

**Record Number:** 25470  
**Author, Monographic:** Gingras, H.//Ouarda, T. B. M. J.//Ghedira, H.//Bobée, B.  
**Author Role:**  
**Title, Monographic:** Correction du débit en présence d'un effet de glace : développement du logiciel CORRECT  
**Translated Title:**  
**Reprint Status:**  
**Edition:**  
**Author, Subsidiary:**  
**Author Role:**  
**Place of Publication:** Québec  
**Publisher Name:** INRS-Eau, Terre & Environnement  
**Date of Publication:** 2002  
**Original Publication Date:** Août 2002  
**Volume Identification:**  
**Extent of Work:** viii, 48  
**Packaging Method:** pages incluant un annexe  
**Series Editor:**  
**Series Editor Role:**  
**Series Title:** INRS-Eau, Terre & Environnement, rapport de recherche  
**Series Volume ID:** 626  
**Location/URL:**  
**ISBN:** 2-89146-494-X  
**Notes:** Rapport annuel 2002-2003  
**Abstract:** Numéro de rapport et ISBN demandés par Johanne et fournis par Genevieve le 12 août 2002  
**Call Number:** R000626  
**Keywords:** rapport/ ok/ dl

***Correction du débit en présence d'un effet  
de glace : développement du logiciel  
CORRECT***

***Rapport de recherche No R-626***

***Août 2002***

# **CORRECTION DU DÉBIT EN PRÉSENCE D'UN EFFET DE GLACE : DÉVELOPPEMENT DU LOGICIEL *CORRECT***

*Rapport préparé à l'attention de:*

Messieurs Van-Diem Hoang et François Godin

Ministère de l'Environnement du Québec

Direction de l'hydraulique et de l'hydrique

Édifice Marie-Guyart, Aile René-Lévesque, 1<sup>er</sup> étage, case 20

675, boul. René-Lévesque Est

Québec (Québec), G1R 5V7

*par:*

**Hugo Gingras**

**Taha B.M.J. Ouarda**

**Hosni Ghedira**

**Bernard Bobée**

Chaire Hydro-Québec/CRSNG/Alcan en Hydrologie Statistique

Institut National de la Recherche Scientifique, INRS-ETE

2800, rue Einstein, C.P. 7500, Sainte-Foy (Québec) G1V 4C7

Rapport de recherche No 626

ISBN : 2-89146-494-x

Août 2002

## **ÉQUIPE DE RECHERCHE**

Ont participé à la réalisation de cette étude:

**Chaire en Hydrologie Statistique**

**Institut National de la Recherche Scientifique, INRS-ETE**

Hugo Gingras

Taha B.M.J. Ouarda

Hosni Ghedira

Bernard Bobée

**Ministère de l'Environnement du Québec**

Van-Diem Hoang

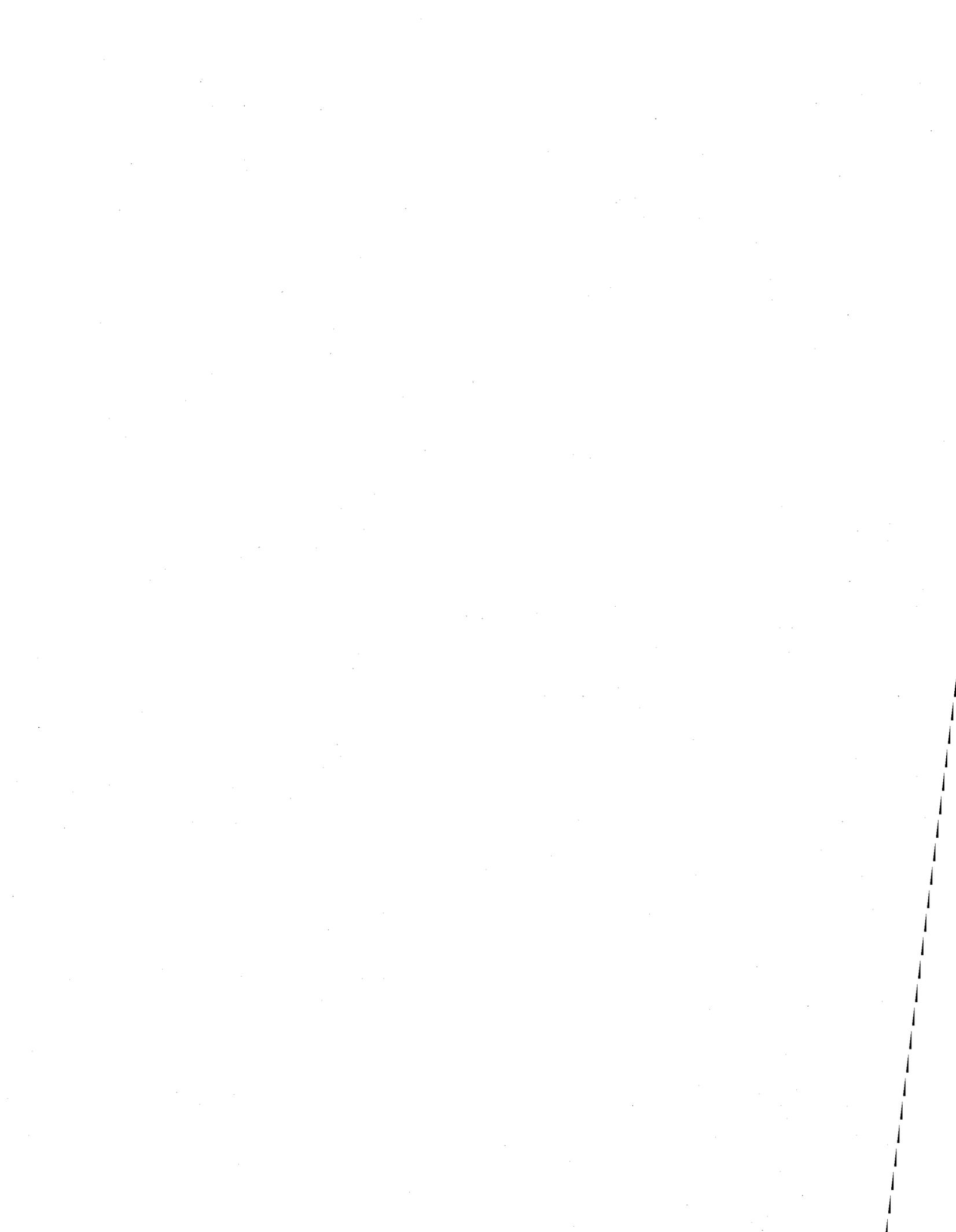
François Godin

Ghislain Chabot

# TABLE DES MATIÈRES

---

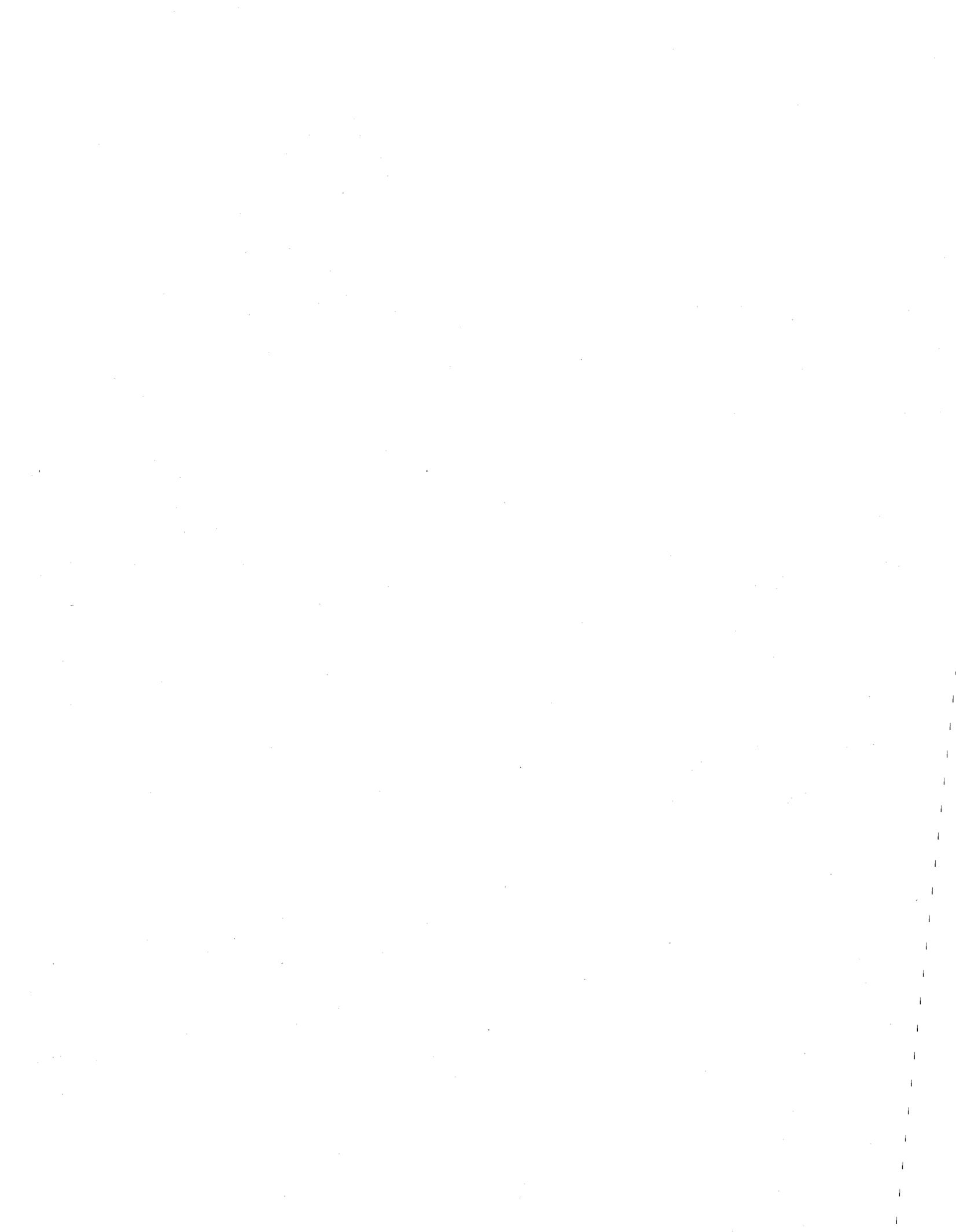
<b>TABLE DES MATIÈRES .....</b>	<b>iii</b>
<b>LISTE DES TABLEAUX .....</b>	<b>v</b>
<b>LISTE DES FIGURES .....</b>	<b>vii</b>
<b>1 INTRODUCTION.....</b>	<b>1</b>
1.1 PROBLÉMATIQUE .....	1
1.2 IMPORTANCE DE LA CORRECTION DU DÉBIT .....	2
1.3 MÉTHODOLOGIE.....	2
<b>2 RÉGRESSION MULTIPLE .....</b>	<b>5</b>
2.1 INTRODUCTION .....	5
2.2 DÉTERMINATION DU MODÈLE OPTIMAL.....	6
<b>3 RÉSEAUX NEURONAUX.....</b>	<b>17</b>
3.1 INTRODUCTION .....	17
3.2 CARACTÉRISTIQUES D'UN RÉSEAU DE NEURONES .....	18
3.3 APPRENTISSAGE DU RÉSEAU .....	20
3.4 VALIDATION DU RÉSEAU .....	23
<b>4 CONCLUSIONS .....</b>	<b>31</b>
<b>RÉFÉRENCES .....</b>	<b>33</b>
<b>ANNEXE : GUIDE DE L'USAGER DU LOGICIEL <i>CORRECT</i>.....</b>	<b>35</b>
A.1 DESCRIPTION ET APERÇU DU LOGICIEL.....	35
A.2 AJOUT D'UNE NOUVELLE STATION.....	37
A.3 SUPPRESSION D'UNE STATION EXISTANTE.....	38
A.4 CALIBRAGE DES MODÈLES RÉGRESSIF ET NEURONAL POUR UNE STATION DONNÉE.....	38
A.5 VISUALISATION GRAPHIQUE DE LA QUALITÉ DE L'AJUSTEMENT .....	43
A.6 CORRECTION AUTOMATIQUE DU DÉBIT .....	44
A.7 FERMETURE DU LOGICIEL <i>CORRECT</i> .....	47



# LISTE DES TABLEAUX

---

<b>Tableau 2.1</b>	Matrice des corrélations entre les variables Q, N, T, T <sub>10</sub> et P <sub>N</sub> .....	7
<b>Tableau 2.2</b>	Sélection du meilleur modèle sur la base du R <sup>2</sup> ajusté (R <sup>2</sup> <sub>a</sub> ).....	8
<b>Tableau 2.3</b>	Comparaison des débits prédits aux débits mesurés par la méthode du "Leave one out" (pour le modèle de régression log(Q) vs. log(N) et log(T)).....	12
<b>Tableau 2.4</b>	Modèles de régression retenus pour l'ensemble des stations de l'étude.....	14
<b>Tableau 3.1</b>	Comparaison des débits prédits aux débits mesurés par la méthode du "Leave one out" (pour le modèle neuronal).....	25



# LISTE DES FIGURES

---

<b>Figure 2.1</b>	Comparaison des débits prédits aux débits mesurés pour (a) les données de calibrage et (b) les données de validation de la rivière Dartmouth .....	11
<b>Figure 2.2</b>	Comparaison de la courbe des débits prédits par régression à la courbe des débits corrigés du MEQ établie pour la rivière Dartmouth (sur la période allant du 24/11/98 au 31/01/99).....	13
<b>Figure 3.1</b>	Architecture d'un réseau multicouches .....	18
<b>Figure 3.2</b>	Connections d'un élément processeur (nœud j).....	19
<b>Figure 3.3</b>	Fonctions d'activation .....	20
<b>Figure 3.4</b>	Évolution de l'erreur au cours de la phase d'apprentissage .....	23
<b>Figure 3.5</b>	Description schématique de la phase d'apprentissage.....	24
<b>Figure 3.6</b>	Comparaison de la courbe des débits prédits par réseau de neurones à la courbe des débits corrigés du MEQ établie pour la rivière Dartmouth (sur la période allant du 24/11/98 au 31/01/99) .....	26
<b>Figure 3.7</b>	Comparaison des courbes de débits prédits par les approches régressive et neuronale à la courbe des débits corrigés du MEQ établie pour la rivière Nouvelle (sur la période allant du 13/12/96 au 19/04/97).....	27
<b>Figure 3.8</b>	Comparaison des courbes de débits prédits par les approches régressive et neuronale à la courbe des débits corrigés du MEQ établie pour la rivière Du Sud (sur la période allant du 26/12/98 au 31/03/99) .....	28
<b>Figure A.1</b>	Interface principale du logiciel <i>CORRECT</i> .....	36
<b>Figure A.2</b>	Fenêtre d'ajout d'une nouvelle station.....	37

<b>Figure A.3</b> Exemple d'un fichier de calibrage .....	39
<b>Figure A.4</b> Fenêtre des paramètres du calibrage .....	40
<b>Figure A.5</b> Graphique d'ajustement du modèle régressif .....	42
<b>Figure A.6</b> Graphique d'ajustement du modèle neuronal .....	42
<b>Figure A.7</b> Exemple d'un fichier d'intrants pour la procédure de correction du débit.....	45
<b>Figure A.8</b> Fenêtre d'affichage des débits corrigés.....	46

# 1 INTRODUCTION

---

## 1.1 Problématique

Pour assurer une bonne gestion des ressources hydriques, et un développement durable dans les secteurs associés à l'eau, il est nécessaire de posséder une base de données hydrométriques fiable et de bonne qualité. Cependant, les séries de débits de rivières correspondant à la période hivernale sont souvent de qualité inférieure à celle correspondant au reste de l'année. En effet, une proportion importante des rivières québécoises et canadiennes sont affectées par l'effet de glace ; i.e. le débit estimé par la courbe de tarage ne correspond pas au débit réel dans la rivière à cause de la présence de glace dans la rivière (glace de surface, glace de fond, glace en aiguilles, etc.). La courbe de tarage définit la relation entre le niveau d'eau dans une section transversale et le débit correspondant. En général, elle est construite dans une section stable de la rivière à partir de plusieurs observations niveau-débit. Idéalement, ces observations doivent inclure des valeurs extrêmes pour assurer la bonne extrapolation de la courbe. Cependant cette courbe de tarage ne peut pas être représentative de la période hivernale à cause du changement de la section d'écoulement même et des conditions très variables qui peuvent exister quand l'écoulement est affecté par la présence de glace.

