HAL
open science

Caractérisation des performances énergétiques des systèmes thermiques innovants pour le bâtiment au travers d'essais de courte durée en régime dynamique

Amine Lazrak

▶ **To cite this version:**

Amine Lazrak. Caractérisation des performances énergétiques des systèmes thermiques innovants pour le bâtiment au travers d'essais de courte durée en régime dynamique. Génie civil. Université Grenoble Alpes, 2015. Français. NNT : 2015GREAA031 . tel-01301671

HAL Id: tel-01301671

https://theses.hal.science/tel-01301671v1

Submitted on 12 Apr 2016

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

UNIVERSITÉ GRENOBLE ALPES

THÈSE

Pour obtenir le grade de

DOCTEUR DE L'UNIVERSITÉ GRENOBLE ALPES

Spécialité : **Génie Civil et sciences de l'habitat**

Arrêté ministériel : 7 août 2006

Présentée par

Amine LAZRAK

Thèse dirigée par **Gilles FRAISSE**

préparée au sein des **laboratoires LSBT du CEA et LOCIE de l'Université Savoie Mont Blanc**
dans **l'École Doctorale SISEO**

Caractérisation des performances énergétiques des systèmes thermiques innovants pour le bâtiment au travers d'essais de courte durée en régime dynamique

Thèse soutenue publiquement le **3 décembre 2015**,
devant le jury composé de :

Mr, Christian INARD
Professeur, Université de la Rochelle, président du jury
Mr, Joseph VIRGONE
Professeur, Université Lyon 1, rapporteur
Mr, Sylvain LALOT
Professeur, Université de Valenciennes, rapporteur
Mr, Gilles FRAISSE
Professeur, Université Savoie Mont Blanc, directeur de thèse
Mr, Bernard SOUYRI
Maître de conférences, Université Savoie Mont Blanc, invité
Mr, Philippe PAPILLON
Docteur, CEA, examinateur
Mr, Antoine LECONTE
Docteur, CEA, examinateur

Remerciements

Trois ans avant, précisément le 10 octobre 2012, j'ai commencé cette « aventure » de la recherche scientifique. Le début, comme la fin d'ailleurs, d'une thèse de doctorat sont toujours difficiles et nécessitent beaucoup de patience, d'organisation, de courage, d'encouragement, parfois de la foi mais surtout un bon encadrement. Je souhaite en conséquence présenter mes remerciements les plus sincères à toutes les personnes ayant contribué de près ou de loin, directement ou indirectement à l'accomplissement de ce travail et notamment à l'Agence de la Maîtrise de l'Energie et de l'Environnement et l'Institut National des Sciences et Techniques Nucléaires qui ont financé conjointement mes recherches.

Je tiens à remercier d'abord et de tout cœur mon directeur de thèse, M. Gilles Fraisse, Professeur à l'Université Savoie Mont Blanc. Grâce à son encadrement sans faille, ses précieux conseils et sa disponibilité malgré ses sollicitations diverses, j'ai pu avancer sereinement et dépasser les difficultés rencontrées tout au long de la période de la thèse.

Mes grands remerciements vont également à Dr. Philippe Papillon, directeur du laboratoire des systèmes thermiques à hautes températures au sein du CEA, qui est à l'origine de ce projet de recherche. Je lui suis profondément reconnaissant de m'avoir fait confiance en m'octroyant la possibilité de travailler sur ce sujet. Par son admirable esprit d'analyse ainsi que ses connaissances approfondies du secteur de l'énergie solaire thermique, il a été décisif pour la bonne orientation de mes travaux de thèse.

Je tiens à remercier chaleureusement Dr. Antoine Leconte, chercheur au sein du laboratoire des systèmes thermiques pour le bâtiment au sein du CEA. Il m'a fait beaucoup profiter de sa formidable expertise et maîtrise de la modélisation physique des systèmes solaires combinés. Je lui suis très reconnaissant car c'est grâce à sa supervision de mon travail et son aide de tous les jours, que j'ai pu atteindre les objectifs de la thèse dans les temps.

Je remercie profondément Dr. Bernard Souyri, Maître de conférences à l'Université Savoie Mont Blanc, de m'avoir honoré en faisant partie du comité d'encadrement de la thèse. Ses conseils très pertinents ainsi que sa disponibilité pour l'évaluation de mon manuscrit m'ont beaucoup aidé à proposer un mémoire cohérent et bien structuré.

Je souhaite également adresser ma grande gratitude aux deux rapporteurs, les Professeurs M. Sylvain Lalot de l'Université de Valenciennes et M. Joseph Virgone de l'Université Lyon 1 de m'avoir accordé de leur temps très précieux pour l'évaluation de mon travail de thèse. J'étais également honoré par la présence du Professeur M. Christian Inard de l'Université de La Rochelle et sa présidence du jury.

Je suis reconnaissant aussi envers MM. David Chèze et François Boudéhenn, chercheurs au sein du CEA, pour leur précieuse aide dans l'application de la méthodologie, développée dans le cadre de la thèse, dans le cas des systèmes solaires combinés à une pompe à chaleur et une machine à absorption.

Je souhaite remercier également tous les collègues du CEA ainsi que du laboratoire LOCIE pour leur sympathie, pour le respect qu'ils m'ont accordé mais surtout pour la très bonne ambiance qu'ils ont su créer au sein de nos espaces de travail. C'était vraiment pour moi un

très grand plaisir de travailler avec vous. Je tiens à remercier tout particulièrement les collègues du service avec qui je partageais le repas de tous les jours ; je pense notamment à Mathieu, Franck, Adrien, Didier, Joël et Philippe. Merci beaucoup pour les moments très agréables que j'ai pu partager avec vous mais aussi pour vos conseils et discussions très constructifs et utiles : c'était vraiment très intéressant !

Enfin, je souhaite remercier du fond du cœur tous les membres de ma famille tout particulièrement ma mère et mon père. Malgré les grandes distances qui nous séparaient ils étaient proches et j'ai toujours senti leur support et affection. C'est grâce à leurs prières et amour inconditionnel que j'ai pu surmonter les moments les plus difficiles et persévérer pour finaliser ce long travail que je leur dédie.

RESUME

Les systèmes solaires thermiques combinés à un système d'appoint tel que une chaudière, une pompe à chaleur ou intégrant une machine à absorption, peuvent jouer un rôle important dans la réduction des consommations des bâtiments pour les besoins de chauffage, de climatisation et de production d'eau chaude sanitaire. Dans ce sens, la caractérisation des performances énergétiques des systèmes thermiques est un enjeu crucial.

Les méthodes de caractérisation actuellement disponibles sont soit basées sur plusieurs essais physiques séparés des composants du système à évaluer, chose qui ne prend pas en compte les vraies interactions entre ces derniers, soit sur des modèles physiques qui peuvent être complexes et difficilement identifiables notamment du fait que les systèmes actuels sont compacts et préfabriqués en usine. En l'absence de méthode fiable pour estimer les performances des systèmes solaires thermiques avant leur intégration au bâtiment, le marché de ses derniers subit de fortes contraintes pour son développement.

Dans ce contexte il devient indispensable de développer une méthodologie générique qui peut être appliquée à différentes typologies de systèmes et qui pallie les difficultés rencontrées par les méthodes actuelles.

L'approche d'évaluation proposée dans ce mémoire est constituée de quatre étapes principales : la détermination d'une séquence de test, l'essai du système dans un banc d'essai semi-virtuel selon la séquence déterminée, l'acquisition des données et l'identification d'un réseau de neurones artificiels (RNA) du système et enfin la simulation du modèle en vue de l'estimation de la consommation du système dans l'environnement désiré. L'avantage d'utiliser un modèle complètement « boîte noire » du système complet à l'aide des RNA la rend totalement « non intrusive ». La connaissance des paramètres internes aux systèmes (rendements, conductivités thermiques, régulation etc.) n'est en conséquence pas nécessaire pour l'application de la méthodologie.

La validation de la méthodologie a été réalisée à travers plusieurs expérimentations numériques, pour 7 systèmes issus de 3 typologies différentes, durant lesquelles les estimations des RNA ont été comparées aux calculs des modèles physiques dans plusieurs conditions différentes (qualité du bâtiment, climat et surface du capteur). Une application de l'approche développée dans le cas de 5 systèmes réels a permis la confirmation de la pertinence de la méthodologie.

MOTS-CLÉS : Modélisation, Réseaux de neurones, Systèmes Thermiques, Bâtiment, Energie solaire.

ABSTRACT

Solar thermal systems combined with a backup system such as a boiler, a heat pump or incorporating an absorption chiller, can play an important role in reducing buildings energy consumption for heating, cooling and hot water production needs. In this sense, characterizing the energy performance of thermal systems is crucial.

Currently available methods of system characterization are either based on several separate physical tests of system components to be evaluated, which do not take into account the real interactions between them, or on physical models that can be complex and difficult to identify especially because systems nowadays are compact and prefabricated in the factory. Due to the lack of a reliable method to estimate the performance of solar thermal systems before their integration into buildings, their market faces a lot of impediment to be developed.

In this context, it becomes essential to develop a generic methodology that can be applied to different types of systems which overcomes the difficulties encountered by the current ones.

The proposed evaluation approach in this manuscript is composed of four main steps: determining a test sequence, testing the system in a semi-virtual test bench according to predetermined sequence, data acquisition and identifying an artificial neural network (ANN) of the system and finally the model simulation in order to estimate the system consumption in the desired boundary condition. Using a completely "black box" model of the whole system using the ANN makes the methodology totally "non-intrusive". No prior knowledge about the systems internal parameters (yields, thermal conductivities, regulation etc.) is necessary to apply the proposed approach.

The methodology validation was performed through several numerical experiments for seven systems coming from three different typologies. During the validation process, ANN estimates were compared with calculations of physical models in several different conditions (quality of building, climate and collector area). The developed approach was applied to five real systems as well. The application results allowed the confirmation of the methodology relevance.

KEYWORDS: Modeling, Artificial neural networks, Thermal systems, Building, Solar energy.

TABLE DES MATIERES

Résumé _____ *i*

Abstract _____ *ii*

Table des matières _____ *iii*

Nomenclature _____ *vii*

Abréviation _____ *x*

Introduction générale _____ *- 13 -*

Chapitre I. Enjeu de l'évaluation des performances énergétiques des systèmes thermiques à énergies renouvelables _____ *- 15 -*

I.1. Enjeu des systèmes thermiques innovants _____ - 17 -

I.1.1. Aperçu de la situation énergétique dans le monde _____ - 17 -

I.1.2. Besoins et consommations d'énergie dans les bâtiments _____ - 19 -

I.1.3. Focus sur la France _____ - 21 -

I.1.4. Les systèmes thermiques innovants: une solution ? _____ - 23 -

I.2. Les méthodes d'évaluation des performances énergétiques des SSC _____ - 38 -

I.2.1. Contexte environnemental et économique des SSC et PAC _____ - 39 -

I.2.2. Systèmes et performances énergétiques _____ - 42 -

I.2.3. Critères de performances _____ - 43 -

I.2.4. Méthodes de caractérisation des performances _____ - 43 -

I.2.5. Description de la méthode SCSPT _____ - 48 -

Chapitre II. Développement d'une méthodologie générique d'évaluation des SSC _____ - 56 -

II.1. Présentation de la méthodologie _____ - 58 -

II.1.1. Exigences vis-à-vis de la méthodologie _____ - 58 -

II.1.2. Les étapes de la méthodologie proposée _____ - 59 -

II.2. Modélisation adaptée à la problématique du sujet _____ - 60 -

II.2.1. Définition et principes de la modélisation _____ - 60 -

II.2.2. Structure de modèle selon une approche « boîte noire » _____ - 65 -

II.2.3. Réseaux de neurones artificiels et systèmes thermiques _____ - 70 -

II.3. Processus de modélisation avec les RNA _____ - 71 -

II.3.1. Les RNA formels _____ - 71 -

II.3.2. Caractéristiques du modèle neuronal _____ - 73 -

II.3.3. Techniques d'apprentissage _____ - 82 -

II.3.4. Données d'apprentissage _____ - 85 -

II.3.5. Processus complet d'entrainement et sélection du modèle _____ - 86 -

II.4. Configuration E/S et évaluation de la méthodologie _____ - 88 -

II.4.1. Les deux approches : « E/S physiques » et « flux énergétiques » _____ - 88 -

II.4.2. Evaluation de la méthodologie _____ - 90 -

II.5. Finalisation de la méthodologie par une application à un SSC _____ - 99 -
II.5.1. Description du système étudié _____ - 99 -
II.5.2. Lissage des données _____ - 100 -
II.5.3. Sélection de la configuration E/S de modélisation _____ - 101 -
II.5.4. Critères de sélection des RNA _____ - 110 -
II.5.5. Analyse des résultats d'application de la méthodologie _____ - 124 -

Chapitre III. Extension de la méthodologie à d'autres typologies de systèmes_____ - 134 -

III.1. Première typologie : systèmes solaires combinés à une pompe à chaleur _____ - 136 -
III.1.1. Description des systèmes et de leurs modèles physiques _____ - 136 -
III.1.2. Application de la méthodologie développée _____ - 141 -
III.1.3. Résultats et discussions _____ - 144 -

III.2. Deuxième typologie : systèmes solaires intégrant une machine à absorption _____ - 159 -
III.2.1. Description des systèmes et de leurs modèles physiques _____ - 159 -
III.2.2. Application de la méthodologie _____ - 162 -
III.2.3. Résultats et discussions _____ - 164 -

III.3. Bilan _____ - 166 -

Chapitre IV. Expérimentation physique : Application de la méthodologie dans le cas de SSC réels _ - 170 -

IV.1. Présentation détaillée du banc d'essai semi-virtuel _____ - 172 -
IV.1.1. Caractéristiques techniques du banc d'essai semi-virtuel _____ - 172 -
IV.1.2. La méthode SCSPT théorique et l'essai physique des systèmes _____ - 176 -

IV.2. Test du premier système commercial _____ - 177 -
IV.2.1. Description du système et de son modèle physique _____ - 177 -
IV.2.2. Résultats _____ - 181 -

IV.3. Test du système SSC-PROTO1 _____ - 188 -
IV.3.1. Description du SSC-PROTO1 _____ - 188 -
IV.3.2. Résultats _____ - 189 -

IV.4. Test du système SSC-PROTO2 _____ - 191 -
IV.4.1. Description du SSC-PROTO2 _____ - 191 -
IV.4.2. Résultats _____ - 192 -

IV.5. Exploitation des données de suivi in situ d'un SSC intégrant une machine à absorption _____ - 197 -
IV.5.1. Description du système _____ - 197 -
IV.5.2. Modes de fonctionnement _____ - 198 -
IV.5.3. Résultats _____ - 200 -

IV.6. Bilan _____ - 204 -

Conclusion générale _____ - 206 -

Références _____ - 210 -

Annexes _____ - 220 -

Annexe A. Tableaux récapitulatifs des indicateurs et critères de performances_____ - 222 -

Annexe B. Tentative d'amélioration de la séquence de test _____ - 224 -

B.1. Approche_____ - 224 -

B.2. Résultats _____ - 225 -

Annexe C. Profils hebdomadaire des gains de chaleur internes au bâtiment_____ - 227 -

Annexe D. Comparaison entre un entrainement en boucle ouverte et un entrainement en boucle fermée - 228 -

NOMENCLATURE

Variable	Désignation	Unité
	Par alphabet Grec	
α	Premier paramètre de la fonction objectif (cas de la régularisation bayésienne)	[-]
β	Inclinaison du capteur	[°]
γ	Orientation du capteur par rapport au sud	[°]
γ_S	Azimut solaire par rapport au sud	[°]
δ	Deuxième paramètre de la fonction objectif (cas de la régularisation bayésienne)	[-]
$\eta_{capt,par}$	Le rendement du capteur dans le système parallèle	[-]
θ	Angle d'incidence	[°]
θ_S	Angle d'incidence sur une surface horizontale	[°]
σ	Ecart-type	[-]
σ_{sb}	Constante de Stephan-Boltzman	[W.m^{-2}K^{-4}]
$(\tau\alpha)_{en}$	Produit transmittance-absorbance effective du capteur pour un rayonnement solaire direct à incidence normale	[-]
$\boldsymbol{\varphi}$	Vecteur de régression du modèle	[-]
ω	Un poids synaptique du RNA	[-]
	Par alphabet Français	
A	Surface du capteur	[m²]
b	Biais d'un RNA	[-]
b_0	Constante pour le calcul du facteur d'angle d'incidence	[-]
C	Nombre de neurones dans la couche cachée	[-]
c	Indice indiquant le numéro du neurone considéré dans la couche cachée	[-]
c_1	Coefficient de pertes du capteur linéaires	[W.m^{-2}K^{-1}]
c_2	Coefficient de pertes thermiques du capteur non linéaires	[W.m^{-2}K^{-2}]
c_3	Coefficient de pertes thermiques du capteur en fonction de la vitesse du vent	[W.m^{-3}K^{-1}]
c_4	Coefficient de pertes thermiques du capteur en fonction de la température du ciel	[W.m^{-2}K^{-1}]
c_5	Capacité thermique du capteur	[J.m^{-2}.K^{-1}]
c_6	Rendement optique en fonction du vent	[s.m^{-1}]
$COP_{pac,par}$	COP de la PAC dans le système parallèle	[-]
D_N	Ensemble des données d'apprentissage	[-]
d_{off}	Déphasage de la fonction sinusoïdale décrivant T_{tap}	[-]
$\boldsymbol{e}$	Perturbations (symbole en gras)	[-]
e	Indice indiquant le numéro de la variable d'entrée	[-]
E	Nombre d'entrées du RNA	[-]
E_L	Eclairement pour les longueurs d'onde supérieures à $3\mu m$	[W.m^{-2}]
$E[X]$	Fonction espérance mathématique	[-]
f	Fonction d'activation d'un neurone	[-]
F	Fonction mathématique représentant un système réel	[-]
$\hat{F}$	Estimation de F	[-]
F'	Facteur d'efficacité du capteur	[-]

G	Eclairement global	[W.m^{-2}]
G_b	Eclairement direct	[W.m^{-2}]
G_d	Eclairement diffus	[W.m^{-2}]
G_{grad}	Vecteur gradient	[-]
H	Matrice hessienne	[-]
K	Constante caractérisant le fonctionnement de l'émetteur	[W.K^{-n}]
$K_{\theta b}$	Facteur d'angle d'incidence pour le rayonnement direct	[-]
$K_{\theta d}$	Facteur d'angle d'incidence pour le rayonnement diffus	[-]
L	Fonction de Vraisemblance	[-]
$\dot{m}_{dhw}$	Débit de puisage d'ECS	[kg.h^{-1}]
$\dot{m}_{sh}$	Débit dans l'émetteur de chaleur	[kg.h^{-1}]
n	Exposant caractéristique du radiateur	[-]
N	Nombre d'échantillons de données	[-]
$N_{day,seq}$	Nombre de jours de la séquence de test	[-]
Obj	Fonction objectif	[-]
p	Fonction densité de probabilité	[-]
P_{aux}	Puissance du système d'appoint	[W]
$P_{capt,out}$	Puissance fournie par le capteur	[W]
P_{dhw}	Puissance des besoins d'ECS	[W]
P_{sc}	Puissance des besoins de climatisation	[W]
P_{sh}	Puissance des besoins de chauffage	[W]
$P_{sh,d}$	Puissance de dimensionnement de l'émetteur de chaleur	[W]
$P_{sol,net}$	Apports solaires nets sur le plan du capteur	[W.m^{-2}]
q	Nombre de poids synaptiques	[-]
$\mathcal{R}$	Ensemble des nombres réels	[-]
$R_{X,Y}$	Fonction d'inter-corrélation entre X et Y	[-]
S	Nombre de sorties	[-]
s	Indice indiquant le numéro de la variable de sortie	[-]
$tanh$	Fonction tangente hyperbolique	[-]
T_{capt}	Température du capteur solaire	[°C]
T_{dhw}	Température d'ECS	[°C]
$T_{dhw,cons}$	Température de consigne pour l'ECS	[°C]
T_{ext}	Température extérieure	[°C]
$T_{ext,d}$	Température extérieure de base de dimensionnement des émetteurs	[°C]
$time$	Temps	[h]
T_{int}	Température intérieure du bâtiment	[°C]
$T_{sh,cons}$	Température de consigne à l'intérieur du bâtiment	[°C]
$T_{sh,in}$	Température d'eau en entrée de l'émetteur	[°C]
$T_{sh,in,d}$	Température d'eau de dimensionnement en entrée de l'émetteur	[°C]
$T_{sh,out}$	Température d'eau en sortie de l'émetteur	[°C]
$T_{sh,out,d}$	Température d'eau de dimensionnement en sortie de l'émetteur	[°C]
T_{tap}	Température d'eau froide	[°C]
$T_{tap,av}$	Température moyenne de l'eau froide	[°C]
$\boldsymbol{u}$	Vecteur des entrées	[-]
v	Vitesse du vent	[m.s^{-1}]
$\boldsymbol{w}$	Vecteur des paramètres d'un RNA	[-]
$\boldsymbol{w_0}$	Vecteur des paramètres optimaux d'un RNA	[-]
$\widehat{\boldsymbol{w}}$	Estimation de $\boldsymbol{w}$	[-]

x	Désigne une variable mathématique standard	[-]
y	Variable de sortie d'un système	[-]
$\hat{y}$	Estimation de y	[-]
z	Désigne une variable mathématique standard pour représenter la sortie d'un neurone	[-]
	Autres symboles	
$\Delta\eta_{capt}$	La différence entre le rendement du capteur des systèmes parallèle et en série	[-]
ΔCOP_{pac}	La différence entre le COP de la PAC des systèmes parallèle et en série	[-]
ΔT_{tap}	Amplitude de T_{tap}	[°C]
	Principaux critères statistiques	
AIC	Critère d'information d'Akaike	[-]
BIC	Critère d'information bayésien	[-]
ErrTot	Erreur totale	[-]
R^2	Coefficient de détermination	[-]
RMSE	Racine de l'erreur moyenne quadratique	[-]
R2bar	Coefficient de détermination ajusté	[-]
R2nrj	R^2 appliquée aux données d'énergie prédites par le RNA	[-]

ABREVIATION

ADEME	Agence de l'Environnement et de la Maîtrise de l'Energie
AIE	Agence Internationale de l'Energie
BF	Boucle fermée
BO	Boucle ouverte
capt	Capteur
CCT	Concise Cycle Test
Clim'	Climatisation
Cons	Consigne
COP	Coefficient de performance
CTSS	Component Testing System Simulation
dhw	Domestic Hot Water
DJU	Degré-jour unifié
DST	Dynamic System Test
ECS	Eau chaude sanitaire
EF	Eau froide
EMQ	Erreur moyenne quadratique
ERA	Erreur relative absolue
ES	Early stopping
E/S	Entrées-Sortie
FA	Fonction d'activation
GDM	Gradient descent with momentum
GES	Gaz à effet de serre
GIEC	Groupe d'Experts Intergouvernemental sur l'Évolution du Climat
INES	Institut National de l'Energie Solaire
INSEE	Institut National de la Statistique et des Etudes Economiques
IIASA	International Institute for Applied Systems Analysis
KFCV	k-Fold cross validation
LM	Levenberg-Marquardt
MISO	Multiple input single output
Mtep	Million de tonnes équivalent pétrole
MV	Maximum de vraisemblance
NARMAX	Modèle non-linéaire autorégressif à moyenne ajustée et variable exogène
NARX	Modèle non-linéaire autorégressif à variable exogène
NDN	Nombre de neurones (dans la couche cachée)
NFIR	Modèle non-linéaire à réponse impulsionnelle finie
NNP	Neural network pruning
NOE	Modèle non-linéaire à erreur en sortie
nom	Nominal
NORM	Intervalle de Normalisation
PAC	Pompe à Chaleur
Par	Parallèle
RNA	Réseau de neurones artificiel
RP	Rétropropagation résilient
Sc	Space Cooling

SCG	Scaled gradient conjugué
SCP	Moyenne de la somme des carrés des poids synaptiques
SFH	Single Family House
sh	Space Heating
SHC	Solar Heating and Cooling
SISO	Single input single output
SPF	Seasonal performance factor
SSE	Sum of square errors
SSC	Système solaire combiné
SCSPT	Short Cycle System Performance Test
TD	Time delay
TRNSYS	Transient system simulation

INTRODUCTION GENERALE

Les besoins de chauffage, de climatisation et d'eau chaude sanitaire au sein des bâtiments, représentent environ 30% de la consommation en énergie finale au niveau mondial, selon l'AIE.

Le secteur du bâtiment est un poste important de consommation d'énergie finale et en conséquence constitue un potentiel important de réduction de la consommation d'énergie et d'émission de gaz à effet de serre puisque les sources d'énergie utilisées sont pratiquement toutes fossiles. Aussi, les énergies renouvelables ont connues un développement massif lors des dernières décennies. Les développements technologiques récents permettent de valoriser ces ressources via des systèmes solaires thermiques innovants transformant l'énergie solaire pour répondre aux besoins du bâtiment. Ces systèmes sont souvent combinés à un système d'appoint pour compléter la ressource solaire lorsqu'elle est insuffisante.

Dans ce contexte les systèmes thermiques innovants peuvent jouer un rôle déterminant pour atteindre les objectifs de réduction des consommations d'énergie des bâtiments et d'émissions des gaz à effet de serre.

Cependant, il a été remarqué sur le terrain que beaucoup d'installations ont des performances en deçà des performances attendues (mauvais dimensionnement, inadéquation à l'usage réel, mauvaise mise en œuvre, entretien/maintenance défaillant etc.), ce qui fait que le système ne permet pas de réduire effectivement la consommation de l'énergie comme souhaité. Or, une mauvaise maîtrise de l'intégration des systèmes hydrauliques, de la régulation et surtout des connexions entre les composants du système peut augmenter considérablement sa consommation. La caractérisation des performances énergétiques des systèmes thermiques devient en conséquence un enjeu crucial.

Actuellement, les méthodes de caractérisation sont principalement basées sur des essais en régime statique, sur une extrapolation à un régime dynamique, et sur une juxtaposition des différentes fonctions (chauffage, climatisation, production d'eau chaude sanitaire, ventilation) ou sources d'énergie (renouvelable, traditionnelle, intermittente, …). Il apparaît cependant que cette méthodologie, en limitant les effets des interactions entre fonctions et sources d'énergie, et en ne prenant en compte que partiellement les lois de contrôle-commande, est imprécise et empêche toute comparaison directe entre systèmes différents. Face à cette situation il devient impossible de garantir aux utilisateurs les performances du système pour l'environnement désiré (climat, qualité du bâtiment). Les conséquences de cette situation font que le marché des systèmes solaires thermiques subit de très fortes contraintes pour son développement.

Pour pallier cette difficulté, l'objectif de ce travail est de définir une méthodologie générique de caractérisation des performances énergétiques des systèmes thermiques innovants pour le bâtiment au travers d'essais de courte durée en régime dynamique.

La méthodologie proposée dans ce mémoire est constituée de quatre étapes principales : (i) la détermination d'une séquence de test, (ii) l'essai du système dans un banc d'essai semi-virtuel selon la séquence déterminée, (iii) l'acquisition des données et (iv) l'identification d'un modèle du système et enfin la simulation du modèle en vue de l'estimation de la consommation du système dans l'environnement désiré.

L'intérêt de cette nouvelle méthodologie par rapport aux méthodes existantes concerne sa capacité à s'adapter à différentes typologies de systèmes et la possibilité de caractériser les systèmes compacts (produits préassemblés en usine) qui sont de plus en plus présents sur le marché. En effet, la méthodologie est basée sur l'identification d'un modèle complètement « boîte noire » du système complet avec des réseaux de neurones artificiels, ce qui la rend totalement « non intrusive ». La connaissance des paramètres internes aux systèmes (rendements, conductivités thermiques, régulation etc.) n'est en conséquence pas nécessaire pour l'application de la méthodologie. Aussi, la méthodologie reste objective c'est-à-dire que son application ne dépend pas de l'utilisateur et peut être appliquée par un non expert ce qui la rend compatible avec un moyen de certification de systèmes.

Le présent mémoire de thèse est organisé en quatre chapitres et trois annexes.

Le premier chapitre présente quelques solutions de systèmes thermiques innovants ainsi que le contexte énergétique mondial et français et y situe la place du secteur du bâtiment. Il aborde également l'enjeu de l'évaluation des performances énergétiques des systèmes thermiques à énergies renouvelables et le rôle qu'ils peuvent jouer dans ce contexte.

Ensuite le deuxième chapitre aborde le processus de développement de la méthodologie. Les conditions auxquelles la méthodologie doit se conformer sont présentées. Selon l'étude réalisée une modélisation « boîte noire » par des réseaux de neurones artificiels apparait comme la technique de modélisation la plus adaptée à la problématique de la thèse.

Le processus de développement des modèles neuronaux, du choix du nombre de neurones dans la couche cachée et des entrées pertinentes à la question de la sélection du modèle, a été étudié et commenté dans ce deuxième chapitre. Enfin, une première application de la méthodologie a été réalisée dans le cas d'un système solaire avec un appoint réalisé par une chaudière à gaz.

Le troisième chapitre s'intéresse à l'extension de la méthode proposée pour de nouveaux systèmes à l'aide d'expérimentations numériques. Comme la méthodologie d'évaluation des systèmes solaires combinés doit être générique, elle a été appliquée à d'autres systèmes de deux typologies différentes. Ceci a permis d'investiguer son caractère générique.

Enfin le quatrième chapitre permet d'évaluer la méthodologie dans le cas de systèmes réels. En effet, la méthodologie a été appliquée à plusieurs systèmes testés selon la méthode de test SCSPT sur le banc d'essai semi-virtuel. Les estimations des modèles développés sont comparées aux mesures.

Ce travail se termine par une conclusion générale présentant la contribution de la thèse, les enseignements qu'il est possible d'en tirer et les perspectives d'amélioration et d'application.

Chapitre I. Enjeu de l'évaluation des performances énergétiques des systèmes thermiques à énergies renouvelables

INTRODUCTION

L'objectif de la thèse est de développer une méthodologie permettant l'estimation des performances énergétiques des systèmes solaires combinés pour les applications dans le bâtiment (chauffage, climatisation et production d'eau chaude sanitaire). La méthodologie doit être à la fois générique, valable pour différents types de systèmes, tout en gardant un niveau de complexité minimale pour qu'elle reste compatible avec une intégration dans une norme ou certification. Il s'agit dans ce premier chapitre de présenter, brièvement, le contexte énergétique actuel, le rôle que ces systèmes peuvent jouer dans ce contexte, et un état de l'art des méthodes actuellement disponibles.

Ce chapitre se décompose en deux parties principales. La première dresse tout d'abord un résumé de la question énergétique mondiale, et montre qu'il y a un fort potentiel d'économie d'énergie dans le secteur du bâtiment. Ensuite une étude bibliographique, non exhaustive, des travaux récents ayant pour objet l'étude des systèmes solaires combinés (SSC) est réalisée. Les SSC combinent la ressource solaire et une deuxième source d'énergie. Cette dernière peut être de nature non renouvelable (cas des SSC classiques) ou en partie renouvelable (cas des SSC combinés à une pompe à chaleur ou à une machine à absorption). Ces systèmes thermiques innovants apparaissent comme une solution alternative aux systèmes utilisés actuellement.

La deuxième partie recense, dans un premier temps, les contraintes majeures freinant le développement du marché des SSC. Il s'agit notamment du manque d'une approche globale et générique d'évaluation des performances de ces systèmes. Une présentation des différentes approches et méthodes développées avec leurs limites et avantages, est abordée dans un second temps.

Il ressort de ce chapitre, qu'une nouvelle méthodologie est indispensable. Cette méthodologie doit permettre l'estimation de la consommation annuelle des SSC, dans différentes conditions (climats, bâtiments etc.). Et ce, en se basant sur un seul test physique court selon la méthode de test SCSPT développée au sein de l'INES.

I.1. Enjeu des systèmes thermiques innovants

I.1.1. Aperçu de la situation énergétique dans le monde

La révolution industrielle, caractérisant le passage d'une société à dominante agricole et artisanale à une société commerciale et industrielle [1], a transformé de façon radicale les habitudes, le mode ainsi que les conditions de vie de l'homme contemporain. La découverte des ressources énergétiques fossiles, notamment le pétrole, semblait être une opportunité tant cherchée.

Les mesures de la température réalisées dans différents endroits du globe, au cours du 20ème siècle montrent une augmentation de la température moyenne par rapport au siècle précédent. Cette augmentation se serait déroulée en deux étapes, la première de 1910 à 1945, la deuxième de 1976 à aujourd'hui [2]. Ces deux étapes sont séparées par une période de légère baisse du niveau de température. En outre, les travaux de plusieurs chercheurs indépendants du GIEC montrent l'existence d'une corrélation entre la concentration en CO_2 et

la température à la surface de la terre [3]. Suite à ces constatations, l'exploitation massive des combustibles fossiles a été pointée du doigt comme étant certainement la responsable. Actuellement, les affirmations du GIEC, exprimées dans les différents rapports qu'il a produits [4], ont tranchées sur la responsabilité des énergies fossiles dans l'augmentation des gaz à effet de serre (GES) lors des dernières décennies. Actuellement la majorité des décideurs reconnaissent que le réchauffement climatique est d'origine anthropogène.

Tableau 1- 1 : Réserves et consommations des principales ressources énergétiques fossiles correspondants aux années *2013, **2012 et ***2011. Source : US Energy Information Administration

	Pétrole (milliards de barils)	Gaz (billions de litres)	Charbon (millions de kg)
Réserves	1 645,98*	193 866,60*	888 866 395,20***
Consommation	32,97*	3 398,88**	7 665 382,77**

Outre le caractère nuisible à l'environnement des énergies fossiles, les réserves énergétiques non renouvelables (actuellement connues) sont très limitées comme le montre le Tableau 1- 1 et n'arriveront pas à suivre la très forte consommation énergétique remarquée ces dernières années. En effet, avec le rythme actuel de la consommation mondiale et en se basant sur les données du tableau précédent, il en résulte qu'il reste seulement 50 années d'exploitation du pétrole, 57 années d'exploitation du gaz et 116 années pour l'exploitation du charbon. Les réserves de gaz de schiste sont abondantes mais restent également limitées: elles représentent, selon l'AIE, environ 120 à 150 ans de la consommation actuelle de gaz naturel.

La situation énergétique et environnementale globale est encore plus complexe. La population mondiale ne cesse d'augmenter de façon presque linéaire pendant les trois dernières décennies (Figure 1- 1). La croissance de la population génère naturellement plus d'activités et crée plus de besoins. Ainsi, entre 1980 et 2012, la consommation en énergie primaire a augmenté de 83% et de 70% pour le CO_2, avec une augmentation moyenne annuelle d'environ 2,6% et 2,2% respectivement.

Les récentes prévisions montrent que la demande d'énergie continuera d'augmenter. Selon [5] la consommation de l'énergie dans les pays en voie de développement augmentera avec une moyenne annuelle de 3.2%, et dépassera celle des pays développés à partir de l'année 2020 dont le ratio moyen annuel sera d'environ 1,1%.

Face à cette situation inédite et en dehors d'une percée scientifique majeure permettant le captage du CO_2, deux possibilités s'offrent à l'humanité pour garantir un environnement durable et sain : favoriser l'utilisation d'énergie provenant de sources renouvelables ou tout simplement consommer moins d'énergie soit par sobriété, soit en améliorant l'efficacité des systèmes énergétiques. L'apport du présent travail, présenté dans les sections suivantes, touche à la fois ces deux aspects.

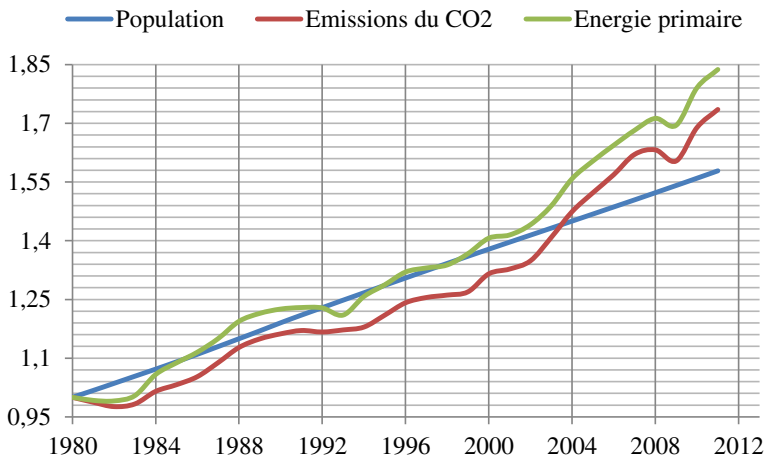

Figure 1- 1 : Evolution relative de la consommation mondiale en énergie primaire, des émissions du CO_2 et de la population mondiale. L'année de référence est 1980. Source : AIE et The World Bank

I.1.2. Besoins et consommations d'énergie dans les bâtiments

La consommation en énergie finale est souvent attribuée à quatre principaux secteurs économiques : l'industrie, le transport, le résidentiel et le tertiaire. La Figure 1- 2 montre que le secteur industriel est le plus énergivore avec environ 51% de consommation suivi du secteur du bâtiment avec une valeur de 30% (résidentiel et tertiaire). En conséquence, 30% d'énergie finale consommée dans le monde est due aux besoins dans les bâtiments (bureaux, commerces, hôpitaux, écoles, hôtels etc.). La recherche permanente du confort, la croissance de la population ainsi que l'augmentation du temps passé dans les bâtiments sont les principales raisons qui ont fait que ce secteur dépasse le niveau de consommation du secteur des transports [5].

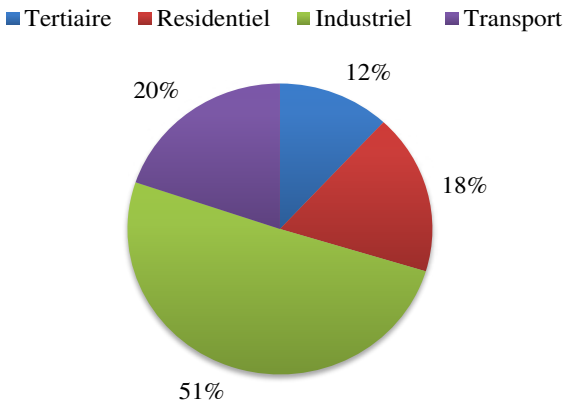

Figure 1- 2 : Consommation en énergie finale mondiale par secteur en 2011. Source : AIE

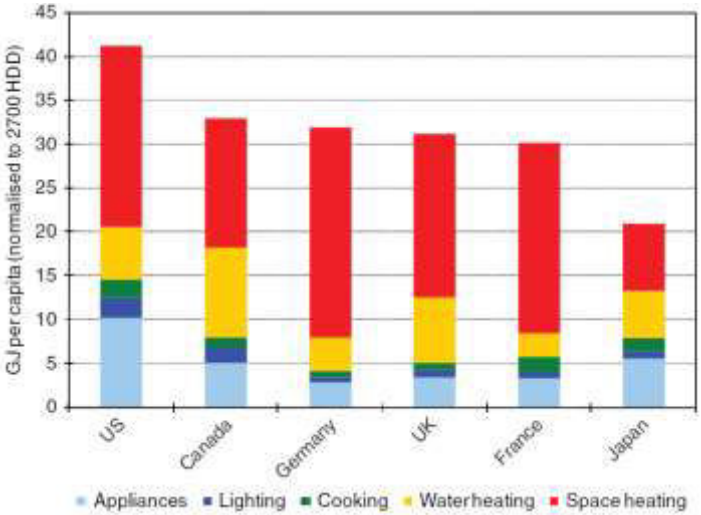

Figure 1- 3 : L'utilisation d'énergie par service dans les pays développés en 2007. Source : rapport d'IIASA [6] (HDD = heating degree day)

Les besoins pour le chauffage, la production de l'eau chaude sanitaire (ECS) et pour la climatisation (dans les climats chauds) sont les services responsables de la grande partie de consommation d'énergie dans les bâtiments (tous types de bâtiment confondus).

Comme le montre la Figure 1- 3, l'utilisation de l'énergie attribuée à ces services, dans les pays développés, représente approximativement entre 60% et 87% d'énergie consommée dans le bâtiment. Une grande partie de cette énergie est de source fossile (Figure 1- 4). Les logements et les bâtiments de type tertiaire sont ainsi à l'origine d'environ 33% des émissions de CO_2, 66% de chlorofluorocarbures et entre 25% et 33% de noir de carbone [6]. Dans les pays du sud, la consommation d'énergie dans le bâtiment est également prépondérante. Elle représente plus de 45% de l'énergie totale consommée par tous les secteurs dans ces pays en 1992 selon les statistiques des Nations Unies. Dans les zones climatiques chaudes les besoins de climatisation prennent la place des besoins de chauffage. Avec une population extrêmement élevée et des ressources énergétiques uniquement basées sur des énergies émettrices de GES, la situation énergétique et environnementale dans ces pays pourrait être plus préoccupante dans le futur.

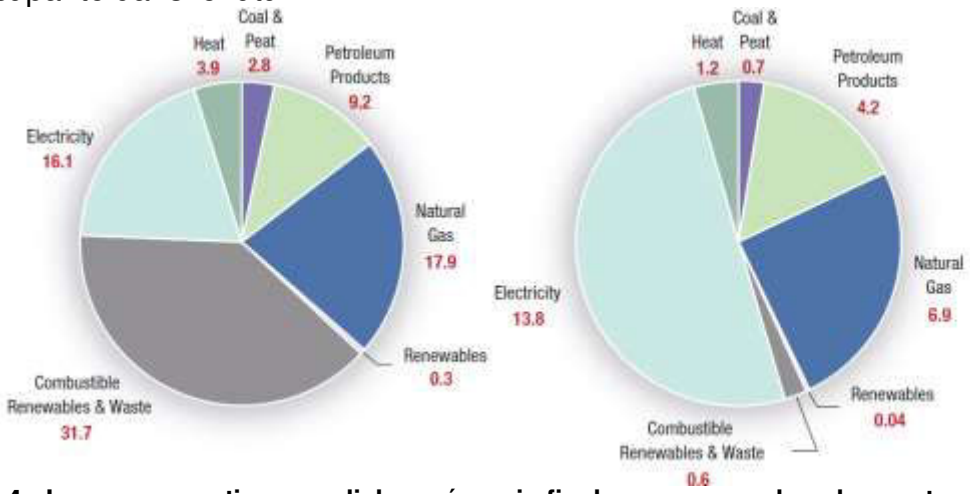

Figure 1- 4 : La consommation mondiale en énergie finale par source dans les secteurs résidentiels (gauche), commerciaux et publics (droite) en 2007 [6]

I.1.3. Focus sur la France

Il est vrai que les besoins unitaires des bâtiments en France diminuent dans le temps (grâce à l'isolation thermique, la prise en compte des aspects énergétiques durant la construction, la réhabilitation des logements etc.). Cependant la demande en énergie dans ce secteur reste très importante notamment en raison du stock important de bâtiments existants fortement consommateurs.

L'analyse des données statistiques concernant le secteur du bâtiment révèle son poids, en tant que poste de consommation, dans le contexte énergétique français. Globalement la consommation du secteur du bâtiment (résidentiel et tertiaire) est pratiquement stable depuis 2003. Avec 68,7 Mtep de consommation en énergie finale en 2012, il représente 44,5% de la consommation totale, loin devant les transports (31,8 %), l'industrie (20,8 %) et l'agriculture (2,9%) (Figure 1- 5).

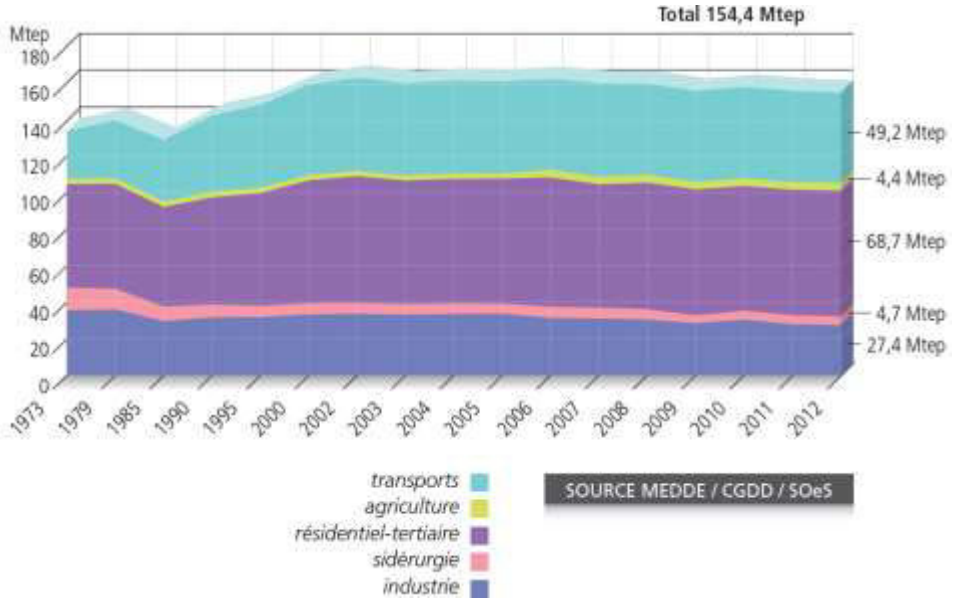

Figure 1- 5 : Évolution de la consommation finale énergétique nationale par secteurs d'activités à climat normal (pas d'effet de température). Source : chiffres clés ADEME, 2013

La consommation totale se décompose en deux tiers pour les bâtiments d'habitation (résidences principales et secondaires) et un tiers pour le secteur tertiaire [7].

En 2012, les besoins de chauffage, de production de l'ECS et de la climatisation représentaient environ 73,4% de la consommation énergétique du secteur résidentiel et 66% dans le secteur tertiaire (chiffres clés ADEME, 2013). Les émissions unitaires de CO_2 attribuées au secteur résidentiel (principalement pour les besoins de chauffage et d'ECS) ont diminué considérablement depuis les années 1970 (Figure 1- 6). Néanmoins, la quantité des émissions des GES reste importante (77,44Mt de CO_2, 1218,56kt de CO et 28,05kt de SO_2 en 2011).

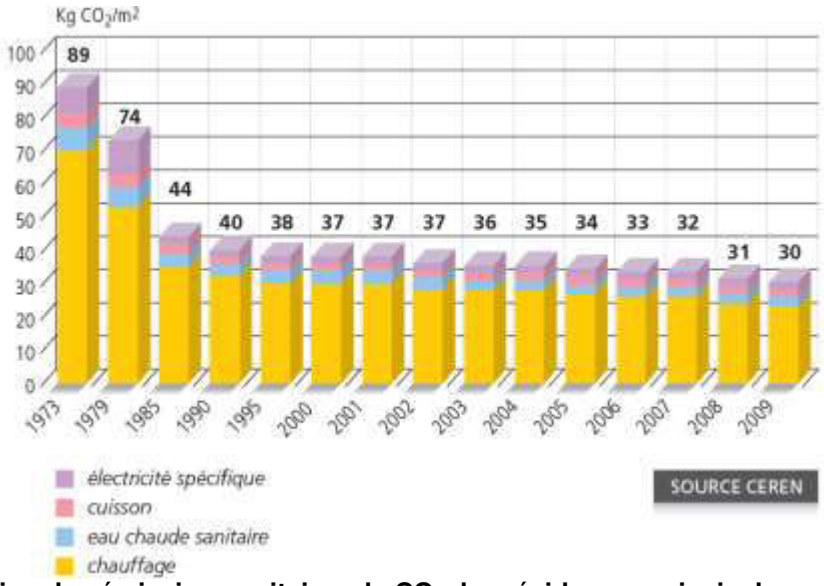

Figure 1- 6 : Évolution des émissions unitaires de CO_2 des résidences principales par usage à climat normal. Source : chiffres clés ADEME, 2013

Selon la Figure 1- 7, le gaz occupe une place prépondérante parmi les types d'énergie utilisés pour le chauffage des résidences principales (36,1%). L'électricité chauffe 31,5 % de l'ensemble des logements. Selon l'ADEME, le mazout est le troisième combustible utilisé dans l'ensemble des logements avec un taux d'équipement de 14,6 %. Le réseau de chaleur, avec 5%, prévaut essentiellement dans les résidences collectives. L'analyse de la consommation d'énergie du secteur tertiaire entre 2001 et 2010 (Figure 1- 8) montre que cette dernière est principalement issue de sources fossiles. Les énergies renouvelables sont donc très minoritaires dans le secteur du bâtiment. Les logements et les bâtiments de type tertiaire sont ainsi à l'origine de 22% des émissions nationales de gaz à effet de serre en 2012 (selon l'ADEME). Ce pourcentage est encore plus prononcé pour les pays dont les ressources énergétiques sont fortement émettrices des GES (contrairement à la France pour laquelle 75% de la production en énergie électrique est d'origine nucléaire).

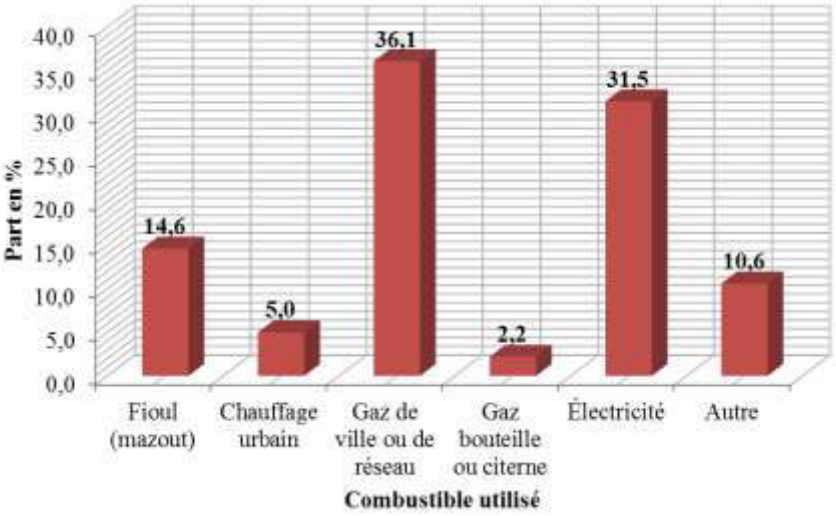

Figure 1- 7 : Résidences principales selon le combustible utilisé pour le chauffage en 2010 (INSEE)

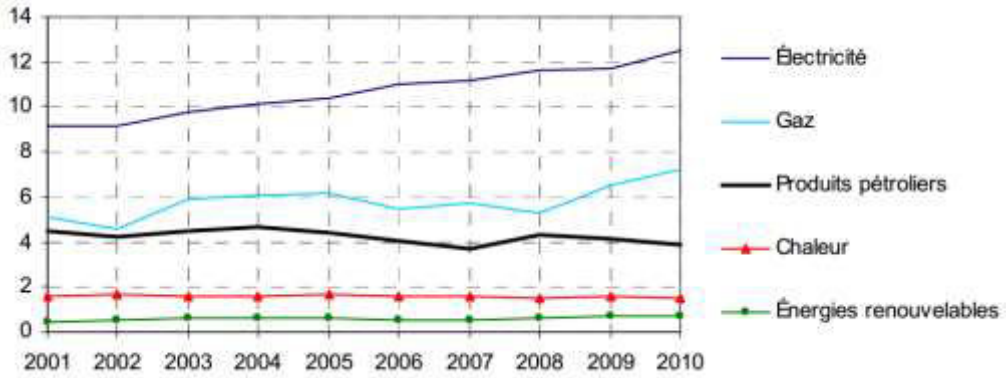

Figure 1- 8 : La consommation d'énergie du secteur tertiaire entre 2001 et 2010 en millions de tep. Source : Commissariat général au développement durable

Face à cette situation inédite, la France s'est engagée, selon l'INSEE, à diviser par quatre ses émissions de GES d'ici à 2050, afin de contenir le réchauffement climatique à un niveau d'élévation de 2°C par rapport à 1990. Elle s'est engagée également à passer d'ici 2020 à 20% de part des énergies renouvelables dans la consommation énergétique globale, à réduire les émissions de GES d'au moins 20% par rapport à 1990, et à réaliser 20% d'économies d'énergie par rapport aux projections réalisées pour l'année 2020.

Le respect de ces engagements, au moins, en ce qui concerne le secteur du bâtiment est, selon l'ADEME, conditionné par trois éléments :

1. Un plan ambitieux de construction et de rénovations thermiques.
2. Des équipements plus performants (fonctionnant principalement avec des énergies renouvelables) pour un niveau de confort au moins équivalent.
3. Des usages spécifiques rationalisés et sobriété énergétique.

Le présent travail de recherche s'inscrit dans le deuxième volet d'action, à savoir, le développement d'une méthode de caractérisation des systèmes thermiques pour améliorer leur intégration au bâtiment et par conséquent, réduire la consommation d'énergie dans ce secteur.

I.1.4. Les systèmes thermiques innovants: une solution ?

Le rayonnement solaire atteignant la surface de la Terre a une valeur maximale d'environ 1000W/m². Certes, cette valeur est atténuée par les nuages, néanmoins la quantité d'énergie reçue et réellement exploitable, reste suffisante pour satisfaire une grande partie des besoins énergétiques mondiaux. En effet, l'énergie solaire reçue au sol est 10000 fois supérieure à la consommation mondiale actuelle d'énergie [7]. L'énergie solaire peut être exploitée de plusieurs manières : (i) Utilisation passive de l'énergie solaire pour minimiser les besoins de chauffage, (ii) Utilisation active de l'énergie solaire via des systèmes thermiques innovants, (iii) Conversion thermodynamique de l'énergie solaire en électricité et (iv) Conversion photovoltaïque de l'énergie solaire.

Les travaux de recherche de cette thèse s'intéressent à la deuxième manière de valorisation des énergies renouvelables. En effet, des systèmes thermiques innovants capables de transformer l'énergie solaire en énergie utile, existent. En outre de l'énergie solaire, ces systèmes sont également capables de puiser l'énergie stockée dans l'air ou le sol (gratuite et permanente). Dans la littérature (par exemple dans [8]) il est possible de trouver la

description de différents systèmes thermiques à énergie renouvelable pour les applications dans le bâtiment.

Dans le but de faire apparaitre le rôle important que ces systèmes innovants peuvent jouer dans la réduction de la consommation en énergie finale dans le secteur du bâtiment, une présentation de ces derniers a été réalisée. Il s'agit d'une présentation des différentes catégories de systèmes, leurs composants et des travaux récents sur ce sujet.

Les informations présentées dans les trois paragraphes suivants sont basées sur les travaux présentés dans [9-35].

I.1.4.1 Les systèmes solaires combinés

Un système solaire combiné est un système thermique actif qui a comme fonction la transformation du rayonnement solaire disponible en énergie thermique utile, et ce pour répondre, en partie, aux besoins en chauffage et en eau chaude sanitaire (combinaison de deux fonctions) d'un bâtiment.

Le rayonnement solaire annuel est si abondant qu'il est possible, théoriquement, de répondre aux besoins thermiques d'un bâtiment rien qu'avec de l'énergie solaire. En effet, sous l'hypothèse d'une irradiation de 1000kWh/m² annuelle, l'énergie solaire captée par un panneau solaire thermique de 25m² de surface, bien orienté, pourrait couvrir les besoins thermiques d'un bâtiment moyennement isolé accueillant une famille de 4 personnes à Stockholm [9]. Aussi, pour les habitations labellisées BBC (bâtiment base consommation), une surface de seulement 10m², soumise aux mêmes conditions d'irradiation, pourrait être suffisante [9]. Cependant, cette abondante irradiation annuelle globale ne peut être totalement exploitée en pratique, pour répondre aux besoins des bâtiments. En effet, l'inconvénient des systèmes solaires thermiques est que l'évolution du rayonnement solaire n'est pas en phase avec celle des besoins en chauffage (Figure 1- 9). Les besoins en ECS sont sensiblement constants sur toute l'année.

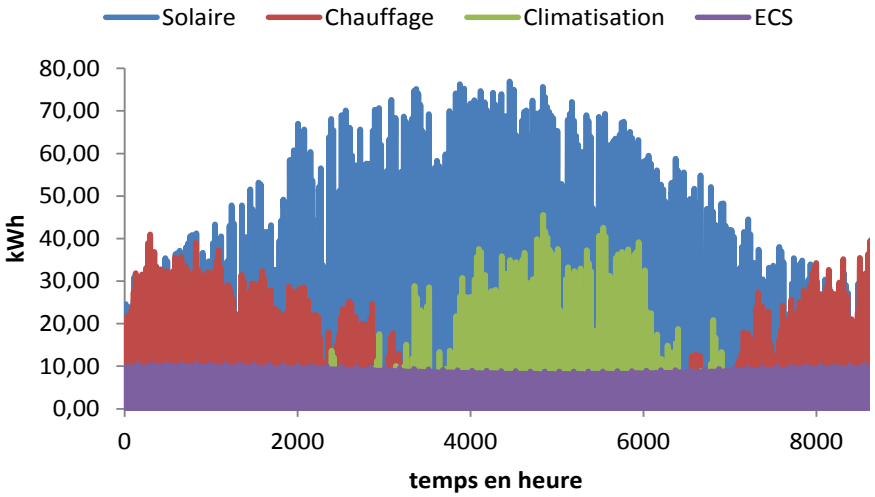

Figure 1- 9 : Evolutions au cours de l'année de l'énergie solaire (capteurs solaires de 8m²), des besoins en chauffage, en ECS et en climatisation (dans la ville de Réveillon)

La figure précédente montre que l'énergie solaire est disponible en abondance uniquement pendant l'été alors que pendant cette période les besoins en chauffage sont très faibles voire

inexistants. L'évolution inversée entre les besoins et la ressource rend l'utilisation d'un système d'appoint indispensable pendant la période hivernale. En outre, il est nécessaire d'utiliser un système de stockage de l'énergie thermique pour résoudre le problème de déphasage entre la ressource solaire et les besoins de chauffage (notamment pendant la nuit) à l'échelle journalière. Pour toutes ces raisons, le dimensionnement d'un SSC n'est pas juste une question de choix de la surface du capteur solaire. Ce type de système assez complexe, nécessite la prise en compte dans le dimensionnement de tous les aspects entrant en jeu dans le processus de la transformation des ressources (solaire et appoint) en énergie thermique utile : captage de l'énergie solaire, stockage de l'énergie, gestion de l'appoint [9]. Un tel dimensionnement doit résulter en un système qui fait appel le moins possible au système d'appoint et valorise au maximum l'énergie solaire sans pour autant compromettre la couverture des besoins du bâtiment.

Sur la Figure 1- 10 est représenté un exemple type d'un système solaire combiné avec un appoint par une chaudière. Tous les SSC sont constitués de quatre composants principaux (en dehors des organes hydrauliques et de protection) :

- Capteur solaire thermique : élément nécessaire pour capter le rayonnement solaire et le transformer en énergie thermique. La chaleur est ensuite transférée, au sein du capteur même, à un fluide caloporteur. Le rendement d'un capteur solaire (rapport entre l'énergie extraite en sortie du capteur et l'énergie reçue à sa surface) est une fonction décroissante de la différence de température entre la température du capteur et de l'air ambiant [10]. Il est donc indispensable de faire fonctionner le capteur solaire à des niveaux de températures les plus proches possibles de la température de l'air ambiant ce qui réduit les pertes de chaleur du capteur vers l'extérieur. Plusieurs technologies ont été développées pour optimiser le captage de l'énergie solaire. Elles diffèrent dans les matériaux utilisés pour absorber le rayonnement solaire ainsi que par leur architecture. Les capteurs les plus utilisés dans les installations thermiques pour le bâtiment sont les capteurs plans vitrés et à tubes sous vide (Figure 1- 11). Les derniers fonctionnent avec des niveaux de températures qui peuvent être supérieurs à 100°C.

- Système de stockage : la fonction de ce composant consiste à stocker toute l'énergie thermique disponible (fournie par le capteur solaire) dans un ballon hydraulique ou un plancher chauffant pour son utilisation au moment opportun. En effet, les horaires d'ensoleillements sont très souvent en déphasage avec les heures de consommation. Les systèmes de stockage actuels sont utilisés uniquement pour résoudre le problème de déphasage journalier. Des travaux de recherches sont en cours dans le but de développer des systèmes de stockage inter-saisonnier (stockage thermochimique par exemple [11]). Ce type de système est très prometteur car, en théorie, ils peuvent stocker toute l'énergie solaire disponible pendant la période estivale pour la restituer lors de la période hivernale.

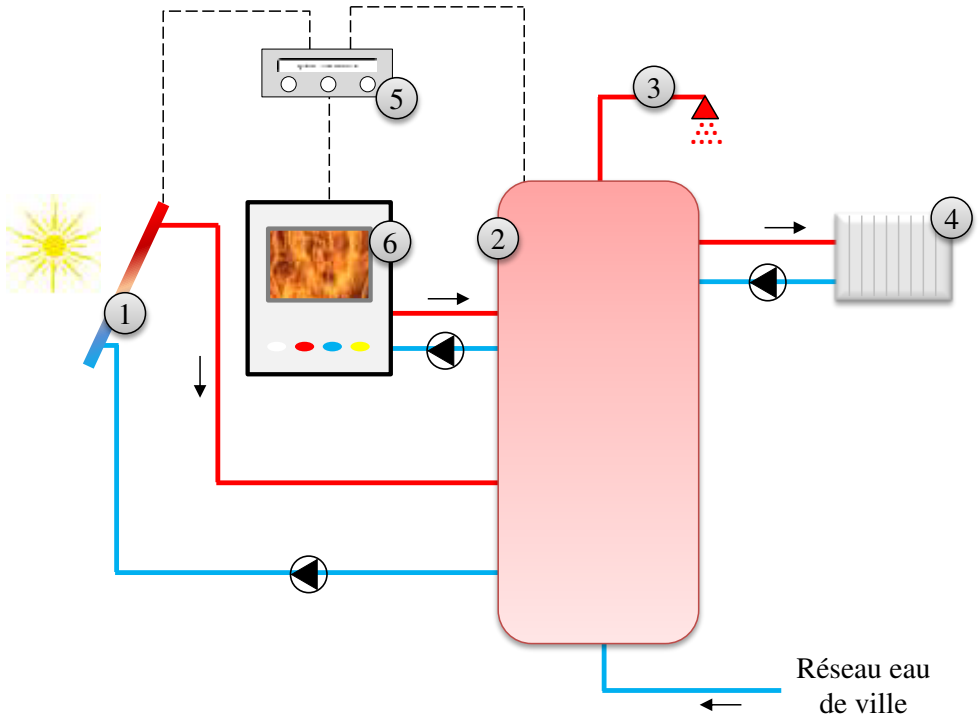

Figure 1- 10 : Schéma de principe Système solaire combiné avec un appoint par une chaudière à gaz : chauffage et ECS. (1) Capteur solaire, (2) système de stockage, (3) ECS, (4) émetteur, (5) régulation, (6) système d'appoint

Figure 1- 11 : Capteurs solaires les plus utilisés. Sources : Home Ecoenergie (capteur tube sous vide) à gauche et Solarskies (capteur plan vitré) à droite

- Systèmes d'appoint : son rôle consiste à fournir la quantité d'énergie nécessaire au système pour répondre aux besoins du bâtiment dans le cas de ressources solaires insuffisantes. L'énergie d'appoint peut être de différents types : électricité, fioul, gaz ou bois. Les conditions de fonctionnement de la chaudière dépendent de la source d'énergie sélectionnée [9]. En effet, dans le cas d'une chaudière à condensation, il est nécessaire de faire en sorte que la température de retour soit la plus basse possible pour récupérer l'énergie de condensation en plus de l'énergie sensible. Tandis que cette température dans le cas d'une chaudière à bois doit être élevée pour pallier les problèmes de corrosion et de dépôt liés à la fumée de combustion. Le système de contrôle de la chaudière doit être adapté au type de chaudière car certaines ont la possibilité de varier la puissance en fonction de la demande alors que d'autres fonctionnent uniquement en « Marche-Arrêt ». Pour éviter les situations d'inconfort dans le bâtiment, liées à une quantité insuffisante de chaleur solaire, il est

conseillé de dimensionner le système de telle sorte que la chaudière peut répondre à elle seule aux besoins du bâtiment.

- Le système de régulation : il contrôle le système solaire pour gérer au mieux les flux énergétiques entre ses composants (stockage, source d'appoint et source solaire). Le contrôle de l'installation se fait à travers les différentes vannes et pompes de circulation au niveau de chaque circuit (solaire, chauffage, appoint).

Les systèmes disponibles dans le commerce se distinguent les uns des autres par la manière dont les flux d'énergie au sein de chaque système sont gérés et du principe de stockage utilisé. En effet, il existe deux grandes catégories de SSC :

- Systèmes stockant l'énergie dans un ballon hydraulique d'un volume allant de 500l à 1500l en fonction des besoins énergétiques du bâtiment et de la taille de l'installation. Il est toujours nécessaire de garder une stratification thermique au sein du ballon (cela permet d'augmenter de 15% les performances du système [12]). En effet, la stratification permet de garder une température élevée dans la partie haute du ballon pour les besoins en ECS et une température moins élevée dans la partie basse du ballon nécessaire pour un meilleur fonctionnement du capteur solaire. Une telle gestion de l'énergie au sein du stockage permet de réduire l'appel au système d'appoint et ainsi de consommer moins d'énergie.

- Systèmes stockant l'énergie dans une dalle épaisse (plancher chauffant d'une douzaine de centimètres d'épaisseur). L'intérêt de ce type de système réside dans le fait qu'il n'est pas nécessaire d'intégrer un ballon de stockage à l'installation solaire. Le plancher chauffant joue à la fois le rôle du système de stockage et de l'émetteur de chaleur. Sans élément intermédiaire entre le capteur solaire et le bâtiment, les pertes thermiques sont moins importantes ce qui permet d'améliorer les performances du système. En l'absence de besoins en chauffage, le capteur solaire chauffe uniquement un ballon de petite taille pour les besoins d'ECS. La programmation du système de régulation doit prendre en compte l'inertie du bâtiment pour gérer le déphasage entre les horaires de l'ensoleillement maximal et les besoins de chauffage. Ainsi, il sera possible d'éviter des températures excessives qui peuvent générer de l'inconfort au sein du bâtiment.

Au sein de ces deux catégories, il existe plusieurs façons de gérer le décalage entre la disponibilité de l'énergie solaire et les besoins de chaleur. Les différents systèmes se différencient selon les stratégies adoptées pour aborder notamment les deux points évoqués ci-dessous [9] :

- L'énergie thermique nécessaire pour répondre aux besoins de chauffage et d'ECS est stockée dans un ou plusieurs ballons hydrauliques.
- L'utilisation de l'appoint : en parallèle avec le capteur solaire pour chauffer l'eau dans le ballon de stockage, séparé du ballon de stockage (l'énergie d'appoint est transférée directement aux besoins) ou intégré au ballon.

Enfin il est important d'attirer l'attention du lecteur que les performances des systèmes solaires thermiques sont très dépendantes de la charge à laquelle ils sont confrontés. En effet, il a été

montré dans [13] que pour qu'un SSC fonctionne de manière optimale, il est plus important d'utiliser des composants bien dimensionnés plutôt que des composants performants.

Une description détaillée des différentes classes de systèmes est donnée dans les livrables du projet européen Combisol ([14] et [15]).

I.1.4.2 Pompe à chaleur et SSC

Parmi les différents types de SSC existants, certains fonctionnent avec une PAC comme système d'appoint. Leur fonctionnement est le même que les systèmes solaires combinés classiques. Seul le type d'appoint est différent Figure 1- 12.

Une pompe à chaleur est un système thermodynamique (Figure 1- 13) qui puise des calories de l'air ou du sol (en fonction de la source de chaleur considérée) pour les transférer via un cycle thermodynamique au milieu à traiter. Elle est composée de quatre éléments :

- Evaporateur : un échangeur de chaleur qui sert à prélever de la chaleur au niveau de l'environnement (source de chaleur : air, sol, eau etc.) qui sera captée par le fluide caloporteur (à faible point d'ébullition). Le fluide change d'état et se transforme en vapeur.
- Le compresseur : comme son nom l'indique, sert à compresser la vapeur. Dans le cas d'une compression adiabatique, cette augmentation de pression s'accompagne d'une augmentation de température.
- Le condenseur : un échangeur de chaleur qui sert à prélever de la chaleur au niveau de la vapeur surchauffée qui sera transmise au milieu à chauffer (ballon de stockage d'eau chaude). Le fluide toujours comprimé redevient liquide en se condensant.
- Le détendeur : sert à diminuer la pression du fluide caloporteur. La température de ce dernier s'abaisse fortement le rendant prêt pour une nouvelle absorption de chaleur.

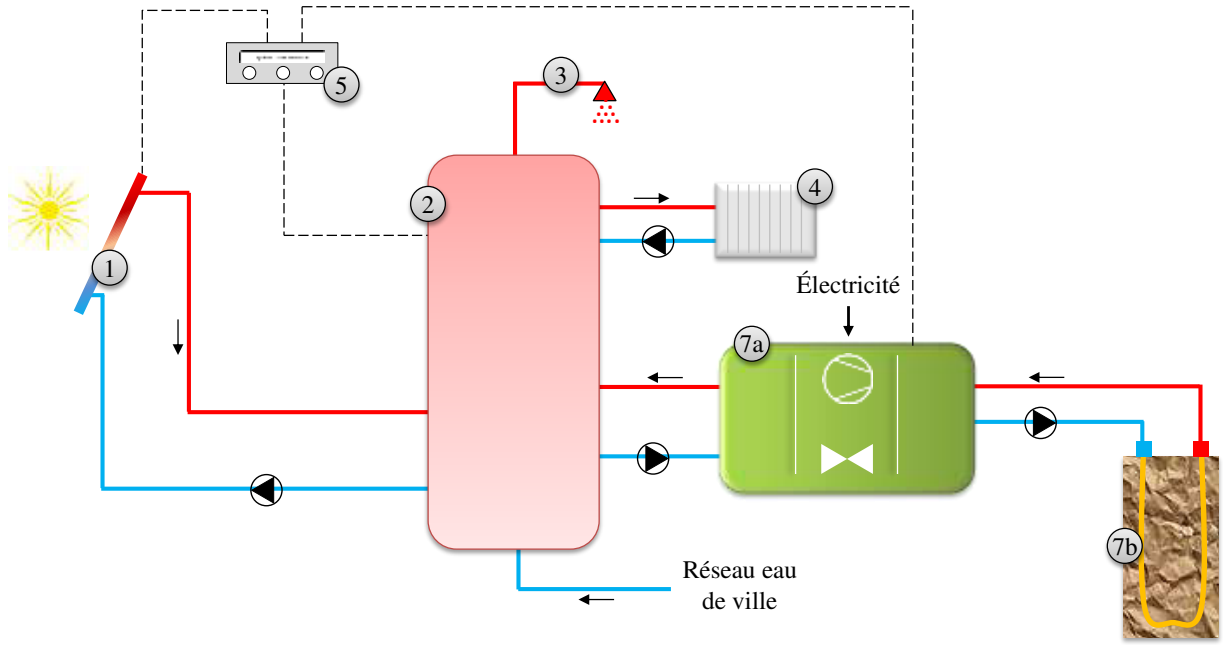

Figure 1- 12 : Schéma de principe Système solaire combiné avec un appoint par une pompe à chaleur à échangeur géothermique: chauffage et ECS. (7a) PAC, (7b) échangeur géothermique

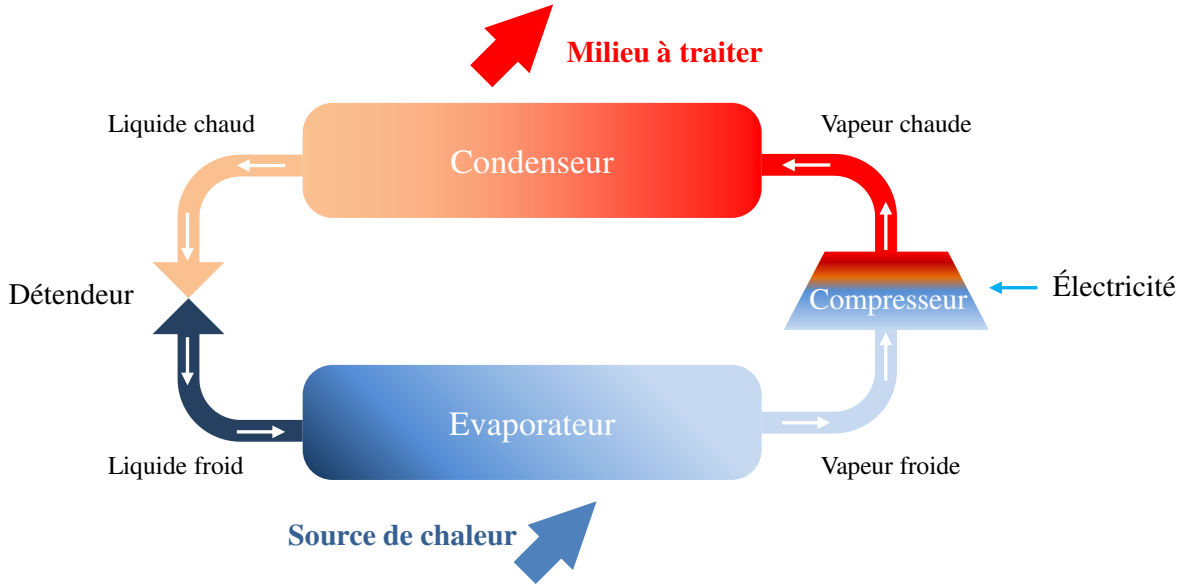

Figure 1- 13 : Schéma de principe d'une pompe à chaleur

Le couplage d'un système solaire avec une PAC est une solution innovante, très attirante pour la réduction de l'utilisation des ressources non renouvelables pour les besoins d'ECS et de chauffage. La consommation du système combiné à une PAC se réduit à la consommation du compresseur en électricité (ou du gaz pour certains types de PAC). Une pompe à chaleur performante et correctement dimensionnée consomme moins d'énergie primaire et émet moins de CO_2 qu'une chaudière à énergie non renouvelable comme l'ont montré certains cas d'études [16], [17], [18], [19]. Une comparaison entre le cycle de vie d'une PAC et d'une chaudière à gaz, ainsi qu'entre leurs impacts sur l'environnement a été réalisée dans [20]. Les résultats du cas d'étude montrent, selon un critère de jugement combinant à la fois les effets sur la santé, l'environnement extérieur et la consommation du système, que la PAC est plus

respectueuse de l'environnement. Malgré le coût d'investissement de départ qui est plus élevé pour les PAC, le temps nécessaire pour la rentabilité du capital investi est de l'ordre de 4 ans [18].

Les premiers développements de SSC à appoint par PAC remontent aux années 70 après la crise pétrolière. Ils resteront toutefois très peu commercialisés. L'engouement explicite à ce type de systèmes date seulement de la dernière décennie [21]. En effet, la majorité des travaux de recherches consacrés à ces systèmes ont été publiés pendant la période de 2009-2014.

Parmi les approches de classification de ce type de systèmes, celle développée dans le cadre de la Tâche 44 de l'AIE est la plus globale [22]. Elle permet de classifier et représenter les systèmes en fonction des sources d'énergie et de la configuration du système sous forme d'un schéma avec des notations génériques.

Les SSC à appoint par PAC se déclinent en trois grandes catégories (Figure 1- 14 et Figure 1- 15). La première regroupe les systèmes dont le capteur solaire ainsi que la PAC alimentent en parallèle le ballon de stockage en énergie thermique. Dans cette configuration le capteur solaire n'a pas d'influence sur le fonctionnement de la PAC. La deuxième catégorie regroupe les systèmes pour lesquels l'évaporateur de la PAC est relié directement en série au capteur solaire ou indirectement via un stockage froid (exemple sur la Figure 1- 16). Ce type de systèmes permet la valorisation du rayonnement solaire capté par les capteurs solaires fonctionnant à basse température. La troisième regroupe les systèmes hybrides qui peuvent fonctionner en mode parallèle ou en série.

Les capteurs solaires plans vitrés et à tube sous vide sont plus souvent utilisés pour les systèmes en parallèle tandis que les capteurs non vitrés sont utilisés uniquement pour les systèmes hybrides et en série.

La conception, le contrôle ainsi que l'installation des systèmes en parallèle sont les moins compliqués. Cette constatation explique leur dominance du marché. En effet, 61% des systèmes commercialisés adoptent un concept en parallèle, 6% en série et 33% hybride [21]. Dans chacune des catégories, plusieurs possibilités de gestion des flux d'énergie, entre les sous-systèmes et les charges, sont possibles comme pour le SSC classique (paragraphe précédent). La PAC peut également, dans chaque catégorie, avoir comme source de chaleur la géothermie ou l'air. A titre d'exemple, les résultats de l'enquête présentés dans [23] montrent qu'environ 48% des systèmes en parallèle sont équipés d'une PAC air-eau et 40% d'une PAC eau glycolée-eau.

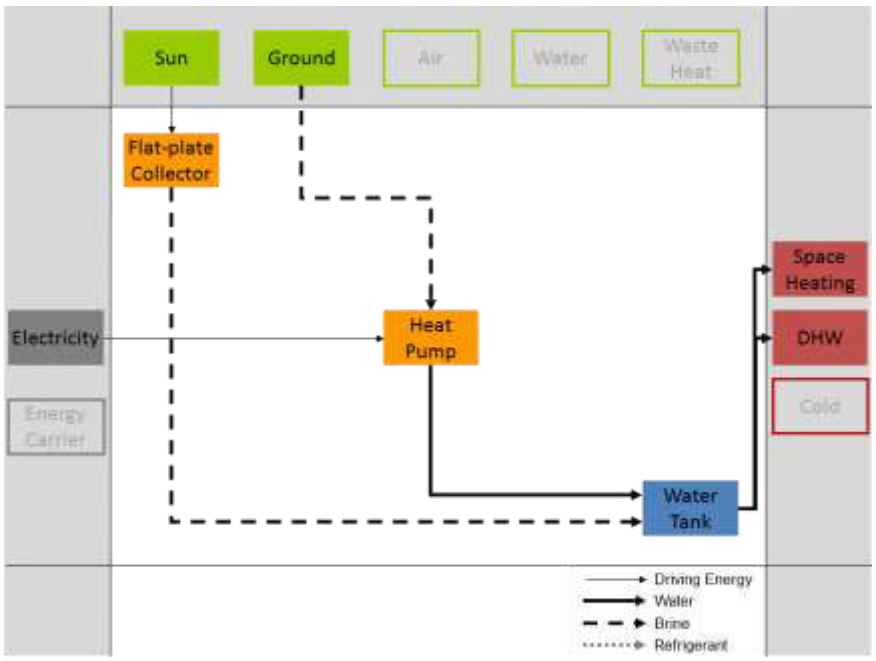

Figure 1- 14 : Représentation, en « *square view* », d'un système solaire combiné à une PAC en parallèle, Tâche 44 de l'AIE

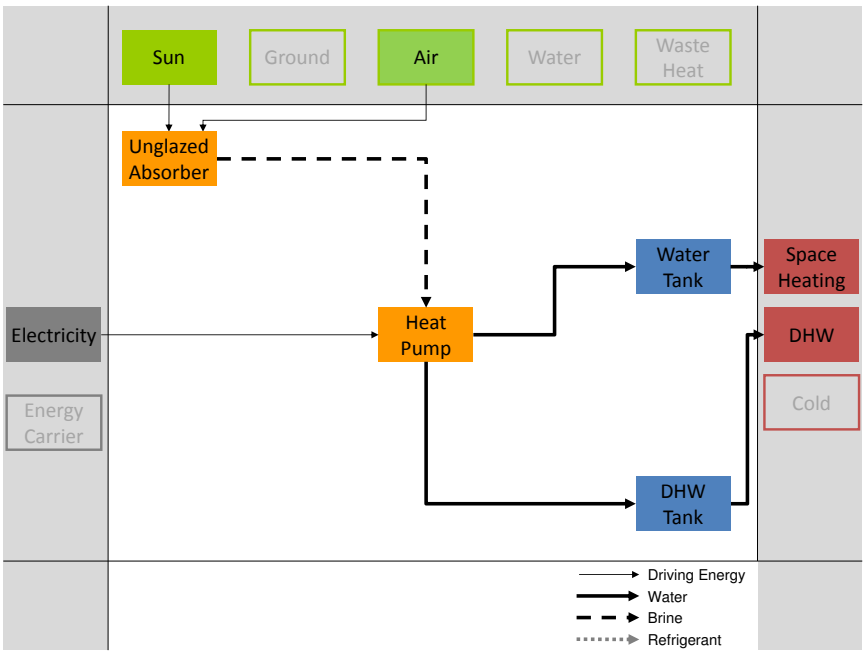

Figure 1- 15 : Représentation, en « *square view* », d'un système solaire combiné à une PAC en série, Tâche 44 de l'AIE

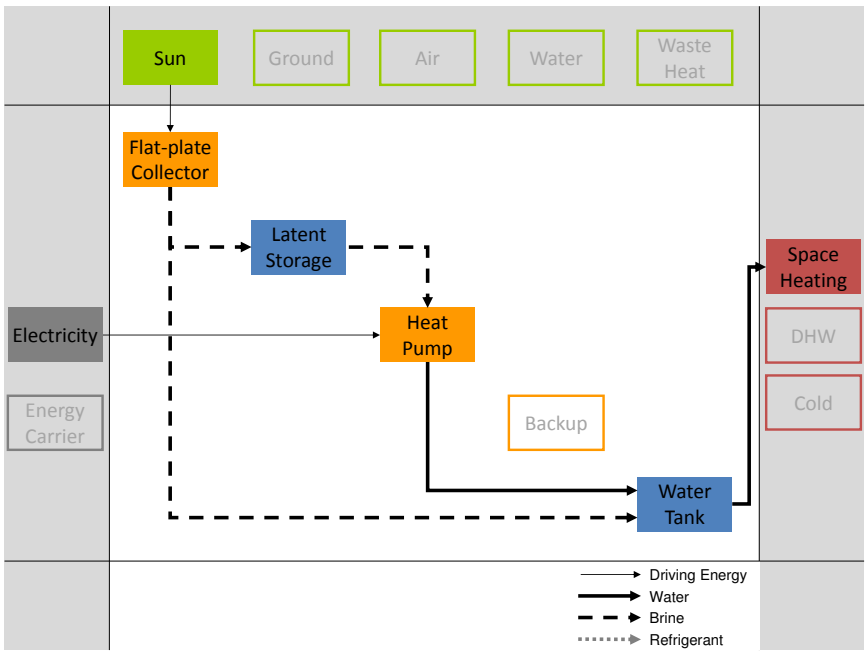

Figure 1- 16 : Représentation, en « *square view* », d'un système solaire combiné à une PAC pouvant fonctionner en parallèle ou en série (via un stockage froid), Tâche 44 de l'AIE

Il existe peu de données de suivi *in situ* ou d'études expérimentales pour les différents concepts de systèmes solaires avec un appoint par PAC. Les études d'évaluation de leurs performances ainsi que la comparaison entre les différentes technologies sont dans la majorité des cas basées sur des simulations numériques. Dans [24] les auteurs comparent les résultats de simulations entre deux systèmes parallèles (deux sources de chaleur : géothermie et air) et un système série avec un grand stockage froid. Les simulations ont été réalisées dans les mêmes conditions c.-à-d. même type de bâtiment, même surface de capteurs solaires et même climat. Les résultats montrent que les systèmes parallèles permettent de significatives économies d'électricité en comparaison avec une solution de référence, c.-à-d. une PAC seule sans ressources solaires. Cette économie relative d'énergie est plus grande dans le cas de la PAC eau glycolée-eau, cependant la consommation en électricité absolue dans le cas d'une PAC air-eau est plus faible. Les auteurs montrent également que le système en série peut atteindre les performances d'un système parallèle avec une PAC géothermique à condition d'augmenter la surface des capteurs solaires. Dans ce cas il est possible de diminuer le volume du stockage froid qui est très coûteux. L'étude [25] similaire à la précédente mais plus générale, intégrant plus de paramètres dans les simulations tels que le type du capteur solaire, la stratégie du système de régulation et le climat, permet de lister les conclusions suivantes :

- L'intérêt de coupler un système solaire à une PAC en parallèle est d'augmenter le coefficient de performance saisonnier SPF (Tableau A- 2) du système et donc l'économie d'énergie. L'économie d'électricité dépend beaucoup de la surface et du type des capteurs, de la source de chaleur utilisée et de la charge (climat et bâtiment).

- L'utilisation d'une vanne thermostatique au niveau des radiateurs permet d'augmenter l'économie d'énergie. En effet, la vanne thermostatique régule le débit à l'entrée des radiateurs en le limitant au débit nécessaire pour répondre aux besoins. Le SPF est cependant légèrement meilleur sans la vanne thermostatique car, dans ce cas de

figure, la puissance fournie pour le chauffage est surestimée notamment pendant les journées avec de forts apports solaires.

- Recommandations à propos de l'intégration du système de stockage : i) le thermocouple contrôlant la production de l'ECS à partir de l'appoint doit être situé, dans le ballon, sur une distance suffisamment éloignée de la zone de chauffage (30cm pour un ballon de stockage de 80cm de diamètre), ii) En mode production d'ECS le retour du ballon vers la PAC doit être tiré au-dessus de la zone de chauffage, iii) contourner le stockage lorsque la PAC fonctionne en mode chauffage, vi) le chargement de la partie supérieure du ballon pour l'ECS ne doit pas dépasser quelques heures et finalement v) la qualité de stratification du système de stockage doit être prise en compte avec différents débits pendant l'installation.

- L'utilisation d'un capteur solaire comme source de chaleur de la PAC pouvant fonctionner avec deux sources de chaleur n'améliore les performances du système que si la relation suivante est vérifiée [26]:

$$\frac{\Delta COP_{pac}}{(COP_{pac,par}-1)} \times \frac{\Delta\eta_{capt}}{\eta_{capt,par}} > 1 \qquad \textbf{(1- 1)}$$

  ΔCOP_{pac} et $\Delta\eta_{capt}$ sont respectivement la différence entre le COP de la PAC et le rendement du capteur des systèmes parallèle et en série. $COP_{pac,par}$ est le COP de la PAC dans le système en parallèle. $\eta_{capt,par}$ est le rendement du capteur dans le cas du système en parallèle.
  Puisque cette relation est difficilement réalisable les systèmes en parallèle sont à privilégier [26].

  - Peu de retour d'expérience ou d'études sont disponibles pour les systèmes exploitant l'énergie solaire comme unique source de chaleur. Cependant, un tel système peut être, en théorie, pertinent dans deux cas : i) pour éviter un forage coûteux ou ii) si la température ambiante est très basse, ce qui veut dire qu'une PAC air-eau n'est pas bien adaptée dans ce cas de figure, l'utilisation d'un stockage froid (de la glace) peut former la source de chaleur de la PAC. Un appoint électrique doit néanmoins être prévu dans ce cas.

  - La régénération (Figure 1- 17) : l'utilisation d'une partie de l'énergie solaire pour régénérer de la chaleur dans les forages est pertinente uniquement dans les cas où ces forages ont été sous-dimensionnés (systèmes déjà en place).

Parmi les quelques études expérimentales qui se sont intéressées aux systèmes solaires couplés aux PAC figurent les travaux suivants :

- Dans l'article [27] S. Stark et al ont exploité des mesures sur terrain d'une durée d'un peu plus d'une année pour analyser les performances de 3 systèmes parallèles. Les trois systèmes se différencient entre eux par le type de l'appoint supplémentaire (chaudière électrique ou gaz) et les caractéristiques des sous-systèmes (surface du capteur, capacité des PAC air-eau etc.). Les conclusions de l'étude réalisée montrent que le couplage d'un système solaire à une PAC air-eau demande encore beaucoup

d'amélioration notamment dans son installation et la gestion intelligente de son système de contrôle pour maximiser les gains solaires, les performances de la PAC mais aussi celles du système global. Une attention très particulière a été donnée à la quantité d'énergie utilisée pour le dégivrage. Les résultats montrent qu'il est essentiel de réduire la consommation d'énergie à cet effet dans le but d'augmenter les performances énergétiques du système. Des propositions pour la réduction de l'énergie utilisée pour le dégivrage ont été également proposées (utilisation de l'énergie solaire par exemple).

- Dans l'article [28] V. Trillat-Berdal et al se sont intéressés à un système hybride avec régénération de la chaleur géothermique grâce au capteur solaire. Après 11 mois de fonctionnement, les résultats montrent que cette configuration est capable de garantir le confort du bâtiment individuel avec une fraction solaire de 68% (part d'énergie solaire gratuite dans la consommation d'énergie globale) et des coefficients COP de la PAC allant de 2,6 à 3,75 (en fonction de la période). La régénération de la chaleur géothermique était déterminante pour garder un COP à des niveaux élevés car elle permet de réduire la fatigue du sol. Les auteurs montrent également que le contrôle des pompes de circulation affecte énormément les performances du système. Il est donc important de s'assurer que les pompes fonctionnent uniquement quand la PAC est en marche.

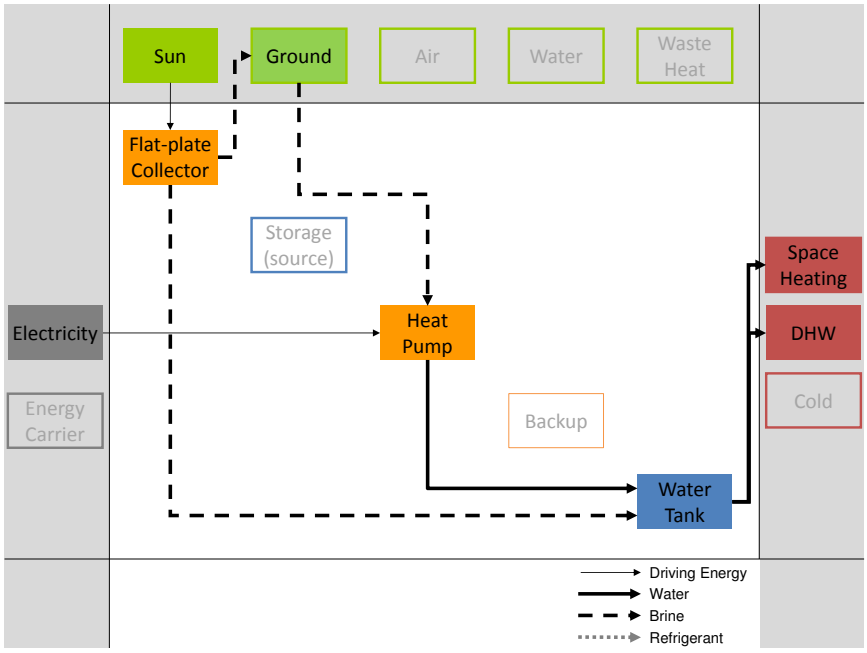

Figure 1- 17 : Représentation, en « *square* view », d'un système solaire combiné à une PAC en parallèle avec une possibilité de régénération de la chaleur du sol, Tâche 44 de l'AIE

I.1.4.3 Le système solaire multifonction

Outre les deux premiers systèmes thermiques à énergies renouvelables décrits ci-dessus, le système solaire couplé à une machine à absorption commence à attirer l'attention de la communauté scientifique et industrielle. L'intégration d'une machine à absorption, Figure 1-18, dans la configuration du système permet de répondre non seulement aux besoins en ECS et en chauffage mais également en climatisation. L'utilisation d'une machine à absorption au lieu d'une PAC à compresseur est plus adaptée car les besoins pour la climatisation sont en

phase avec la disponibilité des ressources solaires indispensables pour le fonctionnement de la machine (pour la partie générateur de la machine).

