

Analyse des mesures d'auscultation des barrages avec les réseaux de neurones

Analysis and Interpretation of Dam Measurements using Artificial Neural Networks

Alexandre, G., Simon¹

EDF-DTG 21 avenue de l'Europe 38040 GRENOBLE
alexandre-gilles.simon@edf.fr

Maxime, J., Royer

ENSE³ 11 rue des mathématiques 38402 SAINT MARTIN D'HERES
maxime.royer@ense3.grenoble-inp.fr

Florian, J., Mauris

EDF-DTG 21 avenue de l'Europe 38040 GRENOBLE
florian.mauris@edf.fr

Jean-Paul, Fabre

EDF-DTG 62Bis rue Raymond IV 31685 TOULOUSE
jean-paul.fabre@edf.fr

MOTS CLÉS

auscultation, réseaux de neurones, sous-pression, débit, déplacement.

RÉSUMÉ

Les mesures d'auscultation des barrages suivent des évolutions dictées par plusieurs phénomènes réversibles (charge hydrostatique, effets thermiques,...) et irréversibles (fluage, gonflement,...). Le modèle HST (Hydrostatique, Saisonnier, Temporel) permet de soustraire aux mesures brutes les influences réversibles pour mettre en évidence le comportement irréversible de l'ouvrage. Les résultats obtenus depuis de nombreuses d'années ont confirmé la pertinence et la robustesse de cette méthode. Néanmoins, certaines limitations du modèle HST ont été clairement identifiées : caractère indépendant des phénomènes explicatifs et lois d'évolution postulées a priori. Une nouvelle approche basée sur l'utilisation des réseaux de neurones va s'affranchir de ces inconvénients. Un réseau de neurones est un modèle permettant, à partir d'un ensemble de mesures en entrée, de modéliser une variable de sortie. Il est possible de modéliser une importante diversité d'évolutions car les réseaux de neurones ont la capacité de modéliser de manière automatique et par apprentissage les liens croisés entre phénomènes explicatifs. De plus, la fonction de transfert permet de s'ajuster à des lois d'évolution très variées (effets à seuil). Les résidus d'un modèle réseau de neurones sont plus faibles qu'avec un modèle HST : la détection éventuelle d'anomalies peut se faire de manière plus précoce. Les réseaux de neurones offrent une possibilité riche pour mettre au point de meilleurs modèles intégrant de nouvelles variables explicatives. La présentation illustre les applications des réseaux de neurones en comparant les performances avec le modèle HST. Les principales difficultés liées à l'utilisation des réseaux de neurones (calage numérique ou le choix de la base d'apprentissage) sont également discutées.

ABSTRACT

Statistical analysis methods for dam monitoring measurements allow separating effects from the various influencing factors, mainly hydrostatic load and thermal strains. The HST (Hydrostatic, Seasonal, Time) was developed by EDF in 1967. This method enables calculating hydrostatic effect, thermal seasonal effect and irreversible changes. Many years of experience confirmed the relevance and efficiency of this method. However, some limitations were also clearly identified: the three calculated effects must be independent and the physical law governing each effect has to be defined before modeling. A new approach based on Artificial Neural Networks (ANN) was tested to analyze dam monitoring measurements (displacement, pressure, flow rate, ...). ANN evidenced the connections between the different effect laws. As an example, hydrostatic effect was modeled both in winter and summer in order to establish the relationships between hydrostatic and thermal loads. A major advantage of ANN relied on the possibility to include any new influencing factor without the need for postulating its physical law. ANN determined and refined the most appropriated

¹ Auteur correspondant

law. In this paper, we use different dam measurement analyses to evaluate ANN and HST methods. Compared to HST, ANN analysis significantly reduces model residuals. Consequently, anomalies (behavior or measurement errors) are detected more rapidly with ANN. In summary, ANN provides a refined and detailed analysis, also consistent with HST. Also discussed, difficulties in applying ANN method.