Pour remédier à cette situation, le Ministère de l'Environnement et de la Faune du Québec (MEQ) et Environnement Canada (EC) effectuent généralement des jaugeages durant la période de présence de glace pour estimer le débit réel qui s'écoule dans les rivières. Ensuite, les débits pendant le reste de la période hivernale sont corrigés par interpolation tout en tenant compte des événements pluvieux ou des réchauffements de température qui peuvent avoir eu lieu ainsi que du comportement hydrologique des autres rivières de la région. Cette approche mène généralement à des résultats satisfaisants mais risque d'introduire des erreurs assez importantes lors de fonte hivernale, embâcles de glace, etc. En effet, l'approche adoptée présentement consiste à interpoler les deux ou trois jaugeages d'hiver pour chaque station, pour reconstituer les débits de toute la période hivernale.

Cette approche est caractérisée par sa subjectivité et sa non reproductibilité. En plus cette approche n'est appliquée que durant le printemps, et donc ne permet pas d'estimer les débits de rivières sur une base journalière durant la période hivernale.

## 1.2 Importance de la correction du débit

L'importance d'effectuer une bonne correction du débit en présence d'un effet de glace est illustrée par l'impact économique des estimations de débits durant la période hivernale. En effet, le débit durant la période hivernale est de plusieurs ordres de grandeur inférieur à celui de la période printanière et estivale. Une erreur dans l'estimation du débit hivernal risque alors d'être critique et d'avoir des conséquences fâcheuses pour la faune aquatique qui devient extrêmement vulnérable durant cette période. De plus, la dilution et la dispersion des effluents acquièrent une importance encore plus grande durant cette période à cause des faibles débits dans les rivières.

D'autre part, les embâcles et les débâcles de glace sont directement reliés à la formation et à la fonte de la glace dans les rivières, et à la valeur du débit durant les périodes critiques. Les crues causées par les embâcles de glace causent plus de 60 millions de dollars en dommages chaque année au Canada. Tout ceci a amené Ouarda *et al.* (2000) à effectuer une étude critique de l'approche adoptée présentement pour l'estimation des débits en période de présence de glace et à proposer le développement d'une approche efficace, objective et reproductible.

## 1.3 Méthodologie

L'étude menée par Ouarda *et al.* (2000) offrait une revue complète des différentes méthodes utilisées pour la correction du débit en présence de glace. Elles étaient classées essentiellement en deux catégories, soit les méthodes subjectives (i.e. qui font appel à une certaine forme de jugement de la part de l'utilisateur) et les méthodes analytiques (i.e. indépendantes du jugement de l'utilisateur). Il va sans dire que pour l'étude de cas présentée ici, on s'intéresse aux méthodes analytiques, plus particulièrement à la régression

---

multiple (chapitre 2) ainsi qu'aux modèles de réseaux de neurones artificiels (RNA) (chapitre 3). L'objectif de ce travail consiste donc à tester et comparer ces deux méthodes sur des données en provenance de stations gérées par le MEQ. On y présente aussi en annexe un logiciel de correction automatique du débit (dénommé *CORRECT*) dans lequel ces deux types de modèles ont été implantés.



## 2 RÉGRESSION MULTIPLE

---

### 2.1 Introduction

De façon générale, la régression multiple permet de construire un modèle reliant la variable réponse à certaines variables explicatives à partir des données historiques. Le modèle alors construit peut être utilisé pour faire une interpolation afin d'obtenir des valeurs non mesurées, mais il peut aussi être utilisé dans une optique prédictive, c'est à dire qu'il permet, pour une autre période que celle qui a servi à calibrer le modèle, d'utiliser l'information mesurée des variables explicatives pour estimer la valeur de la variable réponse. La méthode consiste donc à quantifier la corrélation entre les diverses variables explicatives et la variable réponse, pour la construction du modèle prédictif. Ce modèle peut donc s'exprimer ainsi :

$$Y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \dots + \beta_p X_p + \varepsilon$$

où  $Y$  est la variable réponse,  $\{\beta_0, \beta_1, \beta_2, \dots, \beta_p\}$  sont les paramètres de la régression à estimer,  $\{X_1, X_2, \dots, X_p\}$  sont les variables explicatives et  $\varepsilon$  est l'erreur du modèle. Dans le cas présent, la variable réponse est le débit sous la glace, tandis que les variables explicatives considérées sont le niveau d'eau et la température moyenne journalière, en plus de la neige au sol.

De façon classique, le modèle de régression est calibré à partir des données historiques à l'aide de la méthode des moindres carrés qui consiste à trouver la combinaison de valeurs pour chacun des paramètres  $\beta_i$  qui permet de minimiser l'erreur du modèle  $\varepsilon$ . Une estimation de l'erreur du modèle nous permet de juger de la corrélation entre les variables explicatives et la variable réponse. Si la corrélation est faible, on peut dire que la ou les variables explicatives ne permettent pas d'expliquer suffisamment la variabilité de la variable réponse et que le modèle ne peut être utilisé dans un contexte prédictif. Dans le cas d'une régression linéaire simple, c'est à dire une régression du premier ordre n'impliquant qu'une seule variable explicative  $X$  et la variable réponse  $Y$ , on estime les

paramètres  $\beta_0$  et  $\beta_1$  ainsi que l'erreur du modèle et on peut directement juger de la corrélation entre les deux variables. Par contre, dans le cas d'une régression multiple, l'erreur globale du modèle est composée de l'erreur induite par chacune des variables explicatives et la corrélation globale peut être influencée par la corrélation pouvant exister entre certaines variables explicatives. Ainsi, à cause de la possible corrélation entre les variables explicatives, on pourrait être amené à éliminer quelques-unes des variables considérées au départ afin de ne conserver dans le modèle que des variables peu corrélées entre elles, mais qui sont en revanche fortement corrélées à la variable réponse, de façon à minimiser l'erreur.

## 2.2 Détermination du modèle optimal

L'exemple retenu pour illustrer cette technique est le cas de la rivière Dartmouth, pour laquelle on dispose des données de jaugeages (i.e. les niveaux et les débits associés) allant de 1971 à 2001 (environ deux ou trois mesures par hiver). Cependant, les données météo disponibles (tirées de stations météorologiques avoisinant la station hydrométrique) ne couvrent que la période 1992-1999. Puisqu'on ne peut utiliser que les données météo recueillies lors des journées correspondant aux mesures de jaugeages, le nombre d'observations disponibles pour construire le modèle est donc peu élevé, soit 14. Idéalement, comme dans le cas de l'étude de Ouarda *et al.* (2000), il aurait fallu disposer de données de jaugeages pour une année complète recueillies lors d'une campagne de terrain, permettant ainsi d'utiliser une moitié de l'échantillon pour calibrer le modèle et l'autre moitié pour le valider. On a tenté malgré tout ici quelque chose de similaire en utilisant 10 observations (tirées au hasard dans l'échantillon complet) pour calibrer le modèle et les quatre observations restantes pour le valider. Il est clair toutefois que l'évaluation de la performance du modèle ainsi obtenu n'en demeure que très sommaire.

Voici la notation employée pour désigner la variable réponse et les variables explicatives:

Q : débit ( $\text{m}^3/\text{s}$ )

N : Niveau d'eau (m)

$T$  : Température moyenne journalière (°C) (donnée en fait par la moyenne du minimum et du maximum enregistrés pour la journée)

$T_{10}$  : Température moyenne des 10 jours précédents (°C)

$P_N$  : Hauteur de neige au sol (mm)

On a ajouté la variable  $T_{10}$  puisqu'elle a été retenue dans le modèle de régression développé dans Ouarda *et al.* (2000). Il sera donc intéressant de voir si cette même variable pourra s'avérer assez importante pour être retenue ici aussi. Une autre variable qui aurait pu être considérée est la précipitation liquide, quoique celle-ci soit généralement nulle pour les dates de jaugeages qui se situent au beau milieu de l'hiver. Le Tableau 2.1 présente la matrice des corrélations pour les différentes variables. On constate d'emblée que le niveau  $N$  est la variable la plus corrélée au débit  $Q$ . De plus, les corrélations plus ou moins élevées existant entre les variables explicatives permet déjà de croire qu'il ne sera pas nécessaire de toutes les inclure dans le modèle de régression.

**Tableau 2.1** Matrice des corrélations entre les variables  $Q$ ,  $N$ ,  $T$ ,  $T_{10}$  et  $P_N$

	$Q$	$N$	$T$	$T_{10}$	$P_N$
$Q$	1	<b>0.9481</b>	0.0386	0.3149	-0.3606
$N$	<b>0.9481</b>	1	-0.1051	0.3364	-0.4271
$T$	0.0386	-0.1051	1	0.3244	0.6187
$T_{10}$	0.3149	0.3364	0.3244	1	0.3955
$P_N$	-0.3606	-0.4271	0.6187	0.3955	1

Avant de pouvoir valider le modèle, il importe tout d'abord de faire la sélection des variables explicatives qui seront retenues. Une façon de procéder consiste à considérer toutes les combinaisons de variables explicatives possibles et de calculer pour chaque cas la statistique  $R^2$  qui mesure la proportion de la variabilité totale expliquée par le modèle (le reste étant attribué à l'erreur). Dans le cas où l'on a qu'une seule variable explicative, le  $R^2$  correspond alors au carré du coefficient de corrélation entre la variable et la réponse. Le  $R^2$  a un défaut cependant ; il augmente automatiquement dès qu'on ajoute une variable

explicative au modèle. C'est pourquoi on utilise en pratique le  $R^2$  ajusté, que l'on dénote  $R_a^2$  et que l'on définit comme suit :  $R_a^2 \equiv 1 - \{[(n-1) / (n-p)] \times (1-R^2)\}$ , où  $n$  est le nombre d'observations et  $p$  représente le nombre de prédicteurs (i.e. les variables explicatives). Dans le cas d'un seul prédicteur, on note que  $R_a^2 = R^2$ . On retiendra en bout de ligne le modèle pour lequel la valeur du  $R_a^2$  est la plus grande.

**Tableau 2.2** Sélection du meilleur modèle sur la base du  $R^2$  ajusté ( $R_a^2$ )

Prédicteur(s)	$R_a^2$ (espace réel)	$R_a^2$ (espace logarithmique)
N	0.8989	0.9727
T	0.0015	0.0243
T <sub>10</sub>	0.0991	0.1224
P <sub>N</sub>	0.1301	0.3090
N, T	<b>0.9114</b>	<b>0.9732</b>
N, T <sub>10</sub>	0.8905	0.9707
N, P <sub>N</sub>	0.8930	0.9704
T, T <sub>10</sub>	0.0290	0.1358
T, P <sub>N</sub>	0.1778	0.3160
T <sub>10</sub> , P <sub>N</sub>	0.3264	0.4383
N, T, T <sub>10</sub>	0.9079	0.9726
N, T, P <sub>N</sub>	0.9065	0.9727
N, T <sub>10</sub> , P <sub>N</sub>	0.8858	0.9682
T, T <sub>10</sub> , P <sub>N</sub>	0.3576	0.4034
N, T, T <sub>10</sub> , P <sub>N</sub>	0.8993	0.9714

Les résultats de cet exercice sont reportés au Tableau 2.2. La première colonne indique les prédicteurs considérés tandis que les valeurs de la statistique  $R_a^2$  pour chaque modèle sont reportées dans la deuxième colonne. La troisième colonne donne quant à elle la valeur du

$R^2_a$  dans l'espace logarithmique, i.e. lorsqu'on construit le modèle après avoir effectué au préalable une transformation logarithmique sur les données (le modèle prenant par exemple la forme  $\log(Q) = \beta_0 + \beta_1 \log(N) + \beta_2 \log(T)$ , en ne considérant que les prédicteurs N et T). Le choix de la transformation logarithmique vient du fait que la relation niveau-débit s'exprime généralement mieux en coordonnées logarithmiques.

Il ressort clairement de l'analyse du Tableau 2.2 que le niveau N est la variable la plus importante pour expliquer le débit, ce qui ne surprend guère en raison de la forte corrélation entre ces deux variables (voir Tableau 2.1). En effet, on observe que le  $R^2_a$  est déjà presque égal à 0.9 (i.e. le carré du coefficient de corrélation entre Q et N) avec cette variable seulement dans le modèle. La contribution des autres prédicteurs au modèle est donc moins prépondérante et, tel qu'anticipé, le meilleur modèle sur la base du  $R^2$  ajusté élimine quelques unes des variables de départ, soit  $T_{10}$  et  $P_N$ , pour ne retenir que les variables N et T, et ce autant dans l'espace réel que logarithmique. Étant donné que l'ajustement est meilleur en coordonnées logarithmiques, on privilégiera cette option.

Il existe néanmoins d'autres méthodes permettant de sélectionner un modèle optimal. L'une d'entre elles est la procédure "stepwise" (voir, e.g., Weisberg, 1985). Pour en faire un bref résumé, cette procédure consiste à sélectionner les variables en tenant compte de l'importance de leur contribution au modèle de régression. Ainsi, à la première étape, on ajoute tout d'abord la variable explicative la plus corrélée à la variable réponse. À l'étape suivante, on cherche parmi les prédicteurs restants celui qui améliore le modèle de la façon la plus significative, ce que l'on détermine à l'aide d'un test statistique (dit de Student) qui mesure l'apport marginal de la nouvelle variable. Le prédicteur pour lequel la valeur de la statistique du test est la plus élevée est alors considéré, à la condition que cette valeur dépasse une certaine limite fixée au départ par le choix du seuil  $\alpha$  d'application du test. Chaque nouvelle variable ajoutée au modèle doit donc respecter cette condition. De plus, en vérifiant que cette contrainte soit toujours respectée à chaque étape de la procédure, il est même possible qu'un prédicteur soit exclu du modèle bien qu'il ait été ajouté à une étape précédente. On arrête la procédure lorsqu'il ne reste aucune autre variable ne satisfaisant à la condition d'entrée imposée.