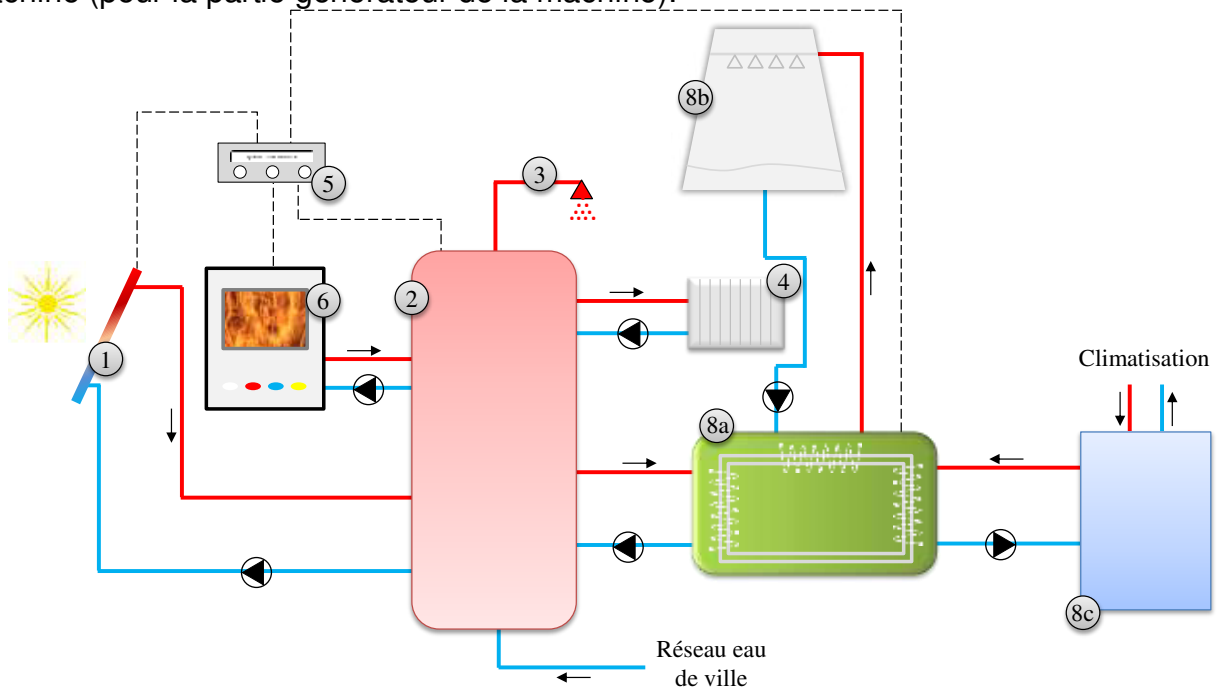

Figure 1- 18 : Schéma de principe d'un système solaire combiné couplé à une machine à absorption : chauffage, ECS et climatisation. (8a) machine à absorption, (8b) tour de refroidissement, (8c) ballon de stockage froid

Le fonctionnement d'une machine à absorption ressemble à celui de la PAC classique pour laquelle la compression mécanique de la vapeur est remplacée par une compression thermique grâce à un couple absorbeur/absorbant. Le fonctionnement des deux éléments caractérisant le cycle de ce type de machine (Figure 1- 19) peut être décrit ainsi :

- L'absorbeur, comme son nom l'indique, absorbe la vapeur d'eau créée dans l'évaporateur dans le but d'y maintenir la basse pression nécessaire à la vaporisation du réfrigérant ;
- Pour éviter la saturation de l'absorbeur, la solution est régénérée dans le concentrateur. Elle est réchauffée par une source de chaleur (un capteur solaire par exemple) pour évaporer l'eau. La solution régénérée retourne à l'absorbeur.

Les systèmes solaires intégrant une machine à absorption sont encore dans un état de développement beaucoup moins mature que les deux systèmes présentés ci-dessus. Jusqu'à présent il n'existe aucune approche de classification ou de recensement des systèmes commercialisés. La majorité des études sur le sujet des systèmes solaires intégrant une machine à absorption s'intéresse uniquement à la fonction de climatisation. En conséquence, les conclusions présentées dans le paragraphe suivant sont tirées de recherches bibliographiques sur des systèmes intégrant une machine à absorption, incluant ou non la fonction chauffage.

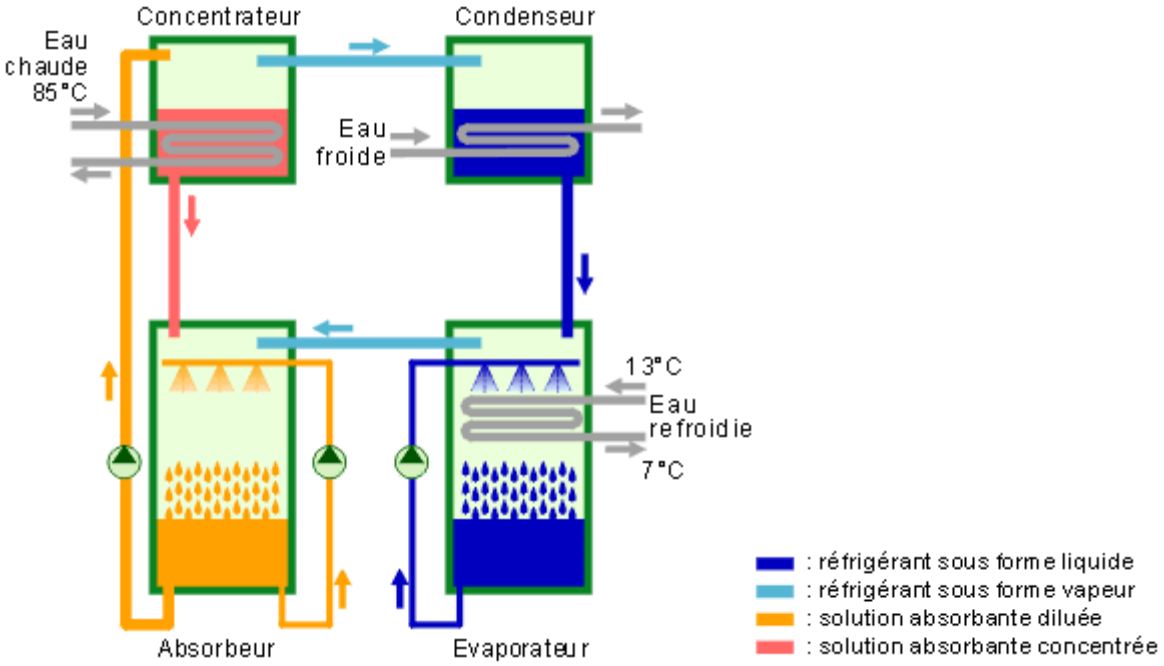

Figure 1- 19 : Schéma de principe Machine à absorption. Source : Energie Plus.

La Figure 1- 20 ([29]) présente les combinaisons possibles entre un système solaire et les différents types de machines à sorption (machines fonctionnant sans compresseur mécanique). La figure montre que les machines à absorption sont plutôt adaptées aux capteurs à tubes sous vides ou aux capteurs plans vitrés. Les températures de fonctionnement se trouvent dans un intervalle allant de 75°C à 110°C pour une machine simple effet. Les capteurs solaires cités ont un rendement supérieur à 50% pour ces niveaux de températures.

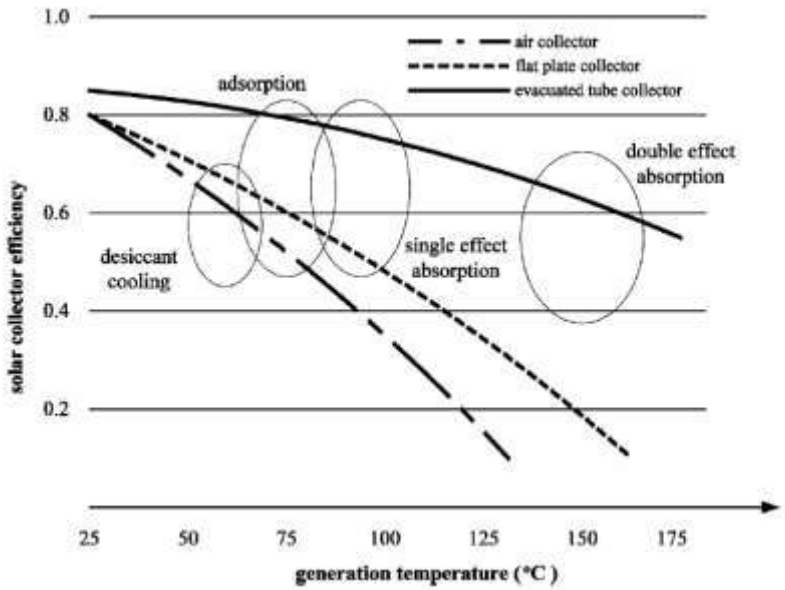

Figure 1- 20 : Différentes combinaisons possibles de couplage entre un système solaire et une machine à sorption [29]

Dans [30], les auteurs présentent une étude expérimentale d'une machine à absorption simple effet, avec une source solaire comme générateur. Les résultats montrent que les performances de la machine dépendent de ses conditions de fonctionnement et qu'il existe pour chaque condition un COP optimal. Les auteurs trouvent que la machine atteint ses performances maximales (COP = 0,7) à basse température au niveau du condenseur et de l'absorbeur (32°C), et ce pour une température du générateur au moins égale à 68°C. Une étude plus récente [31] et plus générale a permis la proposition des Figure 1- 21 et Figure 1- 22, qui donnent la performance des machines en fonction des températures du condenseur et du générateur. La maîtrise de ces deux températures étant impossible à cause de leur variabilité, il s'avère donc nécessaire d'intégrer au système une régulation efficace qui doit permettre le fonctionnement dans les conditions optimales.

Dans [32] et [33], les auteurs proposent de remplacer la tour de refroidissement, généralement utilisée comme condenseur des machines à absorption, par un stockage latent avec matériau à changement de phase. L'étude expérimentale réalisée montre la faisabilité d'un tel système, permettant ainsi d'éviter les inconvénients d'utilisation d'une tour de refroidissement : consommation d'eau, formation de brouillard, prolifération des bactéries, etc…

Une étude expérimentale sur une machine à absorption couplée à un capteur solaire à tube sous vide pour le refroidissement direct de l'air a été présentée dans [34]. Les résultats montrent le potentiel d'un tel système pour répondre aux besoins de climatisation des bâtiments résidentiels.

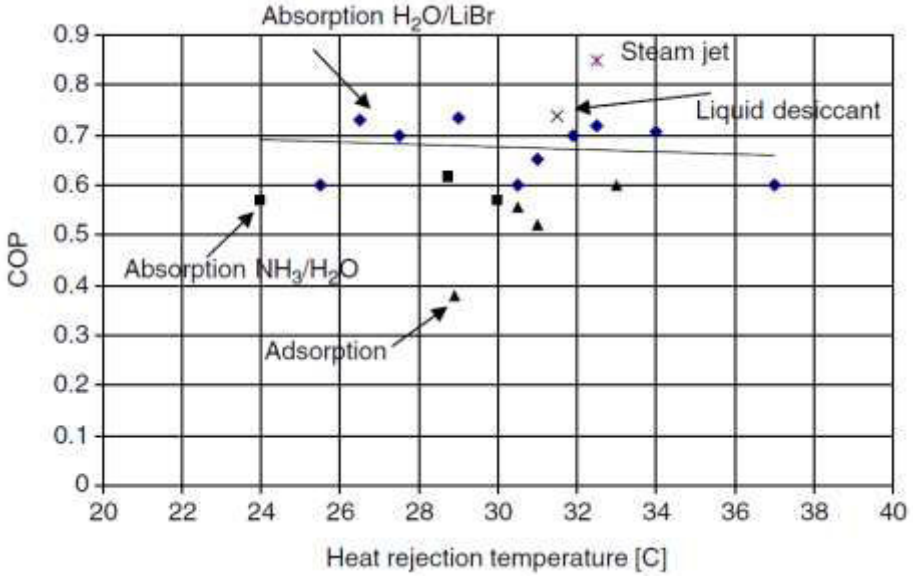

Figure 1- 21 : Le COP en fonction de la température du condenseur [31]

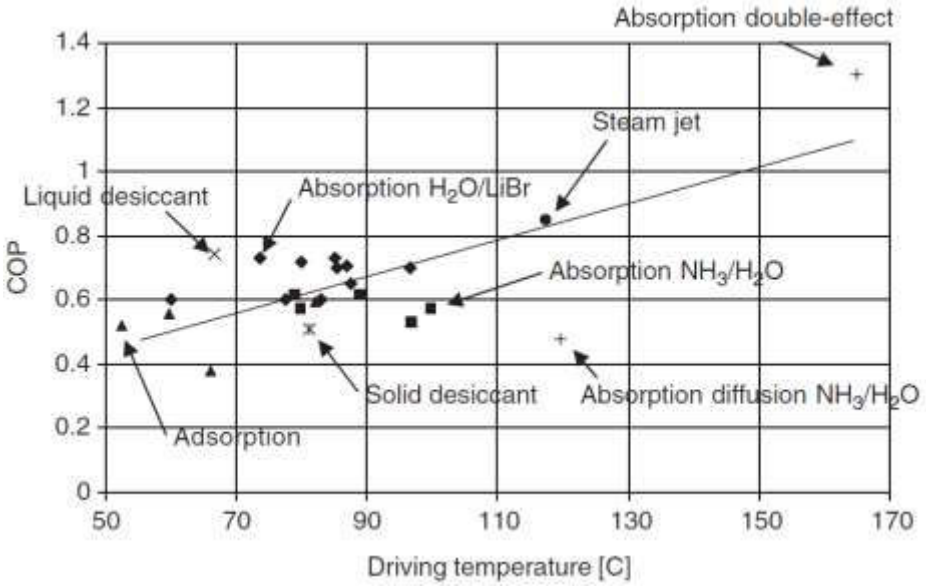

Figure 1- 22 : Le COP en fonction de la température du générateur/concentrateur [31]

Selon [35], le couplage du solaire avec une machine à absorption permet la réduction de la consommation du système d'appoint et donc les émissions de CO_2.

Certaines études se sont intéressées non seulement à l'aspect technologique mais aussi à l'aspect financier du système. Ainsi [36] et [37] préconisent l'utilisation des systèmes solaires uniquement :

i) dans le cas d'un besoin conséquent pour le chauffage, l'ECS et la climatisation (cas des hôtels) ;
ii) dans les climats présentant à la fois des besoins de chauffage et de climatisation (pour la réduction de la taille du stockage) ;
iii) dans les pays où les prix des énergies fossiles sont élevés ;
iv) Les auteurs proposent également, pour les besoins de climatisation, de coupler une machine à absorption et une machine à compression dans le même système. La première pour fonctionner en cas de faibles sollicitations et la deuxième pour le cas des pics de demande de climatisation. De tel système permet, selon les auteurs, d'avoir un gain conséquent en termes d'économie d'énergie.

Quelques exemples d'installations réussies de systèmes thermiques à énergies renouvelables sont présentés dans [38]. Il est possible de constater que l'économie d'énergie est plus importante pour les cas présentant une bonne adéquation entre la disponibilité de la ressource solaire et les besoins.

I.2. Les méthodes d'évaluation des performances énergétiques des SSC

Après avoir présenté le contexte énergétique actuel et illustré la place du secteur du bâtiment dans ce dernier ainsi que le rôle théorique que peuvent jouer les systèmes solaires thermiques innovants, il est important de présenter une analyse économique des SSC ainsi

que l'état de l'art des différentes approches pour l'évaluation de leurs performances énergétiques. Ceci permettra de mieux illustrer l'intérêt de développer une nouvelle méthodologie d'évaluation des SSC.

I.2.1. Contexte environnemental et économique des SSC et PAC

Selon le contexte énergétique considéré ci-dessus, il apparaît clairement que le secteur résidentiel et tertiaire présente un potentiel important d'économie d'énergie et de réduction des émissions de GES. Les systèmes thermiques innovants, moyennant des ressources renouvelables, très abondantes, peuvent jouer un rôle déterminant à ce propos. Selon l'association ENERPLAN, pour que la France respecte ses engagements, notamment vis-à-vis des réductions des émissions de GES, 4 millions de logements devront être équipés de systèmes solaires thermiques. A titre d'exemple, selon l'ADEME, un SSC bien conçu et correctement maintenu permet d'économiser de l'ordre de 350 kWh/an par m² de capteur, soit sensiblement plus que les valeurs obtenues pour un chauffe-eau solaire individuel.

Malgré l'intérêt évident de ces systèmes, leur marché européen peine à retrouver son état de croissance d'avant 2008 (Figure 1- 23 [39]). En effet, selon les résultats de l'enquête menée par l'EurObserv'ER, le marché a subi une nouvelle baisse en 2012, la quatrième consécutive depuis 2009.

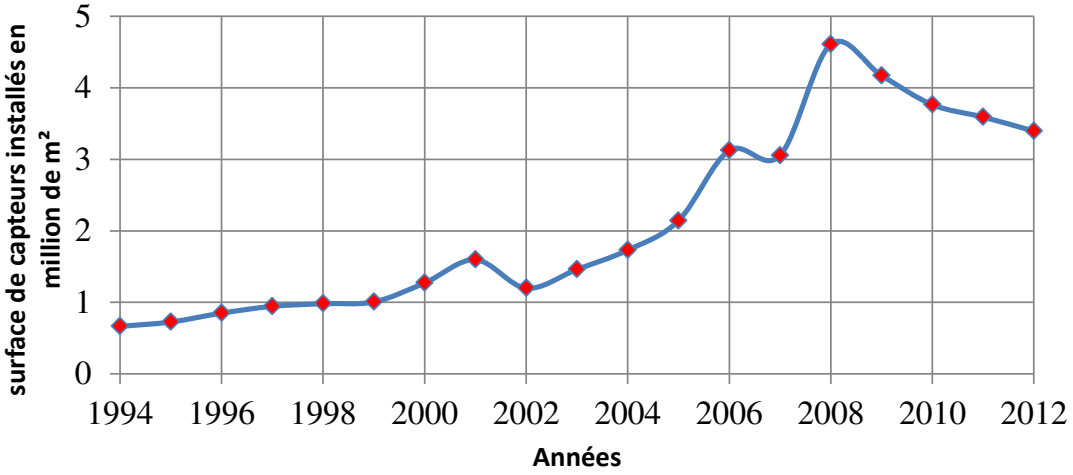

Figure 1- 23 : Evolution annuelle des surfaces solaires thermiques installées dans l'UE depuis 1994 (en m²). Source : Eur'Observ'ER

Avec 1350 SSC installés en 2012, le marché du solaire thermique subit en France également un net recul. Au total, 111700 m² de capteurs thermiques, tous systèmes confondus, ont été installés, soit une baisse de 12% par rapport à 2009 (chiffres clés ADEME, Figure 1- 24).

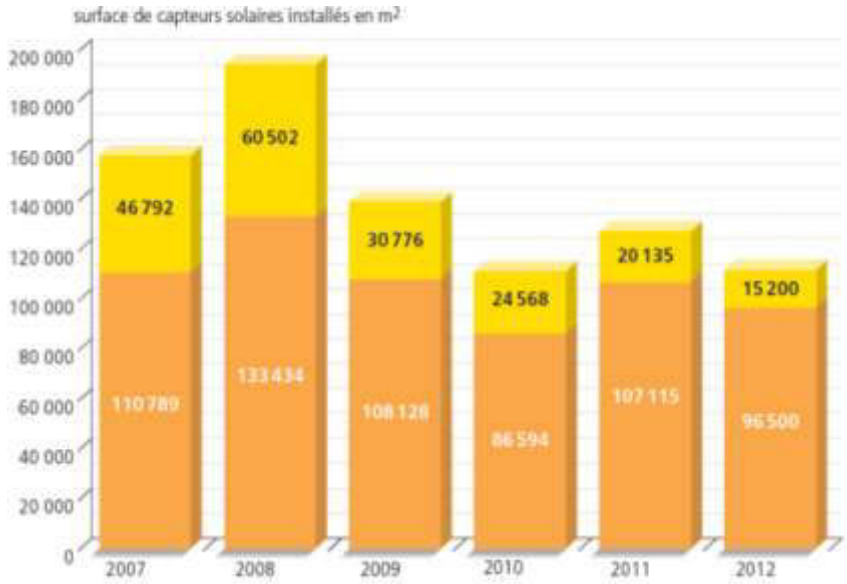

Figure 1- 24 : Evolution du marché des chauffe-eau solaires (en résidentiel individuel, SSC en jaune et CESI en orange). Source : chiffres clés ADEME, 2013

Selon les chiffres clés de l'ADEME de 2013, le nombre de pompes à chaleur vendues était, jusqu'à 1997, de 1500 unités par an seulement. Le marché s'est ensuite développé grâce à l'offre commerciale Vivrélec proposée par EDF. D'après la Figure 1- 25, en 2002, les ventes étaient de 12400 unités. Grâce à la mise en place du crédit d'impôt en 2005, le nombre de PAC vendus a augmenté avant d'atteindre sa valeur maximale en 2008. La réduction du crédit d'impôt comme soutien de cette filière et sa suppression pour l'installation de certains types de PAC ont conduit à une chute des ventes en 2010 d'environ 48%.

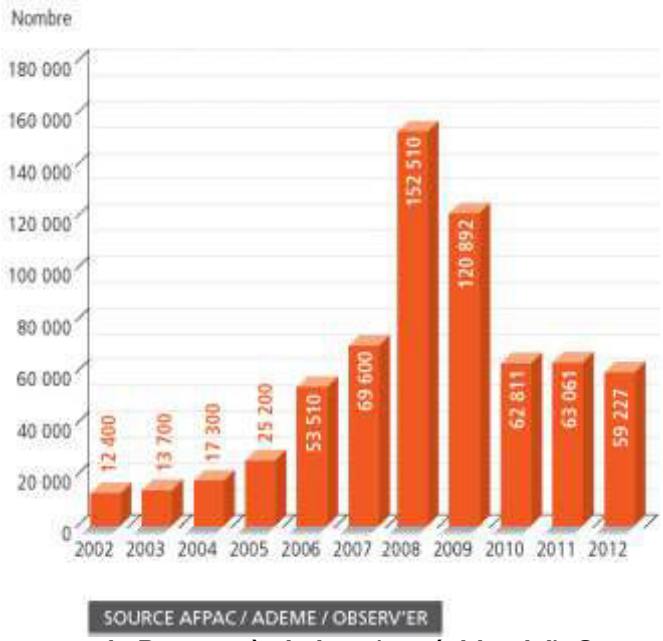

Figure 1- 25 : Evolution de ventes de Pompes à chaleur (en résidentiel). Sources : chiffres clés ADEME, 2013

Un examen du nombre d'entreprises rejoignant le marché des systèmes solaires avec un appoint par PAC (Figure 1- 26, [21]) révèle que le vrai intérêt pour la commercialisation de ce type de systèmes est récent. Le marché des systèmes solaires avec un appoint par PAC est donc encore dans un état embryonnaire.

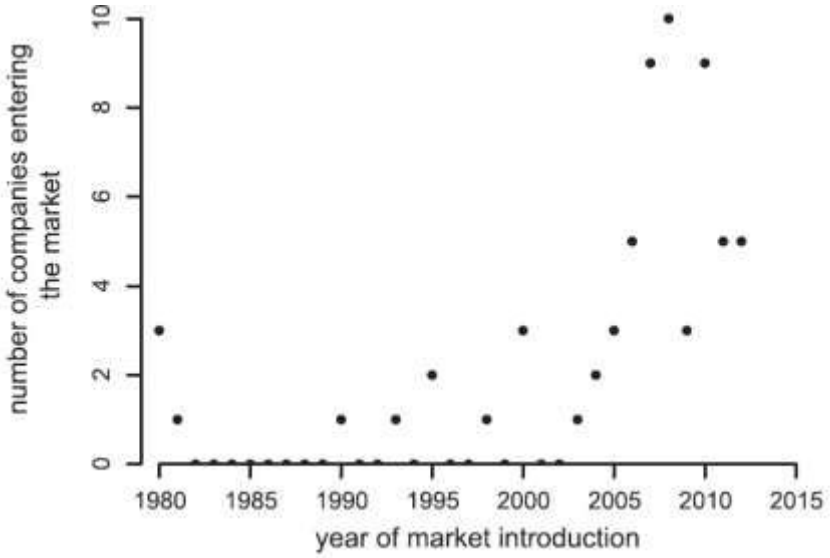

Figure 1- 26 : Nombre d'entreprises ayant pénétré le marché des systèmes solaires combiné au PAC

Selon [40] (Figure 1- 27), l'évolution des installations des systèmes de climatisation solaire est limitée, mais en constante croissance. En effet, le nombre de systèmes installés en 2010 est estimé à 600 systèmes et à 1000 en 2012. Les systèmes solaires considérés ici incluent des systèmes couplés à une PAC à compression alimentée par de l'énergie photovoltaïque (PV), ce qui montre que le nombre d'installations de systèmes solaires couplés à des machines à absorption sont en nombre très réduits.

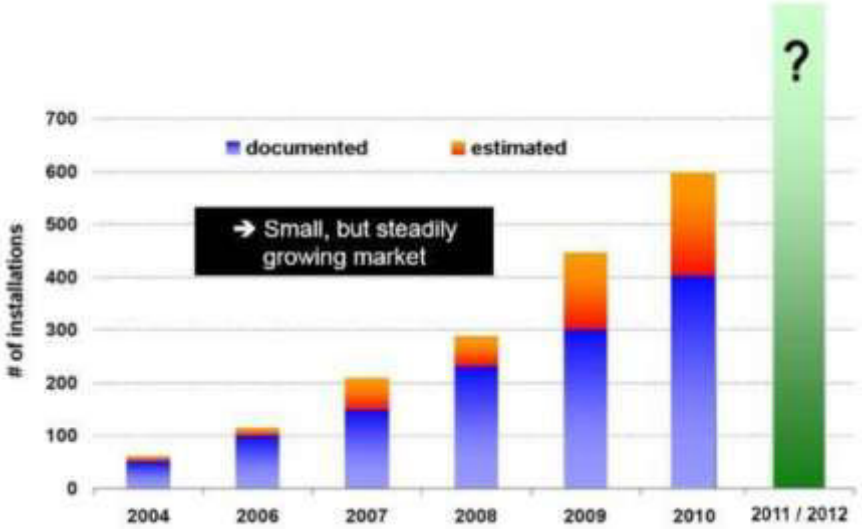

Figure 1- 27 : Estimation du nombre d'installations de climatisation solaire dans le monde

Selon les éléments présentés dans ce paragraphe, il est évident que le marché des systèmes thermiques innovants pour le bâtiment est fortement impacté par le contexte économique mondial et les politiques économiques et environnementales nationales. Une telle

situation n'est guère en faveur ni des objectifs européens (à partir de 2020 tous les bâtiments doivent puiser leur énergie uniquement à partir de ressources renouvelables), ni des objectifs français concernant le contexte énergétique et environnemental expliqués ci-dessus. Il va donc falloir déterminer les obstacles responsables de cette situation pour ensuite proposer une solution capable de remettre en avant ce type de systèmes.

I.2.2. Systèmes et performances énergétiques

Des études et investigations sur les installations thermiques à énergies renouvelables existantes révèlent que beaucoup d'effort d'optimisation, d'amélioration de leur conception et de leur installation sont nécessaires. En effet, il a été remarqué sur le terrain que beaucoup d'installations sont de mauvaise qualité (mauvais dimensionnement, inadéquation à l'usage réel, mauvaise mise en œuvre, entretien/maintenance défaillant etc.), ce qui fait que le système ne permet pas de réduire effectivement la consommation de l'énergie comme souhaité. Une mauvaise maîtrise de l'intégration des systèmes hydrauliques, de la régulation mais surtout des connexions entre le stockage et la PAC peuvent augmenter la consommation électrique de la PAC de 45% ([41], [42]).

Dans le but de faciliter l'installation, de réduire les risques de mauvais raccordements entre les sous-systèmes et de limiter la taille de l'unité complète, un nombre grandissant d'équipements sont préfabriqués et assemblés en usine (des exemples sont donnés dans la Figure 1- 28).

Figure 1- 28 : Exemples des solutions thermiques à énergie renouvelable commercialisables pour le bâtiment (de gauche à droite : Viessmann, De Dietrich, Sonnenkraft et Clipsol)

Le fonctionnement des systèmes dépend beaucoup, comme il a été présenté dans le paragraphe I.1.4. , des conditions de l'environnement dans lequel ils sont installés. Un système installé dans un bâtiment à basse consommation dans un climat avec de faibles ressources renouvelables n'aura pas les mêmes performances que s'il était installé dans un bâtiment à forts besoins dans un climat avec beaucoup de ressources renouvelables. Le comportement des systèmes est donc non-linéaire et très variable en fonction des conditions aux limites (climat, bâtiment etc.). Evaluer, *a priori*, les performances des systèmes avant leur installation est donc indispensable pour définir la meilleure solution technique associée à chaque application (bâtiment résidentiel, tertiaire, etc.), usages (chauffage, ECS, etc.), et conditions environnementales).

Certes, la crise économique, qui n'encourage pas les investisseurs, ainsi que la baisse du dynamisme dans le secteur du bâtiment, sont en grande partie responsables de la régression du marché de ces systèmes thermiques auxquels ce travail s'intéresse. Néanmoins, il existe un deuxième facteur dont le poids est du même ordre de grandeur que le premier. En effet,

des experts en la matière pointent du doigt l'absence de garantie de performances de la part des installateurs, et/ou de suivi des performances une fois que l'installation est réalisée. En outre, il n'y a pas de méthode satisfaisante pour tester au préalable le matériel. Dans ce contexte, l'existence de contre-performances et de systèmes mal installés est un véritable frein au développement commercial des installations solaires thermiques [39].

I.2.3. Critères de performances

La caractérisation des performances énergétiques des systèmes thermiques pour le bâtiment est le processus qui en résulte, la distinction de deux ou plusieurs systèmes assurant la même fonction selon un critère énergétique défini au préalable. La multitude des sources d'énergies, des configurations possibles et des composants des systèmes thermiques innovants pour le bâtiment complique la définition de normes et directives génériques pour l'évaluation des performances. Beaucoup de recherches ont été réalisées à ce propos, d'autres sont en cours, dans des centres européens. Certaines de ces études se sont intéressées à la détermination des critères ou indicateurs pertinents de performances énergétiques. Ces indicateurs prennent souvent la forme d'un rapport entre l'énergie fournie pour satisfaire des besoins (chauffage, ECS, etc.) et l'énergie dépensée par le système pour répondre à ces besoins. Les critères de performances ont été initialement développés pour les composants. Un bon nombre de critères de performances pour les systèmes solaires et les PAC est donné en annexe (Tableau A- 1 et Tableau A- 2). Pour évaluer les performances des systèmes complets, certains de ces critères peuvent être utilisés sous réserve de définir les conditions aux limites adéquates. En effet, dans le cadre des deux projets, QAiST [43], à l'échelle européenne et, MeGaPICS [44], à l'échelle nationale, les critères les plus pertinents ont été sélectionnés et appliqués aux systèmes ayant comme fonctions le chauffage et la climatisation. Cependant, il reste encore difficile de comparer des systèmes ayant des sources d'énergie et des composants différents entre eux, en se basant uniquement sur un seul critère. Une analyse multicritère, intégrant des aspects sources d'énergie mais également environnementaux, est nécessaire pour une meilleure qualification du système considéré.

I.2.4. Méthodes de caractérisation des performances

I.2.4.1 Présentation des différentes méthodes

L'évaluation des performances énergétiques des systèmes thermiques selon les critères définis ci-dessus est toujours précédée d'un essai physique de celui-ci, suivant la norme correspondante, dans un banc expérimental au laboratoire. L'essai physique du système en banc d'essai réel ou semi-virtuel permet de calculer des points de fonctionnement en régime établi (cas des PAC) ou des données entrées/sorties (test dynamique) pour la modélisation selon les différentes approches existantes. En effet, ces données sont nécessaires pour identifier les paramètres des modèles. Dans le Tableau 1- 2 sont synthétisées les méthodes d'évaluation des performances des systèmes solaires, voire des systèmes thermiques à énergie renouvelables de façon générale, tandis que dans le Tableau 1- 3 sont représentés ceux dédiés aux PAC [45].

Les méthodes les plus abouties d'évaluation des performances énergétiques des PAC sont toutes basées sur la méthode des BIN ([45], [46]). Cette méthode permet de calculer le coefficient de performance saisonnier SPF (la définition est donnée dans le Tableau A- 2 en

annexe) moyennant la pondération de quelques points de fonctionnement, fournis par les tests selon une norme donnée, en fonction de la répartition dans le temps des demandes énergétiques du bâtiment. Les méthodes classiques présentées dans le Tableau 1- 3 présentent les limites suivantes :

- L'influence des cycles de la PAC n'est pas ou peu prise en compte. L'inertie du système est ainsi négligée ce qui génère des erreurs d'estimation de performances.
- Ces méthodes supposent que la température interne du bâtiment est fixe ce qui n'est pas le cas en réalité.
- Dans le cas des PAC à échangeur géothermique ces méthodes ne prennent pas en compte le caractère dynamique du système introduit par la forte inertie du sol. Pour pallier cette problématique il est nécessaire d'utiliser des méthodes de simulation dynamique.
- Ces méthodes ne sont pas adaptées à toutes les configurations possibles (ECS, chauffage, climatisation etc.).

Tableau 1- 2 : Aperçu des méthodes de test des systèmes thermiques

Approche :	« Composant »	« Globale »
Définition	Les composants formant le système complet sont testés séparément suivant une norme correspondante. Les paramètres caractéristiques de chacun des composants ainsi déterminés sont introduits dans un logiciel de simulation tel que TRNSYS pour déterminer les performances annuelles du système dans sa globalité.	Contrairement à l'approche 'composant' le système est testé dans sa globalité. Seules les données récupérables aux limites du système sont utilisées pour sa caractérisation. Les mesures aux bornes des composants ne sont pas possibles.
Exemples	• CTSS • Méthode des BIN	• CCT [38] • Combitest [39] • SCSPT • DST (Dynamic System Test) [40]
Avantages	• Possibilité de tester tous les systèmes avec des configurations différentes. • Possibilité d'étudier le dimensionnement des systèmes. • Possibilité de simuler le système dans des conditions différentes	• Les vraies interactions entre les composants sont prises en compte. • Le système de contrôle-commande est testé dans des situations réelles de fonctionnement. • Permet de tester les sous-systèmes dont le développement des modèles est très complexe
Inconvénients	• Longue et peut être coûteuse. • Limité aux composants dont le modèle est réalisable et simple. • Les algorithmes de régulation ne sont pas fournis par le constructeur ce qui est nécessaire pour simuler la gestion des flux énergétiques dans le système. • Les vraies interactions entre les composants ne sont pas prises en compte ; elles sont seulement simulées.	• Les résultats ne sont valables que pour l'environnement du test. • Le banc expérimental est complexe et coûteux • Les résultats sont valables seulement et uniquement pour le système testé. La modification d'un composant du système implique la réalisation d'un nouveau test.

Tableau 1- 3 : Aperçu des méthodes de test des PAC [47].

Méthodes	VDI 2067-6	HTA Lucerne	CEN/TC228/WG4	ASHRAE 116
Champ d'application	PAC eau/eau, eau glycolée/eau, air/eau PAC à compression électrique ou thermique	PAC eau glycolée/eau PAC à compression électrique	PAC air/eau PAC à compression électrique	PAC air/air PAC à compression Electrique
Points de fonctionnement (Bins)	Quatre points recommandés	Trois points standards A-7, A2, A7, 80	Mille points	Selon région climatique Dix-huit points max
Type de fonctionnement	Bivalent	Monovalent	Bivalent	Mono-énergétique
Prise en compte de la production d'ECS	Energie annuelle de la production d'ECS	Energie journalière de la production d'ECS	Non pris en compte	Non pris en compte
Température de source	Air/eau : données météo Eu glycolée/eau : Diagramme de distribution	Air/eau : données météo Eu glycolée/eau : Température constante	Données météo	Données météo
Température de chauffe	Courbe de chauffe	Courbe de chauffe	Courbe de chauffe	
Calcul des besoins d'énergie	Chauffage : DJU, ECS : Energie annuelle transformée en température ambiante fictive plus élevée	Chauffage : DJU, ECS : Energie journalière	Chauffage : Puissance fonction linéaire du produit température ambiante x heures bins	Chauffage : Puissance fonction linéaire du produit température ambiante x heures bins et facteur de correction de 0,77
Evaluation du COP et de la puissance thermique	Données constructeurs ou valeurs par défaut	Test EN 255-2	Test EN 255-2	COP non évalué Evaluation des consommations électriques
Correction du COP et de la puissance thermique pour les conditions de fonctionnements	Interpolation des données des constructeurs	Correction de la température par efficacité exérgétique	Interpolation	Interpolation linéaire
Charge partielle	Non prise en compte	Non prise en compte	Correction avec un coefficient de dégradation dépendant du compresseur et du système de chauffage	Coefficient de dégradation cyclique évalué par la procédure ARI 210/240
Consommation des auxiliaires	Non prises en compte, sauf si intégrée dans le COP	Prises en compte dans le COP	Ratio constant dépendant de la puissance thermique	Non prises en compte
Appoint	Pondération en fonction de l'énergie délivrée au système	Non pris en compte	A partir de la consommation électrique	A partir de la consommation électrique

Une nouvelle méthode d'essai par émulation a été développée pour pallier ces verrous technologiques des PAC [47]. Similairement à la méthode SCSPT (paragraphe I.2.5.), la méthode développée consiste à tester la PAC dans un banc d'essai semi-virtuel où le climat, les puisages ECS, le bâtiment ainsi que le sol sont simulés. L'inconvénient majeur de cette méthode concerne la validité des résultats car ils sont limités aux conditions de test.

Les autres composants du système complet, à savoir le capteur solaire et le ballon de stockage, ainsi que le système de régulation, peuvent également être testés sur des bancs d'essais à la manière des PAC mais selon les normes adaptées.

Les données des essais des composants séparés peuvent être précieuses pour juger et comparer les performances entre composants mais en aucun cas évaluer les performances du système complet (ballon de stockage, régulation appoint etc.). En effet, comme il a été déjà mentionné précédemment, le comportement des composants dans le système varie beaucoup de celui du composant lorsqu'il est pris seul. Pour pallier cette difficulté, des méthodes de test utilisant l'approche « composant » ont été développées (Tableau 1- 2). Les inconvénients de cette approche de test, notamment la non prise en compte des vraies interactions entre les composants, oblige d'envisager le système dans sa globalité, en intégrant les fonctionnements à charge partielle, ainsi que les logiques de régulation dont le rôle est de plus en plus prépondérant dans la recherche de l'efficacité énergétique. D'où l'intérêt des méthodes adoptant une approche « globale ».

L'avantage de l'approche globale réside dans le fait que le système complet est installé sur le banc d'essai comme s'il l'était dans un vrai bâtiment, ce qui fait que les interactions entre composants sont réellement testées et non seulement simulées. Les tests sont réalisés en régime dynamique. Ils prennent donc également en compte le comportement dynamique du système et les logiques de la régulation. Les méthodes les plus abouties utilisant cette approche sont celles proposées par trois laboratoires européens : SCSPT (INES, CEA, France), Combitest (laboratoires SERC et SP, Suède) et CCT (laboratoire SPF, Suisse). La méthode DST (Tableau 1- 2) a été récemment appliquée à un système pour le chauffage de l'eau sanitaire. Il s'agit d'un système solaire avec un appoint par pompe à chaleur en parallèle [48]. Cette méthode semble être valable uniquement pour des systèmes thermiques moins complexes que ceux présentés dans la section I.1.4. L'amélioration proposée par les auteurs consiste à modifier la méthode DST en introduisant une phase de prise en compte des paramètres de la PAC.

D'autres méthodes ont été également abordées dans [49] et [46]. Dans ces deux articles, les méthodes utilisées se basent sur l'identification d'un modèle détaillé sous le logiciel TRNSYS. L'élaboration d'un modèle physique détaillé nécessite la connaissance d'un ensemble d'informations sur les composants et sur le système de régulation du système complet (rendement, coefficients de pertes thermiques, type d'échangeur, principe de régulation, etc.). Or, les industriels rendent très rarement de telles informations disponibles. Dans le deuxième article, les auteurs présentent quelques estimations de performances du système étudié dans différentes conditions, mais sans comparaison avec un cas de référence, ce qui ne permet pas de juger la fiabilité des prédictions.

Une étude comparative des trois méthodes CCT, SCSPT et Combitest peut être consultée dans [50]. L'étude de comparaison réalisée dans cet article révèle qu'elles peuvent donner des résultats très différents notamment à cause du pas de temps du fichier météorologique utilisé et selon la manière dont les conditions aux limites sont prises en compte. Un travail d'harmonisation des méthodes est donc nécessaire.

Selon les éléments cités dans ce paragraphe (prise en compte des vraies interactions, adaptée aux SSC compacts etc.), l'approche de test globale est plus pertinente et sera donc adoptée dans le cadre de ce travail. La méthode de tests SCSPT a été exploitée à plusieurs

reprises pour tester un large panel de systèmes à énergies renouvelables (Systèmes solaires combinés à des PAC dans le projet MACSHEEP [51], SSC classique projet COMBISOL [14], et autres projets bilatéraux). Elle est donc la plus adaptée pour développer une méthodologie générique.

I.2.4.2 Extrapolation des résultats

L'objectif des méthodes d'essais des SSC est évidement de fournir des données entrées-sorties pertinentes pour l'évaluation des performances du système testé à long terme (une année).

Les méthodes Combitest et SCSPT (décrite dans la section I.2.5.) utilise une extrapolation directe des résultats de l'essai en une année pour les mêmes conditions (climat et qualité du bâtiment). En effet, l'estimation de la consommation annuelle du système se fait par une simple multiplication par un facteur dépendant de la durée de l'essai : 365/6 pour Combitest car la durée de l'essai est de 6 jours et 365/12 pour SCSPT car la durée de l'essai est de 12 jours (plus de détail est donné dans le paragraphe I.2.5.2). La procédure d'extrapolation des résultats pour les méthodes CCT et DST est différente car basée sur des simulations annuelles (comme pour les méthodes de l'approche « composant ») du modèle identifié du système (généralement sous le logiciel TRNSYS).

L'extrapolation des résultats pour des conditions aux limites différentes de celles de la séquence de test, est possible avec la méthode CCT à l'aide du modèle identifié (il suffit de renseigner le fichier météo désiré etc.). En ce qui concerne la méthode SCSPT, une procédure d'extrapolation a été récemment développée à l'INES [52]. Elle consiste à exploiter les données mesurées pour identifier un modèle dynamique simplifié de l'ensemble du système. Le modèle simplifié, appelé modèle « boîte grise », combine des équations simplifiées physiques des composants du système et un réseau de neurones artificiels RNA (modèle « boîte noire »). Une fois que le modèle simplifié a été sélectionné et validé, il peut être utilisé pour simuler le fonctionnement du système avec des conditions aux limites différentes de celles du test (qualité du bâtiment et climat) afin d'obtenir des résultats annuels.

L'extrapolation à d'autres conditions aux limites en utilisant ces méthodes est basée, au moins en partie, sur des modèles physiques. Le développement de ces modèles n'est souvent pas facile. En effet, les composants internes (ballon de stockage, système de régulation ou tout autre système spécifique) peuvent être inconnus en particulier parce que les systèmes sont souvent compacts (préfabriqués en usine). Il est important de signaler que l'identification des paramètres physiques des modèles nécessite l'intégration de certains capteurs de mesure au sein du système durant son essai, or cela ne peut se faire sans dégrader le système d'autant plus lorsqu'il est compact. En outre, la fiabilité des résultats dépend en grande partie de la fiabilité et hypothèses utilisés pendant l'élaboration de chacun de ces modèles physiques.

Aussi, ces méthodes devraient être appliquées par un expert en raison de leur complexité, un problème supplémentaire quand ils sont utilisés comme un outil de certification.

L'extrapolation à d'autres conditions aux limites dans le cas de la méthode DST n'a pas encore été étudiée. Les résultats présentés, par exemple, dans [48] et [49] ne concernent que l'estimation des performances à long terme dans un environnement spécifique et pour les systèmes mono-fonction. Ces systèmes sont moins complexes que les systèmes multifonctions. Pour ces derniers le système de commande doit garder différents niveaux de

température dans le réservoir de stockage adapté à chacune des fonctions (ECS, chauffage ou climatisation).

Dans cette thèse une nouvelle méthodologie est proposée. L'idée est de proposer une méthodologie qui dépasse les limitations des méthodes existantes. Elle va également faire partie de la catégorie de l'approche « globale » et sera basée sur la méthode SCSPT.

I.2.5. Description de la méthode SCSPT

La méthode de test SCSPT est une méthode adoptant une approche globale. Elle est composée, comme les méthodes CCT et Combitest, de trois principales phases successives :

1. La détermination des conditions aux limites et du processus de test.
2. Le test du système complet dans le banc d'essais semi-virtuel.
3. Le traitement des données récoltées à l'issue du test pour l'évaluation/ extrapolation des performances de la durée du test à l'année.

Une description détaillée de la méthode peut être consultée dans [53] et [54].

I.2.5.1 Le banc de test semi-virtuel

Le banc d'essai semi-virtuel (situé à l'INES) permet de réaliser les essais de tout système thermique en le confrontant à un environnement virtuel (Figure 1- 29). Dans le banc d'essai, des modules hydrauliques et des résistances électriques sont utilisés pour émuler les conditions aux limites du système à évaluer (températures de retour au niveau de chaque circuit, etc.). Le système complet, avec tous ses sous-systèmes, est installé réellement comme s'il était dans un bâtiment, et ce sont les modules spécifiques (décrits dans le quatrième chapitre) qui garantissent la communication entre le système et le « bâtiment virtuel ». Ces modules hydrauliques sont approvisionnés par un réseau d'eau surchauffée à 180 °C et un réseau d'eau glacée à -12 °C. Les deux réseaux sont alimentés par une chaufferie centrale.

Le transfert des données entre les différents modèles employés pour créer les conditions aux limites virtuelles à l'aide du logiciel TRNSYS (capteur, bâtiment, climat, etc.) et les modules hydrauliques est assuré par l'intermédiaire d'une interface LabVIEW. Cette dernière permet aussi de suivre en temps réel l'évolution de certaines variables pendant l'essai. Les différents circuits du système (circuits d'ECS, de chauffage, de climatisation et le circuit solaire) sont chacun connectés à une unité spécifique. A chaque pas de temps, les variables physiques aux bornes du système testé (températures et débits du fluide dans chaque circuit) sont enregistrées. Ces données sont utilisées par le modèle TRNSYS du bâtiment virtuel pour simuler son interaction avec le système et calculer les débits et les températures de retour vers le système testé. Les modules hydrauliques reçoivent ces données et se chargent du réglage de la température du fluide (selon les consignes du modèle TRNSYS) à envoyer au système.

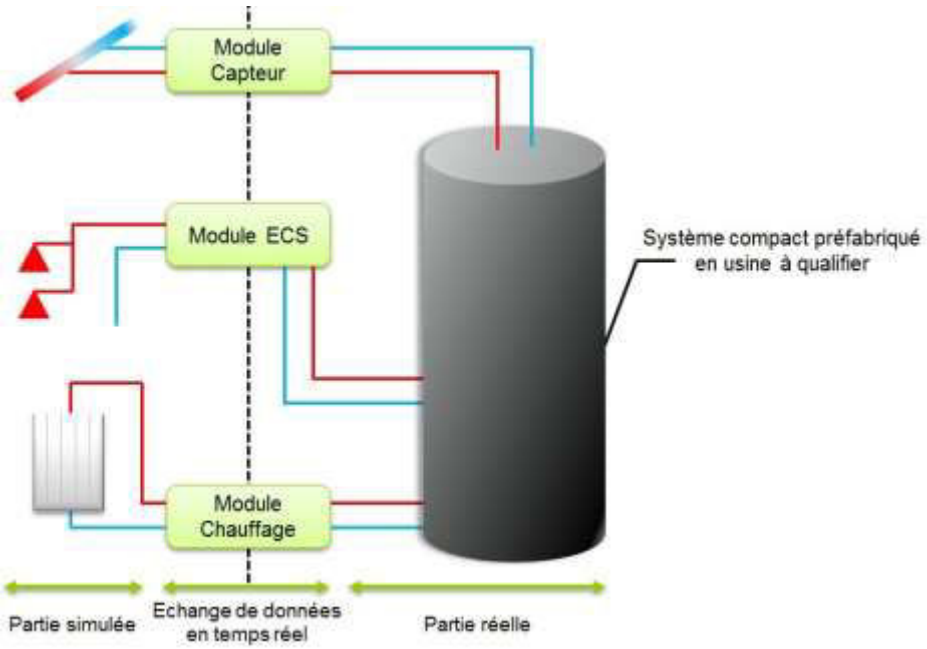

Figure 1- 29 : Illustration d'un système thermique pour les besoins de chauffage et d'ECS installé dans le banc d'essai semi-virtuel

La séquence complète de test consiste en :

1. une phase de conditionnement (32 heures) pour assurer que l'état du stockage est le même au début du test et à la fin du test.
2. une phase centrale de test avec l'environnement décrit ci-dessus (12 jours) ;
3. d'une phase de décharge du stockage (8 heures) voir Tableau 1- 4.

Tableau 1- 4 : Description des phases de la séquence de test complète [9]

N°	Phase	Durée (heures)	Description
1	Conditionnement initial	0	Conditionnement du ballon à 20°C (sans énergie solaire ni appoint)
2	Premier conditionnement	8	Parties basses et hautes du ballon amenées à température raisonnable. Partie haute chauffée par l'appoint selon sa consigne.
3	Deuxième conditionnement	24	Simulation de la dernière journée de la séquence pour amener le ballon dans un état semblable en début et fin de la phase centrale de test
4	Phase centrale de test	288	Réalisation du test de 12 jours dans les conditions déterminées par la méthode SCSPT
5	Décharge finale	8	Décharge du ballon

I.2.5.2 Les conditions aux limites

La Figure 1- 29 donne une représentation des limites entre le système réel et son environnement virtuel.

Avant de procéder au test réel du système sur le banc d'essais, les conditions aux limites de ce dernier doivent être définies et déterminées. Cette étape est cruciale dans le processus de test. C'est elle qui détermine comment le système va être sollicité et donc, en

conséquence, les informations qu'il sera possible de récupérer du système. La richesse de cette base de données est primordiale pour estimer avec une fiabilité suffisante les performances du système sur la période souhaitée (typiquement une année).

- Les conditions aux limites physiques du système testé

La méthode SCSPT étant basée sur une approche globale, tous les composants du système (hormis les capteurs solaires) sont à installer sur le banc d'essai : le, ou les stockages, le groupe solaire (pompes de circulation, vannes, systèmes de régulation etc.), les systèmes d'appoint (PAC, chaudière) et le système de climatisation (PAC ou machine à absorption). La plupart des systèmes actuels sont compacts, préfabriqués en usine.

Le seul composant qui n'est pas réellement testé est le capteur solaire. Le rayonnement solaire est intermittent, variable et surtout incontrôlable. L'intégration du capteur solaire n'est pas envisageable car une reproductibilité du test serait impossible, ou alors il conviendrait d'installer le capteur réel dans un environnement climatique difficile et onéreux à simuler. Par ailleurs, les capteurs solaires sont des composants relativement simples à caractériser. Le modèle de capteur utilisé permet de modéliser les capteurs plans et les capteurs à tubes sous vides. Il utilise des paramètres caractéristiques identifiés à partir d'essais normatifs.

Les autres conditions aux limites sont modélisées ainsi :

- Les besoins de chauffage et de climatisation

Ces besoins sont calculés grâce à des modèles de bâtiments définis lors de la Tâche 32 du programme SHC de l'AIE [55]. Ces bâtiments sont décrits dans le paragraphe II.4.2.2.

- Les besoins d'ECS

Les profils annuels de puisage d'ECS sont générés par l'outil développé par Jordan et al. [56]. Ils sont déterminés à l'aide de certaines méthodes statistiques pour prendre en compte des conditions quasi-réalistes. Dans le site internet du laboratoire Solar Energy Laboratory de l'Université du Wisconsin à Madison, certains profils sont disponibles gratuitement. Un exemple de profil de soutirage est illustré dans la Figure 1- 30.

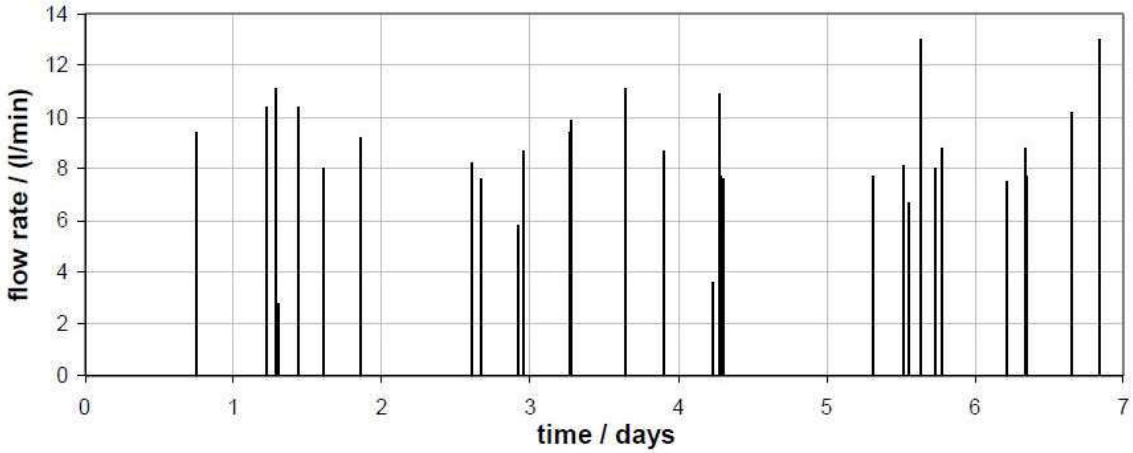

Figure 1- 30 : Profil de puisage d'ECS (exemple issu de [56])

La consigne pour la température de l'eau chaude ($T_{dhw,cons}$) est égale à 45°C. La température d'eau froide est, quant à elle, considérée comme étant une fonction sinusoïdale du temps *time* exprimé en heure (équation (1- 2)). $N_{day,seq}$ est le nombre de jours de la séquence qui est égal à 12 pour la méthode SCSPT. La variable T_{tap} dépend du climat considéré via les autres paramètres (température moyenne $T_{tap,av}$, amplitude ΔT_{tap} et déphasage d_{off}).

$$T_{tap} = T_{tap,av} + \Delta T_{tap} \times sin\left(360\frac{time + (273{,}75 - d_{off}) \times 24}{24N_{day,seq}}\right) \qquad \textbf{(1- 2)}$$

- Le climat de test

La méthode SCSPT utilise une séquence de 12 jours (6 et 12 jours respectivement pour Combitest et CCT) de données météorologiques (position du soleil, ensoleillement, température de l'air extérieur, température sèche et humidité). La séquence est représentative du climat dans lequel nous souhaitons tester le système et est générée par un algorithme présenté dans [53]. Cet algorithme, qui est la particularité de la méthode SCSPT, se base sur une simulation annuelle d'un système de référence.

Dans le Tableau 1- 5 sont données les conditions de la simulation du système de référence, modélisé lors de la Tâche 32 [55], dans le cas des SSC avec un appoint par une chaudière à gaz.

Tableau 1- 5 : Conditions de la simulation du système de référence pour la génération de la séquence de test d'un climat

Bâtiment	SFH60
Puisage ECS	200L/jour Pas de temps de 6minutes
Surface du capteur	12m²
Volume de stockage	600L

L'algorithme sélectionne 12 jours parmi les jours de l'année (pour le climat considéré) tel que [9] :

- Deux critères particuliers, la variation d'énergie stockée dans le ballon (proportionnelle à la variation de sa température) et les besoins de chauffage, aient une évolution

équivalente, jour après jour pendant la séquence de test et mois après mois pendant la séquence annuelle.

- La consommation d'énergie du système testé pendant la séquence « 12 jours » soit proportionnelle (facteur de 365/12) avec celle correspondant à l'année entière pour couvrir les mêmes besoins et pour le même climat (Figure 1- 31).

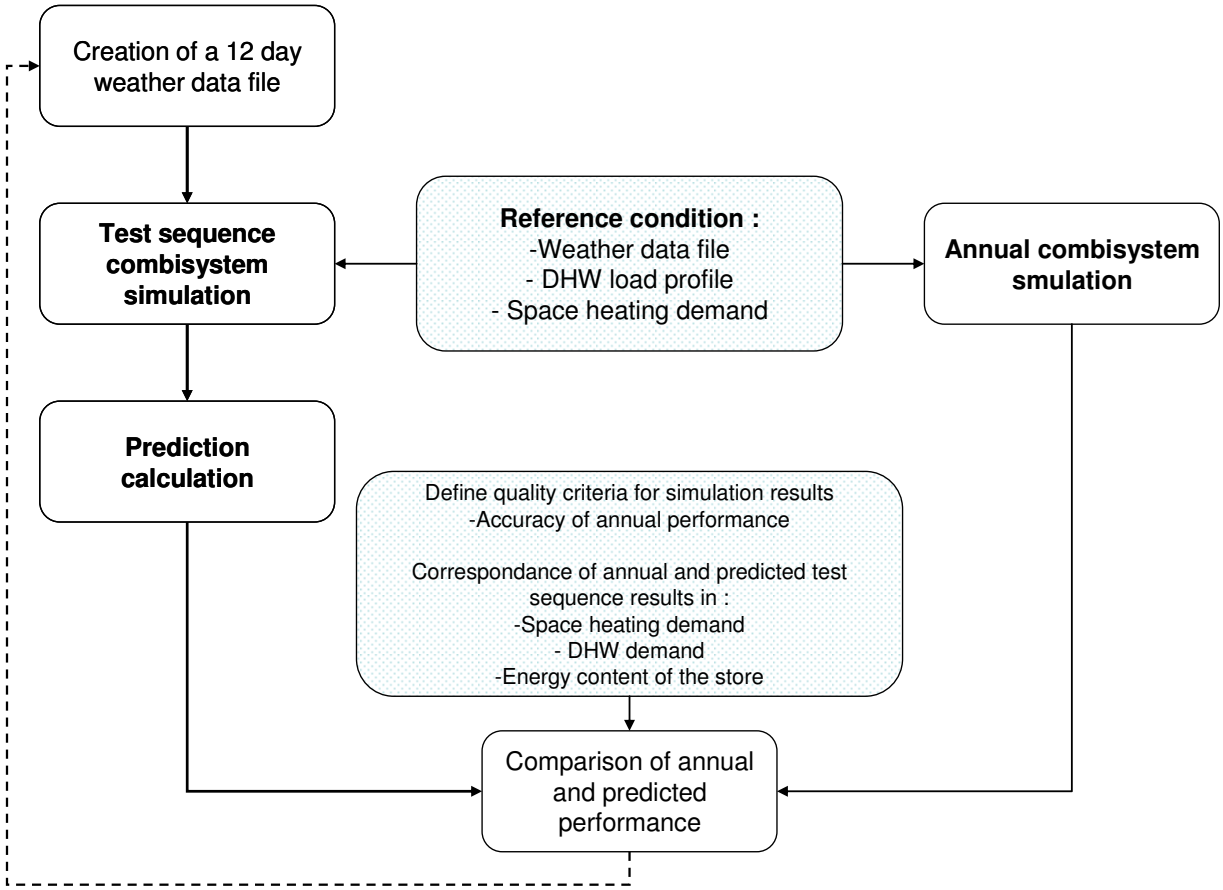

Figure 1- 31 : Processus de sélection des 12 jours de la séquence de test pour un climat et un système donné [53]

A titre d'exemple, les jours sélectionnés par l'algorithme à partir de la séquence annuelle Meteonorm12 de Zurich (CH-Zuerich-SMA-66600.tm2) et Stockholm (SE-Stockholm-Arlanda-24600.tm2) sont présentés dans le Tableau 1- 6.

Tableau 1- 6 : Jours sélectionnés pour constituer les séquences de test correspondant aux climats de Zurich et Stockholm [9]

Climat	N° jour séquence de test	1	2	3	4	5	6	7	8	9	10	11	12
Zurich	N° jour séquence annuelle	20	57	298	116	198	228	244	259	134	57	26	328
Stockholm		64	70	84	137	163	140	140	126	121	113	51	28

Les évolutions de l'énergie stockée dans le ballon et des besoins de chauffage pour un exemple de SSC dans deux climats différents sont données dans la Figure 1- 32.

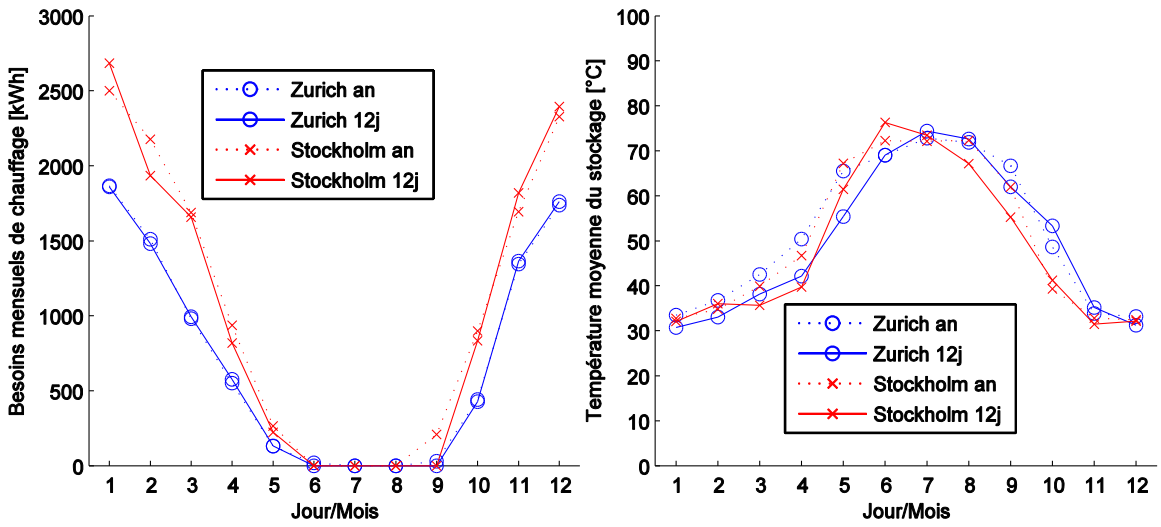

Figure 1- 32 : Comparaison des besoins de chauffage (à gauche) et de variation de l'énergie stockée (à droite) entre la séquence « 12 jours » et la séquence annuelle [9]

La performance du SSC évaluée suite au test selon la méthode SCSPT est valable pour le climat et le type de bâtiment choisis dans cet environnement virtuel.

La méthode de détermination des 12 jours, initialement développée pour les systèmes solaires combinés avec un appoint par une chaudière à gaz, a été étendue également aux systèmes qui intègrent une PAC ou une machine à absorption. L'extension de la méthodologie de sélection des 12 jours pour un SSC classique à un système multifonction (ECS, chauffage et climatisation) consiste à intégrer dans l'algorithme de sélection un autre critère sur les besoins de climatisation. En effet, ce dernier doit également avoir une évolution équivalente, jour après jour pendant la séquence de test et mois après mois pendant la séquence annuelle [57]. Un exemple de l'évolution des 3 critères pour la sélection des 12 jours dans le cas d'un système avec 3 fonctions dans un bâtiment à Aix-en-Provence est donné par la Figure 1- 33, Figure 1- 34 et Figure 1- 35.

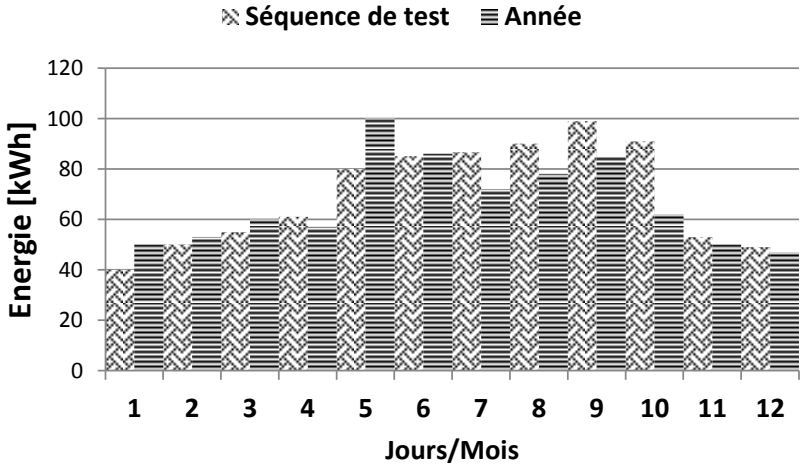

Figure 1- 33 : Comparaison de variation de l'énergie stockée entre la séquence « 12 jours » et la séquence annuelle (système avec rafraichissement solaire)

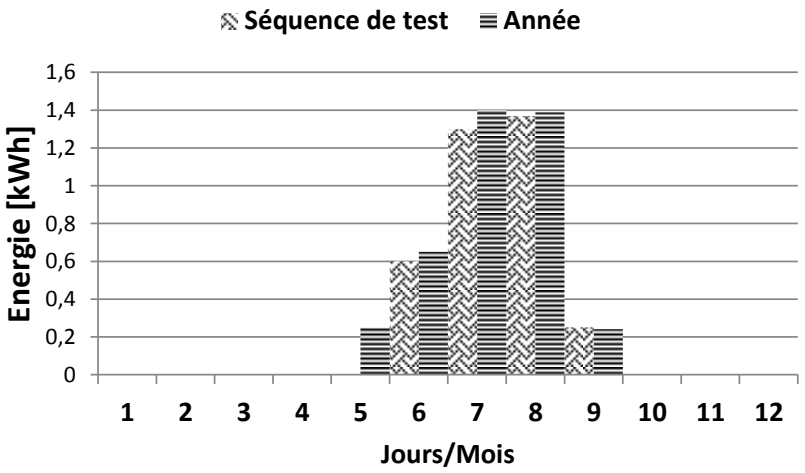

Figure 1- 34 : Comparaison des besoins en climatisation entre la séquence « 12 jours » et la séquence annuelle

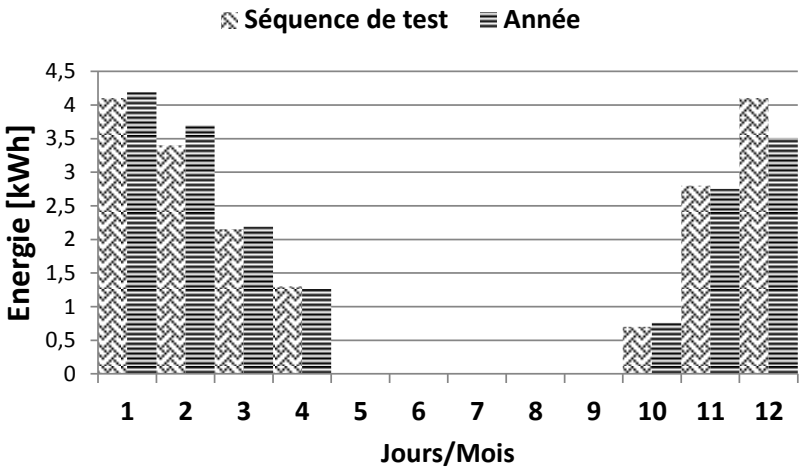

Figure 1- 35 : Comparaison des besoins en chauffage entre la séquence « 12 jours » et la séquence annuelle

CONCLUSION

L'analyse des données concernant la situation énergétique actuelle au niveau mondiale montre qu'il est primordial de privilégier l'utilisation des ressources renouvelables. Dans ce contexte, le secteur du bâtiment apparaît comme un poste important (environ 30%) de consommation d'énergie finale et ce uniquement pour les besoins de chauffage, climatisation et production d'eau chaude sanitaire. Les systèmes thermiques innovants, utilisant des ressources énergétiques renouvelables, constituent une solution alternative aux systèmes actuels.

Il existe de nombreuses configurations de SSC. Chacune se distingue par sa propre conception hydraulique, ses sources d'énergie et sa stratégie de gestion des flux de chaleur en son sein. Cependant, le marché de ces systèmes subit de fortes contraintes dans son développement notamment parce qu'il manque une méthode fiable permettant l'estimation des

performances d'un système donné avant son intégration au bâtiment. Face à cette multitude de solutions technologiques, les utilisateurs potentiels se retrouvent donc sans outil fiable permettant le choix du système adapté à leurs besoins.

L'évaluation des performances des SSC est d'autant plus difficile que leurs comportements sont très dépendants de leur dimensionnement, de leur installation et surtout des conditions énergétiques auxquels ils sont confrontés (climat et qualité du bâtiment). C'est pour cette raison que des travaux de recherches sont actuellement en cours pour développer des méthodes d'évaluation de performances basées sur des essais physiques de courte durée. Les méthodes actuelles se décomposent en deux catégories :

- Approche « composant » : basée sur des essais de composants séparés.
- Approche « globale » : basée sur un essai unique du système dans sa globalité

Les méthodes faisant partie de la deuxième approche sont plus adaptées à l'objectif de la thèse notamment par sa capacité à prendre en compte les vraies interactions entre les composants du système à évaluer. Cependant, ces méthodes présentent quelques inconvénients :

- Soit elles ne permettent pas l'extrapolation des résultats à des conditions différentes de celles de l'essai,
- Soit elles sont basées sur des modèles physiques, ce qui les rend complexes et donc moins adaptées à une perspective de certification ou de norme. En outre, l'application de la méthodologie d'évaluation ainsi que la précision de ses résultats, dépendront de la disponibilité et de la fiabilité des modèles.

Au travers des travaux de recherches de cette thèse une nouvelle méthodologie est proposée. Elle est basée sur un essai physique court selon la méthode SCSPT et sur un modèle complètement « boîte noire » (réseau de neurones) identifié suite aux résultats de l'essai.

Chapitre II. Développement d'une méthodologie générique d'évaluation des SSC

INTRODUCTION

Après avoir présenté les méthodes actuelles de qualification des systèmes solaires thermiques innovants et leurs limites, l'objectif de ce chapitre consiste à présenter le processus de développement d'une nouvelle méthodologie. Cette dernière est basée sur un essai physique selon la méthode SCSPT et sur l'identification d'un modèle « boîte noire » du système complet.

Dans un premier temps, le deuxième chapitre commence par une présentation globale de la méthodologie proposée. Les différentes étapes constitutives de cette dernière sont présentées et commentées. Après une présentation globale des différentes techniques de modélisation, une attention particulière est portée à l'approche basée sur les réseaux de neurones artificiels (RNA). Cette approche est pertinente et adaptée à une méthodologie générique. Le développement du processus de modélisation par RNA est ensuite présenté.

Dans un deuxième temps, ce chapitre s'intéresse à la finalisation de la méthodologie. Une configuration de modélisation définissant les entrées et la sortie du modèle est déterminée. Elle est constituée des variables relatives aux différentes puissances mises en jeu aux bornes du système solaire thermique considéré. La question de la sélection de modèles neuronaux est ensuite abordée et trois critères statistiques sont retenus comme critère de sélection dans le cadre de la méthodologie.

Le deuxième chapitre se termine par une présentation des résultats d'application de la nouvelle méthodologie dans le cas d'un SSC avec un appoint par une chaudière au gaz naturel. Dans cette partie les estimations des modèles RNA sont comparées à un modèle physique développé sous le logiciel TRNSYS.

II.1. Présentation de la méthodologie

II.1.1. Exigences vis-à-vis de la méthodologie

Les enjeux de l'évaluation des performances énergétiques des systèmes thermiques (étiquetage énergétique et garanties de performances par exemple), pour les applications dans le bâtiment, ainsi que la complexité de ces derniers, fixent un certain nombre d'exigences que la méthodologie doit respecter. En effet, pour qu'elle soit pertinente mais aussi pour surmonter les faiblesses des méthodes actuelles, présentées dans le premier chapitre, la méthodologie doit satisfaire les conditions suivantes :

i. Générique : Cette condition implique que la méthodologie puisse être applicable à différents systèmes avec différentes sources d'énergie (Système solaire combiné à une chaudière à gaz ou PAC) et répondant à différentes fonctions (ECS, chauffage et climatisation via une machine à absorption par exemple). Ainsi la méthode doit se présenter comme un outil que les ingénieurs, voire des techniciens, peuvent utiliser sans avoir besoin de connaissances approfondies (modélisation des composants etc.).

ii. Non-intrusive : Une méthode qui a cette caractéristique permet l'évaluation du système en se basant uniquement sur ses entrées-sorties (E/S) standards. Une méthode non-

intrusive ne nécessite aucun démontage du système pour la mise en place de métrologie. Ceci permet d'éviter toute manipulation qui, potentiellement, endommagerait le système ou modifierait son comportement. Cette caractéristique est d'autant plus importante que les systèmes actuels ont tendance à être compacts.

iii. Globale : La méthode doit prendre en compte l'essai du système considéré dans sa globalité (Figure 1- 28), c'est-à-dire, installé comme il le serait dans un bâtiment réel. Avec cette condition, les vraies interactions qui se produisent entre les différents composants (ballon, régulation, appoint etc.) y compris les séquences de régulation sont prises en compte dans le processus de la modélisation. Le test du système global prendra aussi moins de temps par rapport au cas où chaque composant est testé séparément.

iv. Courte : Dans une optique de certification, le processus d'évaluation doit être de courte durée. En outre, un test qui dure moins longtemps coûtera moins cher. La durée de test actuel de 12 jours est considérée comme maximale.

v. Aptitude à généraliser : Cette caractéristique permettra l'extrapolation des résultats des conditions du test à différents environnements (climats et performance du bâtiment).

Les exigences fixées vis-à-vis de la méthodologie ainsi que les perspectives de la thèse vont imposer des contraintes importantes quant à la modélisation envisagée et au processus d'identification.

II.1.2. Les étapes de la méthodologie proposée

L'évaluation, *a priori*, des performances énergétiques (consommation d'énergie) d'un système donné n'est possible qu'à la base d'un modèle simulant le fonctionnement de ce dernier. Une modélisation fiable nécessite généralement d'être calée par des vérifications à l'aide de données réelles (expériences physiques ou numériques). Ces vérifications consistent à identifier et calibrer les modèles utilisés. La méthodologie d'évaluation des performances des systèmes thermiques proposée est naturellement et conjointement basée sur un test du système et sur la construction de son modèle. Elle est constituée de cinq étapes successives. Le processus de la méthodologie est représenté sur la Figure 2- 1.

Figure 2- 1 : Les cinq étapes du processus de la méthodologie d'évaluation proposée

Les trois premières étapes consistent à définir et à réaliser le test du système. La qualité et la richesse des données sont primordiales dans l'identification du modèle. En effet, ces dernières doivent être représentatives du fonctionnement réel du système à modéliser, de la façon la plus complète possible. Par son caractère global (puisqu'elle fait partie de l'approche « globale »), comme par la qualité et la durée de sa séquence de test, la méthode de test SCSPT, décrite dans le premier chapitre, se montre très pertinente pour la méthodologie

proposée (ses avantages par rapport aux autres méthodes ont été présentés dans le paragraphe I.2.4.1).

La quatrième phase consiste à identifier un modèle du système qui a été testé sur le banc d'essai semi-virtuel. Cette phase inclut quelques sous étapes telles que la définition des entrées les plus pertinentes, le traitement et la préparation des données ainsi que la construction du modèle.

Il est possible, grâce à des simulations numériques du modèle identifié, d'évaluer les performances énergétiques du système dans les conditions souhaitées : c'est la cinquième étape.

Tel que la méthodologie est définie, elle est globale, courte et non intrusive (car basée sur la méthode SCSPT). Pour qu'elle soit conforme à la totalité des cinq exigences (paragraphe précédent), une attention particulière a été attachée aux choix du type de la modélisation.

Dans la section suivante une présentation et analyse des différentes méthodes de modélisation possibles, et en particulier de celle qui a été retenue, sont réalisées.

II.2. Modélisation adaptée à la problématique du sujet

II.2.1. Définition et principes de la modélisation

Dans ce paragraphe sont données les définitions des différents concepts qui seront utilisés par la suite.

II.2.1.1 Qu'est-ce que la modélisation d'un système ?

Selon O. Pajonk [58] un système est un objet dans lequel des variables de différentes sortes interagissent (suivant des lois physiques, économiques, etc.) pour produire des résultats observables (sorties) (Figure 2- 2). Le mot système désigne, comme déjà évoqué dans le premier chapitre, les différentes machines thermiques. Dans le cadre de cette étude, nous considérons les définitions suivantes pour les variables d'un système solaire thermique :

- Les entrées : variables à intérêt pour le fonctionnement du système, mesurables, contrôlables ou non contrôlables imposées au système.
- Les sorties : variables mesurables gérées par le système.
- Les perturbations : variables mesurables ou non mesurables, mais non contrôlables, imposées au système. Elles constituent une perturbation du fonctionnement normal du système. Par exemple le rayonnement solaire n'est pas considéré comme perturbation malgré le fait qu'elle ne soit pas contrôlable car c'est une grandeur à intérêt pour le fonctionnement du système.

La transformation des entrées aux sorties par le système est ce qui est communément appelé un processus. Ce dernier est dit dynamique si le temps y joue un rôle fondamental, sinon il est qualifié de statique.

Dans ce contexte, la modélisation est la représentation d'une réalité visible ou observable. Modéliser un processus consiste donc à lui donner une représentation mathématique (modèle, approximation de la fonction '*F*' dans la Figure 2- 2) exprimant la relation entre les entrées et sorties du système. La nature et la forme de cette relation détermine ce qui est appelé dans ce

travail, la structure du modèle (ce point sera traité dans le paragraphe II.2.2.4). Sont appelés paramètres, les constantes du modèle (au sein d'une structure de modèle) à identifier de telle sorte que ce dernier approxime la fonction système.

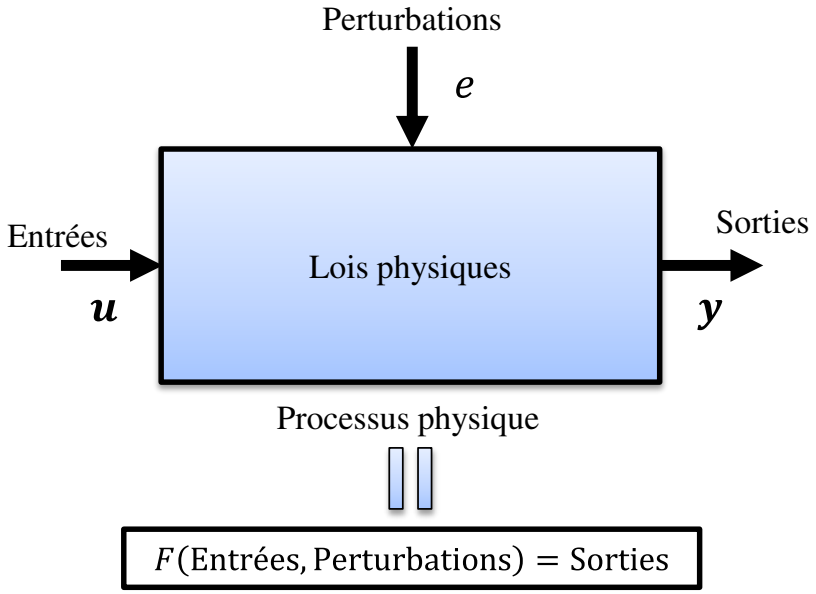

Figure 2- 2: Représentation schématique d'un système avec les interactions avec son environnement

II.2.1.2 Les différents types de modélisation

Trois approches de modélisation sont possibles. Le choix de l'approche est dicté par la connaissance disponible à propos du processus à modéliser.

La première approche consiste à construite un modèle moyennant l'application des lois physiques (ou financière etc. en fonction du domaine où se situe le processus) au système. Généralement, dans ce cas de figure, le modèle est la solution d'un ensemble d'équations mathématiques. Cette solution détermine, à la fois, la structure du modèle mais également ses paramètres (calculés en fonction des constantes caractéristiques du système). Ce type de modèle est qualifié de « boîte blanche » ou de connaissance [59]. Les modèles « boite blanche » sont précis et fiables. Il est donc toujours conseillé de considérer en premier lieu cette approche de modélisation. Cependant, en fonction du processus étudié, la réalisation de tels modèles peut s'avérer très coûteuse (résolution de systèmes d'équations complexes avec plusieurs variables) voire impossible (connaissances insuffisantes pour appréhender le système).

La deuxième approche dite « boîte noire » est une solution alternative à la précédente. Cette approche est nécessaire dans le cas extrême où la structure et les paramètres du modèle sont inconnus. L'objectif dans le cas d'une telle approche consiste à identifier un ensemble de paramètres d'un modèle mathématique dont la structure est choisie préalablement dans une famille de fonctions. Les paramètres sont identifiés de sorte à ce que le modèle colle au mieux à un ensemble de mesures ou d'observations d'entrées-sorties du système. L'intérêt de l'approche « boîte noire » réside dans sa capacité à modéliser des systèmes complexes dont la physique mise en jeu est en grande partie inconnue. Cependant, les paramètres ainsi que la structure du modèle n'ont aucune relation directe avec les constantes physiques du système. En outre, le modèle « boîte noire » dépend beaucoup de la

base de données utilisée pour l'identification des paramètres. La validité du modèle est en conséquence limitée au système considéré et au voisinage des données utilisées pour son élaboration.

La troisième approche de modélisation appelée « boîte grise » se situe entre les deux premières. Dans certains cas de figures, une partie du système est connu et modélisable physiquement tandis que le reste du système est inconnu. L'approche « boite grise » consiste en conséquence à combiner un modèle « boite noire », pour modéliser la partie inconnue, à un modèle « boite blanche » pour modéliser la partie connue. Cette approche permet d'étendre le domaine de validité du modèle mais également de limiter les inconvénients d'une modélisation complètement « boite noire ». Selon [60], il est préférable, dans la mesure où cela est possible, d'utiliser ce type de modèle car l'utilisation de connaissances physiques sur le fonctionnement du système permet de simplifier le modèle global, ce qui facilite beaucoup son entraînement (concept de « l'approximation parcimonieuse » évoquée dans le paragraphe II.2.1.3). L'approche « boîte grise » est certes avantageuse mais n'est pas forcément la plus simple (en termes de coût, du temps et de la complexité des calculs), mais reste dans certains cas quand même moins coûteuse qu'une approche totalement « boîte blanche ». Sur la Figure 2- 3 est représentée la frontière entre les trois approches de modélisation.

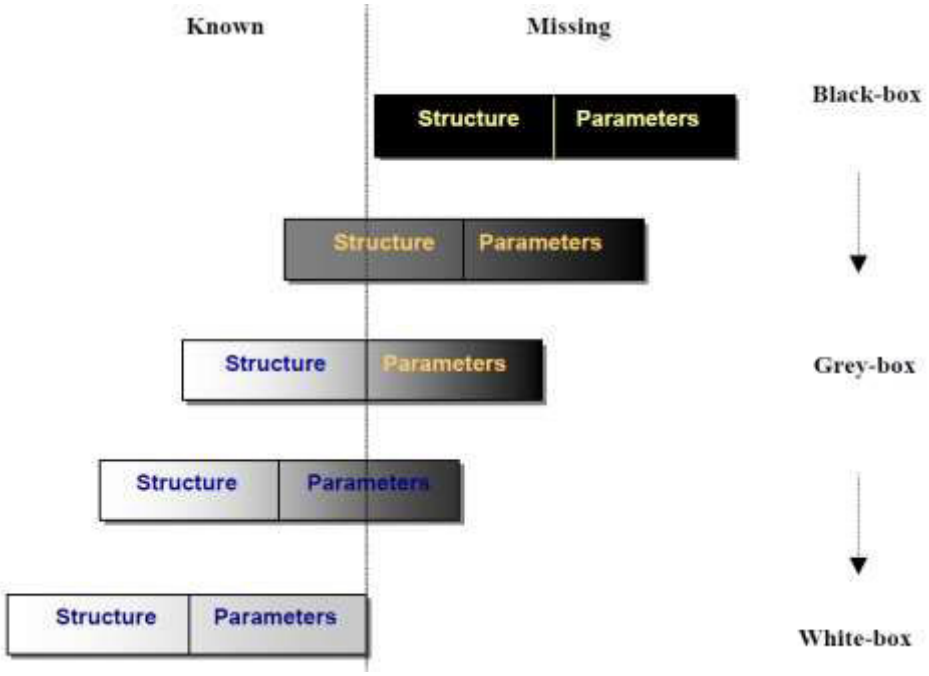

Figure 2- 3 : Frontière entre une modélisation type « boite noire » et « boite blanche » [61]

Des acquis scientifiques majeurs, au travers d'une précédente thèse [9] financée par l'ADEME, permettent de fournir les bases de la construction du travail proposé. En effet, pour les SSC avec un appoint par chaudière gaz, l'obtention d'un modèle caractéristique dynamique du comportement de ces systèmes a été obtenu à partir de modèles de comportement simple (« boite blanche ») et d'un réseau de neurones artificiel (« boite noire », simulant notamment la partie commande et régulation du système), dont l'ensemble des paramètres a été appris lors d'une séquence de test.

Les résultats ont montré que l'approche « boite grise » est pertinente. Cependant, le modèle développé reste intrinsèquement lié au système étudié et aux modèles physiques simplifiés

utilisés. La fiabilité ainsi que le caractère générique de la méthode dépendent beaucoup de la fiabilité et la disponibilité de ces modèles. En outre, les systèmes actuels sont compacts (Figure 1- 28), il devient donc difficile d'accéder à certains paramètres, tels que les rendements des composants et conductivités thermique par exemple (les industriels ne divulguent pas toutes les informations sur leurs systèmes), nécessaires pour le développement de modèles simplifiés mais fiables. En conséquence, ni l'approche « boite grise », ni « boite blanche » répondent aux deux premières exigences (paragraphe II.1.1.) que la méthodologie, proposée dans ce travail, doit satisfaire. La modélisation selon l'approche « boîte noire » vérifie ces conditions. En effet, elle n'est pas intrusive au système et son développement est possible en se basant seulement sur la connaissance des entrées et sorties du système.

Une approche complètement entrées-sorties « boite noire » a donc été retenue dans le cadre de la thèse.

II.2.1.3 Processus de construction d'un modèle selon l'approche « boîte noire »

La modélisation est un moyen très efficace pour l'investigation et la maitrise des systèmes dans tous les domaines. Comme il a été précisé plus haut, modéliser un système revient à rechercher une description mathématique de celui-ci à partir de données observables, pour un but précis. Les études de la modélisation doivent respecter trois principes [61] :

1. Isolation du système (détermination des entrées du modèle) : les limites du système investigué doivent être clairement définies.
2. Sélection des aspects à modéliser : Souvent il existe plusieurs interactions entre les composants du système et son environnement, or il est généralement difficile de les traiter toutes en même temps, il est donc nécessaire d'en négliger quelques-unes.
3. Parcimonie : le modèle le plus performant est celui qui est à la fois le plus simple mais aussi celui qui permet d'atteindre l'objectif de la modélisation [59].

Dans la Figure 2- 4 sont représentées les étapes générales de la modélisation. Ces dernières dépendent énormément du besoin et de comment le modèle sera utilisé. La modélisation peut avoir plusieurs objectifs [58] :

- L'analyse du système : le but dans ce cas est d'explorer le fonctionnement du système.
- L'optimisation du système : très souvent le processus d'optimisation de certains aspects (sortie du modèle) d'un système est très couteux ou difficile (voire présentant des risques) à réaliser en se basant sur le système réel. Une des alternatives consiste à planifier une expérience, dont le coût est acceptable, puis à s'en servir pour construire un modèle. Les algorithmes d'optimisation sont ensuite appliqués au modèle.
- L'élaboration de systèmes de contrôle-commande : dans le domaine du contrôle-commande des systèmes, il est nécessaire d'anticiper la réaction du système pour mieux agir sur ses entrées dans le but de respecter une consigne. Les modèles sont de plus en plus utilisés dans les systèmes de contrôles avancés.
- La détection et le diagnostic : il est possible de détecter les fautes de fonctionnement au sein d'un système via la comparaison entre la sortie du système et celle du modèle.

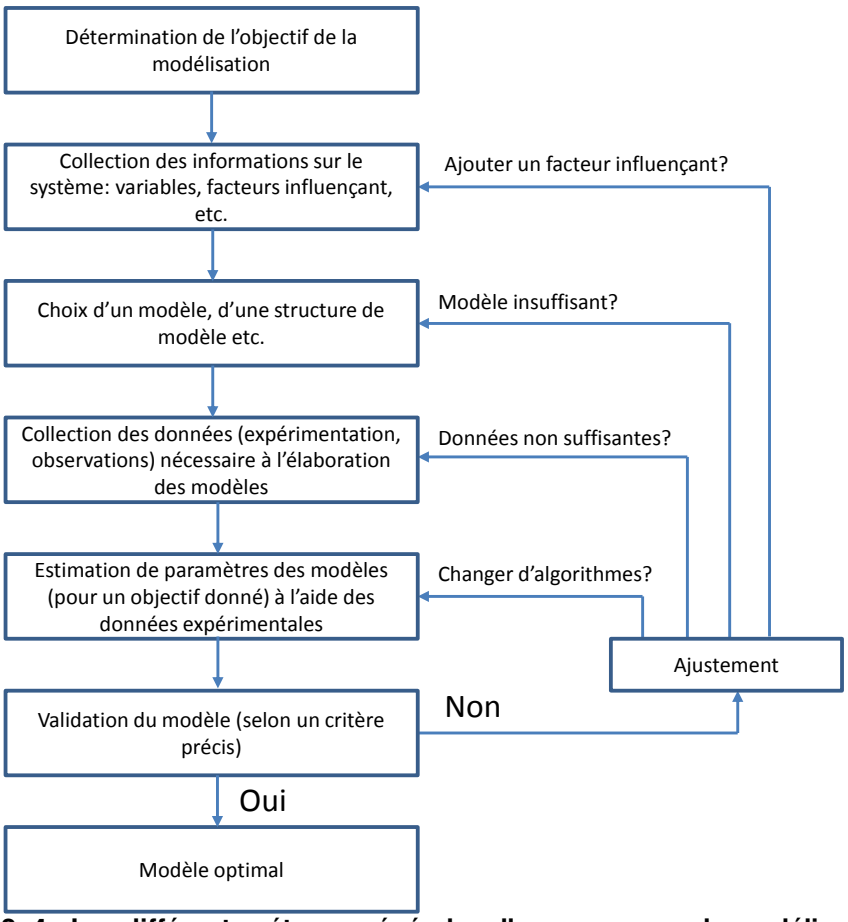

Figure 2- 4 : Les différentes étapes générales d'un processus de modélisation

Les deux premiers exemples d'objectifs nécessitent un modèle simulateur. Ce type de modèle est totalement indépendant du système après sa conception. Seules les entrées et, selon le modèle utilisé, les sorties en certains instants initiaux du système, sont nécessaires pour simuler son fonctionnement. Les deux derniers exemples d'objectifs nécessitent un modèle prédicteur. Pour que ce modèle puisse prédire la réponse du système aux instants futurs il se base sur la connaissance des entrées et des sorties à des instants présent et passés. Le modèle est donc étroitement lié au système et doit être utilisé en parallèle.

L'objectif de la thèse est de pouvoir estimer la consommation d'un système sur une année indépendamment de toute manipulation physique en dehors d'un test court. Les modèles simulateurs seront donc adoptés dans le cadre de cette étude.

La connaissance de l'objectif de la modélisation ainsi qu'une analyse physique du système sont nécessaires pour isoler le système. Ils déterminent également l'ensemble des entrées et sorties à prendre en compte. Mais ces entrées ne sont pas forcément pertinentes pour modéliser le système selon l'objectif prédéfini. Il va donc falloir réduire l'ensemble des entrées en éliminant les entrées perturbatrices.

Le concepteur peut souhaiter explorer l'influence d'un ensemble précis d'entrées sur la sortie. Dans ce cas, seule la connaissance de l'objectif de la modélisation est nécessaire. Dans le cas des modèles dynamiques, il faut également déterminer les entrées retardées à prendre en compte. Ce type de modèle est plus complexe. Cette question est traitée plus loin dans le paragraphe II.3.2.4.

II.2.2. Structure de modèle selon une approche « boîte noire »

II.2.2.1 Formulation mathématique

Dans la suite de ce paragraphe les systèmes sont considérés avec une seule sortie. Le cas des systèmes avec des sorties multiples peut être déduit du cas simplifié.

Selon le paragraphe II.2.1.1, la modélisation est l'approximation mathématique de la fonction réalisée par le système. Le problème de modélisation est donc équivalent à la détermination d'une fonction réelle inconnue $F: \mathcal{R}^E \rightarrow \mathcal{R}$ définissant la relation entre le vecteur des entrées, $\boldsymbol{u}(t)$, et la sortie, $y(t)$, du système considéré. Cette fonction sera appelé le « modèle-hypothèse » du système. Dans le cadre de l'approche « boite noire » l'identification de cette fonction est basée sur un ensemble de données expérimentales : $D_N = \{\boldsymbol{u}(t), y(t)\}_{t=1}^{N}$. E et N désignent respectivement le nombre d'entrées et d'échantillons de données.

Au sein de la catégorie des modèles « boîtes noires », il est possible de représenter le « modèle-hypothèse » (relation mathématique entre les entrées et sorties du système) de deux façons différentes : représentation d'état, selon le système d'équation (2- 1), et représentation E/S, équation (2- 2).

$$\begin{cases} \boldsymbol{x}(t) = F_1(\boldsymbol{x}(t-1), \boldsymbol{\varphi}(t-1)) \\ y(t) = F_2(\boldsymbol{x}(t), \boldsymbol{\varphi}(t)) + \boldsymbol{e}(t) \end{cases} \quad \textbf{(2- 1)}$$

$$y(t) = F(\boldsymbol{\varphi}(t)) + \boldsymbol{e}(t) \quad \textbf{(2- 2)}$$

$\boldsymbol{\varphi}$ est le vecteur de régression du modèle. Il est composé par les entrées $\boldsymbol{u}$ et éventuellement les sorties passées du système. $\boldsymbol{x}$ le vecteur d'état, non mesurable, du système et y sa sortie. F_1, F_2 et F sont les fonctions à approcher décrivant le fonctionnement du système.

