

Université d'avant-garde
INRS
Centre Eau Terre Environnement

Modélisation de l'épaisseur de glace des lacs par les Réseaux de Neurones Artificiels

Par : Imen Zaier

Mémoire présenté pour l'obtention du grade
Maîtrise ès Sciences (M.Sc.)

Jury d'évaluation

Examineur externe	Musandji Fuamba
Examineur interne	André St-Hilaire
Codirecteur	Karem Chokmani
Directeur	Taha B.M.J Ouarda

Novembre 2008

Remerciements

Ce mémoire n'aurait pas pu voir le jour sans la participation et la collaboration d'un grand nombre de personnes, chacun ayant apporté une touche personnelle à la conception de ce mémoire.

Un grand merci à mon directeur de recherche Taha B.M.J OUARDA, pour m'avoir accueilli au sein de son équipe de recherche et pour avoir assuré mon encadrement tout au long de l'élaboration de ce travail, ses directives, ses précieux conseils et son assistance ont été d'une grande utilité.

Je tiens également à exprimer ma reconnaissance et ma gratitude à Fateh CHEBANA et Chang SHU, stagiaires postdoctoraux à l'INRS, pour leur attention et l'intérêt qu'ils ont porté à ce mémoire.

Sincères remerciements à Karem CHOKMANI et Ousmane SEIDOU, professeurs chercheurs, pour tous les conseils qu'ils m'ont donnés au cours de l'année, afin que ce travail fastidieux et intéressant se déroule au mieux.

Merci à Laurent BILODEAU et Georges DESROCHERS, chercheurs à Hydro-Québec, pour m'avoir si bien accueillie, et pour tous leurs conseils et l'aide qu'ils m'ont apportés afin de rendre les recherches moins difficiles.

Je tiens également à remercier Hydro-Québec et le CRSNG pour leur soutien financier dans la collaboration de ce mémoire.

Des remerciements tous particuliers à mon "double" Mohamed SEHIL, pour ses connaissances en informatique, ses conseils, ses idées nombreuses, sa patience et sa gentillesse.

Des remerciements à mes amis, et surtout à Patrick GAGNON, étudiant au doctorat à l'INRS, pour l'intérêt qu'ils ont porté à l'accomplissement et à la bonne mise en œuvre du mémoire et leur soutien moral constant.

Un énorme merci à toute ma famille, et tout particulièrement à ma maman Fatma, mon papa Taher et mes deux frères Anis et Hmida, qui m'ont apporté une aide prodigieuse dans bien des domaines, pour leur soutien moral, et pour l'intérêt et l'attention qu'ils ont porté à la conception de ce mémoire.

Résumé

En milieu nordique, un des problèmes auquel les ingénieurs barragistes doivent faire face est celui de la présence de glace dans les réservoirs. L'estimation et la prévision des épaisseurs de glace en réservoir s'avèrent nécessaires pour permettre la prise en compte de cette variable de première importance dans la conception et la gestion des ouvrages hydrauliques.

Ce mémoire de maîtrise est basé sur une technique de modélisation éprouvée en science et génie : les réseaux de neurones artificiels. L'originalité du travail repose sur une première application des ensembles de réseaux de neurones pour la prédiction des épaisseurs de glace.

L'approche des prédictions en utilisant l'ensemble des réseaux de neurones artificiels est une technique dans laquelle les sorties des réseaux de neurones artificiels entraînés séparément sont combinées, dans le but de former une prédiction unique. L'ensemble des réseaux de neurones artificiels est développé dans cet article, afin d'améliorer les résultats de l'approche de réseaux de neurones artificiels simples pour l'estimation de l'épaisseur de la glace dans des nombreux lacs canadiens durant l'hiver précoce pour la période de la croissance de la glace. Un ensemble efficace se compose de plusieurs réseaux de neurones artificiels qui pourraient ne pas être très performants quand ils sont entraînés séparément, mais, une fois combinés, leur erreur de prédiction est grandement réduite. Ce mémoire évalue l'efficacité de plusieurs techniques incluant l'approche aléatoire, le bagging et le boosting qui contribuent à la création des membres de l'ensemble et les techniques de la moyenne et de l'empilage qui permettent la combinaison de ces membres. Les expériences montrent que, dans le contexte de l'estimation de l'épaisseur

de glace des lacs, le boosting est plus performant que l'approche aléatoire et parfois meilleur que le bagging. L'empilage est plus compétitif que la moyenne. Au final, les modèles de l'ensemble de réseaux de neurones artificiels pour l'estimation de l'épaisseur de glace sont plus efficaces que les modèles de réseaux de neurones simples, en particulier, quand le boosting est utilisé pour combiner les membres de l'ensemble et l'empilage pour la combinaison des sorties des membres individuels. L'ensemble des réseaux de neurones artificiels accomplit une meilleure performance quand la taille de l'ensemble atteint les 20 membres.

En ce qui concerne la seconde partie de ce mémoire, l'épaisseur de la glace dans le passé a été évaluée par les deux modèles LSR et RNA en utilisant les données climatiques observées sur le réservoir la Grande IV. La comparaison de ces modèles a montré que le modèle RNA est plus performant que celui de LSR. Des scénarios de changements climatiques sont ensuite mis à contribution afin de produire une estimation du changement du régime des glaces dans un site du bassin de la rivière La Grande, région importante pour la production hydroélectrique québécoise. Les changements anticipés sont significatifs. Ces résultats devraient interpeller la communauté des ingénieurs hydrauliciens qui se penchent sur la conception des ouvrages hydrauliques dans ce milieu.

Tables des matières

Remerciements.....	iii
Tables des matières	vii
1. Synthèse et connaissances pré-requises.....	1
1.1 Introduction	1
1.2 Les Réseaux de Neurones Artificiels.....	5
1.2.1 Les Réseaux de Neurones Artificiels Simples.....	5
1.2.2 Ensemble de Réseaux de Neurones Artificiels	10
1.3 La Loi de Stefan Révisée:	12
1.4 Contribution de l'étudiante à ce mémoire	14
2. Estimation de la glace des lacs en utilisant l'ensemble de réseaux de neurones artificiels	21
3. Modélisation de l'épaisseur de glace des lacs en utilisant les réseaux de neurones artificiels et la Loi de Stefan Révisée dans le contexte de changement climatique (RAPPORT)	63
Liste des figures	64
Liste des tableaux	65
3.1 Introduction	66
3.2 Modélisation.....	68
3.2.1 Modèle déterministe : Loi de Stefan Révisée	70
3.2.2 Modèle statistique non paramétrique: Réseau de Neurones Artificiels ...	71
3.3 Application.....	75
3.3.1 Site d'application	75

3.3.2 Données utilisées	75
3.3.3 Description des caractéristiques des deux modèles	77
3.4. Résultats	78
3.4.1 Situation passée : Évaluation avec les données observées.....	78
3.4.2 Situations passée et future : Évaluation avec les données simulées du MRCC (1961-2070).....	80
3.5. Conclusion.....	97
3.6 Références bibliographiques	99
4. Conclusion générale	101

1. Synthèse et connaissances pré-requises

1.1 Introduction

Le changement climatique constitue l'une des plus grandes menaces pour le bien être futur de toute la planète. Il s'agit d'une menace non seulement pour l'environnement, mais également pour l'économie (production hydro-électrique, navigation,...) et le mode de vie. Parmi les nombreux dangers qui en découlent, il y a la modification du régime thermique des plans d'eau, et la mise en péril de toutes les espèces qui en dépendent. Ces espèces sont extrêmement sensibles entre autres à la température de l'eau et au régime des glaces (e.g, Morse et Hicks, 2005).

Dans les régions nordiques, un changement de la température moyenne d'un lac modifiera l'intensité des débâcles de la glace dans un système aquatique (Beltaos et Prowse, 2001). Les perturbations physiques associées par la fonte de la glace et le brassage de l'eau influent directement sur la matière organique, la composition chimique de l'eau, ainsi que la diversification et l'abondance du biotope (Scrimgeour et al., 1994; Milburn et Prowse, 2000). De plus, la fonte incitée encourage la migration des sédiments et des nutriments qui sont des éléments nécessaires pour la qualité de l'eau des écosystèmes (Marsh, 1986; Lesack et al., 1991; Wilkins et Calkins, 1993). Donc, la structure d'un système aquatique nordique dépend étroitement de la durée et de l'intensité des débâcles et des embâcles qui s'y passent (Hirst, 1984; Prowse et Reedyk, 1993). En général, si le climat change, il modifiera à long terme la nature de la dynamique de la glace, et par la suite, les caractéristiques biologiques d'un lac seront significativement altérées. Si le changement de la température moyenne d'un système aquatique tend à

réduire la quantité d'eau qui s'y trouve, des répercussions, comme la sécheresse, pourraient être observées dans les réservoirs (Prowse et Conly, 1998). Par conséquent, ces répercussions touchent l'équilibre des espèces de poissons qui y vivent et par la suite les êtres humains à travers la pêche.

Du point de vue socio-économique, quand la fréquence et l'intensité des embâcles et des débâcles augmentent, comme résultat du changement climatique, les infrastructures seront l'objet d'un nouveau risque d'inondation. Alors, il est important d'ajouter les composantes du changement climatique dans la conception des infrastructures, qui seront capables de bien contrôler la dynamique de la glace dans le futur (i.e. barrages, ponts, ouvrages hydrauliques...). Pour la même raison, il est peut être aussi indispensable de réviser les conceptions des infrastructures qui existent déjà, dans le but d'assurer la sécurité de l'être vivant sous la projection des différents scénarios du changement climatique. Les deux secteurs économiques qui peuvent être affectés par le régime de la glace des rivières et des lacs canadiens, reliés au changement climatique, sont le transport et la génération hydro-électrique. La glace peut créer des problèmes et des difficultés hydrotechniques pour les équipements et les opérations hydro-électriques. Ensuite, la dynamique imprévue de la glace a causé des dommages économiques (e.g. van der Vinne et al., 1991). Par exemple, au Nouveau Brunswick, plus de 13 millions de dollars canadiens ont été dépensés pour réparer les pertes causées par un seul événement exceptionnel dans les trafics maritimes. Les impacts qui ont été discutés précédemment montrent l'importance de l'étude du régime de glace des lacs canadiens sous la contrainte du réchauffement planétaire. Il est à noter que les recherches scientifiques portant sur la

glace des lacs ont fait des avancées considérables et se sont multipliées au cours des dernières décennies (e.g, Caine, 2002; Dugay et al., 2003 et 2005). Elles constituent une part importante des recherches effectuées dans le domaine de l'hydrologie.

Les modèles hydrologiques sont conventionnellement assignés à deux larges catégories : Les modèles déterministes (physiques) et les modèles empiriques (Anderson et Burt, 1985; Watts, 1997). La première catégorie décrit le processus hydrologique en utilisant les lois physiques de transfert de masse et d'énergie. La deuxième catégorie utilise des fonctions de transfert stochastiques (comme les équations linéaires de régression multiple) pour relier des variables météorologiques, qui représentent les variables explicatives, à la variable expliquée, par exemple l'épaisseur de la glace dans les lacs.

Dans la présente étude, on utilise deux modèles hydrologiques, à savoir le modèle des réseaux de neurones artificiels et le modèle de la Loi de Stefan Révisée (Seidou et al., 2006). Le premier modèle est un modèle empirique non paramétrique. Il a été récemment utilisé pour la modélisation de l'épaisseur de glace des lacs dans les pays nordiques tout en considérant les données climatologiques et hydrologiques disponibles dans les lacs étudiés. Le deuxième est un modèle déterministe basé sur les équations physiques.

Dans la première partie, une nouvelle approche, appelée «l'ensemble de réseaux de neurones artificiels », a été appliquée pour la modélisation de l'épaisseur de la glace dans plusieurs lacs canadiens. Le but de cette approche est d'améliorer le modèle de

réseau de neurones simple dont les résultats sont présentés dans Seidou et al. (2006). Cette amélioration a pour but de réduire l'incertitude de prédiction de l'épaisseur de glace.

Dans la deuxième partie, l'approche du réseau de neurones artificiels simple et celle de la Loi de Stefan Révisée ont été appliquées pour le réservoir La Grande IV, qui se trouve dans le nord de la province du Québec. Le but est de prédire l'épaisseur de glace à l'horizon de 2050. Les données climatologiques intégrées dans les modèles prennent en considération le dédoublement de la quantité de CO₂ dans l'atmosphère qui va participer au changement climatique.

Les études portant sur la relation entre les variables climatologiques et le comportement des paramètres de la glace, incluant leur analyse et leur interprétation, ont pris une grande importance dans le concept de changement climatique, vers les années quatre-vingts. À partir de cette période, plusieurs chercheurs ont orienté leurs recherches en ce sens. Les travaux concernant les relations entre les dates de fonte, les dates de formation de la glace ainsi que la température de l'air ont été réalisés dans plusieurs régions du monde en utilisant différents types de modèles. Par exemple, une étude menée sur le lac Baïkal en Russie a été faite par Livingstone (1999). Cette étude a établi une relation entre les débâcles et la température de l'air locale et régionale de la Sibérie. Dans une autre étude, Livingstone (1997) a conclu que la relation entre la température de l'air au printemps et les dates de fonte sur un lac alpin a été confirmée pour le lac Lej da San Murezzan (Alpes Suisses) et d'autres lacs situés au Royaume-Uni et en Hollande. Il a été

démontré que les dates de fonte sont fortement corrélées avec la température de l'air de mi-avril à mi-juin.

Dans ce qui suit, les outils qui ont été utilisés dans ce travail sont décrits, à savoir les réseaux de neurones artificiels simples, les ensembles de réseaux de neurones artificiels ainsi que la Loi de Stefan Révisée.

1.2 Les Réseaux de Neurones Artificiels

Les réseaux de neurones artificiels sont divisés en deux types, simple ou ensemble.

1.2.1 Les Réseaux de Neurones Artificiels Simples

Les réseaux de neurones artificiels (Zurada, 1992) constituent une méthode d'approximation des systèmes dynamiques complexes, par exemple un lac ou une rivière. Les prédictions des aspects hydrologiques de ces systèmes peuvent être à court terme ou à long terme, comme dans le cadre de ce projet. Ces modèles sont particulièrement utiles lorsque ces systèmes sont difficiles à modéliser à l'aide des méthodes statistiques classiques (Dawson et Wilby, 2001). Ils constituent alors une alternative prometteuse aux techniques traditionnelles pour le traitement temporel de l'information.

Du point de vue de l'ingénierie, les réseaux de neurones sont des structures calculatoires qui tentent à reproduire des facultés cérébrales. Un neurone artificiel est une représentation mathématique et informatique d'un neurone biologique. Il représente l'unité élémentaire des réseaux de neurones artificiels dans lesquels il est associé à ses

semblables pour approximer des fonctions complexes, utilisées dans diverses applications en intelligence artificielle.

Une analogie existe entre le fonctionnement du neurone artificiel et le neurone biologique (voir figure 1.1). Elle peut être expliquée comme suit : le neurone artificiel possède généralement plusieurs entrées et une sortie qui correspondent respectivement aux dendrites et au cône d'émergence du neurone biologique (point de départ de l'axone). Les actions excitatrices et inhibitrices des synapses sont représentées, la plupart du temps, par des coefficients numériques (les poids synaptiques) associés aux entrées et qui constituent les degrés de liberté élémentaires du système. Les valeurs numériques de ces coefficients sont ajustées dans une phase d'apprentissage. Dans sa version la plus simple, un neurone artificiel calcule la somme pondérée des entrées reçues. Puis, il applique à cette valeur une fonction d'activation, généralement non linéaire. La valeur finale obtenue est la sortie du neurone.

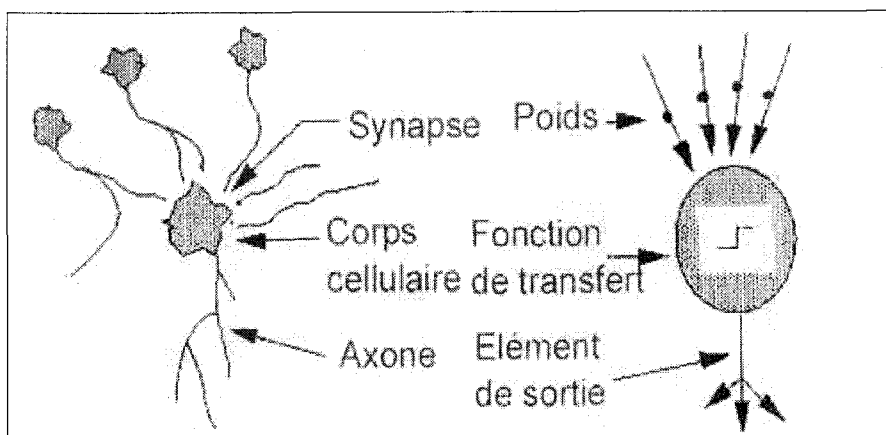


Figure 1.1 Analogie entre le neurone artificiel et le neurone biologique

Diverses architectures de réseaux de neurones diffèrent par la manière dont les neurones sont arrangés, par la méthode avec laquelle les poids sont déterminés (algorithme d'apprentissage) et par le domaine d'application visé (cible).

Haykin (1999) a présenté la définition suivante d'un réseau de neurones: Un réseau de neurones est un processeur qui a une propension naturelle pour emmagasiner la connaissance expérimentale et à la rendre disponible pour utilisation ultérieure. Il ressemble au cerveau selon deux aspects:

- La connaissance est acquise par le réseau à travers un processus d'apprentissage.
- Les forces de connexion inter-neurones appelées poids synaptiques sont utilisées pour l'emmagasinage de l'information.

Dans ce contexte, les réseaux de neurones sont considérés comme « une boîte noire ». Par conséquent, l'utilisation des réseaux de neurones ne présuppose pas la compréhension détaillée des caractéristiques physiques d'un apprentissage. Cela n'exige pas non plus un vaste prétraitement de données, parce qu'un modèle de réseaux de neurones peut, théoriquement, manipuler des données incomplètes, bruyantes et ambiguës (Maier et Dandy, 2000). De plus, les réseaux de neurones artificiels sont souvent plus simples à implémenter par rapport aux autres modèles qui sont basés sur les phénomènes physiques (Campolo et al., 1999). Ils conviennent aussi aux problèmes dynamiques (par

exemple, les lacs et les rivières) et sont parcimonieux en termes de stockage de l'information dans le modèle formé (Thirumalaiah et Deo, 1998).

Dans un contexte général, Haykin (1999) a spécifié que les réseaux de neurones artificiels possèdent les caractéristiques et les propriétés suivantes:

- Non-linéarité : Le neurone élémentaire est un dispositif non-linéaire. Par conséquent, l'interconnexion de plusieurs neurones résulte en un système global lui-même non-linéaire. Cette propriété importante permet l'emploi des réseaux de neurones pour la modélisation de systèmes où ils existent des non-linéarités inhérentes.

- Relation d'entrée/sortie : Un paradigme populaire d'apprentissage est l'entraînement supervisé qui consiste à présenter au réseau des exemples d'apprentissage. Chaque exemple comporte un signal d'entrée et une réponse de sortie désirée. Les poids synaptiques sont modifiés de sorte à minimiser l'erreur entre la réponse voulue et la réponse effective du réseau. En alimentant le réseau avec plusieurs exemples et en répétant l'apprentissage plusieurs fois, il finit par se stabiliser et converger. Le réseau construit donc progressivement une représentation implicite de la relation entrée/sortie. Il a été démontré que les réseaux de neurones artificiels constituent des approximateurs universels Cybenko (1989). Ils peuvent approximer des relations fonctionnelles entrée/sortie non-linéaires arbitraires. Cette approximation se fait avec un certain degré minimal de continuité et avec un nombre suffisant d'unités de traitement (neurones).

- Adaptabilité : Les réseaux de neurones artificiels ont une capacité inhérente d'adaptation en vertu de leurs connexions synaptiques ajustables. Dans un contexte d'environnement

non-stationnaire, un réseau de neurones peut être conçu pour opérer en temps réel. Cela peut constituer un atout avantageux, par exemple, pour le traitement adaptatif du signal.

- Tolérance aux pannes : Un réseau de neurones implémenté a le potentiel d'être tolérant aux pannes dans le sens où la performance se dégrade doucement avec des conditions adverses. Autrement dit, même si la performance du modèle n'est pas bonne, ce dernier produit quand même des résultats qui sont à leurs tours douteux

Le premier modèle de base d'un neurone artificiel a été proposé par McCulloch et Pitts (1943). Depuis ce temps jusqu'à la fin des années quatre vingt-dix, la recherche dans le domaine des réseaux de neurones s'est développée dans trois phases distinctes (Schalkoff, 1997) :

- La première ère s'est concentrée, principalement sur le développement du neurone artificiel. Elle s'étend jusqu'à Minsky et Papert (1969), tout en identifiant plusieurs facteurs de limitation.

- La deuxième ère a commencé par la découverte et la vulgarisation de l'algorithme d'apprentissage de la rétro-propagation (Rumelhart et McClelland, 1986). Auparavant, il était très difficile de former les réseaux neuronaux d'une taille pratique (nombre limité et bien défini des neurones qui constituent le réseau) pour le matériel informatique.

- La troisième ère est caractérisée par la fusion des réseaux neuronaux avec d'autres technologies, comme les algorithmes génétiques et la logique floue.

Mathématiquement, le traitement que fait le neurone élémentaire peut être représenté par un neurone artificiel à m entrées, auquel m entrées notées x_1 à x_m sont soumises. Il

permet d'associer aux m entrées une sortie : c'est donc une fonction à m variables et à valeurs réelles. Dans le modèle de McCulloch et Pitts (1943), à chaque entrée est associé un poids synaptique, c'est-à-dire une valeur numérique notée w_1 pour l'entrée 1 jusqu'à w_m pour l'entrée m . La première opération réalisée par le neurone consiste en une somme des grandeurs reçues en entrées, pondérées par les coefficients synaptiques, c'est-à-dire la somme :

$$w_1x_1 + \dots + w_mx_m = \sum_{j=1}^m w_jx_j \quad (1.1)$$

A cette égalité s'ajoute un seuil w_0 . Le résultat est alors transformé par une fonction d'activation non linéaire F . La sortie Y associée aux entrées x_1 à x_m est ainsi donnée par :

$$Y = F(w_0 + \sum_{j=1}^m w_jx_j) \quad (1.2)$$

1.2.2 Ensemble de Réseaux de Neurones Artificiels

Le réseau de neurone artificiel simple représente l'unité élémentaire de l'ensemble de réseaux de neurones (Shu et Burn, 2004). La formation de cet ensemble comporte deux étapes. La première étape se base sur la création des membres de l'ensemble. La deuxième étape se concentre sur la combinaison des sorties de ces membres. La création des membres peut se faire par trois approches, à savoir :

- La randomisation : C'est la méthode classique qui consiste à la création d'un réseau de neurones artificiels simple et où les poids associés aux entrées sont distribués aléatoirement.

- Le boosting : C'est un algorithme d'apprentissage automatique. C'est un principe qui regroupe de nombreux algorithmes. Ces derniers s'appuient sur des ensembles de prédicteurs. Le boosting optimise les performances de ces prédicteurs. Le principe est issu de la combinaison de prédicteurs. Par itérations successives, la connaissance d'un faible prédicteur est ajoutée au prédicteur final. Un des algorithmes les plus utilisés en boosting s'appelle AdaBoost.R2 et a été appliqué dans le cadre de cette étude (Freund et Schapire, 1996)

- Le bagging : C'est un acronyme de Bootstrap aggregating (Breiman, 1996a). Cette approche est basée sur le bootstrap comme une technique de re-échantillonnage (Efron et Tibshirani, 1993). Le but de cette technique est de générer plusieurs sous-ensembles d'apprentissage qui vont contribuer à l'apprentissage des membres qui forment l'ensemble.

Dans la deuxième étape, une fois que les membres sont formés, deux approches ont été abordées pour la combinaison des sorties de ces membres, à savoir la moyenne et l'empilage qui sont décrits comme suit :

- La moyenne : dans cette approche on trouve deux sortes de moyennes. La première est une moyenne arithmétique, c'est-à-dire la somme des valeurs des sorties de chaque membre divisée par le nombre des membres. Elle s'écrit sous la forme :

$$\bar{x} = \frac{1}{n} \sum_{i=1}^n x_i \quad (1.3)$$

où n est le nombre de membres qui forment l'ensemble.

La deuxième moyenne est une moyenne pondérée. En statistique, si on considère que l'ensemble $X = \{ x_1, x_2, \dots, x_n \}$ représente l'ensemble des sorties des membres de

l'ensemble, et que l'ensemble $W = \{w_1, w_2, \dots, w_n\}$ représente les poids positifs correspondants à chaque sorties. La moyenne pondérée \bar{x} est calculée suivant la formule :

$$\bar{x} = \frac{\sum_{i=1}^m w_i x_i}{\sum_{i=1}^m w_i} \quad (1.4)$$

Dans le cas général, le poids w_i représente l'influence de l'élément x_i par rapport aux autres.

