

TABLE DES MATIERES

TABLE DES MATIERES	i
NOMENCLATURE.....	iv
ACRONYMES	vi
LISTE DES FIGURES	vii
LISTE DES TABLEAUX	xi
INTRODUCTION GENERALE.....	1
CHAPITRE I : LES CYCLONES TROPICAUX.....	3
1.1- Introduction.....	4
1.2- Critères de formation d'un cyclone tropical	5
1.2.1- Les critères énergétiques.....	5
1.2.2- Les critères dynamiques.....	6
1.3- Structure des cyclones tropicaux	6
1.3.1- Equilibre dynamique	7
1.3.1.1- Equilibre cyclostrophique	7
1.3.1.2- Equilibre géostrophique	8
1.3.2- Convergence en basse troposphère	8
1.3.3- Profil du vent	9
1.3.4- Fonctionnement d'un cyclone : Hypothèse de Kerry Emanuel	10
1.4- Classification des cyclones tropicaux	13
1.4.1- Les différentes échelles	13
1.4.1.1- Échelle de Dvorak	13
1.4.1.2- Échelle de Saffir-Simpson	14
1.4.1.3- Échelle de Beaufort	15
1.4.2- Classification des systèmes cycloniques selon l'Organisation Météorologique Mondiale (O.M.M.)	16
1.4.3- Classification des systèmes cycloniques selon la Météo France Réunion	16
1.5- Prévision de l'intensité des cyclones tropicaux	17
1.5.1- Généralité sur la prévision des cyclones tropicaux	17
1.5.2- Quelques modèles de prévision de l'intensité cyclonique	18
1.6- Le bassin Sud-Ouest de l'Océan Indien	18
Chapitre II : LIAISON ENTRE LES VARIABLES - REGRESSION LINEAIRE	
MULTIPLE	21
2.1- Introduction	22
2.2- Liaison entre les variables	22
2.2.1- Corrélation linéaire	22
2.2.1.1- Définitions	22
2.2.1.2- Propriétés	22
2.2.1.3- Interprétation géométrique	23
2.2.2- Rapport de corrélation	23
2.2.2.1- Définition	23

2.2.2.2- Interprétation géométrique	23
2.2.3- Remarques	23
2.3- Régression linéaire multiple (RLM)	24
2.3.1- Définition	24
2.3.2- Estimation des paramètres	24
2.3.3- Estimation de la sortie du modèle et résidus	25
2.3.4- Coefficient de détermination	25
2.3.5- Tests sur les paramètres	26
CHAPITRE III : LES RESEAUX DE NEURONES	27
3.1- Introduction	28
3.2- Réseau de neurones biologiques	28
3.2.1- Le neurone biologique et sa constitution	29
3.2.2- Fonctionnement des neurones	30
3.2.3- Les synapses	31
3.2.4- La règle de Hebb	32
3.3- Réseau de neurones artificiels	32
3.3.1- Modèle d'un neurone	32
3.3.2- Différents types de réseaux	34
3.3.2.1- Réseaux de neurones non bouclés	34
3.3.2.2- Réseaux de neurones bouclés	34
3.3.3- Propriétés fondamentales des réseaux de neurones statiques	34
3.3.3.1- Les réseaux de neurones sont des approximatrices universelles	34
3.3.3.2- Les réseaux de neurones non linéaires par rapport à leurs paramètres sont des approximatrices parcimonieuses	35
3.3.4- Structure du réseau de neurones (nombre de couches, neurones cachés, fonction d'activation)	35
3.3.4.1- Nombre de couches	35
3.3.4.2- Nombre de neurones cachés	35
3.3.4.3- Fonction d'activation	35
3.3.5- Apprentissage	36
3.3.5.1- Principe de l'apprentissage	36
3.3.5.2- Apprentissage supervisé	36
3.3.5.3- Apprentissage non supervisé	37
3.3.5.4- Fonction de coût	37
3.3.5.5- Algorithmes d'optimisation	37
3.3.5.5.1- Principe	37
3.3.5.5.2- Méthode du gradient à pas constant	38
3.3.5.5.3- Méthode du gradient à pas variable	38
3.3.5.5.4- Méthode de Newton	38
3.3.5.5.5- Méthode de quasi-Newton	39
3.3.5.5.6- Méthode du gradient conjugué	40
3.3.5.5.7- Méthode de Levenberg-Marquardt	40
3.3.5.5.8- Remarques	40

TABLE DES MATIERES

3.3.6- Généralisation	41
3.3.6.1- Arrêt prématuré	41
3.3.6.2- Régularisation par modération des poids (Weight Decay)	41
CHAPITRE IV : PREVISION DES INTENSITES CYCLONIQUES PAR LES RESEAUX DE NEURONES	43
4.1- Sélection des prédicteurs	44
4.1.1- Introduction	44
4.1.2- Algorithme de Gram-Schmidt modifié	44
4.1.3- Méthode de sélection de prédicteurs	45
4.2- Application à la sélection des prédicteurs de l'intensité des cyclones tropicaux du BSOOI	46
4.2.1- Les données utilisées	46
4.2.2- Les prédicteurs initiaux	46
4.2.3- Choix du nombre de prédicteurs	47
4.2.4- Rangement des prédicteurs	48
4.2.5- Analyse des prédicteurs retenus	50
4.2.5.1- La vitesse maximale du vent (Vmax)	51
4.2.5.2- Changement d'intensités	53
4.2.5.3- Latitude	56
4.2.5.4- Variation des longitudes	58
4.2.5.5- Jour julien	61
4.2.5.6- Age	61
4.2.5.7- Vitesse de déplacement	64
4.3- Réseau de neurones utilisé	67
4.4- Performance du modèle	71
CONCLUSION GENERALE	87
BIBLIOGRAPHIES	89
PUBLICATIONS	
PUBLICATION A	
PUBLICATION B	
PUBLICATION C	

NOMENCLATURE

Lettres latines

$(O, \vec{e}_r, \vec{e}_\theta, \vec{e}_z)$: repère local cylindrique
\vec{V}	: vecteur vitesse du vent
V_r	: vitesse radiale du vent
V_θ	: vitesse tangentielle du vent
t	: temps
r	: distance au centre du cyclone
f	: force de Coriolis
p	: pression
r_{\max}	: rayon de vent maximum
V_{\max}	: vitesse maximale du vent
z	: altitude
h	: altitude du sommet de la couche limite atmosphérique
T_{out}	: température des particules au niveau de la tropopause
T_B	: température des particules au sommet de la couche limite atmosphérique
X, Y	: séries statistiques
$\text{Cov}(\cdot)$: covariance
$\text{Var}(\cdot)$: variance
$E(\cdot)$: espérance mathématique
a_0, a_1, \dots, a_p	: paramètres du modèle RLM
\hat{a}	: valeurs estimées des paramètres du modèle RLM
e_i	: terme d'erreur
$\text{Vect}(X)$: sous-espace engendré par les vecteurs colonnes de X
R^2	: coefficient de détermination
SSR	: somme des carrés de la régression
SST	: somme totale des carrés
SSE	: somme des carrés des résidus
b	: biais dans le RNA
w_i	: poids des connexions dans le RNA
na	: niveau d'activation d'un neurone
N_s	: nombre de neurones de sortie
$s(i)$: sortie du modèle RNA
$d(i)$: sortie désirée du modèle RNA

$J^i(w)$: fonction du cout pour l'exemple i en fonction du poids de connexion
$J(w)$: moyenne des fonctions du cout sur tous les exemples
w_t	: vecteur des paramètres du RNA à l'itération t
D_t	: direction de descente de gradient
H_t	: matrice du Hessien à l'itération t
M_t	: valeur approchée de la matrice du Hessien à l'itération t
I	: matrice identité
V_{depl}	: vitesse de déplacement du cyclone
V_{max}	: vitesse maximale du vent du cyclone
Lat	: latitude
Lon	: longitude
$DiffLat$: différence des latitudes du cyclone
$DiffLon$: différence des longitudes du cyclone
$DiffVmax$: différence des vitesses maximales du vent du cyclone

Lettres grecques

λ	: angle azimutal
δQ	: chaleur fournie par l'environnement du cyclone, sous forme d'entropie humide
δW	: travail produit par le cyclone, sous forme d'énergie cinétique
ρ_{XY}	: coefficient de corrélation linéaire entre X et Y
σ_x	: écart-type de la série X
σ_y	: écart-type de la série Y
$\eta_{Y X}^2$: rapport de corrélation de Y sur X
σ_e^2	: paramètre de nuisance
μ	: pas d'itération
∇J	: gradient de la fonction du cout

ACRONYMES

RNA	: Réseau de Neurones Artificiels
USA	: United States of America
ZCIT	: Zone de Convergence Inter Tropicale
CT	: Cyclone Tropical
TSM	: Température à la Surface de la Mer
CDO	: Central Dense Overcast
CCC	: Central Cold Cover
CI	: Current Intensity
FI	: Forecast Intensity
CMRS	: Centre Météorologique Régional Spécialisé
TCWC	: Tropical Cyclone Warning Centres
OMM	: Organisation Météorologique Mondiale
SHIFOR	: Statistical Hurricane Intensity FOREcast
SHIPS	: Statistical Hurricane Intensity Prediction Scheme
DSHP	: Decay SHIP
RLM	: Régression Linéaire Multiple
BSOOI	: Bassin Sud-Ouest de l'Océan Indien
GSM	: Gram-Schmidt Modifiée
CGS	: Classical Gram-Schmidt
MGS	: Modified Gram-Schmidt
JTWC	: Joint Typhoon Warning Center
US	: United States
PMC	: Perceptron MultiCouches

LISTE DES FIGURES

<i>Figure 1.1</i>	: Trajectoire des particules de vent en surface pour un cyclone de l'hémisphère Sud.	9
<i>Figure 1.2</i>	: Profil du vent du cyclone Alicia, enregistré par le service météo des USA, à bord d'un avion volant à 1500m.	10
<i>Figure 1.3</i>	: Représentation schématique du cycle de Carnot dans un cyclone tropical. La source chaude ($\delta Q > 0$) est représentée par le sommet de la couche limite atmosphérique ($z = h$) et la source froide ($\delta Q < 0$) est représentée par la tropopause.	11
<i>Figure 1.4</i>	: Evolution de l'erreur de prévision (a) de la position (en miles nautiques) et (b) de l'intensité (en noeuds) des cyclones tropicaux de l'Atlantique concernant la prévision opérationnelle officielle du National Hurricane Center aux Etats-Unis.	17
<i>Figure 1.5</i>	: Carte représentant le bassin Sud-Ouest de l'Océan Indien (BSOOI).	18
<i>Figure 1.6</i>	: Profils saisonniers de l'activité cyclonique (tempêtes et cyclones tropicaux) dans le bassin Sud-Ouest de l'Océan Indien	19
<i>Figure 3.1</i>	: Représentation d'un neurone biologique.	29
<i>Figure 3.2</i>	: Réseau de neurones biologique	30
<i>Figure 3.3</i>	: Représentation du corps cellulaire comme sommateur à seuil.	31
<i>Figure 3.4</i>	: Modèle d'un neurone artificiel	32
<i>Figure 3.5</i>	: Structure typique d'un RNA avec deux couches cachées.	33
<i>Figure 4.1</i>	: Choix du nombre de prédicteurs.	48
<i>Figure 4.2</i>	: Corrélations entre intensité à 0h et intensités de 12h à 72h.	51
<i>Figure 4.3</i>	: Corrélations entre intensité à -6h et intensités de 12h à 72h.	52
<i>Figure 4.4</i>	: Corrélations entre intensité à -24h et intensités de 12h à 72h.	53
<i>Figure 4.5</i>	: Exemple d'allure de la variation de l'intensité du cyclone en fonction du temps.	54
<i>Figure 4.6</i>	: Corrélations entre Différence des intensités à 0h / -6h et intensités de 12h à 72h.	55
<i>Figure 4.7</i>	: Corrélations entre Différence des intensités à 0h / -12h et intensités de 12h à 72h.	55

<i>Figure 4.8</i>	: Corrélations entre Différence des intensités à 0h / -24h et intensités de 12h à 72h.	56
<i>Figure 4.9</i>	: Corrélations entre Latitude à -6h et intensités de 12h à 72h.	57
<i>Figure 4.10</i>	: Corrélations entre Latitude à 0h et intensités de 12h à 72h.	58
<i>Figure 4.11</i>	: Corrélations entre Différence longitudes 0h/-6h et intensités de 12h à 72h.	59
<i>Figure 4.12</i>	: Corrélations entre Différence longitudes 0h/-12h et intensités de 12h à 72h.	59
<i>Figure 4.13</i>	: Corrélations entre Différence longitudes 0h/-18h et intensités de 12h à 72h.	60
<i>Figure 4.14</i>	: Corrélations entre Différence longitudes 0h/-24h et intensités de 12h à 72h.	60
<i>Figure 4.15</i>	: Corrélations entre Jour et intensités de 12h à 72h.	61
<i>Figure 4.16</i>	: Intensité à 72h en fonction de l'âge du phénomène cyclonique.	62
<i>Figure 4.17</i>	: Effectifs (en haut) et écart-types (en bas) des intensités à 72h pour chaque âge.	63
<i>Figure 4.18</i>	: Corrélations entre Age 0h et intensités de 12h à 72h.	64
<i>Figure 4.19</i>	: Corrélations entre vitesse de déplacement 0h/-6h et intensités de 12h à 72h.	65
<i>Figure 4.20</i>	: Corrélations entre vitesse de déplacement -6h/-12h et intensités de 12h à 72h.	66
<i>Figure 4.21</i>	: Corrélations entre vitesse de déplacement -12h/-18h et intensités de 12h à 72h.	66
<i>Figure 4.22</i>	: Réseau de neurones à une seule couche cachée	67
<i>Figure 4.23</i>	: Erreurs de prévision à 12h en fonction du nombre de prédicteurs et du nombre de neurones cachés	68
<i>Figure 4.24</i>	: Erreurs de prévision à 24h en fonction du nombre de prédicteurs et du nombre de neurones cachés	68
<i>Figure 4.25</i>	: Erreurs de prévision à 36h en fonction du nombre de prédicteurs et du nombre de neurones cachés	69
<i>Figure 4.26</i>	: Erreurs de prévision à 48h en fonction du nombre de prédicteurs et du nombre de neurones cachés	69

<i>Figure 4.27</i>	<i>: Erreurs de prévision à 60h en fonction du nombre de prédicteurs et du nombre de neurones cachés</i>	70
<i>Figure 4.28</i>	<i>: Erreurs de prévision à 72h en fonction du nombre de prédicteurs et du nombre de neurones cachés</i>	70
<i>Figure 4.29</i>	<i>: Comparaison des erreurs moyennes de prévision de l'intensité cyclonique utilisant les méthodes RLM (Régression Linéaire Multiple) et RNA (Réseau de Neurones Artificiels).</i>	71
<i>Figure 4.30</i>	<i>: Variations de la différence entre l'intensité observée et l'intensité calculé en fonction l'intensité observée pour la prévision à 12h</i>	72
<i>Figure 4.31</i>	<i>: Variations de la différence entre l'intensité observée et l'intensité calculé en fonction l'intensité observée pour la prévision à 24h</i>	72
<i>Figure 4.32</i>	<i>: Variations de la différence entre l'intensité observée et l'intensité calculé en fonction l'intensité observée pour la prévision à 36h</i>	73
<i>Figure 4.33</i>	<i>: Variations de la différence entre l'intensité observée et l'intensité calculé en fonction l'intensité observée pour la prévision à 48h</i>	73
<i>Figure 4.34</i>	<i>: Variations de la différence entre l'intensité observée et l'intensité calculé en fonction l'intensité observée pour la prévision à 60h</i>	74
<i>Figure 4.35</i>	<i>: Variations de la différence entre l'intensité observée et l'intensité calculé en fonction l'intensité observée pour la prévision à 72h</i>	74
<i>Figure 4.36</i>	<i>: Septième exemple de prévision. L'intensité présente une oscillation</i>	75
<i>Figure 4.37</i>	<i>: Huitième exemple de prévision. L'intensité diminue brusquement puis augmente.</i>	75
<i>Figure 4.38</i>	<i>: Répartition des erreurs absolues pour la prévision à 12h.</i>	76
<i>Figure 4.39</i>	<i>: Répartition des erreurs absolues pour la prévision à 24h.</i>	77
<i>Figure 4.40</i>	<i>: Répartition des erreurs absolues pour la prévision à 36h.</i>	77
<i>Figure 4.41</i>	<i>: Répartition des erreurs absolues pour la prévision à 48h.</i>	78
<i>Figure 4.42</i>	<i>: Répartition des erreurs absolues pour la prévision à 60h.</i>	78
<i>Figure 4.43</i>	<i>: Répartition des erreurs absolues pour la prévision à 72h.</i>	79
<i>Figure 4.44</i>	<i>: Variations de la différence entre l'intensité observée et l'intensité calculé en fonction l'intensité observée pour la prévision à 12h</i>	81
<i>Figure 4.45</i>	<i>: Variations de la différence entre l'intensité observée et l'intensité calculé en fonction l'intensité observée pour la prévision à 24h</i>	82

LISTE DES FIGURES

<i>Figure 4.46</i>	<i>: Variations de la différence entre l'intensité observée et l'intensité calculé en fonction l'intensité observée pour la prévision à 36h</i>	83
<i>Figure 4.47</i>	<i>: Variations de la différence entre l'intensité observée et l'intensité calculé en fonction l'intensité observée pour la prévision à 48h</i>	84
<i>Figure 4.48</i>	<i>: Variations de la différence entre l'intensité observée et l'intensité calculé en fonction l'intensité observée pour la prévision à 60h</i>	85
<i>Figure 4.49</i>	<i>: Variations de la différence entre l'intensité observée et l'intensité calculé en fonction l'intensité observée pour la prévision à 72h</i>	86

LISTE DES TABLEAUX

<i>Tableau 1.1</i>	<i>: Activité cyclonique sur les différents bassins océaniques.</i>	4
<i>Tableau 1.2</i>	<i>: Echelle de Dvorak utilisée par tous les centres de prévision cyclonique pour estimer l'intensité des cyclones tropicaux.</i>	14
<i>Tableau 1.3</i>	<i>: Echelle de Saffir-Simpson utilisée par le CMRS de Miami pour estimer les dégâts associés à des ouragans classés de 1 à 5. Le vent moyen est mesuré sur une minute.</i>	15
<i>Tableau 1.4</i>	<i>: Classification des CT en fonction de l'intensité des vents comparée à l'échelle historique de Beaufort.</i>	15
<i>Tableau 4.1</i>	<i>: Les prédicteurs initiaux.</i>	47
<i>Tableau 4.2</i>	<i>: Liste des prédicteurs classés par ordre de pertinence selon MGS.</i>	49
<i>Tableau 4.3</i>	<i>: Récapitulation des prédicteurs sélectionnés et leurs rangs.</i>	50
<i>Tableau 4.4</i>	<i>: Récapitulation des nombres de prédicteurs et de neurones cachés optimum.</i>	71

INTRODUCTION GENERALE

« Il faudrait ne devenir plus savant que pour changer la vie et augmenter l'espérance »

Jean Guéhenno.

Dans le contexte de l'amélioration de la vie humaine, les scientifiques ne cessent d'explorer des méthodes bien fondées pour résoudre les problèmes. Cependant, on est toujours en train de bien comprendre les phénomènes naturels, de les modéliser afin de prévoir leurs évolutions dans le temps et dans l'espace.

Le mois de Novembre 2013, le typhon Haiyan, l'un des typhons les plus puissants de l'histoire, a balayé les Philippines avec des vents de plus de 300 km/h qui ont fait plus de 1200 morts, selon La Croix-Rouge. Les dégâts produits par ces phénomènes violents restent toujours gravés dans la mémoire de l'humanité.

Des scientifiques cherchaient des moyens pour détruire, atténuer ou au moins dévier les cyclones. La première tentative fut élaborée en 1947 et intitulée "Project Cirrus". L'armée des USA a lâché 35kg de dioxyde de carbone solide émiété dans une tornade. La tornade changea alors de direction, mais atterri à côté de la ville de Savannah, en Géorgie ; il est alors possible de dévier un phénomène cyclonique mais la maîtrise de la nouvelle trajectoire est très difficile.

Toutefois, cette tentative a été reprise onze années plus tard, sous le nom du projet « Stormfury », dans le but de détruire les cyclones alors qu'ils sont encore faibles.

Ce projet se basait sur l'hypothèse qu'en gelant l'eau contenu dans le cyclone grâce à de l'iodure d'argent, cela perturberait la structure interne du cyclone, ce qui entraînerait sa destruction. Ainsi, de nombreux avions furent envoyés afin d'ensemencer les cyclones de l'océan atlantique. Le premier cyclone ensemencé fut le cyclone Beulah. Cela entraîna une baisse des vents de 20%, et un agrandissement de l'œil du cyclone. Cette expérience dut être renouvelée de nombreuses fois, cependant les conditions ne furent pas toujours avantageuses.

Ces expériences ont été poursuivies jusqu'en 1980 où les scientifiques ont été persuadé que les cyclones ont un caractère aléatoire, donc sont très difficiles à maîtriser. Le coût du projet est énorme et la pollution de l'environnement due à l'utilisation des produits chimiques est inimaginable. De plus, les scientifiques comprenaient que le

cyclone est un phénomène indispensable à la régularisation de la température de notre planète. Sa destruction n'est pas nécessaire mais la seule solution est l'augmentation de la qualité des prévisions.

Notre objectif est de construire un modèle capable d'avoir un meilleur résultat sur les prévisions de l'intensité cyclonique. Nous sommes du côté des scientifiques qui pensent que l'exploitation de la mémoire des bassins océaniques en apportera beaucoup d'amélioration. A partir des données statistiques, nous utilisons les réseaux de neurones artificiels (RNA) qui montrent sa réputation dans plusieurs domaines. Ce travail veut explorer la capacité des RNA de résoudre les problèmes non linéaires très complexes comme les cyclones.

Les cyclones tropicaux seront présentés dans le premier chapitre. Il est très important de connaître leur formation et leur structure pour la construction de l'ensemble initial des prédicteurs. C'est à partir de ce dernier que nous allons sélectionner les prédicteurs de l'intensité du cyclone.

La liaison entre variables sera présentée dans la première partie du chapitre II. C'est une partie très importante car le choix des prédicteurs dépend de la liaison considérée. Nous développons dans la deuxième partie de ce chapitre la méthode de régression linéaire multiple, qui consiste à créer un modèle linéaire à partir d'un ensemble de prédicteurs préalablement choisi.

Le troisième chapitre est constitué de la théorie sur les réseaux de neurones. Les découvertes au niveau du réseau de neurones biologique aident beaucoup au perfectionnement du réseau de neurones artificiel. La détermination de l'architecture du RNA à utiliser demande beaucoup d'attention car une petite modification peut changer énormément les résultats.

La sélection des prédicteurs pour la prévision de l'intensité cyclonique sera exposée dans le dernier chapitre, suivie des tests de performance du modèle RNA. On doit retenir les variables les plus corrélées avec la variable à expliquer et les moins corrélées entre elles. Des modèles de prévision seront créés à partir de ces prédicteurs pour en choisir, après, ceux qui présentent des erreurs moyennes minimales.

CHAPITRE I : LES CYCLONES TROPICAUX

1.1- Introduction

Le mot cyclone est en rapport direct avec sa forme presque circulaire qui signifie « *kyklos* » en grec. Particulièrement, les cyclones tropicaux sont des phénomènes qui naissent au-dessus des mers tropicales et se développent en puisant leur énergie dans la chaleur latente de condensation de l'eau. Ce qui caractérise les cyclones tropicaux est le vent violent circulant autour d'une zone à basse pression. Selon l'Organisation Météorologique Mondiale, en 1992, ce phénomène est « une perturbation dépressionnaire d'échelle synoptique non accompagnée d'un système frontal, prenant naissance au-dessus des eaux tropicales et présentant une activité convective organisée et une circulation cyclonique, plus intense en surface qu'en altitude ».

La dénomination des cyclones tropicaux dépend des bassins océaniques où ils se développent : Ouragan sur l'Atlantique Nord, les Caraïbes et le Pacifique Nord-Est ; Typhon dans le Pacifique Nord-Ouest et Cyclone dans le bassin de l'Océan Indien, l'Australie et le Pacifique Sud.

Les cyclones tropicaux sont parmi les phénomènes physiques qui existent toujours chaque année. En moyenne, ils se produisent 80 à 85 phénomènes cycloniques chaque année sur le globe, d'après les études statistiques faites depuis une vingtaine d'années [1,2,3,4] (*Tableau 1.1*).

Tableau 1.1 : Activité cyclonique sur les différents bassins océaniques.

Source : www.cyclonextreme.com

Bassin océanique	Nombre moyen de systèmes	Nombre moyen d'ouragan ou cyclone ou typhon
Atlantique	9,4 (11%)	5,4 (12%)
Pacifique Nord-Est	16,5 (20%)	8,9 (20%)
Pacifique Nord-Ouest	25,7 (31%)	16 (36%)
Océan Indien Nord	5,4 (6%)	2,5 (6%)
Océan Indien Sud-Ouest	10,4 (12%)	4,4 (10%)
Océan Indien Sud-Est	6,9 (8%)	3,4 (8%)
Australie-Pacifique Sud	9 (11%)	4 (9%)

1.2- Critères de formation d'un cyclone tropical

La naissance des cyclones se fait toujours à la surface de l'océan, à partir de perturbations pluvio-orageuses, dont la taille peut varier de quelques dizaines à plusieurs centaines de kilomètres. Ces amas se forment puis se dissipent au cours d'une période de temps pouvant aller de quelques heures à une journée environ [5,6,7]. La persistance de leur activité convective permet à ces systèmes de s'évoluer en perturbations cycloniques.

Il existe sur le globe sept zones susceptibles de réunir les conditions nécessaires à la formation d'un cyclone : le Pacifique Nord-Ouest, le Nord de l'Océan Indien, l'Atlantique-Nord, le Pacifique Nord-Est, le Pacifique Sud-Ouest, le Sud-Est de l'Océan Indien (ou le Nord-Ouest de l'Australie) et le Sud-Ouest de l'Océan Indien.

Des analyses composites de certains paramètres météorologiques et océanographiques [1,8,9] ont permis de mettre en évidence 7 critères d'ordre énergétiques (ou climatologiques) et dynamiques.

1.2.1- Les critères énergétiques

- La température de la mer doit être supérieure à 26,5°C, sur une profondeur d'au moins 60 mètres, dans la zone où règne une dépression. La quantité de chaleur fournie par la mer va alors alimenter le système. Et plus la température de la surface de la mer est élevée, plus l'énergie disponible pour le développement du cyclone est importante. La température élevée de l'océan permettra une évaporation intense. La chaleur et l'humidité de l'océan opèreront un mouvement vers l'atmosphère. Ce transfert sera à l'origine du développement de cumulonimbus qui favorisent l'apparition d'une dépression tropicale, il sera le moteur thermique de la dépression.
- Entre la surface et les couches moyennes de l'atmosphère, il doit y avoir une instabilité convective potentielle suffisamment élevée pour permettre aux particules de monter facilement vers les hautes couches de l'atmosphère.
- L'humidité relative en moyenne troposphère doit être suffisamment élevée. La cyclogenèse est impossible si le taux d'humidité est inférieur à 40%. La formation des cyclones est fréquente lorsque le taux d'humidité atteint 70%. En effet, en

absence d'humidité, les cumulonimbus ne se forment pas, or ce type de nuage est l'une des composantes essentielles du phénomène cyclonique.

En plus de ces conditions énergétiques, la formation d'un cyclone a besoin d'autres critères dynamiques indispensables.

1.2.2- Les critères dynamiques

- La force de Coriolis est nécessaire pour amorcer le mouvement tourbillonnaire du système. Cette force, qui s'annule à l'équateur, prend une valeur suffisante à partir des latitudes généralement supérieures à 5° ;
- Un faible cisaillement vertical du vent horizontal favorise le développement des systèmes convectifs en cyclones tropicaux [10]. Une grande différence entre les vents à la surface et aux couches élevées entraînerait une dissipation, en dehors du système, de l'énergie apportée en altitude par les courants ascendants. Une valeur inférieure à 10m.s^{-1} de la surface au niveau 100hPa est admise [11,12] ;
- Une divergence ou une évacuation généralisée au niveau des couches supérieures de la troposphère, permettant ainsi au flux vertical de s'échapper, ce qui induit la baisse du champ de pression au centre de la perturbation. Ce phénomène est nécessaire à l'entretien du système.
- La préexistence d'une perturbation (amas nuageux par exemple) près de la surface avec un tourbillon relatif positif et une convergence suffisante des flux. Ces amas nuageux se trouvent entre les tropiques au niveau d'une vaste zone de mauvais temps (ZCIT : Zone de Convergence Inter Tropicale).

Il est très important de remarquer que si ces conditions sont indispensables à la naissance d'un cyclone tropical, elles le sont aussi pour son développement. Ces critères ayant donc des rapports directs avec l'intensité cyclonique, le choix des prédicteurs à utiliser dans notre modèle de prévision d'intensité.

1.3- Structure des cyclones tropicaux

Les cyclones tropicaux présentent trois parties distinctes généralement bien visibles :

- L'œil : c'est une zone « calme » de diamètre, généralement, entre 10 et 100km au milieu de la tempête ;

- Le mur de l'œil, situé à quelques dizaines de km du centre, est constitué d'une ceinture de cumulonimbus qui peut s'étendre sur un rayon de 10 à 50km de large;
- Des bandes spirales (5 à 50km de large) convectives et convergentes qui s'enroulent de l'extérieur vers le mur de l'œil.

1.3.1- Equilibre dynamique

Le système de coordonnées le mieux adapté à l'étude dynamique du phénomène cyclonique est le système de coordonnées cylindriques ($O, \vec{e}_r, \vec{e}_\theta, \vec{e}_z$), à cause des mouvements en spiral des particules de vent. O désigne le centre du CT.

La vitesse horizontale peut alors s'écrire :

$$\vec{V} = V_r \vec{e}_r + V_\theta \vec{e}_\theta = \frac{dr}{dt} \vec{e}_r + r \frac{d\lambda}{dt} \vec{e}_\theta \quad (1.1)$$

où V_r la vitesse radiale, V_θ la vitesse tangentielle, t le temps, r la distance au centre du CT et λ un angle.

L'équation du mouvement radial, en absence de frottements s'écrit :

$$\frac{\partial V_r}{\partial t} = \frac{V_\theta^2}{r} + fV_\theta - \frac{1}{\rho} \frac{\partial p}{\partial r} \quad (1.2)$$

où f la force de Coriolis, p la pression et ρ la masse volumique de l'air.

En régime permanent,

$$\frac{V_\theta^2}{r} + fV_\theta - \frac{1}{\rho} \frac{\partial p}{\partial r} = 0 \quad (1.3)$$

1.3.1.1- Equilibre cyclostrophique

La force de Coriolis f devient négligeable devant les autres termes dans l'équation (1.3), lorsqu'on est à l'intérieur du mur de l'œil ($r < 20$ à 40 km). L'équilibre qui existe entre le gradient de pression et la force centrifuge s'appelle « équilibre cyclostrophique » :

$$\frac{V_\theta^2}{r} = \frac{1}{\rho} \frac{\partial p}{\partial r} \quad (1.4)$$

Cet équilibre cyclostrophique peut aussi être atteint lorsque le cyclone se trouve entre les latitudes 5° et 10°, quelle que soit la distance du centre du cyclone considérée [13]. En effet, la force de Coriolis est une fonction croissante de la latitude et sa valeur est toujours négligeable par rapport à la force centrifuge pour les latitudes inférieures à 10°.

1.3.1.2- Equilibre géostrophique

La force centrifuge devient négligeable lorsqu'on se place à l'extérieur du mur de l'œil. C'est l'équilibre entre le gradient de pression et la force de Coriolis qu'on appelle équilibre géostrophique :

$$fV_\theta = \frac{1}{\rho} \frac{\partial p}{\partial r} \quad (1.5)$$

1.3.2- Convergence en basse troposphère

Dans la couche limite atmosphérique, la force de frottement agit sur la vitesse du vent, particulièrement sur la composante V_θ . Cela diminue les valeurs de la force centrifuge V_θ^2/r et de la force de Coriolis fV_θ , qui rend dominant le gradient de pression $(1/\rho)(\partial p/\partial r)$ dans l'équation (1.2). L'accélération normale devient négative, donc la vitesse radiale V_r diminue (en valeur algébrique). Ceci correspond à une vitesse radiale centripète, donc les trajectoires des particules sont en spirales [14] (*Figure 1.1*). Cette convergence en basse couche est suivie d'une ascendance dans le mur de l'œil et une divergence en haute troposphère.

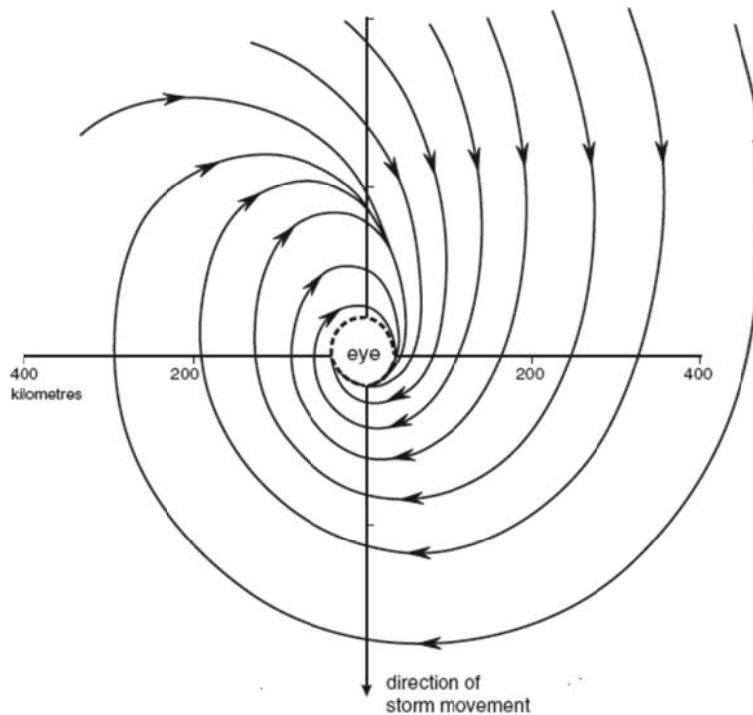


Figure 1.1 : Trajectoire des particules de vent en surface pour un cyclone de l'hémisphère Sud. Extrait de [14]

1.3.3- Profil du vent [13]

La force du vent (composantes tangentielle et radiale) dépend de la zone du cyclone considérée (*Figure 1.2*) :

- entre le centre du cyclone et le mur de l'œil, le vent augmente d'une façon quasi-linéaire que l'on peut approcher par la formule :

$$V(r) = \frac{r}{r_{\max}} V_{\max} \quad (1.6)$$

- dans le mur de l'œil, le vent atteint son maximum d'intensité V_{\max} à une distance r_{\max} appelée « rayon de vent maximum » ;
- au-delà du rayon de vent maximum, le vent diminue plus lentement suivant une loi du type :

$$V(r) = \left(\frac{r_{\max}}{r} \right)^{\alpha} V_{\max} \quad (1.7)$$

avec $\alpha < 1$.

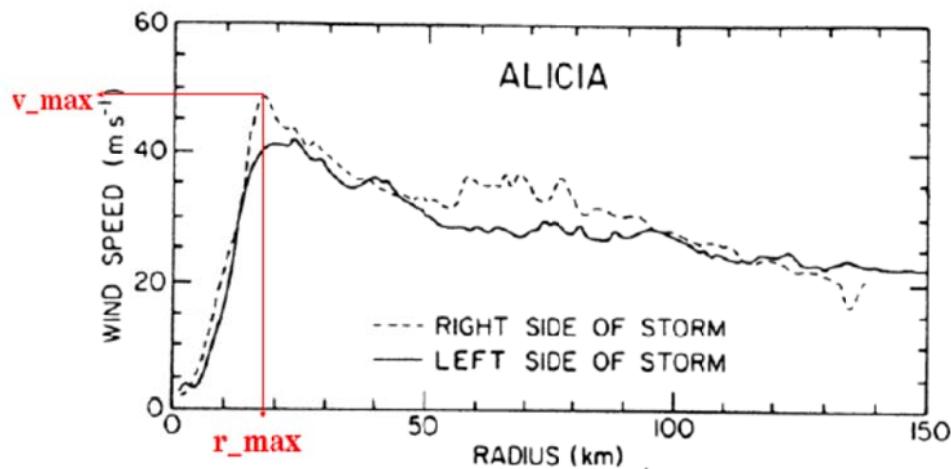


Figure 1.2 : Profil du vent du cyclone Alicia, enregistré par le service météo des USA, à bord d'un avion volant à 1500m. Extrait de [13].

1.3.4- Fonctionnement d'un cyclone : Hypothèse de Kerry Emanuel

Un cyclone tropical peut être considéré, en régime stationnaire, comme une machine thermique alimentée par l'énergie calorifique fournie par l'océan à travers les flux de chaleur sensibles et de chaleur latente échangés à l'interface océan-atmosphère [16]. Ces hypothèses, émises en 1986 par Kerry Emanuel, supposent que les particules suivent des trajectoires formant un cycle thermodynamique, fonctionnant avec deux sources de chaleur (source chaude et source froide).

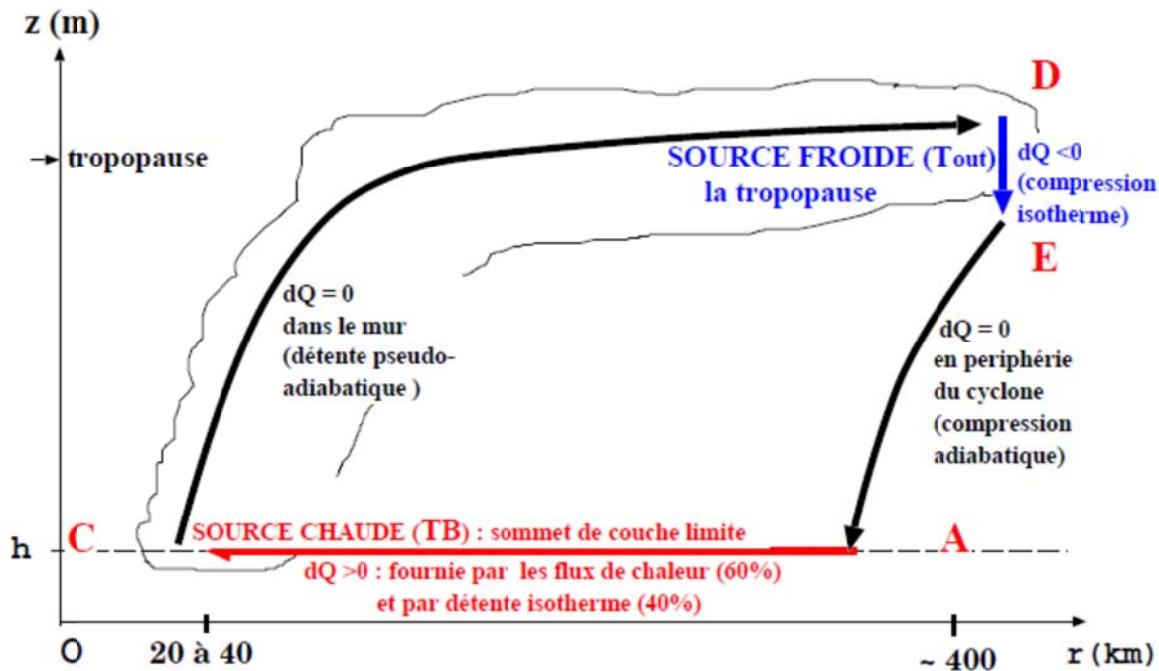


Figure 1.3 : Représentation schématique du cycle de Carnot dans un cyclone tropical. La source chaude ($\delta Q > 0$) est représentée par le sommet de la couche limite atmosphérique ($z = h$) et la source froide ($\delta Q < 0$) est représentée par la tropopause.

Extrait de [13]

Ce cycle, défini par deux isothermes ($\delta T = 0$) et deux adiabatiques ($\delta Q = 0$), produit une puissance motrice. La Figure 1.3 montre les quatre zones décrivant le cycle de Carnot associé au cyclone tropical.

Zone AC:

Lors de la convergence des particules de la périphérie (A) vers le centre du cyclone (C), il y a, d'une part, transfert d'énergie de l'océan vers l'atmosphère (sous forme de chaleur latente de vaporisation, la teneur en vapeur d'eau augmente ; mais aussi par détente isotherme) et d'autre part, transfert d'énergie de l'atmosphère vers l'océan (friction du vent sur la surface de l'océan).

Zone CD:

Au sein du mur de l'œil, les particules s'élèvent en suivant une trajectoire tendant vers l'extérieur due à la conservation du moment angulaire. De plus, elles conservent leur température pseudo-potentielle équivalente et n'échangent pas de chaleur avec l'extérieur (évolution pseudo-adiabatique).

Zone DE:

En haute troposphère, les particules d'air divergent et à quelques centaines de km du centre du cyclone, elles subsident à température constante (compression isotherme car le chauffage par compression est équilibré par refroidissement radiatif). L'énergie gagnée en basse couche est perdu dans cette zone par rayonnement infrarouge.