1. INTRODUCTION

L'analyse des données d'auscultation des barrages (déplacements, fuites, sous-pression, ...) fait appel à des méthodes d'analyse statistique plus ou moins élaborées permettant de séparer les différentes sources de variation du phénomène mesuré. Parmi ces méthodes statistiques, l'analyse par régression linéaire multiple avec le modèle HST (**H**ydrostatique **S**aisonnier **T**emporel) s'est montrée puissante et est largement utilisée à EDF. Ses principaux avantages sont la robustesse ainsi que la facilité d'interprétation et d'utilisation. HST permet de déterminer les effets hydrostatiques (dus à la charge d'eau), les effets saisonniers (dus aux effets thermiques par exemple) en n'en déduire simplement le comportement irréversible de l'ouvrage.

Certaines limitations d'HST ont été identifiées : lorsque les variables explicatives (cote, saison) sont liées par des relations de dépendance, la séparation des effets est imprécise. De plus les lois d'évolution des phénomènes explicatifs (polynomial pour la cote) sont postulés a priori et ne permettent pas toujours de s'adapter à la réalité (effets de seuil entre autres).

Pour pallier ces insuffisances de l'analyse classique, des modèles explicatifs à base de réseaux de neurones ont été testés. Après avoir expliqué le fonctionnement du réseau de neurones, ses avantages par rapport à la régression linéaire seront présentés. Puis des exemples d'application à des données d'auscultation de barrages seront détaillés et comparés avec les résultats d'HST. Les mesures d'auscultation testées avec les réseaux de neurones concernent respectivement les débits de fuite, les sous-pressions et les déplacements.

2. LE CONCEPT DE RESEAUX DE NEURONES

2.1 Rappel sur la méthode HST

La méthode de correction des données HST est le principal modèle d'analyse des mesures d'auscultation utilisé à EDF pour l'analyse des données d'auscultation (déplacements, sous-pressions, fuites, ...). Le modèle HST considère que les mesures brutes peuvent être modélisées par une superposition de trois états :

- Une évolution irréversible du phénomène dans le temps « t », qui peut avoir tendance à s'amortir (adaptation ou consolidation), à s'accélérer (dégradation) ou à se poursuivre régulièrement. La loi correspondant à cet effet est la suivante :

$$(T) = b_1 e^{-t/t_0} + b_2 \cdot t + b_3 \cdot t^2 + b_4 \cdot t^3 + b_5 \cdot t^4 \quad (1)$$

Elle comprend deux termes (exponentiel négatif de temps caractéristique t_0 représentant l'évolution amortie, et un polynôme du temps). En pratique seul le terme linéaire est couramment employé.

- Une influence réversible correspondant à l'effet de la charge hydrostatique « H ». La loi correspondante est du type :

$$(H) = b_6 \cdot Z + b_7 \cdot Z^2 + b_8 \cdot Z^3 + b_9 \cdot Z^4 \quad (2)$$

Où

$$Z = \frac{RN-H}{RN-Rvide} \quad (3)$$

avec H cote de retenue du jour de la mesure, RN cote de retenue normale et Rvide cote à vide. (A barrage plein $Z=0$)

- Une influence réversible saisonnière qui inclut différents phénomènes cycliques (température, rayonnement, ...). La loi saisonnière est une fonction périodique du temps (période d'un an). La variable retenue est alors la saison S assimilée à un angle valant 0° le 1er janvier et 360° le 31 décembre ; on a alors :

$$(S) = b_{10} \cdot \cos(S) + b_{11} \cdot \sin(S) + b_{12} \cdot \cos(2S) + b_{13} \cdot \sin(2S) \quad (4)$$

On cherchera à caler avec une régression linéaire multiple et par les moindres carrés les coefficients b_i (i variant de 0 à 13) permettant d'expliquer la mesure X_j à l'instant t_j , avec un remplissage Z_j , et une saison S_j :

$$X_j = f(t_j) + f(Z_j) + f(S_j) + b_0 + \epsilon_j \quad (5)$$

Le terme ε_j représente le résidu du modèle, intégrant les erreurs expérimentales et l'imprécision du modèle. b_0 est la constante de la régression linéaire. Ainsi, les mesures dites « corrigées » par le modèle HST sont désinfluencées des effets réversibles (Z et S). De ce fait, on note :

$$\text{Mesure corrigée (f(t) +)} = \text{Mesure brute (X)} - f(Z) - f(S) \quad (6)$$

Le modèle fait une hypothèse forte : les influences sont indépendantes (additives). La saison a une même amplitude à barrage plein ou vide et réciproquement. De ce fait les éventuels effets croisés ne sont pas pris en compte.