La grandeur non mesurable $\boldsymbol{e}$ modélisant le bruit affectant le système est souvent supposé être « normal », au sens loi de probabilité, avec une moyenne nulle.

La réalisation d'une représentation entrées-sortie est toujours possible. En absence d'état mesuré du système la représentation d'état peut ne pas être possible [62]. Une recherche bibliographique a révélée qu'il n'existe pas un réel avantage de la dernière sur une approche entrées-sortie. Aussi, l'implémentation de l'approche entrées-sortie est plus simple [59]. Pour ces raisons une représentation entrées-sortie a été privilégiée dans le cadre de ce travail.

Le modèle exploitable du système est une approximation de la relation donnée par l'équation (2- 3) à l'aide d'une fonction paramétrique $\hat{F}$:

$$\hat{y}(t, \boldsymbol{w}) = \hat{F}(\boldsymbol{\varphi}(t); \boldsymbol{w}) \quad \textbf{(2- 3)}$$

$\hat{y}$ est la sortie du modèle et $\boldsymbol{w}$ est un vecteur de paramètres qu'il faut identifier. On suppose que la fonction F coïncide avec un membre de l'ensemble des fonctions $\hat{F}(.; \boldsymbol{w})$ pour un vecteur inconnu $\boldsymbol{w} = \boldsymbol{w}_0$ ($\hat{F}(.; \boldsymbol{w}_0) = F$).

Pour un vecteur $\boldsymbol{\varphi}$ et une famille de fonction $\hat{F}$ (structure du modèle) donnés, le processus d'identification de $\boldsymbol{w}$ de telle sorte que $\hat{F}$ approxime F (approximation de $\boldsymbol{w}_0$) connaissant D_N est appelé l'apprentissage ou entrainement du modèle.

II.2.2.2 Modèles linéaires et non-linéaires dynamiques

Dans cette partie on s'intéresse uniquement à la présentation des modèles dynamiques. En effet, les systèmes thermiques auxquels la thèse s'intéresse intègrent des composants qui ont un comportement dynamique introduit par l'inertie thermique de ses composants (notamment le ballon de stockage).

La plupart des systèmes sont non linéaires : leurs entrées et sorties sont liées par des fonctions non-linéaires. Ils nécessitent donc des modèles non linéaires pour les décrire de façon précise dans la totalité de leur domaine de fonctionnement. Cependant, dans certains cas les modèles linéaires peuvent constituer une approximation suffisante au voisinage de ce domaine.

Les modèles peuvent présenter deux types de linéarité : linéarité par rapport aux paramètres et linéarité par rapport aux entrées. Une fonction est dite linéaire par rapport aux entrées si sa sortie vérifie le principe de linéarité par rapport aux entrées.

$$\forall \pi, \mu \in \mathcal{R}, \hat{F}(\pi.u_1 + \mu.u_2; \boldsymbol{w}) = \pi.\hat{F}(u_1; \boldsymbol{w}) + \mu.\hat{F}(u_2; \boldsymbol{w}) \quad \textbf{(2- 4)}$$

Une fonction est dite linéaire par rapport à ses paramètres si sa sortie vérifie le principe de linéarité par rapport aux paramètres. Les variables π et μ appartiennent à l'ensemble des variables réelles.

$$\forall \pi, \mu \in \mathcal{R}, \hat{F}(u; \pi.\boldsymbol{w}_1 + \mu.\boldsymbol{w}_2) = \pi.\hat{F}(u; \boldsymbol{w}_1) + \mu.\hat{F}(u; \boldsymbol{w}_2) \quad \textbf{(2- 5)}$$

Dans la théorie d'identification des systèmes, la forme générale de la structure des « modèles hypothèses » linéaires est donnée par l'équation (2- 6) ([63]).

$$A(q^{-1}).y(t) = q^{-\zeta}\frac{B(q^{-1})}{F(q^{-1})}u(t) + \frac{C(q^{-1})}{D(q^{-1})}\boldsymbol{e}(t) \quad \textbf{(2- 6)}$$

Où :
$A(q^{-1}) = 1 + a_1 q^{-1} + \cdots + a_\tau q^{-\tau}$
$B(q^{-1}) = b_0 + b_1 q^{-1} + \cdots + b_m q^{-m}$
$C(q^{-1}) = 1 + c_1 q^{-1} + \cdots + c_k q^{-k}$
$D(q^{-1}) = 1 + d_1 q^{-1} + \cdots + d_l q^{-l}$
$F(q^{-1}) = 1 + f_1 q^{-1} + \cdots + f_r q^{-r}$

Dans le cas de modèles « multi-entrées / multi-sorties », ces polynômes sont des matrices polynomiales. q^{-1} est l'opérateur de décalage temporel, il vérifie :

$$\forall \zeta > 0, q^{-\zeta}u(t) = u(t - \zeta) \quad \textbf{(2- 7)}$$

La façon avec laquelle les perturbations du système sont prises en compte dans le modèle (bruit de sortie ou d'entrée) résulte en un type donné du « modèle hypothèse ». Dans le Tableau 2- 1 sont présentés les cinq modèles les plus utilisés ainsi que les prédicteurs associés avec le vecteur $\boldsymbol{\varphi}$ correspondant.

Les modèles non-linéaires se déclinent simplement des modèles linéaires en choisissant une fonction $\hat{F}$ non-linéaire.

Les systèmes solaires thermiques pour l'habitat sont régis par un ensemble de phénomènes physiques non linéaires (transferts thermiques par rayonnement et par convection, changement de phase du fluide réfrigérant etc.). Ainsi l'estimation de leurs performances quelles que soient les conditions nécessite l'emploi d'une structure non linéaire dans le modèle.

Une discussion concernant le choix de la famille de fonction auquel appartient le « modèle hypothèse » retenu pour les travaux réalisés est donnée dans le paragraphe II.2.2.4.

Tableau 2- 1 : Les différents modèles linéaires avec leurs modèle associé, vecteur de régression et vecteur de paramètres à identifier

Nom du modèle-hypothèse	Réponse impulsionnelle finie (FIR) : $A = C = D = F = 1$	Modèle Autorégressif à variable eXogène (ARX) : $C = D = F = 1$	Modèle Autorégressif à Moyenne Ajustée et variable eXogène (ARMAX) : $D = F = 1$	Modèle d'erreur en sortie (Output Error OE) : $A = C = D = 1$
Vecteur de régression $\boldsymbol{\varphi}$	$[u(t-\zeta),\dots,u(t-\zeta-m)]'$	$[y(t-1),\dots,y(t-\tau),u(t-\zeta),\dots,u(t-\zeta-m)]'$	$[y(t-1),\dots,y(t-\tau),u(t-\zeta),\dots,u(t-\zeta-m),y(t-1)-\hat{y}(t-1;\boldsymbol{w}),\dots,y(t-k)-\hat{y}(t-k;\boldsymbol{w})]'$	$[\hat{y}(t-1;\boldsymbol{w}),\dots,\hat{y}(t-r;\boldsymbol{w}),u(t-\zeta),\dots,u(t-\zeta-m)]'$
Vecteur à identifier $\boldsymbol{w}$	$[b_0 \dots b_m]'$	$[-a_1,\dots,-a_\tau,b_0,\dots,b_m]'$	$[-a_1,\dots,-a_\tau,b_0,\dots,b_m,c_1,\dots,c_k]'$	$[-f_1,\dots,-f_r,b_0,\dots,b_m]'$
Modèle associé	$\hat{y}(t) = q^{-\zeta}B(q^{-1})u(t)$	$\hat{y}(t) = q^{-\zeta}B(q^{-1})u(t) + [1-A(q^{-1})]y(t)$	$\hat{y}(t) = q^{-\zeta}\frac{B(q^{-1})}{C(q^{-1})}u(t) + \left(1 - \frac{A(q^{-1})}{C(q^{-1})}\right)y(t)$	$y(t) = q^{-\zeta}\frac{B(q^{-1})}{F(q^{-1})}u(t)$

II.2.2.3 Méthode d'estimation des paramètres

Dans cette partie une description des deux méthodes les plus utilisées pour l'estimation des paramètres est donnée. Les références [60] et [64] décrivent avec détail ces méthodes ainsi que de nombreuses autres.

Le processus d'estimation du vecteur $\mathbf{w}$ est représenté sur la Figure 2- 5. Les méthodes les plus utilisées pour estimer ce vecteur sont la méthode basée sur la minimisation de l'erreur moyenne quadratique et la méthode du maximum de vraisemblance. Ces méthodes se distinguent par la « fonction objectif » utilisée pour mesurer l'écart entre le modèle et le système au moment de l'apprentissage. Elle est fonction de l'erreur entre la sortie du modèle $\hat{y}$ et la sortie du système y :

$$Obj(\boldsymbol{w}) = Obj(y - \hat{y}(D_N, \boldsymbol{w})) \quad \textbf{(2- 8)}$$

Le choix de la fonction objectif dans les méthodes de l'estimation des paramètres dépend des connaissances disponibles sur le processus à modéliser notamment sur les perturbations.

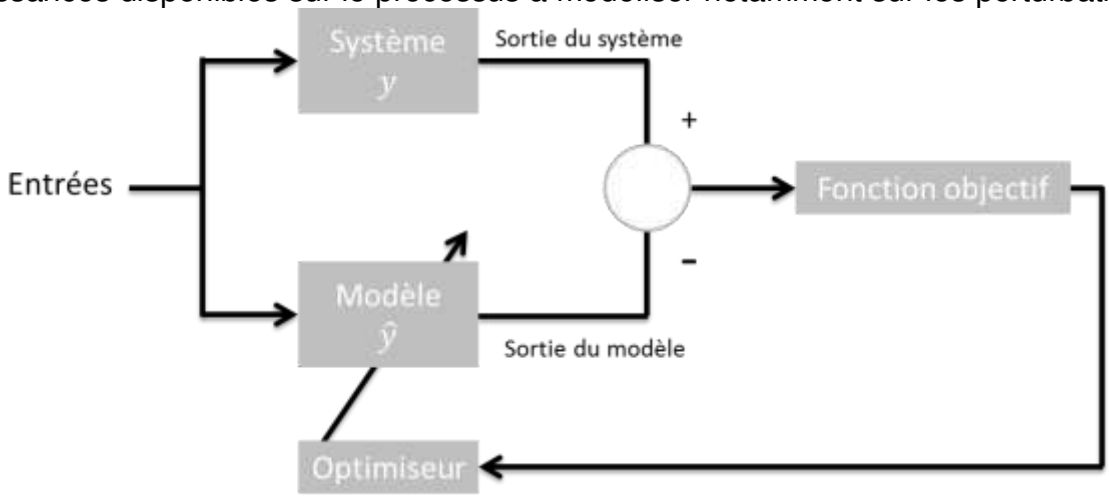

Figure 2- 5 : Représentation schématique du processus d'estimation de paramètres d'un modèle

- La méthode basée sur l'erreur moyenne quadratique :

La fonction objectif dans ce cas de figure est mesurée par la fonction d'erreur moyenne quadratique (EMQ) :

$$Obj(w; D_N) = \frac{1}{N}\sum_{t=1}^{N}\left(y(t) - \hat{y}(t, \boldsymbol{w})\right)^2 \quad \textbf{(2- 9)}$$

L'application de cette méthode est toujours possible car elle nécessite uniquement la connaissance des mesures des sorties du système et celles du modèle pour les mêmes données d'entrée. La solution du problème dans ce cas de figure est donnée par le minimum de la fonction objectif, équation (2- 10) :

$$\widehat{\boldsymbol{w}} = arg\min_{\boldsymbol{w}} Obj(\boldsymbol{w}; D_N) \quad \textbf{(2- 10)}$$

Dans certains cas, l'écart-type du bruit qui affecte les mesures $y(t)$ est calculable. Pour prendre en compte les incertitudes des mesures dans le modèle, une variante de la fonction objectif EMQ consiste à pondérer les erreurs par l'écart-type correspondant $\sigma(t)$ (équation (2-11)). Une mesure aura d'autant plus de poids que son incertitude sera faible.

$$Obj(\boldsymbol{w}; D_N) = \frac{1}{N}\sum_{t=1}^{N}\left(\frac{y(t) - \hat{y}(t, \boldsymbol{w})}{\sigma(t)}\right)^2 \quad \textbf{(2- 11)}$$

- Méthode du maximum de vraisemblance (MV):

Le principe de la méthode du MV est basé sur une vision probabiliste. Son principe est différent de celui de la précédente qui consiste à minimiser une fonction d'erreur sur la base d'observations. La méthode MV considère les observations $y(t)$ comme étant des variables

aléatoires d'une population inconnue. La solution au sens du MV est le vecteur $\widehat{\boldsymbol{w}}$ des paramètres maximisant la probabilité d'apparition des observations $\boldsymbol{y}$. Notons $p_{\boldsymbol{y}|\boldsymbol{u}}(x, \boldsymbol{w}_0)$ la densité de probabilité conditionnelle de la variable aléatoire $\boldsymbol{y}$, contenant les N observations, connaissant le vecteur $\boldsymbol{u}$. Cette densité de probabilité est également considérée comme faisant partie d'une famille paramétrée de fonctions densité de probabilité $p_{\boldsymbol{y}|\boldsymbol{u}}(x, \boldsymbol{w})$. La fonction de vraisemblance est définie par :

$$L(\boldsymbol{w}) = p_{\boldsymbol{y}|\boldsymbol{u}}(x = \boldsymbol{y}, \boldsymbol{w}) \qquad \textbf{(2- 12)}$$

La solution du problème au sens du MV est donnée par :

$$\widehat{\boldsymbol{w}} = arg \max_{\boldsymbol{w}} L(\boldsymbol{w}) \qquad \textbf{(2- 13)}$$

Pour trouver une telle solution les densités de probabilité des variables mises en jeu doivent être connues. Sous l'hypothèse de variables indépendantes et identiquement distribuées (ayant la même loi de probabilité) il est possible de calculer la densité de probabilité $p_{\boldsymbol{y}|\boldsymbol{u}}$ en fonction des densités de probabilité conjointe:

$$p_{\boldsymbol{y}|\boldsymbol{u}}(\boldsymbol{x}, \boldsymbol{w}_0) = \prod_{t=1}^{N} p_{y(t)|\boldsymbol{u}(t)}(x(t), \boldsymbol{w}_0) \qquad \textbf{(2- 14)}$$

Moyennant l'équation (2- 2) il est possible de calculer les densités de probabilité conjointe selon :

$$p_{y(t)|\boldsymbol{u}}(x(t), \boldsymbol{w}_0) = p_{\boldsymbol{e}}\left(x(t) - F\big(\boldsymbol{u}(t)\big)\right) \qquad \textbf{(2- 15)}$$

Ce qui donne en introduisant $\hat{y}(t, \boldsymbol{w}_0)$, l'estimation idéale de $y(t)$, et en utilisant l'équation (2- 14):

$$p_{\boldsymbol{y}|\boldsymbol{u}}(\boldsymbol{x}, \boldsymbol{w}_0) = \prod_{t=1}^{N} p_{\boldsymbol{e}}\big(x(t) - \hat{y}(t, \boldsymbol{w}_0)\big) \qquad \textbf{(2- 16)}$$

La grandeur à maximiser est donc :

$$L(\boldsymbol{w}) = \prod_{t=1}^{N} p_{\boldsymbol{e}}\big(y(t) - \hat{y}(t, \boldsymbol{w})\big) \qquad \textbf{(2- 17)}$$

En considérant que $\boldsymbol{e}$ suit une loi normale de moyenne nulle et d'écart type σ (hypothèse souvent admise, comme précisé dans le paragraphe II.2.2.1), l'équation (2- 17) devient :

$$L(\boldsymbol{w}) = \prod_{t=1}^{N} \frac{1}{\sigma\sqrt{2\pi}} exp^{-\frac{1}{2}\left(\frac{y(t) - \hat{y}(t, \boldsymbol{w})}{\sigma}\right)^2} \qquad \textbf{(2- 18)}$$

Si de plus la variance de $\boldsymbol{e}$ est supposée égale à l'unité, le développement de l'équation (2-18) permet de montrer que :

$$L(\boldsymbol{w}) = \left(\frac{1}{\sqrt{2\pi}}\right)^N exp\left(-N\frac{EMQ}{2}\right) \quad \textbf{(2- 19)}$$

Dans ce cas précis les méthodes MV et la première sont équivalentes puisque le maximum de $L(\boldsymbol{w})$ correspond au minimum de $EMQ(\boldsymbol{w})$.

II.2.2.4 Exemples de familles des structures de modèles

Le choix d'une famille de structure du « modèle hypothèse » $\hat{F}$ est une étape cruciale dans le processus de la modélisation. Il est donc primordial de choisir une structure adaptée au problème étudié. L'utilisation d'une structure sous forme de polynômes peut être suffisante dans certains cas. L'identification de paramètres sera aisée et moins coûteuse en temps de calcul. Pour des problèmes complexes, auxquels il n'est pas possible d'utiliser une forme polynomiale, le développement en séries de Fourrier, de Volterra, l'utilisation d'un réseau d'ondelettes ou de systèmes d'inférence floue, par exemple, peuvent être la solution adaptée. Parmi ces familles les réseaux de neurones artificiels (RNA) ont une place très particulière.

La raison d'être des RNA réside dans leur faculté d'approximer toute fonction continue et bornée [65], voire toute fonction continue par morceaux et bornée [66]. Il a été montré qu'un RNA (fonction non-linéaire par rapport aux entrées mais aussi par rapport aux paramètres) avec une configuration adaptée (i. e. nombre de neurones et fonctions d'activations) est un approximateur universel de fonction à un degré de précision arbitraire fixé. En outre, la famille de fonctions des réseaux de neurones est parcimonieuse. En effet, les RNA peuvent approximer toute fonction avec moins de paramètres que les autres familles.

La modélisation par RNA est largement appliquée, notamment pour répondre à un besoin toujours croissant concernant la prédiction des performances et le contrôle-commande dans le secteur des énergies renouvelables et systèmes thermiques. Les RNA ont montré leur pertinence notamment par leur capacité à résoudre des problèmes complexes, non-linéaires et multi-variables. Un résumé de quelques travaux significatifs et récents est donné dans le paragraphe suivant. Les modèles neuronaux peuvent apprendre à partir d'exemples et tolèrent l'imprécision des mesures et des perturbations, ainsi que les données incomplètes [67]. En outre, les systèmes que nous souhaitons modéliser dans le cadre de cette thèse présentent des caractères complexes, à la fois dynamiques, non-linéaires et discontinues, introduits par les outils de contrôle commande. Les réseaux de neurones artificiels sont adaptés à notre problématique. Ils ont donc été retenus comme structure de modèle.

II.2.3. Réseaux de neurones artificiels et systèmes thermiques

Le travail bibliographique réalisé révèle que les RNA ont été largement utilisés avec succès pour modéliser quasiment tous les composants d'un système solaire combiné. Cela confirme l'intérêt à utiliser un modèle neuronal dans le processus de la méthodologie proposée dans ce travail. Dans ce qui suit un aperçu succinct de quelques travaux impliquant les RNA.

Dans [68], S. Kalogirou et al ont étudié la capacité d'un RNA statique à modéliser le comportement énergétique d'un système solaire thermique à grande échelle. D'après les résultats présentés, il peut être conclu que le RNA prédit efficacement la production d'énergie quotidienne du système. Les entrées nécessaires au modèle sont l'irradiation solaire quotidienne totale incidente au niveau du capteur, la température ambiante moyenne de la

journée et de l'eau du réservoir de stockage au début de la journée. Les prédictions du RNA ont été comparées à des prédictions d'une autre méthode préalablement validée. La concordance entre les résultats est satisfaisante.

Dans [69], C. Burattia et al ont développé un outil, intégrant un modèle neuronal, pour vérifier la performance énergétique annuelle des bâtiments ainsi que les certifications associées. Les entrées du modèle sont composées de paramètres géométriques du bâtiment, d'informations sur le climat et les types des systèmes thermiques intégrés au bâtiment. Parce que le RNA n'est compatible qu'avec des données numériques, tous les paramètres non quantifiables ont été codés en utilisant des nombres entiers. Les résultats de l'outil développé montrent une bonne corrélation avec les données de test.

A. Leconte et al ont développé, dans [52], un modèle selon l'approche « boîte grise » qui prédit, avec une bonne précision, les performances énergétiques annuelles des SSC dans de nombreuses conditions. Le modèle mis au point est composé d'une partie formée par des modèles physiques simples, et une partie formée avec un RNA statique pour modéliser le système de régulation et gérer les flux et échanges d'énergie entre tous les composants.

Dans [70], W. Yaici et al ont mené une étude afin d'évaluer la performance d'un système solaire thermique similaire. Le système a été modélisé à l'aide d'un RNA dynamique. Les variables d'entrée utilisées pour construire la configuration du modèle étaient, en plus du rayonnement solaire et la température ambiante, les températures retardées de six emplacements dans le réservoir de stockage d'eau chaude. Les sorties du modèle sont les mêmes six températures mais à l'instant futur, l'apport de chaleur du capteur solaire à travers l'échangeur de chaleur dans le système et l'apport de chaleur du système d'appoint. Les résultats montrent que le modèle RNA permet d'évaluer les performances énergétiques du système et la stratification du réservoir tout en commettant de faibles erreurs.

Dans [71], S. Rosiek et al ont développé un modèle basé sur RNA d'une machine à absorption entrainée par l'énergie solaire. Le modèle développé est capable de prédire avec une erreur faible le COP ainsi que la capacité de refroidissement du système. Pour cette étude, seules cinq entrées étaient pertinentes pour modéliser l'ensemble du système à savoir les températures d'entrée et de sortie de l'évaporateur et du générateur ainsi que la température moyenne du ballon de stockage chaud. Cependant, ces entrées utilisées ne sont pas adaptées pour évaluer les performances d'une machine à absorption lorsque seules les températures et les débits d'entrées sont disponibles. Une étude similaire est présentée dans [72].

Plusieurs autres études de modélisation avec des RNA ont été réalisées récemment, mais ne sont consacrées qu'à la modélisation de composants de système [73], [74], [75], [76].

II.3. Processus de modélisation avec les RNA

II.3.1. Les RNA formels

Les RNA trouvent leurs racines dans la neurobiologie. Cependant, l'origine biologique des réseaux de neurones artificiels a une importance et des conséquences très limitées dans leurs applications en ingénierie.

Les neurones formels sont des fonctions mathématiques paramétrables mimant le fonctionnement des neurones biologiques. Un neurone, élément de base d'un RNA, réalise une fonction d'une sommation pondérée, par des coefficients appelés poids synaptiques ω_j^l, de ses entrées x_j pour donner une sortie z_l avec l le numéro du neurone. L'opération réalisée par un neurone est traduite par l'équation (2- 20). La fonction f_l est appelé fonction d'activation du neurone.

$$z_l = f_l\left(\sum_{j=1}^{K} \omega_j^l x_j\right) \qquad \textbf{(2- 20)}$$

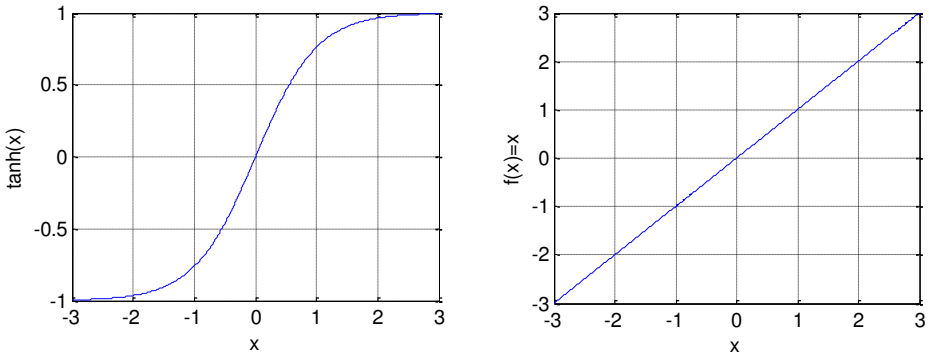

Figure 2- 6 : Quelques exemples de fonctions d'activation couramment utilisées

Quelques exemples de fonctions d'activation (FA) couramment utilisées sont représentés sur la Figure 2- 6. Elles sont généralement des fonctions bornées sauf pour les neurones de sortie.

Des neurones interconnectés forment ce qui est communément appelé un RNA. Il existe différentes architectures neuronales [77]. Parmi celles-ci figure le perceptron multicouche (PMC). Cette architecture est la plus utilisée. Un PMC est un réseau formé de neurones arrangés en couches : une couche d'entrée (formée de variables d'entrées), une ou plusieurs couches cachées et une couche de sortie (Figure 2- 7). Les neurones sont complètement connectés dans un seul sens, des entrées aux sorties. Pour un RNA avec E entrées, C neurones dans la couche cachée et S sorties, la sortie numéro s du RNA peut être obtenue selon l'équation (2- 21):

$$y_s = f_s^S\left(\sum_{c=1}^{C} \omega_{c,s}^S f_c^C\left(\sum_{e=1}^{E} \omega_{e,c}^C u_e\right)\right) \qquad \textbf{(2- 21)}$$

Une description détaillée des RNA peut être consultée dans les ouvrages [59] et [63].

Par analogie avec ce qui a été présenté dans le paragraphe II.2.2.1, la construction d'un modèle d'un système avec les RNA consiste à identifier le « bon » vecteur $\boldsymbol{w}$ formé ici par l'ensemble des poids synaptiques. Mais avant cette étape il faut déterminer tous les éléments caractéristiques du réseau : vecteur de régression, nombre de couches, nombre de neurones et les FA dans chaque unité.

A l'état actuel il n'existe aucune méthode déterministe générique permettant de déterminer les caractéristiques optimales d'un réseau de neurones pour un problème donné. Les caractéristiques du modèle sont inhérentes au système étudié. Ainsi, les caractéristiques d'un RNA différent beaucoup d'un problème à un autre. Dans ce contexte, il devient primordial de

développer une méthodologie de construction de modèles neuronaux adaptée aux systèmes auxquels la thèse s'intéresse.

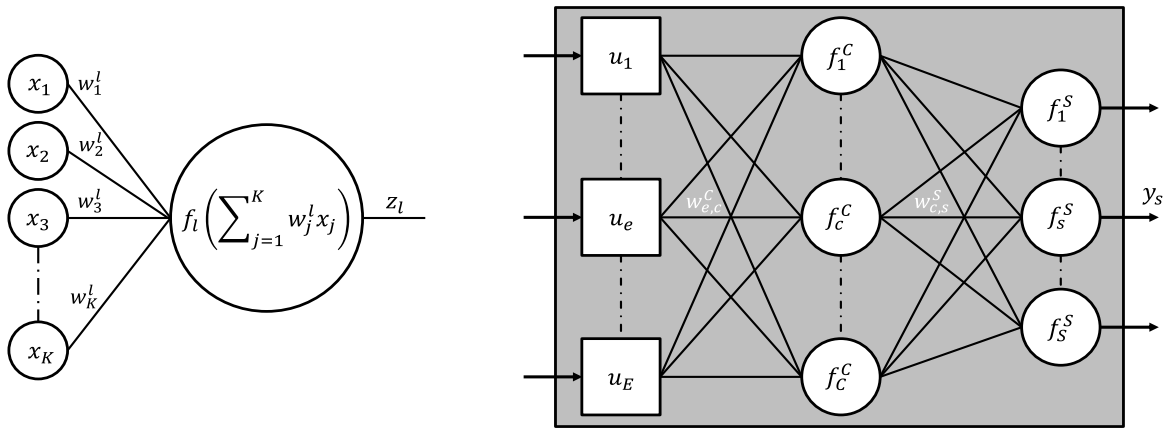

Figure 2- 7 : Représentation d'un neurone formel à gauche, exemple d'un réseau de neurones à une seule couche à droite

Dans les paragraphes suivants est présentée la démarche empruntée, étape par étape, pour déterminer un processus robuste pour la création des modèles neuronaux.

II.3.2. Caractéristiques du modèle neuronal

II.3.2.1 Fonctions d'activations

Selon les théorèmes d'approximation de fonctions cités dans le paragraphe II.2.2.4, une fonction d'activation non linéaire bornée, quelle qu'elle soit, dans les couches cachées et une fonction linéaire (fonction identité $f(x) = x$) dans la couche des sorties sont suffisantes. Aucune unanimité n'a été prononcée envers une fonction précise pour son utilisation [78]. Le choix de la fonction d'activation pourrait se faire en prenant en compte la sortie du système. En effet, si elle est positive par exemple, il pourrait être judicieux de choisir une FA dans la couche de sortie qui reste positive quelle que soit son entrée. Le choix de la FA peut également dépendre du traitement des données réalisé. Ce point est discuté plus loin dans le paragraphe II.3.4.2.

Dans le cadre de la thèse, la fonction tangente hyperbolique (fonction non linéaire et bornée, voir les théorèmes d'approximation de fonctions) a été retenue pour tous les neurones de la couche cachée. L'utilisation de neurones linéaires dans tout le RNA a été écartée. En effet, les systèmes étudiés sont complexes et présentent un comportement non-linéaire. Il faut en conséquence un modèle non-linéaire pour modéliser les non-linéarités dans les systèmes.

En fonction des éléments cités dans ce paragraphe le choix de la FA à utiliser dans la couche de sortie est plus délicat. La solution retenue consiste à tester deux FA : fonction linéaire et fonction tangente hyperbolique. Enfin, ces fonctions sont infiniment dérivables, il est donc possible d'appliquer les algorithmes d'optimisation classiques.

II.3.2.2 Entrées unitaires fixes

Il est toujours utile d'ajouter à l'entrée de chaque neurone (c'est à dire une FA), une valeur fixe équivaut à un multipliée par un poids synaptique. Cela est communément appelé un biais. En effet, cette entrée permet de déplacer, via le poids synaptique associé, la FA vers la

gauche ou vers la droite, ce qui donne plus de liberté pour le réseau et facilite ainsi l'apprentissage. L'intérêt d'utiliser le biais est justifié sur les Figure 2- 8 et Figure 2- 9. Par exemple, quel que soit le poids synaptique w la fonction $tanh(w.x)$ (un neurone) ne peut pas approximer la fonction bleue sur la Figure 2- 9 sans un biais.

L'équation (2- 22) devient dans ce cas :

$$y_s = tanh\left(b_s^S + \sum_{c=1}^{C} \omega_{c,s}^S \, tanh\left(b_c^C + \sum_{e=1}^{E} \omega_{e,c}^C u_e\right)\right) \qquad \textbf{(2- 22)}$$

Avec : b_c^C et b_s^S sont les poids synaptiques des biais pour la couche cachée et la couche de sortie respectivement.

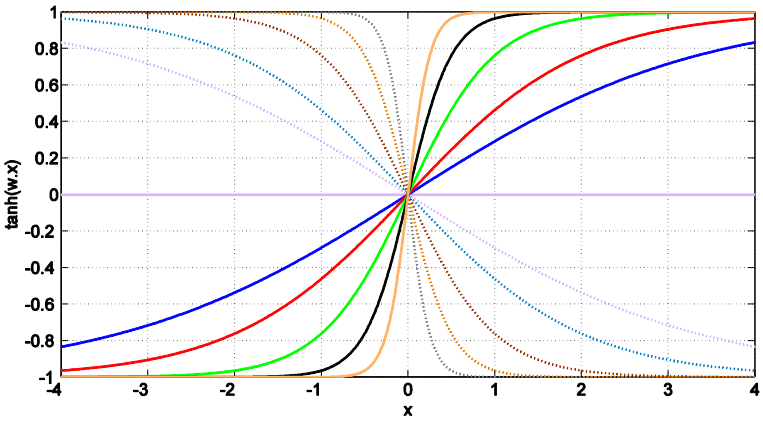

Figure 2- 8 : Représentation de la fonction « tanh » pour plusieurs valeurs de w (poids synaptiques). Quel que soit w la fonction $tanh(w.x)$ passe par l'origine.

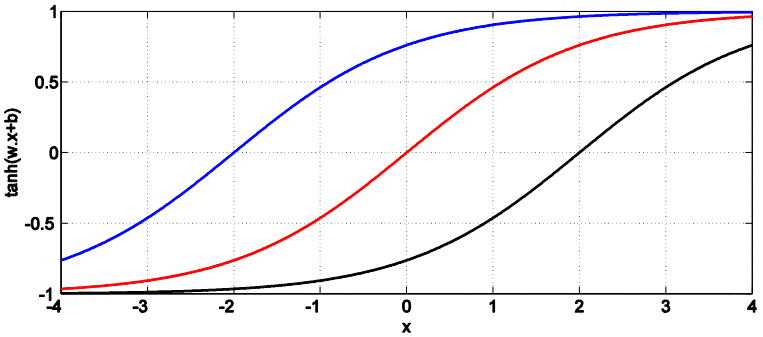

Figure 2- 9 : Illustration de l'ajout d'un biais dans la fonction « tanh » : possibilité de la translation à droite/gauche en outre de la rotation (effet du biais synaptique).

II.3.2.3 Configuration du modèle en entrainement et en exploitation

La question des entrées physiques utilisées pour la formation des modèles est abordée dans la section II.4. . Dans ce paragraphe nous nous intéressons uniquement aux dimensions du vecteur de régression.

Les modèles hypothèses non linéaires basés sur les modèles linéaires abordés précédemment (Tableau 2- 1) sont notés, dans le cadre de l'utilisation des RNA comme famille de fonction non linéaire, par l'ajout de « N » à l'abréviation de chaque structure. En effet, un modèle NARX est un modèle dont la fonction est réalisée par un RNA mais dont le vecteur de régression $\boldsymbol{\varphi}$ est celui d'un modèle ARX linéaire.

Les conditions du test des systèmes sont normalement bien maitrisées lors des essais en laboratoire. Elles sont donc répétables. En outre, des modèles physiques détaillés développés dans le logiciel commercial TRNSYS seront utilisées pour générer les données annuelles en vue d'une validation de la méthodologie. Ces modèles sont déterministes. Les données correspondantes ne sont donc pas entachées de perturbations. En conséquence nous ne considérons aucune entrée stochastique dans les modèles neuronaux qui seront développés. Le modèle hypothèse NARMAX a donc été écarté, d'autant plus qu'il a été signalé que ce type de modèle peut présenter des problèmes de stabilité [63].

Le modèle NFIR ne prend pas en compte la dynamique de la sortie du système. Or cette sortie peut apporter beaucoup d'information sur le fonctionnement du système. En effet, la consommation des systèmes thermiques pour les applications dans le bâtiment, dépend des entrées externes (conditions aux limites : bâtiment et climat) mais également de certaines variables internes tel que l'énergie stockée dans le ballon et de son inertie. L'accès à cette variable sera impossible au moment de l'exploitation du modèle. Une solution consiste à prendre en compte la sortie du modèle, retardée, en tant qu'entrée. A l'aide des données d'apprentissage, le modèle neuronal sera capable de remplacer le manque de connaissance sur les variables internes via la connaissance des sorties retardées. Enfin, seuls les modèles hypothèse NARX et NOE seront considérés dans ce manuscrit. Le modèle NARX peut être exploité de deux façons différentes. La première consiste à l'utiliser pour faire de la prédiction à court terme (utilisation en boucle ouverte) ou en simulation (utilisation en boucle fermée). C'est cette dernière utilisation qui nous intéresse dans le cadre de la méthodologie proposée. Quand le modèle NARX est entrainée en boucle fermée il devient équivalent au modèle NOE. Sur la Figure 2- 10 sont représentés graphiquement les deux modes d'utilisation du NARX.

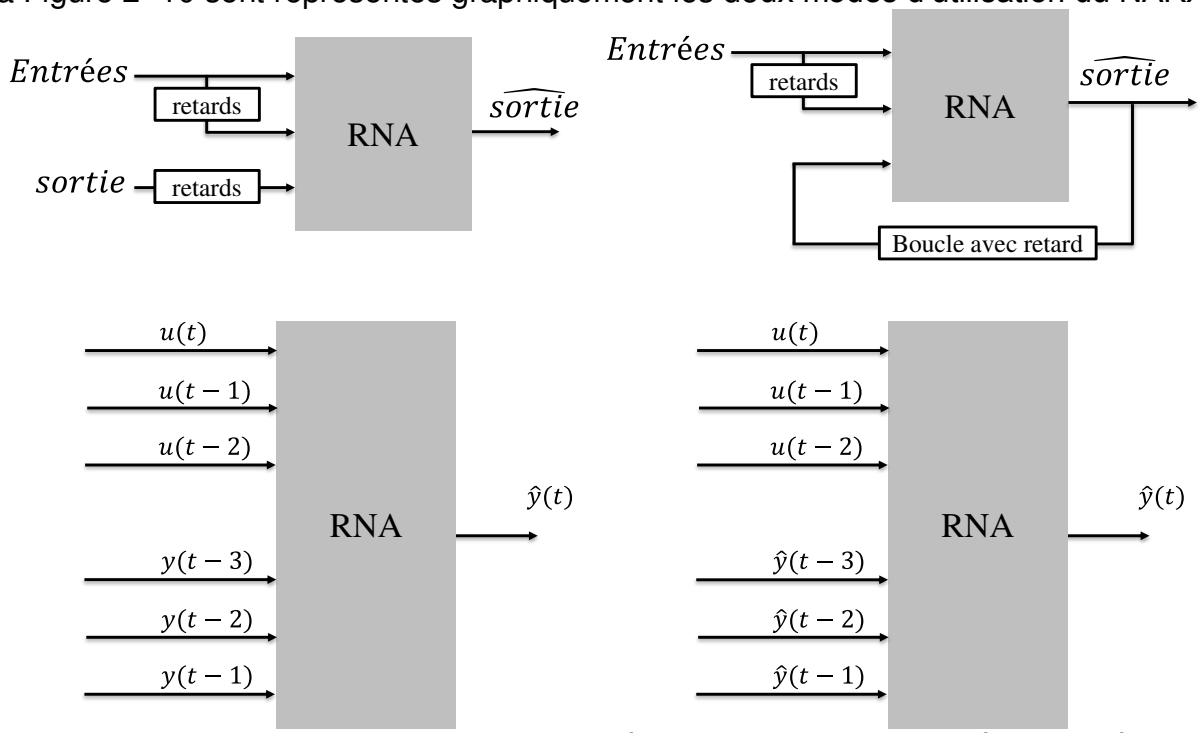

Figure 2- 10 : Configuration en boucle ouverte en haut à gauche, et en boucle fermée en haut à droite. Les deux figures en bas représentent un exemple dans le cas d'un Time Delay (TD) égal à 3 pas de temps et dans le cas d'une seule entrée et une seule sortie.

Durant l'entrainement toutes les données (entrées et sortie) sont disponibles. Il est donc avantageux d'entrainer les RNA selon la configuration en boucle ouverte (BO). Ainsi, durant

l'apprentissage le vecteur de régression du RNA est constitué, en outre des entrées, par des signaux retardés mesurés de la sortie du système, au lieu des signaux retardés estimés par le modèle lui-même (cas de la boucle fermée, voir également Figure 2- 10). Selon [79], [80] et [81] les paramètres du RNA seront ainsi identifiés avec des données précises puisqu'elles sont réelles (et non estimées par le RNA), ce qui va augmenter les chances de créer un modèle plus précis. Aussi, les contraintes de la méthodologie nous imposent une base de données assez réduite. Il n'est pas possible en conséquence de réserver un ensemble de données d'apprentissage à la validation des modèles RNA. Un entrainement en BO, selon la stratégie décrite dans le paragraphe II.3.3.4, puis une exploitation en boucle fermée (BF) permettra d'avoir un premier test du RNA. En effet, tous les modèles ayant de faibles performances en BF seront fortement pénalisés.

Les résultats de la comparaison entre un entrainement en BF et en BO présentés dans l'Annexe D. confirment ce constat. L'exploitation des modèles neuronaux et le processus de sélection seront réalisés en BF.

II.3.2.4 Ordres de régression

Après avoir sélectionné le modèle hypothèse à utiliser, il est également primordial de déterminer les ordres de régression : m, τ, ζ, dans le vecteur φ (voir Tableau 2- 1 paragraphe II.2.2.2). Ce sont ces paramètres qui déterminent quelles sont les entrées et sorties retardées à prendre en compte. En conséquence, ce sont ces paramètres qui déterminent, en partie, la complexité du modèle, mesurée par le nombre de paramètres à identifier. En effet, pour un RNA avec une seule couche cachée contenant C neurones, E entrées et une seule sortie, le nombre de paramètres dans le modèle est égale à : $(E+1).C+(C+1)$. Pour avoir un modèle parcimonieux il est important d'avoir le nombre minimal d'entrée permettant la description de la sortie. Trouver l'ordre de régression des entrées et sorties dans un modèle est une tâche particulièrement complexe notamment pour des données E/S issues d'un système dynamique non-linéaire. Plusieurs auteurs ont tenté de proposer une solution permettant de déterminer les bons ordres de régression. Un examen des méthodes existantes est donné dans [82]. Aucune de ces méthodes n'est générique et ne pourrait être appliquée à tous les problèmes.

Certains chercheurs utilisent des techniques empiriques consistant à tester plusieurs ordres de régression et d'essayer de choisir le meilleur selon un critère précis. D'autres utilisent des techniques plus élaborées pour les optimiser comme les algorithmes génétiques [83]. Pour des problèmes linéaires l'utilisation des fonctions d'autocorrélation ou d'inter-corrélation est très efficace pour la détermination des signaux retardés. Une fonction de corrélation mesure en quelque sorte le degré de similarité entre deux signaux en fonction du retard temporel de l'un sur l'autre. Le calcul de cette fonction permet donc de déterminer les potentiels signaux retardés (corrélés avec la variable à estimer) à prendre en compte dans la configuration E/S du modèle. Lorsque deux variables X et Y sont considérées comme deux réalisations de processus stationnaires (moyenne et variance constantes dans le temps) aléatoires de moyenne μ_X et d'écart-type σ_X (respectivement μ_Y et σ_Y), la fonction de corrélation croisée, entre deux variables X et Y est définie par:

$$R_{X,Y}(\tau) = \frac{E[(X(t)-\mu_X)*(Y(t-\tau)-\mu_Y)]}{\sigma_X.\sigma_Y} \qquad \textbf{(2- 23)}$$

La fonction $E[X]$ désigne ici l'espérance mathématique. La fonction d'autocorrélation d'une variable X est la corrélation croisée avec elle-même. Elle s'obtient en remplaçant Y par X. Il se peut que deux variables soient fortement corrélées, mais que cette corrélation soit due à l'influence d'un facteur extérieur et non pas à un fort lien entre les deux variables. La fonction d'autocorrélation partielle permet de calculer la corrélation entre $X(t)$ et $X(t-k)$ tout en faisant abstraction de l'influence des variables intermédiaires ($X(t-k+i)$ avec i variant entre 1 et $k-1$).

Dans le but de déterminer une stratégie d'identification des retards à prendre en compte dans la configuration des modèles une étude des corrélations entre les différentes variables d'un système solaire combiné a été réalisée. Les résultats sont donnés dans la Figure 2- 11 et la Figure 2- 12. La description du système ainsi que les résultats de sa modélisation sont présentés dans la section II.5. . Le choix des variables considérées dans cette partie est également justifié dans le paragraphe II.5.3. . Les droites en bleues délimitent l'intervalle de confiance à 95% qu'il n'y pas de corrélation au décalage temporel considéré, dans ce cas de figure les barres ne dépassent pas les droites. Les barres rouges dépassant les deux limites indiquent qu'il y a corrélation. Plus l'amplitude des barres est grande plus la corrélation est importante.

Sur la Figure 2- 11 sont tracées les différentes fonctions de corrélation entre la sortie (puissance de l'appoint) et les entrées (puissance des besoins d'ECS, des besoins de chauffage et de la ressource solaire) du système. Les deux premières figures font apparaitre des autocorrélations significatives sur un grand nombre de décalages temporels. Ceci montre que la puissance de l'appoint est instationnaire et qu'il est difficile en conséquence de déterminer une estimation fiable des retards à prendre en compte pour cette variable. La Figure 2- 12 montre également qu'il est difficile de tirer des conclusions fiables sur les bons décalages temporels à prendre en compte pour les variables d'entrées. Ces résultats confirment le fonctionnement non-linéaire des systèmes étudiés où les fonctions de corrélation linéaire trouvent leurs limites.

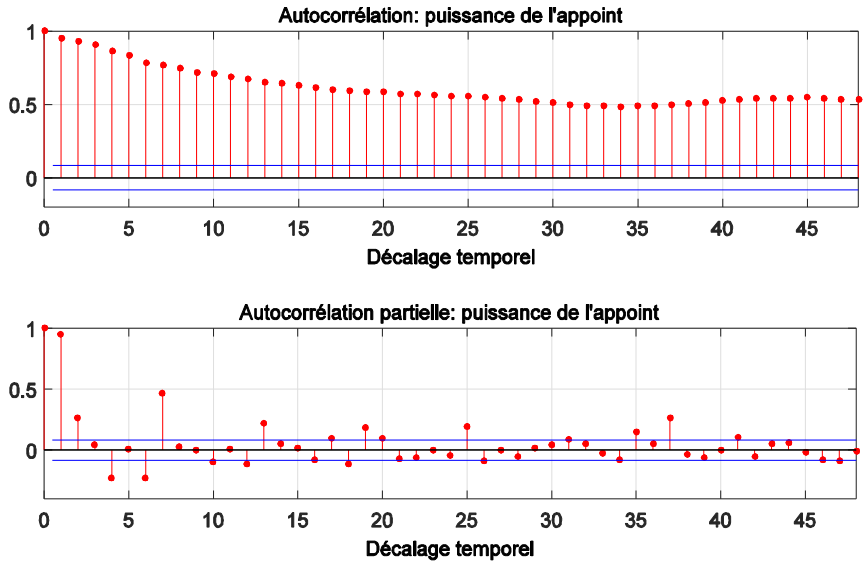

Figure 2- 11 : Corrélation pour la variable puissance de l'appoint (sortie du modèle).

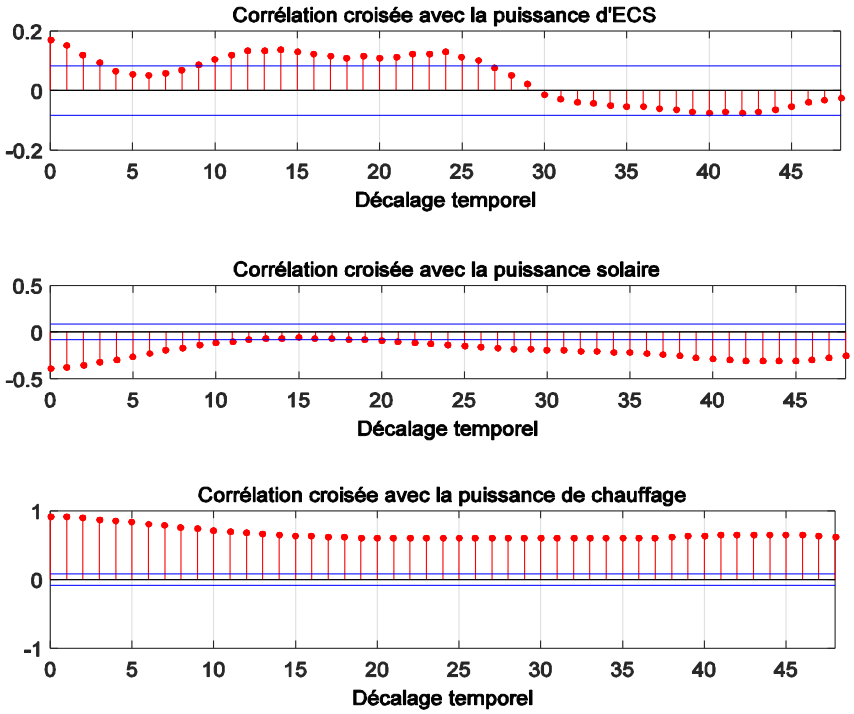

Figure 2- 12 : Corrélations croisées entre la sortie et les entrées.

Une méthode heuristique présentée dans [84] permet de calculer l'ordre de régression sur l'entrée et la sortie d'un système dynamique non-linéaire. Cette méthode est basée sur le calcul du coefficient de Lipschitz:

$$q^{(E)} = \left(\prod_{k=1}^{p_{lip}} \sqrt{E}\, q^{(E)}(k) \right)^{\frac{1}{p_{lip}}} \qquad \textbf{(2- 24)}$$

E représente le nombre total de variables d'entrées du modèle du système considéré, dont éventuellement les entrées retardées à considérer. p_{lip} est un nombre typiquement égal à $0{,}01 - 0{,}02E$. $q^{(E)}(k)$ est le k-ième plus grand quotient de Lipschitz (il est défini par l'équation (2- 25) parmi tous les quotients calculés avec les données entrées/sorties disponibles.

$$q_{ij} = \left| \frac{y(i) - y(j)}{u(i) - u(j)} \right|, (i \neq j) \qquad \textbf{(2- 25)}$$

Il a été montré que si E est le vrai nombre d'entrées du système (et donc l'ordre de régression optimal), q^{E} sera très proche de q^{E+1} par contre q^{E-1} sera très grand par rapport à q^{E}. De même q^{E+2} sera très proche de q^{E+1} mais q^{E-2} sera très grand par rapport à q^{E-1} et ainsi de suite. Le tracé de ces coefficients en fonction de E permettra de détecter à partir de quelle valeur de E ce coefficient se stabilise et ainsi pourra-t-on déterminer l'ordre de régression optimal. L'inconvénient de cette méthode réside dans le fait qu'elle est applicable seulement aux systèmes SISO (single input single output) et qui présentent un comportement continu et régulier. Dans notre cas de figure les systèmes sont de type MISO (multiple input single output) et intègrent des systèmes de régulation qui font que certaines variables ne sont pas régulières (mise en arrêt et en marche).

Parmi les modèles dynamiques rencontrés dans la littérature pour des applications ressemblant à celles de cette thèse ([85], [86], [87] et [88]), les auteurs supposent connaitre à l'avance l'ordre de régression tandis que dans [89] les auteurs ont utilisé le calcul des coefficients de corrélations linéaires. Selon l'étude réalisée et en prenant en compte les inconvénients et avantages de chaque méthode, il est judicieux de procéder autrement. La méthode que nous avons retenue pour la sélection des variables décalées dans le temps consiste à tester plusieurs pas de temps dans un ordre croissant puis sélectionner le modèle ayant les meilleures performances. Il est intéressant de signaler qu'une sous-estimation de l'ordre de régression dans le modèle neuronal implique un modèle incomplet car il va lui manquer au moins une variable d'entrée. Dans ce cas de figure l'apprentissage du RNA ne permettra pas d'atteindre les performances souhaitées. Cependant, une surestimation de l'ordre de régression pourrait être ajustée. En effet, le RNA pourrait s'ajuster à travers ses poids synaptiques pour prendre en compte uniquement les variables nécessaires pour calculer la sortie. Il suffit donc de tester un nombre suffisant de retards.

II.3.2.5 Nombre de neurones

Le nombre de neurones dans un RNA détermine les performances de celui-ci. Un nombre de neurones important est souvent la cause du problème de sur-apprentissage (sans parler du temps de calcul qui va s'accroitre) et inversement un réseau avec un nombre de neurones insuffisant ne sera pas capable de remonter à la relation reliant les entrées aux sorties (peu de paramètres est synonyme d'un degré de liberté réduit).

Plusieurs études ont été réalisées pour tenter de proposer une méthode pour déterminer le nombre optimal de neurones pour une architecture ou un modèle donné. Néanmoins aucune méthode n'est générique (voir [90] et [91]). Certains auteurs donnent ce nombre en fonction de la taille de l'ensemble de données disponibles pour l'apprentissage et le nombre d'entrées et de sorties [90]. D'autres essaient de donner une borne supérieure du nombre de neurones qui ne compromettent pas la capacité de généralisation du réseau [92]. Aucune de ces méthodes n'a réellement prouvé son efficacité et il est facile de trouver un contre-exemple pour chacune d'entre elles [93]. Certains auteurs proposent de mettre le nombre de neurones comme paramètre d'entrée d'un algorithme d'optimisation évolutionniste comme les algorithmes génétiques [91]. Cependant en utilisant ce type de méthodes, des problèmes similaires à ceux des RNA peuvent être rencontrés tels que le choix des paramètres de l'algorithme (la fonction objectif à minimiser, le nombre de population, les probabilités de mutation et de croisement etc.) et le temps de calcul souvent très lourd.

La méthode la plus sure et la plus simple (mais qui reste longue) consiste à entrainer plusieurs réseaux, d'architecture différentes, en partant d'un seul neurone dans la couche cachée, jusqu'à un nombre maximal de neurone, arbitrairement fixé. L'architecture optimale devra être déterminée dans un second temps, selon un critère à définir (ce point est traité dans le paragraphe II.5.4.). C'est cette approche qui est adoptée pour la méthodologie développée dans cette thèse.

II.3.2.6 Nombre de couches

Un RNA sans couche cachée est équivalent, dans le cas d'une FA linéaire, à un modèle linéaire. Il serait donc trop simple pour modéliser des systèmes complexes. L'utilisation de deux couches donnera plus de liberté au RNA pour approximer la fonction du système à

modéliser et pourrait en conséquence réduire le nombre de neurones utilisés. Un nombre de neurones réduit implique un RNA parcimonieux et potentiellement doté d'un pouvoir de généralisation important. Malgré cela, peu d'auteurs optent pour un réseau multicouche. En outre, il n'existe aucune méthode systématique pour déterminer le nombre de couches cachées [78]. Enfin selon le théorème cité dans le sous-paragraphe II.2.2.4, une seule couche cachée est théoriquement suffisante pour modéliser tout système. Le choix de considérer une seule couche cachée est pertinent et justifié. Ce choix permet également de réduire le nombre de cas à explorer et donc le temps de calcul.

II.3.2.7 Algorithmes de minimisation

Après la détermination de l'architecture du RNA, la dernière étape de son élaboration consiste à estimer ses paramètres. L'identification de ses poids synaptiques revient à minimiser une quantité dépendante de ses derniers.

Il existe deux catégories de méthodes pour trouver ce minimum. Les méthodes de la première catégorie sont en mesure de se rapprocher du minimum sans aucun calcul du gradient. Ils utilisent uniquement les valeurs de la fonction objectif. Les algorithmes génétiques et la méthode « simplex » sont deux exemples de cette catégorie. Les méthodes de la seconde sont basées sur le calcul direct ou indirect des dérivés de la fonction objectif. Les RNA étant des fonctions infiniment dérivables, seul les algorithmes appartenant à la deuxième catégorie ont été explorés.

Le développement en série de Taylor est généralement le point de départ de toutes les méthodes basées sur le gradient. En effet, le développement du second ordre en séries de Taylor au voisinage de $\boldsymbol{w}^*$ est :

$$Obj(\boldsymbol{w}; D_N) = Obj(\boldsymbol{w}^*; D_N) + (\boldsymbol{w} - \boldsymbol{w}^*)^T G_{grad}(\boldsymbol{w}^*) + \frac{1}{2}(\boldsymbol{w} - \boldsymbol{w}^*)^T H(\boldsymbol{w}^*)(\boldsymbol{w} - \boldsymbol{w}^*) \quad \textbf{(2- 26)}$$

Où G et H sont respectivement le vecteur gradient et la matrice hessienne de la fonction objectif.

Les conditions nécessaires et suffisantes pour que $\mathbf{w}^*$ soit une solution (le minimum de la fonction objectif) sont :

- le gradient en ce point est nul : $G_{grad}(\boldsymbol{w}^*) = 0$.
- la matrice hessienne est définie positive : $v^T H(\boldsymbol{w}^*) v > 0$ pour tout vecteur v non nul.

La recherche d'un minimum se fait en utilisant une méthode de recherche itérative. Le point de départ est une valeur arbitraire initiale $\boldsymbol{w}^{(0)}$. Les paramètres ($\boldsymbol{w}$) sont ensuite mis à jour à l'aide de la méthode utilisée selon la relation suivante:

$$\boldsymbol{w}^{(i+1)} = \boldsymbol{w}^{(i)} + \mu^{(i)} \boldsymbol{d}^{(i)} \quad \textbf{(2- 27)}$$

Où $\boldsymbol{w}^{(i)}$ représente l'itéré courant, $\mu^{(i)}$ la taille du pas de la mise à jour et $\boldsymbol{d}^{(i)}$ la direction de recherche. Le processus de recherche est ainsi répété jusqu'à ce qu'un critère d'arrêt soit satisfait.

Les algorithmes du gradient diffèrent les uns des autres dans la manière selon laquelle $\boldsymbol{d}^{(i)}$ est défini et calculé. Les méthodes de la descente du gradient utilisent l'opposé de la pente alors que les algorithmes inspirés de la méthode de Newton utilisent l'opposé du gradient multiplié par l'inverse de la matrice hessienne. La taille du pas de la mise à jour $\mu^{(i)}$ qui détermine de combien le vecteur de paramètres sera modifié, est généralement calculé en utilisant l'une des méthodes de recherche linéaire. Dans ce cas μ est mis à jour également à chaque itération. Dans les paragraphes suivants, les algorithmes les plus utilisés sont brièvement décrits.

Gradient descent with momentum (GDM) fait partie des méthodes de descente du gradient. Il a été développé afin de résoudre certains problèmes de convergence des algorithmes classiques. En utilisant cet algorithme, les poids synaptiques sont mis à jour comme suit :

$$\boldsymbol{w}^{(i+1)} = \boldsymbol{w}^{(i)} - \mu\, G_{grad}\left(\boldsymbol{w}^{(i)}\right) + mo\,\left(\boldsymbol{w}^{(i)} - \boldsymbol{w}^{(i-1)}\right) \qquad \textbf{(2- 28)}$$

Où mo est appelé moment.

GDM est sensible non seulement à la direction du gradient de l'itération courante, mais aussi à l'amplitude de la variation de $\boldsymbol{w}$ entre deux itérations successives. Cela permet à l'algorithme de détecter les faibles changements dans la surface de la fonction objectif afin de les ignorer et éviter d'être bloqué dans des minimums locaux [94]. Il a été montré aussi que l'ajout du terme du moment augmente le taux de convergence de façon spectaculaire [95].

Scaled gradient conjugué (SCG) est aussi un algorithme basé sur le calcul du gradient mais pour lequel la mise à jour des poids synaptiques est réalisée en utilisant les directions conjuguées qui forment une base de l'espace de recherche. Cet algorithme n'effectue pas une recherche linéaire par itération. Il est entièrement automatisé. L'utilisateur n'a besoin de spécifier aucun paramètre important [96]. Il existe une autre variante de cette méthode décrite dans [97].

L'algorithme de *Levenberg-Marquardt* (LM) [98] est une combinaison entre la célèbre méthode de Newton et celle de la descente du gradient. Il a été développé pour surmonter le calcul de la matrice hessienne très coûteuse à calculer. L'algorithme de Levenberg-Marquardt utilise une approximation de la matrice hessienne $\widetilde{H}$ pour la mise à jour des paramètres. La relation de mise à jour est semblable à celle de Newton:

$$\boldsymbol{w}^{(i+1)} = \boldsymbol{w}^{(i)} - \left[\widetilde{H}(\boldsymbol{w}^{(i)}) + \mu^{(i)}\, I\right]^{-1} G_{grad}(\boldsymbol{w}^{(i)}) \qquad \textbf{(2- 29)}$$

La valeur initiale de μ doit être indiquée par l'utilisateur. Lorsque le scalaire $\mu^{(i)}$ est nul, la relation dans l'équation (2- 29) n'est autre que la méthode de Newton, en utilisant la matrice approximative $\widetilde{H}$. Lorsque $\mu^{(i)}$ est grand, la même relation devient équivalente à celle d'un algorithme de descente du gradient avec un pas très faible. La méthode de Newton étant plus rapide et plus précise à proximité d'un minimum, l'astuce de LM est donc de basculer vers la méthode de Newton aussi rapidement que possible.

Rétropropagation résilient (RP) est un autre algorithme du gradient mais qui met à jour les paramètres en utilisant uniquement le signe du gradient. La valeur de mise à jour pour chaque paramètre est augmentée d'un facteur prédéfini par l'utilisateur à chaque fois que la dérivée de la fonction objectif par rapport à ce poids a le même signe pour deux itérations successives. Si

le cas inverse se présente la valeur de mise à jour est diminuée par un facteur également prédéfini. Si la dérivée est nulle, la valeur de la mise à jour reste inchangée. Si les paramètres oscillent, l'amplitude de mise à jour est réduite. Dans le cas où ils continuent d'évoluer dans la même direction pendant plusieurs itérations, l'amplitude de mise à jour est augmentée. Une description complète de cet algorithme est donnée dans [99].

L'algorithme d'optimisation qui a été retenu pour la méthodologie est celui de Levenberg-Marquardt qui, par sa stabilité et sa notoriété, est un algorithme relativement fiable. Il présente également l'avantage de trouver une solution même avec un point de départ loin de l'extrémum de la fonction à minimiser. Une étude comparative [100] entre les 4 algorithmes présentés dans ce paragraphe a été réalisée. Il apparait selon l'étude que les deux algorithmes LM et RP sont les plus performants. Ce résultat confirme notre choix.

II.3.2.8 Réinitialisation

L'algorithme LM reste un algorithme local malgré sa capacité à éviter certains minimums locaux. Dans le cas des fonctions objectifs non-linéaires complexes présentant beaucoup d'extrémums, l'algorithme LM peut être piégé par un minimum local et ainsi perturber le processus d'apprentissage. Un moyen efficace pour pallier ce problème consiste à appliquer le processus de minimisation de la fonction objective plusieurs fois en commençant à chaque fois d'un point de départ différent (c'est la réinitialisation). Cependant ces réinitialisations doivent être efficaces et choisis de telle sorte qu'elles soient un bon point de départ. C'est pour cette raison qu'il est indispensable de suivre une stratégie permettant de générer des poids synaptiques initiaux balayant tout l'espace du vecteur $\boldsymbol{w}$. L'algorithme développé par Nguyen-Widrow [101] permet d'atteindre cet objectif. Il a été montré qu'il est meilleur qu'une initialisation arbitraire aléatoire [101].

II.3.3. Techniques d'apprentissage

L'objectif d'algorithmes d'optimisation est de déterminer les poids du RNA qui minimisent la fonction de l'erreur sur la base des données d'apprentissage (observations). Très souvent, les modèles qui résultent de ce processus de minimisation ont de très bonnes performances sur l'ensemble de données d'apprentissage. Cependant, la précision de la prédiction du modèle peut être très mauvaise pour des données nouvelles. Dans certains cas, cela signifie que le modèle résultant a appris seulement les caractéristiques de l'ensemble des données d'apprentissage et non pas le comportement du système. Dans d'autres cas, il se peut que l'algorithme ait été piégé dans un minimum local. Les stratégies d'apprentissage sont des moyens pour piloter le processus de minimisation ou la structure de réseau neuronal afin d'obtenir un modèle qui ait une bonne capacité de généralisation, c'est-à-dire qui fournisse des estimations précises même pour de nouvelles données. Quatre stratégies d'apprentissage ont été explorées dans le but de déterminer celle qui correspond le mieux à notre problématique.

II.3.3.1 L'élagage des connexions (NNP)

L'élagage des connexions (NNP) est une technique qui consiste à former un réseau avec un grand nombre d'entrées et de neurones dans la couche cachée, puis éliminer, progressivement, les connexions qui sont superflues. La méthode utilise un critère d'augmentation minimale de l'erreur au cours de l'apprentissage pour l'élimination des connexions (rendre égal à zéro le poids synaptique correspondant). Ce critère est appelé

saillance et mesure l'augmentation de l'erreur qui se produit lorsque le poids est éliminée ([102], [63]). Par conséquent les connexions à faible saillance sont supprimées. A la fin du processus d'élagage, le RNA résultant est ré-entrainé. Pour cette méthode il est possible de diviser la base de données en deux ensembles et d'utiliser l'un d'eux pour calculer la saillance et l'autre pour la minimisation. L'architecture du RNA résultante est ensuite ré-entrainé sur la totalité de la base. L'application de cette stratégie d'apprentissage dans le cadre de la thèse a révélé un inconvénient de cette dernière dans le cas d'un problème physique [100]. En effet, sur la Figure 2- 13 est représentée l'architecture du modèle résultant après avoir appliqué cet algorithme sur une base de données du système décrit dans le paragraphe II.5.1. . Cette figure montre que l'algorithme a supprimé beaucoup de connexions y compris celles de l'entrée correspondante à l'ensoleillement ($A.G$). Selon le modèle, la ressource solaire n'a donc aucune influence sur la consommation du système, ce qui n'est pas vrai physiquement. Il est donc difficile d'expliquer ce type de résultats. En conséquence malgré son grand intérêt (il est capable de gérer en même temps l'apprentissage, la sélection du nombre de neurones, couches etc.), cette stratégie d'apprentissage ne sera pas retenue pour notre application.

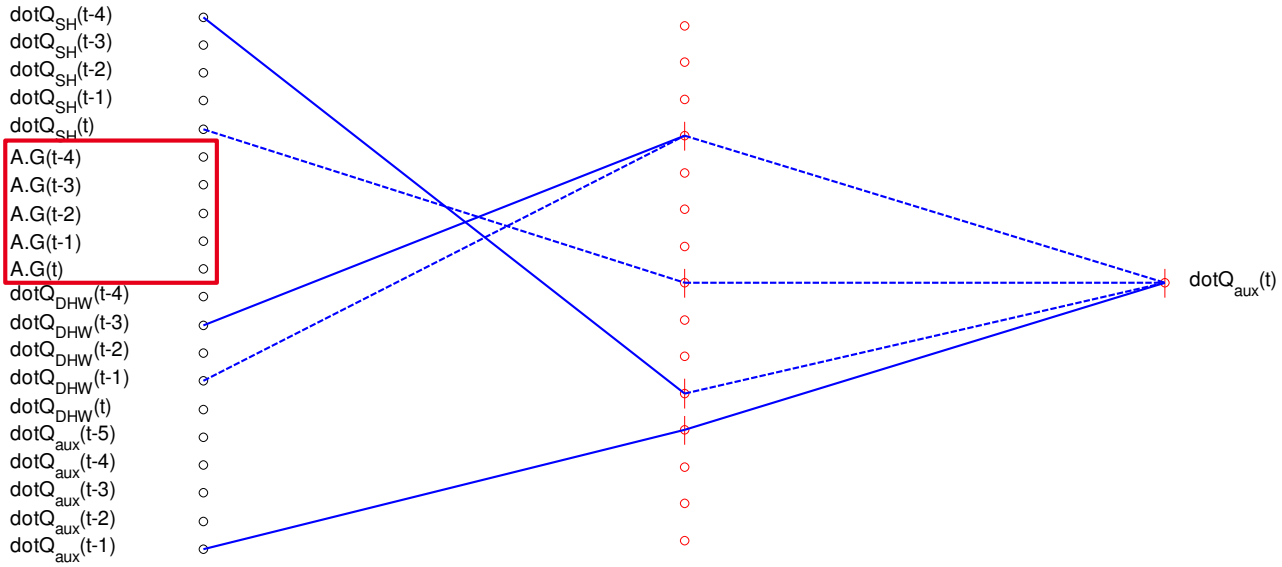

Figure 2- 13 : L'architecture du RNA résultante après avoir appliqué l'algorithme de l'élagage des connexions. Le symbole « t » désigne le temps courant et les chiffres désignent le nombre de pas de temps utilisés pour prendre en compte les variables aux instants retardés. dotQ$_{SH}$, dotQ$_{DHW}$ et dotQ$_{aux}$ désignent ici respectivement les puissances des besoins de chauffage, d'ECS et de la consommation du système.

II.3.3.2 L'arrêt prématuré de l'apprentissage (ES)

L'arrêt prématuré de l'apprentissage (ES) est la stratégie la plus utilisée. Elle consiste à diviser les données en deux ensembles. Le premier, appelé ensemble d'apprentissage, servira à mettre à jour les paramètres en utilisant un algorithme de minimisation et le second, appelé ensemble de validation, servira à évaluer l'entrainement du modèle. En effet, au cours de chaque itération, l'erreur quadratique moyenne (EMQ, la fonction objective dans l'équation (2-9)) est calculée pour les deux ensembles de données. Normalement, l'EMQ pour les données d'entrainement diminue progressivement au fur et à mesure de l'avancement de l'algorithme d'optimisation. Par contre, ce n'est pas toujours le cas pour l'EMQ sur la base de validation.

Lorsque celui-ci commence à augmenter, cela signifie que le RNA commence à être surentrainé. Le processus d'apprentissage doit être arrêté, malgré que le minimum n'ait pas encore été atteint sur l'ensemble d'apprentissage.

L'inconvénient de cette méthode réside dans le fait qu'il est indispensable pour pouvoir l'appliquer d'avoir suffisamment de données pour l'apprentissage. Or, dans certains cas, les données mesurées sont très coûteuses à récolter. Les données ne sont pas forcément suffisantes pour permettre la création d'une base de validation en plus d'une base d'apprentissage. Par exemple, dans notre cas de figure, le peu de temps de l'essai du système, qui est de 12 jours, limite la quantité de données disponibles pour l'apprentissage. Cette méthode ne sera donc pas retenue.

II.3.3.3 La k-validation croisée (KFCV)

La k-validation croisée (KFCV) est utile pour déterminer une structure potentiellement optimale du RNA à utiliser pour modéliser un système. Pour une structure donnée, ce procédé consiste à diviser les données en k sous-ensembles. Chaque sous ensemble est appelé « pli ». Parmi les k sous-ensembles, un seul ensemble est attribué en tant qu'ensemble de validation (similairement à la technique ES), et les $k-1$ restant sont utilisés comme données d'entraînement. Le RNA est entrainé jusqu'à la convergence de l'algorithme de minimisation à un minimum. Après cette étape, une estimation de la capacité de généralisation du modèle est mesurée sur la base de l'ensemble de validation. Le processus de KFCV est ensuite répété k fois, pour qu'en fin de compte, chaque sous-ensemble ait été utilisé exactement une fois comme ensemble de validation. Les k résultats des plis sont ensuite moyennés pour produire une seule estimation de la capacité de généralisation du modèle pour la structure courante du modèle. La structure qui présente la meilleure capacité de généralisation moyenne est ensuite entrainée sur l'ensemble des données disponibles jusqu'à convergence. Cette méthode se distingue donc de l'ES par sa capacité à être appliquée même en cas d'une base de données assez réduite. Cependant, pour un RNA dynamique contenant des sorties bouclées, les données d'apprentissage ne sont pas indépendantes. La division des données en plusieurs sous-ensembles n'est pas possible sans dégrader cette dépendance. La k-validation croisée est donc valable uniquement au modèle statique. Une tentative d'adaptation de cette stratégie au cas dynamique a pourtant été explorée dans [100].

II.3.3.4 La régularisation

L'apprentissage par régularisation peut utiliser un ensemble de données restreint sans compromettre la capacité de généralisation du modèle. C'est pour cette raison qu'elle a été retenue pour le développement de la méthodologie sujet de ce projet de thèse. Typiquement, l'entrainement vise à minimiser EMQ. La régularisation modifie la fonction objectif en y ajoutant un terme supplémentaire: la moyenne de la somme des carrés des poids du réseau :

$$SCP = \frac{1}{N}\boldsymbol{w}^T\boldsymbol{w} \qquad \textbf{(2- 30)}$$

La fonction objectif devient :

$$Obj(\boldsymbol{w}; D_N) = \alpha\, EMQ(\boldsymbol{w}; D_N) + \delta\, SCP(\boldsymbol{w}) \qquad \textbf{(2- 31)}$$

Où α et δ sont des paramètres constants calculés selon une méthode de régularisation bayésienne (BR). Quand α est égal à 1 et δ est choisi par l'utilisateur, l'apprentissage est appelé en « *weight decay* ».

En limitant l'amplitude des poids synaptiques, le processus d'apprentissage produit un RNA avec une bonne capacité de généralisation [103]. En gardant les poids synaptiques avec de faibles amplitudes, la fonction RNA reste lisse ce qui permet d'éviter le surentrainement.

II.3.4. Données d'apprentissage

II.3.4.1 Séquence de test

La qualité des données d'apprentissage occupe une place capitale dans l'élaboration de modèles pertinents. Une base de données pour l'apprentissage non représentative du domaine de variation des entrées du modèle et des non-linéarités du système complexifie beaucoup l'apprentissage. La capacité de généralisation du RNA pour des données non vues sera dans ce cas très faible. Pour des problèmes statiques, il existe des méthodes d'échantillonnage permettant de former une base de données riche en information (Latin hyper-cube etc.). Cette représentativité des données est encore plus demandée et difficile à acquérir dans le cas d'un modèle dynamique ayant comme entrée une sortie bouclée. En effet, cette dernière est inconnue et conditionnée par les entrées. Il devient donc impossible de l'échantillonner. La séquence de test décrite dans le premier chapitre prend en compte le fonctionnement dynamique et physique du système. Elle est donc plus adaptée pour le développement de la méthodologie d'évaluation des systèmes solaires thermiques. Dans l'annexe Annexe B. une possibilité d'amélioration de cette séquence a été également explorée.

II.3.4.2 Traitement de données

Quelle que soit le pas de temps d'enregistrement des données pendant le test du système considéré selon la méthode SCSPT (il est généralement de l'ordre de la minute), celle-ci est remise à un pas de temps de 30 minutes, en prenant les moyennes des différentes valeurs sur cet intervalle temporaire. En considérant l'inertie du ballon de stockage (l'élément présentant le plus d'inertie dans le système), ce pas de temps est suffisamment court pour traduire correctement le comportement dynamique du système. Il est à noter que nous ne cherchons pas ici à reproduire exactement les ordres du système de contrôle qui peuvent agir à haute fréquence sur les systèmes mais uniquement à reproduire les échanges de flux de chaleurs entre leurs composants. Comme il a été précisé précédemment, plusieurs modèles seront créés en modifiant les différents éléments du RNA. Ce choix du pas de temps aura donc l'avantage de réduire le temps de calcul global.

L'utilisation des données brutes pour l'apprentissage des modèles neuronaux est déconseillée même si, théoriquement, le RNA est capable de s'adapter à tout type de données. Un traitement des données d'entrainements est indispensable [78]. En effet, les variables d'entrées peuvent varier dans des intervalles assez différents. Puisque les ordres de grandeurs des variables ne sont pas les mêmes, elles ne seront pas équitablement prises en compte dans la fonction objective EMQ. Les variables de grande amplitude auront donc plus d'influence dans la mise à jour des poids synaptiques. Pour prévenir ce problème, mais

également pour faciliter le travail des algorithmes d'optimisation [104], les données d'apprentissage seront normalisées dans un intervalle unique selon la formule suivante :

$$X = \frac{X_0 - X_{min}}{X_{max} - X_{min}}(nor_{max} - nor_{min}) + nor_{min} \quad \textbf{(2- 32)}$$

Où $[nor_{min}, nor_{max}]$ est l'intervalle de normalisation, X_0 la variable brute et X la même variable ramenée à une variation dans $[nor_{min}, nor_{max}]$. X_{min} et X_{max} sont les valeurs minimal et maximal de X_0.

II.3.5. Processus complet d'entrainement et sélection du modèle

Il est primordial de disposer d'un moyen efficace pour sélectionner le RNA optimal parmi ceux qui seront créés pendant le processus de la modélisation. L'optimalité d'un RNA concerne, évidement, le pouvoir de généralisation du modèle. Plusieurs critères ont été utilisés dans la littérature pour évaluer les performances d'un RNA donné [78], [105]. Plusieurs critères ont été analysés dans le cadre de la thèse et c'est le critère d'information bayésienne qui a été retenu car il est le plus robuste par rapport aux autres. Les résultats de cette étude seront présentés plus loin dans le paragraphe II.5.4.2.

Le processus complet résumant les différentes étapes de conception du modèle neuronal, dans le cadre de la méthodologie proposée, est représenté sur la Figure 2- 14. Ce processus consiste à créer plusieurs modèles neuronaux avec différentes caractéristiques : FA dans la couche de sortie, intervalle de normalisation, nombre de signaux retardés considérés TD (Time Delay), nombre de neurones dans la couche cachée et le numéro de la réinitialisation. Pour un retard TD donné le vecteur d'entrée sera : $[\boldsymbol{u}(t), \boldsymbol{u}(t-1), \dots, \boldsymbol{u}(t-TD+1), y(t-1), y(t-2), \dots, y(t-TD)]$, cela revient à prendre $\zeta = 0$, $m = TD - 1$ et $\tau = TD$ dans la colonne du modèle Autorégressif à variable eXogène du Tableau 2- 1 (voir également un exemple sur la Figure 2- 10).

Chaque modèle doit donc subir au travers ce processus les étapes suivantes :

- Entrainement en boucle ouverte
- Test en boucle fermée : la sortie estimée par le modèle est calculée en boucle fermée avec les données d'apprentissage.
- Calcul du critère de sélection.
- Enregistrement du modèle.

A l'issue de ces étapes un seul modèle sera sélectionné, c'est le modèle minimisant le critère d'information bayésienne.

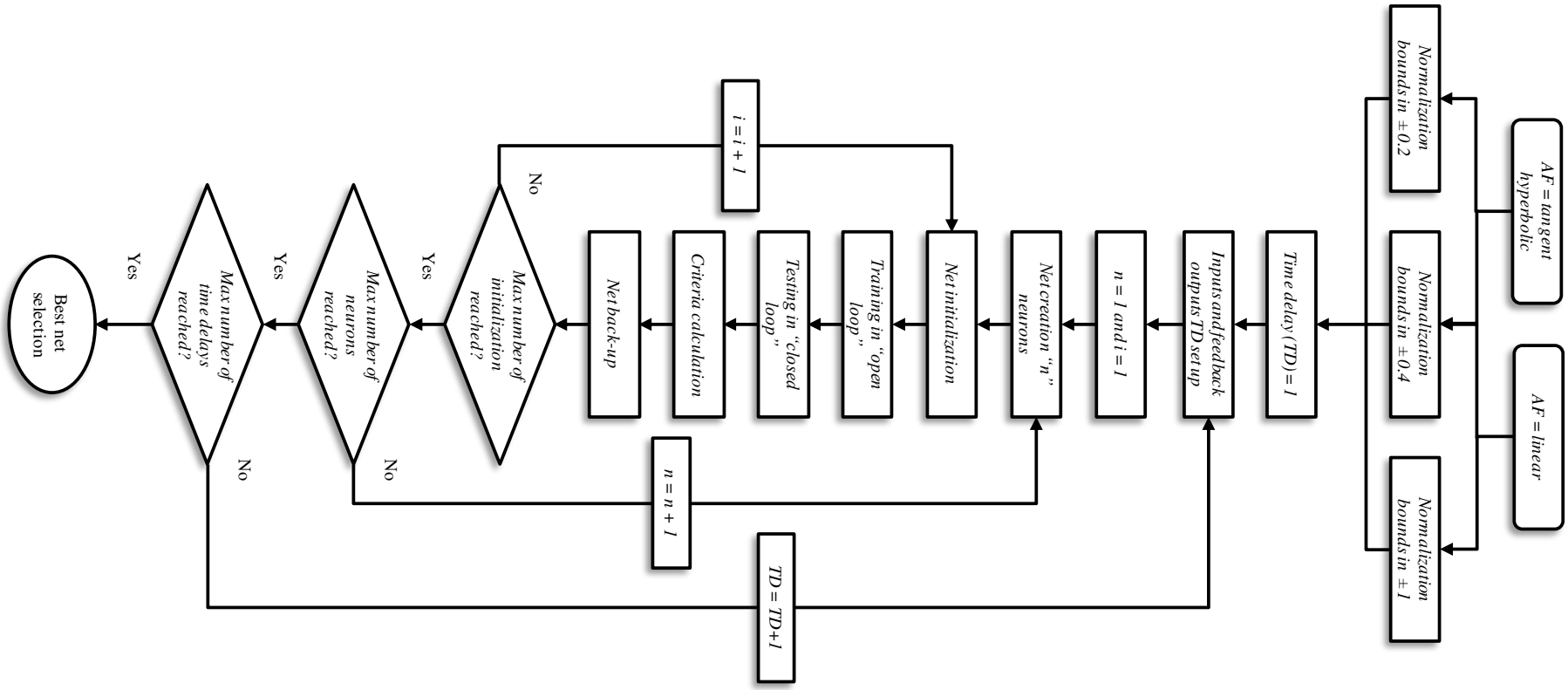

Figure 2- 14 : Processus complet de l'entrainement et de la sélection des modèles RNA

II.4. Configuration E/S et évaluation de la méthodologie

II.4.1. Les deux approches : « E/S physiques » et « flux énergétiques »

Pour modéliser les SSC selon l'approche « boîte noire » qui suppose qu'on n'ait pas accès aux paramètres internes, deux configurations ont été définies.

La première approche consiste à considérer toutes ses entrées et sorties physiques à savoir le rayonnement solaire, les températures du fluide caloporteur à ses extrémités qui le séparent de son environnement ainsi que le débit dans chaque circuit : chauffage, rafraichissement et eau chaude sanitaire. Les températures extérieures et ambiantes sont aussi à prendre en compte car elles sont des entrées pour le système de régulation. Les limites du système à modéliser selon la première approche sont représentées sur la Figure 2-15.

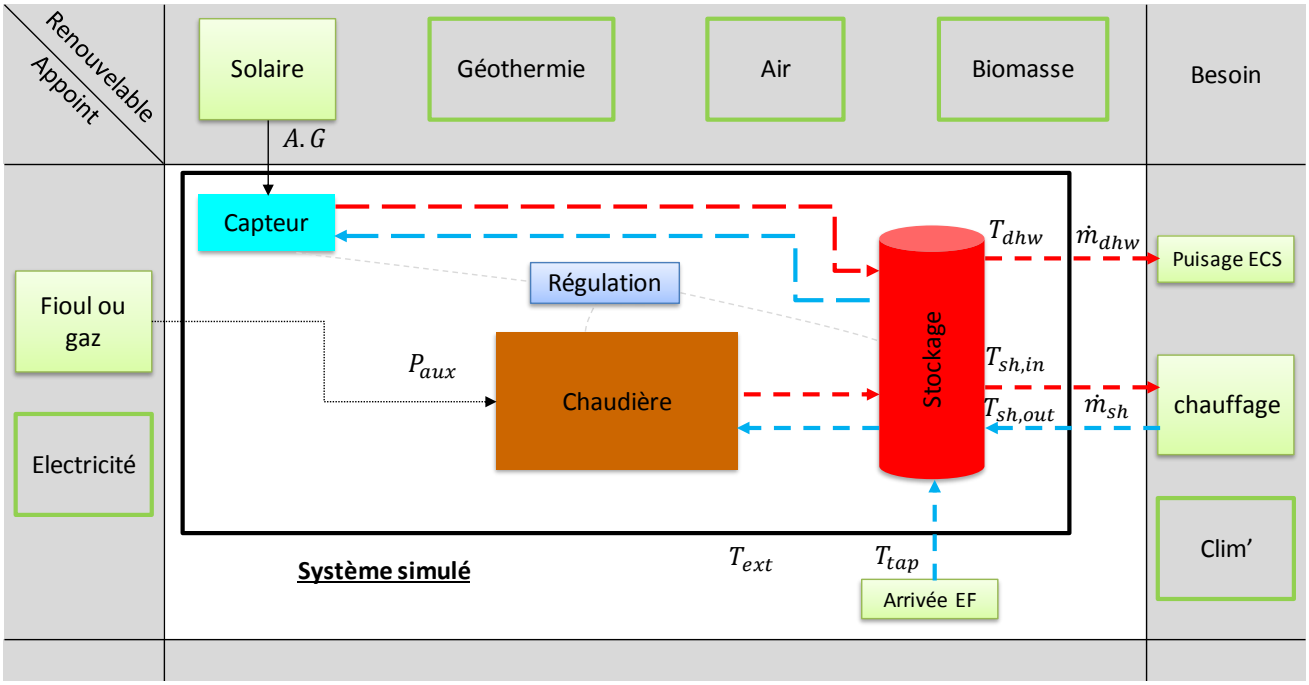

Figure 2- 15 : Mise en évidence des limites du système et des grandeurs échangées avec son environnement selon l'approche « entrées-sorties physiques ». Cas du SSC avec un appoint par une chaudière à gaz

Le modèle entrées/sorties à considérer est donc celui représenté sur la Figure 2- 16.

Cette approche correspond parfaitement à la représentation réelle du système. Aussi, une telle configuration, si elle fournit des résultats satisfaisants, donnerait non seulement accès à une estimation de la puissance de la chaudière mais aussi à celles des différentes températures et débits dans les circuits du système. La connaissance de ces variables est précieuse pour l'optimisation ainsi que le dimensionnement des systèmes. Nous allons donc la considérer comme une potentielle solution à étudier parmi d'autres.

La configuration selon l'approche « entrées-sorties physiques » mettant en jeu plusieurs variables est complexe. En effet, un nombre important de variables dans le modèle est synonyme d'un nombre important de paramètres à identifier ce qui nécessitera normalement une base de données conséquente alors que la durée du test dans le cadre de la méthodologie est limitée.

Il est à rappeler ici que la température d'ECS est variable (comme pour la température intérieure au bâtiment) malgré que la consigne soit constante. En effet, le système solaire thermique, supervisé par son système de régulation, tente de respecter cette consigne en fonction des conditions aux limites qui sont variables dans le temps.

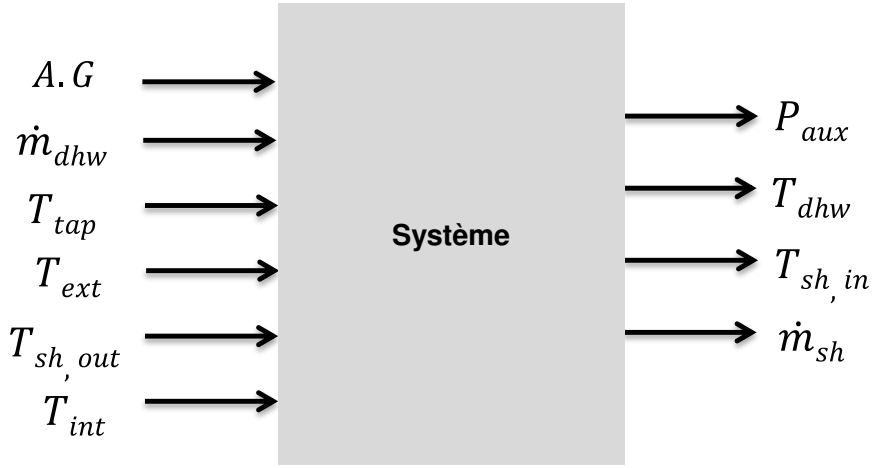

Figure 2- 16 : Configuration entrées-sortie du modèle selon l'approche E/S physiques. Cas du SSC avec un appoint par une chaudière à gaz

Pour avoir un modèle plus compact que le précédent, la deuxième approche considère les flux d'énergie entre le système et son environnement, à savoir : les puissances de chauffage, de climatisation et d'eau chaude sanitaire, le rayonnement solaire à la surface des capteurs et la puissance de la chaudière. Similairement au cas précèdent, les températures intérieure et extérieure sont à prendre en compte également pour le système de régulation (Figure 2- 17 et Figure 2- 18).

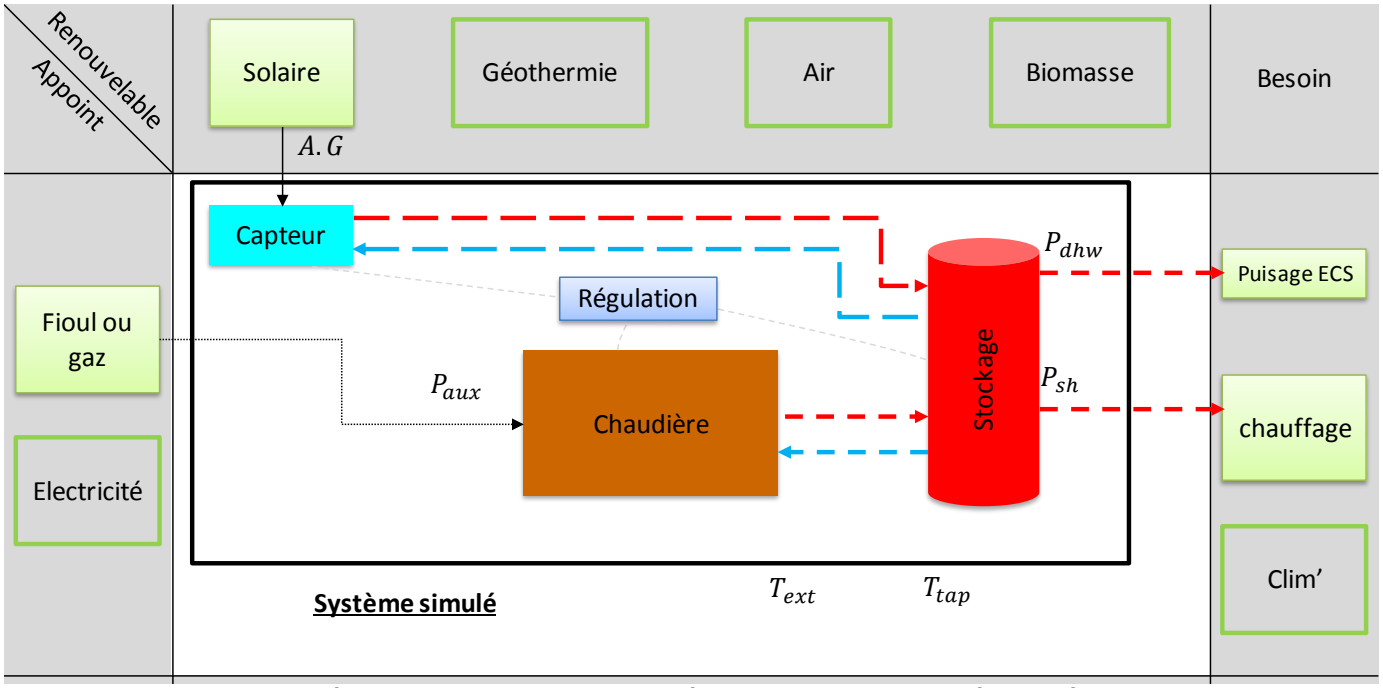

Figure 2- 17 : Mise en évidence des limites du système et des grandeurs échangées avec son environnement selon l'approche flux énergétiques. Cas du SSC avec un appoint par une chaudière à gaz.

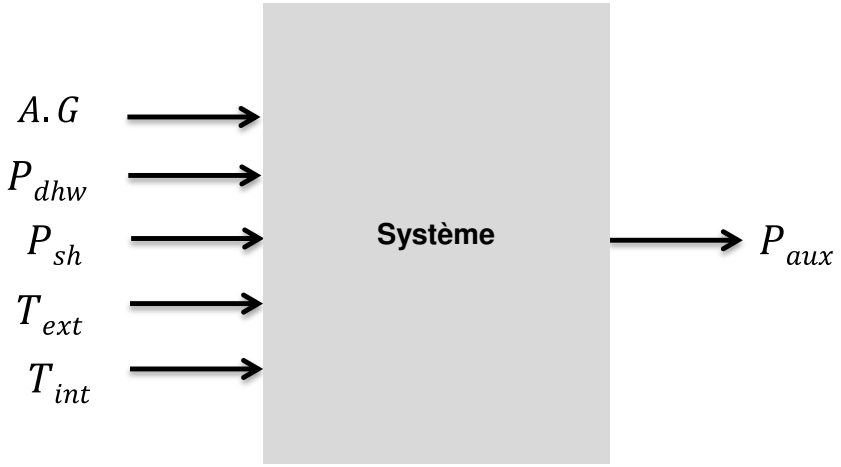

Figure 2- 18 : Configuration entrées-sortie du modèle selon l'approche flux énergétiques. Cas du SSC avec un appoint par une chaudière à gaz

Certes la représentation du système selon la configuration en « flux énergétiques » nous fait perdre en qualité d'information en ce qui concerne son fonctionnement interne. Par exemple, à puissance égale, un puisage à fort débit n'aura pas le même effet sur l'état du stockage qu'un puisage à faible débit. Ce constat est vrai également pour les besoins de chauffage. Deux états internes du stockage impliquent forcément des pertes et des stratifications différentes dans le ballon, ce qui va influencer la puissance de l'appoint. Cependant, la deuxième configuration est plus simple par rapport à la première dans la mesure où elle fait intervenir moins de variables. Elle est aussi parfaitement compatible avec l'objectif de caractérisation puisqu'en sortie du modèle nous aurons accès à la puissance fournie par la chaudière. Nous ne cherchons pas à avoir un modèle reproduisant en détail le comportement du système. L'estimation globale de l'énergie consommée sur une longue période est la seule grandeur qui nous intéresse pour pouvoir donner une garantie de performance (contexte de la thèse présenté dans le premier chapitre). Aussi, nous supposons que les puisages réels d'ECS restent plus ou moins réguliers dans le cadre des systèmes thermiques pour l'habitat.

