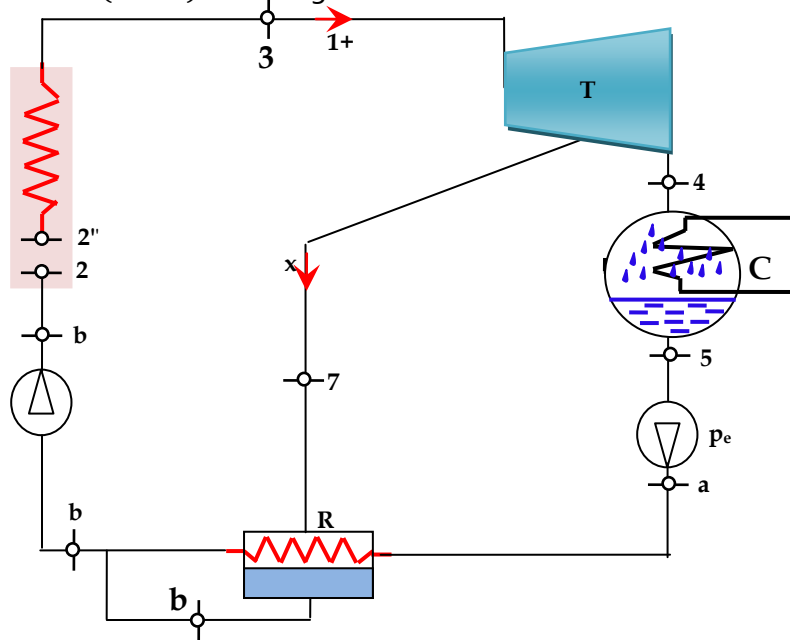


Figure I- 19:Modélisation d'un cycle de soutirage

Le cycle de la partie non soutiré :

3-4	Détente adiabatique, de HP à BP, dans la turbine
4-5	Condensation isobare et isotherme de mélange jusqu'à eau liquide
5-a	Compression d'eau liquide, passage de BP à MP
a-b	Chauffage de l'eau dans le mélangeur au contact de la fraction x soutiré

Le cycle de la fraction x soutiré :

3-7	Détente de la fraction x de vapeur, de HP à MP, dans la turbine
7-b	Condensation de la fraction x soutirée jusqu'à eau liquide, à la pression MP la quantité de chaleur nécessaire au changement de phase sert à chauffer l'eau de a à b

Fin de cycle pour la quantité (1 + x) de fluide :

b-b	Compression de la quantité (1 + x) d'eau liquide, passage de MP à HP
b-2'	Chauffage isobare de la quantité (1 + x) d'eau
2'-2''	Evaporation isobare de la quantité (1 + x) dans le GV
2''-3	Chauffage isobare de la quantité (1 + x) de vapeur

I-5-2-5-BILAN ENERGETIQUE D'INSTALLATIONS A VAPEUR :

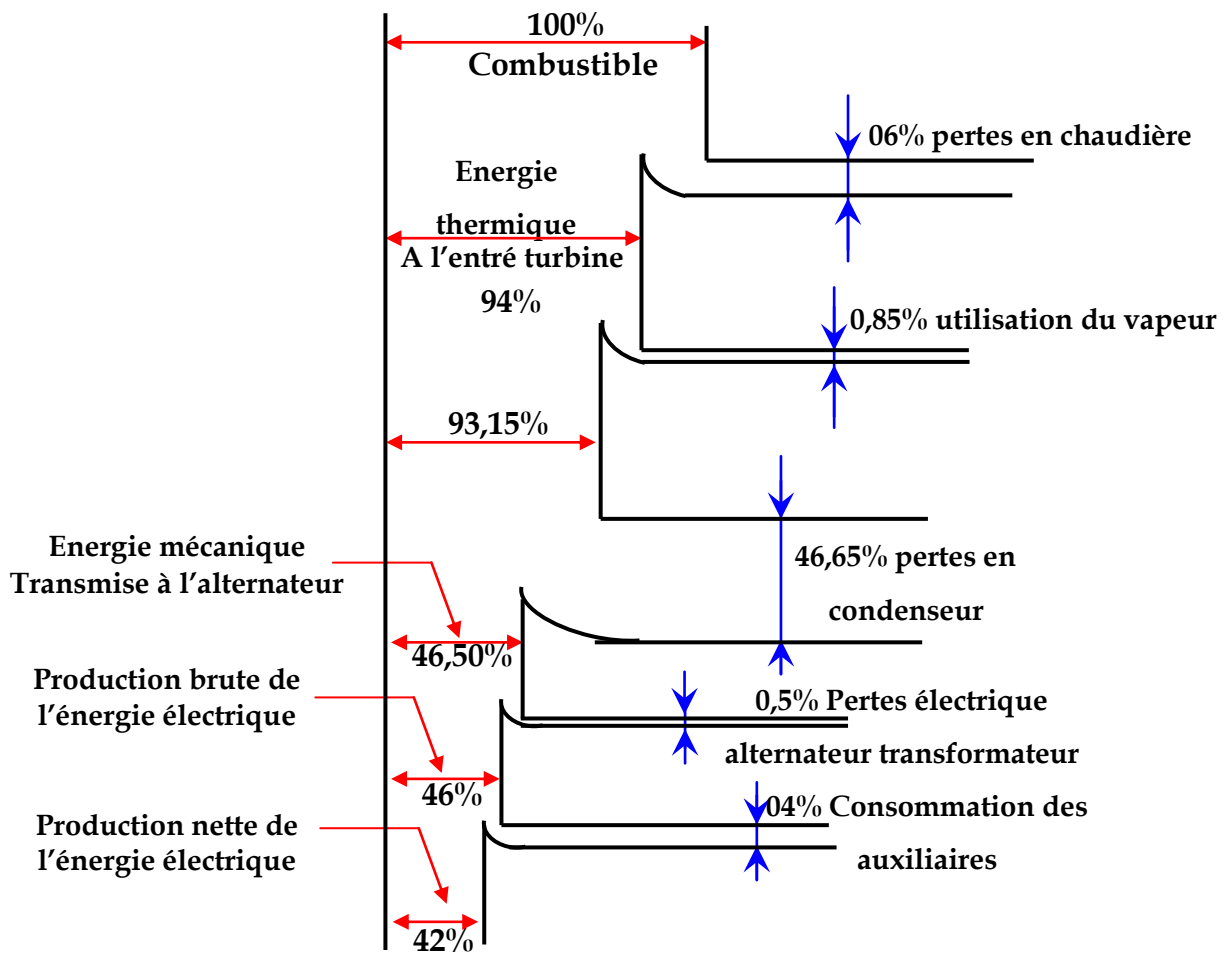


Figure I- 20: Bilan énergétique global d'une installation simple à vapeur



Pour produire 1000 kWh, il faut

- 350 Kg de charbon
- 250 L de fuel
- 300 m³ de gaz
- 4 g d'uranium enrichi

I-5-3-LES TURBINE A GAZ :

I-5-3-1-GENERALITES:

Les turbines à gaz sont séparées en deux catégories selon la nature de la puissance récupérée du fluide en sortie de la turbine :

❖ Les turbomoteurs et turbopropulseurs :

La turbine à gaz génère de la puissance mécanique extraite d'un arbre. Cet arbre peut entraîner un rotor d'hélicoptère, une hélice d'avion (turbopropulseur), un alternateur (groupe électrogène).

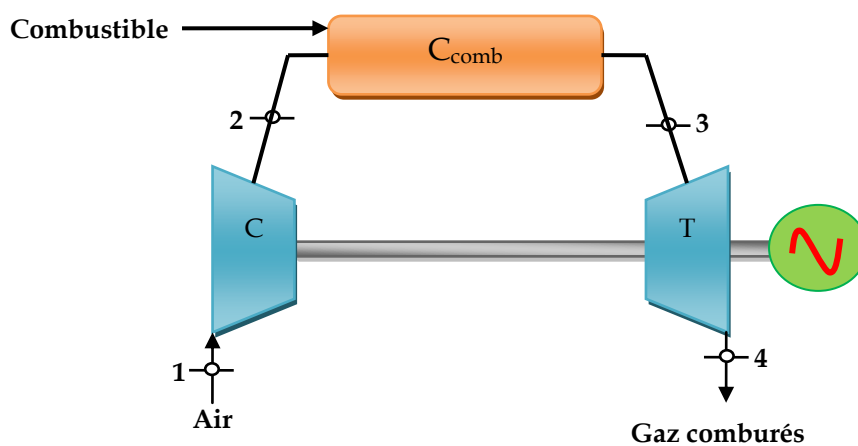


Figure I- 21: Turbomoteur

❖ Les turboréacteurs :

La turbine à gaz génère de l'énergie cinétique sous forme d'un jet à haut vitesse qui sert à la propulsion des avions.

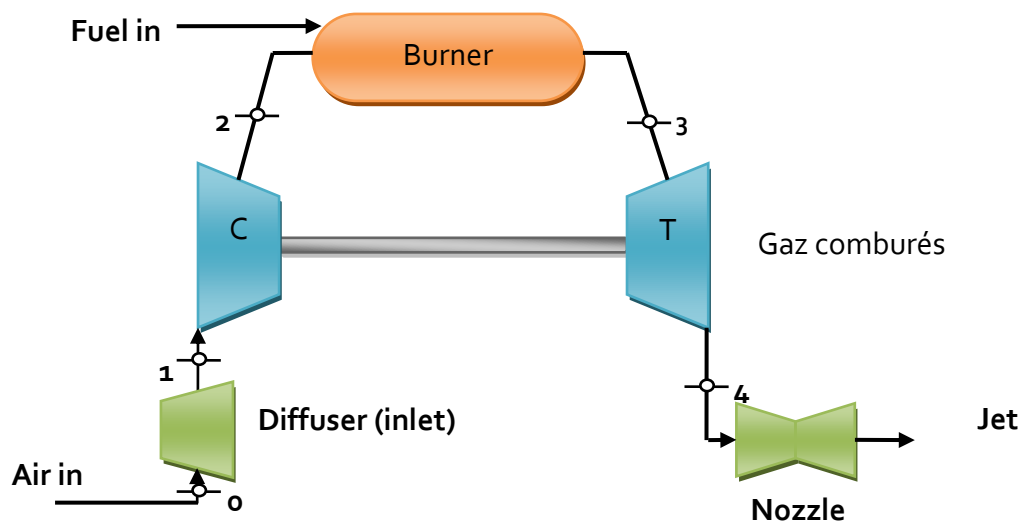


Figure I- 22: Turboréacteur

Le cycle thermodynamique qui représente bien les transformations des turbines à gaz est le cycle de **Brayton**.

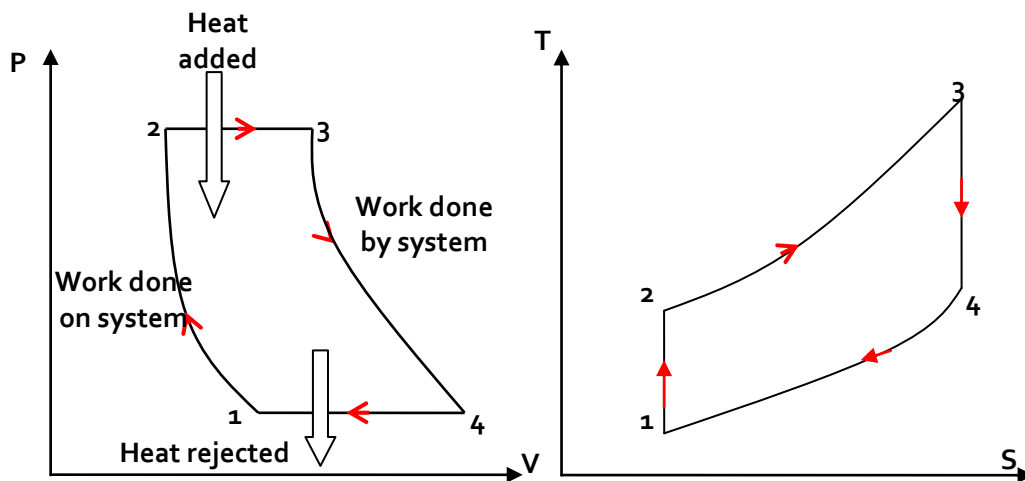


Figure I- 23: Cycle de Brayton

I-5-3-2-CYCLE REVERSIBLE DE BRAYTON :

Il existe plusieurs voies d'améliorations du cycle de **Brayton**.

❖ **FRACTIONNEMENT DE LA COMPRESSION ET DE LA D'ETENTE :**

Principe : introduire un refroidissement intermédiaire entre plusieurs compressions adiabatiques de taux de compression moindre.

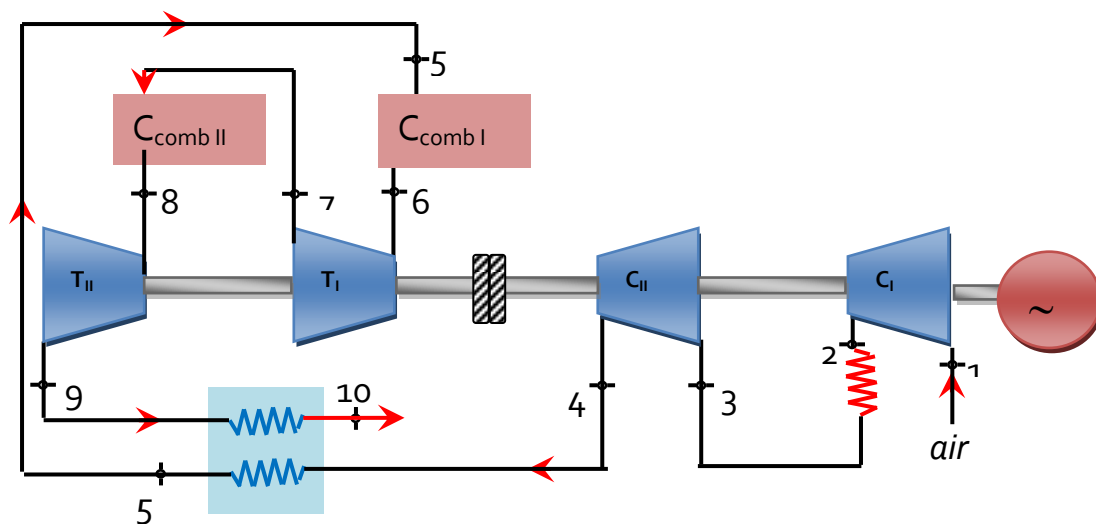


Figure I- 24: Fractionnement de la compression

❖ **REGENERATION:**

Principe : récupérer la chaleur des gaz chauds en sortie turbine, par l'intermédiaire d'un échangeur interne, pour chauffer les gaz après la compression et avant la combustion.

I-5-4-VALORISATION DE L'ENERGIE : COGENERATION - CYCLES COMBINES :

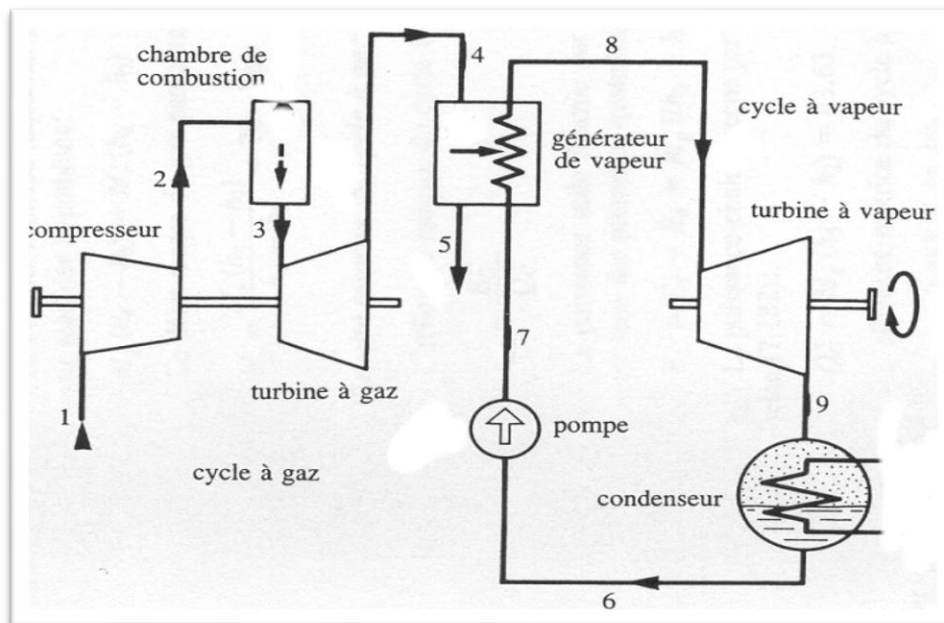


Figure I- 25: Cycle Combiné

I-6-ENERGIE :

I-6-1-ENERGIE - CHAINE ET SYSTEME ENERGETIQUE :

I-6-1-1-ENERGIE ET CONVERTISSEUR :

Energie brute ou primaire : source d'énergie avant toute intervention de l'homme pour la rendre utilisable (vent, soleil, hydrocarbures bruts...).

Densité énergétique d'un corps : un corps est dit énergétiquement dense s'il contient beaucoup d'énergie sous un faible volume. Il est dit diffus dans le cas contraire.

Classement des énergies selon trois types :

- ❖ *Les énergies diffuses et intermittentes :* rayonnement solaire, vent, vagues...elles sont cycliques et irrégulières et donc difficile à utiliser. En revanche, elles se renouvellent en permanence.
- ❖ *les énergies intermédiaires ou semi-denses :* biomasse, hydraulique, géothermie elles sont plus denses, plus régulières et se renouvèlent à une vitesse moyenne.
- ❖ *les énergies denses et stockées :* tourbe, charbon, pétrole, gaz naturel, uranium...ce sont principalement des matières organiques (hydrocarbures), très denses, stockables et

facilement utilisables par l'homme. Les matières fissibles comme l'uranium ou le thorium sont des métaux de densité énergétique énorme. Le grand défaut des énergies denses est qu'elles ne sont pas renouvelables (du moins sur des échelles de temps humaines).

Energie efficace ou finale : énergie qui a subi toutes les transformations nécessaires pour répondre à un besoin précis (essence, électricité...).

Convertisseur d'énergie : tout appareil ou toute technique qui permet de passer d'une énergie brute à une énergie domestiquée. Le convertisseur, c'est l'indispensable pont entre la source d'énergie et son utilisation.

I-6-2-CHAINE ENERGETIQUE :

Chaîne énergétique : chemin qui mène d'une source d'énergie brute à un usage et qui comprend tous les convertisseurs nécessaires.

Trois contraintes :

- ❖ *le lieu* : disponibilité de l'énergie à un endroit donné, transport
Difficulté de faire correspondre le lieu de production avec le lieu d'utilisation.
Recours à des convertisseurs qui assurent le transport. Ex : chaque tonne de pétrole consommée dans le monde parcourt en moyenne 4000km.
- ❖ *le temps* : disponibilité de l'énergie à un instant donné, stockage
Difficulté de faire correspondre le moment où la source est disponible et celui où on en a besoin.
Recours à des convertisseurs qui permettent de stocker l'énergie et de réguler son utilisation en fonction des besoins.
- ❖ *La qualité* : conditionnement de l'énergie pour son usage, transformation
Difficulté de mettre une source dans la forme qui correspond à son emploi.

Quatre opérations :

- ❖ *capter l'énergie* : extraction de gisement, capteur solaire, barrage...
- ❖ *transporter l'énergie* : transport, réseau de distribution, circulation d'eau chaude...
- ❖ *transformer l'énergie* : raffinage, production d'électricité...
- ❖ *utiliser l'énergie* : la dernière étape chez le consommateur.

I-7-LA CONSOMMATION D'ÉNERGIE PRIMAIRE :

Les énergies fossiles représentent près de 88% (chiffre de 2004) de la consommation d'énergie primaire. Le pétrole arrive en tête des énergies consommées. La consommation d'énergie primaire permet de mesurer le taux d'indépendance énergétique d'un pays.

Remarque : On utilise le joule ou la tonne d'équivalent pétrole (tep) pour comparer des formes d'énergie différentes.

Région	Consommation	Pétrole	Gaz-N	Charbon	Nucléaire	Hydro élec
<i>Amérique de nord</i>	<i>2784,40</i>	<i>1122,40</i>	<i>705,90</i>	<i>603,80</i>	<i>201,40</i>	<i>141,90</i>
<i>Etats unis</i>	<i>2331,60</i>	<i>937,60</i>	<i>582,0</i>	<i>564,30</i>	<i>187,90</i>	<i>59,8</i>
<i>Asie-Pacifique</i>	<i>3198,80</i>	<i>1090,50</i>	<i>330,90</i>	<i>1506,60</i>	<i>118,90</i>	<i>634,40</i>
<i>Chine</i>	<i>1386,20</i>	<i>308,60</i>	<i>35,10</i>	<i>956,90</i>	<i>11,30</i>	<i>74,20</i>
<i>Europe</i>	<i>2964</i>	<i>957,30</i>	<i>997,70</i>	<i>537,20</i>	<i>287,20</i>	<i>184,70</i>
<i>France</i>	<i>262,90</i>	<i>94,00</i>	<i>40,20</i>	<i>12,50</i>	<i>101,40</i>	<i>14,80</i>
<i>Allemagne</i>	<i>330,40</i>	<i>123,60</i>	<i>77,30</i>	<i>85,70</i>	<i>37,80</i>	<i>6,10</i>
<i>Moyen-Orient</i>	<i>481,90</i>	<i>250,90</i>	<i>218,0</i>	<i>9,10</i>	<i>--</i>	<i>04</i>
<i>Amérique latine</i>	<i>483,10</i>	<i>221,70</i>	<i>106,20</i>	<i>18,70</i>	<i>4,40</i>	<i>132,10</i>
<i>Afrique</i>	<i>312,10</i>	<i>124,30</i>	<i>61,80</i>	<i>102,80</i>	<i>3,40</i>	<i>19,80</i>

Tableau I- 3: Consommation d'énergie primaire par région, en millions de tep (Mtep)

I-8-CONCLUSION :

Aujourd'hui et pour les cinquante années à venir, de quelles énergies disposerons-nous ? Selon toute vraisemblance, la gamme des sources d'énergie que nous pourrons utiliser restera conventionnelle avec les combustibles fossiles (charbon, gaz, pétrole), l'énergie nucléaire (de fission) et les énergies renouvelables (hydraulique, solaire, éolien, biomasse et géothermie), la fusion thermonucléaire contrôlée restant un objectif beaucoup plus lointain.

Actuellement, la consommation mondiale d'énergie requiert les combustibles fossiles à hauteur de 80%, le nucléaire pour 07% et les énergies renouvelable pour 13%.

La population mondiale devrait atteindre 9 milliards d'habitants d'ici cinquante ans et entraînera une augmentation de la consommation d'énergie, notamment dans les pays en développement. À titre d'exemples, le parc automobile devrait passer de 400 millions à 1 milliard d'unités dès 2020 ; le nombre de villes de plus de 10 millions d'habitants, deux fois plus avides d'énergie que les régions rurales, doubler vraisemblablement.

Les estimations prévoient une consommation d'énergie multipliée par 1,4 d'ici à 2020. Dans ce contexte, quelle sera la place et le rôle des énergies dans l'avenir ?

SOMMAIRE :

I-1- Introduction :	8
I-2- De L'alternateur De Velo Vers Les Centrale :	9
I-2-1- Experience N°01 :	9
I-2-1-1- Observations :	9
I-2-1-2- Conclusion :	9
I-2-2- Experience N°02 :	9
I-2-2-1- Observations :	10
I-2-2-2- Conclusion :	10
I-2-3- Experience N°03 :	10
I-2-3-1- Observations :	11
I-2-3-2- Conclusion :	11
I-3- La Production De L'electricite :	11
I-4- Les Centrales Thermique :	12
I-4-1- Les Centrales Nucleaires :	13
I-4-1-1- Description :	13
I-4-1-2- Fonctionnement Technique :	15
I-4-1-3- Rendement D'une Centrale Nucleaire :	17
I-4-2- Les Centrales A Flamme:	17
I-4-3- Centrales Recuperant De La Chaleur Preexistante (Solaire, Geothermique...) :	18
I-4-3-1- Centrale Solaire Thermodynamique :	18
I-4-3-2- Centrale Geothermique:	20
I-5- Les Turbines :	21
I-5-1- Generalites :	21
I-5-2- Les Turbine A Vapeur :	23

I-5-2-1- Le Cycle De Hirn :.....	25
I-5-2-2- Cycle A Resurchauffe :.....	25
I-5-2-3- Cycle A Soutirage :.....	26
I-5-2-4- Modelisation Usuel D'un Cycle A Soutirage :.....	27
I-5-2-5- Bilan Energetique D'installations A Vapeur :.....	28
I-5-3- Les Turbine A Gaz :.....	29
I-5-3-1- Generalites:.....	29
I-5-3-2- Cycle Reversible De Brayton :.....	30
I-5-4- Valorisation De L'energie : Cogeneration - Cycles Combines :.....	31
I-6- Energie :.....	31
I-6-1- Energie - Chaine Et Systeme Energetique :.....	31
I-6-1-1- Energie Et Convertisseur :.....	31
I-6-2- Chaine Energetique :.....	32
I-7- La Consommation D'energie Primaire :.....	33
I-8- Conclusion :.....	34

LISTE DES FIGURES :

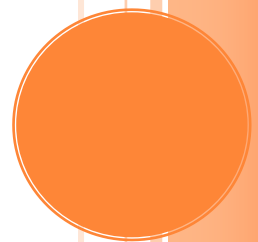
Figure I- 1 : Schéma d'une dynamo de vélo	9
Figure I- 2 : Aimants	9
Figure I- 3 : Expérience N°02	10
Figure I- 4 : Expérience N°03	10
Figure I- 5 : Centrale nucléaire.....	14
Figure I- 6:Centrale à flamme	18
Figure I- 7:Centrale Solaire de THERMIS	19
Figure I- 8:Centrale constituée de capteurs cylindro paraboliques	19
Figure I- 9:Centrale solaire parabolique (projet DISH-STIRLING à Font-Romeu-Odeillo)	19
Figure I- 10:Centrale solaire à tour (SolarTower en Californie).....	20
Figure I- 11: Centrale géothermique de Nesjavellir en Islande.....	21
Figure I- 12:Schéma de principe d'une turbine	23
Figure I- 13:Diagramme (T, S) du cycle	24
Figure I- 14:Cycle de Rankine	24
Figure I- 15:Cycle de Hirn	25
Figure I- 16:Diagramme (T, S) du cycle	25
Figure I- 17:Cycle à Resurchauffé	25
Figure I- 18:Cycle à soutirage	26
Figure I- 19:Modélisation d'un cycle de soutirage	27
Figure I- 20:Bilan énergétique global d'une installation simple à vapeur	28
Figure I- 21:Turbomoteur	29
Figure I- 22:Turboréacteur	29
Figure I- 23:Cycle de Brayton	30
Figure I- 24:Fractionnement de la compression	30
Figure I- 25:Cycle Combiné	31

LISTE DES TABLEAUX :

Tableau I- 1:Capacité géothermique installée (2002).....	21
Tableau I- 2: Echange de masse et d'énergie entre le système et le milieu extérieur	22
Tableau I- 3: Consommation d'énergie primaire par région, en millions de tep (Mtep)	33

Chapitre II

**Modèle de représentation
des systèmes dynamiques
non linéaire**



Chapitre II

Modèle de représentation des systèmes dynamiques Non linéaires



To catch a mouse alive, try this trick: When you see the beast running around a room, open a black umbrella and hold it on the floor at an angle that allows the mouse to run in (it helps if someone else can chase the mouse). Your prey, drawn to dark places, will probably head straight for the umbrella. Then quickly close it, go outside, and set your prisoner free.

—Reader's Digest Practical Problem Solver [241]

II-1-Introduction :

L'identification d'un système dynamique non-linéaire MISO (multi-entrées, mono-sortie) consiste à construire un modèle mathématique de ce système à partir des mesures de ses entrées et de sa sortie. Soit un ensemble de données d'apprentissage $\mathcal{D}_N = \{y_s(t), U(t), t = 1..N\}$ où $U(t)$ et $y_s(t)$ désignent respectivement les entrées et la sortie instantanées du système. L'identification d'un système MISO à partir de ces mesures expérimentales revient à trouver une relation mathématique de la forme :

$$y(t) = F[\varphi(t), \theta]$$

II – 1

Qui donne à chaque instant une estimation satisfaisante de la sortie y_s du système. La fonction multi variable F est le modèle du système et $y(t)$ la sortie instantanée de ce modèle. La fonction F dépend du vecteur de paramètres θ et d'un vecteur de variables expliquant le comportement du système appelé vecteur de régression et noté $\varphi(t)$. Le vecteur de régression est formé à partir des entrées U , de la sortie y_s du système ou de la sortie y du modèle à des instants antérieurs à l'instant courant t . Le problème d'élaboration du modèle décrit par l'équation (II-1) se décompose en 4 tâches

- ❖ Choix des éléments du vecteur de régression $\varphi(t)$,
- ❖ Détermination de la structure de la fonction F ,
- ❖ Estimation du vecteur des paramètres θ ,
- ❖ Evaluation des performances du modèle F .

La qualité de l'approximation du comportement du système dépendra de la structure retenue, du nombre de paramètres qu'elle contient, des variables d'entrée sélectionnées et de la quantité d'information contenue dans les données d'apprentissage. L'un des problèmes majeurs en identification est la détermination de la structure du modèle. Dans le cas linéaire, les techniques de recherche de structure sont bien connues et très développées [Ljung, 1987]. Mais l'hypothèse d'un comportement linéaire du système n'est souvent vérifiée que dans le voisinage immédiat d'un point de fonctionnement du système. Pour peu que le système s'en éloigne, son comportement devient non-linéaire. Pour le décrire, plusieurs types de modèles non-linéaires sont proposés dans la littérature. Les plus anciens sont basés sur la recherche d'un modèle global et les plus récents reposent sur le principe de décomposition d'un problème global en plusieurs problèmes simples. Dans les paragraphes qui suivent, nous dressons l'état de l'art de la modélisation des systèmes dynamiques non-linéaires. Après un tour d'horizon des différents modèles non-linéaires, nous abordons dans un cadre général les aspects pratiques liés à leur identification.

II-2-MODELES CLASSIQUES DES SYSTEMES NON-LINEAIRES :

L'un des plus anciens modèles de représentation des systèmes non-linéaires est celui de Volterra-Wiener. Pour un système à une entrée, il est donné par la décomposition en série :

$$y(t) = \sum_{i=1}^{\infty} y_i(t) \quad \text{II - 2 - a}$$

$$y_i(t) = \sum_{\tau_1=0}^{t-1} \dots \sum_{\tau_i=0}^{t-1} h_i(\tau_1, \dots, \tau_i) u(t - \tau_1) u(t - \tau_i) \quad \text{II - 2 - b}$$

$h_i(\tau_1, \dots, \tau_i)$ est la réponse impulsionnelle d'ordre i du système. Dans la pratique, la série est tronquée à l'ordre n mais du fait que le modèle utilise exclusivement les entrées décalées $u(t - \tau)$, il nécessite un grand nombre de paramètres pour représenter convenablement le système. Cette description présente donc peu d'intérêt en pratique.

Une autre catégorie de modèles est celle des modèles sous forme de blocs structurés, constitués par des blocs distincts de modèles dynamiques linéaires et de non-linéarités statiques.

Ces blocs interconnectés sont placés en série et/ou en parallèle pour former les modèles de **Wiener** et de **Hammerstein** généralisés [Haber et Unbehauen, 1990]. La forme simple de ces structures est illustrée sur les figures II-1 et II-2. Sur ces schémas, le signal intermédiaire $v(t)$ entre les deux blocs n'est pas accessible.

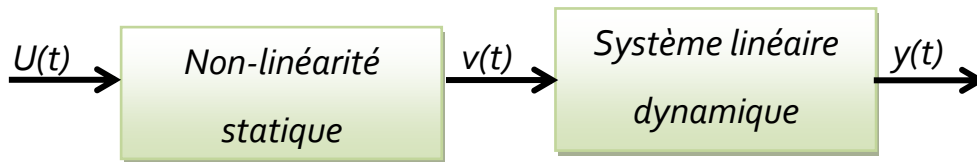


Figure II- 1: Modèle de Hammerstein

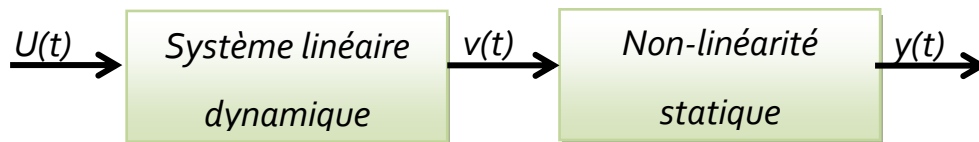


Figure II- 1 : Modèle de Wiener

Les techniques d'identification de ces modèles qu'on peut trouver notamment dans les références [Ouladsine et al. 1993], [Wigren, 1993], [Boutayeb et al, 1993] et [Boutayeb et Darouach, 1995] approchent souvent la non-linéarité par un polynôme de degré fini. Le modèle de Hammerstein d'un système mono-entrée et mono-sortie est exprimé dans ce cas par les équations suivantes :

$$v(t) = c_0 + c_1 u(t) + c_2 u^2(t) + \dots + c_p u^p(t) \quad \text{II} - 3$$

$$y(t) = \frac{b_1 q^{-1} + b_2 q^{-2} + \dots + b_m q^{-m}}{1 + a_1 q^{-1} + a_2 q^{-2} + \dots + a_n q^{-n}} v(t) \quad \text{II} - 4$$

Ou q^{-1} est l'opérateur retard.

II-3-APPROCHES RECENTES DE REPRESENTATION DES SYSTEMES NON-LINEAIRES :

Elles suscitent beaucoup d'intérêt ces dernières années et présentent des propriétés attractives pour la modélisation des systèmes présentant des caractéristiques fortement non linéaires.

Leur philosophie est basée sur la stratégie "**diviser pour régner**" : le système étudié est représenté par un ensemble de modèles locaux $f_i(\varphi(t))$ valides dans une zone de l'espace de fonctionnement du système. La validité locale de chaque modèle est spécifiée par une fonction à support borné (une fonction d'allure gaussienne par exemple). La sortie du modèle du système s'exprime comme la somme pondérée des sorties $y_i(t)$ des modèles locaux, soit :

$$y(t) = \sum_i \rho_i(t) y_i(t) = \sum_i \rho_i(t) f_i(\varphi(t)) \quad \text{II - 5}$$

Les fonctions de pondération $\rho_i(t)$ constituent les fonctions de validité associées aux modèles locaux.

La construction de modèles basés sur cette relation a donné lieu à différents types de structures dont les réseaux à fonctions de base radiale (RFBR), les modèles flous, les modèles multi-experts, ... Les réseaux de neurones qui modélisent le système comme une succession de couches de neurones interconnectés sont généralement rangés dans cette catégorie de modèles non-linéaires. Notons que les modèles sous forme de blocs structurés précédemment présentés adhèrent au principe de décomposition sans toutefois faire appel à des fonctions de validité locale. Dans les paragraphes suivants, nous présentons ces modèles non-linéaires ainsi que la problématique de leur identification.

II-3-1-MODELES FLOUS :

Introduits par **Zadeh** en 1965, les concepts flous permettent de caractériser un système ou un phénomène par un ensemble de règles composées de propositions linguistiques exprimant les imprécisions du raisonnement humain. Modéliser un système à partir de ces concepts revient à générer une base de règles de la forme :

Si prémisse alors conséquence

Un exemple de règle est : **si la température est basse alors mettre le chauffage**. La variable température est appelée variable de prémisse. Etant donnée une mesure de la température, pour évaluer le degré de véracité de la proposition sur la température, on se sert d'un ensemble flou. Cet ensemble flou correspond à un qualificatif ou un label linguistique (*faible, moyen, élevé*) de la variable température. Il joue le rôle d'interface entre la valeur numérique et les qualificatifs linguistiques de cette variable. Le nombre de termes linguistiques (encore appelé nombre de modalités) associés à une variable détermine l'importance de cette variable dans la caractérisation de la non-linéarité du système.

Soit x une variable de prémisses. A un ensemble flou A (représentant par exemple le qualificatif *faible*) de la variable x , on associe une fonction dite d'appartenance qui fournit le degré d'appartenance d'une mesure de x à cet ensemble flou. Cette fonction d'appartenance est définie sur l'univers de discours X de la variable x (c'est-à-dire son intervalle de variation) par :

$$\mu_A(x): X \rightarrow [0,1]$$

II – 6

Elle détermine de manière graduelle l'appartenance d'une mesure de x à A . En comparaison, les ensembles mathématiques classiques ont des fonctions d'appartenance qui prennent deux valeurs : 0 (non appartenance) ou 1 (appartenance). Les fonctions d'appartenance sont des fonctions à support borné, c'est-à-dire des fonctions qui sont significativement différentes de zéro sur un domaine limité de la variable x . Les fonctions d'appartenance usuellement employées (leur définition mathématique, leur forme) sont regroupées dans le tableau II-1.

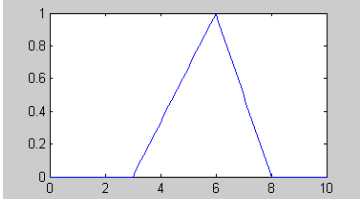
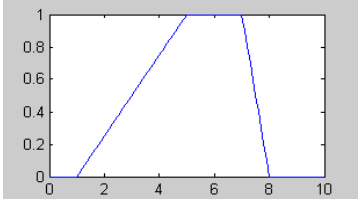
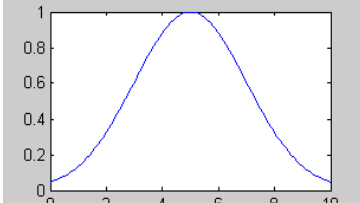
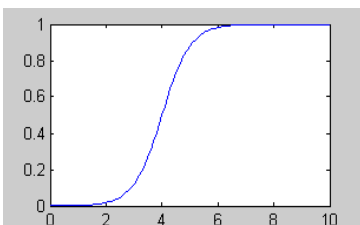
Fonction	Formule	Forme de la fonction
Triangle	$\max\left(\min\left(\frac{x - b_1}{c_1 - b_1}, \frac{x - b_2}{c_2 - b_2}, 0\right)\right)$	
Trapèze	$\max\left(\min\left(\frac{x - b_1}{c_1 - b_1}, 1, \frac{x - b_2}{c_2 - b_2}, 0\right)\right)$	
Gaussienne	$\exp\left(-\frac{(x - c)^2}{2\sigma^2}\right)$	
Sigmoïde	$\frac{1 + \tanh\left(\frac{x - c}{\sigma}\right)}{2}$	

Tableau II- 1 : Fonctions d'appartenance les plus utilisées en modélisation floue

La structure générale d'un modèle flou présentée sur la figure II-3 comporte quatre parties principales :

- ❖ La fuzzification : elle transforme les valeurs numériques d'entrée en ensemble flou,
 - la base de connaissance constituée d'une :
 - base de règles qui contient toutes les règles nécessaires pour décrire le système,
 - base de données qui comprend les paramètres définissant les fonctions d'appartenance associées aux ensembles flous et les paramètres définissant la partie conséquence des règles,
- ❖ le moteur d'inférence : il déduit l'ensemble flou de sortie correspondant à une entrée du système étant donnée la base de connaissance,
- ❖ la défuzzification : elle transforme l'ensemble flou de sortie en une valeur numérique si nécessaire. Plusieurs techniques sont applicables mais la plus usitée est celle du centre de gravité.

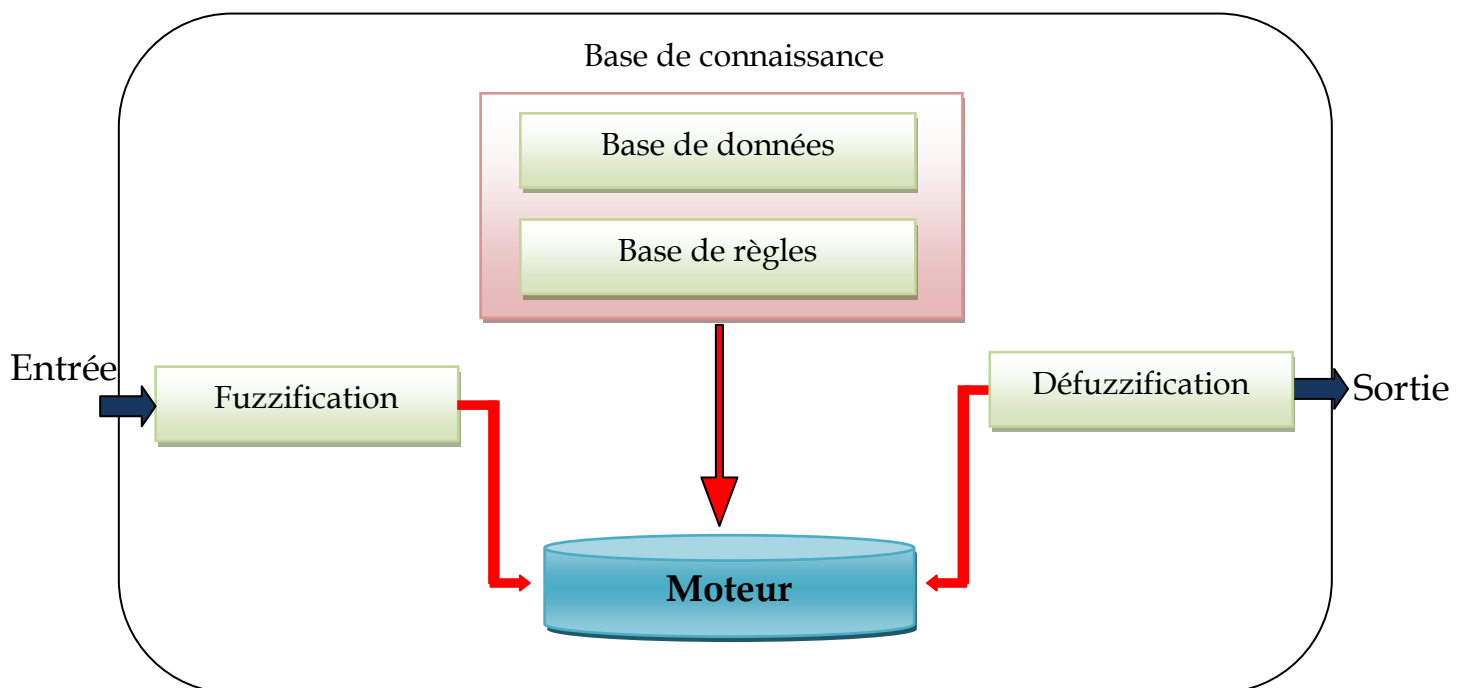


Figure II- 2 : Structure générale d'un modèle flou

La prémisse d'une règle est toujours exprimée sous la forme de propositions portant sur les variables de prémisse. Par contre, l'expression de la partie conséquence peut avoir différentes variantes. Suivant la forme de la partie conséquence, on distingue principalement trois types de modèles flous [Babuška, 1998] :

- ❖ Les modèles flous linguistiques proposés par Zadeh en 1973 et Mamdani en 1977 (voir [Babuška, 1998]) : la partie conséquence est un ensemble flou et un même ensemble flou de sortie peut être associé à plusieurs ensembles flous de prémisse,
- ❖ Le modèle à relations floues proposé par Pedrycz en 1984 (voir [Babuška, 1998]) : c'est une généralisation du modèle précédent car une prémisse est associée à plusieurs propositions de conséquence par l'intermédiaire d'une relation floue,

Le modèle de Takagi-Sugeno [Takagi et Sugeno, 1985] : la conséquence est une fonction affine de variables de conséquence.

Les modèles flous linguistiques et les modèles à relation floue forment une classe de modèles adaptés pour la représentation des systèmes non-linéaires en se basant sur des concepts proches du langage humain. Comparativement, le modèle de Takagi-Sugeno relève plus de la modélisation classique des systèmes car sa partie conséquence n'est pas floue et n'introduit pas d'imprécision. Nous présentons dans les paragraphes suivants les deux types de modèles flous les plus fréquemment rencontrés dans la littérature : modèle de Mamdani et celui de Takagi-Sugeno.

II-3-1-1-MODELE DE MAMDANI :

Soit le modèle (II-1) ; le modèle de Mamdani exprime la fonction F sous la forme d'un ensemble de règles R_i :

$$R_i : \text{Si } z(t) \text{ et } A_i \text{ alors } y(t) \text{ et } B_i \quad i = 1, \dots, M \quad \text{II} - 7$$

A_i est l'ensemble flou d'entrée et B_i l'ensemble floue de sortie pour le règle R_i . M est le nombre de règle de la base de connaissance. Le vecteur $z(t)$, ($z(t) \in \mathbb{R}^{n_z}$) est le vecteur de des variables des prémisses. Ce vecteur peut être la totalité ou une partie du vecteur de régression ou comprendre des variables auxiliaires susceptibles de caractériser les non-linéarités du système.

L'ensemble flou A_i est un ensemble multi variable. Dans la pratique, comme il est difficile de raisonner sur un ensemble multi variable et surtout pour avoir une interprétation linguistique des règles, il est parfois préférable d'exprimer la partie prémisse comme une combinaison logique (conjonction, disjonction ou négation) de propositions portant sur chaque élément du vecteur $z(t)$. Suivant cette idée, la partie prémisse de la règle A_i se décompose, par exemple, comme suit :

$$R_i : \text{Si } z_1(t) \text{ est } A_{i,1} \text{ ou } z_2(t) \text{ est } A_{i,2} \dots \text{ et } z_{n_z}(t) \text{ est } A_{i,n_z} \text{ alors } y(t) \text{ est } B_i \quad \text{II} - 8$$

Où les $A_{i,j}$ sont les modalités associées à chacune des variables de prémisse $Z_j(t)(j = 1, \dots, n_z)$ la règle R_i . L'ensemble flou multivariable A_i est ensuite construit à partir des ensembles flous mono variables $A_{i,j}$. La forme résultante de l'ensemble flou A_i dans l'espace de prémisse (espace engendré par le vecteur de prémisse \mathbf{z}) dépend des opérateurs utilisés pour traduire les combinaisons logiques. Les opérateurs les plus courants sont présentés dans le tableau II-2.

Opérateurs logiques	Symbole	Opérateurs probabilistes	Opérateurs de Lukasiewicz	Opérateurs de Zadeh
A_1 ET A_2	$A_1 \wedge A_2$	$\mu_{A_1} \times \mu_{A_2}$	$\max(\mu_{A_1} + \mu_{A_2} - 1, 0)$	$\min(\mu_{A_1}, \mu_{A_2})$
A_1 OU A_2	$A_1 \vee A_2$	$\mu_{A_1} + \mu_{A_2} - \mu_{A_1} \times \mu_{A_2}$	$\min(\mu_{A_1} + \mu_{A_2}, 1)$	$\max(\mu_{A_1}, \mu_{A_2})$
Non A_1	\bar{A}_1	$1 - \mu_{A_1}$	$1 - \mu_{A_1}$	$1 - \mu_{A_1}$

Tableau II- 2:Tableau des opérateurs logiques et des principaux opérateurs algébriques correspondants

Le degré de véracité de la règle R_i est déterminé à partir des degrés d'appartenance respectifs des variables $Z_j(t)$ et des opérateurs logiques. Généralement, la partie prémisse est formée par la conjonction des propositions individuelles. Si l'on considère l'opérateur produit comme opérateur de conjonction, le degré de véracité de la règle R_i ou de façon équivalente, le degré d'appartenance du vecteur $\mathbf{z}(t)$ à l'ensemble flou multi variable A_i , s'exprime dans ce cas :

$$v_i(\mathbf{z}(t)) = \mu_{A_i}(\mathbf{z}(t)) = \prod_{j=1}^{n_z} \mu_{A_{i,j}}(z_j(t)) \tag{II - 9}$$

Le mécanisme d'inférence consiste à déterminer l'ensemble flou de sortie à partir de la base de règles et des entrées du système. Pour cela, on recherche pour chacune des règles R_i , une relation de déduction \mathfrak{h} de l'ensemble flou de sortie \mathbb{B}_i connaissant l'ensemble flou de prémisse \mathbb{A}_i de ces règles. Ces relations étant construites, l'ensemble flou \mathbb{B}'_i image d'un ensemble flou d'entrée \mathbb{A}' pour la règle R_i est fournie par la relation de composition :

$$\mathbb{B}'_i = \mathbb{A}' \circ \mathfrak{h}_i$$

L'ensemble flou de sortie \mathbb{B}' image de \mathbb{A}' par le modèle de Mamdani est formé alors par l'union des ensembles flous individuels de sortie, soit :

$$\mathbb{B}' = \bigcup_{i=1}^M \mathbb{B}'_i$$

Pour obtenir une sortie numérique, l'ensemble flou \mathbb{B}' est défuzzifié. La technique de défuzzification usuelle est celle de centre de gravité qui exprime la sortie numérique comme le centre de masse de l'ensemble \mathbb{B}' :

$$y(t) = \frac{\int_Y y(t) \mu_{\mathbb{B}'}(y(t)) dy}{\int_Y \mu_{\mathbb{B}'}(y(t)) dy} \quad \text{II - 10}$$

II-3-1-2-MODELE DE TAKAGI-SUGENO (MODELE TS):

Ce modèle exprime la partie conséquence des règles comme une fonction affine des variables d'entrée :

$$R_i: \text{si } z(t) \text{ est } \mathbb{A}_i \text{ alors } y_i(t) = \varphi^T(t) \theta_{i1} + \theta_{i0} \quad \text{II - 11}$$

Géométriquement, il équivaut à associer un hyperplan à chaque zone de l'espace de prémisse. Le mécanisme d'inférence de ce modèle est simple car il consiste en une relation barycentrique pondérée par le poids de chaque règle. La sortie globale du système est donc :

$$y(t) = \frac{\sum_{i=1}^M v_i(z(t)) y_i(t)}{\sum_{j=1}^M v_{ij}(z(t))} \quad \text{II - 12}$$

Posons :

$$\omega_i(z(t)) = \frac{v_i(z(t))}{\sum_{j=1}^M v_j(z(t))} \quad \text{II - 13}$$

$\omega_i(z(t))$ est appelé degré de véracité normalisé de la règle R_i . L'équation II-12 s'écrit :

$$y(t) = \sum_{i=1}^M \omega_i(z(t)) y_i(t) \quad \text{II - 14}$$

On constate d'après la définition des degrés de véracité $\omega_i(z(t))$ que ceux-ci vérifiant la contrainte de partition unité, c'est-à-dire, leur somme vaut 1 pour tout échantillon de la base de données :

$$y(t) = \sum_{i=1}^M \omega_i(z(t)) = 1 \quad \forall t$$

II – 15

La définition d'un modèle TS introduit la notion de degré de véracité normalisé lié au mécanisme d'inférence du modèle TS. On constate, d'après l'équation II-13, que le degré de véracité normalisé $\omega_i(z(t))$ est en réalité le degré d'activation du modèle local correspondant : c'est cette fonction qui détermine l'importance relative du modèle local associé dans l'expression de la sortie $y(t)$ du modèle global suivant la zone où évolue le système. De par la contrainte de partition unité, les fonctions $\omega_i(z(t))$ sont assimilables à des fonctions d'appartenance relatives ou probabilistes [Baraldi et Blonda, 1999]. Comparativement, les degrés de véracité $v_i(z(t))$ sont des fonctions d'appartenance absolues ou possibiliste [Krishnapuram et Keller, 1996] en ce sens qu'elles évaluent la véracité d'une règle particulière en ignorant tout ce qui est relatif aux autres règles. Pour mieux illustrer la différence entre $v_i(z(t))$ et $\omega_i(z(t))$, nous considérons un modèle TS à une variable de prémisse. Deux ensembles flous sont définis sur le support de cette variable et sont caractérisés par des fonctions d'appartenance trapézoïdales. A gauche sur la figure II-3, sont représentés les degrés de véracité $v_i(z(t))$ alors qu'à droite sont tracés les degrés de véracité normalisés $\omega_i(z(t))$.