Dans le cas présent, l'application de la procédure "stepwise" avec un seuil  $\alpha = 0.05$  (i.e. pour un niveau de confiance à  $(1-\alpha)\% = 95\%$ ) ne retient que le niveau N dans le modèle, que l'on se place dans l'espace réel ou bien logarithmique. L'utilisation de seuils plus permissifs (tels 0.15 dans l'espace réel et 0.35 en coordonnées logarithmiques) provoque l'ajout de la température T au modèle. On constate ainsi que le choix du modèle à l'aide de cette méthode n'est pas systématique puisqu'il dépend du seuil fixé par l'utilisateur. En fait, l'utilité de la procédure est mieux mise en valeur lorsque l'on doit faire un choix parmi un grand nombre de variables explicatives, dans le but d'obtenir un modèle à la fois efficace et parcimonieux.

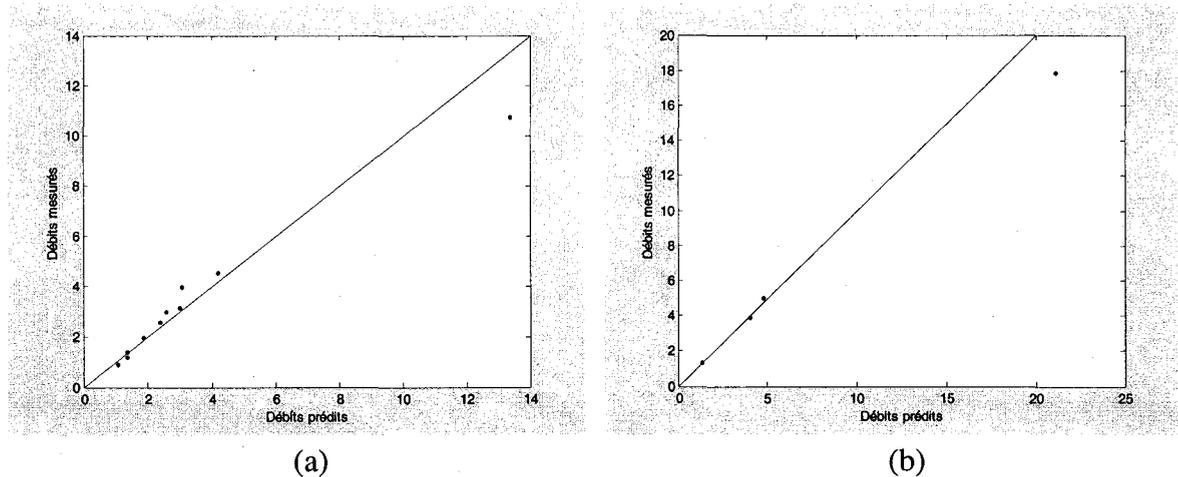
La méthode se basant sur le  $R^2$  ajusté pour le choix du modèle apparaît donc comme étant mieux appropriée pour ce cas d'étude. Il demeure néanmoins que les résultats de la "stepwise" rejoignent dans une certaine mesure la sélection des variables N et T faite sur la base du  $R^2$  ajusté. Il est bon de noter également que l'ajout de la variable T au modèle n'est pas vraiment superflu considérant que les variables N et T sont les variables explicatives les moins corrélées (voir Tableau 2.1). Enfin, comme on l'a mentionné plus tôt, il est plus intéressant de développer ce modèle dans l'espace logarithmique en raison de la meilleure qualité de l'ajustement. Donc, après avoir estimé par la méthode des moindres carrés les paramètres  $\beta_0$ ,  $\beta_1$ , et  $\beta_2$ , le modèle retenu prend la forme suivante:

$$\log(Q) = 149.33 \log(N) - 2.38 \log(T) - 489.52$$

où la température moyenne journalière T est exprimée en degrés Kelvins (le logarithme n'étant défini que pour des valeurs positives). Dans un contexte prédictif, un utilisateur obtient alors le débit journalier en remplaçant dans l'équation précédente les valeurs de N et de T mesurées pour la journée en question.

Afin de valider ce modèle de régression, on peut considérer la possibilité de séparer l'échantillon en deux, i.e qu'on calibre tout d'abord le modèle uniquement à partir d'un sous-échantillon aléatoire de 10 observations provenant de l'échantillon complet, les quatre observations non utilisées étant par la suite comparées aux prédictions données par le modèle calé sur les 10 autres points. La Figure 2.1 présente les graphiques des débits

prédits vs les débits mesurés pour (a) l'échantillon de calibrage et (b) l'échantillon de validation. C'est l'alignement des points sur ces graphiques qui permet de juger de la qualité du modèle. On constate donc que celle-ci semble plutôt bonne, excepté pour les valeurs extrêmes qui sont surestimées dans ce modèle. Il est difficile toutefois d'en dire plus étant donné le nombre limité d'observations pour ce cas-ci.



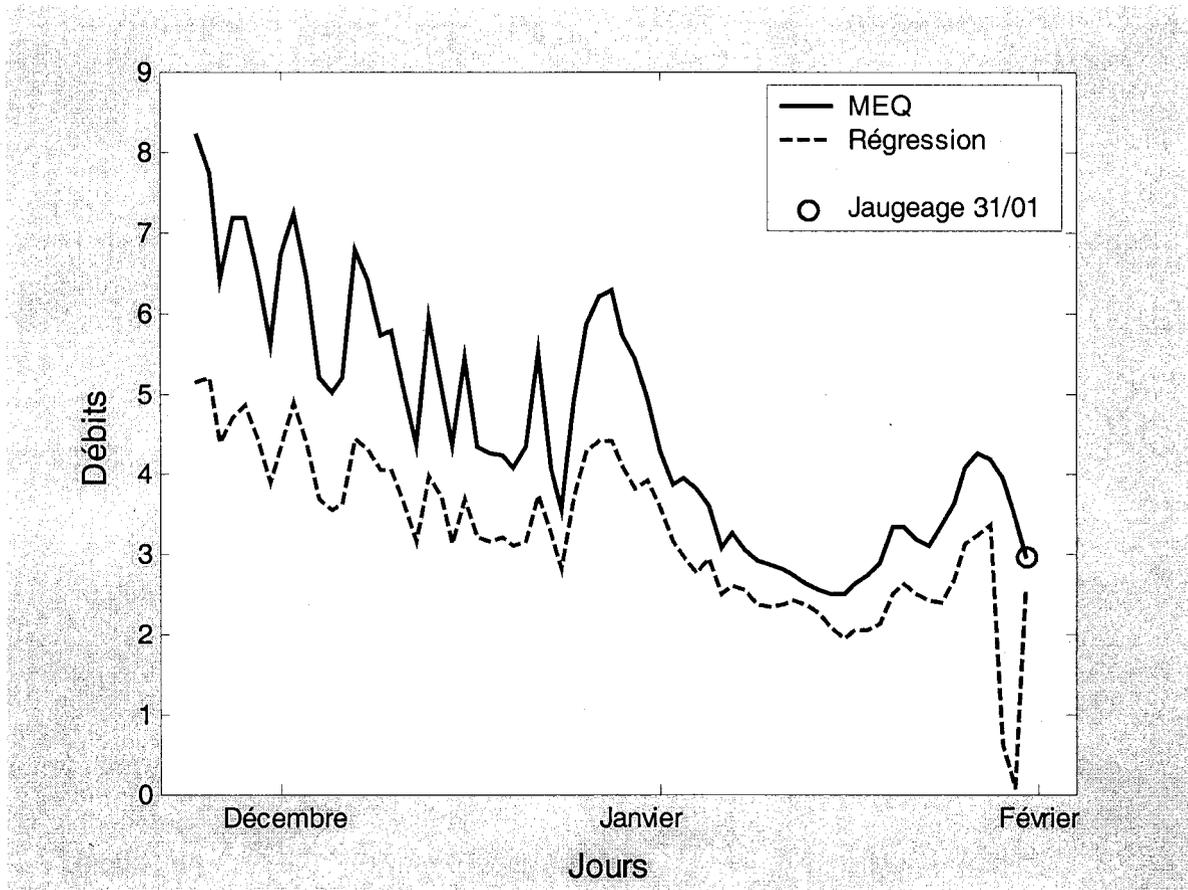
**Figure 2.1** Comparaison des débits prédits aux débits mesurés pour (a) les données de calibrage et (b) les données de validation de la rivière Dartmouth

Une façon sans doute plus adéquate de valider le modèle de régression, vu le petit nombre d'observations dont on dispose, est d'utiliser la méthode du "Leave one out". Celle-ci consiste à calibrer le modèle sur l'ensemble des observations excepté une, puis d'examiner la prédiction obtenue pour cette observation à l'aide du modèle tout juste calibré. En répétant cette procédure pour chaque observation de l'échantillon, on obtient alors la prédiction du modèle pour chacun des débits mesurés, ce qui permet de juger de la performance de celui-ci. Pour l'exemple de la rivière Dartmouth, on a reporté les résultats fournis par cette méthode au Tableau 2.3. On en tire le même constat que pour l'analyse de la Figure 1, soit que l'ajustement semble correct sauf dans le cas des grandes valeurs (i.e. les observation 8 et 13) qui sont surestimées. L'erreur relative pour chaque prédiction a été reportée dans la dernière colonne du Tableau 2.3. Celle-ci dépasse occasionnellement les 20%, bien qu'en moyenne elle se situe à 13.9%.

**Tableau 2.3** Comparaison des débits prédits aux débits mesurés par la méthode du "Leave one out" (pour le modèle de régression  $\log(Q)$  vs.  $\log(N)$  et  $\log(T)$ )

Observation	Débit mesuré	Débit prédit	Erreur relative (%)
1	3.86	4.15	7.52
2	2.57	2.37	8.07
3	1.16	1.42	22.30
4	0.88	1.17	33.77
5	3.95	2.96	25.05
6	3.12	2.93	6.33
7	1.96	1.85	5.51
8	10.76	13.50	25.55
9	5.01	4.63	7.55
10	4.53	3.98	12.23
11	1.36	1.38	1.17
12	1.34	1.40	4.44
13	17.88	21.66	21.13
14	2.96	2.55	13.94
<b>Erreur relative moyenne :</b>			<b>13.90</b>

Enfin, il serait intéressant d'étudier l'évolution des débits prédits par la courbe de régression sur une période continue et de la comparer à la courbe de débits corrigés établie par le MEQ pour la même période. Cet exercice a été effectué pour la période allant du 24 novembre 1998 au 31 janvier 1999 (les données disponibles pour la rivière Dartmouth ne nous permettant pas d'aller plus loin) et le résultat est illustré à la Figure 2.2. On y remarque donc que bien que les deux courbes soient quelque peu décalées, leurs fluctuations quant à elles s'avèrent similaires.



**Figure 2.2** Comparaison de la courbe des débits prédits par régression à la courbe des débits corrigés du MEQ établie pour la rivière Dartmouth (sur la période allant du 24/11/98 au 31/01/99)

Le cas de la rivière Dartmouth, avec ses 14 observations, s'avère donc être le meilleur parmi tous ceux mis à notre disposition pour ce travail. Toutefois, cela n'empêche en rien de répéter la démarche établie avec Dartmouth afin de déterminer les équations de régression pour les autres stations de l'étude. Le Tableau 2.4 résume le tout en présentant pour chaque station le modèle de régression retenu, la valeur du  $R^2$  ajusté obtenu (avec le nombre d'observations disponibles pour calibrer le modèle indiqué entre parenthèses) ainsi que l'erreur relative moyenne associée (telle qu'obtenue par la méthode du "Leave one out"). Par contre, ces équations n'ont pas été déterminées pour les stations dont le nombre d'observations est inférieur à 5 (ce qui est rapporté dans le Tableau 2.4 par la mention "Équation non établie").

**Tableau 2.4** Modèles de régression retenus pour l'ensemble des stations de l'étude

Numéro de la station	Nom de la station	Modèle régressif optimal	R <sup>2</sup> ajusté	Erreur relative moyenne
011201	Nouvelle	$Q = 12.27 N - 334.41$	0.7604 (12)	35.05%
020602	Dartmouth	$\log(Q) = 149.33 \log(N) - 2.38 \log(T) - 489.52$	0.9732 (14)	13.90%
023106	Du Sud	$Q = 9.72 N + 0.16 T - 269.13$	0.8101 (8)	41.38%
023402	Chaudière	$Q = 114.85 N - 2420.67$	0.9134 (7)	50.91%
030103	Nicolet	$\log(Q) = 86.15 \log(N) - 277.13$	0.9208 (6)	19.35%
030234	Eaton	<i>Équation non établie</i>	–	–
040129	Doncaster	$\log(Q) = 117.84 \log(N) - 397.53$	0.9897 (10)	5.53%
041902	Du Moine	<i>Équation non établie</i>	–	–
043012	Kinojevis	$\log(Q) = 65.88 \log(N) + 0.51 \log(T_{10}) - 218.16$	0.9814 (8)	4.96%
050409	Sainte-Anne (bras du Nord)	$\log(Q) = 89.15 \log(N) + 2.61 \log(T_{10}) - 310.33$	0.9680 (10)	9.77%
050801	Jacques-Cartier	$Q = 24.32 N - 0.37 T + 1.03 T_{10} - 676.00$	0.7745 (11)	19.88%
051001	Montmorency	$\log(Q) = 62.80 \log(N) + 8.15 \log(T_{10}) - 249.01$	0.7121 (10)	20.99%
061020	Aux Écorces	$Q = 28.61 N - 0.12 T - 789.56$	0.9560 (8)	16.89%
062102	Mistassini	$Q = 66.57 N - 1732.80$	0.8461 (9)	15.24%
072301	Moisie	$Q = 100.71 N - 0.74 T + 3.30 T_{10} - 2499.48$	0.9570 (5)	47.10%
073801	Romaine	$Q = 126.39 N - 20315.80$	0.7641 (5)	26.61%
080101	Harricana	<i>Équation non établie</i>	–	–
093801	Grande rivière de la Baleine	$Q = 97.89 N - 1.45 T + 3.80 T_{10} - 2569.40$	0.9986 (5)	2.54%

Il va sans dire qu'une réévaluation du choix des variables explicatives s'imposera dans l'éventualité où l'on disposera d'une base de données plus complète. Peut-être alors que le rôle des prédicteurs laissés de côté dans cet exemple s'avérera plus important.