- L'empilage : Il s'agit d'un moyen d'assemblage d'estimateurs. Le but est de faire en sorte que ces derniers produisent leurs propres biais tout en conservant un apprentissage particulier. Par la suite, on procède au filtrage de ces biais prend lieu (Wolpert, 1992). Quand l'empilage est utilisé pour la prédiction d'une nouvelle valeur, en premier lieu cette valeur est mise au niveau zéro des prédicteurs. Chacun de ces prédicteurs calcule sa valeur correspondante. Ces valeurs sont incorporées dans le niveau 1, qui assure leur combinaison pour donner la prédiction finale. Selon Breiman (1996b), l'algorithme d'apprentissage se base sur la minimisation de la fonction G par rapport à C_N :

$$G = \sum_{i=1}^m \left[x_i - \sum_{k=1}^N c_k \hat{x}_i^k \right]^2, \quad c_N > 0. \quad (1.5)$$

Les coefficients $\hat{c}_1, \hat{c}_2, \dots, \hat{c}_K$ dans l'équation (5) sont estimés dans le but de former la sortie finale de l'ensemble, donnée par

$$\hat{x}_i = \sum_{k=1}^N c_k \hat{x}_i^k \quad i = 1 \dots m. \quad (1.6)$$

1.3 La Loi de Stefan Révisée:

La Loi de Stefan qui évalue l'épaisseur de glace est basée sur la somme degrés-jours. Cette dernière est calculée à partir de la température moyenne journalière. Elle

représente la somme de la température moyenne journalière au dessus de zéro durant l'hiver. Si la valeur de degré-jours quotidienne obtenue est négative, on lui affecte la valeur zéro. Toutes les valeurs de degrés-jours quotidiennes sont accumulées pendant la saison de croissance de la glace.

Dans ce contexte, selon Lock (1990), l'équation la plus utilisée est celle de la Loi de Stefan est sous la forme suivante:

$$E = k\sqrt{Dj} \quad (1.7)$$

avec E est l'épaisseur de glace, Dj est la somme degrés-jours au-dessous du point de congélation depuis la naissance de la glace dans une année donnée et k est une constante.

La date de naissance de la glace est nécessaire comme paramètre dans la Loi de Stefan afin de déterminer la date où l'accumulation de la congélation des degrés-jours a commencé, pour un hiver donné. Cependant, cette date est généralement inconnue. Par conséquent, Seidou et al. (2006) ont utilisé la Loi de Stefan Révisée, qui a été considérée pour cette étude. Cette loi se base sur la variable Djs , qui représente l'accumulation des degrés-jours commençant par le premier jour où la température de l'air est en dessous de zéro dans une saison donnée. Cette loi a la forme suivante :

$$E = \begin{cases} k\sqrt{Djs - C} & \text{si } Djs \geq C \\ 0 & \text{si } Djs < C \end{cases} \quad (1.8)$$

Où C représente le retard entre le jour de la naissance de la glace et le premier jour de congélation de la température moyenne de l'air.

1.4 Contribution de l'étudiante à ce mémoire

Dans le cadre de son mémoire, l'étudiante a effectué de nouveaux développements dans l'application de l'approche des réseaux de neurones artificiels. Plus spécifiquement, l'étudiante a adapté l'approche des ensembles réseaux de neurones pour l'estimation de l'épaisseur de glace dans plusieurs lacs canadiens. L'apport de cette recherche est de prouver une amélioration significative pour l'estimation de l'épaisseur de glace des lacs en utilisant le modèle de l'ensemble de réseaux de neurones. La candidate a également comparé cette approche à celle du réseau de neurones simple et la formule empirique de Stefan révisée.

La candidate a été encadrée au sein de l'équipe de la Chaire en estimation des variables hydrométéorologiques à l'INRS-ÉTÉ. Cependant, elle a développé ses propres programmes et a démontré un bon niveau d'autonomie dans l'obtention des résultats.

1.5 Références

- Anderson, M.G. et Burt, T.P. (1985) Hydrological Forecasting. *Chichester: Wiley*. pp.149–60.
- Anderson, W.L., Robertson D.M. et Magnuson J.J. (1996) Evidence of recent warming and ENSO—variation in ice breakups of Wisconsin lakes, *Limnology and Oceanography* 41 (5), pp. 815–821.
- Beltaos, S. et Prowse, T.D. (2001) Climate impacts on extreme ice-jam events in Canadian rivers. *Hydrological Sciences Journal*. 46(1) pp. 157.
- Breiman, L. (1996a) Bagging predictors, *Machine Learning*, 24, pp. 123–140.
- Breiman, L. (1996b) Stacked regressions, *Machine Learning*, 24, pp. 49–64.
- Caine, N. (2002) Declining ice thickness on an alpine lake is generated by increased winter precipitation. *Climatic Change*, 54: pp. 463-470.
- Campolo, M., Andreussi, P. et Soldati, A. (1999) River flood forecasting with a neural network model. *Water Resources Research* 35, pp. 1191–97.
- Cybenko G. (1989). Approximation by Superpositions of a Sigmoidal Function, *Mathematics of Control, Signals and Systems*, Vol. 2, pp. 303-314.
- Dawson, C.W. et Wilby, R.L. (2001) Hydrological modelling using artificial neural networks. *Progress in Physical Geography*, 25(1), pp. 80-108.
- Duguay C.R., Flato G. M., Jeffries M.O., Ménard P., Morris K., et Rouse W. R. (2003) Ice-cover variability on shallow lakes at high latitudes: model simulations and observations. *Hydrological Processes*. 17, pp. 3465–3483
- Duguay, C. Kim M., Martin J. (2005) Model simulation of the effects of climate variability and change on lake ice in central Alaska, USA. *International Glaciological Society*. 40 (1) pp. 113-118
- Efron, B. et Tibshirani, R. (1993) *An introduction to the bootstrap*. Chapman and Hall.
- Freund, Y. et Schapire R. E. (1996) Experiments with a new boosting algorithm, in Proceedings of the Thirteenth International Conference on Machine Learning, Morgan Kaufmann, Burlington, Mass pp. 148– 156.
- Haykin S. (1999) Neural Networks-A Comprehensive Foundation, *Prentice-Hall*, New York.

- Hirst, S. M. (1984) Effects of spring breakup on microscale air temperatures in the Mackenzie Delta. *Arctic* 37, pp. 263-269.
- Lesack, L. F. W., Hecky, R. E. et Marsh, P. (1991) The influence of frequency and duration of flooding on the nutrient chemistry of Mackenzie Delta lakes. In: *Mackenzie Delta, Environmental Interaction and Implications of Development*, (ed. by P. Marsh & C. S. L. Ommanney) (NHRI Symp. no. 4), National Hydrology Research Institute, Environment Canada, Saskatoon, Saskatchewan, Canada. pp. 19-36
- Livingstone, D.M. (1997) Break-up dates of alpine lakes as proxy data for local and regional mean surface air temperatures, *Climatic Change*, 37, pp. 407-439.
- Livingstone, D.M. (1999) Ice break-up on southern Lake Baikal and its relationship to local and regional air temperatures in Siberia and the North Atlantic Oscillation, *Limnology and Oceanography*, 44(6), pp. 1486-1497.
- Lock, G. S. H. (1990) *The Growth and Decay of Ice*, Cambridge Univ. Press, New York
- Maier, H.R. et Dandy, G.C. (2000) Neural networks for the prediction and forecasting of water resources variables: *review of modelling issues and applications*. *Environmental Modelling & Software*, 15, pp. 101-124.
- Marsh, P. (1986) Modelling water levels for a lake in the Mackenzie Delta. In: *Cold Regions Hydrology Symp.*, American Water Resources Association, Fairbanks, Alaska, USA pp. 23-29.
- McCulloch, W.S. et Pitts W. (1943) *A Logical Calculus of the Ideas Immanent in Nervous Activity*. *Bulletin of Mathematical Biophysics* Vol. 5, pp. 115-133.
- Milburn, D. et Prowse, T. D. (2000) Observations on some physical-chemical characteristics of ice breakup. *Journal of Cold Regions* 4(14), 214-223
- Minsky, M.L. et Papert, S.A. (1969) *Perceptrons*. Cambridge, MA: MIT Press.
- Morse, B. and Hicks, F. (2005) Advances in river ice hydrology 1999-2003. *Hydrological Processes*, 19(1), pp. 247-263.
- Prowse, T. D. et Reedyk, S. (1993) Thermal and climatologic effects. In: *Environmental Aspects of River Ice* (ed. by T. D. Prowse & N. C. Gridiey),. NHRI Science Report no. 5, National Hydrology Research Institute, Environment Canada, Saskatoon, Saskatchewan, Canada. pp. 60-75
- Prowse, T. D. et Conly, M. (1998) Impacts of climatic variability and Dow regulation on ice-jam flooding of a northern delta. *Hydrologic Processes* 12(10-11), pp. 1589-1610.

- Rumelhart, D.E. et McClelland, J.L. (1986) editors, *Parallel distributed processing: explorations in the microstructures of cognition*. Vol. 1. Cambridge, MA: MIT Press.
- Scrimgeour, G. J., Prowse, T. D., Culp, J. M. et Chambers, P. A. (1994) Ecological effects of river ice breakup: a review and perspective. *Freshwater Biol.* 32, pp. 261-275.
- Schalkoff, R.J. (1997) *Artificial neural networks*. New York: McGraw-Hill.
- Seidou, O., Ouarda, T.B.M.J., Bilodeau, L., Hessami, M., St-Hilaire, A. et Bruneau, P. (2006) Modelling ice growth on Canadian lakes using artificial neural networks. *Water Resources Research*, 42(W11407, doi:10.1029/2005WR004622).
- Shu, C., et D. H. Burn (2004) Artificial neural network ensembles and their application in pooled flood frequency analysis, *Water Resources Research*, 40, W09301, doi:10.1029/2003WR002816
- Thirumalaiah, K. et Deo, M.C. (1998) Real-time flood forecasting using neural networks. *Computer-Aided Civil and Infrastructure Engineering* 13, pp. 101–110.
- van der Vinne, G., Prowse, T. D. et Andres, D. (1991) Economic impact of river ice jams in Canada. In: *Northern Hydrology, Selected Perspectives* (ed. by T. D. Prowse & C. S. 1. Ommanney) (NHRI Symp. no. 6), 3. National Hydrology Research Institute, Environment Canada, Saskatoon, Saskatchewan, Canada. pp. 33-352
- Watts, G. (1997) Hydrological modelling. In Wilby, R.L., editor, *Contemporary hydrology: towards holistic environmental science*, Chichester: Wiley, pp. 151–93.
- Wilkins, S. P. et Calkins, D. J. (1993) Winter hydrology. In: *Environmental Aspects of River Ice*. (ed. by T. D. Prowse & N. C. Gridley), NHRI Science Report no.5, National Hydrology Research Institute, Environment Canada, Saskatoon, Saskatchewan, Canada, pp. 23-31
- Wolpert, D. H. (1992) Stacked generalization, *Neural Networks*, 5, pp. 241– 259.
- Zurada J. (1992) Introduction to artificial neural systems, *West Publishing Co.*, St. Paul, MN.

NOTE : Étant donné que le présent travail est composé principalement d'un article et d'un rapport, afin de garder l'autonomie de chacune de ces deux parties, certains éléments sont répétés mais d'une manière adaptée au contexte de chacune des parties.

Estimation of ice thickness on lakes using artificial neural network ensembles

I. Zaier¹, C. Shu^{1*}, T.B.M.J. Ouarda¹, O. Seidou² and F. Chebana¹

¹ NSERC/Hydro-Quebec Statistical Hydrology Chair,
Canada Research Chair on the Estimation of Hydrometeorological Variables,
INRS-ETE, University of Quebec
490 de la Couronne
Quebec, Qc,
Canada, G1K 9A9
Tel: (418) 654-3842,
Fax: (418) 654-2600,

² University of Ottawa
550 Cumberland St., Ottawa, Ontario
Canada, K1N 6N5
Tel: (613) 562-5800 ext. x6143

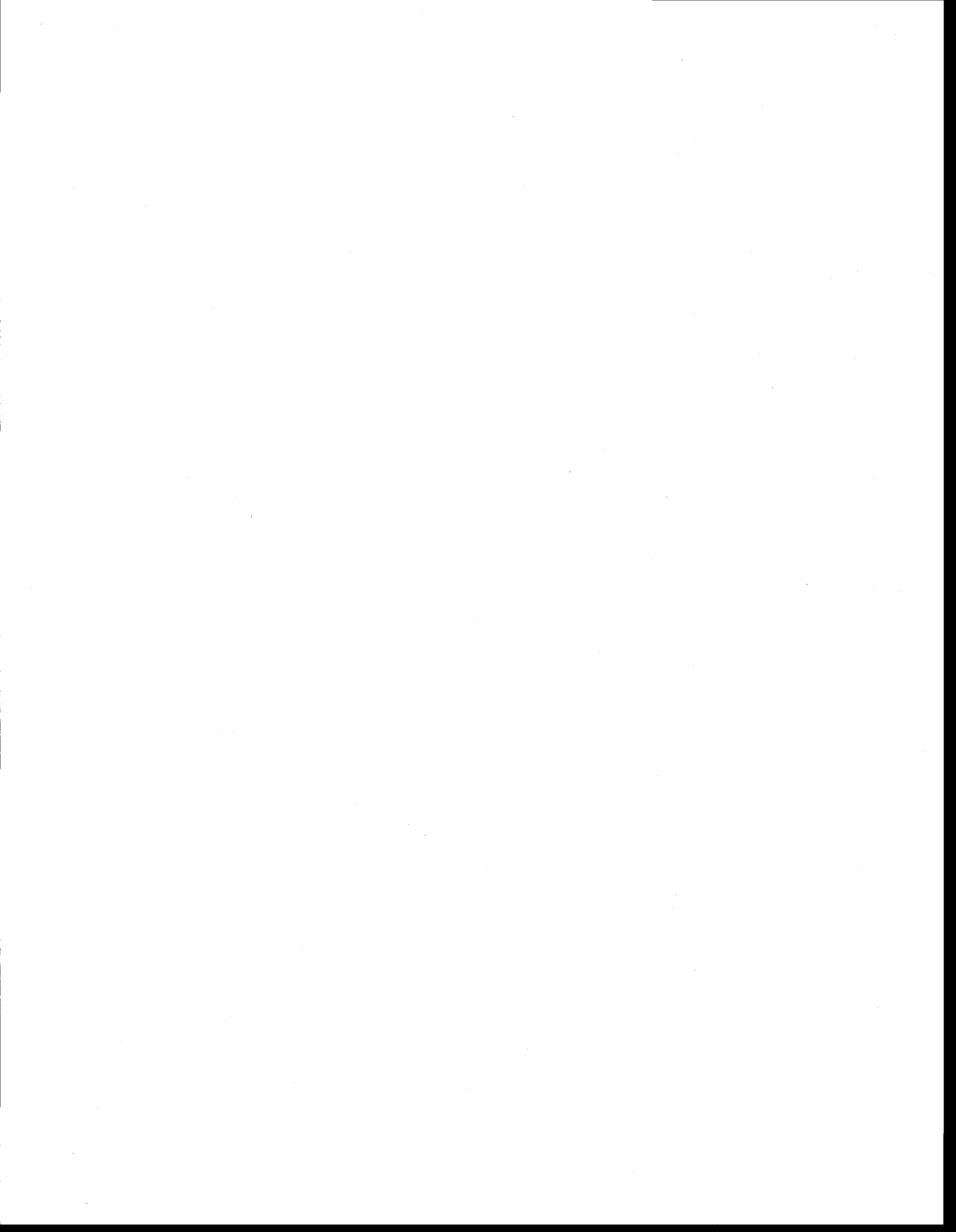
* Corresponding author
E-mail: Chang_shu@ete.inrs.ca

Submitted for publication in the Journal of Hydrology

June 2008

Abstract

Artificial Neural Network Ensemble (ANN Ensemble) prediction is a technique in which the outputs of a set of separately trained ANNs are combined to form one unified prediction. ANN Ensemble models are developed in this paper to improve the results of Single Artificial Neural Network (Single ANN) for the estimation of the ice thickness in a number of selected Canadian lakes during the early winter ice growth period. An effective ensemble consists of a set of ANNs that may not be highly performing when they are used separately, but have their prediction errors greatly reduced once combined. This paper evaluates the effectiveness of a number of ensemble techniques including randomization, bagging and boosting for creating members of an ensemble, then averaging and stacking techniques for combining ensemble members. The experiments show that, in the context of estimation of lake ice thickness, boosting is much better than randomization, and sometimes better than bagging. Stacking was found to be more competitive than averaging. Overall, ANN Ensemble models for the estimation of ice thickness proved to be more accurate than Single ANN models, especially when boosting is used for combining ensemble members and when stacking is used to combine the outputs from individual members. ANN ensembles achieve the best generalization performance when the ensemble size is increased to around 20.



2. Estimation de la glace des lacs en utilisant l'ensemble de réseaux de neurones artificiels

Résumé

L'approche des prédictions en utilisant l'ensemble des réseaux de neurones artificiels est une technique dont les sorties des réseaux de neurones artificiels entraînés séparément sont combinées, dans le but de former une prédiction unique. L'ensemble des réseaux de neurones artificiels est développé dans cet article, afin d'améliorer les résultats de l'approche de réseaux de neurones artificiels simples pour l'estimation de l'épaisseur de la glace dans des nombreux lacs canadiens durant l'hiver précoce pour la période de la croissance de la glace. Un ensemble efficace se compose de plusieurs réseaux de neurones artificiels qui pourraient ne pas être très performants quand ils sont entraînés séparément, mais, une fois combinés, leur erreur de prédiction est grandement réduite. Cet article évalue l'efficacité de plusieurs techniques incluant l'approche aléatoire, le bagging et le boosting qui contribuent à la création des membres de l'ensemble et les techniques de la moyenne et de l'empilage qui permettent la combinaison de ces membres. Les expériences montrent que, dans le contexte de l'estimation de l'épaisseur de glace des lacs, le boosting est plus performant que l'approche aléatoire et parfois meilleure que le bagging. L'empilage est plus compétitif que la moyenne. En tout, les modèles de l'ensemble de réseaux de neurones artificiels pour l'estimation de l'épaisseur de glace sont plus efficaces que les modèles de réseaux de neurones simples, en particulier, quand le boosting est utilisé pour combiner les membres de l'ensemble et l'empilage pour la combinaison des sorties des membres individuels. L'ensemble des

réseaux de neurones artificiels accomplie une meilleure performance quand la taille de l'ensemble atteint les 20 membres.

2.1 Introduction

Ice is present in nearly every Canadian lake for a period that ranges from days to several months every year. Extreme events resulting from ice-jamming are the major causes of much economic damage to properties and infrastructures. Models of ice thickness in lakes provide useful information to deal with these problems, and serves our general aim to better understand lake ice processes. For instance, lake ice thickness, date of ice break-up and other ice characteristics are useful indices of climate change which can be modeled and forecasted.

The evolution of ice thickness in lakes is influenced by many interrelated processes. However, the site-specific nature of these complex processes makes ice thickness difficult to predict using physically-based models. The main drawback of the numerical physically-based models is that they require many physical parameters that are hard to collect. Most numerical ice growth models adopt versions of energy budget with different complexity such as the Canadian Lake Ice Model CLIMO used by *Ménard et al.* [2002a, 2002b], which is a modified version of a one-dimensional sea ice model [*Flato and Brown, 1996*] and has been described in detail by *Duguay and al.* [2003]. Some models are applied to a specific aspect of ice development such as, ice cover initiation [*Schulyakovskii, 1966*], border ice formation [*Matousek, 1984; Svensson et al., 1989*], frazil ice formation [*Omstedt, 1985a, 1985b; Svensson and Omstedt, 1994*] and ice cover growth [e.g., *Schulyakovskii, 1966; Lock, 1990*]. Other models are more complete and may simulate ice formation, transport, growth and decay [*Shen and Chiang, 1984; Shen and Ho, 1986; Shen et al., 1990, 1995*]. As a result, the development of more analytical models is required. A previous study by *Seidou et al.* [2006] has shown that artificial

neural networks (ANN) can be a valuable alternative to complex thermodynamic lake ice growth models, especially when data are not available in sufficient quality and quantity.

ANN models are generally considered as 'black box' models that are able to capture underlying relationships when presented with input and output data. They have been successfully used in hydrology for solving various problems, such as data classification [e.g., *Liang and Hsu*, 1994], river discharge prediction [e.g., *Shamseldin*, 1997], regional flood frequency analysis [*Shu and Burn*, 2004; *Shu and Ouarda*, 2007, 2008], water quality evaluation and forecasting [*Zhang et al.*, 1994], estimating river streamflow affected by ice conditions [*Chokmani et al.*, 2008], rainfall estimation [e.g., *Xiao and Chandrasekar*, 1997] and stream flow under ice estimation [*Ouarda et al.*, 2003; *Chokmani et al.*, 2008]. The suitability of ANNs for modelling complex systems has resulted in an increase in the popularity of ANN models and their use in an ever increasing number of applications.

Recent studies show that a new approach called ANN Ensemble which utilizes multiple ANNs can improve the generalization ability of a single ANN. In an ANN ensemble, a number of ANNs trained for the same purpose as a single ANN are combined to generate a unique output [*Shu and Burn*, 2004]. ANN Ensemble approaches have been used successfully in several domains, such as time series modeling, chemistry, robotics, automatic control and medical diagnosis. In the area of forecasting, it has been shown that better results can be achieved by combining forecasts than by choosing the best one [*Bates and Granger*, 1969]. For details concerning the theoretical studies of the ensemble approaches, the reader is referred to the works by *Krogh and Vedelsby* [1995] and *Hansen and Salamon* [1990]. *Cannon and Whitfield* [2002] and *Shu and Burn* [2004]

provide a general overview of popular ensemble methods. In this paper, six ANN ensemble models are implemented by using a combination of five ensemble modeling techniques in order to model the ice thickness in a number of selected Canadian lakes, and the results are compared with those of the single ANN models.

The remainder of this paper is composed of six parts: a general introduction of the single ANN model for ice growth modelling (Section 2), an overview of general artificial neural network ensemble approaches (Section 3), a description of the data used in this study (Section 4), a description of the methodology adopted for this study, including single and ensemble ANN models for ice growth estimation, performance criteria and evaluation procedure (Section 5), results and discussion (Section 6), and finally, conclusions and future work (Section 7).

2. Single Artificial Neural Network

This section presents an overview of the architecture and characteristics of the single ANN proposed for ice growth modelling. The type of ANN selected in this paper is a multilayer perceptron (MLP) feed-forward network which maps sets of input data onto a set of appropriate outputs. MLP is the most popular ANN architecture in use today. Reviews of the ANN from a statistical perspective have been given by a number of authors [e.g. *Ripley*, 1993; *Cheng and Titterton*, 1994; *White*, 1994]. The most widely used training algorithm for a MLP is the error back-propagation algorithm. This popular algorithm was described firstly by *Werbos* [1974]. However, it was only in 1986 that it was introduced and popularized by *Rumelhart and McClelland* [1986]. The back-propagation algorithm requires that the transfer function used by the artificial neurons be

differentiable. It works by iteratively changing the network's interconnecting weights such that the overall error between observed values and network outputs is reduced. Network geometry is generally defined by the number of hidden layer nodes and the number of nodes in each of these layers. It determines the number of model parameters that need to be estimated. The single ANN used in this study is a one-hidden-layer MLP with sigmoid neurons in the hidden layer and a linear neuron in the output layer. It has been shown that ANNs with one hidden layer can approximate any continuous function, given sufficient degrees of freedom [Funahashi, 1989 and Hornik *et al.*, 1989]. The optimum number of neurons in the hidden layer was identified using a trial and error approach.

3. General Ensemble Approaches

Recent studies have shown that the robustness and reliability of an ANN can be significantly improved by appropriately combining several ANN models into an ANN ensemble [Jacobs *et al.*, 1991; Wolpert, 1992; Perrone and Cooper, 1993; Jordan and Jacobs, 1994; Sridhar *et al.*, 1996; Zhang *et al.*, 1997]. The construction of an ANN ensemble requires two major steps. The first step is to create individual ensemble members and the second step is to find the appropriate combination of outputs from the ensemble members to produce the unique ensemble output [Sharkey, 1999]. Various methods have been developed for creating ensembles, such as bagging and boosting. For general information and the comparison of these methods, the reader is referred to Opitz and Maclin [1999], Sharkey [1999] and Shu and Burn [2004].

3.1. Approaches for generating individual ensemble members

The main objective of combining ANNs in an ensemble is to improve the generalization ability over the single ANN. If the networks in an ensemble share the same characteristics, the ensemble will show a similar performance to the component single ANNs. Thus it is crucial to create individual networks with diverse characteristics while maintaining their individual generalization ability. Various approaches have been proposed to generate ensemble members, and the following four approaches are the most commonly used [*Sharkey, 1999*]:

- (1) Creating a set of networks by varying the initial random weights, while keeping the training data unchanged.
- (2) Creating a set of networks by varying the architecture and the number of hidden units, while keeping the training data unchanged.
- (3) Altering the training algorithm, while keeping the training data unchanged.
- (4) Altering the training data set.