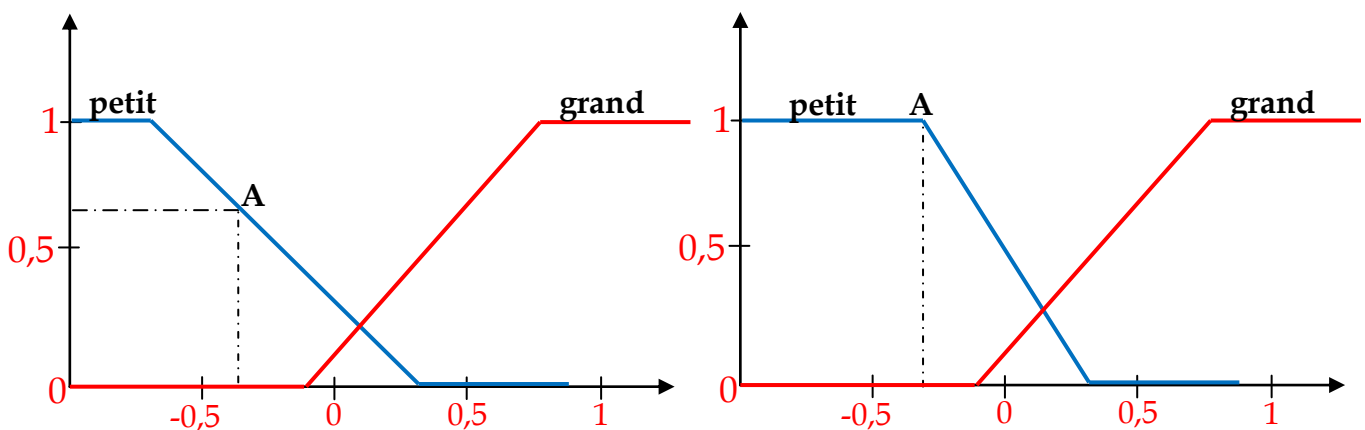


Figure II- 3 : Illustration de la différence entre les degrés de véracité des règles (partie gauche) et les degrés d'activation (partie droite) des modèles locaux associés.

On s'aperçoit que les degrés d'activation $\omega_i(z(t))$ diffèrent des degrés de véracité $v_i(z(t))$. En particulier, le point A relève de la modalité petite avec un degré de véracité de 0.6. Par suite de la normalisation, il est "promu" à un degré de 1. Ceci signifie que le modèle local associé à la modalité petit est activé à 100% au point A bien que le degré de véracité soit de 0.6. Il y a donc discordance entre l'activation réelle d'un modèle local et la véracité de la prémisse qui a généré ce modèle local. Or comme le mécanisme d'inférence du modèle TS réalise une coopération et non

une compétition entre les modèles locaux, ce sont les fonctions d'activation $\omega_i(z(t))$ qui fixent la contribution au modèle global des modèles locaux associés. Nous qualifierons donc les ensembles flous générés par ces fonctions, d'ensembles flous relatifs ou normalisés par opposition aux ensembles flous absolus décrits par les degrés de véracité. Les deux types d'ensembles flous coïncident si les degrés de véracités forment naturellement une partition unité :

C'est le cas des partitions floues fortes [Lindskög, 1996], [Glorennec, 1999] ou des ensembles flous issus d'une procédure de classification [Babuška, 1998].

II-3-2-PROPRIETES DES MODELES FLOUS :

On peut constater après cette description des deux types de modèles flous courants, que la représentation des systèmes par les modèles flous, en particulier le modèle de Takagi-Sugeno, offre un formalisme intéressant permettant d'intégrer facilement des connaissances a priori et des expériences acquises sur le système. En effet, si des connaissances physiques sur le comportement du système dans certaines zone de fonctionnement sont disponibles, certaines règles peuvent être définies a priori et ne seront pas modifiées lors de la phase d'identification du modèle [Bortolet, 1998]. Les modèles locaux associés à ces règles peuvent être issus des considérations physiques. A défaut du modèle physique du système, si des informations a priori sur le gain statique, les temps de réponse et les conditions de stabilité du système dans certaines zones de fonctionnement sont connues, leur incorporation dans le modèle flou pourra se traduire par des contraintes inégalités sur les paramètres de la partie conséquence [Abonyi et al., 2000-a,2000-b], [Lindskög, 1996].

En plus de cette possibilité d'intégration de connaissances a priori, les modèles flous sont des approximateurs universels. Les propriétés d'approximation des modèles flous de type Mamdani ont été étudiées entre autres par Wang et Mendel [1992], Kosko [1992] qui ont établi que ces modèles peuvent approcher avec une précision arbitraire toute fonction continue définie sur un domaine fini. Plus récemment, Ying [1998] a formulé, sur la base du théorème de Weierstrass, les conditions suffisantes garantissant les propriétés d'approximateurs universels des modèles de type Takagi-Sugeno.

II-3-3-PROBLEMATIQUE D'IDENTIFICATION DES MODELES FLOUS :

La représentation d'un système par un modèle flou soulève des problèmes pratiques. En particulier, il faut décider du type de modèle flou (Mamdani, TS), déterminer les variables de prémisse. La

prochaine étape est la décomposition de l'espace de prémisse, la caractérisation de la partie conséquence. La décomposition de l'espace de prémisse peut être réalisée par une partition grille, une partition suivant un arbre de décision ou à travers une technique de classification.

II-4-RESEAUX DE NEURONES :

L'un des modèles non-linéaires les plus populaires après les modèles flous est sans conteste les réseaux de neurones. Les premiers travaux sur les réseaux de neurones avaient pour objectif la modélisation du fonctionnement des neurones biologiques. Le réseau de neurones biologiques est une structure de traitement parallèle et distribué de l'information constituée par un ensemble d'unités de traitement (les neurones) interconnectées entre elles. L'utilisation de ces concepts pour la représentation des systèmes non-linéaires a connu un grand succès et de nombreuses applications dans divers domaines (l'automatique, l'économie, la reconnaissance de formes, ...) sont rapportées dans la littérature.

Dans sa structure, un réseau de neurones comprend une couche d'entrée, une couche de sortie, une ou plusieurs couches intermédiaires appelées couches cachées. Ces couches comportent des neurones qui reçoivent en entrée des signaux qui subissent un traitement. Le signal résultant est ensuite transmis aux neurones d'une autre couche ou de la même couche. Les connexions entre couches peuvent être unidirectionnelles ou bidirectionnelles.

On distingue différents types de réseaux de neurones mais en identification des systèmes non-linéaires, la structure reine est le perceptron multicouches MLP (acronyme de Multi Layer Perceptron). A chaque neurone d'une couche cachée, une fonction d'activation agit sur la somme des entrées du neurone pondérées par des poids w_i (voir figure II-5). La fonction d'activation est souvent une sigmoïde d'équation :

$$g(x) = \frac{1}{1 + e^{-x}} \quad \text{II - 16}$$

La sortie d'un MLP à n_c couches cachés comportant n neurones est exprimés par :

$$y(t) = w_0^{(s)} + \sum_{j=1}^n w_0^{(s)} \phi_j(x(t), W_j^{(c)}) \quad \text{II - 17}$$

Où la composante (s) fait référence aux paramètres de la couche de sortie et l'exposant (c) aux termes et paramètres des couches cachés. Le terme et paramètres des couches cachés. Le terme ϕ_j

est la sortie du neurone j de la couche cachée n_c . W_j est un vecteur regroupant les poids agissant sur les entrées du neurone j . De façon générale, la sortie $\phi_j(t)$ d'un neurone d'une couche cachée est calculée à partir des entrées $x_i(t)$ provenant de la couche cachée précédente par la relation :

$$\phi_j(x(t), W_j^{(c)}) = g \left(w_{0,j}^{(c)} + \sum_{i=1}^n w_{i,j}^{(c)} x_i(t) \right) \tag{II - 18}$$

Une extension du réseau MLP est de considérer la sortie du réseau décalée dans le temps comme un signal d'entrée des neurones des couches cachées. On parle alors de MLP récurrent.

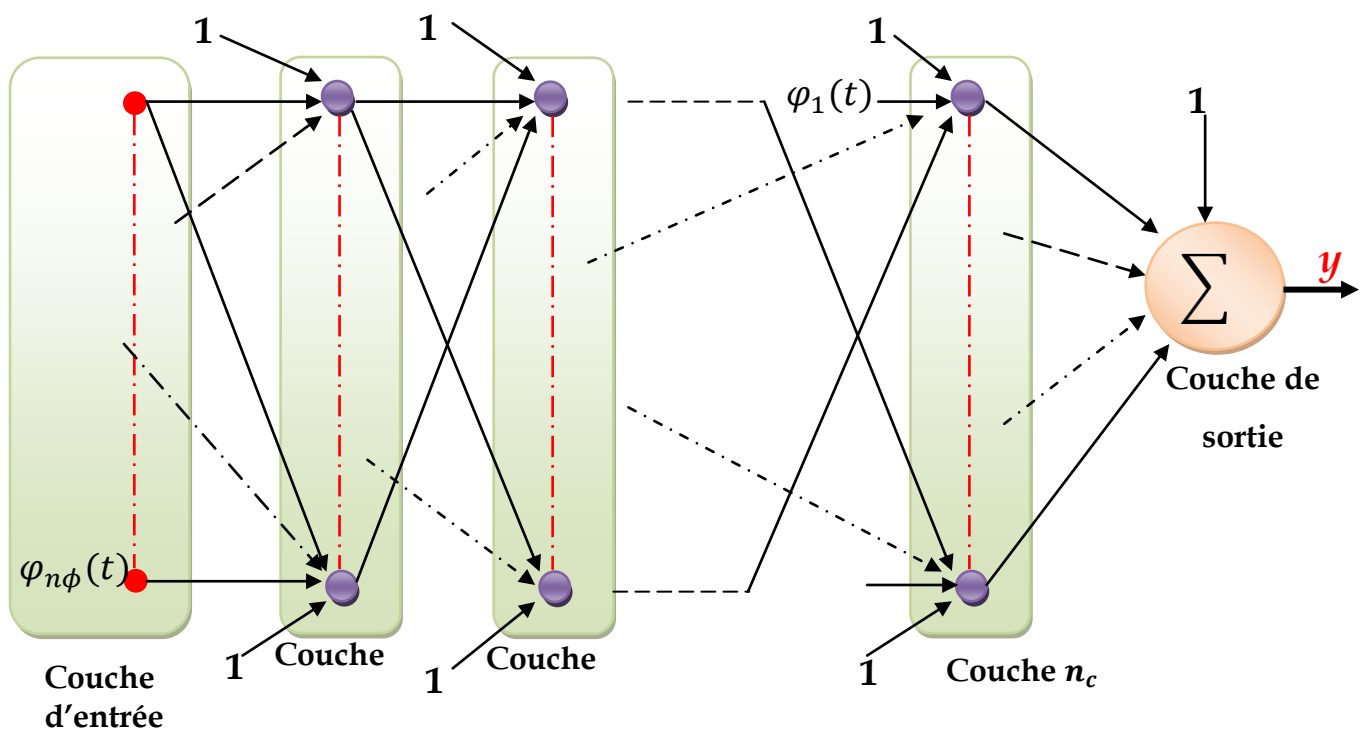


Figure II- 4: Structure générale d'un réseau MLP

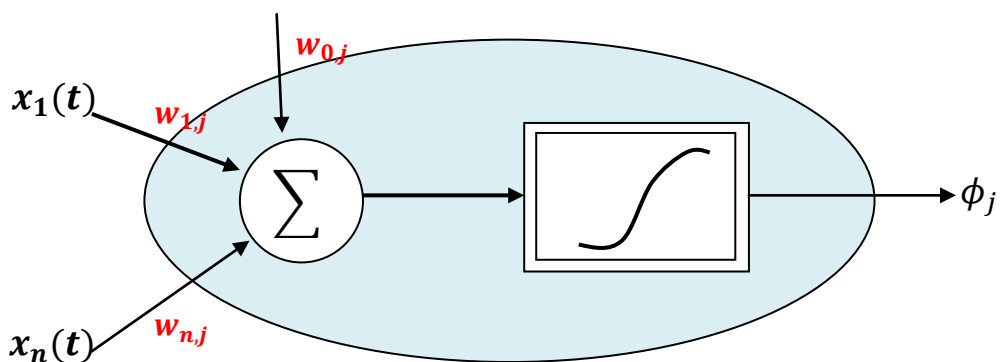


Figure II- 5: Structure d'un neurone d'une couche cachée

Les propriétés d'approximateurs universels des MLP ont été établies notamment par Funahashi [1989] : elles découlent de la flexibilité des réseaux de neurones, flexibilité qui est liée aux degrés de liberté procurés par le nombre de couches cachées et le nombre de neurones par couche.

L'identification d'un MLP comporte la détermination des entrées $\phi_i(t)$ du réseau, le choix du nombre de couches cachées et du nombre de neurones par couche. Pour une structure fixée, des techniques d'optimisation non-linéaire (gradient, Gauss-Newton, ...) et d'autres dérivées de la structure connexionniste des réseaux de neurones (algorithme de propagation arrière) sont utilisées pour estimer les poids afin d'adapter la sortie du réseau à celle du système. Afin de déterminer la taille optimale du réseau de neurones, deux approches heuristiques sont applicables : la première consiste à augmenter de façon graduelle le nombre de couches et le nombre de neurones par couche. Le contrôle de la croissance du réseau est réalisé en testant les capacités d'approximation du réseau obtenu sur des données de validation. La deuxième stratégie démarre avec un réseau important puis procède par suppression successive des connexions entre neurones (élimination des poids des connexions) jusqu'à l'obtention d'une structure satisfaisante. Une synthèse des techniques d'élimination des poids est présentée dans l'article de Kerling [1999].

II-5-LES RESEAUX DE FONCTIONS A BASE RADIALE (RFBR) :

Un RFBR réalise mathématiquement une décomposition sur une base de fonctions radiales :

$$y(t) = \sum_{i=1}^M w_i g(\|\varphi(t) - C_i\|_{\Sigma_i}) \quad \text{II - 19}$$

Avec :

g : fonction d'activation à base radiale,

w^i : Poids,

C_i : Centre de la fonction d'activation (vecteur multidimensionnel),

Σ_i : Matrice qui détermine l'étendue et l'orientation de la fonction d'activation dans l'espace des variables de régression,

$\|\varphi(t) - C_i\|_{\Sigma_i}$: distance du point courant $\phi(t)$ au centre de la fonction d'activation et exprimée par l'équation :

$$\|\varphi(t) - C_i\|_{\Sigma_i} = (\varphi(t) - C_i)^T \sum_i^{-1} (\varphi(t) - C_i) \quad \text{II} - 20$$

La fonction g est caractérisée par un support borné ; elle est uni modale, monotone, décroissante quand le point courant s'éloigne de son centre. Il existe plusieurs types de fonction de base mais on associe généralement aux RFBR la fonction gaussienne de centre C_i et de matrice de dispersion Σ_i .

A cause de la similitude des équations (II-17) et (II-20), on assimile les RFBR à des réseaux de neurones à une couche cachée à la différence qu'un neurone de la couche cachée ne comporte pas de sommateur et que la fonction d'activation $g(\cdot)$ détermine une zone d'influence à partir de son centre C_i . Les RBFR sont également des d'approximateurs universels ; ces propriétés ont été établies entre autres par Powell [1987], Park et Sandberg [1991].

Deux approches sont souvent considérées pour identifier un RFBR :

- ❖ Chaque point de l'espace multidimensionnel est considéré comme un centre potentiel. En fixant les paramètres de dispersion des fonctions d'activation, l'équation (I-19) est alors totalement linéaire par rapport aux poids. Une procédure de sélection de variables est mise en œuvre pour déterminer les centres utiles. Chen et al. [1990], Orr [1996] ont proposé par exemple une procédure basée sur les moindres carrés orthogonaux (MCO) : de nouvelles fonctions d'activation sont ajoutées successivement jusqu'à l'obtention d'une structure satisfaisante.
- ❖ Au lieu de fixer a priori les centres et les dispersions des fonctions d'activation, des chercheurs à l'instar de Karyiannis et Mi [1997], construisent progressivement le RFBR en divisant en deux parties, le domaine d'influence de la fonction de base où l'erreur d'approximation est la plus élevée. Des techniques comme l'algorithme de classification FCM (Fuzzy C-Means) servent à déterminer la position des centres qui ne coïncident plus alors avec les points du jeu d'identification. Une fois son centre fixé, la dispersion d'une fonction de base est déduite de la matrice de variance-covariance calculée à partir des K plus proches voisins de son centre. Les poids w_i sont ensuite optimisés par la méthode des moindres carrés. Le procédé est réitéré jusqu'à la satisfaction d'un critère d'arrêt.

II-6-LES MODELES MULTI-EXPERTS :

Une autre catégorie de modèles non-linéaires est constituée par les modèles multi-experts (mixture of experts) [Jacobs et al. 1991] qui ont fait leur apparition dans la littérature au début des années 1990. Cette approche est basée sur l'idée que les données analysées sont générées par un ensemble de M experts, un expert étant une fonction d'un vecteur d'entrée $\varphi(t)$ et d'un vecteur de paramètres θ_i . Un vecteur d'entrée $\varphi(t)$ est traité simultanément par tous les modules experts et chaque expert fournit une sortie y_i supposée corrompue par un bruit gaussien e_i de moyenne nulle et de variance σ_i :

$$y_i(t) = f(\varphi(t), \theta_i) + e_i(t) \quad i = 1, \dots, M \quad \text{II} - 21$$

Le champ d'action d'un expert i est déterminé par une fonction d'activation $g_i(\varphi(t))$ à support limité. Ces fonctions d'activation pondèrent les contributions de chaque expert à la sortie du modèle global ; elles vérifient les contraintes suivantes :

$$0 \leq g_i(\varphi(t)) \leq 1 \quad \sum_{i=1}^M g_i(\varphi(t)) = 1 \quad i = 1, \dots, M \quad \text{II} - 22$$

Dans la formulation de cette structure non-linéaire, la fonction $g_i(\varphi(t))$ s'interprète comme la probabilité que la sortie du système relève de l'expert i associé, étant donnée le vecteur d'entrée $\varphi(t)$. Initialement, les fonctions d'activation étaient des sigmoïdes normalisées. Plus tard a été proposée l'utilisation de fonctions de densité gaussiennes. La sortie globale générée par le modèle multi-experts est une somme pondérée des sorties fournies par chaque expert, soit :

$$y(t) = \sum_{i=1}^M g_i(\varphi(t)) y_i(t) \quad \text{II} - 23$$

A cause de l'interprétation probabiliste de cette approche, les paramètres mis en jeu pour décrire un modèle multi-experts sont estimés en optimisant un critère de maximum de vraisemblance par l'algorithme EM (Expectation - Minimisation) de Dempster [Meila et Jordan, 1997]. Cet algorithme comporte deux étapes :

- ❖ étape d'Espérance : on y évalue la probabilité qu'une donnée observée relève d'un module expert particulier,
- ❖ étape de Maximisation : les paramètres des fonctions d'activation et ceux des modèles experts sont calculés de manière à maximiser le critère de maximum de vraisemblance.

Il est nécessaire avant de procéder à l'optimisation paramétrique de déterminer la structure et le nombre d'experts. Pour ce dernier problème, Ramamurti et Ghosh [1999] ont proposé une méthode qui consiste à démarrer avec un expert puis à ajouter successivement de nouveaux experts. S'il n'y a plus lieu d'inclure un nouvel expert, on teste la possibilité d'éliminer certains experts afin d'annuler l'effet des éventuelles mauvaises décisions prises lors de la phase de croissance du modèle.

Une généralisation des modèles multi-experts est le modèle multi-experts de Markov (Markov mixtures of experts) [Meila et Jordan, 1997] qui modélise le système sous la forme d'une chaîne de Markov dont les états sont les experts du système. Le passage d'un état à un autre est lié à une probabilité de transition dépendant de l'état précédent et des entrées à l'instant courant.

II-7-AUTRES STRUCTURES NON-LINEAIRES :

Une structure non-linéaire issue de la communauté du traitement de signal est le réseau d'ondelettes. La décomposition sur une base d'ondelettes autorise l'étude simultanée des propriétés temporelles et fréquentielles d'un signal par dilatation et translation d'une ondelette mère. Des informations sur des références bibliographiques traitant de l'identification des systèmes sur la base d'un réseau d'ondelettes peuvent être trouvées dans [Sjöberg et al., 1995].

Une autre approche est la représentation du système par des "hinging hyperplanes". Initialement proposée par Breiman en 1993, son application à la modélisation des systèmes non- linéaires a été approfondie par Pucar en 1995 [Pucar, 1995]. Mathématiquement, un "hinging hyperplane" à deux hyperplans est définie par l'une des deux équations suivantes :

$$h(t) = \max(\varphi^T(t)\theta^+, \varphi^T(t)\theta^-) \quad \text{II - 24 - a}$$

$$h(t) = \min(\varphi^T(t)\theta^+, \varphi^T(t)\theta^-) \quad \text{II - 24 - b}$$

Afin d'éviter le passage brutal (switching) d'un hyperplan à l'autre, les opérateurs min et max sont généralement remplacés par une fonction sigmoïde.

II-8-ASPECTS PRATIQUES DE L'IDENTIFICATION DES SYSTEMES :

Pour un problème particulier, quel que soit le type de modèle choisi, il faut déterminer correctement sa structure et estimer de façon efficace les paramètres mis en jeu de manière à réaliser une bonne adéquation entre la sortie du modèle et celle du système. La démarche usuelle

consiste à fixer une structure, à estimer ses paramètres sur la base d'un jeu de données d'identification, puis passer à l'étape de validation. Cette étape est destinée à évaluer les

II-8-1-ESTIMATION DES PARAMETRES :

Soit \mathcal{S} une structure fixée. Elle contient forcément des paramètres qui sont regroupés dans un vecteur θ . La valeur de ces paramètres est estimée, à partir d'un ensemble de données expérimentales $\mathcal{D}_N = \{(y_s(t), \varphi(t))\}_{t=1}^N$, par minimisation d'un critère, fonctionnelle de l'écart entre la sortie du système $y_s(t)$ et celle du modèle $y(t)$

$$\hat{\theta} = \arg \min_{\theta} J(\theta, \mathcal{D}_N) \quad \text{II} - 25$$

$$J(\theta, \mathcal{D}_N) = \frac{1}{2} \sum_{t=1}^N \varepsilon^2(t, \theta) \quad \text{II} - 26 - a$$

$$\varepsilon(t) = y(t, \theta) - y_s(t) \quad \text{II} - 26 - b$$

La complexité de l'optimisation dépend de la structure du modèle.

II-8-1-1-ESTIMATEUR DES MOINDRES CARRES :

Supposons que le modèle du système est linéaire par rapport aux paramètres, c'est-à-dire de la forme :

$$y(t) = F[\varphi(t), \theta] = \Phi^T(\varphi(t))\theta$$

Où $\varphi(\cdot)$ est une fonction qui réalise une transformation non-linéaire du vecteur de régression. La solution au problème (II-25) est analytique et est fournie par l'estimateur des moindres carrés :

$$\theta = R^{-1}\Phi^{-1}Y_s \quad \text{II} - 27 - a$$

$$R = \Phi^T \Phi \quad \text{II} - 27 - b$$

$$Y_s = [y_s(1) \dots y_s(N)]^T \quad \Phi = [\phi(\varphi(1)) \dots \phi(\varphi(N))]^T \quad \text{II} - 27 - c$$

Le cas typique est celui des RFBR dont les centres et la dispersion sont fixés. Les paramètres restants interviennent alors de manière linéaire dans le modèle avec le vecteur $\phi(\varphi(t))$ formé par la concaténation des fonctions de base.

❖ PROBLEME DE CONDITIONNEMENT DE L'ESTIMATION

PARAMETRIQUE :

L'inversion de la matrice d'information R peut être sujette à des problèmes numériques liés à son mauvais conditionnement. Une matrice carrée est dite mal-conditionnée si le rapport entre sa plus grande valeur propre et sa plus petite valeur propre est élevé (supérieur par exemple à un seuil de 10^6). D'un point de vue numérique, l'estimation $\hat{\theta}$ n'est pas stable et est sensible à des perturbations même faibles affectant la matrice Φ ou le vecteur Y_s .

❖ SOLUTION : LA REGULARISATION :

La théorie de la régularisation est née des travaux de Tikhonov sur la résolution des problèmes mathématiques mal posés, c'est-à-dire, des problèmes pour lesquels il n'existe pas une solution unique ou la solution obtenue est instable et très sensible à de faibles perturbations.

Dans le cadre de l'identification des systèmes, les techniques de régularisation consistent à ajouter des pénalités sur le problème d'estimation des paramètres de façon à réduire l'extrême sensibilité du modèle par rapport aux données d'identification. Dans le cas qui nous intéresse, ces pénalités se traduisent par l'adjonction à la matrice d'information R d'une matrice λK qui lui confère un meilleur conditionnement. Ceci équivaut à minimiser le critère modifié suivant :

$$J_{reg}(\theta, \mathcal{D}_N) = J(\theta, \mathcal{D}_N) + \frac{\lambda}{2} \theta^T K \theta \quad \text{II - 28}$$

Le compromis entre la pénalisation $\theta^T K \theta$ et le critère $J(\theta, \mathcal{D}_N)$ est réglé par le paramètre scalaire positif λ appelé coefficient de régularisation. D'un point de vue bayésien, ce critère modifié correspond à l'estimateur du maximum a posteriori avec l'hypothèse a priori que les paramètres suivent une loi normale de moyenne nulle et de variance $(\lambda K)^{-1}$.

Toute la problématique de la régularisation réside d'une part dans la détermination du coefficient de régularisation et d'autre part dans le choix de la matrice K . Suivant le type de pénalités qu'on désire coder dans la matrice K , différentes façons de la choisir existent [Bossley 1997], [Johansen, 1996, 1997]. La matrice K peut être, par exemple, une approximation de la dérivée seconde $\partial^2 F[\varphi, \theta] / \partial \varphi \partial \varphi^T$ afin de conférer au modèle des propriétés de lissage. Mais la forme courante de régularisation est la « ridge regularisation » ou « weight decay » qui correspond au choix $K = I$. Son but est de limiter la norme du vecteur de paramètres dans des proportions raisonnables. Avec cette formulation, le vecteur de paramètres régularisé est fourni par :

$$\hat{\theta}_{reg} = (R + \lambda I)^{-1} \Phi^T Y_s \quad \text{II - 29 - a}$$

Ou de façon équivalent par la relation suivante :

$$\hat{\theta}_{reg} = \sum_{i=1}^{n_\theta} \frac{\xi_i}{\xi_i^2} \mathcal{V}_i \mathcal{U}_i^T Y_s \quad \text{II - 29 - b}$$

basée sur une décomposition en valeurs singulières de la matrice Φ . Les ξ_i sont les valeurs singulières, les \mathcal{V}_i les vecteurs propres associés aux valeurs propres ξ_i^2 et les \mathcal{U}_i les vecteurs propres de la matrice $\Phi \Phi^T$. On constate d'après (I-29-b) que le coefficient de régularisation élimine l'influence des plus petites valeurs propres qui sont responsables de l'instabilité de la solution des moindres carrés. Pour une valeur propre $\xi_i^2 \ll \lambda$, l'axe de recherche suggéré par le vecteur propre associé n'est pas fiable : il a peu d'influence sur le critère et son effet est donc annihilé par λ . Par contre si $\xi_i^2 \gg \lambda$, la direction de recherche déterminée par cette valeur propre est importante : elle est donc peu affectée par le coefficient de régularisation.

❖ DETERMINATION DU PARAMETRE DE REGULARISATION :

Le choix du coefficient de régularisation est délicat : s'il est trop grand $\xi_i^2 \ll \lambda \quad \forall i$ le vecteur de paramètres tendra vers $\mathbf{0}$ et s'il est trop petit $\xi_i^2 \gg \lambda \quad \forall i$ l'effet stabilisant de la régularisation est inopérant. Comme on désire améliorer le conditionnement de la matrice R , le coefficient λ est choisi de sorte que le conditionnement $\tau_{R+\lambda I}$ de la matrice d'information régularisée $R + \lambda I$:

$$\tau_{R+\lambda I} = \frac{\xi_{max}^2 + \lambda}{\xi_{min}^2 + \lambda}$$

Soit inférieur à un seuil limite τ_{lim} . Si le conditionnement actuel de R dépasse cette valeur, on calcule λ par la formule :

$$\lambda = \frac{\xi_{max}^2 - \tau_{lim} \xi_{min}^2}{\tau_{lim} - 1} \simeq \frac{\xi_{max}^2}{\tau_{lim}} - \xi_{min}^2 \quad \text{II - 30}$$

Autrement, le problème n'est pas régularisé. Dans nos simulations, nous calculons le coefficient de régularisation de manière à garantir un taux de conditionnement entre 10^5 et 10^6 . Des valeurs

similaires ont été utilisées aussi par [Fiordaliso, 1999]. Elles sont certes arbitraires mais dans nos simulations, elles se sont révélées suffisantes pour obtenir des résultats satisfaisants.

II-8-1-2-OPTIMISATION NON-LINEAIRE :

Lorsque le modèle du système est non-linéaire par rapport aux paramètres, il n'existe pas de solution analytique au problème (II-25). On recourt à des techniques itératives d'optimisation non-linéaire : partant d'une estimation courante $\hat{\theta}_{(k)}$ des paramètres, on recherche la direction dans l'espace paramétrique et le déplacement à effectuer suivant cette direction afin de diminuer le critère. Il existe une large variété de techniques mais dans le cadre de ce document, nous privilégions celles basées sur un développement limité du critère $J(\theta, \mathcal{D}_N)$, au voisinage du point $\hat{\theta}_{(k)}$ (pourvu que le critère soit différentiable par rapport à θ , ce que nous supposons tout le long de ce mémoire). La mise à jour du vecteur des paramètres est alors dérivée de la formule générale suivante :

$$\hat{\theta}_{(k+1)} = \hat{\theta}_k - \eta_k D_k \quad \text{II - 31}$$

η_k est le pas de recherche ou le coefficient de relaxation (suivant les cas) à l'itération (k) et $D_{(k)}$ la direction de recherche dans l'espace paramétrique. Selon la façon dont $D_{(k)}$ est calculée, on distingue différentes méthodes d'optimisation dont les principales sont rappelées ci-dessous [Walter et Pronzato, 1994].

❖ ALGORITHME DU GRADIENT :

Cette méthode est basée sur un développement du critère au 1^{ière} ordre. La direction de recherche à l'itération (k) est spécifiée par le gradient du critère $G_{(k)}$:

$$G(\theta_{(k)}) = \frac{\partial J}{\partial \theta} \Big|_{\theta=\hat{\theta}_{(k)}} = \sum_{t=1}^N \frac{\partial \varepsilon(t, \theta)}{\partial \theta} \varepsilon(t, \theta) \Big|_{\theta=\hat{\theta}_{(k)}} \quad \text{II - 32}$$

Le pas de recherche η est calculé par une interpolation quadratique ou cubique du critère autour du point $\hat{\theta}_{(k)}$ ou par une heuristique consistant à augmenter η si le critère décroît et à le réduire si le critère augmente.

❖ ALGORITHME DE NEWTON :

Il repose sur un développement au 2^{ième} ordre du critère. La direction et le pas de recherche sont spécifiés simultanément par l'équation :

$$D_{(k)} = H_{(k)}^{-1} G_{(k)}$$

Où $H_{(k)}$ est la matrice Hessienne du critère définie par :

$$H_{(k)} = \sum_{t=1}^N \frac{\partial \varepsilon(t, \theta)}{\partial \theta} \frac{\partial \varepsilon(t, \theta)}{\partial \theta^T} + \sum_{t=1}^N \frac{\partial^2 \varepsilon(t, \theta)}{\partial \theta^2} \varepsilon(t, \theta) |_{\theta = \hat{\theta}_{(k)}} \quad \text{II - 33}$$

Le pas $\eta_{(k)} = 1$. Cet algorithme nécessite à chaque itération l'inversion du hessien du critère, ce qui est coûteux. De plus le hessien contient des dérivées du 2^{ième} ordre qui sont lourdes à calculer.

Par ailleurs, l'algorithme de Newton peut converger indifféremment vers un minimum ou un maximum car rien ne garantit que le Hessien $H_{(k)}$ est défini positif à l'optimum. Pour éviter une divergence de l'algorithme, on utilise η comme un coefficient de relaxation afin d'assurer un déplacement dans l'espace paramétrique qui conduise à une diminution du critère. Le coefficient de relaxation est déterminé suivant l'heuristique guidant le choix du pas dans l'algorithme du gradient.

❖ ALGORITHME DE GAUSS-NEWTON :

Simplification de la méthode de Newton, il utilise une expression approchée du Hessien en négligeant les termes du 2^{ième} ordre :

$$H_a = \sum_{t=1}^N \frac{\partial \varepsilon(t, \theta)}{\partial \theta} \frac{\partial \varepsilon(t, \theta)}{\partial \theta^T} \quad \text{II - 34}$$

Le Hessien approché H_a étant défini positif, cet algorithme garantit la convergence vers un minimum. La procédure de calcul du coefficient de relaxation η reste identique.

❖ ALGORITHME DE LEVENBERG-MARQUARDT :

De façon similaire à l'estimation des moindres carrés, le Hessien peut être mal conditionné. Pour éviter une singularité de la matrice H_a , l'algorithme de **Levenberg-Marquardt** remplace le Hessien approché par un Hessien régularisé

$$H_R = H_a + \lambda_{(k)} I \quad \text{II - 35}$$

Où I est la matrice identité de dimension appropriée et $\lambda_{(k)}$ un coefficient de régularisation. $\lambda_{(k)}$ est mis à jour par un schéma heuristique qui consiste à augmenter sa valeur si le critère diminue et à

réduire sa valeur s'il y a divergence du critère. Pour de faibles valeurs de $\lambda_{(k)}$, l'algorithme de **Levenberg-Marquardt** se réduit à celui de **Gauss-Newton** et pour de fortes valeurs de $\lambda_{(k)}$, il devient une méthode du gradient avec un pas

$$\eta_{(k)} = 1/\lambda_{(k)}.$$

❖ ALGORITHME DE QUASI-NEWTON :

Comparable à la méthode de **Newton** ou **Gauss-Newton** à la différence près que l'inverse de la matrice Hessienne n'est pas calculée directement mais déduite de la relation de mise à jour suivante

$$H_{(k)} = H_{(k-1)} + C_{(k-1)} \quad \text{II - 36}$$

où $C_{(k-1)}$ est une matrice de correction calculée à partir de $H_{(k-1)}$ et des informations sur la variation du gradient et des paramètres au point courant $\hat{\theta}_{(k)}$. Au début, l'algorithme se comporte comme celui du gradient, puis se rapproche de celui de Gauss-Newton au fur et à mesure que l'estimation de l'inverse du Hessien devient précise.

II-8-2-RECHERCHE DE LA STRUCTURE DU MODELE :

Etant choisie une famille de modèle non-linéaire (RFBR, MLP, modèle flou, ...), la recherche des paramètres structuraux d'un modèle de cette famille est guidée par deux soucis : le modèle élaboré doit être simple (c'est-à-dire comporter le moins de paramètres possibles) et précis afin d'avoir de bonnes propriétés de généralisation. Par généralisation, nous entendons de bonnes capacités d'approximation du modèle quand il est appliqué sur des données autres que celles qui ont servi à son identification. Malheureusement, ces deux souhaits sont antagonistes comme le montre le développement suivant.

❖ COMPROMIS BIAIS/VARIANCE :

Considérons que la qualité du modèle élaboré sur un jeu d'identification \mathcal{D}_N est mesurée par ses performances sur des jeux de validation \mathcal{D}_{N_v} comportant N_v données. L'erreur quadratique moyenne de généralisation (MSGE) du modèle $F[\varphi(t), \hat{\theta}_{n_\theta}]$ comportant n_θ paramètres est exprimés par la formule [Larsen et Hansen, 1994] :

$$\begin{aligned}
 MSGE &= Esp \left\{ \lim_{N_v \rightarrow \infty} \frac{1}{N_v} \sum_{t=1}^{N_v} \left(y_s(t) - F[\varphi(t), \hat{\theta}_{n_\theta}] \right)^2 \right\}_{\mathcal{D}_{N_v}} \\
 &= Esp \left\{ Esp \left(y_s(t) - F[\varphi(t), \hat{\theta}_{n_\theta}] \right)^2 \right\}_{\mathcal{D}_{N_v}}
 \end{aligned}
 \tag{II - 37}$$

L'objectif de l'identification est de déterminer à partir du jeu de données \mathcal{D}_N , un modèle qui fournit de bonnes performances sur tous les jeux de validation possibles possédant les mêmes propriétés statistiques que les données qui ont servi à l'élaboration du modèle. Le meilleur modèle possible) $F^*[\varphi(t), \theta_{n_\theta}^*]$ avec n_θ paramètres est celui qui minimise le critère MSGE.

Supposons que les données analysées sont générées par le "vrai" modèle suivant qui est évidemment inconnu :

$$y_s(t) = F_0(t) + e(t)$$

Où $e(t)$ est une séquence aléatoire de moyenne nulle et de variance σ^2 . On montre que le MSGE se décompose de la façon suivante :

$$\begin{aligned}
 MSGE &= \sigma^2 + Esp \left\{ Esp \left[F_0(t) - F^*[\varphi(t), \hat{\theta}_{n_\theta}^*] \right]^2 \right\}_{\mathcal{D}_{N_v}} \\
 &\quad + Esp \left\{ Esp \left[F[\varphi(t), \hat{\theta}_{n_\theta}] - F^*[\varphi(t), \hat{\theta}_{n_\theta}^*] \right]^2 \right\}_{\mathcal{D}_{N_v}}
 \end{aligned}
 \tag{II - 38}$$

Le terme de biais mesure l'inadéquation entre le "vrai" modèle et le meilleur modèle F^* . C'est un indicateur des limites de la structure de modèle choisie. Le terme de variance donne des indications sur la sensibilité du modèle par rapport aux différents jeux de validation. Il est directement lié à la variance de l'estimation $\hat{\theta}_{n_\theta}$ c'est-à-dire à la distance entre $\hat{\theta}_{n_\theta}$ et $\theta_{n_\theta}^*$. Ce terme de variance est d'autant plus important que le modèle comporte beaucoup de paramètres. En effet, l'expression asymptotique du terme de variance est donnée par [Ljung, 1987] :

$$Variance \approx \sigma^2 \frac{n_\theta}{N} \tag{II - 39}$$

La minimisation du biais implique l'augmentation de la flexibilité du modèle F . La flexibilité ou la complexité du modèle est déterminée par le nombre de paramètres (nombre de degrés de liberté) dans le modèle. Plus F comportera de paramètres, plus le biais sera réduit. La conséquence est

l'augmentation du terme de variance dans le MSGE d'après (I-39) : c'est le dilemme biais/variance [Ljung, 1987], [Geman et al., 1992] qui impose de trouver un compromis entre la précision du modèle et sa flexibilité. La meilleure structure F est celle qui minimise le MSGE.

❖ APPROXIMATIONS DU MSGE :

Dans la pratique, le calcul direct du MSGE n'est pas possible. D'autres critères ont été proposés dans la littérature comme outils d'aide à la sélection de la structure d'un modèle.

La façon directe d'approcher le MSGE est d'évaluer la structure identifiée sur des données de test : c'est la validation croisée qui nécessite au moins un deuxième jeu de validation.

En vue de réduire la sensibilité du modèle par rapport aux données, une démarche consiste à diviser les données disponibles en V ensemble \mathcal{D}_j de taille identique et ayant, si possible, les même caractéristique. A chaque fois, un ensemble \mathcal{D}_j différent est utilisé comme ensemble de validation alors que les données des $V - 1$ ensembles restants servent à l'identification d'un modèle. V modèles F_j sont ainsi construits et le critère de généralisation est approché par :

$$MSGE \approx \frac{1}{V} \sum_{j=1}^V \sum_{t=1}^{N_j} (y_s(t) - F_j(t))^2 \quad \text{II - 40}$$

Dans certaines applications, la rareté des données contraint à évaluer la structure choisie par une approximation du MSGE calculée sur les données d'identification. Ceci a conduit à des critères de sélection qui essaient de trouver un compromis entre la précision du modèle et sa complexité.

Parmi ces critères, on peut citer :

$$AIC = N \log(J(\theta, \mathcal{D}_N)) + 2n_\theta \quad \text{II - 41}$$

$$FPE = J(\theta, \mathcal{D}_N) \frac{N + n_\theta}{N - n_\theta} \quad \text{II - 42}$$

$$AIC = \log(J(\theta, \mathcal{D}_N)) + n_\theta \frac{\log(N)}{N} \quad \text{II - 43}$$

II-8-3-VALIDATION DU MODELE :

La procédure de sélection de structure permet de retenir le meilleur modèle dans une famille donnée de modèles, en termes de compromis entre la précision et la complexité. Mais cette

procédure ne garantit pas que le modèle trouvé est adéquat eu égard aux données traitées ou à l'utilisation envisagée du modèle. Une nécessaire étape de validation doit être réalisée.

Une méthode directe de validation d'un modèle est de tester ses capacités de généralisation en validant le modèle sur des données de test c'est-à-dire des données entièrement différentes de celles ayant servi à l'estimation des paramètres ou lors de la phase de recherche de structure. En particulier, pour les modèles destinés à la simulation, cette validation se traduit par la simulation du modèle sur les données de test.

Une autre technique de validation est le test de corrélation des résidus (écarts ε entre la sortie du système et celle du modèle) largement appliqué dans l'identification des systèmes linéaires. Le modèle linéaire élaboré est valide si la fonction d'autocorrélation des résidus $\hat{r}_{\varepsilon\varepsilon}$ et si la fonction d'intercorrélation $\hat{r}_{\varepsilon u}$ entre les résidus et les entrées se trouvent dans l'intervalle de confiance à 95% défini par $\pm 1,96\sqrt{N}$ (voir [Ljung, 1987]). La fonction d'intercorrélation $\hat{r}_{x_1x_2}(\tau)$ entre deux variables x_1 et x_2 se calcule par la formule :

$$\hat{r}_{x_1x_2}(\tau) = \frac{\sum_{t=1}^{N-\tau} (x_1(t) - \bar{x}_1)(x_1(t) - \bar{x}_1)}{\sqrt{\sum_{t=1}^N (x_1(t) - \bar{x}_1)^2} \sqrt{\sum_{t=1}^N (x_2(t) - \bar{x}_2)^2}} \quad \text{II - 44}$$

Le test des résidus peut être étendu aux systèmes non-linéaires [Leontaritis et Billings, 1987] : pour examiner si des effets non-linéaires n'ont pas été modélisés, on détermine si la fonction d'intercorrélation entre les résidus et des monômes formés à partir des éléments du vecteur de régression est en dehors de l'intervalle de confiance à 95%.

II-9-CONCLUSION :

L'objectif de chapitre est de dresser un état de l'art de l'identification des systèmes dynamiques non-linéaires. Les principales structures non-linéaires couramment rencontrées dans la littérature ont été présentées. Elles relèvent globalement de deux visions de représentation des systèmes : l'approche globale et l'approche locale. Toutes les structures non-linéaires décrites sont essentiellement des modèles boîte noire et leurs propriétés d'approximateurs universels sont bien établies. Elles peuvent donc approcher une large gamme de systèmes dynamiques non-linéaires.

Indifféremment du type de modèle non-linéaire choisi, son identification requiert la recherche de sa structure optimale, l'estimation de ses paramètres et la validation du modèle final afin d'identifier un modèle parcimonieux (c'est-à-dire un modèle précis et comportant un faible nombre de paramètres). La deuxième partie du chapitre a été consacrée à la présentation de ces principes généraux de l'identification des systèmes. La forme particulière de leur mise en œuvre dépendra du modèle non-linéaire choisi.

Sommaire

II-1- Introduction :	39
II-2- Modeles Classiques Des Systemes Non-Lineaires :	40
II-3- Approches Recentes De Representation Des Systemes Non-Lineaires :	41
II-3-1- Modeles Flous :	42
II-3-1-1- Modele De Mamdani :	45
II-3-1-2- Modele De Takagi-Sugeno (Modele Ts):	47
II-3-2- Proprietes Des Modeles Flous :	49
II-3-3- Problematique D'identification Des Modeles Flous :	49
II-4- Reseaux De Neurones :	50
II-5- Les Reseaux De Fonctions A Base Radiale (Rfbr) :	52
II-6- Les Modeles Multi-Experts :	54
II-7- Autres Structures Non-Lineaires :	55
II-8- Aspects Pratiques De L'identification Des Systemes :	55
II-8-1- Estimation Des Parametres :	56
II-8-1-1- Estimateur Des Moindres Carres :	56
II-8-1-2- Optimisation Non-Lineaire :	59
II-8-2- Recherche De La Structure Du Modele :	61
II-8-3- Validation Du Modele :	63
II-9- Conclusion :	65

LISTE DES FIGURES :

FIGURE II- 2 : Modele De Weiner	41
FIGURE II- 3 : Structure Generale D'un Modele Flou	44
FIGURE II- 4 : Illustration De La Difference Entre Les Degres De Veracite Des Regles (Partie Gauche) Et Les Degres D'activation (Partie Droite) Des Modeles Locaux Associes.	48
FIGURE II- 5: Structure Generale D'un Reseau Mlp	51
FIGURE II- 6: Structure D'un Neurone D'une Couche Cachee	51

LISTE DES TABLEAUX

TABLEAU II- 1 : Fonctions D'appartenance Les Plus Utilisees En Modelisation Floue.....	43
TABLEAU II- 2:Tableau Des Operateurs Logiques Et Des Principaux Operateurs Algebriques Correspondants.....	46

Chapitre III

Approche multi-modèle



Chapitre III

Approche multi-modèle



To be sure, mathematics can be extended to any branch of knowledge, including economics, provided the concepts are so clearly defined as to permit accurate symbolic representation.

That is only another way of saying that in some branches of discourse it is desirable to know what you are talking about.

—James R. Newman

III-1-INTRODUCTION :

Désignée encore sous la dénomination de "operating regime approach", l'approche multi-modèle a connu un intérêt certain depuis la publication des travaux de Johansen et Foss en 1992 [Johansen et Foss, 1992]. En réalité, les travaux de ces deux chercheurs ont concerné la formalisation mathématique d'un concept qui existait sous différents noms dans plusieurs domaines. L'idée de cette approche est d'appréhender le comportement non-linéaire d'un système par un ensemble de modèles locaux (généralement de structure simple) caractérisant le fonctionnement du système dans différentes zones de fonctionnement. Le modèle global du système est une combinaison des modèles locaux. La motivation de cette décomposition découle du constat qu'il est souvent difficile d'élaborer un modèle global susceptible de rendre compte de toutes les particularités et de toute la complexité d'un système. Cette approche intuitive de la modélisation a été appliquée par exemple par Tong [Fiordaliso, 1999] qui a proposé des modèles autorégressifs à seuils (modèles TAR) pour la modélisation de séries temporelles.