Zone EA:

En descendant vers la surface, la compression des particules ne se fait plus à température constante à cause de l'importance du chauffage par compression adiabatique par rapport à la perte de chaleur par rayonnement. Il s'agit alors d'une compression adiabatique sèche.

La fermeture du cycle est assurée par la conservation du moment angulaire.

Le rendement ε d'un cycle thermodynamique est le rapport entre le travail produit par la machine thermique et la chaleur fournie par l'extérieur. Dans notre cas, la machine thermique est le cyclone et ce rendement ne dépend que de la température T_B du sommet de la couche limite et de la température T_{out} de la tropopause :

$$\varepsilon = \frac{|\delta W|}{\delta Q} = \frac{T_B - T_{out}}{T_B} \quad (1.8)$$

- δW : travail produit par le cyclone, sous forme d'énergie cinétique ;
- δQ : chaleur fournie par l'environnement du cyclone, sous forme d'entropie humide ;
- T_{out} : température des particules au niveau de la tropopause ;
- T_B : température des particules au sommet de la couche limite atmosphérique.

Un cyclone tropical fonctionne alors, en régime stationnaire, comme une machine transformant l'entropie humide fournie par l'environnement en énergie cinétique.

Cette théorie nous donne une idée sur l'importance de la température de la surface de la mer (TSM) et de la température de la tropopause dans la prévision de l'intensité cyclonique ; l'efficacité de notre machine thermique (le cyclone) dépend de ces deux températures.

Cependant, l'utilisation originale de cette théorie est la détermination du maximum d'intensité que peut atteindre chaque cyclone en fonction de la TSM et de la température de la tropopause, sans considérer les facteurs défavorables (cisaillement vertical de vent trop fort, absence de divergence d'altitude, tourbillon trop faible, ...) au développement du système.

1.4- Classification des cyclones tropicaux

1.4.1- Les différentes échelles

1.4.1.1- Échelle de Dvorak

L'échelle de Dvorak est une méthode initialement mise au point par Dvorak en 1974 [15] pour estimer, à partir des images satellites dans le canal visible, l'intensité et l'évolution des dépressions et cyclones tropicaux. La méthode fut améliorée en 1982 par l'utilisation d'images prises dans le canal infrarouge, disponibles de jour comme de nuit.

La méthode de Dvorak consiste à comparer les caractéristiques nuageuses du « système tropical » à des modèles prédefinis. Elle utilise plusieurs critères comme la structure des bandes spirales, l'œil, la configuration du cisaillement vertical du vent, la couverture nuageuse centrale dense (CDO : Central Dense Overcast), la couverture nuageuse centrale froide (CCC : Central Cold Cover).

Quand le cyclone possède un œil, on utilise l'image infrarouge pour calculer la différence de température entre l'œil qui est chaud et son environnement qui est froid, et plus cette différence est importante plus le cyclone est intense. Cette étude permet de calculer le nombre de Dvorak (CI : Current Intensity) et de faire une prévision d'intensité (FI : Forecast Intensity) du cyclone tropical à une échéance de 24 heures. La technique de Dvorak ayant ses limites de prévision, la pression estimée peut être assez différente de la pression mesurée en temps réel par des sondes lâchées en avion.

Cette méthode est utilisée par tous les centres de prévision cycloniques (6 CMRS : Centre Météorologique Régional Spécialisé et 5 TCWC : Tropical Cyclone Warning Centres). Il existe deux échelles de Dvorak différentes : une première pour les bassins cycloniques sous la surveillance des CMRS de Miami et Honolulu (où des mesures avion ont permis d'étalonner cette estimation), et une seconde pour tous les autres CMRS et TCWC.

Tableau 1.2 : Echelle de Dvorak utilisée par tous les centres de prévision cyclonique pour estimer l'intensité des cyclones tropicaux. Source : Météo-France.

Classification	Nombre de DVORAK CI	Vent soutenu sur 1 mn nœuds	Pression centrale minimale estimée (hPa)	
			Bassins cycloniques sous surveillance des CMRS de Miami et Honolulu	Bassins cycloniques sous surveillance des autres CMRS et TCWC
Perturbation tropicale	1	25	> 1009	> 1002
Perturbation tropicale	2	30	1009	1000
Dépression tropicale	2,5	35	1005	997
CT de classe 1	3,5	64	994 – 987	980 – 976
CT de classe 2	4,5	83	979 – 970	966 – 958
CT de classe 3	5	96	970 – 960	954 – 944
CT de classe 4	5,5	102	960	941 – 928
CT de classe 5	6,5	127	935	910
CT de classe 5	7	140	921	898
CT de classe 5	8	170	890	858

Le *Tableau 1.2* nous indique que pour un même nombre de Dvorak CI, la pression estimée au centre du cyclone varie selon les bassins cycloniques car elle dépend de la climatologie des températures de surface de la mer et des conditions synoptiques. Les estimations de vent et de pression sont calculées à partir de relations empiriques qui lient de façon constante les deux variables. D'autres facteurs comme la taille des cyclones, leur asymétrie, l'évolution interne, les conditions environnementales, l'arrivée sur les terres émergées, modulent également l'intensité des vents et des pluies.

1.4.1.2- Échelle de Saffir-Simpson

L'échelle de Saffir-Simpson qui n'est utilisée que par le CMRS de Miami classe les ouragans de l'Atlantique Nord et du Pacifique Nord-Est en 5 catégories en fonction de leur violence (*Tableau 1.3*). Elle donne également une estimation des dégâts et des inondations le long des côtes en fonction de l'intensité de l'ouragan.

Tableau 1.3 : Echelle de Saffir-Simpson utilisée par le CMRS de Miami pour estimer les dégâts associés à des ouragans classés de 1 à 5. Le vent moyen est mesuré sur une minute. Source : <http://www.nhc.noaa.gov>

Echelle Saffir-Simpson	Pression mesurée (hPa)	Vent en nœuds	Dégâts
Ouragan de classe 1	> 980	64-82	Faibles
Ouragan de classe 2	979-965	83-95	Modérés
Ouragan de classe 3	964-945	96-113	Importants
Ouragan de classe 4	944-920	114-134	Très importants
Ouragan de classe 5	< 920	> 135	Désastreux

1.4.1.3- Échelle de Beaufort

L'échelle de Beaufort constitue une excellente manière d'évaluer la force du vent par la simple observation de son effet. Il n'est pas crucial de savoir sa vitesse exacte, il est par contre très utile de savoir reconnaître les échelons significatifs de sa force (*Tableau 1.4*).

Tableau 1.4 : Classification des CT en fonction de l'intensité des vents comparée à l'échelle historique de Beaufort.

Degré Beaufort	Terme descriptif	Vitesse moyenne en nœuds	Vitesse moyenne en km/h
0	Calme	<1	<1
1	Très légère brise	de 1 à 3	de 1 à 5
2	Légère brise	de 4 à 6	de 6 à 11
3	Petite brise	de 7 à 10	de 12 à 19
4	Jolie brise	de 11 à 16	de 20 à 28
5	Bonne brise	de 17 à 21	de 29 à 38
6	Vent frais	de 22 à 27	de 39 à 49
7	Grand frais	de 28 à 33	de 50 à 61
8	Coup de vent	de 34 à 40	de 62 à 74
9	Fort coup de vent	de 41 à 47	de 75 à 88
10	Tempête	de 48 à 55	de 89 à 102
11	Violente tempête	de 56 à 63	de 103 à 117
12	Ouragan	64 et plus	118 et plus

1.4.2- Classification des systèmes cycloniques selon l'Organisation Météorologique Mondiale (O.M.M.)

L'O.M.M. a défini 3 classes de systèmes en fonction de la vitesse du vent (vent le plus fort en valeur soutenue pendant 1 minute appelé vent maximum soutenu):

- **dépression tropicale** : si le vent est inférieur ou égal à force 7 Beaufort, inférieur ou égal à 61 km/h, ou inférieur ou égal à 33 nœuds ;
- **tempête tropicale** : si le vent est compris entre force 8 et force 11 Beaufort, ou entre 62 et 117 km/h, ou entre 34 ou 63 nœuds ;
- **cyclone** : si le vent est supérieur à force 12 Beaufort, ou supérieur à 117 km/h, ou supérieur à 63 nœuds.

1.4.3- Classification des systèmes cycloniques selon la Météo France Réunion

La Météo France Réunion a aussi sa propre classification utilisée pour le bassin Océan Indien Sud-Ouest :

- **zone perturbée** : zones de basses pressions sans véritable organisation ;
- **perturbation tropicale** : zone dépressionnaire présentant une convection renforcée et une circulation tourbillonnaire dans laquelle les vents n'excèdent pas 50 km/h (force 6 Beaufort) ;
- **dépression tropicale** : système où la convection est organisée et la circulation cyclonique caractérisée ; les vents près du centre sont compris entre 51 et 62 km/h (grand frais, force 7 Beaufort) et la pression supérieure à 996 hPa ;
- **tempête tropicale modérée** : système où les vents près du centre sont compris entre 63 et 88 km/h (coup de vent, force 8 ou 9 Beaufort) et la pression entre 986 et 995 hPa ;
- **forte tempête tropicale** : système où les vents près du centre sont compris entre 89 et 117 km/h (tempête ou forte tempête, force 10 ou 11 Beaufort) et la pression entre 985 et 971 hPa ;
- **cyclone tropical** : système où les vents près du centre sont compris entre 118 km/h (ouragan, force 12 Beaufort) et 165 km/h et la pression comprise entre 971 hPa et 941 hPa ;
- **cyclone tropical intense** : système où les vents près du centre sont compris entre 166 et 212 km/h et la pression entre 941 hPa et 910 hPa ;

- cyclone tropical très intense : système où les vents près du centre sont supérieurs à 212 km/h et la pression inférieure à 910 hPa.

1.5- Prévision de l'intensité des cyclones tropicaux

1.5.1- Généralité sur la prévision des cyclones tropicaux

L'amélioration de la qualité de prévision des caractéristiques des cyclones tropicaux est le meilleur moyen pour réduire le nombre de victimes sur leur passage. La connaissance de sa trajectoire quelques heures avant, par exemple, permet aux autorités de prendre des décisions importantes comme l'évacuation des habitants aux endroits plus sûrs. Faute de renseignements sur l'évolution du système cyclonique, on risque de perdre plusieurs vies humaines. Ce qui est le cas au Bangladesh en 1970, lors du passage du cyclone Bhola, avec un bilan humain estimé entre 300 000 et 500 000 morts.

Lié à la complexité du phénomène, la prévision de l'intensité des cyclones tropicaux est très difficile par rapport à celle de leur trajectoire. Une amélioration sur la prévision de la trajectoire est bien visible au cours des 30 dernières années (*Figure 1.4a*). Par contre, on n'observe qu'un faible progrès pour la prévision de l'intensité cyclonique (*Figure 1.4b*).

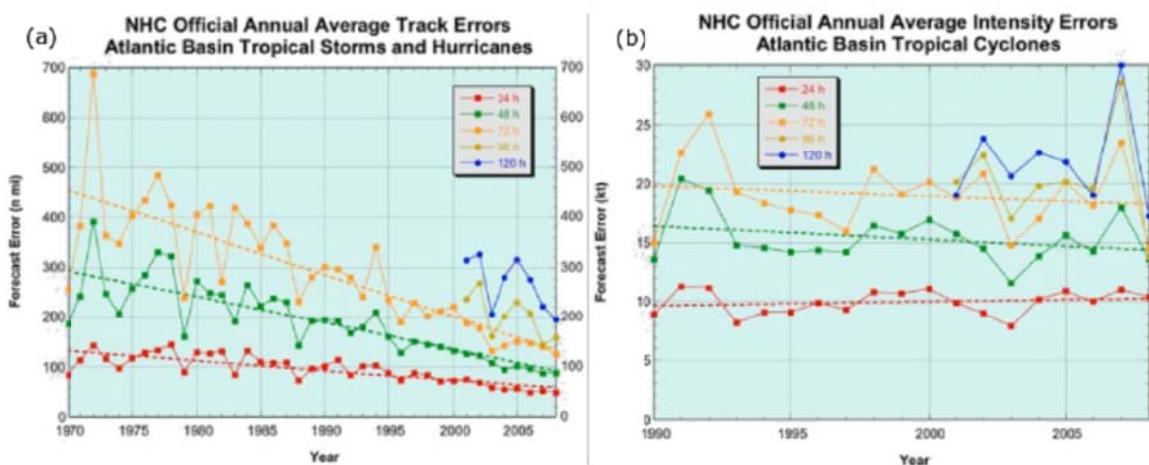


Figure 1.4 : Evolution de l'erreur de prévision (a) de la position (en miles nautiques) et (b) de l'intensité (en noeuds) des cyclones tropicaux de l'Atlantique concernant la prévision opérationnelle officielle du National Hurricane Center aux Etats-Unis. Extrait de [17]

1.5.2- Quelques modèles de prévision de l'intensité cyclonique

Il existe des méthodes de prévision d'intensité dynamique et statistiques qui ont à peu près les mêmes performances de prévisions. Les modèles de prévision de l'intensité cyclonique sont moins nombreux que ceux de la trajectoire de cyclone.

Voici quelques modèles utilisés dans le bassin Atlantique :

- SHIFOR (Statistical Hurricane Intensity FORecast) est un modèle statistique, à régressions multiples, de prévision d'intensité, utilisé pour les prévisions de changement de l'intensité. Ce modèle utilise la persistance et incorpore des données climatologiques ;
- SHIPS (Statistical Hurricane Intensity Prediction Scheme) : c'est un modèle statistico-synoptique, dont parmi les sources d'informations synoptiques sont la température de la mer, le cisaillement vertical du vent, qu'il combine avec la tendance de l'intensité du cyclone ;
- DSHP (Decay SHIP) est un modèle identique au SHIPS mais il diminue l'intensité lorsque le cyclone passe sur terre ;

Dans ce travail, nous avons utilisé les modèles statistiques (RLM et RNA) en exploitant la mémoire du bassin Sud-Ouest de l'Océan Indien.

1.6- Le bassin Sud-Ouest de l'Océan Indien

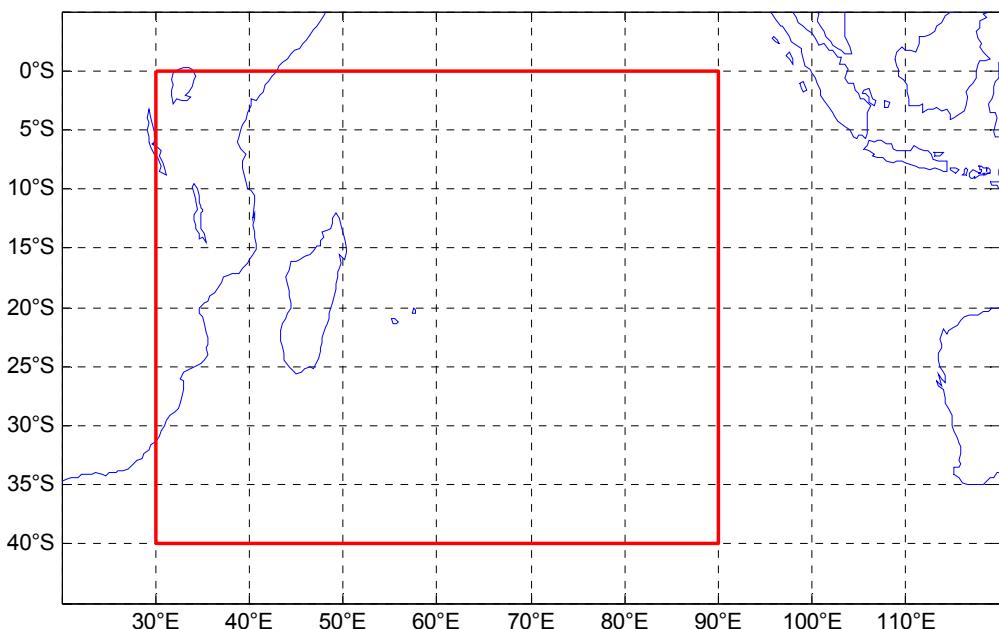


Figure 1.5 : Carte représentant le bassin Sud-Ouest de l'Océan Indien (BSOOI).

Le bassin Sud-Ouest de l'Océan Indien (BSOOI) est la zone s'étendant de 30°E à 90°E en longitude et de 0°S à 40°S en latitude (*Figure 1.5*), selon le CMRS (Centre Météorologique Régional Spécialisé). En moyenne, 12% de l'activité cyclonique annuelle globale se produit dans le BSOOI avec une dizaine de tempêtes tropicales dont 4 aboutissent au stade de cyclone tropical (*Tableau 1.1*).

Les perturbations cycloniques se forment sur le Canal de Mozambique entre 15°S et 20°S et sur l'Océan Indien entre 10°S et 15°S dans la Zone de Convergence Intertropicale (ZCIT). Ces phénomènes se développent dans le BSOOI en se dirigeant généralement vers le sud.

La saison cyclonique est étroitement liée à la saison chaude australe (Novembre-Avril). Ainsi, 90 % des perturbations se forment entre le 15 novembre et le 30 avril (*Figure 1.6*). Les mois les plus actifs sont janvier et février qui regroupent presque la moitié de l'activité cyclonique. Deux maxima secondaires regroupant environ 15 % de l'activité cyclonique chacun, apparaissent en décembre et mars, avec en moyenne des systèmes plus intenses en mars. Cependant, il existe une forte variabilité interannuelle de l'activité cyclonique en ce qui concerne le nombre de cyclones et tempêtes dans le bassin et le nombre de jours d'activité des systèmes [3].

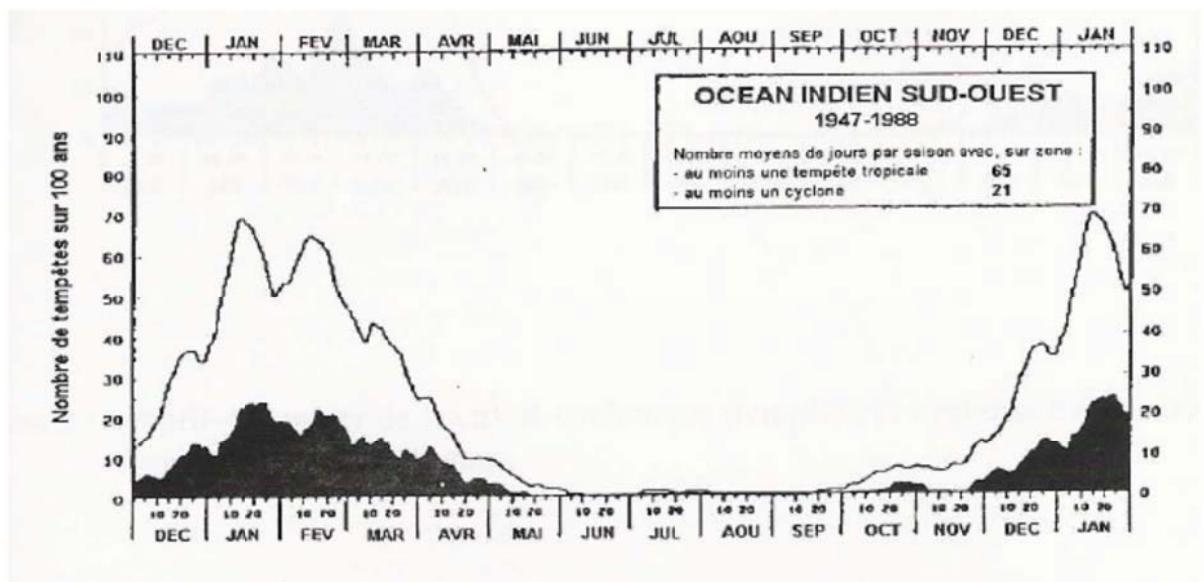


Figure 1.6 : Profils saisonniers de l'activité cyclonique (tempêtes et cyclones tropicaux) dans le bassin Sud-Ouest de l'Océan Indien. Extrait de [3]

La vitesse de déplacement des perturbations sur l'Océan Indien est assez lente, de l'ordre de 15 km/h à 25 km/h . Généralement, elles suivent une trajectoire parabolique décomposable en 3 mouvements successifs :

- mouvement en direction de l'ouest-sud-ouest,
- incurvation vers le sud à la faveur d'un couloir dépressionnaire,
- et enfin, évacuation vers l'est-sud-est aux latitudes extratropicales par le courant d'ouest.

Les trajectoires de début de saison sont plutôt zonales d'est en ouest. En pleine saison prédominent plutôt les trajectoires à recourbement de type parabolique. Un peu plus tard dans la saison on observe encore une proportion importante de trajectoires à recourbement de type parabolique, mais la présence grandissante de couloirs dépressionnaires d'origine polaire offre statistiquement davantage de chance aux perturbations de plonger vers le sud avant d'atteindre les îles habitées. En ce qui concerne la fin de saison, surtout pour celles qui se terminent tardivement, les trajectoires reprennent un caractère zonal. En effet à cause de cyclogenèses tardives se produisant à des latitudes plus basses, les perturbations restent généralement prises dans le flux d'est bien établi au sud de la ZCIT.

Chapitre II : LIAISON ENTRE LES VARIABLES - REGRESSION LINEAIRE MULTIPLE

2.1- Introduction

La première partie de ce chapitre représente les deux types de liaison qui peut exister entre deux séries de données : la liaison linéaire et la liaison non linéaire. On utilise souvent le coefficient de corrélation linéaire pour quantifier la relation entre deux variables. Le problème est qu'on ne peut rien dire lorsque sa valeur absolue est faible (proche de zéro). Deux variables fortement liées peuvent avoir un coefficient de corrélation linéaire nul. Un autre coefficient, capable de détecter la liaison fonctionnelle entre variables, sera exposé dans ce chapitre, c'est le rapport de corrélation.

Dans la deuxième partie, nous verrons la méthode de régression linéaire multiple. Bien que le cyclone soit un phénomène très complexe, nous créerons, avec cette technique, un modèle linéaire utilisant plusieurs prédicteurs. L'objectif est de la comparer avec d'autre méthode non linéaire (le RNA).

2.2- Liaison entre les variables [18]

2.2.1- Corrélation linéaire

2.2.1.1- Définitions

X et Y étant deux variables aléatoires de carrés intégrables.

- La covariance entre X et Y est la valeur

$$\text{cov}(X, Y) = E((X - E(X))(Y - E(Y))) = E(XY) - E(X)E(Y) \quad 2.1$$

Si $\text{cov}(X, Y) = 0$ on dit que X et Y sont orthogonales ($X \perp Y$).

- La corrélation linéaire entre X et Y est quantifiée par

$$\rho_{XY} = \frac{\text{cov}(X, Y)}{\sigma_X \sigma_Y} \quad 2.2$$

2.2.1.2- Propriétés

- si X et Y sont indépendantes, $\text{cov}(X, Y) = 0$ donc $\rho_{XY} = 0$. La réciproque n'est pas vraie.
- $|\rho_{XY}| = 1$ si et seulement si il existe un lien linéaire déterministe entre X et Y .

2.2.1.3- Interprétation géométrique

Comme

$$\text{cov}(X, Y) = \langle X - E(X), Y - E(Y) \rangle \quad 2.3$$

alors

$$\rho_{XY} = \frac{\langle X - E(X), Y - E(Y) \rangle}{\|X - E(X)\| \|Y - E(Y)\|} \quad 2.4$$

Autrement dit, ρ_{XY} est le cosinus de l'angle formé par $X - E(X)$ et $Y - E(Y)$ donc $0 \leq \rho_{XY}^2 \leq 1$.

2.2.2- Rapport de corrélation

2.2.2.1- Définition

X et Y étant deux variables aléatoires. Le rapport de corrélation de Y sur X est la quantité positive $\eta_{Y|X}^2$:

$$\eta_{Y|X}^2 = \frac{\text{var}[E(Y|X)]}{\text{var}(Y)} = \frac{E((E(Y) - E(Y|X))^2)}{\text{var}(Y)} = 1 - \frac{E((Y - E(Y|X))^2)}{\text{var}(Y)} \quad 2.5$$

2.2.2.2- Interprétation géométrique

$\eta_{Y|X}^2$ est le cosinus carré de l'angle formé par $Y - E(Y)$ et $E(Y|X) - E(Y)$ donc $0 \leq \eta_{Y|X}^2 \leq 1$.

2.2.3- Remarques

- Généralement, $\eta_{Y|X}^2 \neq \eta_{X|Y}^2$.
- $\eta_{Y|X}^2 = 1 \Rightarrow E((Y - E(Y|X))^2) = 0 \Rightarrow E(Y|X) = Y$; cette dernière relation nous indique qu'il existe une liaison fonctionnelle entre Y et X .
- $\eta_{Y|X}^2 = 0 \Rightarrow E((E(Y) - E(Y|X))^2) = 0 \Rightarrow E(Y|X) = E(Y)$, la régression entre Y et X est alors constante. Cela implique que $X \perp Y$ (la covariance entre Y et X est nulle).

- $\eta_{Y|X}^2 \geq \rho_{XY}^2$. Le rapport de corrélation apporte une information de plus, sur la liaison de deux variables, par rapport au coefficient de corrélation linéaire.
- $\eta_{Y|X}^2 = 0 \Rightarrow \rho_{XY}^2 = 0$. La réciproque est généralement fausse.
- La régression de Y sur X est linéaire, si et seulement si $\eta_{Y|X}^2 = \rho_{XY}^2$.

2.3- Régression linéaire multiple (RLM)

2.3.1- Définition

Dans ce modèle, une variable quantitative à expliquer Y est mise en relation avec p variables quantitatives explicatives X_1, X_2, \dots, X_p par la relation :

$$y_i = a_0 + a_1 x_{i1} + a_2 x_{i2} + \dots + a_p x_{ip} + e_i \quad i = 1, 2, \dots, n \quad 2.6$$

où a_0, a_1, \dots, a_p sont les paramètres inconnus et $e_i, i = 1, 2, \dots, n$ les termes d'erreur.

Les données sont rangées dans une matrice $\mathbf{X}(n \times (p+1))$ de terme général x_{ij} , dont la première colonne contient le vecteur $\mathbf{1}$ (les n éléments sont tous égaux à 1), et dans un vecteur \mathbf{y} de terme général y_i . Avec les notations vectorielles $\mathbf{e} = [e_1 \dots e_p]'$ et $\mathbf{a} = [a_0 a_1 \dots a_p]'$, le modèle s'écrit matriciellement :

$$\mathbf{y} = \mathbf{X}\mathbf{a} + \mathbf{e} \quad 2.7$$

2.3.2- Estimation des paramètres

Les paramètres inconnus du modèle sont le vecteur \mathbf{a} et σ_e^2 (paramètre de nuisance). Ils peuvent être estimés par minimisation du critère des moindres carrés. L'expression à minimiser sur $\mathbf{a} \in \mathbb{R}^{p+1}$ s'écrit :

$$\sum_{i=1}^n (y_i - a_0 - a_1 x_{i1} - a_2 x_{i2} - \dots - a_p x_{ip})^2 = \|\mathbf{y} - \mathbf{X}\mathbf{a}\|^2 \quad 2.8$$

$$\sum_{i=1}^n (y_i - a_0 - a_1 x_{i1} - a_2 x_{i2} - \dots - a_p x_{ip})^2 = (\mathbf{y} - \mathbf{X}\mathbf{a})' (\mathbf{y} - \mathbf{X}\mathbf{a}) \quad 2.9$$

$$\sum_{i=1}^n (y_i - a_0 - a_1 x_{i1} - a_2 x_{i2} - \dots - a_p x_{ip})^2 = \mathbf{y}'\mathbf{y} - 2\mathbf{a}'\mathbf{X}'\mathbf{y} + \mathbf{a}'\mathbf{X}'\mathbf{X}\mathbf{a} \quad 2.10$$

En dérivant matriciellement la dernière équation par rapport à \mathbf{a} et en égalant à zéro le résultat, on obtient les « équations normales » :

$$\mathbf{X}'\mathbf{y} - \mathbf{X}'\mathbf{X}\hat{\mathbf{a}} = \mathbf{0} \quad 2.11$$

La solution $\hat{\mathbf{a}}$ de cette équation correspond bien à un minimum car la matrice hessienne $2\mathbf{X}'\mathbf{X}$ est définie-positive. Si la matrice $\mathbf{X}'\mathbf{X}$ est inversible, c'est-à-dire que la matrice \mathbf{X} est de rang $(p+1)$ et donc qu'il n'existe pas de colinéarité entre ses colonnes.

L'estimation des paramètres a_j est alors donnée par :

$$\hat{\mathbf{a}} = (\mathbf{X}'\mathbf{X})^{-1} \mathbf{X}'\mathbf{y} \quad 2.12$$

2.3.3- Estimation de la sortie du modèle et résidus

Les valeurs estimées de \mathbf{y} sont obtenues par :

$$\hat{\mathbf{y}} = \mathbf{X}\hat{\mathbf{a}} = \mathbf{X}(\mathbf{X}'\mathbf{X})^{-1} \mathbf{X}'\mathbf{y} = \mathbf{H}\mathbf{y} \quad 2.13$$

où $\mathbf{H} = \mathbf{X}(\mathbf{X}'\mathbf{X})^{-1} \mathbf{X}'$ est la matrice de projection orthogonale dans \mathbb{R}^n sur le sous-espace $\text{Vect}(\mathbf{X})$ engendré par les vecteurs colonnes de \mathbf{X} .

Le vecteur des résidus est :

$$\mathbf{u} = \mathbf{y} - \hat{\mathbf{y}} = \mathbf{y} - \mathbf{X}\hat{\mathbf{a}} = (\mathbf{I} - \mathbf{H})\mathbf{y} \quad 2.14$$

Géométriquement, \mathbf{u} est la projection de \mathbf{y} sur le sous-espace orthogonal de $\text{Vect}(\mathbf{X})$ dans \mathbb{R}^n .

2.3.4- Coefficient de détermination

Le coefficient de détermination R^2 est défini par :

$$R^2 = \frac{\text{SSR}}{\text{SST}} \quad 2.15$$

où SSR est la somme des carrés de la régression et SST la somme totale des carrés avec :

$$\text{SSR} = \|\hat{\mathbf{y}} - \bar{\mathbf{y}}\mathbf{1}\|^2 \quad 2.16$$

et

$$SST = \|\mathbf{y} - \bar{\mathbf{y}}\mathbf{1}\|^2 \quad 2.17$$

On remarque que :

$$SST = SSR + SSE \quad 2.18$$

avec

$$SSE = \|\mathbf{y} - \hat{\mathbf{y}}\|^2 \quad 2.19$$

la somme des carrés des résidus.

Le coefficient de détermination R^2 est donc la part de variation de Y expliquée par le modèle de régression. Géométriquement, c'est le rapport de carrés de longueur de deux vecteurs. C'est le cosinus carré de l'angle entre le vecteur \mathbf{y} et sa projection $\hat{\mathbf{y}}$ sur $\text{Vect}(\mathbf{X})$.

La quantité R est aussi appelée coefficient de corrélation multiple entre Y et les variables explicatives, c'est le coefficient de corrélation usuel entre \mathbf{y} et sa prédiction $\hat{\mathbf{y}}$.

Il est très important de souligner que l'ajout de variables explicatives fait croître le coefficient de détermination, même si ces variables ne sont pas liées à Y .

2.3.5- Tests sur les paramètres

Pour chaque paramètre a_j , on peut montrer que la statistique

$$\frac{\hat{a}_j - a_j}{\sigma_{\hat{a}_j}}$$

suit une loi de Student à $(n-p-1)$ degrés de liberté. $\sigma_{\hat{a}_j}^2$ pour $j=1, \dots, p+1$, est la variance de \hat{a}_j et égale au terme diagonal de la matrice $\sigma^2 (\mathbf{X}'\mathbf{X})^{-1}$.

L'estimateur non biaisé de σ^2 vaut :

$$\hat{\sigma}^2 = \frac{\|\mathbf{y} - \hat{\mathbf{y}}\|^2}{n-p-1} \quad 2.20$$

CHAPITRE III : LES RESEAUX DE NEURONES

3.1- Introduction

L'utilisation des réseaux de neurones artificiels commence depuis 1943. A cette époque, Mc Culloch et Pitts étudièrent un ensemble de neurones formels interconnectés et montrèrent leurs capacités à calculer certaines fonctions logiques. C'était en 1949 que Hebb fournit une explication synaptique du mécanisme d'apprentissage. Mais le modèle de perceptron ne fut présenté qu'en 1962, par Rosenblatt.

Les principales qualités des RNA sont leur capacité d'adaptabilité et d'auto-organisation, d'une part, et la possibilité de résoudre des problèmes non linéaires avec une bonne approximation [19,20]. Les RNA permettent de réaliser plusieurs applications comme la classification, la modélisation, le filtrage, la compression de données, etc...

Le plus grand problème des RNA, aujourd'hui, est la détermination de leur architecture optimale (nombre de couches cachées, nombre de neurones cachés, fonctions d'activation). Cependant, nous allons utiliser des techniques (comme la validation simple) permettant d'avoir des architectures capables de prévoir, avec des erreurs acceptables, le phénomène cyclonique.

Dans ce chapitre, nous allons présenter les propriétés biologiques des réseaux de neurones. Cette partie est très intéressante car la connaissance du cerveau et de son fonctionnement améliore beaucoup sa modélisation. Ensuite, les différents types de RNA, leurs propriétés fondamentales et les techniques courantes d'apprentissage seront exposés.

3.2- Réseau de neurones biologiques

Généralement, l'évolution d'un être humain est caractérisée par le fonctionnement de son cerveau. Cet élément du système nerveux central contient environ 100 milliards de neurones et environ 1000 à 10000 contacts synaptiques par neurone [21]. Au total, le nombre de connexions dans le cerveau est estimé à environ 10^{15} .

Nous allons décrire, ici, les propriétés essentielles des neurones biologiques et de leurs connexions. Notre objectif est de comprendre le fonctionnement de ce système très complexe et de voir les correspondances avec les RNA.

3.2.1 Le neurone biologique et sa constitution

Les éléments de base du système nerveux central sont les neurones, selon Stevens [22]. Ces cellules jouent des rôles très complexes dans la transmission des informations en provenance des autres neurones. Un neurone reçoit les signaux venant des neurones voisins, les intègre, engendre un influx nerveux, le conduit et le transmet à un autre neurone capable de le recevoir.

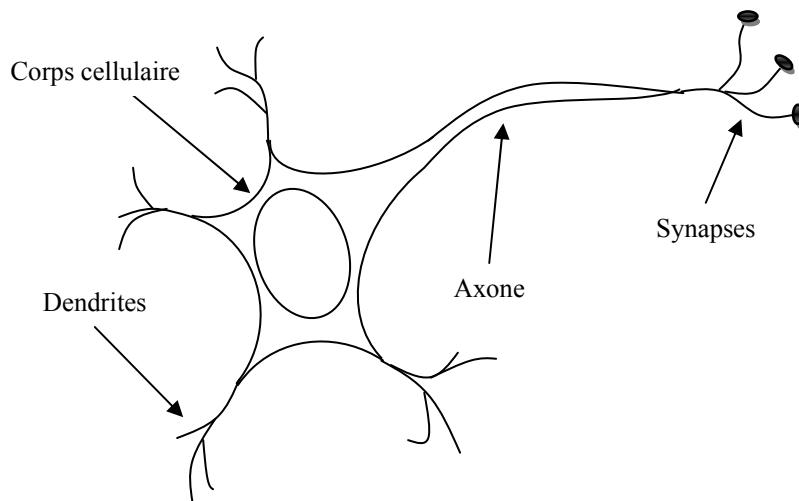


Figure 3.1 : Représentation d'un neurone biologique.

La *Figure 3.1* montre les trois principales parties d'un neurone :

- ❖ Le corps cellulaire, qui contient le noyau du neurone et effectue les transformations biochimiques nécessaires à la synthèse des enzymes et des autres molécules qui assurent la vie du neurone ;
- ❖ Les dendrites, qui sont de fines extensions tubulaires. En se ramifiant, elles forment une espèce d'arborescence autour du corps cellulaire. Elles sont les récepteurs principaux du neurone pour capter les signaux qui lui parviennent ;
- ❖ Et enfin, l'axone qui est la fibre nerveuse servant de moyen de transport pour les signaux émis par le neurone. Il se distingue des dendrites par sa forme et par les propriétés de sa membrane externe. En effet, il est généralement plus long que les dendrites, et se ramifie à son extrémité, là où il communique avec d'autres neurones, alors que les ramifications des dendrites se produisent plutôt près du corps cellulaire.

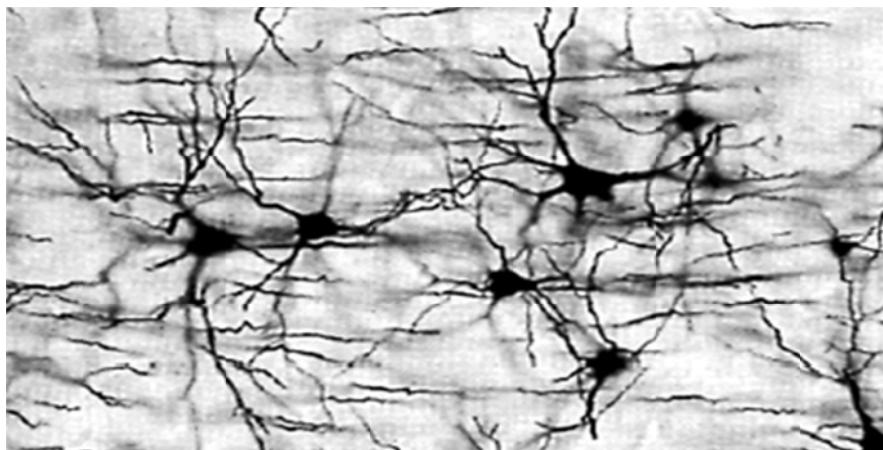


Figure 3.2 : Réseau de neurones biologique

Pour former le système nerveux, les neurones sont connectés les uns aux autres suivant des répartitions spatiales complexes (*Figure 3.2*). La connexion entre deux neurones se fait à l'aide d'un espace de quelques dizaines d'Angströms appelé synapse [21].

3.2.2 Fonctionnement des neurones

C'est la membrane externe d'un neurone qui réalise essentiellement ses fonctions. Voici ses cinq fonctions principales :

- Elle sert à propager des impulsions électriques tout au long de l'axone et des dendrites ;
- Elle libère des médiateurs à l'extrémité de l'axone ;
- Elle réagit à ces médiateurs au niveau des dendrites ;
- Elle réagit au niveau du corps cellulaire aux impulsions électriques que lui transmettent les dendrites pour générer ou non une nouvelle impulsion ;
- Elle permet au neurone de reconnaître les autres neurones afin qu'il puisse se situer au cours de la formation du cerveau et trouver les cellules auxquelles il doit être connecté.

La transmission d'informations entre deux neurones se fait par un phénomène électrique. Le corps cellulaire du neurone traite les courants électriques qui lui proviennent de ses dendrites, et qu'il transmet le courant électrique, si un certain seuil

est atteint, résultant de ce traitement aux neurones auxquels il est connecté par l'intermédiaire de son axone.

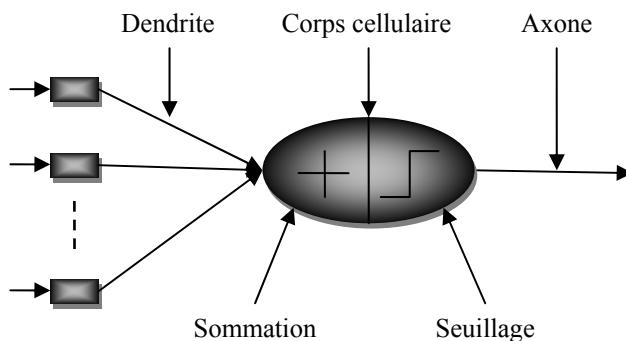


Figure 3.3 : Représentation du corps cellulaire comme sommateur à seuil.

Le schéma classique présenté par les biologistes et illustré par la *Figure 3.3* est celui d'un corps cellulaire effectuant une sommation des influx nerveux transmis par ses dendrites. Si la sommation dépasse un seuil, le neurone répond par un influx nerveux ou potentiel d'action qui se propage le long de son axone. Si la sommation est inférieure à ce seuil, le neurone reste inactif. L'influx nerveux qui se propage entre différents neurones est, au niveau de ces neurones, un phénomène électrique.

3.2.3 Les synapses

Le rôle des synapses est fondamental pour permettre aux cellules nerveuses de communiquer entre elles [23]. Les signaux qui se propagent dans les neurones sont de nature électrique. Cependant, il n'existe pas de liaisons directes entre deux cellules nerveuses. Celles-ci sont séparées par un espace appelé fente synaptique que l'influx électrique ne peut traverser. Le relais s'effectue à ce niveau par l'intermédiaire d'un médiateur chimique, c'est à dire, l'arrivée du potentiel d'action à l'une des extrémités de l'arborisation terminale, déclenche la libération d'une substance chimique appelé neurotransmetteur qui est soit excitateur car il excite le neurone cible et produit dans celui-ci une nouvelle impulsion, soit inhibiteur car il bloque l'impulsion nerveuse aux endroits synaptiques où il sert de relais. Cette substance diffuse dans l'espace synaptique et vient se fixer sur des récepteurs spécifiques appelés neurorécepteurs, situés sur la terminaison du neurone cible. Cette fixation provoque l'ouverture des canaux ioniques, ce qui donne naissance à un nouveau signal électrique.