Among these approaches, the first approach is also known as randomization [*Shu and Burn, 2004*]. It is the easiest way to construct different ensemble members. However, the more sophisticated approaches of altering the training data set using resampling techniques, such as bagging [*Breiman, 1996a*] and boosting [*Schapire, 1990; Freund and Schapire, 1996*], have gained most attention by the researchers. Bagging and boosting are two general techniques for building predictors based on samples from a dataset with the intention that each component network contains different training sets. The empirical study by *Dietterich* [2000] shows that boosting often provides better results than bagging and randomization, while bagging and randomization perform equally well.

3.1.1. Boosting

Boosting is a general method which attempts to boost the performance of a given learning algorithm [Schapire, 1990; Freund and Schapire, 1996]. The main idea of boosting is to produce a sequence of ANNs so that each subsequent ANN concentrates more on the training cases which are not well predicted by the previous one. This is achieved by assigning a probability label to each training case of the database and maintaining it over the whole training phase. Probability labels are updated according to certain rules which are fully described in the Adaboost.R2 algorithm below. Generally, the worse performance of the previous built ANN on a particular training case is, the higher the probability label to be assigned. Training cases that are persistently incorrectly estimated become increasingly likely to be included in new training samples. In this paper, the boosting algorithm ADABOOST.R2 proposed by Drucker [1999] is adopted. ADABOOST.R2 is a variation of the adaptive boosting algorithm ADABOOST.R proposed by Freund and Schapire [1996]. Drucker [1999] showed that, in most cases, the ADABOOST.R2 algorithm provided better estimation than bagging. The ADABOOST.R2 algorithm [Drucker,1999] is described in detail below.

Assume that the training dataset T consists of N instances $[x_1, y_1], \dots, [x_N, y_N]$, where x and y are input and output variables respectively, the probability that the i th training case in T being sampled at step s is $D_s(i)$.

1. At the first step, each element of the initial data set has the same chance to be included in the training sets of the first predictor. Thus, for $s=1$,

$$D_1(i) = 1/m, \text{ over all } i. \quad (1)$$

2. Generate the new training set T_s from the initial training dataset T based on the distribution D_s .

3. Generate new network k_s , and set new T_s as its training set.

4. Compute the maximum loss, L_{max} , between the actual value and the network output $k_s[x_i, y]$, over the initial training set T where:

$$L_{max} = \sup(|k_s(x_i, y) - y_i|), \text{over all } i. \quad (2)$$

5. Compute the individual L_i , loss for each example in the training set:

$$L_i = 1 - \exp\left[-\frac{|k_s(x_i, y) - y_i|}{L_{max}}\right] \quad (3)$$

6. Calculate the weight average loss \bar{L} :

$$\bar{L} = \sum_{i=1}^N L_i D_s(i) \quad (4)$$

7. Set the value of B_s

$$B_s = \frac{\bar{L}}{1 - \bar{L}} \quad (5)$$

8. Update the distribution D_s :

$$D_{s+1}(i) = \frac{D_s(i) B_s^{(1-L_i)}}{Z_s} \quad (6)$$

where Z_s is a normalization factor chosen such that D_{s+1} is a distribution.

9. Set $s = s + 1$

10. Repeat steps 2-9 until the average loss L is less than 0.5.

3.1.2. Bagging

Bagging (bootstrap aggregation) is an algorithm originally proposed by *Breiman* [1996a]. It is a technique that tries to improve a learning algorithm's performance by using bootstrap replicates of the training set [*Efron and Tibshirani*, 1993; *Efron*, 1979]. Under bagging, multiple realizations of the original training dataset are generated and they are subsequently used to train different ANN models. The outputs from each of the ANN models are combined together to give a unique output.

In bagging, each training set is constructed by forming a bootstrap replicate of the original training set. In other words, suppose the training dataset T consists of N instances $[x_1, y_1], \dots, [x_N, y_N]$, where x and y are input and output variables, respectively. To generate a bootstrap member T_B from T , each instance in T is assigned a probability of $1/N$ and generated by sampling with replacement N times from the original dataset T using the above probabilities. Hence each bootstrap dataset T_B may have many instances in T duplicated several times, while other instances may not be selected at all. The process is repeated until a desirable number of T_B is reached. Individual ANN models can then be trained on the multiple T_B generated by this process.

The main difference between bagging and boosting algorithms is as follows: in the boosting algorithm, the distribution of the training set changes adaptively based on the performance of the previously created network, while the bagging algorithm changes the distribution of the training set stochastically. Although the boosting algorithm has better generalization ability than the bagging algorithm in a number of applications [*Drucker*, 1999; *Shu and Burn*, 2004], the latter algorithm has the advantage of training the member networks in an ensemble independently, hence in parallel.

3.2. Approaches for combining the component member networks

As mentioned above, there are two steps to construct an ANN ensemble. Once a set of ANNs has been created, the second step is to combine the different outputs of the member networks constructing the ensemble. The two most commonly used methods are averaging and stacking. *Ahmad and Zhang* [2002] and *Shu and Burn* [2004] provided review and comparison of the two methods.

3.2.1. Averaging

By using the averaging method, the output of the ensemble is obtained by computing the mean of the output of the member networks. Suppose that N is the number of individual ANN members in an ensemble, the combination function f is:

$$\hat{x}_i = f(\hat{x}_i^k) \quad i=1\dots m \quad (7)$$

where \hat{x}_i is a predicted value of the instance i obtained from the k^{th} network and the form of the function f is:

$$f(\hat{x}_i^k) = \frac{1}{N} \sum_{k=1}^N \hat{x}_i^k \quad (8)$$

The implementation of the averaging approach is easy, and it has been shown to be an effective approach to improve the performance of the single ANN model [*Perrone and Cooper*, 1993; *Bishop*, 1995].

3.2.2. Stacking

Stacking or stacked generalization is a general method of using the combination of a higher-level model and the lower-level models in order to achieve a greater

predictive accuracy. Stacking can be viewed as means of collectively using several estimators to estimate their own generalizing biases with respect to a particular learning set, and then filter out those biases [Wolpert, 1992]. The method was shown to be able to improve the accuracy of a model when only a limited number of experimental data points in the training data set are available [Wolpert, 1992]. Stacking tries to learn which models are reliable ones, and use a higher-level learning algorithm (Level 1 generalizer) to discover the best way to combine the outputs of the base models (Level 0 generalizer) (Witten and Frank, 1999). When using the stacking to predict a new instance, the instance is first fed into the level 0 predictors. Each of these predictors calculates a correspondent value. These values are fed into the level 1 predictor which combines them and computes the final prediction. The inputs to the level 1 predictor are the outcomes of the level 0 predictors. This method has been used by *English* [1996], *Drucker* [1997], *Ting and Witten* [1999] and *Hu and Tsoukalas* [2003] in order to improve the generalization capability of ANNs. *Breiman* [1996b] suggests minimizing the function G :

$$G = \sum_{i=1}^m \left[x_i - \sum_{k=1}^N c_N \hat{x}_i^k \right]^2 \quad c_N > 0. \quad (9)$$

The coefficients $\hat{c}_1, \hat{c}_2, \dots, \hat{c}_K$ in Equation (9) are estimated in order to construct the final output of the ensemble:

$$\hat{x}_i = \sum_{k=1}^N c_k \hat{x}_i^k \quad i = 1 \dots m. \quad (10)$$

The ensemble model performance could be adversely affected by minimizing the function G in Equation (9) due to scale effect, thus a better option suggested by Shu and Burn [2004] to minimize the following function in the level 1 generalizer is adopted in this paper:

$$H = \sum_{i=1}^m \left[\frac{x_i - \sum_{k=1}^N c_k \hat{x}_i^k}{x_i} \right]^2 \quad c_N > 0. \quad (11)$$

4. Data

The ANN Ensemble models proposed in this paper are applied to 17 lakes located in Canada. The selection of these lakes is mostly based on the availability and quality of the data. The measuring stations for the lakes are listed in Table 1, and the geographic locations of these stations are shown in Figure 1. Three types of data, including ice thickness data from the Canadian Ice Service [2005], daily meteorological data from Environment Canada, and incident solar radiation at the top of the atmosphere using the formulas presented by Solar Radiation Monitoring Laboratory (2004), are used in this study [Seidou *et al.*, 2006]. According to the data availability and the relative importance of the data in interpreting the heat budget involved in ice growth, a number of candidate variables are retained for the prediction of ice thickness [Seidou *et al.*, 2006]. These variables are daily snow depth on the ground, daily rainfall measurements, daily mean air temperature and daily total solar radiation. Details on formulating the ANN inputs for ice thickness prediction based on these variables are discussed in detail in Section 5.1. For further information regarding the study area, readers are referred to Seidou *et al.* [2006].

Insert table 1

Insert figure 1

5. Methodology

5.1 Single ANN Model for Ice Growth Modelling

The major areas that should be considered in order to build a successful ANN model include the choice of an appropriate network geometry, the selection of appropriate model inputs and outputs, data transformation, training algorithm, and overfitting control. Details regarding these areas of the Single ANN model for ice growth modelling are discussed in the remaining parts of this section.

The Single ANN model used in this study is a one hidden layer neural network with sigmoid neurons in the hidden layer and a linear neuron in the output layer. This kind of structure is shown to be able to approximate any bounded continuous function, and one hidden layer is sufficient for an ANN to approximate any complex nonlinear function [Cybenko, 1989; Hornik et al., 1989]. In this paper, the number of neurons in the hidden layer of the single ANN model for a given lake adopts the optimal number given by Seidou et al., [2006]. These numbers of optimal hidden neurons are listed in column six of Table 1.

Choosing appropriate inputs is crucial for the ANN models. In this paper, the inputs selected should be meaningful for the ice growth process and available at all ice measurement stations. The candidate variables used to construct the input variables for this study are described in Section 4. These variables can be used directly or linearly combined to form the ANN input variables. The type of input variables considered in this paper include [Seidou et al., 2006] (1) the sum of degree-days below the freezing point since the onset of the ice cover in a given year, (2) the sum of solar radiation during the period of ice growth for days with precipitation ($W \text{ day/m}^2$) divided by the sum of degree-days Rad_c , (3) the sum of solar radiation during the period of ice growth for days without precipitation ($W \text{ day/m}^2$) divided by the sum of degree-days Rad_{nc} , (4) the

average daily rainfall (over time) during the ice growth period P (mm), (5) the average on-ground snow depth (over time) during the ice growth period S (cm), and (6) five sets of combinations of Rad_{nc} and Rad_c including Rad_{nc} , $Rad_{nc}+0.25 Rad_c$, $Rad_{nc}+0.5 Rad_c$, $Rad_{nc}+0.75 Rad_c$ and $Rad_{nc}+ Rad_c$.

The best combination of input variables is different from site to site. *Seidou et al.*, [2006] used a leave-one-out cross-validation procedure to find the best combination of input variables for every station. The results for the stations considered in this work are shown in the last column of Table 1. These results are adopted for the single and ensemble ANN models in this paper.

Seidou et al., [2006] suggested using I^2/ D_d , where I is the ice thickness and D_d is the sum of degree-days below the freezing point since the onset of the ice cover in any given year, as the output of the ANN instead of directly using I . The main advantage of using I^2/ D_d as the output is that it contributes in the reduction of the weight of the degree-days in the input variable [*Seidou et al.*, 2006]. This choice of output also gives a better physical comprehension of the influence of the other variable in ice thickness estimation [*Seidou et al.*, 2006]. In this paper, I^2/ D_d is selected as the ANN output due to the above reasons.

Data pre-processing are also essential for the effectiveness of ANN training algorithms. Without appropriate preprocessing, input variables measured on different scales will dominate training to a greater or lesser extent. In this paper, all the input variables are standardized to guarantee they receive equal interest during the training process (*Maier and Dandy*, 2000). The standardization procedure transforms the inputs so that they have zero mean and unity standard deviation.

The training algorithm used in this study is the Levenberg-Marquardt (LM) algorithm. For function approximation problems, the LM algorithm is the fastest method for training moderate size networks which have less than a few hundred weights [Demuth *et al.*, 2008]. This advantage is especially noticeable if very accurate training is required. In many cases, the LM algorithm is able to obtain lower mean square errors than any of the other training algorithms tested [Demuth *et al.*, 2008]. The maximum number of training epochs is set to 500.

An ANN may have poor generalization ability due to the overfitting problem, which means it can be fitted precisely to the set of training data while losing its generalization ability on the unseen data. In order to improve the generalization capabilities of the ANN models, there are two widely used methods known as early stopping and Bayesian regularization [Bishop, 1995]. The former method requires a validation set besides the calibration and test sets. In the case of this study where the number of data is small, this approach can be problematic. The latter technique, which is used here, encourages smoother network mappings by favouring small values for the ANNs parameters. Indeed, it has been shown by Bishop [1995] that small values for the weights decrease the tendency of the model to overfit. The training function adopted in this paper, which is called *trainbr* function in the Matlab environment, uses Bayesian regularisation [Mackay, 1992, 1995] to improve the generalization capabilities of an ANN. This approach minimizes the over-fitting problem by taking into account the goodness-of-fit as well as the network architecture. For more details about the use of Bayesian regularisation in combination with Levenberg-Marquardt training, the reader is referred to *Foresee and Hagan* [1997].

5.2 ANN Ensemble Models for Ice Growth Modelling

In this paper, the three methods for creating the members of an ANN Ensemble including boosting, bagging and randomization and the two methods for combining the outputs of the member networks including stacking and averaging are evaluated. Six ANN ensemble models can be obtained by the mixture of these ensemble techniques. Details regarding the six models are as follows:

1. Boost_Median ANN ensemble model. In this model, boosting is used for the creation of the member ANNs. *Drucker* [1997] suggests that the averaging approach is not suitable for combining outputs of ensemble members generated using the ADABOOST.R2 algorithm. Instead, the ensemble output can be computed using the following equation which is essentially the weighted median of the outputs of the member networks [*Drucker*, 1997]:

$$o_{final} = \inf \left[z \in Z : \sum_{s: o_s \leq z} \log \left(\frac{1}{B_s} \right) \geq \frac{1}{2} \sum_s \log \left(\frac{1}{B_s} \right) \right] \quad (12)$$

For a particular input i , each prediction $o_s(y_i, z)$ has an associated B_s , which is computed by equation (5). The predictions are relabelled according to the following order: $o_1 < o_2 < \dots < o_m$ while retaining the correspondence between B_s and o_s . Then, the B_s are summed until reaching the smallest z such that the inequality in Equation (12) is achieved.

2. Boost_Stack ANN ensemble model. In this model, the boosting algorithm is used to create the members of the ensemble and the stacking is used to combine the outputs of these members.

3. Bag_Mean ANN ensemble model. In this model, bagging is used to generate members of an ensemble and averaging is used to combine the outputs of those member networks.
4. Bag_Stack ANN ensemble model. In this model, bagging is used to generate members of an ensemble and stacking is used to combine the outputs of those member networks.
5. Random_Mean ANN ensemble model. In this model, member networks of an ensemble are generated by using the randomization approach, and the ensemble output is obtained by averaging the outputs of those member networks.
6. Random_Stack ANN ensemble model. In this model, member networks of an ensemble are generated by using the randomization approach, and the ensemble output is obtained by stacked generalization.

5.3. Performance Criteria

The performance of the proposed models is evaluated using five performance criteria: root-mean-square error [RMSE], relative root-mean-square error [RRMSE], model explained variance [r^2], Nash criterion [NASH], and bias [BIAS]. The criteria are defined as follows:

$$\text{RMSE} = \left(\frac{1}{n} \sum_{k=1}^n (I^k - \hat{I}^k)^2 \right)^{\frac{1}{2}} \quad (13)$$

$$\text{RRMSE} = \left(\frac{1}{n} \sum_{k=1}^n \left(\frac{I^k - \hat{I}^k}{\bar{I}} \right)^2 \right)^{\frac{1}{2}} \quad (14)$$

$$\text{NASH} = 1 - \frac{\sum_{i=1}^n (I_i - \hat{I}_i)^2}{\sum_{i=1}^n (I_i - \bar{I}_i)^2} \quad (15)$$

$$r^2 = \frac{\text{cov}\left([I^1, \dots, I^n], [\hat{I}^1, \dots, \hat{I}^n]\right)^2}{\text{var}\left([I^1, \dots, I^n]\right)\text{var}\left([\hat{I}^1, \dots, \hat{I}^n]\right)} \quad (16)$$

$$\text{BIAS} = \frac{1}{n} \sum_{k=1}^n (I^k - \hat{I}^k) \quad (17)$$

where n is the number of catchments in the prediction set, and $I^k, i = 1, \dots, n$ and $\hat{I}^k, i = 1, \dots, n$ are, respectively, the observed and simulated ice thicknesses. The coefficient of determination and the criterion of NASH are statistical measure of how well the regression line approximates the real data points. Values of r^2 and NASH close to 1.0 indicate that the regression line perfectly fits the data.

5.4. Evaluation procedure

A split sample validation procedure is used in this paper to evaluate the performance of the proposed ice thickness prediction models. The data are randomly divided into two sets, the calibration set and the test set. The calibration set which contains 80% of the data is used to train the ANN based models, and the validation set which contains 20% of the data is used to assess the model performance.

6. Results and Discussion

In section 5.2, six ANN ensemble approaches are developed for ice thickness prediction. In this paper, the ANN Ensembles are compared with the Single ANN model

developed by *Seidou et al.*, [2006]. A detailed comparison of the various methods for creating and combining the ensemble members is also carried out in this paper. Member networks of the ANN ensembles and the Single ANNs use the same network design and training characteristics as described in Section 5.1. Programs are written using the Matlab programming language. The base ANN models are created using the Matlab Neural Network Toolbox [*Demuth et al.*, 2008].

The coefficient of determination and the NASH are computed for all the approaches of the ANN Ensemble to assess their goodness of fit. Table 2 lists the mean of r^2 and NASH of the six ensemble methods at each station. It shows that r^2 varies between 0.8 and 0.96 for all stations excluding the station YPY whose r^2 is equal to 0.75. The performance index of NASH is in the range of 0.8 to 0.96 excluding the YPY and YGK stations which have NASH values of 0.74 and 0.79 respectively. Overall, these results indicate that the ANN ensembles provide satisfactory estimation for the ice thickness at most stations.

Insert table 2

The following parts of this section begin with the analysis of the effect of ensemble size on the performance of ANN ensembles. Then, the Single ANN model is compared to the ANN Ensemble models. Next, the three methods for creating ensemble members including randomization, bagging, and boosting are compared. Finally, the two methods for combining the ensemble members including averaging and stacking are compared based on experiment results.

6.1 Ensemble size

Hansen and Salamon [1990] suggested that ensembles with 10 members are adequate to reduce model classification error, a result confirmed by *Agrafiotis et al.* [2002]. An empirical study by *Opitz and Maclin* [1999] showed that for both bagging and boosting, significant reduction in error occurred when using 10 to 15 members. *Tangang et al.* [1996] used neural network ensembles with 20 members to seasonally forecast the tropical Pacific sea surface temperature anomalies of ENSO (El Niño Southern Oscillation). Since most of these studies are based on classification problems, a similar experiment is conducted in the present paper to examine the effect of ensemble size on the performance of the ensemble models.

The performance of the six ensemble methods are evaluated at a typical station WFN (Cree Lake) by sequentially increasing the number of members from 2 to 20. The RMSEs of the six approaches at each ensemble size are plotted in Figure 2. From Figure 2, we can observe that the RMSE decreases slowly when the ensemble size increases from 2 to 6. The RMSE of most ensemble approaches decreases quickly when ensemble size increases from 6 to 10. After 10 members are included in an ensemble, the decrease in RMSE slows down for most approaches. All the ensemble approaches reach the best performance when the ensemble size increases to 20. The ANN ensembles are generally stabilized after 20 networks, and no significant changes are observed with further increase of the ensemble size. For the Random_Mean approach, with the increase of the ensemble size, the magnitude and speed of the decrease in RMSE are not as important for the Random_Mean method as for other ensemble approaches. The experiment is also conducted at other stations, and similar patterns are observed.

Insert figure 2

From the analysis above, we can conclude that at least 10 networks are required in an ANN ensemble to gain significant reduction in estimation error, while best generalization ability of the ANN ensembles can be achieved with an ensemble size around 20.

6.2 Single ANN versus ANN Ensemble

A rank score technique is used to rank the different ANN based models. This technique was used by several researchers, such as *Pandey and Nguyen* [1999], *Grover et al.* [2002] and *Shu and Burn* [2004], to compare the performance of multiple models. Suppose there are V performance indices and W models to be assessed, the models are ranked according to each performance index. A score of 1 is given to the best performing model and W to the worst performing model. The range of the overall rank score S_o of a model is $[V, WV]$. For convenience, the overall rank scores S_o are converted into standardized rank scores S_s which are in the range of $[0, 1]$ according to Equation (18):

$$S_s = \frac{WV - S_o}{WV - V} \quad (18)$$

As a consequence, a model with good performance has a rank score close to 1, while a model with poor performance has a rank score close to 0. The standardized rank scores S_s are used in the discussion for the rest of the paper.

The performance of the ANN ensemble models with 10, 15 and 20 member networks for ice thickness modelling are summarized in tables 3, 4 and 5 respectively. The performance of the single ANN models is also presented in these tables. For each evaluated approach, the performance indices RMSE, RRMSE and BIAS in these tables are computed as the mean of the RMSE, RRMSE and BIAS for all the studied stations.

The better generalization ability of an ANN based model is indicated by a higher value in the rank score. From these tables, it is observed that the rank score of the Single ANN is equal to 0.22 which is close to that of the Random_Mean method. However, the remainder of the rest ANN ensemble models have much higher rank scores varying between 0.39 and 0.88. This indicates that, for ice thickness estimation, the average generalization ability of the ANN ensembles is always better than that of individual predictors regardless of the methods or the number of members used in the construction of the ensemble. Compared to the Single ANN model, more accurate ice thickness prediction can be obtained by using ANNs in an appropriately constructed ensemble.

Insert table 3

Insert table 4

Insert table 5

As indicated in Section 6.1, the ANN Ensemble shows the best performance when 20 members are included. Figures 3, 4, and 5 compare the performance of the six ANN ensemble models (each has 20 member networks) with the single ANN in the three performance indices RMSE, BIAS and RRMSE respectively at each station. From Figure 3, we can observe that the RMSE of these methods varies between 1 and 5 cm, while the RMSE of the Single ANNs varies between 5 and 18 cm. Figure 4 shows that the BIAS of the ANN ensemble models is very close to zero. However, the BIAS varies between -3.7 and 4 cm for the single ANN model. From Figure 5, it is observed that the stacking based approaches outperform the Single ANN approach at all stations, while mixed results are observed with other approaches. There is also one outlier, station HA1 in Figure 5, for

which the performance index RRMSE reaches 60% when the Boost_Median approach is used. Overall, compared to the Single ANN approach, most ANN Ensemble approaches can perform significantly better in the two performance indices RMSE and BIAS, and stacking based ensemble approaches can improve the RRMSE significantly.

Insert figure 3

Insert figure 4

Insert figure 5

6.3 Randomization versus bagging and boosting

According to tables 3, 4 and 5, ANN ensemble models using randomization to generate ensemble members have the lowest rank scores (less than 0.5) among the six ensemble approaches. Thus, randomization is not the best method to create the members of an ANN ensemble for ice thickness estimation. In addition, the Random_Mean approach shows a similar performance to the Single ANN model demonstrated by the similar rank scores in Tables 3, 4 and 5. Excluding the Bag_Mean method with 20 members which has a rank score of 0.38, all the ANN ensemble models using bagging to generate ensemble members have rank scores varying between 0.57 and 0.66. Thus we can conclude that bagging consistently outperforms randomization as evidenced by the rank scores for ice thickness estimation.

From the rank scores in tables 3, 4 and 5, we can observe that Boost_Median has better performance than Random_Mean and Bag_Mean. Boost_Stack also has higher rank scores compared with Bag_Stack and Random_Stack. Thus, Boosting outperforms randomization and bagging when the same method is used for combining the ensemble

members. However, unlike bagging and randomization methods, RMSE, RRMSE and BIAS of the boosting based approaches do not necessarily decrease with the increase of the number of members in an ensemble. For instance, at stations YGK, YGM and YYR, the boosting based approaches show very similar performances in the three performance criteria regardless of the number of members in the ensembles. This phenomenon could be due to the stopping criteria which is the loss for each example in the training set (the parameter L in equation (3)) used in the boosting algorithm. In several cases, boosting can actually degrade the performance of the ANN Ensembles. For example, at station HA1 in figure 5, the RRMSE of Boost_Median is much larger than that of the Single ANN model. This could be due to the major drawback of the boosting algorithm which can be adversely impacted by the outliers. Freund and Shapire [1996] suggested that the sequentially generated training sets by the boosting algorithm may be over-emphasizing examples that are noise (thus creating extremely poor predictors).