Une troisième configuration dérivée de la deuxième est également à considérer. Elle consiste à garder uniquement les puissances. Cette configuration présente l'avantage d'être encore plus compacte tout en gardant une relation physique réelle entre la sortie (consommation de l'appoint) et les entrées (ressource renouvelable, besoin de chauffage/climatisation et d'ECS). En outre au moment de l'exploitation du modèle aucune information sur la température intérieure au bâtiment ne sera disponible à moins d'intégrer un modèle du bâtiment au modèle RNA, ce qui n'est pas envisagé dans le cadre de cette méthodologie.

II.4.2. Evaluation de la méthodologie

II.4.2.1 Protocole de validation

Dans le but de valider rigoureusement la méthodologie proposée, celle-ci doit être appliquée à différents systèmes et dans différentes conditions. Les estimations de consommation annuelle des systèmes calculées par les modèles neuronaux doivent être ensuite comparées aux consommations mesurées. Une telle validation suppose d'avoir réalisé un suivi *in situ* de plusieurs systèmes en différents sites climatiques et pour différents

bâtiments. Or ceci est très couteux et difficilement réalisable en trois années. L'approche que nous avons retenue pour valider la méthodologie développée dans la thèse est constituée de trois éléments :

- Une validation numérique (expérimentation numérique) qui consiste à utiliser un modèle physique détaillé du système solaire thermique pour réaliser à la fois une simulation du test physique du système dans le but de créer son modèle neuronal et des simulations annuelles dans différentes conditions. Le modèle neuronal développé sera ensuite utilisé pour prédire les consommations annuelles du système dans différentes conditions qui seront ensuite comparées aux simulations du modèle physique. Cette approche de validation est représentée dans la Figure 2- 19.

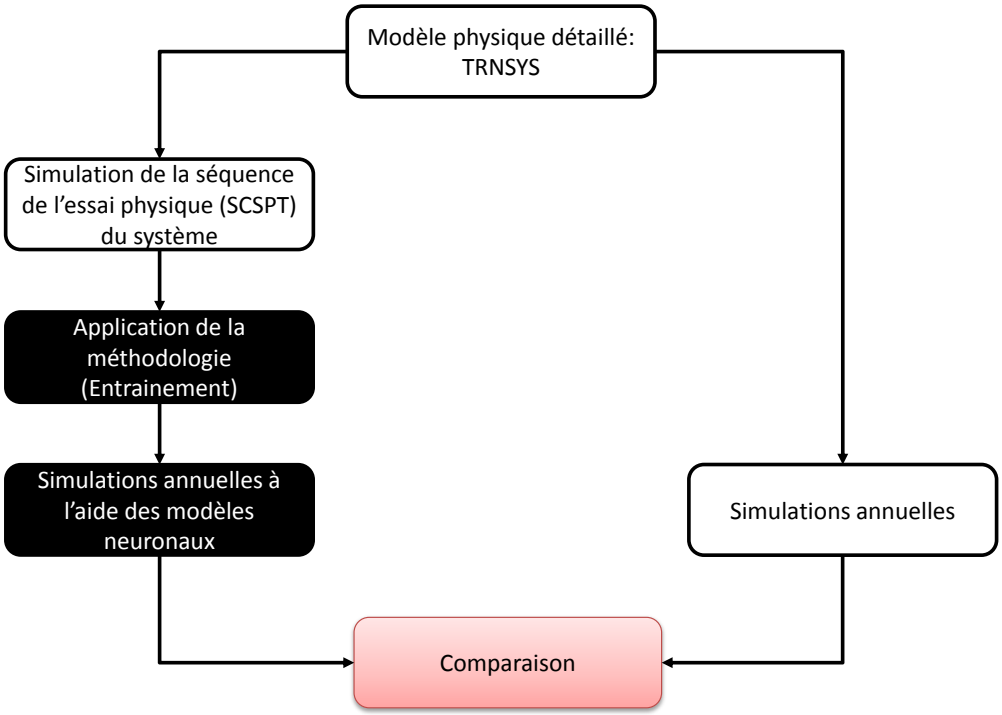

Figure 2- 19 : Schéma représentant le protocole de validation de la méthodologie à partir uniquement d'un modèle détaillé du système

- Une validation semi-expérimentale (expérimentation numérique-physique), qui diffère de l'approche précédente par la séquence de test. En effet, le test est réellement réalisé et non simulé. Ce processus est représenté dans la Figure 2- 20. Les données de l'essai servent donc pour la méthodologie RNA mais aussi pour construire et valider le modèle détaillé dans TRNSYS pour pouvoir par la suite comparer les simulations annuelles.
- Une validation expérimentale (expérimentation physique). Elle consiste à réaliser deux ou trois tests, de courte durée (toujours selon la méthode SCSPT), dans différentes conditions. Les données des tests sont ensuite utilisées à la fois en tant que base d'apprentissage et de test pour les modèles RNA. Par exemple, sur l'une des bases de données un RNA sera identifié et sur le reste il sera évalué. Cette approche est représentée sur la Figure 2- 21.

A défaut de données annuelles complètes de différents systèmes pour différentes conditions, la combinaison de ces trois approches de validation permet d'évaluer la robustesse

ainsi que la pertinence de la méthodologie proposée. Cependant, il serait très intéressant pour des travaux futurs d'envisager un suivi *in situ* de quelques systèmes sur du long terme pour compléter la validation.

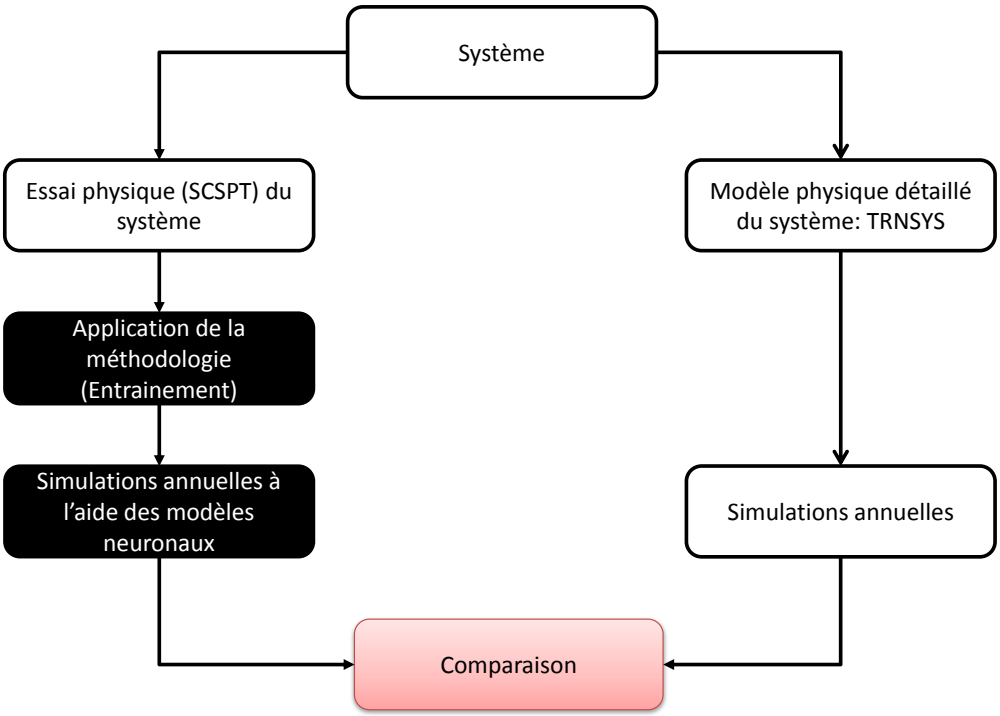

Figure 2- 20 : Schéma représentant le protocole de validation de la méthodologie à partir d'un modèle détaillé du système et de son test physique réel

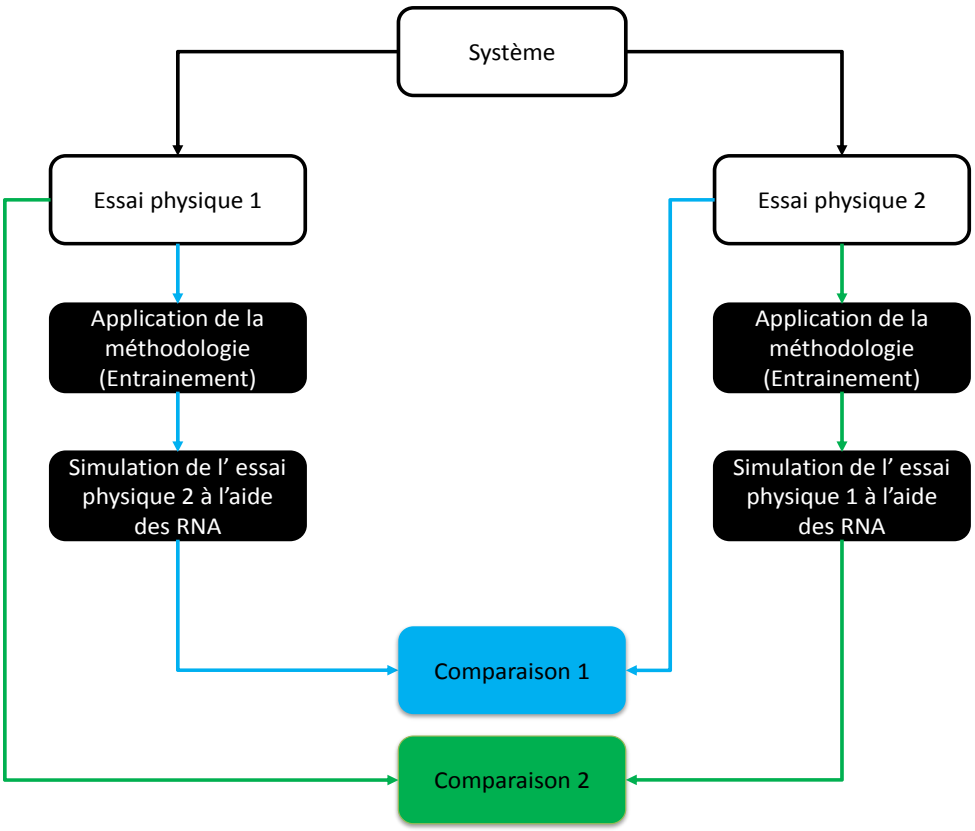

Figure 2- 21 : Schéma représentant le protocole de validation de la méthodologie à partir de deux ou trois test physique réel du système

II.4.2.2 Conditions aux limites de références

Dans ce paragraphe sont présentés les environnements de références définissant les conditions aux limites pour les simulations et tests à réaliser, que ce soit pour la formation de la base d'apprentissage des modèles neuronaux ou pour leur validation. Les conditions aux limites des systèmes sont formées par les puisages en ECS, le bâtiment et le climat dans lequel ce dernier est situé.

- Les climats :

Les climats sélectionnés pour constituer les différents environnements de référence dépendent de la fonction du système considéré. Ainsi, pour les systèmes ayant uniquement comme fonction de répondre aux besoins de chauffage et d'ECS, trois climats très différents sont considérés : Barcelone, Zurich et Stockholm. Ces climats ont été utilisés lors de la Tâche 32 de l'AIE [106]. Ils sont représentatifs des principaux marchés européens. A titre d'exemple, le climat de Stockholm présente un faible rayonnement solaire mais avec des besoins en chauffage très importants. Tandis que le climat de Barcelone est complètement différent de ce dernier (Figure 2- 22). Pour le système intégrant une machine à absorption, et donc répondant également aux besoins de climatisation, les climats considérés sont : Washington, Séville, Palerme, Philadelphie, Marrakech, Las Vegas et Aix-en-Provence. Ces climats ont été utilisés dans le cadre du projet PACSOL consistant à développer un système solaire thermique de climatisation, chauffage et production d'eau chaude sanitaire pour l'habitat individuel, utilisant notamment une machine à absorption. Les marchés visés concernaient trois zones

géographiques : Sud de l'Europe (France, Italie et Espagne), Pays du Maghreb et les Etats-Unis (côte Est). En plus des besoins de chauffage, ces climats peuvent également présenter des besoins pour la climatisation (Figure 2- 23).

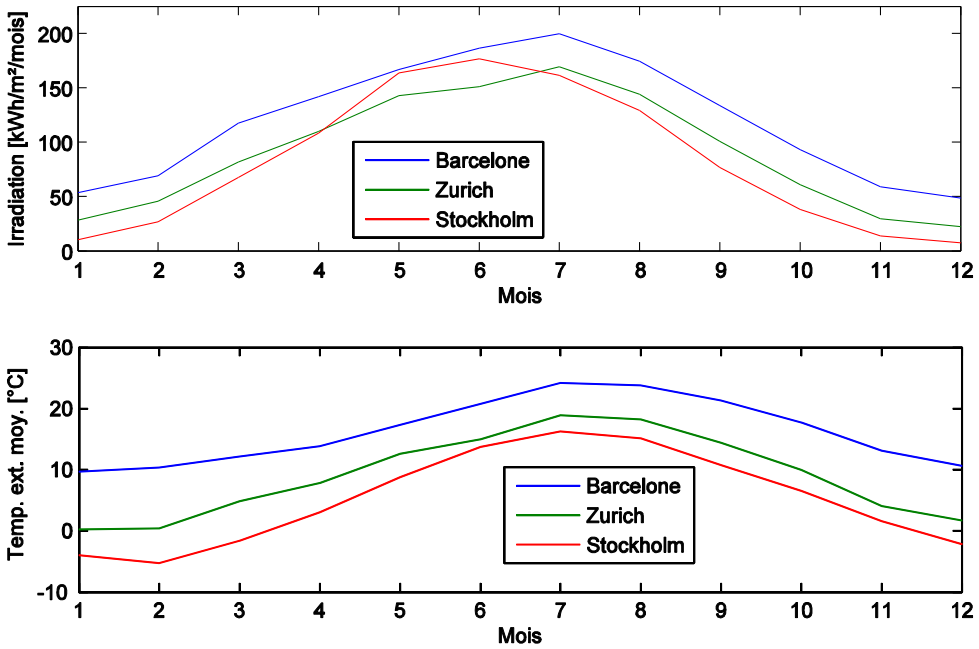

Figure 2- 22 : Evolutions de l'irradiation solaire (sur une surface horizontale) et de la température extérieure moyenne en fonction des mois de l'année pour trois climats différents [9]

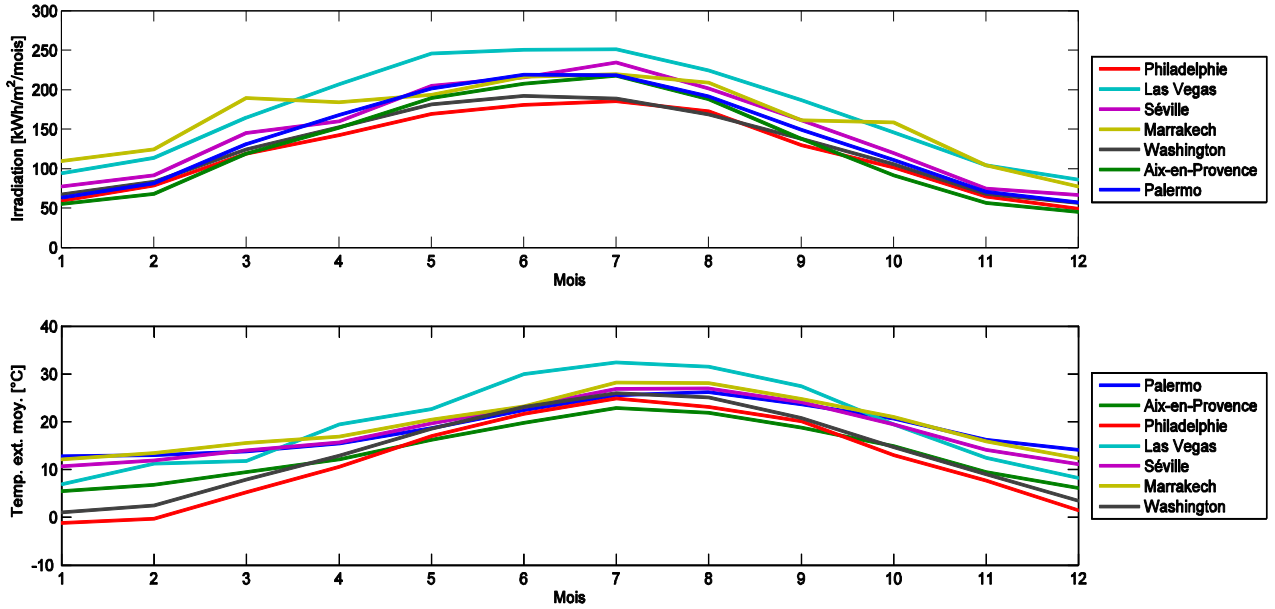

Figure 2- 23 : Evolutions de l'irradiation solaire (sur une surface horizontale) et de la température extérieure moyenne en fonction des mois de l'année pour les climats utilisés dans le cas du SSC intégrant une machine à absorption

Les données climatiques utilisées dans les différentes simulations annuelles sont issues de la plateforme numérique METEONORM. Ce logiciel calcule les données nécessaires à partir

de moyennes mensuelles d'ensoleillement et de température (entre autres), enregistrées sur du long-terme pour chaque site.

Dans le Tableau 2- 2 sont donnés les paramètres qui caractérisent les climats considérés. Ces paramètres sont utilisés dans les différents modèles TRNSYS pour dimensionner les circuits de chauffage ou de climatisation (avec la température extérieure de base $T_{ext,d}$, sauf pour le système intégrant la machine à absorption) et pour la définition du profil de la température d'eau froide (avec sa température moyenne $T_{tap,av}$, son amplitude ΔT_{tap} et son décalage temporel d_{off}, équation (1- 2)).

Tableau 2- 2 : Paramètres caractéristiques utilisés pour chaque climat.

Climat	$T_{ext,d}$ [°C]	$T_{tap,av}$ [°C]	ΔT_{tap} [°C]	d_{off} [jour]
Barcelone	-1	13,5	4,5	19
Zurich	-10	9,7	6,3	60
Stockholm	-17	8,5	6,4	80
Autres climats : SSC absorption	-5 (chauffage) 35 (climatisation)	13,5	4,5	19

- Les puisages d'ECS

Pour générer un profil de puisage d'ECS réaliste, nous utilisons l'outil développé par [56]. Le profil retenu est défini avec un pas de temps de 6 minutes. La consommation journalière d'eau chaude correspondante est de 200 litres à 45°C. Il est possible de télécharger le fichier de données correspondant sur le site du laboratoire Solar Energy Laboratory[1].

La température d'eau froide du réseau est considérée comme ayant une évolution sinusoïdale (équation (1- 2)), en fonction du temps *time* exprimé en heure. Elle est définie par les trois paramètres $T_{tap,av}$, ΔT_{tap} et d_{off}, dépendants du climat et définis dans le paragraphe précédent. La période de la sinusoïde correspond à la durée de la séquence : le nombre de jours de la simulation $N_{day,seq}$ est égal à 365 et 12 pour, respectivement, une simulation annuelle et une simulation de 12 jours.

- Les bâtiments

Similairement aux climats, les bâtiments de références sélectionnés pour constituer les différents environnements de références dépendent du système considéré. Pour les systèmes avec un appoint par une chaudière ou une PAC, nous avons considéré trois bâtiments de qualité d'isolation thermique différente. Ces bâtiments ont été définis dans le cadre de la Tâche 32 de l'IEA. Ils sont appelés SFH30, SFH60 et SFH100. Ils sont conçus de manière à ce qu'ils aient des besoins de chauffage respectifs de 30kWh.m^{-2}, 60kWh.m^{-2} et 100kWh.m^{-2} sur une année pour le climat de Zurich.

L'architecture de ces bâtiments ainsi que leurs caractéristiques sont définies dans [106]. La forme géométrique de ces maisons individuelles (Single Family House) est présentée sur la Figure 2- 24.

[1] https://sel.me.wisc.edu/

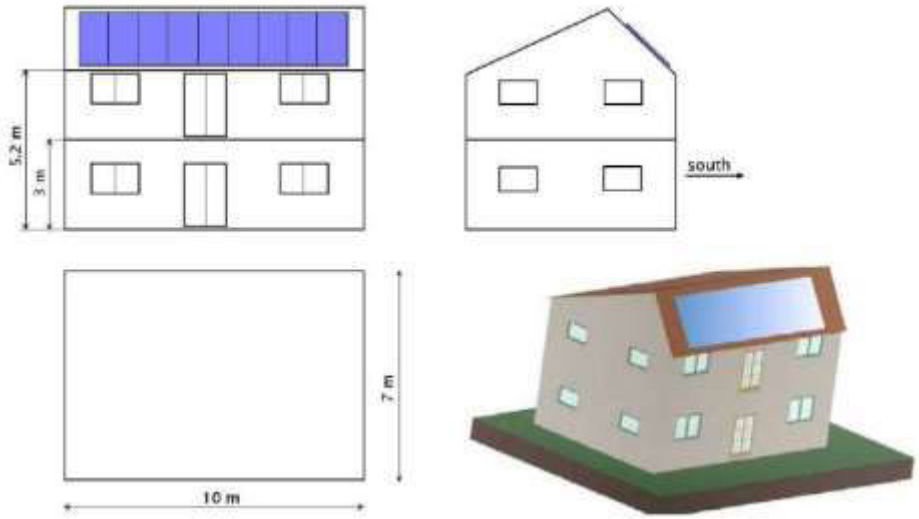

Figure 2- 24 : Caractéristiques géométriques des bâtiments SFH

Les modèles des bâtiments considérés prennent en compte les apports de chaleur internes (générés par les habitants et les appareillages électriques, voir les figures Figure C- 1et Figure C- 2 en annexe). Les simulations annuelles (protocole de validation dans les Figure 2- 19 et Figure 2- 20) utilisent des modèles détaillés de ces bâtiments. Pour certains systèmes ils sont définis sous le logiciel TRNSYS à l'aide du Type 56 [107]. Leurs programmes (*.bui) exploitables dans TRNSYS sont accessibles, avec le modèle du système de référence de la Tâche 32, via le site de l'AIE[2]. Ces bâtiments sont équipés de protections solaires (casquettes sur les fenêtres sud, ventilation nocturne, etc.), programmés de manière précise. Pour d'autres systèmes (systèmes solaires combinés à une PAC) c'est un modèle mathématique du comportement thermique du bâtiment du type ISO 13790-2008 [108] qui est utilisé. Les paramètres de ces modèles ont été identifiés de telle sorte que le modèle mathématique soit équivalent à celui du Type 56 de TRNSYS.

Pour les systèmes intégrant une machine à absorption, les bâtiments de références utilisés ont été développés sur la base du modèle initial du bâtiment de type SFH100 de la Tâche 32 de l'IEA, puis modifié dans le but d'être conforme avec les caractéristiques des bâtiments existants dans le sud de l'Europe, en Afrique du nord et aux Etats-Unis. Contrairement au bâtiment SFH100, les bâtiments considérés pour les systèmes intégrant une machine à absorption ne disposent pas de protection solaire et ont des caractéristiques d'isolation, de surface et de vitrage différentes. Sur le Tableau 2- 3 sont présentés un ensemble de caractéristiques pour le bâtiment considéré pour chaque zone géographique. Ces bâtiments ont été modélisés à l'aide du Type 56 de TRNSYS.

Tableau 2- 3 : quelques caractéristiques des bâtiments utilisés dans le cas du SSC intégrant une machine à absorption

Région	Europe	USA	Maroc
Surface	140m²	300m²	300m²
Consommation de référence	130 kWh/m² sur une année à Rennes	130 kWh/m² sur une année à Washington	60.7 kWh/m² sur une année à Marrakech
Matériau de construction	Brique	Ossature bois	Brique

[2] http://www.iea-shc.org/task32

- Les émetteurs de chaleur

Dans le cas des systèmes combinés à une chaudière ou PAC, les modèles de bâtiments sont couplés à des radiateurs à eau chaude pour répondre aux besoins de chauffage. Les modèles utilisés sont basés sur les paramètres K et n, définis dans la norme européenne EN442-2 [109] et caractérisant le radiateur. Pour les émetteurs choisis (radiateurs) on considère le même exposant 1,3. Le paramètre K quant à lui va changer selon le dimensionnement à considérer en fonction du climat et du bâtiment.

Le dimensionnement des radiateurs dépend à la fois du bâtiment et du climat. Pour simplifier le paramétrage, un premier dimensionnement est défini en fonction du climat de Zurich : les radiateurs de chaque bâtiment sont ainsi caractérisés par une puissance de charge $P_{sh,d,Zurich}$ ainsi que par les températures d'entrée $T_{sh,in,d}$ et de sortie $T_{sh,out,d}$ correspondantes (Tableau 2- 4) [9].

Tableau 2- 4 : Exemples des paramètres de dimensionnement des radiateurs en fonction du bâtiment dans le cas du climat de Zurich

Bâtiment	$P_{sh,d,Zurich}$ [W]	$T_{sh,in,d}$ [°C]	$T_{sh,out,d}$ [°C]	K
SFH30	3100	35	30	2939,77
SFH60	4605	40	35	2819,78
SFH100	6548	60	50	661,32

Pour adapter la puissance de dimensionnement à d'autres climats, il suffit d'utiliser la température extérieure de base comme le montre l'équation (2- 33) [9].

$$P_{sh,d} = P_{sh,d,Zurich} \frac{T_{sh,cons} - T_{ext,d}}{T_{sh,cons} - T_{ext,d,Zurich}} \quad \textbf{(2- 33)}$$

Pour les modèles de radiateurs, on utilise la plupart du temps comme paramètre leur puissance nominale $P_{sh,nom}$ pour des températures données (ici 90°C en entrée, 70°C en sortie avec une consigne pour la température intérieure à 20°C). Le passage de $P_{sh,d}$ à cette puissance nominale dépend du modèle souhaité.

Dans le cas des simulations annuelles de validation, le modèle détaillé de radiateur (Type 362 sous TRNSYS) est utilisé. Il est basé sur un échange de chaleur proportionnel à une différence de température logarithmique [9]. De plus, selon les simulations réalisées dans la Tâche 32 de l'IEA [55], la puissance nominale de ce type de radiateur est basée sur sa température moyenne et peut ainsi s'exprimer par l'équation (2- 34).

$$P_{sh,nom} = P_{sh,d} \left(\frac{\frac{90+70}{2} - 20}{\frac{T_{sh,in,d} + T_{sh,out,d}}{2} - T_{sh,cons}} \right)^n \quad \textbf{(2- 34)}$$

Pour le cas des systèmes solaires intégrant une machine à absorption, ce sont des ventilo-convecteurs qui ont été utilisés en tant qu'émetteur de chaleur et de froid (Type 91 et Type 112 du logiciel TRNSYS). Une étude des besoins de chauffage et de climatisation des 3 bâtiments considérés comme environnement de référence pour ce type de système a révélé qu'ils sont plus importants pour les climats de Marrakech et Philadelphie. Le débit d'air nominal du ventilo-convecteur a donc été dimensionné pour ces deux climats extrêmes. Il est

de 4680kg/h avec une puissance électrique absorbée de 393W. Tandis que le débit d'eau dans l'échangeur est de 1311 l/h.

- Le capteur solaire

Les principaux paramètres des modèles des capteurs solaires utilisés pour les simulations (annuelles ainsi que pour la séquence de test SCSPT) sous le logiciel TRNSYS seront présentés pour chacun des systèmes étudiés. Le fonctionnement de ces capteurs solaires est régi par l'équation suivante, exprimant le bilan d'énergie global du capteur :

$$c_5 \frac{dT_{capt}}{dt} = P_{sol,net} - \frac{P_{capt,out}}{A} \qquad \textbf{(2- 35)}$$

Le premier terme du bilan correspond à l'inertie du capteur, le deuxième aux apports solaires nets (les pertes thermiques sont déduites des apports solaires : équation (2- 36)), tandis que le troisième terme correspond à la chaleur récupérée.

$$\begin{aligned} P_{sol,net} = F'(\tau\alpha)_{en}(K_{\theta b}(\theta).G_b + K_{\theta d}(\theta).G_d) - c_6 v G - c_1(T_{capt} - T_{ext}) \\ - c_2(T_{capt} - T_{ext})^2 - c_3 v(T_{capt} - T_{ext}) + c_4(E_L - \sigma_{sb} T_{ext}^4) \end{aligned} \qquad \textbf{(2- 36)}$$

Tous les paramètres mis en jeu dans ces équations sont identifiés grâce à des tests normatifs de capteur.

Le facteur d'angle d'incidence $K_{\theta b}$ dépend de l'angle d'incidence θ des rayons du soleil sur le plan des capteurs. Il est souvent exprimé par l'équation (2- 37).

$$K_{\theta b}(\theta) = 1 - b_0\left(\frac{1}{\cos(\theta)} - 1\right) \qquad \textbf{(2- 37)}$$

L'angle d'incidence est calculé selon la position du soleil (angle zénithal θ_S et azimut solaire γ_S par rapport au sud), l'inclinaison β et l'azimut par rapport au sud γ du plan des capteurs à l'aide de l'équation (2- 38).

$$\cos(\theta) = \cos(\theta_S)\cos(\beta) + \sin(\theta_S)\sin(\beta)\cos(\gamma_S - \gamma) \qquad \textbf{(2- 38)}$$

Le capteur solaire est considéré, dans le cadre de la méthodologie, comme faisant partie du système (paragraphe II.4.1.). Cependant le rayonnement solaire disponible pour le système dépend de la surface du capteur A. Cette dernière aura un impact sur la consommation d'appoint du système et ainsi sur sa performance énergétique.

Ce paramètre constitue en conséquence une condition aux limites du système autant que le type du bâtiment et le climat. Des valeurs entre 10m² et 19m² ont été considérés pour les simulations annelles.

Dans la suite, les conditions aux limites seront notées pour chaque système SFH**xx**-**zzzz**-**yy** m² signifiant que le système est intégré à un bâtiment de type SFH**xx** dans le climat de **zzzz** et que le capteur solaire a une surface de **yy** m².

Pour que les modèles neuronaux aient un bon pouvoir de généralisation il faut que la base de données contienne des informations représentant un fonctionnement « moyen » fréquent du système. C'est pour cette raison que l'environnement « moyen » SFH60-Zurich-16m² a été

considéré pour l'apprentissage. L'entrainement sur des cas extrêmes tels que SFH30-Barcelone-19m² ou SFH100-Stockholm-10m² résulterait en des modèles avec un faible pouvoir d'extrapolation à d'autres conditions plus fréquentes.

II.5. Finalisation de la méthodologie par une application à un SSC

Après avoir développé le processus de modélisation par les RNA (section II.3.) et déterminé comment la méthodologie sera évaluée numériquement et expérimentalement (section II.4.), il reste à affiner certains éléments dans la méthodologie. En effet, dans cette partie sont abordés le choix des variables d'entrées, le lissage des données d'apprentissage et les critères de sélection des modèles RNA les plus pertinents. En outre, une analyse des résultats de l'application de la méthodologie dans le cas d'un SSC est également réalisée.

II.5.1. Description du système étudié

Le premier système solaire combiné auquel la méthodologie a été appliquée est un système commercialisé principalement en France et en Allemagne. Ce système a deux fonctions, le chauffage et la production d'ECS. Il a été étudié dans plusieurs projets nationaux et européens [9], [110], [111]. Le schéma hydraulique du système est représenté sur la Figure 2- 25.

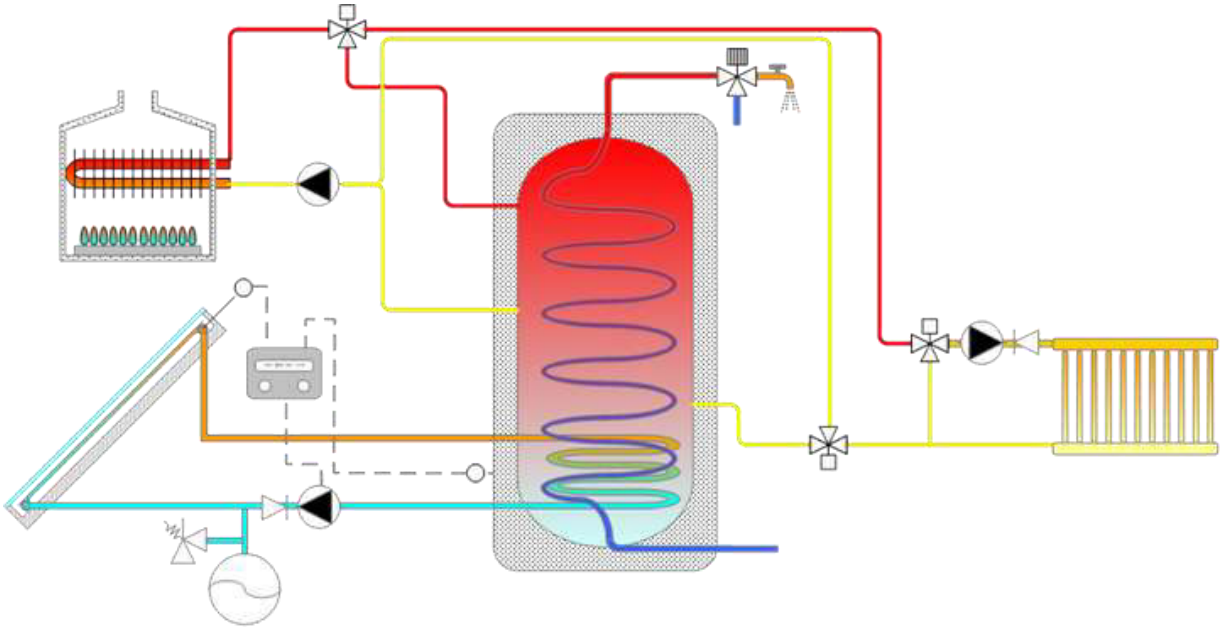

Figure 2- 25 : Schéma hydraulique du premier système étudié SSC

Les transferts des flux énergétiques au sein du système se caractérisent par :

- L'utilisation d'un échangeur interne au ballon de stockage (partie basse) pour la récupération de la chaleur du capteur solaire.
- L'utilisation d'un échangeur interne au ballon de stockage (partie haute) pour le transfert de la chaleur vers les besoins d'ECS.
- La température de retour du circuit de chauffage est augmentée par un piquage direct sur le ballon de stockage.

- L'absence d'échangeur thermique entre le ballon de stockage et la chaudière. L'échange de la chaleur se fait directement au travers de deux piquages.

Le modèle physique exploité de ce système a été développé à partir de son schéma hydraulique et de sa documentation technique (en particulier pour les systèmes de contrôle-commande). Les paramètres du ballon de stockage, spécifiques au Type 340 sous TRNSYS, ont été identifiés par des tests selon la norme EN 12977-3 par le centre d'essai TZS de l'institut ITW de l'Université de Stuttgart. Aucune perte thermique dans les circuits n'est considérée dans ce modèle.

Les caractéristiques physiques de ce système utilisées dans les simulations sous TRNSYS, sont données dans le Tableau 2- 5.

Tableau 2- 5 : Caractéristiques physiques utilisées pour le premier système

Remarques	Type TRNSYS	Caractéristiques
Appoint		
Modèle de chaudière il calcul l'énergie consommée en prenant en compte les pertes de combustion, les pertes thermiques et l'énergie récupérée par condensation.	Type 370 utilisé dans la tâche 32 de l'IEA et développé par [112]	Puissance nominale de 24kW, Chaudière à gaz, la température fournie par la chaudière est variable.
Ballon de stockage		
Modèle détaillé du stockage en eau, échangeur solaire + échangeur ECS immergé	Type 340	720 litres ; 1,93m de hauteur Coefficient de pertes latérales de 8,59 W/K et pour les parties supérieure et inférieure une valeur de 0,76 W/K
Capteur solaire		
	Type 232	$F'(\tau\alpha)_{en}$=0,77; $K_{\theta d}$=0,96 c_1=3,478 (W/m²K); β=45°;c_5=7422 (J.m^{-2} K^{-1}) c_2=0,015 (W/m²K²) ; b_0=0,18

II.5.2. Lissage des données

En analysant les courbes représentatives de l'évolution de la puissance de l'appoint calculée respectivement par le modèle neuronal et le modèle TRNSYS (Figure 2- 26), nous remarquons de façon claire que la courbe de puissance calculée par le RNA suit globalement l'évolution de la courbe de référence. Par conséquent, cette remarque montre que le RNA a saisi globalement le fonctionnement du système. Néanmoins nous constatons aussi que cette concordance entre les deux courbes se détériore aux niveaux de pics qui correspondent aux mises en marche de la chaudière pour la préparation de l'ECS (Figure 2- 26). La puissance dans ce cas peut passer brusquement d'une valeur quasiment nulle à une valeur très élevée.

Ces pics de puissance correspondent au fonctionnement de la chaudière en mode ECS quand la température mesurée dans le ballon à une certaine hauteur passe en dessous d'un certain seuil. En conséquence pour que le RNA reproduise parfaitement ce comportement il faudrait avoir en entrée cette température mesurée dans le ballon. Or, cette donnée n'est pas disponible en entrée. En effet, le test selon la méthode SCSPT (décrit dans le premier chapitre) est non intrusif (pas de mesure dans le ballon), seules les variables aux bornes du système au niveau de chaque circuit sont considérées. En conséquence nous n'avons pas en entrée du modèle les informations suffisantes pour reproduire fidèlement le comportement du

ballon et donc de la chaudière. Dans le but de réduire les effets des pics qui peuvent perturber l'entrainement nous devons traiter ces signaux pour avoir une information en termes d'énergie qui soit cohérente et plus facile à apprendre dans notre cas d'extrêmes contraintes.

L'idée de filtrer les données de façon à avoir des signaux plus réguliers sans pour autant perdre d'information (quantité d'énergie à respecter) constitue une solution intéressante. Réaliser un changement de pas de temps (passage de 30 minutes à 1h par exemple) atténuera les pics sans doute mais ceci affectera la quantité d'information disponible dans l'échantillon de départ et la taille des données d'entrainement.

Il est plus judicieux de réaliser un lissage de données via une moyenne mobile (ou moyenne glissante). Le lissage supprimera quelques fluctuations transitoires trop importantes (Figure 2- 27). Il est indispensable de choisir une fenêtre de lissage qui réalise le compromis entre le lissage des données et la préservation de l'information physique et dynamique. Le choix d'une fenêtre de lissage inférieure à 10 pas de temps répond raisonnablement à ce compromis. Un lissage sur cinq pas de temps a été retenu pour la méthodologie.

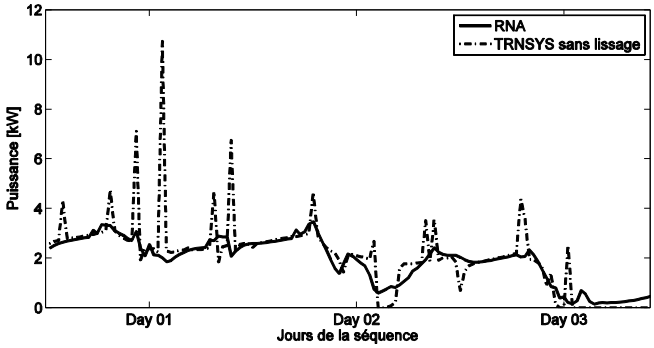

Figure 2- 26 : Comparaison entre la puissance de la chaudière avant l'application de la moyenne mobile sur les données d'apprentissage

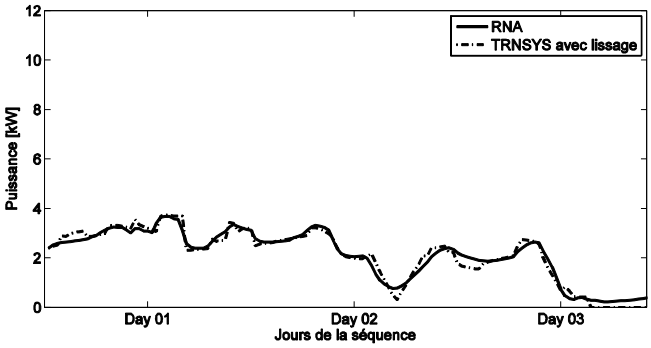

Figure 2- 27 : Comparaison entre la puissance de la chaudière après l'application de la moyenne mobile sur les données d'apprentissage

II.5.3. Sélection de la configuration E/S de modélisation

Dans le but de déterminer l'approche à adopter pour le développement de la méthodologie, le système présenté précédemment a été modélisé selon les trois approches présentées dans le paragraphe II.4.1.

Tableau 2- 6 : Les nombres d'éléments relatifs à la construction des modèles neuronaux explorés

FA	Nombre maximal des neurones	TD	Intervalle de normalisation	Nombre maximal des réinitialisations
2	15	3	1	15

Le test du système (décrit dans le paragraphe II.5.1.) selon la méthode SCSPT a été simulée dans l'environnement de Zurich-SFH60-16m² à l'aide de son modèle TRNSYS. La méthodologie de modélisation par RNA a été ensuite appliquée sur la base des données récupérées suite à la simulation du test.

Dans le Tableau 2- 6 sont donnés les nombres d'éléments relatifs à la construction des modèles neuronaux explorés. L'application du processus de modélisation développé (sans prendre en compte dans cette partie l'aspect sélection de modèle) résulte en 90 modèles entrainés 15 fois chacun (réinitialisation) soit 1350 processus d'entrainement.

Une analyse statistique, uniquement sur la base des données d'apprentissage et pour chaque configuration (retards temporel TD, fonction d'activation FA), a été réalisée. Elle permet de comparer la pertinence des trois configurations de modélisation.

Sur la Figure 2- 28, Figure 2- 29, Figure 2- 30, Figure 2- 31, Figure 2- 32 et Figure 2- 33 sont présentées les boites à moustaches donnant une information sur la distribution des racines des erreurs moyennes quadratiques pour chaque nombre de neurones dans la couche cachée quand l'initialisation est modifiée plusieurs fois selon la technique cité dans le paragraphe II.3.2.8. L'erreur est calculée uniquement pour les données d'apprentissage en boucle fermée et pour la variable qui nous intéresse ici, à savoir la consommation de l'appoint. Ci-après sont regroupées les principales remarques de cette étude en fonction de la fonction d'activation (FA) utilisée.

- Avec « *tanh* » comme FA en sortie :

Pour un TD de 1 ou 2 pas de temps et pour la configuration « entrées sortie physique », les erreurs pour la grande majorité des RNA sont entre 2,5 et 4 kW. Cependant, Il existe quelques RNA avec de faibles erreurs de moins de 0,5 kW mais également des RNA avec des erreurs de plus de 5,5 kW. Pour un TD de 3 pas de temps la moyenne des erreurs tend vers 4-3,5 kW tandis que les extrêmes restent équivalents aux deux premières configurations (TD = 1 ou 2), ils avoisinent 6 kW/0,5 kW.

Globalement les mêmes remarques peuvent être tirées de la configuration « flux énergétiques » avec les deux températures (5 entrées). La différence entre les deux réside dans les ordres de grandeurs des erreurs. En effet, pour TD = 1 et 2 les erreurs pour la grande majorité des RN sont autour de 2,5 kW. Il existe quelques RNA avec des erreurs de moins de 0,5 kW et d'autres avec plus de 5 kW d'erreurs. La moyenne tend vers 2 kW pour TD = 3 mais les extrêmes gardent les mêmes ordres de grandeurs.

Pour la dernière configuration « flux énergétiques », les résultats sont très similaires à ceux de la dernière configuration intégrant la température intérieure au bâtiment et la température extérieure mais avec des erreurs légèrement plus faibles à 0,5 kW prés. Au niveau global il y a plus de RNA avec une faible erreur pour cette configuration que pour les deux premières.

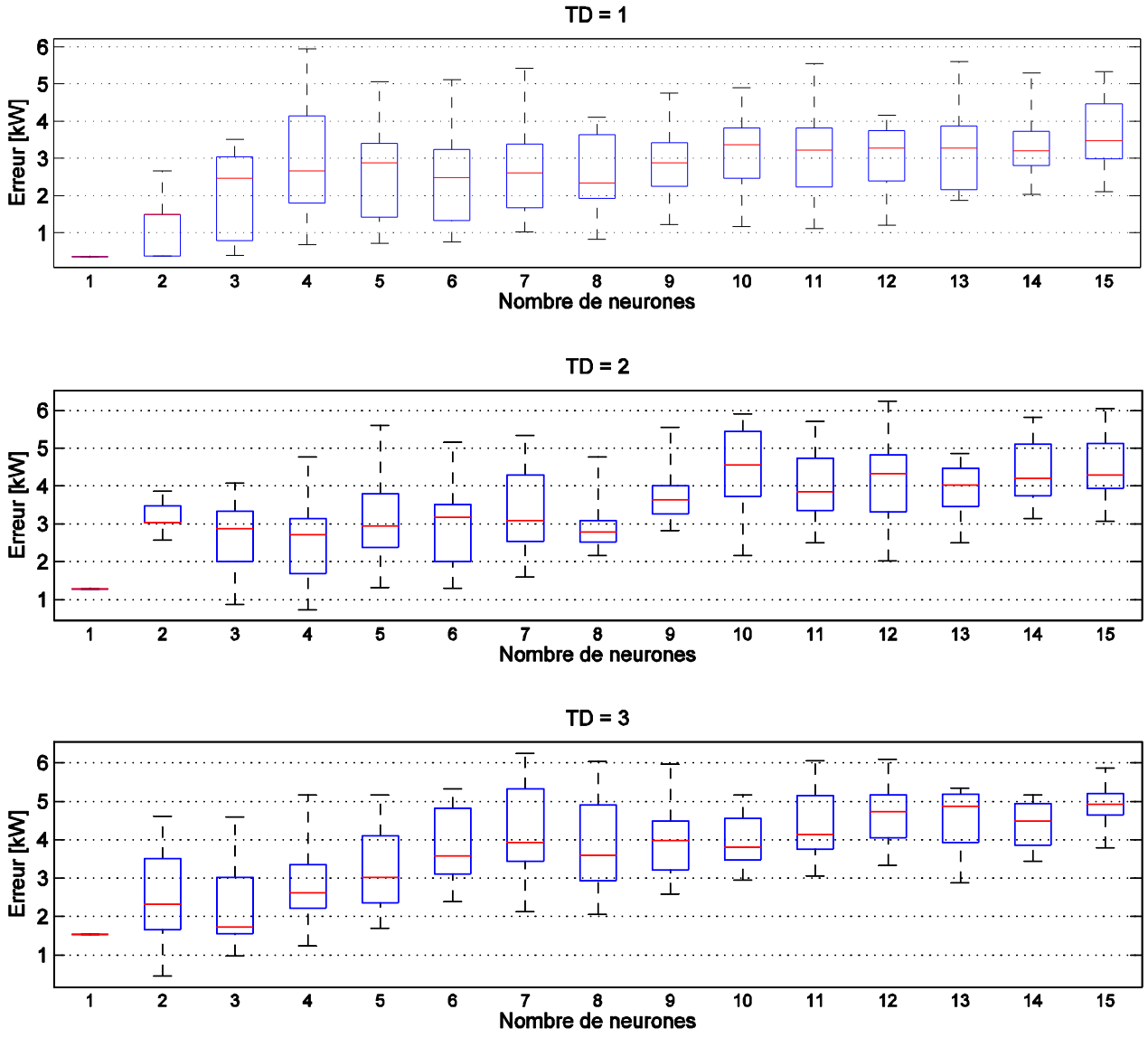

Figure 2- 28 : Distribution des racines des erreurs moyennes quadratiques pour chaque nombre de neurones dans la couche cachée quand l'initialisation est modifiée plusieurs fois. Cas de la FA $tanh$ en sortie et de la configuration « entrées-sortie physiques »

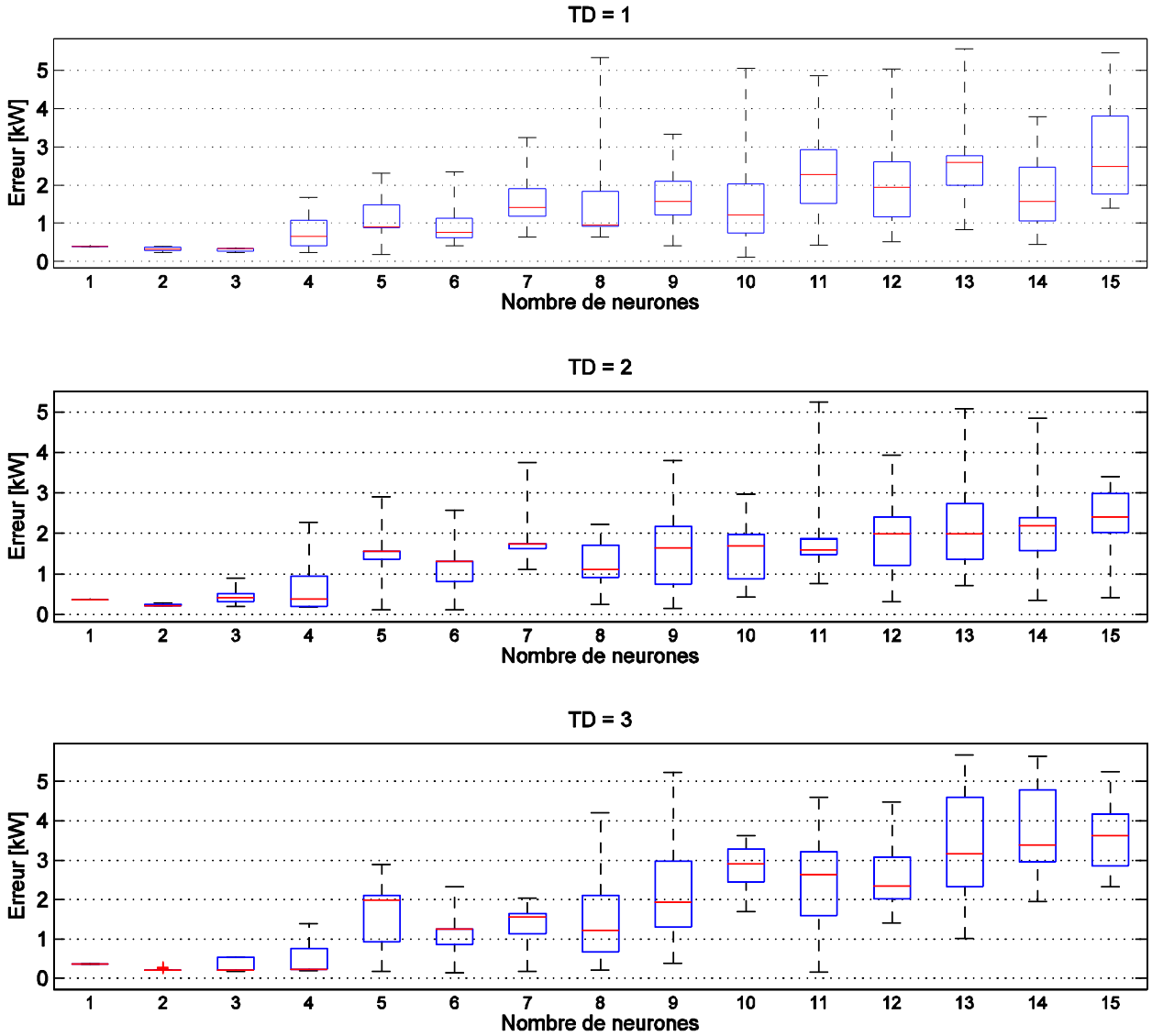

Figure 2- 29 : Distribution des racines des erreurs moyennes quadratiques pour chaque nombre de neurones dans la couche cachée quand l'initialisation est modifiée plusieurs fois. Cas de la FA $tanh$ en sortie et de la configuration « flux énergétique » avec les deux températures

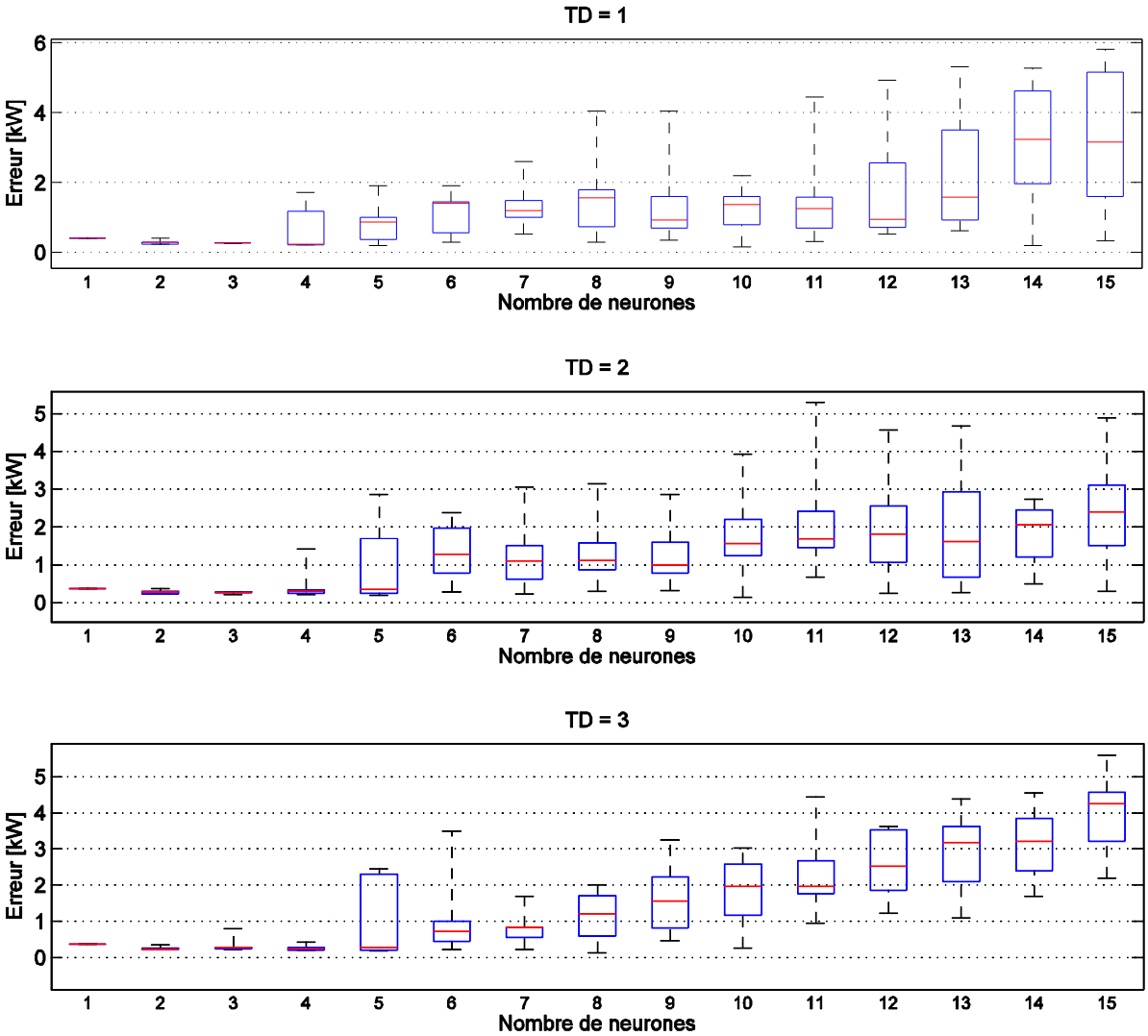

Figure 2- 30 : Distribution des racines des erreurs moyennes quadratiques pour chaque nombre de neurones dans la couche cachée quand l'initialisation est modifiée plusieurs fois. Cas de la FA $tanh$ en sortie et de la configuration « flux énergétique »

- Avec une FA linéaire en sortie :

Les mêmes remarques que précédemment peuvent être citées pour les RNA dont la fonction d'activation de sortie est linéaire mais la tendance des erreurs correspondantes est largement plus importante. Nous constatons aussi que les erreurs les plus faibles sont acquises pour la dernière configuration « flux énergétique ».

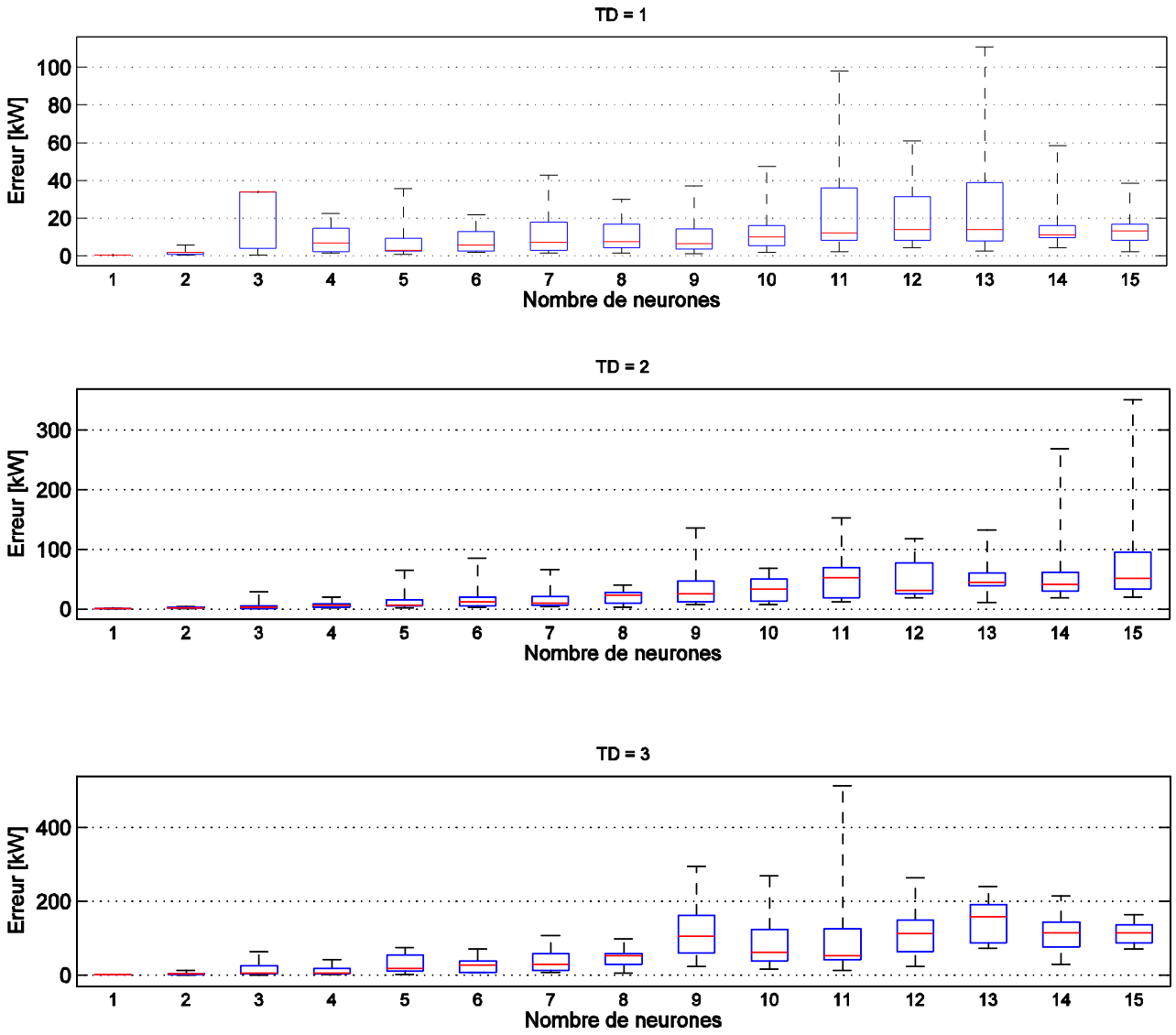

Figure 2- 31 : Distribution des racines des erreurs moyennes quadratiques pour chaque nombre de neurones dans la couche cachée quand l'initialisation est modifiée plusieurs fois. Cas de la FA linéaire en sortie et de la configuration « entrées sortie physiques »

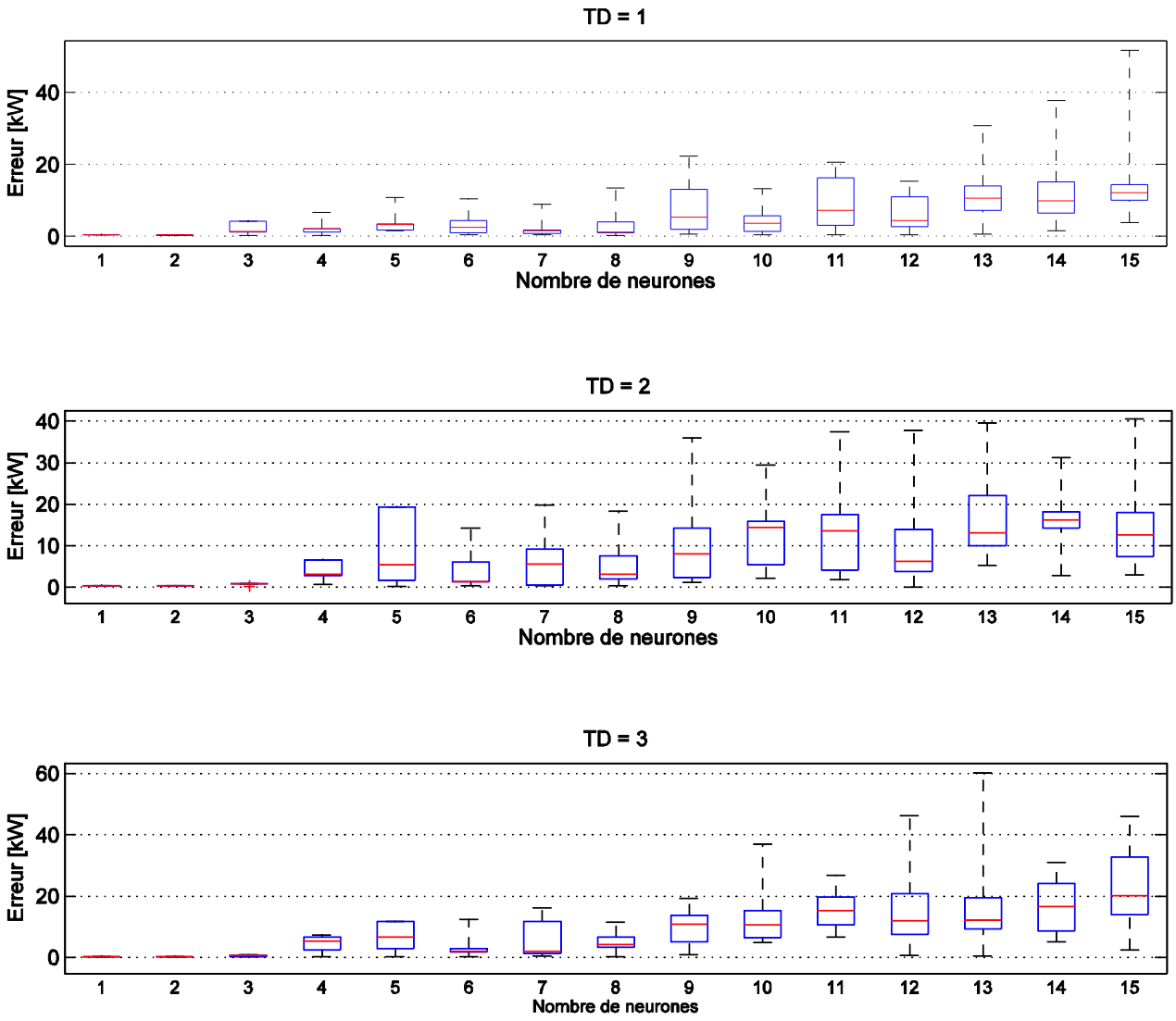

Figure 2- 32 : Distribution des racines des erreurs moyennes quadratiques pour chaque nombre de neurones dans la couche cachée quand l'initialisation est modifiée plusieurs fois. Cas de la FA linéaire en sortie et de la configuration « flux énergétique » avec les deux températures

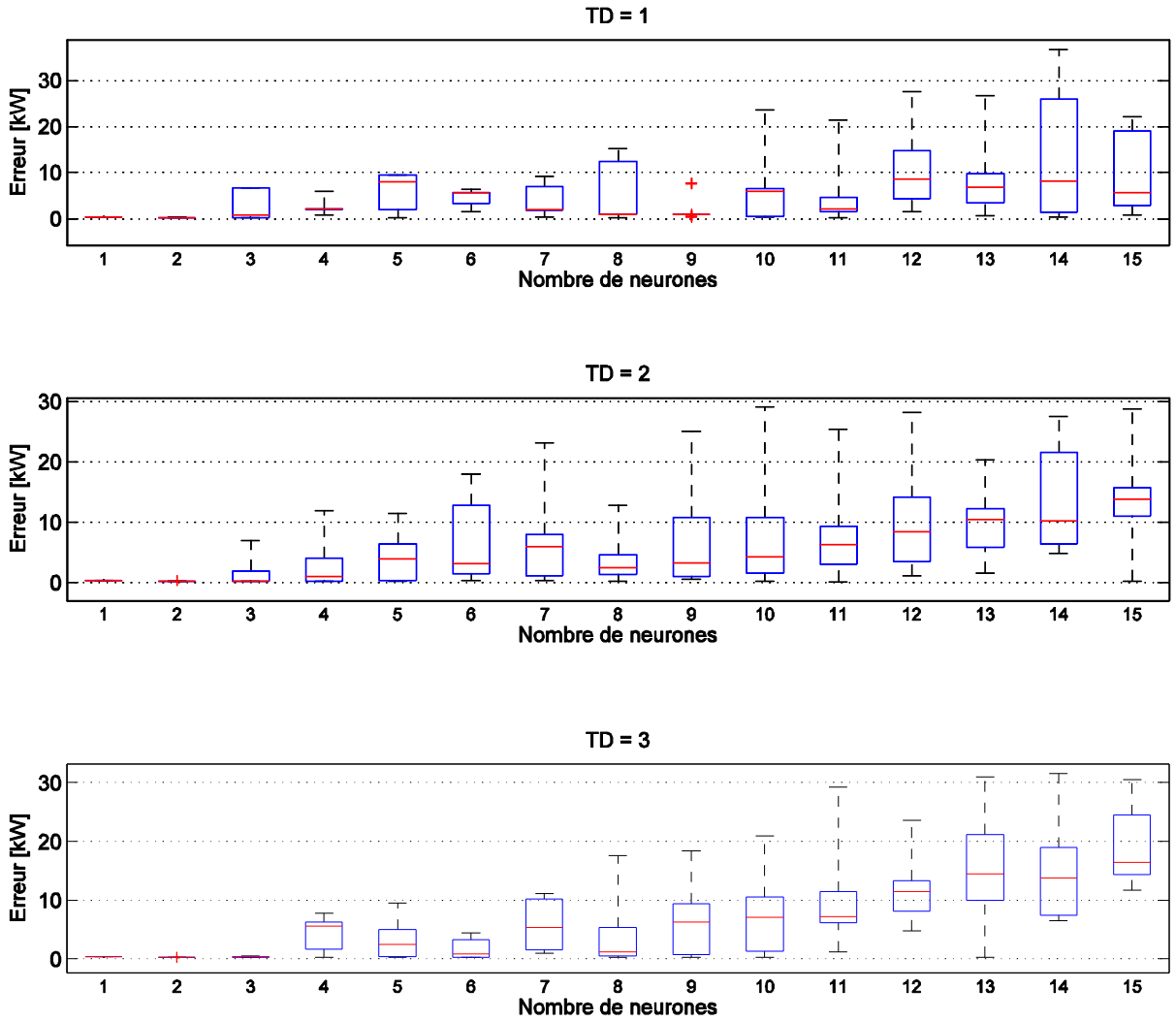

Figure 2- 33 : Distribution des racines des erreurs moyennes quadratiques pour chaque nombre de neurones dans la couche cachée quand l'initialisation est modifiée plusieurs fois. Cas de la FA linéaire en sortie et de la configuration « Flux énergétique »

- Bilan :

Les résultats présentés montrent que la configuration « flux énergétique » avec seulement 3 entrées est la plus pertinente. Pour cette dernière les niveaux d'erreurs sont les plus faibles et ce, quel que soit la FA utilisée. D'autre part, le temps de calcul est relativement moins long par rapport aux autres configurations (il est respectivement 7 fois et 3 fois inférieur aux deux autres configurations, sachant que l'application de la méthodologie prend entre 2 et 4 jours). La configuration « flux énergétique » met en jeu moins de variables (la fonction RNA est en conséquence moins complexe ce qui fait que la minimisation de la fonction objectif prend moins de temps de calcul). En effet, la méthodologie développée étant destinée à être utilisée comme un moyen de certification ou de labellisation énergétique (utilisation par des personnes non expertes), le temps nécessaire pour son application est naturellement souhaité être le plus court possible. Aussi, dans la configuration « flux énergétique » avec 3 entrées, les variables mises en jeu sont homogènes ce qui n'est pas le cas des autres configurations. Les erreurs sont plus faibles et moins dispersées que pour l'approche « entrées sortie physique » mais

restent du même ordre de grandeur que l'autre version de la configuration « flux énergétique ». Il est donc plus facile d'élaborer et sélectionner un modèle pertinent avec cette approche en puissance. En outre, les résultats présentés dans Figure 2- 34, Figure 2- 35 et Figure 2- 36 montrent qu'il est difficile pour les RNA de modéliser les autres sorties du système (les températures), sauf le débit dans l'émetteur de chaleur. Ces sorties étant utilisées également comme entrées retardées vont certainement amplifier la propagation des erreurs dans le modèle.

En prenant en compte les résultats présentés dans cette partie et l'analyse réalisée préalablement dans le paragraphe II.4.1. nous avons retenu la configuration en puissances « flux énergétique » avec seulement 3 entrées, malgré le fait qu'elle ne correspond pas à une représentation physique du système, pour la méthodologie. En effet, c'est cette configuration qui réalise le meilleur compromis entre simplicité, adéquation avec l'objectif de la méthodologie et précision souhaitée.

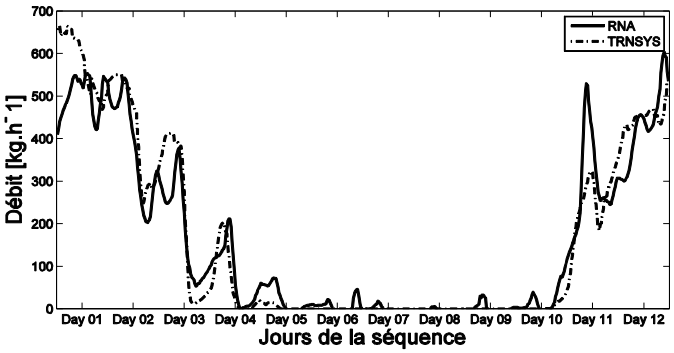

Figure 2- 34 : Comparaison entre le débit dans l'émetteur de chaleur calculée par TRNSYS et prédit par le RNA

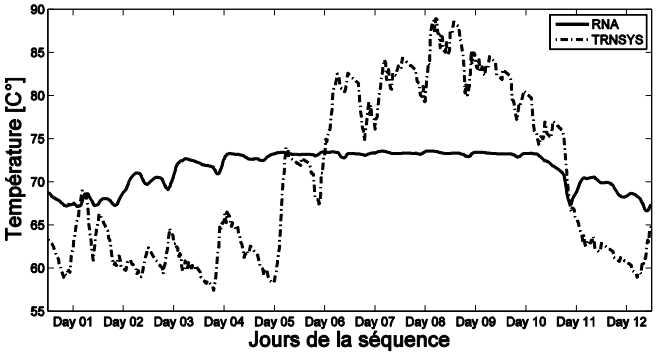

Figure 2- 35 : Comparaison entre la température d'ECS calculée par TRNSYS et prédite par le RNA

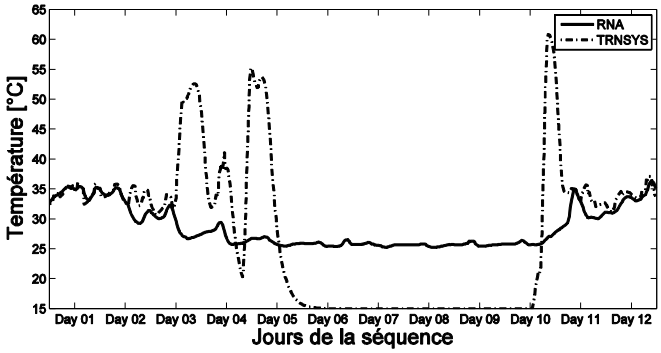

Figure 2- 36 : Comparaison entre la température en entrée de l'émetteur de chaleur calculée par TRNSYS et prédite par le RNA

Les figures des résultats correspondant à la configuration retenue (Figure 2- 30 à Figure 2- 33) montrent que, pour les deux FA, les erreurs les plus faibles sont acquises avec un nombre de neurones réduit dans la couche cachée (inférieur à 6). Cependant, certains modèles avec un nombre de neurones important (14 ou 15) ont également des erreurs très faibles. Cette remarque est en cohérence avec l'analyse du sous paragraphe II.3.2.5. La totalité des nombres de neurones dans la couche cachée (1 à 15) sera donc exploré avant de sélectionner le RNA le plus pertinent.

II.5.4. Critères de sélection des RNA

II.5.4.1 Critère de sélection

La méthodologie prévoit d'entrainer de nombreuses architectures de réseaux de neurones et de sélectionner le meilleur parmi tous les réseaux entrainés. La sélection de ce réseau doit se faire en se basant sur un critère fiable, calculé à partir uniquement des données d'entrainement (seules données disponibles). Or, il existe un grand nombre de critères pour estimer la qualité d'un modèle. Selon le critère choisi, le réseau sélectionné ne sera pas forcément le même, d'où la nécessité de répertorier plusieurs critères et d'étudier leur pertinence dans le cadre de cette méthodologie (Tableau 2- 7). Il est possible de regrouper ces critères en deux catégories : les critères dépendants uniquement de la sortie du modèle et ceux qui dépendent également de sa complexité, c'est à dire le nombre de ses paramètres à identifier.

Sur le Tableau 2- 7, y_i désigne la sortie du système (mesurée ou calculée à l'aide d'un modèle physique détaillé développé sous TRNSYS) à l'instant i, $\overline{y}$ sa moyenne sur toute la période considérée et $\widehat{y}_i$ la sortie du modèle RNA au même instant. N et q sont respectivement le nombre d'échantillons et de paramètres du modèle.

Tous les critères sont à minimiser (« ErrTot » et « ME » doivent être minimisés en valeur absolue) sauf « *R²* », « *Sign* », « *R2bar* » et « *DA* ». Pour étudier la pertinence des critères répertoriés, le réseau sélectionné (minimise ou maximise le critère considéré) pour chaque critère est utilisé pour l'estimation des performances du système dans 36 conditions annuelles différentes (3 climats, 3 types de bâtiment et 4 surfaces de capteur).

Ces estimations sont comparées aux résultats de simulations du modèle détaillé du même système pour les mêmes conditions. Il est à rappeler que ce modèle détaillé est disponible ici

dans le cadre du développement de la méthodologie mais il ne sera pas disponible lors de l'implémentation finale de la méthodologie (dans le cas de garantie des performances énergétiques par exemple). Pour évaluer la pertinence des réseaux de neurones sélectionnés par rapport aux modèles détaillés, on utilise le coefficient de corrélation appliqué aux évaluations des énergies annuelles de chaque simulation (36 conditions sauf pour les SSC couplés à une machine à absorption). Il sera noté, pour éviter toute confusion avec le « *R2* » calculé avec des données de puissance (et uniquement pour les données d'apprentissage), « *R2nrj* ». Nous rappelons ici que le calcul de « *R2nrj* » est possible uniquement si un modèle détaillé est disponible.

Tableau 2- 7 : Critères de sélection des modèles se basant uniquement sur les données d'apprentissage

Critère de sélection	Définition	Critères de sélection **dépendants de la complexité du modèle**	Définition
R²	$1-\frac{\sum(y_i-\hat{y}_i)^2}{\sum(y_i-\bar{y})^2}$	*BIC*	$\log\left(\frac{\sum(y_i-\hat{y}_i)^2}{N}\right)+\frac{q\,log(N)}{N}$
RMSE	$\sqrt{\frac{1}{N}\sum(y_i-\hat{y}_i)^2}$	*BIC1*	$\log\left(\frac{\sum(y_i-\hat{y}_i)^2}{N}\right)+\frac{q^2\,log(N)}{N}$
ErrTot	$\frac{\sum(y_i-\hat{y}_i)}{\sum y_i}$	*BIC2*	$\log\left(\frac{\sum(y_i-\hat{y}_i)^2}{N}\right)+\frac{\sqrt{q}\,log(N)}{N}$
MAE	$\frac{1}{N}\sum\lvert y_i-\hat{y}_i\rvert$	*AIC*	$\log\left(\frac{\sum(y_i-\hat{y}_i)^2}{N}\right)+2\frac{q}{N}$
Sign	$\frac{1}{N}\sum z_i$ où : $z_i=\begin{cases}1\ si\ y_{i+1}.\hat{y}_{i+1}>0,\\0 \qquad sinon.\end{cases}$	*AIC1*	$\log\left(\frac{\sum(y_i-\hat{y}_i)^2}{N}\right)+2\frac{q^2}{N}$
DA	$\frac{1}{N}\sum a_i$ où : $a_i=\begin{cases}1\ si\ (y_{i+1}-y_i)(\hat{y}_{i+1}-\hat{y}_i)>0,\\0\ sinon.\end{cases}$	*AIC2*	$\log\left(\frac{\sum(y_i-\hat{y}_i)^2}{N}\right)+2\frac{\sqrt{q}}{N}$
ME	$\frac{1}{N}\sum(y_i-\hat{y}_i)$	*AICC*	$\log\left(\frac{\sum(y_i-\hat{y}_i)^2}{N}\right)+\frac{2q}{N-q-1}$
d	$\sqrt{(\frac{RMSE}{5})^2+(\frac{ErrTot}{0,9})^2}$	$\bar{R}$ *(R2bar)*	$1-\frac{\sum(y_i-\hat{y}_i)^2}{\sum(y_i-\bar{y})^2}\times\frac{N-1}{N-q}$

II.5.4.2 Analyse des résultats

Le test physique du système (paragraphe II.5.1.), selon la méthode SCSPT décrite dans le premier chapitre, a été simulé à l'aide du modèle présenté dans ce même paragraphe. Les puissances nécessaires pour la construction des entrées et la sortie des modèles neuronaux sont calculées à l'aide des différentes variables récupérées de la simulation du test. Elles sont représentées sur la Figure 2- 37. Le même modèle utilisé pour simuler la séquence de test est exploité pour réaliser plusieurs simulations annuelles pour évaluer les estimations des modèles neuronaux (voir protocole de validation dans la Figure 2- 19).

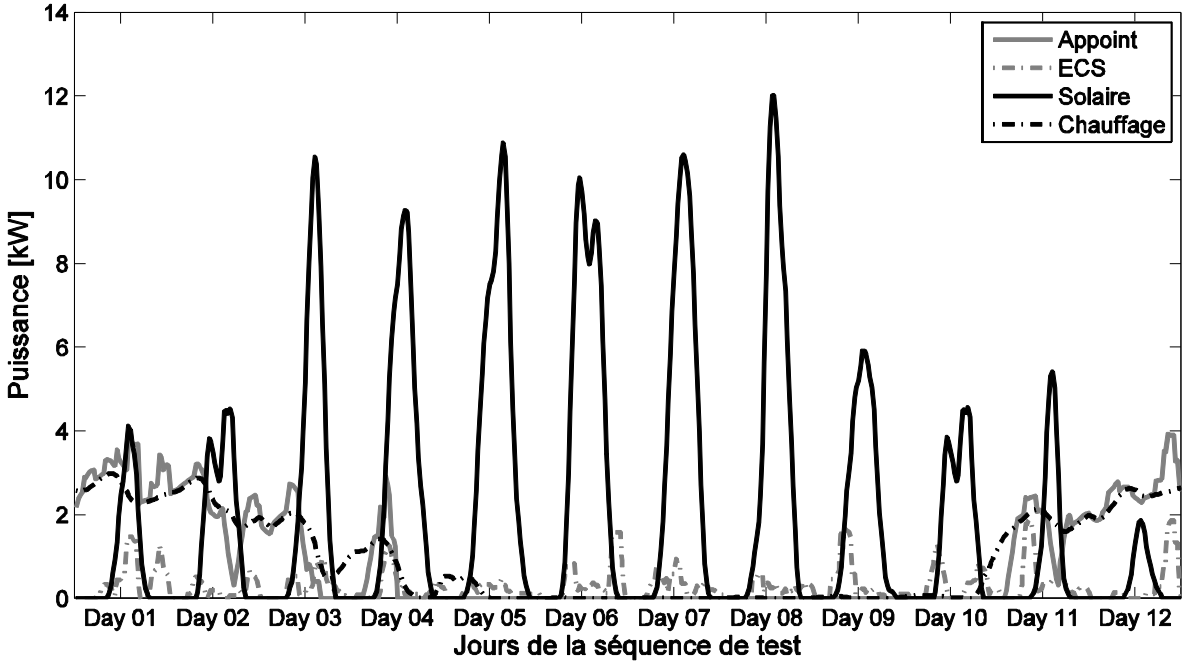

Figure 2- 37 : Données d'apprentissage

A l'issue de l'application de la méthodologie de modélisation développée, 2700 (15x15x3x(3+1)) modèles (Il est possible que deux modèles issus de deux initialisations différentes soient exactement les mêmes) avec différentes caractéristiques sont créés :

- Nombre de neurones : 1 à 15 (pour les deux FA)
- Nombre de réinitialisations : 15 (pour les deux FA)
- Nombre de TD : 3 (pour les deux FA)
- Nombre d'intervalle de normalisation :
  - 3 dans le cas de FA tanh
  - un seul dans le cas de FA linéaire

Dans les paragraphes suivants, nous allons investiguer leurs performances ainsi que la possibilité de sélection des RNA les plus pertinents à l'aide des critères statistiques présentés dans le paragraphe II.5.4.1. Les résultats sont présentés en fonction de la FA utilisée dans la couche de sortie.

Dans le Tableau 2- 8 et le Tableau 2- 9 sont présentés les modèles sélectionnés selon les différents critères statistiques ainsi que leur performance de prédiction, après les avoir testés dans différentes conditions, évalué à l'aide de « *R2nrj* ». Les modèles étant censés être utilisés en mode simulation dans le cadre de la méthodologie d'évaluation, les résultats présentés ici sont issus d'un calcul en boucle fermée. En outre, un tel calcul est intéressant car il permettra de repérer les modèles qui ne sont pas capables de gérer la propagation d'erreurs en fonction du temps. En effet, en mode simulation, parmi les entrées du modèle figurent des sorties retardées calculées précédemment par le modèle lui-même. Ces entrées étant entachées d'erreurs (car prédites), il faut que le modèle soit capable de les gérer sinon les erreurs vont se propager au cours du temps et ainsi fausser les prédictions des valeurs futures de la consommation du système.

Tableau 2- 8 : Modèles sélectionnés selon les différents critères statistiques ainsi que leur performance de prédiction évaluée à l'aide de « R2nrj » après les avoir testés dans différentes conditions. Cas de la FA $tanh$

Critère de sélection	*BIC*	*BIC1*	*BIC2*	*AIC*	*AIC1*	*AIC2*	*AICC*	*ME*	*MAE*	*ErrTot*	*R2bar*	*R2*	*Sign*	*d*	*DA*	*RMSE*
Nombre de neurones	14	1	14	14	2	14	14	14	14	14	14	14	1	14	8	14
« R2nrj »	0,960	0,886	0,960	0,960	0,462	0,960	0,960	0,888	0,960	0,888	0,960	0,960	0,940	0,960	0,944	0,960
TD	1	1	1	1	1	1	1	3	1	3	1	1	1	1	3	1
Intervalle de normalisation	±1	±1	±1	±1	±0,4	±1	±1	±1	±1	±1	±1	±1	±0,2	±1	±1	±1

Tableau 2- 9 : Modèles sélectionnés selon les différents critères statistiques ainsi que leur performance de prédiction évaluée à l'aide de « R2nrj » après les avoir testés dans différentes conditions. Cas de la FA linéaire

Critère de sélection	*BIC*	*BIC1*	*BIC2*	*AIC*	*AIC1*	*AIC2*	*AICC*	*ME*	*MAE*	*ErrTot*	*R2bar*	*R2*	*Sign*	*d*	*DA*	*RMSE*
Nombre de neurones	10	1	11	10	2	11	10	10	15	10	11	11	1	11	15	11
« R2nrj »	0,607	0,924	0,976	0,607	0,765	0,976	0,607	0,953	0,931	0,953	0,976	0,976	0,923	0,976	0,688	0,976
TD	1	1	3	1	1	3	1	1	2	1	3	3	1	3	2	3
Intervalle de normalisation	±1	±1	±1	±1	±1	±1	±1	±1	±1	±1	±1	±1	±1	±1	±1	±1

Les résultats montrent que les modèles sélectionnés sont différents, que ce soit pour les modèles avec une FA linéaire ou non en sortie. Certains ensembles de critères, par exemple : (« *BIC* », « *AIC* », « *AICC* »), (« *BIC2* », « *AIC2* », « *R2bar* », « d », « *RMSE* » et « *R2* ») dans le cas d'une FA linéaire et (« *BIC* », « *AIC* », « *AICC* », « *BIC2* », « *AIC2* », « *R2bar* », « *R2* », « *RMSE* », « *d* » et « *MAE* ») dans le cas d'une FA tangente hyperbolique, ont sélectionné le même modèle. Les performances réelles des modèles sélectionnés, mesurées à l'aide de « *R2nrj* », comme expliqué dans le sous-paragraphe II.5.4.1, sont très satisfaisantes dans le cas des critères « *BIC2* », « *AIC2* », « *R2bar* », « *d* », « *RMSE* » et « *R2* » contrairement à celles des modèles sélectionnés par les autres critères. En effet, pour les critères restants (« *BIC* », « *BIC1* », « *AIC* », « *AIC1* », « *AICC* », « MAE », « ME », « ErrTot », « *Sign* » et « DA ») les performances des modèles sélectionnés sont soit bonnes mais uniquement pour l'une des deux FA, soit leur « *R2nrj* » est inférieur à 0,96.

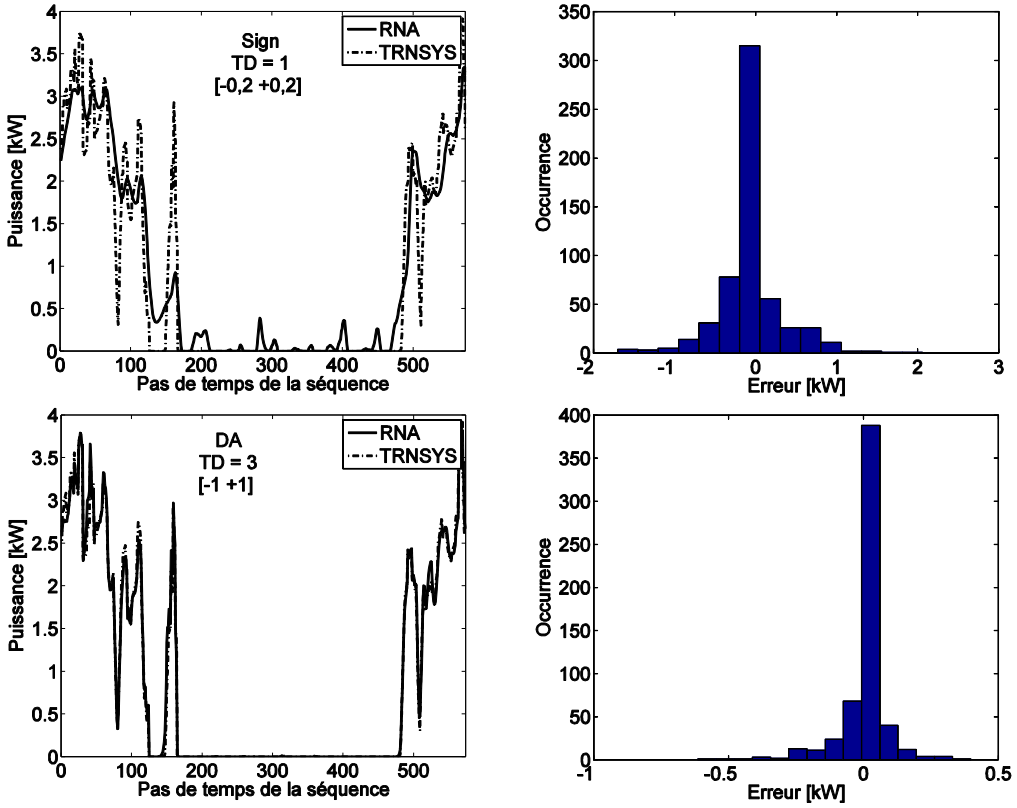

Figure 2- 38 : Comparaison entre les prédictions du RNA sélectionné en fonction de chaque critère (indiqués sur chaque figure) et la simulation du test du système sous TRNSYS (à gauche). L'histogramme des erreurs est donné à droite. Sur chaque figure est indiqué également l'intervalle de normalisation utilisé (±0,2 ; ±0,4 ou ±1,0). Cas de la FA $tanh$

L'analyse des modèles sélectionnés sur la base des données d'apprentissage (voir figures Figure 2- 38, Figure 2- 39, Figure 2- 40 et Figure 2- 41) montre qu'il y a globalement une très bonne concordance entre les prédictions des RNA sélectionnés et les calculs du modèle détaillé TRNSYS, sauf pour ceux sélectionnés à l'aide des critères « *ME* » ou « *ErrTot* » dans le cas de la FA tanh où le RNA présente des oscillations à la fin de la séquence. Ce modèle a été néanmoins sélectionné car ces deux critères ne calculent pas la somme des erreurs en valeur absolue, ce qui fait que les erreurs ayant la même amplitude peuvent se compenser. Les histogrammes des erreurs pour tous les modèles sélectionnés sont globalement centrés sur la valeur 0, ce qui montre que les erreurs de ces modèles sélectionnés par les différents

critères suivent une loi normale. Il est à noter que cette hypothèse est nécessaire pour l'application de la régularisation bayésienne, et rentre en jeu pour le calcul des paramètres α et δ dans l'équation (2- 31).

Les critères « *BIC1* », « *AIC1* » présentent dans leurs expressions un terme de pénalité quadratique par rapport à la complexité du modèle (nombre de paramètres) : $\frac{q^2 \, log(N)}{N}$ et $2\frac{q^2}{N}$ respectivement pour « *BIC1* » et « *AIC1* ». Pour deux RNA ayant un nombre de neurones ou d'entrées différents, la différence entre leur nombre de paramètres est au minimum égale à 4 (elle est égale à $E + 1$ dans le cas d'une différence d'un neurone et $E \times C$ dans le cas d'une différence d'un seul retard). Ce constat explique pourquoi le critère « BIC1 » a sélectionné un modèle avec seulement 1 neurone et le critère « AIC1 » un RNA à 2 neurones. Les autres modèles avec plus de neurones ou d'entrées ont été fortement pénalisés à cause de leur complexité.

Evidemment à partir d'un certain nombre de paramètres, le terme pénalisant la complexité du modèle devient prépondérant par rapport au terme représentant les erreurs du modèle. Or ce nombre de paramètre critique est atteint assez rapidement dans le cas étudié. Il est donc impossible via « AIC1 » et « BIC1 » de sélectionner des réseaux ayant plus de neurones ou d'entrées. Les performances des réseaux sélectionnés par ces critères ne sont pas bonnes (« *R2nrj* » inférieure à 0,92 pour les deux FA). Il est donc primordial de pouvoir choisir des RNA ayant des complexités plus élevés si nécessaire.

Compte tenu de cette analyse et en gardant seulement les critères ayant sélectionné des modèles avec de bonnes performances pour les deux FA, les critères les plus pertinents sont : « *BIC2* », « *AIC2* », « *R2bar* », « *d* », « *RMSE* » et « *R2* ».

Dans le but d'investiguer la possibilité de garder un seul critère pour la sélection des modèles, une étude statistique de la population des modèles créés en fonction de chaque critère candidat et du « *R2nrj* » a été réalisée. Les résultats de cette étude sont représentés sur les Figure 2- 42, Figure 2- 43, Figure 2- 44 et Figure 2- 45.

Les figures montrent qu'il y a une tendance pour chaque critère à sélectionner un bon modèle (« *R2nrj* » proche de 1). En effet, la plupart des RNA susceptibles d'être sélectionnés (minimisant ou maximisant le critère en question) sont agglomérés au niveau des valeurs les plus proches de 1 pour « *R2nrj* ».