## 3 RÉSEAUX NEURONAUX

---

### 3.1 Introduction

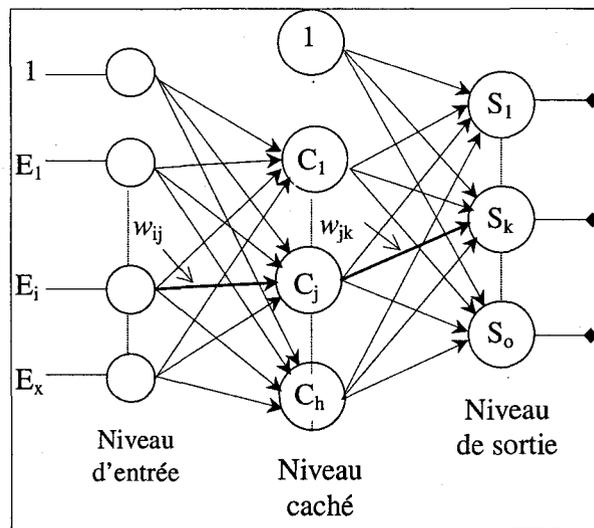
Les réseaux de neurones sont constitués d'un ensemble de neurones artificiels ou nœuds qui sont analogues aux neurones biologiques. Ils sont issus d'une tentative de conception d'un modèle mathématique très simplifié du cerveau humain en se basant sur notre façon d'apprendre et de corriger nos erreurs. Les premiers travaux sur les réseaux de neurones ont été réalisés en 1943 par Mc Culloch et Pitts. Ces deux chercheurs sont les premiers à montrer théoriquement que des réseaux de neurones formels et simples peuvent réaliser des fonctions logiques, arithmétiques et symboliques. Ils ont ainsi présenté un modèle assez simple pour les neurones permettant d'explorer différentes possibilités d'applications.

L'évolution phénoménale des outils informatiques a largement contribué au développement des réseaux de neurones. Les réseaux de neurones font actuellement l'objet de beaucoup de recherches, en raison de leurs propriétés intéressantes d'apprentissage de modèles non linéaires et leurs possibilités d'application à des problèmes de classification, de diagnostic, de prédiction et de contrôle de procédés. En plus, un réseau de neurones permet d'optimiser la meilleure approximation non linéaire basée sur la structure complexe du réseau, et ceci sans aucune contrainte sur la linéarité ou sur la non linéarité spécifiée a priori comme dans les méthodes usuelles de régression.

Il existe plusieurs types de réseaux de neurones tels que les "perceptrons", les réseaux à fonctions de base radiales et les réseaux récurrents. Parmi eux, les perceptrons à alimentation directe (*feed-forward*) et entraînés par rétropropagation (*backpropagation*) ont eu un succès important dans plusieurs applications. Leur intérêt provient de la simplicité de leur utilisation ainsi que de la rapidité et de l'efficacité de leur algorithme de rétropropagation. Cet algorithme d'apprentissage a été proposé par Werbos (1974) et diffusé par Rumelhart *et al.* (1986).

### 3.2 Caractéristiques d'un réseau de neurones

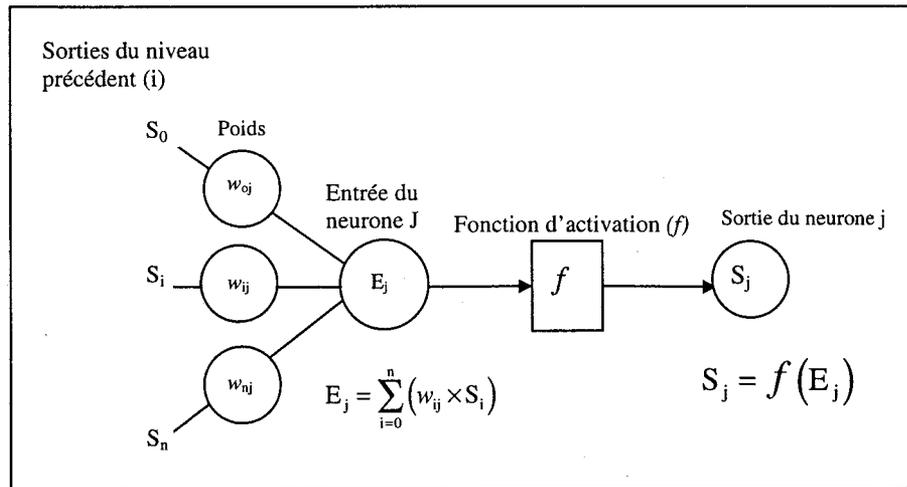
Le nombre de niveaux cachés et le nombre de neurones par niveau représentent les paramètres de l'architecture d'un réseau de neurones (Figure 3.1). La valeur de ces paramètres dépend principalement de la quantité et de la complexité des données. Cependant, une architecture qui donne de bons résultats pour une application donnée ne peut être déterminée que d'une façon expérimentale. En outre, une architecture optimale trouvée pour une application spécifique ne garantit pas des résultats similaires dans d'autres applications. Toutefois, un nombre élevé de neurones dans les niveaux intermédiaires augmente le temps de calcul et diminue la généralisation du réseau, d'où la nécessité de trouver le meilleur compromis possible entre le nombre de niveaux et de neurones cachés.



**Figure 3.1** Architecture d'un réseau multicouches

Les nœuds sont considérés comme éléments processeurs d'un réseau de neurones, chaque nœud permettant la transformation de l'information contenue dans les entrées ( $E_j$ ) par une fonction non linéaire dite d'activation (Figure 3.2). La valeur de la sortie d'un neurone quelconque ( $j$ ) est calculée à partir des entrées qu'il reçoit, ces entrées correspondant aux sorties de la couche précédente.

La réponse d'un neurone dépend des entrées qu'il reçoit, les entrées d'un neurone étant données par les sorties des neurones des couches précédentes pondérées par un facteur de poids ( $w$ ) qui caractérise le lien entre deux neurones. La configuration et le fonctionnement de base pour chaque neurone intermédiaire sont présentés à la Figure 3.2.



**Figure 3.2** Connexions d'un élément processeur (nœud j)

La méthode utilisée pour le transfert de l'information entre deux neurones  $i$  et  $j$  appartenants à deux couches successives est basée sur les trois équations suivantes:

$$S_j = f(E_j)$$

$$E_j = \sum_{i=0}^n (w_{ij} \times S_i)$$

$$f(x) = \tanh(x) \approx \frac{e^x - e^{-x}}{e^x + e^{-x}}$$

où

$S_i$  : La valeur à la sortie du neurone  $i$

$S_j$  : La valeur à la sortie du neurone  $j$

$f$  : La fonction d'activation (Exemple : tangente hyperbolique)

$w_{ij}$  : Le coefficient de pondération (poids) entre les neurones  $i$  et  $j$

$E_j$  : La valeur à l'entrée du neurone  $j$

Dans la plupart des applications, les fonctions d'activation utilisées sont soit la sigmoïde soit la tangente hyperbolique. Ces deux fonctions sont non linéaires et ont une forme asymptotique (Figure 3.3). Elles travaillent comme des amplificateurs non linéaires du signal. Les fonctions d'activation permettent de compresser la sortie d'un neurone dans un intervalle  $[0,1]$  pour la fonction sigmoïde et dans un intervalle  $[-1,1]$  pour la tangente hyperbolique afin d'éviter la saturation du signal.

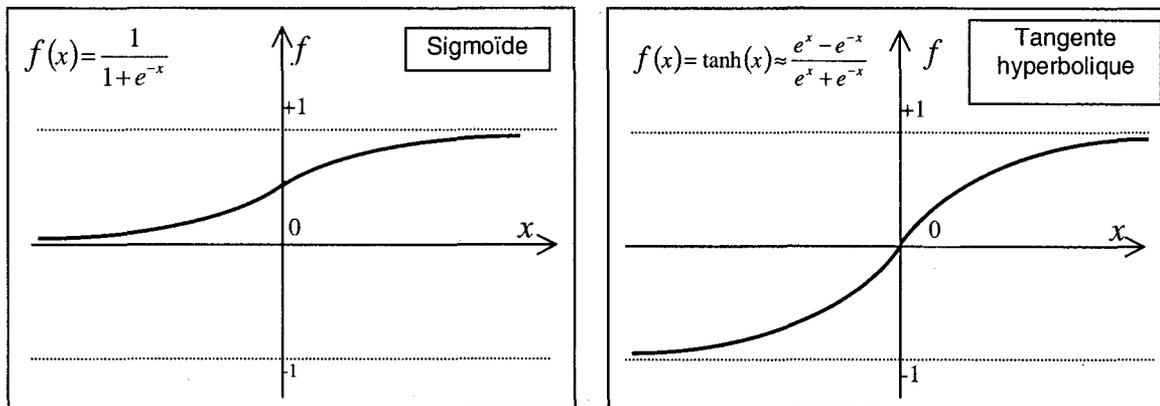


Figure 3.3 Fonctions d'activation

### 3.3 Apprentissage du réseau

Les réseaux de neurones sont des outils de modélisation numérique qui tentent de prédire les sorties d'un système à partir de la connaissance des entrées. Cette prédiction est réalisée en construisant au cours d'une phase d'apprentissage (ou d'entraînement) un modèle non linéaire entre des couples entrées-sorties. Les poids ( $w$ ) précisent le lien entre deux neurones appartenant à deux niveaux successifs (Figure 3.2). Leurs valeurs sont ajustées et affinées continuellement tout au long de la phase d'apprentissage. Pendant cette phase, un certain nombre de couples entrées-sorties sont fournis au réseau. Ces données représentent le groupe d'apprentissage et elles sont constituées des informations disponibles pour chaque station (température, niveau d'eau, etc.).

Dans un premier temps, les poids sont fixés aléatoirement pour permettre au réseau de calculer ses propres sorties à partir des entrées déjà fournies. Les poids sont alors corrigés

de manière à minimiser la différence entre les sorties ainsi calculées et les sorties réelles. Cette phase de minimisation correspond à l'apprentissage ; elle est primordiale à l'efficacité du réseau. L'ensemble des données utilisées pour cette étape doit donc être représentatif des situations qui seront rencontrées ultérieurement, lors de l'utilisation réelle. En effet, le réseau ne peut fournir de réponses correctes si les valeurs présentées lui paraissent inconnues.

Afin d'assurer un bon fonctionnement du réseau, les données présentées à l'entrée doivent être normalisées. Cette opération garantit une réponse significative de la fonction d'activation. C'est à dire que, pendant l'ajustement des poids, la sortie ajustée de chaque neurone doit refléter les ajustements initiaux. Ceci nous permet d'éviter que de petits changements dans l'entrée du réseau génèrent des grands changements à la sortie entraînant la saturation du réseau.

Contrairement au nombre de neurones des niveaux cachés (qui doivent être déterminés expérimentalement), le nombre de neurones du niveau d'entrée et du niveau de sortie est directement lié aux informations disponibles et aux résultats attendus du réseau.

Pour le niveau d'entrée, on affecte généralement un neurone pour chaque information fournie au réseau. L'ordre de présentation des données d'entrée n'est pas important. Par contre, le format de valeur présentée au réseau a un effet primordial sur les phases d'entraînement et de classification. Les informations présentées à l'entrée seront filtrées par le réseau en donnant des poids différents pour chaque information. Comme cela, seules les données utiles seront prises en considération pour calculer la sortie.

L'architecture choisie pour ce projet est composée de trois couches de neurones : un niveau d'entrée composé de trois neurones (un neurone pour chaque information fournie au réseau), un niveau caché de 15 neurones et un seul neurone au niveau de sortie (qui correspond au débit calculé). Les informations fournies au réseau sont les suivantes:

- ◆ Niveau d'entrée (neurone 1) : niveau d'eau (m).
- ◆ Niveau d'entrée (neurone 2) : température moyenne journalière obtenue en faisant la moyenne du minimum et du maximum enregistrés pour la journée (°C).

- ◆ Niveau d'entrée (neurone 3) : température moyenne des 10 jours précédents (°C).
- ◆ Niveau de sortie : le débit  $Q$  mesuré sur le terrain ( $m^3/s$ ).

L'apprentissage a été effectué par un algorithme de rétropropagation avec un taux d'apprentissage variable et une fonction d'activation sigmoïde.

Au début de la phase d'apprentissage, les groupes apprentissage et validation sont présentés au réseau avec les valeurs de sortie correspondantes. Les poids sont ajustés et affinés continuellement tout au long de la phase d'apprentissage. La correction des poids au cours de l'entraînement ne tient compte que des données appartenants au groupe d'apprentissage. Au cours de cette phase, les poids du réseau sont corrigés de manière à minimiser l'erreur au carré entre la réponse calculée par le réseau et la réponse attendue.

Généralement, l'erreur calculée sur le groupe d'apprentissage diminue continuellement au cours de l'entraînement. Toutefois, une longue phase d'entraînement diminue la capacité de généralisation du réseau en l'adaptant uniquement aux données de l'apprentissage (Figure 3.4). Ce phénomène est appelé le surentraînement ou «*overfitting*» en anglais. À cet effet, nous avons ajouté un autre groupe de données (groupe de validation) pour déterminer à quel moment l'apprentissage doit être arrêté. Les données appartenants à ce groupe servent uniquement à vérifier le comportement du réseau au cours de l'entraînement face à des données qui lui sont étrangères. Contrairement à l'erreur calculée sur le groupe d'apprentissage qui diminue continuellement au cours de l'entraînement, celle calculée sur le groupe de validation diminue dans la première phase d'entraînement en suivant une allure semblable à celle du groupe d'apprentissage avant de commencer à s'accroître (Figure 3.4). Ceci s'explique par le fait que le réseau commence à perdre son pouvoir de généralisation en adaptant ces neurones uniquement au groupe d'apprentissage.

L'entraînement du réseau est donc arrêté dès que cette erreur commence son ascension. Toutefois, afin d'éviter un arrêt prématuré de l'apprentissage causé par une augmentation ponctuelle de l'erreur du groupe de validation, nous avons introduit un seuil de décision qui tolère de légères ascensions successives de l'erreur. Si cette erreur continue son ascension au delà de ce seuil, on arrête l'apprentissage du réseau et on conserve les valeurs des poids

qui correspondent à l'itération qui précède cette ascension. Après plusieurs tests, nous avons trouvé qu'un seuil de 50 itérations est largement suffisant pour contourner ce genre de situation.