3.2.4 La règle de Hebb

Hebb a proposé en 1949 une règle où la force de la connexion entre deux neurones augmente s'il y a corrélation d'activité : « Quand une cellule A excite par son axone une cellule B et que, de manière répétée et persistante, elle participe à la genèse d'une impulsion dans B, un processus de croissance ou un changement métabolique a lieu dans l'une ou dans les deux cellules, de telle sorte que l'efficacité de A à déclencher une impulsion dans B est, parmi les autres cellules qui ont cet effet, accrue ».

3.3- Réseau de neurones artificiels

Les réseaux de neurones artificiels (RNA) sont, par définition, des modélisations du réseau de neurones biologiques qui est beaucoup plus complexe.

3.3.1- Modèle d'un neurone

Un RNA est un système composé de plusieurs petites unités appelées neurones artificiels. Mathématiquement, un neurone artificiel est une fonction, généralement, non linéaire, paramétrée, à valeurs bornées.

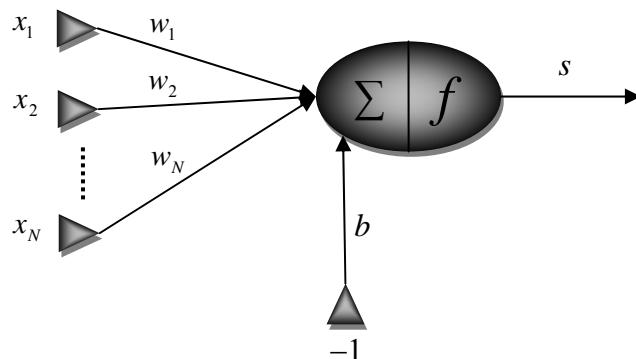


Figure 3.4 : Modèle d'un neurone artificiel

Le modèle mathématique d'un neurone artificiel est illustré à la Figure 3.4. Un neurone est essentiellement constitué d'un intégrateur qui effectue la somme pondérée (les poids sont $w_i, i = 1 \dots N$, b est le biais du neurone appelé aussi seuil d'activation) de ses entrées $x_i, i = 1 \dots N$. Le résultat de cette somme est ensuite transformé par une fonction d'activation f qui produit la sortie s du neurone. La sortie na de l'intégrateur est donnée par la relation :

$$na = \sum_{i=1}^N w_i x_i - b \quad 3.1$$

na est appelé le niveau d'activation du neurone. Lorsque le niveau d'activation atteint ou dépasse le seuil b , alors a devient positif ou nul. Sinon, il est négatif.

$$s = f\left(\sum_{i=1}^N w_i x_i - b\right) \quad 3.2$$

Ainsi, chaque neurone peut faire de petites opérations mathématiques pour les transmettre aux cellules de la couche suivante. Un groupe de neurones effectuant les mêmes opérations mathématiques, constitue un groupe appartenant à la même couche. La couche d'entrée est constituée directement des variables (régresseurs), tandis que la couche de sortie est composée des neurones calculant l'estimateur de la variable dépendante. Une ou plusieurs couches cachées est nécessaire pour les opérations internes (*Figure 3.5*).

Les neurones sont reliés d'une couche à l'autre de façon unidirectionnelle par des connexions représentant les synapses dans le réseau de neurones biologique. L'ensemble des neurones et des connexions constitue le réseau de neurones artificiels.

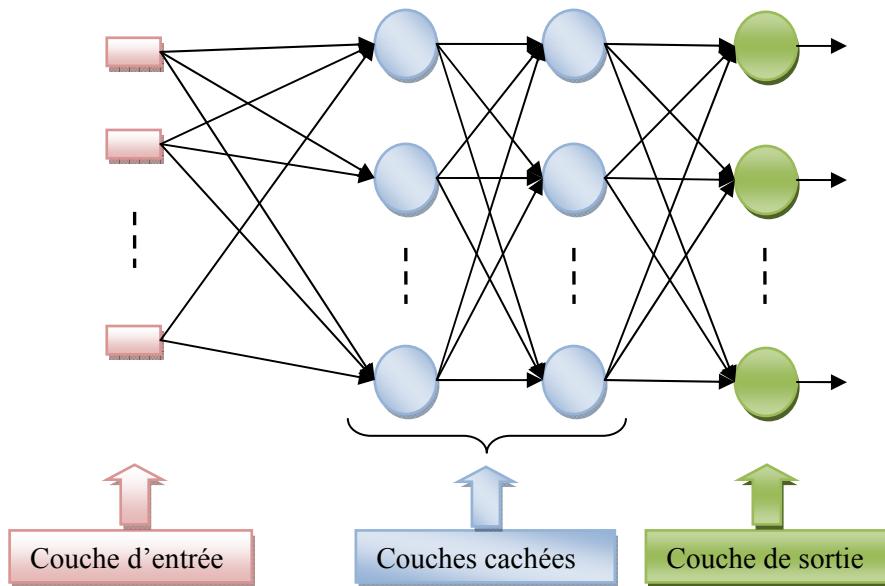


Figure 3.5 : Structure typique d'un RNA avec deux couches cachées.

3.3.2- Different types de réseaux

3.3.2.1- Réseaux de neurones non bouclés [24]

Un réseau de neurones non bouclé réalise une ou plusieurs fonctions algébriques de ses entrées, par composition des fonctions réalisées par chacun de ses neurones. Un réseau de neurones non bouclé est représenté graphiquement par un ensemble de neurones connectés entre eux, l'information circulant des entrées vers les sorties sans retour en arrière.

La *Figure 3.5* représente un réseau de neurones non bouclé qui a une structure particulière. Il comprend une couche d'entrée, deux couches de neurones cachés et une couche de sortie. Remarquons que les neurones d'une couche cachée ne sont pas connectés entre eux.

Ce type de réseau est utilisés principalement pour effectuer des tâches d'approximation de fonction non linéaire, de classification ou de modélisation de processus statiques non linéaires (c'est notre cas).

Les réseaux de neurones non bouclés à couches dont les neurones cachés ont une fonction d'activation sigmoïde, sont souvent appelés des perceptrons multicouches (PMC).

3.3.2.2- Réseaux de neurones bouclés [24]

Ce type de réseaux peut avoir une topologie de connexions quelconque, comprenant notamment des boucles qui ramènent aux entrées la valeur d'une ou plusieurs sorties. Un réseau de neurones bouclé est schématisé par un graphe des connexions qui est cyclique.

Les réseaux de neurones bouclés sont utilisés pour effectuer des tâches de modélisation de systèmes dynamiques, de commande de processus, ou de filtrage.

3.3.3- Propriétés fondamentales des réseaux de neurones statiques

3.3.3.1- Les réseaux de neurones sont des approximatrices universels

Toute fonction bornée suffisamment régulière peut être approchée uniformément, avec une précision arbitraire, dans un domaine fini de l'espace de ses variables, par un réseau de neurones comportant une couche de neurones cachés en nombre fini, possédant tous la même fonction d'activation, et un neurone de sortie linéaire [25,26,27].

3.3.3.2- Les réseaux de neurones non linéaires par rapport à leurs paramètres sont des approximateurs parcimonieux

Dans la pratique, le nombre de fonctions nécessaires pour réaliser une approximation est un critère important dans le choix d'un approximateur de préférence à un autre. La complexité d'un modèle est liée au nombre de ses paramètres : pour contrôler le surajustement, on doit toujours faire en sorte que ce nombre soit le plus petit possible. En d'autres termes, on cherche l'approximation la plus *parcimonieuse*. Les réseaux de neurones possèdent cette propriété de parcimonie : *c'est en cela que réside leur intérêt par rapport aux modèles linéaires en leurs paramètres tels que les polynômes.*

On montre [28] que, si l'approximation dépend des paramètres ajustables de manière non linéaire, elle est plus parcimonieuse que si elle dépend linéairement des paramètres.

3.3.4- Structure du réseau de neurones (nombre de couches, neurones cachés, fonction d'activation)

3.3.4.1- Nombre de couches

La détermination du nombre de couches cachée pour un cas donné reste encore un grand problème pour l'optimisation du réseau de neurones. Cependant, il est possible à partir d'une seule couche cachée d'avoir un résultat acceptable selon la propriété citée dans le paragraphe 3.3.3.1-

3.3.4.2- Nombre de neurones cachés

Le nombre de neurones nécessaires pour la couche cachée dépend des données. Nous utilisons la technique considérant une partie des données pour la validation.

3.3.4.3- Fonction d'activation

Le choix des fonctions d'activation dépend de leur utilisation. Les plus fréquentes sont les fonctions seuil, linéaire et sigmoïde. La fonction seuil est utilisée généralement pour les données catégorielles (classification). Comme nous avons vu au paragraphe 3.3.3.1-, la fonction d'activation du neurone de sortie de notre réseau est linéaire. Pour les neurones cachés, nous avons utilisé la fonction sigmoïde qui a l'avantage d'être non linéaire. La sortie du réseau est alors une fonction non linéaire de ses paramètres. Ce qui offre au RNA un avantage par rapport au modèle polynomial, contre l'inconvénient d'une procédure d'optimisation non linéaire des paramètres. Le cout de calcul pour le RNA est plus important.

3.3.5- Apprentissage

Par définition, l'apprentissage est une phase du développement d'un réseau de neurones durant laquelle le comportement du réseau est modifié jusqu'à l'obtention du comportement désiré. Lors d'un apprentissage, on a besoin d'un ensemble d'exemples de comportement.

3.3.5.1- Principe de l'apprentissage

Durant cette phase, le réseau adapte ses poids des connexions afin de fournir sur ses neurones de sortie les valeurs désirées (cas de l'apprentissage supervisé du paragraphe suivant). Après initialisation des poids du réseau (en général des valeurs aléatoires), on présente les exemples au réseau et on calcule les sorties correspondantes. Une valeur d'erreur ou de correction est calculée et une correction des poids est appliquée.

Il y a deux grandes classes d'algorithme d'apprentissage selon qu'il est supervisé ou non supervisé. Nous allons voir dans les paragraphes suivants la différence entre ces deux approches qui dépend surtout de la forme des exemples d'apprentissage utilisés.

3.3.5.2- Apprentissage supervisé

Ce type d'apprentissage nécessite N couples de la forme (entrée, sortie désirée). Il consiste à déterminer les poids synaptiques w de telle sorte que les sorties calculées par le RNA et les sorties désirées sont aussi proches que possibles.

Si on note par R_w le réseau de neurones et par (x_i, y_i) les couples d'observations, alors on cherche w de R_w capable de mettre x_i et y_i en correspondance. Autrement dit on veut avoir :

$$R_w(x_i) = y_i \text{ avec } i = 1, \dots, N \quad 3.3$$

Nous allons utiliser ce type d'apprentissage pour caractériser le Bassin Sud-Ouest de l'Océan Indien du point de vue intensité cyclonique.

3.3.5.3- Apprentissage non supervisé

L'apprentissage est dit non supervisé lorsque seules les valeurs d'entrée sont disponibles. Dans ce cas, les exemples présentés à l'entrée provoquent une auto-adaptation du réseau afin de produire des valeurs de sortie qui soient proches en réponse à des valeurs d'entrée similaires (de même nature).

3.3.5.4- Fonction de cout

Pour un exemple i d'un ensemble d'observation, la fonction de cout des moindres carrés est égale à la somme, sur les N_S neurones de la couche de sortie, des carrés des écarts entre la sortie du modèle $s(i)$ et la sortie désirée $d(i)$. La fonction de cout dépend du vecteur des paramètres w , et on note par $J^i(w)$:

$$J^i(w) = \sum_{q=1}^{N_S} (d_q^i - s_q^i)^2 \quad 3.4$$

Sur l'ensemble d'exemples, la fonction de cout est notée $J(w)$ et elle est définie par la moyenne des carrés des écarts sur les N exemples de cet ensemble :

$$J(w) = \frac{1}{N} \sum_{i=1}^N J^i(w) \quad 3.5$$

Cette fonction dépend du vecteur de paramètres w et de l'ensemble d'exemples considéré.

3.3.5.5- Algorithmes d'optimisation

3.3.5.5.1- Principe

Les algorithmes d'apprentissage des réseaux de neurones formels sont des algorithmes d'optimisation. On cherche à minimiser une fonction coût qui constitue une mesure de l'écart entre les réponses réelles du réseau et ses réponses désirées.

Notre objectif, pour l'apprentissage, est d'avoir une fonction de cout minimale pour un vecteur de paramètres w estimé à partir des algorithmes d'optimisation.

Le modèle réseau de neurones est un modèle non linéaire par rapport aux paramètres. La recherche de ces derniers s'effectue à partir des méthodes itératives, qui nécessite des fonctions de cout dérivable par rapport à w .

Ces techniques consistent à se placer en un point de l'espace des paramètres, trouver une direction de descente du cout et à se déplacer d'un pas suivant cette direction. On arrive à un nouveau point et on recommence les étapes précédentes jusqu'à satisfaction d'un critère d'arrêt.

Soit w_{t+1} le vecteur des paramètres à l'itération $t+1$. Sa relation avec le vecteur des paramètres w_t de l'itération précédente est donnée par :

$$w_{t+1} = w_t + \mu \cdot D_t \quad 3.6$$

où μ est le pas et D_t la direction de descente, qui dépend des w_t .

Ainsi, la modification des paramètres dépend de la valeur du pas μ et celle de la direction de descente D_t .

Il y a plusieurs méthodes d'optimisation des RNA dont nous allons voir dans les paragraphes suivants les plus fréquemment utilisées.

3.3.5.5.2- Méthode du gradient à pas constant

Dans cette méthode, le pas μ est constant. Le problème consiste alors à calculer le gradient qui donne la direction de descente. Les paramètres à l'itération $t+1$ sont donnés par :

$$w_{t+1} = w_t - \mu \cdot \nabla J_t \quad 3.7$$

On remarque que $D_t = -\nabla J_t$.

Cette méthode est très simple d'utilisation et elle est efficace loin d'un minimum (à condition de ne pas être sur un plateau de la fonction de cout). En revanche, lorsque l'on s'approche du minimum, le gradient tend vers 0 et la vitesse de convergence diminue très fortement.

3.3.5.5.3- Méthode du gradient à pas variable

Pour une direction de descente choisie par le gradient par exemple, il est possible d'asservir le pas μ de telle sorte que la fonction de cout diminue à chaque modification des paramètres. On s'intéresse donc à la fonction $g(\mu)$ unidimensionnelle définie par :

$$g(\mu) = J(w_t + \mu \cdot D_t) \quad 3.8$$

A partir de la fonction g , il faut trouver une valeur convenable (ni trop petit pour assurer une convergence rapide, ni trop grande pour ne pas être confronté à un comportement oscillatoire). Les méthodes les plus efficaces sont les méthodes de dichotomie mais nécessitent souvent trop de calculs.

3.3.5.5.4- Méthode de Newton

Cette méthode utilise la courbure de la fonction de cout pour atteindre le minimum plus rapidement. Les paramètres à l'itération $t+1$ sont donnés par :

$$w_{t+1} = w_t - H_t^{-1} \nabla J_t \quad 3.9$$

Dans ce cas $D_t = -H_t^{-1} \nabla J_t$ et $\mu = 1$ où H_t^{-1} est l'inverse de la matrice du Hessien à l'itération t .

Si $J(w)$ est une quadrique, l'algorithme atteint la solution en une seule itération. Sinon, cette méthode est très efficace au voisinage d'un minimum. Cependant, pour que la méthode converge vers le minimum, le Hessien doit être défini positif. Dans le cas général d'un modèle non linéaire, cette hypothèse de convergence n'est pas toujours respectée et la méthode peut ne pas converger. En pratique elle est peu employée car elle nécessite, de plus, le calcul du Hessien à chaque itération. On préfère des méthodes plus économiques dites de « quasi-Newton ».

3.3.5.5.5- Méthode de quasi-Newton

Dans cette méthode, l'inverse de l'Hessien est approché par une matrice M_t , définie positive modifiée à chaque itération. La suite des matrices (M_t) est construite de manière à converger vers l'inverse de l'Hessien lorsque la fonction de cout $J(W)$ est une quadrique. La modification des paramètres est donnée par :

$$w_{t+1} = w_t - \mu \cdot M_t \nabla J_t \quad 3.10$$

Dans ce cas $D_t = -M_t \nabla J_t$ et μ est évalué avec une méthode de minimisation unidimensionnelle.

A la première itération, la matrice M_0 est prise égale à la matrice identité. Parmi toutes les méthodes de quasi-Newton existantes, l'outil de simulation et d'apprentissage de réseaux de neurones du laboratoire propose la méthode BFGS, développée indépendamment par Broyden, Fletcher, Goldfarb et Shanno dont la vitesse de convergence est beaucoup plus grande que celle de la méthode du gradient.

3.3.5.5.6- Méthode du gradient conjugué

Cette méthode fait partie des techniques classiques d'optimisation d'ordre 1. Le principe de celle-ci est de calculer une direction le long de laquelle la valeur de la fonction objective diminue. Il s'agit ensuite de calculer le minimum de la fonction objective le long de cette direction de descente. Cette étape, appelée méthode de recherche d'une ligne, est une minimisation à une dimension de la fonction objective. On obtient ainsi un nouveau point où l'on recalcule une nouvelle direction de descente et ainsi de suite jusqu'à la convergence.

3.3.5.5.7- Méthode de Levenberg-Marquardt

La méthode de Levenberg-Marquardt consiste à modifier les paramètres selon la relation suivante :

$$w_{t+1} = w_t - [H_t + \lambda_t \cdot I]^{-1} \nabla J_t \quad 3.11$$

I étant la matrice identité et λ_t un coefficient.

Cette méthode est particulièrement astucieuse car elle s'adapte d'elle-même à la forme de la fonction de cout. Elle effectue un compromis entre la direction du gradient et la direction donnée par la méthode de Newton. En effet, si λ_t est grand, on reconnaît la méthode du gradient (dans ce cas la valeur du pas est donné par $1/\lambda_t$) et si λ_t est petit, la modification des paramètres correspond à celle de la méthode de Newton.

3.3.5.5.8- Remarques

Pour rendre l'optimisation plus performante, on peut utiliser des méthodes de second ordre. Le calcul est très efficace, mais lourd. Elles ont de nombreuses limitations, quant aux conditions de convergence, sur les dérivées seconde. Des corrections sont proposées pour éviter ce problème, et sont prises en compte par les méthodes dites de Quasi-Newton ou de Newton modifiée. Il a été observé que les poids calculés par la méthode de rétropropagation sont plus faibles que ceux obtenus par la technique de Quasi-Newton, ce qui semblerait montrer que la recherche d'un minimum par rétropropagation est restreinte à un voisinage immédiat des poids initiaux, d'où une dépendance plus forte de cette méthode par rapport à l'initialisation. C'est pour ces raisons que l'optimisation par une méthode Quasi-Newton a été largement utilisée dans certains travaux récents de modélisation et a été choisie également dans la suite de nos travaux. La boîte à outils NETLAB est utilisée pour la création et l'optimisation du RNA.

3.3.6- Généralisation

3.3.6.1- Arrêt prématué

Comme l'apprentissage consiste à minimiser, grâce à un algorithme itératif, une fonction de cout calculée sur la base d'apprentissage. La méthode de l'arrêt prématué (early stopping) consiste à arrêter les itérations avant la convergence de l'algorithme. Si la convergence n'est pas menée à son terme, le modèle ne s'ajuste pas trop finement aux données d'apprentissage : le surajustement est limité.

Pour mettre en œuvre cette méthode, il faut déterminer le nombre d'itérations à utiliser pendant l'apprentissage. La méthode la plus classique consiste à suivre l'évolution de la fonction de cout sur une base de validation, et à arrêter les itérations lorsque le cout calculé sur cette base commence à croître. Cependant, en pratique, cette méthode peut être inapplicable, car il est difficile de déterminer avec précision le moment exact où il faut arrêter l'apprentissage puisque les performances sur la base de validation ne se dégradent pas nettement.

3.3.6.2- Régularisation par modération des poids (Weight Decay)

Lorsque les poids du réseau sont grands en valeur absolue, les sigmoïdes des neurones cachés sont saturées, si bien que les fonctions modélisées peuvent avoir des variations brusques. Pour obtenir des fonctions régulières, il faut travailler avec la partie linéaire des sigmoïdes, ce qui implique d'avoir des poids dont la valeur absolue est faible.

La méthode de régularisation du Weight decay limite la valeur absolue des poids en utilisant :

$$\Omega = \frac{1}{2} \sum_{i=1}^p \omega_i^2 \quad 3.12$$

L'apprentissage s'effectue en minimisant :

$$J' = J + \frac{\alpha}{2} \sum_{i=1}^p \omega_i^2 \quad 3.13$$

avec p le nombre de poids que comporte le réseau.

α est un hyperparamètre qui détermine l'importance relative des deux termes dans la nouvelle fonction de cout. Si α est trop grand, les poids tendent rapidement vers zéro, le modèle ne tient plus compte des données. Si α est trop petit, le terme de régularisation perd de son importance et le réseau de neurones peut être surajusté. Dans le cas intermédiaire, les poids après l'apprentissage ont des valeurs modérées.

Cette méthode présente l'avantage d'être très simple à mettre en œuvre, puisque le gradient de J' se calcule très simplement à partir du gradient de J et du vecteur des poids du réseau W :

$$\nabla J' = \nabla J + \alpha W \quad 3.14$$

Il suffit alors d'ajouter la quantité αW au vecteur ∇J calculé par l'algorithme de rétropropagation.

Un modèle qui a bien appris la base d'apprentissage correspond à une valeur faible de J , alors qu'une fonction régulière correspond à une fonction W faible : l'apprentissage doit trouver une solution qui satisfasse ces deux exigences.

CHAPITRE IV : PREVISION DES INTENSITES CYCLONIQUES PAR LES RESEAUX DE NEURONES

CHAPITRE IV : PREVISION DES INTENSITES CYCLONIQUES PAR LES RESEAUX DE NEURONES

4.1- Sélection des prédicteurs

4.1.1- Introduction

Notre objectif est de n'utiliser que les prédicteurs capables de donner au modèle réseau de neurones la meilleure performance possible. Pour cela, nous allons utiliser une technique permettant de sélectionner à partir d'un ensemble donné un sous ensemble de prédicteurs.

L'idée de cette méthode est, dans un premier temps, d'ordonner les prédicteurs par ordre de pertinence. Au départ, nous disposons d'un ensemble de p prédicteurs parmi lesquels nous cherchons à l'aide de la méthode de Gram-Schmidt Modifiée (GSM) celui qui décrit le mieux la sortie désirée, puis le deuxième et ainsi de suite. Nous obtenons finalement un classement des prédicteurs.

Dans un deuxième temps, nous allons utiliser, outre les p prédicteurs initiaux, des prédicteurs aléatoires, pour déterminer le nombre final de prédicteurs à utiliser dans le réseau de neurones.

4.1.2- Algorithme de Gram-Schmidt modifié

Dans notre cas, nous allons exploiter cet algorithme pour ordonner les prédicteurs par ordre de pertinence.

Il existe deux manières de mettre en œuvre l'algorithme de Gram-Schmidt :

- L'algorithme de Gram-Schmidt classique (CGS ou Classical Gram-Schmidt) ;
- L'algorithme de Gram-Schmidt modifié (MGS ou Modified Gram-Schmidt).

CGS est plus économique en termes d'occupation de mémoire mais elle est très sensible aux erreurs d'arrondi. La méthode MGS est numériquement plus stable. Cependant, il faut noter que ces deux méthodes seraient strictement équivalentes en l'absence d'erreurs d'arrondi [1]. Nous utilisons alors l'algorithme de Gram-Schmidt modifié.

Principe [2] :

L'algorithme d'orthogonalisation de Gram-Schmidt considère les valeurs prises par les prédicteurs et la sortie désirée comme des vecteurs. Soient

$$X = \begin{pmatrix} x_1^1 & \cdots & x_p^1 \\ \vdots & \ddots & \vdots \\ x_1^N & \cdots & x_p^N \end{pmatrix} = (X_1 \ X_2 \ \cdots \ X_p) \text{ la matrice des entrées,}$$

où $X_p = \begin{pmatrix} x_p^1 \\ \vdots \\ x_p^N \end{pmatrix}$ le vecteur de l'entrée p , et

$$Y = \begin{pmatrix} y^1 \\ \vdots \\ y^N \end{pmatrix} \text{ le vecteur de la sortie.}$$

N étant le nombre d'échantillons.

A chaque itération, il nous faut trouver le vecteur d'entrée qui explique le mieux la sortie. Pour cela, nous calculons le maximum du carré des cosinus des angles que font les vecteurs d'entrée avec le vecteur de sortie :

$$\max_{1 \leq i \leq p} [\cos^2(X_i, Y)] = \max_{1 \leq i \leq p} \left[\frac{(X_i^T Y)^2}{(X_i^T X_i)(Y^T Y)} \right] \quad 4.1$$

Et après chaque itération, nous éliminons le meilleur vecteur d'entrée, et sa contribution en projetant tous les vecteurs d'entrée restants et le vecteur de sortie sur son sous espace orthogonal.

Ce processus se termine lorsque tous les vecteurs d'entrée sont ordonnés.

4.1.3- Méthode de sélection de prédicteurs [1, 3]

Le principe de cette méthode est de comparer les prédicteurs initiaux avec des prédicteurs pris aléatoirement.

D'abord, il nous faut lancer l'algorithme de Gram-Schmidt sur la totalité des prédicteurs (100 prédicteurs aléatoires avec les prédicteurs initiaux). A chaque itération, nous choisissons le meilleur prédicteur en ne tenant pas compte des variables aléatoires. Nous détectons par la suite les variables aléatoires qui expliquent mieux la

sortie que le meilleur prédicteur, nous les comptions et nous les extrayons de l'ensemble des prédicteurs. Il ne reste plus qu'à projeter la sortie, les prédicteurs initiaux et aléatoires restants sur le sous espace orthogonal au prédicteur sélectionné. Et ainsi de suite jusqu'à ce que tous les prédicteurs initiaux sont ordonnés.

4.2- Application à la sélection des prédicteurs de l'intensité des cyclones tropicaux du BSOOI

4.2.1- Les données utilisées

Le JTWC (Joint Typhoon Warning Center) est un centre inter-armés de prévision des cyclones tropicaux entre la US Navy et la US Air Force. Il est installé au centre de prévision maritime de Pearl Harbor, Hawaii. Il a pour mission d'émettre les alertes pour les cyclones tropicaux pouvant affecter les bases militaires et aériennes américaines ainsi que la flotte de ce pays à travers les océans Pacifique et Indien. Il émet également des bulletins pour les populations civiles des territoires américains de Micronésie.

Les données que nous avons utilisées proviennent du site web du JTWC. Ce dernier maintient une archive de données de trajectoire de cyclones tropicaux, communément appelés les « meilleures trajectoires » (Best-Tracks). Les données contiennent les emplacements du centre des cyclones tropicaux et leurs intensités (c'est à dire, les vents maximums soutenus en vertu d'une moyenne sur une minute, mesurée à 10 mètres au-dessus de la surface) à des intervalles de six heures.

Seules les données cycloniques à partir de l'année 1985 sont considérées de meilleure qualité par le site du JTWC.

4.2.2- Les prédicteurs initiaux

Au départ, nous avons considéré 33 prédicteurs (*Tableau 4.1*) dérivés des variables climatologiques disponibles dans les données cycloniques (jour, longitude, latitude, vitesse de vent maximale). Le processus d'élimination décrit précédemment est appliqué à ces prédicteurs.

Tableau 4.1 : Les prédicteurs initiaux.

N°	Prédicteurs	N°	Prédicteurs
1	Jour Julien	18	Différence longitude 0h -- 6h
2	Age	19	Différence longitude 0h -- 12h
3	Longitude à -24h	20	Différence longitude 0h -- 18h
4	Longitude à -18h	21	Différence longitude 0h -- 24h
5	Longitude à -12h	22	Différence latitude 0h -- 6h
6	Longitude à -6h	23	Différence latitude 0h -- 12h
7	Longitude à 0h	24	Différence latitude 0h -- 18h
8	Latitude à -24h	25	Différence latitude 0h -- 24h
9	Latitude à -18h	26	Vitesse de déplacement 0h--6h
10	Latitude à -12h	27	Vitesse de déplacement -6h--12h
11	Latitude à -6h	28	Vitesse de déplacement -12h--18h
12	Latitude à 0h	29	Vitesse de déplacement -18h--24h
13	Vitesse de vent max à -24h	30	Différence Vitesse max 0h -- 6h
14	Vitesse de vent max à -18h	31	Différence Vitesse max 0h -- 12h
15	Vitesse de vent max à -12h	32	Différence Vitesse max 0h -- 18h
16	Vitesse de vent max à -6h	33	Différence Vitesse max 0h -- 24h
17	Vitesse de vent max à 0h		

4.2.3- Choix du nombre de prédicteurs

La *Figure 4.1* montre le nombre des prédicteurs aléatoires meilleurs que les prédicteurs initiaux. Après avoir choisi le nombre de prédicteur (valeur lue sur l'axe des abscisses) à utiliser dans le modèle (RNA et RLM dans notre cas), il y aura un risque (valeur lue en % sur l'axe des ordonnées) de choisir au moins un prédicteur moins potentiel qu'un prédicteur aléatoire.

On a ici un risque de 0% de choisir les 11 premiers prédicteurs ordonnés par la méthode de Gram-Schmidt pour 12h, les 9 premiers prédicteurs pour 24h, 36h et 48h et les 8 premiers prédicteurs pour 60h et 72h.

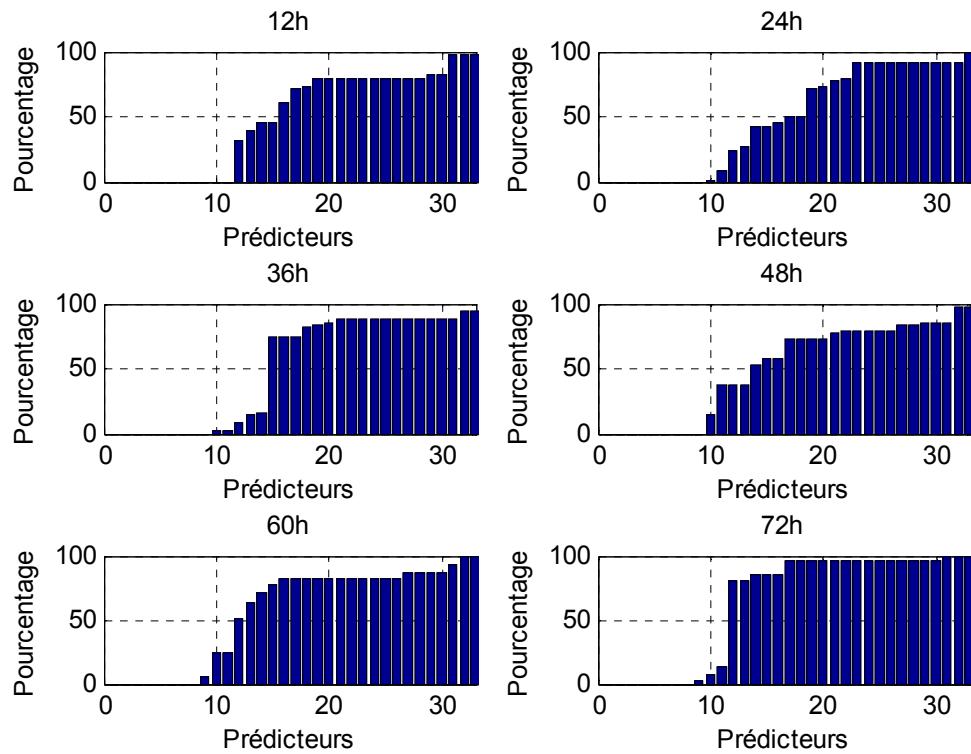


Figure 4.1: Choix du nombre de prédicteurs.

4.2.4- Rangement des prédicteurs

Le tableau suivant (*Tableau 4.2*) montre les 33 prédicteurs ordonnés pour chaque temps de prévision (12h, 24h, 36h, 48h, 60h, 72h).

CHAPITRE IV : PRÉVISION DES INTENSITÉS CYCLONIQUES PAR LES RÉSEAUX DE NEURONES

Tableau 4.2. Liste des prédicteurs classés par ordre de pertinence selon MGS.

Rang	12h	24h	36h	48h	60h	72h
1	Vmax 0h	Vmax 0h	Vmax 0h	Vmax 0h	DiffVmax 0h—24h	DiffVmax 0h—24h
2	Vmax -12h	Vmax -12h	Vmax -12h	DiffVmax 0h—12h	Vmax 0h	Vmax 0h
3	Vmax -6h	Vmax -6h	Lat 0h	Lat 0h	DiffLon 0h—6h	DiffLon 0h—6h
4	Lat -6h	Lat 0h	DiffVmax 0h—6h	DiffLon 0h—12h	Lat 0h	Lat 0h
5	DiffLon 0h—18h	DiffLon 0h—18h	DiffLon 0h—12h	Vmax -6h	Vmax -6h	DiffVmax 0h—6h
6	Vmax -24h	Age	Age	Age	Jour	Jour
7	Age	Vdepl -12h—18h	Vdepl -6h—12h	Jour	Age	Age
8	Vdepl -12h—18h	Jour	Jour	Vdepl -6h—12h	Vdepl -0h—6h	Vdepl -0h—6h
9	Jour	DiffVmax 0h—24h	DiffVmax 0h—24h	Vmax -24h	DiffVmax 0h—12h	Vmax -12h
10	DiffLat 0h—18h	DiffLat 0h—6h	Lat -6h	Vdepl -0h—6h	Lat -24h	DiffLat 0h—24h
11	DiffLon 0h—6h	Vdepl -6h—12h	Lat -24h	DiffLat 0h—6h	Lat -18h	Lat -18h
12	DiffVmax 0h—18h	Lon -24h	Vdepl -0h—6h	DiffLat 0h—24h	DiffLat 0h—6h	DiffLon 0h—24h
13	DiffLon 0h—24h	DiffVmax 0h—18h	Lat -18h	DiffLon 0h—6h	DiffLon 0h—18h	Vmax -18h
14	Vdepl -6h—12h	DiffLat 0h—24h	Lon 0h	Lon 0h	Vdepl -6h—12h	Vdepl -18h—24h
15	Vdepl -0h—6h	Lat -18h	DiffLon 0h—18h	Lon -18h	Lat -12h	DiffLon 0h—12h
16	DiffLat 0h—24h	Vdepl -0h—6h	Vdepl -12h—18h	DiffLat 0h—18h	DiffLon 0h—24h	DiffLon 0h—18h
17	Lon 0h	DiffLon 0h—24h	DiffLon 0h—6h	Lon -24h	Vdepl -18h—24h	Lat -24h
18	Lat 0h	Lon -6h	Vdepl -18h—24h	Vdepl -12h—18h	DiffVmax 0h—18h	Lon -24h
19	Vdepl -18h—24h	Lat -12h	Lon -24h	Vdepl -18h—24h	Lon -12h	Lon -6h
20	DiffLon 0h—12h	Lon -12h	DiffVmax 0h—18h	Vmax -18h	Lat -6h	Lon -18h
21	DiffLat 0h—12h	Lon -18h	Vmax -18h	Lat -6h	DiffLat 0h—12h	Lon 0h
22	Lat -12h	DiffLon 0h—12h	Vmax -24h	DiffLon 0h—18h	DiffVmax 0h—6h	DiffVmax 0h—12h
23	Lon -6h	DiffLon 0h—6h	DiffVmax 0h—12h	Lat -18h	DiffLat 0h—18h	DiffVmax 0h—18h
24	DiffVmax 0h—12h	DiffLat 0h—12h	Vmax -6h	Lon -12h	DiffLat 0h—24h	Vmax -24h
25	DiffLat 0h—6h	DiffLat 0h—18h	DiffLat 0h—6h	DiffLon 0h—24h	Vmax -24h	Vmax -6h
26	DiffVmax 0h—24h	Lon 0h	DiffLat 0h—24h	DiffVmax 0h—6h	Vmax -12h	DiffLat 0h—6h
27	Lat -24h	DiffVmax 0h—12h	Lon -6h	DiffVmax 0h—18h	Vmax -18h	Lat -6h
28	Lon -24h	Vmax -18h	DiffLon 0h—24h	Lat -24h	Lon -24h	Lon -12h
29	Vmax -18h	Lat -6h	Lon -12h	DiffVmax 0h—24h	Lon 0h	DiffLat 0h—18h
30	Lon -18h	Vmax -24h	DiffLat 0h—18h	Vmax -12h	Lon -18h	Vdepl -12h—18h
31	Lon -12h	DiffVmax 0h—6h	DiffLat 0h—12h	DiffLat 0h—12h	Lon -6h	DiffLat 0h—12h
32	Lat -18h	Lat -24h	Lat -12h	Lat -12h	DiffLon 0h—12h	Lat -12h
33	DiffVmax 0h—6h	Vdepl -18h—24h	Lon -18h	Lon -6h	Vdepl -12h—18h	Vdepl -6h—12h

4.2.5- Analyse des prédicteurs retenus

La sélection des prédicteurs dépend de la technique utilisée pour quantifier la liaison entre ces variables et la sortie. La liaison linéaire est la plus simple et la plus courante. Cependant, il est très important de contrôler ces liaisons par d'autres méthodes non linéaires. Le *Tableau 4.3* récapitule les prédicteurs choisis pour chaque échéance avec leur rang obtenu par la technique MGS.

Tableau 4.3. Récapitulation des prédicteurs sélectionnés et leurs rangs.

Prédicteurs	Rang					
	12h	24h	36h	48h	60h	72h
Vmax 0h	1	1	1	1	2	2
Vmax -6h	3	3		5	5	
Vmax -12h	2	2	2			
Vmax -24h	6			9		
Diffvmax 0h--6h			4			5
Diffvmax 0h--12h				2		
Diffvmax 0h--24h		9	9		1	1
Lat 0h		4	3	3	4	4
Lat -6h	4					
Difflon 0h--6h	11				3	3
Difflon 0h--12h			5	4		
Difflon 0h--18h	5	5				
Difflat 0h--18h	10					
Jour	9	8	8	7	6	6
Age	7	6	6	6	7	7
Vdepl 0h--6h					8	8
Vdepl -6h--12h			7	8		
Vdepl -12h--18h	8	7				

4.2.5.1- La vitesse maximale du vent (Vmax) :

Le premier prédicteur ordonné par la méthode MGS est la vitesse du vent maximal à 0h pour 12h à 48h (*Tableau 4.3*). Il se trouve au deuxième rang pour les prévisions à 60h et 72h. Ce résultat confirme la corrélation qui existe entre la vitesse du vent maximale à 0h et celles de 12h à 72h (*Figure 4.2*). Cependant, le coefficient de corrélation diminue lorsqu'on s'éloigne de 0h. Il est devenu très faible à 72h.

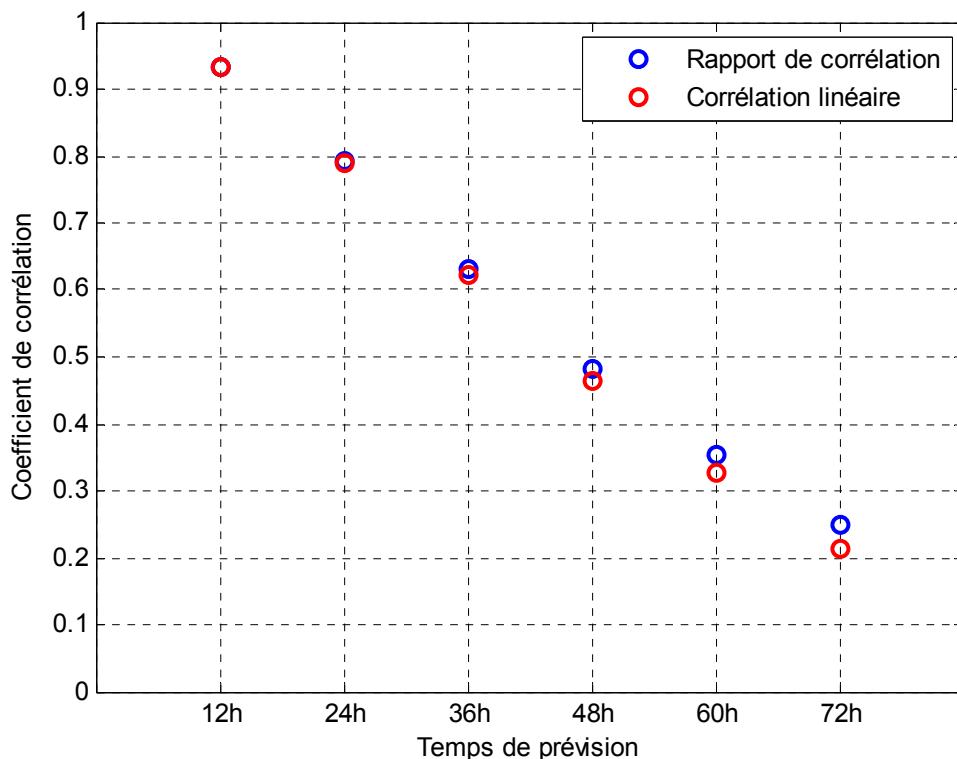


Figure 4.2 : Corrélations entre intensité à 0h et intensités de 12h à 72h.