Il est intéressant de signaler que plusieurs modèles, de l'ordre de 2%, susceptibles d'être sélectionnés par « *R2* », « *RMSE* » et « *d* » (valeur proche de 1, 488 ou 57 en fonction du critère) ont de très mauvaises performances de généralisation (« R2nrj » proche de 0). Ce phénomène est beaucoup moins présent chez les critères dépendant du nombre de paramètres du RNA. La sélection de modèles selon « *R2* », « *RMSE* » ou « *d* » présente en conséquence un risque élevé de sélectionner un mauvais modèle. Il est toujours possible de créer un RNA avec de très faibles erreurs (et donc un très bon « *R2* », « *RMSE* » et « d ») à condition d'utiliser un nombre suffisamment élevé de neurones dans la couche cachée. Par contre, ces modèles auront de fortes chances de présenter un problème de sur-apprentissage. C'est pour cette raison qu'il est primordial de choisir un critère intégrant un terme de pénalisation pour les modèles complexes (ayant beaucoup de paramètres c.-à-d. de neurones dans la couche cachée). Les critères « *R2* », « *RMSE* » « d » sont donc à écarter parmi les autres candidats.

Enfin, en l'absence d'une réelle supériorité d'un des trois critères restant (« *AIC2* », « *BIC2* » et « *R2bar* ») nous allons tous les garder. Les résultats des applications de la méthodologie pour de nouveaux systèmes permettront d'affiner le choix du critère le plus pertinent. Les critères retenus, « *AIC2* » et « *BIC2* » pénalisent la complexité des RNA avec un terme en racine carrée. La seule différence entre ces deux critères réside dans la pondération de ce terme. La pénalisation va être légèrement plus faible avec « AIC2 ». Puisque la plage de variation du terme logarithmique est plus importante (variation entre 12 et 22 pour le SSC présenté dans ce chapitre) que celle du deuxième terme dépendant des paramètres du RNA (variation entre 0,03 et 0,16 pour « BIC2 » ; 0,009 et 0,05 pour « AIC2 »), il est fort probable que ces deux critères sélectionnent les mêmes modèles, comme pour ce premier SSC étudié. Dans l'application de la méthodologie dans le cas des autres systèmes nous allons pouvoir apprécier l'influence de ce terme de pénalisation.

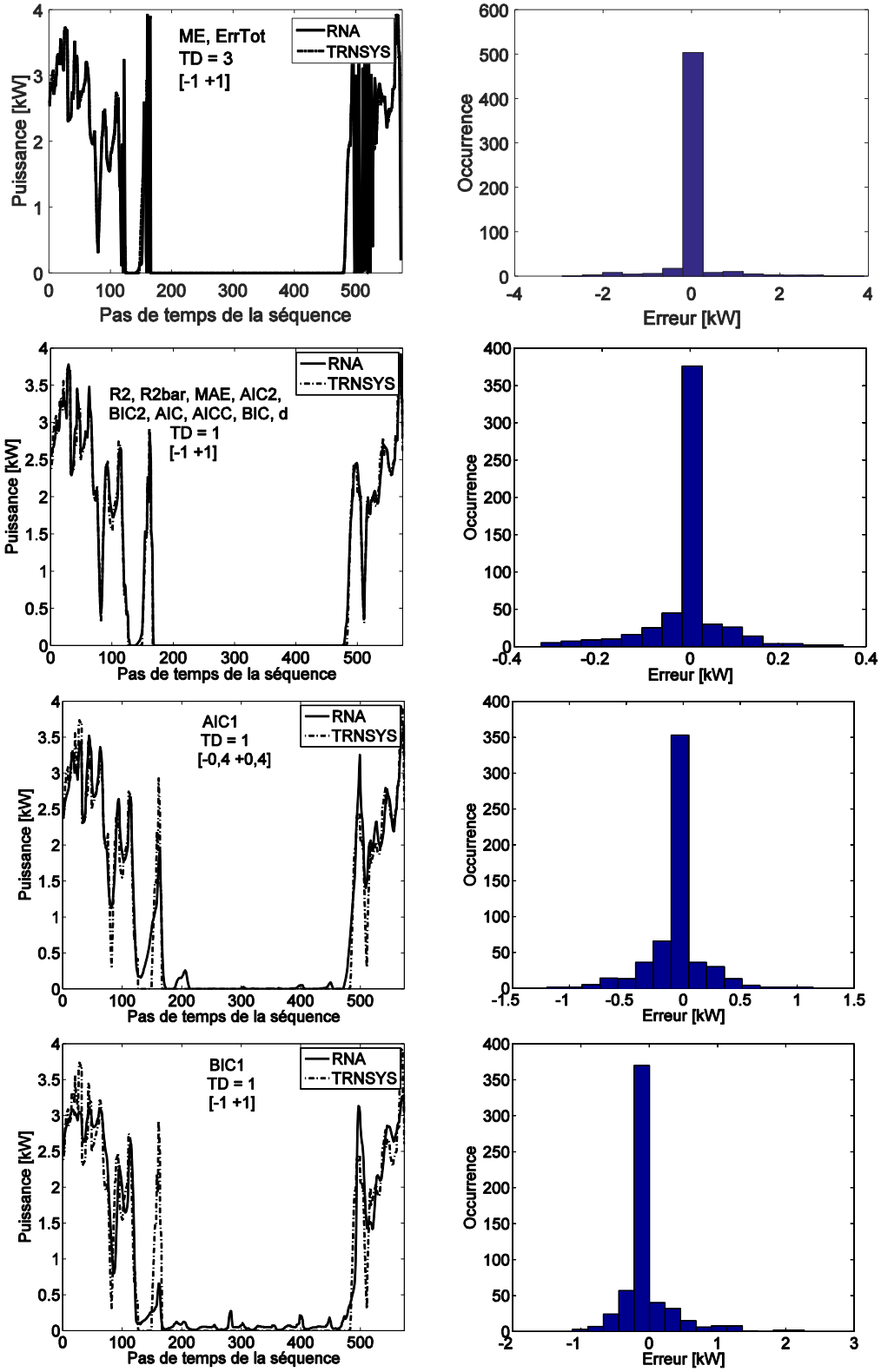

Figure 2- 39 : Comparaison entre les prédictions du RNA sélectionné en fonction de chaque critère (indiqués sur chaque figure) et la simulation du test du système sous TRNSYS (à gauche). L'histogramme des erreurs est donné à droite. Sur chaque figure est indiqué également l'intervalle de normalisation utilisé (±0,2 ; ±0,4 ou ±1,0). Cas de la FA $tanh$. Suite de la Figure 2- 38

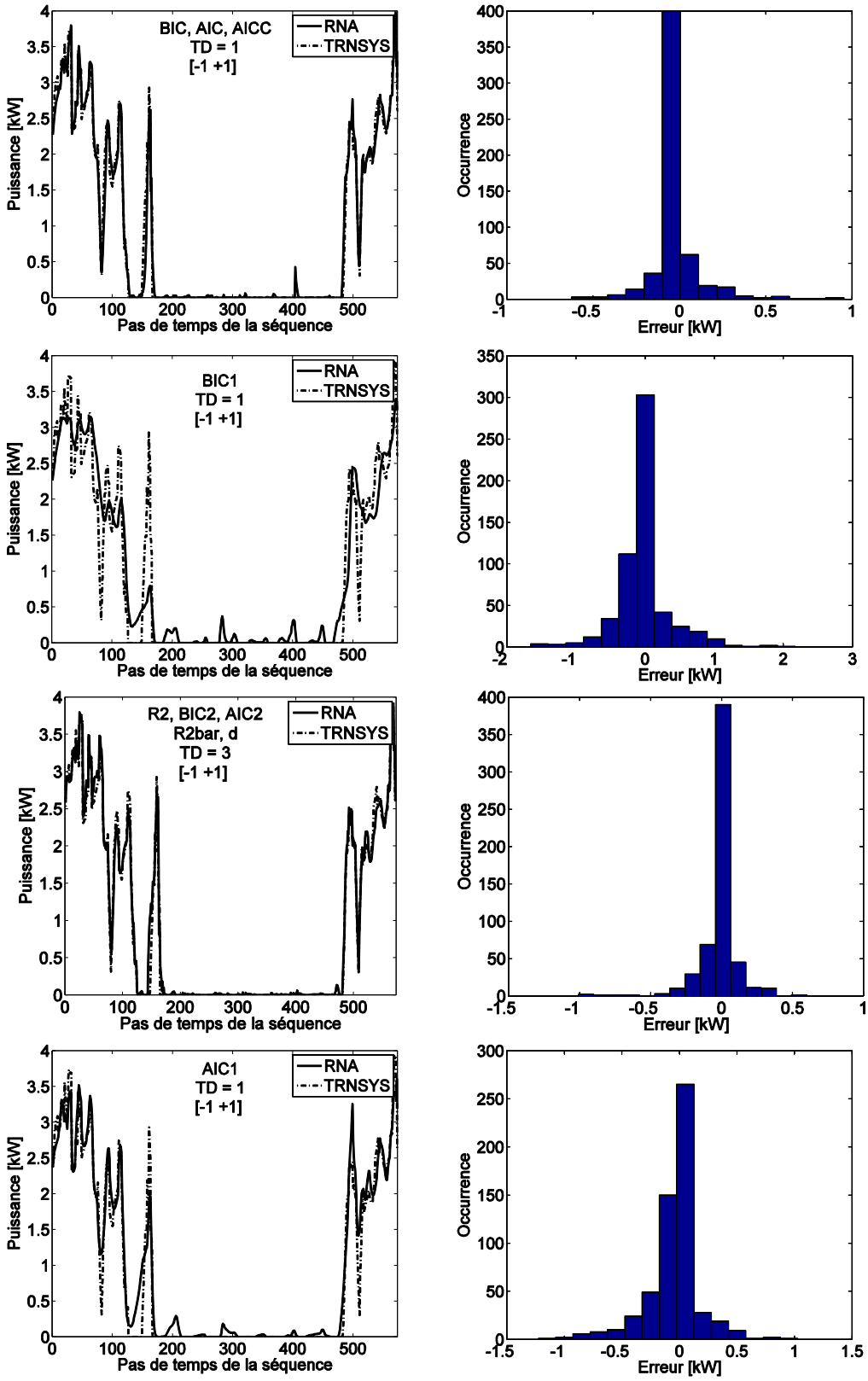

Figure 2- 40 : Comparaison entre les prédictions du RNA sélectionné en fonction de chaque critère (indiqués sur chaque figure) et la simulation du test du système sous TRNSYS (à gauche). L'histogramme des erreurs est donné à droite. Sur chaque figure est indiqué également l'intervalle de normalisation utilisé (±0,2 ; ±0,4 ou ±1,0). Cas de la FA $linéaire$

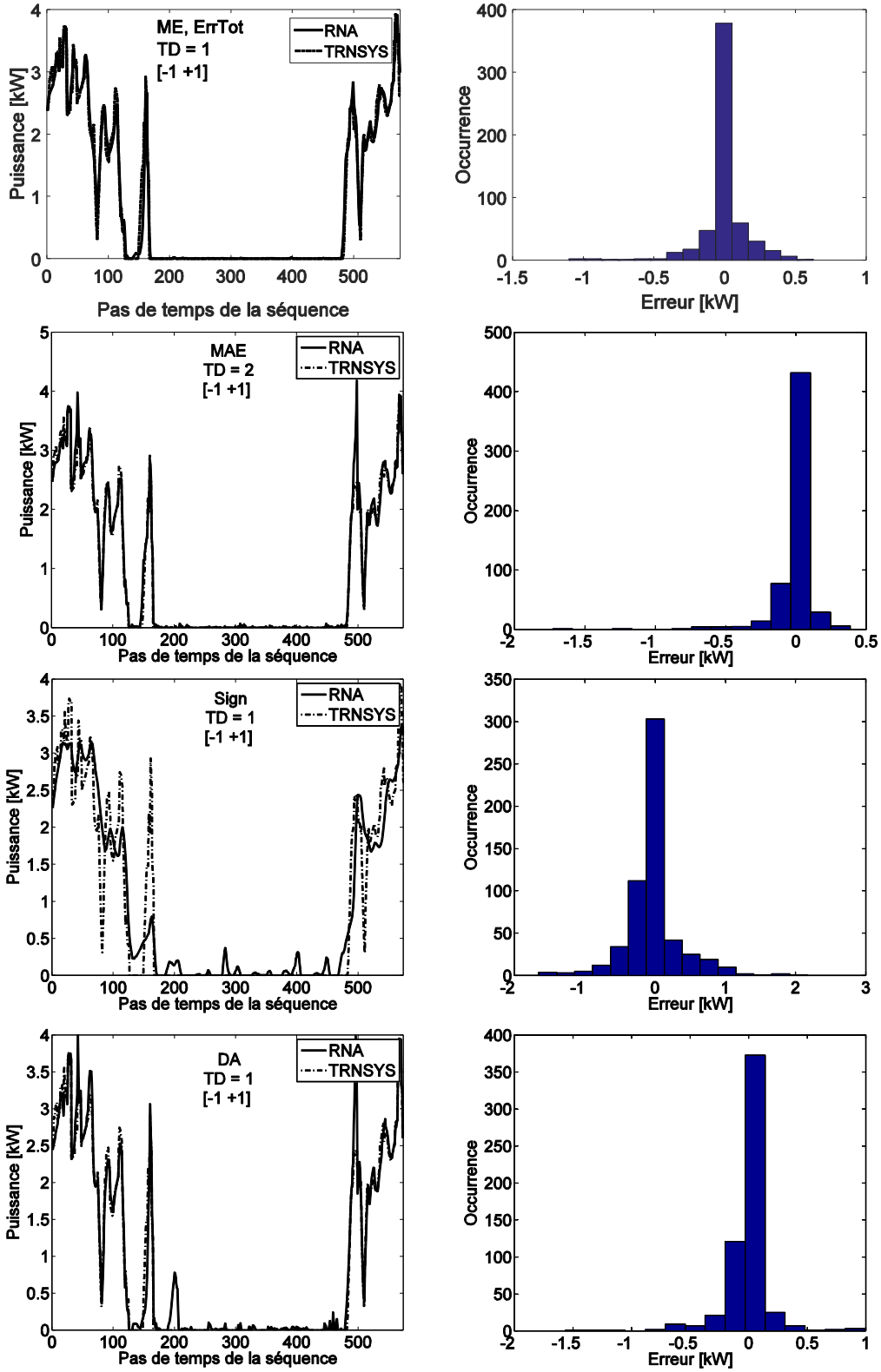

Figure 2- 41 : Comparaison entre les prédictions du RNA sélectionné en fonction de chaque critère (indiqués sur chaque figure) et la simulation du test du système sous TRNSYS (à gauche). L'histogramme des erreurs est donné à droite. Sur chaque figure est indiqué également l'intervalle de normalisation utilisé (±0,2 ; ±0,4 ou ±1,0). Cas de la FA $linéaire$. Suite de la Figure 2- 40.

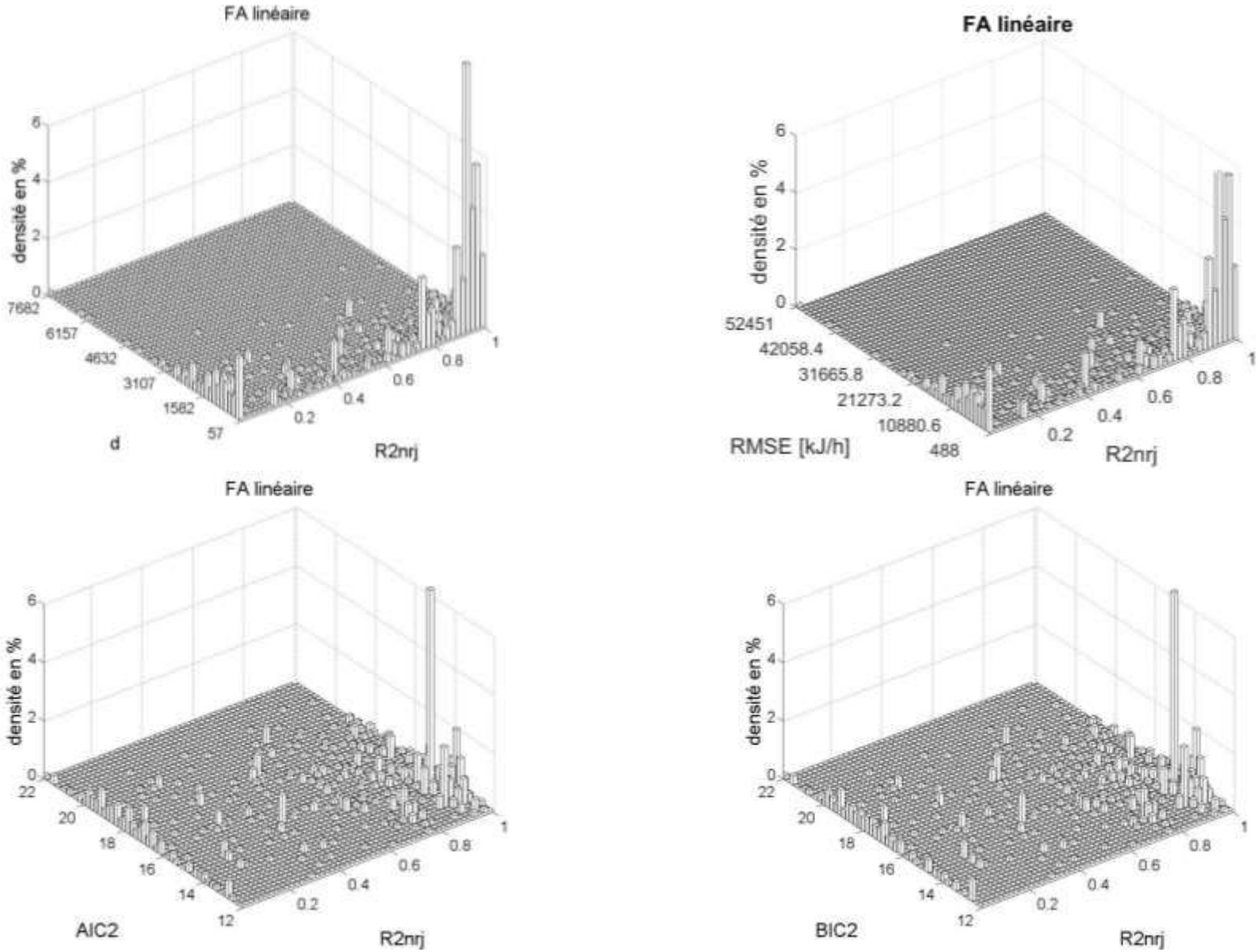

Figure 2- 42 : Histogramme illustrant la densité des RNA en fonction de chaque couple de critères. Cas de la FA *linéaire* et des critères *AIC2, d, RMSE et BIC2.* Tous les critères sont sans dimension sauf RMSE.

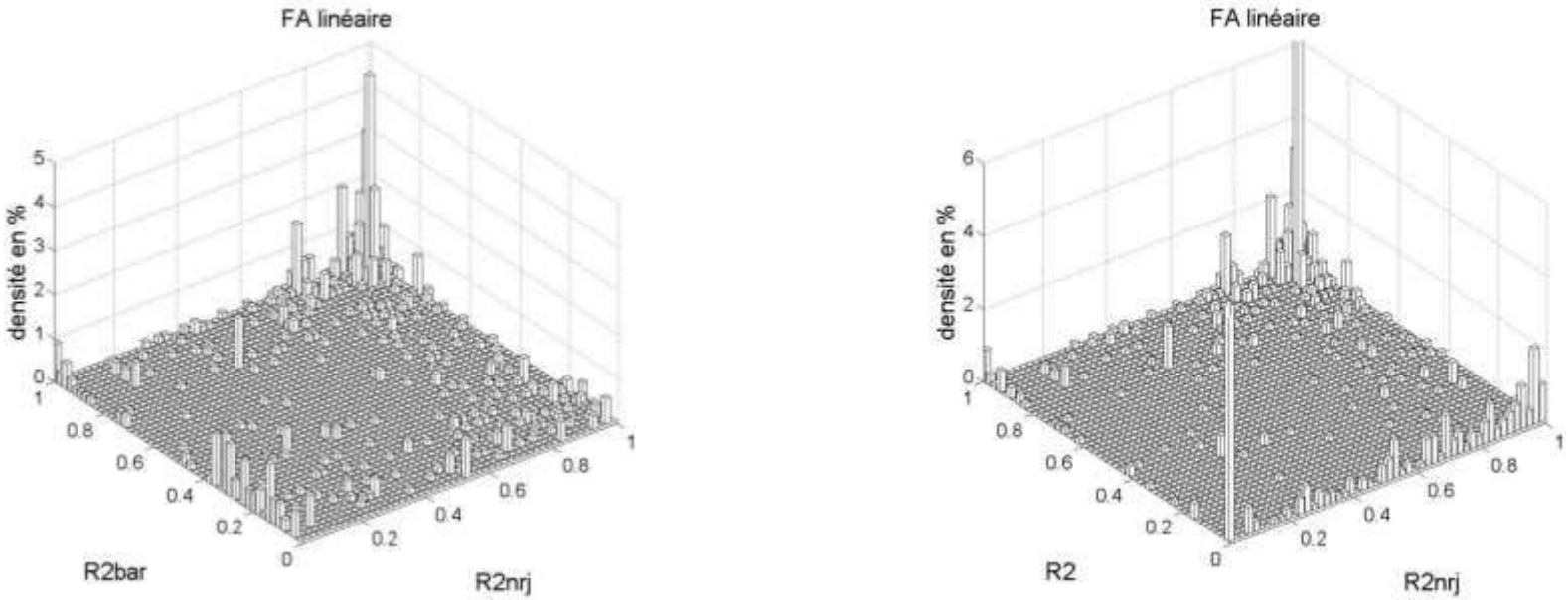

Figure 2- 43 : Histogramme illustrant la densité des RNA en fonction de chaque couple de critères. Cas de la FA $linéaire$ et des critères R2bar et R^2. Tous les critères sont sans dimension sauf RMSE.

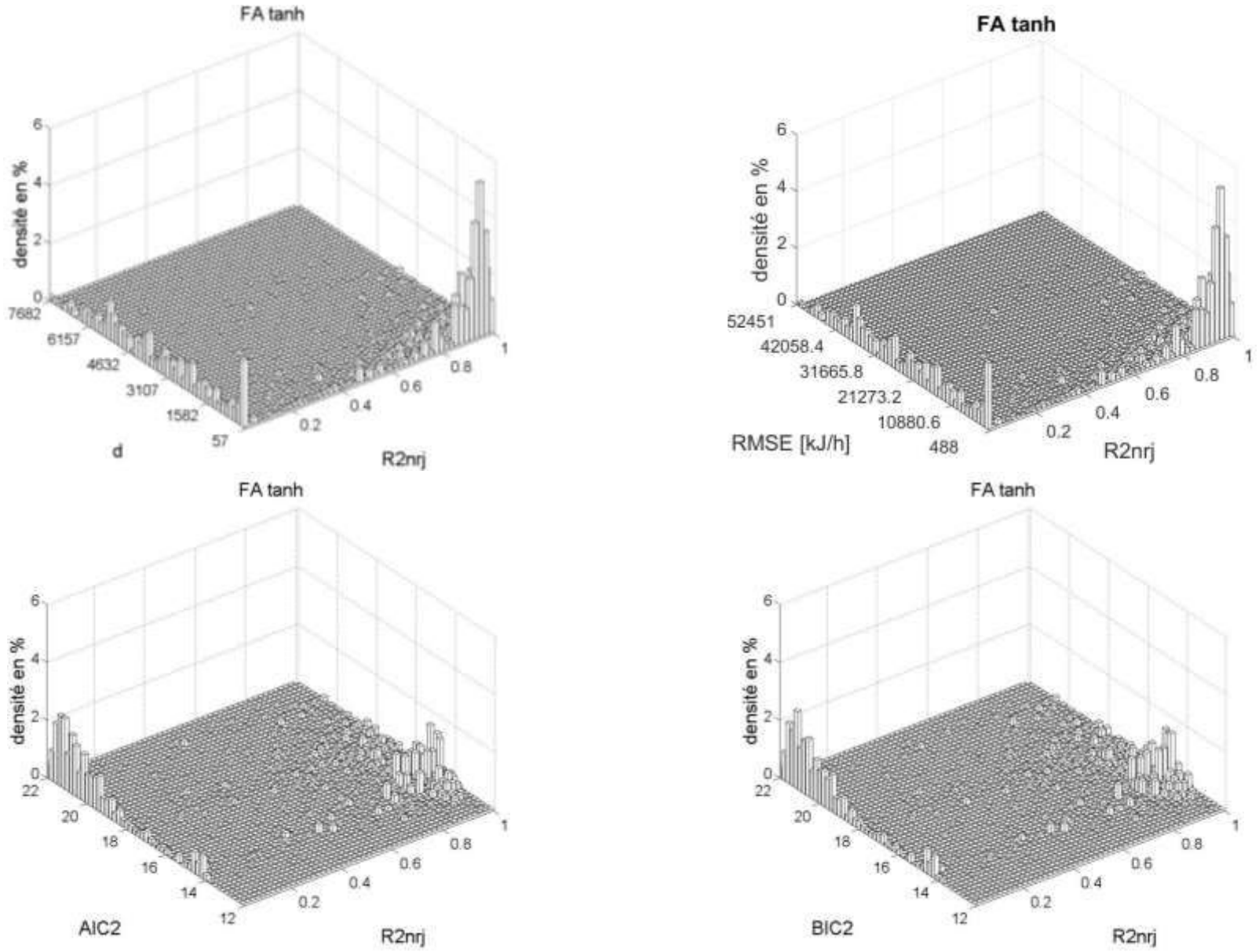

Figure 2- 44 : Histogramme illustrant la densité des RNA en fonction de chaque couple de critères. Cas de la FA tanh et des critères *d*, *AIC2*, *RMSE* et *BIC2.* Tous les critères sont sans dimension sauf RMSE.

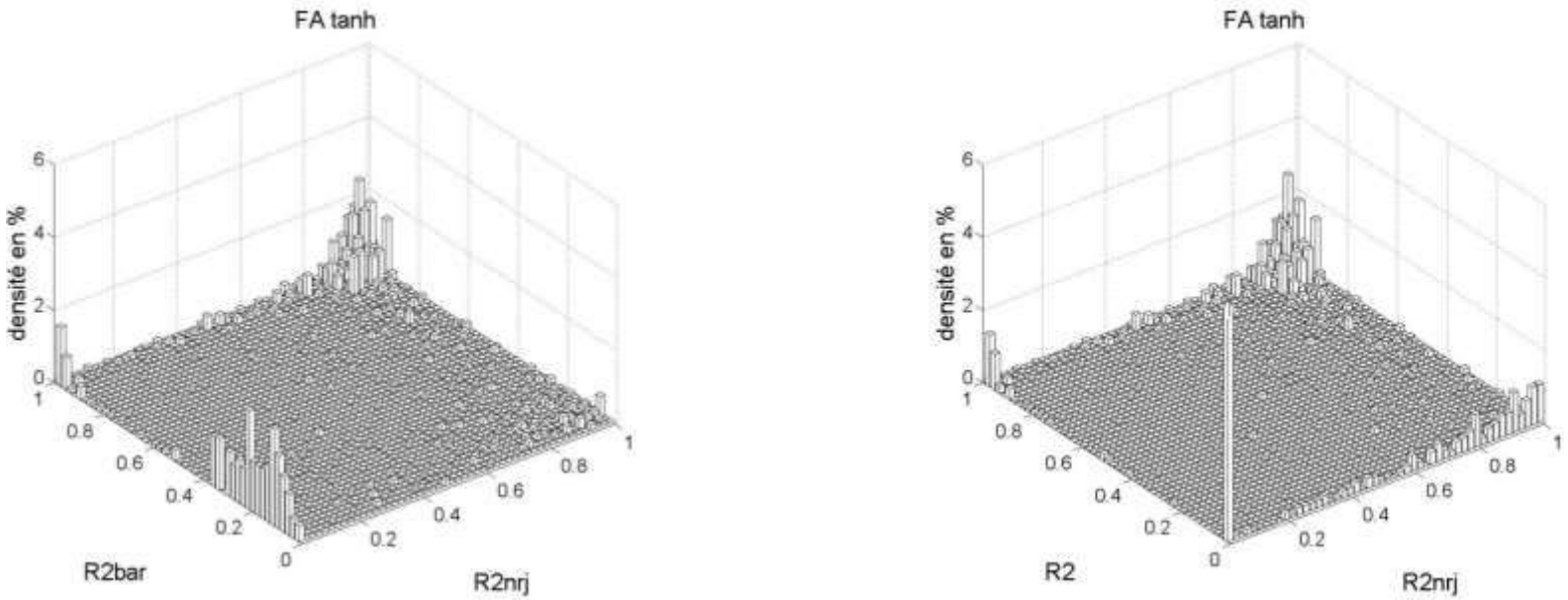

Figure 2- 45 : Histogramme illustrant la densité des RNA en fonction de chaque couple de critères. Cas de la FA *tanh* et des critères R2bar et R^2. Tous les critères sont sans dimension sauf RMSE.

La Figure 2- 46 permet de résumer les différences entre les deux catégories de modèles en fonction de leur FA. Il ressort que le pourcentage du nombre de modèles avec un bon pouvoir de généralisation (« *R2nrj* »>0,95) dans le cas de la FA linéaire est supérieur par rapport à l'autre cas. Par contre, le nombre de modèles avec un très bon pouvoir de généralisation (« *R2nrj* »>0,99) est plus élevé dans le cas de la tangente hyperbolique. Une fonction linéaire comme FA permet donc de minimiser le risque de sélectionner un mauvais RNA mais il réduit les chances de sélectionner un très bon modèle. Nous gardons donc les deux fonctions d'activation et nous sélectionnons les modèles selon les critères statistiques sans privilégier aucune des deux a priori.

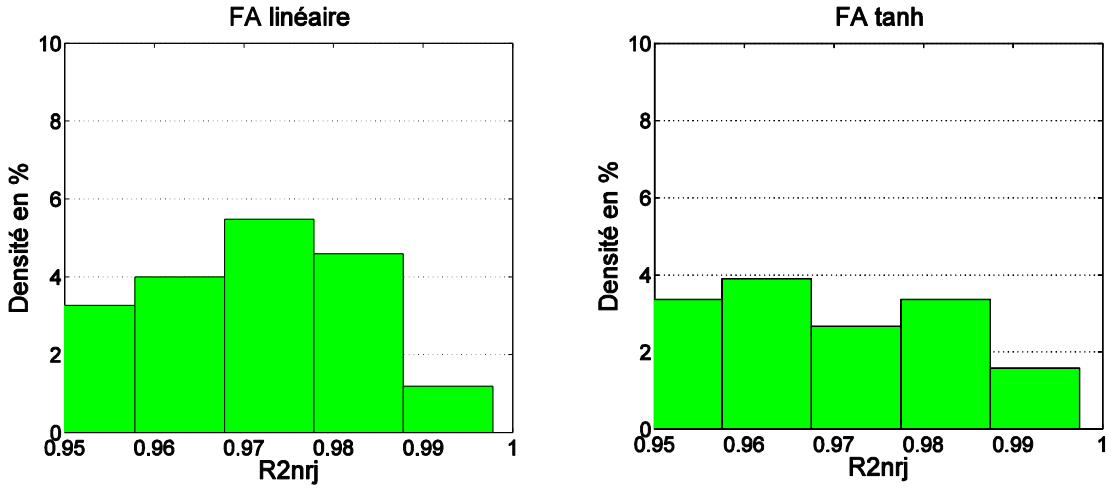

Figure 2- 46 : Histogramme illustrant la densité des RNA (nombre de neurones dans la couche cachée) et la valeur du « *R2nrj* » correspondante.

II.5.5. Analyse des résultats d'application de la méthodologie

II.5.5.1 Résumé du processus de l'application de la méthodologie

Le système à évaluer est, dans un premier temps, testé sur le banc d'essai semi-virtuel dans des conditions bien spécifiques. Ces conditions doivent permettre « l'excitation » du système de façon à ce qu'une grande partie de son domaine de fonctionnement soit explorée. Dans le cadre de ces travaux de recherche, les conditions de test qui répondent à cette exigence sont « Zurich-SFH60-16m² ». D'autres conditions ont été également explorées. Elles seront précisées le cas échéant. Les tests réels des systèmes étant complexes et couteux notamment en termes de temps, ces derniers sont simulés (expérimentation numérique) sauf dans le chapitre expérimentations physiques (quatrième chapitre).

Les variables qu'il est possible de mesurer aux bornes du système pendant le test sont données sur le Tableau 2- 10.

Tableau 2- 10 : Grandeurs nécessaire pour la formation des entrées et sortie du modèle neuronal

Climat	ECS	Appoint	Emetteur
G	$\dot{m}_{dhw}$ T_{tap} T_{dhw}	P_{aux}	$\dot{m}_{sh}$ $T_{sh,in}$ $T_{sh,out}$

Les différentes variables nécessaires pour former les entrées et la sortie de chaque RNA sont donc calculées avec les données mesurées (avec les données de la simulation du test dans le cas échéant).

Le processus d'apprentissage (présenté dans le paragraphe II.3.5.) est ensuite appliqué aux données ainsi formées. Une sélection est ensuite réalisée selon les critères statistiques. Les paramètres utilisés dans ce processus sont donnés dans le Tableau 2- 11. Sauf mention contraire, ces valeurs sont utilisées pour tous les systèmes.

Tableau 2- 11 : Paramètres utilisés dans le programme de modélisation par RNA

Pas de temps du modèle	Nombre de réinitialisation	Nombre d'intervalle de normalisation	Nombre de TD	Nombre de neurones maximal	Nombre de FA en sortie explorés
30 min	15	3 (sauf pour AF linéaire pour lequel un seul intervalle est considéré)	3 (1, 2 et 3)	15	2 (*linéaire* et $tanh$)

Tout le processus de modélisation abordé dans ce mémoire a été développé dans MATLAB, 2012b avec l'aide de l'outil Neural Network Toolbox. Les algorithmes d'optimisation ainsi que de réinitialisation utilisent des générateurs de nombre aléatoires. Pour que le processus d'entrainement soit reproductible nous réinitialisons ce générateur, de la même façon pour tous les systèmes et avant chaque application du processus.

II.5.5.2 Résultats

Dans la suite du document, deux catégories de modèles seront considérées pour présenter les résultats de l'application de la méthodologie :

- En premier lieu, le modèle qui présente les meilleurs performances selon le « R2nrj » (telles que définies dans le sous-paragraphe II.5.4.1). Il sera noté RNA-R2nrj.
- Puis les modèles sélectionnés selon les critères « BIC2 », « AIC2 » et « R2bar ». Ils seront notés respectivement RNA-BIC2, RNA-AIC2 et RNA-R2bar.

Ces modèles sont présentés à la fois pour les réseaux ayant une fonction d'activation tanh et linéaire. La FA est précisée entre parenthèse dans le nom du réseau présenté.

Les autres caractéristiques (nombre de neurones, TD etc.) des modèles RNA seront également indiquées.

La présentation des résultats correspondants aux modèles RNA-R2nrj (meilleurs modèles parmi tous les RNA créés) nous permettra de savoir si la méthodologie est capable de produire des modèles pertinents et si les modèles sélectionnés par les critères statistiques ont les mêmes capacités de prédiction que les meilleurs modèles. Nous rappelons ici que « R2nrj » est calculé sur la base de comparaisons avec un modèle détaillé TNRNSYS qui n'est pas prévu dans la méthodologie finale mais il sert ici uniquement à l'analyse et évaluation de la méthodologie. En revanche les modèles déterminés à l'aide des critères statistiques correspondent à ce qu'on obtiendrait en suivant la méthodologie à la lettre.

Sur la Figure 2- 47 sont présentées les évolutions de la puissance de la chaudière en fonction du temps, calculées par le modèle TRNSYS et estimées par les modèles neuronaux

pour les données d'apprentissage. Les caractéristiques des quatre modèles sont données dans le Tableau 2- 12.

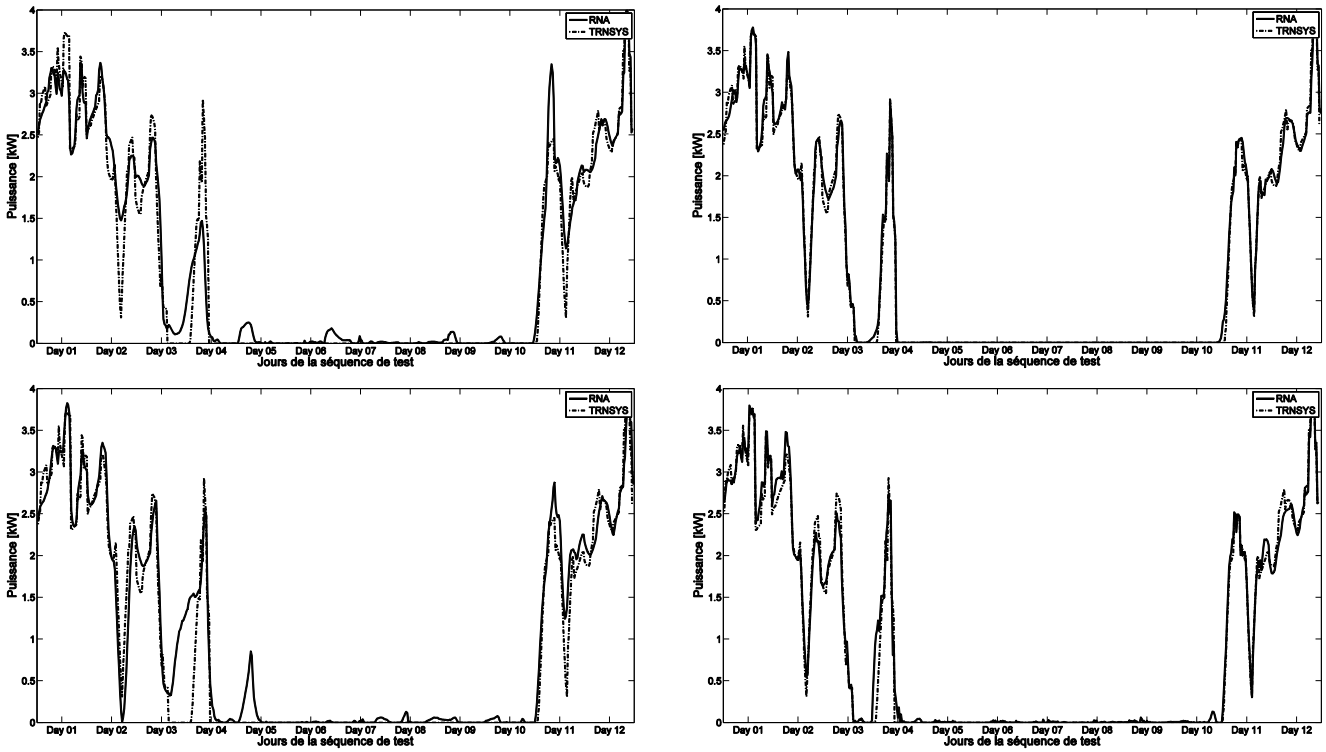

Figure 2- 47 : Performances des modèles sélectionnés sur les données d'apprentissage. A gauche les modèles RNA-R2nrj(tanh) en haut et RNA-R2nrj(linéaire) en bas, A droite les modèles RNA-AIC2(tanh) en haut et RNA- AIC2(linéaire) en bas

Tableau 2- 12 : Caractéristiques des RNA sélectionnés

Modèles	NORM	NDN	TD	« R2nrj »	« R2bar »	« BIC2 »	« AIC2 »
RNA-R2nrj(tanh)	±0,2	3,0	2,0	0,9974	0,9572	13,7402	13,6981
RNA-R2bar(tanh)	±1,0	14,0	1,0	0,9603	0,9965	11,4141	11,3444
RNA-BIC2(tanh)	±1,0	14,0	1,0	0,9603	0,9965	11,4141	11,3444
RNA-AIC2(tanh)	±1,0	14,0	1,0	0,9603	0,9965	11,4141	11,3444
RNA-R2nrj(linéaire)	±1,0	5,0	1,0	0,9978	0,9406	14,0674	14,0253
RNA-R2bar(linéaire)	±1,0	11,0	3,0	0,9765	0,9910	12,5218	12,4276
RNA-BIC2(linéaire)	±1,0	11,0	3,0	0,9765	0,9910	12,5218	12,4276
RNA-AIC2(linéaire)	±1,0	11,0	3,0	0,9765	0,9910	12,5218	12,4276

Quel que soit le modèle neuronal, la corrélation entre les deux courbes (RNA et TRNSYS sur la Figure 2- 47) est très bonne. L'histogramme des erreurs, Figure 2- 48, pour chaque cas peut très bien être approximé par une loi normale de moyenne nulle ce qui montre la pertinence des deux modèles RNA-R2nrj et les RNA sélectionnés par les trois critères pour les deux FA. Les critères statistiques ont sélectionnés le même modèle. Il est possible de constater aussi que sur la base des données d'apprentissage, le RNA sélectionné par ces critères pour les deux FA est plus précis que les modèles RNA-R2nrj. Ceci est certainement dû au nombre de neurones élevé au sein des deux RNA (Tableau 2- 12). En effet, avec un nombre de neurones élevé et donc de paramètres élevés (poids synaptiques), le modèle neuronal a plus de liberté pour se caler sur les données d'apprentissage. Cependant, un RNA avec de bonnes performances sur les données d'apprentissage n'est pas synonyme d'un RNA avec un bon pouvoir de généralisation.

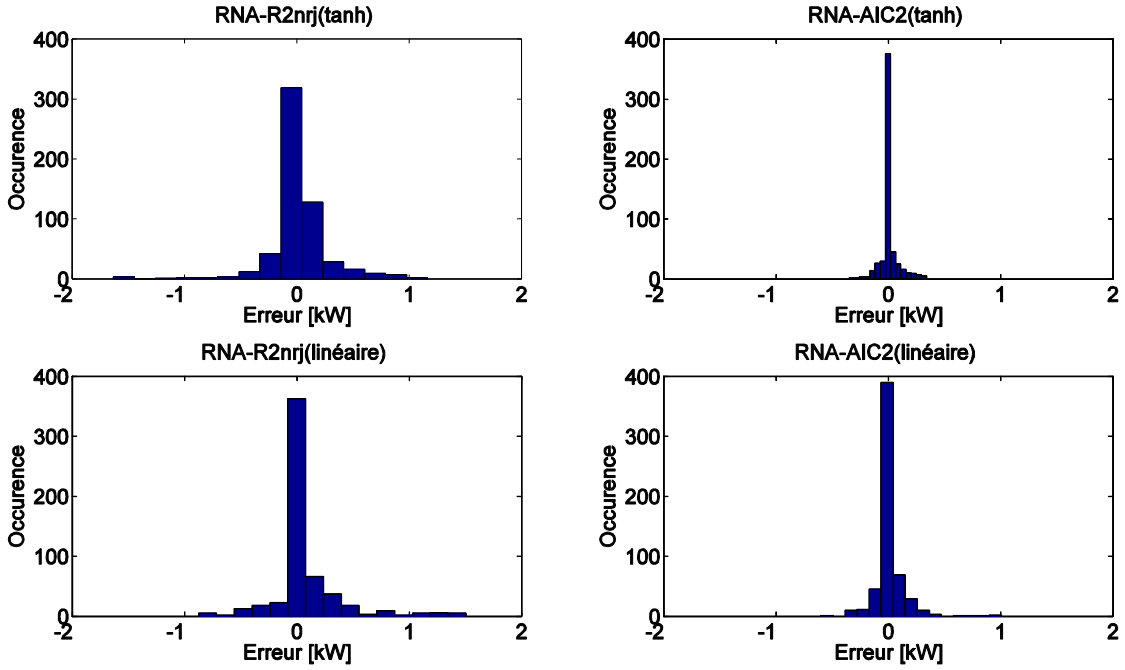

Figure 2- 48 : Histogrammes des erreurs des modèles sélectionnés. Le calcul est fait sur les données d'apprentissage

Les modèles présentés ci-dessus ont été utilisés pour prédire la consommation du système, par m² de surface de bâtiment, dans différentes conditions et pendant une année (voir protocole de validation dans la Figure 2- 19). Sur la Figure 2- 49 sont présentées les prédictions annuelles dans le cas de la FA tangente hyperbolique. Pour le modèle RNA-R2nrj(tanh) les consommations prédites sont très proches des valeurs de références calculées par le modèle TRNSYS. Le coefficient de détermination est supérieur 0,99. L'erreur relative commise reste inférieure à 8% pour toutes les conditions sauf pour Zurich-SFH30 où elle est légèrement supérieure à 10% (mais uniquement pour les deux surfaces de capteurs 10 m² et 13 m²).

Outre ses capacités à estimer la consommation annuelle du système, le modèle neuronal gère correctement l'influence de la surface du capteur sur cette dernière. En effet, plus la surface du capteur est élevée moins le système consomme de l'énergie.

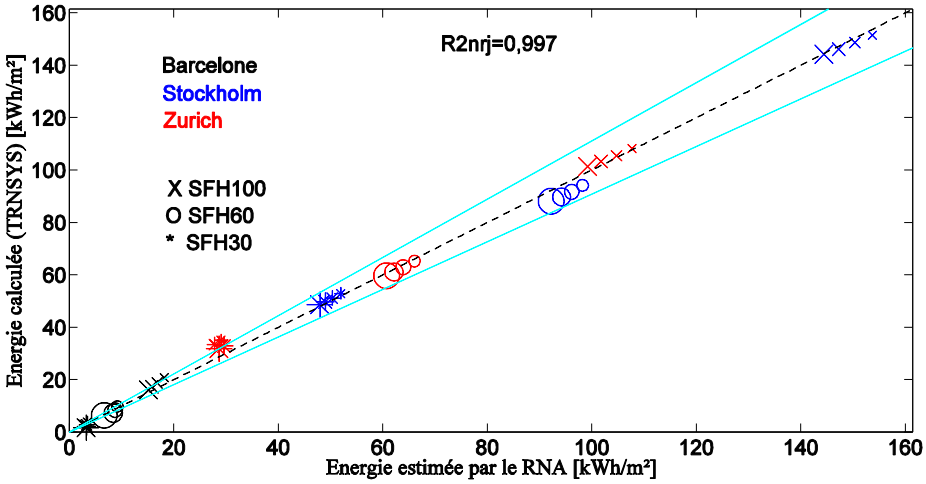

Figure 2- 49 : Comparaison entre la consommation annuelle estimée par le RNA-R2nrj(tanh) et calculée par le modèle physique TRNSYS dans différentes conditions. La taille des symboles indique la surface du capteur utilisée (10 m², 13 m², 16 m² ou 19 m²).

Sur la Figure 2- 50 sont représentées les deux courbes de puissance de la chaudière, calculées avec les deux modèles RNA-R2nrj(tanh) et TRNSYS pour l'environnement « Barcelone-SFH100-19m² ». Même pour ce climat éloigné des conditions de test, la concordance globale entre les deux courbes est bonne.

Malgré le fait que le modèle neuronal soit utilisé pour la prédiction à long terme en boucle fermée, il est capable de bien gérer ses propres erreurs (étant donné que la sortie bouclée en entrée est une estimation).

Puisque l'objectif de la méthodologie est d'utiliser les modèles RNA pour calculer l'énergie sur la période de la simulation, les différences que nous pouvons constater en fonction du temps sont donc acceptables.

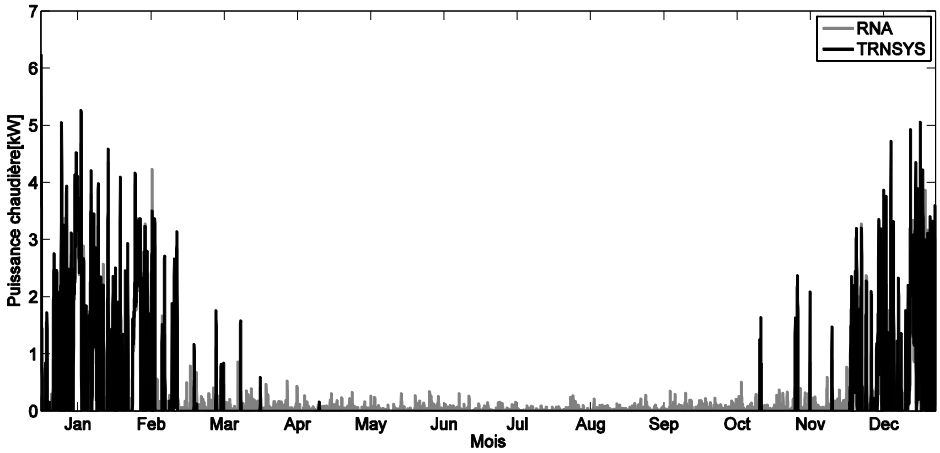

Figure 2- 50 : Evolution de la puissance de la chaudière calculée par le modèle neuronal et TRNSYS pour le cas de Barcelone-SFH100-19m². La FA est $tanh$.

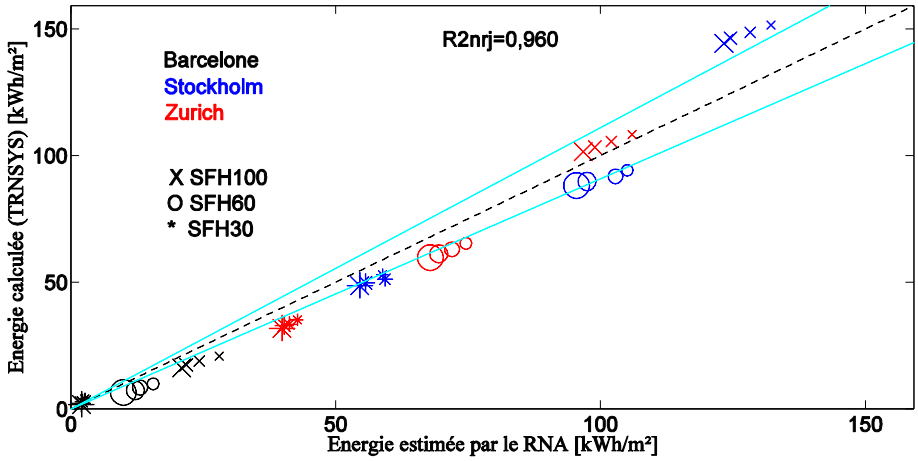

Figure 2- 51 : Comparaison entre la consommation annuelle estimée par le RNA-BIC2(tanh) et calculée par le modèle physique TRNSYS dans différentes conditions. La taille des symboles indique la surface du capteur utilisée (10 m², 13 m², 16 m² ou 19 m²).

Concernant le modèle RNA-BIC2(tanh) (Figure 2- 51), les résultats sont moins bons que pour le modèle RNA-R2nrj(tanh) mais restent tout à fait satisfaisants. Quelques écarts non négligeables cependant apparaissent notamment pour le climat de Stockholm.

Pour mieux comprendre la raison de cette amplification de l'erreur dans le cas de Stockholm, nous avons tracé les deux courbes de l'évolution de la puissance dans l'environnement « Stockholm-SFH100-16m² » sur la Figure 2- 52. La puissance de la chaudière prédite par le modèle RNA-BIC2(tanh) est majorée par une valeur d'environ 4kW. C'est pour cette raison que la consommation annuelle de l'énergie est sous-estimée. L'utilisation de la fonction tangente hyperbolique est la responsable de cette valeur limite en sortie du RNA. En effet, moyennant l'équation (2- 22) et l'inégalité triangulaire, et puisque la fonction tangente hyperbolique est bornée entre +1 et -1, l'estimation de la puissance consommée par l'appoint est dans ce cas bornée par deux valeur ψ et ϑ (relation (2- 39)) :

$$\psi < \hat{P}_{aux} < \vartheta \qquad \textbf{(2- 39)}$$

Où :

$$\psi = \frac{P_{aux}^{max} - P_{aux}^{min}}{nor_{max} - nor_{min}} \left(tanh\left(b_s^S - \sum_{c=1}^{C} \left|\omega_{C,s}^S\right| \right) - nor_{min} \right) + P_{aux}^{min} \qquad \textbf{(2- 40)}$$

$$\vartheta = \frac{P_{aux}^{max} - P_{aux}^{min}}{nor_{max} - nor_{min}} \left(tanh\left(b_s^S + \sum_{c=1}^{C} \left|\omega_{C,s}^S\right| \right) - nor_{min} \right) + P_{aux}^{min} \qquad \textbf{(2- 41)}$$

P_{aux}^{max} et P_{aux}^{min} sont respectivement les valeurs minimale et maximale de la variable P_{aux} dans les données issues de la séquence de test. L'intervalle de normalisation utilisé pour le modèle sélectionné par « *BIC2* » est ±1. Selon la relation (2- 39) et en élargissant encore les bornes puisque la tangente hyperbolique est bornée entre +1 et -1 nous obtenons :

$$P_{aux}^{min} < \hat{P}_{aux} < P_{aux}^{max} \qquad \textbf{(2- 42)}$$

Il est primordial de faire attention au fonctionnement en mode extrapolation des RNA. En conséquence il est indispensable d'explorer plusieurs intervalles de normalisation dans le but d'élargir les bornes ψ et ϑ même en extrapolation.

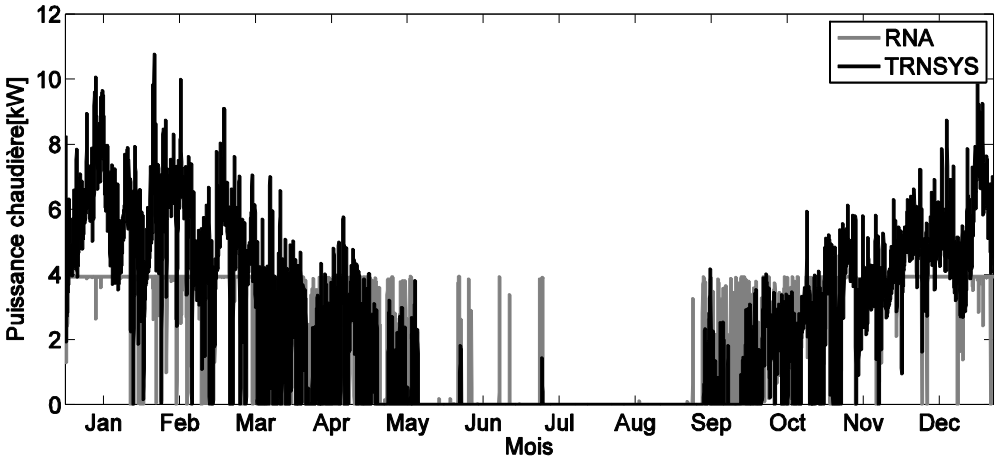

Figure 2- 52 : Evolution de la puissance de la Chaudière calculée par le modèle neuronal RNA-BIC2(tanh) et TRNSYS cas de Stockholm SFH100 16m².

L'intervalle de normalisation pour le modèle RNA-R2nrj(tanh) est ±0,2. La sortie du RNA sera donc dans l'intervalle :

$$-2.P_{aux}^{max} + 3.P_{aux}^{min} < \widehat{P}_{aux} < 3.P_{aux}^{max} - 2.P_{aux}^{min} \qquad \textbf{(2- 43)}$$

Le modèle neuronal a donc plus de liberté pour l'estimation $\widehat{P}_{aux}$. Sur la Figure 2- 53 sont représentées les deux courbes de puissances dans les mêmes conditions que la figure précédente. La concordance entre les deux courbes s'est beaucoup améliorée. Un zoom sur la période des mois d'octobre et novembre, Figure 2- 54, confirme que globalement le modèle neuronal arrive bien à reproduire l'évolution de la consommation de la chaudière même au niveau mensuel avec une bonne précision.

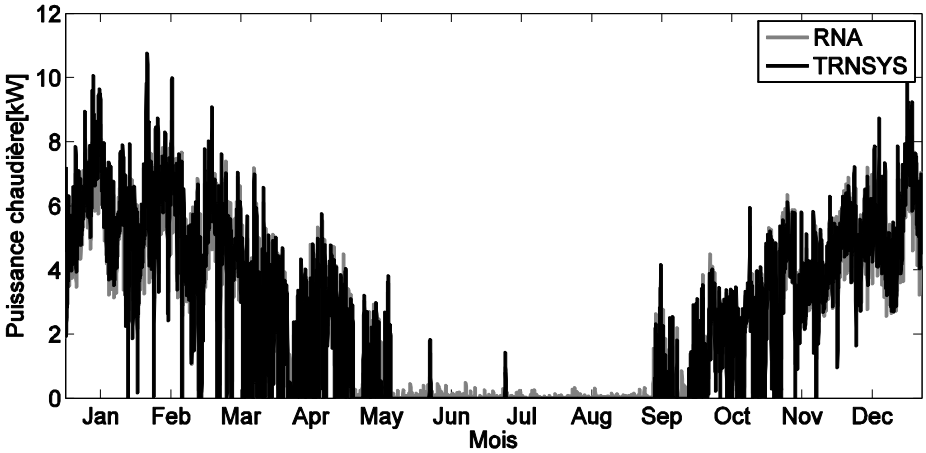

Figure 2- 53 : Evolution de la puissance de la Chaudière calculée par le modèle neuronal RNA-R2nrj(tanh) et TRNSYS pour le cas de Stockholm SFH100 16m². La FA est $tanh$ mais l'intervalle de normalisation est [-0,2 0,2]

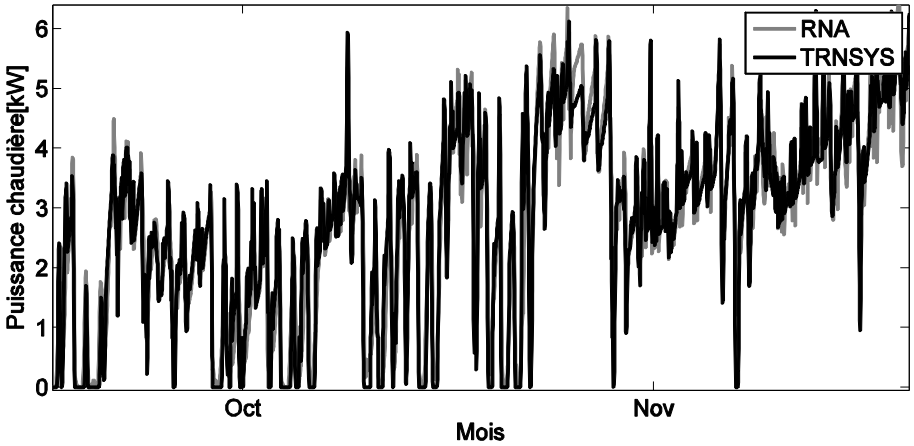

Figure 2- 54 : Zoom sur la Figure 2- 53

Les modèles neuronaux dans le cas de la fonction d'activation linéaire donnent aussi des résultats assez intéressants, Figure 2- 55 et Figure 2- 56. Le modèle sélectionné par « BIC2 » (et aussi avec les deux autres critères) dans ce cas est encore plus performant que celui avec une FA tangente hyperbolique. En effet, son « R2nrj » est légèrement supérieur. Les erreurs relatives commises pour l'estimation de la consommation énergétique annuelle, pour ce modèle, sont pour la plupart des conditions inférieures à 10%.

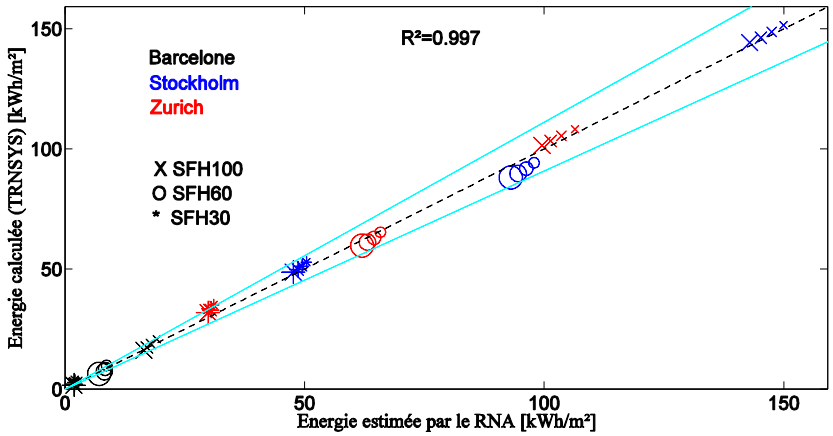

Figure 2- 55 : Comparaison entre la consommation annuelle estimée par le RNA-R2nrj(linéaire) et calculée par le modèle physique TRNSYS dans différentes conditions. La taille des symboles indique la surface du capteur utilisée (10 m², 13 m², 16 m² ou 19 m²).

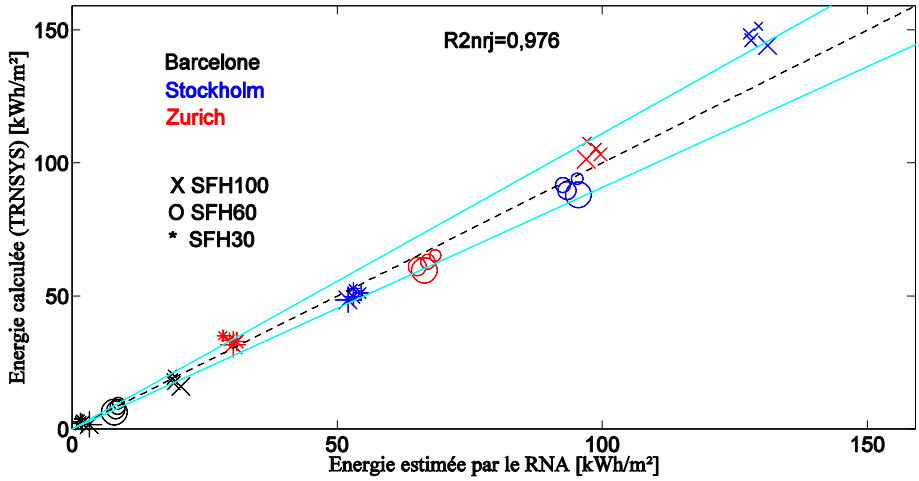

Figure 2- 56 : Comparaison entre la consommation annuelle estimée par le RNA-BIC2(linéaire) et calculée par le modèle physique TRNSYS dans différentes conditions. La taille des symboles indique la surface du capteur utilisée (10 m², 13 m², 16 m² ou 19 m²).

La FA en sortie du modèle étant linéaire, le modèle n'est pas borné et dépend uniquement des poids synaptiques notamment les biais. C'est pour cette raison que pour une telle FA il n'est pas nécessaire d'explorer plusieurs intervalles de normalisation, un seul intervalle comme ±1 est suffisant. Cependant, selon la courbe de la Figure 2- 57, les modèles neuronaux avec une fonction d'activation linéaire sont moins précis car ils présentent des oscillations.

Comme le premier objectif des RNA est d'estimer l'énergie consommée sur l'année, ses oscillations ne portent pas préjudice à la qualité attendue du modèle puisque les oscillations sont compensées par intégration. Par contre, ces oscillations poseraient problème si on souhaite utiliser le RNA à des fins de contrôle commande.

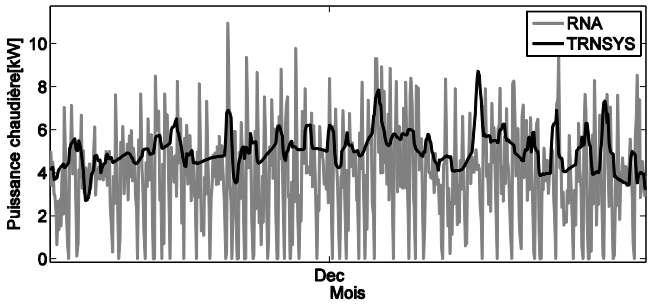

Figure 2- 57 : Evolution de la puissance de la Chaudière calculée par le modèle neuronal et TRNSYS cas de Stockholm SFH100 16m² et de RNA-BIC2(linéaire)

CONCLUSION

Une nouvelle méthodologie d'évaluation des performances des systèmes solaires thermiques innovants a été proposée dans ce deuxième chapitre. Etant basée sur un test du système complet et sur une approche de modélisation de type « boîte noire » par RNA, la méthodologie est indéniablement générique. En outre, les variables mises en jeu dans sa configuration de modélisation sont toutes des puissances (puissance de chauffage, puissance

d'ECS et ensoleillement, etc.). Elle n'est donc pas limitée à une unique typologie de système. La méthode proposée est constituée de quatre phases principales :

i. La détermination d'une séquence de test selon la méthode SCSPT
ii. L'essai du système sur un banc d'essai semi virtuel pendant 12 jours
iii. L'identification d'un modèle RNA du système complet
iv. L'estimation de la consommation annuelle du système dans différentes conditions à l'aide du RNA identifié.

Le processus de modélisation retenu consiste à identifier plusieurs modèles RNA avec différents paramètres et architectures (intervalle de normalisation, FA, nombre de neurones dans la couche cachée, etc.) sur la base des données d'apprentissage. Cette base de données est récupérée suite au test physique. Trois critères statistiques (« BIC2 », « AIC2 » et « R2bar ») ont été retenus pour sélectionner les modèles neuronaux les plus proches du modèle RNA optimal.

Les résultats de la première application de la méthodologie révèlent son intérêt. Qu'elle que soit la fonction d'activation utilisée, le meilleur modèle parmi tous les modèles estime avec précision la consommation du système. Et ce même pour certains environnements très différents des conditions d'apprentissage (Zurich-SFH60-16m²), notamment Barcelone. Cependant, malgré la qualité très satisfaisante des modèles sélectionnés, il n'est pas possible de sélectionner le modèle optimal selon les trois critères.

Dans le but de valider le caractère générique de la méthodologie ainsi que les premières conclusions, il est indispensable d'examiner ses performances dans le cas de nouveaux systèmes. L'extension de la méthodologie développée est abordée dans le chapitre suivant.

Chapitre III. Extension de la méthodologie à d'autres typologies de systèmes

INTRODUCTION

La méthodologie d'évaluation des systèmes solaires combinés, objet de la thèse, doit être générique. Dans le but d'investiguer le caractère générique de la méthodologie présentée dans le deuxième chapitre, celle-ci a été appliquée à des systèmes de typologies différentes :

- Système solaire combiné à une pompe à chaleur. Quatre systèmes ont été exploités (deux systèmes avec une PAC air-eau et deux systèmes avec une PAC eau glycolée-eau);
- Système solaire intégrant une machine à absorption. Deux systèmes différents ont été exploités. Ces systèmes peuvent répondre aux besoins de chauffage, d'ECS, mais aussi de climatisation.

La validation de la méthodologie dans ce chapitre est réalisée à l'aide d'expérimentations numériques selon le protocole présenté dans le deuxième chapitre (sous-paragraphe II.4.2.1).

Un modèle physique détaillé a été développé, sous le logiciel TRNSYS, pour chacun des systèmes solaires. Ces modèles ont été utilisés pour la simulation de la séquence de test selon la méthode SCSPT et pour les simulations annuelles dans différentes conditions.

Les données de la simulation des séquences de test serviront à l'apprentissage des modèles neuronaux, tandis que les données annuelles serviront à l'évaluation des estimations de la consommation d'énergie annuelle de chaque système dans différentes conditions par les RNA.

Dans ce troisième chapitre la description des différents systèmes, ainsi que les résultats de l'extension de la méthodologie, sont présentés et analysés.

III.1. Première typologie : systèmes solaires combinés à une pompe à chaleur

III.1.1. Description des systèmes et de leurs modèles physiques

III.1.1.1 Systèmes dits de référence

Les systèmes de référence définis dans le cadre du projet européen MacSheep [51], sont un exemple des systèmes actuellement disponibles dans le commerce. Ces deux systèmes servaient de référence pour la comparaison avec de nouveaux systèmes de rupture, développés dans le cadre du projet MacSheep. En effet, le but du projet consiste à développer des systèmes permettant une économie d'énergie d'environ 25% par rapport au cas de référence tout en restant rentable à long terme.

Le schéma du premier système de référence, le système solaire combiné à une pompe à chaleur avec un échangeur géothermique est donné sur la Figure 3- 1. Ce système sera noté dans la suite SSCPACEE-REF. Il s'agit selon la classification de la tâche 44 de l'IEA (premier chapitre) d'un système solaire « en parallèle » puisque la PAC est couplée au système de stockage en parallèle avec le capteur solaire. Ce dernier échange de la chaleur avec le ballon de stockage via un échangeur interne situé dans la partie basse. La préparation de l'ECS se

fait également via un échangeur interne mais traversant toute la hauteur du ballon. La conception de ce système est inspirée des systèmes qui sont actuellement proposés et promus par les entreprises du secteur du chauffage en Europe. La pompe à chaleur est connectée au système de stockage via trois tubes. Les deux tubes supérieurs sont utilisés pour charger la partie supérieure du ballon, qui est utilisée pour la préparation d'eau chaude sanitaire. Pour charger le volume tampon pour les besoins de chauffage, situé au milieu du ballon, la connexion peut être commutée par l'intermédiaire de deux vannes 3 voies afin de connecter la ligne d'écoulement en sortie de la pompe à chaleur à l'orifice du milieu et le retour de la PAC à la plus basse connexion. Avec ce schéma hydraulique il est possible de répondre aux besoins de chauffage directement par la PAC sans passer par le ballon de stockage. Lorsque la PAC chauffe le ballon pour les besoins de chauffage, une partie du flux passera par la boucle de distribution de chauffage et le reste ira dans le ballon. Les proportions exactes dépendent des conditions de fonctionnement de la boucle de distribution de chauffage (via une vanne thermostatique) et des températures au sein de la PAC. Ce système n'inclut pas la régénération de la chaleur du sol via l'énergie solaire.

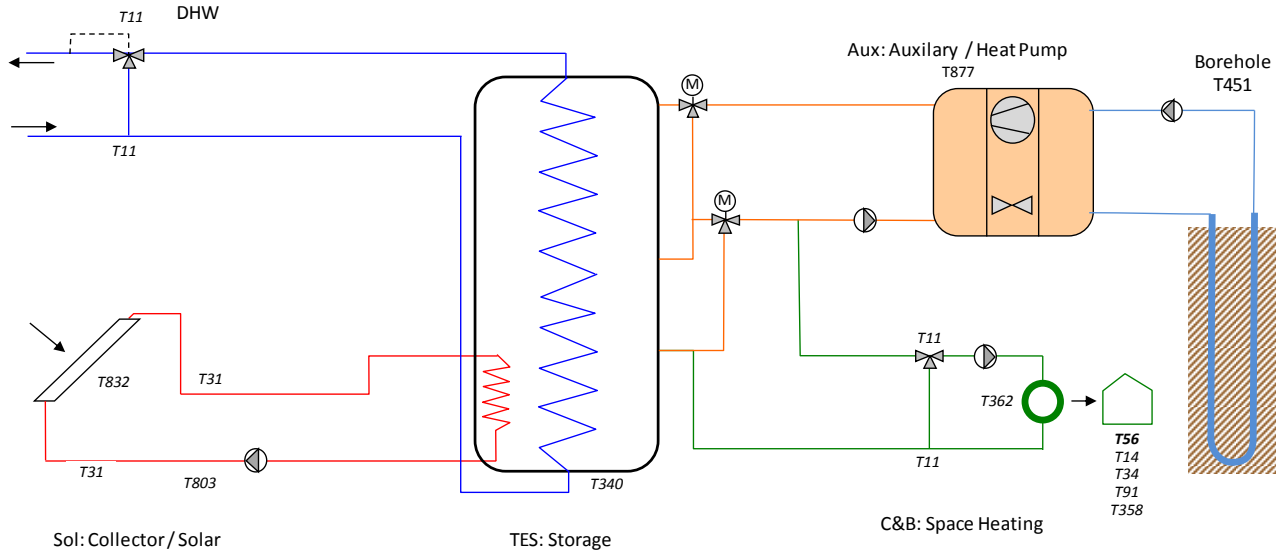

Figure 3- 1 : Schéma hydraulique du système SSCPACEE-REF

Le schéma hydraulique du deuxième système de référence est similaire au premier (Figure 3- 2). La seule différence entre les deux se situe au niveau de la PAC : le deuxième système utilise une PAC air-eau dont la vitesse du compresseur est variable. Ce système sera noté dans la suite par SSCPACAE-REF.

Pour les deux systèmes, les consignes pour la température intérieure du bâtiment et celle de l'ECS fournie aux habitants sont respectivement 20°C et 45°C. La PAC doit également garder constamment une température dans la partie supérieure du ballon à une température proche de 50°C.

Les modèles détaillés de ces deux systèmes ont été réalisés à partir de leurs schémas hydrauliques. Les caractéristiques physiques des deux systèmes utilisés dans les simulations sous le logiciel TRNSYS, sont données dans le Tableau 3- 1. Les paramètres caractérisant les pertes et le cycle thermodynamique des deux PAC sont différents d'une PAC à l'autre.

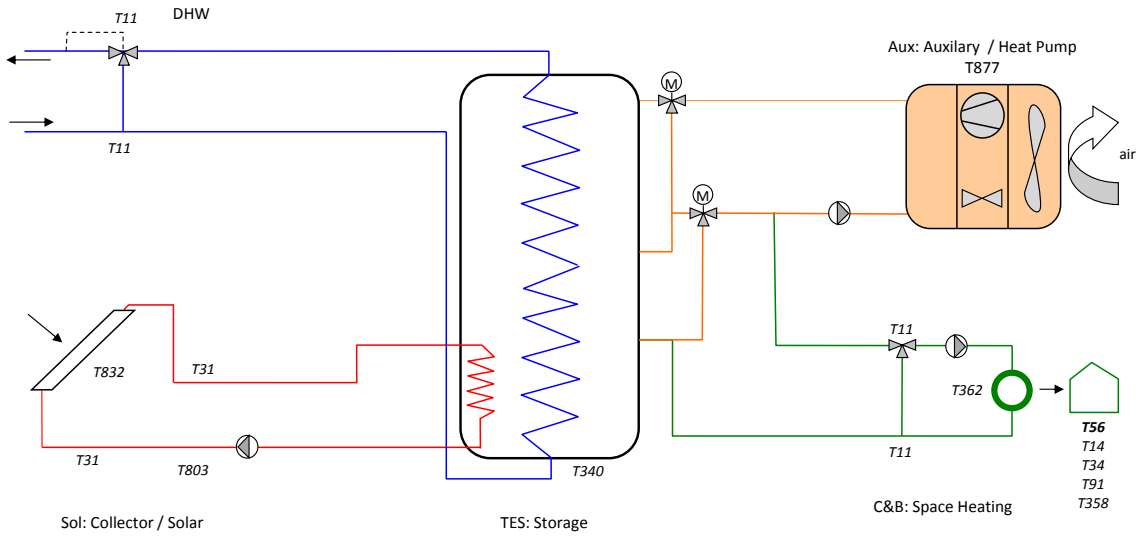

Figure 3- 2 : Schéma hydraulique du système SSCPACAE-REF

Tableau 3- 1 : Caractéristiques physiques utilisées pour la modélisation des systèmes SSCPACAE-REF et SSCPACEE-REF dans TRNSYS

Remarques	Type TRNSYS	Caractéristiques
	Pompe à chaleur	
Ce Type TRNSYS permet de modéliser à la fois une PAC eau glycolée -eau et air-eau.	TYPE 877 [113]	Puissance nominale de 8,24kW Vitesse de rotation du compresseur non variable pour la PAC eau glycolée-eau et variable pour la PAC air-eau.
	Echangeur géothermique	
Modèle permettant la simulation des sondes géothermiques verticales sous forme de U. Ce modèle est dynamique et prend en compte le régime transitoire.	Type 451 [114]	Fluide caloporteur : Glycol. Longueur des sondes : 50-123m en fonction des besoins du bâtiment. Diamètre des sondes : 2,6-3,2cm en fonction des besoins du bâtiment
	Ballon de stockage	
	Type 340 [115]	750 litre de volume 1,74m de hauteur Coefficient de pertes latérale de 2,81 W/K et pour les parties supérieure et inférieure une valeur de 0,3 W/K, 0,67 W/K respectivement.
	Capteur solaire	
	TYPE 832 [116]	$F'(\tau\alpha)_{en}$=0.793; $K_{\theta d}$=0.96 c_1=3.95 (W/m^{2}K); c_5=7000 (J.m^{-2} K^{-1}); c_2=0.0122 (W/m^{2}K^2) β=45° ; b_0=0.18

III.1.1.2 Systèmes dits améliorés

Comme pour les systèmes de références les systèmes considérés ici (Figure 3- 3) sont également couplés à des pompes à chaleur (PAC eau glycolée-eau ou air-eau). Les deux systèmes font partis des propositions, dans le cadre du projet Macsheep, de solutions

innovantes pour le développement de systèmes performants. Les deux systèmes sont également des systèmes en parallèle. Ils seront notés dans la suite par SSCPACEE-AME et SSCPACAE-AME respectivement. Les améliorations clés apportées aux deux systèmes par rapport à ceux de références sont :

- La pompe à chaleur comporte un économiseur et un compresseur à injection de vapeur. Les performances de la PAC sont améliorées grâce à l'économiseur qui permet la réduction de la consommation électrique du compresseur par rapport au cas du système de référence. En effet, le système économiseur utilise un échangeur de chaleur entre une fraction du fluide réfrigérant en sortie du condenseur et la deuxième fraction de ce dernier ayant suivi un deuxième chemin i. e. après le passage par un détendeur voir Figure 3- 4. La deuxième fraction du fluide étant évaporé à une pression intermédiaire (entre la pression basse et haute) est injectée directement dans le compresseur. Ce qui résulte en une économie d'énergie puisque le compresseur consomme moins d'électricité pour compresser la vapeur injectée.
- Un ballon de stockage performant utilisant un système innovant de stratification et des dispositifs anti-thermosiphon pour mieux garder la stratification dans le ballon. Dans le but de réduire au maximum les pertes thermiques, l'isolation du ballon a été également améliorée via l'utilisation d'un *vacuum insulation palen (VIP)*.
- Le chauffage par le capteur solaire ainsi que l'ECS sont réalisés via des échangeurs externes dans le but d'intensifier la stratification.

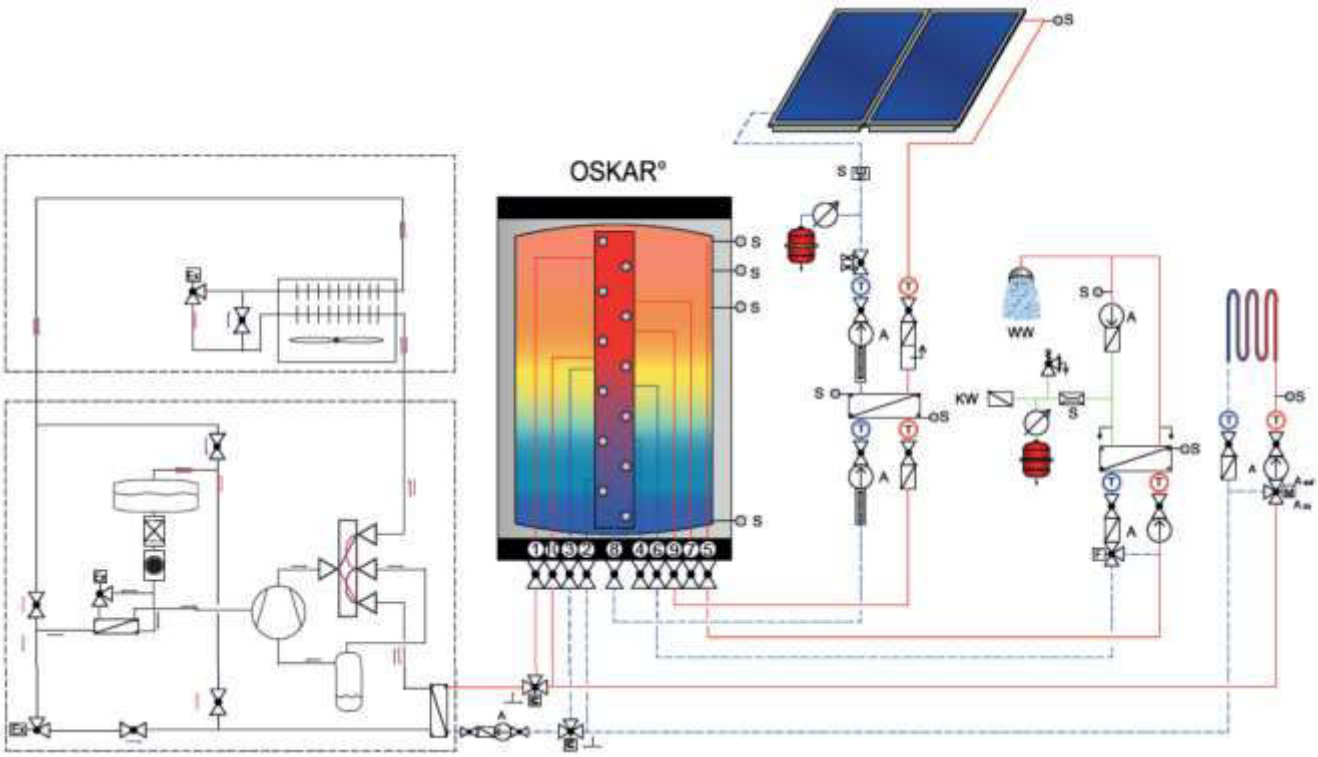

Figure 3- 3 : Schéma hydraulique du système SSCPACEE-AME (développé par RATIOTHERM et SERC)

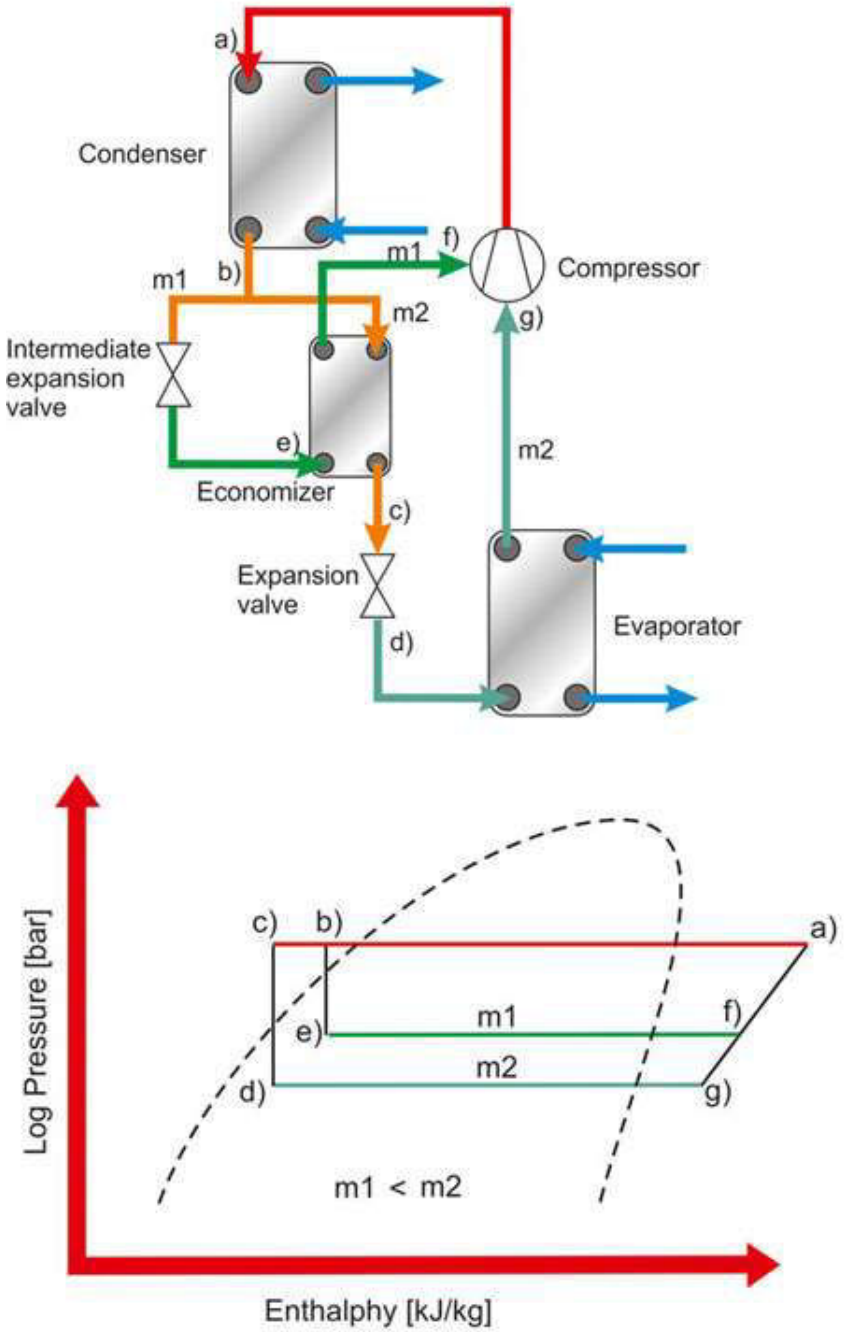

Figure 3- 4 : Principe de fonctionnement et cycle thermodynamique d'une PAC avec économiseur (source [117])

Les modèles détaillés de ces deux systèmes ont été réalisés à partir de leurs schémas hydrauliques. Les caractéristiques physiques des deux systèmes utilisés dans les simulations sous TRNSYS sont données dans le Tableau 3- 2. Les paramètres caractérisant les pertes et le cycle thermodynamique des deux PAC sont différents d'une PAC à l'autre.

Tableau 3- 2 : Caractéristiques physiques utilisées pour la modélisation des systèmes SSCPACAE-AME et SSCPACEE-AME dans TRNSYS

Remarques	Type TRNSYS	Caractéristiques
Pompe à chaleur		
Ce Type TRNSYS permet de modéliser à la fois une PAC eau glycolée -eau et air-eau, avec le système économiseur.	TYPE 887 [118]	Puissance nominale de 8,24kW Vitesse de rotation du compresseur non variable pour la PAC eau-eau et variable pour la PAC air-eau.
Echangeur géothermique		
Modèle permettant la simulation des sondes géothermiques verticales sous forme de U. Ce modèle est dynamique et prend en compte le régime transitoire.	Type 451	Fluide caloporteur : Glycol. Longueur des sondes : 50-123m en fonction des besoins du bâtiment. Diamètre des sondes : 2,6-3,2cm en fonction des besoins du bâtiment
Ballon de stockage		
	Type 340	920 litre de volume 2,1 m de hauteur Coefficient de pertes latérale de 1,38 W/K et pour les parties supérieure et inférieure une valeur de 0,153 W/K, 0,34 W/K respectivement.
Capteur solaire		
	TYPE 832	$F'(\tau\alpha)_{en}$=0.793; $K_{\theta d}$=0.96 c_1=3.95 (W/m²K); c_5=7000 (J.m^{-2} K^{-1}); c_2=0.0122 (W/m²K²) β=45° ; b_0=0.18 Surface du capteur = 16 m²

III.1.2. Application de la méthodologie développée

La méthodologie proposée dans ce mémoire a été appliquée aux différents systèmes combinés à des pompes à chaleur, décrits précédemment, de la même façon que dans le deuxième chapitre. Le processus de modélisation selon la méthodologie développée dure environ 4 jours par système. Les paramètres (nombre de neurones maximum, nombre de réinitialisation etc.) du programme développé pour la modélisation sont les mêmes que ceux présentés dans le Tableau 2- 11 (deuxième chapitre). Les données d'apprentissage issues de la simulation du test de chaque système selon la méthode SCSPT sont représentées en fonction du temps, pour chaque système, sur les Figure 3- 5, Figure 3- 6, Figure 3- 7 et Figure 3- 8. Les conditions de test sont les mêmes pour tous les systèmes : Zurich-SFH60-16m². L'évolution du rayonnement solaire ainsi que les besoins de chauffage au cours des 12 jours de la séquence de test ne sont pas forcément les mêmes. En effet, les jours sélectionnés selon la méthode SCSPT dépendent de la nature et du comportement de chacun des systèmes.

La consommation annuelle de chaque système (consommation électrique du compresseur) est calculée dans différentes conditions par leur modèle physique détaillé sous TRNSYS et par les modèles neuronaux entrainés par la méthodologie mise en place (voir protocole de

validation dans la Figure 2- 19). Les résultats des modèles physiques détaillés seront comparés similairement au deuxième chapitre (sous-paragraphe II.5.5.2) :

- Aux résultats des modèles neuronaux les plus pertinents (selon l'analyse de la méthodologie) sélectionnés par les critères statistiques (voir le deuxième chapitre).
- Aux résultats des modèles les plus performants (selon le critère « *R2nrj* ») parmi tous les modèles créés.

Dans les sections suivantes, une analyse des résultats par système est faite. Les points forts ainsi que des explications des limites de la méthodologie proposée sont discutés et appréhendés.