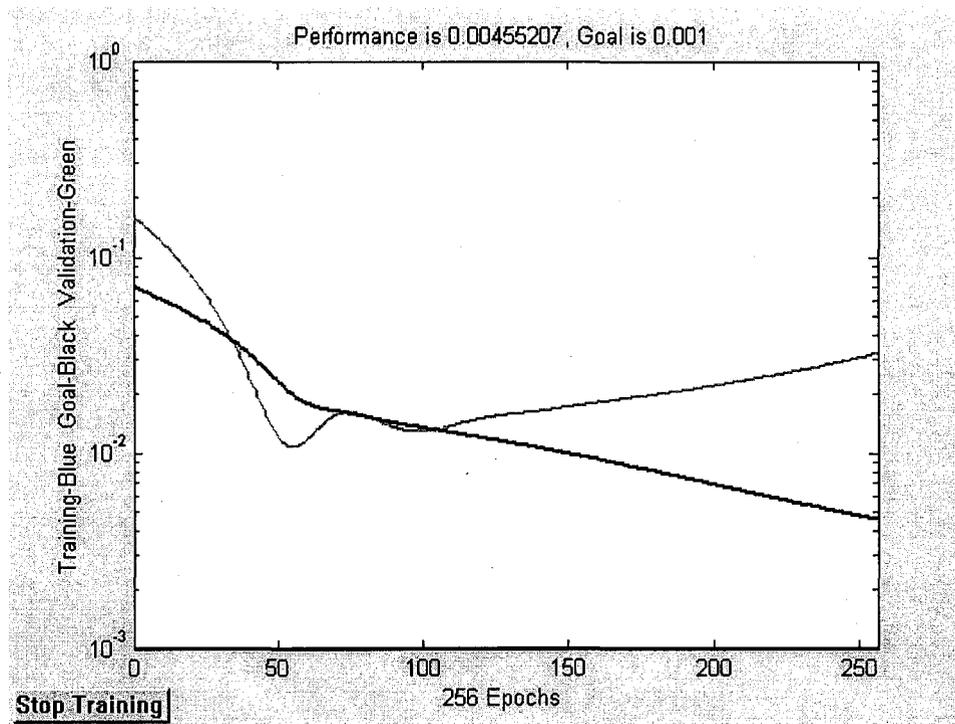
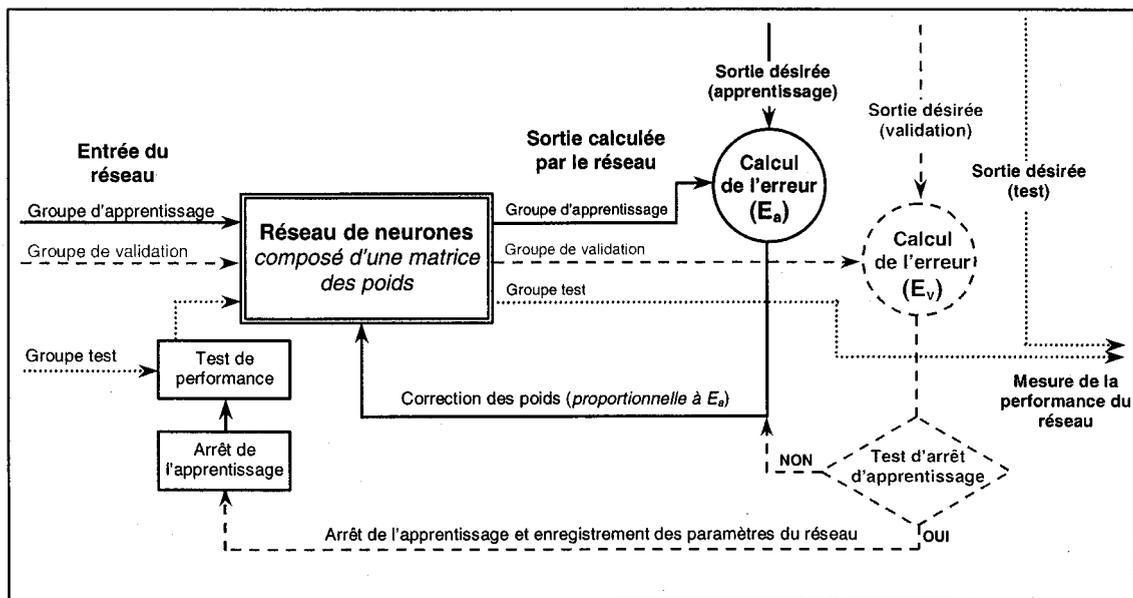


Figure 3.4 Évolution de l'erreur au cours de la phase d'apprentissage

### 3.4 Validation du réseau

Après l'arrêt de l'apprentissage, il est toujours préférable de vérifier la performance du réseau avec un troisième groupe de données (groupe test). Ce groupe doit être constitué d'un ensemble de données qui n'ont pas servi à l'apprentissage et qui n'ont joué aucun rôle dans le choix du moment de l'arrêt de l'apprentissage. Le groupe test est utilisé uniquement pour mesurer la performance du réseau après l'arrêt de l'apprentissage. Si le réseau arrive à prédire correctement les débits contenus dans ce groupe de données avec une précision « acceptable », on peut dire que le réseau est opérationnel. Dans le cas contraire, il faut réviser les intrants du réseau et recommencer l'apprentissage.



**Figure 3.5** Description schématique de la phase d'apprentissage

L'organigramme illustré à la Figure 3.5 résume les principales étapes suivies pendant l'apprentissage et la validation du réseau de neurones et montre le cheminement de chaque groupe de données.

En calquant la méthode de validation du modèle régressif utilisée au chapitre précédent avec l'exemple de la rivière Dartmouth (i.e. l'application d'une procédure de "Leave one out"), on a reporté au Tableau 3.1 l'ensemble des valeurs prédites à l'aide du modèle neuronal pour chacune des 14 observations disponibles pour ce cas de figure. On y constate d'emblée que les écarts entre les valeurs mesurées et prédites sont plus importants que ceux observés avec le modèle régressif (voir Tableau 2.3). Il en résulte donc des erreurs relatives plus élevées, i.e. que la performance globale du modèle neuronal (avec une erreur relative moyenne de 43.95%) semble moins bonne que celle du modèle régressif (dont l'erreur relative moyenne a été estimée à 13.9%).

La principale explication de ce moindre rendement réside d'abord et avant tout dans la taille restreinte de l'échantillon employé pour la configuration du réseau de neurones. En effet, la taille de l'échantillon fourni en entrée ne devrait jamais être inférieure au nombre de

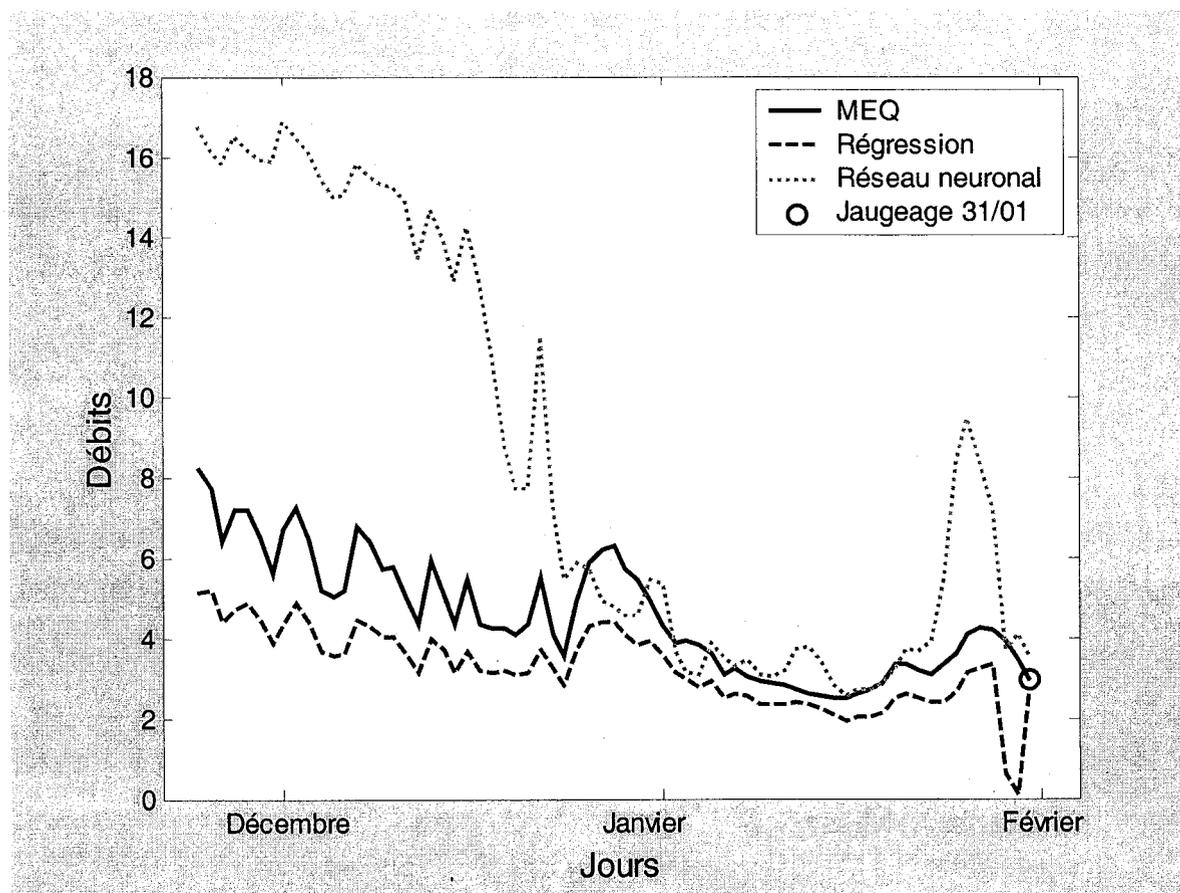
connexions entre les différents neurones, afin d'être en mesure d'entraîner adéquatement le réseau. Dans ce cas-ci, avec 3 neurones au niveau d'entrée, 15 neurones au niveau intermédiaire plus un neurone au niveau de sortie, on a un total de  $(3 \times 15) + (15 \times 1) = 60$  connexions. On comprend alors qu'il est difficile d'obtenir de bons résultats avec seulement 14 observations.

**Tableau 3.1** Comparaison des débits prédits aux débits mesurés par la méthode du "Leave one out" (pour le modèle neuronal)

Observation	Débit mesuré	Débit prédit	Erreur relative (%)
1	3.86	5.29	37.05
2	2.57	3.48	35.41
3	1.16	1.83	57.76
4	0.88	1.27	44.32
5	3.95	4.10	3.80
6	3.12	1.93	38.14
7	1.96	1.51	22.96
8	10.76	12.18	13.20
9	5.01	1.80	64.07
10	4.53	10.44	130.46
11	1.36	1.48	8.82
12	1.34	2.32	73.13
13	17.88	10.08	43.62
14	2.96	1.70	42.57
<b>Erreur relative moyenne :</b>			<b>43.95</b>

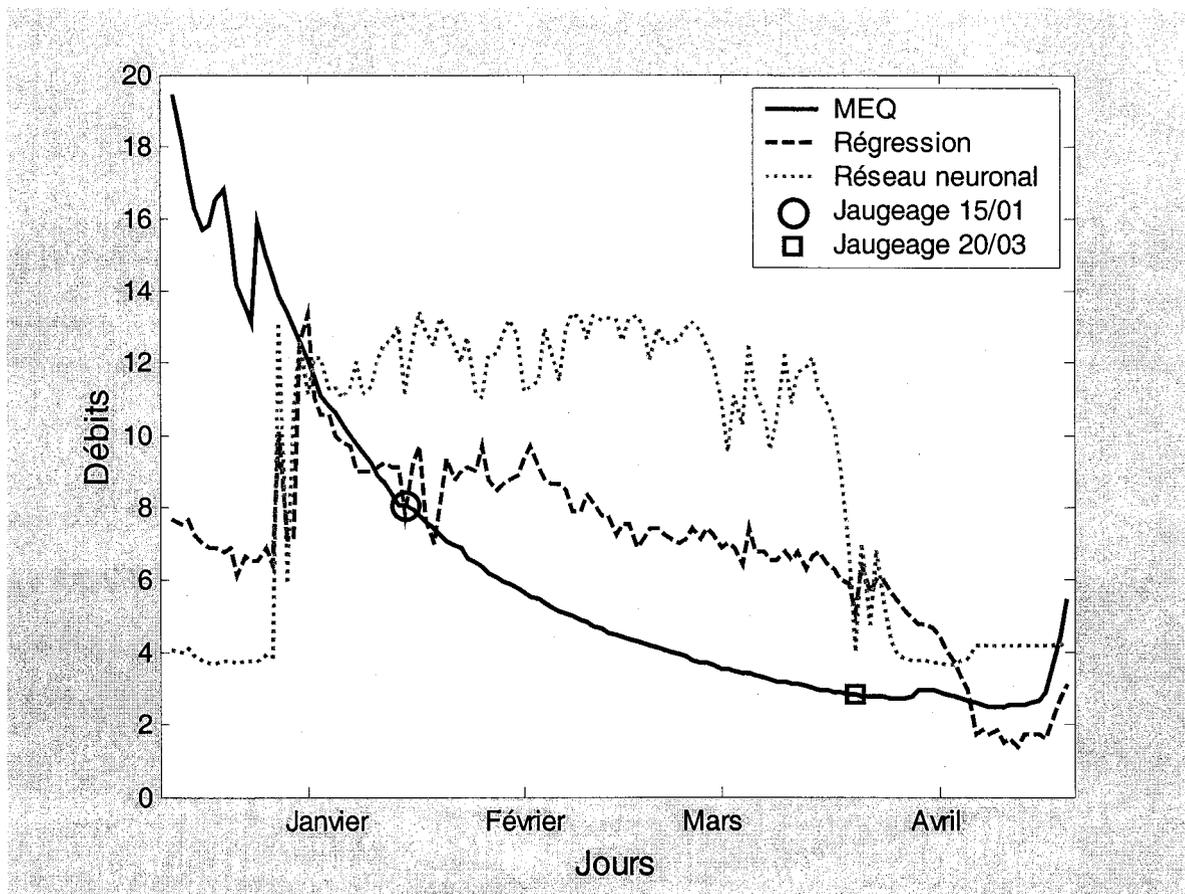
Un autre facteur pouvant affecter la performance du modèle est relié l'algorithme d'entraînement qui se base sur la différence observée entre les débits minimum et maximum

de l'échantillon de façon à reproduire une gamme de débits couvrant bien cet écart. Hors, dans le cas retenu ici, la plus grande valeur de débit enregistrée (soit  $17.88 \text{ m}^3/\text{s}$ ) se démarque du reste du lot ; la question qui se pose à partir de là consiste à déterminer s'il s'agit d'une valeur singulière. Dans l'affirmative, la réponse du modèle neuronal va éventuellement donner des résultats dont l'ordre de grandeur semble irréaliste, tel qu'on pourrait le supposer par exemple pour l'observation 10 au Tableau 3.1. En revanche, si la valeur n'est pas singulière et qu'elle n'est pas incluse dans la phase d'apprentissage du réseau, celui-ci ne sera pas en mesure de prédire adéquatement le débit lorsqu'il sera confronté aux conditions ayant mené à un tel événement (comme on peut le constater pour l'observation 13 au Tableau 3.1).