Le modèle permet de comparer des données à conditions identiques de saisons et de cotes et ainsi de mettre en évidence les effets temporels lents et les anomalies.

2.2 Principe de fonctionnement du réseau de neurones

Le réseau de neurones est un outil de modélisation puissant. Il est utilisé dans des domaines aussi divers que les estimations boursières, la météorologie, le tri postal ou l'approximation de fonctions. Des applications [3],[4],[5] pour l'auscultation des barrages ont été testées, essentiellement pour l'analyse des déplacements des barrages en béton.

Un réseau de neurones se compose de plusieurs neurones interconnectés entre eux. Les neurones sont répartis sur des couches, en général trois (figure 2). Pour comprendre le fonctionnement du réseau de neurones il faut d'abord s'intéresser au fonctionnement d'un unique neurone. Chaque neurone fonctionne de la même manière (figure 1). Sur les schémas neuronaux, les entrées sont à gauche et la sortie (unique) est à droite. S'il y a plusieurs branches en sortie d'un neurone, c'est la même valeur numérique qui est copiée dans chacune des branches. A partir des valeurs numériques en entrée, le neurone fait dans un premier temps une combinaison linéaire des entrées. Les coefficients de la combinaison linéaire sont appelés poids. Une fonction de transfert (figure 3) est ensuite appliquée à cette valeur numérique. Cette fonction peut être classiquement la fonction sigmoïde (fonction tangente hyperbolique) ou une fonction linéaire. La valeur numérique en sortie est ensuite propagée vers un ou plusieurs neurones de la couche suivante.

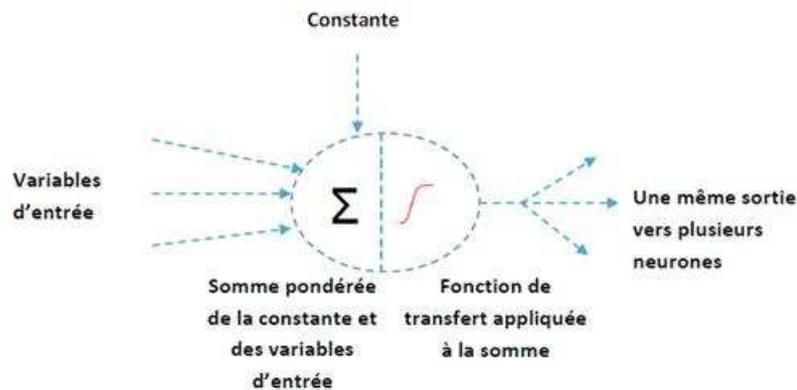


Figure 1: Fonctionnement d'un neurone

Le réseau de neurones à trois couches utilisé dans cet article se compose d'une couche d'entrée, d'une couche cachée et d'un neurone de sortie. Chaque neurone de la première couche a pour variables d'entrée toutes les variables explicatives du modèle (cote, saison, temps, pluie, neige...). De même, chaque neurone de la deuxième couche reçoit la sortie de tous les neurones précédents. Enfin, le neurone de sortie fait la somme pondérée du résultat des neurones de la couche cachée et obtient la prévision du modèle.