6.4 Stacking versus simple averaging

Results in tables 3, 4 and 5 demonstrate that, when the same method is used for creating the members of an ensemble, ANN ensemble approaches using stacking to combine the member networks generally have higher rank scores than those using averaging. Figure 3 shows that when randomization is used to generate the members of an ensemble, at the majority of stations, stacking shows better performance in RMSE compared with averaging. From Figure 4, we can observe that all the ANN ensemble approaches have similar performances in the BIAS criterion. Figure 5 shows that

significant improvement in RRMSE can be achieved using stacked generalization at all stations, when bagging and boosting are used to create the members of an ensemble. In the case of using randomization for creating ensemble members, for the majority of the stations, stacking shows better performance in the RRMSE criterion than averaging. In conclusion, using stacking for the combination of the outputs of the ANN ensembles for ice thickness estimation is more effective than averaging when the same method for creating ensemble members is used.

7. Conclusions and future work

The main purpose of this paper is to apply ANN ensemble models to improve ice thickness estimation on lakes. Different methods for creating and combining ensemble members are evaluated, and the results of the ANN ensemble models are compared to those of the Single ANN models.

In comparison with the single ANN approach, significant improvement in the generalization ability can be achieved by applying the ANN Ensemble approaches, especially when those ensemble approaches using stacking to combine member networks. It is shown that boosting is the method that best reduces the estimation error in most cases. Using stacking to combine the outputs of the ANN ensembles for ice thickness estimation proved to be more efficient than averaging when the same method for creating ensemble members is used. At least 10 networks in an ensemble are required to gain significant reduction in estimation error, while the best generalization performance of ANN Ensembles can be achieved with an ensemble size of about 20.

The ANN models developed in this paper mainly use the meteorological variables as inputs. Geomorphologic data were not available to use in these models despite their known relevance to ice growth processes. Ice thickness growth can be related to lake characteristics such as depth, perimeter and shape. For instance, it is well known that ice thickness growth differs from the borders to the center of most lakes. Introduction of these variables should be considered in future work to improve the accuracy of the developed ice growth models.

8. Conclusion

L'article présenté dans ce mémoire apporte des contributions importantes au domaine de recherche relatif à l'utilisation des ensembles des réseaux de neurones dans l'estimation des épaisseurs de glace des lacs. En effet, le but principal de cet article est d'appliquer les ensembles des réseaux de neurones pour améliorer l'estimation de l'épaisseur de glace des lacs canadiens. Différentes méthodes utilisées pour la création et la combinaison des membres de l'ensemble ont été évaluées. Les résultats de ce modèle ont été comparés à ceux du modèle de réseau de neurones simple. Dans cette comparaison, des améliorations significatives ont été apportées à la performance du modèle en appliquant l'approche de l'ensemble des réseaux des neurones, plus spécifiquement, quand la méthode de l'empilage est utilisée pour la combinaison des membres du modèle.

Dans la plupart des cas, il a été constaté que le boosting est la meilleure méthode pour réduire l'estimation de l'erreur. L'application de l'empilage a montré une meilleure performance que l'utilisation de la méthode de la moyenne, dans le cas où la même méthode est appliquée pour créer les membres de l'ensemble. L'utilisation d'un

minimum de dix membres dans l'ensemble est requise pour apporter une réduction significative pour l'estimation de l'erreur. Cependant, des meilleures performances peuvent être obtenues en utilisant vingt membres. Il est nécessaire de noter que les résultats présentés dans cet article scientifique sont importants pour l'avancement des connaissances dans le domaine scientifique étudié.

REFERENCES

- Ahmad Z., and Zhang J. (2002) A comparison of different methods for combining multiple neural networks models, in *Proceedings of the 2002 World Congress on Computational Intelligence*, pp. 12–117
- Agrafiotis D. K., Cedeno W. and Lobanov V. S. (2002) On the use of neural network ensembles in QSAR and QSPR, *J. Chem. Inf. Comput. Sci.*, 42, pp. 903–911.
- Bates J.M and Granger C.W.J. (1969) The combination of forecasts. *Operations Research Quarterly*, 20 pp. 451-468.
- Bishop C. M. (1995) *Neural Networks for Pattern Recognition*, Oxford Univ. Press, New York.
- Breiman L. (1996a) Bagging predictors, *Mach. Learning*, 24, pp. 123–140.
- Breiman L. (1996b) Stacked regressions, *Mach. Learning*, 24, pp. 49–64.
- Cheng B. and Titterington D.M., (1994) Neural network A review from a statistical perspective, *St& Sci.* 0 (1), pp. 2-54,
- Cho S.B. and Won H.H. (2007) Cancer classification using ensemble of neural networks with multiple significant gene subsets. *Applied intelligence.* (26)-3, pp.243-250
- Chokmani K., Ouarda T.B.M.J., Hamilton S. Ghedira M.H. and Ginras H. (2008) Comparison of ice-affected streamflow estimates computed using artificial neural networks and multiple regression techniques. *Journal of hydrology* 349, pp. 383-396
- Cybenko G. (1989) Approximation by superposition of as sigmoidal function, *Math. Control Signals Syst.* 2 pp. 303–314.
- Demuth H., Beale, M. and Hagan, M. (2008) *Matlab Neural Network Toolbox 6, User's Guide*. The MathWorks Inc.
- Dietterich T. G. (2000) An experimental comparison of three methods for constructing ensembles of decision trees: Bagging, boosting, and randomization, *Mach. Learning*, 40(2), pp. 139–157
- Drucker H. (1997) Improving regressors using boosting techniques, in *Machine Learning: Proceedings of the Fourteenth International Conference*, pp. 107–115,

- Drucker H. (1999) Boosting using neural networks, in *Combining Artificial Neural Nets: Ensemble and Modular Multi-Net Systems*, edited by A. J. C. Sharkey, Springer-Verlag, New York. pp. 51–78,
- Duguay C. R., Flato G. M., Heffries M. O, Ménard P., Morris K., and Rouse W. R. (2003) Ice covers variability on shallow lakes at high altitudes: Model simulation and observations, *Hydroogica Processes*17, pp. 3465–3483.
- Efron, B. (1979) Computers and the theory of statistics: thinking the unthinkable. *SIAM Review*, pp. 21–460.
- Efron, B. and Tibshirani, R. (1993) An introduction to the bootstrap. *Chapman and Hall*.
- English TM.,(1996). Stacked_generalization and simulated evolution. *Biosystems*, 39 (1) pp. 3-18
- Fernandez M, Carreiras MC, Marco J.L.(2006) Modeling of acetylcholinesterase inhibition by tacrine analogues using Bayesian-regularized Genetic Neural Networks and ensemble averaging. *Journal of enzyme inhibition and medicinal chemistry.* (21)-6, pp. 647-661
- Flato G. M., and Brown R. (1996) Variability and climate sensitivity of land fast arctic sea ice, *J. Geophys. Res.*, 101, pp. 25767–25777.
- Foresee, F.D. and Hagan M.T. (1997) Gauss-Newton approximation to Bayesian regularization, *Proceedings of the 1997 International Joint Conference on Neural Networks*pp. pp. 1930–1935.
- Freund Y. and Schapire R. E. (1996) Experiments with a new boosting algorithm, in *Proceedings of the Thirteenth International Conference on Machine Learning*. pp. 148– 156, Morgan Kaufmann, Burlington, Mass.
- Friedman, J., Hastie, T. and Tibshirani, R. (1998) Additive logistic regression, *A statistical view of boosting*. (<http://www-stat.stanford.edu/~jhf>).
- Funahashi K.,(1989) On the approximate realization of continuous mappings by neural networks, *Neuml Networks* 2, pp. 183-192,
- Grover P. L., Burn D. H. and Cunderlik J. M. (2002) A comparison of index flood estimation procedures for ungauged catchments, *Can. J. Civ. Eng.*, 29(5), pp. 734–741.
- Hagan, M. T. (1996) Neural Network Design, *PWS*, Boston
- Hansen L. and Salamon P. (1990) Neural network ensembles, *IEEE Trans. Pattern Anal. Mach. Intell.*, 12, pp. 993–1001.

- Hornik K., Stinchcombe M. and White H. (1989) Multilayer feedforward neural networks are universal approximators, *Neural Networks*, 2, pp. 359– 366.
- Hu M. Y. and Tsoukalas C. (2003) Explaining consumer choice through neural networks: The stacked generalization approach, *Eur. J. Oper. Res.*, 146(3), pp. 650– 660.
- Jacobs R.A., Jordan M.I., Nowlan S.J. and Hinton G.E. (1991) Adaptive mixture of local experts, *Neural Computation* 3, pp. 79–87.
- Jordan M.I. and Jacobs R.A., (1994) Hierarchical mixture of experts and the EM algorithm, *Neural Computation* . 6, pp. 181–214.
- Lachtermacher G. and Fuller J.D. (1994) Backpropagation in hydrological time series forecasting, In *Stochastic and Statistical Methods in Hydrologic and Environmental Engineering*, (Edited by K.W. Hipel et al.), pp. 229-242, Kluwer Academic, Dordrecht,
- Liang R. H., and Hsu Y. (1994) Scheduling of hydroelectric generation using artificial neural networks, *IEE Proc., Part C*, 145, pp. 452– 458.
- Lock, G. S. H. (1990) *The Growth and Decay of Ice*, Cambridge Univ. Press, New York
- Maier, H.R. and Dandy, G.C. (2000) Neural networks for the prediction and forecasting of water resources variables: a review of modelling issues and applications. *Environmental Modelling and Software* 15, pp. 101–23.
- Masters, T. (1995) *Neural, novel and hybrid algorithms for time series prediction*. New York: Wiley.
- Matousek V. (1984) Regularity of the freezing up of the water surface and heat exchange between water body and water surface, *Int. Assoc. of Hydraul. Eng. And Res. Hamburg*, Germany.
- Ménard P., Duguay C. R., Flato G. M., and Rouse W. R. (2002a) Simulation of ice phenology on Great Slave Lake, Northwest Territories, Canada, *Hydrol. Processes*, 16, pp. 3691– 3706.
- Ménard P., Duguay C. R., Flato G. M., and W. R. Rouse (2002b) Simulation of ice phenology on a large lake in the Mackenzie River basin, *Proc. Annu. Meet. Eastern Snow Conf.*, 59th (3), 12.
- Omstedt, A. (1985a) On supercooling and ice formation in turbulent seawater, *J. Glaciol.*, 31, pp. 272–280.

- Omstedt, A. (1985b) Modeling frazil ice and grease ice formation in the upper layers of the ocean, *Cold Reg. Sci. Technol.*, 11, pp. 87– 98.
- Opitz D. and Maclin R. (1999) Popular ensemble methods: An empirical study, *J. Artif. Intell. Res.*, 11, pp. 169–198.
- Ouarda T. B. M. J., Gingras H., Hamilton S., Ghedira H. and Bobée B. (2003) Estimation of streamflow under ice, *paper presented at the 12th Workshop on the Hydraulics of Ice Covered Rivers, Committee on River Ice Processes and the Environment, Canadian Geophysical Union [CGU-HS], Edmonton*, pp. 19– 21.
- Pandey G. R. and Nguyen V.-T.-V. (1999) A comparative study of regression based methods in regional flood frequency analysis, *J. Hydrol.*, 225, pp. 92– 101.
- Perrone M. P., and Cooper L. N. (1993) When networks disagree: Ensemble methods for hybrid neural networks, in *Artificial Neural Networks for Speech and Vision*, edited by R. J. Mammone, pp. 126–142, Chapman and Hall, New York.
- Ripley B.D., (1993) Statistical aspect.6 of neural networks, In *Networks and Chaos-Statistical and Probabilistic Aspects*, (Edited by O.E. Barndorff-Nielsen, J.L. Jensen and W.S. Kendall), pp. 40-123, Chapman and Hell, London,
- Rumelhart D.E. and McClelland J.L.(1986) Parallel Distributed Processing: Explorations in the Microstructure of Cognition, *MIT Press* (1), USA,
- Satyanarayana D., Kannan K. and Manavalan R. (2006) Genetic algorithm optimized neural networks ensemble as calibration model for simultaneous spectrophotometric estimation of atenolol and losartan potassium in tablets. *South African Journal of Chemistry-Sud-Afrikaanse Tydskrif Vir Chemie* (59), pp. 3-11.
- Shamseldin A. Y.(1997) Application of neural network technique to rainfall-runoff modeling, *J. Hydrol.*, 199, pp. 272– 294.
- Schapire R. E.(1990) The strength of weak learnability, *Mach. Learning*, 5, pp. 197–227.
- Schulyakovskii L. G. (Ed.) (1966) Manual of Ice-Formation Forecasting for Rivers and Inland Lakes, *Isr. Prog. for Sci. Transl.*, Jerusalem.
- Seidou O., Ouarda, T.B.M.J., Bilodeau, L., Hessami, M., St-Hilaire, A. and Bruneau, P. (2006) Modelling ice growth on Canadian lakes using artificial neural networks. *Water Ressources Research*, pp. 42.
- Sharkey A. J. C. (Ed.) (1999) Combining Artificial Neural Nets: Ensemble and Modular Multi-Net Systems, *Springer-Verlag*, New York.

- Shen H. T., and Chiang L. A. (1984) Simulation of growth and decay of river ice cover, *J. Hydraul. Eng.*, pp. 110, 958–971.
- Shen, H. T., and C. F. Ho (1986) Two-dimensional simulation of ice cover formation in a large river, paper presented at IAHR Ice Symposium, *Int. Assoc. of Hydraul. Eng. And Res.*, Iowa City, Iowa.
- Shen H. T., Shen H. H., and Tsai S. M. (1990) Dynamic transport of river ice, *J. Hydraul. Res.*, 28, pp. 659–671.
- Shen H. T., Wang D. S., and Lal A. M. W. (1995) Numerical simulation of river ice processes, *J. Cold Reg. Eng.*, 107, pp. 107–118.
- Shu C., Burn, D.H. (2004) Artificial neural network ensembles and their application in pooled flood frequency analysis. *Water Resources Research* 40 (9).
- Shu, C. and Ouarda T. B. M. J. (2007) Flood frequency analysis at ungauged sites using artificial neural networks in canonical correlation analysis physiographic space. *Water Resources Research* 7, pp. 43
- Shu, C. and Ouarda T. B. M. J. (2008) Regional Flood Frequency Analysis at Ungauged Sites Using the Adaptive Neuro-Fuzzy Inference System. *Journal of Hydrology*. (349), iss. 1-2, pp. 31-43.
- Solar Radiation Monitoring Laboratory (2004) Solar radiation basics, *report, Univ. of Oreg., Corvallis.* (Available <http://solardat.uoregon.edu/SolarRadiationBasics.html>).
- Sollich, P. and Krogh, A. (1996) Learning with ensembles: How over-fitting can be useful. In Touretsky, D., Mozer, M., & Hasselmo, M. (Eds.), *Advances in Neural Information Processing Systems*, 8, pp. 190–196 Cambridge, MA. MIT Press.
- Sridhar D.V., Seagrave R.C. and Bartlett E.B. (1996) Process modelling using stacked neural networks, *AIChE Journal* 42, pp. 2529–2539.
- Statsoft, Statistica Neural Network, *Statsoft*, Tulsa, OK, 1998.
- Svensson U. and Omstedt A. (1994) Simulation of supercooling and size distribution of frazil ice dynamics, *Cold Reg. Sci. Technol.*, 22, pp. 221–233.
- Svensson, U., L. Billfalk and Hammar L. (1989) A mathematical model of border-ice formation in rivers, *Cold Reg. Sci. Technol.*, 16, pp. 179–189.
- Tangang, F. T., Tang, B., Monahan, A. H., and Hsieh, W. W. (1998) Forecasting ENSO events--A neural network-extended EOF approach. *Journal of Climate*, 11, pp. 29-41.

- Ting K.M. and Witten I.H. (1999) Issues in stacked generalization. *Journal of artificial Intelligence Research*, 10, pp. 271-189.
- Werbos, P.J., (1974). Beyond regression: New tools for Prediction and Analysis in the Behavioral Sciences. *Thèse de Ph.D., Harvard University*, Cambridge, États-Unis.
- White H. (1989) Learning in artificial neural networks: *A statistical perspective*, *Neural Comp.* 1, pp. 425-464,
- Witten I.H. and Frank E. (1999) Data Mining: Practical Machine Learning Tools and Techniques with Java Implementations, *Morgan Kaufmann*, San Francisco
- Wolpert D. H. (1992) Stacked generalization, *Neural Networks*, 5, pp. 241–259.
- Xiao R., and V. Chandrasekar (1997) Development of a neural network based algorithm for rainfall estimation from radar observation, *IEEE Trans. Geosci. Remote Sens.*, 35, pp. 160–171.
- Zhang J., Martin E.B., Morris A.J. and Kiparissides C., (1997) Inferential estimation of polymer quality using stacked neural networks, *Computers & Chemical Engineering* 21, pp. 1025–s1030.
- Zhang S. P., Watanabe H. and Yamada R. (1994) Prediction of daily water demand by neural networks, in *Stochastic and Statistical Methods in Hydrology and Environmental Engineering*, 3, edited by K. W. Hipel et al., pp. 217–227, Springer, New York.

Table 1: Ice Thickness Measurement Stations

Station Number	Code station	Water body	Longitude °W	Latitude °N	Number of hidden neurons	Best combination of input variables
1	HA1	Quaqtaq	68.36	61.03	1	$D_d; Rad_{nc} + Rad_c; P; S$
2	WFN	Cree Lake	105.15	57.33	9	$D_d; Rad_{nc} + Rad_c$
3	WLH	Lansdowns house	86.09	52.21	4	$D_d; P; S$
4	WTL	Big trout Lake	88.11	53.81	1	$D_d; Rad_{nc} + 0.75 Rad_c; P; S$
5	YAH	La grande IV	72.36	53.76	3	$D_d; P$
6	YBK	Baker Lake	95.96	64.30	1	$D_d; Rad_{nc} + Rad_c; P$
7	YBT	Brochet	100.31	57.86	1	$D_d; Rad_{nc} + 0.75 Rad_c; S$
8	YBX	Blanc Sablon	56.81	51.45	10	$D_d; Rad_{nc} + 0.75 Rad_c$
9	YEI	Ennadai Lake	99.09	61.11	1	$D_d; Rad_{nc} + Rad_c; S$
10	YGK	Kingston	75.5	44.2	9	$D_d; Rad_{nc} + 0.25 Rad_c; S$
11	YGM	Gimli	95.01	50.61	3	$D_d; Rad_{nc} + 0.75 Rad_c; S$
12	YIV	Island Lake	93.31	53.84	2	$D_d; Rad_{nc} + Rad_c; S$
13	YKL	Schefferville	65.19	54.78	9	$D_d; Rad_{nc}$
14	YPY	Fort chipwyan	110.83	58.7	9	$D_d; Rad_{nc}; P$
15	YVP	Kuujuuaq	67.53	58.11	2	$D_d; Rad_{nc} + Rad_c; S$
16	YYR	Goose bay	59.58	53.33	9	$D_d; Rad_{nc} + 0.75 Rad_c;$
17	YZF	Yellowknife	114.34	62.45	2	$D_d; Rad_{nc} + Rad_c; S$

Table 2: Coefficient of determination and NASH of ANN Ensembles

Station Number	r^2	Nash
1	0.96	0.96
2	0.88	0.88
3	0.95	0.93
4	0.92	0.92
5	0.93	0.92
6	0.97	0.97
7	0.93	0.92
8	0.80	0.80
9	0.92	0.90
10	0.84	0.78
11	0.92	0.90
12	0.84	0.82
13	0.91	0.90
14	0.75	0.74
15	0.96	0.90
16	0.84	0.82
17	0.93	0.92

Table 3: Performance indices for ice thickness estimation using 10 members in the ANN Ensemble

Method	RMSE(cm)	RRMSE	Bias(cm)	Rank Score
Single ANN	8.42	0.15	0.52	0.22
Random_Mean	7.45	0.13	0.49	0.23
Random_Stack	4.03	0.35	0.09	0.39
Bag_Mean	3.96	0.32	0.07	0.50
Bag_Stack	5.09	0.07	0.08	0.39
Boost_Median	2.28	0.22	-0.01	0.83
Boost_Stack	2.39	0.03	0.01	0.85

Table 4: Performance indices for ice thickness estimation using 15 members in the ANN Ensemble

Method	RMSE(cm)	RRMSE	Bias (cm)	Rank Score
Single ANN	8.42	0.15	0.52	0.22
Random_Mean	7.33	0.12	0.48	0.25
Random_Stack	2.92	0.24	0.06	0.50
Bag_Mean	3.01	0.24	0.08	0.57
Bag_Stack	3.29	0.04	0.01	0.67
Boost_Median	1.71	0.18	0.01	0.67
Boost_Stack	1.80	0.02	0.02	0.87

Table 5: Performance indices for ice thickness estimation using 20 members in the ANN Ensemble

Method	RMSE	RRMSE	Bias	Rank Score
Single ANN	8.42	0.15	0.52	0.22
Random_Mean	6.98	0.11	0.47	0.29
Random_Stack	2.10	0.16	0.05	0.44
Bag_Mean	1.88	0.16	0.04	0.38
Bag_Stack	2.19	0.02	0.01	0.66
Boost_Median	1.41	0.14	0.01	0.61
Boost_Stack	1.53	0.02	-0.04	0.88