Différents modèles de la série sont construits et le passage d'un modèle à un autre est déclenché par une fonction de type échelon (fonction de Heaviside). Cette fonction est indexée sur les états antérieurs de la série temporelle. Dans le même ordre d'idée, Strömberg et al. [1991] ont proposé de représenter des systèmes non-linéaires avec des modèles linéaires par morceaux construits à partir d'un arbre de décision binaire. Le résultat de ces modèles de commutation est une approximation discontinue du système. Mais ces discontinuités peuvent être indésirables dans certaines

applications. Il est plus intéressant d'assurer un passage progressif d'un modèle à l'autre au lieu d'une commutation brutale. Pour ce faire, on substitue aux fonctions de commutation à front raide des fonctions à pente douce. Les zones de validité des modèles ne sont plus des partitions disjointes totalement isolées les unes des autres mais on autorise un recouvrement entre elles. Les fonctions de commutation sont alors des fonctions à dérivée continue dont la pente détermine la vitesse de transition d'un modèle à l'autre. L'intégration de ce principe aux modèles TAR a engendré les modèles autorégressifs à seuils doux (ou modèles STAR) en 1986 (voir [Fiordaliso, 1999]). Indépendamment, à la même époque, Takagi et Sugeno [1985] ont présenté leur modèle flou dont le point essentiel est la représentation du système étudié par un ensemble de règles "si prémisse alors conséquence". La conséquence d'une règle est un modèle local affine et le modèle global s'obtient par agrégation des modèles locaux. Quelques années plus tard, Jacob et al. [1991] ont présenté les modèles multi-experts dont la philosophie repose sur la combinaison de différents experts (un expert est un modèle local décrivant le comportement du système dans une zone limitée) par l'entremise de fonctions d'activation.

Hathaway et Bezdek [1993] ont développé les Fuzzy c -Regression Models (FCRM) comme un problème de classification consistant à identifier c classes floues dont les prototypes sont des modèles linéaires. La résultante de toutes ces techniques est un modèle global du système qui est une combinaison de modèles localement valables. Cette notion a été reprise dans un cadre mathématique général par Johansen et Foss [1992], ce qui a conduit à l'approche multi-modèle de représentation des systèmes non-linéaires.

III-2-FORMULATION MATHÉMATIQUE :

Considérons le problème de représentation d'un système non-linéaire dynamique par le modèle entrée-sortie général suivant :

$$y(t) = F[\varphi(t)] \quad \text{III} - 1$$

Où $\varphi(t)$ est le vecteur de régression. Supposons qu'on dispose d'un ensemble de M Modèles locaux $f_i(\varphi(t))$ descriptif du comportement du système dans différentes zones de fonctionnement. Ces modèles peuvent être construits par exemple à partir de connaissances physiques sur le fonctionnement du système dans ces zones.

La validité locale de chaque modèle f_i est indiquée par une **fonction de validité** $\rho_i(\varphi(t), \beta_i)$ paramétrée par le vecteur β_i et tel que $\rho_i(\varphi(t), \beta_i) \geq 0, \forall i$.

Cette fonction est significativement différente de 0 si le point courant $\varphi(t)$ est proche du "centre" φ_i de la zone et décroît plus ou moins rapidement vers 0 dès qu'on s'en éloigne. Pour un nombre suffisant M de modèles locaux judicieusement positionnés dans l'espace de régression, on peut exprimer le modèle global recherché $F[\varphi(t)]$ comme la combinaison de ces modèles locaux, plus précisément comme le barycentre des modèles locaux pondérés par leur fonction de validité.

Sous la condition :

$$\sum_{i=1}^M \rho_i(\varphi(t), \beta_i) > 0 \quad \forall \varphi(t) \quad \text{III - 2}$$

On déduit le modèle global :

$$y(t) = \frac{\sum_{i=1}^M \rho_i(\varphi(t), \beta_i) f_i(\varphi(t))}{\sum_{j=1}^M \rho_j(\varphi(t), \beta_j)} \quad \text{III - 3}$$

En posant :

$$\omega_i = \frac{\sum_{i=1}^M \rho_i(\varphi(t), \beta_i)}{\sum_{j=1}^M \rho_j(\varphi(t), \beta_j)} \quad \text{III - 4}$$

On obtient l'expression générale d'une architecture multi-modèle [Johansen et Foss, 1993]

$$y(t) = \sum_{i=1}^M \omega_i(\varphi(t), \beta) f_i(\varphi(t)) \quad \text{III - 5}$$

La fonction de pondération $\omega_i(\varphi(t))$ détermine le degré d'activation du modèle local associé. Selon la zone où évolue le système, cette fonction indique la contribution plus ou moins importante du modèle local correspondant dans le modèle global. Elle assure un passage progressif et doux de ce modèle aux modèles voisins. Nous désignerons la fonction ω_i également par **fonction d'activation ou fonction d'interpolation**. La zone délimitée par ω_i dans l'espace de régression est appelée **zone d'interpolation**. Les fonctions ω_i dépendent du vecteur global $\beta = [\beta_1^T, \beta_2^T \dots \beta_M^T]^T$ formé par la concaténation des vecteurs de paramètres β_i des fonctions de validité ρ_i . Remarquons que les fonctions d'activation sont normalisées au sens où elles vérifient la contrainte :

$$\sum_{i=1}^M \omega_i(\varphi(t), \beta) = 1 \quad \forall \varphi(t) \quad \text{III} - 6$$

On parle alors de partition unité [Murray-Smith, 1994].

La définition de l'approche multi-modèle introduit la notion de zone de fonctionnement qu'il est utile de préciser pour clarifier la présentation. En automatique classique, la zone de fonctionnement d'un système désigne un sous-espace de l'espace d'état "centré" autour d'un point, généralement un point stationnaire (ou point d'équilibre) du système. Dans le cadre de l'approche multi-modèle, une zone de fonctionnement doit être entendue comme une région de l'espace de fonctionnement où le système étudié est représenté par un modèle local donné. Le point de fonctionnement afférent à cette zone n'est pas obligatoirement un point d'équilibre du système. Les limites d'une zone de fonctionnement sont définies par la **fonction de validité** associée. D'après Johansen et Foss [1993], une fonction de validité typique est une fonction gaussienne dont la dispersion détermine l'étendue de la zone.

III-3-STRUCTURE DES MODELES LOCAUX :

Il existe une large variété de choix pour définir les modèles locaux associés aux zones de fonctionnement. La complexité de leur structure (modèle constant, modèle linéaire ou non-linéaire) est laissée au libre arbitre de l'utilisateur. Néanmoins dans la pratique, des modèles locaux de structure simple, principalement des modèles linéaires ou affines, sont privilégiés afin de pouvoir appliquer les techniques d'analyse de l'automatique linéaire aux modèles locaux (analyse des pôles et des zéros, gain statique, ...). Différentes représentations des modèles locaux sont envisageables ; certaines sont décrites dans les sous-sections qui suivent.

En considérant le choix des modèles locaux suggéré par Johansen et Foss [1993] , on définit les modèles locaux f_i comme les p premiers termes du développement en série de Taylor du vrai modèle $F_{0[\varphi(t)]}$ autour des points φ_i . En se limitant à un ordre $p = 1$, on obtient :

$$f_i(\varphi(t)) = F_0(\varphi_i) - (\varphi(t) - \varphi_i)^T \cdot \frac{\partial F_0[\varphi]}{\partial \varphi} \Big|_{\varphi=\varphi_i} \quad \text{III} - 7$$

Cette expression s'arrange sous la forme :

$$f_i(\varphi(t)) = \varphi^T(t) \cdot \theta_{i1} + \theta_{i0} \quad \text{III} - 8$$

Où :

$$\theta_{i1} = \frac{\partial F_0}{\partial \varphi} \Big|_{\varphi=\varphi_i} \quad \theta_{i0} = F_0(\varphi_i) - \varphi_i^T \theta_{i1} \quad \text{III} - 9$$

La forme affine (III-8) du modèle local f_i ($i=1, \dots, M$) s'interprète comme la linéarisation de la fonction multi variable $F_0[\varphi(t)]$ autour de φ_i . Si on injecte cette expression dans l'équation (III-5), la sortie du multi-modèle devient :

$$y(t) = \sum_{i=1}^M \omega_i(z(t), \beta) (\varphi^T(t) \cdot \theta_{i1} + \theta_{i0}) \quad \text{III} - 10$$

Elle peut également s'écrire :

$$y(t) = \sum_{i=1}^M \varphi^T(t) [\omega_i(z(t), \beta) \theta_{i1}] + [\omega_i(z(t), \beta) \theta_{i0}] = \varphi^T(t) \theta_1(t) + \theta_0(t) \quad \text{III} - 11$$

L'interprétation directe de cette relation est qu'une architecture multi-modèle avec des modèles locaux affines est un modèle affine à paramètres variables au cours du temps. La variation des paramètres est fonction de la zone de fonctionnement dans laquelle évolue le système. Il est alors possible d'associer à l'espace caractéristique \mathbf{Z} , un polytope de l'espace des paramètres décrivant le lieu d'évolution des paramètres du modèle affine (III-11) [Babuška, 1998].

Le modèle local de l'équation (III-8) s'écrit sous la forme compacte suivante :

$$f_i(\varphi(t), \theta_i) = \varphi_a^T(t) \theta_i \quad \text{III} - 12$$

Où θ_i est le vecteur de paramètres local formé par la concaténation de θ_{i1} et θ_{i0} . $\varphi_a(t)$ est le vecteur de régression augmenté défini par :

$$\varphi_a(t) = [\varphi^T(t) \ 1]^T \quad \text{III} - 13$$

La forme des modèles locaux étant définie, intéressons-nous maintenant aux éléments du vecteur de régression car suivant la constitution de ce vecteur, on peut distinguer des formes particulières de multi-modèle. Trois cas seront analysés étant entendu que d'autres types de vecteur de régression sont applicables. Pour la simplicité de la présentation, on supposera que tous les modèles locaux possèdent les mêmes paramètres structuraux (entrées, ordres, retards).

❖ CAS 1 : MULTI-MODELE SERIE-PARALLELE : MULTI-MODELE A ERREUR D'EQUATION :

Le vecteur de régression est formé par la sortie et les entrées du système décalées dans le temps.

Pour un système à m entrées, il a la forme suivante :

$$\varphi(t) = [-y_s(t-1) \cdots -y_s(t-n_y) u_1(t-nk_{u_1}) \cdots u_1(t-nk_{u_1}-n_{u_1}) \cdots u_m(t-nk_{u_m}) \cdots u_m(t-nk_{u_m}-n_{u_m})] \quad \text{III - 14}$$

Dans cette équation, y_s représente la sortie du système, n_y désigne l'ordre de la partie autorégressive, n_{u_j} et nk_{u_j} désignent respectivement l'ordre et le retard associés à l'entrée u_j ($j=1, \dots, m$). En adoptant une écriture polynomiale, la sortie du modèle local $f_i(\varphi(t), \theta_i)$ explicitée par $y_i(t)$ s'exprime :

$$y_i(t) = -A_i(q)y_s(t) + \sum_{j=1}^m B_{ij}(q)u_j(t-nk_{u_j}) + \theta_{i0} \quad \text{III - 15}$$

Où :

$$A_i(q) = \sum_{k=1}^{n_y} a_k^{(i)} q^{-k} \quad \text{III - 16 - a}$$

$$B_{ij}(q) = \sum_{k=0}^{n_{u_j}-1} b_{jk}^{(i)} q^{-k} \quad \text{III - 16 - b}$$

q^{-1} désigne l'opérateur retard. Les coefficients des polynômes A_i et B_{ij} et θ_{i0} sont les paramètres des modèles locaux à estimer. La structure multi-modèle obtenue pour un système SISO est schématisée sur la figure III-1. On constate aisément qu'elle est de type série-parallèle (ou erreur d'équation) car faisant intervenir la sortie du processus décalée dans le temps.

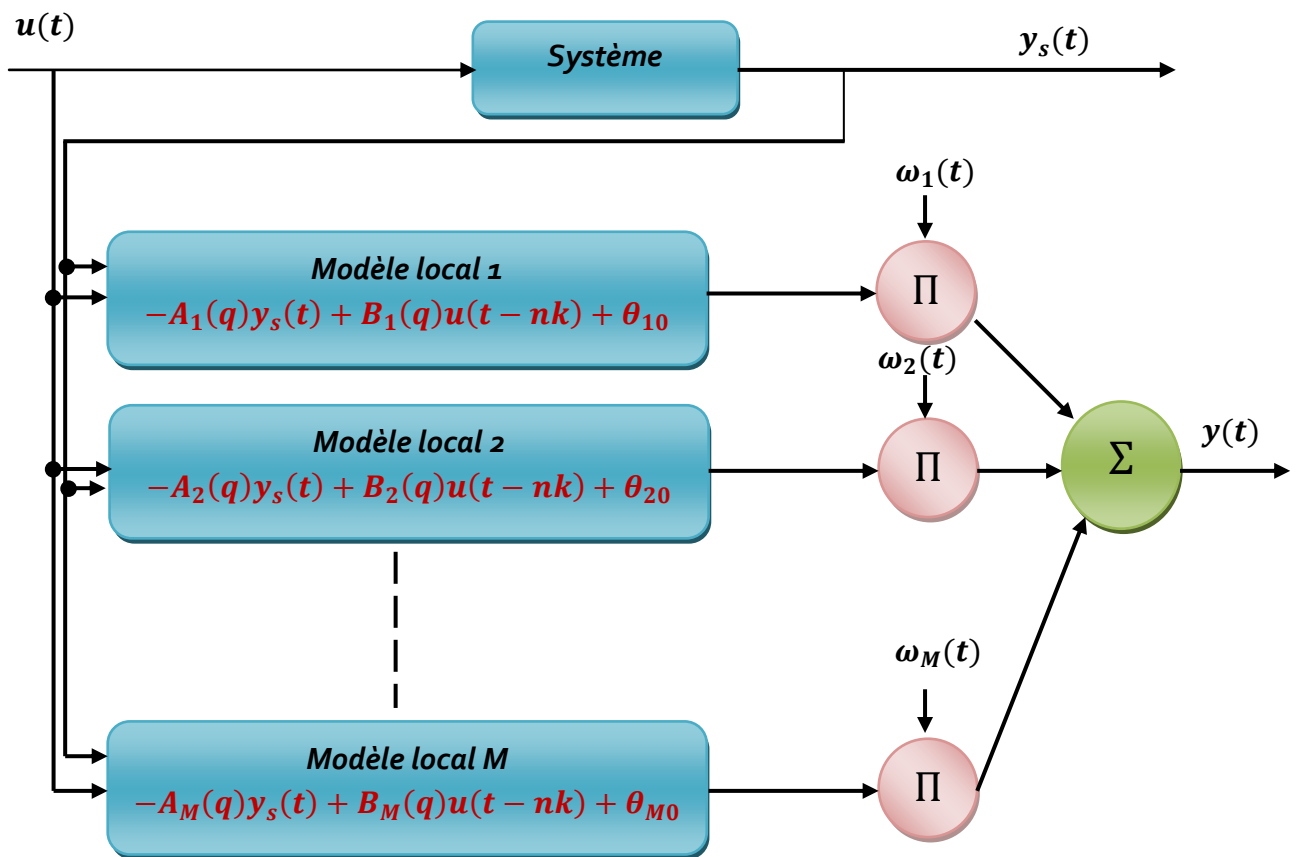


Figure III- 1: Multi-modèle série-parallèle avec des modèles locaux couplés

❖ CAS 2 : MULTI-MODELE PARALLELE AVEC DES MODELES LOCAUX COUPLES : MULTI-MODELE DE TYPE ERREUR DE SORTIE (OE) :

Le vecteur de régression contient non plus la sortie du système mais la sortie du multi-modèle décalée dans le temps :

$$\varphi(t) = [-y(t - 1) \dots - y(t - n_y) u_1(t - nk_{u1}) \dots u_1(t - nk_{u1} - n_{u1}) \dots u_m(t - nk_{um}) \dots u_1(t - nk_{um} - n_{um})]^T \tag{III - 17}$$

Le modèle local $f_i(\varphi(t), \theta_i)$ s'écrit sous la forme polynomiale suivante :

$$y_i(t) = -A_i(q)y(t) + \sum_{j=1}^m B_{ij}(q)u_j(t - nk_{uj}) + \theta_{i0} \tag{III - 18}$$

Avec les définitions précédentes des polynômes $A_i(q)$ et $B_{ij}(q)$. On obtient une structure multi-modèle parallèle avec le système. Son schéma est présenté sur la figure III-2.

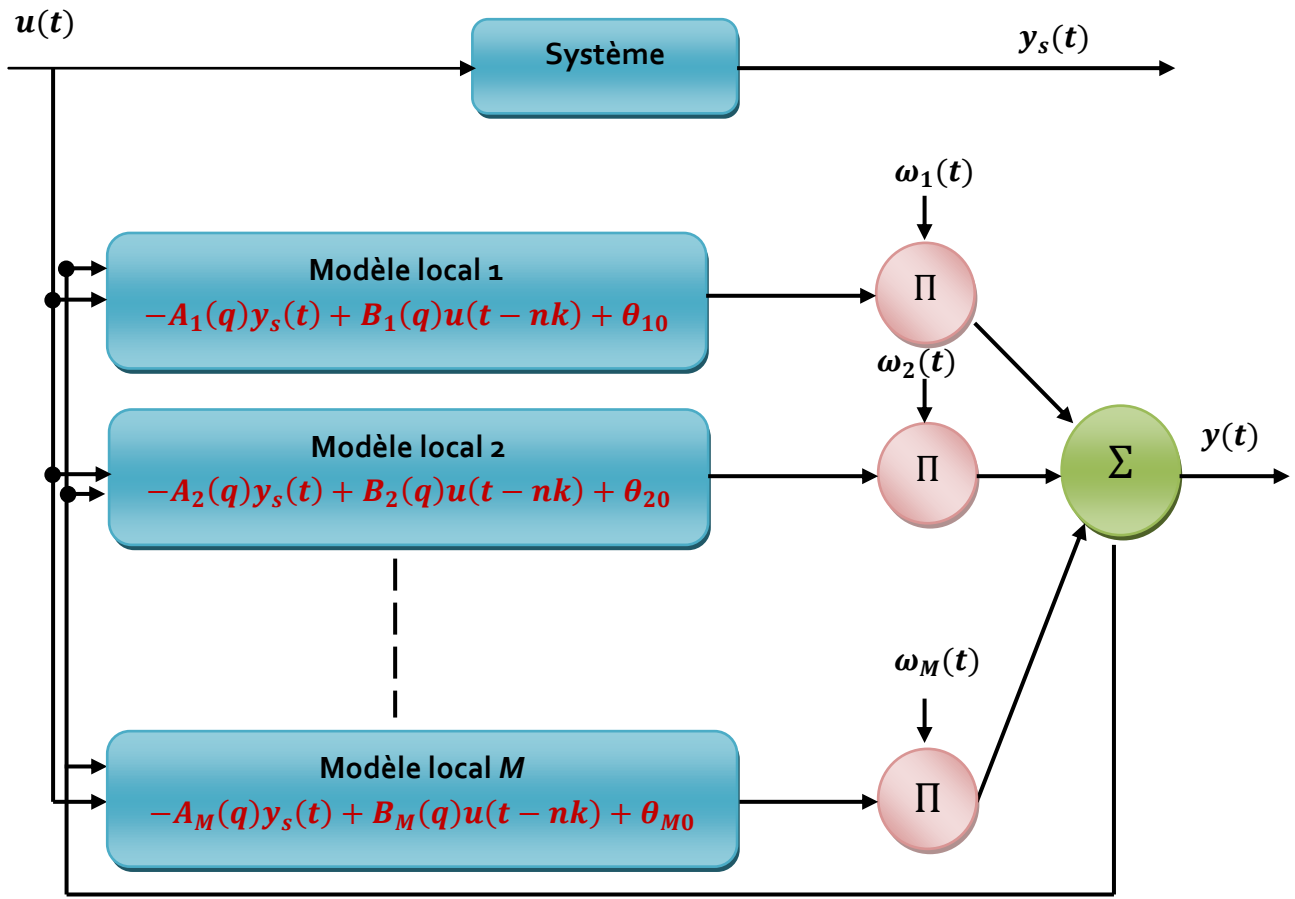


Figure III- 2: Multi-modèle parallèle avec des modèles locaux couplés

❖ CAS 3 : MULTI-MODELE PARALLELE AVEC DES MODELES LOCAUX TOTALEMENT DECOUPLES :

Dans les deux cas précédents, les modèles locaux sont partiellement couplés en raison de l'utilisation des mêmes termes décalés $y_s(t - k)$ ou $y(t - k)$ dans leur expression. Une autre façon d'examiner le problème est de considérer que les modèles locaux ne font pas usage des informations passées sur la sortie du système ou du modèle et qu'ils sont totalement découplés.

Pour un système à m entrées et une sortie, le modèle local $f_i(\varphi(t), \theta_i)$ s'exprime :

$$y_i(t) = \sum_{j=1}^m \frac{B_{ij}(q)}{1 + A_{ij}(q)} u_j(t - nk_{uj}) + \theta_{i0} \tag{III - 19}$$

Chaque modèle local est une interconnexion de m blocs parallèles dont les sorties sont sommées pour fournir celle du modèle local. Les sorties locales sont ensuite pondérées par les fonctions d'activation pour fournir la sortie de l'architecture multi-modèle. Cette dernière est une interconnexion de M modèles locaux en parallèle. L'avantage de ce découplage est de spécifier des

dynamiques différentes (codées par les coefficients des polynômes A_{ij}) pour chaque entrée dans chaque modèle local. L'inconvénient est un nombre de paramètres plus important à estimer.

Une simplification du problème est atteinte en considérant que les polynômes A_{ij} sont identiques pour toutes les entrées u_j ($j = 1, \dots, m$). Cette structure multi-modèle est décrite sur la figure III-3.

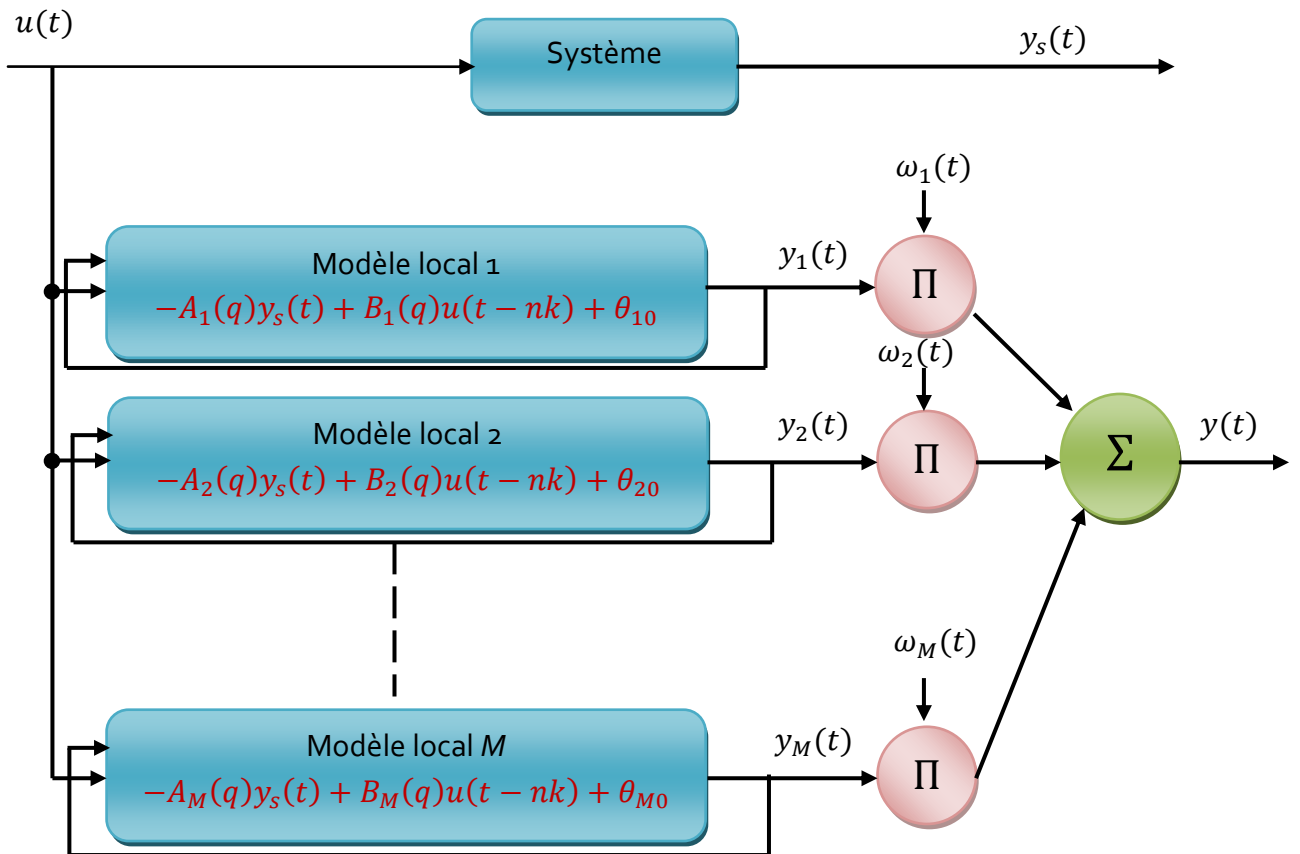


Figure III- 3: Multi-modèle parallèle avec des modèles locaux totalement découplés

III-4-FILTRES ORTHOGONAUX :

A la place des modèles locaux sous forme de relations entrée-sortie, on peut utiliser des filtres issus de la Base Orthogonale Généralisée [Malti, 1999]. Un modèle local décrit par un banc de filtres orthogonaux s'exprime par :

$$y_i(t) = \sum_{j=0}^{n_f-1} a_j g_j * u(t)$$

III – 20

Où « * » désigne l'opérateur de convolution. Les paramètres a_j sont les coefficients de la décomposition sur la base des n_f filtres g_j . La fonction de transfert d'un filtre g_j de la BOG qui s'écrit :

$$G_j(z) = \sqrt{1 - |\xi_j|^2} \frac{z}{z - \xi_j} \prod_{k=0}^{j-1} \left(\frac{1 - \xi_{k-1}^* z}{z - \xi_k} \right) \quad \text{III - 21}$$

Est fonction des pôles ξ_j et ξ_j^* du filtre. Ces pôles déterminent la dynamique incorporée dans le filtre et il y a autant de paires de pôles (à trouver) que de filtres. Si cette dynamique est proche de celle du système localement, le nombre n_f de filtres nécessaires est réduit de façon drastique [Heuberger et al., 1995]. Par ailleurs, si le comportement local du système est caractérisé par un pôle dominant réel ou une paire de pôles complexes, les filtres de Laguerre ou de Kautz (qui nécessitent en tout un pôle ou une paire de pôles complexes) peuvent remplacer ceux de la BOG.

Les chercheurs qui ont utilisé les filtres orthogonaux dans la représentation multi-modèle fixent les pôles et le nombre de filtres [Sbarbaro, 1997] sur la base des connaissances a priori. Si ces dernières sont indisponibles, certains à l'instar de Nelles [1997-b] identifient séparément chacun des modèles locaux sous forme de modèle ARX. La dynamique ainsi suggérée sert à spécifier les pôles des filtres et à déterminer le nombre de filtres [Nelles, 1997-b]. Par contre, Malti [1999] a identifié un multi-modèle par une méthode à deux niveaux qui estime les pôles des filtres par un algorithme non-linéaire et les coefficients a_i par la méthode des moindres carrés, le nombre de filtres de chaque sous-modèle ayant été fixé a priori.

III-5-REPRESENTATION D'ETAT :

Au lieu d'un modèle sous forme entrée-sortie, il est intéressant dans certaines applications (par exemple l'identification de systèmes MIMO) d'utiliser une représentation d'état. D'ailleurs, les chercheurs dont les travaux portent sur l'analyse de la stabilité du multi-modèle adoptent cette représentation. De même, la commande des processus modélisés par l'approche multi-modèle fait souvent appel à la représentation d'état, afin d'étendre au cas non-linéaire des techniques de commande par retour d'état. La construction d'observateurs non-linéaires à partir de structures multi-modèle requiert aussi des modèles locaux exprimés sous la forme de modèle d'état.

Utiliser la représentation d'état revient à exprimer le modèle du système sous la forme :

$$x(t + 1) = g(x(t), u(t)) \quad \text{III - 22}$$

$$y(t) = h(x(t), u(t)) \quad \text{III - 23}$$

$x \in \mathcal{R}^n$ et $u \in \mathcal{R}^m$ sont respectivement les vecteurs d'état et de commande. En supposant que les modèles locaux sont issus d'une linéarisation des fonctions $g(x, u)$ et $h(x, u)$ autour de M points (x_i, u_i) , la formulation multi-modèle aboutit aux équations suivantes :

$$x(t + 1) = \sum_{i=1}^m \omega_i(z(t), \beta) [A_i x(t) + B_i u(t) + d_i] \tag{III - 24 - a}$$

$$y(t) = \sum_{i=1}^m \omega_i(z(t), \beta) [C_i x(t) + d'_i] \tag{III - 24 - b}$$

Avec :

$$A_i = \left. \frac{\partial g}{\partial x} \right|_{\substack{x=x_i \\ u=u_i}} \tag{III - 25 - a}$$

$$B_i = \left. \frac{\partial g}{\partial u} \right|_{\substack{x=x_i \\ u=u_i}} \tag{III - 25 - b}$$

$$C_i = \left. \frac{\partial h}{\partial x} \right|_{\substack{x=x_i \\ u=u_i}} \tag{III - 25 - c}$$

$$d_i = g(x_i, u_i) - (A_i x_i + B_i u_i) \tag{III - 26 - a}$$

$$d'_i = h(x_i, u_i) - C_i x_i \tag{III - 26 - b}$$

A_i, B_i, C_i sont respectivement les matrices d'état, de commande et de sortie du modèle local f_i . Le vecteur des variables caractéristiques $z(t)$ est formé par les états et la commande ou par une combinaison de ces variables. De façon similaire au multi-modèle entrée-sortie, les équations (III- 24) s'arrangent sous la forme :

$$x(t + 1) = \left[\sum_{i=1}^M \omega_i(z(t), \beta) A_i \right] x(t) + \left[\sum_{i=1}^M \omega_i(z(t), \beta) B_i \right] u(t) + \left[\sum_{i=1}^M \omega_i(z(t), \beta) d_i \right] \tag{III - 27 - a}$$

$$y(t) = \left[\sum_{i=1}^M \omega_i(z(t), \beta) C_i \right] x(t) + \left[\sum_{i=1}^M \omega_i(z(t), \beta) d'_i \right] \dots \dots \dots \tag{III - 27 - b}$$

Qui exprime le multi-modèle comme un modèle d'état dont les matrices d'état, de commande et de sortie sont dépendantes du temps : le modèle du système s'assimile à un modèle d'état affine variant dans le temps.

Les fonctions g et h étant en pratique inconnues, les matrices A_i, B_i, C_i et les d_i, d'_i coefficients mis en jeu dans le multi-modèle sont à déterminer à partir de mesures expérimentales.

III-6-LIENS AVEC LES MODELES NON-LINEAIRES EXISTANTS :

III-6-1-MODELE FLOU DE TAKAGI-SUGENO (MODELE TS) :

Il est évident de ranger le modèle flou TS dans la catégorie des multi-modèles. En effet, un modèle TS réalise une partition floue de l'espace caractéristique Z . Les zones de fonctionnement sont définies en termes de propositions sur les variables de prémisses. En prenant l'opérateur produit comme t-norme, le modèle flou TS coïncide exactement avec le multi-modèle. Si à la base, Takagi et Sugeno dans leur article de référence datant de 1985 [Takagi et Sugeno, 1985] n'ont pas été précis sur le type d'opérateur utilisé, toute ambiguïté a été levée avec la sortie de l'article de Sugeno et Kang [1988] où explicitement l'opérateur produit est l'opérateur de conjonction. Le caractère flou disparaît de la formulation du modèle TS car la partie conséquence d'une règle est parfaitement déterminée. Le modèle TS a donc plus de rapports avec le multi-modèle qu'avec les modèles linguistiques de Mamdani ou les modèles à relation floue. Comparativement au multi-modèle, la seule contrainte du modèle TS est sa lecture sous forme de propositions sur les variables de prémisses. Les notions suivantes sont par conséquent équivalentes et expriment les mêmes idées :

- ❖ degré de véracité normalisé \Leftrightarrow fonction d'activation ou d'interpolation,
- ❖ espace de prémisses \Leftrightarrow espace caractéristique,
- ❖ variables de prémisses \Leftrightarrow variables caractéristiques,
- ❖ partie conséquence \Leftrightarrow modèle local.

III-6-2-MODELES MULTI-EXPERTS :

Ces modèles sont partie intégrante de l'approche multi-modèle car ils sont constitués par un ensemble d'experts qui collaborent afin de fournir le modèle global du système. La seule différence avec le multi-modèle est la notion de probabilité associée aux fonctions d'interpolation. Pour ces raisons, on dit que les modèles multi-experts sont des structures multi-modèles dont les

fonctions d'interpolation réalisent une décomposition stochastique de l'espace de fonctionnement. Les modèles locaux sont ici les experts.

III-6-3-LES RESEAUX DE FONCTION A BASE RADIALE (RFBR) :

Strictement, les RFBR tels que définis par l'équation (II-18) au deuxième chapitre ne coïncident pas avec le multi-modèle car les fonctions de base qui agissent sur les poids ne sont pas normalisées (c'est-à-dire leur somme n'est pas égale à 1 pour tout échantillon de la base de données) contrairement aux fonctions d'interpolation qui agissent sur les modèles locaux dans l'approche multi-modèle. Néanmoins, le lien est possible si les fonctions de base sont normalisées : on parle dans ce cas de RFBR généralisés [Hunt et al. 1996]. Les modèles locaux sont alors des constantes (développement d'ordre 0 de la fonction à approcher autour des centres). Soulignons qu'avec des modèles locaux constants, il faudrait un nombre important de zones de fonctionnement pour représenter un système complexe car peu de phénomènes sont pris en compte par une constante. Comparativement un multi-modèle avec des modèles locaux affines nécessiterait.

III-6-4-PROBLEMATIQUE D'IDENTIFICATION D'UNE ARCHITECTURE MULTI-MODELE :

Comme tout problème de modélisation, l'identification d'une architecture multi-modèle comprend deux tâches importantes : la recherche de la structure optimale et l'estimation des paramètres. Nous limitons le cadre de l'analyse aux modèles locaux sous forme de relation entrée-sortie, l'extension à une représentation d'état étant immédiate.

III-6-4-1-IDENTIFICATION STRUCTURELLE :

Au titre de l'identification structurelle, les points essentiels à étudier sont :

- ❖ la spécification des zones de fonctionnement qui comprend :
 - ❖ la sélection des variables caractéristiques les plus représentatives des non-linéarités du système (de façon équivalente, la définition de l'espace caractéristique),
 - ❖ la décomposition de cet espace en zones de fonctionnement c'est-à-dire la détermination de la forme, de la position et du nombre de fonctions de validité,
- ❖ la spécification de la structure de chaque modèle local (variables d'entrée, ordres et retards associés).

L'objectif final est d'identifier un multi-modèle parcimonieux (bonnes qualités d'approximation pour un nombre minimal de modèles locaux).

III-6-4-1-a-COMPROMIS NOMBRE DE MODELES LOCAUX – COMPLEXITE DES MODELES LOCAUX :

Trouver le nombre et la structure adéquats des modèles locaux ne peut se faire de façon simultanée. La raison est que dans l'approche multi-modèle, il y a un compromis à déterminer entre le nombre de modèles locaux et la complexité de leur structure. Si cette structure est complexe, l'architecture multi-modèle nécessitera peu de modèles locaux. Inversement, il faudra plusieurs modèles locaux de structure simple pour obtenir de bonnes capacités d'approximation. La complexité des sous-modèles est directement liée aux ordres n_y et n_{uj} des éléments du vecteur de régression. Elle est également liée au choix d'une structure linéaire ou non-linéaire pour les modèles locaux. Il y a donc interdépendance entre la structure des modèles locaux et la décomposition de l'espace de fonctionnement qui impose de procéder de façon séquentielle.

Deux approches de solution sont possibles :

- ❖ fixer une structure pour les modèles locaux, trouver ensuite la décomposition adéquate de l'espace de fonctionnement puis raffiner ultérieurement la structure des sous-modèles,
- ❖ positionner les zones de fonctionnement et rechercher alors les paramètres structuraux des modèles locaux.

Dans la première option, on recherche dans les données récoltées, des comportements spécifiques correspondant à la structure des modèles locaux. Ceci nécessite forcément de préciser les variables d'entrée, les ordres et les retards de ces modèles avant de prétendre rechercher le nombre et la position des zones de fonctionnement.

Le choix des variables d'entrée utiles et des ordres adéquats peut être guidé par les connaissances disponibles sur le système. Autrement, il faut utiliser une technique heuristique.

La solution directe est de considérer différentes structures pour les modèles locaux, de construire à chaque fois un multi-modèle et de retenir finalement celui qui donne les meilleures performances. A cette solution lourde en temps de calcul, on peut préférer des outils permettant d'identifier les variables d'entrée et les ordres utiles sans recourir au calcul d'un modèle. L'une de ces techniques est la méthode de He et Asada [1993], améliorée par Boukhris et al. [1999] qui est basée sur l'analyse de la sensibilité par rapport aux entrées d'un indicateur évalué à partir de quotients de Lipschitz. Mais la méthode ne fournit de résultats pertinents que si les données traitées sont peu bruitées [Espinosa et Vandewalle, 1997] ou si la variance du bruit de mesure est connue

[Boukhris et al., 1999], ce qui la rend d'une utilisation délicate dans les applications pratiques. Une autre solution certes sous-optimale mais d'une mise en œuvre simple consiste à identifier le meilleur modèle linéaire : sa structure sera celle des modèles locaux [Tanaka et al, 1995]. Une amélioration consiste à déterminer des modèles linéaires optimaux autour de différents points de fonctionnement "judicieusement" choisis [Yu et al., 1997]. Les variables d'entrée et les ordres suggérés par ces différents modèles serviront à spécifier la structure des modèles locaux.

Cette dernière solution rejoint d'une certaine manière la deuxième option qui consiste à positionner les zones de fonctionnement et à essayer de trouver le meilleur modèle local pouvant être affecté à chacune des zones. Cette solution a été appliquée entre autres par [Wang et Langari, 1995], [Nelles, 1997-a].

III-6-4-1-b-DECOMPOSITION DE L'ESPACE DE FONCTIONNEMENT :

La taille du réseau de modèles locaux est non seulement liée à la complexité des sous- modèles mais aussi au type de décomposition adopté pour l'espace caractéristique. On distingue principalement trois techniques de partition : la partition grille, la partition basée sur un arbre de décision et la classification.

❖ PARTITION GRILLE :

Elle réalise un maillage de l'espace caractéristique Z . Le support (ou encore intervalle de variation) de chaque variable caractéristique $Z_j(t) = (j = 1 \dots n_z)$ est divisé en p_j partitions individuelles. La partition de Z est obtenue en prenant l'ensemble des combinaisons des partitions individuelles. Le nombre de modèles locaux générés est alors :

$$M = \prod_{j=1}^{n_z} p_j \quad \text{III - 28}$$

Ce nombre devient très vite explosif si la dimension de l'espace Z est élevée ou le nombre de partitions sur le support des variables caractéristiques est important. Par exemple, pour un système à $n_z = 4$ variables caractéristiques avec $p_j = 5$ divisions par variable, le nombre de modèles locaux est $4^5 = 1024$. Pour ces raisons, la partition grille n'est adaptée qu'aux systèmes de faible dimension (3 à 5 variables caractéristiques). Une illustration de la grille est présentée sur la figure III-4

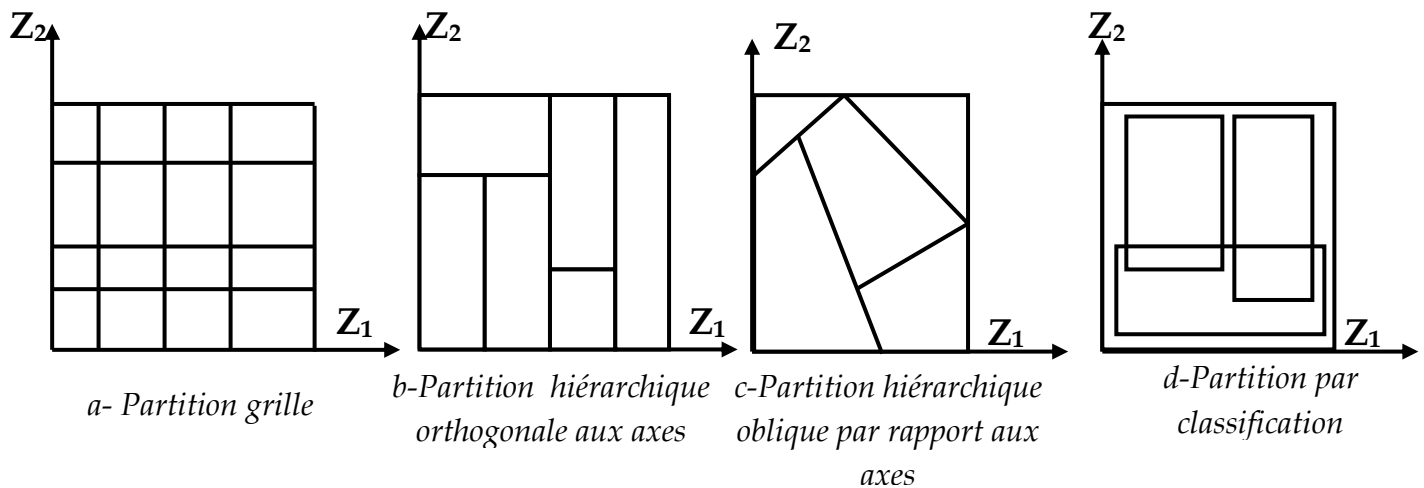


Figure III- 4: Différents types de partition de l'espace caractéristique Z

❖ PARTITION SUIVANT UN ARBRE DE DÉCISION :

On peut réduire le nombre de modèles locaux en adoptant une partition suivant un arbre de décision. Utilisée notamment par Sugeno et Kang [1988], Sun [1994], Johansen et Foss [1995], Nelles [1997-a] et Ernst [1998], elle permet de décomposer l'espace caractéristique selon une procédure hiérarchique. A l'étape v de la décomposition, l'espace Z est déjà découpé en v zones, soit une structure multi-modèle à v modèles locaux. Le découpage d'une zone est envisagé pour obtenir une structure à $v + 1$ modèles locaux. La division de cette zone peut être réalisée par un hyper-plan orthogonal aux axes [Sugeno et Kang, 1988], [Johansen et Foss, 1995], [Nelles, 1997-a] ou se faire suivant un hyper-plan oblique par rapport aux axes [Murray-Smith, 1994], [Ernst 1998]. Le résultat de ces techniques de partition est montré sur les schémas -b- et -c- de la figure III-4. Le modèle final comporte souvent moins de zones que la partition sous forme de grille mais il nécessite plus de fonctions de validité pour les décrire.

❖ PARTITION PAR CLASSIFICATION :

Un autre moyen de partition est l'utilisation d'ensembles multidimensionnels issus d'une procédure de classification. L'objectif est de rechercher des formes ellipsoïdales (par l'algorithme de Gustaffson-Kessel) dans l'espace produit de la sortie et de l'espace de régression. Chaque ellipse correspond à une description du comportement local du système par un hyperplan [Babuška et Verbruggen, 1995]. Les fonctions d'activation sont ensuite déduites par projection des ellipses sur l'espace de régression. La forme des zones de fonctionnement est illustrée sur le schéma d) de la figure III-4. Une excellente référence de l'application des techniques de classification à l'identification d'un multi-modèle est l'ouvrage de Babuška [1998]. Soulignons qu'avec cette approche, l'espace caractéristique est forcément l'espace de régression. Or la reconnaissance des formes devient difficile si la taille de l'espace produit est grande. Pour cette raison, la méthode ne

donne des résultats probants que pour des systèmes de petite dimension et d'ordre faible [Babuška, 1998]. Un autre désavantage de la technique de classification est qu'elle est seulement adaptée pour les multi-modèles à erreur d'équation. Elle est difficilement applicable à un multi-modèle à erreur de sortie : à cause de la présence des termes décalés de la sortie du multi-modèle dans le vecteur de régression, la distribution des données dans l'espace de régression évolue d'une itération à l'autre, ce qui rend difficile la reconnaissance des comportements locaux du système.

III-6-4-1-c-CHOIX DES VARIABLES CARACTERISTIQUES :

Le choix des variables caractéristiques est un problème aussi épineux que celui de la détermination de la structure des modèles locaux. Dans ce cas également, de bonnes connaissances sur le système peuvent aider à spécifier les bonnes variables. Autrement, il faut considérer un ensemble de variables candidates, puis appliquer des procédures heuristiques d'essais-erreurs pour identifier les variables les plus représentatives des non-linéarités du système.

III-6-4-1-d-STRATEGIES D'IDENTIFICATION STRUCTURELLE :

Pour finir notre discussion sur l'identification structurelle, nous présentons trois stratégies de recherche de la structure optimale : ascendante, descendante et alternée. Ces stratégies sont génériques aux problèmes d'identification. Elles sont certes sous-optimales mais elles constituent de bonnes alternatives à la recherche combinatoire. La première méthode procède d'une philosophie de *coarse-to-fine* (du simple au complexe). On démarre avec une structure multi-modèle simple (généralement un modèle linéaire). Elle est ensuite complexifiée par ajouts successifs de nouveaux modèles locaux et la prise en considération de nouvelles variables caractéristiques. Un indicateur permet de spécifier les zones où le multi-modèle a de mauvaises capacités d'approximation et qui nécessitent un raffinement. La procédure continue jusqu'à l'obtention d'un résultat satisfaisant. L'intérêt de cette procédure est que la structure du multi-modèle s'adapte progressivement compte tenu de la complexité du système. Par ce biais, on espère éviter le phénomène de sur-apprentissage.

La deuxième stratégie est l'inverse de la première : elle est de type *fine-to-coarse* (du complexe au simple). Une structure multi-modèle initiale complexe est considérée. Elle est ensuite simplifiée par élimination des modèles locaux ayant un faible pouvoir d'explication et/ou fusion des modèles locaux voisins susceptibles de décrire des comportements similaires du système. L'intérêt est que la structure initiale fine peut potentiellement mettre en évidence toutes les particularités du

système. La simplification ne servira en fait qu'à réduire la finesse de cette structure initiale. Sun [1994] parle de compression de la base de modèles locaux.

La troisième approche est une solution intermédiaire entre les deux premières. Elle consiste à ajouter de nouveaux modèles locaux jusqu'à un point d'arrêt. Ensuite, cette structure est simplifiée. On procède ainsi de façon alternative entre la stratégie incrémentale et décrémente jusqu'à la satisfaction d'un critère d'arrêt. Les critères d'arrêts sont soit des critères de validation croisée soit les critères de type **FPE, AIC, MDL**.

III-6-4-2-TECHNIQUES D'APPRENTISSAGE :

L'optimisation paramétrique consiste à estimer les paramètres des fonctions de validité et des modèles locaux. Dans le cadre de l'approche multi-modèle, trois types d'apprentissage sont envisageables : l'apprentissage local, l'apprentissage global et l'apprentissage avec un critère mixte.

III-6-4-2-a-APPRENTISSAGE LOCAL :

Cette méthode vise à minimiser l'écart entre la sortie du système y_s et celle des modèles locaux [Murray-Smith, 1994], [Murray-Smith et Johansen, 1997-b]. Cette démarche est légitime, car l'idéal dans la représentation des systèmes par un multi-modèle locaux serait de caractériser localement le fonctionnement du système. De plus si les modèles locaux sont issus de la linéarisation d'un modèle physique autour de certains points de fonctionnement ou sont construits à partir de connaissances comportementales du processus étudié, il est préférable d'adopter une approche locale d'optimisation des paramètres des modèles locaux. On définit donc pour chaque modèle local, un critère quadratique pondéré appelé critère local. Le critère local associé au modèle $f_i(\varphi(t), \theta_i)$, est exprimé par :

$$J_i = \frac{1}{2} \sum_{t=1}^N \rho_i(t) (y_i(t) - y_s(t))^2 \quad \text{III - 29}$$

Où N représente la taille du jeu de données d'identification. La contribution d'une observation particulière $\varphi(t)$ au modèle local $f_i(\varphi(t), \theta_i)$, est mesurée par la valeur de la fonction de validité $\rho_i(t)$. Par conséquent, les observations relevant fortement de la zone de validité du modèle f_i sont prépondérantes dans l'estimation de son vecteur de paramètres θ_i . Le critère d'apprentissage local est la somme de tous les critères locaux, soit :

$$J_L = \frac{1}{2} \sum_{i=1}^M \sum_{t=1}^N \rho_i(t) (y_i(t) - y_s(t))^2 \quad .III - 30$$

Si ce critère permet l'optimisation des paramètres des modèles locaux sans difficultés, l'optimisation des paramètres des fonctions de validité sur la base du critère (III-30) pose quelques problèmes. Telles que présentées par Johansen et Foss [1993], les fonctions de validité $\rho_i(t)$ sont des fonctions définies de façon absolue : ce sont des fonctions d'allure gaussienne indépendantes les unes des autres. Il en découle qu'une minimisation du critère (III-30) par rapport aux paramètres des fonctions de validité équivaut à la minimisation de chacun des termes J_i de façon indépendante. Deux ou plusieurs fonctions de validité peuvent être positionnées en un même point suite à l'optimisation. Ce problème est comparable à celui rencontré avec la technique de classification possibiliste PCM [Krishnapuram et Keller, 1996]. A cause de l'indépendance des composantes du critère, il arrive que plusieurs classes identifiées soient confondues [Barni et al., 1996] surtout si l'initialisation de l'algorithme n'a pas été faite de façon judicieuse. Pour des systèmes de petite dimension sur lesquels de bonnes connaissances physiques sont disponibles, on peut positionner a priori les points de fonctionnement et éviter le problème. Faute d'informations a priori, il est préférable d'utiliser les fonctions d'activation ω_i qui dépendent de toutes les zones et garantissent que tous les points de l'espace Z seront couverts à cause de la contrainte de normalisation.