Les deux mesures de corrélation (le coefficient de corrélation linéaire ρ_{xy}^2 et le rapport de corrélation $\eta_{y|x}^2$) sont confondues pour 12h et 24h. Cela signifie que la liaison entre les intensités à 0h, 12h et 24h est purement linéaire. Il est plus facile de prévoir une variable à partir des prédicteurs liés linéairement avec elle. Pour la variable Vmax à 0h, l'écart entre ρ_{xy}^2 et $\eta_{y|x}^2$ croît faiblement avec le temps de prévision.

Les *Figures 4.3* et *4.4* montrent les liaisons des prédicteurs Vmax à -6h et à -24h avec les intensités de 12h à 72h. Les valeurs des coefficients diminuent lorsqu'on utilise

les prédicteurs loin des prédictants. Les écarts entre entre ρ_{XY}^2 et $\eta_{Y|X}^2$ croit rapidement pour Vmax à -24h par rapport à Vmax à 0h.

Cependant, les variables intensités cycloniques présentent des coefficients de corrélation assez élevés par rapport aux autres prédicteurs.

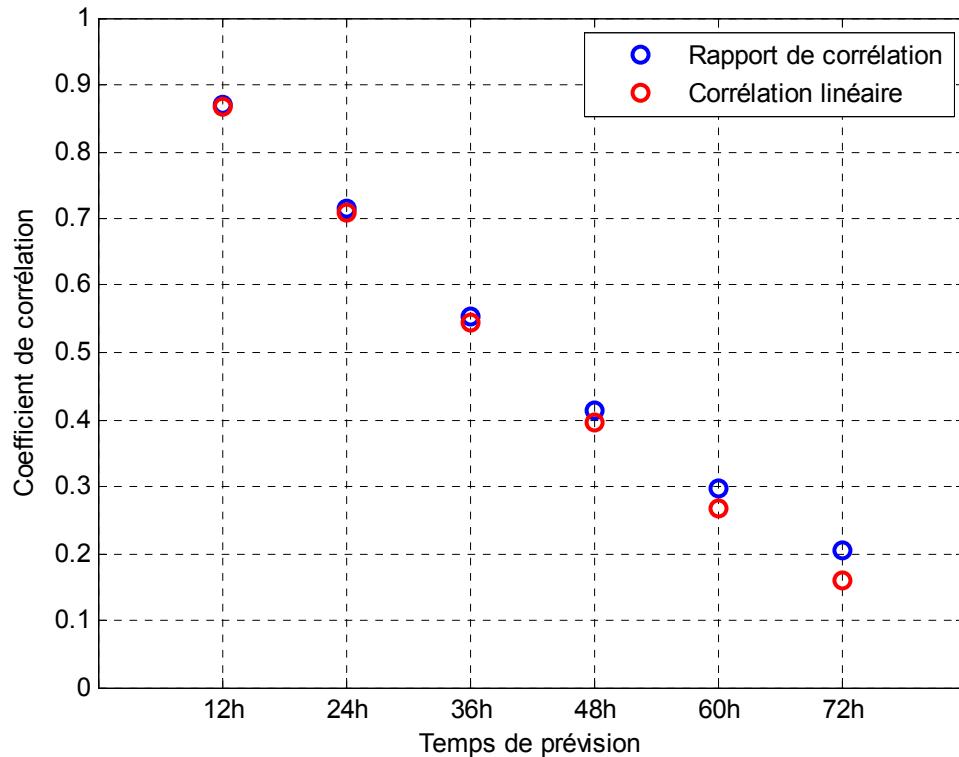


Figure 4.3 : Corrélations entre intensité à -6h et intensités de 12h à 72h.

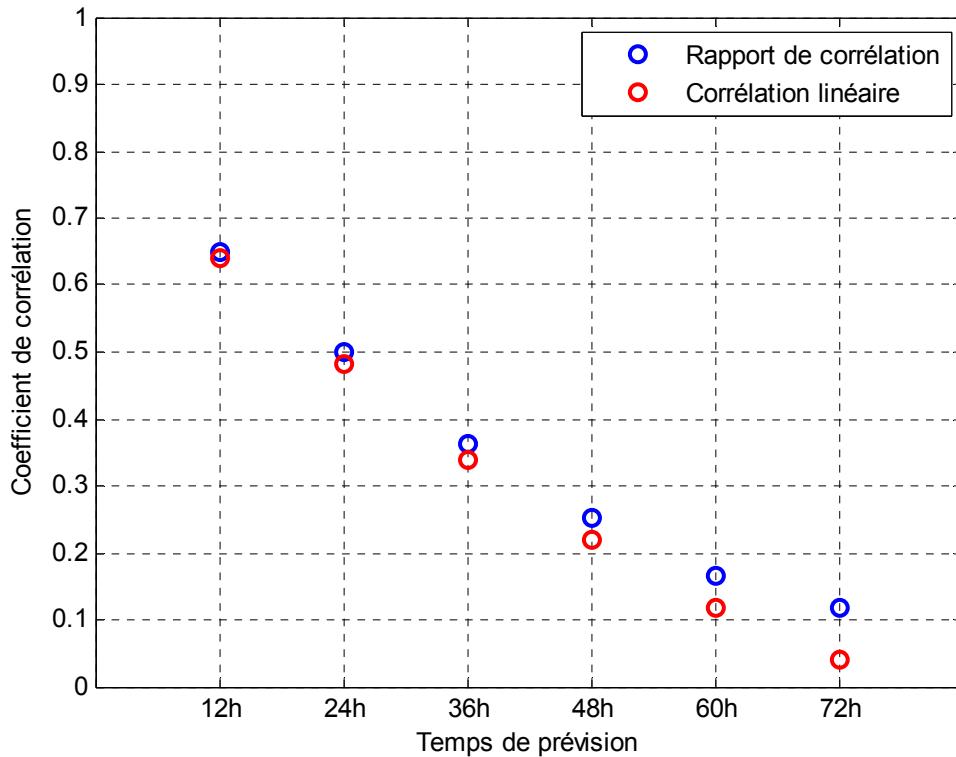


Figure 4.4 : Corrélations entre intensité à -24h et intensités de 12h à 72h.

4.2.5.2- Changement d'intensités :

Le changement d'intensité de -24h à 0h est le meilleur prédicteur pour les prévisions à 60h et 72h (*Tableau 4.3*). C'est la différence entre Vmax à 0h et Vmax à -24h. La *Figure 4.5* montre l'importance de l'utilisation de cette variable. En effet, la connaissance de Vmax à 0h ne suffit pas toujours pour dire la future intensité ; une valeur de Vmax (à 0h) égale à 80 noeuds, par exemple, ne permet pas de dire la valeur de Vmax après 24h. On a ici deux valeurs : 95 noeuds et 35 noeuds.

Par contre, on a deux valeurs de changements d'intensité (-24h à 0h) pour 80 noeuds : +50 noeuds et -35 noeuds ; le signe indique généralement le sens de la variation de l'intensité et la valeur désigne la pente de cette variation, permettant ainsi de prévoir l'évolution future de l'intensité cyclonique.

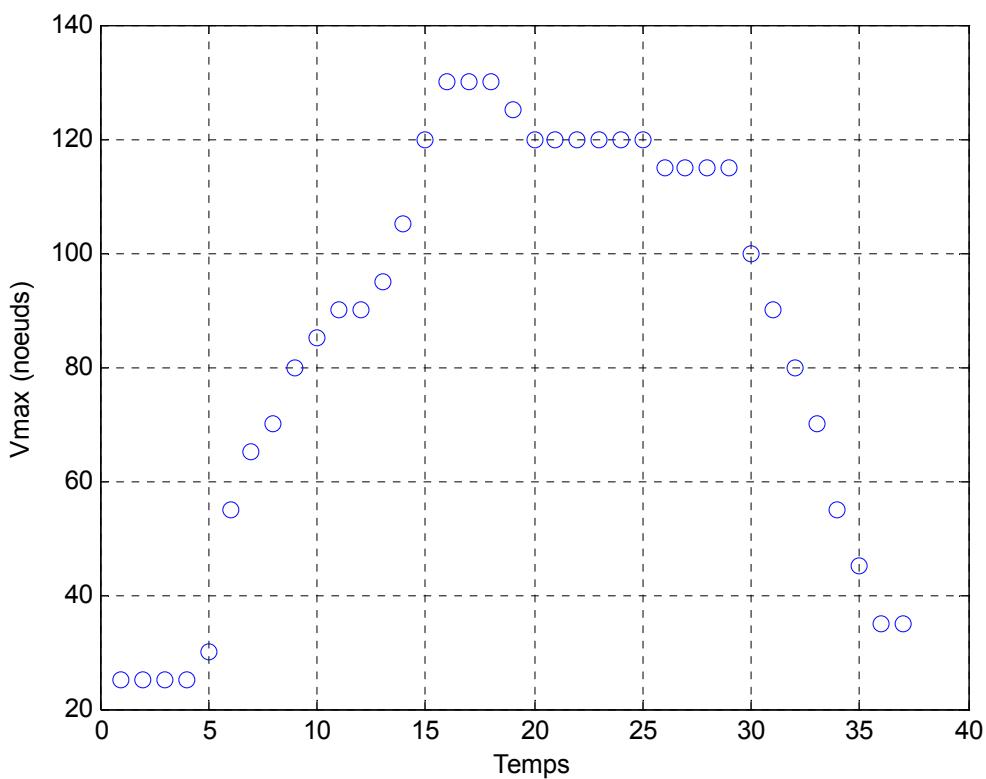


Figure 4.5 : Exemple d'allure de la variation de l'intensité du cyclone en fonction du temps.

Les Figures 4.6, 4.7 et 4.8 montrent l'évolution des coefficients de corrélation entre les variations d'intensités (0h/-6h, 0h/-12h, 0h/-24h) et les intensités de 0h à 72h. Les courbes présentent des valeurs maximales pour 24h ou 36h. C'est là que les liaisons sont importantes. Mais il faut aussi tenir compte les écarts entre les deux coefficients de corrélation. On voit une diminution des écarts (de 12h à 72h) pour les trois prédicteurs. L'écart le plus faible est à 72h. Il est alors plus facile de prévoir les intensités à longues échéances (60h et 72h) avec ces prédicteurs obtenus à partir de la différence des intensités.

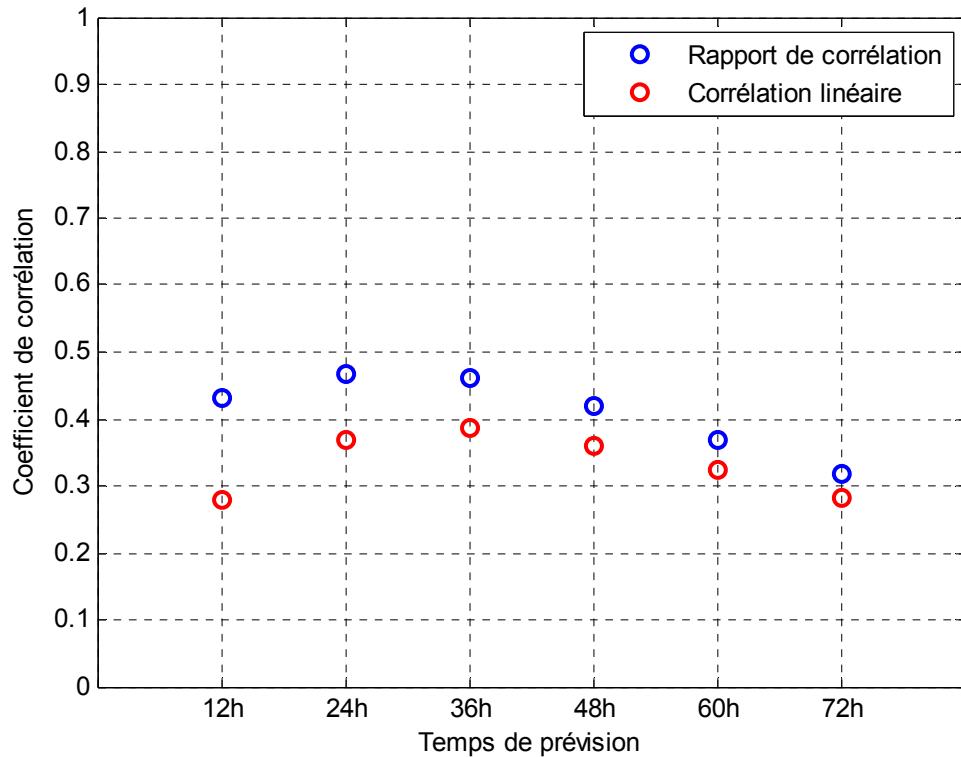


Figure 4.6 : Corrélations entre Différence des intensités à 0h / -6h et intensités de 12h à 72h.

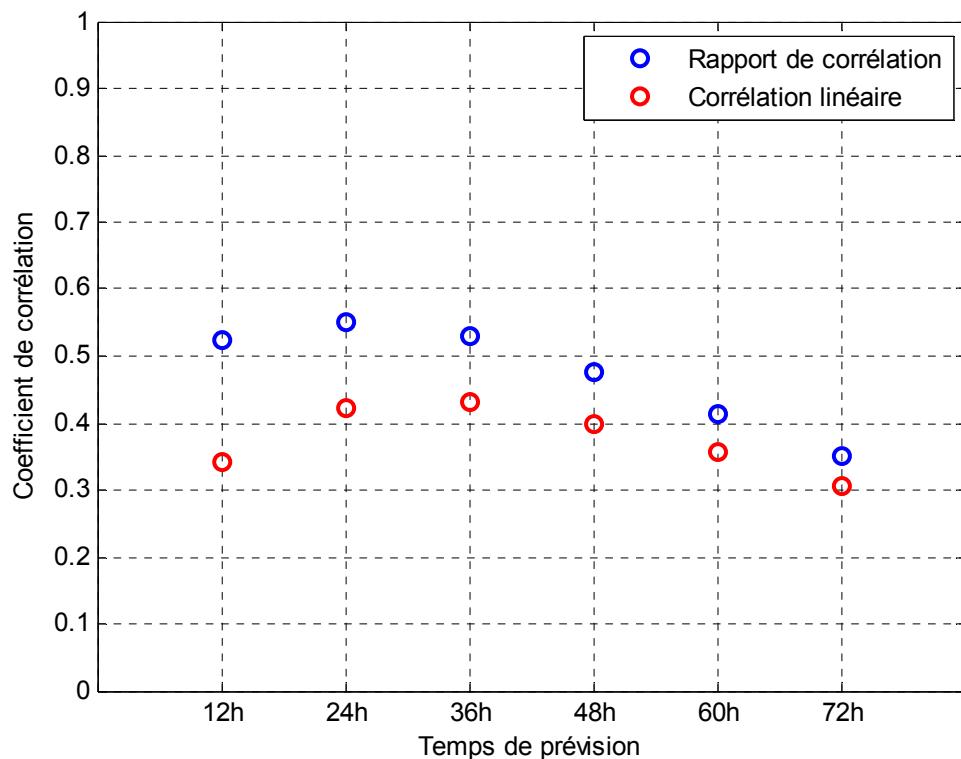


Figure 4.7 : Corrélations entre Différence des intensités à 0h / -12h et intensités de 12h à 72h.

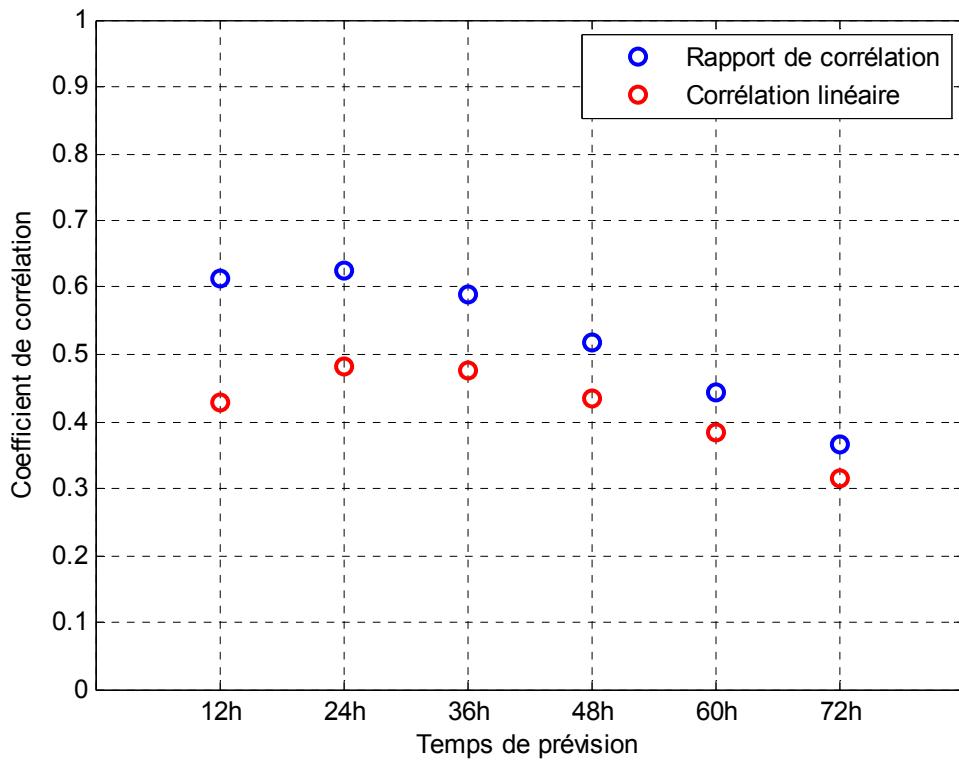


Figure 4.8 : Corrélations entre Différence des intensités à 0h / -24h et intensités de 12h à 72h.

4.2.5.3- Latitude :

Comme la mer agit telle qu'une source chaude qui alimente le cyclone [4] (comme une machine thermique), la variation de la température à la surface de la mer (TSM) influence son intensité. En effet, on sait que la température maximale est au niveau de l'équateur. Elle diminue lorsqu'on s'approche des pôles. Le déplacement du cyclone vers le sud (dans l'hémisphère sud) devra alors influencer sur son intensité. On a vu que la température de démarrage des cyclones au niveau de la surface de la mer est environ 26°C. La vie des cyclones commence auprès de l'équateur dans la zone chaude et se termine dans la zone plus froide, loin de l'équateur. La relation qui existe entre la latitude et l'intensité du cyclone tropical est alors due à celle qui existe entre la TSM et la latitude. La latitude est un bon prédicteur pour toutes les échéances.

Les *Figure 4.9* et *4.10* montrent les coefficients de corrélation entre les latitudes (-6h et 0h) et les intensités de 12h à 72h. Ces coefficients présentent des valeurs minimales à 36h et des valeurs maximales à 72h. Les écarts sont minimaux à 72h. Ces

prédicteurs sont alors bons pour la prévision à longue échéance. Cependant, la différence avec les autres échéances est relativement faible.

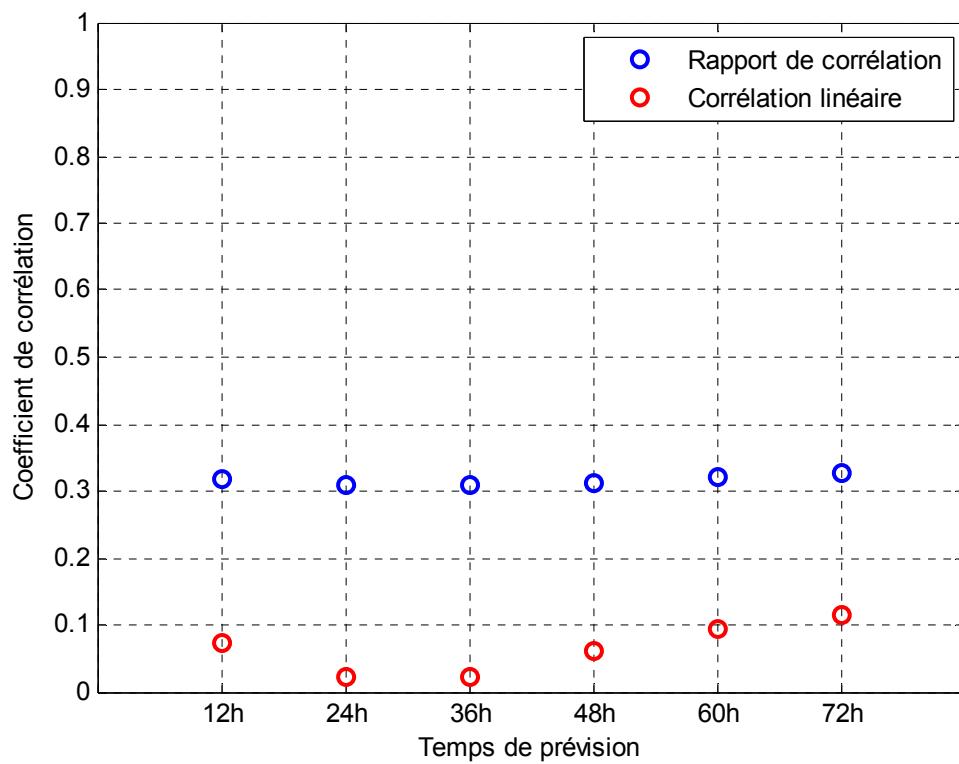


Figure 4.9 : Corrélations entre Latitude à -6h et intensités de 12h à 72h.

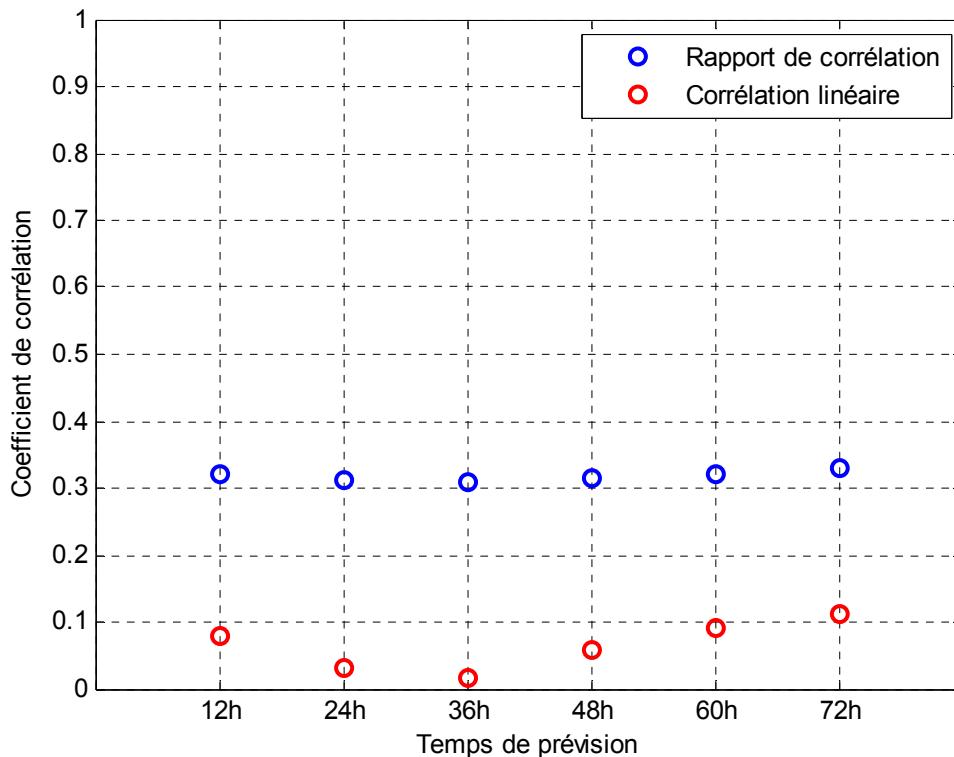


Figure 4.10 : Corrélations entre Latitude à 0h et intensités de 12h à 72h.

4.2.5.4- Variation des longitudes :

C'est la différence entre les longitudes à 0h et à -6h, -12h, -18h, -24h. Ces variables caractérisent le déplacement du cyclone suivant la longitude. En effet, ces différences sont calculées sur une durée de 6h, 12h, 18h ou 24h. Leurs valeurs absolues ne sont autres que les vitesses de déplacement des cyclones et leurs signes indiquent le sens du déplacement vers l'Ouest ou vers l'Est.

Les Figures 4.11, 4.12, 4.13 et 4.14 montrent l'augmentation des coefficients de corrélation. Le déplacement du cyclone suivant la longitude est très important surtout pour les prévisions à longues échéances. Ce déplacement situe le phénomène cyclonique dans sa phase de développement. Généralement, les cyclones suivent une trajectoire parabolique décomposable en 3 mouvements successifs : mouvement en direction de l'ouest-sud-ouest, puis incurvation vers le sud et enfin, évacuation vers l'est-sud-est aux latitudes extratropicales.

Les écarts entre le coefficient de corrélation linéaire et le rapport de corrélation restent presque constants.

CHAPITRE IV : PRÉVISION DES INTENSITÉS CYCLONIQUES PAR LES RÉSEAUX DE NEURONES

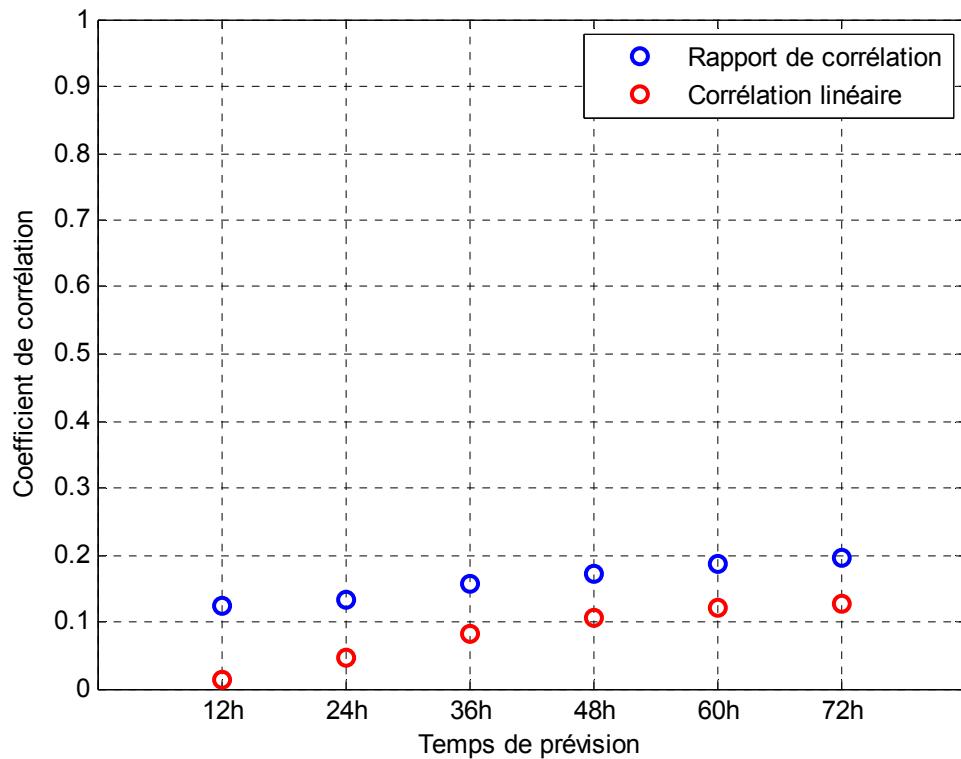


Figure 4.11 : Corrélations entre Différence longitudes 0h/-6h et intensités de 12h à 72h.

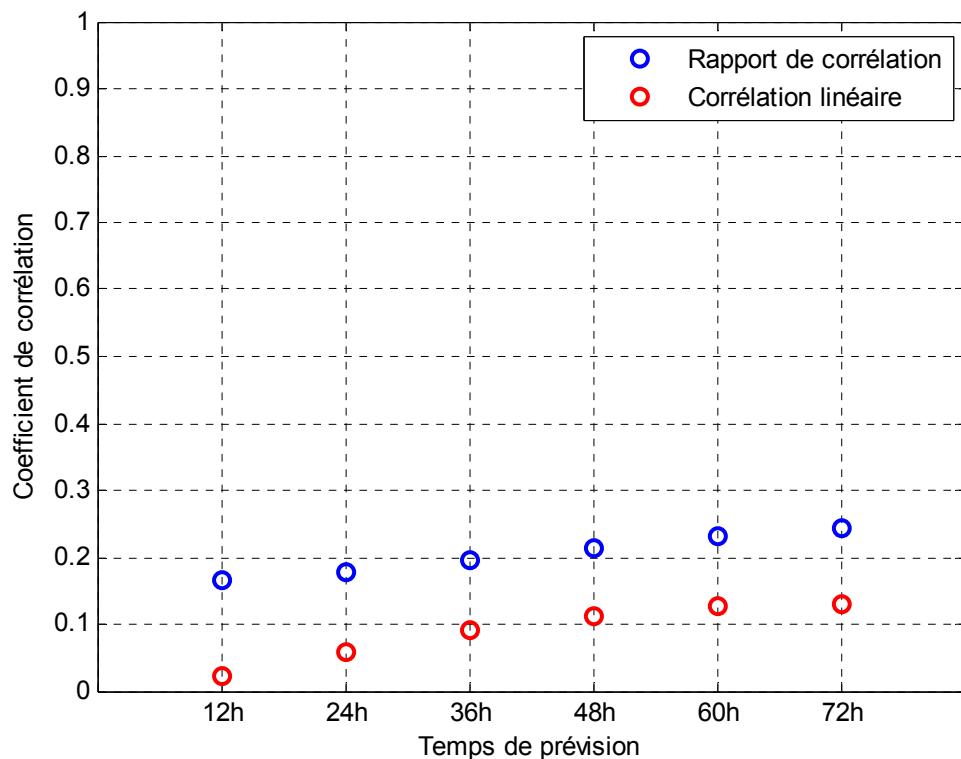


Figure 4.12 : Corrélations entre Différence longitudes 0h/-12h et intensités de 12h à 72h.

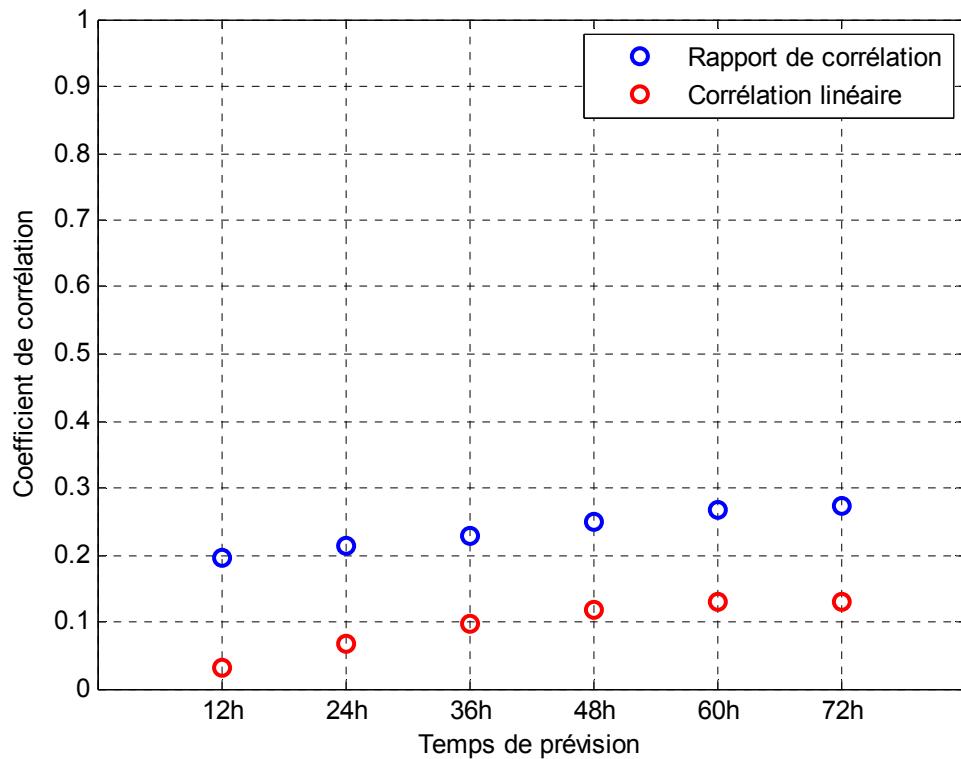


Figure 4.13 : Corrélations entre Différence longitudes 0h/-18h et intensités de 12h à 72h.

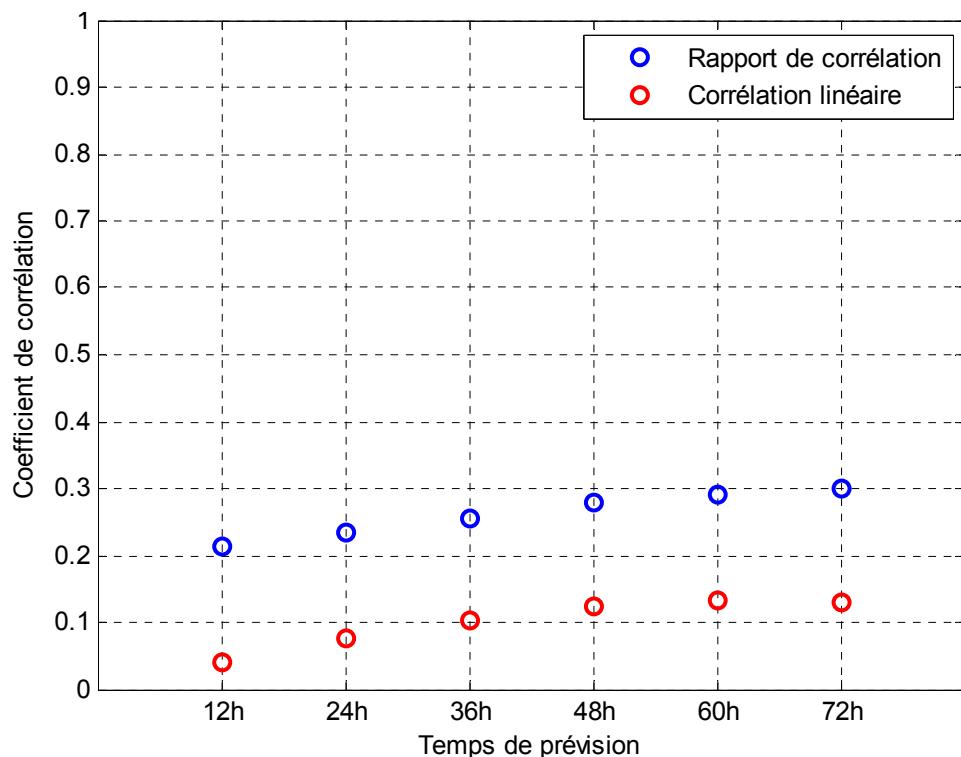


Figure 4.14 : Corrélations entre Différence longitudes 0h/-24h et intensités de 12h à 72h.

4.2.5.5- Jour julien :

C'est le rang du jour dans l'année dont on commence le comptage le 1^{er} septembre jusqu'au 31 aout de l'année suivante. Dans cette prévision, ce prédicteur correspond au rang du jour où se situe le 0h du cyclone.

Le jour julien a une contribution relativement significative par rapport aux autres prédicteurs pour la prévision de l'intensité cyclonique. Cette variable situe le phénomène dans le temps où d'autres phénomènes comme le passage de la Zone de Convergence Inter Tropicale (ZCIT) a lieu.

Le coefficient de corrélation linéaire est presque constant de 12h à 72h (*Figure 4.15*). L'écart entre les coefficients de corrélation est minimum pour 12h. Cependant, le jour julien est aussi bon pour les autres temps de prévision (*Tableau 4.3*).

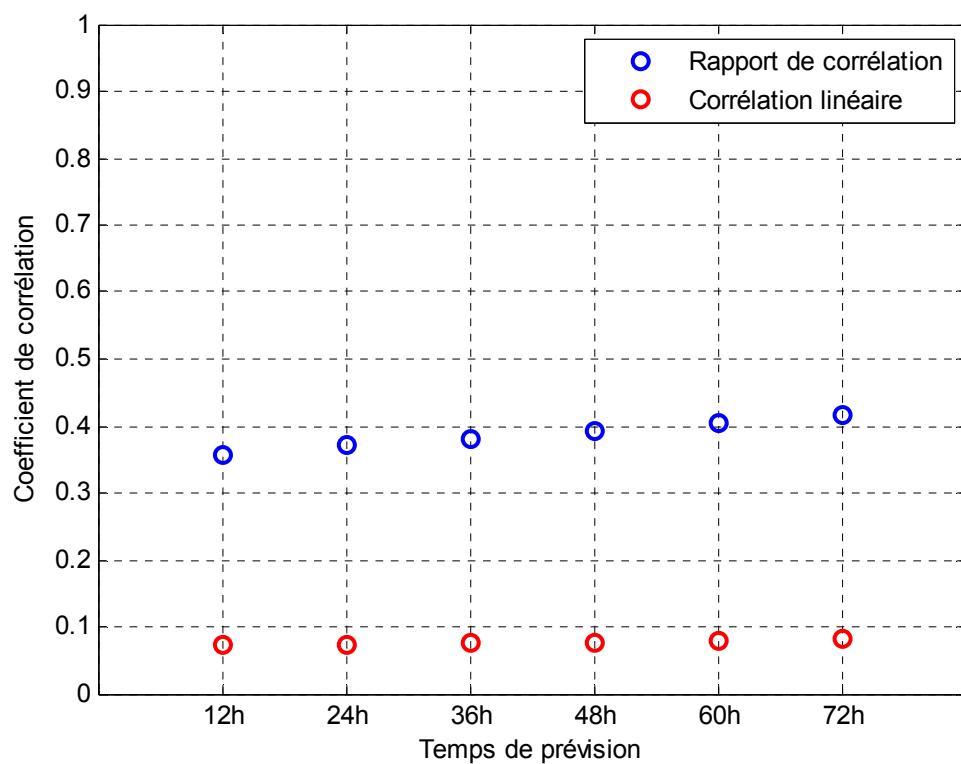


Figure 4.15 : Corrélations entre Jour et intensités de 12h à 72h.

4.2.5.6- Age :

Cette variable est obtenue en choisissant une origine des dates par rapport à l'intensité du phénomène cyclonique. Les temps où l'intensité passe de 35 noeuds sont choisis comme origines des dates. On compte alors négativement les temps passés avant cette origine et positivement ceux écoulés après. L'allure de la variation de l'intensité cyclonique à 72h en fonction de l'âge est représentée sur la *Figure 4.16*. En général, l'intensité est maximale quelques heures après l'origine choisie (courbe en rouge sur la figure).

La *Figure 4.17* montre les effectifs et les écart-types de cette distribution à chaque valeur de l'âge. L'effectif maximal correspond à l'âge où l'intensité est maximale.

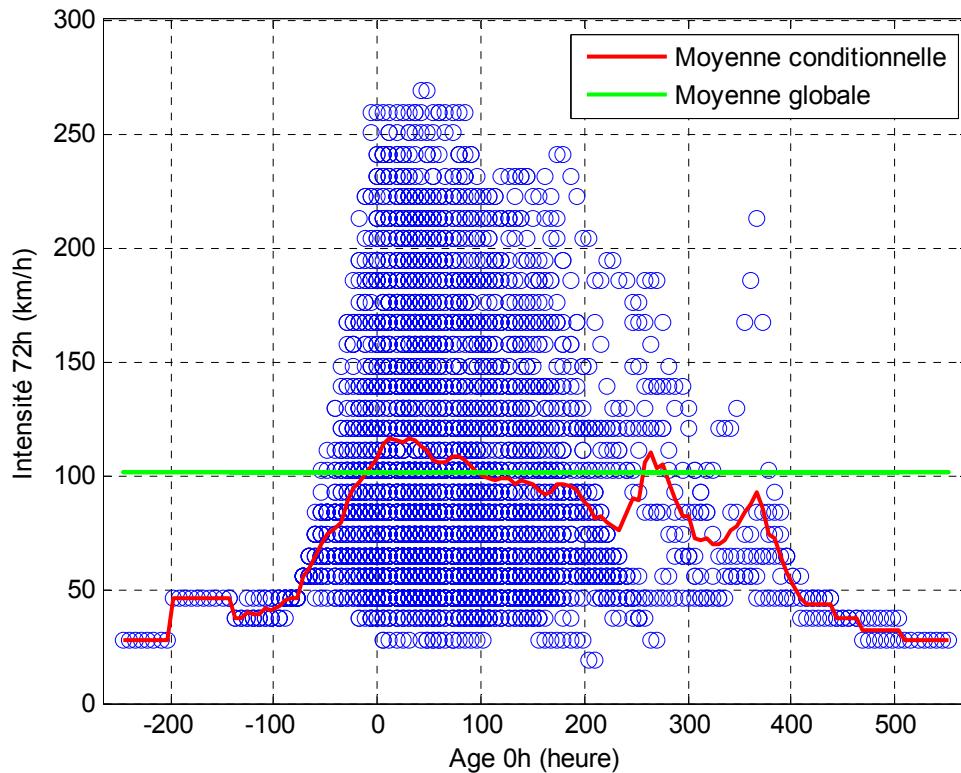


Figure 4.16 : Intensité à 72h en fonction de l'âge du phénomène cyclonique.