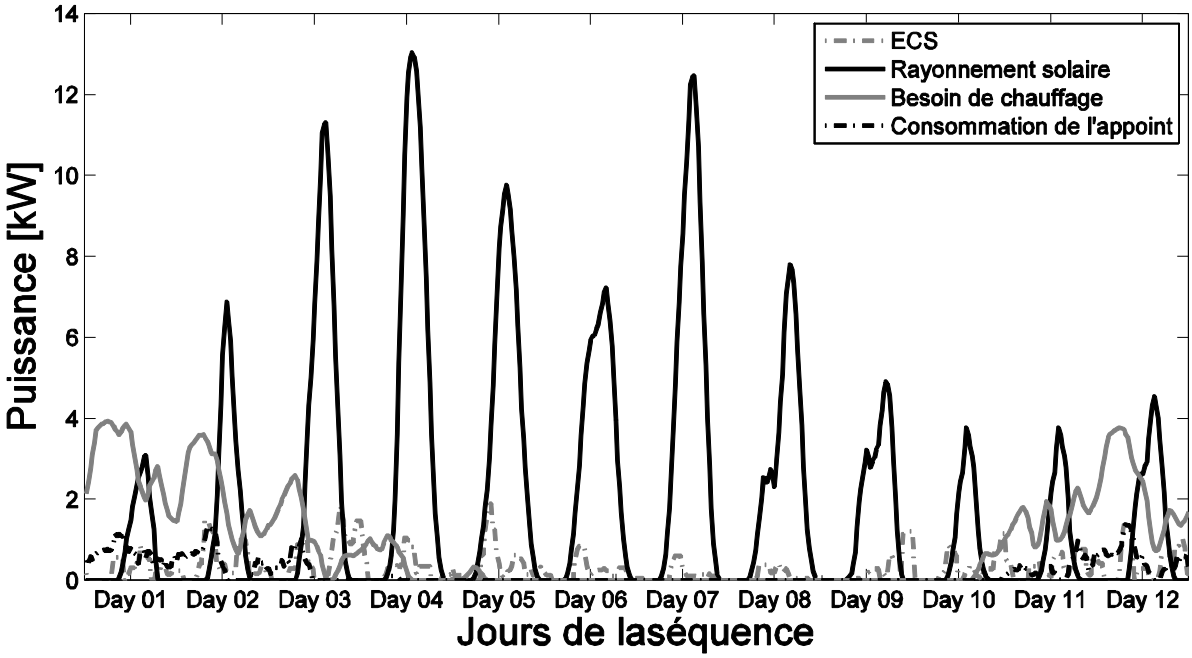

Figure 3- 5 : Séquence d'apprentissage pour la modélisation du système SSCPACEE-REF. Conditions dans lesquelles le test selon la méthode SCSPT a été simulée: Zurich-SFH60-16m²

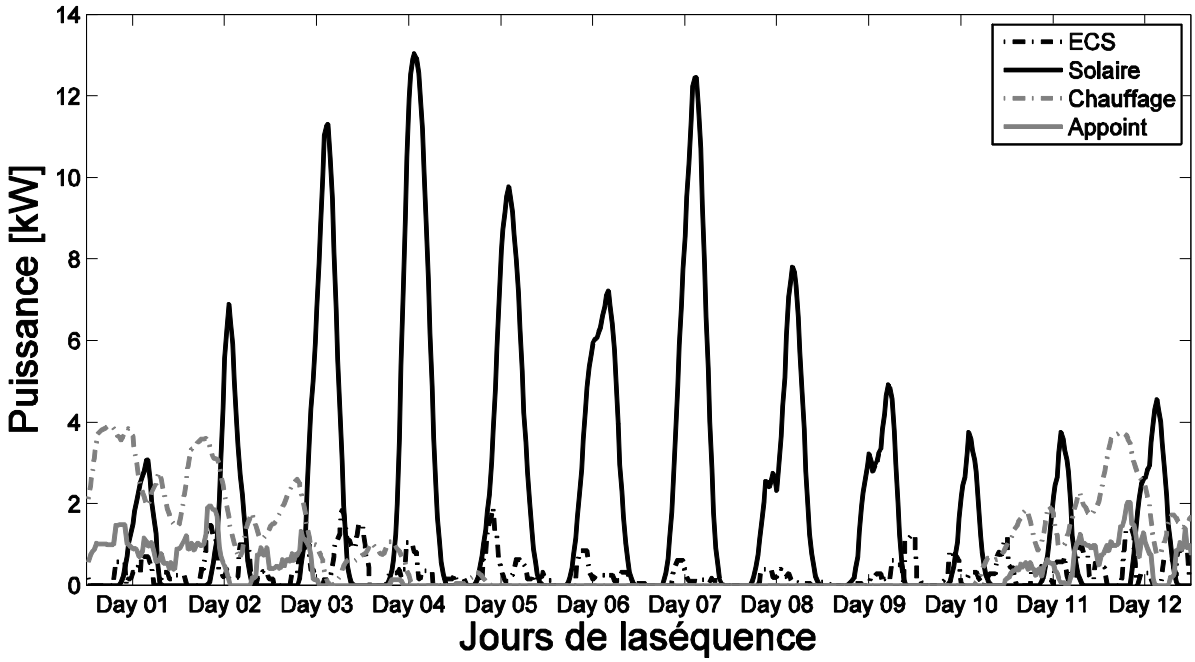

Figure 3- 6 : Séquence d'apprentissage pour la modélisation du système SSCPACAE-REF. Conditions dans lesquelles le test selon la méthode SCSPT a été simulée: Zurich-SFH60-16m²

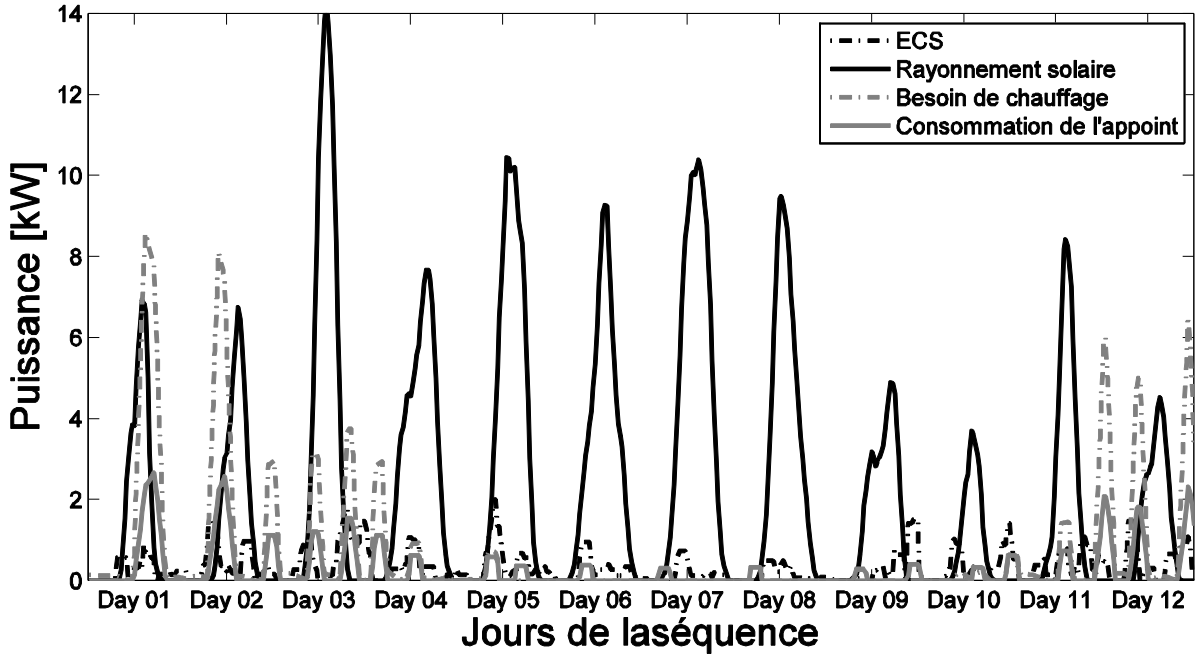

Figure 3- 7 : Séquence d'apprentissage pour la modélisation du système SSCPACEE-AME. Conditions dans lesquelles le test selon la méthode SCSPT a été simulée: Zurich-SFH60-16m²

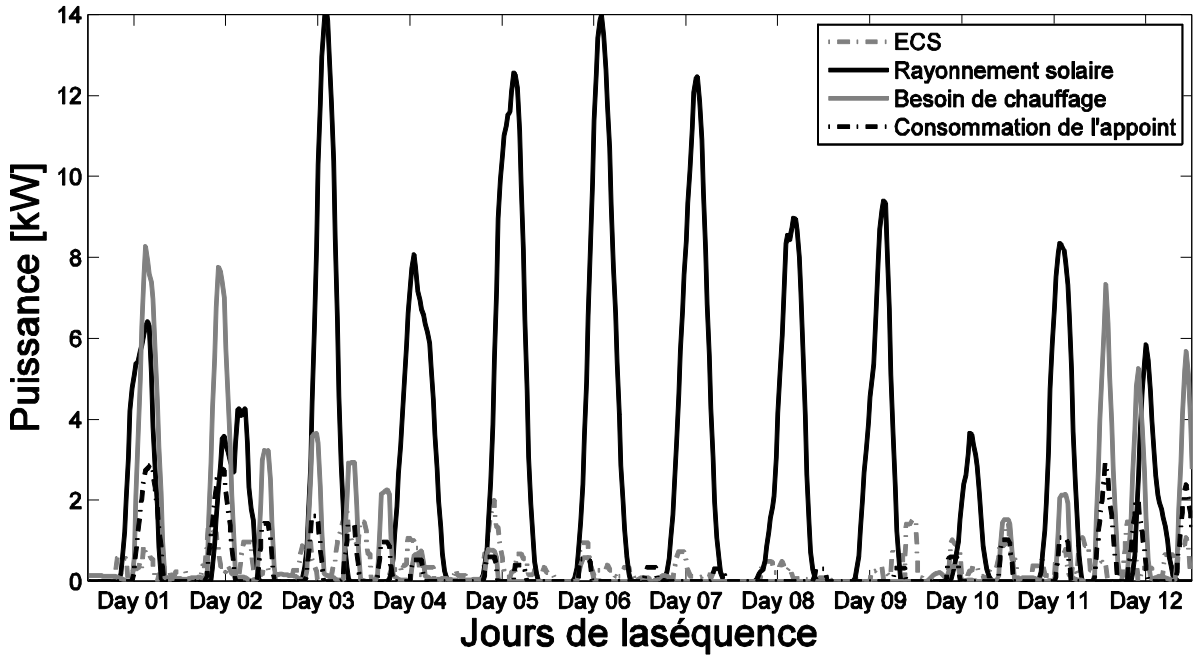

Figure 3- 8 : Séquence d'apprentissage pour la modélisation du système SSCPACAE-AME. Conditions dans lesquelles le test selon la méthode SCSPT a été simulée: Zurich-SFH60-16m²

III.1.3. Résultats et discussions

Similairement aux SSC avec un appoint par une chaudière à gaz, objet du deuxième chapitre, toutes les simulations, que ce soit avec les modèles neuronaux ou à l'aide de TRNSYS, ont été réalisées dans différentes conditions. Les tendances générales des résultats de l'application de la méthodologie dépendent du modèle considéré.

III.1.3.1 SSCPACEE-REF

Les résultats des prédictions des meilleurs modèles pour les deux Fonctions d'Activation (FA) de sortie (RNA-R2nrj(tanh) et RNA-R2nrj(linéaire) sur le Tableau 3- 3), parmi tous les modèles créés sont donnés sur les Figure 3- 9 et Figure 3- 10. Les coefficients de détermination pour ces modèles sont supérieurs à 0,98. Les erreurs relatives des prédictions des modèles neuronaux, inférieurs à 10% dans la plupart des cas, montrent la pertinence et le pouvoir de généralisation assez important des modèles issus de la méthodologie. En effet, il est possible, grâce à ces modèles, d'estimer la consommation annuelle qu'aura le système dans diverses conditions, différentes de celle du test de 12 jours.

Tableau 3- 3 : Caractéristiques des modèles neuronaux étudiés (système SSCPACEE-REF)

Modèles	NORM	NDN	TD	« R2nrj »	« R2bar »	« BIC2 »	« AIC2 »
RNA-R2nrj(tanh)	±0,2	9,0	2,0	0,9905	0,9189	11,8290	11,7568
RNA-R2bar(tanh)	±1,0	11,0	2,0	0,8824	0,9814	10,4111	10,3315
RNA-BIC2(tanh)	±1,0	8,0	2,0	0,9257	0,9812	10,3448	10,2767
RNA-AIC2(tanh)	±1,0	8,0	2,0	0,9257	0,9812	10,3448	10,2767
RNA-R2nrj(linéaire)	±1,0	2,0	3,0	0,9830	0,9497	11,1852	11,1445
RNA-R2bar(linéaire)	±1,0	12,0	2,0	0,9361	0,9830	10,3516	10,2684
RNA-BIC2(linéaire)	±1,0	12,0	2,0	0,9361	0,9830	10,3516	10,2684
RNA-AIC2(linéaire)	±1,0	12,0	2,0	0,9361	0,9830	10,3516	10,2684

Des erreurs relatives supérieures à 10%, malgré la qualité des deux RNA, ont été enregistrées pour les conditions présentant des besoins faibles : pour le RNA-R2nrj(tanh) cela concerne le climat de Barcelone et pour le RNA-R2nrj(linéaire) cela concerne en particulier les conditions avec un bâtiment SFH30. Le modèle RNA-R2nrj(tanh) surestime la consommation annuelle du système pour le climat de Barcelone tandis que le RNA-R2nrj(linéaire) la sous-estime pour le cas du bâtiment SFH30.

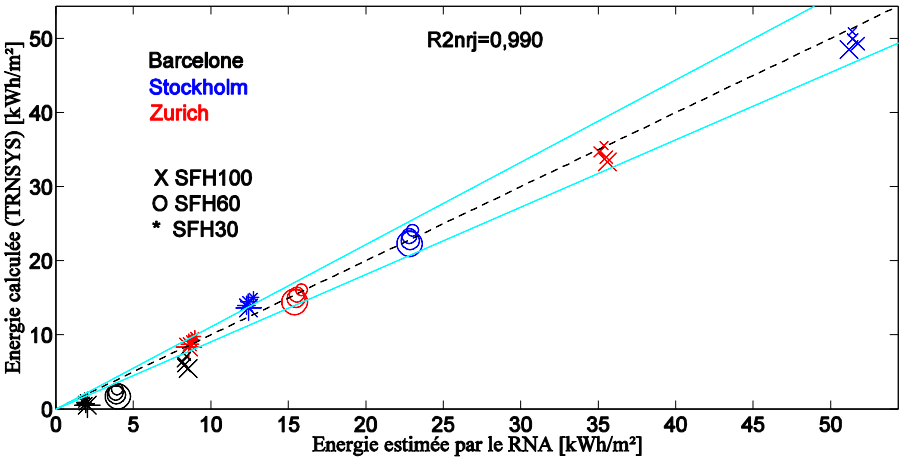

Figure 3- 9 : Comparaison entre la consommation annuelle estimée par le RNA-R2nrj(tanh) et calculée par le modèle physique TRNSYS dans différentes conditions. Cas du système SSCPACEE-REF.

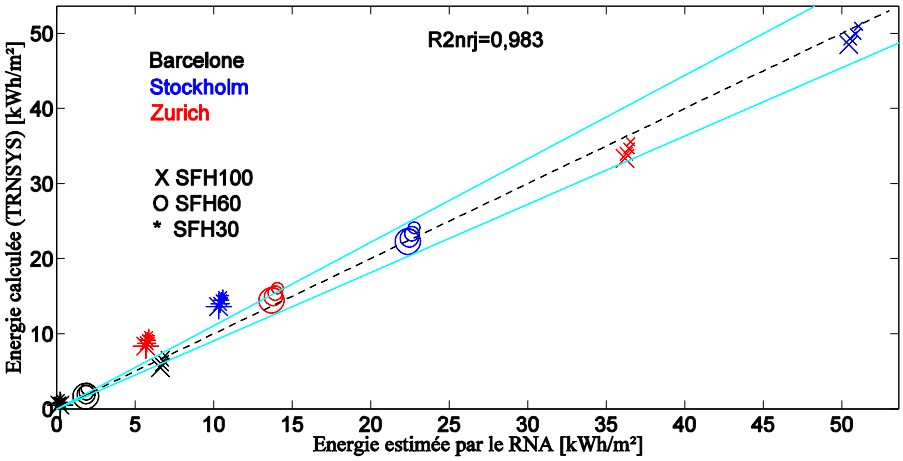

Figure 3- 10 : Comparaison entre la consommation annuelle estimée par le RNA-R2nrj(linéaire) et calculée par le modèle physique TRNSYS dans différentes conditions. Cas du système SSCPACEE-REF.

Dans le but d'expliquer ces phénomènes, les estimations pour les conditions où l'écart absolu est le plus important à savoir Barcelone-SFH100-19 m² (cas du RNA-R2nrj(tanh)) et Stockholm-SFH30-10 m² (cas du RNA-R2nrj(linéaire)) sont à analyser de plus près.

Sur la Figure 3- 11 sont présentées les consommations mensuelles du système, calculées par TRNSYS ou prédites par le RNA-R2nrj(tanh) pour la condition Barcelone-SFH100-19 m². Il apparait clairement sur la figure que les écarts les plus importants se produisent pendant la saison de chauffage, de janvier à avril et puis d'octobre à décembre. Pendant la période intermédiaire les besoins de chauffage sont très faibles voire inexistante. Cette période coïncide avec l'été où l'ensoleillement atteint ses valeurs maximales.

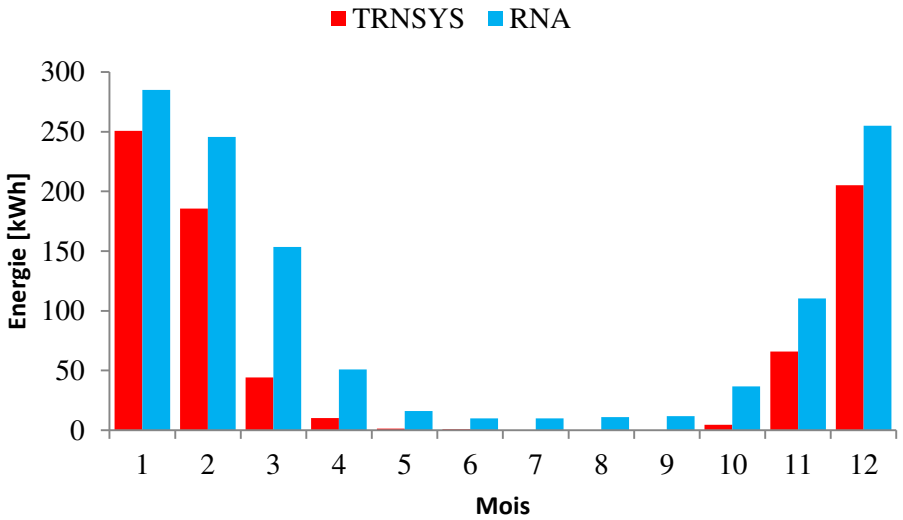

Figure 3- 11 : Consommations mensuelles du système SSCPACEE-REF calculés par TRNSYS et prédites par le RNA-R2nrj(tanh) pour l'environnement Barcelone-SFH100-19 m².

La comparaison entre l'ensoleillement journalier de la séquence d'apprentissage (Zurich) et l'ensoleillement journalier moyen pour chaque mois de la séquence annuelle du climat de Barcelone (Figure 3- 12) explique les écarts constatés. En effet, pendant la période de chauffage, l'ensoleillement à Barcelone est presque le double de celui de Zurich. Le RNA-R2nrj(tanh) ne peut donc pas réagir correctement à des phénomènes très différents des conditions d'apprentissage (les valeurs des variables ne varient pas sur le même domaine de variation des données d'apprentissage), qui lui sont inconnus. Malgré un ensoleillement élevé pendant la période estivale, le RNA-R2nrj(tanh) a su estimer avec un degré de précision très satisfaisant la consommation mensuelle du système. Ceci est certainement dû à la forte sensibilité du modèle aux valeurs nulles de la variable « besoin de chauffage ». En effet, dans les données d'apprentissage (Figure 3- 5), il est facile de constater qu'il y a plus d'exemples de situations où les besoins de chauffage sont nuls, par rapport aux autres situations, et ce pour différentes valeurs de l'ensoleillement.

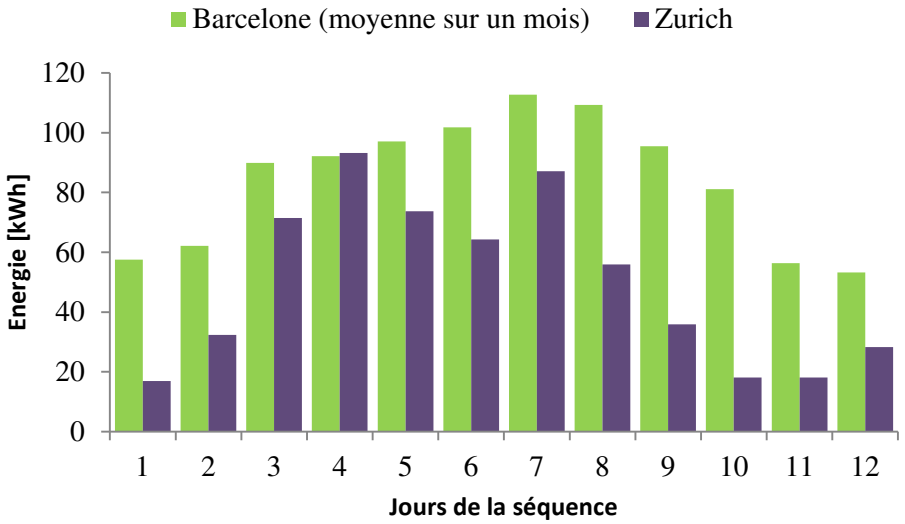

Figure 3- 12 : Comparaison entre l'ensoleillement journalier de la séquence de test pour Zurich et l'ensoleillement moyen journalier pour chaque mois à Barcelone.

L'écart entre le modèle RNA-R2nrj(linéaire) et le modèle physique TRNSYS correspondant à l'environnement Stockholm-SFH30-10 m² sont également dû au fait qu'elles sont éloignées des conditions d'entrainement mais selon un autre aspect. En effet, l'ensoleillement à Zurich est plus important qu'à Stockholm (Figure 3- 13), le RNA-R2nrj(linéaire) a donc forcément rencontré pendant le processus de son élaboration des situations similaires. L'écart constaté ne peut donc être expliqué que par l'influence des besoins de chauffage.

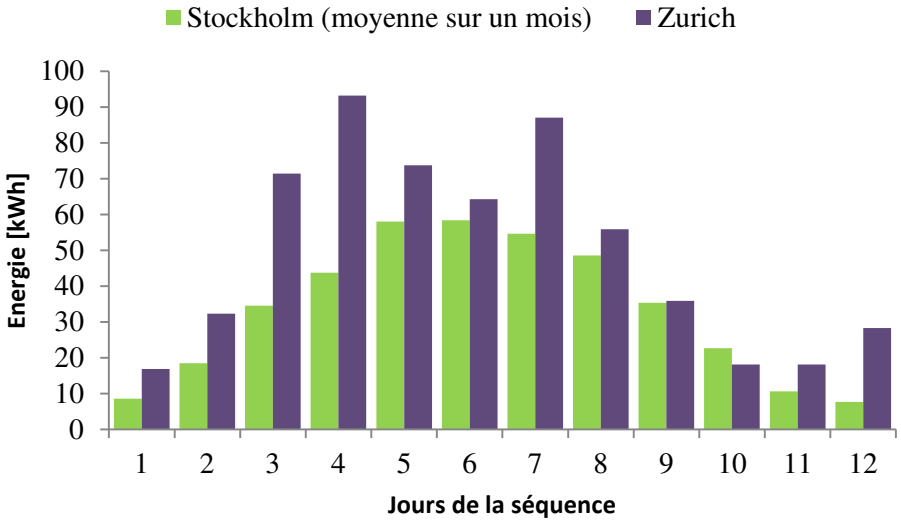

Figure 3- 13 : Comparaison entre l'ensoleillement journalier de la séquence de test pour Zurich et l'ensoleillement moyen journalier pour chaque mois à Stockholm.

Pour une distribution d'ensoleillement donnée, les besoins de chauffage d'un bâtiment SFH30 sont beaucoup plus faibles qu'un bâtiment SFH60. Pendant le mois de décembre l'écart entre le RNA-R2nrj(linéaire) et TRNSYS est le plus élevé (Figure 3- 14). Pour ce mois-ci, l'ensoleillement à Stockholm est très faible. Des intensités du même ordre de grandeur peuvent être rencontrées pendant le premier jour de la séquence pour le climat de Zurich (jour avec l'ensoleillement le plus faible). Théoriquement, pour que le modèle neuronal estime correctement la consommation du système pendant décembre, il faut que les besoins de chauffage avoisinent les 70 kWh/jour (Zurich-SFH60, janvier, Figure 3- 15) or les besoins de

chauffage à Stockholm-SFH30 pendant le mois de décembre sont seulement de l'ordre de 45 kWh/jour. Cette explication est globale et révèle que le modèle rencontrera forcément des situations inconnues pouvant être la cause d'un écart important même sur une durée limitée. Cependant, il est aussi important de rappeler que les modèles utilisés sont dynamiques (le RNA-R2nrj(linéaire), en l'occurrence, prend en compte en entrée les 3 valeurs de la sortie du modèle estimées aux instants précédents), un écart important pendant une période de plus de trois pas de temps peut donc fausser les estimations futures du modèle.

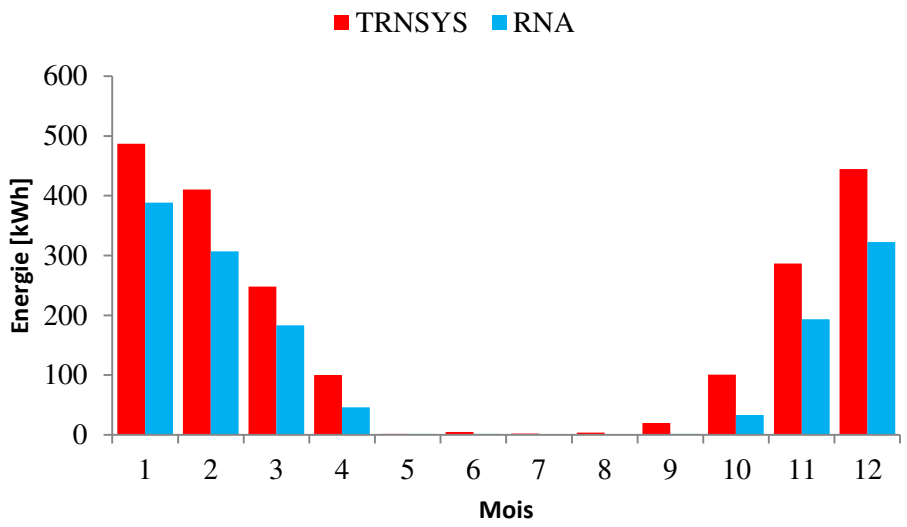

Figure 3- 14 : Consommations mensuelles du système SSCPACEE-REF calculés par TRNSYS et prédites par le RNA-R2nrj(linéaire) pour l'environnement Stockholm-SFH30-10 m².

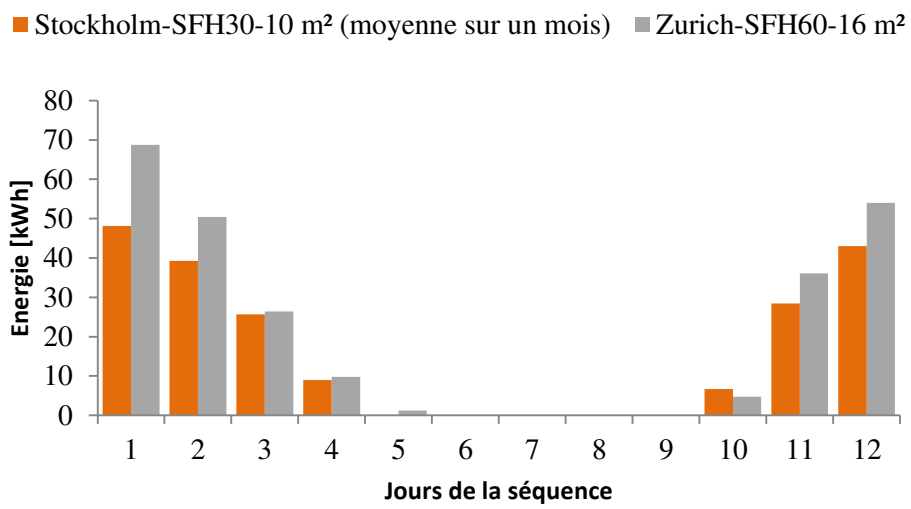

Figure 3- 15 : Comparaison entre les besoins de chauffage journaliers de la séquence de test pour Zurich-SFH60-16 m² et les besoins de chauffage mensuels à Stockholm-SFH30-10 m² moyennés sur un mois pour se ramener à des besoins journaliers équivalents.

Contrairement au modèle RNA-R2nrj(tanh), le modèle RNA-R2nrj(linéaire) sait gérer correctement l'influence de la surface du capteur (la consommation diminue quand la surface augmente) malgré le fait qu'elle n'ait pas été incluse dans le processus de la modélisation en tant que paramètre.

Le RNA-R2nrj(linéaire) a pu modéliser le fonctionnement physique du système même si la base d'apprentissage ne comporte pas suffisamment d'information.

Pour effectivement prendre en compte le paramètre surface du capteur dans le modèle neuronal, il faudrait lui présenter des exemples avec différentes valeurs de ce paramètre. Or, la variation de la surface du capteur dans la méthode de test SCSPT, telle qu'elle a été développée jusqu'à présent, n'est pas intégrée.

Les modèles neuronaux sont des fonctions analytiques. Leurs paramètres n'ont pas d'explications physiques. Il a été difficile de montrer pourquoi le RNA surestime ou sous-estime la consommation du système pour les conditions extrêmes analysées dans ce paragraphe. Il est donc difficile d'analyser plus en détail les problèmes d'estimation de consommation en énergie d'appoint des systèmes pour ces conditions extrêmes.

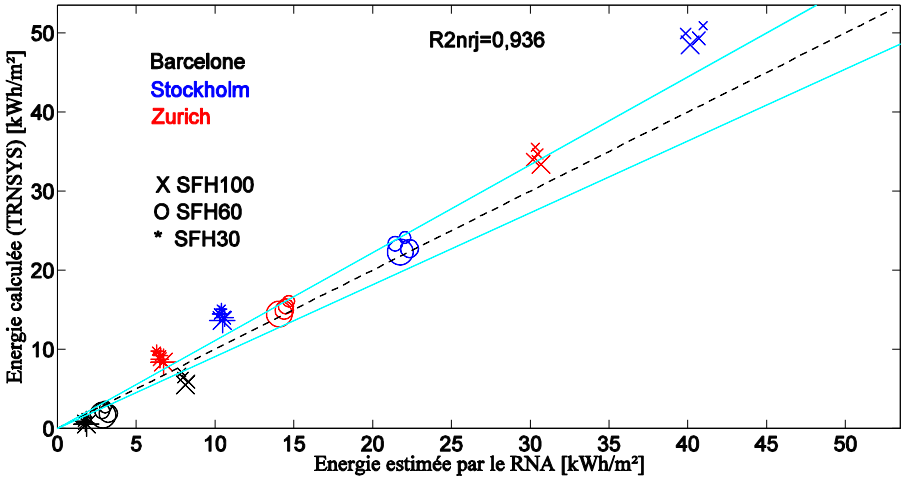

Figure 3- 16 : Comparaison entre la consommation annuelle estimée par le RNA-BIC2(linéaire) et calculée par le modèle physique TRNSYS dans différentes conditions. Cas du système SSCPACEE-REF. La taille des symboles indique la surface du capteur utilisée (10 m², 13 m², 16 m² ou 19 m²).

Dans le cas de la FA linéaire les trois critères ont sélectionnés le même RNA. Les résultats de simulations de ce modèle sont présentés sur la Figure 3- 16. Globalement ce modèle donne de bonnes estimations (erreurs relatives inférieure à 20%) sauf pour Stockholm-SFH100 et certaines conditions dans lesquelles la consommation du système est inférieure à 12kWh/m². Dans la plupart des cas ce modèle sous-estime la consommation du système. Les modèles présentant ce défaut sont facilement détectables d'autant plus que la FA de sortie correspondante est la fonction *tanh*. En effet, la cause de cette sous-estimation de la consommation du système est le phénomène d'extrapolation déjà décrit et expliqué dans le deuxième chapitre (sous-paragraphe II.5.5.2).

Sur la Figure 3- 17 est présenté un exemple montrant la bonne concordance entre la puissance du compresseur calculée par TRNSYS et estimée par le RNA-BIC2(linéaire) pour Stockholm-SFH60-19 m² pendant le mois de novembre.

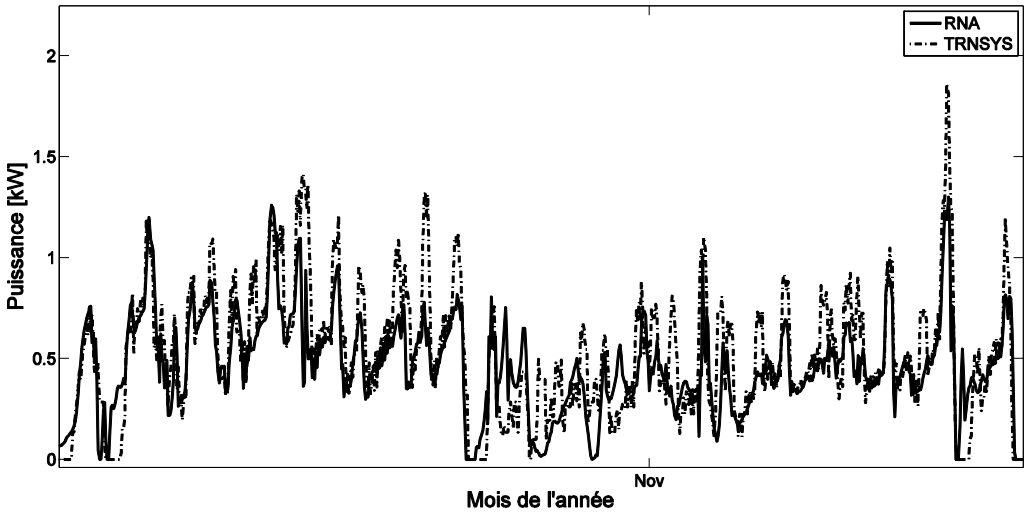

Figure 3- 17 : Zoom sur l'évolution de la puissance de l'appoint (PAC) du système SSCPACEE-REF au cours des mois de novembre et décembre. Comparaison entre les estimations du RNA-BIC2(linéaire) et les calculs du modèle TRNSYS. Cas de Stockholm-SFH60-19 m².

Pour la FA tanh les critères statistiques n'ont pas sélectionné les mêmes modèles. Les performances de ces derniers sont moins bonnes que dans le cas de la FA linéaire (notamment pour le RNA-R2bar(linéaire)).

Les estimations des modèles sélectionnés selon les critères statistiques sont évidemment moins bonnes que les deux premiers modèles. Néanmoins, les résultats correspondants montrent qu'il est toujours possible de déterminer des modèles ayant de bonnes performances mais pour un nombre réduit de conditions et avec des écarts relatifs un peu plus élevés.

III.1.3.2 SSCPACAE-REF

Les résultats de la modélisation du système SSCPACAE-REF sont légèrement mieux que ceux du premier système. En effet, les coefficients de détermination (« R2nrj ») des meilleurs modèles sélectionnés sont égaux à 0,992 et 0,993 respectivement pour le modèle avec une FA tanh (RNA-R2nrj(tanh)) et linéaire (RNA-R2nrj(linéaire)), Figure 3- 18 et Figure 3- 19. Les caractéristiques des modèles considérés dans cette partie sont données sur le Tableau 3- 4.

Tableau 3- 4 : Caractéristiques des modèles neuronaux étudiés (système SSCPACAE-REF)

Modèles	NORM	NDN	TD	« R2nrj »	« R2bar »	« BIC2 »	« AIC2 »
RNA-R2nrj(tanh)	±0,4	8,0	3,0	0,9923	0,9035	12,8095	12,7291
RNA-R2bar(tanh)	±1,0	5,0	3,0	0,8231	0,9760	11,3084	11,2447
RNA-BIC2(tanh)	±1,0	5,0	3,0	0,8231	0,9760	11,3084	11,2447
RNA-AIC2(tanh)	±1,0	5,0	3,0	0,8231	0,9760	11,3084	11,2447
RNA-R2nrj(linéaire)	±1,0	8,0	1,0	0,9928	0,0836	15,3445	15,2916
RNA-R2bar(linéaire)	±1,0	8,0	3,0	0,9882	0,9768	11,3876	11,3072
RNA-BIC2(linéaire)	±1,0	6,0	3,0	0,8626	0,9763	11,3326	11,2628
RNA-AIC2(linéaire)	±1,0	6,0	3,0	0,8626	0,9763	11,3326	11,2628

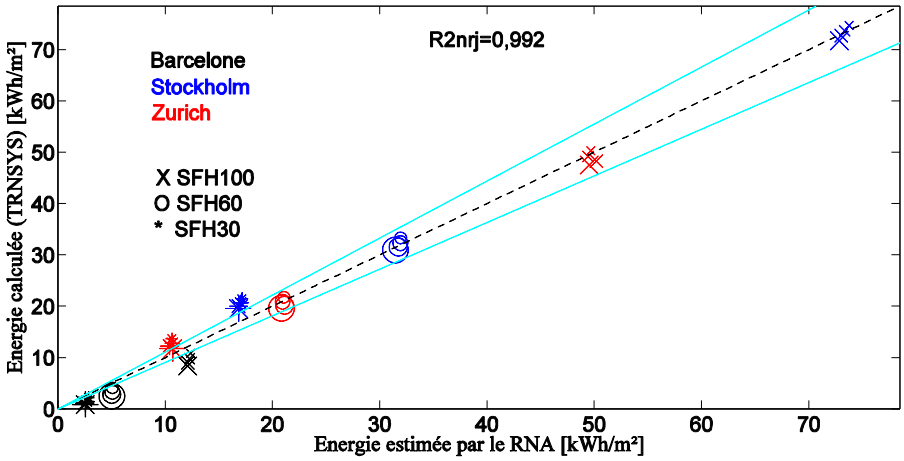

Figure 3- 18 : Comparaison entre la consommation annuelle estimée par le RNA-R2nrj(tanh) et calculée par le modèle physique TRNSYS dans différentes conditions. Cas du système SSCPACAE-REF. La taille des symboles indique la surface du capteur utilisée (10 m², 13 m², 16 m² ou 19 m²).

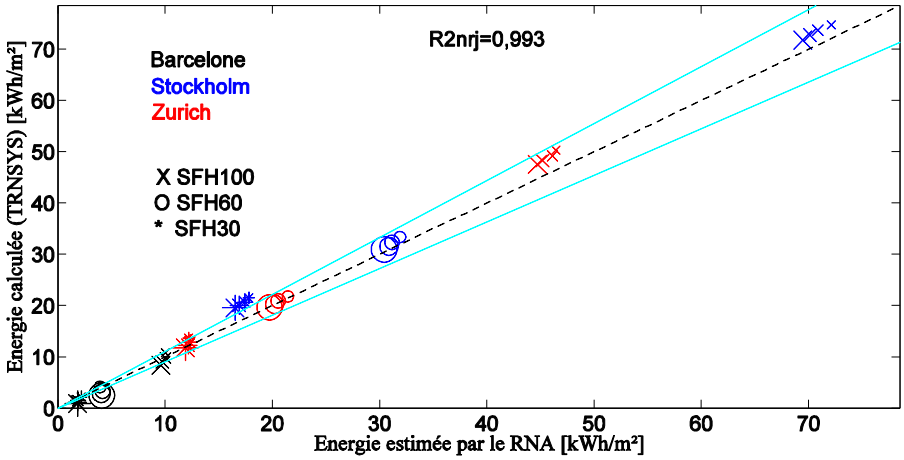

Figure 3- 19 : Comparaison entre la consommation annuelle estimée par le RNA-R2nrj(linéaire) et calculée par le modèle physique TRNSYS dans différentes conditions. Cas du système SSCPACAE-REF. La taille des symboles indique la surface du capteur utilisée (10 m², 13 m², 16 m² ou 19 m²).

Les deux modèles estiment avec un degré de précision très satisfaisant la consommation annuelle du système dans différentes conditions. Les écarts relatifs sont inférieurs à 10%. Similairement au cas du système SSCPACEE-REF les modèles neuronaux présentent des difficultés pour estimer la consommation du système dans les conditions éloignées de celles de l'apprentissage : climat de Barcelone et le bâtiment de type SFH30. Malgré cela, les écarts relatifs dans ces conditions restent inférieurs à 20% et ce pour les deux modèles. Aussi, les

écarts absolus sont faibles. Il est important de signaler également que ces deux modèles, notamment le RNA-R2nrj(linéaire), gèrent beaucoup mieux l'influence de la surface du capteur sur la consommation du système.

Pour ce système, le critère « R2bar » s'est distingué des autres en sélectionnant un RNA avec de très bonnes performances : R2nrj = 0,988 (Figure 3- 20). Ce modèle reste néanmoins moins performant que les deux premiers : RNA-R2nrj(tanh) et RNA-R2nrj(linéaire). Encore une fois, les estimations pour les conditions extrêmes sont les moins bonnes. Le modèle a des difficultés à estimer la consommation dans le cas des conditions présentant de faibles besoins. Cependant, il gère correctement l'influence de la surface du capteur.

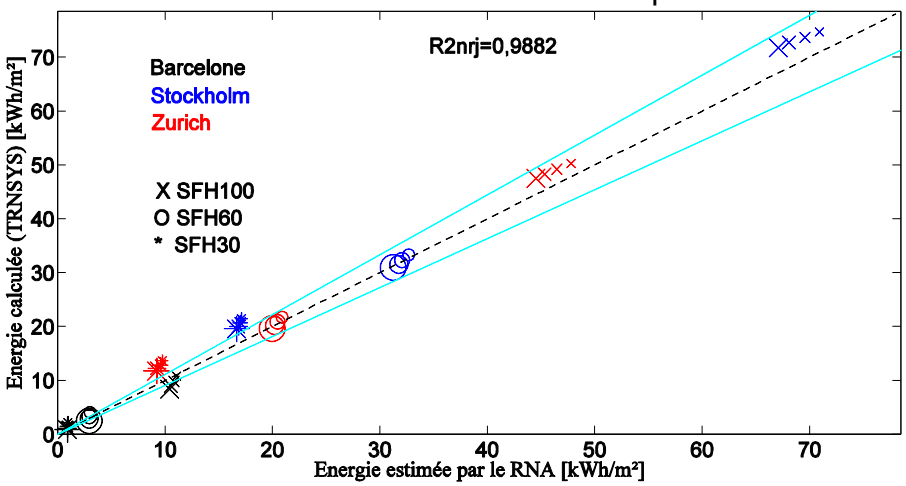

Figure 3- 20 : Comparaison entre la consommation annuelle estimée par le RNA-R2bar(linéaire) et calculée par le modèle physique TRNSYS dans différentes conditions. Cas du système SSCPACAE-REF.

La Figure 3- 21 permettent d'apprécier la capacité du modèle du RNA sélectionné par le « R2bar » à reproduire l'évolution temporelle de la puissance du compresseur dans l'environnement Barcelone-SFH100-19m² (conditions très différentes des conditions de la séquence de test).

Les deux autres critères ont encore sélectionnés le même modèle mais le pouvoir de généralisation de ce dernier est moins important (Tableau 3- 4).

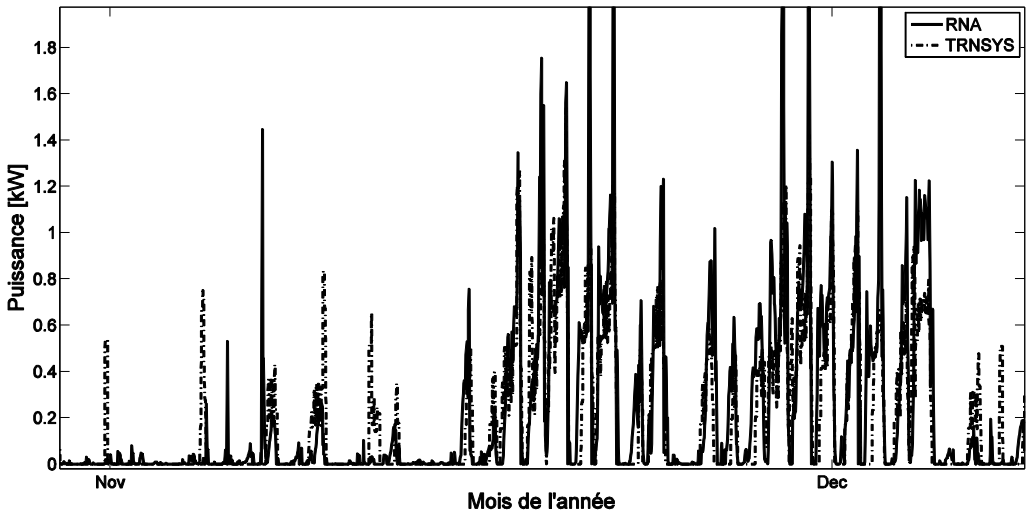

Figure 3- 21 : Zoom sur l'évolution de la puissance de l'appoint (PAC) du système SSCPACAE-REF au cours des mois de novembre et décembre. Cas de Barcelone-SFH100-19 m².

La modélisation des systèmes intégrant une PAC air/eau a été réalisée avec une deuxième configuration entrées-sortie de modèles neuronaux. En effet, le comportement d'une PAC air-eau est dépendant de la température de l'air, variable au cours de l'année (Figure 3- 22). Il est donc indispensable d'investiguer l'influence de la prise en compte de cette variable dans les modèles. Sur la Figure 3- 23 est représentée la configuration entrées-sortie correspondante.

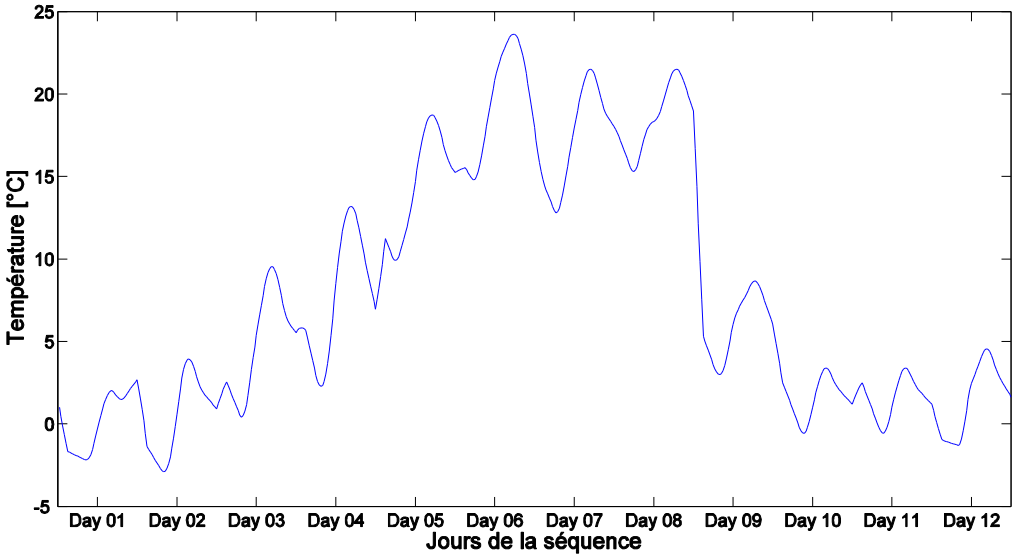

Figure 3- 22 : Evolution de la température extérieure au cours des 12 jours de la séquence de test (Apprentissage)

Le processus de la méthodologie a été appliqué au système SSCPACAE-REF, selon la nouvelle configuration, similairement aux cas précédents. Les résultats de l'application de la méthodologie ainsi que les caractéristiques des modèles étudiés dans cette partie sont données sur le Tableau 3- 5.

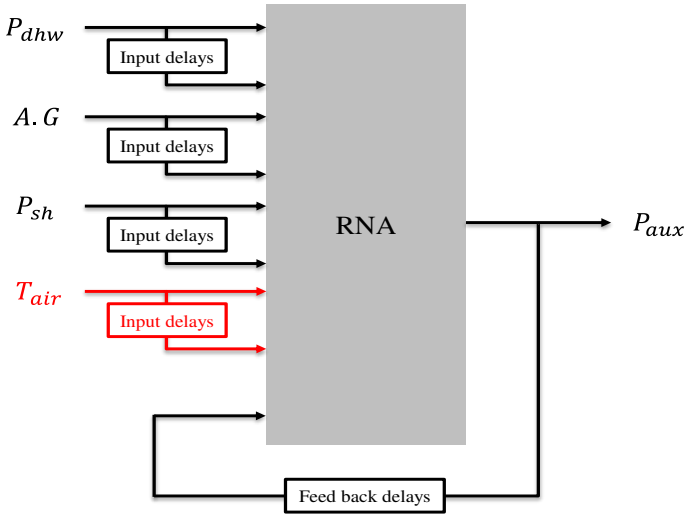

Figure 3- 23 : Configuration entrées-sortie, prenant en compte la température d'air extérieur (T_{air}), utilisée pour la modélisation des systèmes SSCPACAE-REF et SSCPACAE-AME

Tableau 3- 5 : Caractéristiques des modèles neuronaux (avec la configuration prenant en compte la température extérieure) étudiés (système SSCPACAE-REF)

Modèles	NORM	NDN	TD	« R2nrj »	« R2bar »	« BIC2 »	« AIC2 »
RNA-R2nrj(tanh)	±0,4	6,0	3,0	0,9900	0,8903	12,9109	12,8341
RNA-R2bar(tanh)	±1,0	6,0	3,0	0,7554	0,9832	11,0364	10,9596
RNA-BIC2(tanh)	±1,0	6,0	3,0	0,7554	0,9832	11,0364	10,9596
RNA-AIC2(tanh)	±1,0	6,0	3,0	0,7554	0,9832	11,0364	10,9596
RNA-R2nrj(linéaire)	±1,0	9,0	1,0	0,9786	0,9519	11,9848	11,9243
RNA-R2bar(linéaire)	±1,0	7,0	3,0	0,8773	0,9832	11,0811	10,9982
RNA-BIC2(linéaire)	±1,0	7,0	3,0	0,8773	0,9832	11,0811	10,9982
RNA-AIC2(linéaire)	±1,0	7,0	3,0	0,8773	0,9832	11,0811	10,9982

Sur les Figure 3- 24 et Figure 3- 25 sont présentés les résultats d'estimations des deux RNA RNA-R2nrj(tanh) et RNA-BIC2(linéaire). Globalement il n'y a pas de différences majeures à signaler en termes de qualité des résultats par rapport à la configuration entrées-sortie sans la prise en compte de la température de l'air. En effet, le coefficient de détermination est du même ordre de grandeur pour le modèle RNA-R2nrj(tanh) mais il est légèrement moins bon que les meilleurs modèles du cas précédent (RNA-R2nrj).

Pour chacune des deux FA, les critères ont sélectionnés le même RNA. Le modèle sélectionné pour cette nouvelle configuration est moins performant (« R2nrj » = 0,877) que dans le premier cas.

Il est possible de constater que les estimations de l'énergie dans le cas des conditions avec le bâtiment SFH30 se sont améliorées mais au détriment d'autres conditions (Bâtiments SFH60 pour RNA-R2nrj(tanh)).

En l'absence d'une nette amélioration des résultats, la configuration initiale sans prise en compte de la température est donc suffisante. Il n'est pas nécessaire de complexifier le modèle avec une nouvelle entrée.

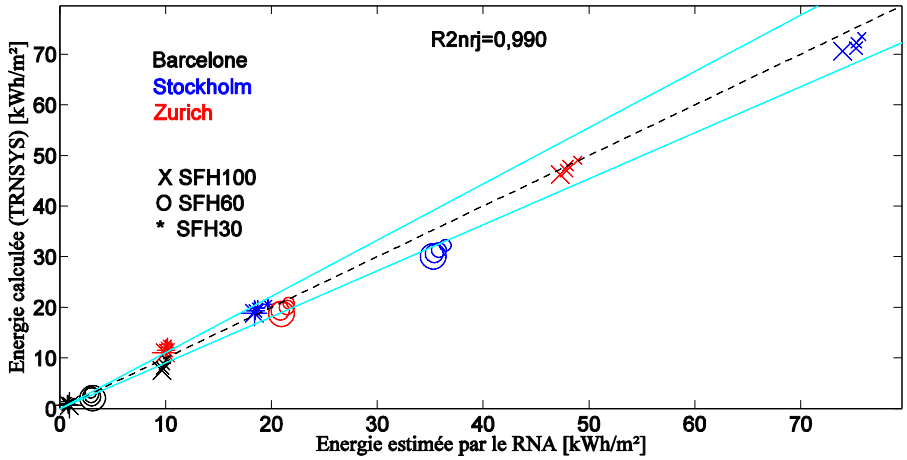

Figure 3- 24 : Comparaison entre la consommation annuelle estimée par le RNA-R2nrj(tanh) (avec la configuration prenant en compte la température extérieure) et calculée par le modèle physique TRNSYS dans différentes conditions. Cas du système SSCPACAE-REF. La taille des symboles indique la surface du capteur utilisée (10 m², 13 m², 16 m² ou 19 m²).

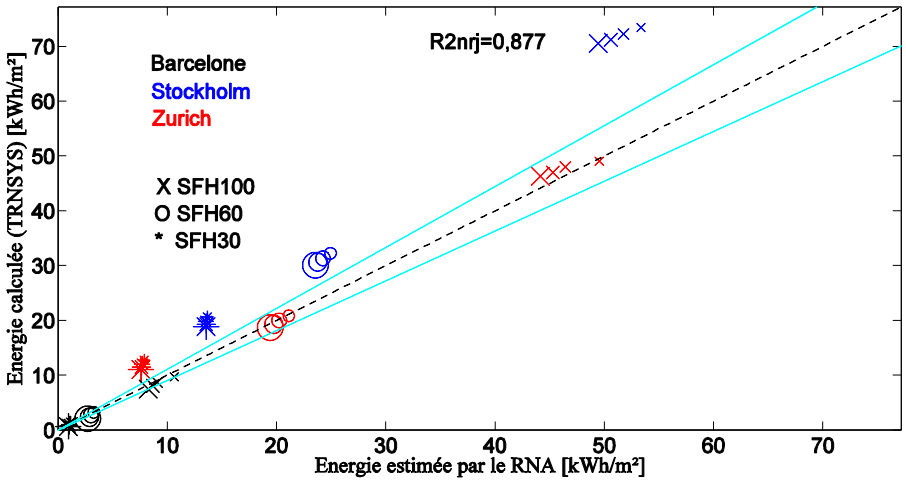

Figure 3- 25 : Comparaison entre la consommation annuelle estimée par le RNA-BIC2(linéaire) (avec la configuration prenant en compte la température extérieure) et calculée par le modèle physique TRNSYS dans différentes conditions. Cas du système SSCPACAE-REF. La taille des symboles indique la surface du capteur utilisée (10 m², 13 m², 16 m² ou 19 m²).

III.1.3.3 SSCPACEE-AME

Les résultats de l'application de la méthodologie au SSCPACEE-AME sont représentés sur la Figure 3- 26. Les caractéristiques des RNA étudiés dans cette partie sont données dans le Tableau 3- 6.

Tableau 3- 6 : Caractéristiques des modèles neuronaux étudiés (système SSCPACEE-AME)

Modèles	NORM	NDN	TD	« R2nrj »	« R2bar »	« BIC2 »	« AIC2 »
RNA-R2nrj(tanh)	±1,0	5,0	3,0	0,9947	0,4710	14,7056	14,6419
RNA-R2bar(tanh)	±1,0	11,0	1,0	0,7262	0,9674	11,9095	11,8476
RNA-BIC2(tanh)	±1,0	11,0	1,0	0,7262	0,9674	11,9095	11,8476
RNA-AIC2(tanh)	±1,0	11,0	1,0	0,7262	0,9674	11,9095	11,8476
RNA-R2nrj(linéaire)	±1,0	2,0	3,0	0,9591	0,9207	12,6949	12,6542
RNA-R2bar(linéaire)	±1,0	2,0	1,0	0,7334	0,9487	12,2099	12,1827
RNA-BIC2(linéaire)	±1,0	2,0	1,0	0,7334	0,9487	12,2099	12,1827
RNA-AIC2(linéaire)	±1,0	2,0	1,0	0,7334	0,9487	12,2099	12,1827

Les modèles TRNSYS utilisés pour évaluer les estimations des RNA ont été développés et validés uniquement pour une seule surface du capteur. C'est pour cette raison qu'une seule surface de capteur (Tableau 3- 2) a été retenue pour la définition des différentes conditions de test des modèles neuronaux. Nous avons fait varier en conséquence uniquement le climat et le type du bâtiment.

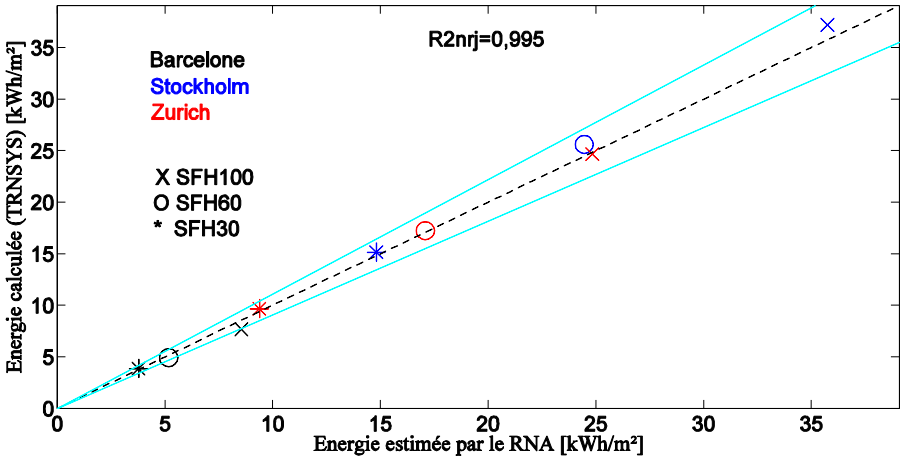

Figure 3- 26 : Comparaison entre la consommation annuelle estimée par le RNA-R2nrj(tanh) et calculée par le modèle physique TRNSYS dans différentes conditions. Cas du système SSCPACEE-AME.

Le meilleur modèle (RNA-R2nrj(tanh)) parmi tous les modèles créés estime avec de très faibles écarts relatifs la consommation du système pour toutes les conditions. En effet, ces derniers sont de moins de 5% sauf pour Barcelone-SFH100 où l'écart est de l'ordre de 10%.

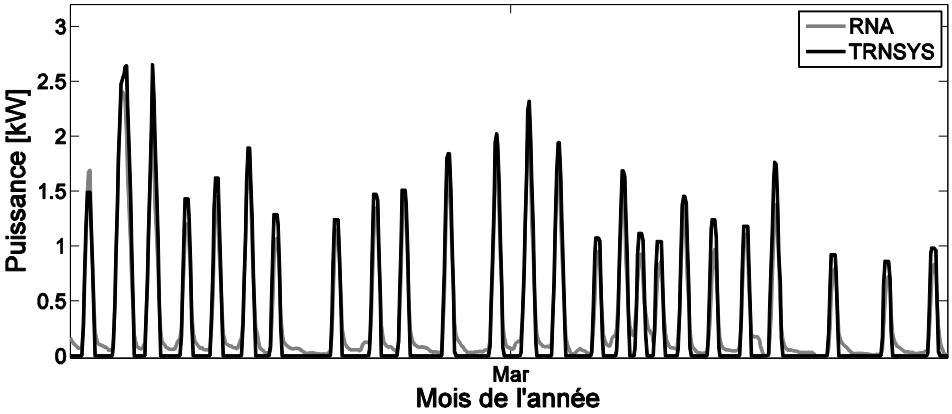

Figure 3- 27 : Zoom sur l'évolution de la puissance du l'appoint (PAC) du système SSCPACEE-AME au cours du mois de février. Comparaison entre les estimations du RNA et les calculs du modèle TRNSYS. Cas de Stockholm-SFH30-16 m².

Les modèles sélectionnés par les critères statistiques dans le cas du système SSCPACEE-AME ont des performances globales très modestes. Les coefficients de corrélations correspondants sont de l'ordre de 0,73.

La Figure 3- 27 montre qu'il y a pourtant une très bonne concordance entre la prédiction du modèle neuronal et les calculs de TRNSYS pour Stockholm-SFH30-16m². Les modèles RNA reproduisent mieux le vrai fonctionnement du système par rapport au cas du système de référence. L'évolution de la puissance dans ce cas est plus facile à maitriser par le RNA que dans le cas du système de référence. Cela est surement dû au fait que le fonctionnement de la PAC pour le système amélioré est plus uniforme (exemple sur la Figure 3- 27).

III.1.3.4 SSCPACAE-AME

Les mêmes conclusions que précédemment (cas du SSCPACEE-AME) restent valable pour SSCPACAE-AME. En effet, les prédictions du RNA-R2nrj(linéaire) (Tableau 3- 7) sont précises pour toutes les conditions, le coefficient de détermination correspondant est de l'ordre de 0,994 (Figure 3- 28).

Le modèle sélectionné par le critère « R2bar » pour la FA tanh a des performances très pauvres. En effet, le modèle ayant la valeur du R2bar la plus élevée, possède un coefficient de corrélation nul ce qui montre qu'il n'y a pas de corrélation entre ses estimations et les valeurs calculées par TRNSYS. Ce type de modèle avec des performances si faibles peut facilement être repéré et écarté parmi les modèles sélectionnés.

La concordance entre les prédictions du RNA et TRNSYS dans le cas des modèles sélectionnés par « BIC2 » avec une FA tanh est également très satisfaisante. Un exemple est donné sur la Figure 3- 29 pour l'environnement Stockholm-SFH60-16m².

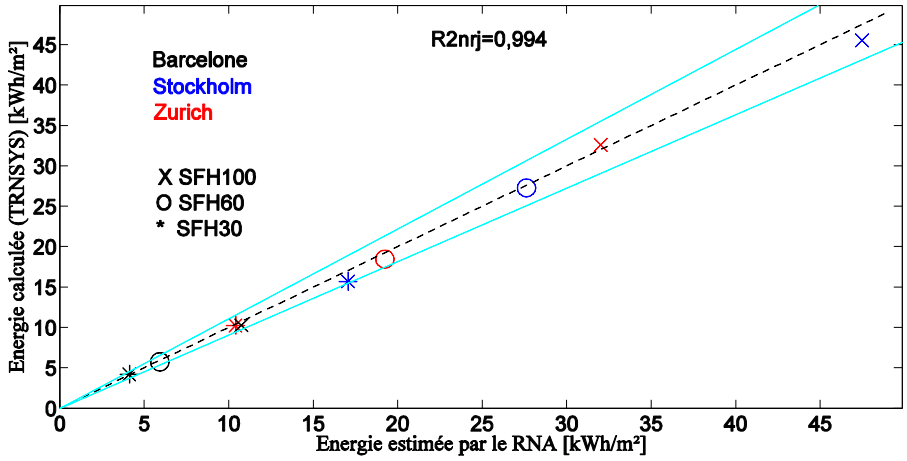

Figure 3- 28 : Comparaison entre la consommation annuelle estimée par le RNA-R2nrj(linéaire) et calculée par le modèle physique TRNSYS dans différentes conditions. Cas du système SSCPACAE-AME.

Tableau 3- 7 : Caractéristiques des modèles neuronaux étudiés (système SSCPACAE-AME)

Modèles	NORM	NDN	TD	« R2nrj »	« R2bar »	« BIC2 »	« AIC2 »
RNA-R2nrj(tanh)	±0,4	5,0	3,0	0,9875	0,9605	12,3738	12,3101
RNA-R2bar(tanh)	±0,4	10,0	3,0	0,0001	0,9810	11,8308	11,7410
RNA-BIC2(tanh)	±1,0	8,0	3,0	0,9478	0,9804	11,7874	11,7070
RNA-AIC2(tanh)	±1,0	8,0	3,0	0,9478	0,9804	11,7874	11,7070
RNA-R2nrj(linéaire)	±1,0	10,0	2,0	0,9939	0,5925	14,7826	14,7066
RNA-R2bar(linéaire)	±1,0	8,0	3,0	0,8440	0,9779	11,9050	11,8246
RNA-BIC2(linéaire)	±1,0	8,0	3,0	0,8440	0,9779	11,9050	11,8246
RNA-AIC2(linéaire)	±1,0	8,0	3,0	0,8440	0,9779	11,9050	11,8246

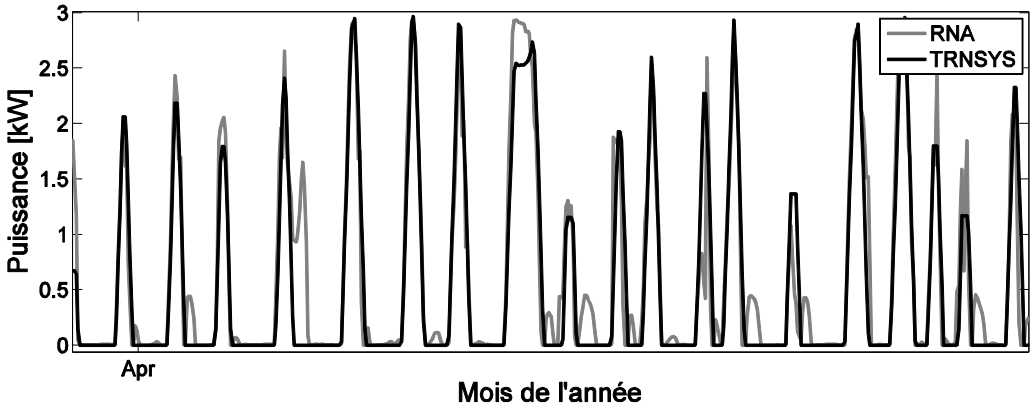

Figure 3- 29 : Zoom sur l'évolution de la puissance de l'appoint (PAC) du système SSCPACAE-AME au cours du mois de février. Comparaison entre les estimations du RNA et les calculs du modèle TRNSYS. Cas de Stockholm-SFH60-16 m².

La configuration entrées-sortie avec la prise en compte de la température de l'air extérieur a été également étudiée pour ce système. Les résultats des estimations des RNA (Tableau 3- 8) sont présentés sur la Figure 3- 30. Les coefficients de détermination pour les RNA sélectionnés par les différents critères (notamment RNA-BIC2 et RNA-AIC2) sont clairement moins bons que ceux de la configuration précédente. Cela montre que l'ajout de la température en entrée du modèle n'apporte pas d'information utile pour le RNA. Cette conclusion est en conformité avec celle du SSC de référence.

Tableau 3- 8 : Caractéristiques des modèles neuronaux (avec la configuration prenant en compte la température extérieure) étudiés (système SSCPACAE-AME)

Modèles	NORM	NDN	TD	« R2nrj »	« R2bar »	« BIC2 »	« AIC2 »
RNA-R2nrj(tanh)	±1,0	14,0	2,0	0,9818	0,7033	14,6523	14,5540
RNA-R2bar(tanh)	±0,4	15,0	3,0	0,9122	0,9910	11,4338	11,3128
RNA-BIC2(tanh)	±0,4	15,0	1,0	0,0001	0,9897	11,1271	11,0492
RNA-AIC2(tanh)	±0,4	15,0	1,0	0,0001	0,9897	11,1271	11,0492
RNA-R2nrj(linéaire)	±1,0	11,0	2,0	0,9828	0,9780	11,9522	11,8650
RNA-R2bar(linéaire)	±1,0	15,0	3,0	0,8295	0,9925	11,2596	11,1386
RNA-BIC2(linéaire)	±1,0	12,0	2,0	0,4946	0,9912	11,0772	10,9862
RNA-AIC2(linéaire)	±1,0	12,0	2,0	0,4946	0,9912	11,0772	10,9862

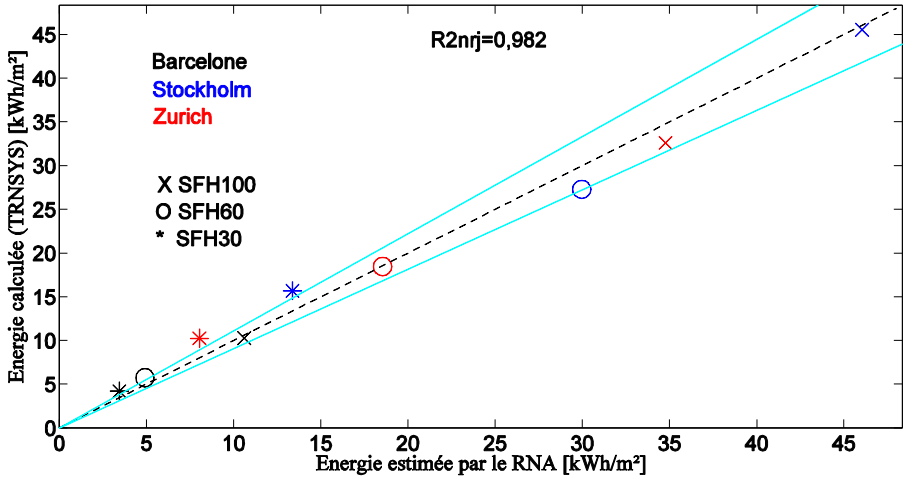

Figure 3- 30 : Comparaison entre la consommation annuelle estimée par le RNA-R2nrj(tanh) (avec la configuration entrées-sortie prenant en compte la température extérieure) et calculée par le modèle physique TRNSYS dans différentes conditions. Cas du système SSCPACAE-AME.

III.2. Deuxième typologie : systèmes solaires intégrant une machine à absorption

III.2.1. Description des systèmes et de leurs modèles physiques

L'extension de la méthodologie d'évaluation des SSC a été étudiée également pour une autre typologie de SSC. Il s'agit de SSC intégrant une machine à absorption pour répondre aux besoins de climatisation. Deux SSC de ce type de système ont été étudiés durant ce projet de recherche. Les deux systèmes ont été développés et dimensionnés dans le cadre du projet PACSOL. Ils ont les mêmes fonctions mais diffèrent dans leurs architectures hydrauliques. En effet, deux architectures répondant au cahier des charges suivant ont été déterminées :

- Systèmes destinés à certains types de bâtiments des pays du Maghreb, du sud de l'Europe et des Etats Unis.
- Systèmes avec un nombre limité d'auxiliaires (vannes, circulateurs et échangeurs) pour limiter les risques de défaillance.
- Système performant.
- Coût réduit

Dans le cas de la première architecture, Figure 3- 31, il est possible d'évacuer la chaleur en cas de surchauffes sur la tour de refroidissement. Elle permet également de répondre aux besoins de chauffage directement via le capteur solaire sans passage par la machine à absorption. En cas de ressources solaires insuffisantes, l'appoint en ligne prend la relève. La première architecture utilise uniquement des vannes à trois voies pour le passage d'un mode de fonctionnement à un autre (climatisation, ECS, chauffage). Un des inconvénients de cette architecture concerne le risque de gel du fluide dans les circuits du système. En effet, ces derniers sont tous connectés ; ainsi, il est impossible de mettre du fluide antigel dans les parties exposées au gel (la tour de refroidissement) sans en mettre dans tous les circuits du système, distribution dans le bâtiment incluse. En conséquence, il est indispensable d'utiliser des kits anti-glycol en guise de protection contre les risques de gel. Ce système sera noté dans la suite par SSCMA1.

La deuxième architecture, Figure 3- 32, permet de répondre également aux différentes conditions et exigences prescrites dans le cahier des charges. Contrairement à l'architecture précédente, celle-ci intègre un découplage physique entre la partie composant de stockage de la chaleur / ventilo-convecteur / distribution du fluide caloporteur et le reste de l'installation. Ainsi, il est possible de mettre du fluide antigel uniquement dans les circuits au niveau de l'évaporateur et de l'absorbeur/condenseur qui se trouvent à l'extérieur et donc soumis au risque de gel. Pour évacuer la chaleur en cas de surchauffes, pour cette architecture, le passage par la machine à absorption est obligatoire. Ainsi, le froid produit à l'évaporateur et le chaud à l'absorbeur/condenseur sont mélangés en sortie de la machine à absorption avant d'être évacués dans la tour de refroidissement. Des vannes à trois voies sont nécessaires pour permettre cette opération. Le découplage physique entre la partie distribution vers le bâtiment et le reste de l'installation implique l'installation d'une pompe et d'un échangeur à plaques du côté de l'évaporateur de la machine à absorption. Ce système sera noté dans la suite par SSCMA2.

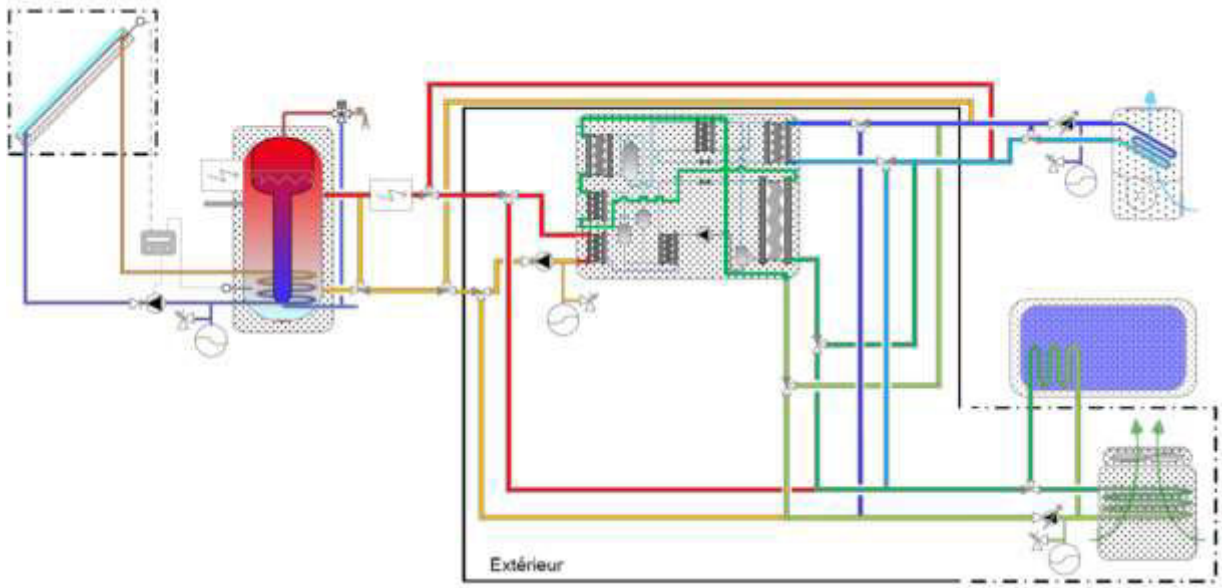

Figure 3- 31 : Schéma hydraulique du système SSCMA1

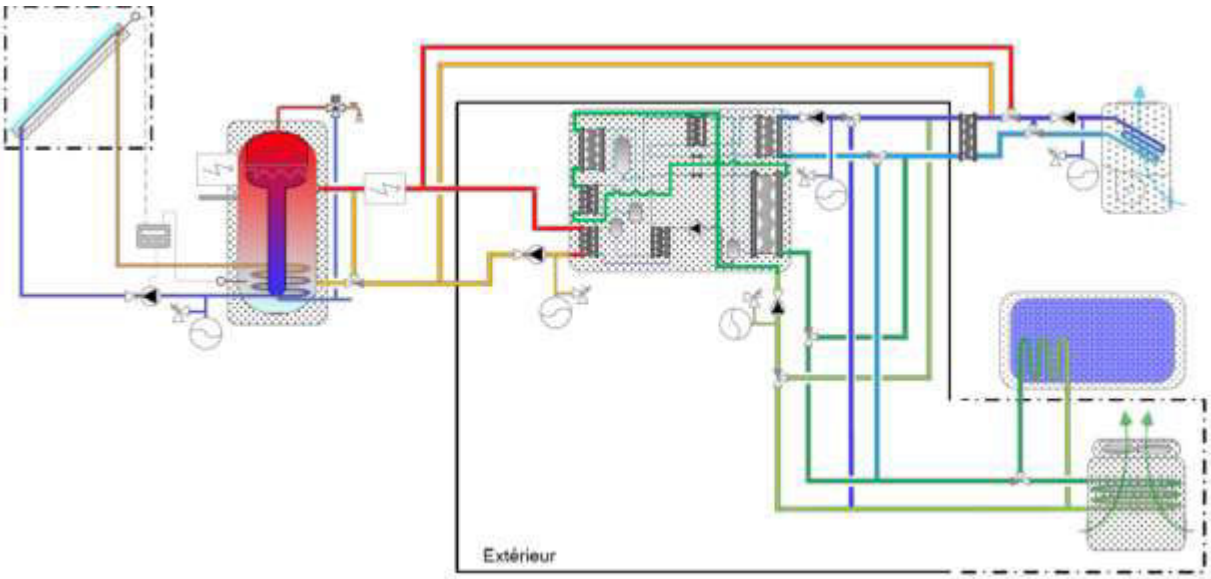

Figure 3- 32 : Schéma hydraulique du système SSCMA2

Les caractéristiques physiques des deux systèmes utilisés dans les simulations sous TRNSYS, sont données dans le Tableau 3- 9.

Tableau 3- 9 : Caractéristiques physiques utilisées pour la modélisation des systèmes SSCMA1 et SSCMA2 sous le logiciel TRNSYS

Remarques	Type TRNSYS	Caractéristiques
Machine à absorption		
Ce Type a été développé spécifiquement dans le cadre du projet. Il est basé sur les travaux de Ziegler [119]	TYPE 881	Machine à absorption NH3/H2O Puissance nominale de 7 kW.
Appoint électrique		
Intégré dans le ballon	Bloc d'équation	Puissance nominale de 9 kW
Une puissance constante ajoutée au fluide en sortie du ballon de stockage.	Bloc d'équation	Puissance nominale de 24 kW.
Tour de refroidissement		
Echangeur + ventilateur	Type 112 et Type 91	Débit nominal = 6000kg/h, consommation du ventilateur = 190W
Ballon de stockage		
Modèle détaillé stockage en eau tampon échangeur solaire + échangeur ECS immergé (tank-in tank)	Type 340	1000 litre de volume 2 m de hauteur 10 cm d'isolant. Coefficient de pertes latérales de 6,76 W/K et pour les parties supérieure et inférieure une valeur de 0,62 W/K
Capteur solaire		
Surface du capteur = 23 m²	TYPE 832	$F'(\tau\alpha)_{en}$=0,793; $K_{\theta d}$=0,9 c_1=0,749 (W/m²K); c_5=9180(J.m^{-2} K^{-1}); c_2=0,005 (W/m²K²) β=45° ; b_0=0,18

III.2.2. Application de la méthodologie

Comme il a été présenté précédemment, les deux systèmes solaires étudiés dans cette partie sont destinés à des environnements présentant un besoin en climatisation. Ils ont été en conséquence dimensionnés pour des conditions spécifiques différentes de ceux utilisés précédemment. Comme abordé dans le deuxième chapitre, les conditions de la séquence d'apprentissage doivent être des conditions « moyennes » en termes d'amplitudes par rapports aux différentes conditions que le système peut rencontrer. L'ensoleillement d'Aix-En-Provence est plutôt modéré par rapport aux autres climats (Philadelphie, Marrakech, etc.). Le bâtiment du Maghreb est un bâtiment dont les besoins de chauffages et de climatisation se situent entre ceux des bâtiments du sud de l'Europe et des Etats Unis. Il est donc judicieux d'utiliser le climat d'Aix-En-Provence combiné au bâtiment du Maghreb (conditions mixtes différentes des conditions auxquels les SSC sont destinés) pour les simulations des tests des deux systèmes selon la méthode SCSPT. Sur les Figure 3- 33 et Figure 3- 34 sont représentés les résultats des simulations des deux tests selon la méthode SCSPT.

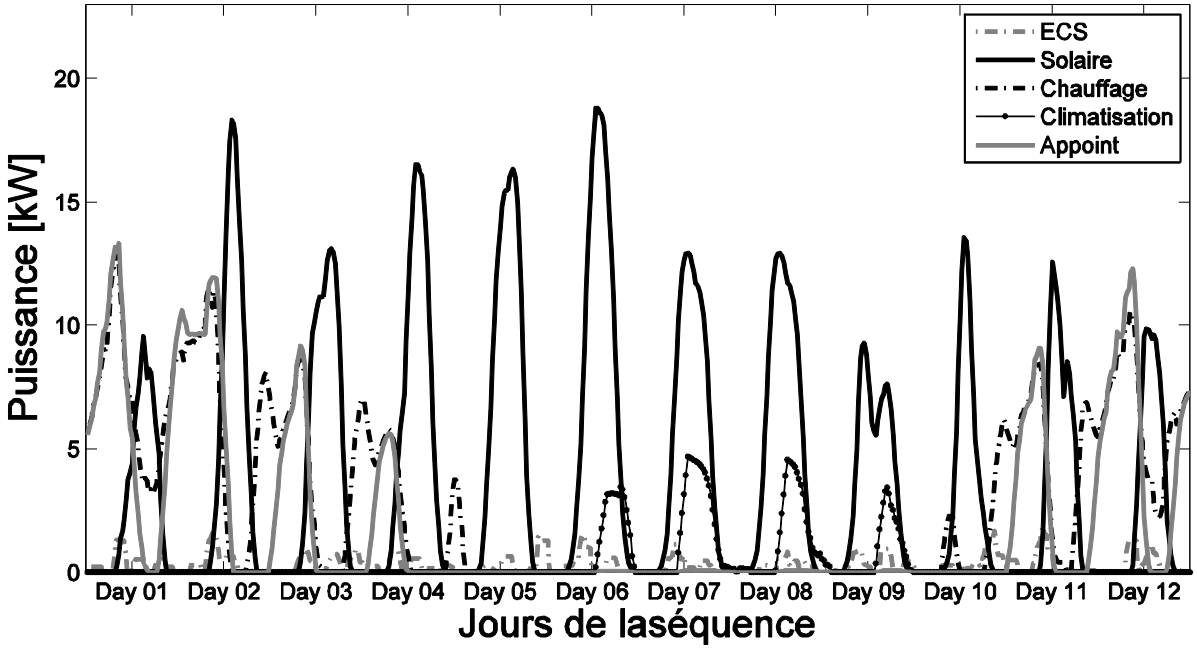

Figure 3- 33 : Séquence d'apprentissage pour la modélisation du système SSCMA1.

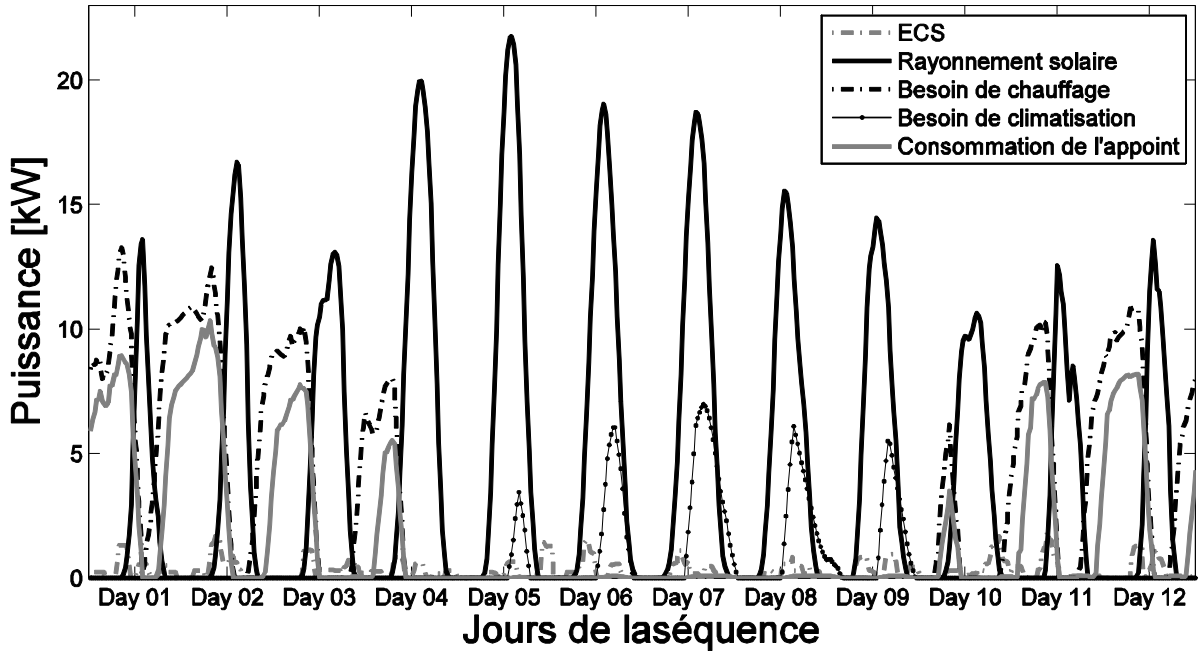

Figure 3- 34 : Séquence d'apprentissage pour la modélisation du système SSCMA2.

Les besoins de climatisation (variable notée P_{sc}), en termes de puissance, ont été rajoutés à la configuration de modélisation entrées-sortie. La configuration résultante est représentée sur la Figure 3- 35. Le processus de l'application de la méthodologie et de sa validation (voir le protocole dans la Figure 2- 19) est similaire aux systèmes précédents.

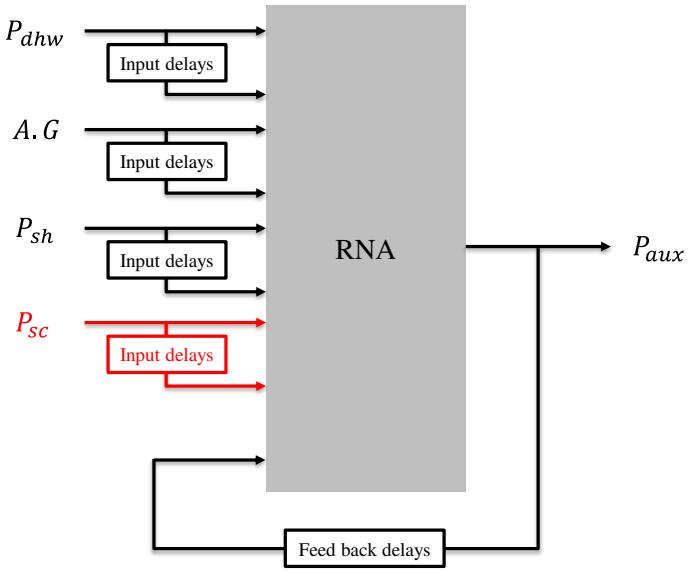

Figure 3- 35 : Configuration entrées-sortie utilisée pour la modélisation des systèmes SSCMA1 et SSCMA2

III.2.3. Résultats et discussions

III.2.3.1 SSCMA1

Les prédictions du meilleur modèle (RNA-R2nrj(tanh), voir également le Tableau 3- 10) parmi ceux créés durant le processus de la méthodologie sont précisées dans Figure 3- 36. Les erreurs sont très faibles (inférieurs à 5%) sauf pour certains climats de faibles besoins (notamment de chauffage) : Marrakech, Palerme et San Diego. Cependant, les écarts absolus sont faibles (maximum de 7,3 kWh/m² atteint à Palerme). Les raisons des difficultés du RNA pour ces conditions sont similaires à ceux évoqués dans le cas des systèmes précédents à savoir les grandes différences entre les conditions d'apprentissage et ceux des trois derniers climats.

Tableau 3- 10 : Caractéristiques des modèles neuronaux étudiés, cas du système SSCMA1

Modèles	NORM	NDN	TD	« R2nrj »	« R2bar »	« BIC2 »	« AIC2 »
RNA-R2nrj(tanh)	±1,0	8,0	3,0	0,9978	0,9951	-2,3772	-2,4658
RNA-R2bar(tanh)	±0,4	15,0	3,0	0,5684	0,9995	-4,3026	-4,4236
RNA-BIC2(tanh)	±0,4	15,0	3,0	0,5684	0,9995	-4,3026	-4,4236
RNA-AIC2(tanh)	±0,4	15,0	3,0	0,5684	0,9995	-4,3026	-4,4236
RNA-R2nrj(linéaire)	±1,0	7,0	3,0	0,9976	0,7288	1,5675	1,4847
RNA-R2bar(linéaire)	±1,0	12,0	2,0	0,1237	0,9988	-3,7487	-3,8397
RNA-BIC2(linéaire)	±1,0	12,0	2,0	0,1237	0,9988	-3,7487	-3,8397
RNA-AIC2(linéaire)	±1,0	12,0	2,0	0,1237	0,9988	-3,7487	-3,8397

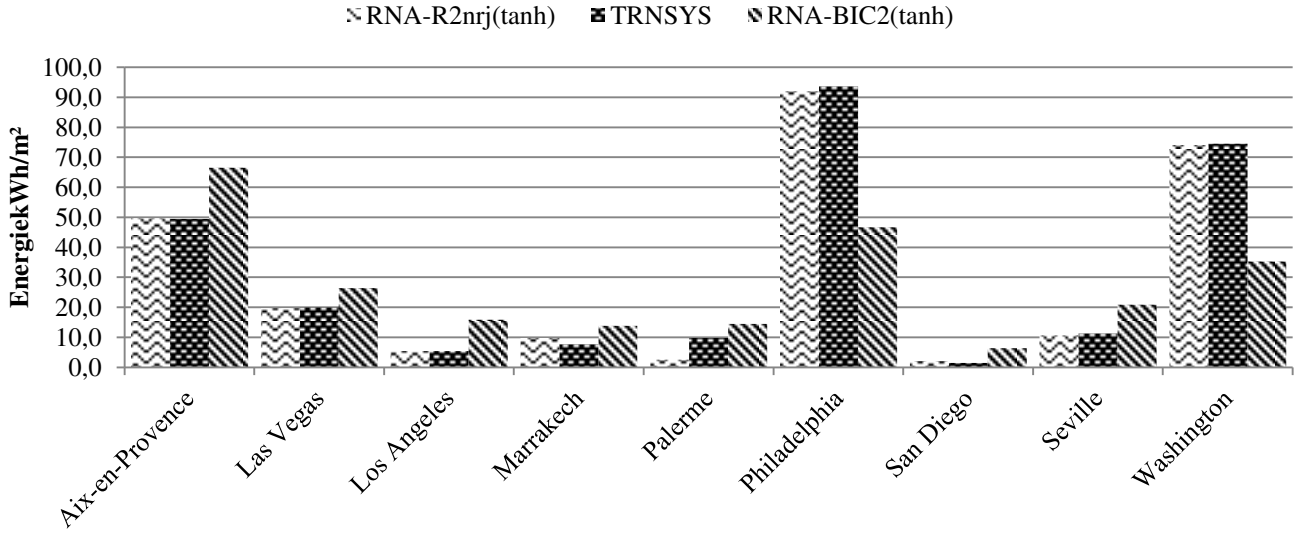

Figure 3- 36 : Comparaison entre la consommation annuelle estimée par le RNA-BIC2(tanh) et RNA-R2nrj(tanh) et calculée par le modèle physique TRNSYS dans différentes conditions. Système SSCMA1.

Les résultats de prédictions des modèle RNA sélectionnés par les critères « *BIC2* », « AIC2 » et « R2bar » révèlent la médiocrité de certains modèles neuronaux qu'un critère statistique peut être amené à sélectionner. Cependant, ce modèle est facilement détectable parmi les modèles à écarter. En effet, selon le modèle RNA-BIC2(tanh) la consommation du système SSCMA1 à Aix-En-Provence est comparable à sa consommation à Philadelphie ou

Washington. Or, cela est impossible. La recherche d'un autre RNA respectant les tendances des consommations en fonction des climats, toujours parmi les modèles sélectionnés (par ces critères) selon chaque configuration (intervalle de normalisation, retards et FA), n'a pas permis de sélectionner un modèle plus performant.

III.2.3.2 SSCMA2

Globalement les résultats (Figure 3- 37) pour le système SSCMA2 sont similaires à ceux du système précédent. Un écart très notable pour le cas du RNA-R2nrj(tanh) (ses caractéristiques sont données sur le Tableau 3- 11) concerne le climat d'Aix-En-Provence. La combinaison de ce climat avec le bâtiment ciblé sud de l'Europe doit certainement présenter des séquences extrêmes que le RNA-R2nrj(tanh) n'a pas rencontrées durant son apprentissage. En revanche, en l'état actuel des travaux, nous n'avons pas pu déterminer précisément les raisons de l'écart constaté. D'autant plus que les prédictions du RNA-R2nrj(tanh) dans le cas du SSCMA1 étaient très précises en particulier pour Aix-En-Provence.

Le modèle RNA-BIC2(tanh), sélectionné parmi tous les modèles par « BIC2 », donne globalement des estimations acceptables de la consommation du système pour certaines conditions. Les écarts les plus faibles sont constatés pour les climats ayant de faibles besoins de chauffage.

Sur la Figure 3- 38 est donné un exemple de la puissance de la consommation du système en fonction du temps pendant la période de climatisation (cas de Philadelphie et du RNA-BIC2(tanh)). Globalement le modèle suit correctement la vraie évolution de la puissance, ce qui montre que les écarts constatés dans le cas de certains climats se produisent pendant la période de chauffage (hiver). La création d'un modèle par période aurait peut être donné des résultats meilleurs. Dans le quatrième chapitre cette remarque a été prise en compte pour l'exploitation des données de suivi *in situ* d'une installation solaire intégrant une machine à absorption.

Tableau 3- 11 : Caractéristiques des modèles neuronaux étudiés, cas du système SSCMA2.

Modèles	NORM	NDN	TD	« R2nrj »	« R2bar »	« BIC2 »	« AIC2 »
RNA-R2nrj(tanh)	±0,4	12,0	1,0	0,9909	0,9011	0,1286	0,0589
RNA-R2bar(tanh)	±1,0	7,0	3,0	0,0001	0,9972	-3,3321	-3,4149
RNA-BIC2(tanh)	±1,0	4,0	3,0	0,8906	0,9968	-3,3359	-3,3987
RNA-AIC2(tanh)	±1,0	7,0	3,0	0,0001	0,9972	-3,3321	-3,4149
RNA-R2nrj(linéaire)	±1,0	12,0	2,0	0,9876	0,9863	-1,6858	-1,7768
RNA-R2bar(linéaire)	±1,0	14,0	3,0	0,8730	0,9980	-3,2848	-3,4017
RNA-BIC2(linéaire)	±1,0	14,0	3,0	0,8730	0,9980	-3,2848	-3,4017
RNA-AIC2(linéaire)	±1,0	14,0	3,0	0,8730	0,9980	-3,2848	-3,4017

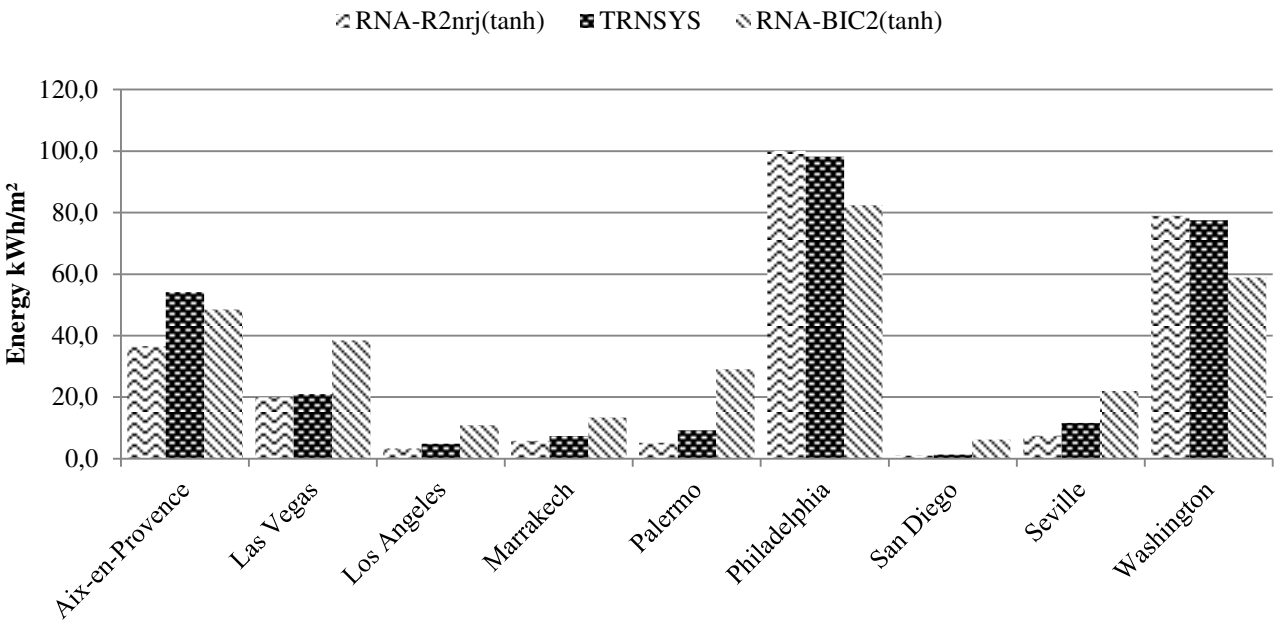

Figure 3- 37 : Comparaison entre la consommation annuelle estimée par le RNA-BIC2(tanh) et RNA-R2nrj(tanh) et calculée par le modèle physique TRNSYS dans différentes conditions. Système SSCMA2.