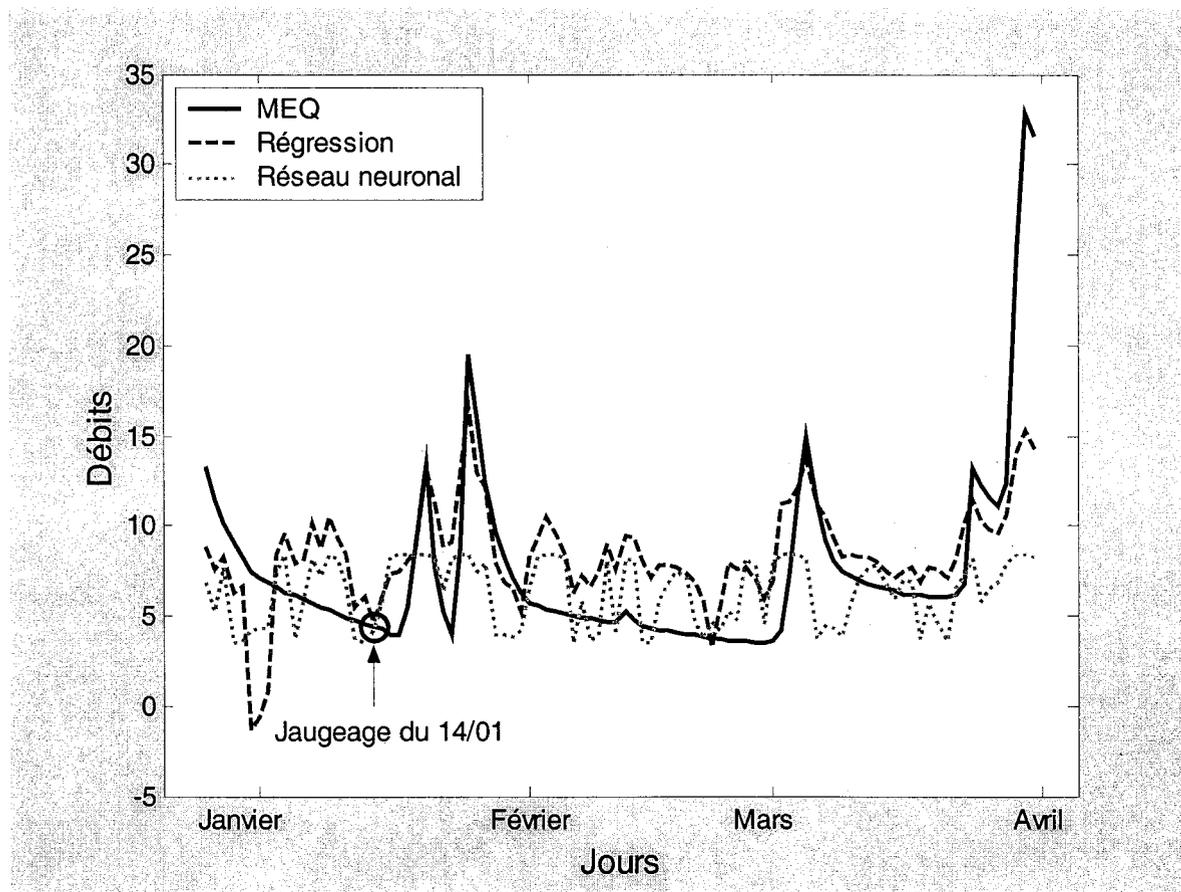


**Figure 3.6** Comparaison de la courbe des débits prédits par réseau de neurones à la courbe des débits corrigés du MEQ établie pour la rivière Dartmouth (sur la période allant du 24/11/98 au 31/01/99)

À la lumière de ces résultats, on se s'attend donc pas à ce que la courbe de débits prédits à l'aide du réseau neuronal sur une période continue se compare avantageusement à la courbe de débits corrigés établie par le MEQ pour la même période. C'est effectivement ce qui ressort à première vue de l'étude de la Figure 3.6 (où les débits ont été établis pour la même période qu'à la Figure 2.2), où l'on compare les trois types d'approches, i.e. celle du MEQ, la régression et le réseau de neurones. On constate néanmoins que les fluctuations de la courbe du modèle neuronal respectent dans une certaine mesure celles de la courbe du MEQ. Quant à l'ordre de grandeur des valeurs prédites par le réseau, on peut sans doute le relier à l'usage de la valeur maximale observée de  $17.88 \text{ m}^3/\text{s}$  pour la station Dartmouth au cours de l'entraînement du réseau (tel qu'expliqué au paragraphe précédent).



**Figure 3.7** Comparaison des courbes de débits prédits par les approches régressive et neuronale à la courbe des débits corrigés du MEQ établie pour la rivière Nouvelle (sur la période allant du 13/12/96 au 19/04/97)



**Figure 3.8** Comparaison des courbes de débits prédits par les approches régressive et neuronale à la courbe des débits corrigés du MEQ établie pour la rivière Du Sud (sur la période allant du 26/12/98 au 31/03/99)

En ce qui concerne les autres stations de l'étude, les résultats ne sont guère plus probants (Dartmouth constituant le meilleur cas avec ses 14 observations). Bien que l'on puisse calibrer le réseau de neurones pour la plupart d'entre elles, l'emploi d'une procédure de "Leave one out" aura tôt fait de révéler une grande variabilité au niveau des écarts entre les débits mesurés et prédits. L'évaluation de la performance des modèles par le critère de l'erreur relative moyenne est donc plutôt futile dans ce contexte puisque le petit nombre d'observations disponibles pour chaque station ne permet pas de refléter le vrai potentiel de ce type de modèle.

---

Ainsi, à défaut de pouvoir produire une bonne étude comparative des trois approches dont il est question ici (soit MEQ, régression et réseau de neurones), on complète alors simplement cette section en présentant graphiquement (de façon analogue à ce qui a été fait pour Dartmouth à la Figure 3.6) les résultats obtenus pour deux cas supplémentaires, i.e les stations Nouvelle (Figure 3.7) et Du Sud (Figure 3.8). On remarque immédiatement que les trois courbes ne concordent pas du tout à la Figure 3.7, tandis que cette disparité est moins prononcée à la Figure 3.8.

Bien que l'approche neuronale ne semble pas présenter d'avantages par rapport au modèle de régression sur ces deux figures, il est bon de rappeler que la performance de ce dernier est en fait moins bonne dans ces deux cas que dans celui de Dartmouth (en se basant sur les valeurs du  $R^2$  ajusté et de l'erreur relative moyenne reportées au Tableau 2.4), ceci étant principalement dû à la corrélation plus faible existant entre les variables de débit et de niveau pour ces stations. Sachant cela, on pourrait en temps normal s'attendre à ce que le modèle neuronal, plus flexible dans sa structure, puisse en arriver à de meilleures estimations. Excepté que pour l'instant, il s'agit là d'une hypothèse qui ne pourra être vérifiée que lorsqu'on aura en main de meilleurs jeux de données à fournir en entrée au réseau de neurones...



## 4 CONCLUSIONS

---

L'objectif visé dans cette étude était de développer et de comparer deux types d'approches pour la correction du débit en présence d'un effet de glace, la première étant basée sur la technique de régression multiple et la seconde sur les réseaux de neurones artificiels, ces approches étant principalement caractérisées par leur objectivité et leur reproductibilité (par opposition aux méthodes présentement en vigueur au MEQ et à EC qui sont plutôt caractérisées par leur subjectivité).

Pour l'approche régressive, on s'est d'abord intéressé à une combinaison de différents modèles faisant intervenir les variables explicatives suivantes : le niveau d'eau, la température moyenne journalière, la température moyenne des 10 jours précédents ainsi que la hauteur de neige au sol. La méthode proposée pour le choix du modèle optimal s'appuie sur la valeur du  $R^2$  ajusté. Cette technique a d'abord été validée, avec des résultats satisfaisants, pour le cas de la rivière Dartmouth à l'aide d'une procédure de "Leave one out" (l'erreur relative moyenne ayant été estimée à 13.9%), qui avec seulement 14 observations disponibles constituait le meilleur cas de figure parmi l'ensemble des stations de l'étude (en raison de données météo insuffisantes). Pour les autres stations, les équations de régression optimales ont également été déterminées, quoique le peu d'observations disponibles pour calibrer ces modèles fait en sorte que la performance de ceux-ci demeure bien relative pour l'instant.

En ce qui concerne l'approche neuronale, les mêmes variables explicatives (excepté la hauteur de neige) ont été considérées. L'architecture proposée pour le réseau est constituée de trois couches de neurones se déclinant ainsi : un niveau d'entrée composé de trois neurones (un neurone pour chaque variable explicative), un niveau caché de 15 neurones et un seul neurone au niveau de sortie (qui correspond au débit calculé). L'apprentissage a été effectué par un algorithme de rétropropagation avec un taux d'apprentissage variable et une fonction d'activation sigmoïde. L'échantillon de calibrage est séparé en deux groupes, soit l'un pour l'apprentissage et l'autre pour la validation. Les données appartenant au groupe de validation servent à vérifier le comportement du réseau au cours de

l'entraînement face à des données qui lui sont étrangères et permettent donc de déterminer à quel moment l'apprentissage doit être arrêté.

La validation du modèle neuronal sur le cas de la rivière Dartmouth a donné de moins bons résultats (avec une erreur relative moyenne ayant été estimée cette fois à 43.95%) que ceux obtenus au moyen de l'approche régressive. Cela est principalement dû à la petite taille de l'échantillon qui devrait en fait être supérieure ou égale au nombre de connexions entre les différents neurones (60 dans ce cas-ci) afin de pouvoir bien entraîner le réseau. Ce même problème persiste alors pour l'ensemble des stations, ce qui ne permet pas de conclure sur les véritables capacités du modèle.

Ainsi, à défaut de pouvoir quantifier adéquatement la performance du réseau neuronal, on s'est limité pour la comparaison des approches à une étude graphique entre les courbes de débits prédits selon la méthode du MEQ, puis selon les approches régressive et neuronale. Cela a permis de constater que l'allure des courbes découlant des modèles analytiques peut se rapprocher de belle façon de celle des courbes établies par le ministère, ce qui laisse entrevoir un potentiel justifiant une exploration plus approfondie du sujet, pour autant que le tout soit conditionné à un accès à des données plus substantielles.

## RÉFÉRENCES

---

Mc Culloch, W.S., Pitts, W. (1943). A logical calculus of the ideas immanent in nervous activity. *Bull. Mathem. Biophys.*, **5**, 115-133.

Ouarda, T.B.M.J., Faucher, D., Coulibaly, P., Bobée, B. (2000). Correction du débit en présence d'un effet de glace; étude de faisabilité pour le développement d'un logiciel. Rapport de recherche No. R-559, INRS-Eau, Québec, 75 pages.

Rumelhart, D.E., Hinton, G.E., Williams, R.J. (1986). *Learning internal representation by error propagation. In: Parallel distributed processing : Exploration in the microstructure of cognition (Vol. 1), ed. Rumelhart, D.E. & McClelland, J.L.* MIT Press, pp. 318-364.

The MathWorks, Inc. (2000). *Using Matlab, Version 6*.

Weisberg, S. (1985). *Applied linear regression (Second Edition)*. Wiley & Sons, New York, 324 pages.

Werbos, P.J. (1974). *Beyond regression : New tools for prediction and analysis in the behavioral sciences*. Thèse de doctorat non publiée, Université Harvard.



# ANNEXE : GUIDE DE L'USAGER DU LOGICIEL *CORRECT*

---

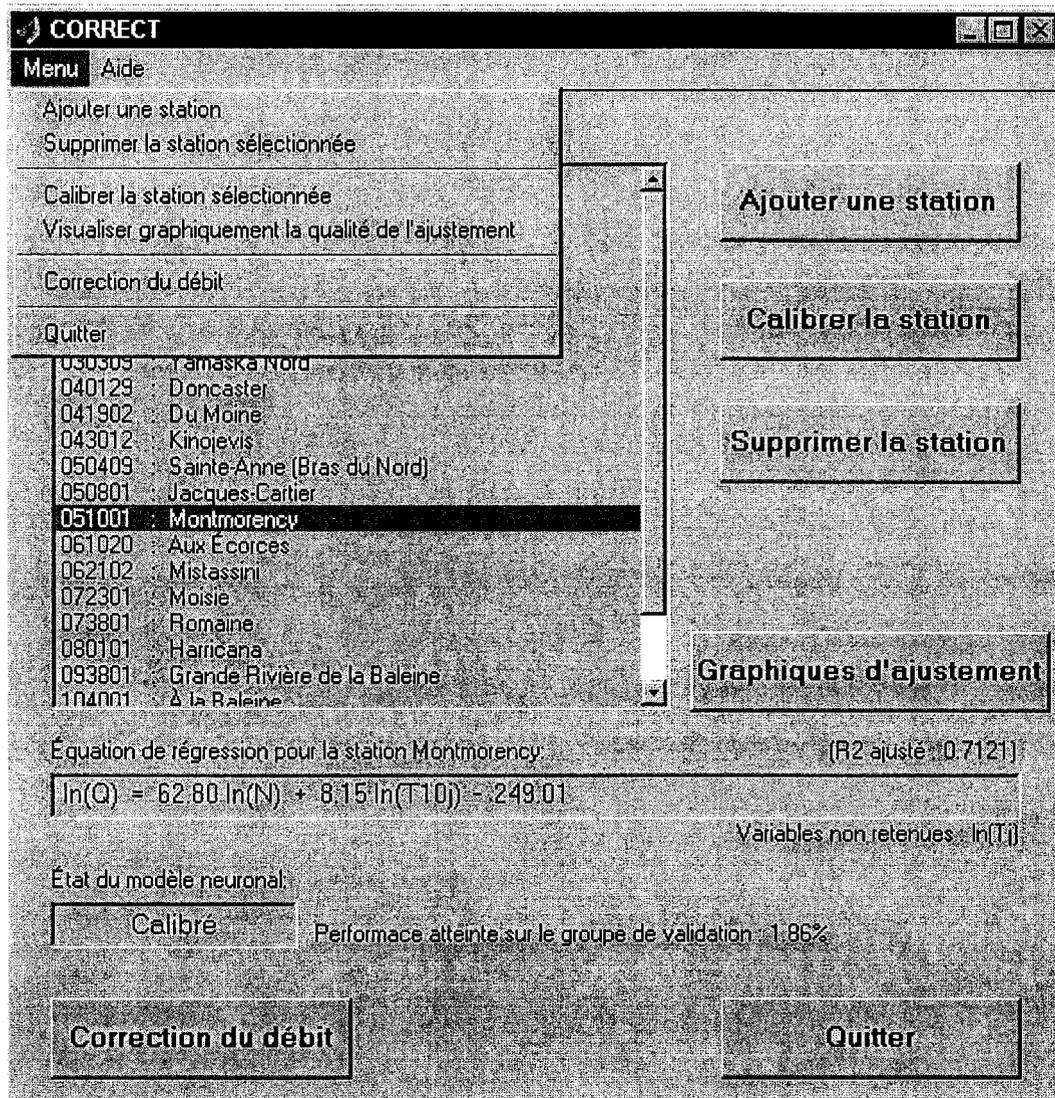
## A.1 Description et aperçu du logiciel

Le logiciel *CORRECT* a été conçu dans le but d'automatiser les étapes menant au calibrage des modèles régressif et neuronal, en vue de leur utilisation subséquente dans une procédure de correction automatique du débit compensant pour l'effet de glace. Il a été développé dans l'environnement *MATLAB* (version 6.1), i.e. que l'utilisateur potentiel doit être détenteur d'une licence *MATLAB* (version 6 et plus, avec en prime les boîtes d'outils "Statistics" et "Neural Networks") pour être en mesure d'exécuter *CORRECT*. On l'installe en copiant simplement quelques fichiers (certains avec l'extension ".m", pour les programmes, et d'autres avec l'extension ".fig" pour les interfaces) sur la machine de l'utilisateur. Puis, on l'ouvre en tapant le mot "Correct" dans la fenêtre de commandes de *MATLAB* (et après s'être assuré que le répertoire courant soit celui où les différents fichiers ont été copiés), ce qui a pour effet de faire apparaître l'interface principale, telle que reproduite à la Figure A.1 (il apparaît aussi avant la fenêtre d'introduction "À propos de CORRECT" qui se ferme automatiquement au bout de 10 secondes).