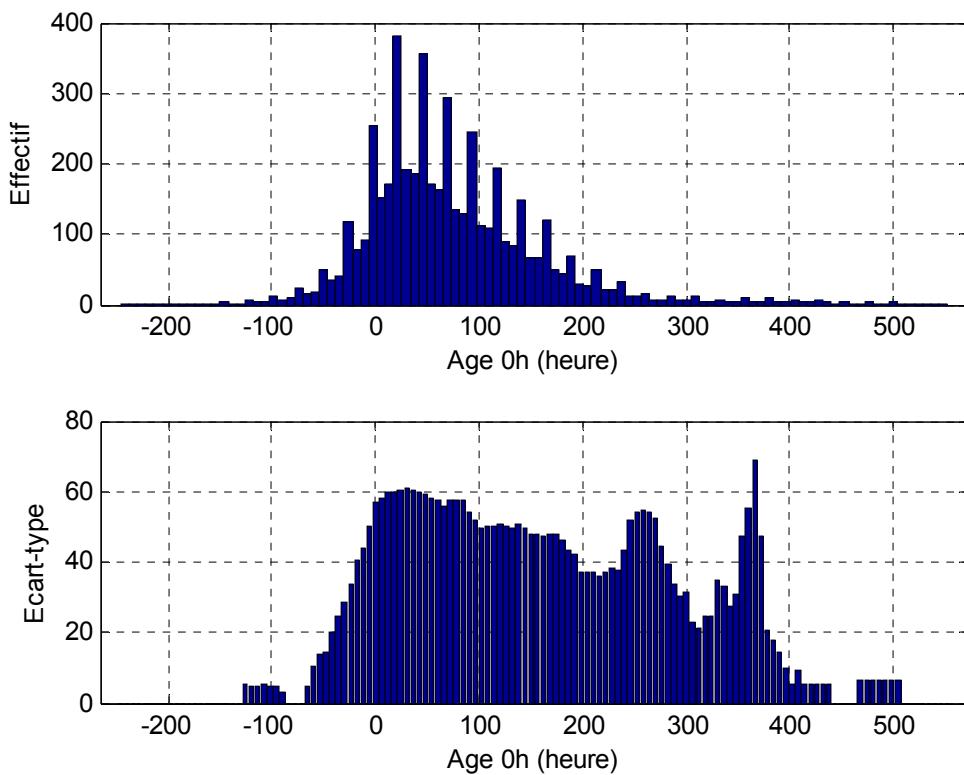


Figure 4.17 : Effectifs (en haut) et écart-types (en bas) des intensités à 72h pour chaque âge.

L'âge situe le phénomène dans le temps. Nous savons que l'intensité du phénomène cyclonique croît, généralement, avant d'atteindre un maximum, puis diminue. En utilisant cette variable, on peut savoir la phase de développement du phénomène, et cela va faciliter sa prévision.

La *Figure 4.18* montre la variation du coefficient de corrélation en fonction des temps de prévision. Le coefficient de corrélation linéaire est minimum pour 36h. Après 36h, ce coefficient augmente et l'écart entre le coefficient de corrélation linéaire et le rapport de corrélation diminue. Il est alors plus facile de prévoir l'intensité à 72h en utilisant l'âge à 0h du cyclone. Le *Tableau 4.3* montre que ce prédicteur présente presque le même rang pour tous les temps de prévision.

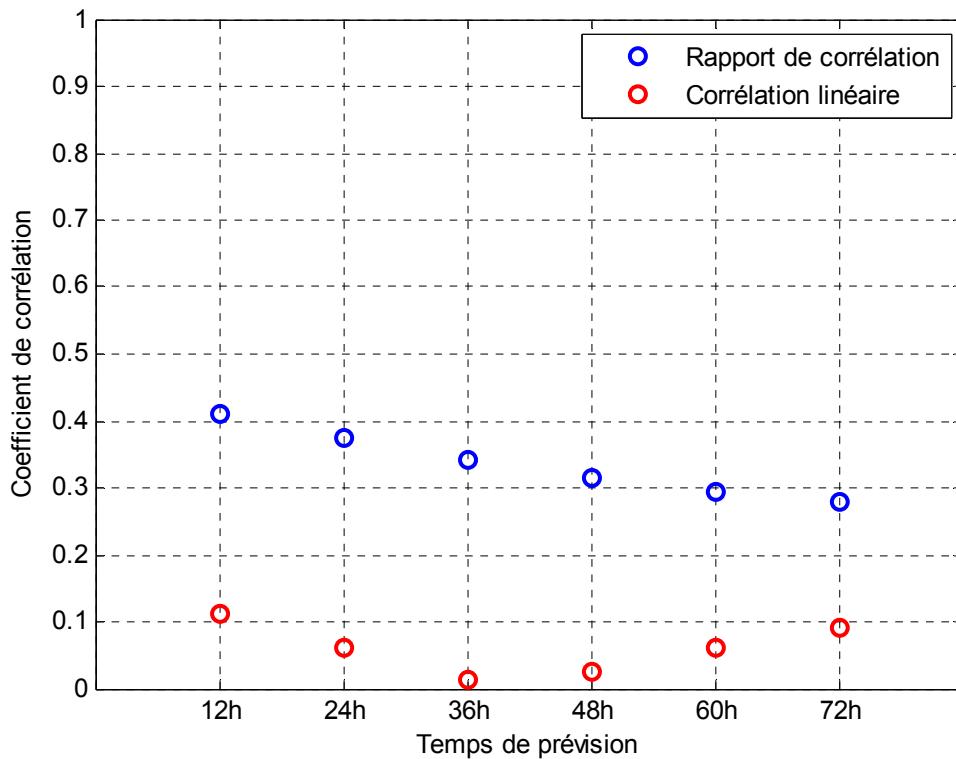


Figure 4.18 : Corrélations entre Age 0h et intensités de 12h à 72h.

4.2.5.7- Vitesse de déplacement :

Cette variable est calculée en suivant un arc de la surface de la Terre sur un intervalle de 6h. La vitesse de déplacement d'un cyclone détermine le temps de son contact avec une partie de la surface de la mer. Plus il se déplace assez vite, moins ce temps de contact est faible. Cela peut augmenter ou diminuer l'intensité du cyclone selon la quantité d'énergie utilisable dans la mer en dessous.

Les Figures 4.19, 4.20 et 4.21 montrent la variation des coefficients de corrélation entre les vitesses de déplacement du phénomène cyclonique et les intensités de 12h à 72h. Les coefficients de corrélation linéaires sont très faibles. Cependant, on voit une légère augmentation de ce coefficient jusqu'à atteindre une valeur maximale pour 72h.

L'écart entre le coefficient de corrélation et le rapport de corrélation est énorme. Les rapports de corrélation sont proches de 1. Cela signifie que l'intensité et la vitesse de déplacement présente une liaison fonctionnelle assez importante. A chaque valeur de la vitesse de déplacement correspond presque une valeur de l'intensité. Mais cette relation

est très complexe. C'est pourquoi les coefficients de corrélation linéaire sont très faibles et les écarts entre ces deux coefficients sont élevés.

Cependant, ces prédicteurs sont parmi les sélectionnés par la méthode MGS et le Tableau 4.3 montre que Vdepl 0h/-6h est bon pour les prévisions à 60h et 72h, Vdepl -6h/-12h pour 36h et 48h et Vdepl -12h/-18h pour 12h et 24h.

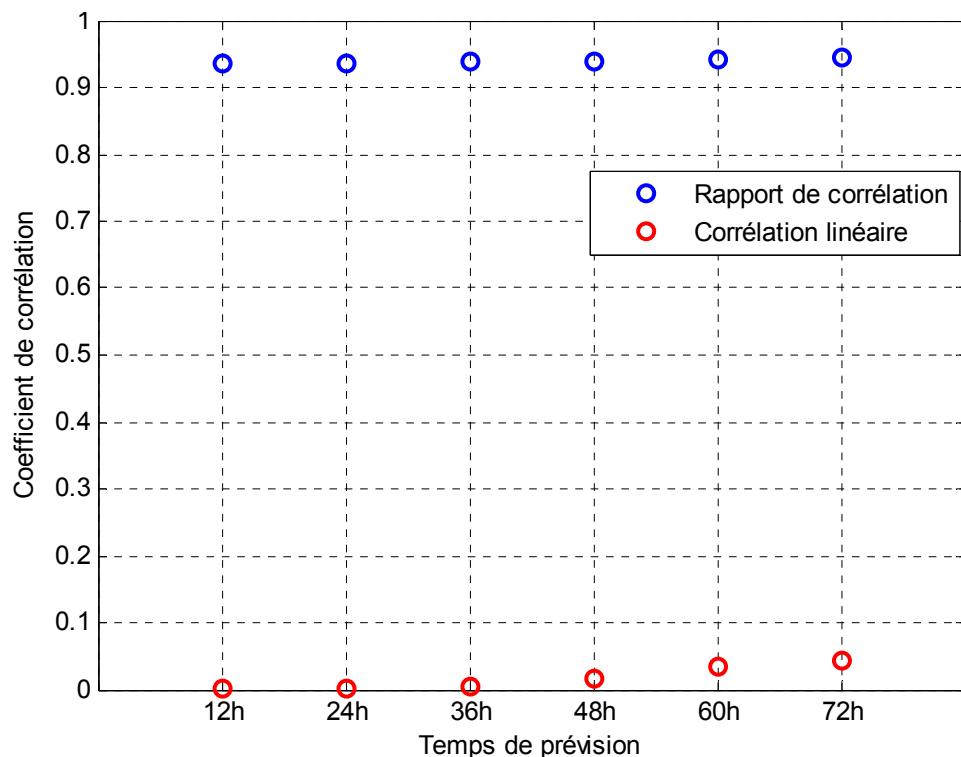


Figure 4.19 : Corrélations entre vitesse de déplacement 0h/-6h et intensités de 12h à 72h.

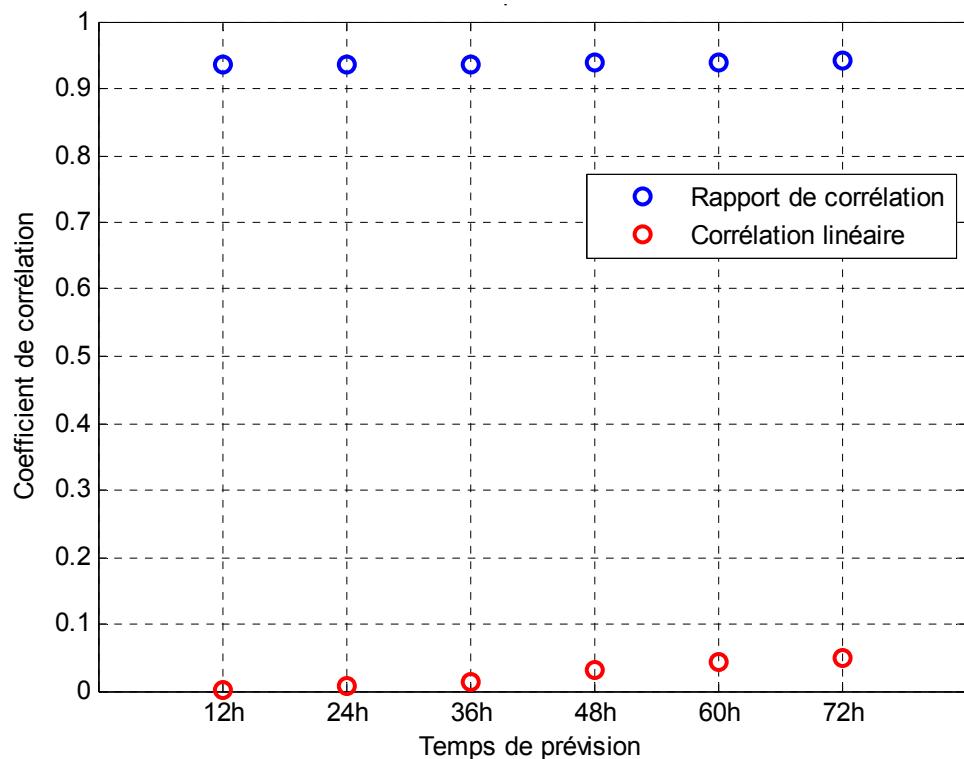


Figure 4.20 : Corrélations entre vitesse de déplacement -6h/-12h et intensités de 12h à 72h.

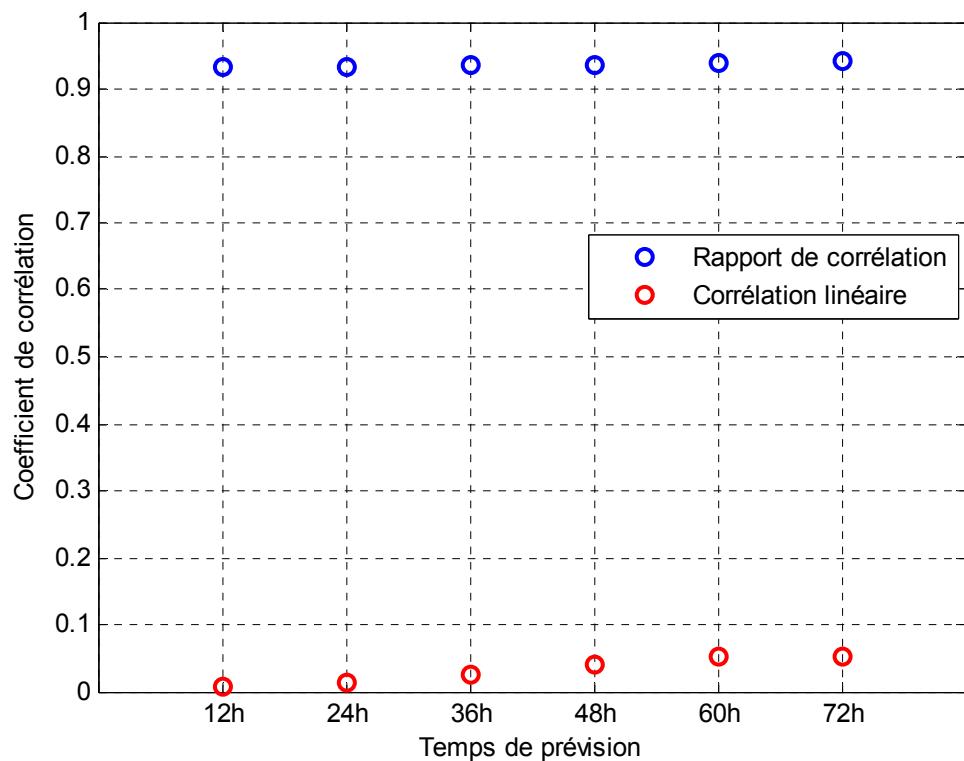


Figure 4.21 : Corrélations entre vitesse de déplacement -12h/-18h et intensités de 12h à 72h.

4.3- Réseau de neurones utilisé

Dans ce travail, un réseau de neurones à une seule couche cachée est utilisé (*Figure 4.22*). Les fonctions d'activation dans la couche cachée est la fonction sigmoïde tandis que pour la couche de sortie, on utilise une fonction linéaire.

La sortie du réseau est l'intensité cyclonique à 12h, 24h, 36h, 48h, 60h ou 72h. On utilise les prédicteurs sélectionnés dans les paragraphes précédents pour déterminer les entrées du modèle. Au départ, on n'utilise que le premier prédicteur choisi et on fait varier le nombre de neurones de la couche cachée. Puis, on prend les deux premiers prédicteurs choisis et on fait varier le nombre de neurones de la couche cachée, et ainsi de suite jusqu'aux douze premiers prédicteurs.

On cherche alors l'architecture optimale du réseau de neurones dans l'espace de 12 prédicteurs, 15 neurones cachés et une sortie.

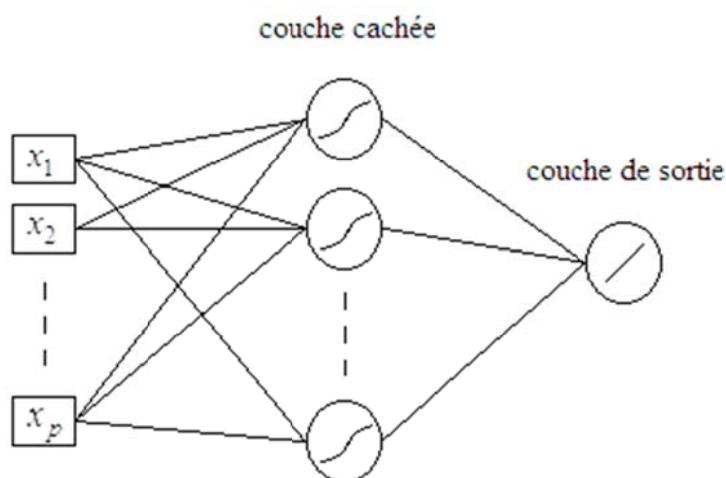


Figure 4.22 : Réseau de neurones à une seule couche cachée

Nous avons choisi la méthode quasi-Newton pour l'optimisation de notre réseau.

Les Figures 4.23 à 4.28 montrent les erreurs quadratiques en fonction du nombre de prédicteurs et du nombre de neurones cachés. On voit qu'il est difficile de préciser à priori la structure du réseau de neurones. Il y a une forte variabilité de l'erreur lorsqu'on modifie les paramètres. C'est pour cela que nous utilisons cette technique de varier les nombres de prédicteurs et neurones cachés.

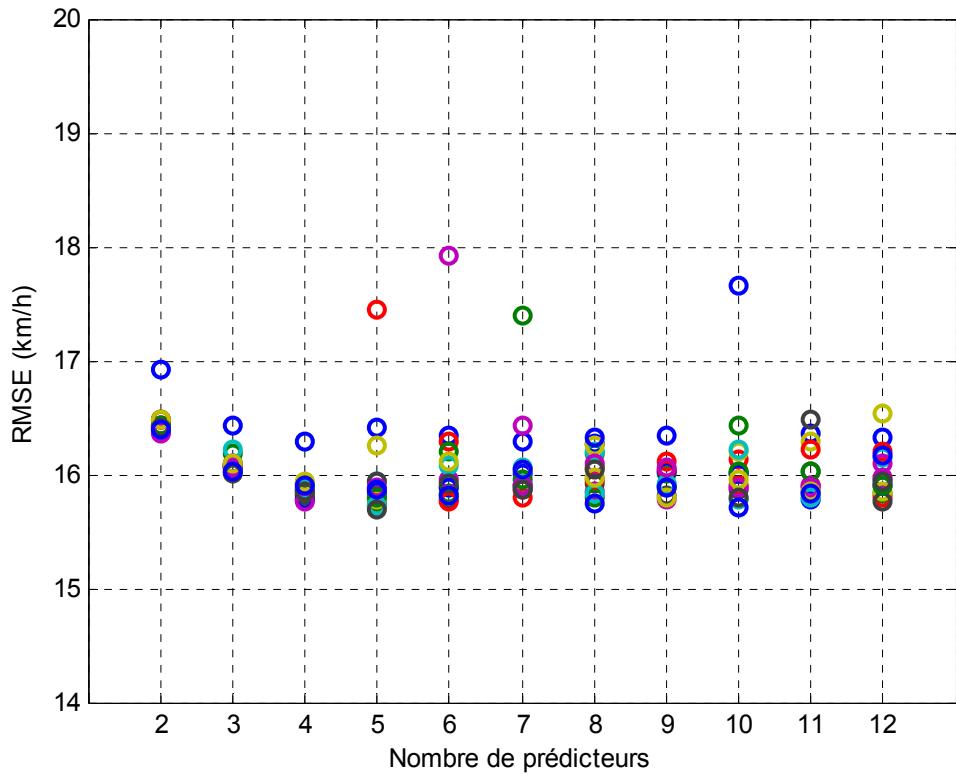


Figure 4.23 : Erreurs de prévision à 12h en fonction du nombre de prédicteurs et du nombre de neurones cachés

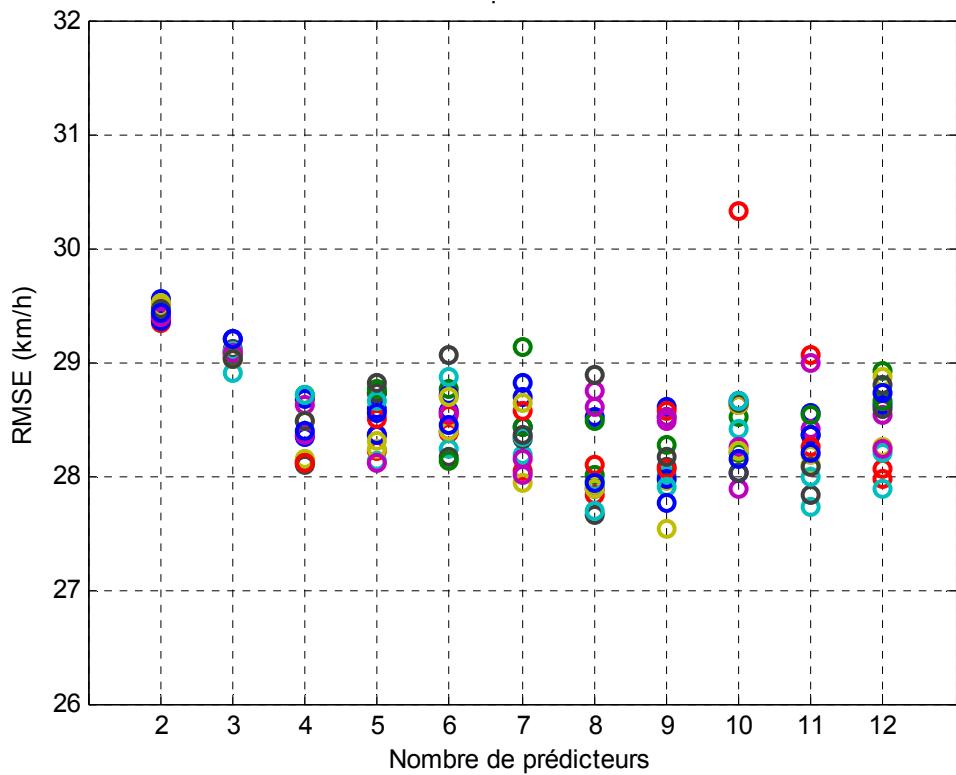


Figure 4.24 : Erreurs de prévision à 24h en fonction du nombre de prédicteurs et du nombre de neurones cachés

CHAPITRE IV : PRÉVISION DES INTENSITÉS CYCLONIQUES PAR LES RÉSEAUX DE NEURONES

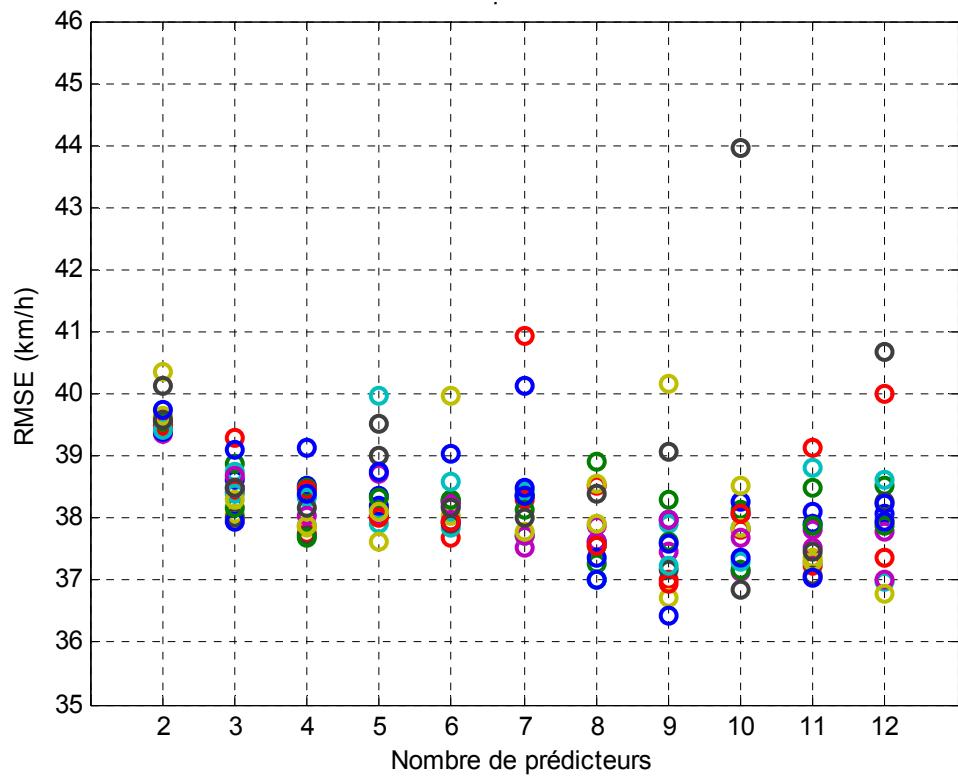


Figure 4.25 : Erreurs de prévision à 36h en fonction du nombre de prédicteurs et du nombre de neurones cachés

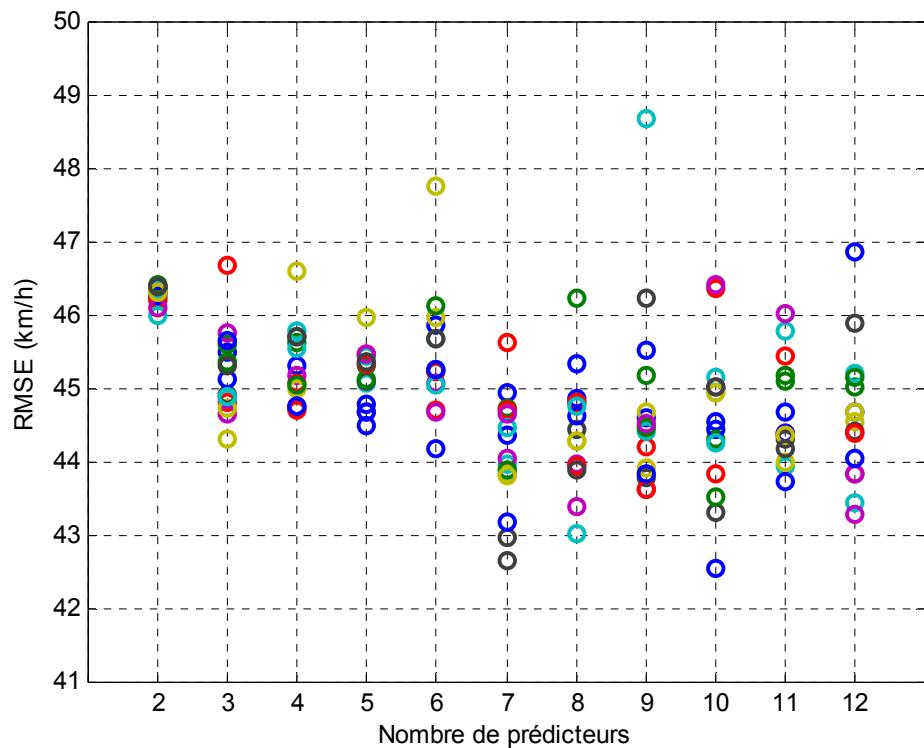


Figure 4.26 : Erreurs de prévision à 48h en fonction du nombre de prédicteurs et du nombre de neurones cachés

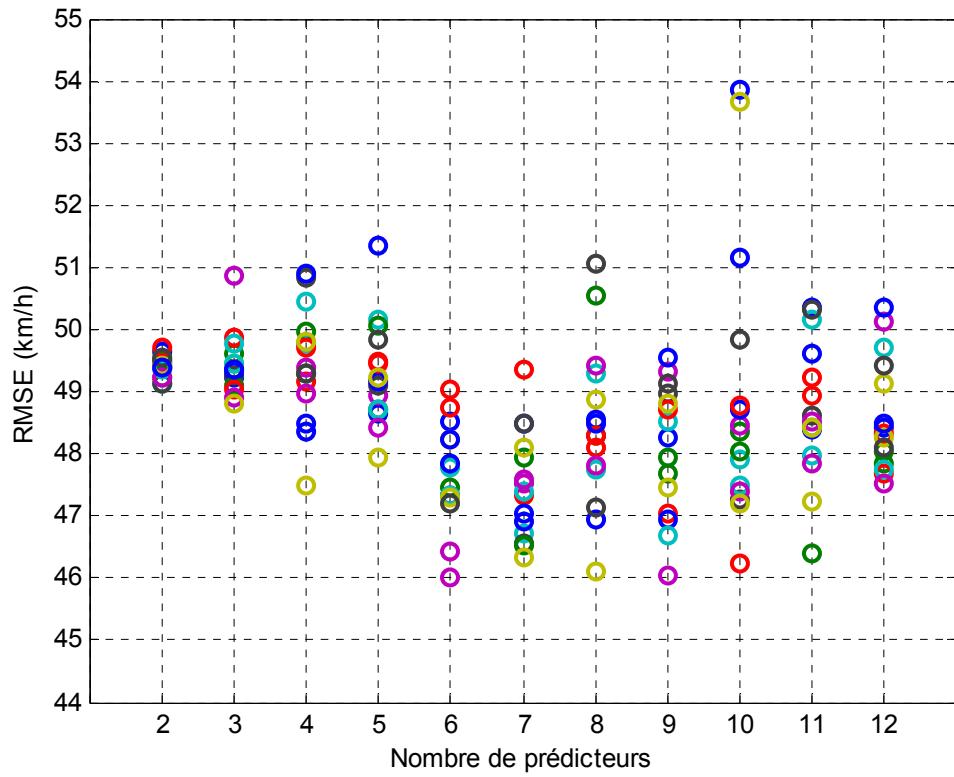


Figure 4.27 : Erreurs de prévision à 60h en fonction du nombre de prédicteurs et du nombre de neurones cachés

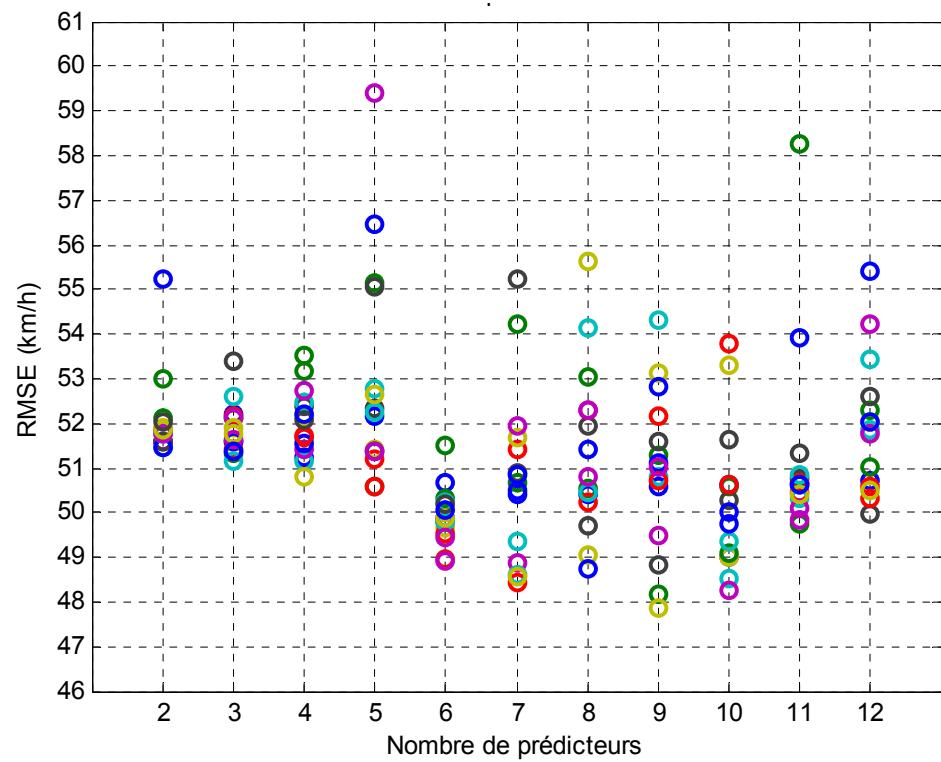


Figure 4.28 : Erreurs de prévision à 72h en fonction du nombre de prédicteurs et du nombre de neurones cachés

Le *Tableau 4.4* montre les nombres de prédicteurs et nombres de neurones cachés donnant des erreurs quadratiques minimales.

Tableau 4.4: Récapitulation des nombres de prédicteurs et de neurones cachés optimum.

Temps de prévision	12h	24h	36h	48h	60h	72h
Nombre de prédicteurs	5	9	9	10	6	9
Nombre de neurones cachés	14	13	8	15	5	13

4.4- Performance du modèle

Dans ce travail, nous avons choisi le modèle neuronal pour prévoir l'intensité cyclonique dans le bassin sud-ouest de l'océan indien. La *Figure 4.29* montre que, en moyenne, le RNA est plus performant que le modèle linéaire pour tous les temps de prévision.

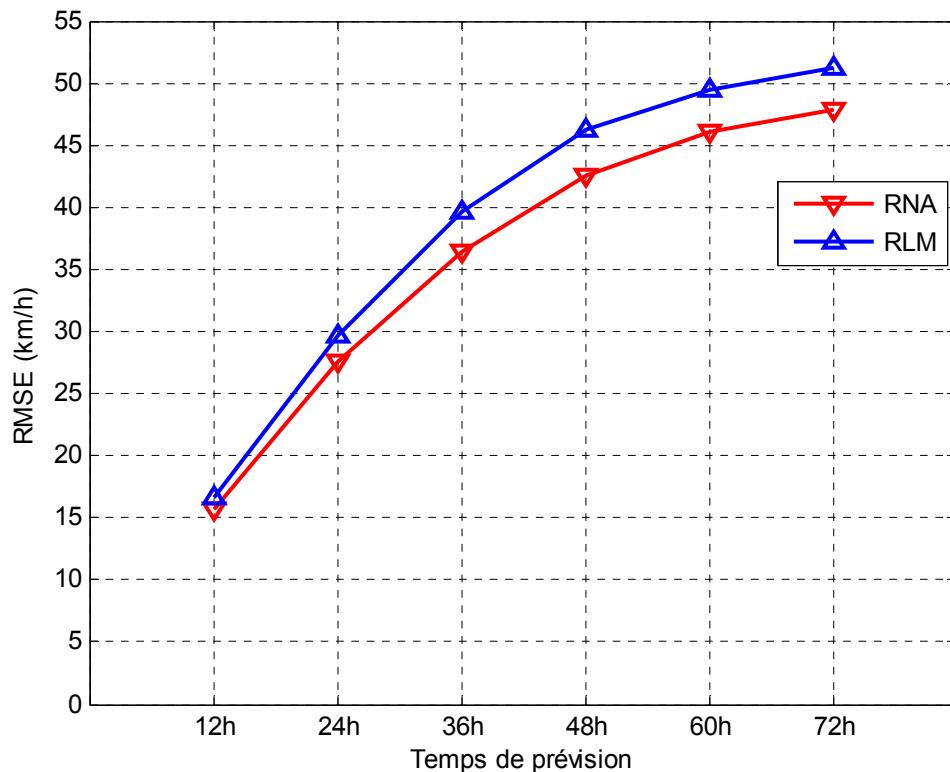


Figure 4.29 : Comparaison des erreurs moyennes de prévision de l'intensité cyclonique utilisant les méthodes RLM (Régression Linéaire Multiple) et RNA (Réseau de Neurones Artificiels).

Les Figures 4.30 à 4.37 montrent des exemples des intensités cycloniques prévues avec les modèles Réseau de Neurones Artificiels et Régression Linéaire Multiple.

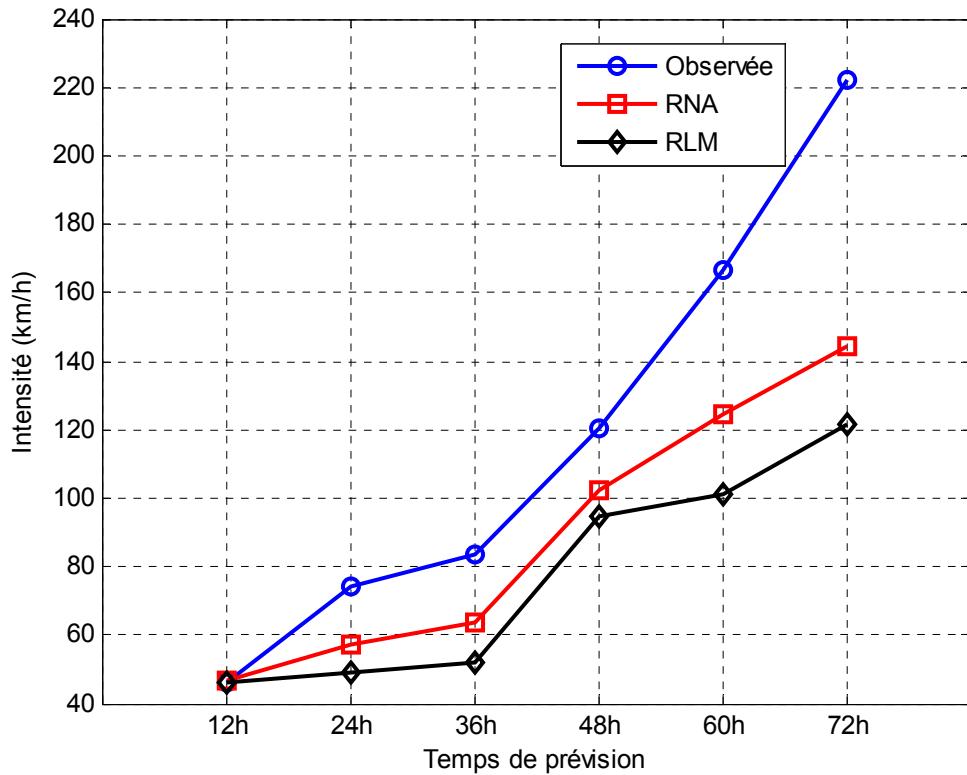


Figure 4.30 : Premier exemple de prévision. L'intensité augmente assez vite.

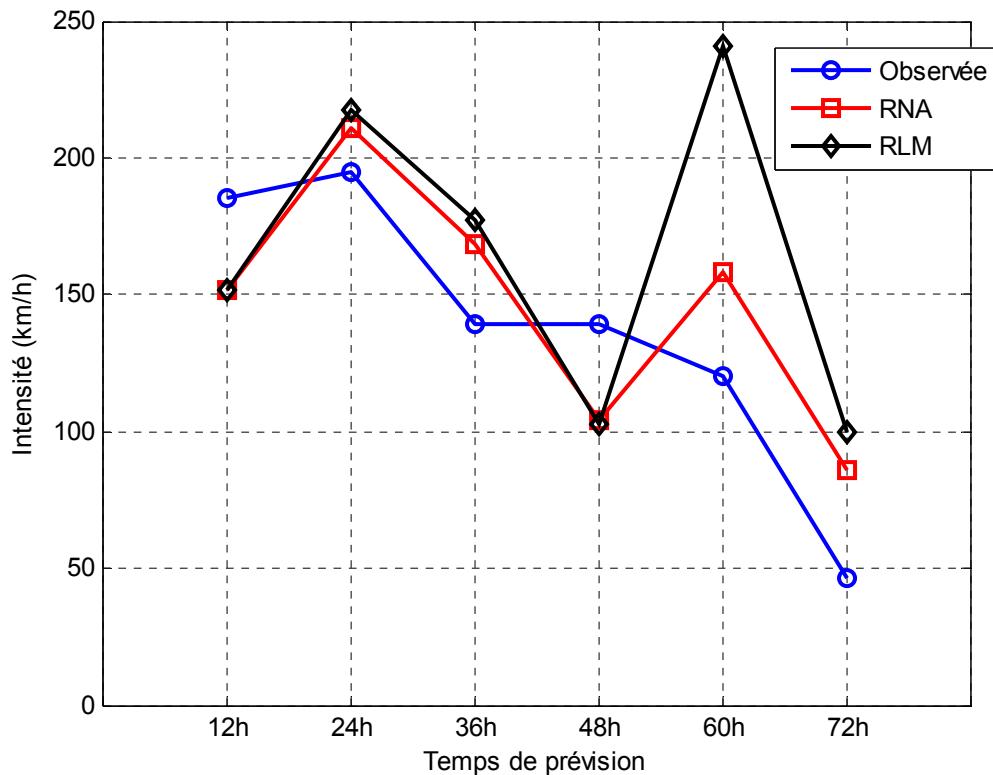


Figure 4.31 : Deuxième exemple de prévision. L'intensité diminue.