Implicitement, les limites des zones de validité sont maintenant fixées par les fonctions ω_i . Le critère d'apprentissage local est alors défini par :

$$J_L = \frac{1}{2} \sum_{i=1}^M \sum_{t=1}^N \omega_i(t) (y_i(t) - y_s(t))^2 \quad III - 31$$

Pour les fonctions d'activation ω_i fixées, les paramètres des modèles locaux sont identifiés indépendamment les uns des autres. Ce découplage présente l'avantage d'éviter des problèmes numériques lors de l'estimation des paramètres des modèles locaux [Murray-Smith, 1994]. Toutefois, l'apprentissage local nécessite plusieurs modèles locaux, comparativement à l'apprentissage global, pour décrire le comportement du système. En outre, le recouvrement entre les modèles locaux doit être limité c'est-à-dire que les fonctions d'interpolation ω_i doivent avoir un support assez localisé pour garder une interprétation vraiment locale des modèles obtenus.

III-6-4-2-b-APPRENTISSAGE GLOBAL :

L'objectif est d'ajuster les paramètres de manière à minimiser l'écart entre la sortie y_s du système et celle y du multi-modèle. Le critère minimisé est :

$$J_g = \frac{1}{2} \sum_{i=1}^M (y(t) - y_s(t))^2 \quad \text{III} - 31$$

L'apprentissage global recherche une adéquation entre le système et le modèle global sans forcément se soucier de la qualité de l'approximation locale produite par les modèles locaux. Dans l'apprentissage global, les modèles locaux sont positionnés de sorte que leur interpolation donne une bonne approximation du comportement du système. Il en résulte quelquefois, une architecture multi-modèle qui fournit une bonne approximation globale du système mais avec des modèles locaux qui ne sont pas représentatifs du comportement observé dans leur zone de validité. Ce problème mis en évidence dans [Murray-Smith, 1994] est lié au recouvrement plus ou moins grand qui peut exister entre les zones de fonctionnement : les modèles locaux totalement différents du comportement réel du système se compensent entre eux pour reproduire le comportement global désiré.

III-6-4-2-c-APPRENTISSAGE AVEC UN CRITERE MIXTE :

L'apprentissage local vise à préserver le caractère local de la description du système au détriment de la qualité d'approximation globale alors que l'apprentissage global privilégie la qualité globale au détriment de la signification locale des modèles locaux. Cet antagonisme entre la qualité de la description locale et la précision de l'approximation globale peut être résolu en combinant les deux approches d'identification de manière à bénéficier des avantages procurés par chacune des approches. On peut définir un critère composite ou mixte qui est une combinaison linéaire du critère local et du critère global [Yen et al. 1998-b] :

$$J_m = J_g + \eta J_L \quad \text{III} - 32$$

Selon le choix du paramètre de pondération η , on privilégiera la précision globale du modèle ou la qualité d'approximation locale des modèles locaux. Mais le choix de η est souvent subjectif.

III-7-CONCLUSION :

Ce chapitre a été consacré à la présentation de l'approche multi-modèle. A travers cet exercice, on a montré que le multi-modèle est un concept général regroupant des structures non-linéaires comme les modèles flous de Takagi-Sugeno, les modèles multi-experts, L'idée essentielle est la définition du modèle global du système comme la combinaison de modèles locaux. Les notions à retenir sont : zone de fonctionnement, variables caractéristiques, fonction de validité, fonction d'interpolation (ou d'activation), modèles locaux.

Différentes formes de modèles locaux sont possibles : modèles entrée-sortie, représentation d'état pour ne citer que les plus couramment employées. La structure de ces modèles peut être complexifiée à souhait mais on privilégie généralement des modèles locaux affines afin de pouvoir leur appliquer les techniques d'analyse de l'automatique linéaire.

La problématique d'identification structurelle d'un multi-modèle a été discutée. Il ressort de l'analyse que le nombre de modèles locaux à considérer dépend de la complexité de la structure de ces derniers. Plus leur structure est simple, plus leur nombre sera important et vice-versa. Différentes méthodes de décomposition de l'espace de fonctionnement ont été présentées ; elles montrent les options qui s'offrent à l'utilisateur pour construire un multi-modèle.

Deux approches d'optimisation des paramètres ont fait l'œuvre d'une étude détaillée. La méthode d'apprentissage global assure une bonne qualité d'approximation générale du multi-modèle au détriment de la qualité locale de représentation. Si l'objectif de la modélisation ne requiert pas la caractérisation du comportement local du système (exemples : modèles de redondance analytique en diagnostic, modèles de prévision), cette approche s'applique pleinement. Par contre, si on recherche une bonne description locale du système, il convient d'appliquer la méthode d'apprentissage local.

Sommaire

III-1- Introduction :	69
III-2- Formulation Mathematique :	70
III-3- Structure Des Modeles Locaux :	72
III-4- Filtres Orthogonaux :	77
III-5- Representation D'etat :	78
III-6- Liens Avec Les Modeles Non-Lineaires Existants :	80
III-6-1- Modele Flou De Takagi-Sugeno (Modele Ts) :	80
III-6-2- Modeles Multi-Experts :	80
III-6-3- Les Reseaux De Fonction A Base Radiale (Rfbr) :	81
III-6-4- Problematique D'identification D'une Architecture Multi-Modele :	81
III-6-4-1- Identification Structurelle :	81
III-6-4-1-a- Compromis Nombre De Modeles Locaux – Complexite Des Modeles Locaux :	82
III-6-4-1-b- Decomposition De L'espace De Fonctionnement :	83
III-6-4-1-c- Choix Des Variables Caracteristiques :	85
III-6-4-1-d- Strategies D'identification Structurelle :	85
III-6-4-2- Techniques D'apprentissage :	86
III-6-4-2-a- Apprentissage Local :	86
III-6-4-2-b- Apprentissage Global :	88
III-6-4-2-c- Apprentissage Avec Un Critere Mixte :	88
III-7- Conclusion :	89

LISTE DES FIGURES

Figure III- 1: Multi-modèle série-parallèle avec des modèles locaux couplés.....	75
Figure III- 2: Multi-modèle parallèle avec des modèles locaux couplés.....	76
Figure III- 3: Multi-modèle parallèle avec des modèles locaux totalement découplés.....	77
Figure III- 4: Différents types de partition de l'espace caractéristique Z.....	84

Chapitre IV

Optimisation paramétrique des multi-modèles



Chapitre IV

Optimisation paramétrique des multi-modèles



As philosopher of science Karl Popper has emphasized, a good theory is characterized by the fact that it makes a number of predictions that could in principle be disproved or falsified by observation. Each time new experiments are observed to agree with the predictions the theory survives, and our confidence in it is increased; but if ever a new observation is found to disagree, we have to abandon or modify the theory.

At least that is what is supposed to happen, but you can always question the competence of the person who carried out the observation.

—Stephen Hawking, a Brief History of Time [117]

IV-1-INTRODUCTION :

Le chapitre précédent a été dédié à la présentation de l'approche multi-modèle ainsi qu'à la problématique d'identification qu'elle soulève.

Rappelons que les questions à se poser pour construire un réseau de modèles locaux sont :

- ❖ Combien faut-il de modèles locaux ?
- ❖ Quelle structure de modèles locaux faut-il choisir ?
- ❖ Comment choisir les variables caractéristiques Z_j et comment effectuer la décomposition de l'espace de fonctionnement Z ?
- ❖ Comment estimer les paramètres mis en jeu ?

Les réponses à ces diverses interrogations constituent le menu de ce chapitre.

RAPPELS :

Avant de rentrer dans le vif du sujet, nous faisons ici quelques rappels pour situer le cadre de l'étude. Le multi-modèle est exprimé par la formule générale :

$$y(t) = \sum_{i=1}^M \omega_i(z(t), \beta) f_i(\varphi(t), \theta_i) \quad IV - 1$$

Où les ω_i sont des fonctions d'activation dépendant de variables caractéristiques $z(t)$. Les modèles locaux f_i considérés dans ce chapitre sont des modèles entrée-sortie affines ; la sortie $y_i(t)$ d'un modèle local est exprimée par

$$y_i(t) = \varphi_a^T(t) \theta_i \quad IV - 2$$

$\varphi_a(t)$ est le vecteur de régression augmenté formé par : $\varphi_a(t) = [\varphi^T(t) \quad 1]^T$.

θ_i est le vecteur des paramètres du modèle f_i . Ce vecteur est de dimension :

$$n_i = n_y + \sum_{r=1}^m n_{u_r} + 1 \quad IV - 3$$

avec m le nombre d'entrées du système, n_y, n_{u_r} les ordres. La sortie du modèle local f_i s'écrit également de façon polynomiale

$$y_i(t) = -A_i(q)y^*(t) + \sum_{r=1}^m B_{ir}(q)u_r(t - nk_{u_r}) + \theta_{i0} \quad IV - 4$$

Où $y^*(t)$ est la sortie $y(t)$ du multi-modèle, si ce dernier est de type erreur de sortie (OE), soit la sortie du système, si le multi-modèle considéré est à erreur d'équation. $y^*(t)$ peut également être la sortie du modèle local f_i dans le cas de multi-modèle parallèle à modèles locaux découplés. A_i et B_{ir} sont des polynômes en q^{-1} (q^{-1} : *opérateur retard*). Pour des raisons de simplicité des écritures mathématique, nous supposons que tous les modèles locaux ont la même structure (les mêmes entrées u_r , les mêmes ordres n_y, n_{u_r} et les mêmes retards nk_{u_r}).

IV-2-OPTIMISATION D'UN MULTI-MODELE A TRAVERS UNE PARTITION GRILLE :**IV-2-1-FORMULATION MATHEMATIQUE DE LA GRILLE :**

Soit un système non-linéaire qu'on désire représenter avec un multi-modèle comportant n_z variable caractéristique z_1, \dots, z_{n_z} . On envisagé de décomposer l'espace engendré par ces variables sous la

forme d'une grille. La construction de la partition grille repose sur la définition de partitions sur le support (intervalle de variation) de chacune des variables caractéristiques.

Soit p_j le nombre de partitions réalisées sur le support d'une variable z_j . On attribue à chaque partition $p_{l,j}$ de z_j , une fonction de validité individuelle $\mu_{l,j}$ ($l = 1, \dots, p_j$ et $j = 1, \dots, n_z$).

Examinons maintenant comment sont construites les zones de fonctionnement (que nous désignons encore par zones de validité). Par exemple, la zone de fonctionnement Z_1 est construite par combinaison de la 1^{ière} partition ($P_{1,j}$) de chacune des variables z_j .

Les contours de cette zone sont délimités par la fonction de validité ρ_1 , produit des fonctions de validité individuelles $\mu_{l,j}$, c'est-à-dire $\rho_1 = \prod_{j=1}^{n_z} \mu_{1,j}$. L'architecture multi-modèle décrite par une partition grille résulte de l'ensemble de ces combinaisons susceptible d'être générées. Le nombre de zones Z_i est alors :

$$M = \prod_{j=1}^{n_z} p_j$$

Pour un multi-modèle avec deux variables caractéristiques $n_z = 2$, deux partitions sur le support de z_1 , et trois sur z_2 , la procédure de génération de toutes les zones de fonctionnement, à partir des partitions individuelles, est résumée par le tableau IV-1.

one Z_i	Z_1		Z_2			Fonction de validité ρ_i	Doublet \mathfrak{Z}_i
	Partition de Z_1		Partition de Z_2				
	$P_{1,1} \rightarrow \mu_{1,1}$	$P_{2,1} \rightarrow \mu_{2,1}$	$P_{1,2} \rightarrow \mu_{1,2}$	$P_{2,2} \rightarrow \mu_{2,2}$	$P_{3,2} \rightarrow \mu_{3,2}$		
Z_1	1	0	1	0	0	$\mu_{1,1} \times \mu_{1,2}$	(1,1)
Z_2	1	0	0	1	0	$\mu_{1,1} \times \mu_{2,2}$	(1,2)
Z_3	1	0	0	0	1	$\mu_{1,1} \times \mu_{3,2}$	(1,3)
Z_4	0	1	1	0	0	$\mu_{2,1} \times \mu_{1,2}$	(2,1)
Z_5	0	1	0	1	0	$\mu_{2,1} \times \mu_{2,2}$	(2,2)
Z_6	0	1	0	0	1	$\mu_{2,1} \times \mu_{3,2}$	(2,3)

Tableau IV- 1: Table de composition des zones de fonctionnement pour un multi-modèle à 2 variables caractéristiques

A chaque zone Z_i , on associe un n_z -uplet noté \mathfrak{S}_i qui code les partitions des variables z_j intervenant dans la construction de la zone Z_i . La fonction de validité de cette zone est déduite comme précédemment, c'est-à-dire :

$$\rho_i = \prod_{j=1}^{n_z} \mu_{l_j^{(i)}, j} \quad i = 1, \dots, M \quad IV - 5$$

Dans cette équation, l'indice $l_j^{(i)}, j$ désigne la partition particulière de z_j intervenant dans la construction de la zone Z_i . La valeur de la variable $l_j^{(i)}$ est fournie par le $j^{\text{ème}}$ élément de \mathfrak{S}_i en se référant à la table de composition des zones. Par exemple pour la zone Z_3 , $l_1^{(3)} = 1$ et $l_2^{(3)} = 3$.

Les fonctions d'activation ω_i sont issues de la normalisation des fonctions de validité ρ_i

$$\omega_i = \frac{\rho_i}{\sum_{k=1}^M \rho_k}$$

On montre aisément que le terme au dénominateur dans cette équation se factorise sous la forme :

$$\sum_{k=1}^M \rho_k = \prod_{j=1}^{n_z} \left(\sum_{l=1}^{p_j} \mu_{l,j} \right)$$

On en déduit que la fonction d'activation ω_i est équivalente à l'expression suivante :

$$\omega_i = \prod_{j=1}^{n_z} \left(\frac{\mu_{l_j^{(i)}, j}}{\sum_{l=1}^{p_j} \mu_{l,j}} \right) = \prod_{j=1}^{n_z} \overline{\mu_{l_j^{(i)}, j}}$$

Qui est le produit de fonctions de validité individuelles normalisées $\overline{\mu_{l_j^{(i)}, j}}$. Ceci signifie, que dans une partition grille, la contrainte de partition unité sur les fonctions d'activation ω_i induit implicitement une normalisation sur les fonctions individuelles $\mu_{l,j}$ des variables caractéristiques.

IV-2-2-CHOIX DES FONCTIONS DE VALIDITE INDIVIDUELLES :

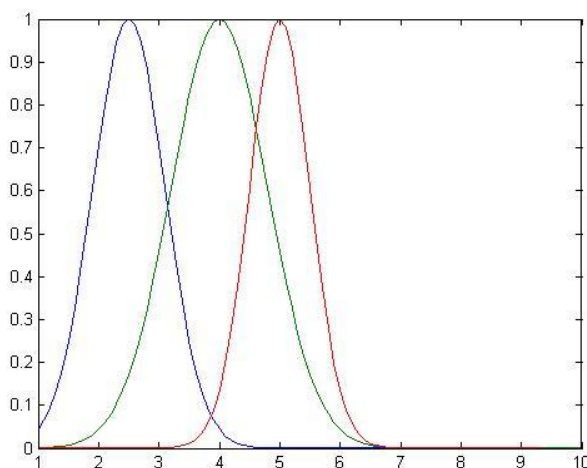
Il existe une kyrielle de choix pour les fonctions de validité. Les paramètres de ces fonctions intervenant de façon non-linéaire dans l'expression du multi-modèle, leur optimisation passe par des techniques de recherche itérative. D'un autre côté, les fonctions de validité doivent être choisies de

manière à assurer des propriétés lisses à l'approximation du système. Un choix particulier sera donc motivé par les algorithmes d'optimisation de paramètres qu'on mettra en œuvre et/ou la qualité d'interpolation désirée. On peut distinguer les situations suivantes :

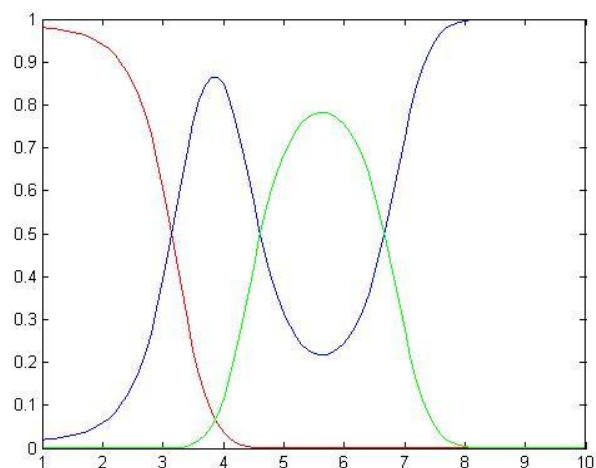
- les fonctions de validité individuelles $\mu_{i,j}$ sont construites à partir de fonctions à dérivées non continues comme les fonctions triangulaires ou trapézoïdales : les algorithmes de type Newton ne peuvent s'appliquer. On a recours à des techniques comme la méthode du simplexe de Nelder-Mead ou à des algorithmes génétiques. Des exemples d'optimisation des paramètres de fonctions triangulaires ou trapézoïdales à partir d'algorithmes génétiques sont reportés dans [Siarry et Guely, 1998], [Glorennec, 1999]. Par ailleurs, les fonctions de validité continues par morceaux réalisent une interpolation linéaire alors que dans certaines applications, il est appréciable d'avoir une transition lisse entre les modèles locaux,
- les fonctions de validité $\mu_{i,j}$ sont à dérivées continues, à l'exemple des fonctions gaussiennes ou sigmoïdes : les techniques d'optimisation de type Newton s'appliquent alors pleinement. Ces fonctions présentent l'avantage de réaliser des interpolations lisses entre les modèles locaux. Pour ces raisons, nous préférons dans la suite ce type de fonctions de validité.

IV-2-3-FONCTIONS DE VALIDITE D'ALLURE GAUSSIENNE :

L'utilisation des fonctions de validité d'allure gaussienne a été popularisée par les techniques "neuro-floues" [Jang, 1993]. Mais elles présentent l'inconvénient d'entraîner des problèmes de réactivation lorsqu'elles sont normalisées [Shorten et Murray-Smith, 1997]. Ce phénomène est illustré sur la figure IV-1.



a- Fonctions de validité d'allure gaussienne



b-Fonctions d'interpolation issue de la normalisation des fonctions de validité

Figure IV- 1:Illustration du phénomène de réactivation

Le phénomène de réactivation est évité si toutes les fonctions d'allure gaussienne définies sur le support d'une variable z_j ont la même dispersion. Pour se prémunir contre une éventuelle réactivation, nous imposons des dispersions identiques pour les fonctions de validité individuelles $\mu_{l,j}$ lorsque nous utilisons les gaussiennes. Par conséquent, les fonctions de validité afférentes à une variable z_j ($j = 1, \dots, n_z$) sont définies par :

$$\mu_{l,j}(z_j(t)) = \exp\left(\frac{(z_j(t) - c_{l,j})^2}{2\sigma_j^2}\right) \quad l = 1, \dots, p_j \quad IV - 6$$

$c_{l,j}$ est le centre. σ_j est la dispersion et elle est commune à toutes les fonctions $\mu_{l,j} \forall l = 1, \dots, p_j$. Avec cette contrainte, la définition des p_j partitions $p_{l,j}$ sur le support d'une variable caractéristique z_j nécessite $p_j + 1$ paramètres (p_j centre et la dispersion commune).

IV-3-ALGORITHMES D'OPTIMISATION PARAMETRIQUE :

On suppose les variables caractéristiques z_j définies ainsi que le nombre p_j de partitions sur leur domaine. La structure des modèles locaux est également supposé connue. La prochaine étape est l'ajustement des paramètres du multi-modèle sur la base des données expérimentales.

Soit $\theta = [\theta_1^T \ \theta_2^T \ \dots \ \theta_M^T]^T$, le vecteur formé par la concaténation des vecteurs de paramètres de tous les modèles locaux. La dimension de ce vecteur est :

$$n_\theta = M \times n \quad IV - 7$$

Où n est la dimension commune à tous les vecteurs de paramètres locaux θ_i donnée par (IV-7).

Soit $\beta = [\beta_1^T \ \beta_2^T \ \dots \ \beta_{n_z}^T]^T$, le vecteur total des paramètres des fonctions de validité individuelles. Sa dimension est :

$$n_\beta = \sum_{j=1}^{n_z} (p_j + 1) \quad IV - 8$$

Notons $\Theta = [\beta^T \ \theta^T]^T$, le vecteur global de paramètres formé par la concaténation des vecteurs β et θ .

Le critère optimisé est le critère d'apprentissage global défini à l'équation suivante :

$$J_g = \frac{1}{2} \sum_{t=1}^N (y(t) - y_s(t))^2 \quad \text{IV - 9}$$

Le problème de recherche du minimum du critère J_g par rapport θ n'ayant pas de solution analytique, on a recours à des techniques itératives d'optimisation non-linéaire. L'estimation des paramètres est dérivée de la formule générale suivante :

$$\widehat{\theta}_{(k+1)} = \widehat{\theta}_{(k)} - \eta_{(k)} H_{(k)}^{-1} G_{(k)} \quad \text{IV - 10}$$

Où $\widehat{\theta}_{(k)}$ est l'estimation obtenue à l'itération (k). $\eta_{(k)}$ est le coefficient de relaxation, $G_{(k)}$ une estimation du gradient du critère J_g à l'itération (k) et $H_{(k)}$ une matrice qui modifie la direction de recherche. La mise en œuvre de recherche itérative décrite par (IV-10) peut être réalisée "en ligne" ou "hors ligne". Nous restreignons notre étude au cas "hors ligne" où la procédure de mise à jour des paramètres fait appel à chaque itération à l'ensemble des données d'identification.

Selon le choix de la matrice H , on distingue différentes méthodes d'optimisation non-linéaire. Nous adoptons la méthode de Levenberg-Marquardt (LM) qui possède de bonnes propriétés de convergence même pour des problèmes mal conditionnés [Sjöberg et Viberg, 1997].

L'optimisation simultanée des paramètres des fonctions de validité et des modèles locaux par une technique itérative est en général lourde. La raison est la dimension élevée de θ . De plus, dans le cas de la grille, il peut exister des zones plus ou moins dépourvues de données d'apprentissage ; l'optimisation directe du vecteur θ est sujette à des problèmes numériques. Il est alors difficile d'assurer de bonnes propriétés de convergence de l'algorithme d'optimisation.

Dans la pratique, on essaie de simplifier le problème, soit en supposant connue la position des zones de fonctionnement, soit en mettant en œuvre une procédure à deux niveaux qui alterne entre l'optimisation du vecteur θ et celle de β . Nous présentons dans les paragraphes suivants, différents algorithmes d'apprentissage, de la méthode la plus simple où les paramètres des fonctions de validité sont fixés, à la méthode générale où tous les paramètres sont optimisés.

IV-3-1-LES PARAMETRES DES FONCTIONS DE VALIDITE SONT FIXES, OPTIMISATION D'UN CRITERE QUADRATIQUE PAR RAPPORT AUX PARAMETRES DES MODELES LOCAUX :

La configuration la plus simple qu'on puisse rencontrer est celle où les fonctions de validité sont fixées a priori. Les paramètres recherchés sont alors ceux des modèles locaux θ .

Si ces modèles sont linéaires par rapport à leurs paramètres (multi-modèle à erreur d'équation), le critère J_g est quadratique par rapport au vecteur θ dont l'estimation est fournie par la solution analytique des moindres carrés. Par contre, pour des modèles locaux dynamiques de type erreur de sortie, le critère n'est pas quadratique par rapport à θ car le vecteur de régression $\phi(t)$ comporte des sorties décalées $y(t-k)$ ($k = 1, \dots, n_y$) du multi-modèle. L'estimation du vecteur θ passe alors par une technique itérative. Néanmoins, en remplaçant dans le vecteur de régression, les sorties $y(t-k)$ par les mesures $y_s(t-k)$ de la sortie du système, le critère J_g devient quadratique par rapport au vecteur θ . On passe d'un critère à erreur de sortie à un critère à erreur d'équation.

Pour un ensemble de $N + n_a$ observations $n_a = \max(n_y, n_{u_1} + nk_{u_1, \dots, u_m} + nk_{u_m})$, on déduit des équations (IV-1) et (IV-2), l'écriture matricielle :

$$Y = \Phi_g \theta \quad \theta \in \mathbb{R}^{n_\theta \times 1} \quad IV - 11$$

Avec les définitions suivantes :

$$\Phi_g = [W_1 \Phi_a \quad W_2 \Phi_a \quad \dots \quad W_M \Phi_a] \Phi_g \in \mathbb{R}^{N \times n_\theta} \quad IV - 12 - a$$

$$\Phi_a = \begin{pmatrix} \varphi_a^T(n_a + 1) \\ \varphi_a^T(n_a + 2) \\ \vdots \\ \varphi_a^T(N + n_a) \end{pmatrix} \Phi_a \in \mathbb{R}^{N \times n} \quad IV - 12 - b$$

$$W_i = \text{diag}[\omega_i(t)]_{t=n_a+1}^{N+n_a} W_i \in \mathbb{R}^{N \times N} \quad i = 1, \dots, M \quad IV - 12 - c$$

W_i est une matrice de pondération diagonale dont les éléments sont constitués par les valeurs de la fonction d'interpolation $\omega_i(t)$ aux différents instants. Φ_a est la matrice de régression augmentée comprenant les variables explicatives de la sortie du système. La minimisation du critère J_g donne la solution analytique :

$$\hat{\theta} = R^{-1}(\Phi_g^T Y) \quad Y \in \mathbb{R}^{N \times 1} \quad IV - 13$$

$$R = \Phi_g^T \Phi_a$$

Compte tenu de la structure de la matrice Φ_g , une vue éclatée de l'estimation $\hat{\theta}$ est donnée par :

$$\hat{\theta} = \begin{bmatrix} \Phi_a^T W_1^2 \Phi_a & \cdots & \Phi_a^T W_1 W_M \Phi_a \\ \vdots & \ddots & \vdots \\ \Phi_a^T W_M W_1 \Phi_a & \cdots & \Phi_a^T W_M^2 \Phi_a \end{bmatrix}^{-1} \begin{bmatrix} \Phi_a^T W_1 Y \\ \vdots \\ \Phi_a^T W_M Y \end{bmatrix} \quad IV - 14$$

Cette équation montre l'interaction qui existe entre les modèles locaux lorsque leurs paramètres sont identifiés par la technique d'apprentissage global. On constate en effet que les éléments diagonaux de la matrice d'information R portent sur les points appartenant aux différentes zones de fonctionnement Z_1 à Z_M . Les termes hors diagonaux représentent le recouvrement qui existe entre les modèles locaux. Pour un multi-modèle à une variable caractéristique dont les partitions sont représentées par des fonctions de validité triangulaires ou trapézoïdales, il est possible d'établir que la matrice R est une matrice tri diagonale par blocs. Ceci simplifie le calcul de l'inverse de R d'une part et permet de montrer d'une façon plus précise, l'interdépendance qui existe entre les vecteurs de paramètres des modèles locaux. Comparons l'équation (IV-14) avec l'estimation donnée par l'apprentissage local, que nous explicitons ci-dessous.

❖ APPRENTISSAGE LOCAL :

On considère que le critère local est quadratique par rapport aux paramètres des modèles locaux.

Pour $N + n_a$ observations, les critères locaux associés aux modèles f_i s'écrivent :

$$J_i = \frac{1}{2} \|\Phi_a \theta_i - Y\|_{W_i}^2 \quad i = 1, \dots, M \quad IV - 15$$

Avec la définition précédente de la matrice de pondération W_i . Les estimateurs des paramètres des modèles locaux sont alors donnés par :

$$\hat{\theta}_i = (\Phi_a^T W_i \Phi_a)^{-1} \Phi_a^T W_i Y \quad i = 1, \dots, M \quad IV - 16$$

A partir de cette série d'équations, on peut écrire le vecteur global de paramètres sous la forme :

$$\hat{\theta} = \begin{bmatrix} \Phi_a^T W_1 \Phi_a & 0 & \cdots & 0 \\ 0 & \ddots & \vdots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & \Phi_a^T W_M \Phi_a \end{bmatrix}^{-1} \begin{bmatrix} \Phi_a^T W_1 Y \\ \vdots \\ \vdots \\ \Phi_a^T W_M Y \end{bmatrix} \quad IV - 17$$

❖ COMPARAISON DES DEUX ESTIMATIONS :

En comparant les équations (IV-14) et (IV-17), on constate que dans l'approche globale, la matrice d'information comporte des termes relatifs au recouvrement entre les modèles locaux. Par ailleurs, les termes diagonaux de la matrice d'information font intervenir le carré de la matrice de pondération contrairement à l'approche locale. Les estimations fournies par les deux méthodes sont quasiment égales s'il y a peu de recouvrement entre les modèles locaux. Compte tenu de la contrainte de partition unité, peu de recouvrement signifie que les fonctions d'activation ont des fronts plutôt raides.

❖ PROBLEME DE CONDITIONNEMENT DE L'ESTIMATION PARAMETRIQUE :

L'inversion de la matrice d'information R dans l'approche globale peut être sujette à des problèmes numériques liés à son mauvais conditionnement.

Dans le cadre de l'identification d'un multi-modèle, un certain nombre de causes responsables du mauvais conditionnement de la matrice d'information peuvent être répertoriées. Ces causes peuvent être :

- numériques : un échantillonnage trop rapide du système peut rendre des colonnes de la matrice de régression augmentée Φ_a quasiment colinéaires. Elle sera mal conditionnée induisant inévitablement le mauvais conditionnement de la matrice de régression globale Φ_g . Les perturbations aléatoires affectant les mesures peuvent également provoquer ce problème. Ce raisonnement reste valable si l'estimation des paramètres est réalisée avec l'approche locale,
- structurelles : la surparamétrisation due au nombre parfois excessif de modèles locaux (ce problème est typique de la grille où le nombre de sous-modèles devient très vite explosif) et/ou l'absence ou l'insuffisance de données d'apprentissage dans certaines zones de fonctionnement sont des causes plausibles du mauvais conditionnement. Si des zones de fonctionnement sont dépourvues de données, les matrices de pondération W_i correspondantes sont quasiment nulles. On remarque, à partir de l'équation IV – 12 – α qu'il va exister des blocs de Φ_g qui seront presque nuls ce qui induira forcément une perte de rang de la matrice d'information,
- liées à un recouvrement trop important entre les modèles locaux. Pour deux zones Z_i et Z_j se recouvrant fortement, on a la matrice de pondération $W_i \cong W_j$. Les blocs matrices $W_i \Phi_a$ et $W_j \Phi_a$ sont quasiment semblables. Il en résulte une "forte" colinéarité entre des

colonnes de la matrice de régression Φ_g et donc là aussi une perte de rang de la matrice d'information.

IV-3-2-OPTIMISATION DE TOUS LES PARAMETRES SUR LA BASE D'UN

CRITERE A ERREUR D'EQUATION:

Si des connaissances a priori sur la position des zones de fonctionnement ne sont pas disponibles, les paramètres de la structure multi-modèle doivent être optimisés au moyen d'une procédure itérative en raison de la non-linéarité du modèle global par rapport à ses paramètres. Cependant, en supposant que le critère J_g est toujours quadratique par rapport aux paramètres des modèles locaux (cas des modèles locaux linéaires par rapport à leurs paramètres ou d'un critère à erreur d'équation), on peut mettre en œuvre un algorithme à deux niveaux : elle alterne entre l'optimisation des paramètres β des fonctions de validité par une procédure non-linéaire, θ étant fixé et l'optimisation des paramètres θ des modèles locaux par l'estimateur des moindres carrés, éventuellement régularisés, β étant fixé. Les étapes de la procédure d'optimisation sont résumées dans l'algorithme suivant.

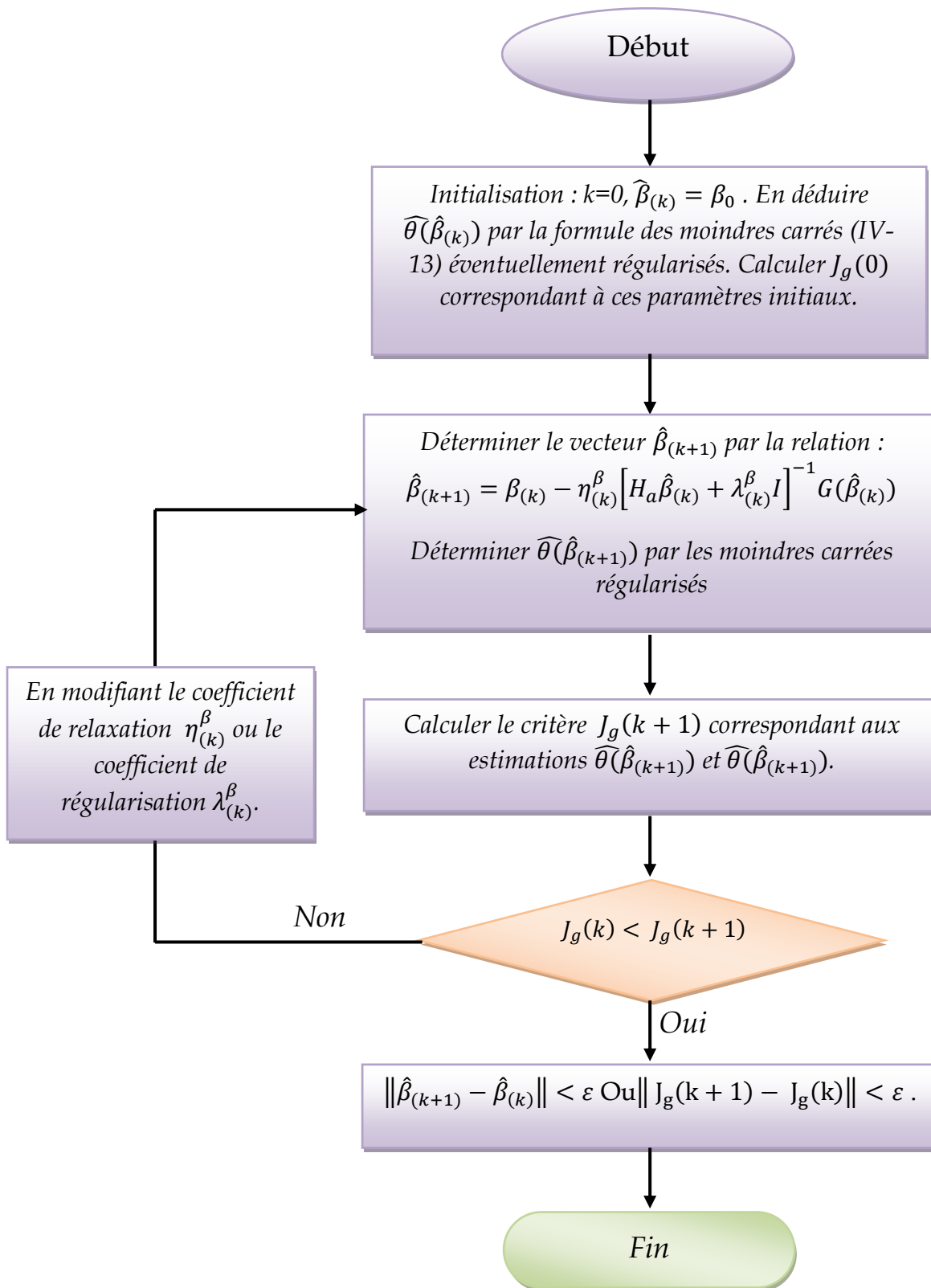


Figure VI- 2: optimisation a deux niveaux MC+LM

Le gradient G et le hessien approché H_a du critère J_g nécessaires à la mise en œuvre de la méthode de Levenberg-Marquardt s'obtiennent par les équations :

$$G(\beta_{(k)}) = \sum_{t=1}^N (y(t) - y_s(t)) \frac{\partial y(t)}{\partial \beta} \Big|_{\beta=\hat{\beta}_{(k)}} \quad IV - 18 - a$$

$$H_a(\beta_{(k)}) = \sum_{t=1}^N \frac{\partial y(t)}{\partial \beta} \frac{\partial y(t)}{\partial \beta^T} \Big|_{\beta=\hat{\beta}_{(k)}} \quad IV - 18 - b$$

A partir de la formulation générale (IV – 1) et (IV – 2), on exprime la fonction de sensibilité $\partial y/\partial \beta$ sous la forme :

$$\frac{\partial y(t)}{\partial \beta} = \sum_{i=1}^N \frac{\partial \omega_i(z(t), \beta)}{\partial \beta} \varphi_a^T(t) \theta_i \quad IV - 19$$

Les fonctions de validité $\mu_{l,j}$ afférentes à une variable $z_j (j = 1, \dots, n_z)$ sont définies par :

$$\mu_{l,j}(z_j(t)) = \exp\left(-\frac{(z_j(t) - c_{l,j})^2}{2\sigma_j^2}\right) \quad l = 1, \dots, p_j$$

$$\omega_i(t) = \prod_{j=1}^{n_z} \frac{\mu_{l_j^{(i)},j}(t)}{\sum_{l=1}^{p_j} \mu_{l,j}(t)} = \prod_{j=1}^{n_z} \bar{\mu}_{l_j^{(i)},j}(t) \quad l = 1, \dots, p_j$$

$c_{l,j}$: Centre de la fonction $\mu_{l,j}$.

σ_j : Sa dispersion ; elle est commune à toutes les gaussiennes $\mu_{l,j} (\forall l = 1, \dots, p_j)$ afin d'éviter des problèmes de réactivation.

IV-3-2-1-DERIVEE DE LA SORTIE DU SYSTEME PAR RAPPORT A UN CENTRE

$c_{r,j} (r = 1, \dots, p_j, j = 1, \dots, n_z)$:

Elle s'écrit :

$$\frac{\partial y(t)}{\partial c_{r,j}} = \frac{\partial}{\partial c_{r,j}} \left(\sum_{i=1}^M \omega_i(t) y_i(t) \right)$$

Pour des raisons de simplification de l'écriture, nous allons omettre le temps t dans les formules qui suivront.

$$\frac{\partial y}{\partial c_{r,j}} = \sum_{i=1}^M \frac{\partial \omega_i}{\partial c_{r,j}} y_i \quad IV - 20$$

Décomposons la dérivée $\partial \omega_i / \partial c_{r,j}$

$$\frac{\partial \omega_i}{\partial c_{r,j}} = \frac{\partial}{\partial c_{r,j}} \left(\prod_{k=1}^{n_z} \bar{\mu}_{l_k^{(i)},k}(z_k) \right) \quad IV - 21$$

Les fonctions individuelles normalisées d'une variable z_j étant indépendantes des fonctions individuelles normalisées des autres variables, nous pouvons écrire :

$$\frac{\partial \omega_i}{\partial c_{r,j}} = \frac{\partial \bar{\mu}_{l_j^{(i)},j}(z_j)}{\partial c_{r,j}} \times \left(\prod_{\substack{k=1 \\ k \neq j}}^{n_z} \bar{\mu}_{l_k^{(i)},k}(z_k) \right) \quad IV - 22$$

Explicitons à présent la dérivée $\partial \bar{\mu}_{l_j^{(i)},j} / \partial c_{r,j}$. Elle s'obtient par :

$$\begin{aligned} \frac{\partial \bar{\mu}_{l_j^{(i)},j}(z_j)}{\partial c_{r,j}} &= \frac{\partial \mu_{l_j^{(i)},j}(z_j)}{\partial c_{r,j} \sum_{l=1}^{p_j} \mu_{l,j}(z_j)} \\ \frac{\partial \bar{\mu}_{l_j^{(i)},j}(z_j)}{\partial c_{r,j}} &= \frac{1}{\sum_{l=1}^{p_j} \mu_{l,j}(z_j)} \times \left[\frac{\partial \mu_{l_j^{(i)},j}(z_j)}{\partial c_{r,j}} - \bar{\mu}_{l_j^{(i)},j} \times \frac{\partial \mu_{r,j}(z_j)}{\partial c_{r,j}} \right] \quad IV - 23 \end{aligned}$$

La dérivée $\partial \bar{\mu}_{l_j^{(i)},j} / \partial c_{r,j}$ est non nulle si $l_j^{(i)} = r$ c'est-à-dire si la partition $P_{r,j}$ de la variable z_j intervient dans la définition de la zone de validité Z_i du modèle local f_i . On exprime donc ce dérivé par la relation suivante :

$$\frac{\partial \mu_{l_j^{(i)},j}(z_j)}{\partial c_{r,j}} = \delta_{l_j^{(i)},r} \times \frac{\partial \mu_{r,j}(z_j)}{\partial c_{r,j}}$$

Avec $\delta_{l_j^{(i)},r}$ l'opérateur de Kronecker tel que :

$$\delta_{l_j^{(i)},r} = \begin{cases} 1 & \text{si } l_j^{(i)} = r \\ 0 & \text{si } l_j^{(i)} \neq r \end{cases}$$

En injectant cette expression dans IV – 23 on obtient :

$$\frac{\partial \bar{\mu}_{l_j^{(i)},j}(z_j)}{\partial c_{r,j}} = \frac{1}{\sum_{l=1}^{p_j} \mu_{l,j}(z_j)} \times \left[\delta_{l_j^{(i)},r} - \bar{\mu}_{l_j^{(i)},j}(z_j) \right] \times \frac{\partial \mu_{r,j}(z_j)}{\partial c_{r,j}} \quad \text{IV – 24}$$

Les fonctions de validité individuelles $\mu_{r,j}$ étant des gaussiennes, la dérivée $\partial \mu_{r,j}/\partial c_{r,j}$ est fournie par :

$$\frac{\partial \mu_{r,j}(z_j)}{\partial c_{r,j}} = \frac{z_j - c_{r,j}}{\sigma_j^2} \times \mu_{r,j}(z_j)$$

Notons $d_{r,j}(z_j) = \frac{z_j - c_{r,j}}{\sigma_j^2}$. La dérivée $\partial \bar{\mu}_{l_j^{(i)},j}/\partial c_{r,j}$ devient alors :

$$\frac{\partial \bar{\mu}_{l_j^{(i)},j}(z_j)}{\partial c_{r,j}} = \frac{\mu_{r,j}(z_j)}{\sum_{l=1}^{p_j} \mu_{l,j}(z_j)} \times \left[\delta_{l_j^{(i)},r} - \bar{\mu}_{l_j^{(i)},j}(z_j) \right] \times d_{r,j}(z_j)$$

$$\frac{\partial \bar{\mu}_{l_j^{(i)},j}(z_j)}{\partial c_{r,j}} = d_{r,j}(z_j) \times \left[\delta_{l_j^{(i)},r} - \bar{\mu}_{l_j^{(i)},j}(z_j) \right] \times \bar{\mu}_{r,j}(z_j) \quad \text{IV – 25}$$

Remontons cette dans l'expression IV – 22 de la dérivée $\partial \omega_i/\partial c_{r,j}$. Cette dernière devient :

$$\frac{\partial \omega_i}{\partial c_{r,j}} = \left[d_{r,j}(z_j) \times \left(\delta_{l_j^{(i)},r} - \bar{\mu}_{l_j^{(i)},j}(z_j) \right) \times \bar{\mu}_{r,j}(z_j) \right] \times \prod_{\substack{k=1 \\ k \neq j}}^{n_z} \bar{\mu}_{l_k^{(i)},k}(z_k)$$

$$\begin{aligned} \frac{\partial \omega_i}{\partial c_{r,j}} &= d_{r,j}(z_j) \times \delta_{l_j^{(i)},r} \times \left[\bar{\mu}_{r,j}(z_j) \times \prod_{\substack{k=1 \\ k \neq j}}^{n_z} \bar{\mu}_{l_k^{(i)},k}(z_k) \right] - d_{r,j}(z_j) \times \bar{\mu}_{r,j}(z_j) \times \bar{\mu}_{l_j^{(i)},j}(z_j) \\ &\quad \times \prod_{\substack{k=1 \\ k \neq j}}^{n_z} \bar{\mu}_{l_k^{(i)},k}(z_k) \end{aligned}$$

Avec :

$$\bar{\mu}_{l_j^{(i)},j}(z_j) \times \prod_{\substack{k=1 \\ k \neq j}}^{n_z} \bar{\mu}_{l_k^{(i)},k}(z_k) = \omega_i$$

Cette expression comporte deux termes : le premier terme est non nul si $l_j^{(i)} = r$. Si c'est le cas l'expression $\bar{\mu}_{l_j^{(i)},j}(z_j) \times \prod_{\substack{k=1 \\ k \neq j}}^{n_z} \bar{\mu}_{l_k^{(i)},k}(z_k)$,est rien d'autres que la fonction d'activation ω_i du modèle f_i . Il s'en suit cette expression simplifiée de la dérivée $\partial \omega_i / \partial c_{r,j}$:

$$\frac{\partial \omega_i}{\partial c_{r,j}} = d_{r,j}(z_j) \times \left(\delta_{l_j^{(i)},r} - \bar{\mu}_{r,j}(z_j) \right) \times \omega_i \quad IV - 26$$

A partir de cette équation, la fonction de la sortie du multi-modèle par rapport au centre $c_{r,j}$ s'explique :

$$\frac{\partial y}{\partial c_{r,j}} = \sum_{i=1}^M d_{r,j}(z_j) \times \left(\delta_{l_j^{(i)},r} - \bar{\mu}_{r,j}(z_j) \right) \times (\omega_i \times y_i) \quad IV - 27$$

Ceci conduit à cette expression compacte de la fonction de sensibilité :

$$\frac{\partial y}{\partial c_{r,j}} = d_{r,j}(z_j) \times \left(\sum_{i=1}^M \delta_{l_j^{(i)},r} \times (\omega_i \times y_i) - \bar{\mu}_{r,j}(z_j) \times y \right) \quad IV - 28$$

IV-3-2-2-DERIVEE DE LA SORTIE DU SYSTEME PAR RAPPORT A LA DISPERSION COMMUNE $\sigma_j(j = 1, \dots, n_z)$:

Elle est fournie par l'équation suivante :

$$\frac{\partial y}{\partial \sigma_j} = \sum_{i=1}^M \frac{\partial \omega_i}{\partial \sigma_j} \times y_i$$

Les calculs sont similaires à ceux de la fonction de sensibilité par rapport au centre $c_{r,j}$. La modification porte uniquement sur la dérivation $\partial \bar{\mu}_{l_j^{(i)},j} / \partial \sigma_j$ qui est développée ici.

$$\frac{\partial \bar{\mu}_{l_j^{(i)},j}(z_j)}{\partial \sigma_j} = \frac{\partial}{\partial \sigma_j} \frac{\mu_{l_j^{(i)},j}(z_j)}{\sum_{l=1}^{p_j} \mu_{l,j}(z_j)} = \frac{1}{\sum_{l=1}^{p_j} \mu_{l,j}(z_j)} \left[\frac{\partial \mu_{l_j^{(i)},j}(z_j)}{\partial \sigma_j} - \bar{\mu}_{l_j^{(i)},j}(z_j) \times \frac{\partial}{\partial \sigma_j} \left(\sum_{l=1}^{p_j} \mu_{l,j}(z_j) \right) \right]$$

Or on a :

$$\frac{\partial \mu_{l,j}(z_j)}{\partial \sigma_j} = e_{l,j}(z_j) \times \mu_{l,j}(z_j) \quad IV - 29 - a$$

$$e_{l,j}(z_j) = \frac{(z_j - c_{l,j})^2}{\sigma_j^3} \quad IV - 29 - b$$

De plus, comme les fonctions de validité de z_j ont la même dispersion, la dérivée $\partial \mu_{l_j^{(i)},j} / \partial \sigma_j$ n'est jamais nulle. Par analogie à l'équation ci-dessus, on l'exprime comme suit :

$$\frac{\partial \mu_{l_j^{(i)},j}(z_j)}{\partial \sigma_j} = \frac{(z_j - c_{l_j^{(i)},j})^2}{\sigma_j^3} \mu_{l_j^{(i)},j}(z_j) = e_{l_j^{(i)},j}(z_j) \times \mu_{l_j^{(i)},j}(z_j) \quad IV - 30$$

En partant de ces considérations, on écrit la dérivée $\partial \bar{\mu}_{l_j^{(i)},j} / \partial \sigma_j$:

$$\begin{aligned} \frac{\partial \bar{\mu}_{l_j^{(i)},j}(z_j)}{\partial \sigma_j} &= \frac{1}{\sum_{l=1}^{p_j} \mu_{l,j}(z_j)} \\ &= \left[e_{l_j^{(i)},j}(z_j) \times \mu_{l_j^{(i)},j}(z_j) - \bar{\mu}_{l_j^{(i)},j}(z_j) \times \left(\sum_{l=1}^{p_j} e_{l,j}(z_j) \times \mu_{l,j}(z_j) \right) \right] \end{aligned}$$

Cette écriture se simplifie :

$$\frac{\partial \bar{\mu}_{l_j^{(i)},j}(z_j)}{\partial \sigma_j} = \bar{\mu}_{l_j^{(i)},j}(z_j) \times \left[e_{l_j^{(i)},j}(z_j) - \sum_{l=1}^{p_j} e_{l,j}(z_j) \times \bar{\mu}_{l,j}(z_j) \right] \quad IV - 31$$

En injectant dans la dérivée $\partial \omega_i / \partial \sigma_j$, on obtient :

$$\begin{aligned} \frac{\partial \omega_i}{\partial \sigma_j} &= \frac{\bar{\mu}_{l_j^{(i)},j}(z_j)}{\partial \sigma_j} = \prod_{\substack{k=1 \\ k \neq j}}^{n_z} \bar{\mu}_{l_k^{(i)},k}(z_k) \\ \frac{\partial \omega_i}{\partial \sigma_j} &= \left[e_{l_j^{(i)},j}(z_j) - \sum_{l=1}^{p_j} e_{l,j}(z_j) \times \bar{\mu}_{l,j}(z_j) \right] \times \bar{\mu}_{l_j^{(i)},j}(z_j) \times \prod_{\substack{k=1 \\ k \neq j}}^{n_z} \bar{\mu}_{l_k^{(i)},k}(z_k) \end{aligned}$$

$$\frac{\partial \omega_i}{\partial \sigma_j} = \left[e_{l_j^{(i)}}(z_j) - \sum_{l=1}^{p_j} e_{l,j}(z_j) \times \bar{\mu}_{l,j}(z_j) \right] \omega_i \quad IV - 32$$

En exploitant cette expression, la fonction de sensibilité s'écrit :

$$\frac{\partial y}{\partial \sigma_j} = \sum_{i=1}^M \frac{\partial \omega_i}{\partial \sigma_j} y_i = \sum_{r=1}^M \left[e_{l_j^{(i)}}(z_j) - \sum_{l=1}^{p_j} e_{l,j}(z_j) \times \bar{\mu}_{l,j}(z_j) \right] \omega_i \times y_i$$

Finalement on obtient la relation donnant la fonction de sensibilité sous la forme :

$$\frac{\partial y}{\partial \sigma_j} = \sum_{i=1}^M e_{l_j^{(i)}}(z_j) \times (\omega_i \times y_i) - \left(\sum_{l=1}^{p_j} e_{l,j}(z_j) \times \bar{\mu}_{l,j}(z_j) \right) \times y \quad IV - 33$$

Dans cet algorithme, toute modification du vecteur β est immédiatement suivie par une mise à jour de l'ensemble des paramètres θ des modèles locaux. L'avantage de cette procédure à deux niveaux est la réduction de la taille de l'espace paramétrique. A chaque itération de l'algorithme, l'estimateur des moindres carrés fournit une solution optimale dans l'espace des paramètres des modèles locaux, ce qui accélère la convergence de l'algorithme. D'autre part, les paramètres de régularisation sont réglés de façons différentes : dans le calcul de $\hat{\theta}$, il est déterminé de façon à assurer un bon conditionnement de la matrice d'information R ; dans l'estimation de β , le paramètre de régularisation est adapté de manière heuristique.