L'interface principale est constituée d'une boîte défilante donnant la liste des stations, de deux boîtes d'édition donnant les informations relatives au calibrage des modèles régressif et neuronal (intitulées respectivement "Équation de régression pour la station ..." et "État du modèle neuronal"), ainsi que des boutons suivants, qui ont chacun leur équivalent dans l'une des options du menu défilant accessible au haut de la fenêtre:

- Ajouter une station (correspondant à l'option "Ajouter une station" du menu défilant)
- Supprimer une station (correspondant à l'option "Supprimer la station sélectionnée" du menu défilant)
- Calibrer la station (correspondant à l'option "Calibrer la station sélectionnée" du menu défilant)

- Graphiques d'ajustement (correspondant à l'option "Visualiser graphiquement la qualité de l'ajustement" du menu défilant)
- Correction du débit (correspondant à l'option "Correction du débit" du menu défilant)
- Quitter (correspondant à l'option "Quitter" du menu défilant)



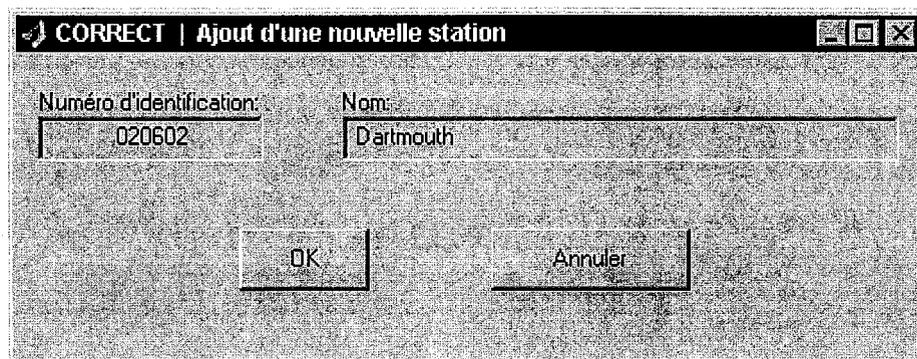
**Figure A.1** Interface principale du logiciel *CORRECT*

Les sections A.2 à A.7 traitent dans cet ordre des actions reliées à ces différents boutons, i.e. que l'ajout et la suppression de stations font l'objet des sections A.2 et A.3, tandis que le calibrage et la visualisation de la qualité de l'ajustement font l'objet des sections A.4. et

A.5, puis que l'on discute de la correction du débit à la section A.6, et qu'enfin l'on décrit comment quitter le logiciel à la section A.7.

## A.2 Ajout d'une nouvelle station

Pour ajouter un nouvel élément à la liste des stations disponibles, il suffit d'appuyer sur le bouton "Ajouter une station" (ou encore de sélectionner l'option correspondante dans le menu défilant), puis d'entrer le numéro ainsi que le nom de la nouvelle station dans la fenêtre prévue à cet effet, telle que représentée à la Figure A.2.



**Figure A.2** Fenêtre d'ajout d'une nouvelle station

Lorsque les informations requises ont été entrées, on doit appuyer sur le bouton "OK" pour confirmer l'ajout. Pour annuler l'opération, l'utilisateur a également le choix d'appuyer sur le bouton "Annuler". La nouvelle station ainsi définie s'insérera automatiquement dans la liste des stations de l'interface principale, qui soit dit en passant, sont triées par ordre croissant selon leur numéro d'identification. Tant que les modèles régressif et neuronal n'auront pas été calibrés pour la nouvelle station, les boîtes d'édition "Équation de régression pour la station ..." et "État du modèle neuronal" afficheront respectivement les mentions "< Équation non établie >" et "Non calibré", tandis que le bouton "Graphiques d'ajustement" de même que l'option correspondante du menu défilant demeureront inactifs.

### A.3 Suppression d'une station existante

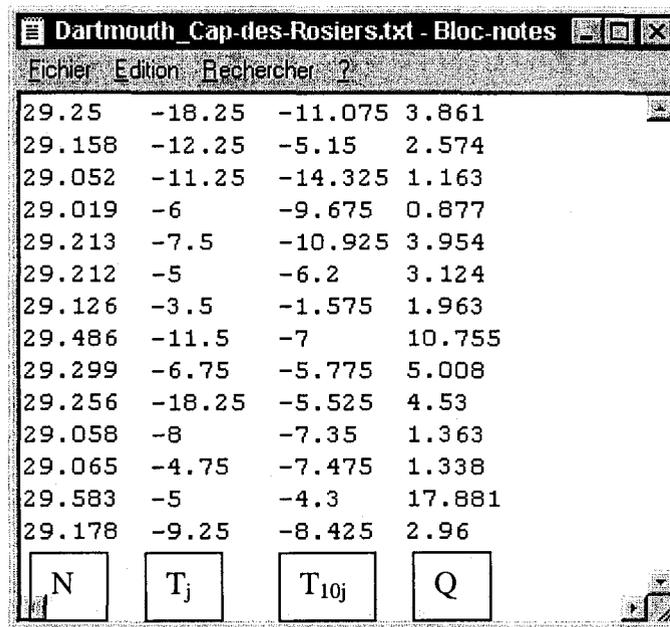
Pour supprimer une station de la liste, on doit tout d'abord la sélectionner (par exemple, sur la Figure A.1, on voit que c'est la station Montmorency (051001) qui est sélectionnée) puis appuyer ensuite sur le bouton "Supprimer la station" (ou bien sélectionner l'option correspondante dans le menu défilant). Une boîte de dialogue apparaîtra alors demandant à l'utilisateur de confirmer (par le bouton "Oui") ou non (par le bouton "Non") la suppression de la station en question.

### A.4 Calibrage des modèles régressif et neuronal pour une station donnée

Avant de passer à l'étape du calibrage, il est nécessaire de préparer un certain nombre (tout dépendant du nombre de stations à calibrer) de fichiers de données dont *CORRECT* a besoin afin d'ajuster les modèles. Il s'agit de fichiers au format texte (voir Figure A.3), constitués de quatre colonnes de valeurs correspondant respectivement aux variables suivantes:

- Niveau d'eau (N)
- Température moyenne journalière ( $T_j$ )
- Température moyenne des 10 jours précédents ( $T_{10j}$ )
- Débit (Q)

Le choix de ces trois variables explicatives du débit découle bien entendu des résultats de la présente étude. Les abréviations pour ces variables, données entre parenthèses, sont celles utilisées dans *CORRECT*. Chaque ligne du fichier de données est alors associée à une certaine date de jaugeage. Notons qu'il n'est pas nécessaire que celles-ci soit classées par ordre chronologique. Un minimum de cinq lignes d'observations est requis pour chaque fichier.



Fichier	Edition	Rechercher	?
29.25	-18.25	-11.075	3.861
29.158	-12.25	-5.15	2.574
29.052	-11.25	-14.325	1.163
29.019	-6	-9.675	0.877
29.213	-7.5	-10.925	3.954
29.212	-5	-6.2	3.124
29.126	-3.5	-1.575	1.963
29.486	-11.5	-7	10.755
29.299	-6.75	-5.775	5.008
29.256	-18.25	-5.525	4.53
29.058	-8	-7.35	1.363
29.065	-4.75	-7.475	1.338
29.583	-5	-4.3	17.881
29.178	-9.25	-8.425	2.96

N T<sub>j</sub> T<sub>10j</sub> Q

**Figure A.3** Exemple d'un fichier de calibration

Une fois que les fichiers de calibration ont été complétés, on peut enclencher la procédure d'ajustement pour la station sélectionnée en appuyant sur le bouton "Calibrer la station" (ou encore en sélectionnant l'option correspondante dans le menu déroulant). Ceci provoque l'ouverture de la fenêtre standard de *Windows* pour la sélection d'un fichier, i.e. que l'on demande à l'utilisateur de se servir de l'outil de navigation afin de spécifier le fichier de calibration à utiliser. Par défaut, seuls les fichiers dont l'extension équivaut à ".txt", ".prn" ou encore ".csv" sont affichés dans la fenêtre. Si l'extension de votre fichier de données ne correspond à aucune de ces trois extensions précédentes, vous devez alors sélectionner l'option "Tous les fichiers" du menu "popup" intitulé "Type de fichiers" et situé tout au bas de la fenêtre.

Après avoir choisi le fichier de calibration, une nouvelle fenêtre s'ouvre demandant cette fois à l'utilisateur de préciser les paramètres du calibration. Cette fenêtre est illustrée à la Figure A.4. Au haut de cette dernière est reporté le nom du fichier de données à employer (si jamais il y a erreur, l'utilisateur peut ainsi annuler immédiatement l'opération). L'utilisateur doit ensuite entrer et/ou vérifier trois paramètres avant de procéder à l'ajustement.

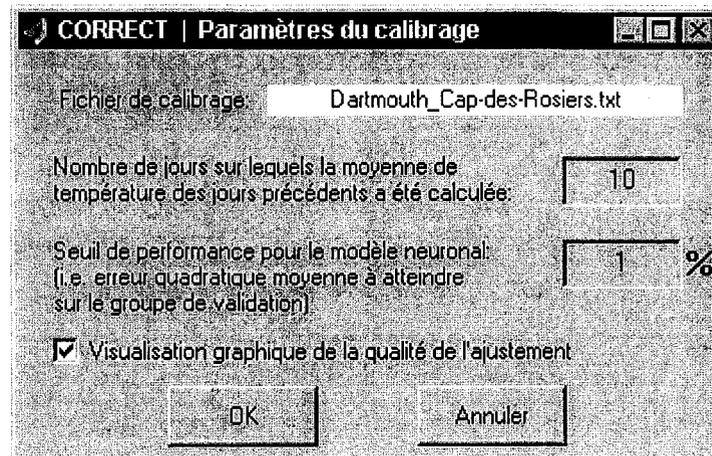


Figure A.4 Fenêtre des paramètres du calibrage

Le premier concerne le nombre de jours sur lesquels la moyenne de température des jours précédents a été calculée. Bien que cette valeur soit fixée par défaut à 10, on donne donc à l'utilisateur la possibilité de faire varier ce nombre de jours (l'important étant de s'assurer que les valeurs reportées dans le fichier de calibrage soient cohérentes avec le nombre de jours inscrit par l'utilisateur). L'abréviation de la variable sera modifiée en conséquence (par exemple, pour 5 jours, on aura  $T_{5j}$  au lieu de  $T_{10j}$ ).

Le deuxième paramètre à préciser est le seuil de performance pour le modèle neuronal, ou plus précisément l'écart quadratique moyen que l'on cherche à atteindre entre les valeurs prédites et observées pour le groupe de validation. Cette valeur est fixée par défaut à 1% (i.e. 0.01), ce qui peut toutefois s'avérer trop sévère à l'occasion, particulièrement dans les cas où l'on dispose de peu de données dans le fichier de calibrage. L'algorithme d'ajustement du modèle neuronal cherchera donc à satisfaire le critère de précision entré, ce qui peut mener éventuellement à un succès tout comme à un échec.

Le dernier paramètre laisse le choix à l'utilisateur de visualiser ou non de façon graphique la qualité de l'ajustement obtenu pour les deux modèles. Cette option est surtout utile dans le cas du modèle neuronal, où l'on voudra s'assurer que l'ensemble des valeurs prédites demeurent dans des limites acceptables. En effet, l'algorithme de convergence du réseau

neuronal peut parfois donner des résultats non réalistes, surtout si l'on dispose d'un nombre limité de données.

Le bouton "OK" de la fenêtre de paramètres permet finalement d'exécuter la procédure de calibrage, à moins qu'il n'y ait une erreur au niveau des paramètres entrés ou bien que le format du fichier de calibrage soit incorrect (e.g. pas un fichier texte, plus ou moins de quatre colonnes, etc.). Sinon, un message d'attente apparaît dans lequel on demande à l'utilisateur de patienter puisque la configuration du réseau neuronal peut s'étendre sur plusieurs minutes (il est à noter que le temps de calcul sera d'autant plus long que la taille du fichier de calibrage sera grande). Lorsque l'opération est complétée, une boîte de dialogue indique si la convergence du réseau a été atteinte ou non (en fonction du seuil de performance entré dans la fenêtre de paramètres). Si tel n'est pas le cas, le même message informe l'utilisateur de la meilleure précision obtenue sur le groupe de validation au cours de la phase d'apprentissage, et lui demande s'il souhaite malgré tout conserver la configuration associée à cette performance.

Si l'utilisateur a laissé cochée l'option de visualisation graphique de la qualité de l'ajustement dans la fenêtre de paramètres, les graphiques d'ajustement (i.e. valeurs prédites vs valeurs observées pour chaque modèle) sont également produits par *CORRECT* (voir Figures A.5 et A.6). Dans le cas où l'utilisateur a choisi de ne pas conserver la configuration proposée pour le réseau neuronal, seul le graphique relié au modèle régressif est affiché.

En retournant à l'interface principale du logiciel (Figure A.1), l'utilisateur verra alors que l'équation de régression  $y$  est maintenant reportée dans la boîte d'édition "Équation de régression pour la station ...". Si celle-ci se retrouve exprimée sous la forme  $\ln(Q) = \beta_1 \ln(N) + \beta_2 \ln(T_j) + \beta_3 \ln(T_{10j}) + \beta_0$ , cela signifie que l'ajustement dans l'espace logarithmique (i.e. en appliquant au préalable une transformation logarithmique sur les données) a donné de meilleurs résultats. S'il y a lieu, les variables explicatives non retenues dans le modèle sont reportées sous la boîte d'édition de l'équation, tandis que la valeur du  $R^2$  ajusté est donnée juste au-dessus du coin supérieur droit de cette même boîte.