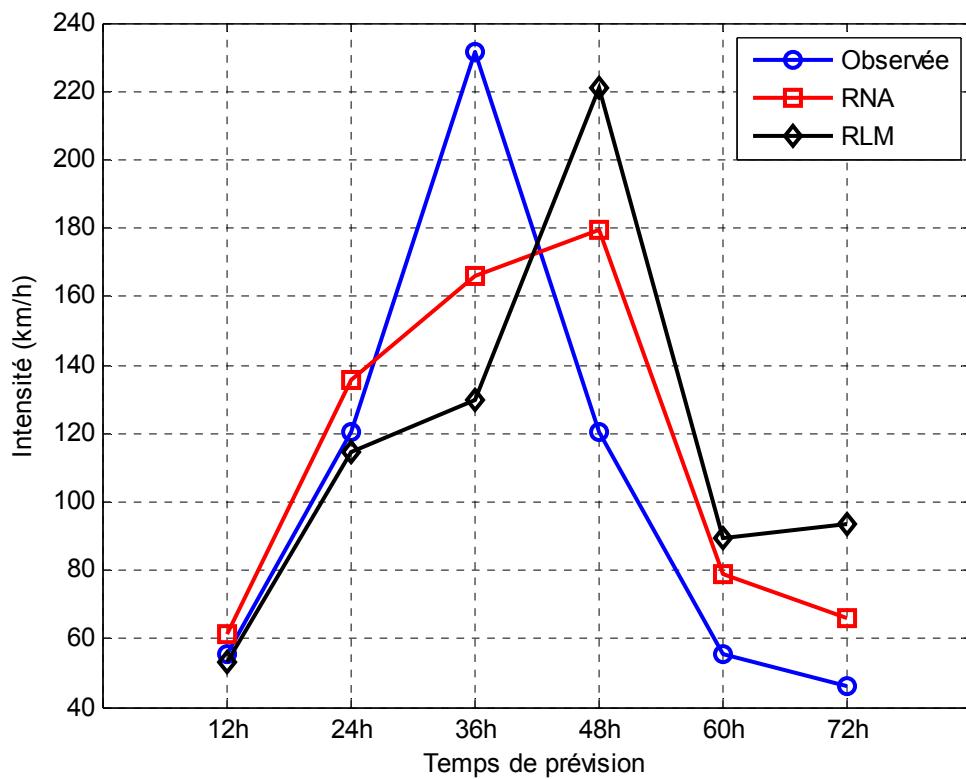


Figure 4.32 : Troisième exemple de prévision. L'intensité présente un maximum.

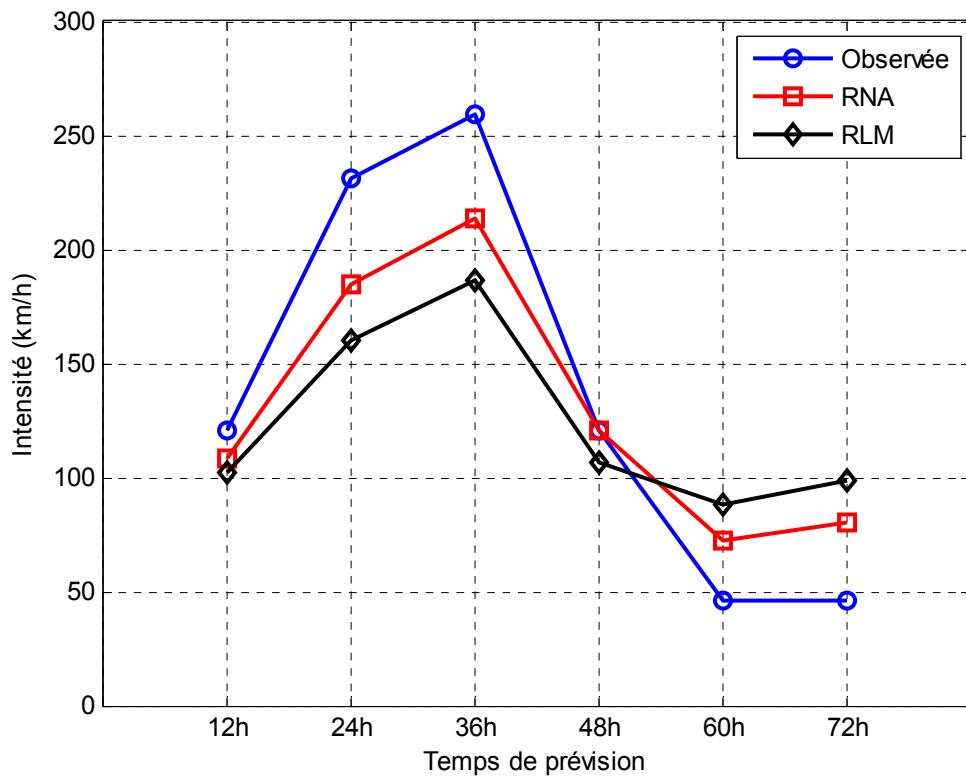


Figure 4.33 : Quatrième exemple de prévision. L'intensité présente un maximum.

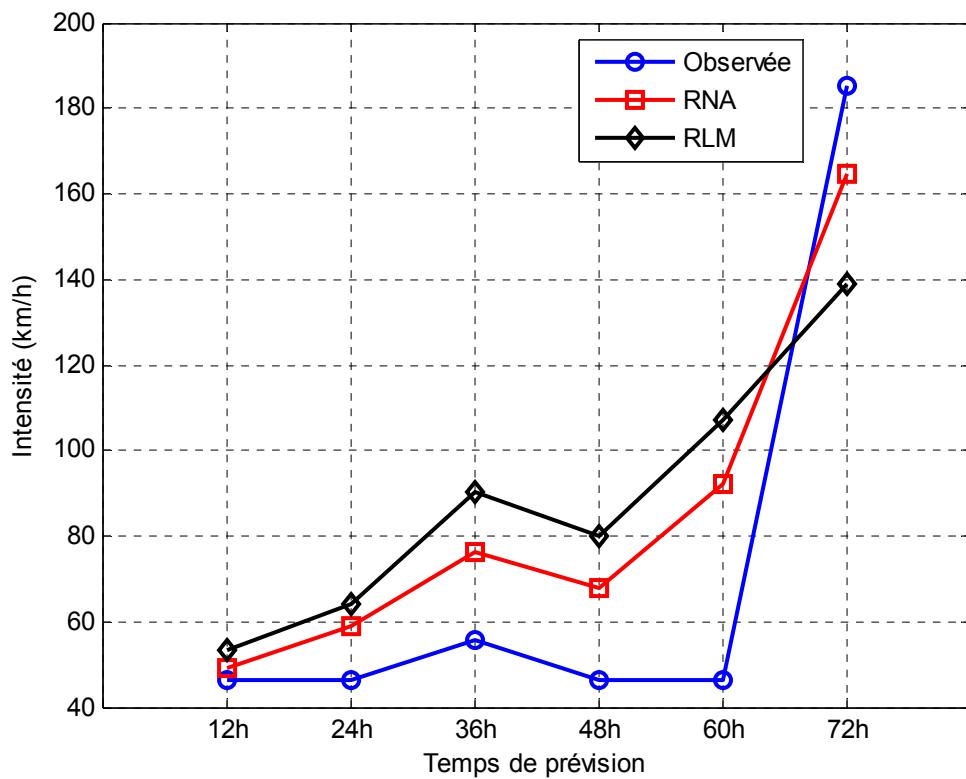


Figure 4.34 : Cinquième exemple de prévision. L'intensité augmente brusquement.

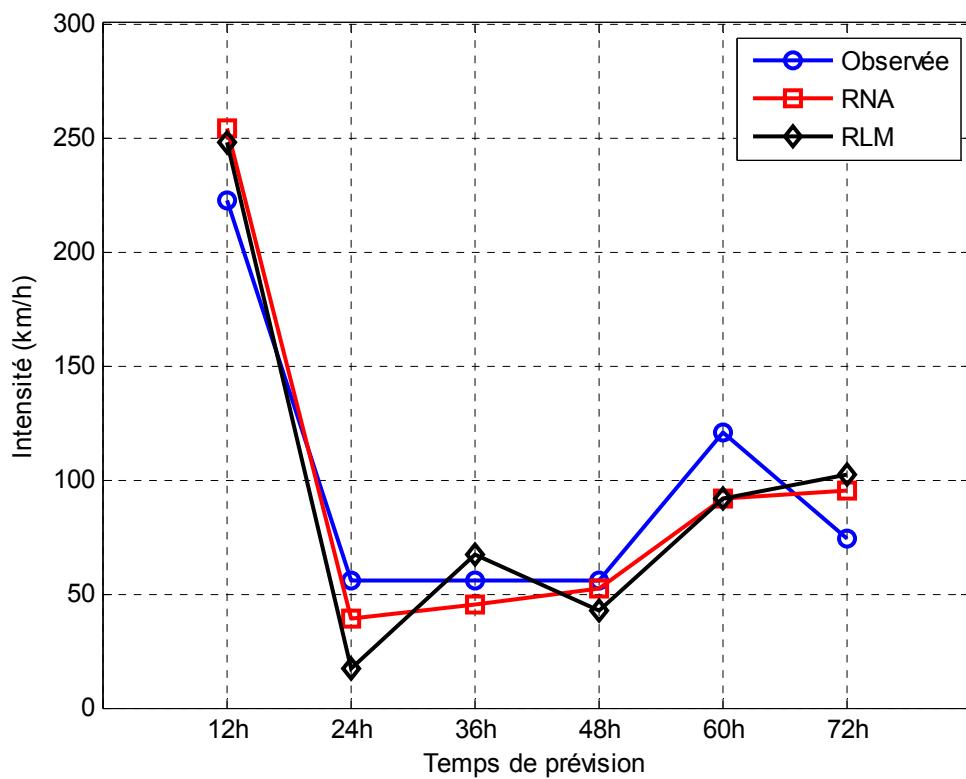


Figure 4.35 : Sixième exemple de prévision. L'intensité diminue brusquement.

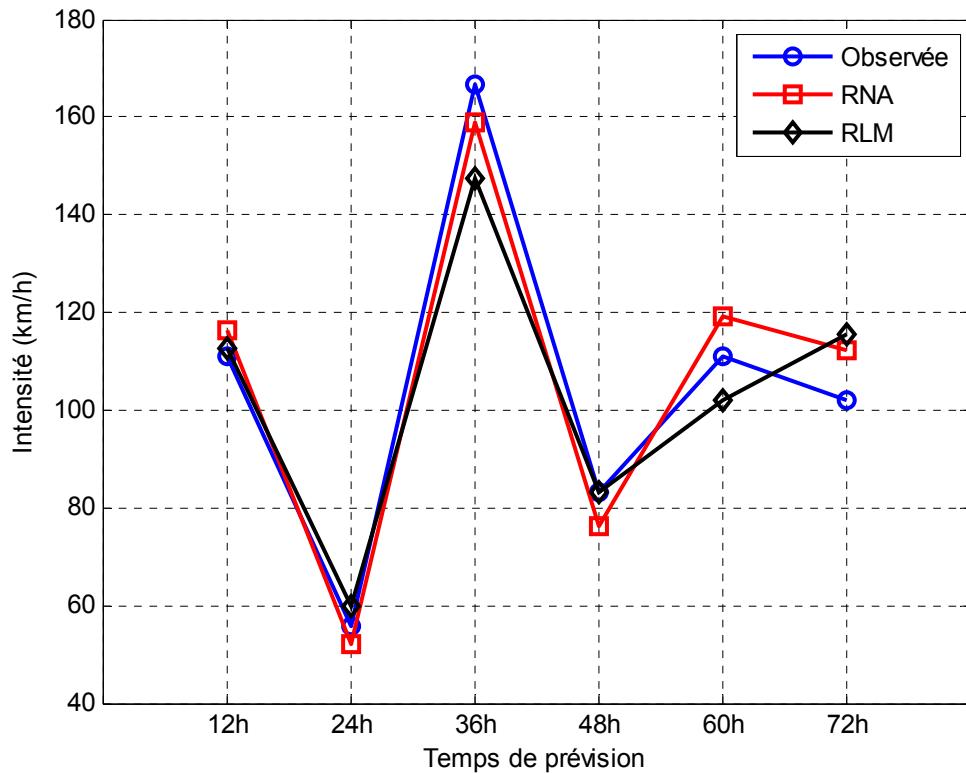


Figure 4.36 : Septième exemple de prévision. L'intensité présente une oscillation.

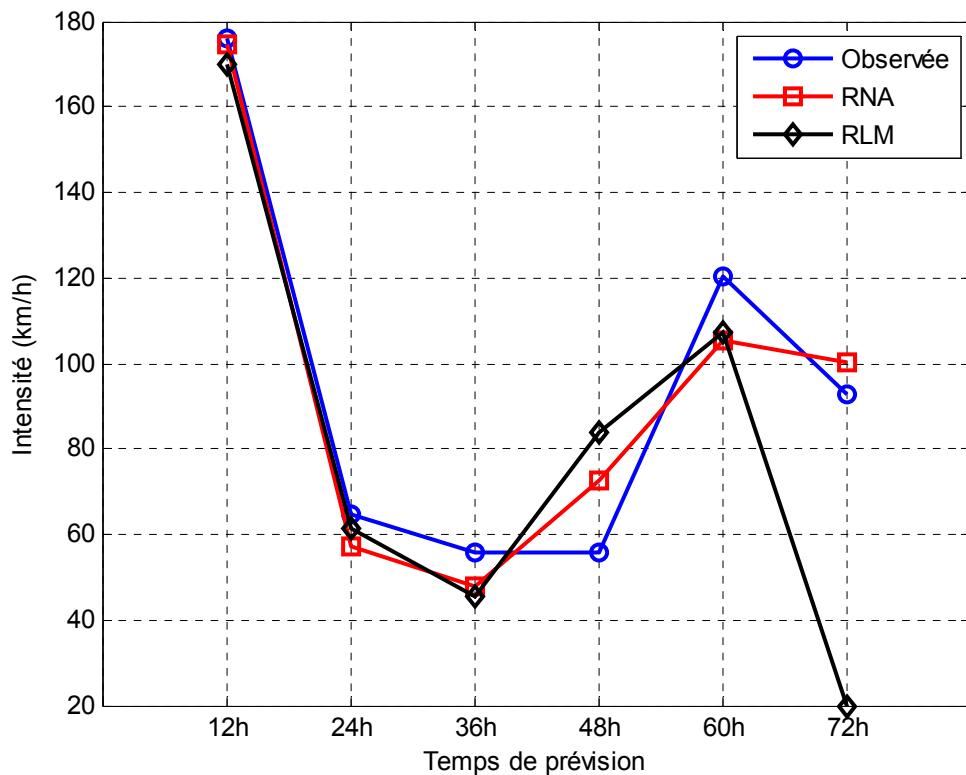


Figure 4.37 : Huitième exemple de prévision. L'intensité diminue brusquement puis augmente.

En général, ces résultats montrent que les allures des intensités obtenues par les deux modèles ne sont pas très différentes de celle de l'intensité observée. Cependant, le RNA à l'avantage de pouvoir suivre au mieux la variation de l'intensité par rapport à la régression linéaire.

D'après ces huit figures, on peut avoir des bons et des mauvais résultats de prévisions pour les deux modèles.

Les *Figures 4.38 à 4.43* comparent les répartitions des erreurs absolues produites par RNA et RLM. Les effectifs présentent des maximums pour chaque échéance de prévision. Ces maximums se trouvent aux alentours de 10km/h pour les prévisions à 12h, 24h et 36h et de 20km/h pour les prévisions à 48h, 60h et 72h.

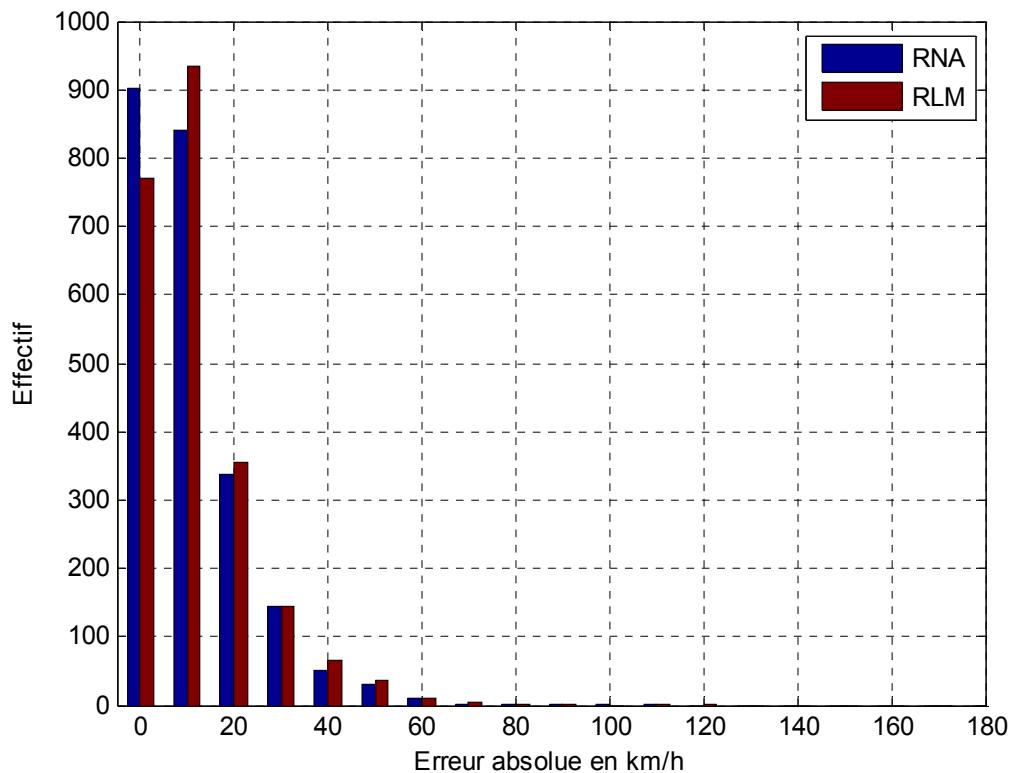


Figure 4.38 : Répartition des erreurs absolues pour la prévision à 12h.

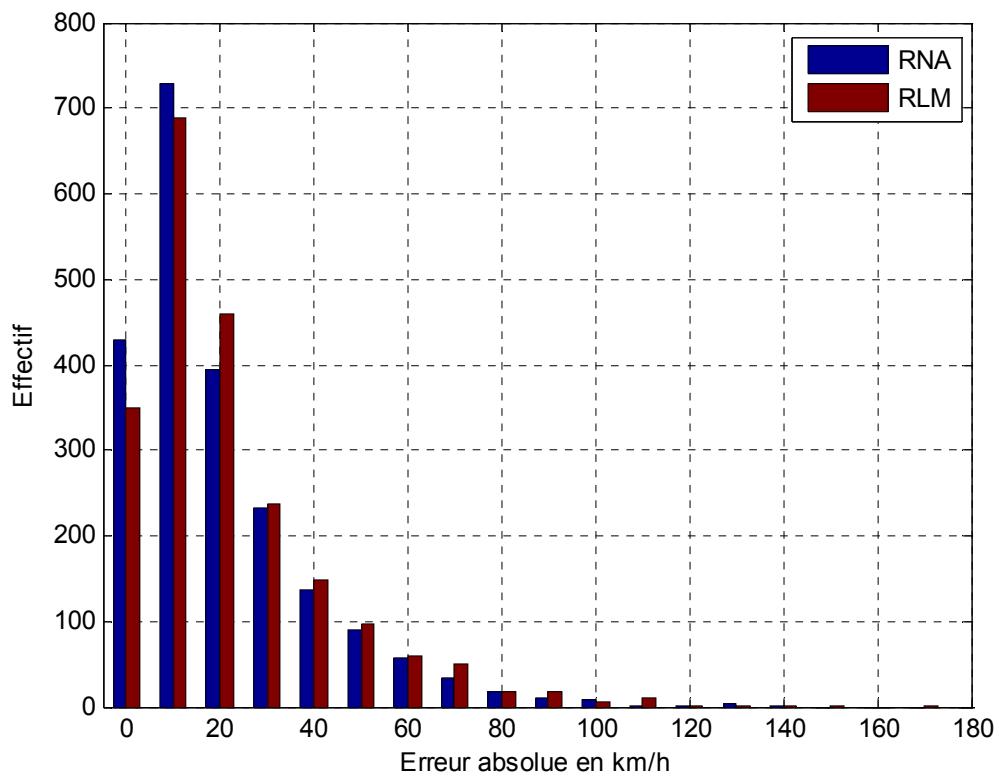


Figure 4.39 : Répartition des erreurs absolues pour la prévision à 24h.

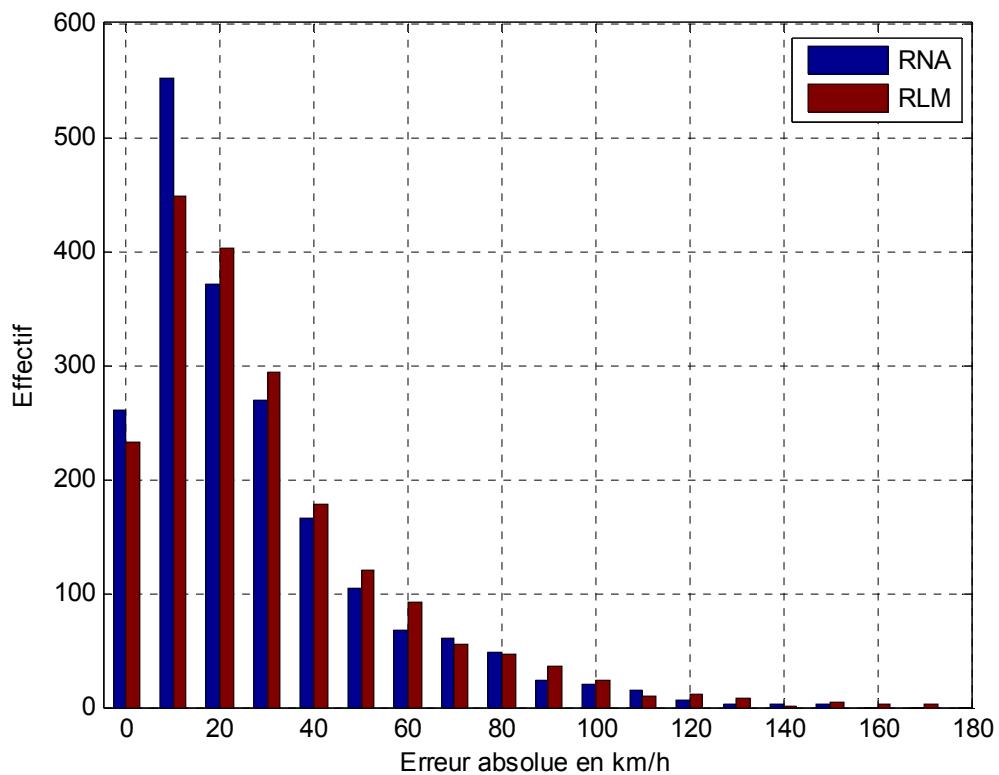


Figure 4.40 : Répartition des erreurs absolues pour la prévision à 36h.

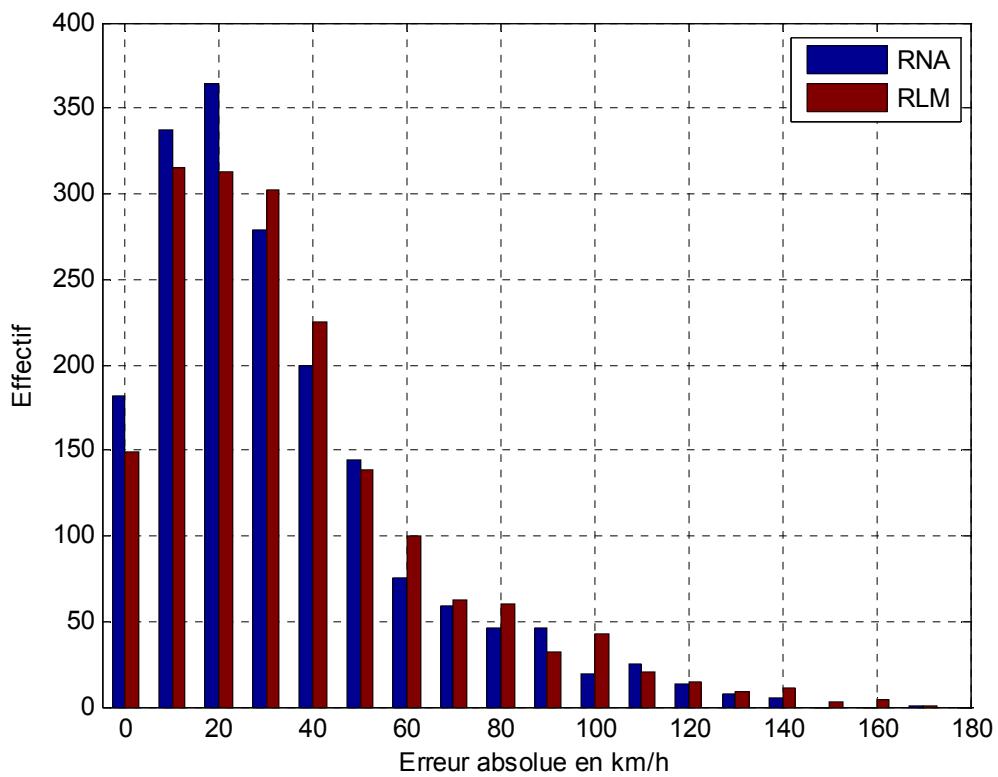


Figure 4.41 : Répartition des erreurs absolues pour la prévision à 48h.

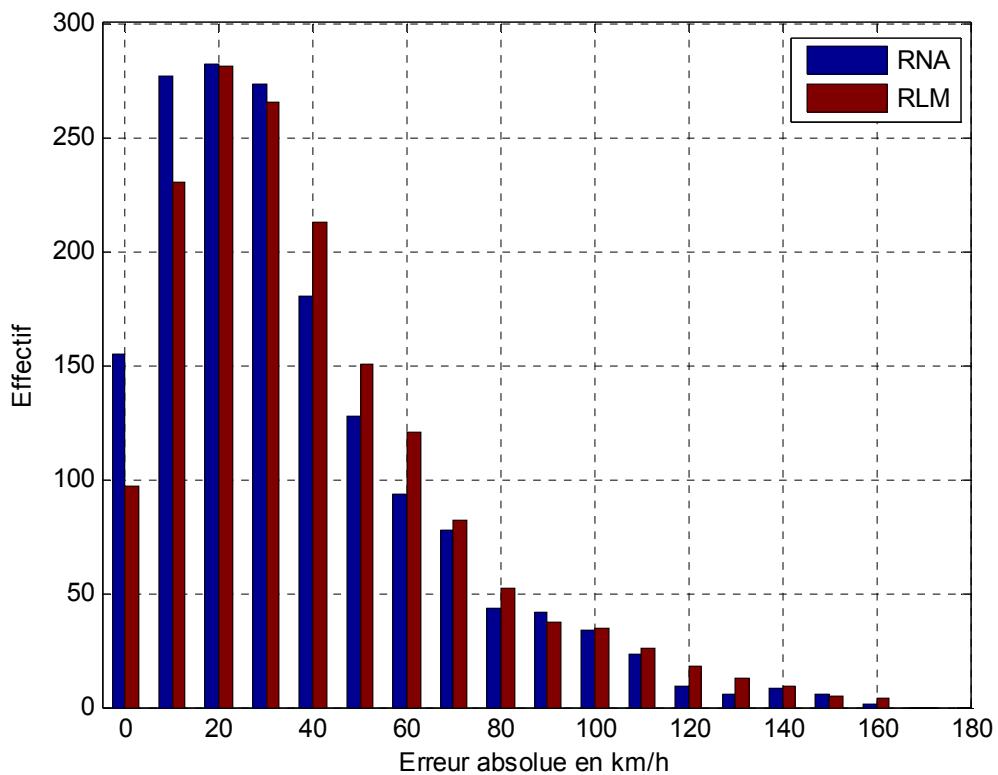


Figure 4.42 : Répartition des erreurs absolues pour la prévision à 60h.

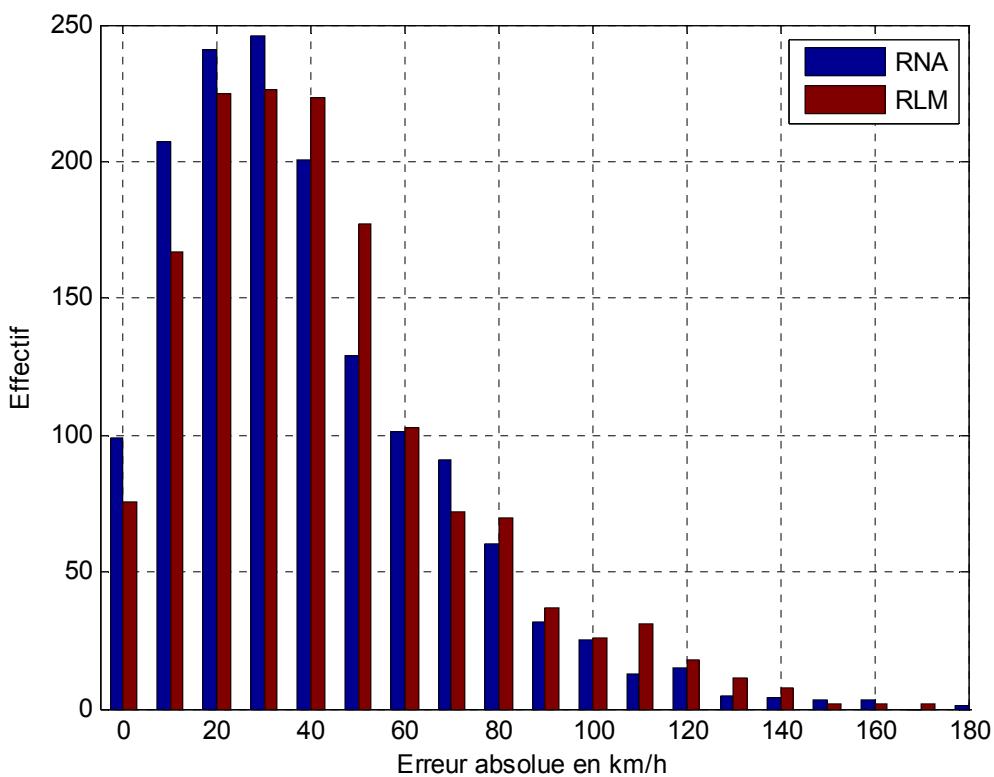


Figure 4.43 : Répartition des erreurs absolues pour la prévision à 72h.

Avant les effectifs maximums, les nombres des erreurs absolues obtenues avec RNA sont supérieurs aux nombres des erreurs absolues calculées avec RLM. Inversement, après les effectifs maximums, les nombres des erreurs absolues trouvées avec RLM sont plus élevés. Ces résultats indiquent que les erreurs relativement faibles sont plus fréquentes et celles relativement élevées sont peu nombreuses avec RNA. Cela montre l'avantage de RNA par rapport au modèle linéaire.

Les *Figures 4.44 à 4.49* représentent les variations de la différence entre l'intensité observée et l'intensité calculé en fonction de l'intensité observée. Ces figures montrent aussi les effectifs et les moyennes conditionnelles des cas où l'intensité observée est supérieure à l'intensité calculée et inversement.

Généralement, il n'y a pas beaucoup de différence entre les deux modèles concernant les effectifs des écarts d'intensités positifs et négatifs. C'est le nombre des écarts positifs, correspondant aux faibles intensités, qui diminue lorsqu'on augmente le temps de prévision. Les deux modèles surestiment alors les faibles intensités surtout pour la prévision à longue échéance.

Par contre, le nombre des écarts positifs, correspondant aux fortes intensités croît lorsqu'on augmente le temps de prévision. Cela signifie que les deux modèles sous-estiment les fortes intensités surtout pour la prévision à longue échéance.

Sur ces figures (*4.44 à 4.49*), la surestimation des faibles intensités est beaucoup plus nombreuse que la sous-estimation des fortes intensités.

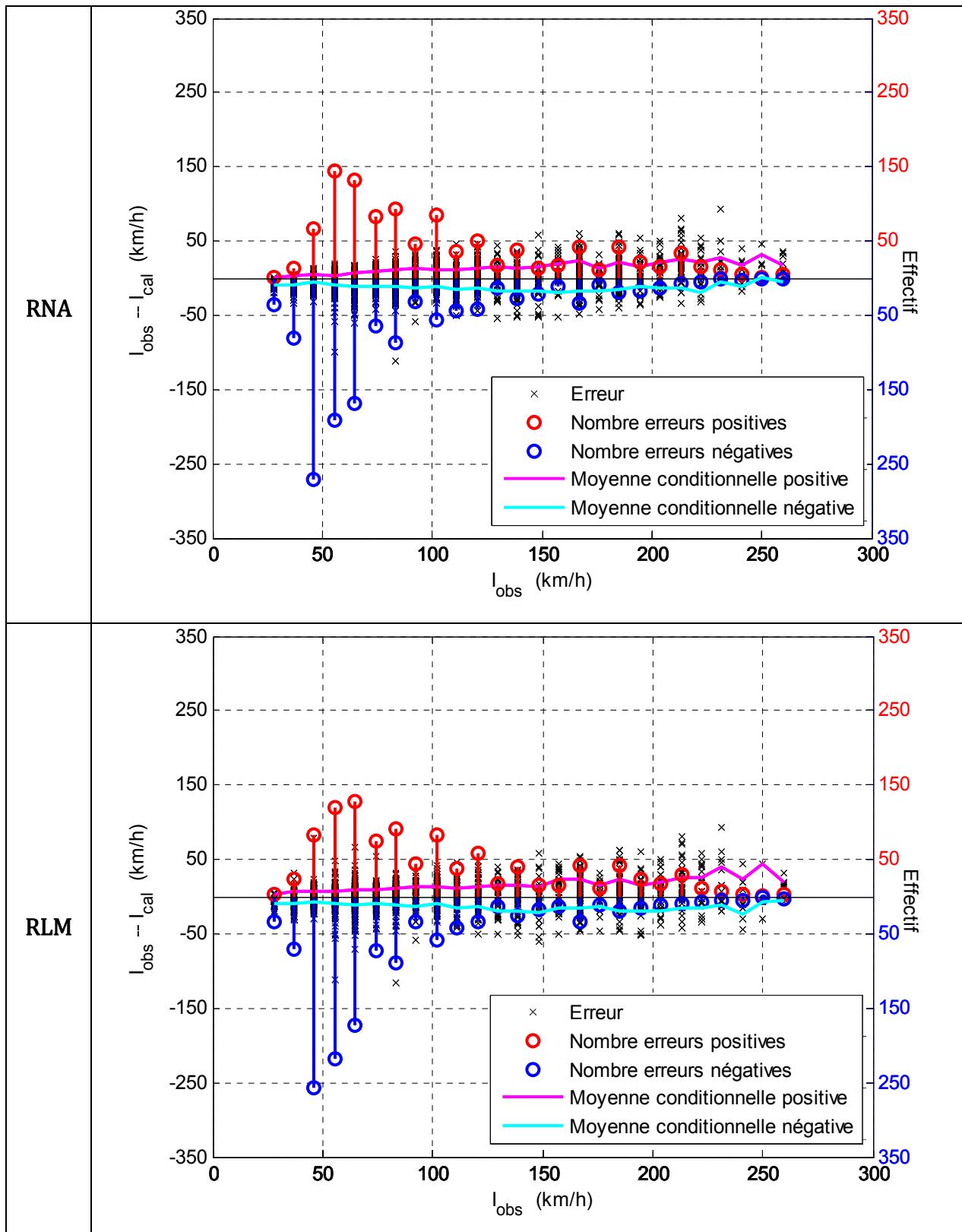


Figure 4.44 : Variations de la différence entre l'intensité observée et l'intensité calculé en fonction l'intensité observée pour la prévision à 12h

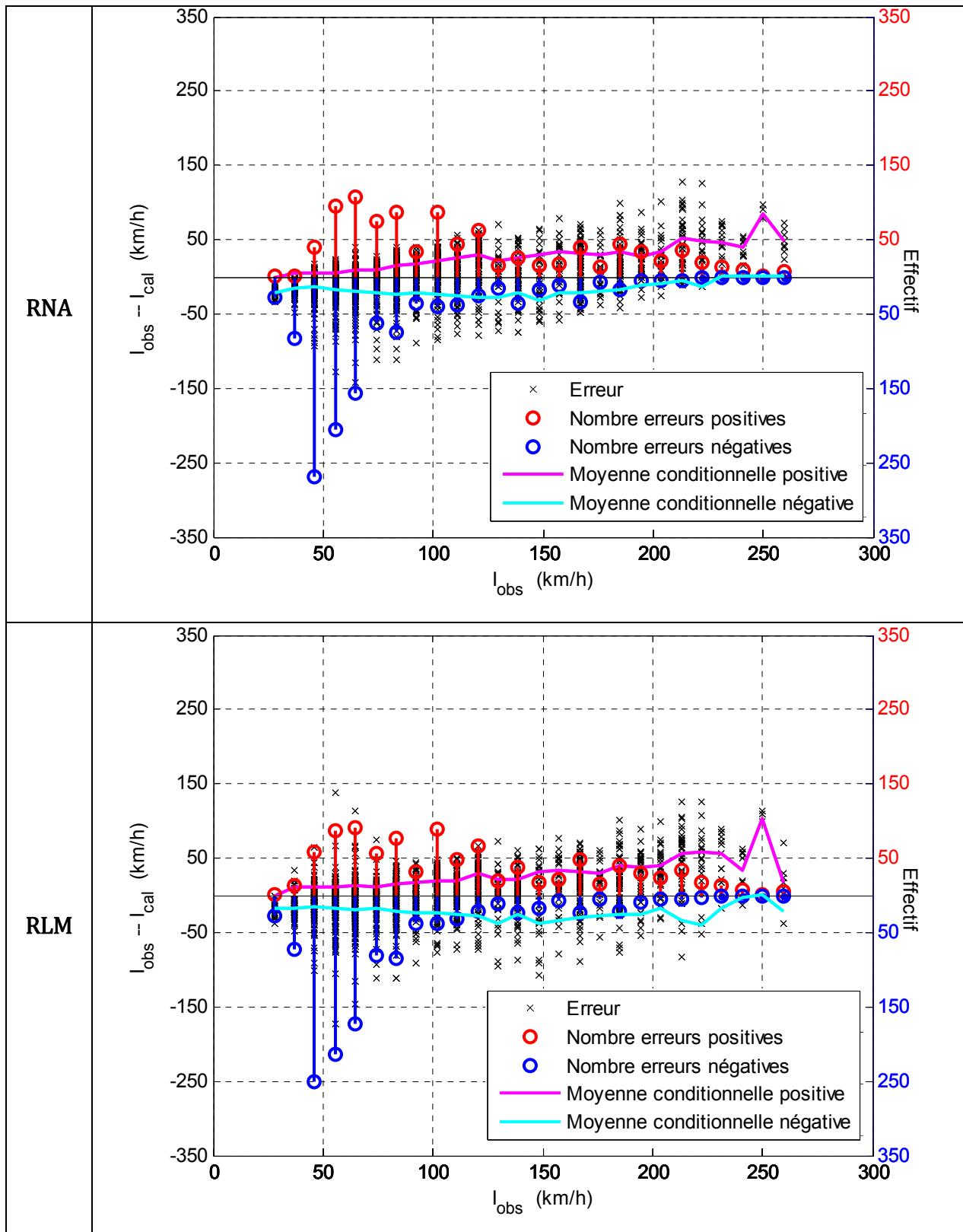


Figure 4.45 : Variations de la différence entre l'intensité observée et l'intensité calculé en fonction l'intensité observée pour la prévision à 24h

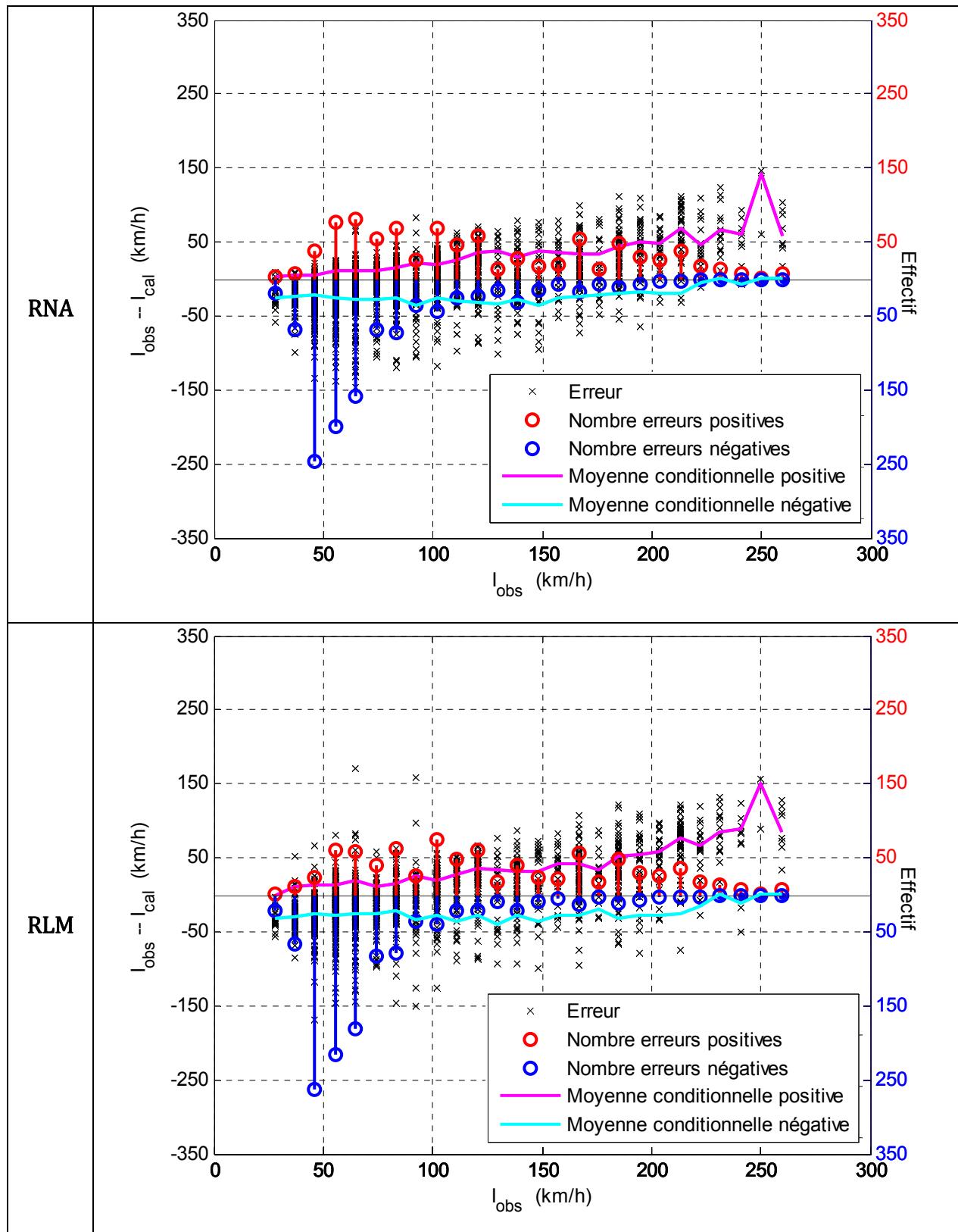


Figure 4.46 : Variations de la différence entre l'intensité observée et l'intensité calculé en fonction l'intensité observée pour la prévision à 36h