IV-4-EXEMPLE:

La procédure d'identification jusqu'alors décrite est illustrée sur un exemple de simulation. Soit à modéliser un système non-linéaire à 4 entrées et une sortie décrit par les données entrées sortie d'une simulation sous Simulink d'une installation carburant.

La sortie représente la température à la sortie de l'installation carburant (unité de mesure est le Kelvin)

Les variables caractéristiques candidates que nous avons choisies sont $u_1(t)$, $u_2(t)$, $u_3(t)$, $u_4(t)$, La structure initiale des modèles locaux est choisie de la forme :

$$y_i(t) = -a_1^{(i)} y_s(t-1) - a_2^{(i)} y_s(t-2) - a_3^{(i)} y_s(t-3) + b_{11}^{(i)} u_1(t-1) + b_{12}^{(i)} u_2(t-1) \\ + b_{13}^{(i)} u_3(t-1) + b_{14}^{(i)} u_4(t-1)$$

- Le nombre de partition sur chaque support caractéristique pour la première itération vaut 1 ($p_1 = p_2 = p_3 = p_4 = 1$).
- Le nombre de modèles initial est $M = p_1 \times p_2 \times p_3 \times p_4$.
- Le nombre de paramètres pour chaque modèle vaut 7.

On augmente le nombre de partition progressivement sur chaque support on obtient les résultats suivantes:

N° ligne	u_1 p_1	u_2 p_2	u_3 p_3	u_4 p_4	Nombre de modèles	Erreur Total
01	1	1	1	1	1	0.069
02	2	1	1	1	2	0.0035
03	3	1	1	1	3	0.0026
04	4	1	1	1	4	0.0016
05	5	1	1	1	5	0.0015
06	6	1	1	1	6	0.0011
07	7	1	1	1	7	0.00096
08	8	1	1	1	8	9.6878e-4
09	9	1	1	1	9	6.2647e-4
10	10	1	1	1	10	6.3647e-4
11	11	1	1	1	10	6.3647e-4
12	10	2	1	1	20	2.8256e-4
13	10	3	1	1	30	2.8295e-4
14	10	2	2	1	40	1.5834e-4
15	10	2	3	1	60	1.1285e-4
16	10	2	4	1	80	8.0075e-5
17	10	2	4	2	160	2.5277e-5
18	10	2	4	3	240	8.1557e-5

Tableau IV- 2: Recherche d'une structure optimale du multi-modèle

Discussion :

Si on augmente le nombre de partitions sur le support de u_1 (ligne 11 du tableau IV-2), l'erreur totale n'évolue pas donc l'augmentation du nombre de partition sur ce support n'améliore pas notre approximation.

Le nombre de partition sur u_1 reste vaut 10 et l'algorithme incrémente le nombre de partition sur le support de u_2 ce qui nous donne une amélioration sur l'erreur totale (ligne 12 du tableau IV-2), l'algorithme incrémente de 1 le nombre de partition sur le support de u_2 pour la prochaine itération ce qui donne une erreur total plus grande que la précédente, l'algorithme prend en compte que l'itération précédente est mieux que la présente.

L'algorithme fixe le nombre de partition sur le support de u_2 à 2 et incrémente le nombre de partition sur le support de u_3 (ligne 15 du tableau IV-2).

Pour l'itération suivante l'erreur totale devient très grand par rapport à la précédente (ligne 16 du tableau IV-2), l'algorithme revient à l'itération précédente et incrémente le nombre de partition de la variable suivante (ligne 17 du tableau IV-2).

L'algorithme augmente progressivement le nombre de partition sur le support de dernier variable jusqu'au l'erreur total devient inferieur à J_{stru} fixé à priori.

La sortie de multi modèle et de système après la dernière itération est représentée dans la figure IV-4

La structure optimale, au vu de ce tableau, correspond à un nombre de partitions de 10 et 2 et 4 et 2 respectivement pour les variables caractéristiques u_1, u_2, u_3, u_4 . Le multi-modèle comporte donc 240 modèles locaux soit 4080 paramètres pour l'ensemble des modèles locaux.

Remarquent que le nombre de modèles deviens très vite explosive, c'est l'un des désavantage de l'approche multi modèle.

Une solution implicite est de fusionner où éliminer les modèle redondante et les zones vide, parce que ces zones n'apportent pas d'information pour l'explication du comportement du système.

Les figures (IV-3), (IV-4) montrent la superposition des composantes du vecteur de sorties du modèle non linéaire de l'installation carburant et leurs approximation par le multi modèle

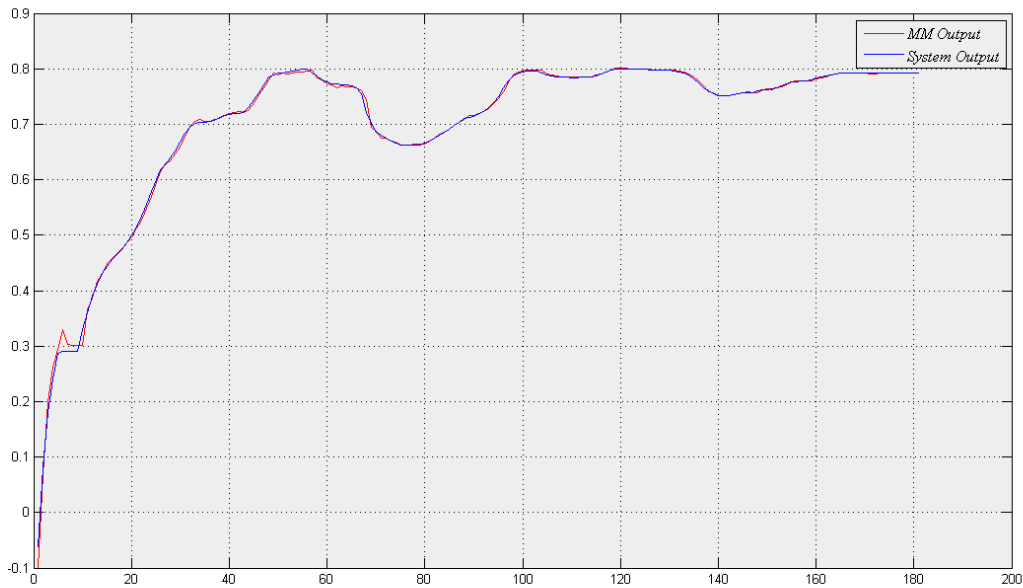


Figure VI- 3: Comparaison des mesures avec la simulation fournie par le multi-modèle dans la première itération

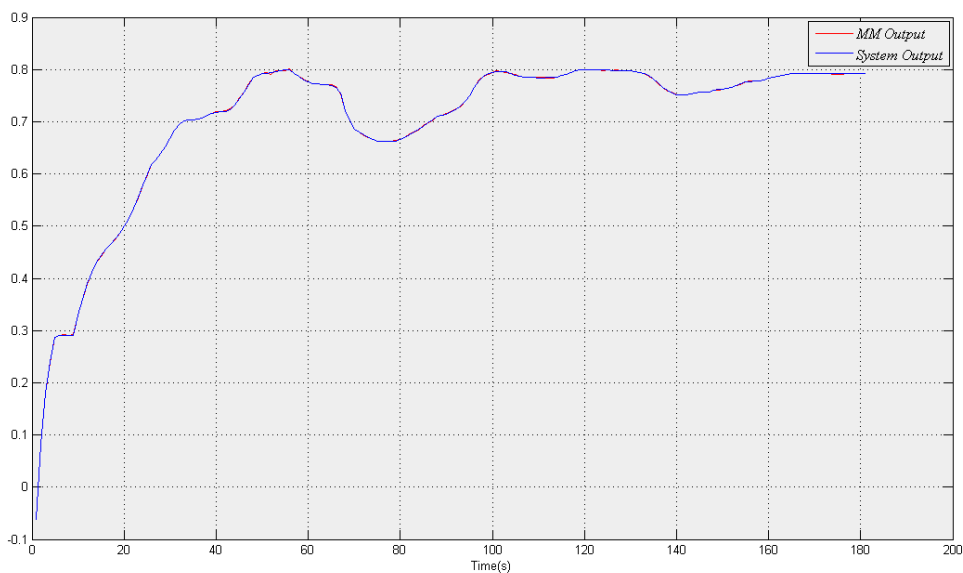


Figure VI- 4: Comparaison des mesures avec la simulation fournie par le multi-modèle dans la dernière itération

IV-5-CONCLUSION :

Ce chapitre a opté pour l'optimisation simultanée des paramètres des fonctions d'activation et des modèles locaux. L'optimisation de tous les paramètres mis en jeu à travers des algorithmes à deux niveaux alternant entre la détermination des paramètres θ des modèles locaux d'une part et des paramètres β des fonctions de validité de l'autre. Selon que le critère d'identification est de type erreur d'équation ou erreur de sortie, l'estimation des paramètres θ est obtenue par les moindres carrés régularisés (la régularisation a consisté à assurer un conditionnement de la matrice d'information) ou par l'algorithme de Levenberg-Marquardt. Les paramètres β intervenant de façon non quadratique dans le critère d'identification, ils sont estimés par l'algorithme de Levenberg-Marquardt.

A la fin de ce chapitre, on introduit un exemple illustre la mise en œuvre de l'approximation d'un modèle non linéaire par plusieurs modèles locaux interpolés par des fonctions d'activation. On applique l'algorithme d'identification décrit dans le schéma synoptique V-9.

L'augmentation du nombre de modèles locaux permet la prise en compte de la complexité plus ou moins grande de système et permet également d'atteindre la précision souhaitée du modèle global selon les objectifs poursuivis.

Les limites de la partition grille sont atteintes lorsque l'espace caractéristique est de grande dimension : le nombre de zones devient important, des zones vides de données apparaissent ainsi que des zones redondantes (zones voisines pouvant décrire un même comportement du système).

Sommaire

IV-1- Introduction :	94
Rappels :	95
IV-2- Optimisation D'un Multi-Modele A Travers Une Partition Grille :	95
IV-2-1- Formulation Mathematique De La Grille :	95
IV-2-2- Choix Des Fonctions De Validite Individuelles :	97
IV-2-3- Fonctions De Validite D'allure Gaussienne :	98
IV-3- Algorithmes D'optimisation Parametrique :	99
IV-3-1- Les Parametres Des Fonctions De Validite Sont Fixes, Optimisation D'un Critere Quadratique Par Rapport Aux Parametres Des Modeles Locaux :	100
IV-3-2- Optimisation De Tous Les Parametres Sur La Base D'un Critere A Erreur D'equation: ..	104
IV-3-2-1- Derivee De La Sortie Du Systeme Par Rapport A Un Centre c_r, j ($r = 1, \dots, p, j =$ $1, \dots, nz$) :	106
IV-3-2-2- Derivee De La Sortie Du Systeme Par Rapport A La Dispersion Commune σ_j ($j =$ $1, \dots, nz$) :	109
IV-4- Exemple:	111
IV-5- Conclusion :	115

LISTE DES FIGURES :

Figure VI- 1:Illustration du phénomène de réactivation	98
Figure VI- 2: optimisation a deux niveaux MC+LM	105
Figure VI- 3:Comparaison des mesures avec la simulation fournie par le multi-modèle dans la première itération.....	114
Figure VI- 4:Comparaison des mesures avec la simulation fournie par le multi-modèle dans la dernière itération.....	114

LISTE DES TABLEAUX :

Tableau IV- 1: Table de composition des zones de fonctionnement pour un multi-modèle à 2 variables caractéristiques.....	96
Tableau IV- 2: Recherche d'une structure optimale du multi-modèle.....	112

Chapitre V

Modélisation Et Simulation



Chapitre V

Modélisation Et Simulation



It is not only old and early impressions that deceive us; the charms of novelty have the same power.

Blaise Pascal

V-1-INTRODUCTION :

La turbine à gaz est un moteur thermique réalisant les différentes phases de son cycle thermodynamique dans une succession d'organes traversés par un fluide moteur gazeux en écoulement continu. C'est une différence fondamentale par rapport aux moteurs à pistons qui réalisent une succession temporelle des phases dans un même organe (généralement un cylindre).

Dans sa forme la plus simple, la turbine à gaz fonctionne selon le cycle dit de Joule (Brayton), figure I-20, comprenant successivement et schématiquement :

- une compression adiabatique qui consomme de l'énergie mécanique ;
- un chauffage isobare.
- détente adiabatique jusqu'à la pression ambiante qui produit de l'énergie mécanique ;
- un refroidissement isobare.

La turbine à gaz est le plus souvent à cycle ouvert et à combustion interne. Dans ce cas, la phase de refroidissement est extérieure à la machine et se fait par mélange à l'atmosphère. La turbine à gaz peut également être à cycle fermé et à combustion externe. Le chauffage et le refroidissement sont alors assurés par des échangeurs de chaleur. Cette disposition plus

complexe permet l'utilisation de gaz particuliers ou de travailler avec une pression basse différente de l'ambiante.

Le cycle de base décrit plus haut peut être amélioré par différents organes complémentaires :

- **récupération de chaleur à l'échappement** : les gaz détendus en sortie de turbine traversent un échangeur pour préchauffer l'air comprimé avant son admission dans la chambre de combustion ;
- **compression refroidie** : la compression comprend deux étages (ou plus) séparés par un échangeur de chaleur (air/air ou air/eau) refroidissant l'air. La puissance nécessaire à la compression s'en trouve réduite au bénéfice du rendement ;
- **combustion étagée** : la détente comprend deux étages (ou plus) séparés par un ou des réchauffages additionnels. La puissance fournie est accrue d'où amélioration du rendement.

V-2-PRINCIPE DE FONCTIONNEMENT :

Le compresseur (repère C), constitué d'un ensemble d'ailettes fixes (stator) et mobiles (rotor), comprime l'air extérieur (repère. A), simplement filtré, jusqu'à 10 à 15 bars, voire 30 bars pour certains modèles.

Du gaz (repère. G), ou un combustible liquide pulvérisé, est injecté dans la (les) chambre(s) de combustion (repère. CH) où il se mélange à l'air comprimé et s'enflamme. Les gaz chauds se détendent en traversant la turbine (repère. T), où l'énergie thermique des gaz chauds est transformée en énergie mécanique. La turbine est constituée d'une ou plusieurs roues également munies d'ailettes précédées d'aubages fixes (directrices). Les gaz de combustion s'échappent par la cheminée (repère. EC) à travers un diffuseur. Le mouvement de rotation de la turbine est communiqué à l'arbre A qui actionne d'une part le compresseur, d'autre part une charge qui n'est autre qu'un appareil (machine) récepteur (ice), (pompe, alternateur ...) accouplé à son extrémité droite. Pour la mise en route, on utilise un moteur de lancement qui joue le rôle de démarreur. Le réglage de la puissance et de la vitesse de rotation est possible en agissant sur le débit de l'air en entrée et sur l'injection du carburant.

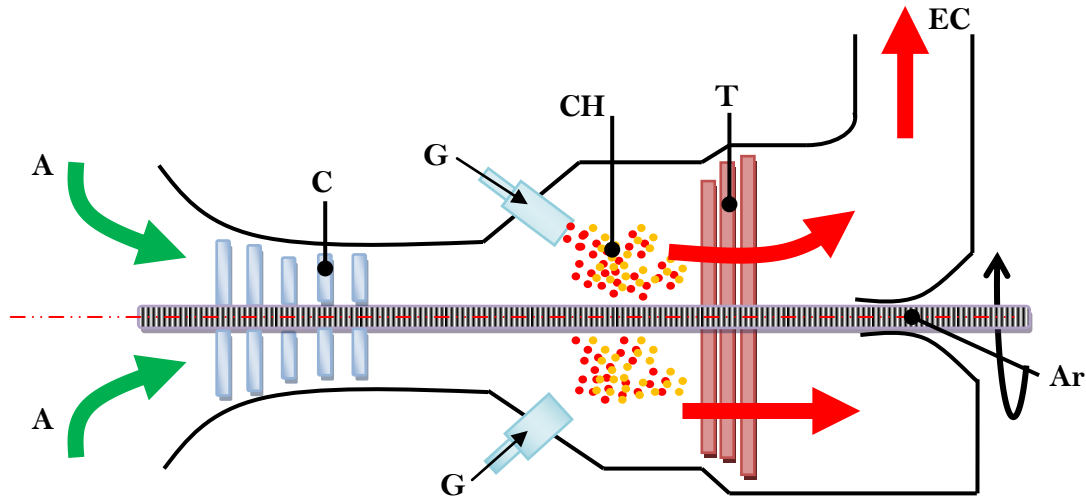


Figure V- 1 : Principe de fonctionnement d'une turbine à gaz

V-3-MODELISATION ET IDENTIFICATION :

La technique d'identification par l'approche multi modèle a été appliquée sur des données obtenues à partir d'une simulation d'une turbine à gaz on utilisant l'outil Logiciel Simulink.

La turbine à gaz travaille sous les conditions ISO de ça fonctionnement, données par le tableau ci-dessous.

Air mass flow rate [Kg/s]	24,4
Cycle pressure ratio (P_{oc}/P_{ic})	9,1
Electrical power (P_e) [KW]	5220
Exhaust temperature (T_{ot})[K]	796
Fuel mass flow rate (M_f) [Kg/s]	17

Tableau V- 1 : Paramètres principaux (ISO) de fonctionnement d'une turbine à gaz

V-4-REPRESENTATION SOUS SIMULINK :

V-4-1-SIMUNLINK :

Le logiciel Simulink est le standard de fait dans l'industrie pour l'activité de modélisation, de spécification et de simulation des systèmes.

Simulink est une extension du logiciel de calcul Matlab. Il permet de modéliser et de simuler des systèmes évoluant dans le temps. Par abus de langage, Simulink désigne également le langage

graphique permettant de représenter ces systèmes. Les technique et méthodes constituant le logiciel Simulink sont issues de l'analyse numérique. Le moteur de simulation appelé **solver**, c'est-à-dire le cœur de l'outil, est centré sur les méthodes d'intégration numérique.

V-4-2-LES BRIQUES DE BASE

Simulink se lance en tapant simplement Simulink sur le terminal de Matlab. Un nouveau modèle se crée avec File -> New -> Model, qui affiche une fenêtre blanche, sur laquelle on pourra déposer les différents blocs (par drag and drop) et les lier entre eux à l'aide de la souris.

Les blocs sont groupés dans différents groupes :

- ❖ **Sources** : permet de spécifier des sources d'input à un système (sinusoïde, constante, rampe, séquence répétée...).
- ❖ **Sinks** : permet de spécifier des outputs. Nous utiliserons principalement To **Workspace**, qui permet d'envoyer le résultat au **workspace** de Matlab, et **Scope** qui permet d'afficher le résultat directement dans un graphique.
- ❖ **Divers** : le bloc **Fcn** (Groupe User Defined Functions) permet de décrire soi-même une fonction, le bloc **State Space** (Groupe Continuous) un système simple d'équations différentielles. Le bloc **DEE** « Differential equation editor » permet d'utiliser des équations différentielles arbitraires. Pour pouvoir l'utiliser, taper **dee** sur le terminal Matlab.

V-4-3-SCHEMA BLOC SOUS SIMULINK DE LA TURBINE :

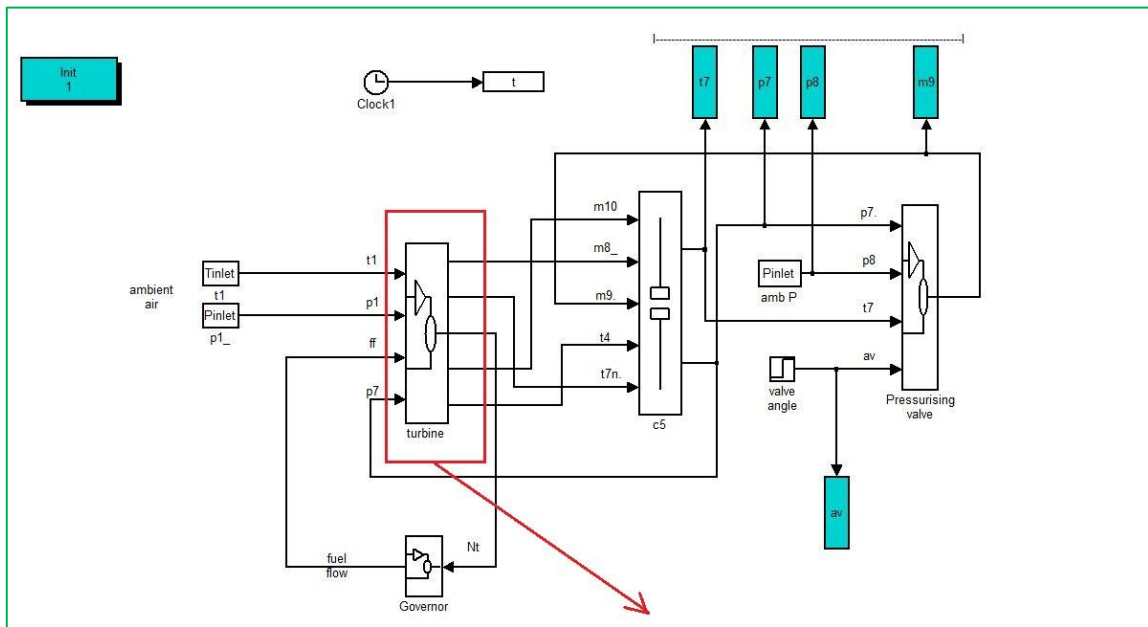


Figure V- 2 : Représentation du bloc turbine à gaz sous Simulink

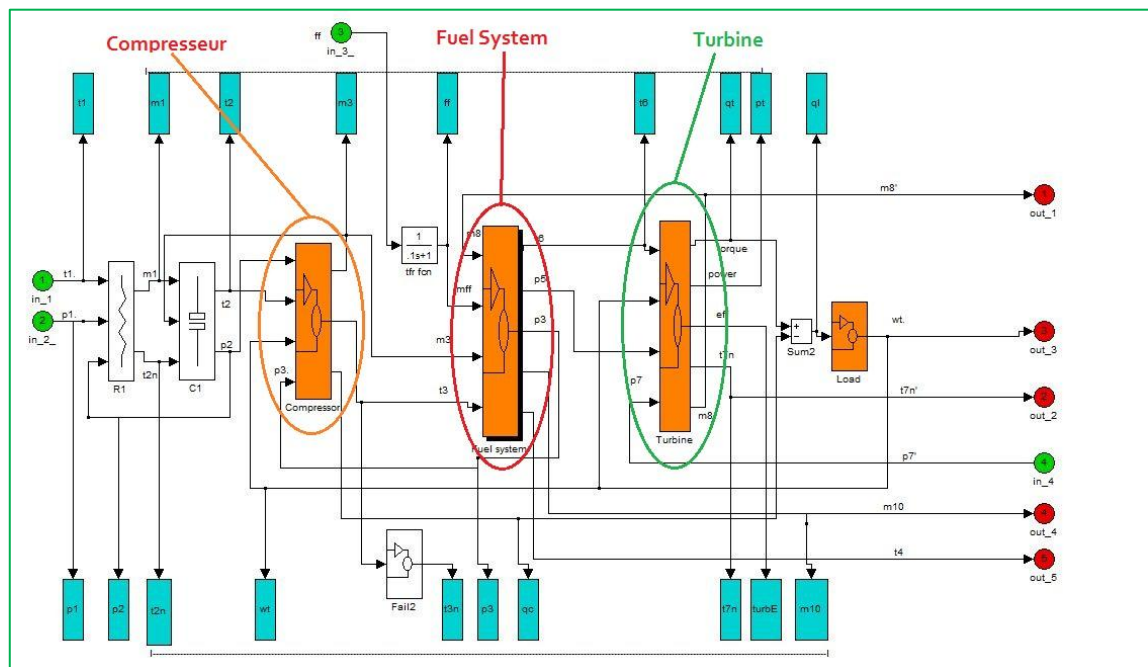


Figure V- 3 : Développement du bloc turbine (compresseur+fuel system+turbine)

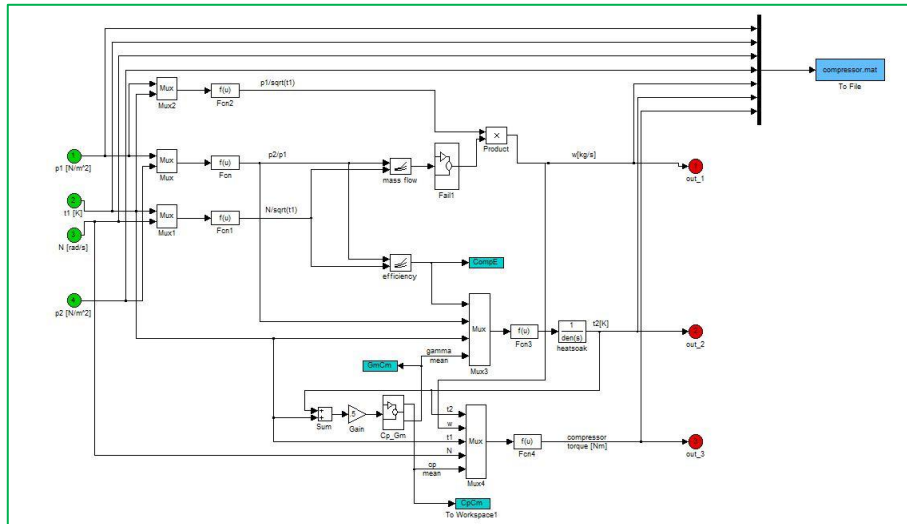


Figure V- 4 : Représentation du bloc « compresseur »

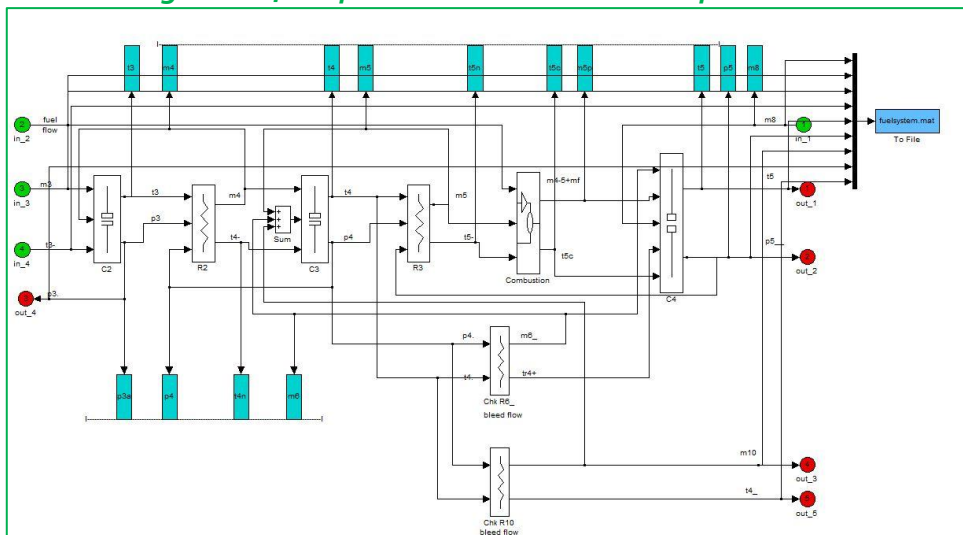


Figure V- 5 : Représentation du bloc « installation carburant »

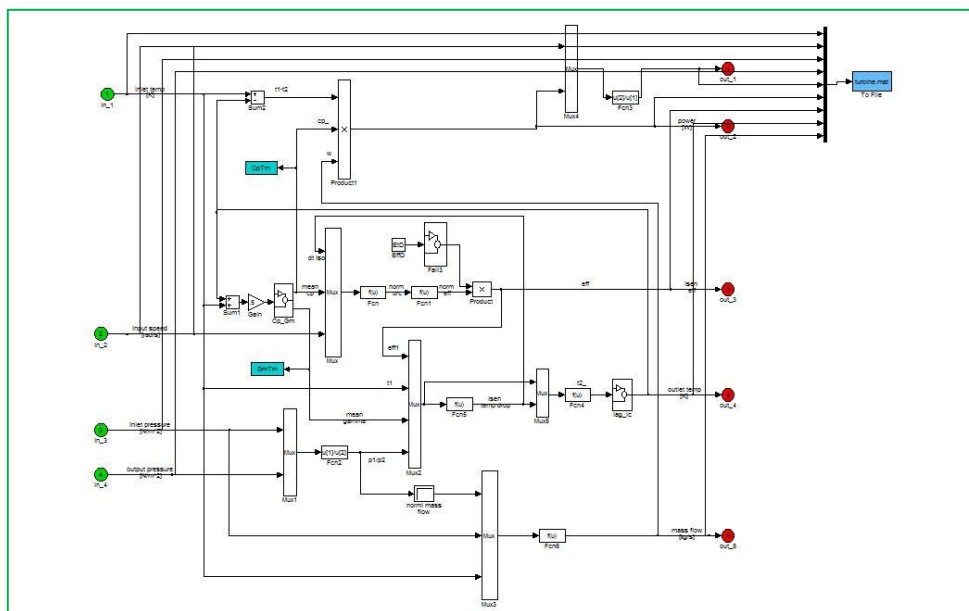


Figure V- 6 : Représentation du bloc « turbine »

V-5-COMPRESSEUR :



Figure V- 7: Le compresseur

$p_2(N/m^2)$: Pression ambiante.

$t_2(k)$: Température ambiante.

$wt(rad/s)$: Vitesse d'axe de turbine.

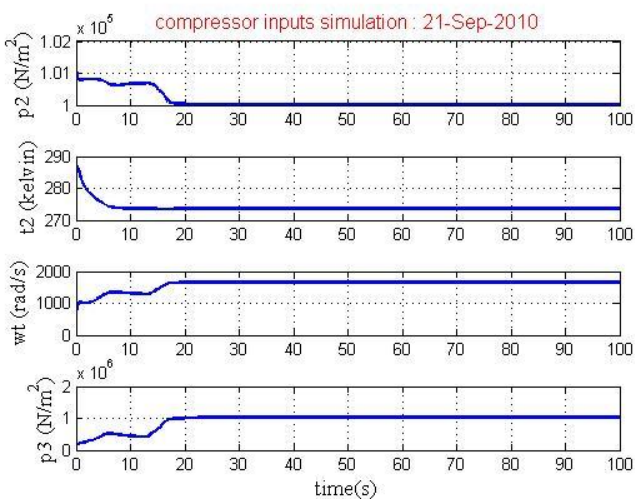
$p_3(N/m^2)$: Pression de retour.

$m_3(kg/s)$: Ecoulement de masse

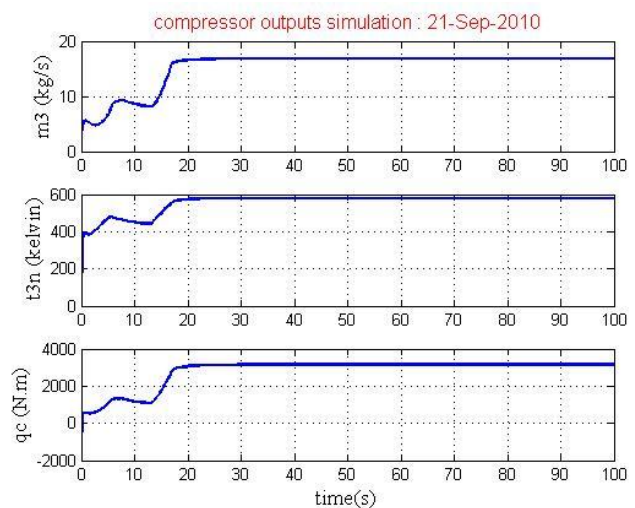
$t_3(k)$: Température de gaz à la sortie du compresseur.

$q_c(N.m)$: Couple du compresseur.

V-6-LES ENTREES-SORTIES DU COMPRESSEUR :



a- Les entrées



b- Les sortie

Figure V- 8: Les entrées-sorties du compresseur

❖ La figure V-8-a montre l'évolution de l'entrée appliquée au compresseur :

Entré P2 : Après un régime transitoire de **17,19 second** la valeur de la grandeur pression se stabilise à la valeur de **10⁵ Pa**, rappelant que **1 Pa=1 N/m²**, c'est bien la pression ambiante.

Entré t_2 : Après un régime transitoire de **07,44 second** la valeur de la grandeur température se stabilise à la valeur de **273,9 Kelvin**, c'est la température ambiante (273,15 kelvin=0 Celsius).

Entré w_t : Après un régime transitoire de **17,31 second** la valeur de la vitesse d'axe du turbine se stabilise à la valeur de **1635 rad/s**, Pour la mise en route du compresseur, on utilise un moteur de lancement qui joue le rôle de démarreur

Entré P_3 : Après un régime transitoire de **17,31 second** la valeur de la grandeur pression se stabilise à la valeur de **9,726x10⁵ Pa**, c'est la pression de retour.

❖ La figure V-8-b montre l'évolution des sorties du compresseur :

Sortie m_3 : Après un régime transitoire de **17,73 second** la valeur de la grandeur de l'écoulement de masse se stabilise à la valeur de **16,27 kg/s**.

Sortie t_{3n} : Après un régime transitoire de **17,73 second** la valeur de la température du gaz à la sortie du compresseur se stabilise à la valeur de **566,70 kelvin**.

Sortie q_c : Après un régime transitoire de **17,73 second** la valeur de la grandeur de l'écoulement de masse se stabilise à la valeur de **16,27 kg/s**.

V-7-TURBINE :



Figure V- 9: La turbine

$t_6(k)$: Inlet temperature.

$q_t(N.m)$: Turbine torque.

$w_t(rad/s)$: Input speed.

$p_t(watt)$: The power.

$p_5(N/m^2)$: Inlet pressure.

$t_7(k)$: Outlet temperature.

$p_7(N/m^2)$: Output pressure.

$m_8(kg/s)$: Mass flow.

V-8-LES ENTREES-SORTIES DE LA TURBINE :

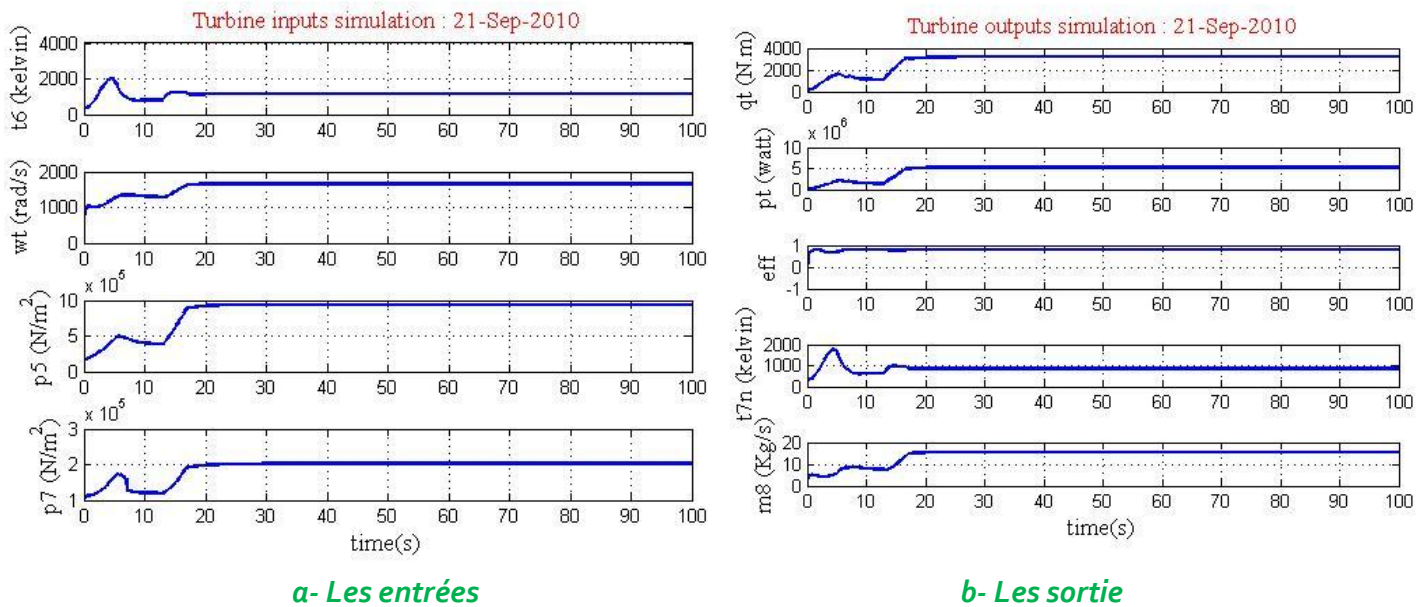


Figure V- 10:Les entrées-sorties de la turbine

❖ La figure V-10-a montre l'évolution de l'entrée appliquée à la turbine :

Entré t_6 : Après un régime transitoire de **16,70 second** la valeur de la grandeur température (inlet température) se stabilise à la valeur de **1224 kelvin**.

Entré w_t : Après un régime transitoire de **17,33 second** la vitesse de rotation de la turbine se stabilise à la valeur de **1630 rad/s**.

Entré p_5 : Après un régime transitoire de **17,31 second** la valeur de la pression à l'entrée de la turbine se stabilise à la valeur de **$8,92 \times 10^5 \text{N/m}^2$** .

Entré P_7 : Après un régime transitoire de **19,17 second** la valeur de la pression à la sortie de la turbine se stabilise à la valeur de **$1,965 \times 10^5 \text{N/m}^2$** .

❖ La figure V-10-b montre l'évolution des sorties du turbine :

Sortie q_t : Après un régime transitoire de **16,86 second** la valeur du couple appliqué par la turbine se stabilise à la valeur de **3065 N.m**.

Sortie p_t : Après un régime transitoire de **17,31 second** la valeur de la puissance à la sortie du turbine se stabilise à la valeur de **$5,003 \times 10^5 \text{ watt}$** (Un watt est la puissance d'un système énergétique dans lequel une énergie de 1 joule est transférée uniformément pendant 1 seconde)

Sortie t7n : Après un régime transitoire de **17,73 second** la valeur de la température du gaz à la sortie du turbine se stabilise à la valeur de **849,40 kelvin**.

Sortie m8 : Après un régime transitoire de **18,06 second** la valeur de la grandeur de l'écoulement de masse se stabilise à la valeur de **15,10 kg/s**.

V-9-INSTALLATION CARBURANT :



Figure V- 11: Fuel system

$m_8(kg/s)$: Mass flow.

$m_{ff}(kg/s)$: Fuel flow.

$m_3(kg/s)$: Mass flow.

$t_3(k)$: Input temperature.

$t_6(k)$: Output temperature.

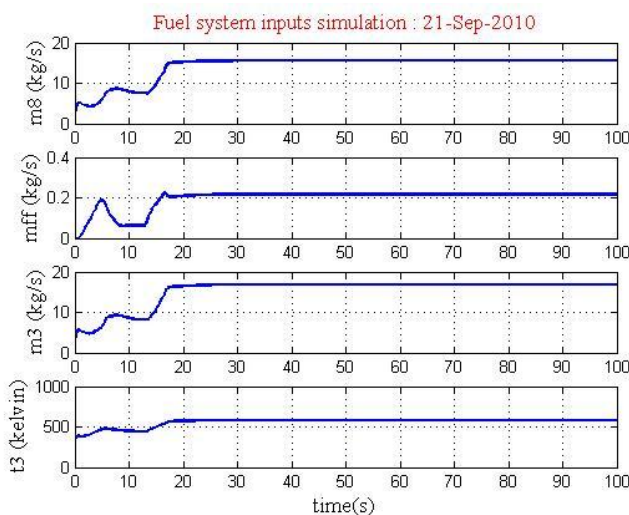
$p_5(N/m^2)$: Output pressure.

$p_3(N/m^2)$: Output pressure.

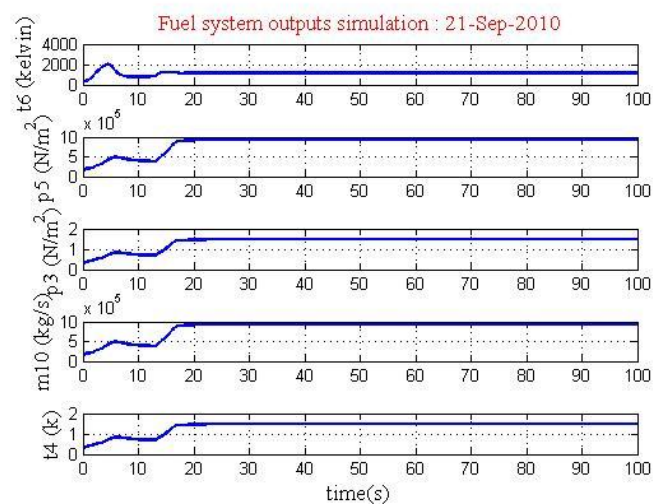
$m_{10}(kg/s)$: Mass flow.

$t_4(k)$: Output pressure.

V-10-LES ENTREES -SORTEES DE L'INSTALLATION CARBURANT :



a- Les entrées



b- Les sortie

Figure V- 12 : Les entrées-sortie de l'installation carburant

❖ La figure V-12-a montre l'évolution de l'entrée appliquée à l'installation carburant :

Entré m8 : Après un régime transitoire de **17,32 second** la valeur de l'écoulement de masse se stabilise à la valeur de **14,89kg/s**.

Entré mff : Après un régime transitoire de **17,33 second** la valeur de l'écoulement de masse du carburant se stabilise à la valeur de **0,206kg/s**.

Entré m3 : Après un régime transitoire de **17,31 second** la valeur de l'écoulement de masse se stabilise à la valeur de **16,10 kg/s**.

Entré t3 : Après un régime transitoire de **19,17second** la valeur de la température à l'entrée de l'installation carburant se stabilise à la valeur de **561,60 kelvin**.

❖ La figure V-12-b montre l'évolution des sorties de l'installation carburant :

Sortie t6 : Après un régime transitoire de **17,31 second** la valeur de la température à la sortie de l'installation carburant se stabilise à la valeur de **1129 kelvin**.

Sortie p5 : Après un régime transitoire de **17,31 second** la valeur de la pression à la sortie de l'installation carburant se stabilise à la valeur de **8,92x10⁵Pa**

Sortie p3 : Après un régime transitoire de **17,31 second** la valeur de la pression à la sortie de l'installation carburant se stabilise à la valeur de **1,424x10⁵Pa**

Sortie m10 : Après un régime transitoire de **18,06 second** la valeur de la grandeur de l'écoulement de masse à la sortie de l'installation carburant se stabilise à la valeur de **8,92x10⁵ kg/s**.

Sortie t4 : Après un régime transitoire de **17,31second** la valeur de la température à la sortie de l'installation carburant se stabilise à la valeur de **1,424 kelvin**.

V-11-PRESENTATION PAR MULTI-MODELES :

V-11-1-ALGORITHME D'IDENTIFICATION :

Le schéma synoptique de la figure V-9 résume la procédure d'identification d'un multi-modèle décrit par une grille (recherche des variables candidates, recherche du nombre de partitions à associer à ces variables, optimisation des paramètres).

La démarche adoptée relève de la stratégie incrémentale. Elle démarre avec un modèle unique linéaire ou affine et un ensemble de variables candidates $z_j, (j=1, \dots, nz)$ dont le nombre de partitions p_j est initialisé à 1. La procédure consiste ensuite à ajuster la structure de façon heuristique en augmentant progressivement le nombre de divisions des différentes variables z_j .

L'itération n°1 est l'itération d'initialisation. Le critère correspondant est celui enregistré avec le meilleur modèle linéaire ou affine. A chaque itération on incrémente le nombre de partitions d'une et une seule variable z_j . Les paramètres de la grille sont optimisés par l'algorithme présenté dans la figure IV-2. On calcule ensuite le critère J_g correspondant à ce multi-modèle. On procède de la même façon pour toutes les autres variables candidates. Le multi-modèle retenu est celui ayant fourni le meilleur critère J_g . On développe ensuite le modèle de meilleure qualité, ce qui évite d'explorer toutes les branches. La recherche s'arrête si on atteint le minimum de J_g ou si ce dernier évolue peu.

L'algorithme est également une technique à deux niveaux. Il consiste à estimer les paramètres β des fonctions de validité par la méthode de Levenberg-Marquardt pour les paramètres θ fixés, puis à ajuster les paramètres θ des modèles locaux également par un algorithme de Levenberg-Marquardt pour les paramètres β précédemment déterminés.