Figure Captions

Figure 1: Geographic location of studied lakes

Figure 2: RMSE vs. the number of networks in the ensembles at WFN station

Figure 3: RMSE of all stations using 20 members in the ANN Ensemble

Figure 4: BIAS of all stations using 20 members in the ANN Ensemble

Figure 5: RRMSE of all stations using 20 members in the ANN Ensemble

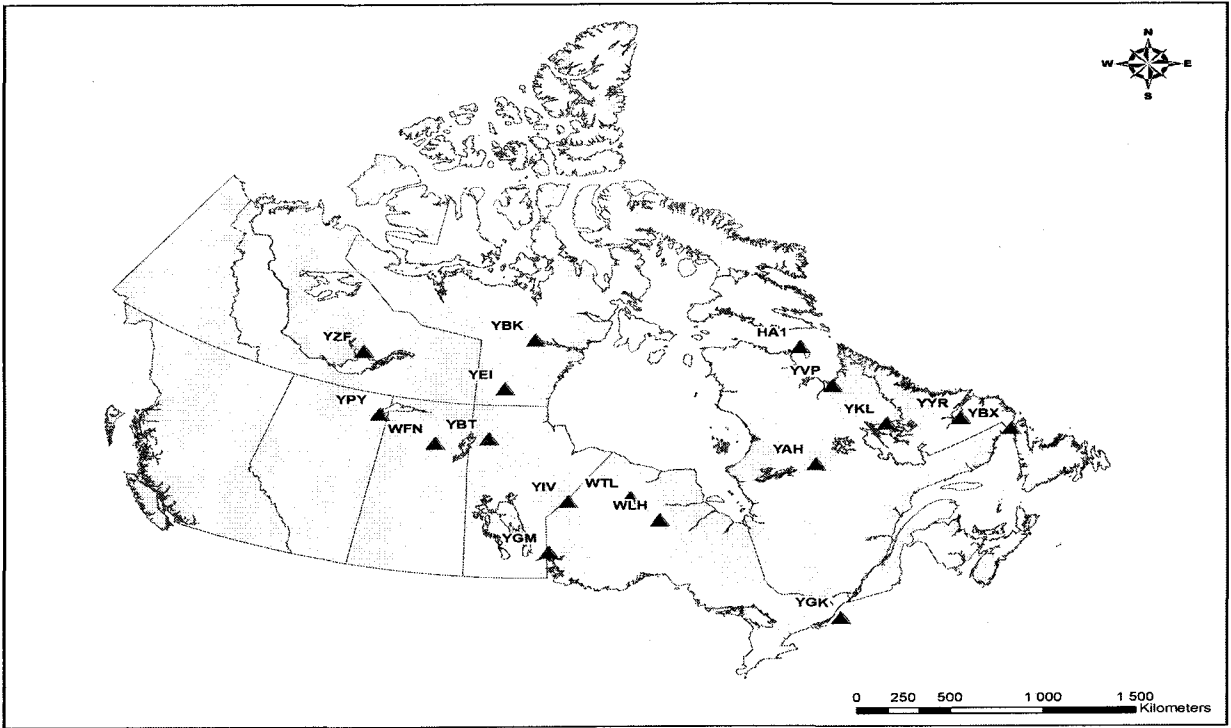


Figure 1: Geographic location of studied lakes

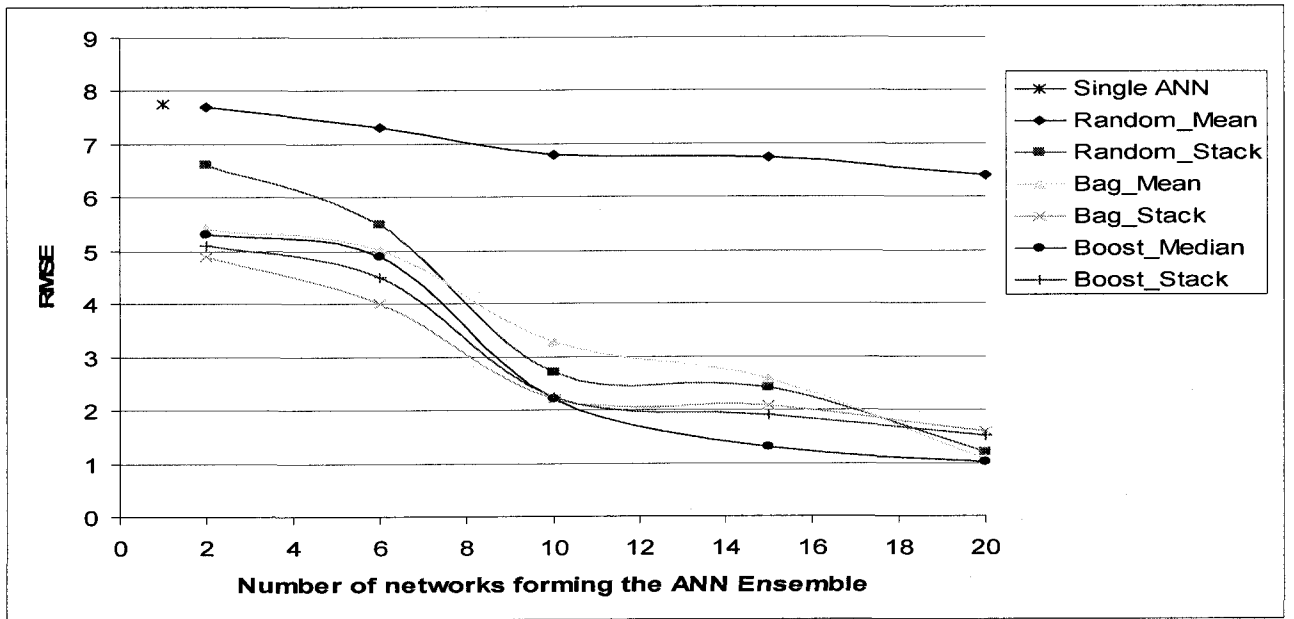


Figure 2: RMSE vs. the number of networks in the ensembles at WFN station

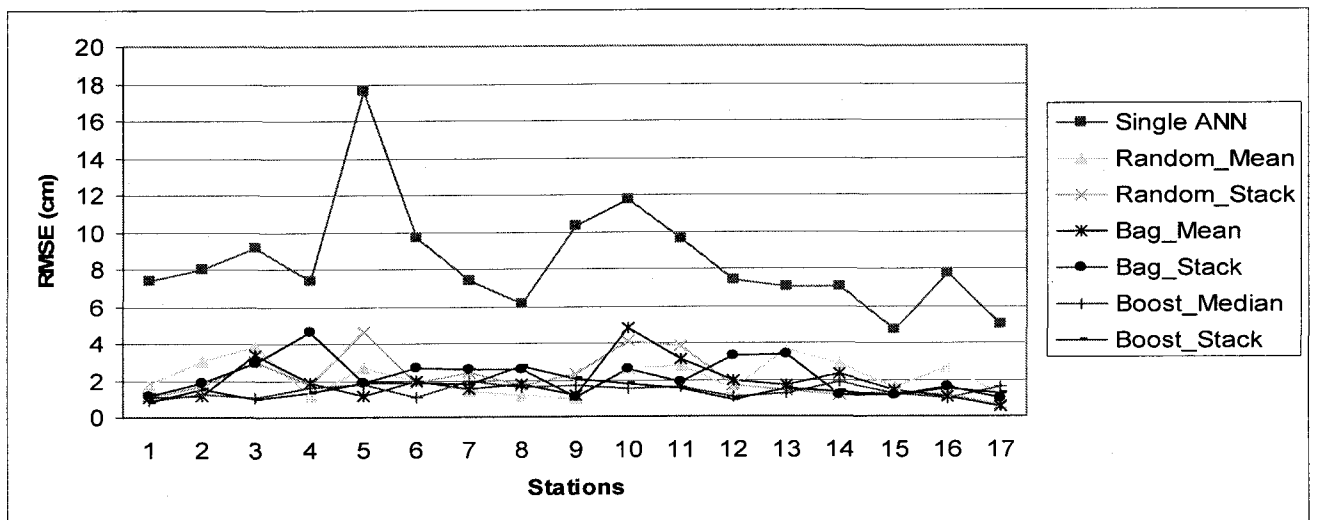


Figure 3: RMSE of all stations using 20 members in the ANN Ensembles

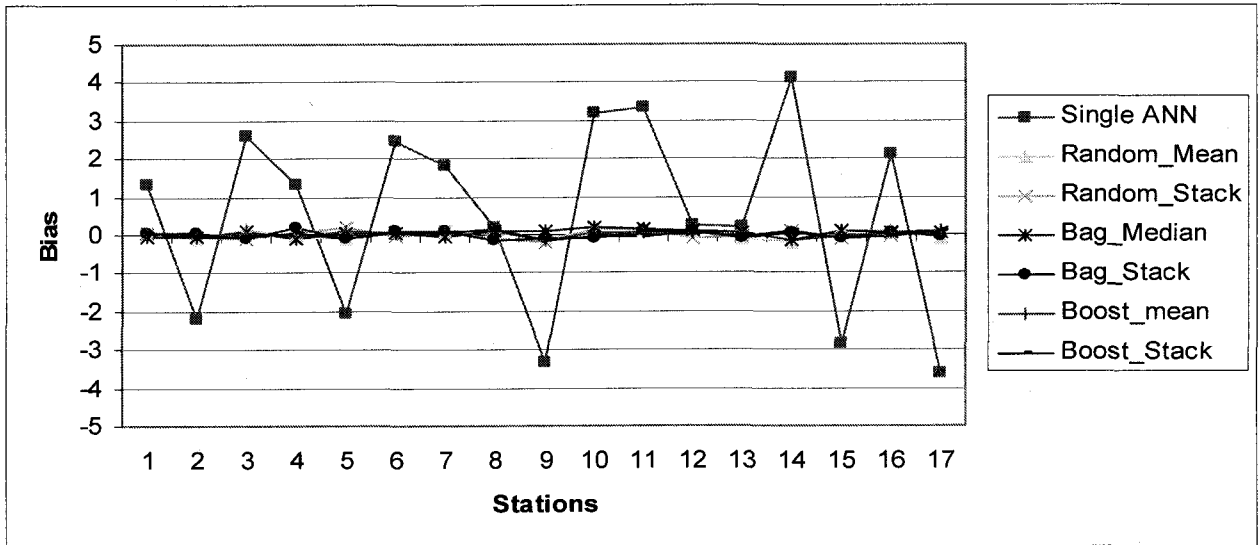


Figure 4: BIAS of all stations using 20 members in the ANN Ensembles

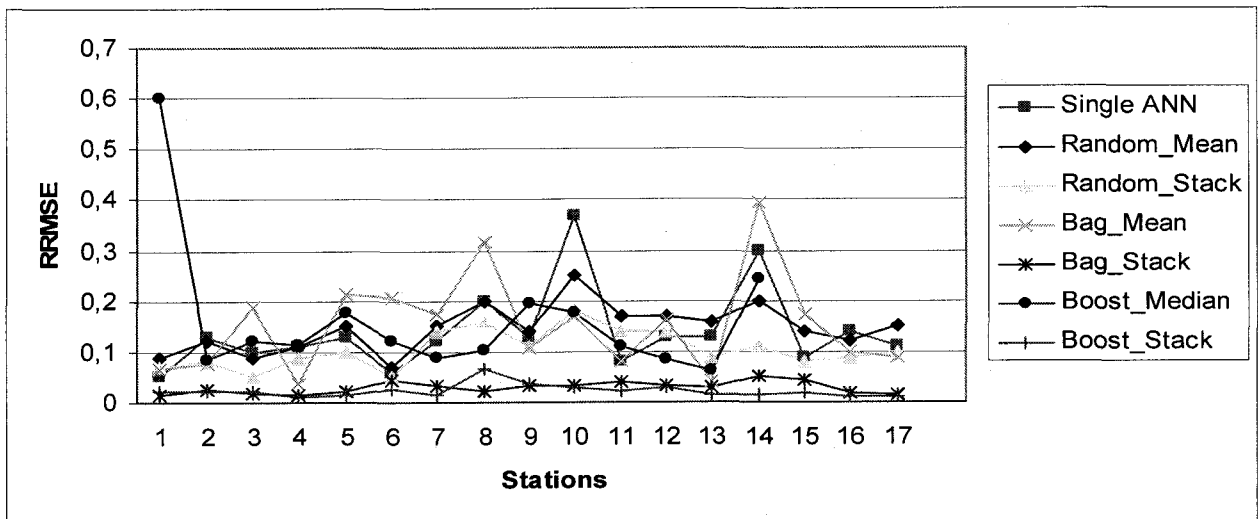


Figure 5: RRMSE of all stations using 20 members in the ANN Ensembles

3. Modélisation de l'épaisseur de glace des lacs en utilisant les réseaux de neurones artificiels et la Loi de Stefan Révisée dans le contexte de changement climatique (RAPPORT)

Préparé par :

Imen Zaier

Fateh Chebana

Taha B.M.J Ouarda

Georges Desrochers

Luc Roy

Chaire industrielle en hydrologie statistique

Institut National de la Recherche Scientifique,

Centre Eau Terre Environnement 490 de la Couronne

G1K 9A9 Québec (QC)

Rapport de recherche

Juillet 2008

Liste des figures

Figure 3.1 : Réseau de neurones artificiel

Figure 3.2 : Localisation du site d'application

Figure 3.3 : Localisation de la station YAH dans la grille

Figure 3.4 : Comparaison de l'épaisseur de glace observée, celle calculée par le RNA et celle calculée par le LSR

Figure 3.5 : Moyenne interannuelle journalière de l'épaisseur de glace de 1961 à 2070

Figure 3.6: Valeurs annuelles journalières de l'épaisseur de glace en utilisant le modèle RNA de 1961 à 2070

Figure 3.7 : Valeurs annuelles journalières de l'épaisseur de glace en utilisant le modèle LSR de 1961 jusqu'à 2070

Figure 3.8 : Valeurs journalières de l'épaisseur de glace pour chaque année du RNA en fonction de la somme de degrés jours de 1961 à 2070

Figure 3.9 : Valeurs journalières de l'épaisseur de glace pour chaque année du LSR en fonction de la somme de degrés jours de 1961 à 2070.

Figure 3.10 : Moyenne interannuelle journalière de l'épaisseur de glace pour chaque décennie en utilisant le RNA

Figure 3.11: Moyenne interannuelle journalière de l'épaisseur de glace pour chaque décennie en utilisant le LSR

Figure 3.12 : Moyenne interannuelle journalière de l'épaisseur de glace pour chaque trentaine d'années en utilisant le modèle RNA

Figure 3.13 : Moyenne interannuelle journalière de l'épaisseur de glace pour chaque trentaine d'années en utilisant le modèle LSR

Figure 3.14 : Moyenne interannuelle journalière de l'épaisseur de glace pour chaque trentaine d'années en fonction de la somme de degrés jours en utilisant le modèle RNA

Figure 3.15 : Moyenne interannuelle journalière de l'épaisseur de glace pour chaque trentaine d'années en fonction de la somme de degrés jours en utilisant le modèle LSR

Figure 3.16 : Moyennes mobiles sur les maximums annuels de l'épaisseur de glace calculée par le modèle RNA

Figure 3.17 : Moyennes mobiles sur les maximums annuels de l'épaisseur de glace calculée par le modèle LSR

Figure 3.18 : Variation des moyennes mobiles sur les maximums annuels de l'épaisseur de glace calculée par le modèle RNA

Figure 3.19: Variation des moyennes mobiles sur les maximums annuels de l'épaisseur de glace calculée par le modèle LSR

Liste des tableaux

Tableau 3.1 : Critères de performance des modèles RNA et LSR

3.1 Introduction

Après plusieurs années d'intenses controverses, le diagnostic scientifique sur le changement climatique ne fait aujourd'hui quasiment plus débat : les émissions de gaz à effet de serre provoquées par les activités humaines menacent dangereusement l'équilibre climatique planétaire. Le mécanisme de l'effet de serre, schématiquement, est le suivant (OCDE, 1995) : le rayonnement solaire incident a une partie absorbé par la terre, principalement à la surface, et une autre renvoyé vers l'espace sous forme de rayonnements de longueur d'onde plus élevée. Une partie de ce rayonnement est à son tour absorbée et réfléchi par les gaz à effet de serre de l'atmosphère, principalement par la vapeur d'eau, les nuages, le CO₂ (dioxyde de carbone), le CH₄ (méthane), le N₂O (protoxyde d'azote) et les CFC (chlorofluorocarbones). Autrement dit, la terre reçoit à la fois un rayonnement provenant directement du soleil et un rayonnement réfléchi par l'atmosphère, ce qui entraîne une élévation des températures moyennes à la surface. Cet effet est amplifié par les formes contemporaines du développement des activités humaines. En effet, la plupart des activités humaines (transports, chauffage, réfrigération, industrie, élevage, déchets...) rejettent des gaz à effet de serre.

Le changement du climat dans le monde et plus particulièrement au Canada a déjà commencé à affecter l'être humain. En particulier, un impact important est observé sur les régimes hydrologiques, tant au niveau de la qualité, qu'au niveau de la quantité. Des mesures et des décisions ont été prises pour combattre les changements climatiques par la réduction des émissions de gaz à effet de serre.

La recherche sur les changements climatiques a orienté les scientifiques dans une direction pour mieux suivre l'évolution du climat. Ces chercheurs devraient avoir le sens de l'analyse et de la compréhension des états passé et actuel des changements afin de réduire l'incertitude dans les prévisions du climat de l'avenir.

La plupart des lacs canadiens, particulièrement québécois, sont recouverts de glaces pendant la majeure partie de l'hiver. Beaucoup de progrès ont été réalisés en ce qui concerne la compréhension et la prévision des problèmes de glace et la diminution au maximum possible les dommages socio-économiques connexes. Vue l'importance de ce phénomène, plusieurs chercheurs ont orienté leurs recherches vers ce domaine. Chacun a entamé le problème selon son expertise, à savoir la compréhension des processus physique, l'analyse des séries hydrologiques ou encore l'étude et la description statistique. Le présent rapport aborde le problème de l'évolution de l'épaisseur de glace dans un lac canadien en utilisant deux approches à savoir, une approche déterministe et autre statistique non paramétrique.

La dernière approche, basée sur la modélisation empirique, permet de passer outre les limites des connaissances physiques du système. Elle prend seulement en compte l'aspect aléatoire des phénomènes. Le modèle de réseaux de neurones artificiels (RNA) est un modèle statistique non paramétrique. Cependant, la modélisation déterministe exige des paramètres physiques, comme c'est le cas de la Loi de Stefan Révisée (LSR). L'analyse des variables hydrologiques nordiques peut être faite par ces

méthodes pour montrer les impacts sur l'évolution de la température de l'eau des lacs et sur les différentes caractéristiques du couvert de glace.

Dans le présent travail, une comparaison est effectuée entre les résultats obtenus par les deux modèles RNA et LSR. Cette comparaison a pour objectif d'évaluer l'évolution de la glace dans la situation passée et future. L'application est faite pour l'horizon 2050 au niveau du réservoir la Grande IV, situé au Québec, Canada.

Dans ce qui suit, on présente des généralités sur la modélisation et une description des deux modèles utilisés (section 2), l'application qui a été faite sur le lac en question (section 3) et les résultats obtenus (section 4).

3.2 Modélisation

La prévision hydrologique est définie comme l'estimation des conditions futures des phénomènes hydrologiques pour une période donnée, à partir des observations passées et actuelles (WMO, 1994). Son objectif général est de fournir les meilleures estimations de ce qui peut arriver en un point donné à une date future précise (Hipel et McLeod, 1994), contrairement à la prédiction qui vise l'estimation des conditions futures sans référence à un temps spécifique (Lettenmaier et Wood, 1993). Par exemple, si t désigne le temps d'origine, le calcul de la prévision de l'épaisseur de glace E d'une rivière, qui peut être appliqué aussi pour le cas d'un lac, au temps $(t + L)$ est de la forme suivante :

$$E_{t+L} = \hat{E}_t(L) + \hat{e}_t(L) \quad (3.1)$$

où $\hat{E}_i(L)$ est la prévision, L est le pas de temps en avant, $\hat{e}_i(L)$ est l'estimation de l'erreur de prévision. La méthode la plus connue pour obtenir les meilleures prévisions possibles consiste à minimiser l'estimation de l'erreur de prévision $\hat{e}_i(L)$ (Coulibaly et al., 1999).

Les mesures d'erreur généralement employées sont : la racine de l'écart quadratique moyen (REQM), la racine de l'écart quadratique moyen relatif (REQMR), le BIAIS, le coefficient NASH et le coefficient de détermination (r^2). Ces indices vont permettre de comparer les performances des deux modèles par rapport aux données observées. Ils sont donnés respectivement par :

$$\text{REQM} = \frac{1}{n} \sum_{i=1}^n (E_i - \hat{E}_i)^2 \quad (3.2)$$

$$\text{REQMR} = \frac{1}{n} \sum_{i=1}^n \frac{(E_i - \hat{E}_i)^2}{E_i^2} \quad (3.3)$$

$$\text{BIAIS} = \frac{1}{n} \sum_{i=1}^n (E_i - \hat{E}_i) \quad (3.4)$$

$$\text{NASH} = 1 - \frac{\sum_{i=1}^n (E_i - \hat{E}_i)^2}{\sum_{i=1}^n (E_i - \bar{E}_i)^2} \quad (3.5)$$

$$r^2 = \frac{\text{cov}([E_i^1, \dots, E_i^n], [\hat{E}_i^1, \dots, \hat{E}_i^n])^2}{\text{var}([E_i^1, \dots, E_i^n]) \text{var}([\hat{E}_i^1, \dots, \hat{E}_i^n])} \quad (3.6)$$

avec n est le nombre d'observations, \hat{E}_i est l'épaisseur de glace estimée au jour i , E_i est l'épaisseur de glace observée au jour i et \bar{E} est la moyenne des épaisseurs de glace observées.

3.2.1 Modèle déterministe : Loi de Stefan Révisée

Dans la pratique, la température de l'air est la variable sur laquelle se base les formules simples du bilan d'énergie. Dans ce contexte, selon Lock (1990), l'équation la plus utilisée est celle de la Loi de Stefan, elle s'écrit sous la forme suivante:

$$E = k\sqrt{Dj} \quad (3.7)$$

avec E est l'épaisseur de glace, Dj est la somme degrés-jours au-dessous du point de congélation depuis la naissance de la glace dans un hiver donné (saison de croissance de glace) et k est une constante.

Dans le cadre de cette étude, la Loi de Stefan Révisée de Seidou et al. (2006) est utilisée au lieu de la Loi de Stefan. Ceci est dû au manque de connaissances au niveau de la date du début de l'accumulation des degrés-jours pour une saison de croissance de glace donnée. Cette loi se base sur la variable Djs , qui représente l'accumulation de la somme des degrés-jours commençant par le premier jour où la température de l'air est en dessous de zéro dans une saison donnée, débutant le 15 août et se terminant 31 mai. Cette loi est de la forme suivante :

$$E = \begin{cases} k\sqrt{Djs - C} & \text{si } Djs \geq C \\ 0 & \text{si } Djs < C \end{cases} \quad (3.8)$$

où C , en °C*jour, représente le retard entre le jour de la naissance de la glace et le premier jour de congélation de la température moyenne de l'air.

3.2.2 Modèle statistique non paramétrique: Réseau de Neurones Artificiels

Dans cette section, le modèle RNA va être décrit, au niveau de sa structure et les étapes à suivre pour son implémentation.

3.2.2.1 Description du modèle

Les RNAs constituent une méthode d'approximation des systèmes complexes, comme les lacs ou les rivières. Ces modèles sont particulièrement utiles lorsque ces systèmes sont difficiles à modéliser à l'aide des équations physiques ou avec des méthodes statistiques classiques (Dawson et Wilby, 2001). La figure 3.1 montre un RNA, qui est un assemblage d'éléments de structure identique appelés neurones (ou cellules) interconnectées à l'instar des cellules du système nerveux du cerveau. Chaque point de connexion entre deux neurones, appelé poids, joue le rôle de l'élément principal entre les neurones.

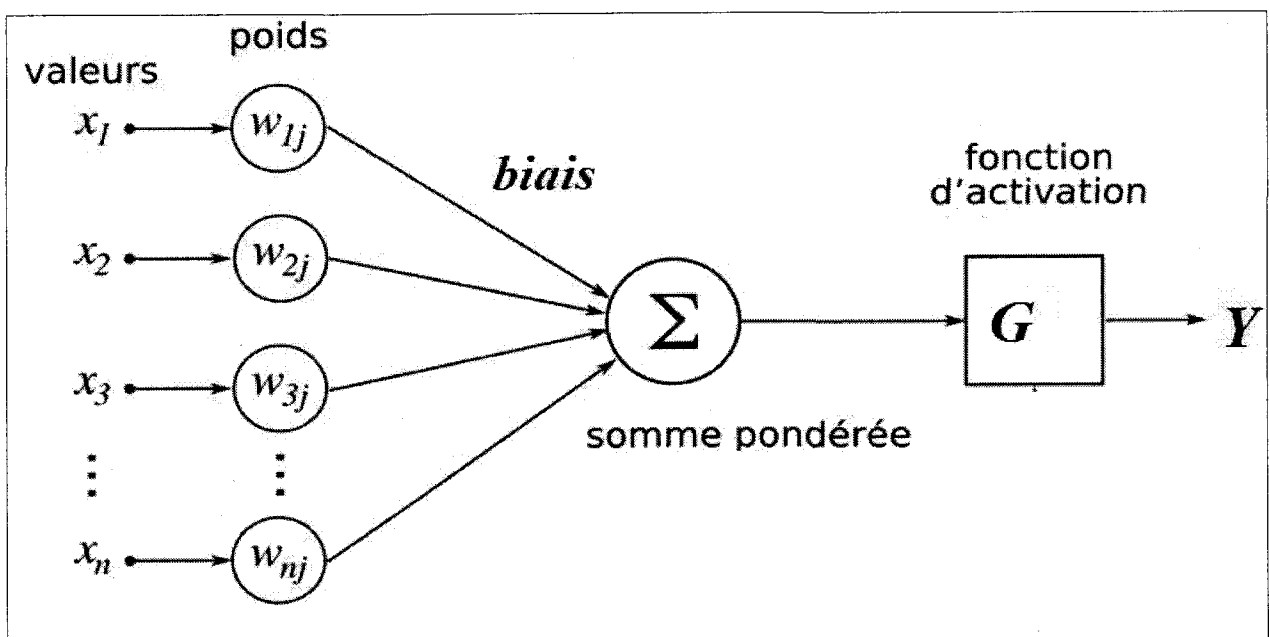


Figure 3.1 : Réseau de neurones artificiel

Les équations qui représentent le modèle RNA se trouvent dans la section 1.2.1. En ce qui concerne la fonction d'activation (ou fonction de transfert), elle peut être de type :

- Sigmoidale : $G(x) = \frac{1}{1 + e^{-x}}$

- Logistique : $G(x) = ax(1 - x)$

- Tangente hyperbolique : $G(x) = \frac{e^x - e^{-x}}{e^x + e^{-x}}$

- Linéaire : $G(x) = ax + b$

L'assemblage de plusieurs neurones permet de former différents types d'architectures du RNA. L'architecture des connexions d'un réseau de neurones multicouches appelée perceptrons multicouches (PMC) a été privilégiée pour la prévision de phénomènes hydrologiques (Dawson et Wilby, 2001). Pour plus de détails sur cette architecture, le lecteur est référé à l'article de Franck (1958).

L'apprentissage est une procédure adaptative par laquelle les connexions des neurones sont ajustées face à une source d'information (Rumelhart et al., 1986). Il existe plusieurs catégories d'apprentissage. La procédure usuelle dans le cadre de la prévision est l'apprentissage supervisé qui consiste à associer une réponse spécifique désirée à chaque signal d'entrée. La mise à jour des poids s'effectue progressivement jusqu'à ce que l'erreur entre les sorties du réseau (ou résultats calculés) et les résultats désirés soit minimisée. L'algorithme de rétro-propagation est l'exemple d'apprentissage supervisé le plus utilisé grâce au succès de certaines applications telle que la détection de la fraude dans les opérations par cartes de crédit (Coulibaly et al., 1999).

3.2.2.2 Structure d'un RNA

Selon Maier et Dandy (2001), un réseau de neurone est en général composé d'une succession d'une ou plusieurs couches d'entrée et d'une ou plusieurs couches cachées. Chacune de ces couches prend ses entrées sur les sorties de la précédente. Chaque couche est composée de neurones, prenant leurs entrées sur les neurones de la couche précédente. À chaque couche, un poids est associé de sorte que les neurones sont multipliés par ce poids, puis additionnés par les neurones du niveau précédant, ce qui est équivalent à multiplier le vecteur d'entrée par une matrice de transformation. Mettre l'une derrière l'autre, les différentes couches d'un réseau de neurones reviendrait à mettre en plusieurs matrices de transformation. Ceci montre l'importance du choix judicieux d'une bonne fonction d'activation.