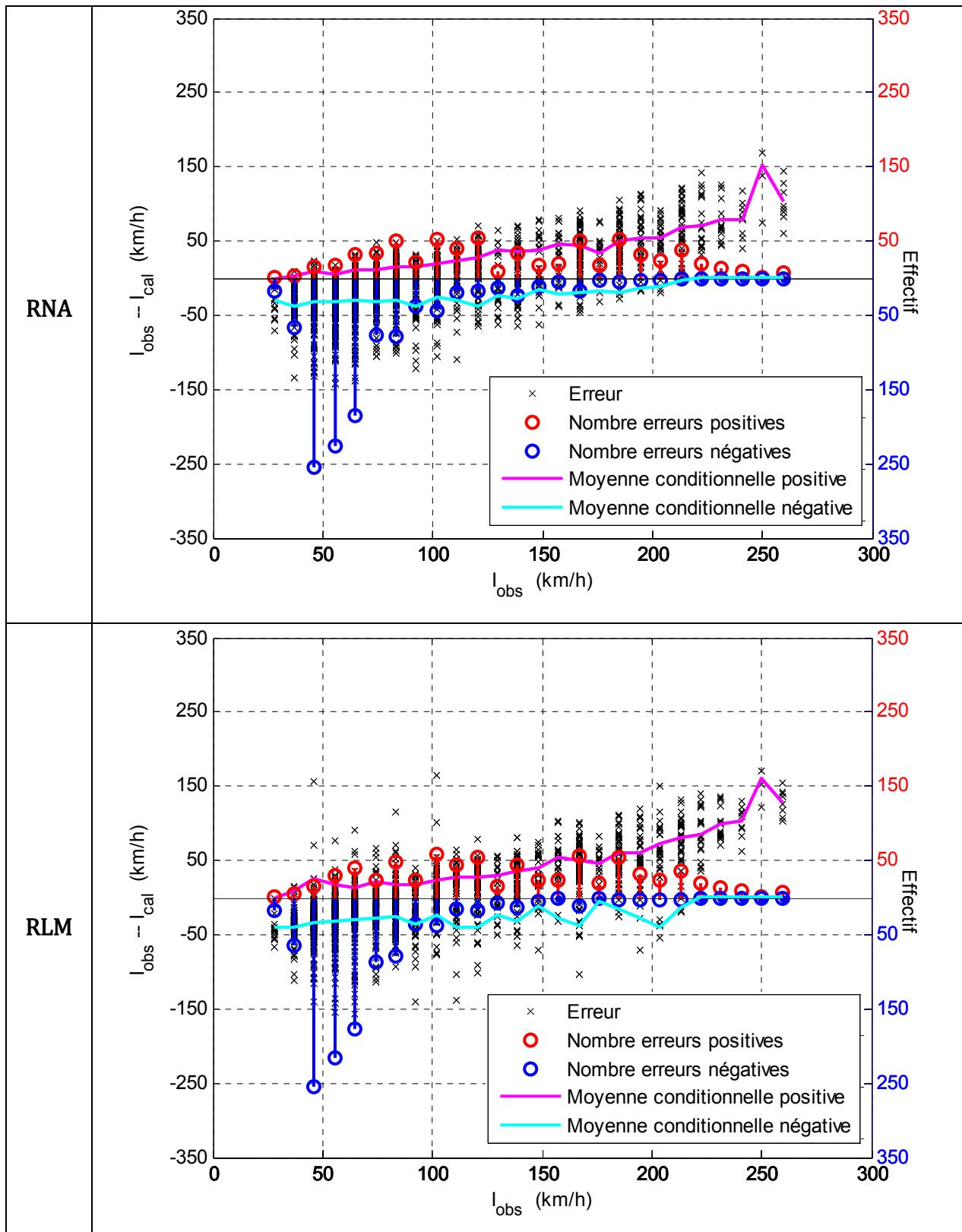


Figure 4.47 : Variations de la différence entre l'intensité observée et l'intensité calculé en fonction l'intensité observée pour la prévision à 48h

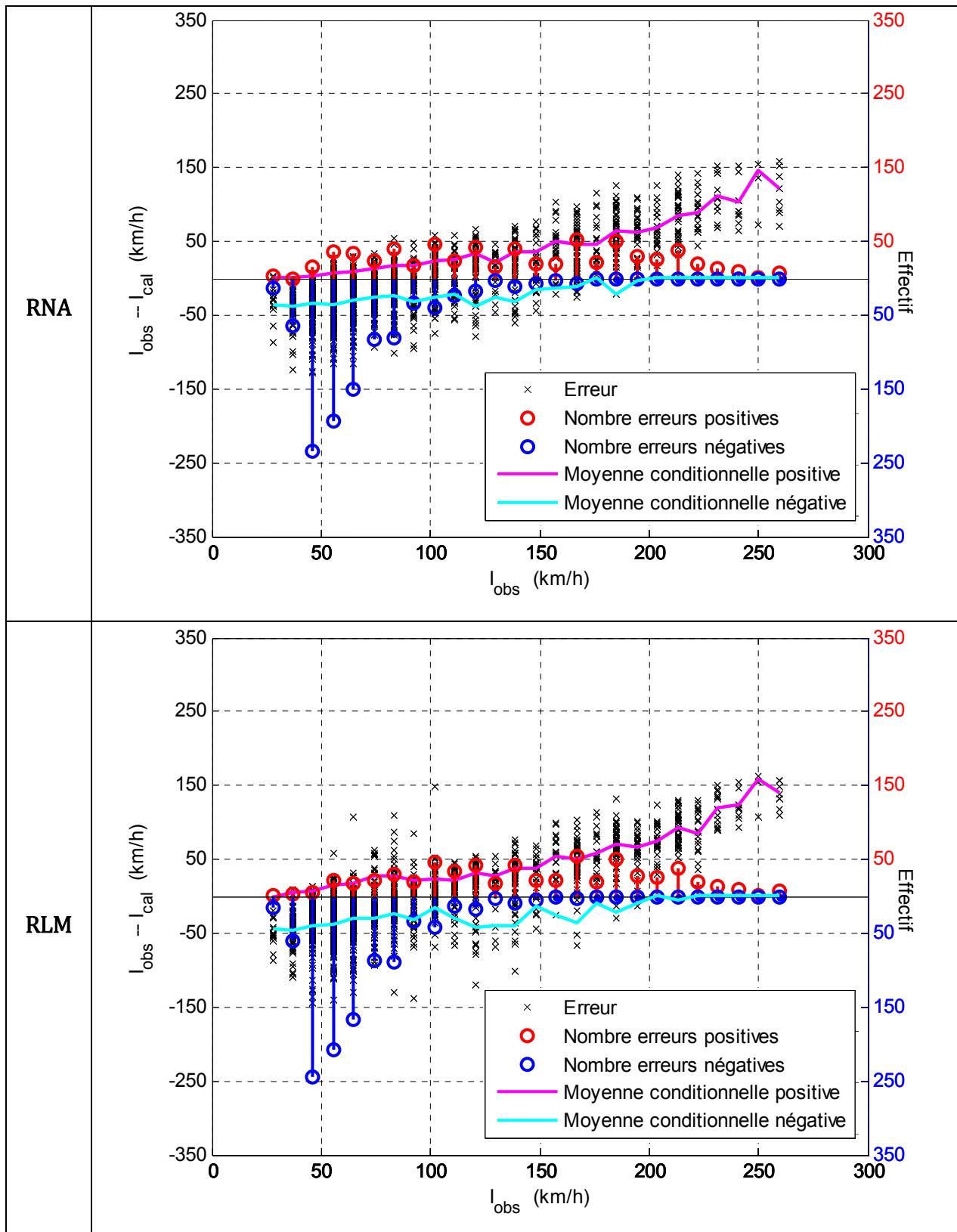


Figure 4.48 : Variations de la différence entre l'intensité observée et l'intensité calculé en fonction l'intensité observée pour la prévision à 60h

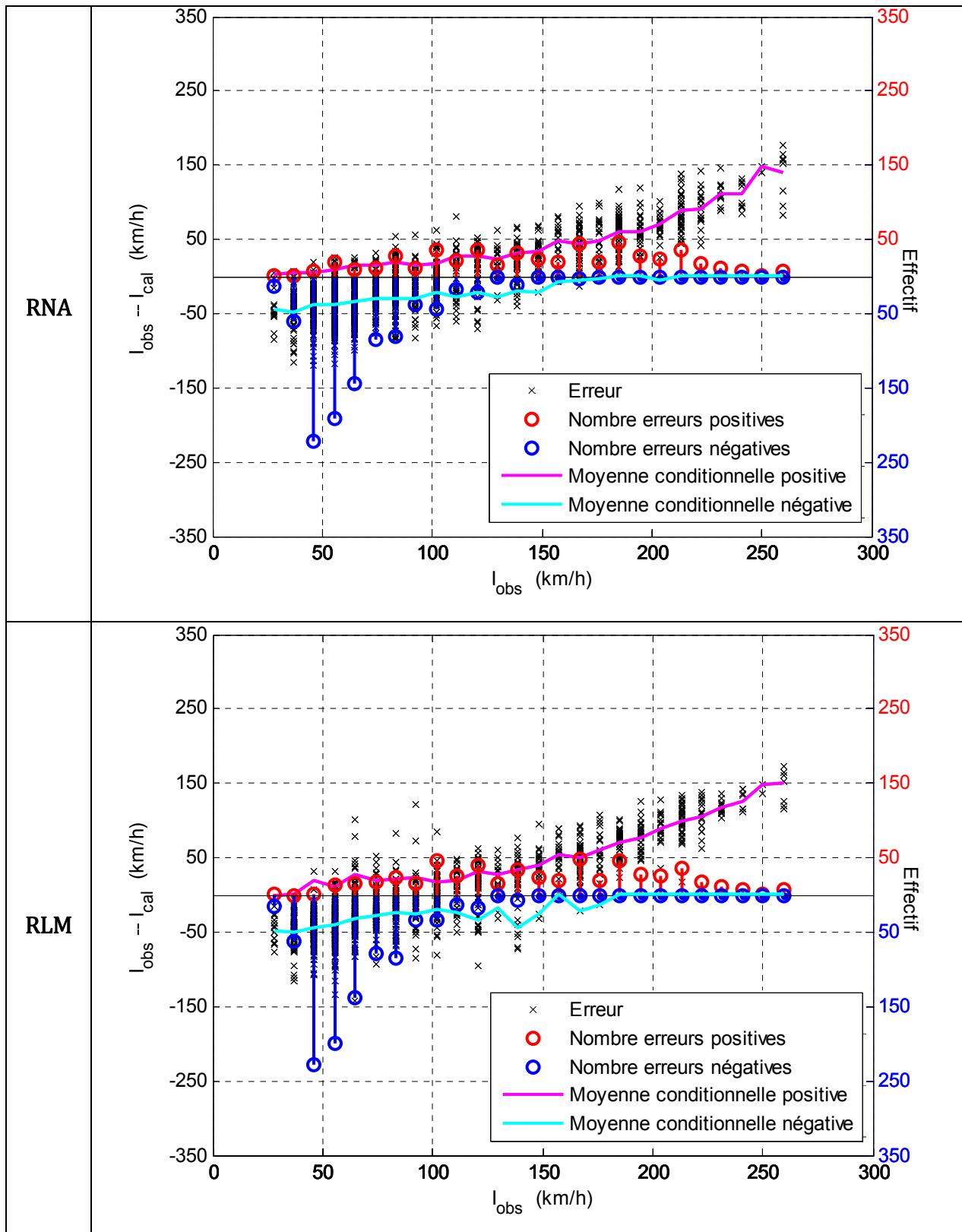


Figure 4.49 : Variations de la différence entre l'intensité observée et l'intensité calculé en fonction l'intensité observée pour la prévision à 72h

CONCLUSION GENERALE

La recherche des méthodes non analytique pour la résolution des problèmes de modélisation a été déjà incontournable depuis longtemps. Ceci est dû à la complexité du phénomène étudié. On a cherché, dans ce travail, à modéliser à l'aide d'une approche statistique l'intensité cyclonique. On a vu que ce phénomène est extrêmement complexe. Il dépend de plusieurs grandeurs physiques que les chercheurs de nos jours sont en train de découvrir.

Dans le premier Chapitre, l'environnement du phénomène cyclonique a été décrit. La connaissance des critères de formation des cyclones a permis de choisir certains prédicteurs comme la latitude, le jour julien, etc. Le développement d'un cyclone dépend beaucoup de ces conditions de formation. Dans ce chapitre, les différentes échelles habituellement utilisées ont été aperçues. L'utilisation de la technique classique de Dvorak nous a incité de chercher et d'explorer d'autre méthode plus précise.

Dans le second Chapitre, deux indices statistiques permettant de détecter « partiellement » la liaison entre les variables physiques ont été exposés à savoir le coefficient de corrélation linéaire et le rapport de corrélation. L'existence ou non d'une liaison fonctionnelle entre deux variables à partir du rapport de corrélation est aperçue. Mais il est remarqué dans ce travail qu'avoir une liaison fonctionnelle importante n'apporte pas toujours des bons résultats de prévision. La modélisation est plus facile à réaliser lorsque la liaison est linéaire. Ce chapitre montre que la limite inférieure de la liaison fonctionnelle est la liaison linéaire ; lorsque le coefficient de corrélation linéaire et le rapport de corrélation sont égaux, alors la liaison entre les variables est purement linéaire.

Dans le troisième chapitre, une ample connaissance sur une partie de la biologie des réseaux de neurones et l'analogie avec les réseaux de neurones artificiels a été acquise. Ce travail est basé sur l'exploitation des capacités des réseaux de neurones artificiels pour la prévision de l'intensité des phénomènes cycloniques. Ce chapitre montre qu'il est difficile de savoir l'architecture optimale d'un réseau de neurones artificiels. Cependant, il y a des propriétés intéressantes donnant des espérances pour leur utilisation comme la parcimonie, la possibilité, avec une seule couche cachée, de modéliser un phénomène physique donné.

Les résultats ont été exposés dans le quatrième Chapitre où les prédicteurs candidats sont présentés. La méthode MGS a permis de ranger ces prédicteurs initiaux. Une technique utilisant des prédicteurs aléatoires a donné le nombre maximal de prédicteurs à utiliser pour chaque temps de prévision. Les prédicteurs potentiels trouvés sont :

- les vitesses maximales du vent qui représentent les intensités antérieures du cyclone ;
- les différences de ces vitesses maximales du vent qui montrent les variations des intensités cycloniques antérieures (augmentation ou diminution) ;
- les latitudes qui positionnent le phénomène cyclonique par rapport à l'équateur (l'endroit le plus chaud du globe) ;
- les différences des longitudes qui décrivent les déplacements zonaux antérieurs du cyclone (vers l'ouest ou vers l'est) ;
- les différences des latitudes qui indiquent les déplacements méridiens antérieurs du cyclone (s'approche ou s'éloigne de l'équateur) ;
- le rang du jour qui situe le phénomène dans la saison cyclonique ;
- l'âge du cyclone qui détermine sa durée de vie par rapport à l'origine des dates ;
- et les vitesses de déplacement antérieur du cyclone qui précisent la durée de contact du système cyclonique sur une partie de la surface de la mer.

Ces prédicteurs ont été utilisés pour la prévision des cyclones à 12h, 24h, 36h, 48h, 60h et 72h.

En moyenne, le RNA est plus performant que le modèle linéaire pour toutes les échéances de prévision.

En guise de perspective, nous pensons que la méthode “ k-means ” (k-moyennes) pourrait améliorer la prévision. C'est une technique de classification par partitionnement. Nous pourrons classer les cyclones similaires dans un groupe avant d'utiliser le RNA. Il suffit de classer les cyclones dans la base de données en k groupe et injecter le nouveau cyclone dans son groupe correspondant.

BIBLIOGRAPHIES

- [1] Gray, W. M., 1979: Hurricanes: their formation, structure and likely role in tropical circulation, in meteorology over the tropical ocean. *Roy. Meteor. Soc.*, 155-218.
- [2] Frank, W. M., 1987: Tropical cyclone development, in a global view of tropical cyclones (ed. R. L. Elsberry). *ONR Marine Meteorology Program, Arlington (Virginie, Etats Unis)*, 53-90.
- [3] Neumann, C. J., 1993: Global guide to tropical cyclone forecasting. Chap 1: Global overview. TD 560 - TCP 31, OMM, Genève, Suisse, 1.1-1.37.
- [4] McBride, J., 1995: Global perspectives on tropical cyclone. Chap. 3: Tropical cyclone formation. TD 693 - TCP 38, OMM, Genève, Suisse, 63-105.
- [5] Gray, W. M., 1975: Tropical cyclone genesis. *Dept. of Atmos. Sci., Colorado State University, Fort Collins (Colorado, Etats Unis)*, Paper n°232, 121pp.
- [6] Frank, W. M., 1987: Tropical cyclone development, in a global view of tropical cyclones (ed. R. L. Elsberry). *ONR Marine Meteorology Program, Arlington (Virginie, Etats Unis)*, 53-90.
- [7] Chan, J. C. L., and R. H. F. Kwok, 1999: Tropical cyclone genesis in a global numerical weather prediction model. *Mon. Wea. Rev.*, 127, 611-624.
- [8] Gray, W. M., 1968: Global view of the origin of tropical disturbances and storms. *Mon Wea. Rev.*, 96, 669-700.
- [9] Gray, W. M., 1975: Tropical cyclone genesis. *Dept. of Atmos. Sci., Colorado State University, Fort Collins (Colorado, Etats Unis)*, Paper n°232, 121pp.
- [10] McBride, J. and R. Zehr, 1981: Observational analysis of tropical cyclone formation. Part II: Comparison of non-developing versus developing systems. *J. Atmos. Sci.*, 38, 1132-1151.

- [11] Zehr, R. M., 1992: Tropical cyclogenesis in the western north pacific. *NOAA Tech. Rep.*, NESDIS 61, 181pp.
- [12] DeMaria, M., J. J. Baik, and J. Kaplan, 1993: Upper level angular momentum fluxes and tropical cyclone intensity. *J. Atmos. Sci.*, 50, 1133-1147.
- [13] Florent Beucher, 2010 : MétéorologieTropicale : des alizés au cyclone. Manuel de Météorologie tropicale – Météo France.
- [14] Terray P., F. Chauvin, and H. Douville, 2007: Impact of southeast indian ocean sea surface temperature anomalies on monsoon-enso-dipole variability in a coupled ocean-atmosphere model. *Climate Dynamics*, 28 :553–580, DOI : 10.1007/s00382-006-0192-y. 246, 251.
- [15] Dvorak Vernon F. La technique de dvorak : procédures et règles d'analyses de l'intensité des perturbations tropicales. Technical report, 1974.
- [16] Emanuel K. A., 1986 An air-sea interaction theory for tropical cyclone. *J. Atmos. Sci.*, 43 :585– 604. 271, 272, 434, 436, 438, 518
- [17] Emmanuel Vincent, Juillet 2012 : Interactions entre les Cyclones Tropicaux et l'Océan: de l'échelle synoptique à l'échelle climatique. Thèse de doctorat présentée à l'Université Pierre et Marie Curie.
- [18] A. Tsybakov, 2006/2007 : Statistique appliquée, Université Pierre et Marie Curie. Maitrise de Mathématiques, 35-40.
- [19] Anand, Mehrotra, Mohan et Ranka «*Intelligent Control Using Neural Networks*», IEEE Control System mag; Avril, pp 11-18, 1992.
- [20] R.L. Watrous «*Learning Algorithms for Connectionist Networks: Applied Gradient Methods of Nonlinear Optimization*», IEEE First International Conference on Neural Networks, 2, pp 619-627, 1987.

- [21] Claude TOUZET, Juillet 1992 : LES RESEAUX DE NEURONES ARTIFICIELS, Introduction au connexionnisme.
- [22] Stevens, C. *Le neurone*. Masson, Paris. 1979.
- [23] Mac Clelland, J. et Rumelhart, D.E. *Explorations in parallel distributed processing: a handbook of models, programs, and exercises*. Vol. 2. The MIT Press, Cambridge, Massachusetts, 1988.
- [24] C. BISHOP, Neural Networks for Pattern Recognition, Oxford University Press, 1995.
- [25] K. HORNIK, M. STINCHCOMBE, H. WHITE, "Multilayer Feedforward Networks are Universal Approximators", *Neural Networks* Vol. 2, pp. 359-366, 1989
- [26] HORNIK K., STINCHCOMBE M., WHITE H. , Universal approximation of an unknown mapping and its derivatives using multilayer feedforward networks, *Neural Networks*, 3, p. 551-560. 1990
- [27] HORNIK K., Approximation capabilities of multilayer feedforward networks, *Neural Networks*, 4, p. 251-257. 1991
- [28] BARRON A. [1993], Universal approximation bounds for superposition of a sigmoidal function, *IEEE Transactions on Information Theory*, 39, p. 930-945.

PUBLICATIONS

PUBLICATION A

PREDICTORS SELECTION FOR THE FORECAST OF TROPICAL CYCLONE INTENSITY IN THE INDIAN OCEAN SOUTH WEST BASIN

Solohery Hajanirina RAKOTOARISON¹, Lahatra RAKOTOJOHN¹, Miloud BESSAFI ²,
Adolphe RATIARISON¹

¹ *Suspensions Rheology Laboratory, Department of Physics,
University of Antananarivo, Madagascar*

² *Laboratory for Energy, Electronics and Processes, La Réunion*

http://www.slac.stanford.edu/econf/C110825/hepmad11_talks/rakotoarison.pdf

PUBLICATION B

TEMPORAL CHARACTERISTIC OF THE MONSOON FLOW IN THE SOUTHERN SUMMER AT THE ENTRY OF THE MOZAMBICA CHANNEL

Maxwell Djaffard a) Bessafi Miloud b) Rakotoarison Solohery Hajanirina a)

1**Rakotovao Niry Arinavalona a) Ramiarinjanahary Olga a) Ratiarison Adolphe A.a)

a) *Department of Physics, University of Antananarivo, Antananarivo, Madagascar*

b) *Energy, Electronics and Processes Laboratory, University of La Réunion, La Réunion*

http://www.slac.stanford.edu/econf/C110825/hepmad11_talks/rakotovao.pdf

PUBLICATION C

CYCLOGENESIS FORECASTING IN THE SOUTH INDIAN OCEAN BASIN

*MAXWELL Djaffard ¹ djafmax@yahoo.fr

Miloud BESSAFI ² bessafi@univ-reunion.fr

RABEHARISOA Jean Marc ¹ haris_marc@yahoo.fr

RAKOTOARISON Solohery Hajanirina ¹ solohery7@yahoo.fr

RAMIHARIJAFY Rodolphe ¹ zafirodolphe@yahoo.fr

RATIARISON Adolphe Andriamanga ¹ aaratiarison@yahoo.fr

¹ Suspensions Rheology Laboratory, Department of Physics, University of Antananarivo,
Madagascar

² Laboratory for Energy, Electronics and Processes, La Réunion

http://www.slac.stanford.edu/econf/C110825/hepmad11_talks/maxwell.pdf

PUBLICATION

PUBLICATION A

**Predictors selection for the forecast of tropical cyclone intensity
in the Indian Ocean South West basin**

**Solohery Hajanirina RAKOTOARISON¹, Lahatra RAKOTOJOHN¹, Miloud BESSAFI²,
Adolphe RATIARISON¹**

*¹ Suspensions Rheology Laboratory, Department of Physics,
University of Antananarivo, Madagascar*

² Laboratory for Energy, Electronics and Processes, La Réunion

ABSTRACT

The problem in prediction, in general, is the search for relevant predictors. We applied the technique of modified Gram Schmidt for putting in order candidates predictors. Then another method allowed us to choose the number of predictors more effective than those taken randomly.

We were able to assess risk as when using a predictor when it is worse than a random predictor.

The selected predictors are used in the neural network with one hidden layer. Network optimization is done with the early stopping technique.

1. Introduction

Our goal is to use only predictors able to give the neural network model better performance. For this, we will use a technique to select from a given set a subset of predictors.

The idea of this method is, at first step, to order the predictors in order of relevance. Initially we have a set of p predictors of which we seek with the Modified Gram-Schmidt method one that best describes the desired output, then the second and so on. We finally obtain a ranking of predictors.

In a second step, we will use in addition to the p original predictors, predictors random to determine the final number of predictors to use in the neural network.

The Modified Gram-Schmidt technique and the method of predictor's number determination will be described later.

2. Modified gram-Schmidt algorithm

In our case, we will use this algorithm to order the predictors in order of relevance.

There are two manners to implement the Gram-Schmidt algorithm:

- The Classical Gram-Schmidt algorithm (CGS), orthogonalization;
- The Modified Gram-Schmidt algorithm (MGS).

Remarks :

CGS is most economical in terms of occupancy of memory but is sensitive to rounding errors. The MGS method is numerically more stable. Note that these two methods would be strictly equivalent in the absence of rounding errors [1]. So we use the modified Gram-Schmidt algorithm.

Principle [2]:

The algorithm of Modified Gram-Schmidt whereas the values of the predictors and output as vectors.

$$X = \begin{pmatrix} x_1^1 & \cdots & x_p^1 \\ \vdots & \ddots & \vdots \\ x_1^N & \cdots & x_p^N \end{pmatrix} = (X_1 \ X_2 \ \cdots \ X_p) \text{ is the matrix of inputs}$$

$$\text{where } X_p = \begin{pmatrix} x_p^1 \\ \vdots \\ x_p^N \end{pmatrix} \text{ is the vector of the input } p, \text{ and } Y = \begin{pmatrix} y^1 \\ \vdots \\ y^N \end{pmatrix} \text{ the output vector,}$$

N is the number of samples.

At each iteration, we must find the best input vector which explains the output. For this, we calculate the maximum of the square cosines of angles which make the input vectors with the output vector.

$$\max_{1 \leq i \leq p} [\cos^2(X_i, Y)] = \max_{1 \leq i \leq p} \left[\frac{(X_i^T Y)^2}{(X_i^T X_i)(Y^T Y)} \right]$$

And after each iteration we eliminate the best input vector and its contribution by projecting all of the remaining inputs vectors and output vector on its orthogonal subspace.

This process ends when all the input vectors are ordered.

3. Predictors selection method [1], [3]

The principle of this method is to compare the initial predictors with predictors taken randomly.

First, we must start the Gram-Schmidt algorithm on all predictors (100 random predictors + the original predictors). At each iteration, we select the best predictor by not taking into account the random variables. Then we detect random variables that explain more output than the best predictor, we count and we extract from the set of predictors. It only remains to project the output, the initial and random remaining predictors on the orthogonal subspace to selected predictor. And so on until all initial predictors are ordered.

4. Application to the tropical cyclone intensity predictors classification

4.1 Introduction

Initially, we considered 32 predictors for forecasting cyclone intensity. The process of elimination that we have described above is applied to these predictors.

First let's number as follows the predictors (Table I):

Table I: Numbering of initial predictors.

1	Modified Julian day at -24h	12	Latitude at -18h	23	Difference in longitudes at 0h - 18h
2	Modified Julian day at -18h	13	Latitude at -12h	24	Difference in longitudes at 0h - 24h
3	Modified Julian day at -12h	14	Latitude at -6h	25	Difference in latitudes at 0h - 6h
4	Modified Julian day at -6h	15	Latitude at 0h	26	Difference in latitudes at 0h - 12h
5	Modified Julian day at 0h	16	Maximum wind speed at -24h	27	Difference in latitudes at 0h - 18h
6	Longitude at -24h	17	Maximum wind speed at -18h	28	Difference in latitudes at 0h - 24h
7	Longitude at -18h	18	Maximum wind speed at -12h	29	Difference in wind speed at 0h - 6h
8	Longitude at -12h	19	Maximum wind speed at -6h	30	Difference in wind speed at 0h - 12h
9	Longitude at -6h	20	Maximum wind speed at 0h	31	Difference in wind speed at 0h - 18h
10	Longitude at 0h	21	Difference in longitudes at 0h - 6h	32	Difference in wind speed at 0h - 24h
11	Latitude at -24h	22	Difference in longitudes at 0h - 12h		

Remark :

Modified Julian day is defined by :

$$jd_modif = \cos(jd_angle + s) \text{ where } jd_angle = \frac{(julian_day - 1) * \pi}{365} \text{ and } s \text{ the shift parameter.}$$

The problem with Julian days is that there is discontinuity in transition from end of December to 1st of January while the cyclone intensity does not have such a strong discontinuity between two successive days. The above transformation enables us to solve this issue.

4.2 Choice of predictors number

Figure 1 shows the number of random predictors better than the original predictors.

We can read this figure as follows: after choosing the number of predictor in the model (neural network in our case), that we can read on the x-axis, there will be a risk (value in % on the y-axis) to select at least one predictor less potential than a random predictor.

Here we have almost 0% risk of choosing the first 8 predictors ordered by the Gram-Schmidt method.

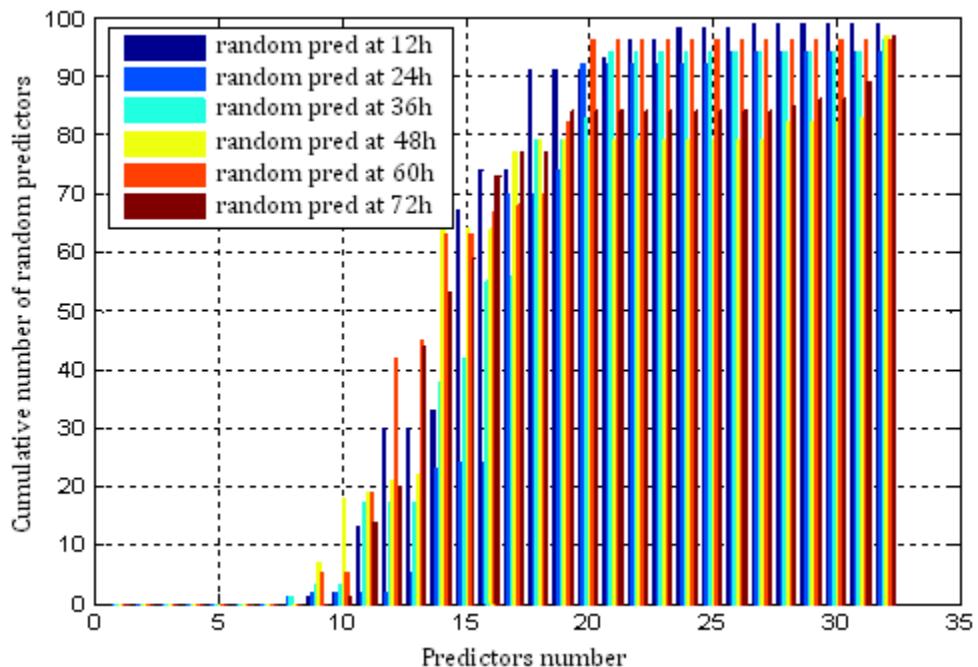


Figure 1: Choice of predictors number

4.3 Ranking of predictors

The following table (*Table II*) shows the 32 ordered predictors for each forecast time (12h, 24h, 36h, 48h, 60h, 72h).

For all time of prediction, the first predictor ordered by the MGS method is the maximum wind speed at 0h. This result confirms the strong correlation between the maximum wind speed at 0h and those at 12h - 72h, compared to other predictors.

The second predictor is the difference between wind speeds at 0h and 12h, except for the prediction at 24 hours.

As the sea acts as a heat source which powers the cyclone [4] (such as a heat engine), the change in temperature at the sea surface (SST) influences his intensity. The relationship between latitude and tropical cyclone intensity is then due to that which exists between SST and latitude. Latitude is an especially good predictor for long-range forecasts (48h, 60h and 72h).

The modified Julian days also have a significant contribution relative to other predictors for predicting cyclone intensity.

The first 8 predictors are shown in the table II.

Table II : Ranking of predictors.

	12h	24h	36h	48h	60h	72h
1st	20	20	20	20	20	20
2nd	30	17	30	30	30	30
3rd	17	19	31	15	15	15
4th	19	1	2	2	2	3
5th	1	14	15	21	11	28
6th	12	22	19	17	21	21
7th	23	18	22	19	31	19
8th	25	5	32	11	29	31
9th	5	15	11	6	6	24
10th	15	11	13	7	7	23
11th	21	32	5	22	22	10
12th	6	26	6	32	32	25
13th	8	3	7	14	4	4
14th	2	6	14	3	13	1
15th	11	7	3	5	12	8
16th	3	9	1	4	1	22
17th	24	12	10	10	25	13
18th	10	2	12	8	5	27
19th	28	4	4	24	3	7
20th	14	25	9	9	9	26
21th	22	29	24	28	10	29
22nd	9	27	25	23	27	18
23rd	27	16	27	31	24	17
24th	7	31	17	29	23	14
25th	4	30	29	18	14	12
26th	32	23	18	25	18	11
27th	13	8	16	16	16	6
28th	26	24	26	26	19	9
29th	29	13	21	12	28	16
30th	31	21	23	27	17	32
31st	18	10	8	1	8	2
32nd	16	28	28	13	26	5

5. Modelisation with neural network [5], [6]

We tested with the neural network these predictors previously ordered. By varying the number of predictors used, we find that the choice of the number of predictors is very important for the quality of forecasts.

After ordering the predictors in order of relevance, if the number of predictors chosen is too low, the model can only predict the intensity with a relatively high error.

On the other hand, if the number of predictors is very high, there will be parameters that will affect the forecast, and the error increases. In addition the calculation cost becomes very important.

Figure 2 shows that the forecast error is minimal at 72 hours for a number of predictors equal to 8.

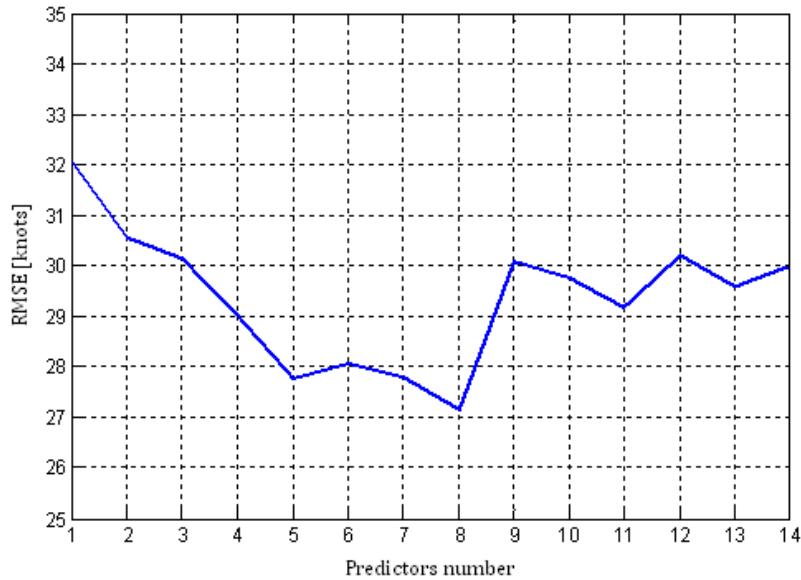


Figure 2: Variation of RMSE versus number of predictors used at 72h.

It is therefore possible to optimize the number of predictors to use. The results found directly with neural networks coincide with that provided by the above selection technique.

In the following, we will choose the first 8 predictors previously selected as inputs of the neural network. A network with one hidden layer is used and the numbers of iterations, the optimal hidden neurons number are determined from the technical "early stopping".

Figure 3 shows the variation of the root mean square error versus the forecast time.

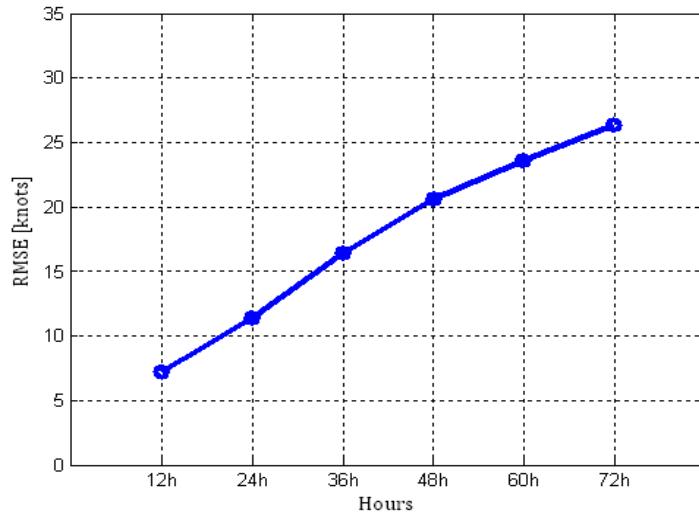


Figure 3: RMSE on 2151 cyclones.

Conclusion

Starting with any set of parameters, it is possible to extract a subset for predicting cyclone intensity. We have ordered the original settings and select potential predictors by the method of modified Gram-Schmidt.

This technique is effective but we must further increase the basis of initial predictors; that is, we need to find candidates predictors able to predict the intensities especially for the long term.

References:

- [1]: H. STOPPIGLIA. Méthodes statistiques de sélection de modèles neuronaux. Applications financières et bancaires. Thèse de l'Université Pierre et Marie Curie, Paris. 1997.
- [2]: Chen S. billings S.A. luo w. orthogonal least squares methods and their application to non-linear system identification. Int J. Control vol 50 n°5 pp. 1873-1896. 1989.
- [3]: L. OUKHELLOU. A new decision criterion for feature selection. France. 1998.
- [4] : Kerry EMMANUEL. Maximum intensity estimation. 1997.
- [5]: Marc PARIZEAU. Réseaux de neurons. Université LAVAL. 2004.
- [6]: Claude TOUZET. Les réseaux de Neurones Artificiels (RNA), Introduction au connexionnisme. Juillet 1992.

PUBLICATION

PUBLICATION B

Temporal characteristic of the monsoon flow in the Southern summer at the entry of the Mozambique Channel

Maxwell Djaffard ^{a)}

Bessafi Miloud ^{b)}

Rakotoarison Soloher Hajanirina ^{a)}

^{1**}Rakotovao Niry Arinavalona ^{a)}

Ramiarinjanahary Olga ^{a)}

Ratiarison Adolphe A. ^{a)}

^{a)}Department of Physics, University of Antananarivo, Antananarivo, Madagascar

^{b)}Energy, Electronics and Processes Laboratory, University of La Réunion, La Réunion

Abstract

The purpose of this study is to propose a documentation on the monsoon flow in the North of the Mozambique Channel and in particular the modes of apparition of this monsoon flow. As a matter of fact, the meridian wind at the 850 hPa pressure levels, issued from the data analysis of the European centre ECWMM has been specifically studied. In a first setting, the superposition of the average and the standard deviation of the meridian wind permits, on one hand, to see the modulation and the significativity of the monsoon flow in the North of the Channel, and on the other hand, to define the strong variability of this flow. This zone is located between the longitudes 39°- 44° E and the latitudes 0°-10°S. The spectral analysis by using the Fast Fourier Transform (FFT) and the Maximum Entropy Method (MEM) has been made in the second instance in order to highlight the 3 modes of the Cross-equatorial monsoon flows: 6-10 days, 15-25 days and 25-45 days.

1. Introduction

In the world, there are several monsoons to know especially the African monsoon, the Indian monsoon and the Australian monsoon. On this work, we are interested in the African monsoon and precisely the oriental African monsoon which generates the seasonal rains in Madagascar during the Southern summer. The implications of this monsoon on the Madagascar economic life are important. The African monsoon is still badly known while certain researchers have already studied this phenomenon. At the present time, the realization AMMA (African Monsoon Multidisciplinary Analysis) is in progress and should permit to improve the standard of the knowledge on the African monsoon. Concerning the East African monsoon, Duchiron [1] mentioned the conditions of the development of the monsoon, the elements that allows on the monsoon genesis and the interannual variability of the monsoon. We decided to undertake the temporal characteristic of this monsoon during the Southern summer. Thus, the goal of this study is to propose a documentation on the monsoon in the North of the Mozambique Channel and particularly the modes of the appearance of this monsoon flow.