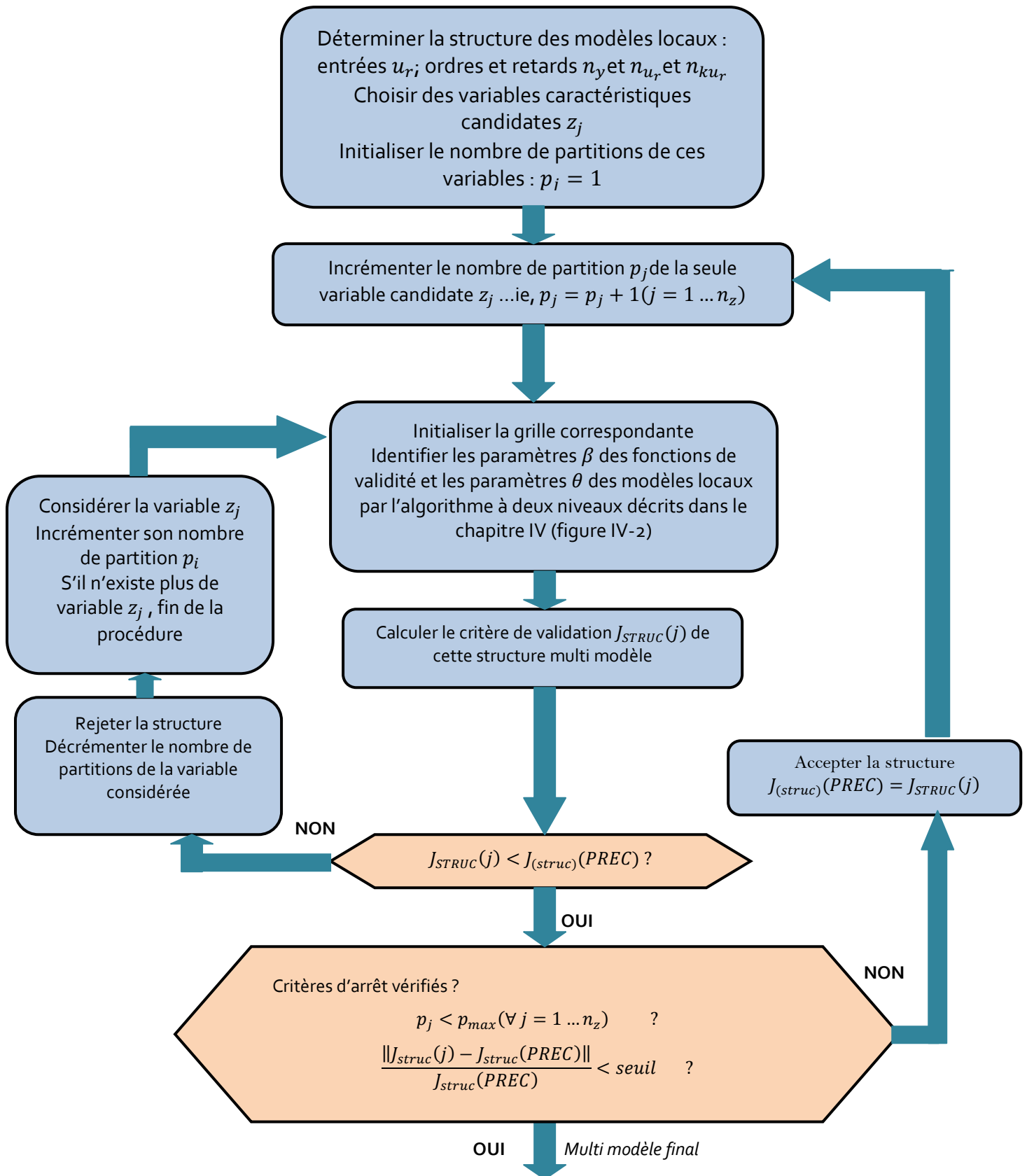


Figure V- 13: Algorithme récapitulant la procédure d'identification d'un multi-modèle

V-11-2-LES VARIABLES CARACTERISTIQUES CANDIDATES:

Les variables caractéristiques candidates que nous avons choisies sont les entrées de chaque éléments du centrale (turbine, compresseur, installation carburant).

La structure initiale des modèles locaux est choisie de la forme :

$$y_i(t) = -a_1^{(i)}y_s(t-1) - a_2^{(i)}y_s(t-2) - a_3^{(i)}y_s(t-3) + b_{11}^{(i)}u_1(t-1) + b_{21}^{(i)}u_1(t-2) \\ + b_{31}^{(i)}u_1(t-3) + b_{12}^{(i)}u_2(t-1) + b_{22}^{(i)}u_2(t-2) + b_{32}^{(i)}u_2(t-3) + b_{13}^{(i)}u_3(t-1) \\ + b_{23}^{(i)}u_3(t-2) + b_{33}^{(i)}u_3(t-3) + b_{14}^{(i)}u_4(t-1) + b_{24}^{(i)}u_4(t-2) + b_{34}^{(i)}u_4(t-3)$$

Notre identification est de type boîte noir, pour la première itération, le nombre de partition sur chaque support caractéristique vaut 1.

Le nombre de partitions p est choisi tel que $1 \leq p \leq 9$; rappelons que pour un nombre de partitions $p=9$ et un nombre de variable $n_z = 3$, le nombre de modèles $M=9^3=729$.

V-12-IDENTIFICATION ET OPTIMISATION PARAMETRIQUES:

V-12-1-COMPRESSEUR:

V-12-1-1-a-La sortie m_3 (écoulement de masse) :

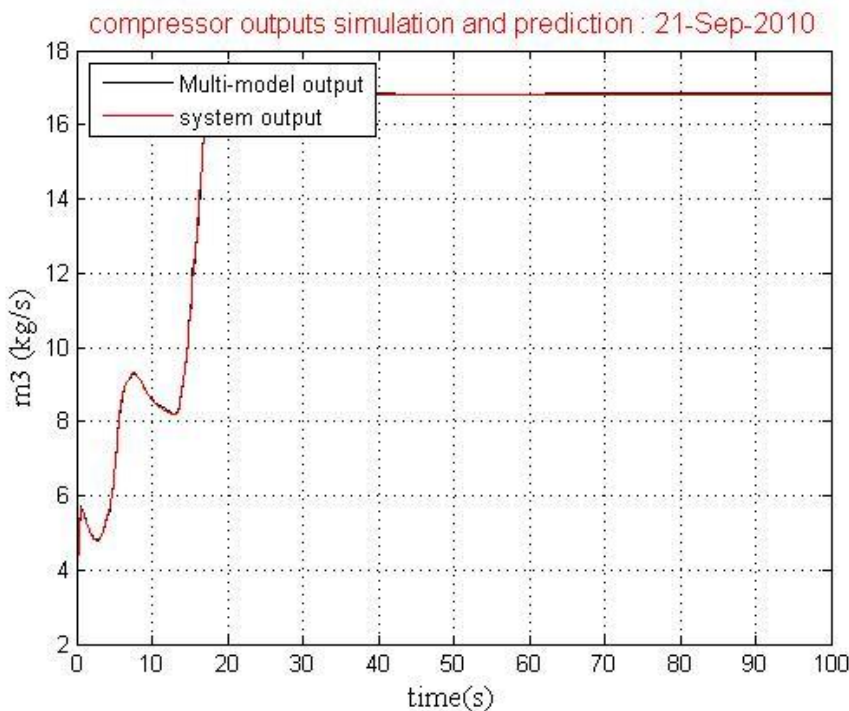
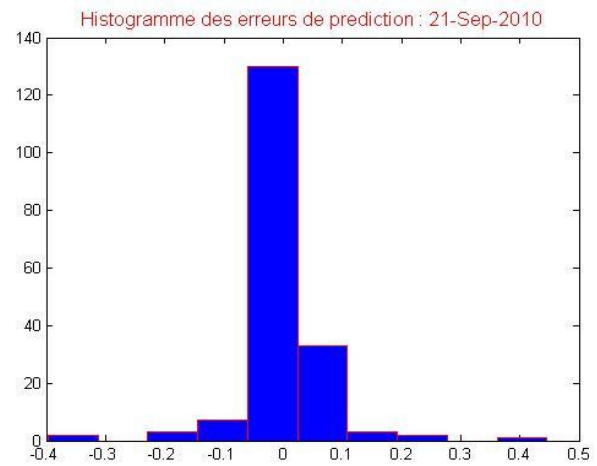
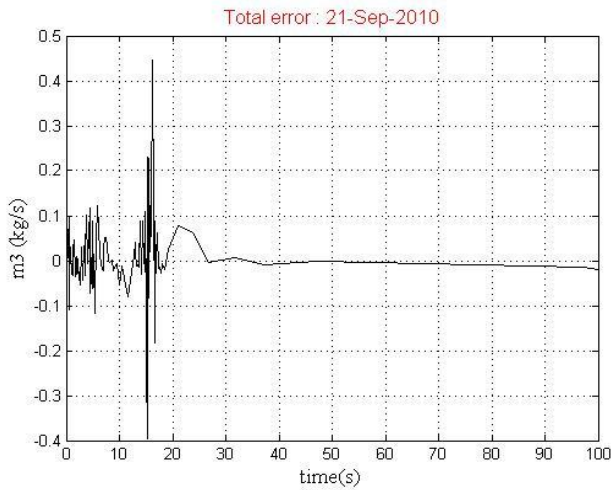


Figure V- 14: La sortie m_3 (mass flow) du compresseur avec la sortie du Multi-modèle



a- Les erreurs

b- Histogramme

Figure V-15: Erreur et Histogramme de prédiction

❖ Discussion :

Le nombre de modèle locaux qui nous a permis d’obtenir cette représentation et après optimisation est **M=16 modèles locaux**, dont les fonctions de validité sont construites à partir de fonctions gaussienne cela se traduit par 16 zones de fonctionnement du système.

Le vecteur « **comb** » qui représente le nombre de partition sur chaque support de variable caractéristique est :

$$comb = [2 \ 4 \ 1 \ 2] ;$$

On a alors 2 partitions sur le support u_1 , 4 sur le support u_2 , 1 sur le support u_3 , et 2 sur u_4 .

Les seize modèles obtenues sont représenté sur les figures V-16 :

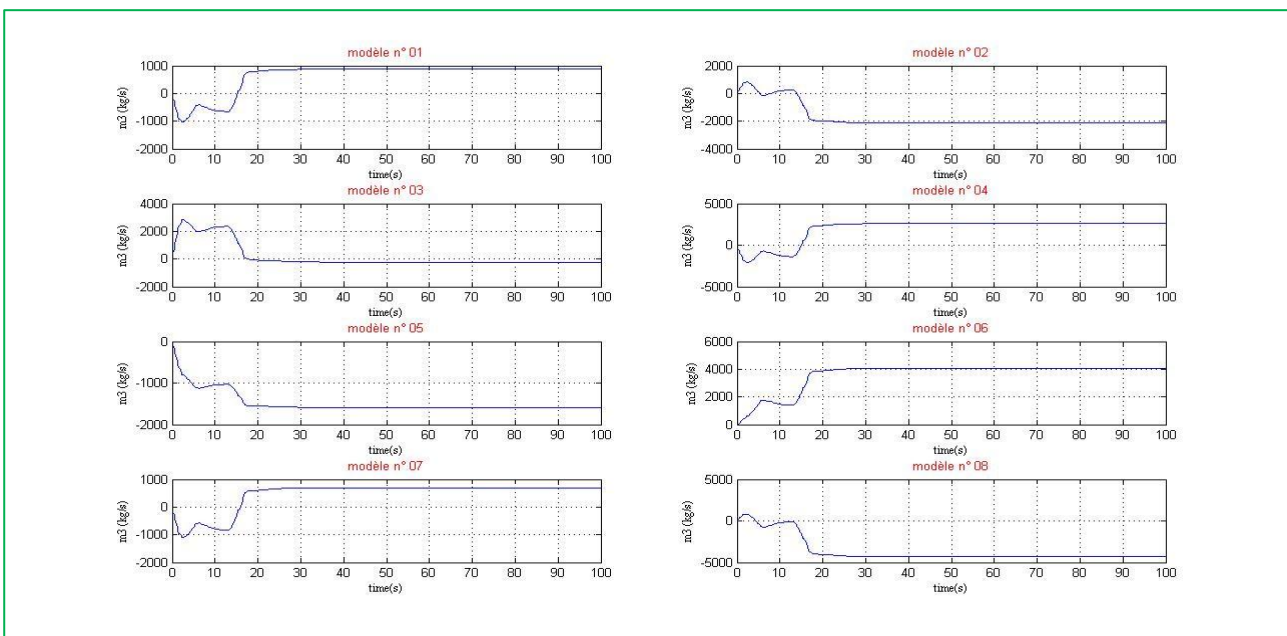


Figure V-16: Les sorties des 8 premiers modèles

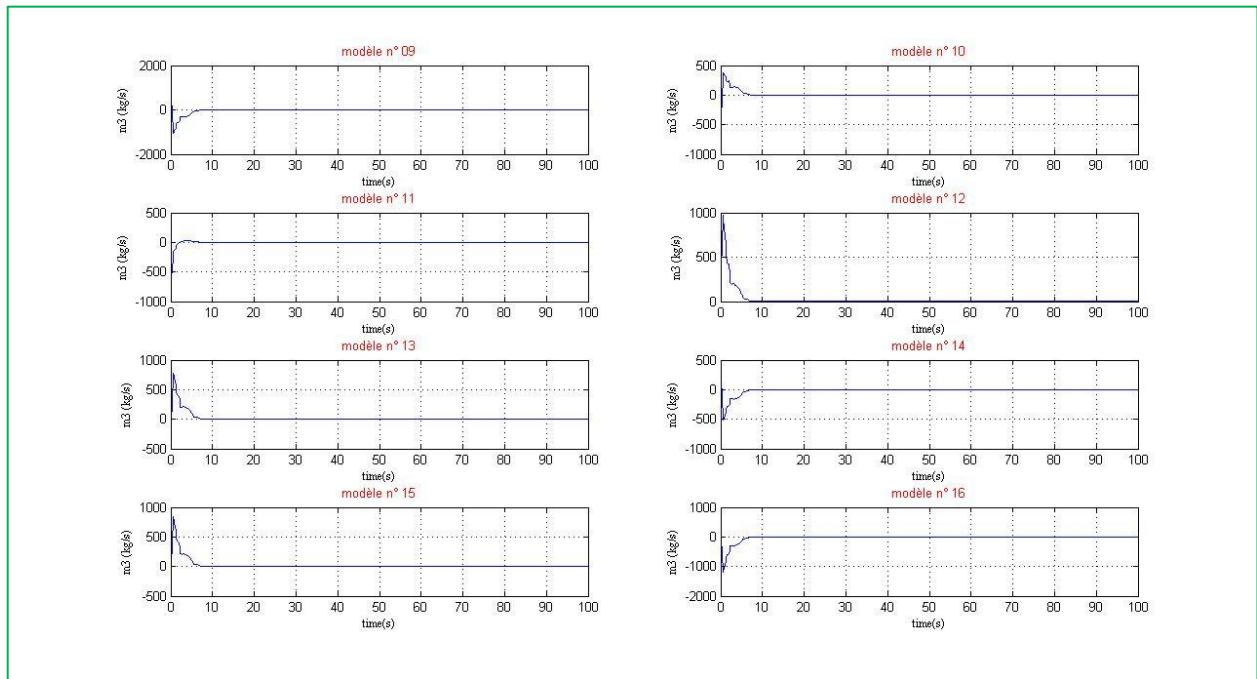


Figure V-17: Les sorties des 8 derniers modèles

Ainsi les paramètres du vecteur θ sont regroupés dans le tableau V-2 :

modèle locale f_i	a_1	a_2	b_{11}	b_{12}	b_{13}	b_{14}	b_0
f_1	-1.2590	0.3818	-0.1690	-0.1710	-0.0475	-0.0256	0.0002
f_2	-1.2637	0.3975	0.1816	-0.1316	-0.0638	0.0156	0.0001
f_3	-1.3062	0.4060	0.3772	-0.1477	-0.0726	0.0541	0.0002
f_5	-1.3105	0.4251	-0.3619	-0.1094	-0.0006	-0.0282	0.0001
f_6	-1.2942	0.4117	-0.0554	-0.1246	-0.0492	-0.0139	0.0001
f_7	-1.2980	0.4330	-0.0345	-0.0844	0.1144	0.0179	0.0001
f_8	-1.2852	0.4112	-0.1699	-0.1188	-0.0373	-0.0210	0.0001
f_9	-1.2888	0.4328	0.2351	-0.0778	0.1486	0.0004	0.0001
f_{10}	0.0537	0.5908	0.3900	-0.0285	-0.2222	-0.2274	-0.0003
f_{11}	0.0493	0.5603	-0.1939	-0.0219	-0.0359	0.1096	-0.0003
f_{12}	0.0523	0.5890	-0.1123	-0.0249	-0.0932	0.0518	-0.0003
f_{13}	0.0480	0.5587	-0.1287	-0.0168	0.0883	0.0916	-0.0003
f_{14}	0.0493	0.5652	-0.2138	-0.0203	0.0120	0.1334	-0.0003
f_{15}	0.0451	0.5361	0.1851	-0.0112	0.1826	-0.1096	-0.0002
f_{16}	0.0483	0.5572	-0.2008	-0.0190	0.0348	0.1278	-0.0003

Tableau V- 2 : Paramètres du multi-modèle descriptif de la sortie m_3 du compresseur

La représentation mathématique de la sortie m_3 du compresseur est de la forme :

$$m_3^i(t) = -a_1^i m_3^i(t-1) - a_2^i m_3^i(t-2) + b_{11}^i p_2(t) + b_{12}^i t_2(t) + b_{13}^i w_t(t) + b_{14}^i p_3(t) + b_0^i$$

Où les paramètres a_i^i , b_i^i , b_0^i sont déduits du tableau V-2.

À partir des figures V-15, nous pouvons constater que l'erreur d'estimation est très petite relativement au signal de sortie m_3 .

Les modèles locaux que nous avons établis dans la section précédente sont présentés sur les figures (V-12), (V-13), en effet ces modèles traduisent le fonctionnement du système dans 16 zones de fonctionnements (modes de fonctionnements).

Concernant les figures V-16, V-17, on remarque l'absence des zones vides, et l'existence des zones redondantes (zones voisines pouvant décrire un même comportement du système), cela apparaît dans les modèles (2,3), (7,4), (14,16), (13,15).

Afin de remédier à cet état de fait on peut fusionner les modèles redondantes (à titre d'exemple on arrive à une réduction de moitié du nombre de modèle).

V-12-1-1-b-La sortie t3 (Température de gaz à la sortie du compresseur) :

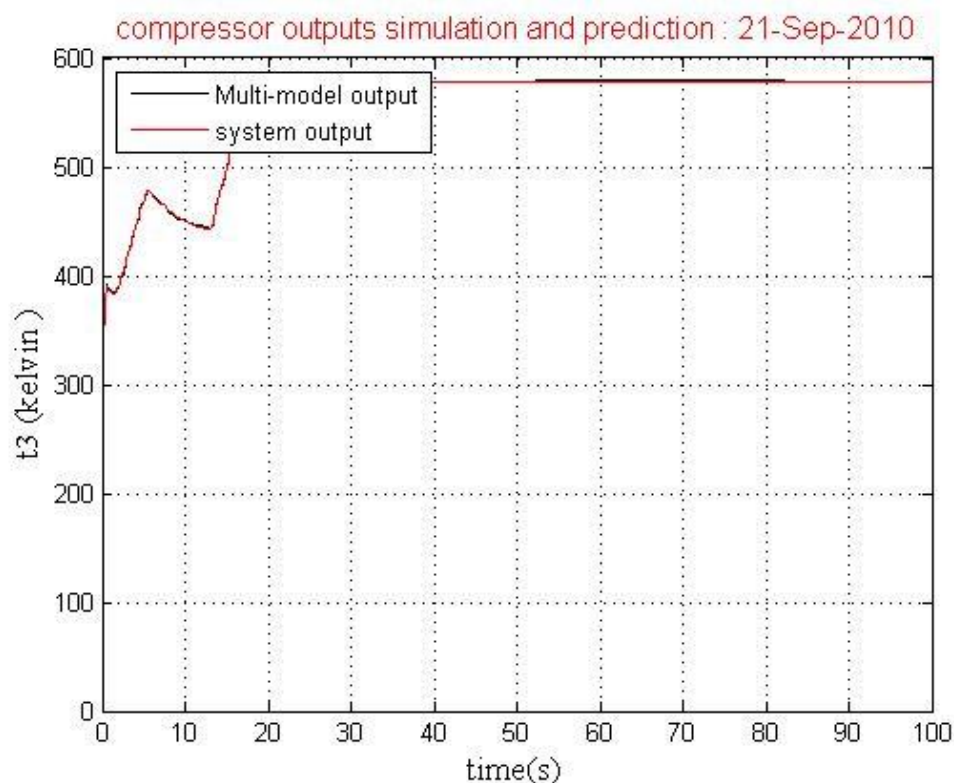
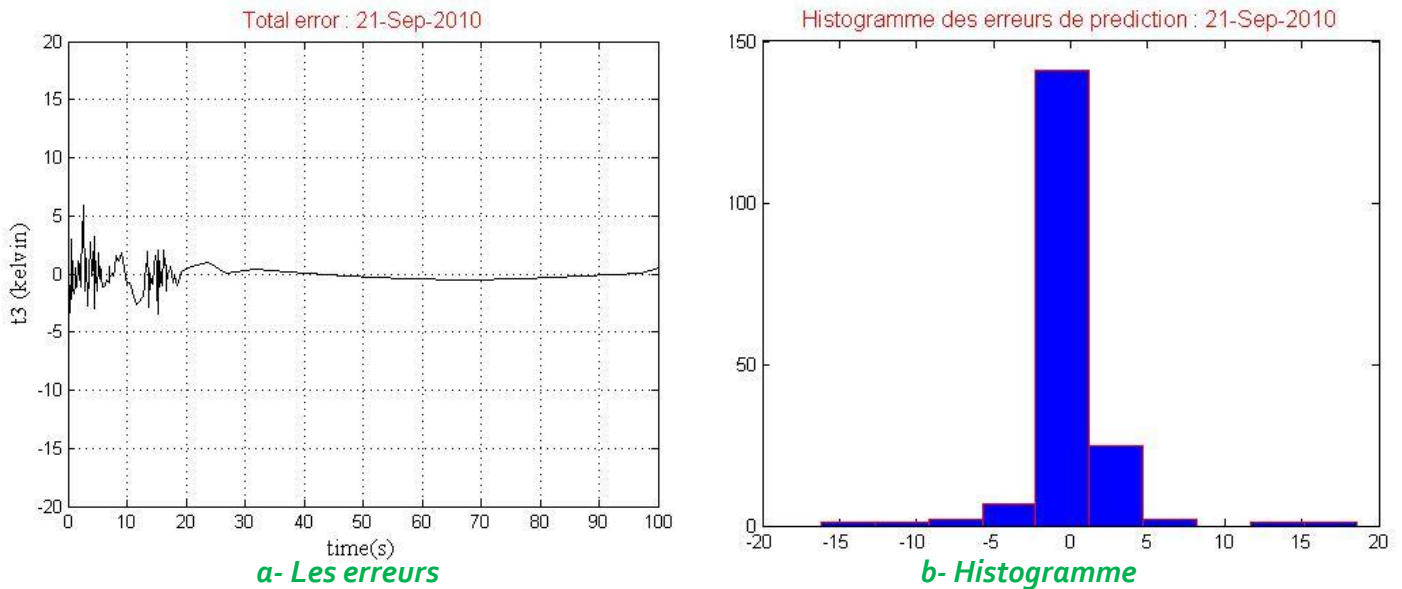


Figure V- 18: La sortie t3 (température) du compresseur avec la sortie du Multi-modèle



a- Les erreurs
b- Histogramme
Figure V-19: Erreur et Histogramme de prédiction

❖ **Discussion :**

Le nombre de modèle locaux qui nous a permis d’obtenir cette représentation et après optimisation est **M=08 modèles locaux**, dont les fonctions de validité sont construites à partir de fonctions gaussienne cela se traduit par 08 zones de fonctionnement du système.

Le vecteur « **comb** » qui représente le nombre de partition sur chaque support de variable caractéristique est : $comb = [2 \ 2 \ 1 \ 2]$;

Donc, 2 partitions sur le support u_1 , 2 sur le support u_2 , 1 sur le support u_3 , et 2 sur u_4 .

Les huit modèles obtenues sont représenté sur les figures V-20 :

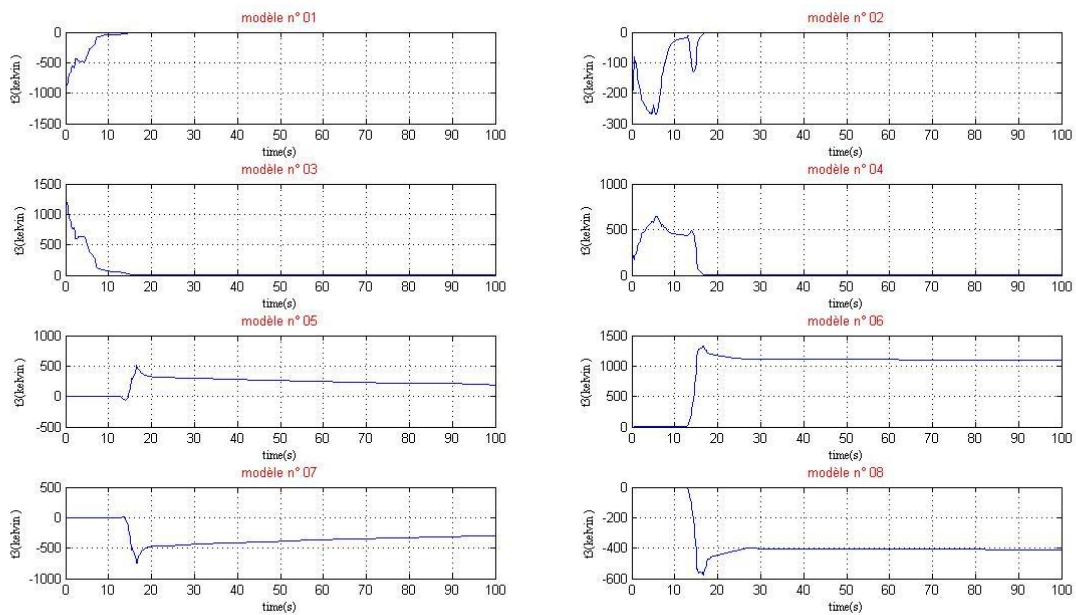


Figure V-20: Les sorties des 8 modèles

Ainsi les paramètres du vecteur θ sont regroupés dans le tableau V-3 :

modèle locale f_i	a_1	a_2	b_{11}	b_{12}	b_{13}	b_{14}	b_0
f_1	14,1452	-09,1426	00,0655	-31,6163	07,8338	-00,0302	-02,7458
f_2	00,5302	02,0892	00,0872	-42,9108	05,2079	-00,0062	-00,2933
f_3	-19,3480	11,8399	-00,1094	49,8621	-08,0092	00,0278	-02,3121
f_5	-02,1304	-01,1369	-00,0299	19,8293	-03,8555	00,0043	-00,1426
f_6	01,0220	-00,1627	-08,4301	-00,0183	00,7318	01,4889	-00,0001
f_7	26,7346	-05,0414	00,0175	04,1853	09,3275	-00,0024	-00,0059
f_8	01,6634	-00,3319	04,2179	00,0213	01,2976	-00,8978	00,0000

Tableau V-3 : Paramètres du multi-modèle descriptif de la sortie t_3 du compresseur

La représentation mathématique de la sortie t_3 du compresseur est de la forme :

$$t_3^i(t) = -a_1^i t_3^i(t-1) - a_2^i t_3^i(t-2) + b_{11}^i p_2(t) + b_{12}^i t_2(t) + b_{13}^i w_t(t) + b_{14}^i p_3(t) + b_0^i$$

Où les paramètres a_i^i , b_i^i , b_0^i sont déduits du tableau V-3.

La sortie t_3 présente la température du gaz à la sortie du compresseur, cette grandeur est accessible (mesurable), l'unité de mesure est le Kelvin.

À partir des figures V-19-a, nous pouvons constater que l'erreur d'estimation est très petite relativement au signal de sortie t_3 .

Les modèles locaux que nous avons établis dans la section précédente sont présentés sur les figures (V-20), en effet ces modèles traduisent le fonctionnement du système dans 08 zones de fonctionnements (modes de fonctionnements).

Le multi modèle que nous avons obtenu représente bien la température t_3 à la sortie du compresseur.

V-12-1-1-c-La sortie q_c (couple du compresseur) :

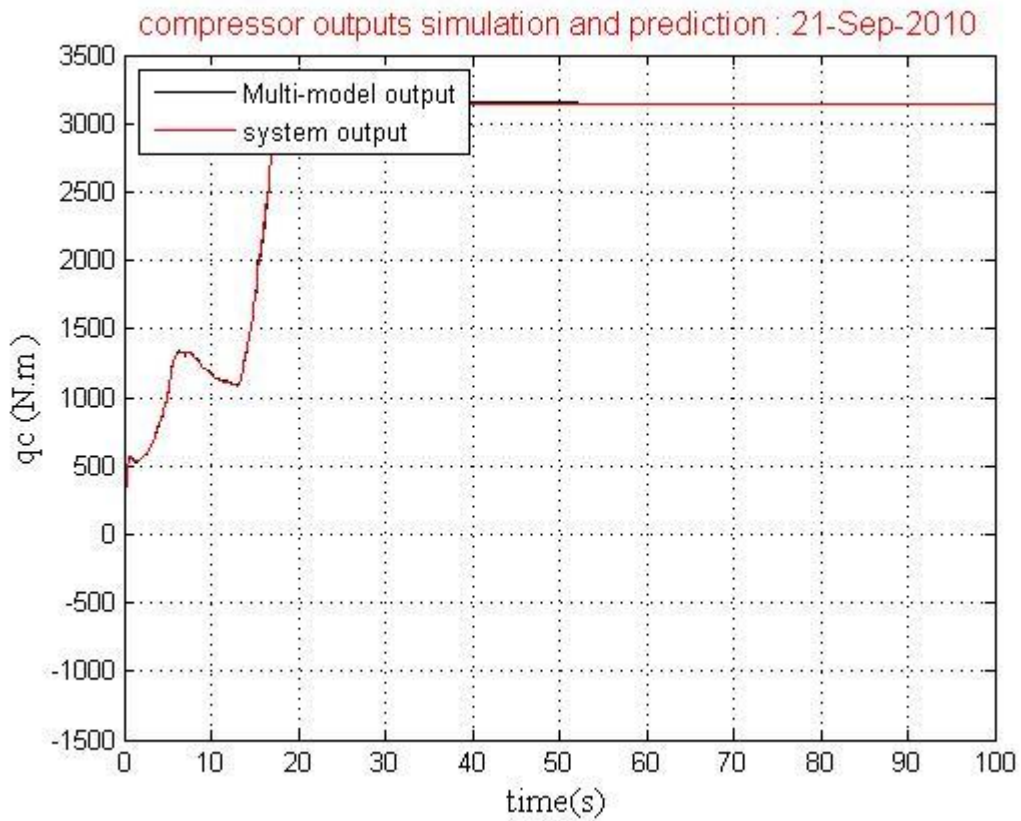


Figure V- 21: La sortie q_c (couple) du compresseur avec la sortie du Multi-modèle

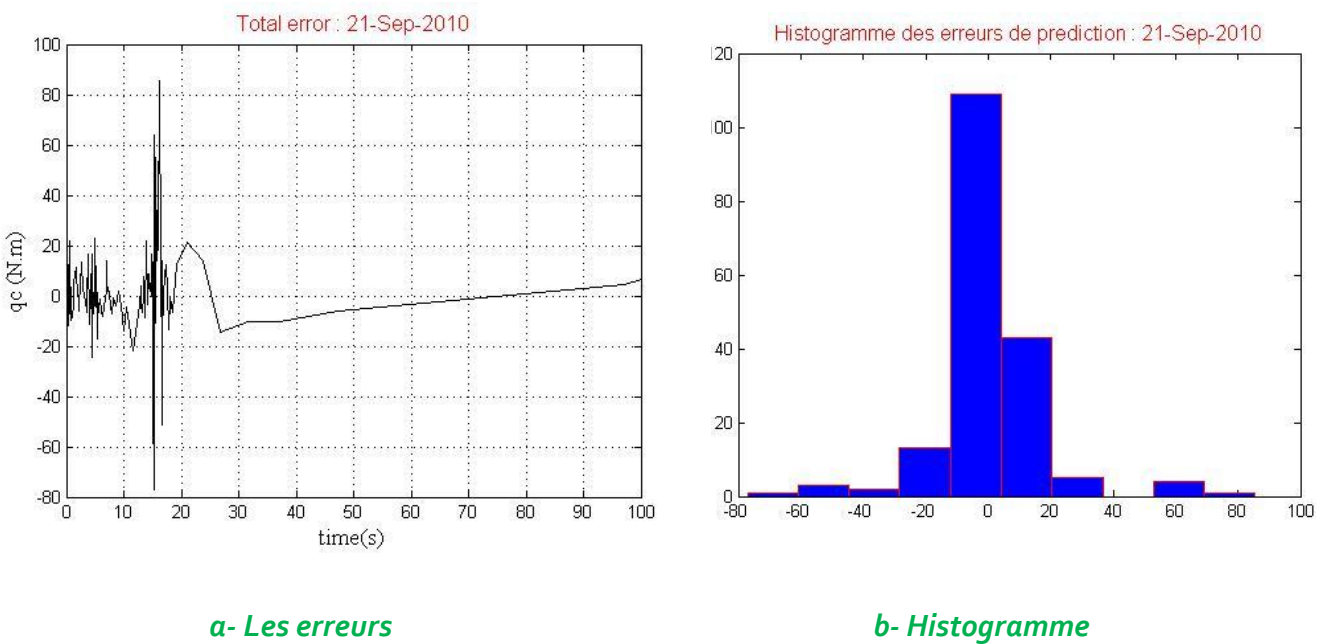


Figure V- 22: Erreur et Histogramme de prédiction

❖ **Discussion :**

Le nombre de modèle locaux qui nous a permis d’obtenir cette représentation et après optimisation est **M=04 modèles locaux**, dont les fonctions de validité sont construites à partir de fonctions gaussienne cela se traduit par 04 zones de fonctionnement du système.

Le vecteur « **comb** » qui représente le nombre de partition sur chaque support de variable caractéristique est :

$$comb = [1 \ 2 \ 2 \ 1] ;$$

On a alors, une partition sur le support u_1 , 2 sur le support u_2 , 2 sur le support u_3 , et une partition sur u_4 .

Les quatre modèles obtenues sont représenté sur la figure V-23 :

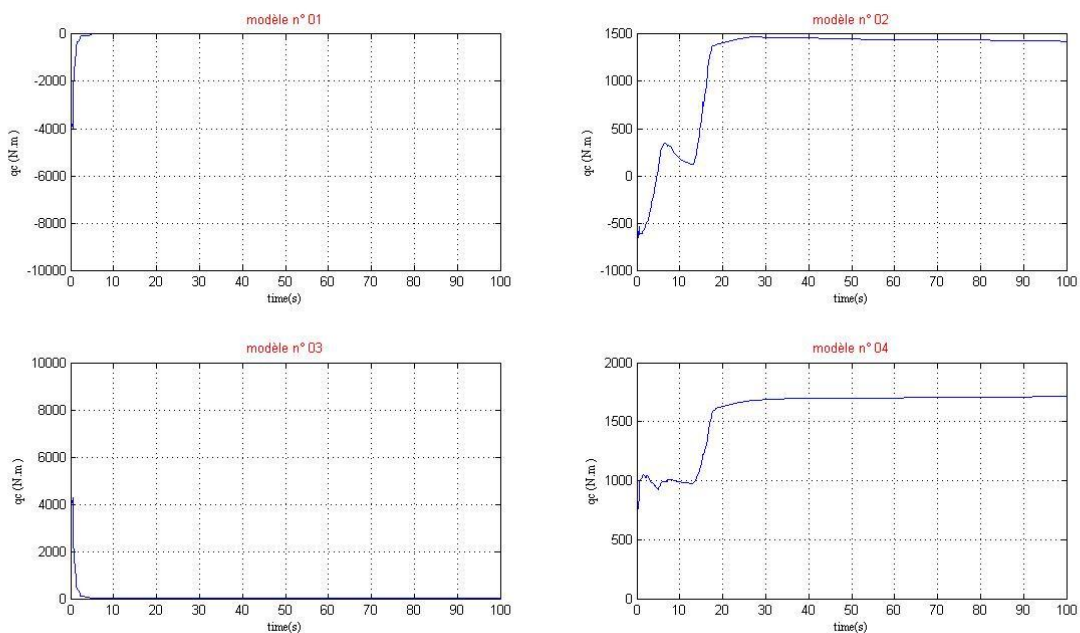


Figure V-23: Les sorties des 4 modèles

Ainsi les paramètres du vecteur θ sont regroupés dans le tableau V-4 :

modèle locale f_i	a_1	a_2	b_{11}	b_{12}	b_{13}	b_{14}	b_0
f_1	-19,6948	18,7702	01,1251	-313,3508	21,1757	-00,3755	-02,1067
f_2	-02,6572	01,1227	00,1156	-51,0456	00,2522	00,0015	-04,4631
f_3	20,5320	-20,2670	00,5479	-268,9121	-08,0195	00,3114	-01,8814
f_4	-00,7027	00,0047	-00,1249	56,4920	-01,6150	00,0001	-08,3682

Tableau V- 4 : Paramètres du multi-modèle descriptif de la sortie qc du compresseur

La représentation mathématique de la sortie q_c du compresseur est de la forme :

$$q_c^i(t) = -a_1^i q_c^i(t-1) - a_2^i q_c^i(t-2) + b_{11}^i p_2(t) + b_{12}^i t_2(t) + b_{13}^i w_t(t) + b_{14}^i p_3(t) + b_0^i$$

Où les paramètres a_i^i , b_i^i , b_0^i sont déduits du tableau V-4.

La sortie q_c présente le couple de l'arbre du compresseur, cette grandeur est accessible (mesurable), l'unité de mesure est le **(N.m)**

À partir des figures V-22-a, nous pouvons constater que l'erreur d'estimation est très petite relativement au signal de sortie q_c .

Les modèles locaux que nous avons établis dans la section précédente sont présentés sur les figures **(V-23)**, en effet ces modèles traduisent le fonctionnement du système dans 04 zones de fonctionnements (modes de fonctionnements).

Concernant les figures V-23, on remarque l'existence des zones vides, cela apparaît dans les modèles **(1,2)** et l'absence des zones redondantes (zones voisines pouvant décrire un même comportement du système).

Afin de remédier à cet état de fait on peut éliminer les zones vides (à titre d'exemple on arrive à une réduction de moitié du nombre de modèle).

.

V-12-2-TURBINE:

V-12-2-1-a-La sortie q_t (couple de la turbine) :

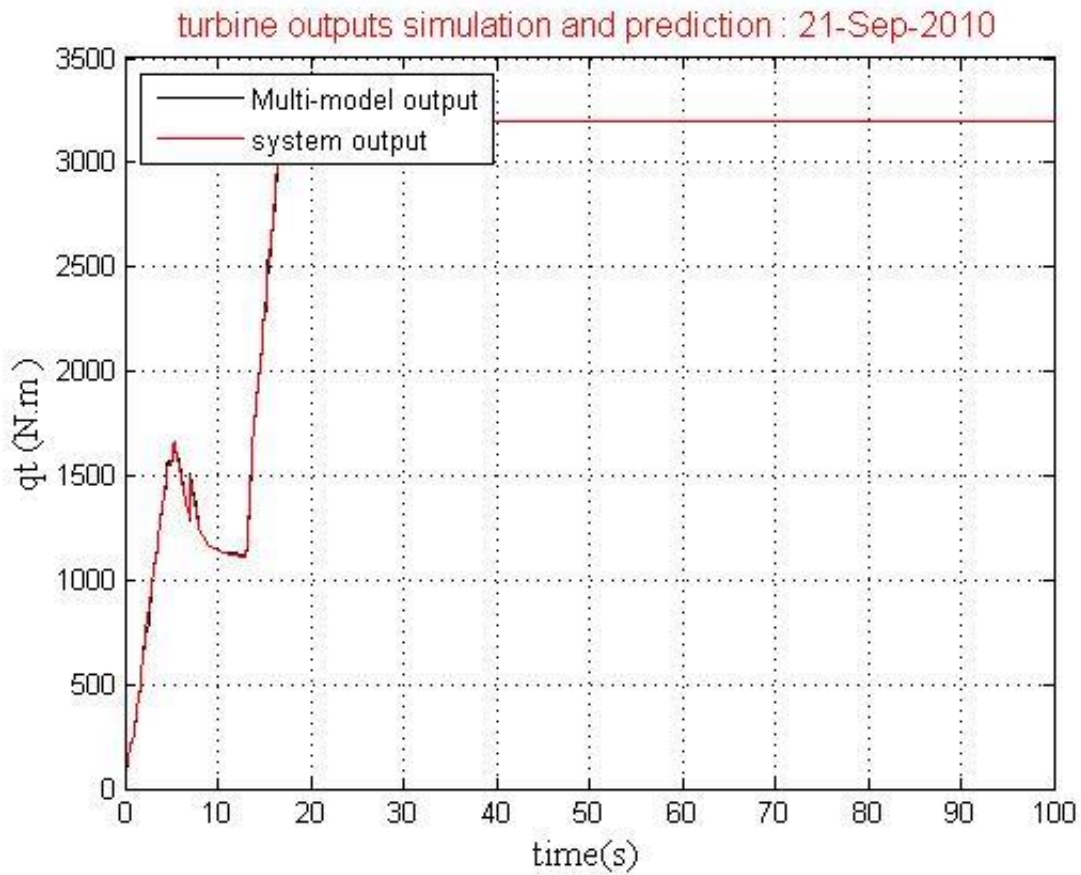
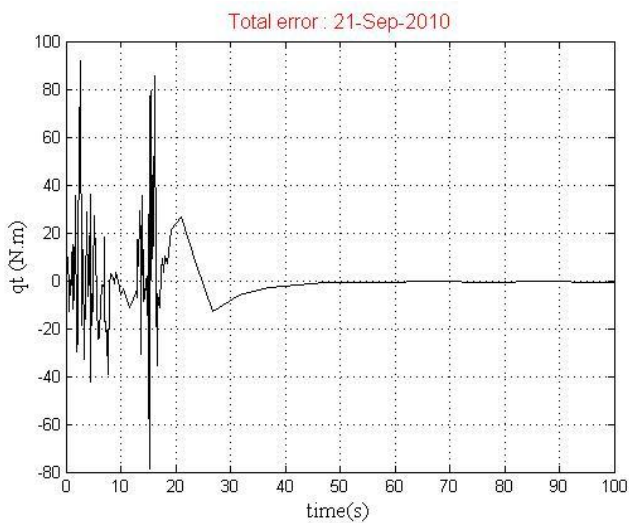
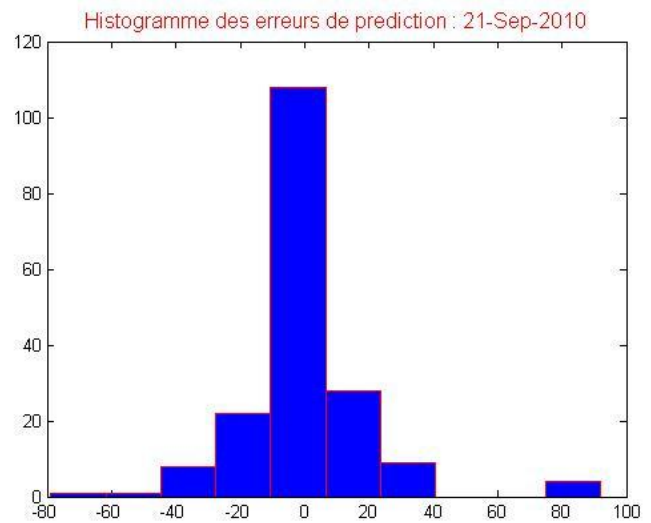


Figure V- 24:La sortie q_t (couple) de la turbine avec la sortie du Multi-modèle



a- Les erreurs



b- Histogramme

Figure V- 25: Erreur et Histogramme de prédiction

❖ Discussion :

Le nombre de modèle locaux qui nous permet d'obtenir cette représentation et après optimisation est : **M=03 modèles locaux**, dont les fonctions de validité sont construites à partir de fonctions gaussienne cela se traduit par 03 zones de fonctionnement du système.

Le vecteur, « **comb** » qui présente le nombre de partition sur chaque support de variable caractéristique est :

$$comb = [3 \ 1 \ 1 \ 1] ;$$

On a alors, trois partition sur le support u_1 , une sur le support u_2 , une sur le support u_3 , et une partition sur u_4 .

Les trois modèles obtenues sont représenté sur la figure V-26 :

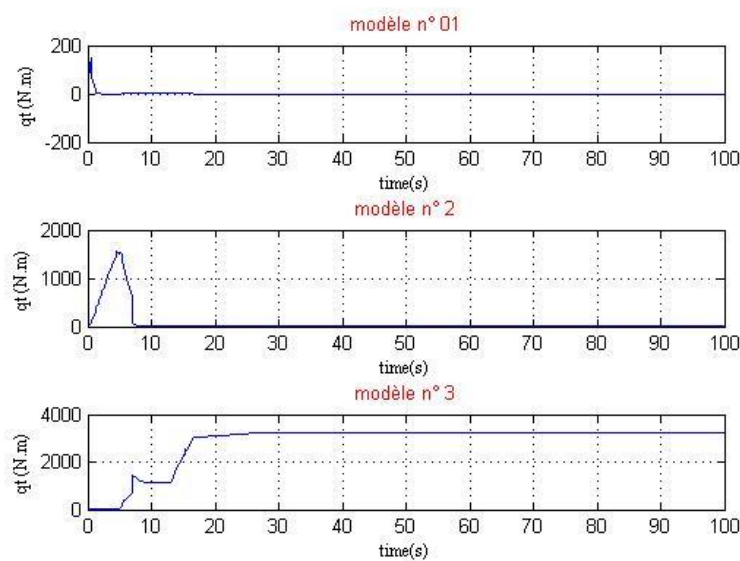


Figure V-26: Les sorties des 3 modèles

Ainsi les paramètres de vecteur θ sont regroupés dans le tableau V-5 :

modèle locale f_i	a_1	a_2	b_{11}	b_{12}	b_{13}	b_{14}	b_0
f_1	-02,4843	01,0521	-01,0716	00,8474	-00,0041	00,0021	-08,2858
f_2	-01,3126	00,5897	00,1228	-00,2114	00,0006	00,0009	31,6436
f_3	-01,6227	00,6950	00,1382	-00,1714	00,0002	00,0006	47,9548

Tableau V- 5 : Paramètres du multi-modèle descriptif de la sortie qt de la turbine

La représentation mathématique de la sortie q_t de la turbine est de la forme :

$$q_t^i(t) = -a_1^i q_t^i(t-1) - a_2^i q_t^i(t-2) + b_{11}^i t_6(t) + b_{12}^i w_t(t) + b_{13}^i p_5(t) + b_{14}^i p_7(t) + b_0^i$$

Où les paramètres a_i^i , b_i^i , b_0^i sont déduits du tableau V-5.

La sortie q_t présente le couple de l'arbre du turbine, cette grandeur est accessible (mesurable), l'unité de mesure est le **(N.m)**.

À partir des figures V-25-a, nous pouvons constater que l'erreur d'estimation est très petite relativement au signal de sortie q_t .

Concernant les figures V-26, on remarque l'existence des zones vides, cela apparaît dans le modèle **(1)** et l'absence des zones redondantes (zones voisines pouvant décrire un même comportement du système).

V-12-2-1-b-La sortie p_t (la puissance) :

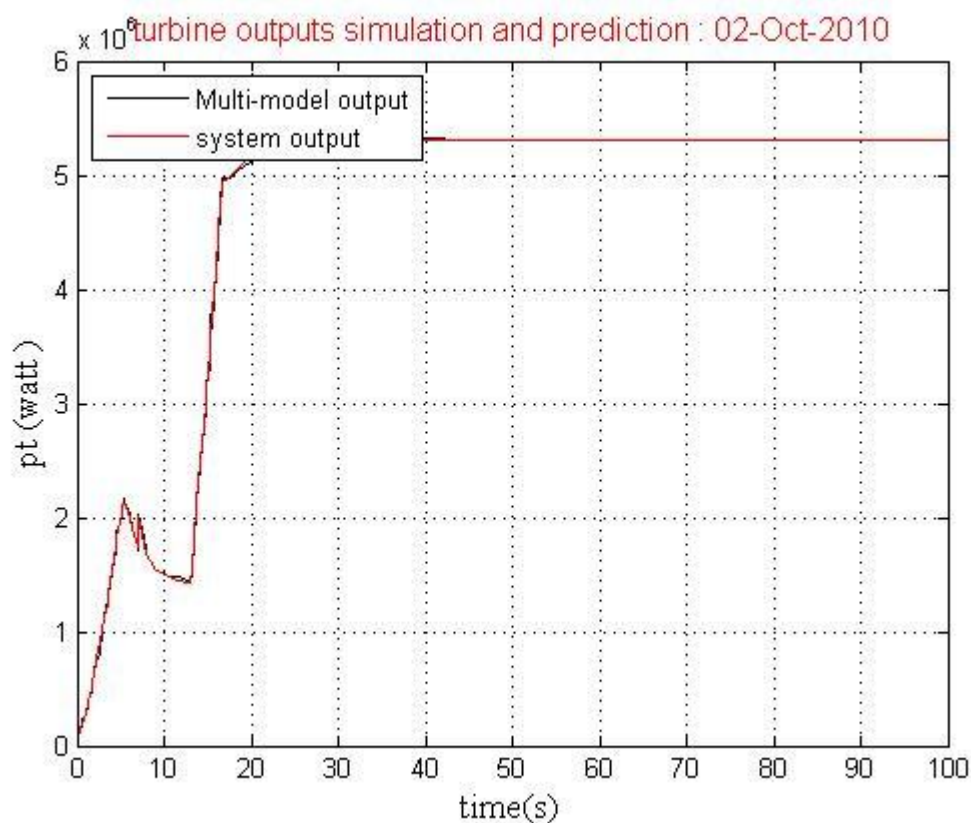


Figure V- 27: La sortie w_t (la puissance) de la turbine avec la sortie du Multi-modèle

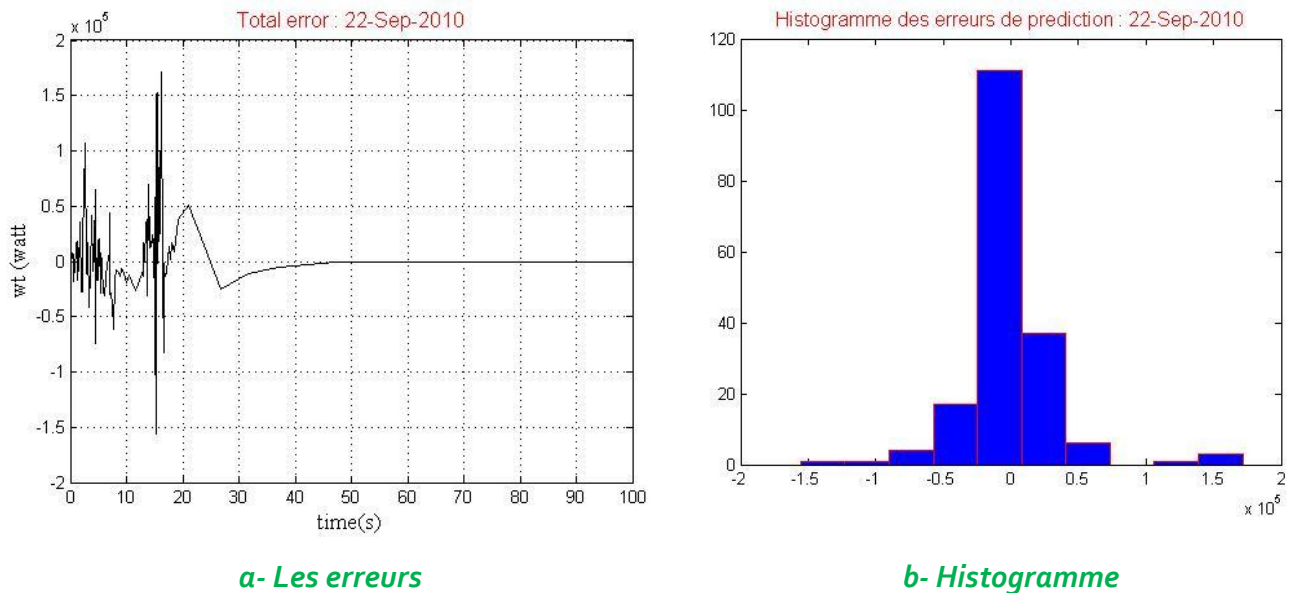


Figure V- 28: Erreur et Histogramme de prédiction

❖ Discussion :

Le nombre de modèle locaux qui nous permet d'obtenir cette représentation et après optimisation est : **M=08 modèles locaux**, dont les fonctions de validité sont construites à partir de fonctions gaussienne cela se traduit par 08 zones de fonctionnement du système.