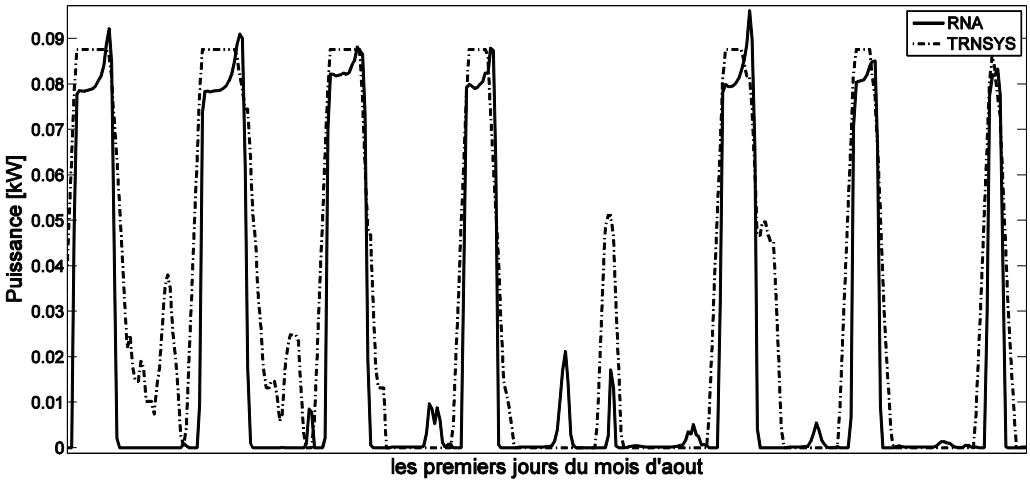

Figure 3- 38 : Zoom sur l'évolution de la puissance du l'appoint (Chaudière et machine à absorption) du système SSCMA2. Comparaison entre les estimations du RNA-BIC2(tanh) et les calculs du modèle TRNSYS. Cas de Philadelphie.

III.3. Bilan

Pour les meilleurs modèles, sélectionnés parmi tous les modèles neuronaux selon leurs qualités à estimer la consommation du système dans les 36 conditions, les prédictions des consommations annuelles sont très proches de celles calculées par les modèles TRNSYS. En effet, les coefficients de corrélations multiples sont entre 0,990 (pour le système SSCPACEE-REF) et 0,995 (pour le système SSCPACEE-AME). Globalement, l'erreur de prédiction est inférieure à 10% pour toutes les conditions et pour tous les systèmes sauf dans certaines

conditions spécifiques pour les deux systèmes de références. En effet, dans le cas du système SSCPACEE-REF l'erreur relative commise par le RNA dans les conditions présentant des besoins de chauffage faibles, en l'occurrence à Barcelone, est relativement importante. Cependant, il est important de signaler que l'écart absolu entre l'estimation du modèle neuronal et le calcul détaillé à l'aide de TRNSYS est assez faible : il est au maximum de l'ordre de 3,75 kWh/m². Tandis que dans le cas du système SSCPACAE-REF, l'erreur relative commise par le RNA est supérieure à 10%, mais reste inférieure à 14% à Stockholm pour le type de bâtiment SFH30. Il ressort en conséquence de cette analyse globale deux constatations :

- La qualité des résultats de l'application de la méthodologie dépend du système modélisé. Il est ainsi plus facile d'apprendre le comportement de certains systèmes que d'autres.
- Les erreurs sont relativement importantes uniquement pour des conditions très différentes des conditions d'apprentissage.

Les résultats apportés par les meilleurs modèles étant très satisfaisants, l'extension de la méthodologie développée au système intégrant une PAC est donc possible.

Il est important aussi d'analyser les résultats des modèles sélectionnés par les critères statistiques « *BIC2* », « *AIC2* » et « *R2bar* ». En effet, il est indispensable de disposer d'un moyen efficace permettant la sélection des modèles les plus pertinents pour que la méthodologie soit utilisée comme un outil de garantie de performances, sûr et fiable, dans un contexte de certification des SSC. Au vu des résultats, il apparait naturellement que les modèles neuronaux correspondants sont moins efficaces que ceux présentés précédemment. Les valeurs du « R2nrj » pour les modèles pertinents varient entre 0,56 (cas du SSCMA1) et 0,98 (cas du SSCPACAE-REF). Les modèles sélectionnés prédisent avec une erreur inférieure à 10% la consommation du SSC uniquement dans certaines conditions limitées qui dépendent de chaque système.

Nous constatons que les critères « AIC2 » et « BIC2 » ont sélectionné le même modèle pour tous les systèmes sauf SSCMA2. Pour ce dernier, la sélection via « BIC2 », dont le pouvoir de pénalisation est plus important, conduit à choisir un modèle à 4 neurones, ayant un meilleur pouvoir de généralisation que le réseau à 7 neurones qui serait sélectionné via le critère « AIC2 ». La pénalisation plus sévère du « BIC2 » est plus adaptée dans ce cas.

Les résultats d'applications ne nous permettent pas de donner un avis tranché sur la pertinence d'un critère sur un autre. Dans certains cas les critères ont sélectionnés des modèles performants, pour d'autres ce n'était pas le cas. La sélection des modèles pour les trois critères statistiques présente des risques de sélectionner des modèles très médiocres (cela a été constaté en particulier dans le cas de SSCM1). Il faut en conséquence ne pas donner une confiance totale à ces critères. Une étude a *posteriori* des modèles sélectionnés est nécessaire pour déceler ces modèles à risque.

Globalement des conclusions similaires sont à rapporter dans le cas des deux systèmes intégrant une machine à absorption (SSCMA1 et SSCMA2). Nous constatons également que les erreurs les plus conséquentes concernent très souvent (pour les meilleurs modèles dans le

sens de « R2nrj ») les conditions pour lesquelles les besoins sont faibles notamment pour le chauffage (à Palerme par exemple).

L'extension de la méthodologie développée au système intégrant une machine à absorption est donc possible.

Par ailleurs, il apparait que les modèles avec une FA tanh sont les plus performants : pour tous les systèmes, le meilleur modèle (selon « R2nrj ») possède un coefficient de détermination supérieure à 0,98. Il apparait aussi que le choix d'un nombre de neurones dans la couche cachée au maximum égal à 15 est juste. En effet, pratiquement pour tous les cas étudiés les modèles sélectionnés ont un nombre de neurones inférieur à 14.

CONCLUSION

Les expérimentations numériques du troisième chapitre ont permis de valider la méthodologie proposée et ainsi montrer qu'il est possible de l'appliquer à différentes typologies de systèmes. En effet, la méthodologie a été appliquée telle que définie dans le deuxième chapitre à six systèmes différents. Ces systèmes diffèrent par leur schéma hydraulique mais aussi par leur système d'appoint.

Les résultats de cette partie montrent qu'il est possible de prédire la consommation annuelle d'un système solaire thermique innovant pour un nombre assez important de conditions. Cependant le pouvoir de généralisation des modèles neuronaux trouve ses limites, dans la plupart des cas, lorsqu'on est en présence de conditions extrêmes très différentes des conditions de l'apprentissage.

Les résultats de l'application de la méthodologie ont révélé qu'il est difficile de sélectionner des modèles ayant des performances très proches des meilleurs modèles en se basant uniquement sur les données d'apprentissage.

Compte-tenu des résultats jugés satisfaisants la méthodologie est finalement appliquée à des cas réels dans le dernier chapitre.

Chapitre IV. Expérimentation physique : Application de la méthodologie dans le cas de SSC réels

INTRODUCTION

Dans le but de compléter le processus de validation de la méthodologie, il est indispensable de la confronter à des situations réelles d'application pour lesquelles les données de mesures peuvent présenter des incertitudes. Des expérimentations physiques ont donc été réalisées. Plusieurs systèmes solaires de différentes typologies, ont été testés physiquement selon la méthode SCSPT et ont été ensuite évalués selon la méthodologie élaborée dans le cadre de la thèse. Des données de suivi *in situ* d'une installation solaire thermique intégrant une machine à absorption ont également été exploitées pour tester la méthodologie développée.

Le quatrième chapitre commence par donner des informations générales sur le banc de test semi-virtuel ainsi que sur le processus de déroulement des tests. Il se termine par la présentation des résultats de l'application de la méthodologie développée dans le cas d'essais physiques de systèmes.

IV.1. Présentation détaillée du banc d'essai semi-virtuel

IV.1.1. Caractéristiques techniques du banc d'essai semi-virtuel

Le mode opératoire de la méthode SCSPT a été préalablement introduit dans le premier chapitre, paragraphe I.2.5. Dans cette partie sont présentées les caractéristiques techniques et physiques du banc d'essai semi-virtuel présent dans les locaux de l'INES. Ce banc a permis l'application de ce genre de test sur plusieurs systèmes. Il est composé de cinq éléments principaux :

- Une cellule de test

Elle permet d'accueillir le système à tester ainsi que tout le matériel nécessaire aux tests (modules hydrauliques, baies de mesures, les différentes alimentations etc.). L'air de la cellule est conditionné par un réseau aéraulique permettant de maintenir une température intérieure de 20°C±1°C. Sur la Figure 4- 1 est représenté un exemple d'un SSC installé sur le banc semi-virtuel.

Tous les composants du système à évaluer, à l'exception du capteur solaire et de l'échangeur géothermique (qui sont émulés), doivent être installés et connectés aux différents modules hydrauliques. Ces derniers simulent la connexion à un bâtiment (dans lequel le système est virtuellement installé). Les températures de retour (des modules hydrauliques aux composants du système) sont contrôlées et régulées par des programmes informatiques (LabVIEW et TRNSYS). Les modules hydrauliques sont approvisionnés par un réseau de distribution d'eau chaude et un autre d'eau froide.

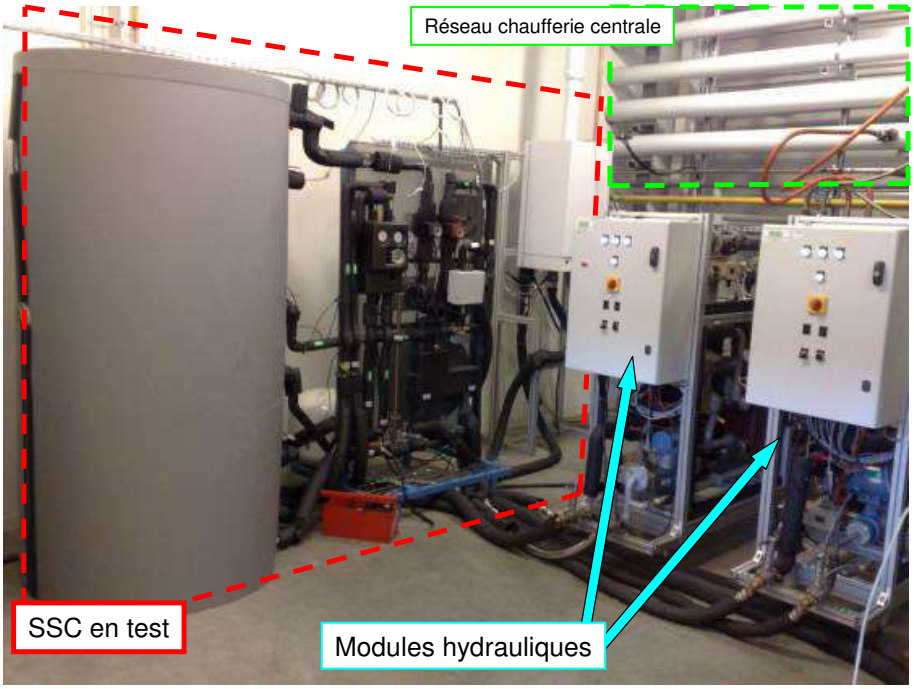

Figure 4- 1 : La cellule de test disponible à l'INES avec un exemple de SSC installé et connecté aux modules hydrauliques

- Une chaufferie centrale

La chaufferie centrale est un système permettant de produire et distribuer de l'eau chaude (avec une puissance de 50 kW et une température maximale de 180°C) et l'eau froide (avec une puissance de 150kW et une température minimale de -10°C) avec deux circuits totalement indépendants.

- Quatre modules hydrauliques de 25kW (Figure 4- 2)

Figure 4- 2 : Le module hydraulique 25kW

Ils servent à émuler la charge et décharge thermique sur tous les circuits du système sauf le circuit d'ECS. Les quatre modules sont pilotés selon le programme TRNSYS (modèles du bâtiment, capteur solaire, radiateurs, échangeur géothermique, etc.) auxquels ils sont connectés via LabVIEW. Le schéma hydraulique de ces modules est représenté sur la Figure 4- 3. En fonction des informations reçues, le module prépare une température de consigne de départ du module vers le système. Ainsi, le régulateur intégré à chaque module 25 kW pilote les vannes VIP1, VIP2, VDP3, VDP4, ainsi que les pompes P07 et P08 pour régler la température de départ.

La plage des températures mesurables via les modules 25 kW est de -10°C à 170°C avec des débits qui vont de 100l/h à 3600l/h.

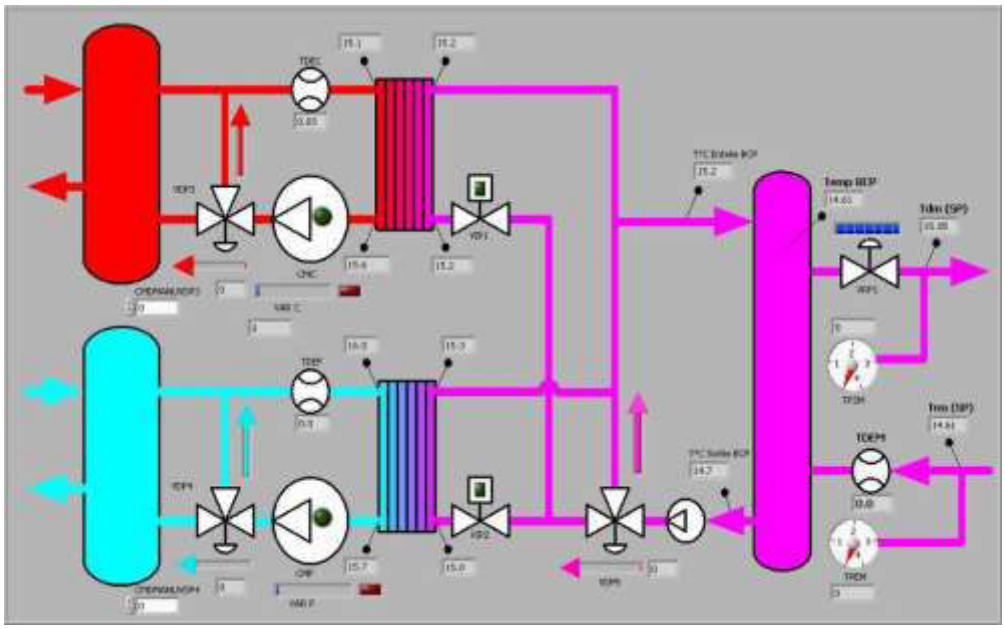

Figure 4- 3 : Schéma hydraulique d'un module 25kW. Le circuit rouge (en haut à gauche) est raccordé au réseau hydraulique de distribution de fluide chaud. Le circuit bleu (en bas à gauche) est raccordé au réseau hydraulique de distribution de fluide froid. Le réseau mauve (à droite) est raccordé à l'équipement à tester.

- Un module hydraulique de 50 kW (Figure 4- 4)

Ce module sert à émuler le puisage d'ECS. Le circuit hydraulique correspondant est représenté sur la Figure 4- 5. Son principe de fonctionnement et similaire aux modules 25 kW. La plage de température accessible via ce module est de 2°C à 25°C avec des débits qui vont de 1l/min à 64l/min.

Figure 4- 4 : Le module hydraulique 50kW

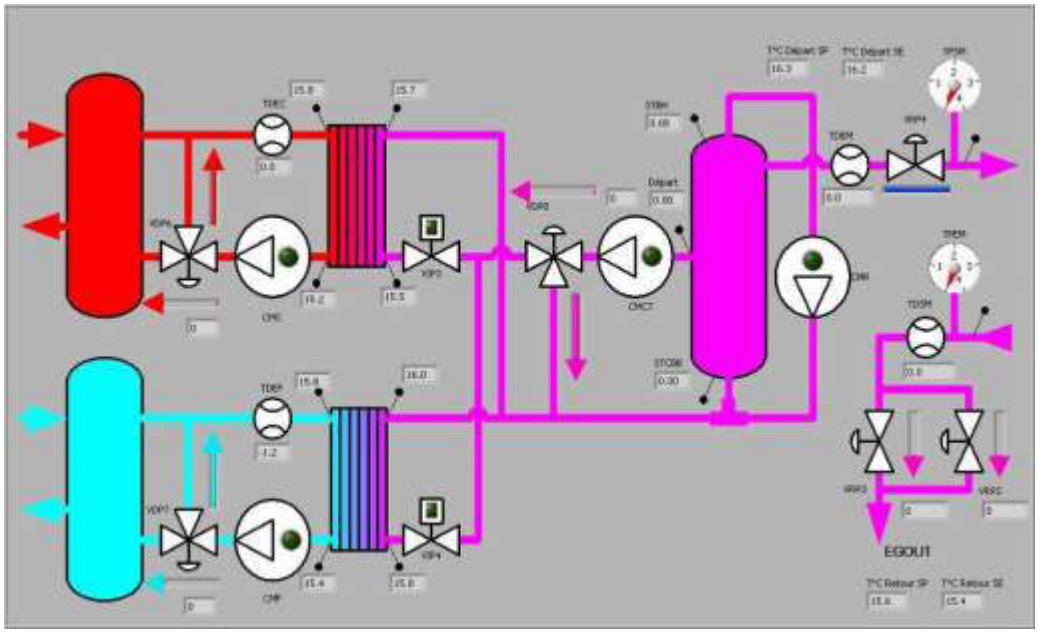

Figure 4- 5 : Schéma hydraulique du module 50kW

- Baies de mesures (Figure 4- 6) assurant :

- L'acquisition des mesures de certaines températures (via des sondes à résistance de platine PT100) et débits (via des débitmètres de type Coriolis ou de type électromagnétique selon les modules) aux bornes du système à tester,
- L'alimentation électrique et l'acquisition de la mesure de la consommation électrique ou en gaz (à travers la mesure de son volume au cours du temps) de l'équipement à tester,
- La génération de signaux résistifs émulant des sondes de température (température extérieure et d'ambiance) pour l'équipement à tester.

Figure 4- 6 : La baie de mesures

Les baies de mesures sont connectées, via un réseau Ethernet, à un poste central de supervision, équipé du logiciel LabVIEW. Ce dernier permet le transfert des données avec les modules, la connexion avec le logiciel TRNSYS (simulant l'environnement virtuel du test) et la visualisation du déroulement de la séquence de test.

IV.1.2. La méthode SCSPT théorique et l'essai physique des systèmes

Les conditions aux limites des systèmes utilisées lors des tests (variables climatiques, bâtiment, apports internes et solaires etc.) sont présentées dans le deuxième chapitre. Elles sont émulées par les modules hydrauliques présentés ci-dessus. Elles sont générées par l'algorithme de la méthode SCSPT présenté dans le premier chapitre. Cependant les limites techniques des modules hydrauliques et de leur régulation imposent quelques légères modifications aux conditions présentées réellement aux systèmes notamment au niveau de la modélisation des besoins en ECS :

- Le débit minimal du puisage d'ECS imposé par les capacités physiques du module 50 kW est limité à 60 kg/h. En conséquence, la consommation journalière d'ECS passe de 201 litres, selon le profil déterminé initialement par la méthode SCSPT, à 203 litres.

- Pour que le module 50 kW puisse réguler correctement la température d'eau froide, le profil théorique de cette dernière, déterminé par l'algorithme de la méthode SCSPT, voir section I.2.5. , est remplacé par un profil en paliers journaliers de température (Figure 4- 7). Ce profil est choisi ainsi pour respecter grossièrement l'évolution sinusoïdale du profil théorique. Les paliers sont calculés sur la base d'une moyenne de 3 jours. Les besoins en ECS passent ainsi de 99,78kWh (profil sinusoïdal) à 99,25kWh (profil par paliers) sur l'ensemble de la séquence de 12 jours pour le climat de Zurich (en prenant en compte la contrainte sur le débit ci-dessus).

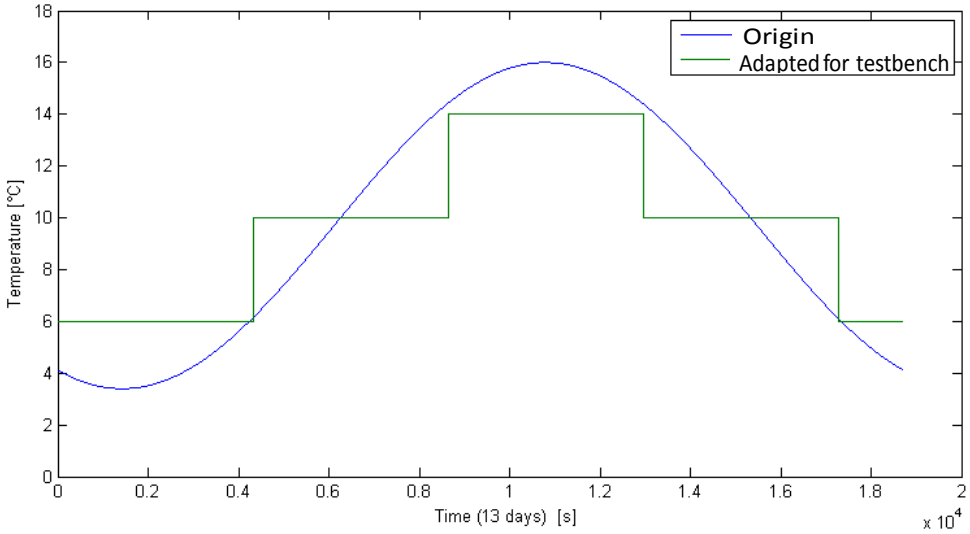

Figure 4- 7 : Evolution théorique et réel de la température d'eau froide lors des essais pour le climat de Zurich

Les deux modifications ci-dessus n'apportent pas de modifications significatives sur les résultats des essais (en comparaison avec la méthode SCSPT théorique).

IV.2. Test du premier système commercial

IV.2.1. Description du système et de son modèle physique

Ce système est un SSC compact avec un appoint par une chaudière à gaz actuellement disponible sur le marché. Le schéma hydraulique du premier système est représenté sur la Figure 4- 8. Il est composé principalement d'un ballon de stockage, d'une chaudière à condensation, d'une panoplie hydraulique permettant le raccordement de l'ensemble des composants et d'un système de régulation innovant développé par l'industriel.

Globalement ce système fonctionne comme le SSC présenté dans le deuxième chapitre. En effet, il est composé de quatre circuits :

- Un circuit solaire avec un échangeur (serpentin) interne dans la partie basse du ballon de stockage.
- Un circuit de chauffage puisant la chaleur via un piquage au milieu de la partie supérieure du ballon.
- Un circuit d'appoint permettant la préparation de l'ECS, via un piquage dans la partie supérieure du ballon, mais également du chauffage direct sans passage par le réservoir d'eau chaude. Le basculement entre les différents modes de chauffage sont réalisés à l'aide d'un jeu de vannes trois voies.
- Un circuit pour la production d'ECS via un échangeur (serpentin) traversant toute la longueur du ballon.

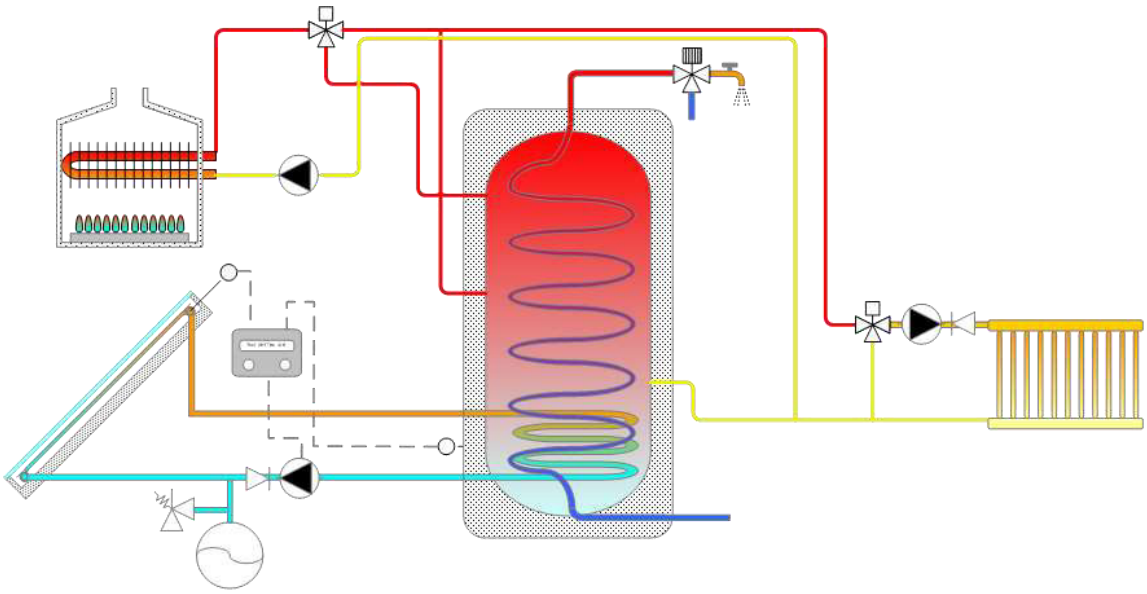

Figure 4- 8 : Schéma hydraulique du premier système

La particularité de ce système par rapport au SSC du deuxième chapitre réside dans les deux éléments principaux de fonctionnement suivants :

- En mode chauffage direct, la chaudière chauffe également la partie du milieu du ballon de stockage. Ce mode de fonctionnement n'existe pas dans le cas du SSC du deuxième chapitre.
- Il est possible de fonctionner en mode chauffage via le ballon de stockage sans passer par la chaudière. Tandis que le fonctionnement dans ce mode, dans le cas du SSC du deuxième chapitre, exige le passage par la chaudière même en arrêt.

Le système a été installé sur le banc d'essai semi-virtuel, en vue de la réalisation des essais, conformément à la méthodologie SCSPT. Son installation (ainsi que pour les autres systèmes) consiste à réaliser :

- Un raccordement des circuits du système aux modules hydrauliques, émulant le fonctionnement d'un circuit de chauffage, d'un circuit d'ECS et d'un circuit solaire.
- L'installation des différents appareils de mesure des débits et températures aux bornes de chacun de ces circuits ainsi que de la consommation en gaz de la chaudière.
- L'installation des résistances émulant les sondes de températures nécessaires à la régulation du système (température intérieure, température extérieure, température capteur).

Tous les circuits du système ont été remplis avec de l'eau pure (fluide caloporteur), contrairement aux installations sous un environnement réel où les fluides glycolés sont privilégiés (notamment dans le circuit solaire) pour éviter les risques de gel. Ce choix a été réalisé pour pallier la difficulté d'avoir des valeurs suffisamment précises pour la capacité calorifique des fluides glycolés.

Le premier système a été testé selon deux configurations physiques. La première avec un système de régulation standard (régulation de base). La deuxième consiste en une configuration, avec un système de régulation optimisé tout en gardant le même schéma hydraulique. Les systèmes résultants sont donc différents malgré leur ressemblance physique. Dans la suite, le premier système sera noté SSC-STD tandis que le deuxième sera noté SSC-OP. Deux tests physiques de 12 jours ont été réalisés pour chaque système dans les conditions suivantes :

- Zurich-SFH60-16 m², l'entrainement des modèles a été réalisé dans cet environnement (Figure 4- 9 et Figure 4- 10).
- Zurich-SFH100-16 m², environnement utilisé pour la validation expérimentale. La loi de chauffe (courbe fixant la température de consigne en entrée des radiateurs en fonction de la température extérieure) utilisée lors du test des deux systèmes est différente de celle utilisée lors du test sous l'environnement Zurich-SFH60-16 m². En effet, ces tests ont été préalablement prévus pour les besoins d'un autre projet. L'ajustement de la loi de chauffe a été réalisé pour que la gestion des besoins de chauffage soit adaptée au bâtiment de l'environnement virtuel correspondant à l'essai en cours. Le deuxième test pour les deux systèmes permettra donc de tester la robustesse de la méthodologie dans le cas d'un système légèrement différent de celui qui a été utilisé lors de l'entrainement.

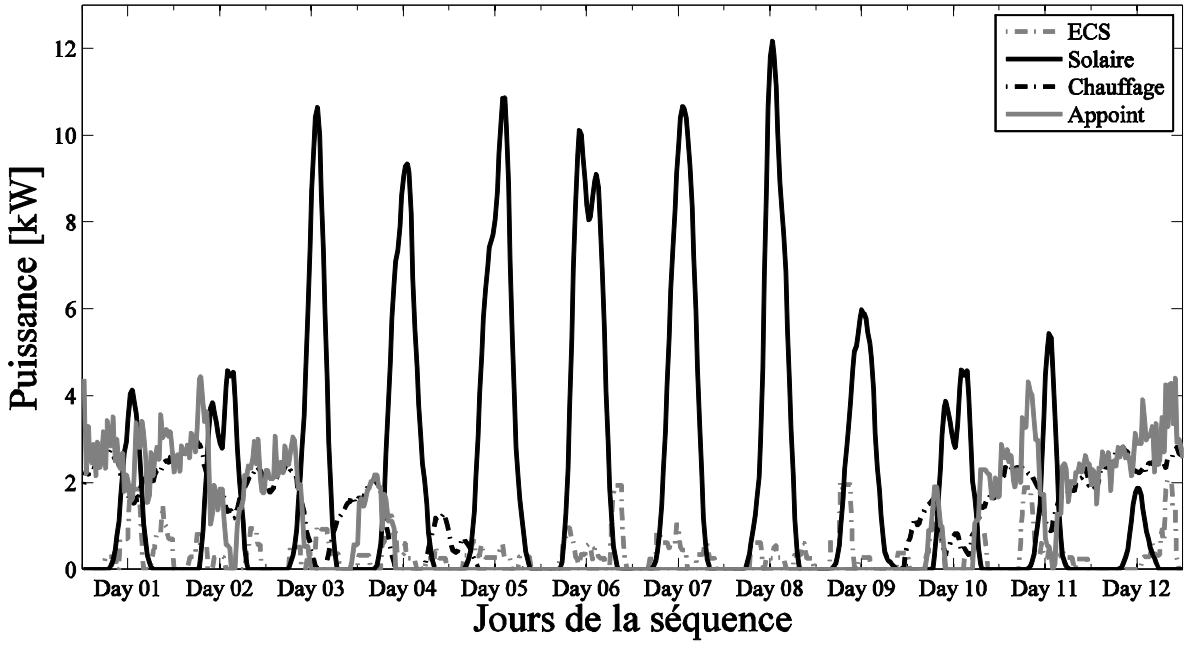

Figure 4- 9 : Séquence d'apprentissage pour la modélisation du système SSC-STD. Conditions dans lesquelles le test selon la méthode SCSPT a été réalisé: Zurich-SFH60-16m²

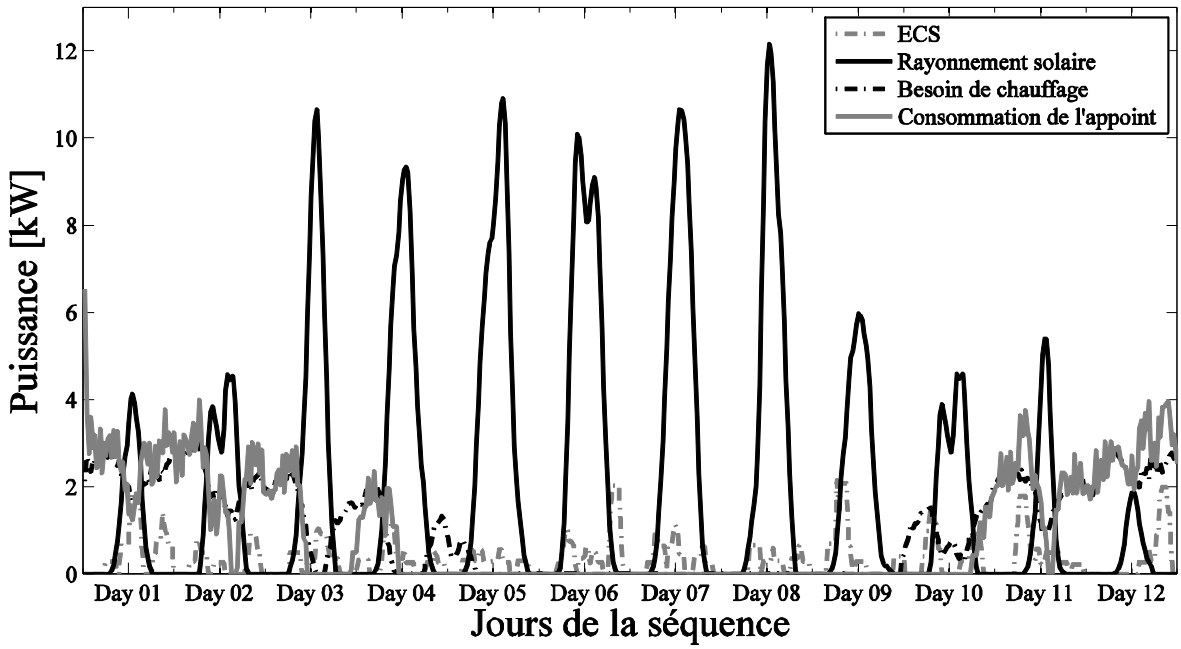

Figure 4- 10 : Séquence d'apprentissage pour la modélisation du système SSC-OP. Conditions dans lesquelles le test selon la méthode SCSPT a été réalisé: Zurich-SFH60-16m²

Un modèle physique pour chacun des deux systèmes a été développé (sous le logiciel TRNSYS) et validé dans le cadre du projet Systheff. Les différents paramètres des composants du système, nécessaires pour la construction du modèle physique, ont été identifiés grâce aux tests réalisés. Les résultats présentés dans la Figure 4- 11 témoignent de la qualité des modèles TRNSYS validés qui seront utilisés dans la suite pour réaliser les simulations annuelles (voir protocole Figure 2- 20). L'identification d'un modèle TRNSYS ne fait pas partie de la méthodologie proposée. Nous en profitons uniquement pour valider les modèles neuronaux.

Les caractéristiques de ce dernier sont présentées sur le Tableau 4- 1. Grâce à ce modèle il sera possible de valider la méthodologie dans différentes conditions à l'aide de simulations annuelles selon le protocole « validation semi expérimentale » décrit dans le deuxième chapitre (paragraphe II.4.2.1).

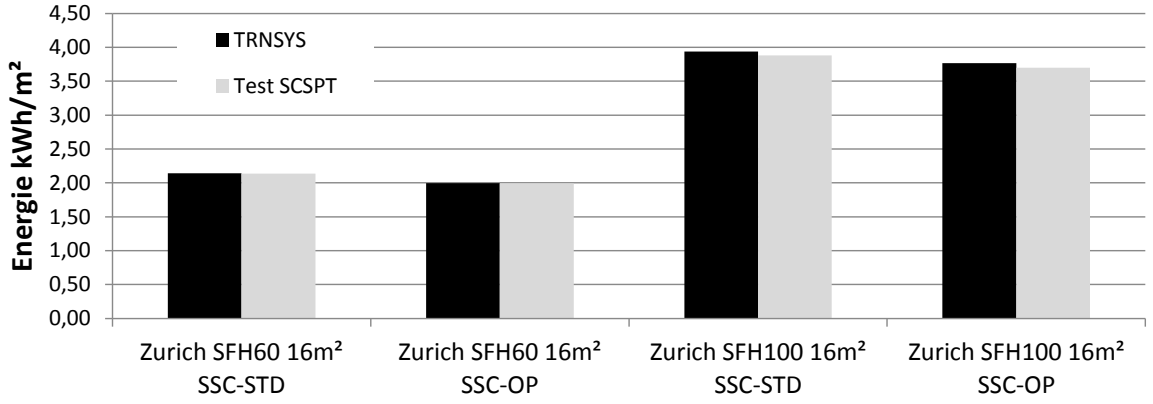

Figure 4- 11 : Comparaison entre la consommation des systèmes SSC-STD et SSC-OP mesurée et calculée à l'aide de TRNSYS dans les deux conditions de test selon la méthode SCSPT

Tableau 4- 1 : Caractéristiques physiques utilisées pour la modélisation des systèmes SSC-STD et SSC-OP dans TRNSYS

Remarques	Type TRNSYS	Caractéristiques
	Appoint	
Type développé en interne calculant les performances de la chaudière à partir d'une table de performances en fonction de la température de la chaudière et de la charge.	Type 5011	Puissance nominale de 19,2 kW, Chaudière à gaz,
	Ballon de stockage	
	Type 340	735,9 litre ; 1,56 m de hauteur Coefficient de pertes latérales de 1,86 W/K et pour les parties supérieure et inférieure une valeur de 1,5 W/K
	Capteur solaire	
	Type 832	$F'(\tau\alpha)_{en}$=0,807; $K_{\theta d}$=0,83; c_1=4,031 (W/m²K); b_0=0.18 c_2=0,012 (W/m²K²) ; β=45°; c_5=7000 (J.m^{-2} K^{-1})

IV.2.2. Résultats

La méthodologie de modélisation objet de la thèse a été appliquée aux systèmes SSC-STD et SSC-OP de la même façon qu'aux systèmes étudiés dans le deuxième chapitre. Dans un premier temps les données des tests (Zurich-SFH60-16 m²) vont servir à la construction des modèles neuronaux tandis que les modèles physiques détaillés (TRNSYS) vont servir à la validation semi-expérimentale de la méthodologie selon le protocole décrit dans le paragraphe II.4.2.1. Dans un second temps, une validation purement expérimentale est réalisée en comparant l'estimation des RNA avec les mesures issues du test Zurich-SFH100-16 m².

IV.2.2.1 SSC-STD

Les différents modèles neuronaux retenus pour l'analyse des résultats sont donnés dans le Tableau 4- 2.

Tableau 4- 2 : Caractéristiques des modèles neuronaux étudiés (système SSC-STD)

Modèles	NORM	NDN	TD	« R2nrj »	« R2bar »	« BIC2 »	« AIC2 »
RNA-R2nrj(tanh)	±1,0	7,0	3,0	0,9870	0,4301	16,5919	16,5166
RNA-R2bar(tanh)	±0,2	13,0	3,0	0,0001	0,9847	13,2114	13,1091
RNA-BIC2(tanh)	±1,0	9,0	3,0	0,9180	0,9843	13,0813	12,9961
RNA-AIC2(tanh)	±1,0	9,0	3,0	0,9180	0,9843	13,0813	12,9961
RNA-R2nrj(linéaire)	±1,0	10,0	2,0	0,9871	0,9682	13,7165	13,6405
RNA-R2bar(linéaire)	±1,0	13,0	3,0	0,9459	0,9836	13,2798	13,1775
RNA-BIC2(linéaire)	±1,0	14,0	2,0	0,7574	0,9820	13,2608	13,1710
RNA-AIC2(linéaire)	±1,0	14,0	2,0	0,7574	0,9820	13,2608	13,1710

Pour le meilleur modèle RNA-R2nrj(linéaire), sélectionné parmi tous les modèles selon le critère « *R2nrj* », les résultats sont très satisfaisants : le coefficient de détermination est de l'ordre de 0,987 (Figure 4- 12). L'erreur relative commise par ce RNA lors de l'estimation de la consommation annuelle du SSC-STD est quasiment, pour toutes les conditions, inférieure à 10% sauf, comme pour le système étudié dans le deuxième chapitre, pour Barcelone-SFH100 et Barcelone-SFH30. Cependant, les écarts absolus restent quand même faibles (inférieurs à 9 kWh/m²). Encore une fois, malgré le fait que le paramètre « surface du capteur » n'ait pas

été pris en compte en tant que variable à l'entrée du modèle, le RNA gère relativement bien l'influence de la surface du capteur sur la consommation du système pour certaines conditions. Il s'agit des conditions présentant des besoins énergétiques inférieurs à 80 kWh/m².

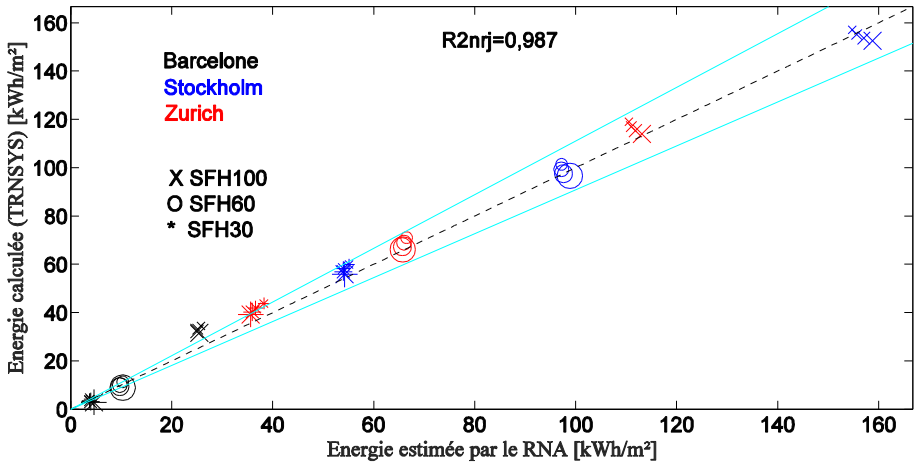

Figure 4- 12 : Comparaison entre la consommation annuelle estimée par le RNA-R2nrj(linéaire) et calculée par le modèle physique TRNSYS dans différentes conditions. Cas du système SSC-STD.

Le modèle RNA-R2bar(linéaire) est le modèle sélectionné par le critère « R2bar » parmi tous les RNA créés avec une FA linéaire. Avec un coefficient de détermination de l'ordre de 0,946 (Figure 4- 13), les performances du modèle RNA-R2bar(linéaire) sont naturellement moins bonnes que ceux du modèle précédent. Les erreurs relatives sont, pour un bon nombre de conditions (Zurich-SFH60, Zurich-SFH100 et Stockholm-SFH60), inférieures à 15%. Pour les autres conditions, l'erreur relative est plus élevée mais reste inférieure à 23% dans la plupart des cas. Un autre point caractéristique du modèle RNA-R2bar(linéaire) par rapport au modèle précédent, réside dans sa capacité à gérer correctement l'influence de la surface pour les conditions extrêmes.

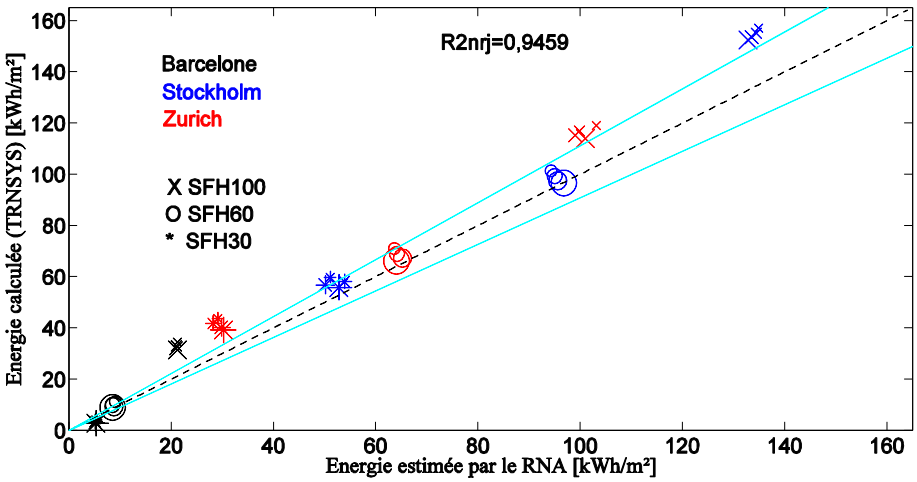

Figure 4- 13 : Comparaison entre la consommation annuelle estimée par le RNA-R2bar(linéaire) et calculée par le modèle physique TRNSYS dans différentes conditions. Cas du système SSC-STD. La taille des symboles indique la surface du capteur utilisée (10 m², 13 m², 16 m² ou 19m²)

Sur la Figure 4- 14 est représentée l'évolution de la puissance au cours de la séquence de test prédite par le RNA-R2bar(linéaire) et réellement mesurée dans les conditions d'apprentissage Zurich-SFH60-16 m². La très bonne concordance générale entre les deux courbes montre que l'apprentissage s'est bien déroulé.

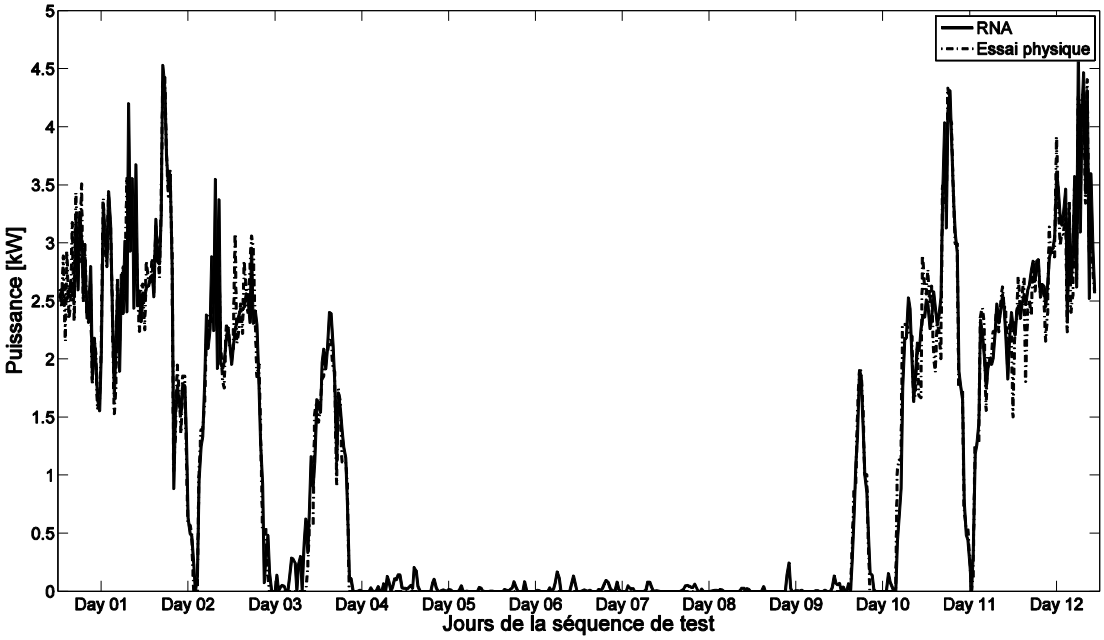

Figure 4- 14 : Evolution de la puissance de l'appoint du système SSC-STD au cours du test de 12 jours. Comparaison entre les estimations du RNA et les mesures. Cas des données d'apprentissage.

Pour valider le modèle, la comparaison doit être faite dans des conditions différentes de celles de l'apprentissage. Les deux modèles précédents ont donc été utilisés pour simuler le test réel du système dans l'environnement Zurich-SFH100-16m². Les erreurs relatives de l'estimation de la consommation du système lors du test sont inférieures à 8% pour les deux modèles (Figure 4- 15). Ceci valide expérimentalement la pertinence de la méthodologie pour la prédiction des performances du système dans des conditions différentes de celles de l'entraînement.

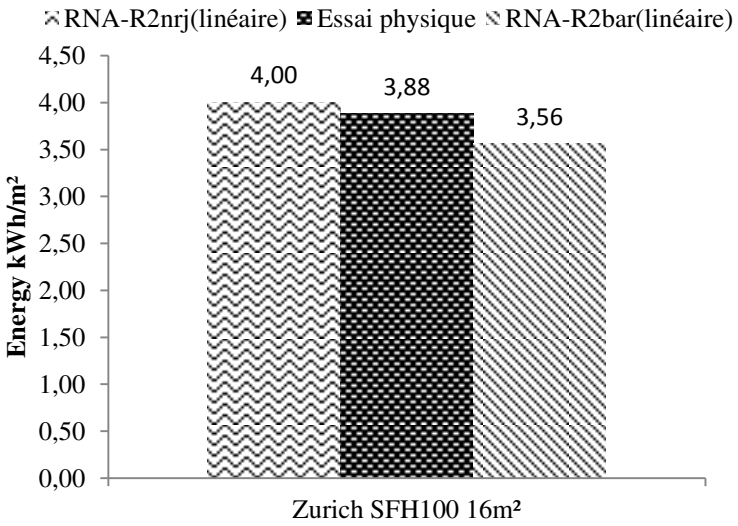

Figure 4- 15 : Comparaison entre la consommation pendant la séquence de test estimée par les RNA et mesurée durant le test physique. Cas du système SSC-STD.

Sur la Figure 4- 16 est représentée l'évolution réellement mesurée de la puissance au cours de la séquence de test dans les conditions de Zurich-SFH100-16 m² et celle prédite par le RNA sélectionné par le critère « *R2bar* ». Globalement le RNA suit correctement l'évolution de la puissance mesurée sauf pendant certains jours. Il est important de signaler que dans le cadre de la méthodologie développée, il n'est pas question de développer des modèles capables de prédire l'évolution précise et détaillée de la puissance du système.

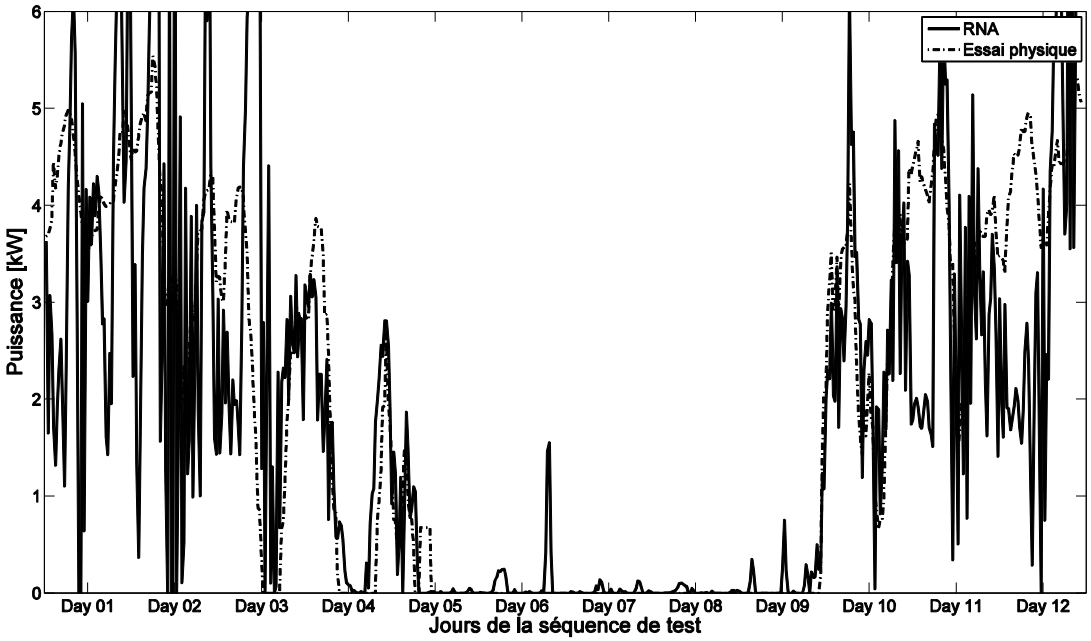

Figure 4- 16 : Evolution de la puissance du l'appoint du système SSC-STD au cours du test de 12 jours. Comparaison entre les estimations du RNA-R2bar(linéaire) et les mesures. Cas de Zurich-SFH100-16 m².

IV.2.2.2 SSC-OP

Les résultats de modélisation du SSC-OP à partir des données expérimentales, selon les deux modèles RNA (sélectionnés selon « R2nrj »), Tableau 4- 3, sont légèrement mieux que

ceux du système SSC-STD : le coefficient de détermination est supérieur à 0,990. Sur la Figure 4- 17 sont comparés les résultats de prédiction du modèle neuronal ayant une FA tanh avec les calculs de TRNSYS.

Tableau 4- 3 : Caractéristiques des modèles neuronaux étudiés, cas du système SSC-OP

Modèles	NORM	NDN	TD	« R2nrj »	« R2bar »	« BIC2 »	« AIC2 »
RNA-R2nrj(tanh)	±0,4	6,0	3,0	0,9910	0,9629	13,7627	13,6930
RNA-R2bar(tanh)	±0,4	13,0	3,0	0,7645	0,9874	12,9583	12,8560
RNA-BIC2(tanh)	±0,4	13,0	3,0	0,7645	0,9874	12,9583	12,8560
RNA-AIC2(tanh)	±0,4	13,0	3,0	0,7645	0,9874	12,9583	12,8560
RNA-R2nrj(linéaire)	±1,0	2,0	2,0	0,9903	0,8198	15,1690	15,1344
RNA-R2bar(linéaire)	±1,0	7,0	3,0	0,8590	0,9826	13,0431	12,9678
RNA-BIC2(linéaire)	±1,0	7,0	3,0	0,8590	0,9826	13,0431	12,9678
RNA-AIC2(linéaire)	±1,0	7,0	3,0	0,8590	0,9826	13,0431	12,9678

Pour la majorité des conditions ayant un besoin supérieur à 25 kWh/m², les erreurs relatives de prédictions sont inférieures à 5% avec une bonne prise en compte de l'impact de la surface du capteur. Pour les conditions présentant un besoin énergétique plus faible, les erreurs relatives sont plus importantes mais les valeurs absolues restent faibles notamment pour le climat de Barcelone.

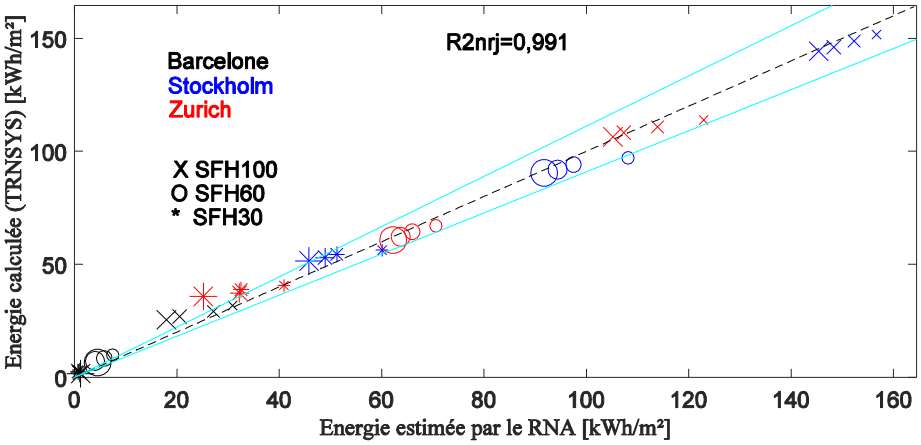

Figure 4- 17 : Comparaison entre la consommation annuelle estimée par le RNA-R2nrj(tanh) et calculée par le modèle physique TRNSYS dans différentes conditions. Cas du système SSC-OP. La taille des symboles indique la surface du capteur utilisée (10 m², 13 m², 16 m² ou 19m²)

Le modèle sélectionné selon le critère « R2bar » (RNA-R2bar(linéaire), Tableau 4- 3) présente des qualités très modestes malgré sa bonne prise en compte de l'influence de la surface du capteur (Figure 4- 18). En effet, les erreurs relatives sont supérieures à 20% pour la majorité des conditions.

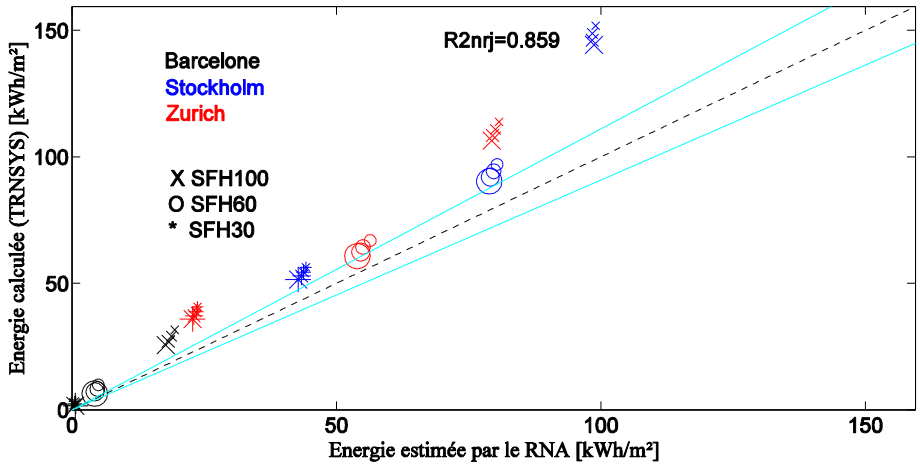

Figure 4- 18 : Comparaison entre la consommation annuelle estimée par le RNA-R2bar(linéaire)et calculée par le modèle physique TRNSYS dans différentes conditions. Cas du système SSC-OP. La taille des symboles indique la surface du capteur utilisée (10 m², 13 m², 16 m² ou 19m²)

Comme pour le système SSC-STD, deux des quatre modèles RNA, modélisant le comportement du système SSC-OP, ont été testés dans l'autre condition réelle.

La Figure 4- 19 révèle que le modèle sélectionné selon le « R2bar » arrive à estimer correctement la puissance du système dans le cas des conditions d'apprentissage. La concordance entre la puissance mesurée et prédite par ce RNA est, naturellement, légèrement moins bonne dans le cas des conditions de validation (Figure 4- 20).

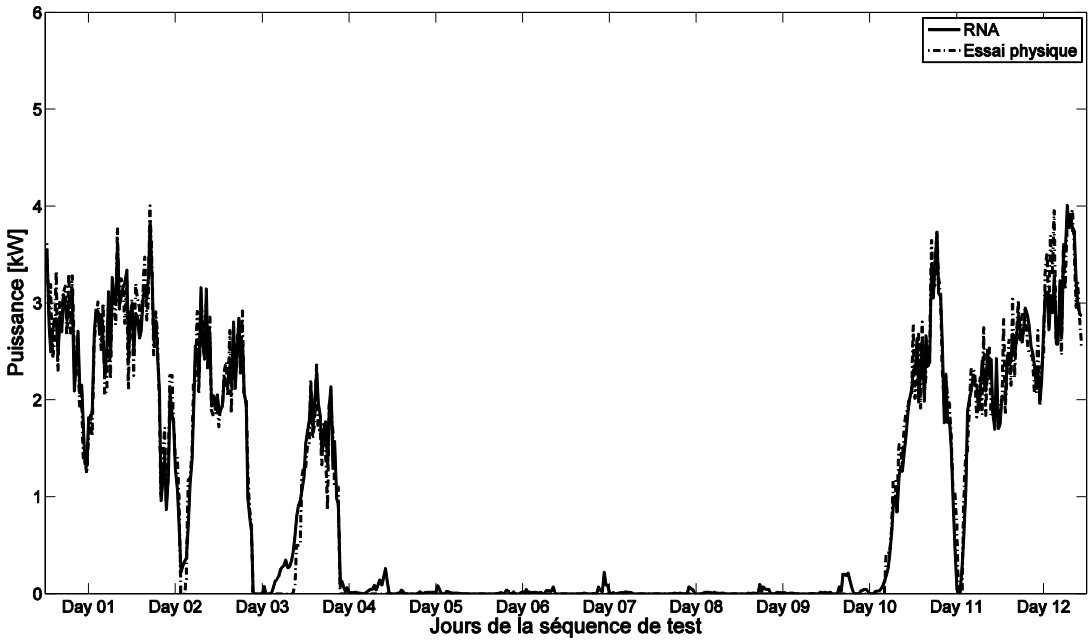

Figure 4- 19 : Evolution de la puissance de l'appoint du système SSC-OP au cours du test de 12 jours. Comparaison entre les estimations du RNA-R2bar(linéaire) et les mesures. Cas des données d'apprentissage.

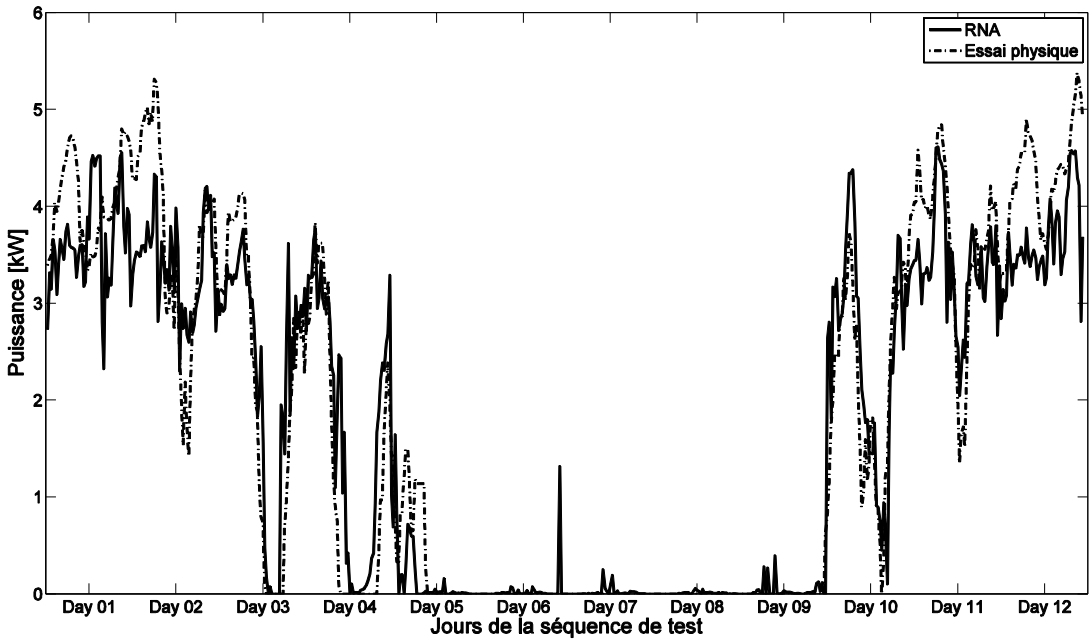

Figure 4- 20 : Evolution de la puissance de l'appoint du système SSC-OP au cours du test de 12 jours. Comparaison entre les estimations du RNA-R2bar(linéaire) et les mesures. Cas de Zurich-SFH100-16 m².

Cependant le calcul de l'énergie consommée par le système testé par les modèles RNA, la quantité recherchée, et la comparaison avec ce qui a été mesuré, montre que les résultats sont satisfaisants (Figure 4- 21). En effet, les erreurs relatives sont inférieures à 10%. Le fait que les RNA soient capables de prédire la consommation du système avec de tels pourcentages d'erreur malgré le fait que la loi de chauffe ait été légèrement modifiée d'un test à l'autre montre que la méthodologie est robuste et a un bon pouvoir d'adaptabilité. Cela reste vrai tant que le système est correctement dimensionné pour l'environnement considéré.

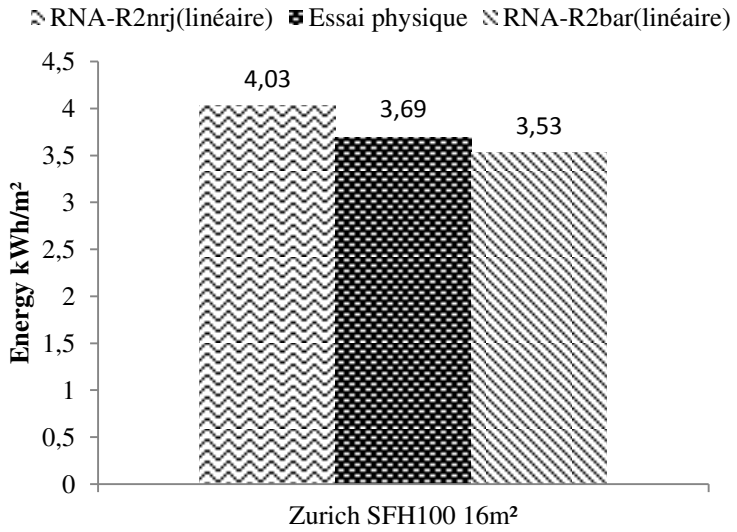

Figure 4- 21 : Comparaison entre la consommation pendant la séquence de test estimée par les RNA et mesurée durant le test physique. Cas du système SSC-OP.

IV.3. Test du système SSC-PROTO1

IV.3.1. Description du SSC-PROTO1

Le troisième système, utilisé pour la validation expérimentale de la méthodologie, est un prototype d'un SSC combiné à une chaudière à gaz qui a été développé par un fabriquant européen. Il sera noté dans la suite par SSC-PROTO1. La différence entre le système SSC-PROTO1 et les systèmes précédents réside dans ses caractéristiques physiques (Tableau 4-4) et dans son schéma hydraulique (Figure 4- 22).

Tableau 4- 4 : Caractéristiques physiques du système SSC-PROTO1

Appoint	Puissance nominale de 24 kW, Chaudière à gaz.
Ballon de stockage	Volume de 1000 litres Coefficient de perte globale égale à 5 W/K
Capteur solaire	$F'(\tau\alpha)_{en}$=0,8 $K_{\theta d}$=0,83 c_1=3,5 (W/m^{2}K); c_5=7000 (J.m^{-2} K^{-1}) c_2=0,015 (W/m^{2}K^2) β=45° b_0=0,2

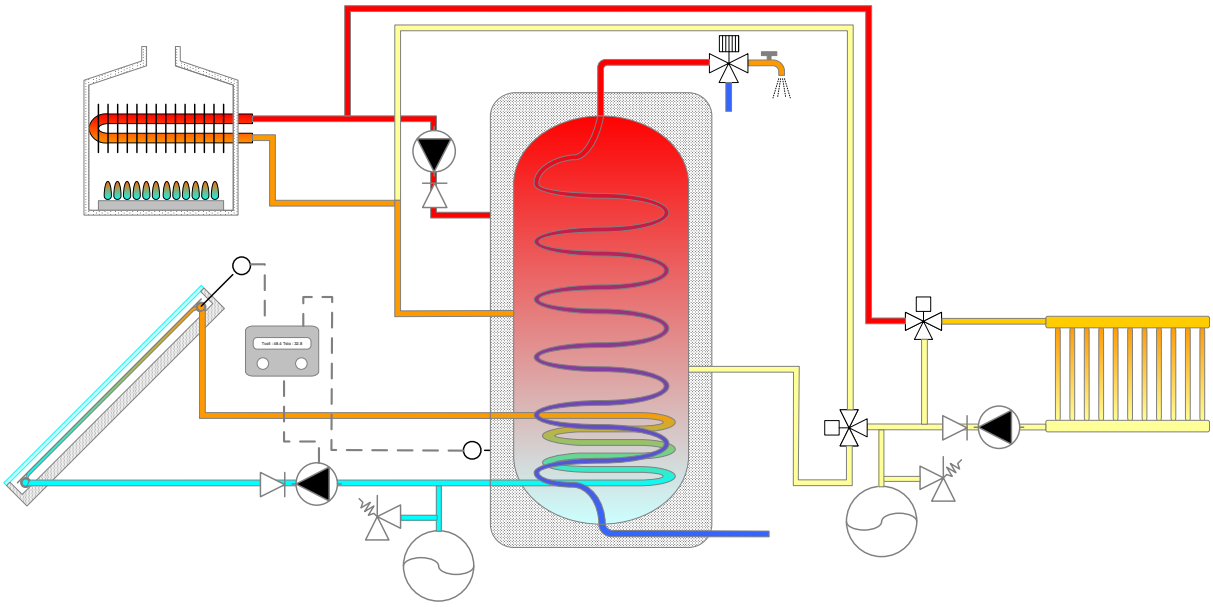

Figure 4- 22 : Schéma hydraulique du système SSC-PROTO1

Le fonctionnement du SSC-PROTO1 en termes de gestion des flux de chaleur est caractérisé par :

- Un transfert de chaleur pour la préparation de l'ECS mais aussi entre le capteur solaire et le ballon se font via un échangeur interne au ballon.

- La réponse aux besoins de chauffage peut être faite directement via la chaudière sans passage par le ballon. Dans ce cas le retour peut se faire à travers le ballon (si la température de retour est inférieure à la température du ballon, c'est le préchauffage) ou sans passage par ce dernier (si la température du ballon est inférieure à la température de retour). Le chauffage peut être également accompli à partir du ballon mais dans ce cas le passage par la chaudière (en arrêt) est obligatoire.

Similairement aux systèmes précédents et pour les mêmes raisons, la loi de chauffe du système est légèrement modifiée pour chaque test. Les deux tests ont été réalisés selon la méthode SCSPT dans les deux conditions suivantes : Zurich-SFH60-16 m² et Stockholm-SFH100-10 m².

IV.3.2. Résultats

Dans le but d'exploiter au maximum les données des deux tests du système SSC-PROTO1, la méthodologie a été validée en deux temps. En effet, chaque environnement servira de conditions de test mais aussi d'entrainement. Ce protocole de validation est représenté sur la Figure 2- 21 dans le deuxième chapitre.

Pour ce système une seule condition est disponible pour évaluer sa consommation estimée par les RNA. Nous allons en conséquence considérer le meilleur RNA celui qui présente la plus faible erreur relative absolue (ERA) au lieu de « R2nrj » utilisé précédemment. Ce modèle sera noté également RNA-ERA. La même notation sera utilisée également pour le système solaire intégrant une machine à absorption.

IV.3.2.1 Entrainement sur Zurich-SFH60-16m²

Les résultats présentés dans cette partie concernent les RNA entrainés sur l'environnement classique de Zurich-SFH60-16 m². La méthodologie a été appliquée de la même façon que pour tous les autres systèmes présentés précédemment. Sur le Tableau 4- 5 sont présentés l'ensemble des modèles sélectionnés.

Tableau 4- 5 : Caractéristiques des modèles neuronaux étudiés (SSC-PROTO1, entrainement sur Zurich)

Modèles	NORM	NDN	TD	ERA	« R2bar »	« BIC2 »	« AIC2 »
RNA-ERA(tanh)	±0,2	2,0	1,0	0,0101	0,6929	16,2478	16,2206
RNA-R2bar(tanh)	±0,2	8,0	3,0	1,0411	0,8745	15,6282	15,5478
RNA-BIC2(tanh)	±0,2	8,0	3,0	1,0411	0,8745	15,6282	15,5478
RNA-AIC2(tanh)	±0,2	8,0	3,0	1,0411	0,8745	15,6282	15,5478
RNA-ERA (linéaire)	±1,0	1,0	1,0	0,0033	0,6803	16,2880	16,2607
RNA-R2bar(linéaire)	±1,0	7,0	3,0	0,1983	0,8120	15,9952	15,9199
RNA-BIC2(linéaire)	±1,0	7,0	3,0	0,1983	0,8120	15,9952	15,9199
RNA-AIC2(linéaire)	±1,0	7,0	3,0	0,1983	0,8120	15,9952	15,9199

Les prédictions de la consommation du système SSC-PROTO1 par les différents modèles neuronaux dans l'environnement de Stockholm-SFH100-10 m² sont comparés aux mesures sur la Figure 4- 23. Le modèle RNA-ERA(tanh) estime très bien la consommation du système. L'erreur de prédiction est faible (de l'ordre de 1%). Par contre, le modèle sélectionné par les critères statistiques présente de faibles performances.

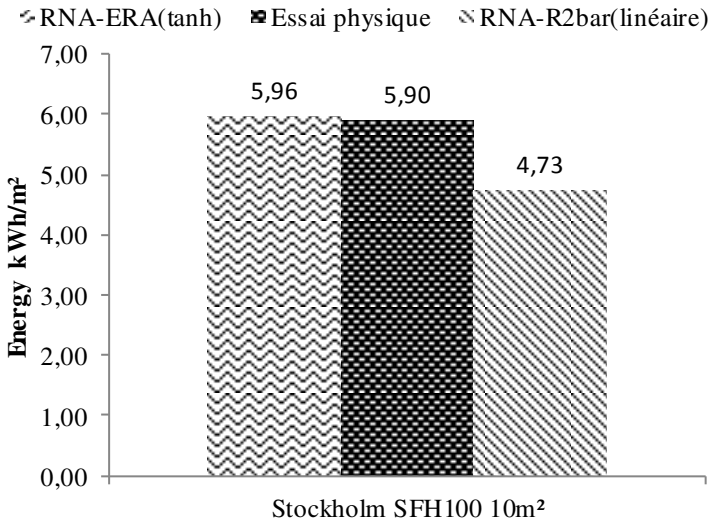

Figure 4- 23 : Comparaison entre la consommation pendant la séquence de test estimée par les RNA et mesurée durant le test physique. Cas du système SSC-PROTO1 et de l'entrainement sur Zurich.

Sur la Figure 4- 24 est présentée l'évolution de la puissance au cours de la séquence estimée par le RNA-ERA. La concordance globale avec la puissance mesurée est plutôt bonne à l'exception de la surestimation de la puissance pendant les 6ème et 7ème jours.

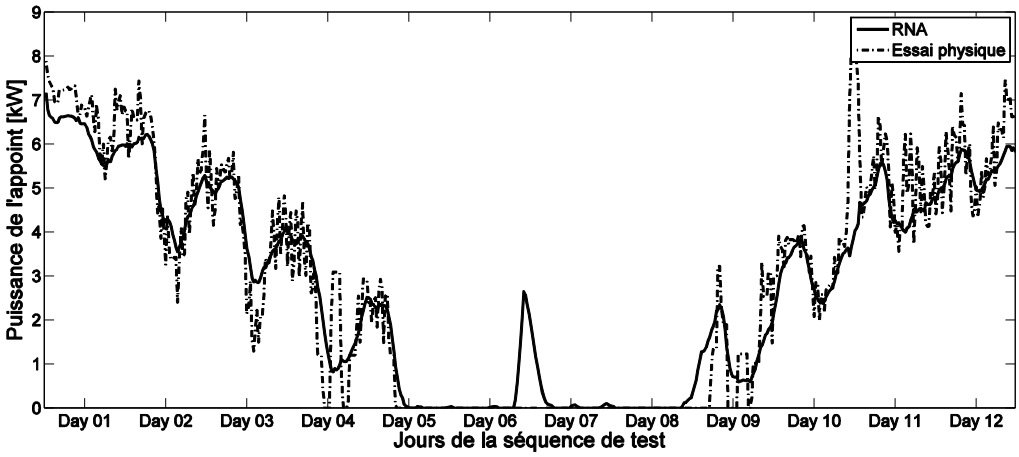

Figure 4- 24 : Evolution de la puissance de l'appoint du système SSC-PROTO1 au cours du test de 12 jours. Comparaison entre les estimations du RNA-ERA(tanh) et les mesures. Stockholm-SFH100-10 m².

IV.3.2.2 Entrainement sur Stockholm-SFH100-10m²

Dans cette partie les conditions précédemment utilisées pour l'apprentissage ont été utilisées ici pour tester les performances des modèles neuronaux.

Tableau 4- 6 : Caractéristiques des modèles neuronaux étudiés (SSC-PROTO1, entrainement sur Stockholm)

Modèles	NORM	NDN	TD	ERA	« R2bar »	« BIC2 »	« AIC2 »
RNA-ERA(tanh)	±1,0	2,0	1,0	0,0489	0,9330	15,6402	15,6130
RNA-R2bar(tanh)	±1,0	8,0	3,0	2,8552	0,9677	15,1748	15,0944
RNA-BIC2(tanh)	±1,0	8,0	3,0	2,8552	0,9677	15,1748	15,0944
RNA-AIC2(tanh)	±1,0	8,0	3,0	2,8552	0,9677	15,1748	15,0944
RNA-ERA (linéaire)	±1,0	2,0	2,0	0,0448	0,9341	15,6439	15,6092
RNA-R2bar(linéaire)	±1,0	8,0	3,0	0,4244	0,9641	15,2805	15,2001
RNA-BIC2(linéaire)	±1,0	8,0	3,0	0,4244	0,9641	15,2805	15,2001
RNA-AIC2(linéaire)	±1,0	8,0	3,0	0,4244	0,9641	15,2805	15,2001

Les estimations de la consommation du système par les RNA sont également très bons à Zurich-SFH60-16 m² (Figure 4- 25). Cependant, les erreurs relatives commises par les modèles neuronaux sont plus élevées que dans le cas de Stockholm-SFH100-10 m² (en particulier pour le modèle sélectionné par les critères statistiques). Ces résultats montrent et confirment la pertinence de réaliser l'entrainement sur les conditions Zurich-SFH60-16m² plutôt que sur des conditions un peu extrêmes.

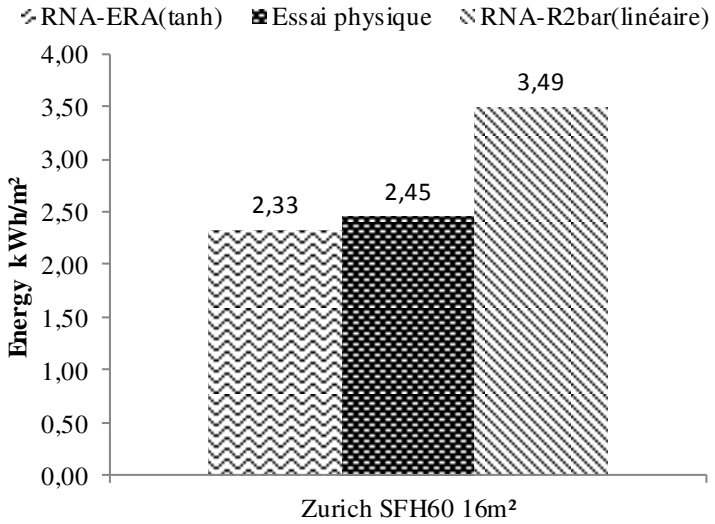

Figure 4- 25 : Comparaison entre la consommation pendant la séquence de test estimée par les RNA et mesurée durant le test physique. Cas du système SSC-PROTO1 et de l'entrainement sur Stockholm.

IV.4. Test du système SSC-PROTO2

IV.4.1. Description du SSC-PROTO2

Le quatrième système (noté SSC-PROTO2) utilisé pour évaluer expérimentalement la méthodologie est un système similaire au troisième. Les caractéristiques physiques de ce dernier sont données sur le Tableau 4- 7.

Le système SSC-PROTO2 a été testé (Figure 4- 26) dans trois conditions différentes selon la méthodologie SCSPT : Zurich-SFH100- 16 m², Barcelone-SFH60- 5 m² et Stockholm-SFH60-16 m².

Tableau 4- 7 : Caractéristiques physiques du système SSC-PROTO2

Appoint	Puissance nominale de 24 kW, Chaudière à gaz.
Ballon de stockage	750 litres de volume 1,56 m de hauteur Coefficient de perte globale égale à 5,58 W/K
Capteur solaire	$F'(\tau\alpha)_{en}$=0,8 $K_{\theta d}$=0,83 c_1=3,5 (W/m^{2}K); c_5=7000 (J.m^{-2} K^{-1}) c_2=0,015 (W/m^{2}K^2) β=45° b_0=0,2

Figure 4- 26 : Photo du système SSC-PROTO2 installé dans le bac d'essai semi-virtuel

IV.4.2. Résultats

Les conditions des trois tests réalisés sont différentes de celles utilisées dans le cas du système SSC-PROTO1. La diversification des conditions des tests permettra de voir l'influence des différentes conditions sur l'apprentissage des modèles.

IV.4.2.1 Entrainement sur Zurich-SFH100-16m²

Les résultats des prédictions des RNA (Tableau 4- 8) sont issus de l'application de la méthodologie avec comme conditions d'apprentissage Zurich-SFH100-16 m². Le test des modèles est ensuite réalisé dans les autres conditions (Barcelone et Stockholm dans ce sous-paragraphe). Même si le bâtiment SFH100 utilisé pour l'apprentissage est un bâtiment plus énergivore que le bâtiment SFH60, les erreurs relatives commises sont faibles pour tous les RNA. En effet, pour Stockholm-SFH60-16 m², les erreurs relatives sont inférieures à 1,1% (cette valeur est atteinte avec le modèle RNA-R2bar(tanh)). Pour Barcelone-SFH60-5 m², les

erreurs sont plus conséquentes mais, encore une fois, les écarts absolus sont très faibles (inférieurs à 0,17 kWh/m²).

Tableau 4- 8 : Caractéristiques des modèles neuronaux étudiés (SSC-PROTO2, entrainement sur Zurich)

Modèles	NORM	NDN	TD	« R2nrj »	« R2bar »	« BIC2 »	« AIC2 »
RNA-R2nrj(tanh)	±0,4	1,0	2,0	0,9999	0,9657	14,3898	14,3647
RNA-R2bar(tanh)	±1,0	8,0	3,0	0,9970	0,9950	12,7458	12,6654
RNA-BIC2(tanh)	±1,0	8,0	3,0	0,9970	0,9950	12,7458	12,6654
RNA-AIC2(tanh)	±1,0	8,0	3,0	0,9970	0,9950	12,7458	12,6654
RNA-R2nrj(linéaire)	±1,0	1,0	2,0	0,9995	0,9662	14,3728	14,3477
RNA-R2bar(linéaire)	±1,0	8,0	3,0	0,9945	0,9945	12,8530	12,7726
RNA-BIC2(linéaire)	±1,0	7,0	3,0	0,9966	0,9943	12,8475	12,7722
RNA-AIC2(linéaire)	±1,0	7,0	3,0	0,9966	0,9943	12,8475	12,7722

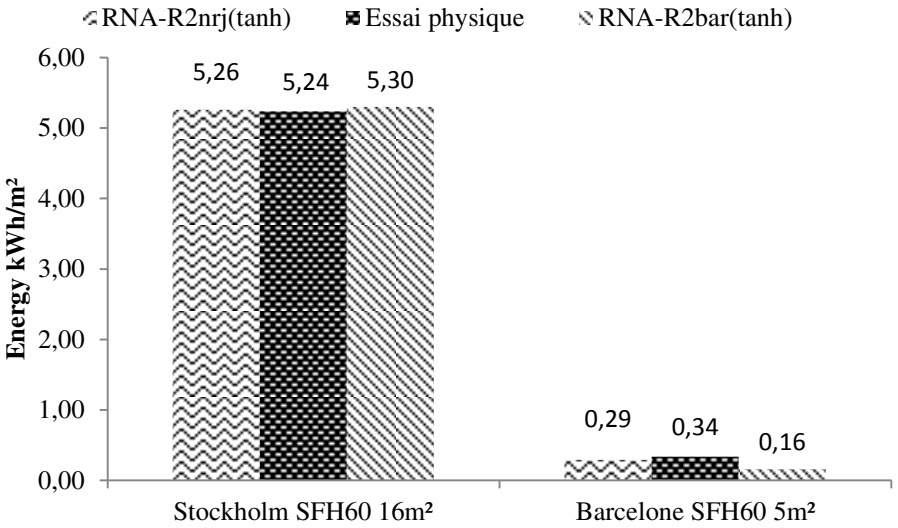

Figure 4- 27 : Comparaison entre la consommation pendant la séquence de test estimée par les RNA et mesurée durant le test physique. Cas du système SSC-PROTO2 et de l'entrainement sur Zurich.

Sur la Figure 4- 28 et Figure 4- 29 est représentée l'évolution de la puissance de l'appoint estimée par le modèle RNA-R2bar(tanh) dans les deux conditions de test. Pour le cas de Stockholm-SFH60-16 m², mise à part une légère sous-estimation de la puissance lors des deux premiers jours de la séquence, la concordance entre les mesures et les prédictions du RNA est assez bonne. Pour le cas de Barcelone-SFH60-5 m², l'évolution globale de la puissance est acceptable (cela correspond au fait que le RNA arrive à détecter la nécessité de mise en marche de la chaudière), mais les écarts en terme d'amplitude sont plus importants que dans l'autre condition de test.

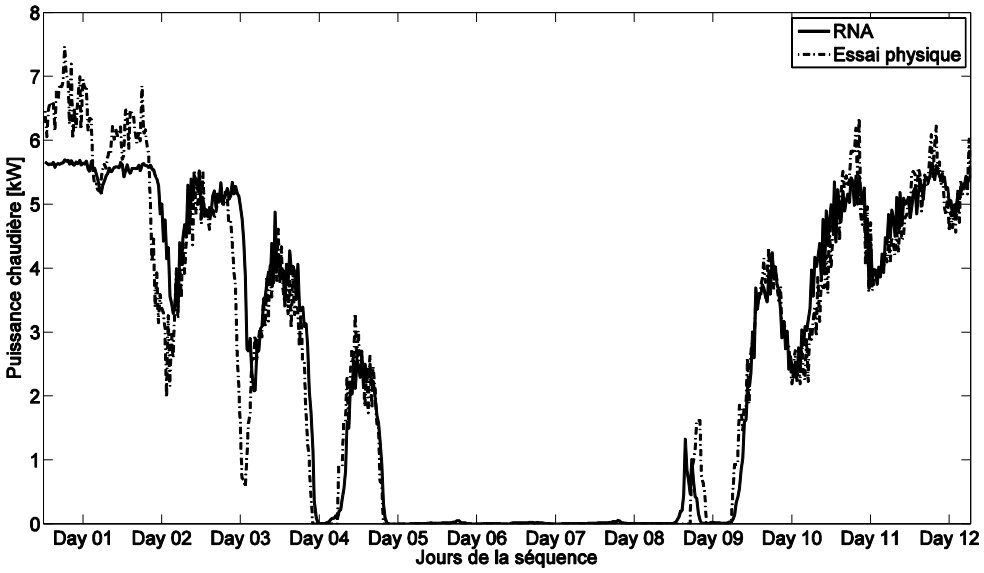

Figure 4- 28 : Evolution de la puissance de l'appoint du système SSC-PROTO2 au cours du test de 12 jours. Comparaison entre les estimations du RNA-R2nrj(tanh) et les mesures. Stockholm-SFH60-16m².

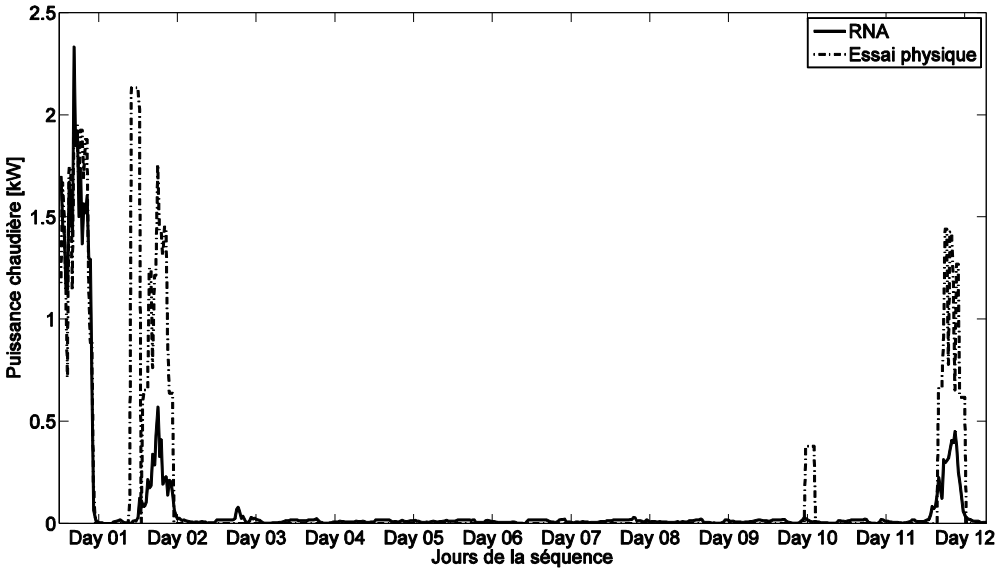

Figure 4- 29 : Evolution de la puissance de l'appoint du système SSC-PROTO2 au cours du test de 12 jours. Comparaison entre les estimations du RNA-R2nrj (tanh) et les mesures. Barcelone-SFH60-5m².

IV.4.2.2 Entrainement sur Stockholm-SFH60-16m²

Une étude similaire à la précédente mais avec, cette fois-ci, l'environnement de Stockholm-SFH60-16 m² comme condition d'apprentissage est présentée dans ce paragraphe.

Tableau 4- 9 : Caractéristiques des modèles neuronaux étudiés (SSC-PROTO2, entrainement sur Stockholm)

Modèles	NORM	NDN	TD	« R2nrj »	« R2bar »	« BIC2 »	« AIC2 »
RNA-R2nrj(tanh)	±0,2	2,0	1,0	1,0000	0,9773	14,4253	14,3980
RNA-R2bar(tanh)	±1,0	8,0	3,0	0,9850	0,9934	13,4716	13,3912
RNA-BIC2(tanh)	±1,0	7,0	3,0	0,9972	0,9933	13,4511	13,3758
RNA-AIC2(tanh)	±1,0	7,0	3,0	0,9972	0,9933	13,4511	13,3758
RNA-R2nrj(linéaire)	±1,0	7,0	1,0	1,0000	0,9854	14,0726	14,0230
RNA-R2bar(linéaire)	±1,0	6,0	3,0	0,9990	0,9921	13,5767	13,5070
RNA-BIC2(linéaire)	±1,0	6,0	3,0	0,9990	0,9921	13,5767	13,5070
RNA-AIC2(linéaire)	±1,0	6,0	3,0	0,9990	0,9921	13,5767	13,5070

Les estimations des RNA (Tableau 4- 9) sont représentées sur la Figure 4- 30. Les erreurs relatives sont très faibles pour tous les modèles (inférieure à 5%). Comme pour les résultats précédents, l'erreur est plus importante (25% au maximum) pour le climat de Barcelone mais reste toujours faible en termes d'écart absolu.

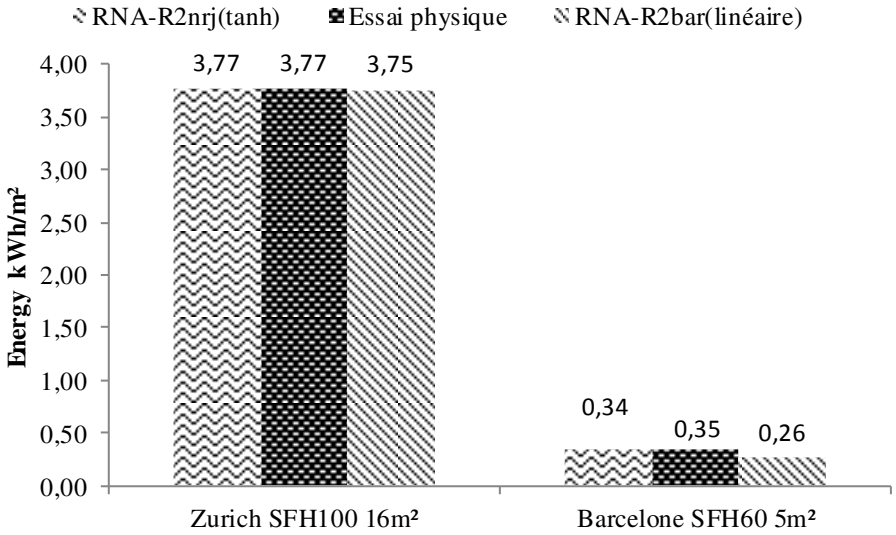

Figure 4- 30 : Comparaison entre la consommation pendant la séquence de test estimée par les RNA et mesurée durant le test physique. Cas du système SSC-PROTO2 et de l'entrainement sur Stockholm.

IV.4.2.3 Entrainement sur Barcelone-SFH60-5m²

La troisième étude concerne l'entrainement sur une condition extrême qu'est Barcelone-SFH60-5 m². Les résultats des modèles neuronaux (Tableau 4- 10) sont représentés sur la Figure 4- 31.

Tableau 4- 10 : Caractéristiques des modèles neuronaux étudiés, (SSC-PROTO2, entrainement sur Barcelone)

Modèles	NORM	NDN	TD	« R2nrj »	« R2bar »	« BIC2 »	« AIC2 »
RNA-R2nrj(tanh)	±0,4	5,0	1,0	0,9854	0,6625	14,0209	13,9788
RNA-R2bar(tanh)	±1,0	4,0	3,0	0,0001	0,9626	11,8504	11,7933
RNA-BIC2(tanh)	±1,0	4,0	3,0	0,0001	0,9626	11,8504	11,7933
RNA-AIC2(tanh)	±1,0	4,0	3,0	0,0001	0,9626	11,8504	11,7933
RNA-R2nrj(linéaire)	±1,0	4,0	2,0	0,8661	0,0697	18,8016	18,7532
RNA-R2bar(linéaire)	±1,0	5,0	1,0	0,4962	0,9195	12,5887	12,5466
RNA-BIC2(linéaire)	±1,0	5,0	1,0	0,4962	0,9195	12,5887	12,5466
RNA-AIC2(linéaire)	±1,0	5,0	1,0	0,4962	0,9195	12,5887	12,5466

Pour le meilleur modèle ayant une FA tanh dans la couche de sortie (RNA-R2nrj(tanh)) les erreurs relatives pour les deux conditions de tests sont très faibles. Le deuxième meilleur modèle mais ayant une FA linéaire dans la couche de sortie (RNA-R2nrj(linéaire)) présente une erreur maximale de l'ordre de 9% dans le cas de l'environnement de test Zurich-SFH100-16 m².