2. Data and methodology

For this study, we used the meridian wind with a pressure level at 850 hPa issued from the daily reanalysis data of the ERA-40 [2] from the European centre ECWMM that is available online at the

^{1**} Speaker

synoptic scale with a grid of $2.5^\circ \times 2.5^\circ$ at the rate of two observations each day and a spatial resolution of $0^\circ\text{-}30^\circ\text{S}/30^\circ\text{-}120^\circ\text{E}$ in a temporal depth covering the 1979-2002 period.

Here are the different methods and techniques which will be used in order so as to analyse and represent the data.

First, we are interested in the meridian composite of the wind at a pressure level of 850 hPa during the Southern summer of the 1979-2002 period. We have proceeded to the calculation of the temporal average at each point on the grid of this meridian wind in order to study the modulation and the significance of this monsoon flow. In order to define the zone of high variability of this monsoon flow, we will superpose the average of the wind to its standard deviation.

The spectral analysis by choosing the fast transformation of Fourier [3] and the maximum entropy method [4] [5] [6] [7] [8] was done in a second period. It permitted to determine the dominant modes of the monsoon flow at the entry of the Channel for each temporal series of the Southern summer (November-April) of the 1979-2002 period. For each temporal series, we extracted the annual and semi-annual composing.

3. Results

Concerning the average values (FIG. 1), we can notice 5 poles with 3 negative nucleus indicating a flow from the North on the eastern coast of Kenya, the mid-western coast of Madagascar and on the South-eastern flank of Madagascar while 2 other nucleus are positive, each centred towards 50°E . These latters indicate a flow from the South, the going down and climbing of the trade wind flow because of the surrounding of the Island from the North and the South. Concerning the variability (FIG. 1), the area with a high variability is situated between the 39° and 44°E longitudes and the 0° and 10°S latitudes, and the monsoon flow skirting the eastern coast of Kenya is thus modulated and significant. Always in this high variability area, this monsoon flow takes place from mid-November with maximum amplitude during the months of January and February (FIG. 2). It also happens right in the middle of high cyclonic activity seasons. This flow changes direction by March. The figure 3 represents in a spectral form the dominant modes of summer monsoon flows at the entry of the Mozambique Channel for each Southern summer (November to April) of the 1979-2002 period, 6 to 10 days, 15 to 25 days, and 25 to 45 days. However, the average values of this flow for a Southern summer point out the 6-10 days and 20-30 day modes (FIG.4).

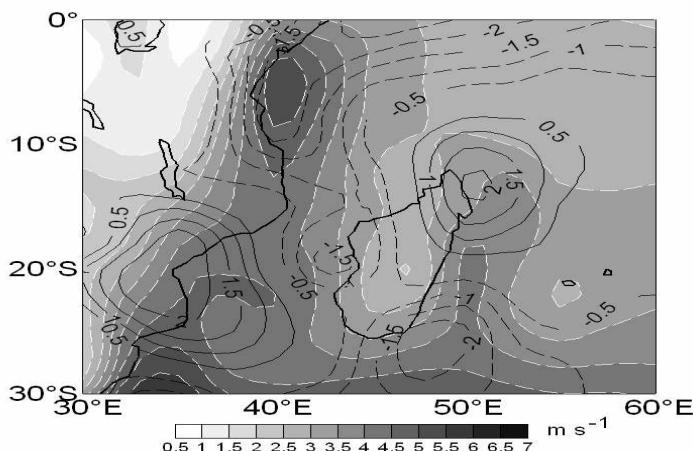


FIG. 1: Geographic configuration of the average (solid and dashed lines), the standard deviation (grey shading) of the meridian wind at the level of 850hPa during the Southern summer (November-April) during the 1979-2002 period.

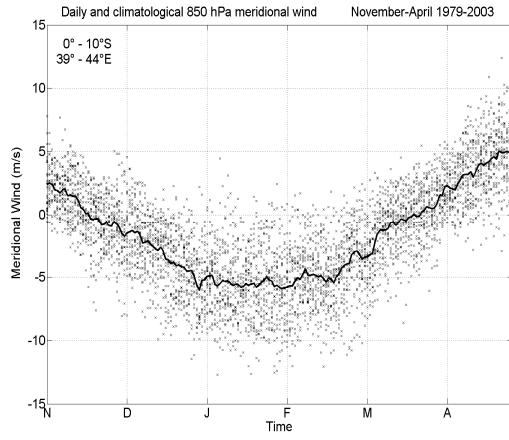


FIG. 2: seasonal repartition of the climatologic averaged (on the curved line) and of the daily averaged (cross) of the meridional wind at 850 hPa between the cyclonic season of 1979 and 2002 in the realm of 0° - 10° S and 39° - 44° E.

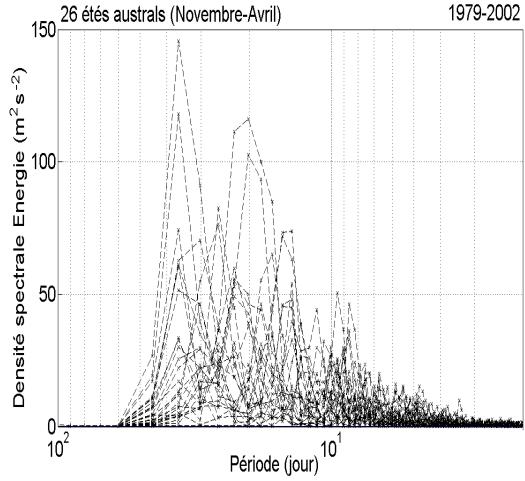


FIG. 3: Appearing modes of the monsoon flow for each southern summer for the 24 years of the 1979-2002 period in the 0° - 10° S and 39° - 44° E zone of the meridional wind at 850 hPa.

4. Conclusion and discussion

For this study, we have been able to complete the information about the monsoon flow at the entry of the Mozambique Channel. As a matter of fact, we have been able to define a zone with a high variability of this monsoon flow. This zone is located between 0° - 10° S and 10° - 44° E. It focuses on the position of the cyclogenesis in the Channel. The latter is located between 7° and 20° S. It seems that this Cross-equatorial monsoon flows is presented following three major modes on a timely interval of 6-10 days, 15-25 days and 25-40 days during the Southern summer of the 1979-2002 period. The temporary signature of 6-10 days lets us perceive an appreciable potential of cyclogenesis events forecast in the Mozambique Channel. These major modes let us to think of the modulation of this monsoon flow by the equatorial waves, especially the MJO and the ER wave. Besides, a considerable difference exists between the Cross-equatorial monsoon and the other monsoons namely the West African monsoon, the Indian monsoon and the Australian monsoon: this monsoon flow has an inverse direction of the usual direction in Southern summer and particularly, it blows the continent toward the ocean following a North-South direction (FIG. 5). So, it is Cross-equatorial. Finally, the existence of these modes let us to glimpse certain perspective. We are being able to study the modulation of this monsoon flow by the equatorial waves, especially the MJO and the ER wave.

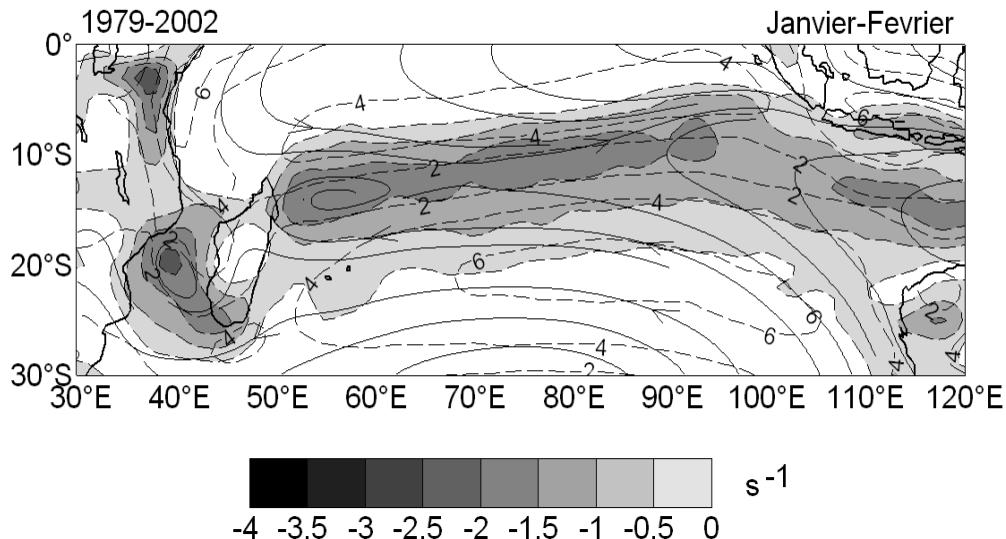


FIG. 5 : Geographic configuration of the atmospheric circulation (solid lines), vorticity (grey shading), module of zonal wind (dashed lines) at 850 hPa during the Southern summer period : January-February

Acknowledgments

I also want to express my gratitude the Physics Departments of Science Faculties in Universities who held the “physicists’ days” here in Madagascar in order to install a physicists’ journal of international level in Madagascar.

References

- Duchiron B., Variabilité interannuelle de la pluviométrie dans l'espace riverain de l'Océan Indien. Thèse de doctorat de l'université de Paris 7, 272p (2002).
- Cohen, T., E.I. Sweetser, 1975: The 'Spectra' of the solar cycle and of data for Atlantic tropical cyclones. *Nature*, **256**, 295 (1975).
- Ramiarinjanahary O., Contribution du flux de mousson de l'été austral sur la cyclogenèse dans le canal de Mozambique. Thèse de Doctorat. Département de Physique, Faculté des Sciences, Université d'Antananarivo, 158p (2007).
- Cadet, D. and H. Teitelbaum, J. Atmos. Sci., **36**, 892 (1979).
- Ferretti R. and D. F., 1986: Maximum Entropy Method applied to meteorological data. Report CNR/IFA n. 86/14 (1986).
- Keppenne, C. L. and M. Ghil, J. Geophys. Res., **97**, 20449 (1992).
- Mo, K., C., Month. Wea. Rev., **192**, 802 (2001).
- Uppala et al., Q. J. Roy. Meteorol. Soc., **131**, 2961 (2005).

PUBLICATION

PUBLICATION C

CYCLOGENESIS FORECASTING IN THE SOUTH INDIAN OCEAN BASIN

¹*MAXWELL Djaffard ¹ djafmax@yahoo.fr

Miloud BESSAFI ² bessafi@univ-reunion.fr

RABEARISOA Jean Marc¹ haris_marc@yahoo.fr

RAKOTOARISOA Soloher¹ soloher7@yahoo.fr

RAMIHARIJAFY Rodolphe¹ zafirodolphe@yahoo.fr

RATIARISON Adolphe Andriamanga¹ aaratiarison@yahoo.fr

¹Suspensions Rheology Laboratory, Department of Physics, University of Antananarivo, Madagascar

²Laboratory for Energy, Electronics and Processes, La Réunion

ABSTRACT.

The zone of cyclonic activity in the Indian Ocean basin lies between 7° to 18° of latitude South and 52° to 103° of longitude East. Study of the parameters affecting the formation of cyclone shows that the second eigenvalue greater than 1 is a rational index precursor of cyclogenesis in this area. We have determined the probable position of cyclone formation when this indication precursor appears.

I. INTRODUCTION.

This work investigates the tropical cyclone in the southwest of the Indian Ocean basin. The cyclones being born in the channel of Mozambique and those that attack the coasts of Australia are not consideration in this survey.

The objective is to find, on the one hand, a precursor index of cyclogenesis in the Indian Ocean basin, on the other hand, to determine the probable position of cyclone formation predicted when the precursor index occurs.

II. DATA SOURCE.

This study uses reanalysis data from France Weather Report. These are satellite data whose spatial resolution is $2.5^{\circ} \times 2.5^{\circ}$ for the OLR (Outgoing Longwave Radiation), $1.5^{\circ} \times 1.5^{\circ}$ for the SST (Sea Surface Temperature), MSLP (Mean Surface Level Pressure) and the data at different levels of the **zonal wind**, the **meridional wind**, the **divergence** of wind, the **vorticity** and the **specific humidity**. Thanks to a program published under the software MATLAB, we made the interpolation for obtaining an identical spatial resolution of $1^{\circ} \times 1^{\circ}$ for all these reanalysis data.

III. CHOICE OF THE STUDY AREA.

The figures 1, 2, 3, and 4 represent the histogram of the distribution numbers and the places of cyclone formation according to the latitude and the longitude in the Indian Ocean basin.

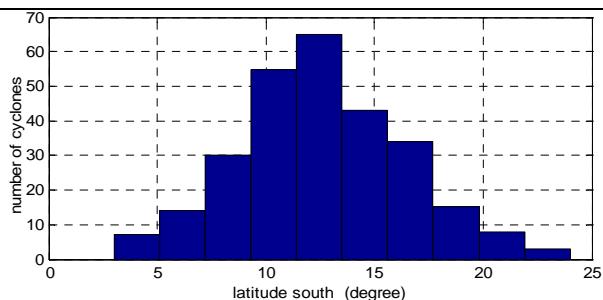


Figure 1: distribution of number and place of cyclone formation following the latitude.

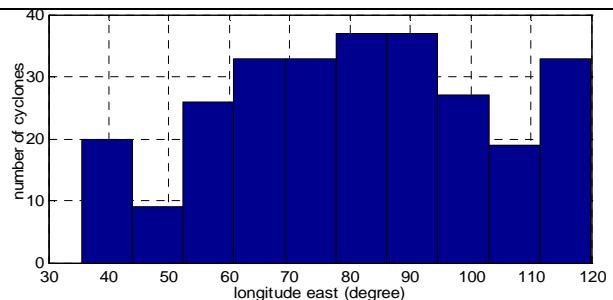


Figure 2 : distribution of number and place of cyclone formation following the longitude.

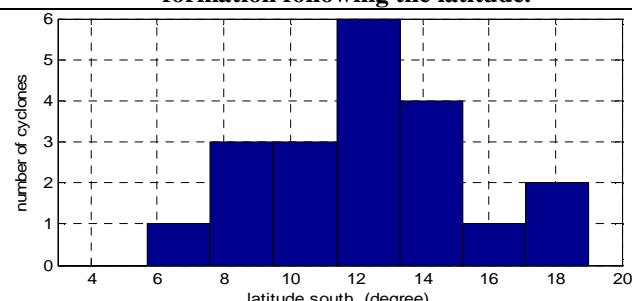


Figure 3: distribution of number and place of formation of second cyclone following the latitude.

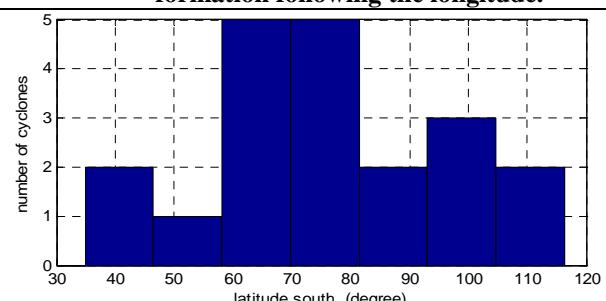


Figure 4: distribution of number and place of formation of second cyclone following the longitude.

We note that, on these figures (1, 2, 3 and 4), the majority of cyclones in the Indian Ocean basin formed themselves between the latitudes -7° and -18° and the longitudes 52° and 103° during the southern summers from November 1989 to April 2010. Among the 304 cyclones that took birth in this basin, 254 formed themselves between the latitudes -7° and -18° either 83.55% and 219 between the longitudes 52° and 103° either 72.04%. These statistical results show that the area bounded by these latitudes -7° to

* speaker

-18° and longitudes 52° to 103° is representative of cyclonic activity in the Indian Ocean basin. For this reason, we focused our survey of cyclogenesis in this zone that we named « zone of cyclogenesis »

IV. SELECTION OF VARIABLES

For selecting the variables, we followed the extreme values of the variables likely to influence the cyclogenesis. The next step was made:

- we followed the changes in their daily values at the point of formation of each cyclone that formed in the zone of cyclogenesis between 16 November 1989 and April 30, 2010, 15 days prior to the day of cyclogenesis [1] [2] ;
- we calculated the average of these variations for 181 cyclones that originated during this period in the zone of cyclogenesis;
- At the same time, we followed the changes in their daily maximum and their daily minimum values in the zone of cyclogenesis ;
- we calculated the average of variations of the maxima and the average of variations of the minimums since 15 days before until the day of cyclogenesis of these 181 cyclones.

Table 1 shows the coefficients of interrelationship of the variations thus gotten.

Variables	Correlation coefficient of minimum with the values at the point of cyclone formation and probability of non-correlation		Correlation coefficient of maximum with the values at the point of cyclone formation and probability of non-correlation	
	Correlation	probability of non-correlation	Correlation	probability of non-correlation
OLR	0.9497	0,0000	0.7971	0.0002
SST	0.3984	0.1265	0.6587	0.0055
MSLP	0.9557	0,0000	0.9308	0,0000
specific humidity at 850hPa level	0.0875	0.7472	0.9525	0.0000
specific humidity at 700hPa level	0.6621	0.0052	0.9358	0.0000
vorticity at 200hPa level	0.4087	0.1160	-0.2637	0.3237
divergence at 200hPa level	-0.8031	0.0002	0.9058	0.0000
vorticity at 850hPa level	0.9172	0.0000	-0.8464	0.0000
divergence at 850hPa level	0.8691	0.0000	-0.8548	0.0000
wind at 850hPa level	0.3142	0.2360	0.8800	0.0000
wind at 700hPa level	0.3142	0.2360	0.8604	0.0000
wind at 200hPa level	0.1777	0.5102	0.0430	0.8742
vertical shear between levels 700hPa and wind 850hPa	-0.1638	0.5444	0.0147	0.9570
vertical shear between levels 200hPa and wind 850hPa	-0.1222	0.6520	-0.3461	0.1891

Table 1: table of correlation coefficient

Our selection criteria is based simultaneously on the correlation coefficient and the behavior of the average variation that each variable must be consistent with the already known results [3] [4] [5]. We will hold the variables with a coefficient of correlation superior to 0.80 and a probability of non-correlated zero. Variables that meet this criterion are:

- ❖ the **mean pressure at the surface of the sea** : minimum of **mean pressure at the surface of the sea** ;
- ❖ the **OLR** : minimum of the **OLR** ;
- ❖ the **divergence** : minimum of **divergence of wind at 850hPa level** and maximum of **divergence of wind at 200hPa level**;
- ❖ the **vorticity** : minimum of **vorticity at 850hPa level** ;
- ❖ the **specific humidity** : maximum **specific humidity at 850hPa level** and maximum **specific humidity at 700hPa level**;
- ❖ the **wind** : maximum **wind at 850hPa level** and maximum **wind at 700hPa level**;

V. STUDY OF EIGENVALUES OF RAW DATA MATRIXES

V-1 Practical approach.

- We have followed the daily position of mobile points where the values of the variables in correlation with the cyclogenesis are extremal.
- We determined, every time, the daily value of the data matrix in a squared zone of $2^\circ \times 2^\circ$. This zone is centered on the daily position of the mobile point. With the spatial resolution of $1^\circ \times 1^\circ$, we then obtained a matrix 3×3 .
- Using a program published under the software MATLAB, we sought the eigenvalues of each data matrix.

- We have studied the spectrum of values thus obtained for the 15 days prior to the formation of cyclones in the zone of cyclogenesis.

V-2 Results and discussions

The figure 5 represents an example of curves, among so many others, representing the second eigenvalues of data matrixes in mobile areas of minimum or maximum, according to the correlation of each variable with the cyclogenesis, before the formation of cyclone CINDA.

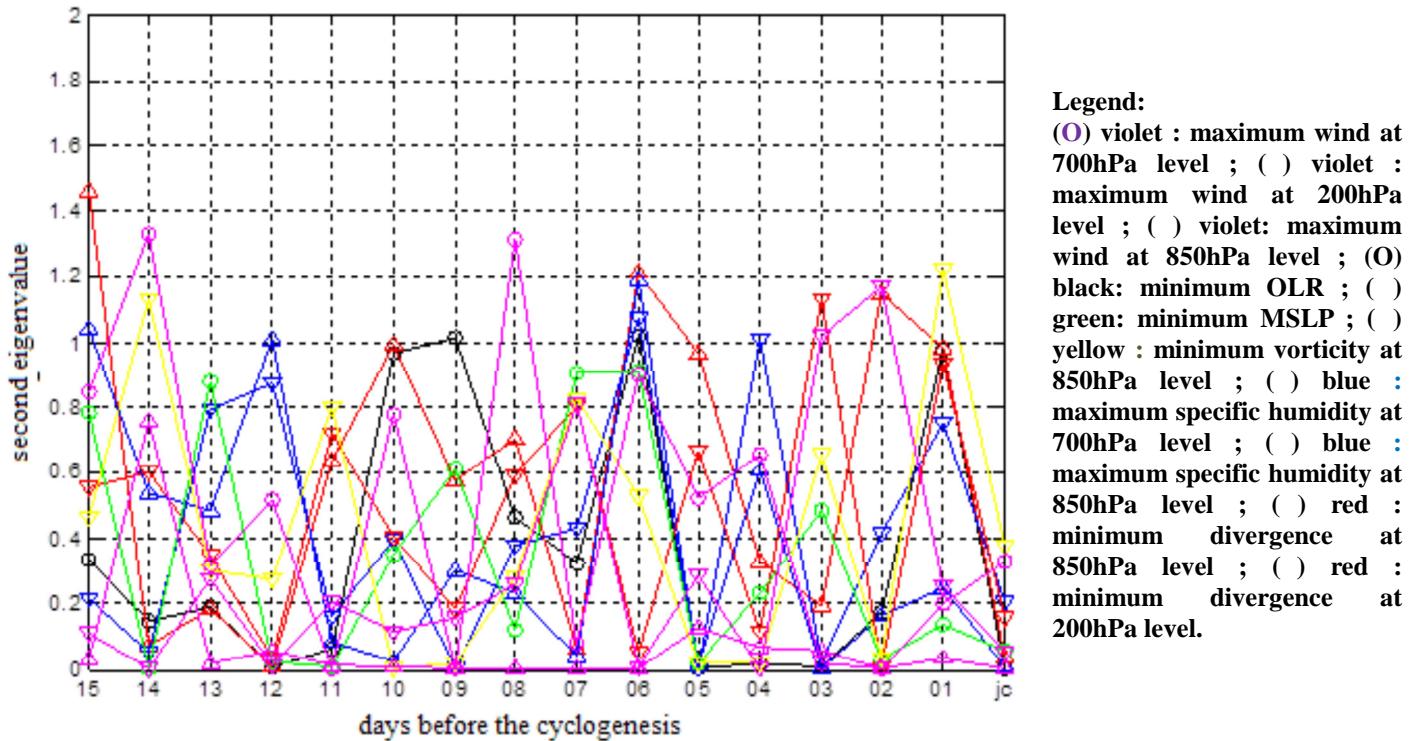


Figure 5: spectrum of second eigenvalues before the formation of cyclone CINDA.

On the figure 5, we see that:

- second eigenvalues of the variables take a value greater than 1 for some days before the cyclogenesis.
- second eigenvalues of the variables do not all superior to 1 at same day.

While considering all cyclones and while limiting in the interval of 15 days before until the day of cyclogenesis, we established the percentage of cyclones preceded of second eigenvalues superior to 1. These cyclones have arisen in the zone of cyclogenesis between 16 November and 30 April of the southern summers covering cyclone periods from 1989 to 2010.

Variables	Percentage of cyclones preceded by second eigenvalue superior to 1
Maximum of divergence at 200hPa level	96,1%
Maximum of specific humidity at 700hPa level	93,9%
Maximum of specific humidity at 850hPa level	93,4%
Maximum of wind at 700hPa level	92,8%
Maximum of wind at 850hPa level	92,3%
Minimum of divergence at 850hPa level	91,7%
Minimum of vorticity at 850hPa level	82,3%
Minimum of OLR	77,3%
Maximum of wind at 200hPa level	77,3%
Minimum of MSLP	64,6%

Table 2: percentage of cyclones preceded by second eigenvalue superior to 1

This result on the eigenvalue superior to 1 is in agreement with the criteria of KAISER in Principal Component Analysis [6] [7] [8].

The table 3 summarizes the percentage of second eigenvalues greater than 1 which have been followed by cyclogenesis, in descending order, at the end of 15 days.

Variables	Percentage of second eigenvalues greater than 1, followed by cyclones
Maximum of wind at 200hPa level	79,1%
Maximum of specific humidity at 850hPa level	73,7%
Maximum of wind at 850hPa level	70,6%
Maximum of specific humidity at 700hPa level	70,1%
Maximum of wind at 700hPa level	66,1%
Minimum of vorticity at 850hPa level	65,2%
Maximum of divergence at 200hPa level	65,1%
Minimum of divergence at 850hPa level	64,1%
Minimum of MSLP	63,9%
Minimum of OLR	63,4%

Table 3: Percentage of second eigenvalues greater than 1, followed by cyclones

We see that the percentage of second eigenvalues greater than 1 followed by cyclogenesis is lower than the percentage of cyclones preceded by second eigenvalue superior to 1 at the end of 15 days. This result can be interpreted by the fact that the condition of cyclone formation depends on several factors. A single variable taken individually would not give a high percentage.

For improve the percentage of second eigenvalue greater than 1 followed by cyclogenesis, we have imposed the second eigenvalues are simultaneously greater than 1 for all variables in correlation with the cyclogenesis. The result has shown us that on the periods of southern summers from 1989 to 2010, there was no day where the second eigenvalues of all variables are all superior to 1.

We passed then to more restricted combinations. We found that the **coupling of maximum wind at 200hPa level and specific humidity at 700hPa level provides the greatest percentage of second eigenvalue followed by cyclone.**

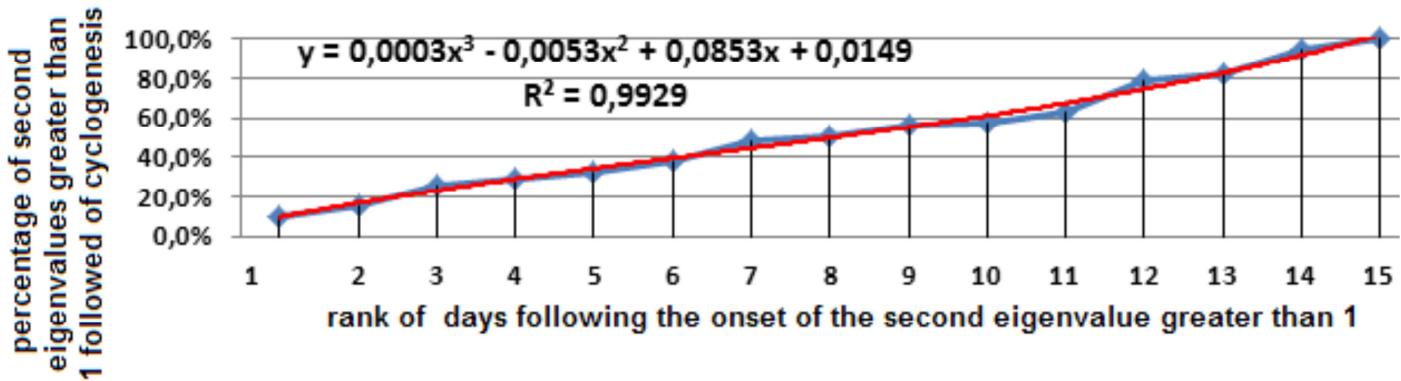


Figure 6: percentage of second eigenvalues greater than 1 followed of cyclogenesis by coupling the maximum wind at 200hPa level and the specific humidity at 700hPa level.

In this case of simultaneous consideration, figure 6 shows a significant improvement in the result : on the southern summers, from 1989 to 2010 periods, the percentage of second eigenvalues simultaneously greater than 1 followed by cyclogenesis reaches **100%** at the end of 15 days. The trendline equation is $y = 0,0003x^3 - 0,0053x^2 + 0,0853x + 0,0149$ where x is the rank of the day following the onset of the second eigenvalue greater than 1. Its coefficient of correlation $R^2 = 0,9929$ is very close to 1, one notices therefore that the linear regression is very relevant.

VI. PREDICTION OF THE POSITION OF CYCLONE FORMATION

VI-1 Practical approach.

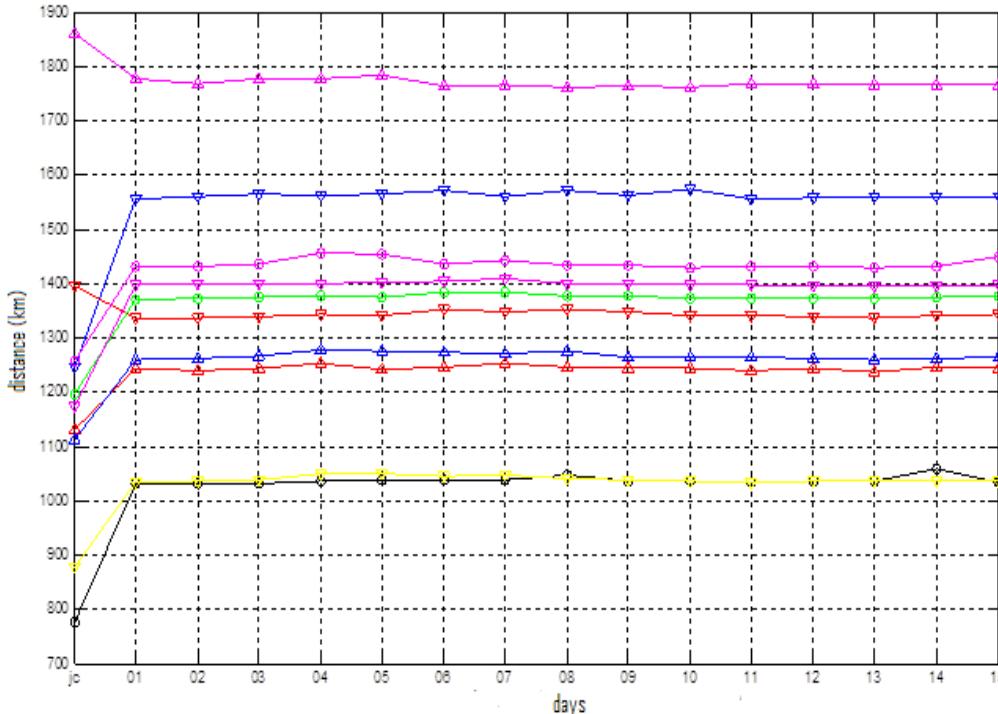
For every variable in interrelationship with the cyclogenesis and since 15 days before until the day of formation of every cyclone:

- we followed the daily position of the mobile point where the value of the variable is minimal or maximal (according to the existing interrelationship).
- we calculated the distance between the position of the mobile point where the value is extremal and the position of cyclone formation . So we got the distance variation since 15 days before until the day of formation of every cyclone.
- we calculated the average of this distance variation for all cyclones.

IV – 2 Results and discussions

The figure 7 represents the distance between positions of extremal values and positions of cyclone formation. This figure shows that:

- the mean distance between the position of the point of cyclone formation and the positions of the points where the values of the OLR and vorticity are minimum is the weakest ;
- on the day of cyclogenesis, the positions of the points of extremal values are closest to the position of the point of cyclone formation, except for the maximum of divergence of wind at 200 hPa level and the maximum of wind 200 hPa at level ;
- on the day of cyclogenesis, the position of the point where the value of the OLR is minimum is closest to the position of the point of cyclone formation ;
- before the cyclogenesis, the distances between the points where observed the extremal values and the position of the cyclone formation, are more or less constant;



Legend:

(O) violet : maximum wind at 700hPa level ; (Δ) violet : maximum wind at 200hPa level ; (∇) violet: maximum wind at 850hPa level ; (O) black: minimum OLR ; (O) green: minimum MSLP ; (Y) yellow : minimum vorticity at 850hPa level ; (Δ) blue : maximum specific humidity at 700hPa level ; (∇) blue : maximum specific humidity at 850hPa level ; (∇) red : minimum divergence at 850hPa level ; (Δ) red : minimum divergence at 200hPa level.

Figure 7 : distance between position of extremal values and position of cyclone formation at the day of cyclogenesis, at 1 day, at two days...at 15 days before the cyclogenesis

These results on the distances variation allow us to:

- ✓ first, take the OLR as reference in research of point of cyclone formation.
- ✓ Second, consider that the **first likely position** to meet the point of cyclone formation is at the center of inertia of the system formed by the point of minimum of OLR and the point of minimum of vorticity at 850hPa level.
- ✓ Thirdly, to consider that the **second likely position** to meet the point of cyclone formation is at the center of inertia of the system formed by the point of minimum of OLR, the point of minimum of vorticity at 850hPa level, the point of maximum of divergence at 200hPa level and the point of pressure favorable to cyclogenesis.
- ✓ quarto, consider that likely positions to meet the point of cyclone formation is within an area bounded by a circle whose center is the first likely position and whose radius is the distance between the first probable position and the second probable position. We will call this area "probable area of cyclogenesis".

The problem is then to find the coefficients that intervene in the calculation of the barycenter. To resolve this problem, we attributed the coefficient 1 to the variable OLR. We allocated to every variable i the coefficient of correlation between the extrémal values of this variable i with its values in the mobile point of minimum of OLR.

Noting by c_i : The coefficient of correlation of extrémal values of a variable i with its values in the mobile point of minimum of OLR, l_o : the latitude of the mobile point of minimum of OLR, l_i : the latitude of the point of extremal values of a variable i , l : the probable latitude of the point of cyclone formation, L_o : the longitude of the point of minimal value of the OLR, L_i : the longitude of the point of extremal values of a variable i , L : the probable longitude of the point of cyclone formation, n : number of variables in correlation, the formula of barycenter gives us the probable positions of cyclone formation when the indication precursor appears.

$$l = \frac{l_o + \sum_{i=1}^n c_i * l_i}{1 + \sum_{i=1}^n c_i} \quad (1)$$

$$L = \frac{L_o + \sum_{i=1}^n c_i * L_i}{1 + \sum_{i=1}^n c_i} \quad (2)$$

While applying these equation (1) and equation (2) in a program published under the software MATLAB, we got the likely positions and the real position where will form itself the INDLADA cyclone, March 12, 2007 at 00 local hour to the latitude

-13.40° and longitude 58.10°.

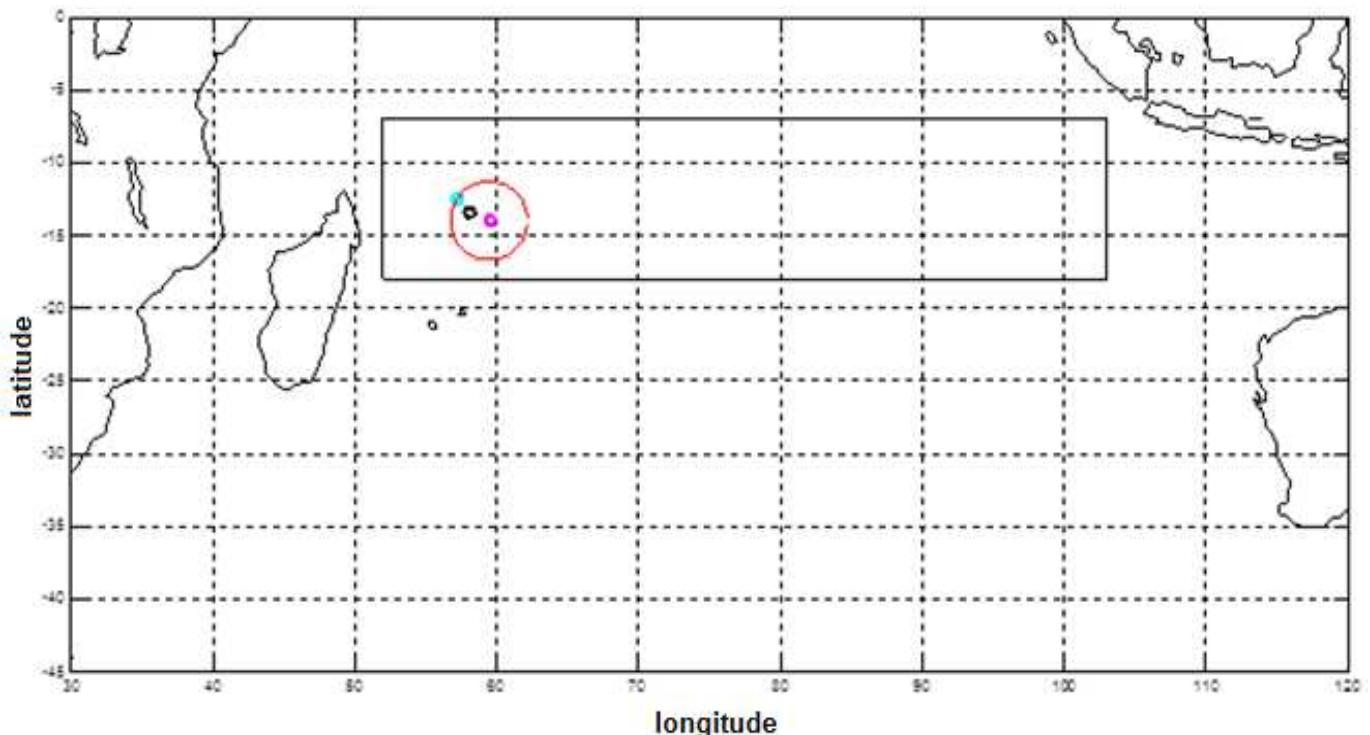


Figure 8 : likely positions and real position of the point of formation of the INDLADA cyclone March 12, 2003 to two days before its formation.

Black point: real position of cyclogenesis; purple point: first probable position of cyclogenesis; green point: second probable position of cyclogenesis.

On the figure 8, the forecasting is done at two days before the cyclogenesis, that means March 10, 2007. The distance between the real position of cyclogenesis and the first probable position is 166km; the one between the real position and the second probable position is 137km. INDLADA is inside the probable zone of cyclogenesis.

Another cyclone is already formed on the day of forecasting of INDLADA, it was the JACOB cyclone. It was close to the Australian coast at the latitude -16.20° and longitude 110.20 °. These two points are separated by a distance of 5620km. Because of this distance, JACOB's presence didn't have an influence on the forecasting of the point of formation of INDLADA.

CONCLUSION

The zone of cyclonic activity in the Indian Ocean basin lies between 7° to 18° of latitude South and 52° to 103° of longitude East. The simultaneous apparition of eigenvalue superior to 1 in an area of maximum wind at 200hPa level and in an area of maximum specific humidity at 700hPa level is an indication precursor of cyclogenesis. The probable position of the foreseen cyclone formation is inside a circle that delimits a probable area of cyclogenesis.

BIBLIOGRAPHY.

- [1] SUZANA J. CAMARGO, ADAM H. SOBEL: « Characteristics of Western North Pacific Model Tropical Cyclogenesis »
- [2] SUZANA J. CAMARGO, ADAM H. SOBEL, ANTHONY G. BARNSTON1 and KERRY A. EMANUEL: « Tropical cyclone genesis potential index in climate models »
- [3] <http://www.cyclonextreme.com/cyclonedicostructure.htm>
- [4] etudiant.univ-lv.fr/.../conditions_de_formation.html
- [5] www.meteo.fr/.../cyclone/.../cyclone/lieux_periodes.htm
- [6] Gilbert SAPORTA : « Probabilités analyses des données et statistique »
- [7] Sylvie VIGUIER : « Analyse des données »
- [8] Dominique DESBOIS : « Une introduction à l'Analyse en Composantes Principales avec SPSS pour Windows »

PREVISION DE L'INTENSITE DES CYCLONES TROPICAUX DU BASSIN SUD-OUEST DE L'OCEAN INDIEN PAR LES RESEAUX DE NEURONES ARTIFICIELS (RNA)

Résumé

Le présent travail a pour objectif d'améliorer la prévision des intensités de cyclones tropicaux par la méthode statistique. La prévision des intensités des cyclones est faite avec le Réseau de Neurones Artificiels (RNA) et la Régression Linéaire Multiple (RLM), sur les données du bassin Sud-Ouest de l'Océan Indien. La méthode de Gram Schmidt Modifié a été utilisée pour ranger les prédicteurs. Puis une autre technique utilisant des prédicteurs aléatoires a été appliquée pour sélectionner les entrées du modèle.

Comparé avec la régression linéaire, le réseau de neurones artificiels a l'avantage d'avoir des erreurs moyennes de prévision inférieures, sur un ensemble de cyclone test.

Mots clés: *cyclone, intensité, réseau de neurones artificiels, Gram-Schmidt modifié, régression linéaire multiple.*

Abstract

The present work has for objective to improve the tropical cyclones intensities forecast by statistical method. Tropical cyclone intensities are predicted with the artificial neural network and the multiple linear regression using the Indian Ocean South West basin data. The modified Gram Schmidt method is applied for putting in order candidates predictors. Then another method allowed us to choose the number of predictors more effective than those taken randomly.

Compared with the linear regression, the artificial neural network has an lower mean errors on the forecast in the cyclones test.

Keywords: *cyclone, intensity, artificial neural network, modified Gram-Schmidt, multiple linear regression.*

Directeur de Thèse

Professeur RATIARISON Adolphe A.

Impétrant

RAKOTOARISON Soloheray Hajanirina

e-mail : soloheray77@gmail.com

Lot IAF 9A Antanetibe Ambohidrapeto Itaosy
102 Antananarivo - Madagascar