L'algorithme le plus utilisée dans les applications est la rétro-propagation. Par définition, cette technique de rétro-propagation de l'erreur consiste à corriger les erreurs selon l'importance des entrées qui ont justement contribué à la réalisation de ces erreurs. En effet, les poids qui contribuent à engendrer une erreur importante se verront modifiés et mis à jour de manière plus significative dans le but d'engendrer une erreur minimale (Coulibaly et al., 1999).

3.2.2.3 Étapes d'implémentation d'un RNA

Les étapes nécessaires à implémenter un modèle de RNA sont énumérées ci-dessous (Dawson et Wilby, 2001):





a) Collecte des données: S'assurer que les données sont disponibles et suffisantes pour une étude significative et que la qualité des données est bonne.

b) Sélection des entrées: Identifier les entrées les plus significatives. Si nécessaire, réduire le nombre de ces entrées au moyen de l'analyse en composantes principales.

c) Prétraitement de données :

- Standardiser les données dans des gammes spécifiques.
- Créer des sous-ensembles de données en données de calibration, de validation et de test.

d) Choix de RNA :

- Choisir le type d'architecture de réseau le plus approprié pour l'application. On commence par le modèle le plus classique, qui est le PMC. L'algorithme de l'apprentissage utilisé est la rétro-propagation
- Choisir la fonction d'activation appropriée pour les neurones.
- Choisir la structure de RNA appropriée pour modifier les poids et les biais et déterminer l'architecture du réseau.

e) Apprentissage du réseau: Spécifier le nombre de couches cachées. Souvent, une seule couche cachée est suffisante (Funahashi, 1989).

f) Évaluation: Choisir des mesures d'erreur qui sont appropriées à la sortie du modèle et à la cible (Seidou et al., 2006).

3.3 Application

3.3.1 Site d'application

Dans cette étude, l'application concerne le lac La Grande IV situé dans la province du Québec. Le code de sa station est YAH, ayant une longitude de 72.36 ouest et une latitude de 53.76 nord. La figure 3.2 montre la localisation de ce lac.

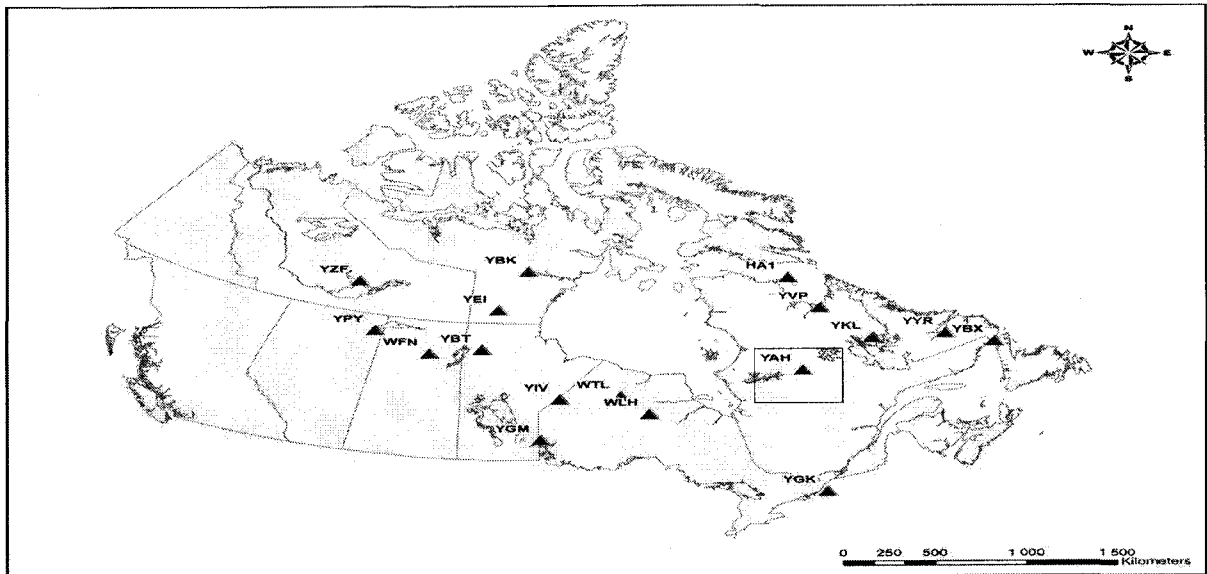


Figure 3.2 : Localisation du site d'application

3.3.2 Données utilisées

La présente étude utilise deux types de données, à savoir :

- les données observées relatives à l'épaisseur de glace qui sont fournies par Environnement Canada selon les coordonnées de la station YAH,
- les données simulées, concernant les variables climatologiques, par le Modèle Régional Climatique Canadien (MRCC) version 4.1.1 (Music et Caya 2007; Brochu et Laprise 2007), sur un domaine centré sur le Québec (112x88 points de grille) avec une grandeur de maillage horizontal de 45 km. Les simulations des deux modèles

climatiques global et régional prennent en considération l'évolution projetée des gaz à effet de serre et des aérosols (Nakicenovic et Swart, 2000). Ces dernières données ont été générées par Hydro-Québec et Ouranos. Pour avoir plus d'informations sur les modèles climatiques mentionnés, le lecteur est référé au site Internet du CCCma (2008). Pour ce type de données, le point le plus proche de la station YAH, comme le montre la figure 3.3, a été considéré dans ce travail.

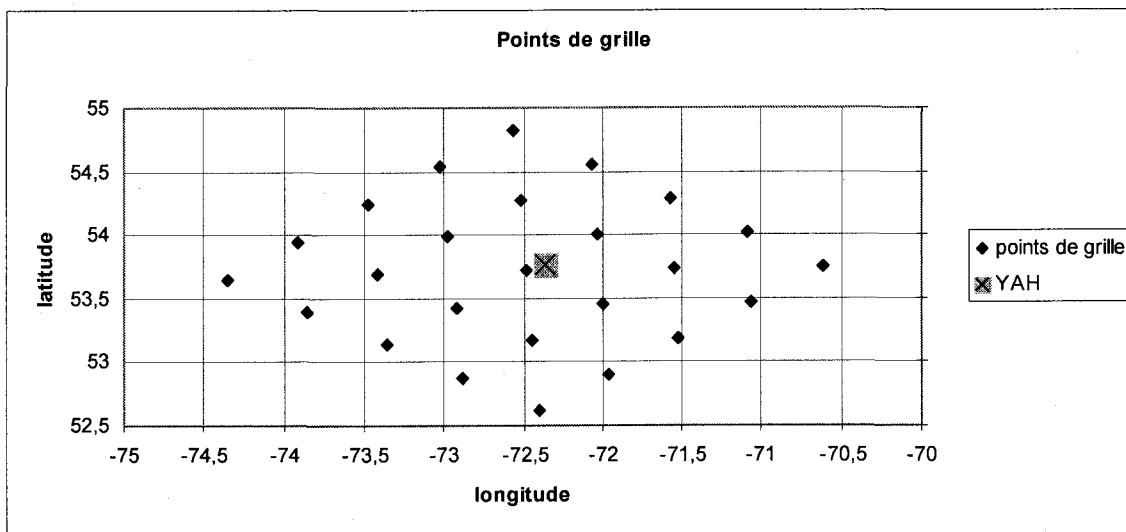


Figure 3.3 : Localisation de la station YAH dans la grille

Les variables utilisées pour la modélisation sont:

- la température moyenne de l'air ($^{\circ}\text{C}$). Elle a servi au calcul de la radiation et de la somme degrés-jours. Pour le calcul de la radiation, le lecteur est référé aux travaux du SRML (2004),
- les précipitations journalières (mm),
- l'épaisseur de la neige journalière au sol (cm).

Comme procédure de validation, le « split sample » a été appliquée. Cette procédure consiste à diviser les données en deux sous-ensembles : 80% des données ont été utilisées pour la calibration du modèle et les 20 % restants ont été appliquées pour la validation. Cette division est faite de façon aléatoire tout en conservant l'ordre séquentiel (Dawson et Wilby, 2001).

3.3.3 Description des caractéristiques des deux modèles

Concernant le modèle LSR, les paramètres k et C de l'équation 3.8 ont été obtenus en minimisant l'erreur quadratique moyenne entre l'épaisseur de la glace observée et celle estimée à partir de la température journalière observée, dans le passé. Selon Seidou et al. (2006), en utilisant les données observées, les paramètres k et C prennent respectivement les valeurs 1,76 et 62,73. Ces mêmes valeurs vont être considérées pour l'estimation de l'épaisseur de glace en utilisant les sorties du MRCC.

Dans le présent travail, le modèle RNA est caractérisé par :

- Perceptron multi-couches :
 - une couche d'entrée, une couche cachée et une couche de sortie,
- Fonction d'activation: tangente hyperbolique,
- Algorithme d'apprentissage : rétro-propagation,
- Nombre de neurones dans la couche cachée: trois.

Dans ce modèle, il y a une seule couche cachée ce qui est suffisant pour faire une application de ce type (Funahashi, 1989). L'apprentissage se fait une seule fois lors de la simulation de la situation passée. Une fois cette dernière réalisée, le modèle mémorise

toutes les données et les caractéristiques de la série hydrologique qui vont être utilisées pour la simulation future. Les variables d'entrées du modèle RNA sont la radiation, la somme degrés-jours, les précipitations et l'épaisseur de glace et de la neige au sol.

Il est à noter que la saison de croissance de glace a été considérée pour la période du 15 août jusqu'au 31 Mai.

3.4. Résultats

Dans cette section, les deux modèles RNA et LSR ont été appliqués pour les situations passée et future. Plus précisément, cette partie se compose de deux sous-parties : (1) une comparaison de l'évolution de l'épaisseur observée, celle calculée par le RNA et celle calculée par le LSR à partir des données observées entre 1985 et 1994 et (2) une comparaison entre l'évolution de l'épaisseur de glace calculée à partir du modèle RNA et celle du modèle LSR en utilisant les sorties du MRCC pour la période de 1961 jusqu'à 2070.

3.4.1 Situation passée : Évaluation avec les données observées

L'évolution de l'épaisseur de la glace dans le lac a été évaluée par les deux modèles mentionnés ci-dessus en utilisant les données climatiques observées dans le passé. Une comparaison a été faite entre les résultats de l'épaisseur de glace obtenues en appliquant ces deux modèles et celle observées entre 1985 et 1994. Le tableau 3.1 présente les valeurs des critères de performance obtenues en appliquant les deux modèles, tant pour la calibration que pour la validation. Ce tableau montre que le modèle RNA performe mieux que le modèle LSR. En effet, au niveau du BIAIS, le modèle RNA

est beaucoup moins biaisé que le modèle LSR. En ce qui concerne la REQMR, celle du modèle RNA est légèrement meilleure que celle du modèle LSR. Concernant le REQM, malgré que le modèle LSR présente une meilleure performance au niveau de la validation, au niveau de la calibration le modèle RNA est plus performant. Le NASH et le r^2 montrent que le modèle RNA est plus adéquat que le modèle LSR.

Tableau 3.1 : Critères de performance des modèles RNA et LSR

	LSR		RNA	
	Calibration 80%	Validation 20%	Calibration 80%	Validation 20%
BIAIS	-4,77	-2,56	0,50	0,67
REQMR	0,17	0,15	0,13	0,11
REQM (cm)	11,00	7,80	8,6	10,11
NASH	0,75	0,75	0,84	0,83
r^2	0,88	0,82	0,91	0,85

La figure 3.4 montre l'évolution de la glace pour le lac entre 1985 et 1994. Les trois courbes de cette figure présentent l'épaisseur de glace observée, celle calculée avec le modèle RNA et celle calculée avec le modèle LSR en utilisant les données climatiques observées comme entrées. Cette évolution considère seulement la période de gel. Les courbes sont assez inclinées vers la droite car il s'agit des valeurs journalières épisodiques pour chaque année. En ce qui concerne l'épaisseur de la glace observée, à l'exception de l'année de 1985 où il y a un manque de données, les valeurs maximales de l'épaisseur varient entre 65 et 108 cm. Les valeurs calculées à partir du modèle RNA sont proche de celles observées et elles varient entre 75 et 105 cm. Par contre, les valeurs maximales de l'épaisseur de glace calculées par le modèle LSR varient entre 85 et 100

cm. Il est à noter que les valeurs maximales de ce dernier modèle dépassent les valeurs observées, pour les années entre 1991 et 1994, de l'ordre d'une dizaine de centimètres.

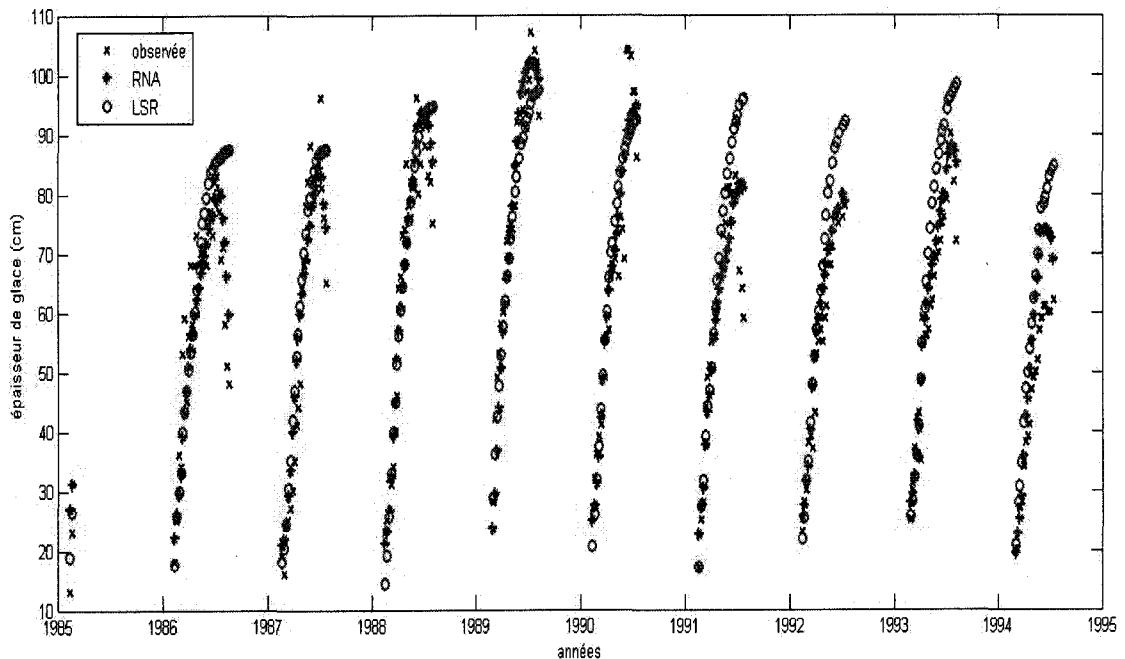


Figure 3.4 : Comparaison de l'épaisseur de glace observée, celle calculée par le RNA et celle calculée par le LSR

Il est à conclure que d'après le tableau 3.1 et la figure 3.4, le modèle RNA est plus performant que le modèle LSR. Malgré cette performance, les résultats des deux modèles vont être présentés pour montrer l'évolution de l'épaisseur de la glace calculée à partir des sorties du MRCC pour la période future.

3.4.2 Situations passée et future : Évaluation avec les données simulées du MRCC (1961-2070)

Dans cette section, les présentations et l'analyse des résultats sont basés sur les moyennes interannuelles classiques et sur les moyennes mobiles.

3.4.2.1 Résultats avec analyse statistique directe

La figure 3.5 présente la moyenne interannuelle journalière de l'épaisseur de glace calculée, sur la période de 1961 jusqu'à 2070, par les deux modèles RNA et LSR. Les deux modèles présentent une augmentation brusque de l'évolution de la glace pour la même date de gel, vers le début du mois de janvier. Avant cette date, l'épaisseur de glace est nulle. Les coefficients directeurs des deux courbes sont assez proches. À une épaisseur de glace de 40 cm, la courbe du modèle RNA admet un point d'inflexion où la croissance continue légèrement jusqu'à la fin du mois de février où elle atteint 58 cm. Cependant, la croissance de l'épaisseur de glace évaluée par le modèle LSR est plus importante que celle du RNA. Elle commence à une épaisseur de 50 cm et s'arrête vers la fin du mois d'avril où elle atteint 95 cm.

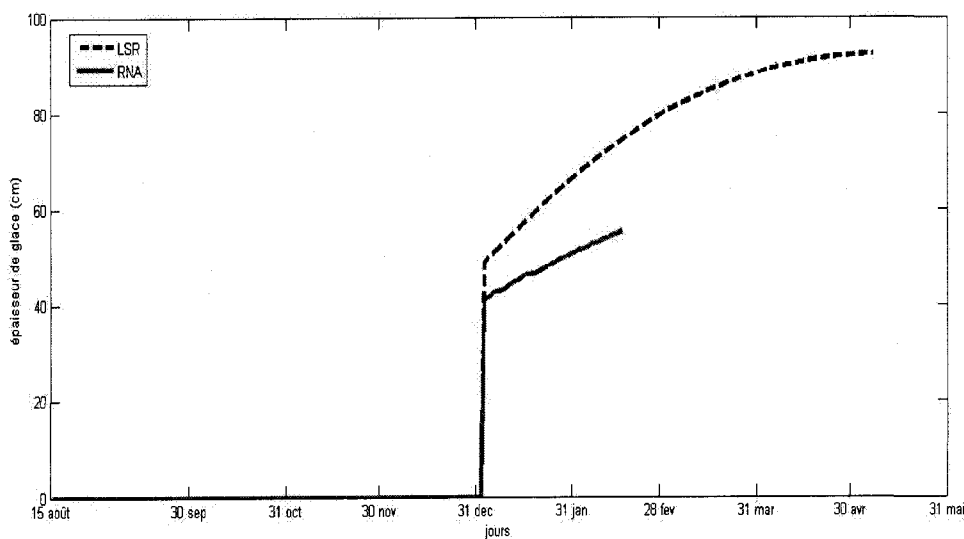


Figure 3.5 : Moyenne interannuelle journalière de l'épaisseur de glace de 1961 à 2070

Les figures 3.6 et 3.7 présentent, respectivement pour le modèle RNA et le modèle LSR, les valeurs annuelles journalières de l'épaisseur de glace du lac étudié entre

1961 et 2070. Chacune des courbes de ces figures représente une année d'évolution d'épaisseur de glace. Pour la figure 3.6, concernant le modèle RNA, l'épaisseur de la glace est négligeable du 15 août au 30 septembre. À partir de cette dernière date et jusqu'au 31 octobre, l'épaisseur de glace continue à croître et elle ne dépasse pas les 40 cm. Vers la mi-novembre, la glace commence à évoluer d'une façon assez linéaire pour toutes les années et elle peut atteindre une épaisseur de 100 cm pour quelques années.

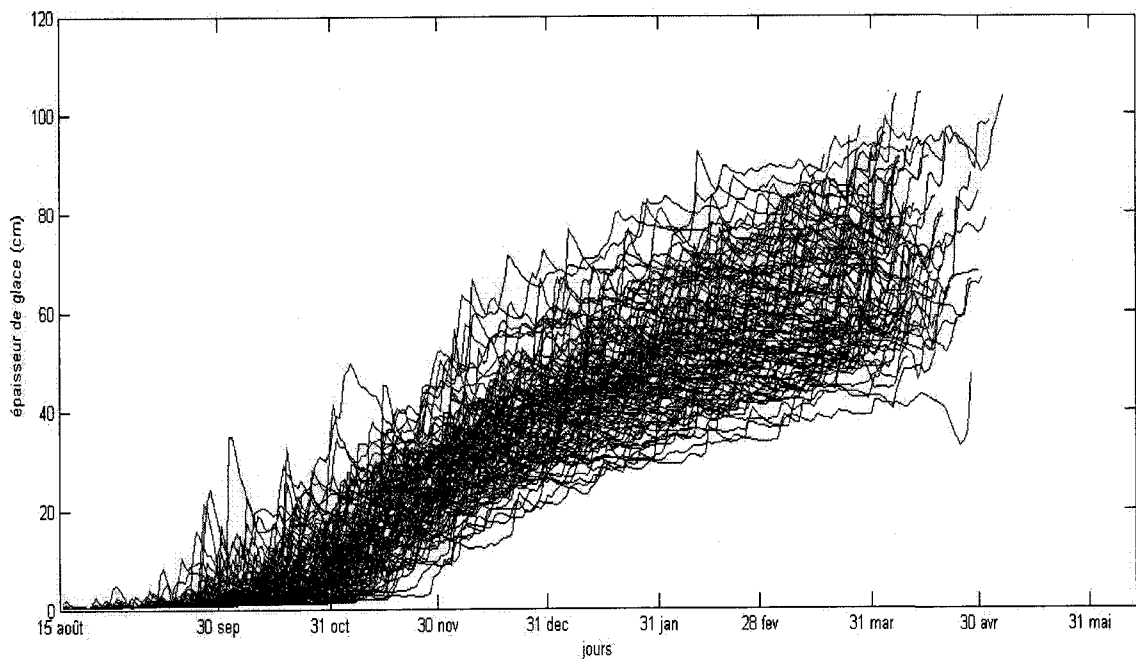


Figure 3.6: Valeurs annuelles journalières de l'épaisseur de glace en utilisant le modèle RNA de 1961 à 2070

En ce qui concerne la figure 3.7, l'épaisseur de glace est calculée en utilisant le modèle LSR. La date du début de l'évolution de la glace est différente d'une année à une autre. Elle varie entre le 30 septembre et le 15 novembre. À partir de ces dates, l'évolution de la glace augmente linéairement jusqu'au 31 mars où elle commence à être constante jusqu'à la fin du mois de mai.

Il est important de noter que l'évolution de la glace pour le modèle RNA est assez irrégulière d'une année à une autre. De plus, les courbes du RNA sont plus dispersées que celles du LSR, en particulier vers la fin de la saison de croissance de la glace. Cette dispersion pourrait être dû à l'interaction de plusieurs variables (la neige, la radiation et les précipitations...) qui interagissent entre elles comme entrées omniprésentes pour le modèle RNA afin de contribuer à la croissance de la glace. Par contre, le modèle LSR fait intervenir une seule variable, à savoir la somme degrés-jours, qui elle-même suit une monotonie ascendante dans le temps.

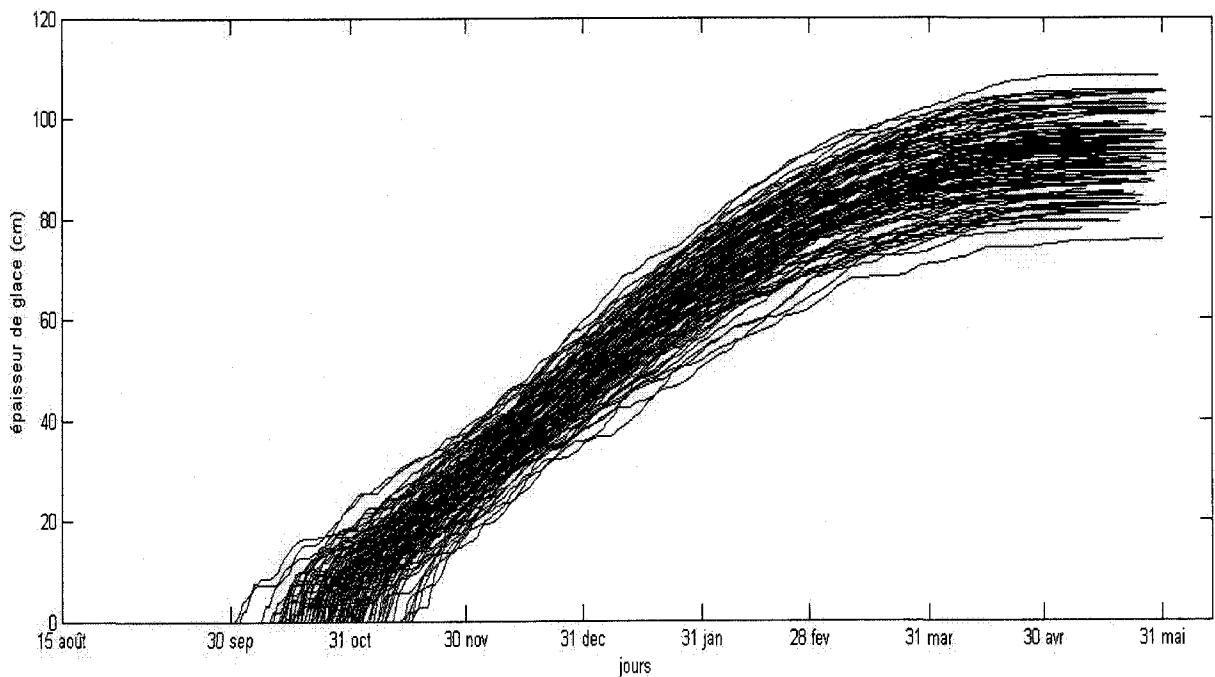


Figure 3.7 : Valeurs annuelles journalières de l'épaisseur de glace en utilisant le modèle LSR de 1961 jusqu'à 2070

Étant donné que la variable de la somme degrés-jours est la seule variable commune entre les deux modèles, il est intéressant de présenter les valeurs de l'épaisseur de la glace en fonction de cette variable. Alors, les mêmes valeurs d'épaisseur de glace

des figures 3.6 et 3.7 sont reproduites dans les figures 3.8 et 3.9 en fonction de la somme degrés-jours pour les modèles RNA et LSR, respectivement. Dans la figure 3.8, l'évolution de la glace est assez brusque pour la plupart des années entre 0 et 500 °C*jour qui correspondent à des épaisseurs de 0 à 20 cm au début et 20 à 70 cm à la fin. À partir de 500 °C*jour, la glace continue à augmenter légèrement pour atteindre des maximums entre 40 et 100 cm qui correspondent à une somme degrés-jours de 2750 °C*jour.