Le vecteur, « **comb** » qui présente le nombre de partition sur chaque support du variable caractéristique est :

$$comb = [2 \ 2 \ 1 \ 2] ;$$

On a alors deux partitions sur le support u_1 , deux sur le support u_2 , une sur le support u_3 , et deux partition sur u_4 .

Les huit modèles obtenues sont représenté sur la figure V-29 :

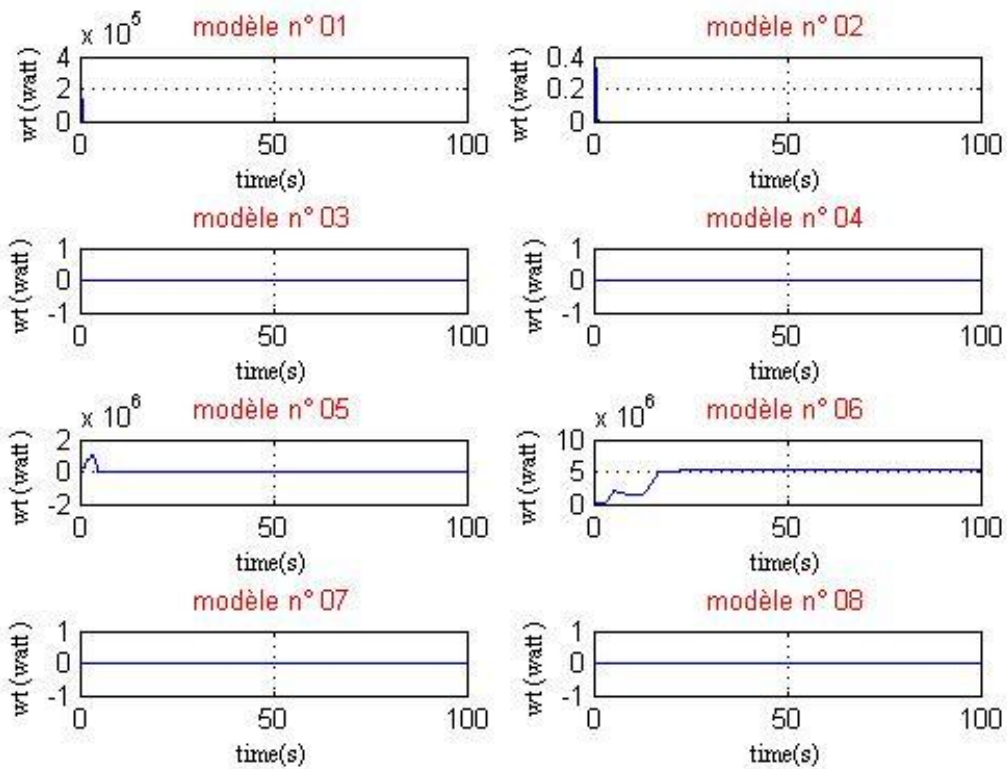


Figure V-29: Les sorties des 8 modèles

Ainsi les paramètres de vecteur θ sont regroupés dans le tableau V-6 :

modèle locale f_i	a_1	a_2	b_{11}	b_{12}	b_{13}	b_{14}	b_0
f_1	-01,3805	00,7367	372,0356	-284,4084	01,9587	-01,3737	-922,3382
f_2	-30,2980	-30,5535	00,0619	00,1527	27,7664	16,4215	00,0001
f_3	00,0000	00,0000	00,0000	00,0000	00,0000	00,0000	00,0000
f_4	00,0000	00,0000	00,0000	00,0000	00,0000	00,0000	00,0000
f_5	-00,4273	00,5245	124,0373	-555,8140	09,3637	-09,2708	01,0952
f_6	-01,8411	00,6892	-19,8933	194,5899	-01,2722	00,4312	1176,9628
f_7	00,0000	00,0000	00,0000	00,0000	00,0000	00,0000	00,0000
f_8	00,0000	00,0000	00,0000	00,0000	00,0000	00,0000	00,0000

Tableau V- 6 : Paramètres du multi-modèle descriptif de la sortie p_t de la turbine

La représentation mathématique de la sortie \mathbf{w}_t de la turbine est de la forme :

$$p_t^i(t) = -a_1^i p_t^i(t-1) - a_2^i p_t^i(t-2) + b_{11}^i t_6(t) + b_{12}^i w_t(t) + b_{13}^i p_5(t) + b_{14}^i p_7(t) + b_0^i$$

Où les paramètres a_i^i , b_i^i , b_0^i sont déduits du tableau V-6.

La sortie p_t présente la puissance du turbine, cette grandeur est accessible (mesurable), l'unité de mesure est le (**watt**).

À partir des figures V-28-a, nous pouvons constater que l'erreur d'estimation est très petite relativement au signal de sortie p_t .

Concernant les figures V-29, on remarque l'existence des zones vides, cela apparaît dans les modèles (1), (2), (3), (4), (7), (8) et l'absence des zones redondantes (zones voisines pouvant décrire un même comportement du système).

La question posée par la teneur de cette sortie est que l'augmentation du nombre de modèle n'améliore pas l'estimation de la sortie p_t .

Les causes possible de ce résultat sont dues à l'absence et l'insuffisance de données d'apprentissage sur cette zones de fonctionnement. Si des zones de fonctionnement sont dépourvues de données, les matrices de pondération \mathbf{W}_i correspondantes sont quasiment nulles. Donc il existe des blocs de Φ_g qui seront presque nuls ce qui induira forcément une perte de rang de la matrice d'information,

V-12-2-1-c-La sortie t_7 (la température à la sortie de la turbine) :

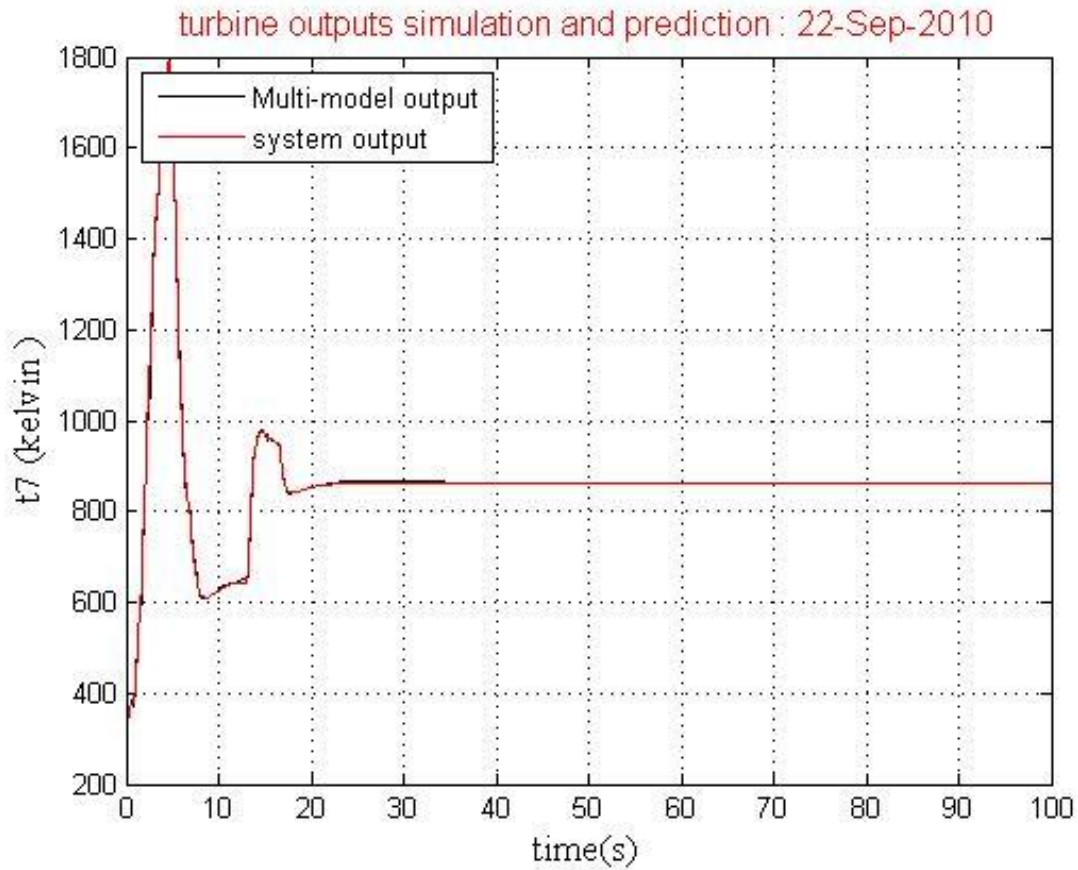


Figure V-30: La sortie t_7 (température) de la turbine avec la sortie du Multi-modèle

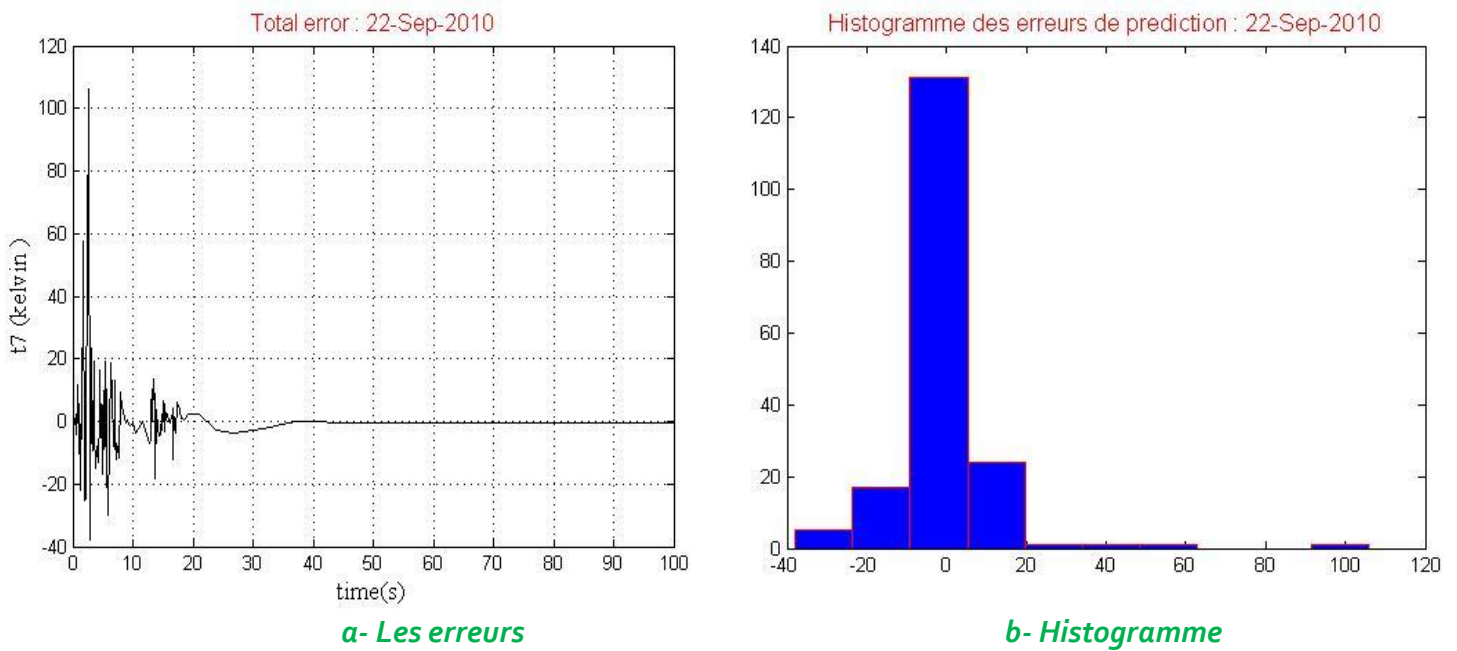


Figure V-31: Erreur et Histogramme de prédiction

❖ Discussion :

Le nombre de modèle locaux qui nous permet d'obtenir cette représentation et après optimisation est : **M=16 modèles locaux**, dont les fonctions de validité sont construites à partir de fonctions gaussienne cela se traduit par 16 zones de fonctionnement du système.

Le vecteur, « **comb** » qui présente le nombre de partition sur chaque support du variable caractéristique est :

$$comb = [4 \ 1 \ 2 \ 2] ;$$

On a alors quatre partitions sur le support u_1 , une sur le support u_2 , deux sur le support u_3 , et deux partition sur u_4 .

Les seize modèles obtenues sont représenté sur la figures V-32 et V-33:

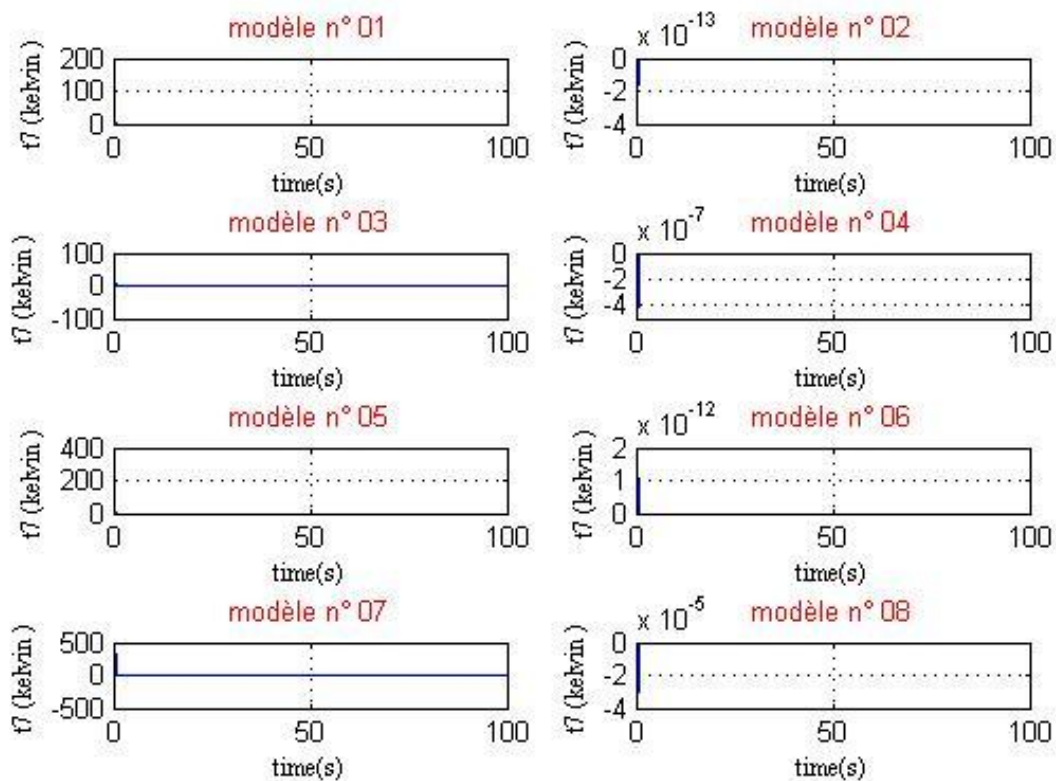


Figure V-32: Les sorties des 8 premiers modèles

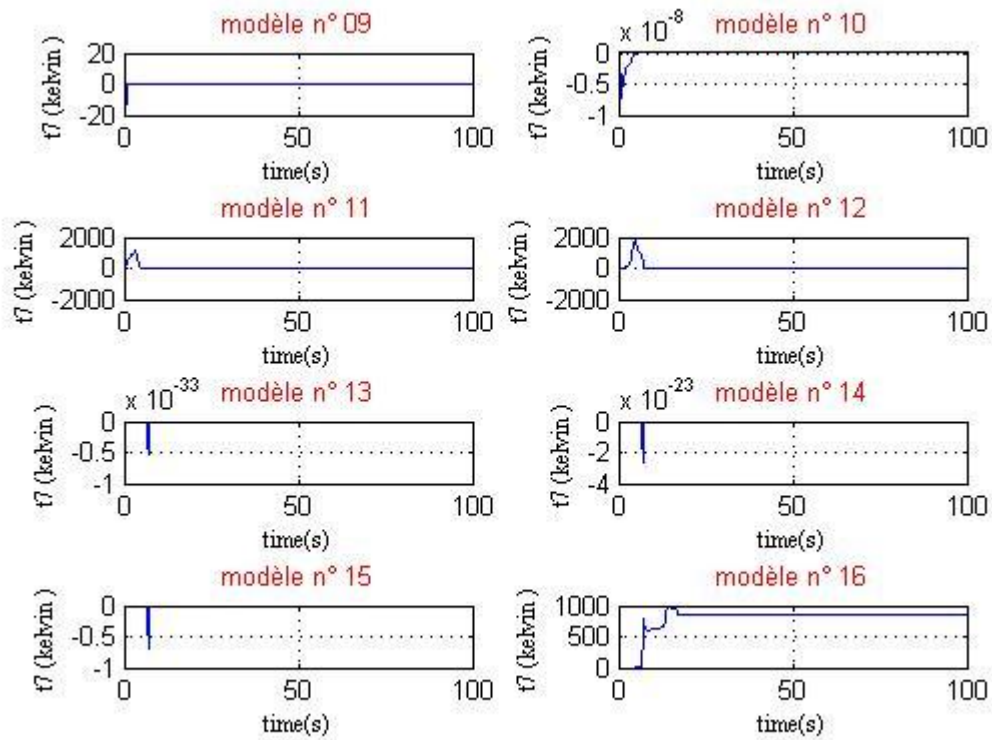


Figure V-33: Les sorties des 8 derniers modèles

Ainsi les paramètres de vecteur θ sont regroupés dans le tableau V-7 :

modèle locale f_i	a_1	a_2	b_{11}	b_{12}	b_{13}	b_{14}	b_0
f_1	-00,1635	00,1063	00,4337	00,0665	-00,0012	00,0022	00,0004
f_2	00,0000	00,0000	00,0000	00,0000	00,0000	00,0000	00,0000
f_3	00,3277	00,1317	-00,2840	-00,6742	00,0284	-00,0383	-00,0001
f_4	00,0000	00,0000	00,0000	00,0000	-00,0005	-00,0003	00,0000
f_5	-00,2890	-00,1168	01,0310	-00,1960	00,0012	-00,0014	00,0019
f_6	00,0000	00,0000	00,0000	00,0000	00,0000	00,0000	00,0000
f_7	00,8488	00,3306	-00,4794	00,6902	-00,0077	00,0162	-00,0001
f_8	00,0000	00,0000	00,0000	00,0000	-00,0044	-00,0025	00,0000
f_9	00,1228	00,0764	-00,1445	00,0208	00,8708	-01,5826	00,0000
f_{10}	00,0000	00,0000	00,0000	00,0000	-00,0002	-00,0001	00,0000
f_{11}	03,7533	00,5207	05,7571	-00,4140	-00,0195	00,0340	00,2772
f_{12}	02,5706	00,7308	03,2665	-00,6634	-00,0208	00,0676	-00,0576

f_{13}	00,0000	00,0000	00,0000	00,0000	00,0000	00,0000	00,0000
f_{14}	00,0000	00,0000	00,0000	00,0000	00,0000	00,0000	00,0000
f_{15}	00,0016	00,0017	-00,0019	-00,0021	-00,8258	-00,3171	00,0000
f_{16}	-01,4686	00,9319	00,3694	-00,0119	-00,0001	00,0005	-00,1009

Tableau V- 7 : Paramètres du multi-modèle descriptif de la sortie t_7 de la turbine

La représentation mathématique de la sortie t_7 de la turbine est de la forme :

$$t_7^i(t) = -a_1^i t_7^i(t-1) - a_2^i t_7^i(t-2) + b_{11}^i t_6(t) + b_{12}^i w_t(t) + b_{13}^i p_5(t) + b_{14}^i p_7(t) + b_0^i$$

Où les paramètres a_i^i , b_i^i , b_0^i sont déduits du tableau V-7.

La sortie t_7 présente la température du gaz à la sortie du turbine, cette grandeur est accessible (mesurable), l'unité de mesure est le (**kelvin**).

À partir des figures V-31-a, nous pouvons constater que l'erreur d'estimation est très petite relativement au signal de sortie t_7 .

Concernant les figures V-32 et V-33, on remarque l'existence des zones vides, cela apparaît dans les modèles (2), (4), (6), (8), (10), (13) et l'absence des zones redondantes (zones voisines pouvant décrire un même comportement du système).

Les causes possible de ce résultat est dû à l'absence et l'insuffisance de données d'apprentissage sur cette zones de fonctionnement. Si des zones de fonctionnement sont dépourvues de données, les matrices de pondération W_i correspondantes sont quasiment nulles. Donc il existe des blocs de Φ_g qui seront presque nuls ce qui induira forcément une perte de rang de la matrice d'information,

V-12-2-1-d-La sortie m_8 (écoulement de masse à la sortie de la turbine) :

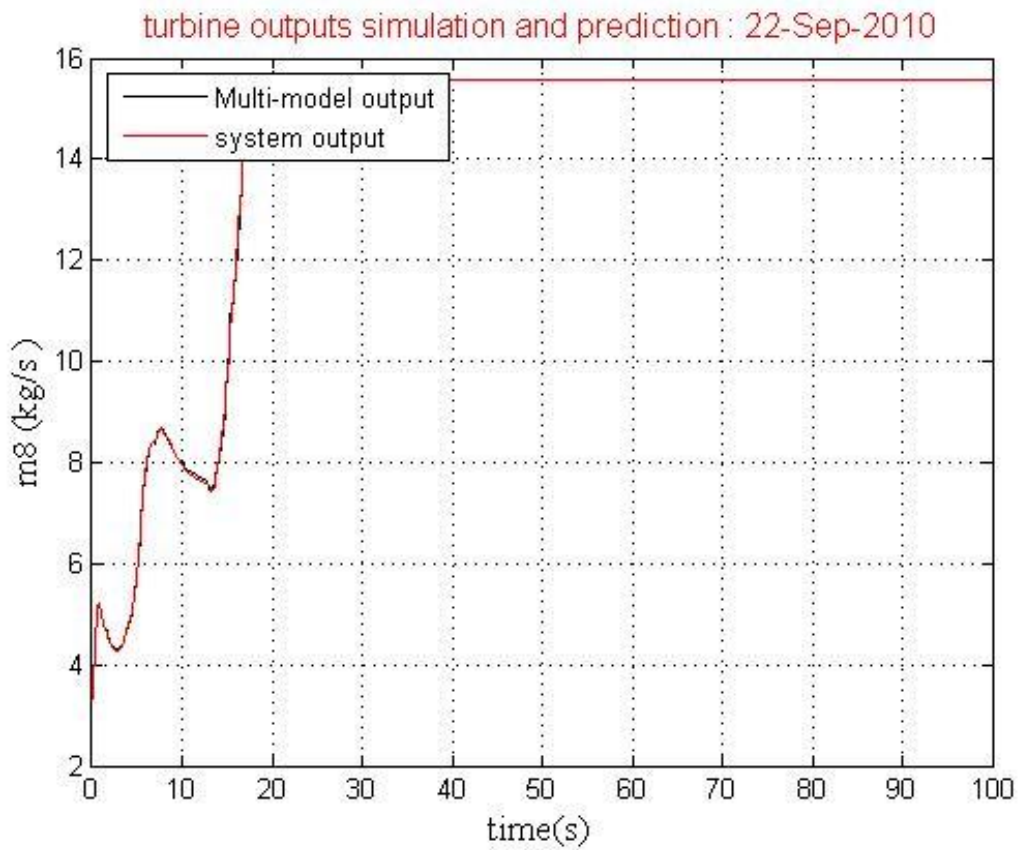
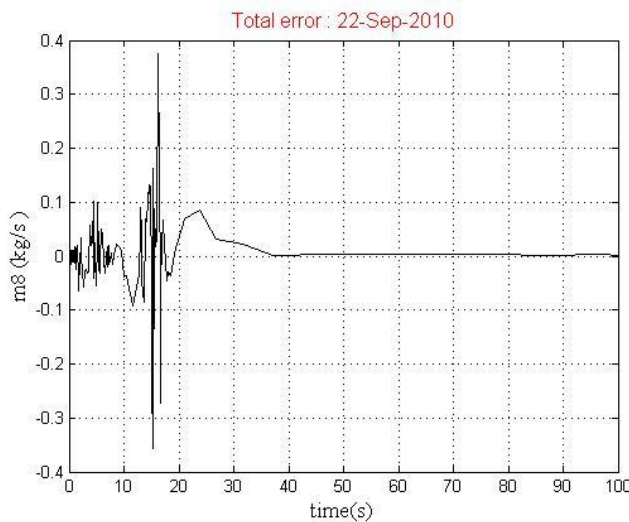
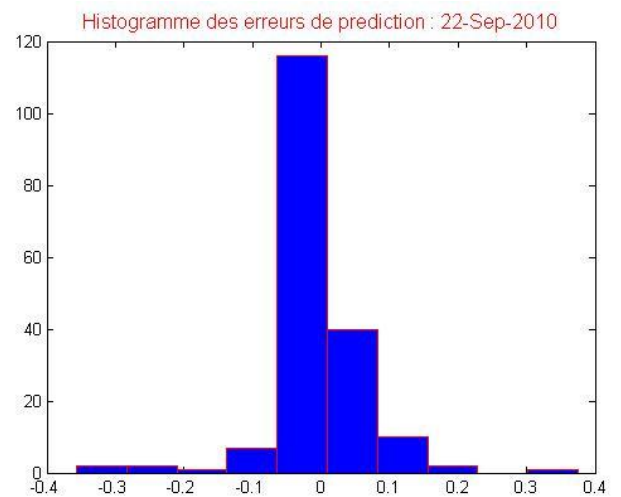


Figure V- 34: La sortie m_8 (mass flow) de la turbine avec la sortie du Multi-modèle



a- Les erreurs



b- Histogramme

Figure V- 35: Erreur et Histogramme de prédiction

❖ Discussion :

Le nombre de modèle locaux qui nous permet d'obtenir cette représentation et après optimisation est : **M=72 modèles locaux**, dont les fonctions de validité sont construites à partir de fonctions gaussienne cela se traduit par 72 zones de fonctionnement du système.

Le vecteur, « **comb** » qui présente le nombre de partition sur chaque support du variable caractéristique est :

$$comb = [4 \ 3 \ 2 \ 3] ;$$

On a alors, quatre partitions sur le support u_1 , trois sur le support u_2 , deux sur le support u_3 , et trois partition sur u_4 .

On présente quelques modèles dans les figures V-36 et V-37.

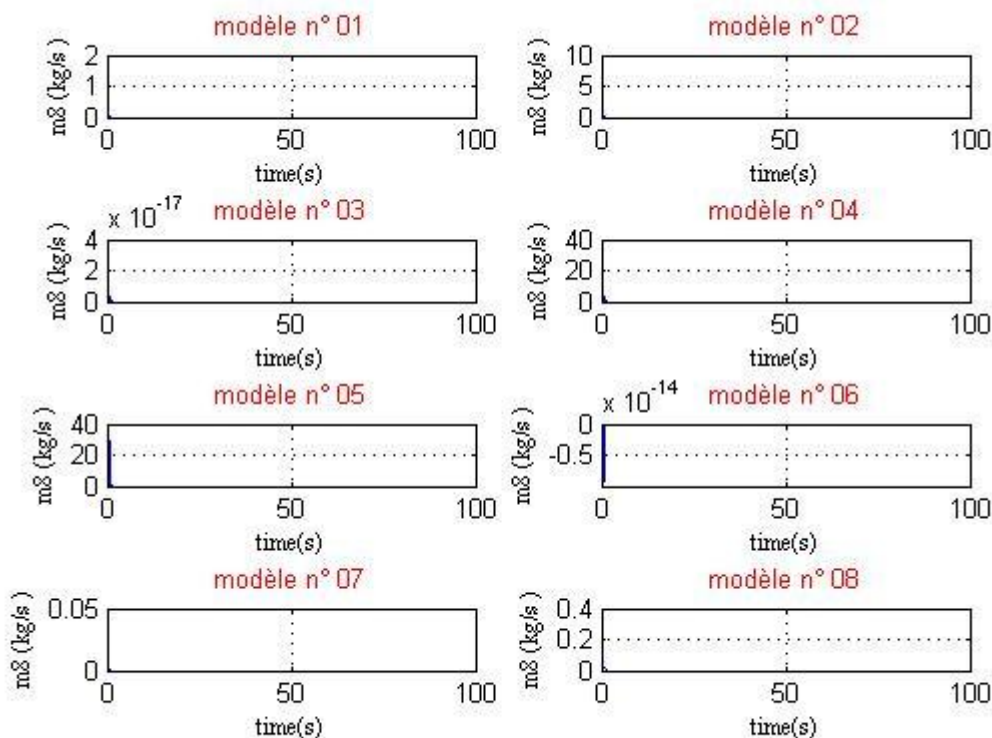


Figure V-36: Les sorties des 8 premiers modèles

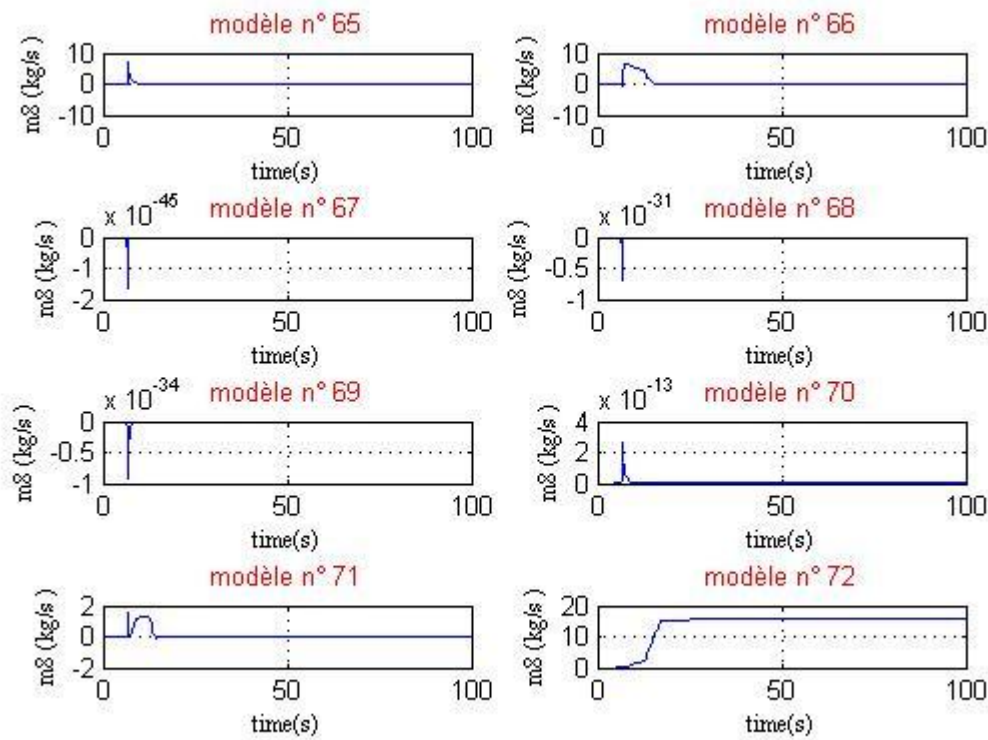


Figure V-37: Les sorties des 8 derniers modèles

Le vecteur θ est n'es pas présenté par ce que son dimension est très grande (dim $\theta = 504$)

La représentation mathématique de la sortie m_8 de la turbine est de la forme :

$$m_8^i(t) = -a_1^i m_8^i(t - 1) - a_2^i m_8^i(t - 2) + b_{11}^i t_6(t) + b_{12}^i w_t(t) + b_{13}^i p_5(t) + b_{14}^i p_7(t) + b_0^i$$

La sortie m_8 présente l'écoulement du gaz à la sortie du turbine, cette grandeur est accessible (mesurable), l'unité de mesure est le (kg/s).

À partir des figures V-35-a, nous pouvons constater que l'erreur d'estimation est très petite relativement au signal de sortie m_8 .

Concernant les figures V-36 et V-37, on remarque l'existence des zones vides, et des zones redondantes (zones voisines pouvant décrire un même comportement du système).

Les causes possible de ce résultat est dû à l'absence et l'insuffisance de données d'apprentissage sur cette zones de fonctionnement. Si des zones de fonctionnement sont dépourvues de données, les matrices de pondération \mathbf{W}_i correspondantes sont quasiment nulles. Donc il existe des blocs de Φ_g qui seront presque nuls ce qui induira forcément une perte de rang de la matrice d'information,

Le multi modèle que nous avons obtenue est malgré tous représente bien la sortie m_8 du turbine.

V-12-3-FUEL SYSTEM:

V-12-3-1-a-La sortie t_6 (température à la sortie de l'installation carburant) :

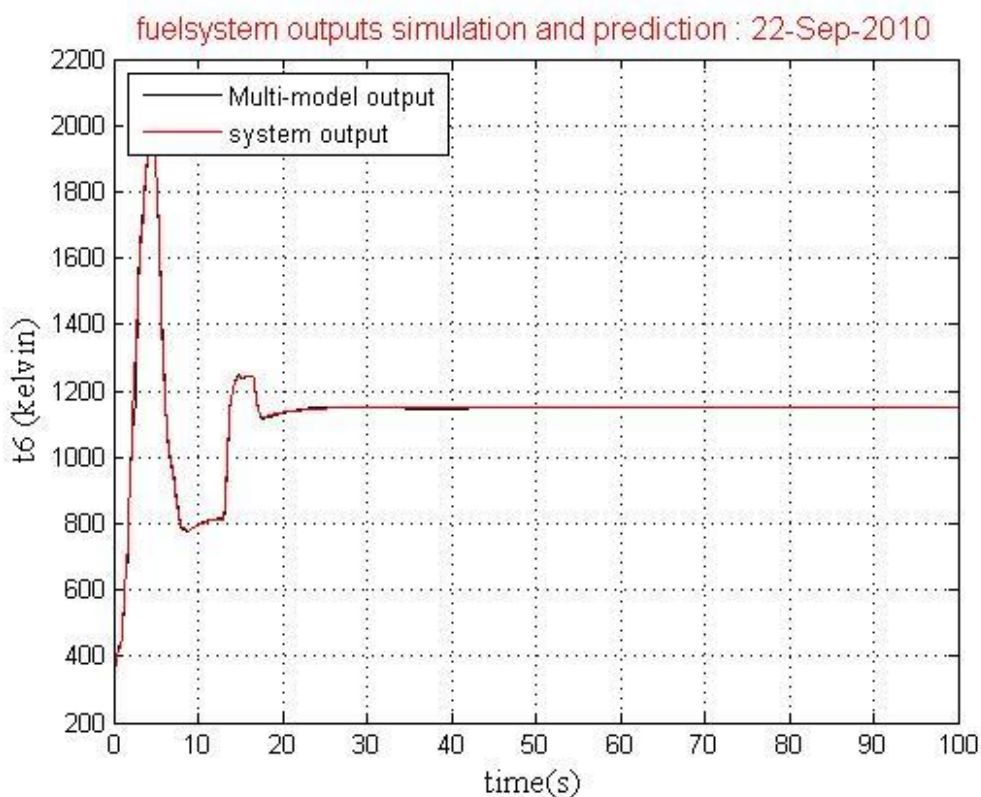
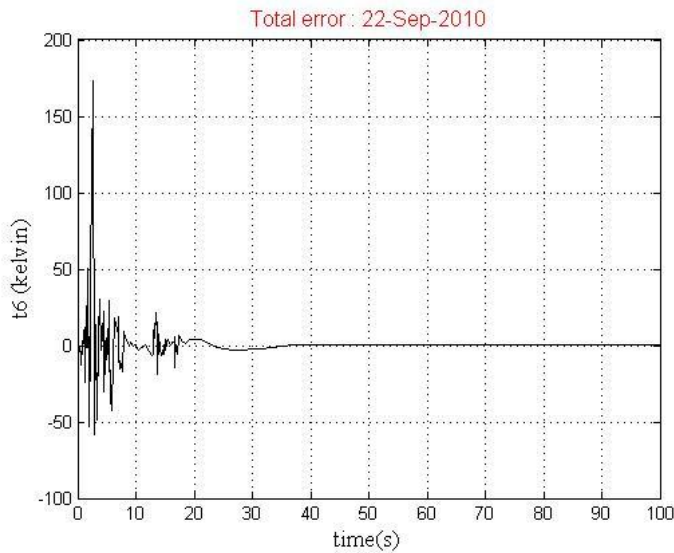
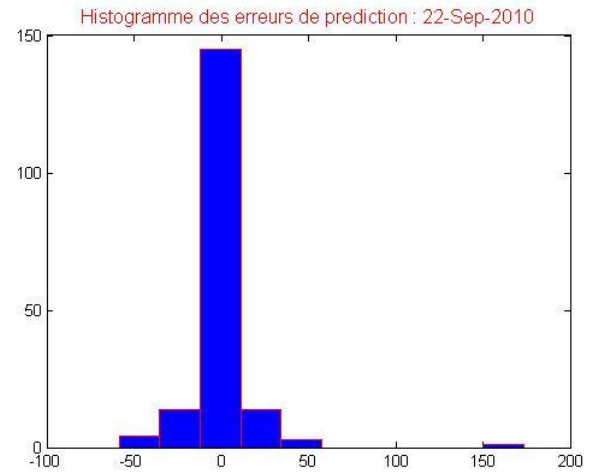


Figure V- 38: La sortie t_6 (température) de l'installation carburant avec la sortie du Multi-modèle



a- Les erreurs



b- Histogramme

Figure V- 39: Erreur et Histogramme de prédiction

❖ Discussion :

Le nombre de modèle locaux qui nous permet d'obtenir cette représentation et après optimisation est : **M=01 modèle local**, dont les fonctions de validité sont construites à partir de fonctions gaussienne cela se traduit par une zones de fonctionnement du système.

Le vecteur, « **comb** » qui présente le nombre de partition sur chaque support du variable caractéristique est :

$$comb = [1 \ 1 \ 1 \ 1] ;$$

On a alors, une partition sur le support u_1 , une sur le support u_2 , une sur le support u_3 , et une partition sur u_4 .

Le modèle obtenue est représenté sur la figure V-40 :

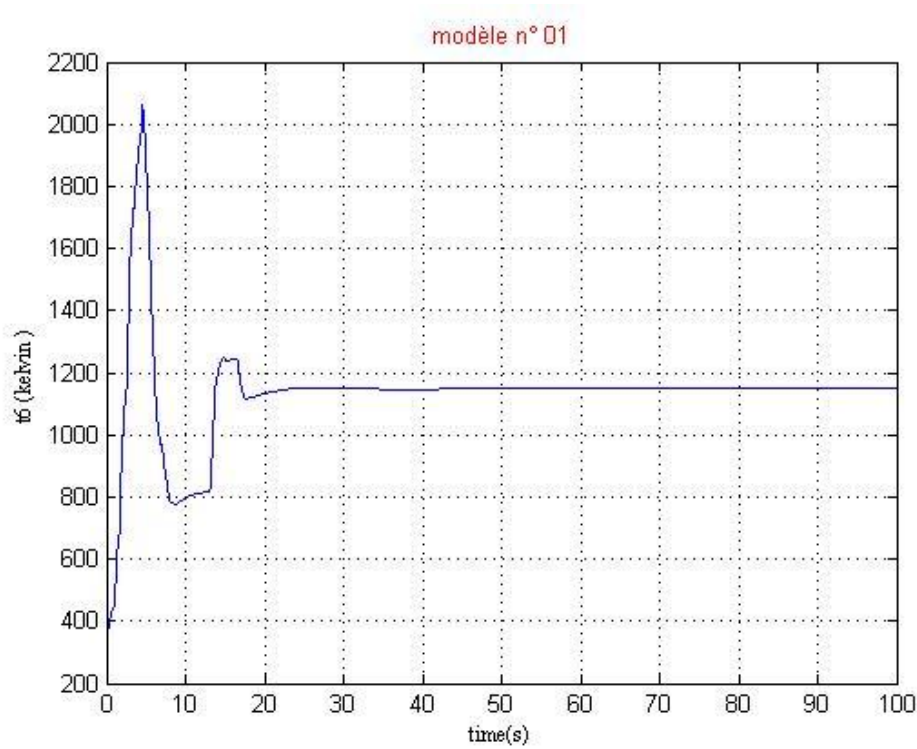


Figure V-40: La sortie du multi modèle

Ainsi les paramètres de vecteur θ sont regroupés dans le tableau V-8 :

modèle locale f_i	a_1	a_2	b_{11}	b_{12}	b_{13}	b_{14}	b_0
f_1	-1.8542	0.8616	4.5912	-4.1365	-4.5202	0.0186	3.0344

Tableau V- 8 : Paramètres du multi-modèle descriptif de la sortie t_6 du fuel system

La représentation mathématique de la sortie t_6 de la turbine est de la forme :

$$t_6^i(t) = -a_1^i t_6^i(t - 1) - a_2^i t_6^i(t - 2) + b_{11}^i m_8(t) + b_{12}^i m_{ff}(t) + b_{13}^i m_3(t) + b_{14}^i t_3(t) + b_0^i$$

Où les paramètres a_i^i, b_i^i, b_0^i sont déduits du tableau V-8.

La sortie t_6 présente la température du gaz à la sortie de l'installation carburant, cette grandeur est accessible (mesurable), l'unité de mesure est le (kelvin).

À partir des figures V-39-a, nous pouvons constater que l'erreur d'estimation est très petite relativement au signal de sortie t_6 .

Concernant les figures V-40, on remarque l'absence des zones vides, et des zones redondantes (zones voisines pouvant décrire un même comportement du système).

V-12-3-1-b-La sortie p_5 (température à la sortie de l'installation carburant) :

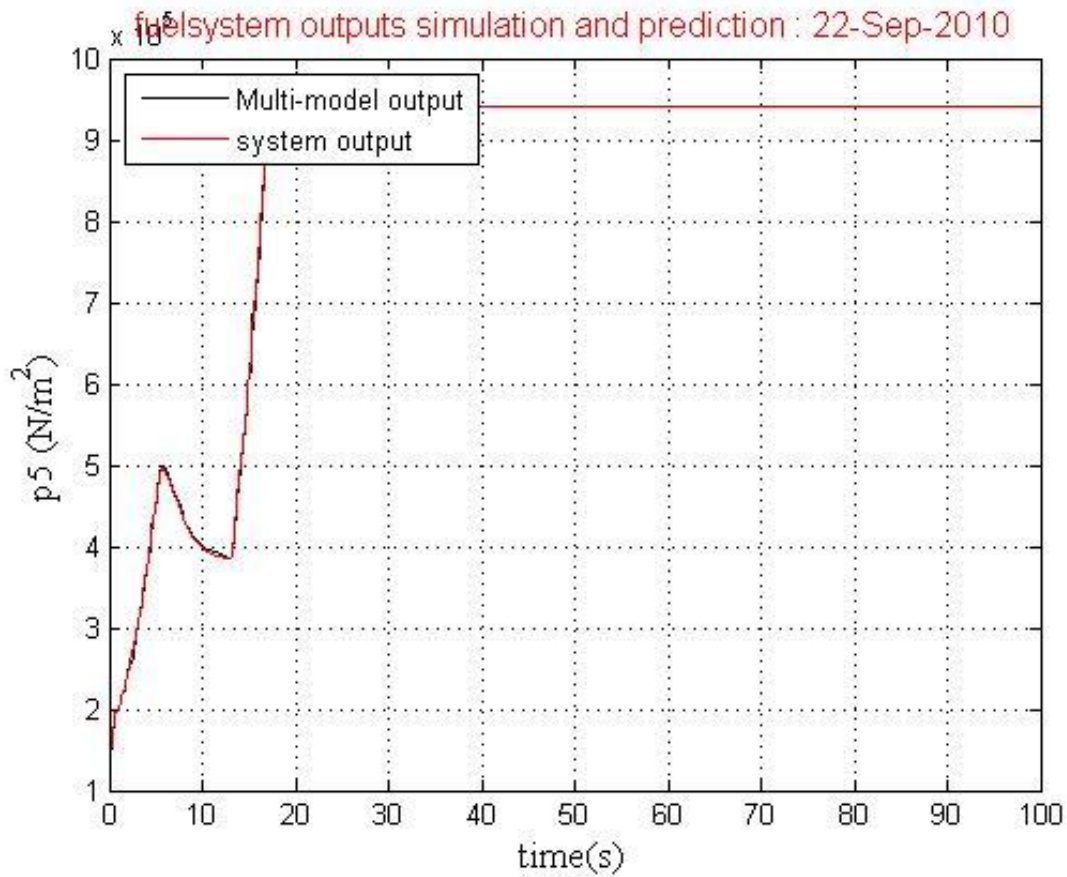


Figure V- 41: La sortie p_5 (pression) de l'installation carburant avec la sortie du Multi-modèle

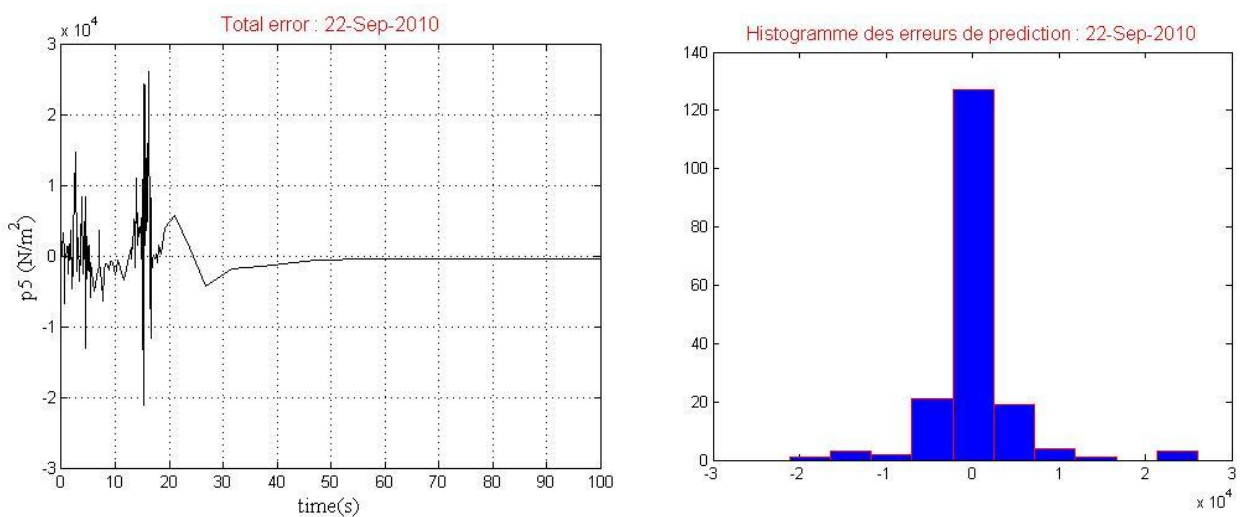


Figure V- 42: Erreur et Histogramme de prédiction

❖ Discussion :

Le nombre de modèle locaux qui nous permet d'obtenir cette représentation et après optimisation est : **M=02 modèles locaux**, dont les fonctions de validité sont construites à partir de fonctions gaussienne cela se traduit par 02 zones de fonctionnement du système.

Le vecteur, « **comb** » qui présente le nombre de partition sur chaque support du variable caractéristique est :

$$comb = [2 \ 1 \ 1 \ 1] ;$$

On a alors, deux partitions sur le support u_1 , une sur le support u_2 , une sur le support u_3 , et une partition sur u_4 .