En ce qui concerne les modèles sélectionnés selon les critères statistiques, les résultats présentent des erreurs relatives assez élevées. Une valeur de plus de 60% a été enregistrée pour Stockholm-SFH60-16 m² (RNA avec une FA linéaire).

Les résultats de l'entrainement sur le climat de Barcelone sont globalement moins bons (en particulier pour les modèles sélectionnés par les critères) qu'un entrainement sur l'environnement de Stockholm ou Zurich. Nous observons que l'identification avec un climat de type « moyen » (Zurich dans notre étude) ne donne pas systématiquement les meilleurs résultats. Par conséquent, il serait intéressant de développer en perspectives une nouvelle séquence d'apprentissage qui puisse s'adapter à tout type de système.

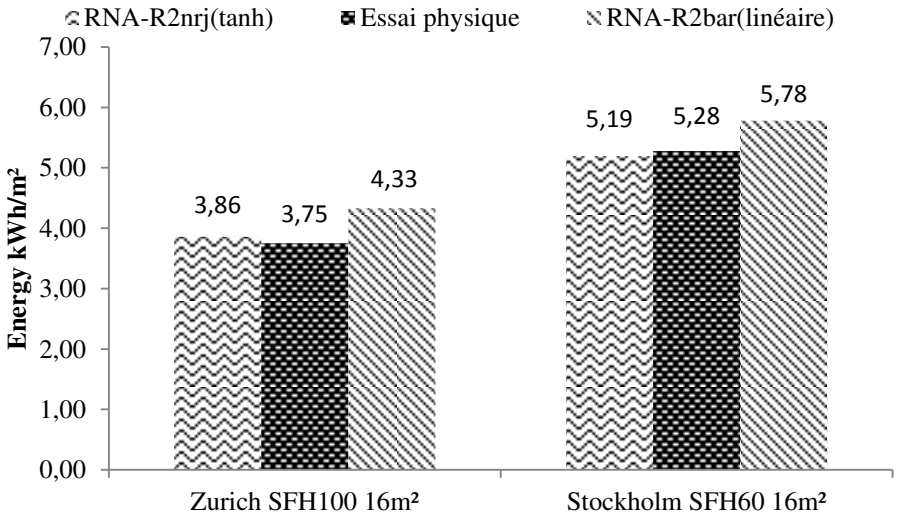

Figure 4- 31 : Comparaison entre la consommation pendant la séquence de test estimée par les RNA et mesurée durant le test physique. Cas du système SSC-PROTO2 et de l'entrainement sur Barcelone.

IV.5. Exploitation des données de suivi in situ d'un SSC intégrant une machine à absorption

IV.5.1. Description du système

Il a été signalé dans le deuxième chapitre, qu'une validation expérimentale rigoureuse nécessiterait la comparaison entre les estimations des modèles neuronaux et des mesures de suivi *in situ* de différentes installations et dans différentes conditions. Malheureusement compte-tenu des coûts élevés que cela génèrerait, cela s'avère impossible. Bien qu'une telle comparaison n'ait pas été prévue dans le cadre de la Thèse, l'exploitation de données de suivi *in situ* d'une installation solaire intégrant une machine à absorption (Figure 4- 32) disponible au sein des locaux de l'INES a été réalisée.

L'installation solaire a été réalisée dans le cadre d'un projet européen dans le but d'avoir un retour terrain sur les performances d'un tel système innovant. Le système global est composé de deux ballons de stockage : un chaud pour l'énergie solaire (avec un appoint électrique intégré) et un tampon pour le froid et le chaud en fonction de la période, d'une machine à absorption performante, d'un capteur solaire plan vitré et d'un système de rejet de chaleur enterré (connecté au niveau du condenseur de la machine à absorption). Le système solaire permet de répondre aux besoins de chauffage et de climatisation de trois bureaux d'une surface totale de 63 m². Quelques caractéristiques physiques du système sont données dans le Tableau 4- 11.

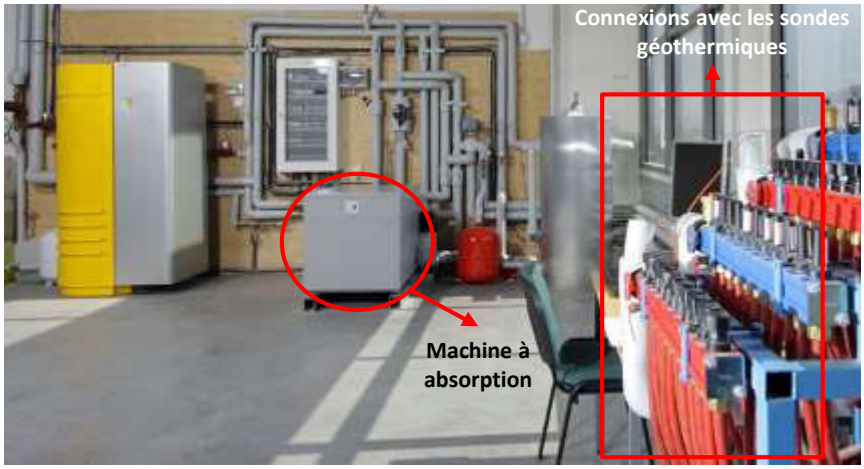

Figure 4- 32 : Le SSC combiné à une machine à absorption installé dans les locaux de l'INES

Tableau 4- 11 : Caractéristiques physiques du système SSC combiné à une machine à absorption

Appoint électrique	Puissance nominale de 12 kW
Ballons de stockage	400 litres de volume pour le ballon de stockage solaire et 300 litres de volume pour le ballon de stockage tampon
Capteur solaire	$F'(\tau\alpha)_{en}$=0,8;$K_{\theta d}$=0,83 c_1=3,5 (W/m²K); c_5=7000 (J.m^{-2} K^{-1}) c_2=0,015 (W/m²K²) ; β=45° ;b_0=0,2 Surface du capteur = 30 m²
Machine à absorption	Puissance nominale de 4,5 kW
Sondes géothermiques	2200 m de ligne horizontal enterré à une profondeur entre 0,75 et 1,1 m

IV.5.2. Modes de fonctionnement

Le système fonctionne selon deux modes distincts. Le premier est le mode chauffage. Dans ce cas le système fonctionne comme un système solaire classique. L'appoint se met en marche dès que l'énergie solaire n'est pas suffisante pour couvrir les besoins de chauffage. Le deuxième mode est le mode rafraichissement (le basculement à ce mode se fait manuellement via les vannes de séparation, Figure 4- 33 et Figure 4- 34), dans ce cas la machine à absorption se met en marche dès que la température des bureaux est supérieure à une température de consigne. Là aussi, si la température du stockage solaire (générateur de la machine) est inférieure à un seuil préfixé, l'appoint électrique se met en marche. La distribution vers les bureaux se fait par des ventilo-convecteurs à partir du ballon tampon (chaud ou froid en fonction du mode).

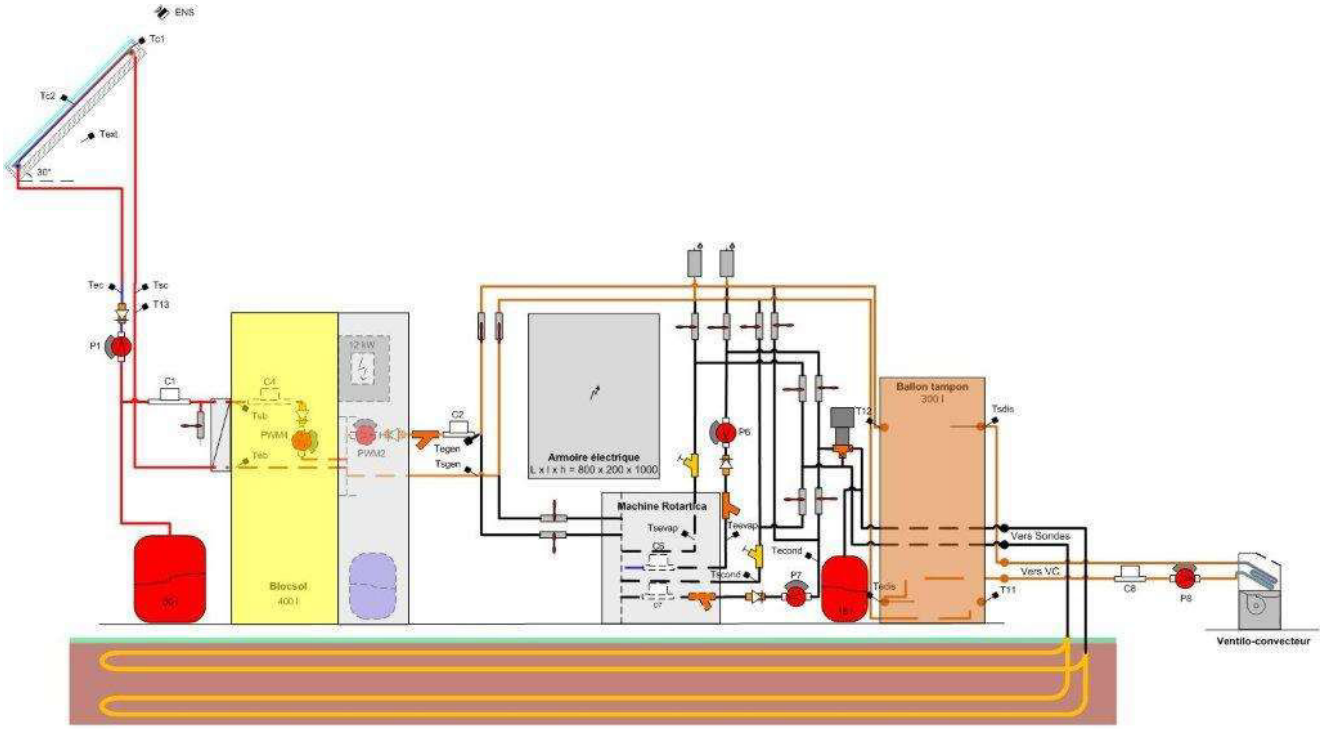

Figure 4- 33 : Schéma hydraulique du SSC combiné à une machine à absorption, mode chauffage

Les deux modes de fonctionnement étant séparés, il est donc judicieux de les modéliser avec des modèles distincts, entrainés sur des données issues des deux modes (période de chauffage et période de climatisation). Les données de suivi *in situ* de l'installation correspondant à la période du 16/11/2011 au 08/01/2012 ont été utilisées pour la saison de chauffage tandis que celles de la période du 02/08/2011 au 21/09/2011 ont été utilisées pour la saison de rafraichissement.

Suite à une contrainte liée aux difficultés de démontage du système en vue de son installation sur le banc d'essai semi-virtuel et à la disponibilité de ce dernier (mutualisé entre divers projets de recherche), il n'a pas été possible de tester le système selon la méthode SCSPT. Pour pallier cette difficulté, il a été décidé de sélectionner les douze premiers jours des données de suivi *in situ* pour l'apprentissage des modèles et garder le reste de la base de données pour leur validation.

La méthodologie d'évaluation des performances n'est pas appliquée rigoureusement comme prévue. Cette validation permet néanmoins de tester la qualité des modèles neuronaux et de leur apprentissage dans le cas d'une base de données (d'entrainement) arbitraire composée de mesures réelles.

Sur les Figure 4- 35, Figure 4- 36 et Figure 4- 37 sont présentées les données utilisées pour l'entrainement. L'ECS ne faisant pas partie des besoins auxquels le système doit répondre, les variables utilisées en entrée des modèles sont la demande de chauffage ou de climatisation (en fonction du mode de fonctionnement du système) et la ressource solaire. En sortie, les modèles devraient prédire la consommation du système en électricité (machine et appoint électrique).

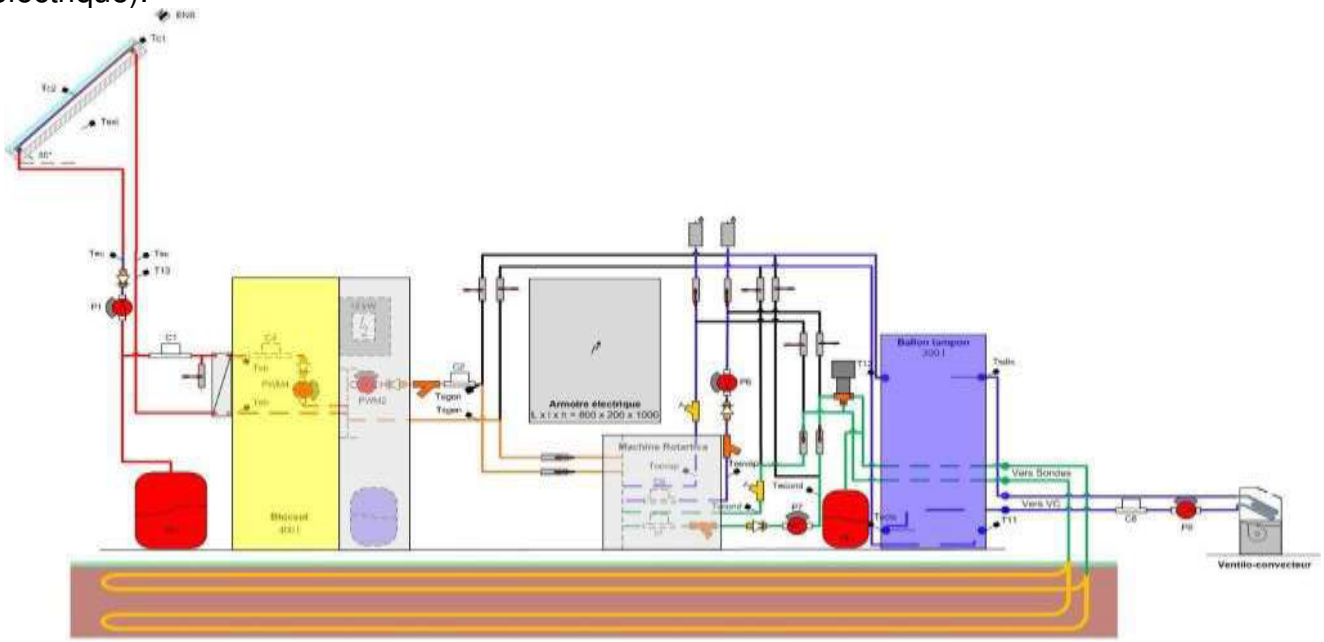

Figure 4- 34 : Schéma hydraulique du SSC combiné à une machine à absorption, mode climatisation

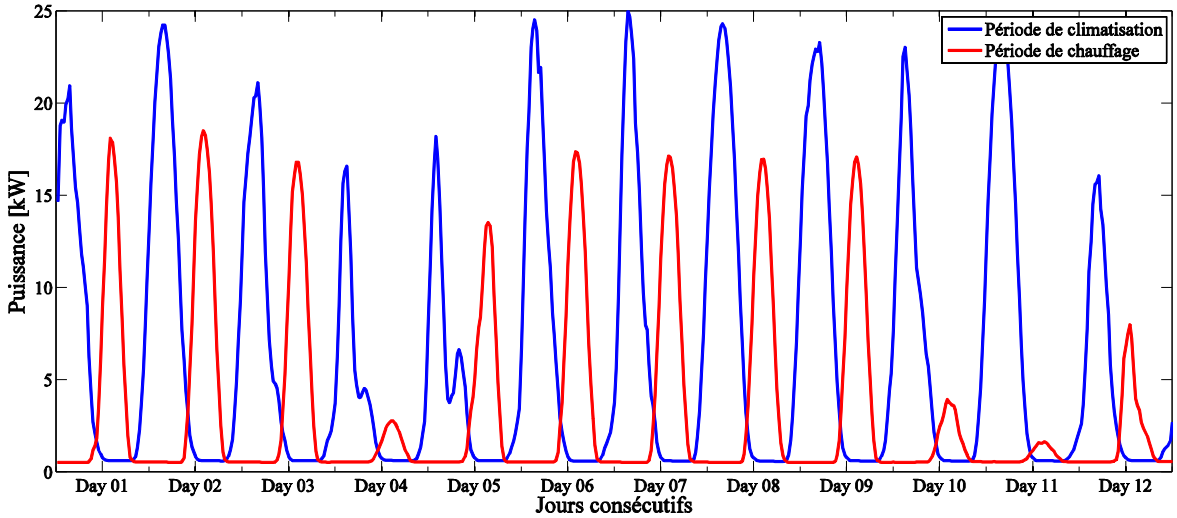

Figure 4- 35 : Rayonnements solaire des deux séquences d'apprentissage pour la modélisation du système combiné à une machine à absorption. Données de suivi in situ.

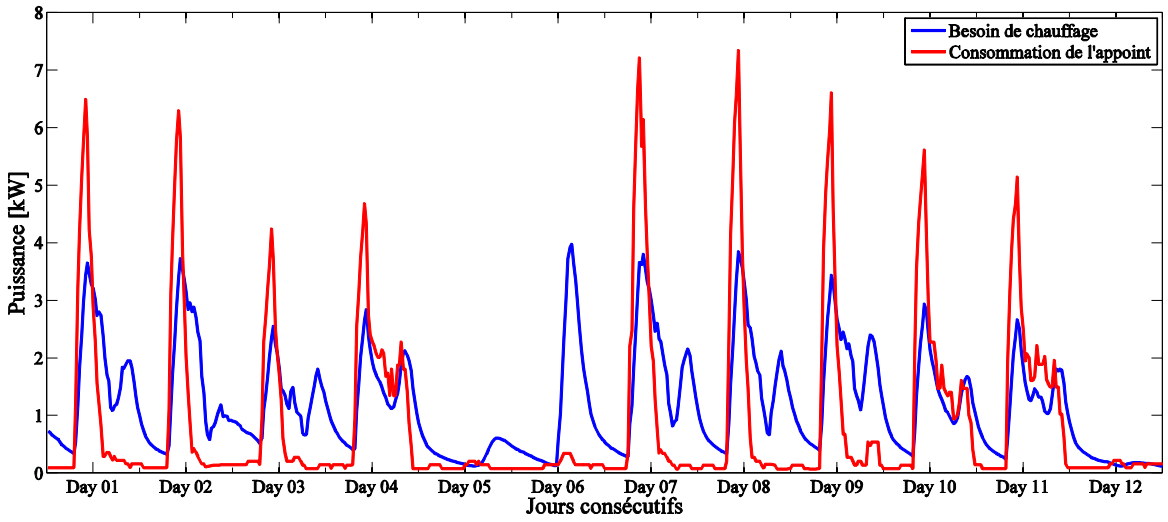

Figure 4- 36 : Besoins de chauffage et consommation du système combiné à une machine à absorption utilisés pour l'apprentissage. Données de suivi in situ. Mode chauffage

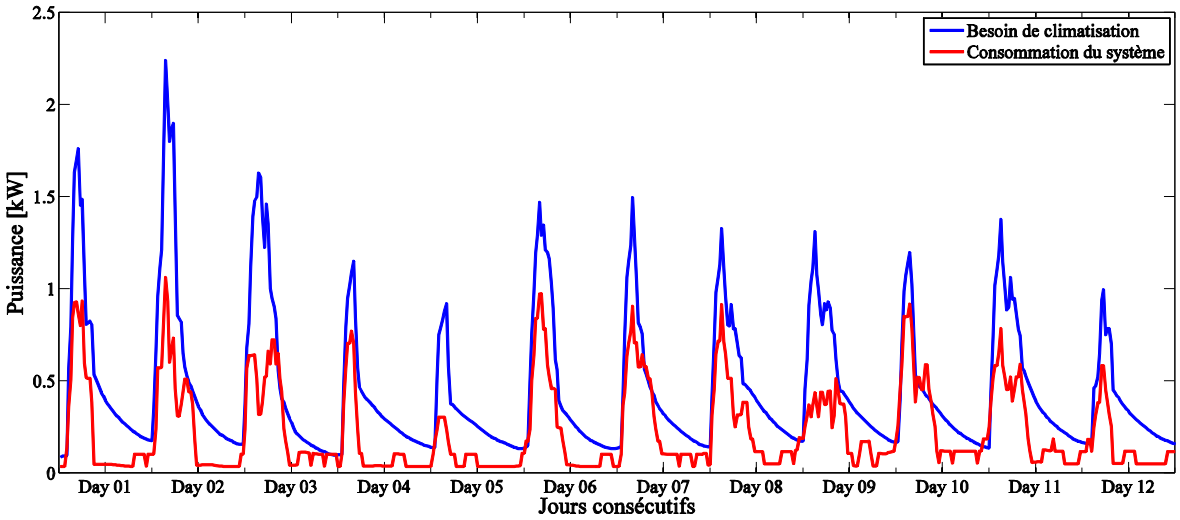

Figure 4- 37 : Besoins de climatisation et consommation du système combiné à une machine à absorption utilisés pour l'apprentissage. Données de suivi in situ. Mode climatisation

IV.5.3. Résultats

La méthodologie de modélisation a été appliquée de la même façon que pour les autres systèmes, même si les données d'apprentissage ne sont pas issues de la méthode SCSPT. Une différence à noter par rapport aux différentes modélisations des systèmes précédents concerne le paramètre du retard utilisé. Il a été constaté durant le processus de modélisation qu'avec les paramètres initiaux du Tableau 2- 11, les erreurs relatives de prédiction sont assez élevées même sur les données d'apprentissage. Nous avons dû augmenter le nombre de retards pour faciliter l'apprentissage du fonctionnement du système. Le nombre de retards maximal a donc été remplacé par 6.

IV.5.3.1 Fonctionnement en mode chauffage

Les modèles sélectionnés pour présenter les estimations de la consommation du système pour la période de chauffe allant du 28/11/2011 au 08/01/2012 sont donnés sur le Tableau 4-12. La définition du modèle RNA-ERA a été introduite dans le paragraphe IV.3.2.

Tableau 4- 12 : Caractéristiques des modèles neuronaux étudiés, cas du système combiné à une machine à absorption, mode chauffage

Modèles	NORM	NDN	TD	ERA	« R2bar »	« BIC2 »	« AIC2 »
RNA-ERA(tanh)	±0,4	8,0	5,0	0,0188	0,9928	-3,7048	-3,7933
RNA-R2bar(tanh)	±0,4	13,0	6,0	0,4338	0,9964	-4,0201	-4,1423
RNA-BIC2(tanh)	±0,4	13,0	3,0	0,0766	0,9959	-4,1396	-4,2419
RNA-AIC2(tanh)	±0,4	13,0	3,0	0,0766	0,9959	-4,1396	-4,2419
RNA-ERA (linéaire)	±1,0	4,0	4,0	0,0872	0,9884	-3,4133	-3,4761
RNA-R2bar(linéaire)	±1,0	12,0	6,0	1,3847	0,9967	-4,1792	-4,2966
RNA-BIC2(linéaire)	±1,0	12,0	3,0	0,3489	0,9961	-4,2140	-4,3123
RNA-AIC2(linéaire)	±1,0	12,0	3,0	0,3489	0,9961	-4,2140	-4,3123

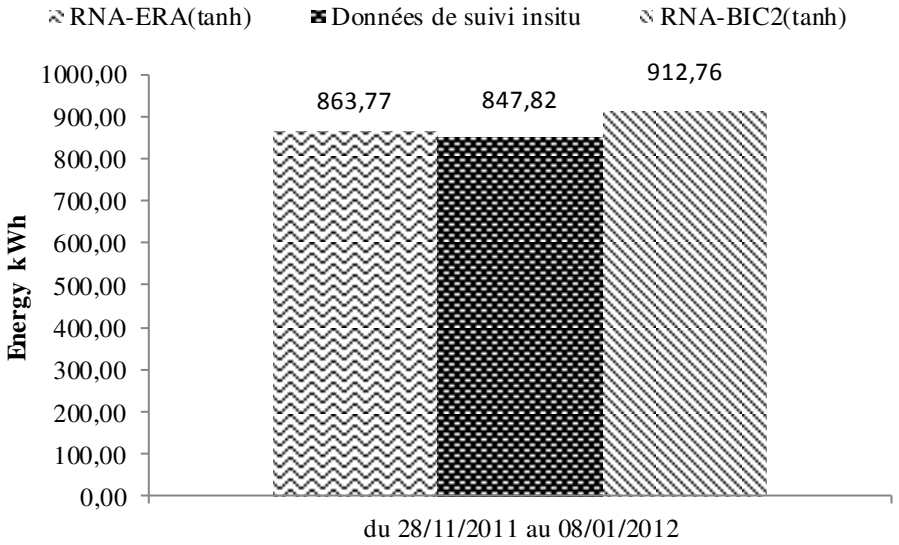

Figure 4- 38 : Comparaison entre la consommation mesurée et celle estimée par les RNA. Cas du système combiné à une machine à absorption en mode chauffage.

Les résultats de prédiction sont très satisfaisants (Figure 4- 38), même le RNA sélectionné selon le critère « BIC2 » avec une FA tanh donne de bons résultats. L'erreur relative de prédiction est inférieure à 8%.

La Figure 4- 39 représente l'évolution de la puissance du système estimée par le modèle RNA-ERA(tanh) pour la période du 28/11/2011 au 08/01/2012. L'évolution globale de la puissance est bien respectée par le modèle neuronal.

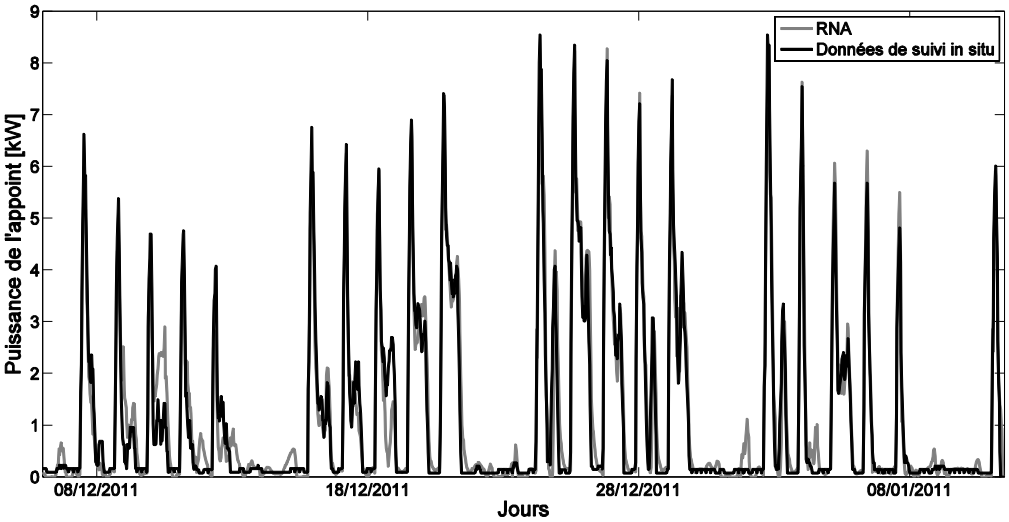

Figure 4- 39 : Evolution de la puissance de l'appoint du système combiné à une machine à absorption. Comparaison entre les estimations du RNA-ERA(tanh) et les mesures. Mode chauffage

IV.5.3.2 Fonctionnement en mode rafraichissement

Le meilleur modèle sélectionné (Tableau 4- 13) parmi tous les modèles créés donnent de très bons résultats (erreur relatives inférieures à 1%) (Figure 4- 40). De même, le modèle sélectionné par les 3 critères, pour la FA tanh, prédit très bien la consommation du système avec une erreur très faible. En effet, elle est seulement de 3,5%.

Tableau 4- 13 : Caractéristiques des modèles neuronaux étudiés, cas du système combiné à une machine à absorption, mode climatisation La définition du modèle RNA-ERA a été introduite dans le paragraphe IV.3.2.

Modèles	NORM	NDN	TD	ERA	« R2bar »	« BIC2 »	« AIC2 »
RNA-ERA(tanh)	±1,0	2,0	3,0	0,0019	0,9734	-2,5672	-2,6340
RNA-R2bar(tanh)	±0,4	9,0	5,0	0,0359	0,9961	-4,2718	-4,3657
RNA-BIC2(tanh)	±0,4	9,0	5,0	0,0359	0,9961	-4,2718	-4,3657
RNA-AIC2(tanh)	±0,4	9,0	5,0	0,0359	0,9961	-4,2718	-4,3657
RNA-ERA (linéaire)	±1,0	2,0	4,0	0,0024	0,9906	-3,5330	-3,6097
RNA-R2bar(linéaire)	±1,0	8,0	5,0	0,4005	0,9959	-4,2560	-4,3445
RNA-BIC2(linéaire)	±1,0	8,0	5,0	0,4005	0,9959	-4,2560	-4,3445
RNA-AIC2(linéaire)	±1,0	8,0	5,0	0,4005	0,9959	-4,2560	-4,3445

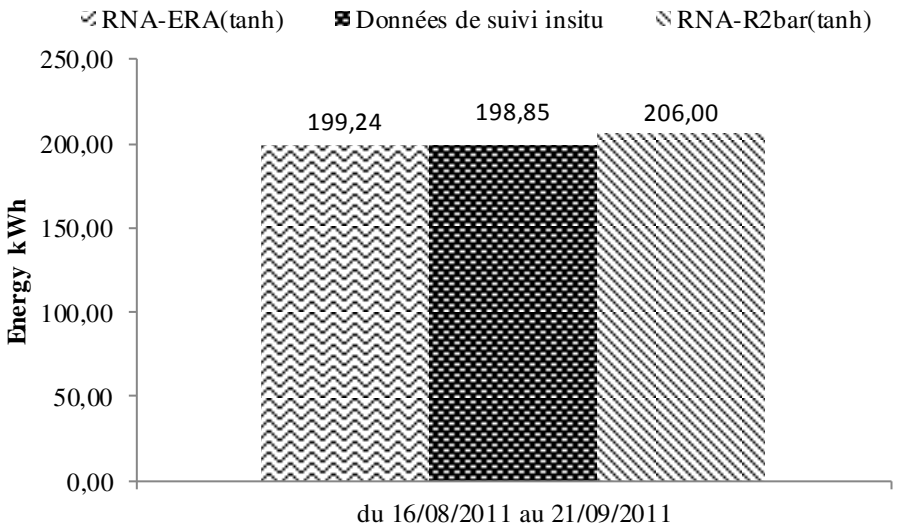

Figure 4- 40 : Comparaison entre la consommation estimée par les RNA et mesurée. Cas du système combiné à une machine à absorption en mode climatisation.

Un exemple de l'évolution de la consommation du système en fonction du temps pendant la période du 16/08/2011 au 21/09/2011 prédite par le RNA-ERA(tanh) est donnée sur la Figure 4- 41.

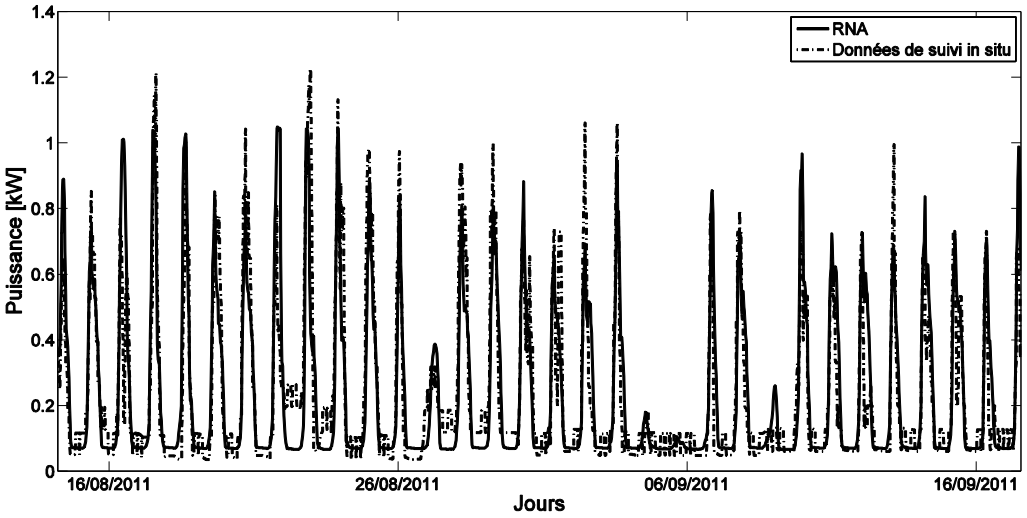

Figure 4- 41 : Evolution de la puissance de l'appoint du système combiné à une machine à absorption. Comparaison entre les estimations du RNA-ERA(tanh) et les mesures. Mode climatisation

Malgré l'utilisation d'une base d'apprentissage arbitraire et quel que soit le mode de fonctionnement, les estimations des RNA sont précises même avec le modèle sélectionné par « BIC2 » « AIC2 » et « R2bar ». Ces résultats prometteurs confirment la pertinence de la méthodologie malgré ses limitations constatées.

IV.6. Bilan

Les résultats de ce chapitre sont en cohérence avec ceux de du chapitre précédent. Les résultats de l'application de la méthodologie sont très satisfaisants pour les meilleurs RNA et ce pour tous les systèmes.

Quant aux modèles sélectionnés par les trois critères statistiques, les résultats ne sont pas toujours bons. En effet, la capacité de chacun des critères à sélectionner un modèle avec un bon pouvoir de généralisation dépend du système étudié. Les résultats dans le cas du système SSCPROTO1 sont loin de ce que nous attendions. Les modèles sélectionnés dans le cas de ce système sont caractérisés par un nombre relativement élevé de neurones dans la couche cachée (7 et 8). Or les meilleurs modèles possèdent un nombre de neurones au maximum égale à 2. Ceci est probablement dû au fait que le terme de pénalisation pour ces critères n'est pas assez important dans le cas du système SSCPROTO1.

Les résultats des modèles sélectionnés selon les différents critères montrent qu'il est impossible de sélectionner un RNA pertinent pour tous les systèmes. Or, pour avoir une méthodologie générique les critères de sélection devraient être capables de repérer un bon modèle indépendamment du système étudié.

CONCLUSION

La méthodologie a été appliquée à plusieurs systèmes réels dans le cadre de ce dernier chapitre. Ces systèmes sont représentatifs des différents systèmes solaires pour le chauffage, la production d'eau chaude sanitaire et la climatisation dans les bâtiments.

Pour chaque système, les résultats de l'application de la méthodologie ont été comparés aux mesures issues des différents essais selon la méthode SCSPT.

Pour les meilleurs modèles, les erreurs d'estimation sont au maximum de l'ordre de 9% quel que soit le système. Pour SSC-PROTO2, ces erreurs sont plus importantes, dans le cas de Barcelone, mais restent tout à fait acceptables puisque les différences en valeur absolue sont faibles.

Les estimations des modèles sélectionnés selon les critères « *BIC2* », « AIC2 » et « R2bar » sont quant à elles légèrement moins bonnes que les premières mais restent très satisfaisantes pour la plupart des systèmes. Aussi, la capacité de chaque critère à sélectionner un modèle pertinent reste très dépendante du système auquel la méthodologie est appliquée.

Les résultats très satisfaisants de ce dernier chapitre révèlent le potentiel de la méthodologie développée et permettent sa validation expérimentale. Cependant, il serait très intéressant, pour une validation complète, d'envisager la comparaison des résultats de son application avec des données de suivi *in situ* dans différentes conditions et sur une année.

Enfin, ce dernier chapitre a permis également la confirmation de la pertinence de l'environnement de Zurich-SFH60-16 m² comme conditions de test pour la formation des données d'entraînement.

CONCLUSION GENERALE

Contributions

Les énergies renouvelables, en particulier l'énergie solaire, s'imposent de jours en jours comme une solution indéniable à la problématique d'énergie actuelle dans le monde. Cependant, si la puissance installée du solaire photovoltaïque a augmenté d'environ 3500% entre 2004 et 2013, celle du solaire thermique a augmenté d'environ 300% seulement sur la même période.

L'une des raisons qui peuvent expliquer ce développement assez modeste du solaire thermique réside dans le manque d'approche fiable et générique pour estimer les performances énergétiques des systèmes solaires thermiques. Or, ceci est indispensable pour garantir les performances du système aux utilisateurs.

Par la présente thèse, nous avons amorcé les voies d'une nouvelle solution complète et générique pour résoudre cette problématique. La solution consiste en une méthodologie de caractérisation des performances énergétiques qui peut être appliquée à différents systèmes solaires combinés. Ces systèmes utilisent de l'énergie solaire pour répondre, en partie, aux besoins de chauffage, de climatisation et d'ECS dans les bâtiments. L'énergie solaire étant intermittente, les SSC sont couplés à des systèmes d'appoint pour prendre le relais si cela est nécessaire.

A travers l'étude réalisée dans le cadre de la thèse, nous avons présenté les limites et inconvénients des méthodes actuelles. Ces méthodes sont soit basées sur plusieurs essais physiques qui caractérisent séparément les composants du système à évaluer (cette méthode ne prend pas en compte les vraies interactions entre ces derniers), soit sur des modèles physiques qui peuvent être complexes et difficilement identifiables notamment du fait que les systèmes actuels sont de plus en plus compacts et préfabriqués en usine.

La méthodologie que nous proposons prend comme entrée les données de l'essai, de 12 jours, du système selon la méthode SCSPT sur le banc d'essais semi-virtuel développée à l'INES. Elle permet ensuite d'élaborer un modèle neuronal modélisant le système complet.

La méthodologie développée cumule donc les points forts de la méthode SCSPT, à savoir la récupération de données issues d'un test unique du système dans sa globalité, et ceux des RNA, qui sont l'apprentissage du comportement du système selon la configuration des entrées et sortie considérées.

L'estimation de la consommation du système à évaluer se fait par une simple simulation à l'aide du RNA dans les conditions désirées.

Le développement des RNA nécessite la détermination d'un ensemble d'éléments qui le constitue (l'architecture du modèle, la FA du neurone en sortie, les variables retardées à prendre en compte, le nombre de neurones et de couches cachées, la stratégie d'apprentissage, le traitement des données etc.). Un processus complet a donc été proposé à ce propos. Une méthode de sélection des modèles pertinents s'appuyant sur des critères

statistiques, calculés uniquement sur la base des données d'apprentissage, a été également proposée.

La thèse propose une configuration entrées-sortie des modèles RNA compatible avec l'aspect générique de la méthodologie. En effet, les entrées retenues pour les modèles neuronaux représentent les puissances des besoins énergétiques du bâtiment (chauffage, ECS, climatisation) et de la ressource solaire tandis que la sortie représente la consommation du système. Cette configuration en « puissance » est facilement extrapolable à différents systèmes dans le bâtiment.

La pertinence de la méthodologie a été validée dans un premier temps à l'aide de plusieurs expérimentations numériques. Elles consistaient en une comparaison entre les résultats des simulations des RNA, entraînés sur une base de données issues d'une simulation de la séquence de test SCSPT du système considéré, et les calculs des modèles physiques détaillés du même système. Les systèmes considérés dans cette partie représentent un panel des systèmes actuellement disponibles sur le marché (système solaire combiné à une chaudière à gaz, à une pompe à chaleur ou intégrant une machine à absorption).

Globalement, une concordance très satisfaisante entre les estimations des meilleurs modèles neuronaux, dans plusieurs conditions différentes (climat, qualité thermique du bâtiment et surface de capteur solaire), avec les calculs des modèles physiques a été remarquée. Pour certaines conditions très différentes des conditions d'apprentissage (à Barcelone dans la plupart des cas), les erreurs relatives sont relativement importantes mais l'amplitude des écarts reste faible.

Dans un second temps, une application de la méthodologie dans le cas de plusieurs systèmes réels a été réalisée. Les résultats de comparaison entre les estimations des modèles RNA et des mesures confirment à la fois la pertinence de la méthodologie proposée ainsi que la qualité des conditions « Zurich-SFH60-16m² » comme étant un environnement d'apprentissage approprié.

Des données de suivi *in situ* d'un système intégrant une machine à absorption ont également été exploitées. Bien que les données utilisées pour l'apprentissage ne soient pas issues d'un essai selon la méthode SCSPT, la consommation du système a été estimée par les RNA avec un très bon degré de précision. Ces résultats encourageants laissent penser qu'il est possible d'utiliser ce type de modèle pour l'exploitation des données de suivi *in situ* pour de nouvelles problématiques comme le développement de systèmes de régulation avancée.

Les résultats des expérimentations numériques et physiques ont permis également de montrer les limites de l'approche de sélection des modèles selon les critères statistiques. En effet, sélectionner un modèle neuronal sur la base des critères considérés dans le mémoire présente un risque non négligeable de sélectionner un modèle avec des performances très modestes. Ces résultats concernent pratiquement la plupart des systèmes étudiés. Aussi, selon ces résultats, la fonction d'activation « tanh » apparait comme étant la plus pertinente par rapport à la fonction d'activation linéaire.

Il est important de signaler qu'aucune hypothèse n'a été faite par rapport aux systèmes auxquels il est possible d'appliquer la méthodologie. Ainsi la méthode est potentiellement

valable à d'autres typologies de systèmes intégrant de nouvelles fonctions comme la ventilation ou la génération de l'électricité via des panneaux solaires photovoltaïques.

Perspectives

Les perspectives de la thèse sont nombreuses, et concernent à la fois des voies d'amélioration mais aussi des débouchés variés.

Dans le but d'améliorer la qualité de la méthodologie et augmenter sa fiabilité nous avons repéré plusieurs possibilités d'amélioration qui concernent :

- **La séquence d'apprentissage** : le pouvoir de généralisation des RNA est très dépendant des données d'apprentissage. Ainsi les erreurs commises par les modèles neuronaux sont plus importantes dans le cas de données ayant une dynamique ou des amplitudes très différentes des conditions d'entraînement. L'amélioration des prédictions des RNA pour ces conditions extrêmes ne peut pas se faire via l'amélioration de son architecture ou du processus de leur entrainement. Cependant, une base de données intégrant toute la dynamique du système avec des amplitudes très variées peut énormément améliorer le pouvoir de généralisation des modèles neuronaux. Les 12 jours de la nouvelle séquence de test peuvent, par exemple, être choisis intelligemment parmi les jours de l'année dans plusieurs régions (Barcelone, Zurich, …) de telle sorte que la séquence soit représentative de tous les climats ou au moins pour les climats auxquels le système est destiné. Cette possibilité d'amélioration a été explorée dans l'Annexe B.
  Une autre méthode intéressante consisterait à utiliser des plans d'expériences optimaux pour l'apprentissage dynamique des RNA. La difficulté de ce type de méthode réside dans la possibilité de réaliser les essais physiques correspondants. En effet, le banc d'essai doit être très réactif pour répondre aux plans d'expériences générés. Aussi, la durée pourrait être supérieure à 12 jours ce qui augmenterait le coût de la méthodologie.

- **La sélection des modèles** : pour augmenter la fiabilité de la méthodologie il est indispensable de disposer d'un moyen efficace et sûr pour sélectionner un modèle, parmi ceux créés, qui se rapproche le plus du meilleur RNA. Il a été remarqué que, parmi l'ensemble des modèles sélectionnés selon les critères statistiques au sein d'une même famille (c'est-à-dire parmi les RNA qui diffèrent uniquement par rapport à leur nombre de neurones dans la couche cachée), il existe souvent un modèle pertinent. L'idée d'amélioration peut donc consister à sélectionner un seul modèle par famille et puis à réaliser une étude de sensibilité pour chacun des RNA sélectionnés. Cette dernière va permettre d'écarter les modèles neuronaux qui ne reproduisent pas le comportement physique du système et de ne garder que le meilleur.
  Une variante serait, dans un premier temps, de sélectionner un ensemble de modèles neuronaux selon un critère donné ou tout simplement suite à une analyse de sensibilité. Ensuite, la deuxième étape consisterait à former un modèle unique

basé sur tous ces modèles sélectionnés. Le modèle résultant serait un modèle constitué d'une combinaison [61] des différents modèles sélectionnés, et s'appuyant sur des pondérations dépendantes des performances de chaque modèle (calculées uniquement sur la base des données d'apprentissage).

- **L'estimation du degré de confiance sur les résultats** : quelle que soit la pertinence de la base d'apprentissage, du processus d'entraînement ou de sélection, il est impossible de développer un modèle précis qui peut se substituer aux systèmes pour toutes les conditions imaginables. Dans ce sens, il serait intéressant de disposer d'une méthode qui permette d'approximer le degré de fiabilité des prédictions des RNA à l'aide par exemple d'intervalles de confiance. L'étude bibliographie réalisée a révélé que de telles méthodes existent [120], [121].

Le développement des pistes proposées en perspective peut résulter en une méthodologie complète et fiable. Cette méthodologie permettra l'estimation des performances énergétiques des systèmes thermiques dans différentes conditions, en se basant sur un test court et unique du système à évaluer. La méthodologie pourra donc ensuite être utilisée comme un moyen pour présenter des garanties de performances aux utilisateurs, de certifications ou d'étiquetage des systèmes.

Il est indispensable de rappeler que les résultats de la thèse sont tous issus de la simulation des RNA en boucle fermée. L'utilisation des modèles en boucle ouverte, comme prévue, permet de réduire l'accumulation des erreurs étant donné que la variable bouclée est réelle et n'est plus estimée. La Figure 4- 42 en est la preuve. Avec de tels degrés de précision il serait possible d'envisager l'utilisation de la méthodologie pour former des modèles réutilisables dans des systèmes de contrôle-commande avancés avec comme objectif, par exemple, la garantie de performances solaires…

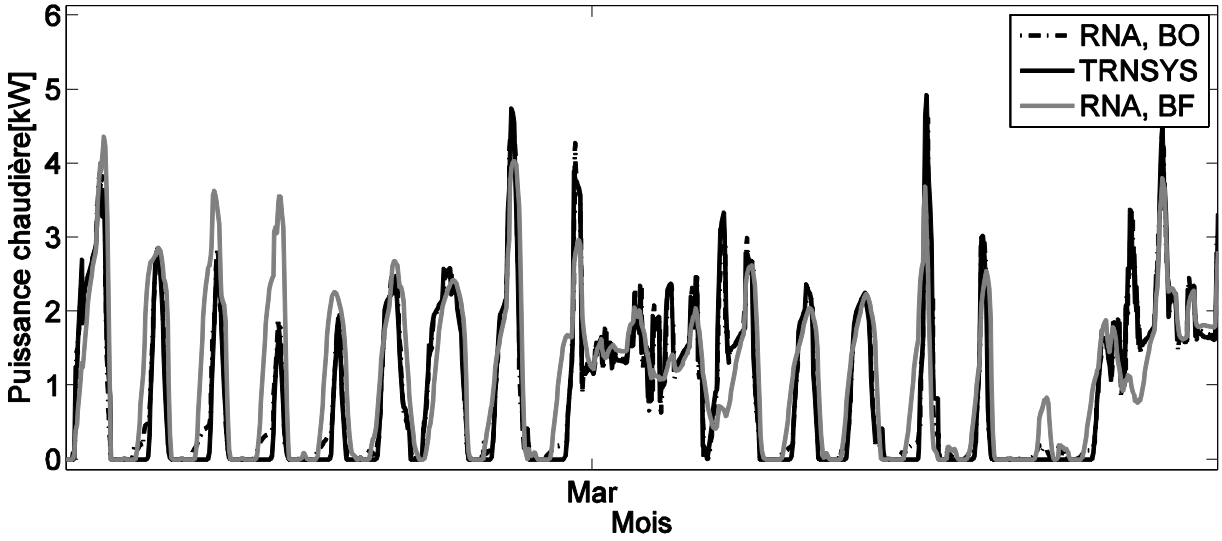

Figure 4- 42 : Exemple de comparaison entre l'évolution de la puissance du système estimée par le RNA-R2nrj(tanh) en boucle fermée et ouverte et calculée par le logiciel TRNSYS. Zurich-SFH60-16m². Cas du système décrit dans le deuxième chapitre.

Références

[1] H. Saint-Simon. Catéchisme des industriels. Paris: Imprimerie de Sétier, 1823.

[2] J. Hansen, R. Ruedy, M. Sato and K. Lo. “Global Surface Temperature Change” *Reviews of Geophysics,* vol. 48, no. 4, 2010.

[3] GIEC, “Le cinquième Rapport d’évaluation” 2013. [En ligne]. Available: https://www.ipcc.ch/home_languages_main_french.shtml. [Accès le 21 09 2015].

[4] GIEC, “Rapports d’évaluation” [Online]. Available: https://www.ipcc.ch/home_languages_main_french.shtml. [Accessed 30 06 2015].

[5] L. Pérez-Lombard, J. Ortiz et C. Pout. “A review on buildings energy consumption information” *Energy and Buildings,* vol. 40, n° 13, pp. 394-398, 2008.

[6] D. Ürge-Vorsatz, N. Eyre, P. Graham, D. Harvey, E. Hertwich, Y. Jiang, C. Kornevall, M. Majumdar, J. E. McMahon, S. Mirasgedis, S. Murakami et A. Novikova, «Chapter 10 - Energy End-Use: Building. In Global Energy Assessment - Toward a Sustainable Future» 2012. [En ligne]. Available: http://www.iiasa.ac.at/web/home/research/Flagship-Projects/Global-Energy-Assessment/Chapte10.en.html. [Accès le 21 09 2015].

[7] H. Ben Ahmed, B. Multan et Y. Thiaux, “Consommation d'énergie, ressources énergétiques et place de l'électricité” *Techniques de l'Ingénieur,* vol. TIB248DUO*,* 2011, pp. D 3 900v2-1 - D 3 900v2-18.

[8] F. Enric, S. Federico et F. Marco. “Integrated HVAC and DHW production systems for Zero Energy Buildings” *Renewable and Sustainable Energy Reviews,* vol. 40, pp. 515-541, 2014.

[9] A. Leconte, “Identification de caractéristiques réduites pour l’évaluation des performances des systèmes solaires combinés (thèse) ” Université de Savoie, Bourget du Lac, 2011.

[10] F. A. Peuser, K.-H. Remmers et M. Schnauss. Installations solaires thermiques: Conception et mise en oeuvre, Le Moniteur, 2005.

[11] F. Marias, P. Neveu, G. Tanguy et P. Papillon. “Thermodynamic analysis and experimental study of solid/gas reactor operating in open mode” *Energy,* vol. 66, pp. 757–765, 2014.

[12] W. Dobbin. “Stratification in solar domestic hot water storage tanks” *Energy Developments, New Forms, Renewables, Conservation,* pp. 501–506, 1984.

[13] E. Kaçan et K. Ulgen. “Energy and exergy analysis of solar combisystems” *Int. J. Exergy,* vol. 14, n° 13, pp. 364-387, 2014.

[14] P. Papillon, “Combisol project: Solar Combisystems Promotion and Standarisation” 2011. [Online]. Available: http://solarthermalworld.org/content/combisol-project-solar-combi-systems-are-gaining-

market-share. [Accessed 27 January 2015].

[15] A. Thür, J. Breidler and G. Kuhness, “D2.4: Updated State of the Art Report Of Solar Combisystems Analysed within CombiSol” CombiSol Project, 2010.

[16] D. Pineau, P. Rivière, P. Stabat, P. Hoang et V. Archambault. “Performance analysis of heating systems for low energy houses” *Energy and Buildings,* vol. 65, pp. 45-54, 2013.

[17] N. Kelly et J. Cockroft. “Analysis of retrofit air source heat pump performance: Results from detailed simulations and comparison to field trial data” *Energy and Buildings,* vol. 43, pp. 239-245, 2011.

[18] A. Tautvydas, “E4EM: heat pump case study” [En ligne]. Available: http://e4em.eu/contents/attached/c1/Case_study_Heat_Pump.pdf. [Accès le 16 11 2014].

[19] J. Brenn, P. Soltic et C. Bach. “Comparison of natural gas driven heat pumps and electrically driven heat pumps with conventional systems for building heating purposes” *Energy and Buildings,* vol. 42, n° 16, pp. 904-908, 2010.

[20] A. Nitkiewicz et R. Sekret. “Comparison of LCA results of low temperature heat plant using electric heat pump, absorption heat pump and gas-fired boiler” *Energy Conversion and Management,* vol. 87, pp. 647-652, 2014.

[21] J. Ruschenburg, S. Herkel et H.-M. Henning. “A statistical analysis on market-available solar thermal heat pump systems” *Solar Energy,* vol. 95, pp. 79-89, 2013.

[22] F. Elimar, H. Michel, S. Herkel et J. Ruschenburg, “Systematic classification of combined solar thermal and heat pump systems” *EuroSun*, Graz, 2010.

[23] M. D’Antoni, R. Fedrizzi et W. Sparber, “nachhaltigwirtschaften” [En ligne]. Available: http://www.nachhaltigwirtschaften.at/iea_pdf/newsletter/iea_shc_task44_annex38.pdf. [Accès le 16 11 2014].

[24] D. Carbonell, M. Y. Haller, D. Philippen et E. Frank. “Simulations of combined solar thermal and heat pump systems for domestic hot water and space heating” *Energy Procedia,* vol. 48, p. 524 – 534, 2014.

[25] M. Y. Haller et D. Carbonell, “Solar and heat pump systems – summary of simulation results of the IEA SHC TASK 44/HPP ANNEX 38” *IEA heat pump coference*, Montréal , 2014.

[26] M. Y. Haller et E. Frank, “On the potential of using heat from solar thermal collectors for heat pump evaporators” *ISES Solar Wolrld Congress*, Kassel, 2011.

[27] S. Stark, A. Loose et H. Drück. “Field test results from combined solar thermal and air-source heat pump systems with a special focus on defrosting” *Energy Procedia ,* vol. 48, p. 654 – 663, 2014 .

[28] V. Trillat-Berdal, B. Souyri et G. Fraisse. “Experimental study of a ground-coupled heat pump combined with thermal solar collectors” *Energy and Buildings ,* vol. 38, p. 1477–1484, 2006.

[29] A. Papadopoulos, S. Oxizidis et N. Kyriakis. “Perspectives of solar cooling in view of the developments in the air-conditioning sector” *Renewable and Sustainable Energy Reviews,* vol. 7, n° 15, pp. 419-438, 2003.

[30] I. Bell, A. Al-Daini, H. Al-Ali, R. Abdel-Gayed et L. Duckers. “The design of an evaporator/absorber and thermodynamic analysis of a vapor absorption chiller driven by solar energy Pages” *Renewable Energy,* vol. 9, n° 11, p. 657–660, 1996.

[31] C. A. Balaras, G. Grossman, H.-M. Henning, C. A. I. Ferreira, E. Podesser, L. Wang et E. Wiemken. “Solar air conditioning in Europe—an overview” *Renewable and Sustainable Energy Reviews,* vol. 11, n° 12, pp. 299-314, 2007.

[32] M. Helm, K. Hagel, W. Pfeffer, S. Hiebler et C. Schweigler. “Solar Heating and Cooling System with Absorption Chiller and Latent Heat Storage – A Research Project Summary” *Energy Procedia,* vol. 48, pp. 837-849, 2014.

[33] M. Helm, C. Keil, S. Hiebler, H. Mehling et C. Schweigler. “Solar heating and cooling system with absorption chiller and low temperature latent heat storage: Energetic performance and operational experience” *International Journal of Refrigeration,* vol. 32, n° 14, pp. 596-606, 2009.

[34] R. Lizarte, M. Izquierdo, J. Marcos et E. Palacios. “An innovative solar-driven directly air-cooled LiBr–H2O absorption chiller prototype for residential use” *Energy and Buildings,* vol. 47, pp. 1-11, 2012.

[35] R. Gomri. “Simulation study on the performance of solar/natural gas absorption cooling chillers” *Energy Conversion and Management,* vol. 65, pp. 675-681, 2013.

[36] H.-M. Henning et J. Döll. “Solar Systems for Heating and Cooling of Buildings” *Energy Procedia,* vol. 30, pp. 633-653, 2012.

[37] V. Boopathi Raja et V. Shanmugam. “A review and new approach to minimize the cost of solar assisted absorption cooling system” *Renewable and Sustainable Energy Reviews,* vol. 16, n° 19, pp. 6725-6731, 2012.

[38] Solair-project. “Here you can find a list of current and past projects in the field of solar cooling” [En ligne]. Available: http://www.solair-project.eu/112.0.html. [Accès le 30 06 2015].

[39] Eur'Observ'ER, “Solar Thermal and Concentrated Solar Power Barometer” *Systèmes Solaires,* Mai 2013.

[40] D. Mugnier, “Work plan for Task 53: New Generation solar cooling & heating systems (PV and solar thermally driven systems)” [En ligne]. Available: http://task53.iea-shc.org/data/sites/1/publications/Task53-Work-Plan.pdf. [Accès le 17 11 2014].

[41] M. Y. Haller, R. Haberl, I. Mojic et E. Frank. “Hydraulic Integration and Control of Heat Pump and Combi-storage: Same Components, Big Differences” *Energy Procedia,* vol. 48, pp. 571-580, 2014.

[42] R. Haberl, M. Y. Haller, A. Reber et E. Frank. “Combining Heat Pumps with Combistores: Detailed Measurements Reveal Demand for Optimization” *Energy Procedia,* vol. 48, pp. 361-369, 2014.

[43] I. Malenković, “Definition of performance figures for solar and heat pump systems” 2012. [En ligne]. Available: http://www.estif.org/fileadmin/estif/content/projects/QAiST/QAiST_results/QAiST%20D5.1%20TR%205.1.3%20Performance%20figures.pdf.

[44] J. Nowag, F. Boudéhenn, A. Le Denn, F. Lucas, O. Marc, M. Raduslescu et P. Papillon. “Calculation of Performance Indicators for Solar Cooling, Heating and Domestic Hot Water System” *Energy Procedia,* vol. 30, pp. 937-946, 2012.

[45] C. Wemhöner et T. Afjei, “Seasonal performance calculation for residential heat pumps with combined space heating and hot water production (FHBB Method)” University of Applied Sciences Basel, Bâle, 2003.

[46] J. Facão and M. J. Carvalho. “New test methodologies to analyse direct expansion solar assisted heat pumps for domestic hot water” *Solar Energy,* vol. 100, pp. 66-75, 2014.

[47] V. Partenay, “Développement d'une méthodologie d'évaluation des performances de systèmes de pompes à chaleur géothermique (Thèse)” Université de Savoie, Bourget-du-Lac, 2010.

[48] G. Panaras, E. Mathioulakis et V. Belessiotis. “A method for the dynamic testing and evaluation of the performance of combined solar thermal heat pump hot water systems” *Applied Energy,* vol. 114, pp. 124-134, 2014.

[49] P. Almeidaa, M. J. Carvalho, R. Amorim, J. F. Mendes and V. Lopes. “Dynamic testing of systems–use of TRNSYS as an Approach for Parameter Identification” *Solar Energy,* vol. 104 , p. 60–70, 2014.

[50] M. Haller, R. Haberl, T. Persson, C. Bales, P. Kovacs, D. Chèze and P. Papillon. “Dynamic whole system testing of combined renewable heating systems – The current state of the art” *Energy and Buildings,* vol. 66, pp. 667-677, 2013.

[51] MacSheep, “MacSheep - New Materials and Control for a next generation of compact combined Solar and heat pump systems with boosted energetic and exergetic performance” 2012-2015. [Online]. Available: http://macsheep.spf.ch/Project.248.0.html. [Accessed 2014].

[52] A. Leconte, G. Achard et P. Papillon. “Global approach test improvement using a neural network model identification to characterise solar combisystem performances” *Solar Energy,* vol. 86, pp. 2001-2016, 2012.

[53] M. Albaric, J. Nowag and P. Papillon, “Thermal performance evaluation of solar combisystems using a

global approach" *Eurosun conference*, Lisbon, 2008.

[54] M. Albaric, B. Mette, J. Ullman, H. Drück and P. Papillon, "Comparison of two different methods for solar combisystems performance testing" *Eurosun conference*, Graz, 2010.

[55] R. Heimrath and M. Haller, "Project Report A2 of Subtack A: The reference heating system, the template solar system. A report of IEA SHC - Task 32" 2007.

[56] U. Jordan et K. Vajen, "Realistic domestic hot water profiles in different time scales" 2001. [En ligne]. Available: http://sel.me.wisc.edu/trnsys/trnlib/library15.htm#IEA26Load. [Accès le 17 11 2014].

[57] D. Triche, "Développement d'une méthode d'essais dynamiques semi-virtuelle pour caractériser les systèmes de climatisation/chauffage solaire (Projet de Master)" CEA, INES, LETh, Bourget du Lac, 2012.

[58] O. Pajonk, "Overview of System Identification with Focus on Inverse Modeling Literature review" Institute of Scientific Computing, Technical University of Braunschweig, 2009.

[59] G. Dreyfus, Neural networks methodology and applications, Springer, 2005.

[60] L. Ljung, System Identification: Theory for the User (2nd Edition), Prentice Hall, 1999.

[61] G. Horvath, "Neural Networks in System Identification, Chapter 3. In: V. Piuri (Ed.) Neural networks in measurement Systems NATO ASI NIMIA" Crema, 2001.

[62] R. Pearsona and Ü. Kottab. "Nonlinear discrete-time models: state-space vs. I/O representations" *Journal of Process Control,* vol. 14, no. 5, p. 533–538, 2004.

[63] M. Norgaard, O. Ravn, N. Poulsen and L. Hansen, Neural networks for modelling and control of dynamic systems, Springer, 2000.

[64] F. Giri and E.-W. Bai, Block-oriented nonlinear system identification, Springer, 2010.

[65] G. Cybenko. "Approximation by Superpositions of a Sigmoidal Function" *Math. Control Signals Systems,* vol. 2, pp. 303-314, 1989.

[66] S. Rastko R. et L. Frank L. "Neural Network Approximation of piecewise continuous functions: Application to friction compensation" *IEEE Transactions on Neural Networks,* vol. 13, n° 13, May 2002.

[67] S. A. Kalogirou. "Artificial neural networks in renewable in renewable energy systems application: a review" *Renewable and Sustainable Energy Reviews,* vol. 5, pp. 373-401, 2001.

[68] S. Kalogirou, E. Mathioulakis and V. Belessiotis. "Artificial neural networks for the performance prediction of large solar systems" *Renewable Energy,* vol. 63, pp. 90-97, 2014.

[69] C. Burattia, M. Barbanerab and D. Palladinoa. “An original tool for checking energy performance and certification of buildings by means of Artificial Neural Networks” *Applied Energy,* vol. 120, p. 125–132, 2014.

[70] W. Yaïci and E. Entchev. “Performance prediction of a solar thermal energy system using artificial neural networks” *Applied Thermal Engineering,* vol. 73 , pp. 1348–1359, 2014.

[71] S. Rosiek and F. Batlles. “Modelling a solar-assisted air-conditioning system installed in CIESOL building using an artificial neural network” *Renewable Energy,* vol. 35 , pp. 2894-2901, 2010.

[72] S. Rosiek and F. Batlles. “Performance study of solar-assisted air-conditioning system provided with storage tanks using artificial neural networks” *International Journal of Refrigeration,* vol. 34, pp. 1446-1454, 2011.

[73] M. Ben Ammar, M. Chaabene et Z. Chtourou. “Artificial Neural Network based control for PV/T panel to track optimum thermal and electrical power” *Energy conversion and management,* vol. 65, pp. 372-380, 2013.

[74] M. Kljajic, D. Gvozdenac and S. Vukmirovic. “Using neural networks for modeling and predicting boiler's operationg performance” *Energy,* vol. 45, pp. 304-311, 2012.

[75] P. Vig and I. Farkas. “Neural network Modeling of thermal stratification in a solar DHW storage” *Solar Energy,* vol. 84, pp. 801-806, 2010.

[76] Y. O. Ozgoren, S. Cetinkaya, S. Saridemir, C. Adem and F. Kara. “Predictive modeling of performance of a helium charged Stinrling engine using and artificial neural network” *Energy conversion and Management,* vol. 67, pp. 357-368, 2013.

[77] P. Wira, *Approche neuromimétique pour l'identification et la commande (mémoire HDR),* Université de Haute-Alsace, 2009.

[78] G. Zhang, B. E. Patuwo et M. Y. Hu. “Forecasting with artificial neural networks: The satte of the art” *International Journal of Forecasting,* vol. 14, pp. 35-62, 1998.

[79] *Neural Network Tool Box,* MATLAB R2012b, 2012.

[80] K. S. Narendra et K. Parthasarathy. “Learning Automata Approach to Hierarchical Multiobjective Analysis” *IEEE Transactions on systems, man, and cybernetics,* vol. 21, n° 11, pp. 263-273, 1991.

[81] N. Zhang, “Urban Stormwater Runoff using Recurrent Neural Networks” *Advances in Neural Networks, ISNN 2011: 8th International Symposium on Neural Networks*, Guilin, China, 2011.

[82] R. May, G. Dandy et H. Maier, “Review of Input Variable Selection Methods for Artificial Neural Networks”, *Artificial Neural Networks*, *Methodological Advances and Biomedical Applications,* Prof. Kenji

Suzuki (Ed.), ISBN : 978-953-307-343-2, in Tech, 2011.

[83] B. Tlili, “Identification et commande des systèmes non linéaires : utilisation des modèles de type NARMA (thèse)” Université de Tunis El Manar, 2008.

[84] X. He and H. Asada, “*A new method for identifying orders of input-output models for nonlinear dynamic systems”* San Francisco: Proceedings of the 1993 American Control Conference, 1993, pp. 2520-2523.

[85] S. Fischer, P. Frey et H. Druck. “A comparison between state-of-the-art and neural network modeling of solar collectors” *Solar energy,* vol. 86, pp. 3268-3277, 2012.

[86] W. Gang and J. Wang. “Predictive ANN models of ground heat exchanger for the control of hybrid ground source heat pump systems” *Applied Energy,* 2013.

[87] B. Roberto, C. Ubaldo, S. Mariotti, R. Innamorati, S. Elisa et P. Mura. “Graybox and adaptative dynamic neural network identification models to infer the steady state efficiency of solar thermal collectors starting from the transient condition” *Solar Energy,* vol. 84, pp. 1027-1046, 2010.

[88] N. Li, L. Xia, D. Shiming, X. Xu et M. Chan. “Dynamic modeling and control of a direct expansion air conditioning system using artificial neural network” *Applied Energy,* vol. 91, pp. 290-300, 2012.

[89] A. Kusiak et G. Xu. “Modeling and optimization of HVAC systems using a dynamic neural network” *Energy,* vol. 42, pp. 241-250, 2012.

[90] K. Gnana Sheela and D. N. S. “Review on Methods to Fix Number of Hidden Neurons in Neural Networks” *Mathematical Problems in Engineering,* vol. 2013, no. Article ID 425740, 2013.

[91] Stathakis D. “How many hidden layers and nodes?” *International Journal of Remote Sensing,* vol. 30, n° 18, pp. 2133-2147, 2009.

[92] B. Widrow and M. Lehr. “30 Years of Adaptive Neural Networks: Perceptron, Madaline, and Backpropagation” *Proceedings of the IEEE,* vol. 78, no. 9, pp. 1415-1442, 1990.

[93] S. Lawrence, C. L. Giles and A. C. Tsoi, “What Size Neural Network Gives Optimal Generalization? Convergence Properties of Backpropagation” Institute for Advanced Computer Studies, University of Maryland, College Park, 1996.

[94] Q. Ning. “On the momentum term in gradient descent learning algorithms” *Neural Networks,* vol. 12, n° 11, p. 145–151, 1999.

[95] M. Z. Rehman et N. M. Nawi. “Improving the Accuracy of Gradient Descent Back Propagation Algorithm (GDAM) on Classification Problems” *International Journal on New Computer Architectures and Their Applications,* vol. 1, n° 14, pp. 838-847, 2011.

[96] M. F. Moller. “A scaled conjugate gradient algorithm for fast supervised learning” *neural networks,* vol. 6, pp. 525-533, 1993.

[97] M. Powell. “Restart procedures for the conjugate gradient method” *Mathematical Programming,* vol. 12, p. 241–254, 1977.

[98] D. Marquardt. “An Algorithm for Least-Squares Estimation of Nonlinear Parameters” *SIAM Journal on Applied Mathematics,* vol. 11, no. 2, p. 431–441, 1963.

[99] M. Riedmiller et H. Braun, “A direct adaptive method for faster backpropagation learning: The RPROP algorithm” *Proceedings of the IEEE International Conference on Neural Networks*, 1993.

[100] A. Lazrak, A. Leconte, G. Fraisse, P. Papillon and B. Souyri, “*A solar combisystem input-output modelling: exploration of several algorithms and learning procedures” COFMER,* Rabat, 2014, p. 10.

[101] D. Nguyen and B. Widrow. “Improving the learning speed of 2-layer neural networks by choosing initial values of the adaptative weights” *Proceeding of the International Joint Conference on Neural Networks,* vol. 3, pp. 21-26, 1990.

[102] B. Hassibi, D. G. Stork et G. J. Wolff, “Optimal brain surgeon and general network pruning” *Proc. IEEE Int. Conf. Neural Networks*, San Francisco, 1992.

[103] D. MacKay. “Bayesian interpolation” *Neural Computation,* vol. 4, no. 3, pp. 415-447, 1992.

[104] J. Sola and J. Sevilla. “Importance of input data normalization for the application of neural networks to complex industrial problems” *IEEE Transactions on Nuclear Science,* vol. 44, no. 3, pp. 1464-1468, 1997.

[105] M. Qia and G. P. Zhangb. “An investigation of model selection criteria for neural network time series forecasting” *European Journal of Operational Research,* vol. 132, no. 3, p. 666–680, 2001.

[106] R. Heimrath et M. Haller, “Project Report A2 of Subtack A: The reference heating system, the template solar system. A report of IEA SHC - Task 32” 2007.

[107] TRANSSOLAR, “TRNSYS 16 User Manual - Multizone building modeling with type 56 and TRNBuild” 2005. [En ligne]. Available: http://sel.me.wisc.edu/trnsys/user-resources/index.html. [Accès le 20 4 2015].

[108] ISO13790:2008, “Performance énergétique des bâtiments -- Calcul des besoins d'énergie pour le chauffage et le refroidissement des locaux” 2008.

[109] E. s. EN442-2, “Radiators and convectors. Part 2 : Test methods and ratings” Comité Européen de Normalisation, Bruxelles, 1996.

[110] B. Mette, J. Ullman, H. Drück, M. Albaric, A. Leconte and P. Papillon, “Combisol project. Solar combisystem promotion and standardisation. D3.1:Comparison of test methods” 2010. [Online].

Available: http://www.combisol.eu/. [Accessed 20 4 2015].

[111] M. Albaric, P. Riederer and S. Boddaert, “Projet EvaluationSSC : Méthode d'évaluation de Systèmes Solaires Combinés, Rapport final n°2, Phase 3/4 - Amélioration de la méthode et comparaisons inter-laboratoire [Confidentiel, pour l'ADEME]” 2009.

[112] T. Fiedler and A. Knirsch, “Add-on to the hitherto existing gas fuelled boiler-model for a wood-fired boiler with calorific value usage” Transsolar, Stuttgart, 1998.

[113] A. Heinz and M. Haller, “Models of Sub-Components and Validation for the IEA SHC Task 44/HPP Annex 38 - Part C: Heat Pump Models - DRAFT - A technical report of subtask C Deliverable C2” Appendix A3 - Description of TRNSYS Type 877 by IWT and SPF, 2012.

[114] M. Wetter and A. Huber, "TRNSYS Type 451 - Vertical Borehole Heat Exchange Model" 1997. [Online]. Available: www.transsolar.com.

[115] H. Drûck, “Multiport store - Model for TRNSYS - Type 340. Version 1.99F” Institut fûr Thermodinamik und Wârmetachnik (ITW), Universitât Stuttgart, 2006.

[116] M. Haller, “TRNSYS Type 832 v5.00 “Dynamic Collector Model by Bengt Perers” Updated Input-Output Reference” Institut für Solartechnik, Rapperswil. (Switzerland), 2012.

[117] SWEP, “Economizers” [Online]. Available: http://www.swep.net/en/products_solutions/solutions/Pages/Economizers.aspx. [Accessed 09 07 2015].

[118] A. Heinz and M. Haller, “Description of TRNSYS Type 887 (v119)” IWT and SPF, 2014.

[119] F. Ziegler, H.-M. Hellmann and C. Schweigler, “An approximative method for modeling the operating characteristics of advanced absorption chillers” in *20th Int. congress Refrigeration*, Sydney, 1999.

[120] A. Khosravi, S. Nahavandi, D. Creighton and A. Atiya. “Comprehensive Review of Neural Network-Based Prediction Intervals and New Advances” *IEEE transactions on neural networks,* vol. 22, no. 9, pp. 1341-1356, 2011.

[121] G. Chryssolouris, M. Lee and A. Ramsey. “Confidence interval prediction for neural network models” *IEEE transactions on neural networks,* vol. 7, no. 1, pp. 229-232, 1996.

[122] I. Malenković, “Definition of Main System Boundaries and performance Figures for Reporting on SHP Systems” IEA SHC Task 44 / HPP Annex 38, A technical report of Subtask B, Deliverable B1.1, 2012.

Annexes

Annexe A. Tableaux récapitulatifs des indicateurs et critères de performances

Tableau A- 1 : Aperçu de différents critères de performances de capteurs solaires et systèmes solaires thermiques définies dans les normes et directives [122]

Norme/Directive	Critère de performance	Définitions
EN 12975-2	η	L'efficacité thermique du capteur solaire est le rapport de la chaleur enlevée par le fluide caloporteur pendant une période de temps donnée au produit de la surface du capteur et le rayonnement solaire sur cette surface, pendant la même période de temps en régime établi ou transitoire (selon ISO 9488).
ISO 9806 ASHRAE 93	η	Pareil qu'en EN 12975
	η_g	L'efficacité thermique du capteur solaire est définie comme le rapport entre l'énergie utile collectée et l'énergie solaire reçue par la section du capteur.
EN 12976, EN12977	f_{sol}	La fraction solaire est le rapport entre la chaleur fournie par le système thermique au système de distribution de chauffage et d'ECS et la charge totale du système.
	f_{sav}	Le taux d'économie d'énergie est défini comme le rapport entre la consommation de l'énergie classique économisée grâce au système thermique (énergie renouvelable) et la consommation du système de référence (sans système thermique solaire ou à une énergie renouvelable) pour répondre au même besoin.
	Performance thermique	Une performance thermique est définie comme un ensemble d'indicateurs de performance. Pour un système solaire sans appoint ils sont : La chaleur fournie par le système solaire, f_{sol} et l'énergie perdue. Pour des systèmes avec appoint : L'énergie nette demandée à l'appoint, f_{sav} et l'énergie perdue.
ISO 9459	Performance thermique	Définition comparable à la norme EN 12976 et EN 12977.
EN 15316-43		Même appellation qu'en EN 12977.

Tableau A- 2 : Aperçu de différents critères de performance de PAC et de systèmes de PAC définis dans les normes et directives [122]

Norme/Directive	Critère de performance	Définitions
EN 14511	COP	Rapport entre la puissance de chauffage en sortie de la PAC et celle à son entrée pour un point de fonctionnement en régime statique. Les énergies en entrée et en sortie sont corrigées par l'énergie nécessaire pour vaincre les pertes de charge dans les échangeurs thermiques.
	EER	Même définition que le COP mais pour la climatisation.
EN 15879-1	COP/EER	Utilise les mêmes définitions qu'en EN 14511 appliqués aux PAC à expansion directe.
EN 16147	COP	Rapport entre l'énergie nécessaire pour la production de l'ECS et l'électricité consommée lors un cycle de puisage complet. Le système se compose de la PAC, les pompes de recirculation et du ballon de stockage. Les pertes de chaleurs dans ce dernier sont comprises dans les calculs.
AHRI 320/325/330	COP	Rapport entre la puissance de chauffage, sans prendre en compte celle de l'appoint, et la puissance à l'entrée de la PAC pour un point de fonctionnement en régime statique.
	EER	Même définition que le COP mais pour la climatisation.
ISO 13256-1 ISO 13256-2	COP	Rapport entre la puissance nette de chauffage, et la puissance à l'entrée du système pour un point de fonctionnement en régime statique. Les puissances en entrées et en sorties sont corrigées de la même manière qu'en EN 14511
	EER	Même définition que le COP mais pour la climatisation
EN 14825	SCOP	Rapport entre l'énergie totale délivrée par le système pour les besoins de chauffage en une année et l'énergie totale à l'entrée du système. Le système est testé selon EN 14511 et les calculs se font selon quelques hypothèses concernant les besoins en chauffage, le climat, le contrôle commande du système etc.
	SEER	Même définition que le SCOP mais pour la climatisation
ASHRAE 116	HSPF	Rapport entre l'énergie totale délivrée par le système pour une saison donnée (<12 mois) de chauffage et l'énergie totale à l'entrée du système. Le système est testé selon une norme donnée et les calculs se font selon quelques hypothèses concernant les besoins en chauffage, le climat, le contrôle commande du système etc.
	HEER	Rapport entre la chaleur totale enlevée durant la saison d'utilisation pour la climatisation (<12 mois) et l'énergie totale en entrée durant la même période. Le calcul est réalisé de la même manière que pour HSPF.
Vdi 4650-1	SPF (β)	Rapport entre l'énergie fournie sur une période donnée (période de chauffage ou de climatisation) à la consommation d'énergie électrique du compresseur et les autres auxiliaires. Les calculs sont basés sur le test selon la norme EN 14511.
EN 15316-4-2	SPF	Rapport entre l'énergie totale en sortie et l'énergie totale en entrée de la PAC pour le chauffage et la production de l'ECS

Annexe B. Tentative d'amélioration de la séquence de test

B.1. Approche

La richesse de la séquence de test (création des données d'apprentissage) est un élément clé dans la modélisation par RNA. Les résultats obtenus en se basant sur la séquence classique sont satisfaisants même si elle n'a pas été développée spécialement pour l'apprentissage des RNA. Elle n'est donc certainement pas la séquence de test optimale pour l'entrainement des modèles neuronaux. Dans cette partie une proposition d'amélioration est proposée.

Il a été remarqué dans le deuxième et le troisième chapitre, que les erreurs de prédictions les plus importantes concernent très souvent le climat de Barcelone. Une possibilité d'amélioration de la séquence de test peut consister donc aux choix de quelques jours représentatifs de l'ensemble de climats que le système peut rencontrer et non seulement le climat de Zurich.

En vue de la configuration de modélisation entrées-sortie retenue, tout climat est caractérisé par l'ensoleillement, la température extérieure et la température d'eau froide. Pour des raisons de simplifications nous avons considéré uniquement les deux premières variables. La représentation des températures extérieures moyennes et l'énergie solaire sur une journée de ces deux variables dans un plan 2D est pertinente pour observer la variété des journées qui se succèdent pendant une année. Sur la Figure B- 1 est représenté un nuage de points sur un plan 2D (en abscisse la température journalière moyenne et en ordonnée l'énergie solaire journalière cumulée) correspondants aux jours de l'année de l'environnement de Zurich et Barcelone. Pour un entrainement optimal des RNA, il faut que les jours choisis soient répartis sur la totalité et de façon uniforme sur le nuage de points. Un tel choix permettra au modèle de s'adapter à des échantillons de données différentes et ainsi la possibilité d'apprendre le comportement du système en des circonstances différentes.

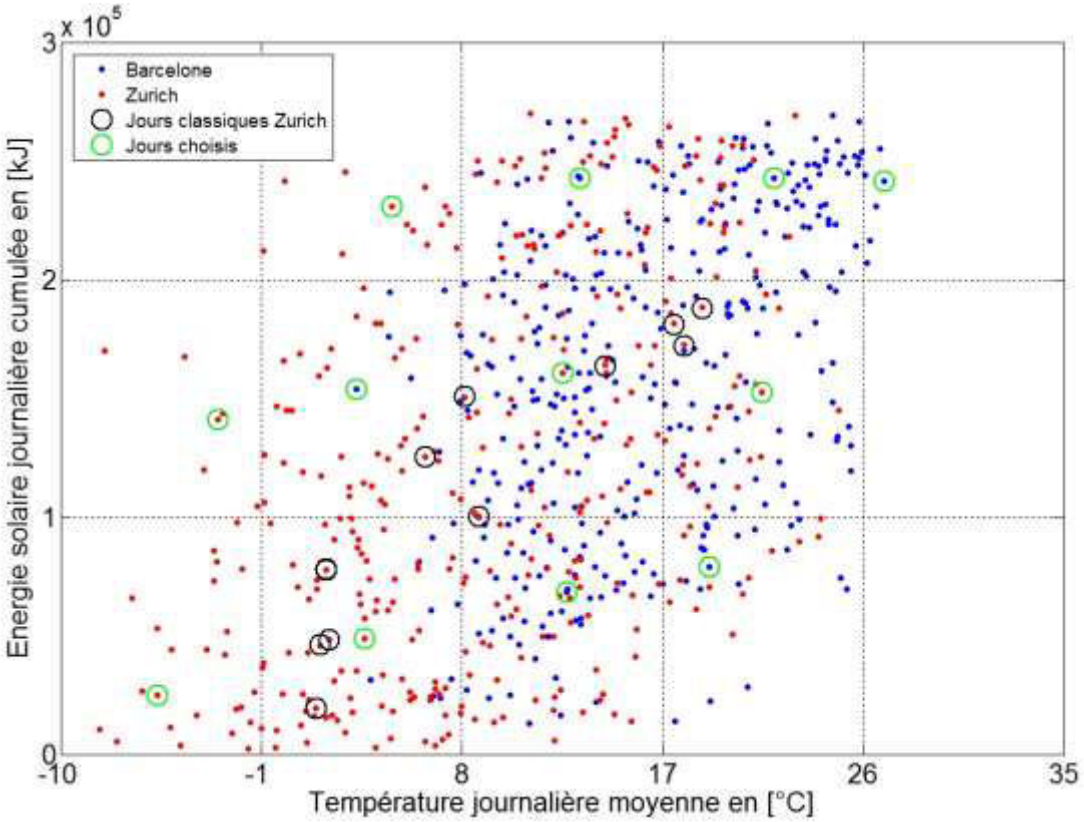

Figure B- 1 : Nuage de points sur un plan 2D, chaque point est équivalent à un jour de l'année

Les 12 jours de la séquence classique et les 12 jours qui vont former la nouvelle séquence prenant en compte les deux fichiers météo sont indiqués par des cercles. Ces derniers ont été choisis de façon à ce que le plan 2D soit, dans la mesure où cela est possible, uniformément représenté par les 12 points (12 jours) encerclés en vert. Le choix des jours dans le cadre de la méthode SCSPT étant dépendant de l'objectif de cette dernière, nous constatons clairement que ces 12 jours ne sont pas répartis uniformément sur le plan 2D formé par le nuage de points. Quelques jours choisis dans la séquence sont très similaires dans le plan 2D considéré. Ceci se voit par la faible distance entre quelques points. Nous pouvons imaginer que supprimer quelques journées de la séquence classique n'aura donc pas de conséquence sur la qualité de l'apprentissage. Ceci pourra être une voie d'amélioration importante pour la méthodologie de caractérisation. En effet, la réduction de la durée de la séquence baisse le coût d'essai en banc semi-virtuel.

B.2. Résultats

Deux modèles neuronaux ont été élaborés, l'un entrainé sur la séquence classique et l'autre sur la nouvelle séquence. Les deux modèles ont été ensuite utilisés pour prédire la consommation de la chaudière en une année d'un SSC dans 9 environnements différents composés d'un type de bâtiment, climat et une surface de capteur de 10m².

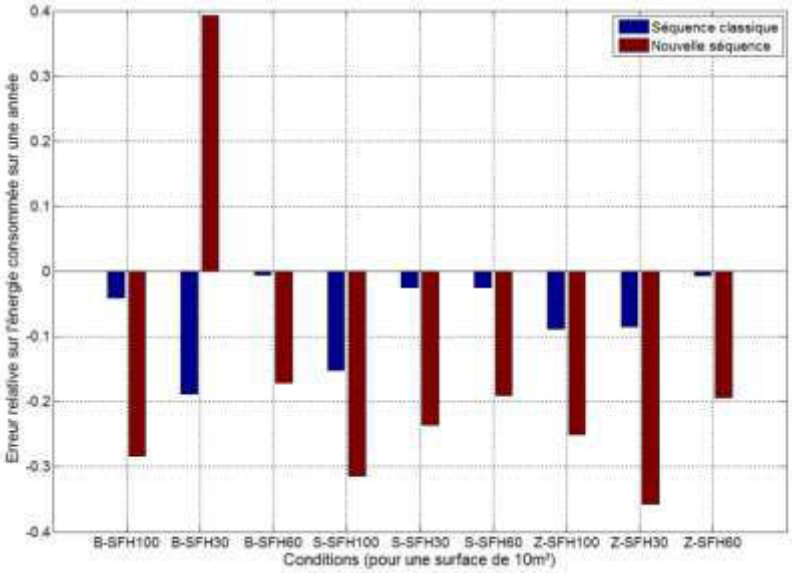

Figure B- 2 : Comparaison entre les résultats de la séquence classique et la nouvelle en termes de l'erreur relative sur l'énergie consommée

Les résultats de prédictions du RNA entrainé sur la nouvelle séquence sont moins performants en termes d'erreur relative sur l'énergie consommée prédite (Figure B- 2). En effet, quelles que soient les conditions, l'erreur relative est au moins 15% supérieure dans le cas d'un entrainement sur la nouvelle séquence.

La nouvelle séquence ne prend pas en compte l'inertie du ballon dans la sélection des jours. Ce qui montre que cette variable n'est pas négligeable dans la sélection de jours pertinents pour l'apprentissage. En conséquence les résultats montrent que la séquence classique est meilleure.

En perspectives, la sélection des jours de la séquence de test peut être faite autrement :

- L'utilisation d'un plan 3D en incluant la variable température de l'eau froide
- L'utilisation des variables en puissances (chauffage) en entrées des modèles RNA au lieu des températures ce qui garantira la prise en compte du bâtiment également.
- La complexification du plan de représentation en prenant en compte, dans le processus de sélection des jours, des caractéristiques statistiques de chaque variables (moyenne, valeur maximale/minimale etc.).
- L'utilisation d'un ensemble de fichiers météo très large (Barcelone, Zurich, Stockholm etc.).

Annexe C. Profils hebdomadaire des gains de chaleur internes au bâtiment

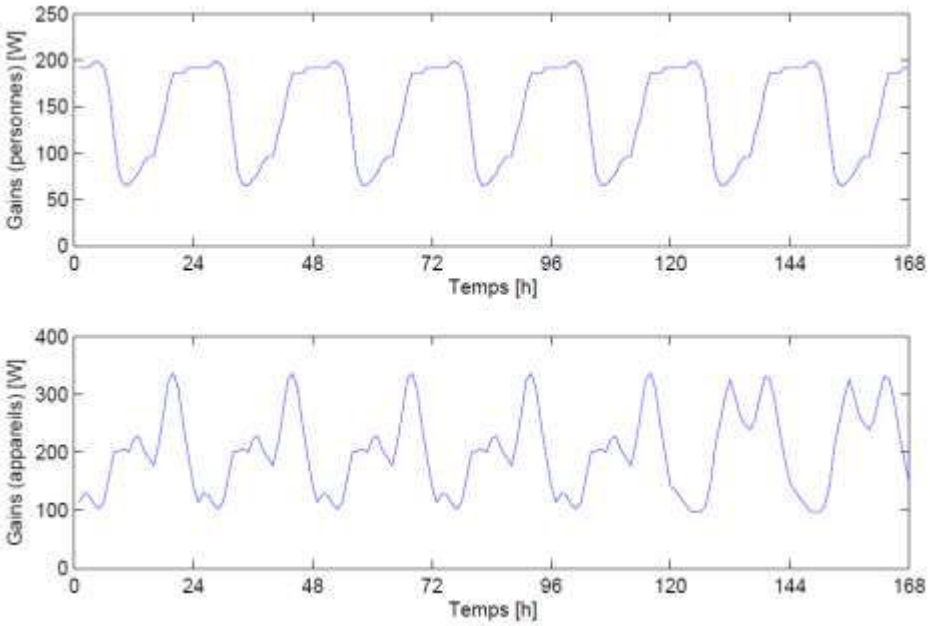

Figure C- 1 : Profil hebdomadaire des gains de chaleur internes au bâtiment dus à la présence des habitants et à l'utilisation d'appareils électriques : cas des SSC intégrant une machine à absorption ou une chaudière

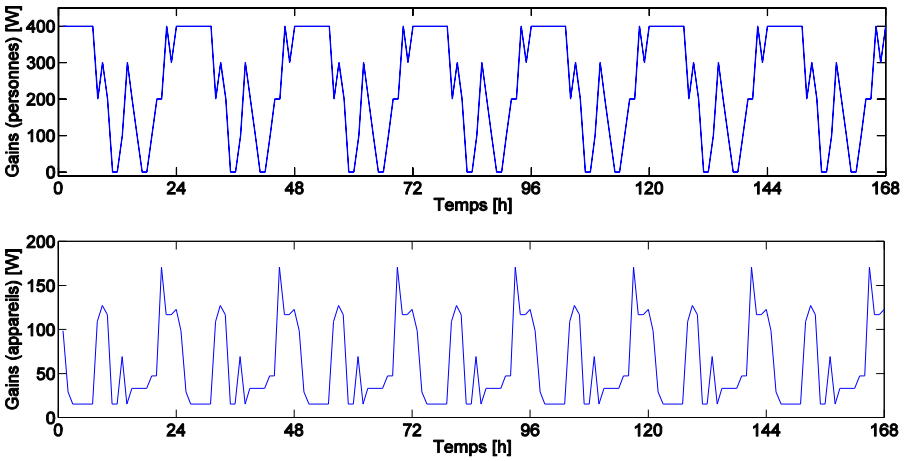

Figure C- 2 : Profil hebdomadaire des gains de chaleur internes au bâtiment dus à la présence des habitants et à l'utilisation d'appareils électriques : cas du SSC intégrant une PAC

Annexe D. Comparaison entre un entrainement en boucle ouverte et un entrainement en boucle fermée

Dans cette annexe sont comparés les résultats de modélisation du système présenté dans le deuxième chapitre selon les deux approches : en boucle fermée et en boucle ouverte. Les données d'apprentissage sont les mêmes que celles présentées dans ce même chapitre.

Plusieurs modèles ont été créés avec différents nombres de neurones dans la couche cachée, différents retards, mais avec une fonction d'activation linéaire en sortie. Pour chacune des deux catégories de modèles, entrainés en boucle ouverte ou fermée, nous avons sélectionnés deux modèles pour analyser les résultats conformément au processus présentés dans le deuxième chapitre (voir paragraphe II.3.5 et le début du paragraphe II.5.5.2).

Les résultats sont présentés dans les tableaux Tableau D- 1 et Tableau D- 2

Tableau D- 1 : Résultats des modèles ayant été entrainés en boucle ouverte

Modèles	NORM	NDN	TD	« *R2nrj* »	« *R2bar* »	« *BIC2* »	« *AIC2* »
RNA-R2nrj(linéaire)	±1,0	5,0	1,0	0,9978	0,9406	14,0674	14,0253
RNA-R2bar(linéaire)	±1,0	11,0	3,0	0,9765	0,9910	12,5218	12,4276
RNA-BIC2(linéaire)	±1,0	11,0	3,0	0,9765	0,9910	12,5218	12,4276
RNA-AIC2(linéaire)	±1,0	11,0	3,0	0,9765	0,9910	12,5218	12,4276

Tableau D- 2 : Résultats des modèles ayant été entrainés en boucle fermée

Modèles	NORM	NDN	TD	« *R2nrj* »	« *R2bar* »	« *BIC2* »	« *AIC2* »
RNA-R2nrj(linéaire)	±1,0	3,0	2,0	0,9941	0,9886	12,3332	12,3201
RNA-R2bar(linéaire)	±1,0	10,0	3,0	0,0001	0,9982	10,5370	10,5131
RNA-BIC2(linéaire)	±1,0	10,0	3,0	0,0001	0,9982	10,5370	10,5131
RNA-AIC2(linéaire)	±1,0	10,0	3,0	0,0001	0,9982	10,5370	10,5131

Concernant les meilleurs modèles au sens du « R2nrj » les performances des deux modèles sélectionnés entrainés en BF ou BO sont très proches avec un léger avantage du modèle entrainé en BO (0,9978>0,9941). Cependant, le modèle sélectionné selon les critères statistiques (calculés en en se basant uniquement sur les données d'apprentissage et en BF) et entrainé en BO est largement meilleur que son homologue entrainé en BF. Le pouvoir de généralisation de ce dernier est très faible. En effet, la valeur de son « R2nrj » est proche de zéro.