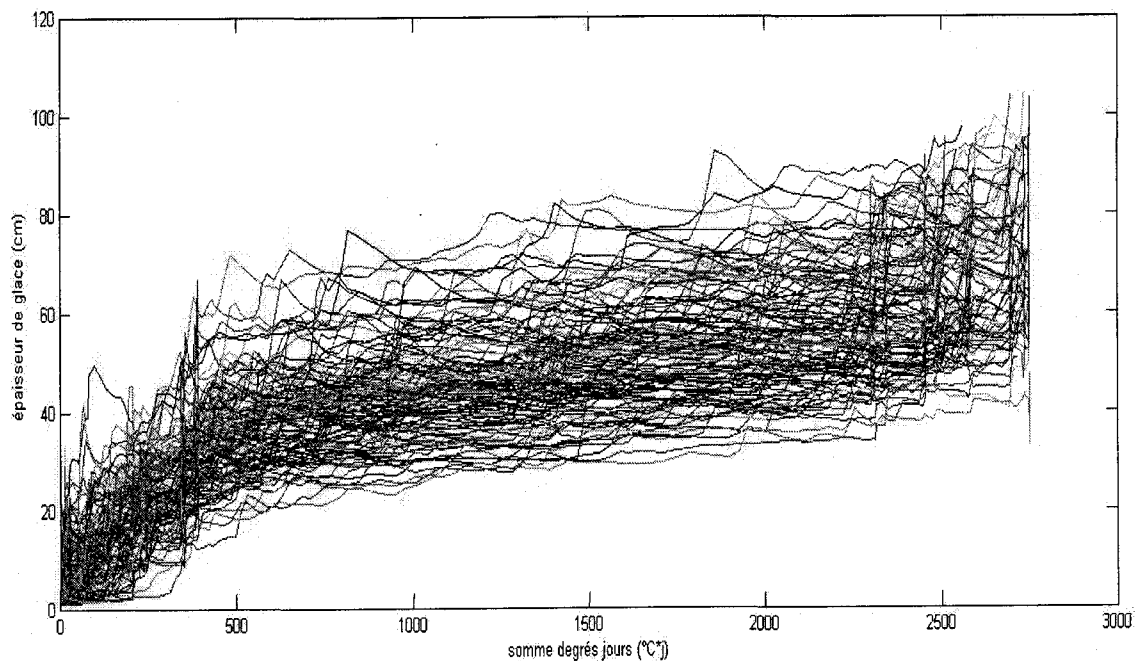


Figure 3.8 : Valeurs journalières de l'épaisseur de glace pour chaque année du RNA en fonction de la somme degrés jours de 1961 à 2070

Au niveau de la figure 3.9, l'évolution de la glace en fonction de la somme degrés-jours se ressemble entre les années. De plus, cette évolution augmente d'une façon linéaire sur deux morceaux. Le premier s'étend entre 0 et 500 °C*jour qui correspondent à 0 et 20 à son début et 25 à 50 cm à sa fin. Le deuxième morceau s'étend de 500 à 2750 °C*jour. Les courbes pour ce morceau ont des pentes moins importantes

que celles du premier, et elles correspondent à des maximums entre 80 et 105 cm. Il est important de noter que les courbes du modèle LSR sont plus régulières et ont moins de variabilité comparées à celle du modèle RNA.

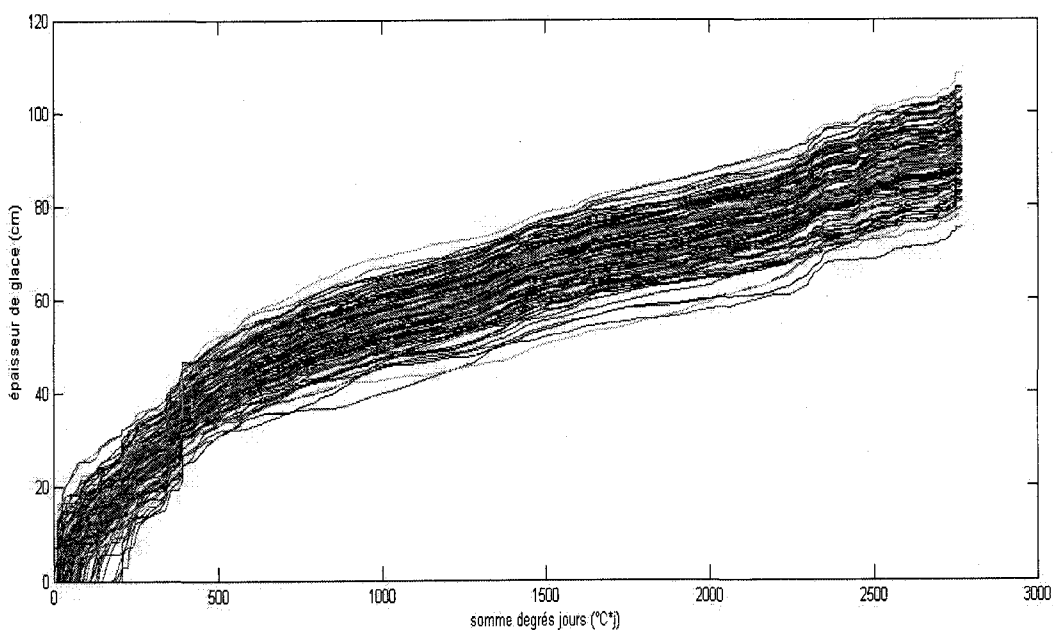


Figure 3.9 : Valeurs journalières de l'épaisseur de glace pour chaque année du LSR en fonction de la somme degrés jours de 1961 à 2070.

Pour une meilleure visualisation de l'évolution de la glace fournie par les deux modèles, des courbes sont représentées pour chaque décennie et pour chaque trente années en allant de 1961 jusqu'à 2070.

Les figures 3.10 et 3.11, respectives aux modèles RNA et LSR, présentent les moyennes interannuelles journalières de l'épaisseur de glace pour chaque décennie de 1961 à 2070. Ces décennies sont notées D1,..., D11, respectivement, de la première à la dernière décennie. Dans la figure 3.10, l'épaisseur de la glace évaluée avec le modèle RNA est négligeable entre le 15 août et le 30 septembre. Au-delà de cette date, la glace commence à augmenter jusqu'à 31 mars où elle peut atteindre les 70 cm. Les courbes de décennies

D7 et D10 représentent les limites inférieure et supérieure dans laquelle se situent les autres courbes. En effet, la courbe D7, qui correspond à la décennie de 2021 jusqu'à 2031, est la décennie où il y a moins de glace et où l'évolution de celle-ci s'arrête vers la fin du mois de février. De plus, la courbe D3 représente la décennie où l'épaisseur de glace est la plus importante, tout au long de l'évolution de la glace. En ce qui concerne les autres décennies, elles se trouvent à l'intérieur des limites des courbes D3 et D7.

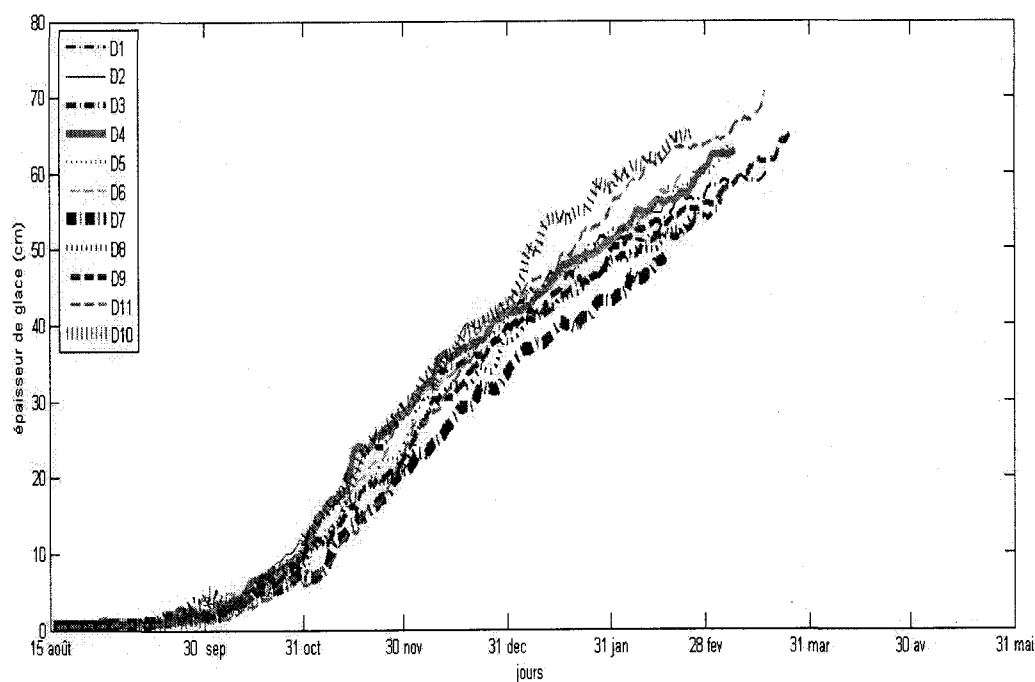


Figure 3.10 : moyenne interannuelle journalière de l'épaisseur de glace pour chaque décennie en utilisant le RNA

En ce qui concerne la figure 3.11, l'épaisseur de glace est calculée par le modèle LSR. Entre la période de la mi-août et la fin octobre, pour toutes les décennies, l'épaisseur de glace est presque nulle. À partir de la fin du mois d'octobre, l'évolution de la glace commence à augmenter jusqu'à la fin du mois de mai pour atteindre des

épaisseurs entre 80 et 100 cm. Vers la fin du mois de mars et jusqu'à la fin de la saison, l'évolution devient presque constante pour toutes les décennies.

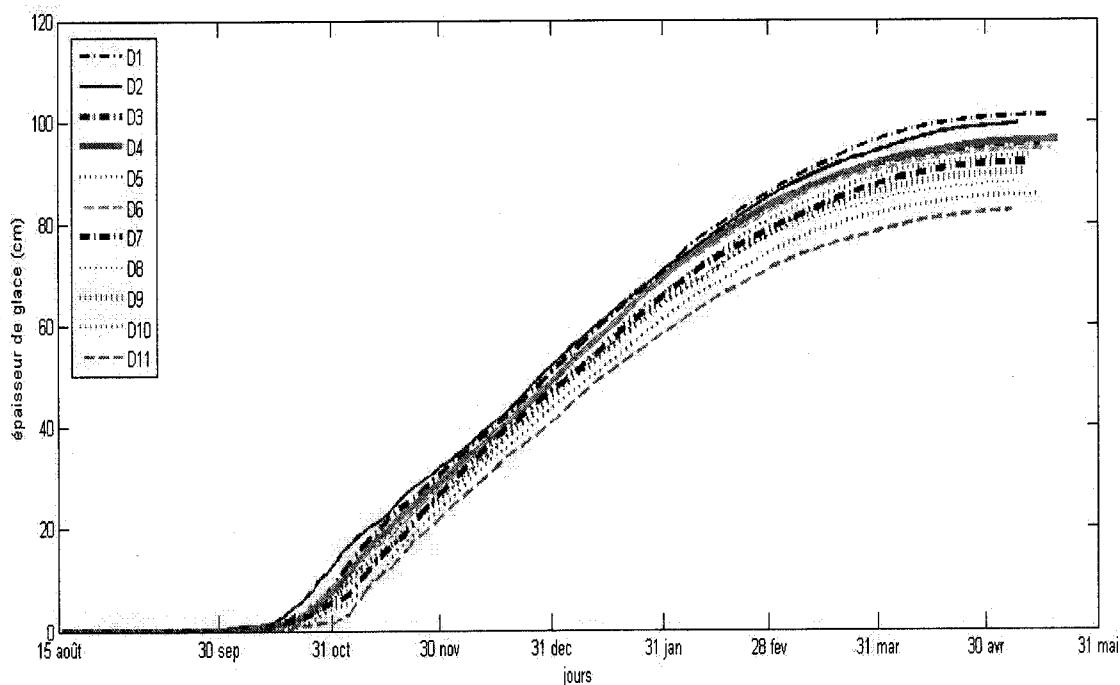


Figure 3.11: Moyenne interannuelle journalière de l'épaisseur de glace pour chaque décennie en utilisant le LSR

Les figures 3.12 et 3.13, relatives respectivement aux modèles RNA et LSR, présentent les moyennes interannuelles journalières calculées sur les périodes successives de 30 ans de 1961 à 2070. Les courbes de ces figures sont notées T1, T2, T3 et T4, de la première 30 ans à la dernière. Il est à noter que celle-ci contient seulement 20 ans. Au niveau de la figure 3.12, il à noter que la courbe T4 représente l'évolution de l'épaisseur de glace la plus importante. Par contre, la courbe de T3 présente l'évolution de glace la moins accentuée. En ce qui concerne les courbes de T1 et T2, elles ont des allures

similaires et elles sont assez rapprochées l'une de l'autre. L'évolution de la glace s'arrête vers la mi-février.

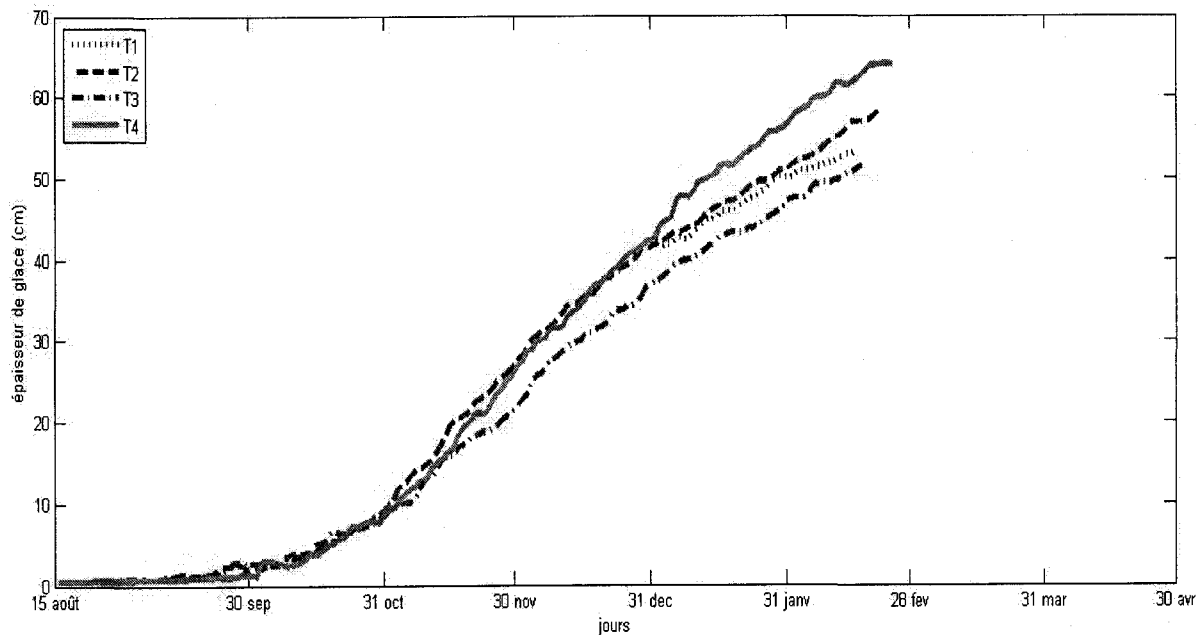


Figure 3.12 : Moyenne interannuelle journalière de l'épaisseur de glace pour chaque trentaine d'années en utilisant le modèle RNA

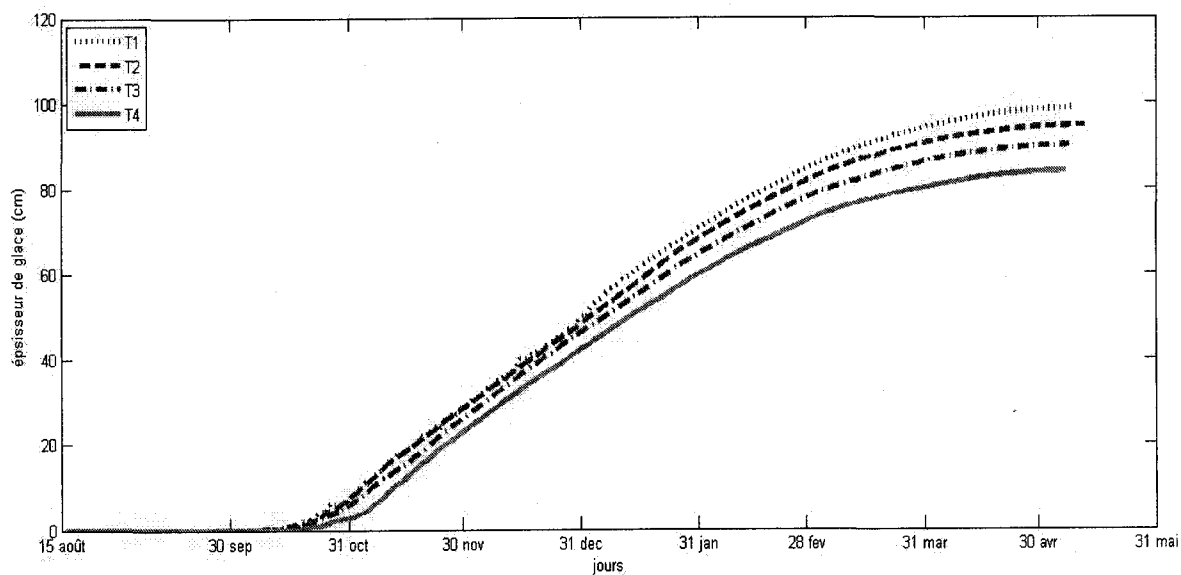


Figure 3.13 : Moyenne interannuelle journalière de l'épaisseur de glace pour chaque trentaine d'années en utilisant le modèle LSR

En ce qui concerne la figure 3.13, les courbes limites inférieure et supérieure sont celles correspondantes respectivement à T1 et T4. L'évolution de la glace, pour toutes ces périodes, s'arrête vers la fin du mois d'avril avec une épaisseur de glace qui peut atteindre 100 cm.

Les figures 3.14 et 3.15 sont similaires aux figures 3.12 et 3.13 respectivement, mais elles reproduisent l'épaisseur de glace en fonction de la somme degrés-jours. Au niveau de la figure 3.14, pour le RNA, l'évolution de la glace est linéaire sur deux morceaux. Le premier morceau s'étend entre 0 et 500 °C*jour et correspond à une épaisseur de glace de 0 à 40 cm. Le deuxième s'étend jusqu'à une somme degrés-jours de 2250, correspondant à une épaisseur de glace commençant entre 30 et 40 cm et se terminant de 50 à 65 cm, dépendamment de la période. D'après les figures 3.5 à 3.13, il est à noter que l'intervalle de croissance de la glace se rétrécit par rapport à la période dans laquelle les moyennes ont été calculées (intervalle de 10 ans, 30 ans, même les valeurs directes). Cela est dû au fait que les moyennes sont calculées sur les dates communes où la valeur qui représente l'épaisseur de glace existe.

L'allure des courbes de la figure 3.15, concernant le LSR, est similaire à celle de la figure 3.14. Autrement dit, toutes les courbes ont une linéarité sur deux morceaux. Le premier morceau des courbes ressemble beaucoup à celui du modèle RNA. Par contre le deuxième morceau s'étend jusqu'à une somme degrés jours de 2750 °C*jour pour toutes les périodes et une épaisseur entre 80 et 100 cm. Il est à noter que l'ordre des courbes des figures 3.14 et 3.15 est le même que les figures 3.12 et 3.13, respectivement.

Des figures similaires aux figures 3.14 et 3.15 pour les décennies pourraient être faites, mais elles sont omises parce qu'elles mènent aux mêmes conclusions.

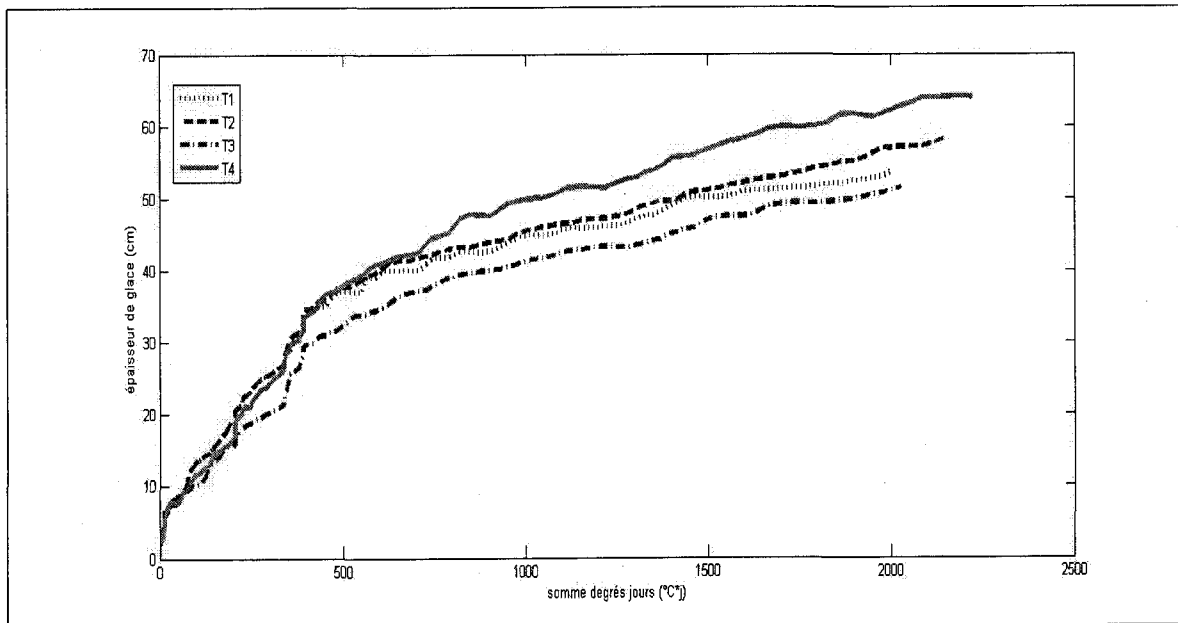


Figure 3.14 : Moyenne interannuelle journalière de l'épaisseur de glace pour chaque trentaine d'années en fonction de la somme degrés jours en utilisant le modèle RNA

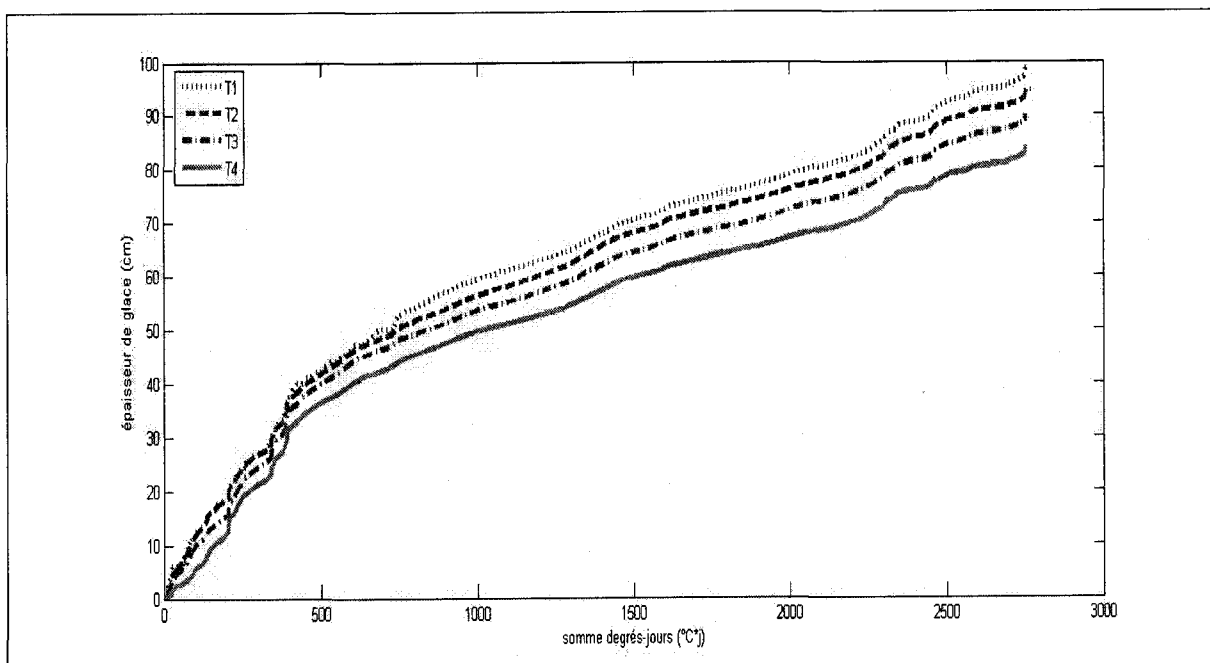


Figure 3.15 : Moyenne interannuelle journalière de l'épaisseur de glace pour chaque trentaine d'années en fonction de la somme degrés jours en utilisant le modèle LSR

3.4.2.2 Résultats avec moyennes mobiles

Selon Brockwell et Davis (1987), la moyenne mobile est définie comme un type de moyenne statistique utilisée pour analyser des séries ordonnées de données, le plus souvent des séries temporelles. Par exemple, dans le domaine de l'hydrologie, en supprimant les fluctuations transitoires de façon à détecter les tendances à plus long terme. Cette moyenne est dite mobile parce qu'elle est recalculée de façon continue, en utilisant à chaque calcul un sous-ensemble d'éléments dans lequel un nouvel élément remplace le plus ancien ou s'ajoute au sous-ensemble. La description de l'évolution de l'épaisseur de glace nécessite successivement la transformation de la chronique brute en série de moyennes mobiles sur une durée bien définie.