Le modèle obtenue est représenté sur la figure V-43 :

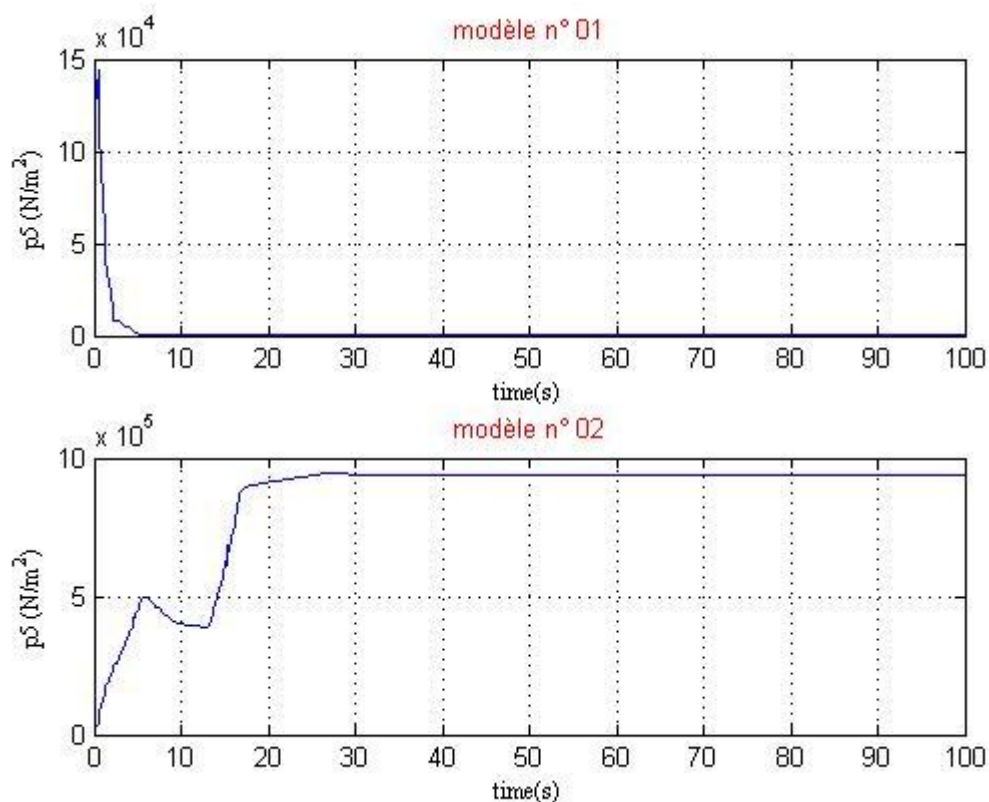


Figure V-43: Les sorties des modèles

Ainsi les paramètres de vecteur θ sont regroupés dans le tableau V-9 :

modèle locale f_i	a_1	a_2	b_{11}	b_{12}	b_{13}	b_{14}	b_0
f_1	-01,8389	00,8365	-149,8756	-18,7444	-506,1859	04,5339	1075,6464
f_2	-01,7554	00,7557	-5941,4870	78,8403	5479,0002	00,4067	763,6153

Tableau V- 9 : Paramètres du multi-modèle descriptif de la sortie p_5 du fuel system

La représentation mathématique de la sortie p_5 de la turbine est de la forme :

$$p_5^i(t) = -a_1^i p_5^i(t - 1) - a_2^i p_5^i(t - 2) + b_{11}^i m_8(t) + b_{12}^i m_{ff}(t) + b_{13}^i m_3(t) + b_{14}^i t_3(t) + b_0^i$$

La sortie p_5 présente la pression du gaz à la sortie de l'installation carburant, cette grandeur est accessible (mesurable), l'unité de mesure est le (N/m^2) .

À partir des figures V-42, nous pouvons constater que l'erreur d'estimation est très petite relativement au signal de sortie p_5 .

Concernant les figures V-43, on remarque l'existence de zones vides (modèle n°01), et l'absence des zones redondantes (zones voisines pouvant décrire un même comportement du système).

V-12-3-1-c-La sortie p_3 (la pression à la sortie de l'installation carburant) :

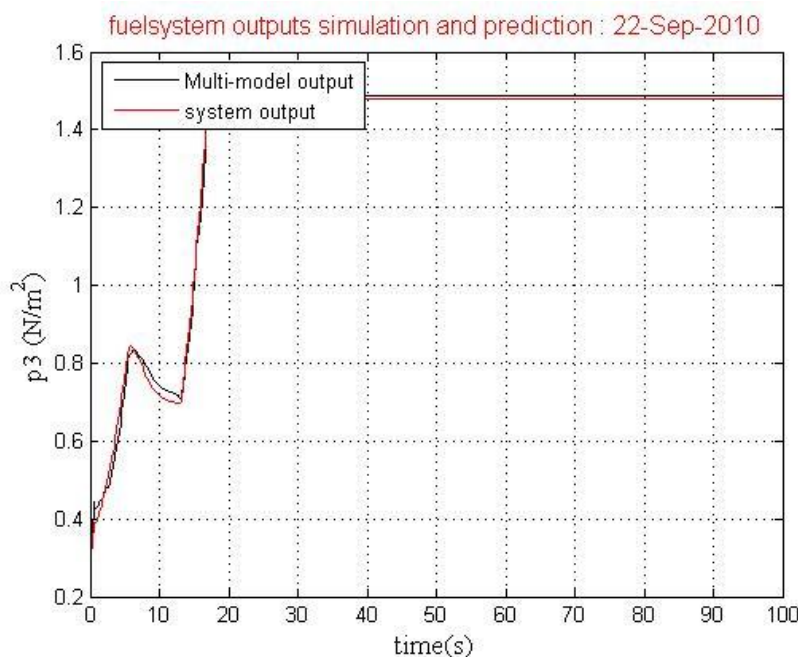
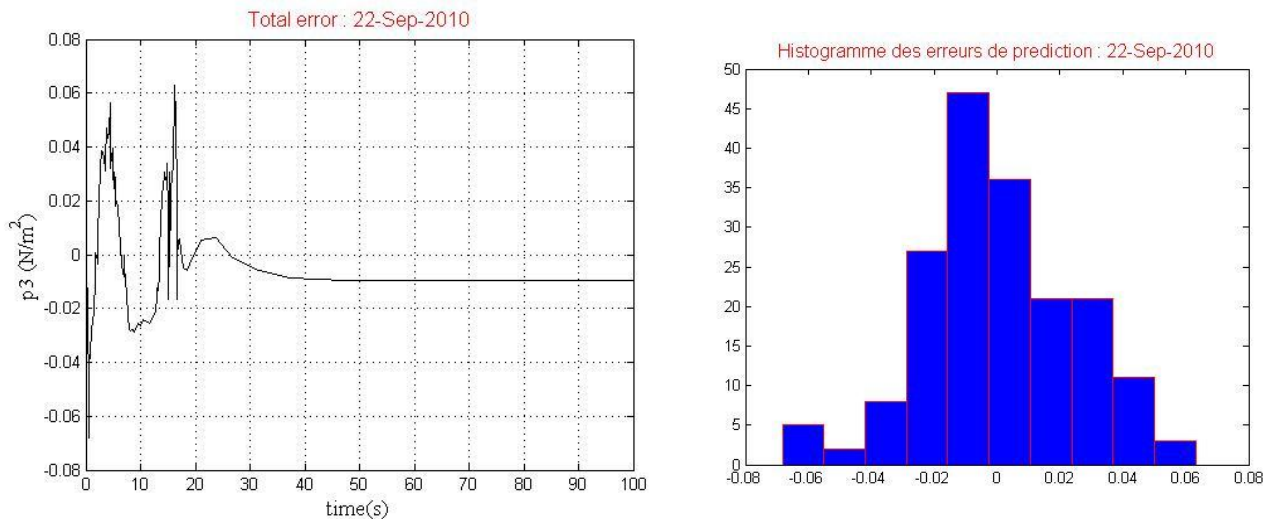


Figure V- 44: La sortie p_3 (pression) de l'installation carburant avec la sortie du Multi-modèle



a- Les erreurs

b- Histogramme

Figure V- 45: Erreur et Histogramme de prédiction

❖ Discussion :

Le nombre de modèle locaux qui nous permet d'obtenir cette représentation et après optimisation est : **M=01 modèle local**, dont les fonctions de validité sont construites à partir de fonctions gaussienne cela se traduit par une zones de fonctionnement du système.

Le vecteur, « **comb** » qui présente le nombre de partition sur chaque support du variable caractéristique est :

$$comb = [1 \ 1 \ 1 \ 1] ;$$

On a alors, une partition sur le support u_1 , une sur le support u_2 , une sur le support u_3 , et une partition sur u_4 .

Le modèle obtenue est représenté sur la figure V-46 :

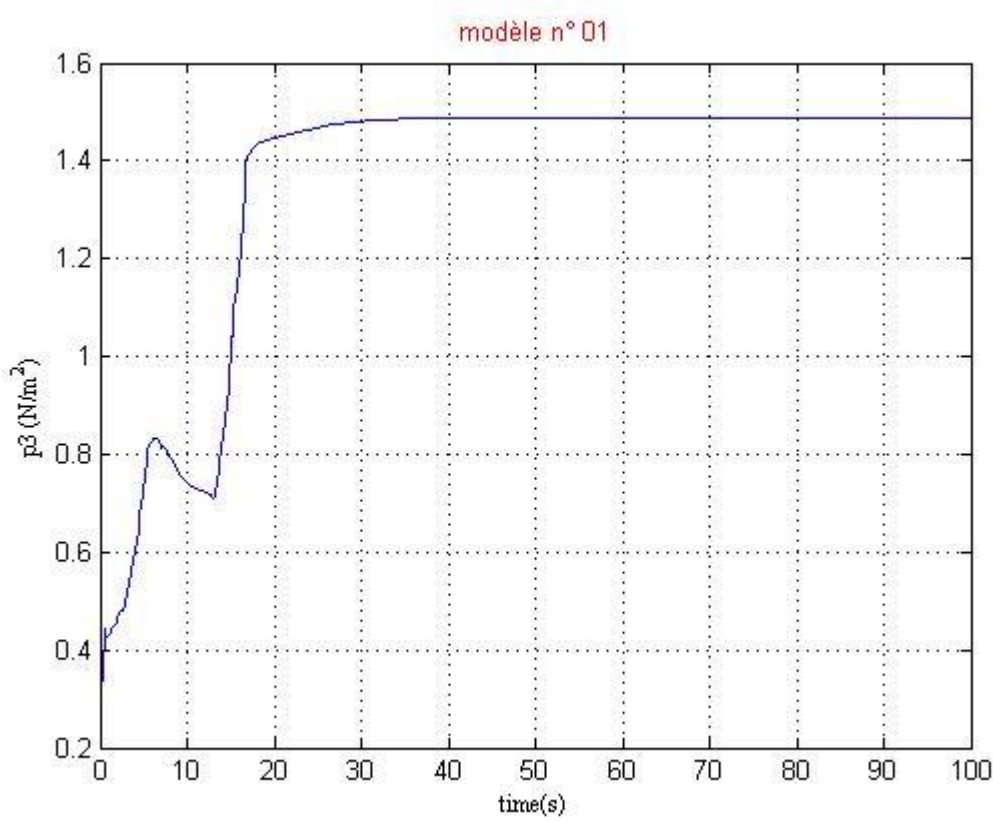


Figure V-46: La sortie du multi modèle

Ainsi les paramètres de vecteur θ sont regroupés dans le tableau V-10 :

modèle locale f_i	a_1	a_2	b_{11}	b_{12}	b_{13}	b_{14}	b_0
f_1	-00,302	-00,274	-00,103	00,172	00,127	00,000	-00,048

Tableau V- 10 : Paramètres du multi-modèle descriptif de la sortie p_3 du fuel system

La représentation mathématique de la sortie q_t de la turbine est de la forme :

$$p_3^i(t) = -a_1^i p_3^i(t - 1) - a_2^i p_3^i(t - 2) + b_{11}^i m_8(t) + b_{12}^i m_{ff}(t) + b_{13}^i m_3(t) + b_{14}^i t_3(t) + b_0^i$$

La sortie p_3 présente la pression du gaz à la sortie de l'installation carburant, cette grandeur est accessible (mesurable), l'unité de mesure est le (N/m^2) .

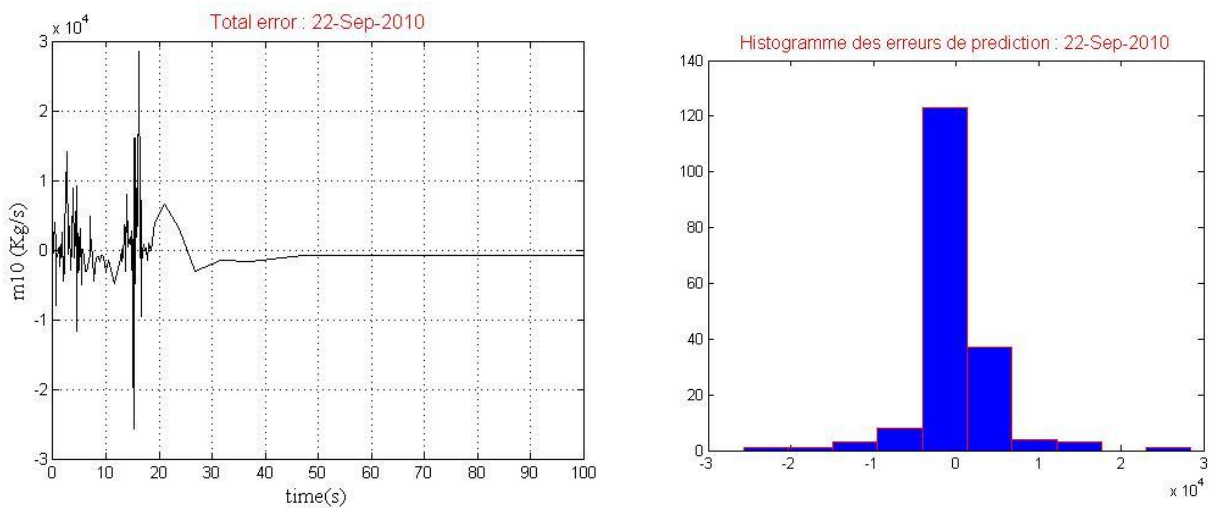
À partir des figures V-45-a, nous pouvons constater que l'erreur d'estimation est importante relativement au signal de sortie p_3 .

Le modèle obtenue ne représente pas bien le système, l'augmentation de nombre de partition sur le support des variables caractéristique ne résoudre pas le problème.

V-12-3-1-d-La sortie m_{10} (l'écoulement de masse à la sortie de l'installation carburant) :



Figure V- 47: La sortie m_{10} (mass flow) de l'installation carburant avec la sortie du Multi-modèle



a- Les erreurs

b- Histogramme

Figure V- 48: Erreur et Histogramme de prédiction

❖ Discussion :

Le nombre de modèle locaux qui nous permet d'obtenir cette représentation et après optimisation est : **M=04 modèle locaux**, dont les fonctions de validité sont construites à partir de fonctions gaussienne cela se traduit par 04 zones de fonctionnement du système.

Le vecteur, « **comb** » qui présente le nombre de partition sur chaque support du variable caractéristique est :

$$comb = [2 \ 1 \ 2 \ 1] ;$$

On a alors, deux partitions sur le support u_1 , une sur le support u_2 , 2 sur le support u_3 , et une partition sur u_4 , les quatre modèles obtenues sont représenté sur les figures suivantes :

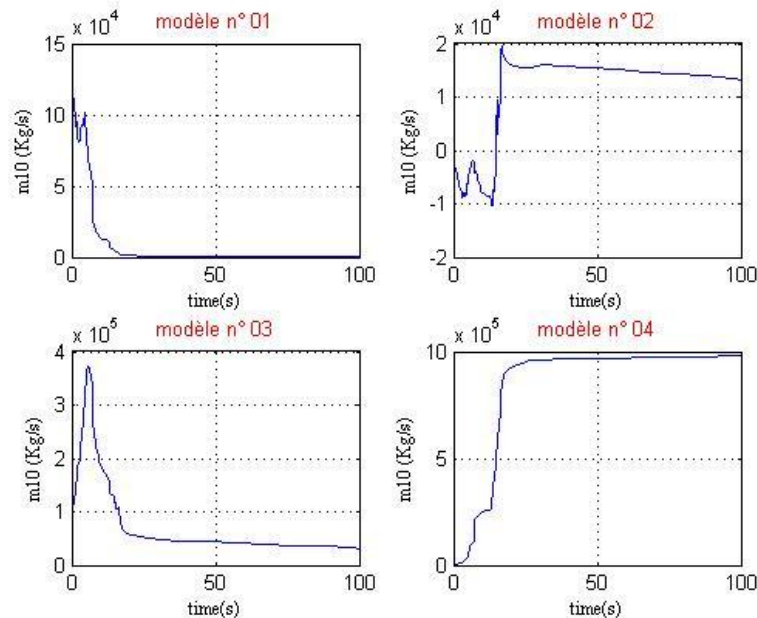


Figure V-49: Les sorties des 4 modèles

La représentation mathématique de la sortie m_{10} de l'installation carburant est de la forme :

$$m_{10}^i(t) = -a_1^i m_{10}^i(t-1) - a_2^i m_{10}^i(t-2) + b_{11}^i p_2(t) + b_{12}^i t_2(t) + b_{13}^i w_t(t) + b_{14}^i p_3(t) + b_0^i$$

La sortie m_{10} présente l'écoulement de masse du carburant, cette grandeur est accessible (mesurable), l'unité de mesure est le (**kg/s**).

À partir des figures V-48, nous pouvons constater que l'erreur d'estimation est très petite relativement au signal de sortie m_{10} .

Concernant les figures V-49, on remarque l'absence des zones vides, et l'absence des zones redondantes (zones voisines pouvant décrire un même comportement du système).

V-12-3-1-e-La sortie t4 (température du carburant à la sortie de l'installation carburant) :

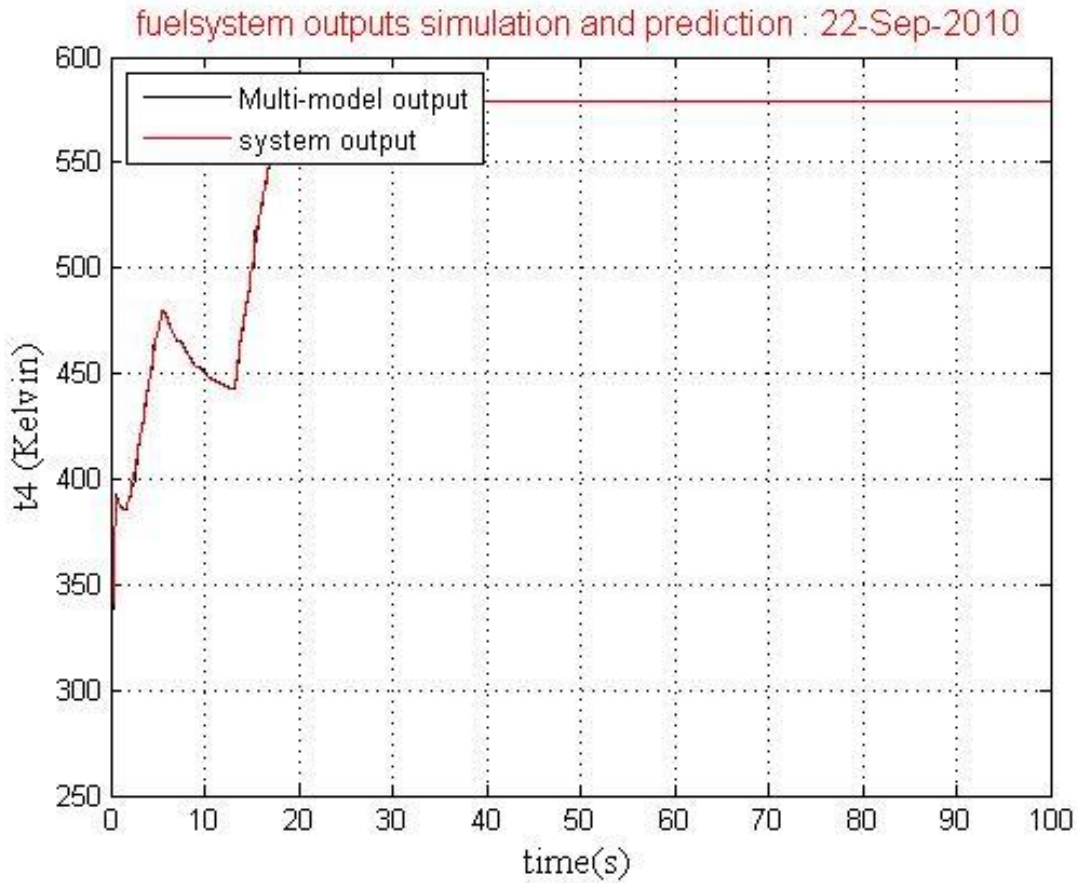
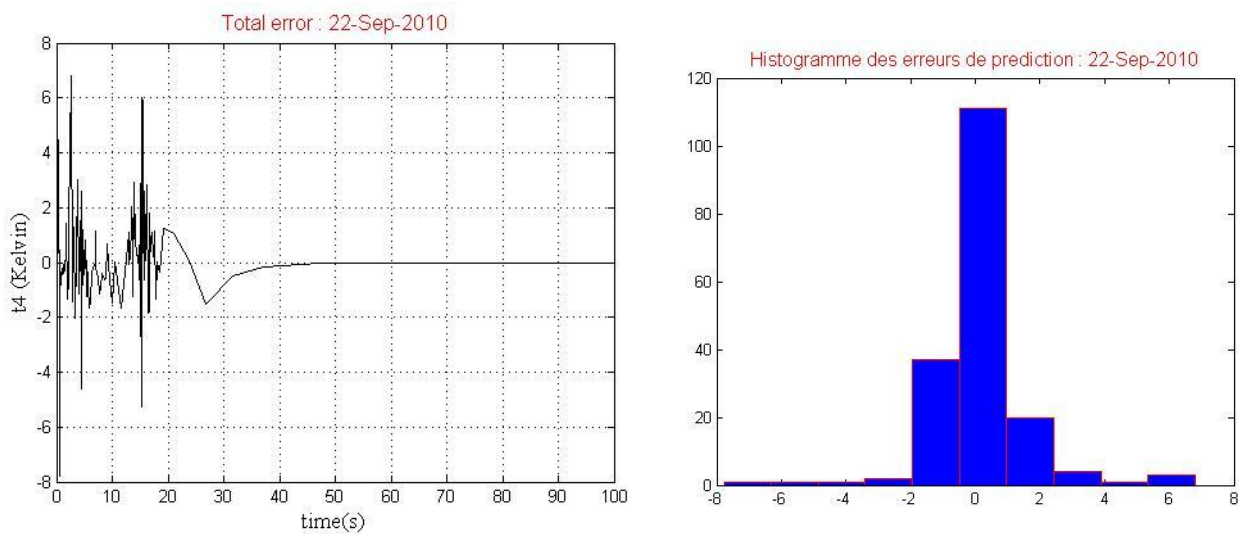


Figure V- 50: La sortie t_4 (température) de l'installation carburant avec la sortie du Multi-modèle



a- Les erreurs

b- Histogramme

Figure V- 51: Erreur et Histogramme de prédiction

❖ Discussion :

Le nombre de modèle locaux qui nous permet d'obtenir cette représentation et après optimisation est : **M=01 modèle local**, dont les fonctions de validité sont construites à partir de fonctions gaussienne cela se traduit par une zones de fonctionnement du système.

Le vecteur, « **comb** » qui présente le nombre de partition sur chaque support du variable caractéristique est :

$$comb = [1 1 1 1] ;$$

On a alors, une partition sur le support u_1 , une sur le support u_2 , une sur le support u_3 , et une partition sur u_4 .

Le modèle obtenue est représenté sur la figure V-52 :

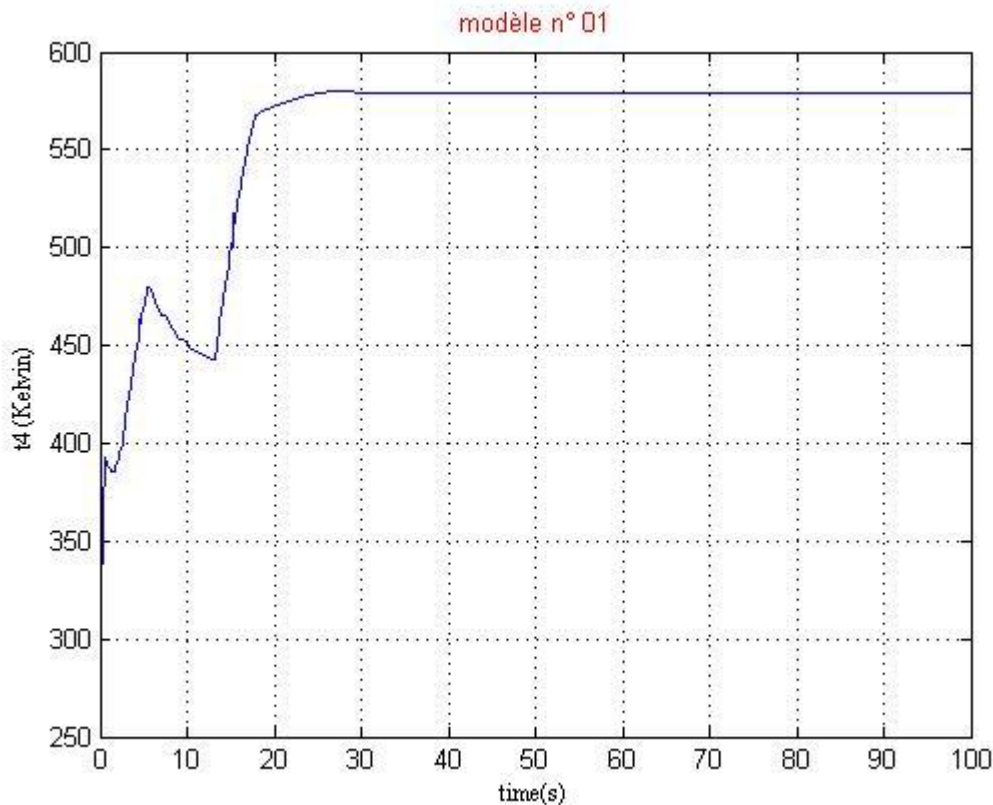


Figure V-52: La sortie du multi modèle

Ainsi les paramètres de vecteur θ sont regroupés dans le tableau V-11 :

modèle locale f_i	a_1	a_2	b_{11}	b_{12}	b_{13}	b_{14}	b_0
f_1	-1.8395	0.8476	-0.3342	0.4313	0.3332	0.0046	1.5148

Tableau V- 11 : Paramètres du multi-modèle descriptif de la sortie t_4 du fuel system

La représentation mathématique de la sortie t_4 de la turbine est de la forme :

$$t_4^i(t) = -a_1^i t_4^i(t-1) - a_2^i t_4^i(t-2) + b_{11}^i m_8(t) + b_{12}^i m_{ff}(t) + b_{13}^i m_3(t) + b_{14}^i t_3(t) + b_0^i$$

La sortie t_4 présente la pression du gaz à la sortie de l'installation carburant, cette grandeur est accessible (mesurable), l'unité de mesure est le (kelvin).

À partir des figures V-51, nous pouvons constater que l'erreur d'estimation est très petite relativement au signal de sortie t_{o4} .

V-13-DUSCUTION ET INTERPRETAION DES RESULTATS:

- ❖ On constate une bonne adéquation entre la sortie réelle du processus modélisé (compresseur, turbine, installation carburant) et le modèle identifié.

On peut conclure que le multi modèle linéaire identifié par l'algorithme à deux niveaux construit un modèle très proche du modèle physique du (compresseur, turbine, installation carburant).

- ❖ L'algorithme à deux niveaux que nous avons développés dans ce mémoire présente une rapidité de calcul assez intéressante de l'ordre de quelques secondes. En outre, nous avons une erreur d'approximation à décroissance rapide.
- ❖ L'inconvénient de la partition grille est son caractère combinatoire qui se traduit d'une part par un nombre important de modèles locaux et de l'autre par l'existence de zones vides qui sont inutiles car elles n'apportent pas d'information pour l'explication du comportement du système.

CONCLUSION:

Ce chapitre a été dédié à l'identification hors ligne et l'optimisation paramétrique et structurelle d'un multi-modèle.

L'algorithme d'identification à deux niveaux décrit dans la figure V-13, alterne à la fois entre la détermination et l'optimisation des paramètres des fonctions d'activation, et la détermination des paramètres du multi-modèle.

L'inconvénient de la partition grille est son caractère combinatoire qui se traduit d'une part par un nombre important de modèles locaux et de l'autre par l'existence de zones vides qui sont inutiles car elles n'apportent pas d'information pour l'explication du comportement du système.

Ex : $n_z=5$, et $p_1=p_2=p_3=p_4=p_5=6$, ce que donne : $5^6 = 15625$ partitions.

Du constat que la partition grille génère une structure multi-modèle avec des modèles inutiles (couvrant des zones de fonctionnement vides) et des modèles adjacents pouvant décrire le même comportement du système, Pour remédier à ce problème une idée consiste à fusionner ou éliminer ces modèles.

Sommaire

V-1- INTRODUCTION :	122
V-2- PRINCIPE DE FONCTIONNEMENT :	123
V-3- MODELISATION ET IDENTIFICATION :	124
V-4- REPRESENTATION SOUS SIMULINK :	124
V-4-1- SIMUNLINK :	124
V-4-2- LES BRIQUES DE BASE :	125
V-4-3- SCHEMA BLOC SOUS SIMULINK DE LA TURBINE :	126
V-5- COMPRESSEUR :	128
V-6- LES ENTREES-SORTIES DU COMPRESSEUR :	128
V-7- TURBINE :	129
V-8- LES ENTREES-SORTIES DE LA TURBINE :	130
V-9- INSTALLATION CARBURANT :	131
V-10- LES ENTREES -SORTEES DE L'INSTALLATION CARBURANT :	131
V-11- PRESENTATION PAR MULTI-MODELES :	132
V-11-1- ALGORITHME D'IDENTIFICATION :	132
V-11-2- LES VARIABLES CARACTERISTIQUES CANDIDATES:	135
V-12- IDENTIFICATION ET OPTIMISATION PARAMETRIQUES:	135
V-12-1- COMPRESSEUR:	135
V-12-1-1-a- La sortie m_3 (écoulement de masse) :	135
V-12-1-1-b- La sortie t_3 (Température de gaz à la sortie du compresseur) :	138
V-12-1-1-c- La sortie q_c (couple du compresseur) :	141
V-12-2- TURBINE:	144
V-12-2-1-a- La sortie q_t (couple de la turbine) :	144
V-12-2-1-b- La sortie w (la puissance) :	146
V-12-2-1-c- La sortie t_7 (la température à la sortie de la turbine) :	150
V-12-2-1-d- La sortie m_8 (écoulement de masse à la sortie de la turbine) :	154
V-12-3- FUEL SYSTEM:	157
V-12-3-1-a- La sortie t_6 (température à la sortie de l'installation carburant) :	157
V-12-3-1-b- La sortie p_5 (température à la sortie de l'installation carburant) :	160
V-12-3-1-c- La sortie p_3 (la pression à la sortie de l'installation carburant) :	162
V-12-3-1-d- La sortie m_{10} (l'écoulement de masse à la sortie de l'installation carburant) :	165
V-12-3-1-e- La sortie t_4 (température du carburant à la sortie de l'installation carburant) :	167

V-13- DUSCUTION ET INTERPRETAION DES RESULTATS:.....	169
CONCLUSION:.....	170

Liste des Figures :

Figure V- 1 : Principe de fonctionnement d'une turbine à gaz	124
Figure V- 2 : Représentation du bloc turbine sous Simulink.....	126
Figure V- 3 : Développement du bloc turbine (compresseur+fuel system+turbine)	126
Figure V- 4 : Représentation du bloc « compresseur »	127
Figure V- 5 : Représentation du bloc « installation carburant »	127
Figure V- 6 : Représentation du bloc « turbine ».....	127
Figure V- 7: Le compresseur	128
Figure V- 8: Les entrées-sorties du compresseur	128
Figure V- 9: La turbine	129
Figure V- 10: Les entrées-sorties de la turbine	130
Figure V- 11: Fuel system	131
Figure V- 12 : Les entrées-sortie de l'installation carburant	131
Figure V- 13: Algorithme récapitulant la procédure d'identification d'un multi-modèle	134
Figure V- 14: La sortie m_3 (mass flow) du compresseur avec la sortie du Multi-modèle	135
Figure V-15: Erreur et Histogramme de prédiction	136
Figure V-16: Les sorties des 8 premiers modèles	136
Figure V-17: Les sorties des 8 derniers modèles	137
Figure V- 18: La sortie t_3 (température) du compresseur avec la sortie du Multi-modèle.....	138
Figure V- 19: Erreur et Histogramme de prédiction	139
Figure V-20: Les sorties des 8 modèles	139
Figure V- 21: La sortie q_c (couple) du compresseur avec la sortie du Multi-modèle	141
Figure V- 22: Erreur et Histogramme de prédiction	141
Figure V-23: Les sorties des 4 modèles	142
Figure V- 24: La sortie q_t (couple) de la turbine avec la sortie du Multi-modèle	144
Figure V- 25: Erreur et Histogramme de prédiction	144
Figure V-26: Les sorties des 3 modèles	145
Figure V- 27: La sortie w_t (power) de la turbine avec la sortie du Multi-modèle	146
Figure V- 28: Erreur et Histogramme de prédiction	147
Figure V-29: Les sorties des 8 modèles	148
Figure V- 30: La sortie t_7 (température) de la turbine avec la sortie du Multi-modèle.....	150
Figure V- 31: Erreur et Histogramme de prédiction.....	150
Figure V-32: Les sorties des 8 premiers modèles	151
Figure V-33: Les sorties des 8 derniers modèles	152
Figure V- 34: La sortie m_8 (mass flow) de la turbine avec la sortie du Multi-modèle	154
Figure V- 35: Erreur et Histogramme de prédiction	154
Figure V-36: Les sorties des 8 premiers modèles	155
Figure V-37: Les sorties des 8 derniers modèles.....	156
Figure V- 38: La sortie t_6 (température) de l'installation carburant avec la sortie du Multi-modèle ..	157
Figure V- 39: Erreur et Histogramme de prédiction	158
Figure V-40: La sortie du multi modèle.....	159
Figure V- 41: La sortie p_5 (pression) de l'installation carburant avec la sortie du Multi-modèle	160

<i>Figure V- 42: Erreur et Histogramme de prédiction</i>	<i>160</i>
<i>Figure V-43: Les sorties des modèles.....</i>	<i>161</i>
<i>Figure V- 44: La sortie p_3 (pression) de l'installation carburant avec la sortie du Multi-modèle</i>	<i>162</i>
<i>Figure V- 41: Erreur et Histogramme de prédiction</i>	<i>163</i>
<i>Figure V-46: La sortie du multi modèle.....</i>	<i>164</i>
<i>Figure V- 47: La sortie m_{10} (mass flow) de l'installation carburant avec la sortie du Multi-modèle..</i>	<i>165</i>
<i>Figure V- 48: Erreur et Histogramme de prédiction</i>	<i>165</i>
<i>Figure V-49: Les sorties des 4 modèles</i>	<i>166</i>
<i>Figure V- 50: La sortie t_4 (température) de l'installation carburant avec la sortie du Multi-modèle .</i>	<i>167</i>
<i>Figure V- 51: Erreur et Histogramme de prédiction</i>	<i>167</i>
<i>Figure V-52: La sortie du multi modèle.....</i>	<i>168</i>

Liste des tableaux :

<i>Tableau V- 1 : Paramètres principaux (ISO) de fonctionnement d'une turbine à gaz</i>	<i>124</i>
<i>Tableau V- 2 : Paramètres du multi-modèle descriptif de la sortie m_3 du compresseur</i>	<i>137</i>
<i>Tableau V- 3 : Paramètres du multi-modèle descriptif de la sortie t_3 du compresseur</i>	<i>140</i>
<i>Tableau V- 4 : Paramètres du multi-modèle descriptif de la sortie q_c du compresseur</i>	<i>142</i>
<i>Tableau V- 5 : Paramètres du multi-modèle descriptif de la sortie q_t du turbine.....</i>	<i>145</i>
<i>Tableau V- 6 : Paramètres du multi-modèle descriptif de la sortie w_t du turbine.....</i>	<i>148</i>
<i>Tableau V- 7 : Paramètres du multi-modèle descriptif de la sortie t_7 du turbine</i>	<i>153</i>
<i>Tableau V- 8 : Paramètres du multi-modèle descriptif de la sortie t_6 du fuel system.....</i>	<i>159</i>
<i>Tableau V- 9 : Paramètres du multi-modèle descriptif de la sortie p_5 du fuel system</i>	<i>162</i>
<i>Tableau V- 10 : Paramètres du multi-modèle descriptif de la sortie p_3 du fuel system</i>	<i>164</i>
<i>Tableau V- 11 : Paramètres du multi-modèle descriptif de la sortie t_4 du fuel system</i>	<i>169</i>

Conclusion générale

Afin de mener à bien une étude en simulation dans le but de commander ou de surveiller le fonctionnement d'un système, l'établissement d'un modèle du système est un passage obligatoire. Dans ce mémoire, notre intérêt s'est porté sur l'identification d'un système dynamique non-linéaire dans le domaine énergétique (turbine à gaz) en utilisant l'approche multi-modèle.

Le principe de l'approche multi-modèle constitue en la définition du modèle global du système comme une combinaison de modèles locaux valables dans certaines zones de fonctionnement de celui-ci.

Notre contribution se situe aux niveaux de l'application de l'approche multi-modèle à un composant stratégique qui est la turbine à gaz, élément essentiel dans la production de l'énergie et aussi dans le transport.

A cet effet nous avons adopté la stratégie suivante à savoir construire les modèles locaux par un algorithme à deux niveaux, qui alterne entre l'identification des paramètres de ces derniers par la méthode des moindres carrés régularisée, et l'estimation des paramètres des fonctions de validités par l'algorithme de Levenberg Marquard .

Sur le plan théorique, après une présentation détaillée du concept de multi-modèle, l'optimisation d'un multi-modèle à travers une partition grille a été traitée. Le principe de la grille est la décomposition de l'espace caractéristique (ou espace de fonctionnement) du système sous la forme d'un maillage régulier. Les zones de fonctionnement résultent de la combinaison de partitions individuelles définies sur le support des variables caractéristiques.

L'inconvénient de la partition grille est le nombre prohibitif de modèles locaux générés lorsque la dimension de l'espace caractéristique est élevée et/ou le nombre de partitions par variable est important.

L'algorithme en question est appliqué sur un ensemble des données obtenu à partir d'une simulation d'une turbine à gaz sous Simulink, développé dans l'université de Ferrara (Italie) ayant subi à une modification de certain bloc par nous-mêmes.

Les résultats en simulation obtenu nous ont permis de dégager certains désavantages dus à la méthode elle-même notamment l'existence des zones redondantes et des zones vide.

A fin de pallier à ces inconvénients des solutions intuitive sont proposés tel la fusion des modèles redondants et l'élimination des modèle vide, dans le cas où nous possédons des informations heuristiques cela peut être traduit par un gain assez appréciable du temps de calcul.

Concernant les perspectives ouvertes par ce travail nous pensons quelles sont prometteuses, en particulier nous mentionnons la possibilité de recherche de structure de commande adaptée à des systèmes complexe et précisément la turbine à gaz ainsi que l'aspect surveillance des systèmes énergétiques.

Annexe

Description of the Simulink Gas Turbine Model

1. Model Overview

Air flows via an inlet duct to the compressor; high pressure air from the compressor is heated in combustion chambers and expands through a single stage compressor turbine. A Butterfly valve provides a means of generating a back pressure on the compressor turbine (there is no power turbine present in the model). Cooling air is bled from the compressor outlet to cool the turbine stator and rotor.

The Governor regulates the combustor fuel flow to maintain the compressor speed at a set-point value. Under steady state conditions the power generated by the compressor turbine is balanced by that absorbed by the compressor, and losses.

The resistive elements represent pressure losses in components, capacitive elements component volumes.

2. Simulink block diagrams

2.1 Top level:

Components between the inlet duct and the compressor turbine outlet including mechanical losses, are represented by the block labelled '**turbine**'. The block inputs are ambient air temperature and pressure '**t1**' and '**p1**', fuel flow '**ff**' and turbine back pressure '**p7**'. Block outputs are turbine cooling air flow and temperature '**m10**' and '**t4**', turbine outlet mass flow and temperature '**m8**' and '**t7n**', and turbine speed '**Wt**'.

The turbine outlet duct volume is represented by the block labelled '**c5**', flows '**m10**' and '**m8**' are mixed in this element. The inputs to block '**c5**' are, turbine outputs '**m8**', '**m10**', '**t4**', '**t7n**' described above, together with pressurising valve mass flow '**m9**'. Block outputs are pressurising valve up-stream pressure and temperature '**p7**' and '**t7**'.

Inputs to the '**Pressurising valve**' block are up-stream pressure and temperature '**p7**' and '**t7**' described above, valve angle '**av**', and valve down steam pressure '**p8**', assumed to be ambient. Block output is valve flow '**m9**'.

The shaded blocks shown in the top left provide model initialisation and processing options, a mouse click on the appropriate block runs the underlying Matlab '*.m' file.

Block '**Init 1**' initialises the model data for investigating dynamic performance. Parameters are set-up to sequence the speed demand and valve angle.

2.2 Turbine:

Inlet Duct:

The inlet ducts losses are represented by resistive element block '**R1**' and the duct volume by capacitive element '**C1**'.

2.2.1 Compressor:

Mass flow and polytropic efficiency are evaluated as functions of inlet pressure, outlet pressures, inlet temperature and speed. Outlet temperature is calculated from pressure ratio, efficiency, inlet temperature, and mean gamma. Compressor shaft torque is calculated from mass flow, temperature drop, mean **Cp** and shaft speed.

2.2.2 Fuel system:

All components from the compressor outlet to turbine inlet are included in this block

Diffuser:

Block '**C2**' represents the gas volume of the compressor, and compressor outlet diffuser.

Blocks '**R2**' represents the diffuser losses.

Combustors:

All combustors are combined and modelled as a single component.

Block '**C3**' represents the volume associated with the outer cases of the combustion chambers.

Block '**R3**' represents the losses associated with air flow into the combustion chambers

Block '**Combustion**' represents the combustion process. Where:-

Output temperature $t_{5c} = \text{inlet temperature } t_5 + \text{temperature rise } dt$

Where dt is a function of air/fuel mass flow ratio.

Combustor outlet mass flow is the sum of the inlet air mass flow ' m_5 ' and fuel flow ' ff '.

Block '**C4**' represents the gas volume of the combustion chamber and the gas path between the combustion chamber and turbine inlet.

Cooling air bleed flows:

Air flow is diverted from the outlet of the compressor to cool the stator and rotor of the power turbine.

Blocks 'Chk R6' and 'Chk R10' represent the cooling air bleed flow paths from the compressor to the turbine. The flow through these paths is considered to be choked and a function of upstream temperatures and pressures. Several paths for cooling air are provided in the engine; these have been combined and reduced to two in the model. The path via block 'Chk R6' is assumed to enter the gas path prior to the power turbine and contribute to the turbine shaft power. The path via 'Chk 10' is assumed to enter the gas path after the power turbine with no contribution to the shaft power.

2.2.3 Compressor turbine:

The compressor turbine is modelled using generic curves to calculate turbine efficiency and mass flow.

Turbine efficiency is calculated from a series of modified parabolic curves which relate isentropic efficiency and velocity ratio. A Hodge curve is used to calculate mass flow as a function of pressure ratio. Both curves are scaled to fit the design point conditions.

From block inputs; inlet temperature, inlet pressure, outlet pressure, and shaft speed; the following block outputs are calculated; shaft torque, shaft power, isentropic efficiency, outlet temperature and mass flow.

2.2.4 Mechanical load:

This block represents the relationship between load torque and shaft speed.

Load torque ie the difference between compressor and turbine torques, is the input to the block, and shaft speed the block output.

Within the block torque losses, assumed proportional to speed², are subtracted from the load torque, divided by the load inertia and the result integrated to give shaft speed.

A feature is also included to ramp the speed linearly during the initial part of a simulation to the point where the component models are valid, in a similar way to the starter on the real engine.

2.3 Pressurising valve:

The pressurising valve is of the butterfly type. The upstream side of the valve is linked to the turbine outlet via duct volume, block 'c5', the downstream side is assumed to exhaust to atmosphere.

The valve is considered as a nozzle with both choked and un-choked regions modelled. The nozzle effective area is considered to be a function of both valve angle and pressure ratio, as defined by graphical information supplied by the valve manufacturer.

Block inputs are; upstream pressure, downstream pressure (ambient), upstream temperature and valve angle; block output is valve flow.

Valve angle is derived from a step function block which can be programmed to introduce a step in valve angle at any time during a simulation.

2.4 Speed governor:

No attempt has been made to model the fuel control system in detail.

Fuel control in the model is performed in the block 'Governor' with turbine speed 'Wt' as input and fuel flow 'ff' as output.

Fuel flow is initially increased as a linear ramp, accelerating the turbine. When the turbine speed reaches a pre-programmed value the ramp rate is increased. When the speed reaches a second pre-programmed value a proportional plus integral speed governor is brought into operation, and the turbine speed is controlled to a set-point. Finally with the turbine under speed control, the speed set-point is ramped to its steady state value.

3. Model Variable Names and Units

Model variables are computed in SI units as follows:

p	-	Pressure [N/m ²]
q	-	Torque [Nm]
m	-	Mass flow [kg/s]
ff	-	fuel flow [kg/s]
t	-	Temperature [K]
Nt	-	Shaft speed [rad/s]
av	-	Valve angle [Deg]
pct	-	Turbine shaft power [W]
turbE	-	Turbine isentropic efficiency [-]

Références bibliographiques

- [1] Abdelkader Akhenak . Mohammed Chadli. Josée Ragot. Didier Maquin. [CIFA 2006] « Synthèse d'un multi observateur robuste pour multi modèle incertain à entrées inconnues : approche LMI ».
- [2] Abdelkader Akhenak [2004]. « Conception d'observateurs non linéaires par approche multi modèle : application au diagnostic ».
- [3] Dalil Ichalal, Benoit Marx, José Ragot, Didier Maquin.[2008] « Diagnostic des systèmes non linéaires par approche multi modèle ».
- [4] T. Takagi and M. Sugeno. IEEE Transactions on Systems, Man, and Cybernetics, 15 :116–132, 1985 . «Fuzzy identification of systems and its applications to modeling and control».
- [5] Rodolfo ORJUELA, Didier MAQUIN et José RAGOT. [2006], « Identification des systèmes non linéaires par une approche multi-modèle à états découplés »
- [6] Liu Jizhen, Chen Yanqiao, Zeng Deliang, Yang Guangjun, [IEEE 2002], « Identification of a boiler-turbine system using T-S Fuzzy Model»
- [7] R.Babuska, J.A.Roubus, H.B.Verbruggen [1996], « Identification of MIMO systems by input-output TS Fuzzy Model»
- [8] Manuel Adam Medina, [2004], « Diagnostic de défauts des systèmes à représentation multi-modèle linéaire invariant dans le temps »
- [9] Mickaël Rodrigues, [2005] « Diagnostic et commande active tolérante aux défauts appliqués aux systèmes décrits par des multi-modèles linéaires »
- [10] Gregor Gregorcic Gordon Lightbodyb, [2008] «Nonlinear system identification: From multiple-model networks to Gaussian processes»

- [11] Rodolfo Orjuela, Benoît Marx, José Ragot et Didier Maquin, [2007], « Estimation d'état des systèmes non linéaires par une approche multi modèle découplé »
- [12] Tor A. Johansen and Bjarne A. Foss, [1995], « Identification of Non-linear System Structure and Parameters using Regime Decomposition »
- [13] Rodolfo Orjuela, Benoît Marx, José Ragot and Didier Maquin, [2008] « Estimating the state and the unknown inputs of nonlinear systems using a multiple model approach »
- [14] Mohammed Chadli, [2002], « Stabilité et commande de systèmes décrits par des multi modèles »
- [15] K.J. Astrom, R.D. Bell, [1999], « Drum-Boiler Dynamics »
- [16] Wen Tan, Horacio J. Marquez, Tongwen Chen, Jizhen Liu, [2005], « Analysis and control of a nonlinear boiler-turbine unit ».
- [17] F.P. de Mello, Fellow, [1991], « Boiler models for system dynamic performance studies ».
- [18] C.M. Cheng and N.W. Rees, [2008], « Application of the Fuzzy Dynamical Model Approach for identification of nonlinear Drum-Boiler Process »
- [19] Wen Tan, Horacio J. Marquez, Tongwen Chen, Jizhen Liu, [2005], « Analysis and control of a nonlinear boiler-turbine unit »
- [20] Andreas Fahlman, « A Modified Marquardt-Levenberg Parameter Estimation Routine for Matlab »
- [21] Oliver Nelles, [2000], « Nonlinear System Identification »
- [22] Rao V. Dukkipati, [2007], « Solving engineering system dynamics problems with MATLAB »
- [23] Ibrahim EKSIN, Osman Kaan EROL [2000], « A fuzzy identification method for nonlinear systems ».
- [24] M. Golob and B. Tovornik, [2000], « Identification of non-linear dynamics systems with decomposed Fuzzy Models ».
- [25] Lennart Ljung, [2005], « Identification of Nonlinear Systems ».
- [26] Imre Benyó, Jenő Kovács, Jari Mononen and Urpo Kortela [2004], « Modeling of steam temperature dynamics of a super heater ».