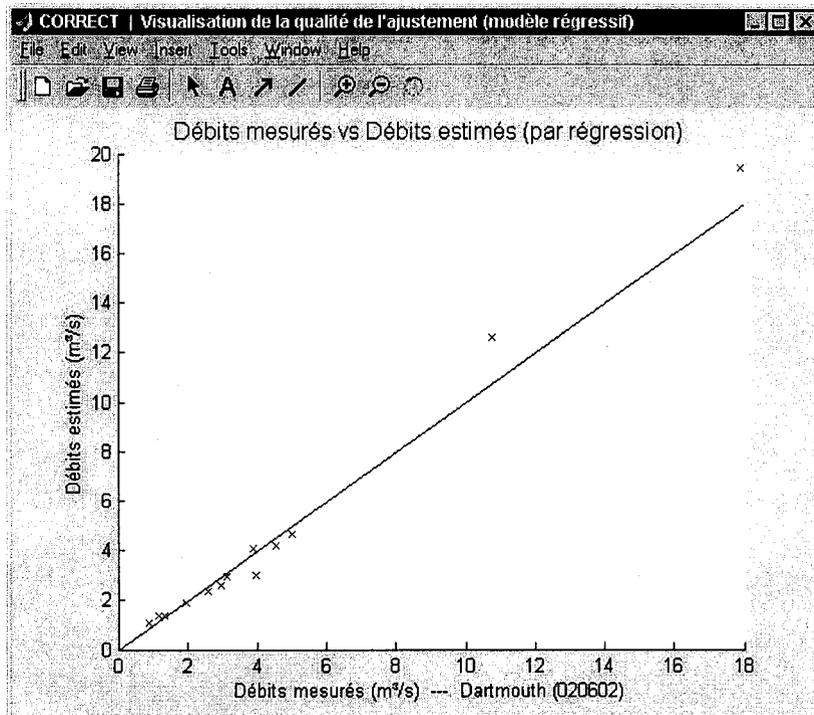


Figure A.5 Graphique d'ajustement du modèle régressif

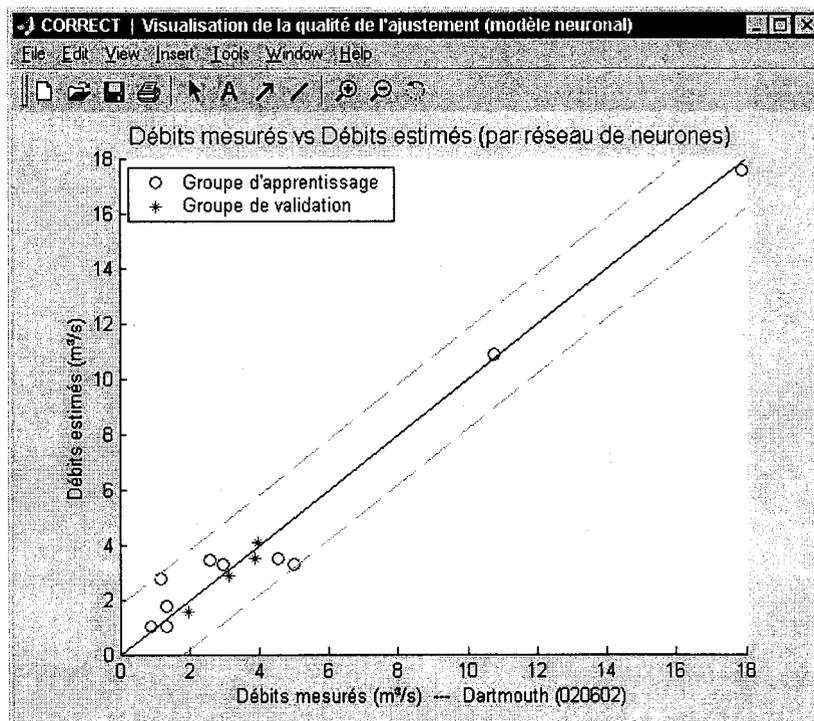


Figure A.6 Graphique d'ajustement du modèle neuronal

Pour ce qui est du modèle neuronal, la mention "Calibré" dans la boîte d'édition "État du modèle neuronal" indique que la convergence du réseau a été atteinte (l'inverse étant donné par la mention "Non calibré"), et l'erreur quadratique moyenne atteinte sur le groupe de validation est à ce moment également inscrite à côté de la boîte d'édition en question.

L'opération de calibrage peut être répétée pour chacune des stations. Elle peut même être répétée pour une station déjà calibrée, soit, par exemple, pour reconfigurer le réseau neuronal suite à un échec (en modifiant la valeur du seuil de performance) ou encore pour améliorer l'ajustement lorsque de nouvelles mesures de débit deviennent disponibles. On peut donc recalibrer à volonté une même station quoique seuls les derniers changements seront conservés. En effet, tous les résultats liés au calibrage des modèles sont sauvegardés dans un fichier dénommé "ParamRegNet.mat" (format *MATLAB*). Comme celui-ci s'enregistre par défaut dans le répertoire courant de *MATLAB*, il est important de spécifier le répertoire d'installation du logiciel comme répertoire courant avant de lancer l'exécution de *CORRECT* (de façon à ne garder qu'une seule copie à jour du fichier "ParamRegNet.mat" sur la machine de l'utilisateur, au même endroit que les fichiers de programmes). Il n'est donc pas recommandé d'ajouter le répertoire d'installation du logiciel *CORRECT* dans la variable "Path" de *MATLAB*.

## **A.5 Visualisation graphique de la qualité de l'ajustement**

Pour revoir les graphiques d'ajustement des modèles régressif et neuronal (tels qu'illustrés aux Figures A.5 et A.6) obtenus avec une station préalablement calibrée, il faut tout d'abord sélectionner la station dans l'interface principale et ensuite appuyer sur le bouton "Graphiques d'ajustement", ou bien sélectionner l'option "Visualiser graphiquement la qualité de l'ajustement" du menu défilant au haut de la fenêtre d'interface. Pour les stations où le calibrage du modèle neuronal a échoué, seul le graphique d'ajustement du modèle régressif est affiché. Pour fermer ces figures après les avoir consultées, il suffit de choisir l'option "Close" du menu "File" que l'on retrouve au haut de ces graphiques.

## A.6 Correction automatique du débit

Précisons tout d'abord que la procédure de correction du débit implantée (et qui peut être appelée en appuyant sur le bouton "Correction du débit" ou encore en sélectionnant l'option du même nom dans le menu défilant de l'interface principale) présentement dans *CORRECT* ne constitue en fait qu'une version préliminaire puisque les modalités d'implantation d'une telle procédure au MEQ, dans une perspective de correction en temps réel des débits pour l'effet de glace, n'ont pas encore été clairement établies (particulièrement au niveau du lien avec la base de données BDH). Pour l'instant, cette fonction agit plutôt à titre de démonstrateur de ce qui pourrait être offert dans le futur par le logiciel.

Avant de procéder toutefois, il est nécessaire de préparer un fichier de données (en format texte) dont chaque ligne, correspondant à une date d'observation, doit contenir les informations suivantes : le numéro de la station, les mesures du niveau, de la température moyenne journalière ainsi que de la température moyenne des 10 jours précédents (du moins le nombre de jours qui a été utilisé lors du calibrage de la station) enregistrées pour cette date. Un exemple d'un tel fichier est présenté à la Figure A.7. Lorsque l'utilisateur appuie sur le bouton "Correction du débit", on lui demande alors, tout comme dans le cas du fichier de calibrage (voir section A.4), de se servir de l'outil de navigation standard de *Windows* afin de spécifier le fichier de données à partir duquel *CORRECT* produira les valeurs de débits corrigés. En faisant appel aux paramètres sauvegardés dans le fichier "ParamRegNet.mat", le logiciel est donc en mesure de retourner les valeurs des débits estimés selon les approches régressive et neuronale.

Le tout est affiché dans une fenêtre (voir Figure A.8) où les quatre premières colonnes correspondent aux données entrées dans le fichier, tandis que les colonnes intitulées "Qreg" et "Qnet" contiennent les valeurs de débits corrigés associées aux modèles régressif et neuronal respectivement. Dans le cas où l'un ou l'autre des modèles n'a pas pu ou n'a pas été calibré pour l'une des stations inscrites dans le fichier, le code -99.99 est retourné. Il est possible d'enregistrer sous forme d'un fichier texte le tableau de résultats affiché à la

Figure A.8 en appuyant sur le bouton "Sauvegarder" puis en spécifiant un nom de fichier dans la fenêtre de sauvegarde standard de *Windows* qui s'ouvrira alors. Le bouton "Fermer" quant à lui sert simplement à la fermeture de la fenêtre de résultats.

011201	27.87	-7.5	-3.95
011201	27.86	-2.25	-4.875
011201	27.87	-2.5	-5
011201	27.84	-2.25	-4.5
011201	27.82	-0.5	-3.75
011201	27.81	2.25	-3.4
011201	27.81	0.5	-3.125
011201	27.8	-6.75	-2.875
011201	27.81	-10.75	-3.1
011201	27.75	-8.5	-3.575
011201	27.79	-8	-3.825
011201	27.78	-4.75	-3.875
011201	27.78	-8.5	-4.125
011201	27.81	-14.75	-4.725
011201	27.77	-13.75	-5.975
011201	28.07	-10.75	-7.3
011201	27.83	-5	-8.6
011201	27.83	-18.5	-9.15
011201	28.27	-22.25	-10.325
023106	28.686	-6.5	-7.325
023106	28.588	-7.5	-8.05
023106	28.601	-4	-8.475
023106	28.511	-10.75	-8.175
023106	28.541	-10.5	-8.35
023106	27.944	-23.25	-8.575
023106	28.037	-24.5	-10.375
023106	28.097	-20	-12.25
023106	28.736	-11.5	-12.7
023106	28.764	-5.75	-12.75
023106	28.818	-19	-12.425
023106	28.78	-15.5	-13.675
023106	28.969	-14.75	-14.425
023106	28.871	-17	-15.55
023106	28.955	-11.25	-16.175
023106	28.889	-14.5	-16.25
023106	28.839	-17.25	-15.375
023106	28.588	-20.5	-14.65
023106	28.608	-18	-14.7
020602	29.339	5.5	-1.37
020602	29.328	-2	-0.545
020602	29.295	-2	-0.395
020602	29.313	1	-0.395
020602	29.315	-2	-0.195
020602	29.296	-2.5	-0.12
020602	29.272	-1	-0.095
020602	29.302	3.75	0.125
020602	29.316	-2	0.4
020602	29.295	-2	0.15
020602	29.258	-2.75	-0.325
020602	29.252	-2.5	-1.15
020602	29.257	-1.75	-1.2
020602	29.302	0.75	-1.175
020602	29.294	-0.5	-1.2
020602	29.276	-3.5	-1.05
020602	29.277	-3	-1.15
020602	29.253	-4.5	-1.35
020602	29.228	-3.25	-2.175

Figure A.7 Exemple d'un fichier d'intrants pour la procédure de correction du débit

Tel que mentionné au début de cette section, cette fonction de correction automatique du débit n'est offerte pour l'instant qu'à titre de prototype, dans le sens où de nombreux ajouts à celle-ci peuvent être envisagés. Notamment, le format suggéré pour le fichier de données est appelé à être revu pour y ajouter par exemple des informations sur la date et l'heure de chaque enregistrement (puisque l'on envisage ici une procédure qui devra fonctionner en temps réel...). Une autre idée intéressante serait d'intégrer un petit module d'aide à la décision (e.g. par le biais de critères de performance) permettant de choisir entre les deux types d'approches celle qui semble la plus appropriée pour chacune des stations.

**CORRECT | Valeurs des débits corrigés**

Station	N	Tj	T10j	Qreg	Qnet
011201	27 87	-7 50	-3 95	7 62	6 64
011201	27 86	-2 25	-4 88	7 49	4 14
011201	27 87	-2 50	-5 00	7 62	4 30
011201	27 84	-2 25	-4 50	7 25	4 09
011201	27 82	-0 50	-3 75	7 00	3 60
011201	27 81	2 25	-3 40	6 88	3 39
011201	27 81	0 50	-3 13	6 88	3 51
011201	27 80	-6 75	-2 88	6 76	5 42
011201	27 81	-10 75	-3 10	6 88	7 10
011201	27 75	-8 50	-3 58	6 14	5 93
011201	27 79	-8 00	-3 83	6 63	6 05
011201	27 78	-4 75	-3 88	6 51	5 07
011201	27 78	-8 50	-4 13	6 51	6 29
011201	27 81	-14 75	-4 72	6 88	9 24
011201	27 77	-13 75	-5 97	6 39	8 98
011201	28 07	-10 75	-7 30	10 07	10 56
011201	27 83	-5 00	-8 60	7 12	6 16
011201	27 83	-18 50	-9 15	7 12	10 26
011201	28 27	-22 25	-10 32	12 52	10 92
023106	28 69	-6 50	-7 33	8 69	7 48
023106	28 59	-7 50	-8 05	7 57	3 81
023106	28 60	-4 00	-8 47	8 27	8 06
023106	28 51	-10 75	-8 18	6 29	2 34
023106	28 54	-10 50	-8 35	6 62	2 37
023106	27 94	-23 25	-8 57	-1 26	2 94
023106	28 04	-24 50	-10 38	-0 56	2 42
023106	28 10	-20 00	-12 25	0 76	2 62
023106	28 74	-11 50	-12 70	8 36	8 61
023106	28 76	-5 75	-12 75	9 57	8 68
023106	28 82	-19 00	-12 43	7 93	7 03
023106	28 78	-15 50	-13 68	8 13	8 58
023106	28 97	-14 75	-14 48	10 09	8 69
023106	28 87	-17 00	-15 55	8 77	8 67
023106	28 95	-11 25	-16 18	10 53	8 68
023106	28 89	-14 50	-16 25	9 35	8 68
023106	28 84	-17 25	-15 38	8 42	8 65
023106	28 59	-20 50	-14 65	5 45	4 42
023106	28 61	-18 00	-14 70	6 05	7 95
020602	29 34	5 50	-1 37	5 15	3 22
020602	29 33	-2 00	-0 55	5 19	3 51
020602	29 30	-2 00	-0 40	4 39	3 51
020602	29 31	1 00	-0 40	4 69	3 45
020602	29 32	-2 00	-0 20	4 86	3 52
020602	29 30	-2 50	-0 12	4 43	3 51
020602	29 27	-1 00	-0 10	3 87	3 47
020602	29 30	3 75	0 13	4 33	3 36
020602	29 32	-2 00	0 40	4 89	3 52

Sauvegarder Fermer

Figure A.8 Fenêtre d'affichage des débits corrigés

---

## **A.7 Fermeture du logiciel *CORRECT***

Pour quitter *CORRECT*, il suffit d'appuyer sur le bouton "Quitter" ou bien de sélectionner l'option du même nom tout au bas du menu défilant de l'interface principale. L'utilisateur verra apparaître à ce moment une boîte de dialogue lui demandant de confirmer s'il désire réellement quitter le logiciel. Rappelons qu'il n'y a pas d'opération de sauvegarde par rapport aux résultats du calibrage des modèles qui soit nécessaire d'effectuer avant de quitter puisque le tout s'enregistre automatiquement au fur et à mesure dans le fichier "ParamRegNet.mat". L'utilisateur retrouvera donc ces mêmes résultats lorsqu'il ouvrira à nouveau *CORRECT* et pourra donc continuer les opérations de calibrage pour les stations n'ayant pas encore été traitées.