Dans ce qui suit, les moyennes mobiles sont basées sur les résultats des maximums annuels de l'épaisseur de glace et une période de 30 ans sur la période entre 1961 et 2070. Dans ce cas, il est nécessaire de noter que les résultats seront présentés entre 1976 et 2054, où chaque année représente le centre d'une période de trente ans. Par exemple, l'année 1976 est le centre de la période de 1961 à 1991. Les figures 3.16 et 3.17 présentent les moyennes mobiles sur les maximums annuels de l'épaisseur de glace calculées à partir des résultats des modèles RNA et LSR, respectivement.

La courbe de la figure 3.16 peut être divisée en trois parties. La première partie est entre 1976 et 2000. Pendant cette période, l'épaisseur de glace augmente légèrement de 69 à 75 cm avec quelques fluctuations par certaines années. La deuxième partie s'étend de 2000 jusqu'à 2030, pendant laquelle, il y a une chute assez importante de

l'épaisseur de glace de 75 à 65 cm. La troisième partie s'étend de 2030 jusqu'à 2054, où l'épaisseur de glace reprend son évolution d'augmentation pour atteindre presque 80 cm. L'intervention de plusieurs variables comme entrées pour le modèle RNA pourrait expliquer cette variation non monotone sur toute la période. Pour avoir une idée quantitative sur la différence globale de l'évolution de la glace sur toute la période, dans cette figure, $m_1=72$ cm et $m_2=71$ représentent, respectivement, les moyennes des moyennes mobiles pour les trente premières et dernières années de la période étudiée. Puisque, il existe des fluctuations assez importantes, la différence entre ces deux moyennes n'est seulement que de 1 cm.

En ce qui concerne la figure 3.17, l'épaisseur de glace, évaluée en utilisant le LSR, diminue légèrement et d'une façon linéaire sur toute la période. Cette dégradation commence avec une épaisseur de 100 cm de glace pour la première année et se termine avec une épaisseur d'environ 85cm. De plus, dans cette figure $m_1=96$ cm et $m_2=90$ représentent, respectivement, les moyennes des moyennes mobiles pour les trente premières et dernières années de la période étudiée. La différence entre ces deux moyennes est de 6 cm.

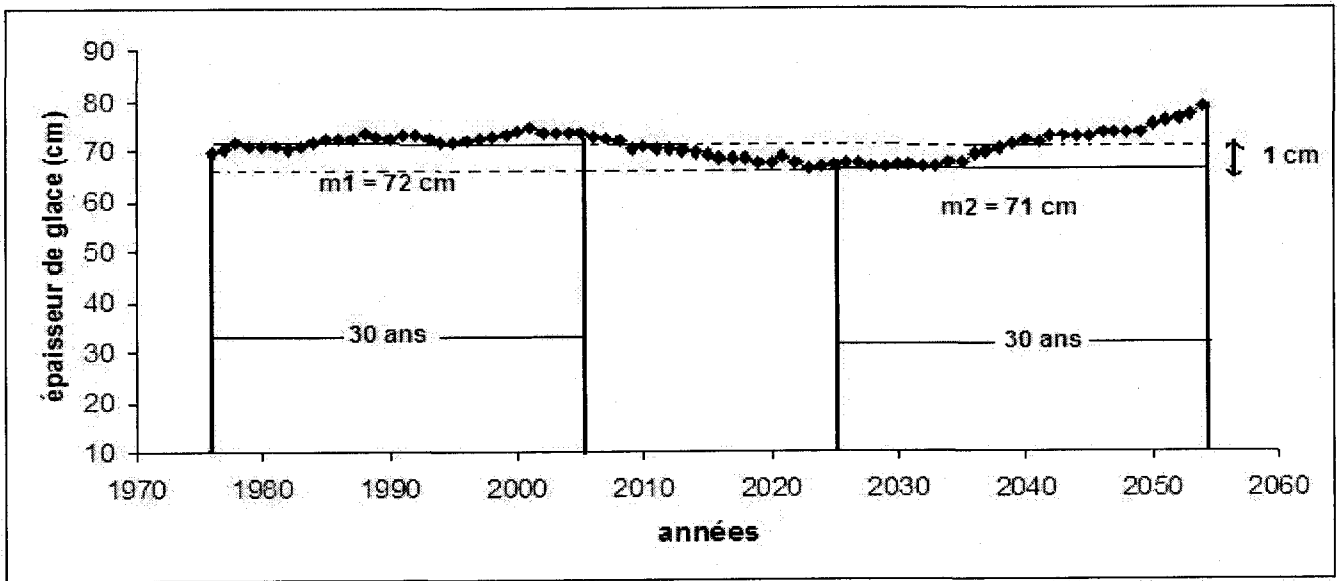


Figure 3.16 : Moyennes mobiles sur les maximums annuels de l'épaisseur de glace calculée par le modèle RNA

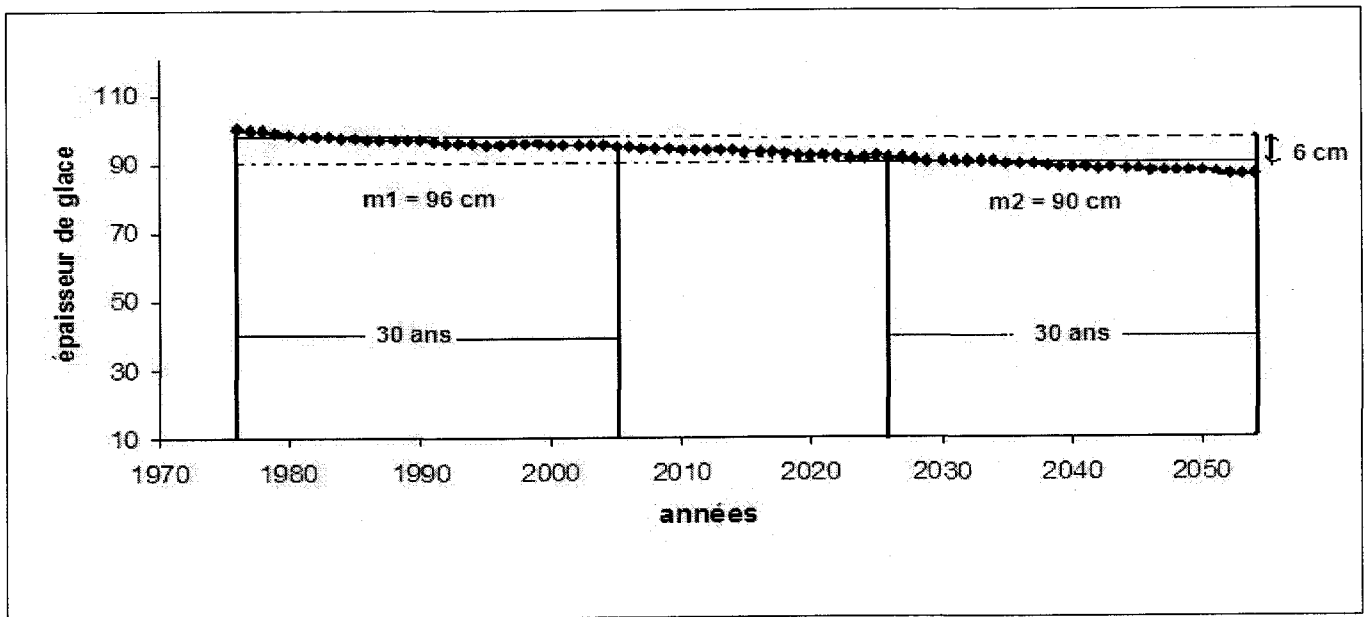


Figure 3.17 : Moyennes mobiles sur les maximums annuels de l'épaisseur de glace calculée par le modèle LSR

Il est important de noter que le modèle RNA présente beaucoup de variations par rapport au modèle LSR. En ce qui concerne le modèle RNA, la figure 3.12 ne présente pas une régularité chronologique dans l'évolution de la glace. Ceci pourrait être confirmé par la figure 3.16 où les moyennes mobiles des maximums annuels de l'épaisseur de glace montrent des fluctuations dans le temps. Cependant, le modèle LSR montre une évolution monotone assez régulière de l'épaisseur de glace. Dans ce contexte, les périodes de trente ans de la figure 3.13 montrent une évolution régressive dans l'ordre chronologie. Cette dégradation régulière pourrait être également observée dans la figure 3.17.

Pour mieux visualiser l'évolution de la glace dans le temps par rapport à la première moyenne mobile, les différences entre les moyennes successives et la première sont calculées. Elles sont représentées dans les figures 3.18 et 3.19 pour le RNA et le LSR respectivement. Ces figures ont presque les mêmes allures que les respectives figures 3.16 et 3.17. Il est important de conclure que pour la période future entre 2010 et 2030, l'épaisseur de glace pourrait être diminuée de 2 cm et 10 cm respectivement pour les modèles RNA et LSR. Ensuite, entre 2030 et 2054, cette diminution pourrait devenir de 6 cm pour le modèle LSR. Par contre, une augmentation probable de 12 cm pour l'épaisseur calculée par le modèle RNA pourrait être observée.

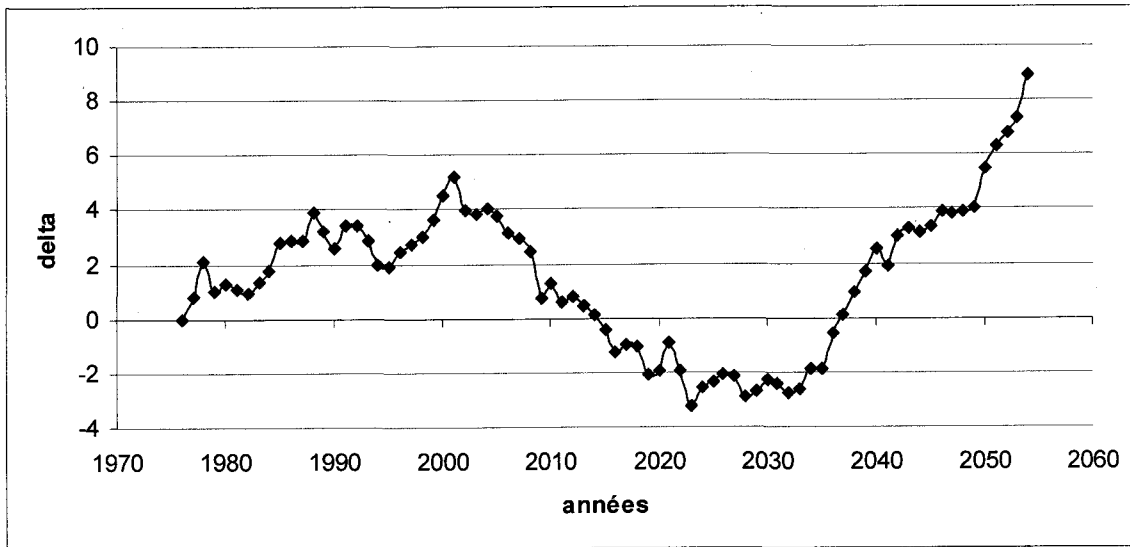


Figure 3.18 : Variation des moyennes mobiles sur les maximums annuels de l'épaisseur de glace calculée par le modèle RNA

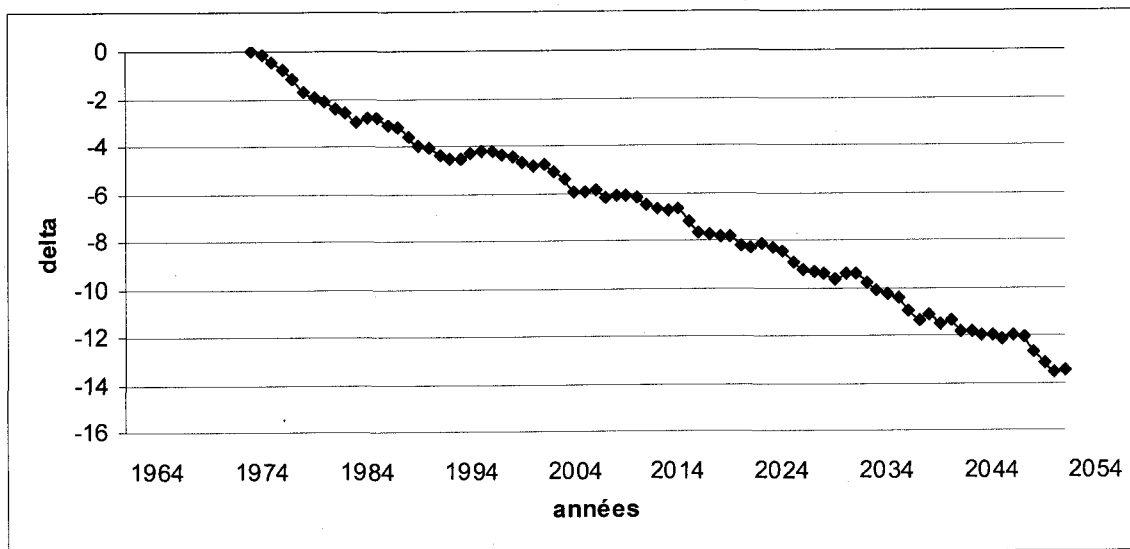


Figure 3.19: Variation des moyennes mobiles sur les maximums annuels de l'épaisseur de glace calculée par le modèle LSR

D'après les figures 3.16 jusqu'à 3.19, malgré que les deux modèles ne présentent pas les mêmes prédictions pour l'épaisseur de glace, il est clair que la situation de la période passée du lac étudié pourrait ne pas ressembler à celle de la période future.

D'après cette constatation, l'évaluation de la glace des lacs dans le nord du Québec pourrait être utilisée comme indicateur de changements climatiques.

Comme la montre les figures 3.17 et 3.19, une tendance de monotonie peut être observée pour les épaisseurs de glace calculées par LSR. Afin de détecter cette tendance dans le temps, le test de Mann-Kendall (Mann, 1945 et Kendall, 1975) a été appliqué pour les moyennes mobiles des maximums des deux modèles.

Soit $X = \{x_1, x_2, \dots, x_n\}$ l'échantillon de taille n qui représente les moyennes mobiles des maximums de l'épaisseur de glace calculées en utilisant les modèles RNA ou LSR de 1961 à 2070. Le test statistique non-paramétrique de Mann-Kendall est défini comme suit :

Pour tester les hypothèses :

H_0 : il n'a y pas de tendance monotone,

H_1 : il y a une tendance monotone,

la statistique de test est

$$Z = \begin{cases} \frac{S-1}{[\text{var}(S)]^{1/2}} & \text{si } S > 0 \\ 0 & \text{si } S = 0 \\ \frac{S+1}{[\text{var}(S)]^{1/2}} & \text{si } S < 0 \end{cases} \quad (3.9)$$

avec

$$S = \sum_{k=1}^{n-1} \sum_{j=k+1}^n \text{sgn}(x_j - x_k) \quad (3.10)$$

où

$$\text{sgn}(x) = \begin{cases} +1 & x > 0 \\ 0 & x = 0 \\ -1 & x < 0 \end{cases} \quad (3.11)$$

et la règle de la décision, pour une erreur de première espèce α , est :

si $Z < Z_{\alpha/2}$ alors on accepte H_0 ,

si $Z > Z_{\alpha/2}$ alors on rejette H_0 ,

où $Z_{\alpha/2}$ est le quantile d'ordre $\alpha/2$ de la loi normale. Dans le cas où $\alpha = 0,05$ alors $Z_{\alpha/2} = 1,96$.

Dans la présente étude, pour le cas du modèle RNA, on obtient $Z = 0,11$ donc pour $\alpha = 0,05$, on accepte H_0 . Par contre, dans le cas du modèle LSR, on a $Z = 12,70$ et, donc, on rejette H_0 pour le même α .

3.5. Conclusion

L'étude de la formation de glace dans les lacs et sa croissance dans les pays nordiques est de première importance pour l'économie (infrastructures, ouvrages hydrauliques, transport). Toutefois, étant donné la complexité de ces processus, la température de l'air est souvent utilisée pour quantifier l'effet des conditions de gel sur le comportement de la glace. C'est le cas du modèle LSR qui fait intervenir la somme des degrés jours accumulés, calculés à partir de la température moyenne journalière de l'air. Le modèle RNA fait intervenir d'autres variables comme les précipitations journalières et l'épaisseur journalière de la neige au sol. Celles-ci pourront présenter une combinaison

de variables efficaces pour expliquer les processus physiques de la croissance de la glace dans les lacs.

L'épaisseur de la glace dans le passé a été évaluée par les deux modèles LSR et RNA en utilisant les données climatiques observées sur le réservoir la Grande IV. La comparaison de ces modèles a montré que le modèle RNA est plus performant que celui de LSR. Malgré la meilleure performance du modèle RNA, les résultats du modèle LSR ont été aussi présentés pour la prédiction.

D'une façon générale, les modèles RNA et LSR présente des résultats avec la statistique descriptive directe qui différent au niveau de la variabilité et la dispersion au cours des années. Également, en utilisant les moyennes mobiles sur les maximums annuels de l'épaisseur de glace pour les deux modèles sur toute la période, une grande différence a été observée. En effet, le modèle RNA présente des fluctuations dans le temps au niveau de l'évolution de l'épaisseur de glace. Contrairement, le modèle LSR possède une certaine monotonie descendante, en avançant dans le temps.

En conclusion, les résultats montrent que le changement climatique influe d'une façon directe sur l'épaisseur de glace du lac La Grande IV, puisque la situation passée pourrait ne pas ressembler à la situation future. Il est important de noter que pour l'épaisseur de glace dans le lac étudié, la prédiction par le LSR montre une diminution probable à l'horizon de 2050. Par contre, à ce même horizon, le modèle RNA a montré qu'une augmentation dans l'épaisseur de glace pourrait être observée.

3.6 Références bibliographiques

- Bishop, C. M. (1995) *Neural Networks for Pattern Recognition*, Oxford Univ. Press, New York
- Brochu, R. et Laprise R. (2007) Surface Water and Energy Budgets over the Mississippi and Columbia River Basins as Simulated by Two Generations of the Canadian Regional Climate Model. *Atmos.-Ocean*, 45(1) pp. 19-35
- Brockwell, P. J. et Davis, R.A. (1987) *Time Series: Theory and Methods*, Springer-Verlag.
- Brutsaert, W. (1982) *Evaporation into the Atmosphere, Theory, History and Applications*, Reidel, Dordrecht, The Netherlands
- CCCma (2008) Centre canadien de la modélisation et de l'analyse climatique (<http://www.cccma.ec.gc.ca/>)
- Coulibaly, P., Anctil, F. et Bobée, B. (1999) Prévion hydrologique par réseaux de neurones artificiels : état de l'art. *Canadian journal of civil engineering* (26), pp. 293-304
- Dawson, C.W. et Wilby, R.L. (2001) Hydrological modelling using artificial neural networks. *Progress in Physical Geography* (25) 1 pp. 80-108
- Franck, R. (1958) The Perceptron : probabilistic model for information storage and organization in the brain. *Psychological Review*, 65 pp. 386-408
- Funahashi, K. (1989) On the approximate realization of continuous mappings by neural networks, *Neuml Networka* 2, pp.183-192.
- Hipel, K.W., et McLeod, A.I. (1994) *Time series modelling of water resources and environmental systems*. Elsevier, Hollande.
- Kendall, M.G. (1975) *Rank Correlation Methods*. Griffin, London, UK.
- Lettenmair, DP., et Wood, E.F. (1993) Hydrologic forecasting. *Handbook of hydrology*. Mc Graw-Hill, New york.
- Lock, G. S. H. (1990) *The Growth and Decay of Ice*, Cambridge Univ.Press, New York.
- Maier, R.H. et Dandy, G.C. (2001) Neural network Based modelling of environmental approach. *Mathematical and computer modelling* 33 pp. 669-682
- Mann, H.B. (1945) Nonparametric tests against trend. *Econometrica* 13 pp. 245-259.

- Music, B., et Caya D. (2007) Evaluation of the Hydrological Cycle over the Mississippi River Basin as Simulated by the Canadian Regional Climate Model (CRCM). *J. Hydromet.*, In Press.
- Nakicenovic, N. et Swart, R. (2000) Special Report on Emissions Scenarios, A Report of the Working Group III, *IPCC (Intergovernmental Panel on Climate Change)* pp. 612
- OCDE (1995) Organisation de Coopération et de Développement Économique (www.oecd.org)
- Ryan, P.J. et Harleman, D.R.F. (1973) An analytical and experimental study of transfer cooling pond behaviour. *Report No.161. Ralph M. Parsons Laboratory, Department of civil engineering, Massachusetts Institute of Technology, Cambridge, MA.*
- Rumelhart, D.E., Hinton, E. et Williams, J. (1986) Learning internal representation by error propagation. Parallel distributed processing. *MIT Press, Cambridge, Massachusetts*, 1 pp. 318-362.
- Scinocca, J.F. et McFarlane N.A. (2004) The Variability of Modelled Tropical Precipitation. *J. Atmos. Sci.*, 61 pp. 1993-2015.
- Seidou, O., Ouarda, T. B. M. J., Bilodeau, Hessami, L. M., St-Hilaire, A., et Bruneau P. (2006) Modeling ice growth on Canadian lakes using artificial neural networks. *Water Ressources Research*, vol. 42.
- SRML (2004) Solar Radiation Monitoring Laboratory Solar radiation basics, *report, Univ. of Oreg. Corvallis.* (<http://solardat.uoregon.edu/SolarRadiationBasics.html>).
- WMO (1994) World Meteorological Organization (www.wmo.org)

4. Conclusion générale

Le climat est un facteur important qui affecte la formation, la croissance et le dégel de la glace. La croissance de la glace au cours de la saison d'hiver détermine la quantité et la force de la glace au moment du dégel. Il est important alors de contrôler le régime des glaces afin de réduire les problèmes qui en résultent. Ces problèmes peuvent être traduits par les inondations, la perte de production d'énergie hydroélectrique et la destruction des ouvrages hydrauliques, Plusieurs chercheurs ont orienté leurs travaux vers ce domaine dans différentes régions du monde. Ces travaux ont donné lieu à diverses contributions significatives à ce sujet. Ainsi, les connaissances scientifiques ont été développées pour diminuer les pertes causées par l'évolution de la glace. Ces connaissances ont démontré que plusieurs facteurs interviennent dans la formation de la glace comme la climatologie, la géomorphologie et la physiographie.

Les variables climatologiques, comme les températures moyennes de l'air, la neige, les précipitations contribuent d'une façon directe ou indirecte dans l'évolution de l'épaisseur de glace. D'une part, l'étude du phénomène de la glace a été traitée par une approche déterministe. Cette approche se base sur les lois de la physique comme le transfert de masse et d'énergie. Dans ce cadre, la Loi de Stefan Révisée a servi comme un outil pour la prédiction de l'épaisseur de la glace des lacs. Elle se base sur la somme degrés-jours qui est calculée à partir de la température moyenne journalière de l'air. D'autre part, une approche stochastique, appelée les réseaux de neurones artificiels simple, a été utilisée pour estimer l'épaisseur de la glace des lacs. Dans les travaux de Seidou et al. (2006), ce modèle a été appliqué pour plusieurs lacs canadiens. Par la suite,

dans le cadre de ce travail, une nouvelle approche, désignée par l'ensemble de réseaux de neurones artificiels a été appliquée. L'objectif de cette dernière est d'améliorer les prédictions obtenues par le modèle simple pour réduire l'incertitude de prédiction.

Malgré que les techniques mentionnées promettent de pouvoir améliorer l'incertitude de la prédiction de la glace des lacs, des travaux futurs sont à envisager dans ce domaine. Les caractéristiques physiques d'un lac jouent un rôle important dans le régime des glaces. Alors, plus les variables climatiques, les variables physiographiques et morphologiques des lacs pourront être intégrées dans les modèles afin de donner plus d'informations sur l'évolution de la glace dans un lac donné.