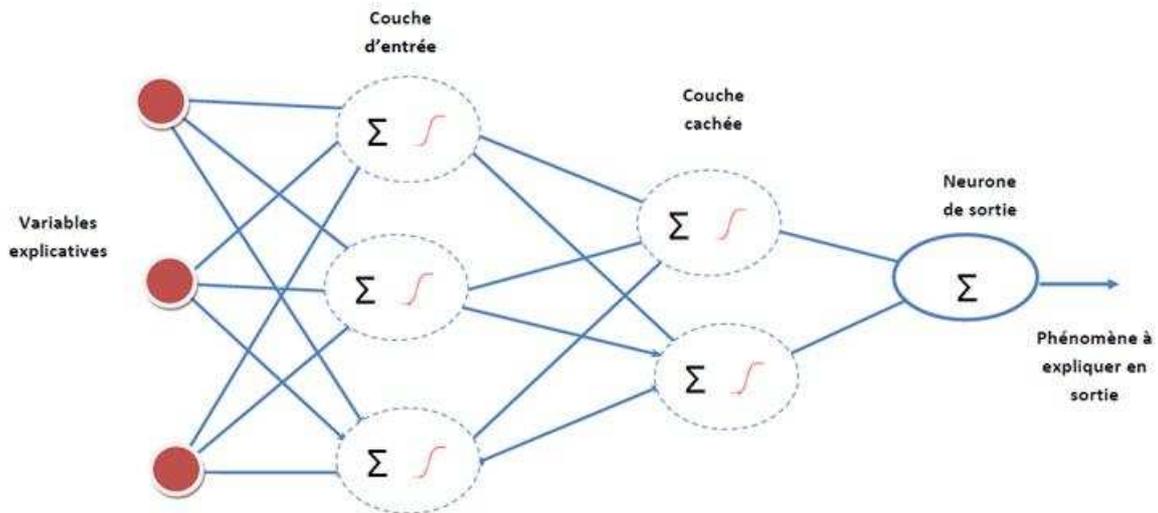


Figure 2 : Réseau de neurones à trois couches (entrée, cachée et de sortie)

Le calage du modèle s'effectue par ajustement itératif des poids dans les branches pour minimiser l'écart entre la valeur modélisée et la mesure. Les poids sont initialisés aléatoirement et ajustés itérativement. Il existe plusieurs algorithmes mathématiques pour ajuster les poids. La modification des poids à chaque itération se fait d'aval vers l'amont en modifiant légèrement les poids par un facteur déduit grâce à la dérivée de la fonction de transfert et d'un terme inertiel pour éviter les minimums locaux.

Le calage s'effectue sur une base d'apprentissage, par exemple une valeur sur deux, et calcule en même temps l'erreur de la base test (l'autre moitié). Lors du calage, les erreurs sur les deux bases sont sensiblement proches. Si l'erreur sur la base de test augmente au fil des itérations alors que l'erreur sur la base d'apprentissage continue à décroître cela signifie que le réseau de neurone est en **sur-apprentissage**. En d'autres termes le réseau de neurones commence à « apprendre du bruit » sur la base d'apprentissage. Lorsqu'il essaie d'appliquer ce bruit sur la base de test, le modèle se dégrade brutalement et l'erreur sur la base de test augmente sensiblement. Une des difficultés consiste à trouver le bon nombre d'itérations qui minimise les deux erreurs. Il s'agit donc d'arrêter le calage avant le sur-paramétrage en surveillant en parallèle l'évolution de l'erreur sur les bases apprentissage et test. En pratique, lorsqu'il y a peu de mesures ou lorsque le coefficient de corrélation est faible, le sur-apprentissage est plus facile à atteindre.

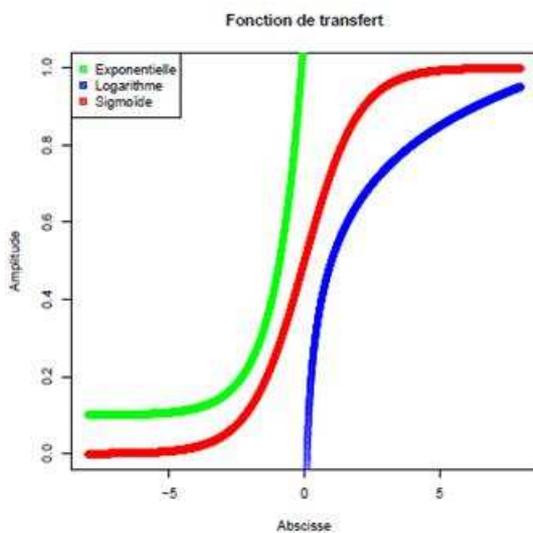


Figure 3 : Propriétés de la fonction sigmoïde de transfert

$$f(x) = \frac{1}{1 + e^{-x}}$$

Avec des micro-ordinateurs de bureautique (non dédiés aux calculs) et en considérant un millier de mesures (réparties pour moitié en base d'apprentissage et pour moitié en base de test) le calage du réseau de

neurones (sur la base de 100000 itérations) dure environ 300 s. Le temps de calcul est sommairement proportionnel au nombre de mesure et au nombre d'itérations.

2.3 Potentialités offertes par les réseaux de neurones

Contrairement aux polynômes, la fonction sigmoïde présente des asymptotes horizontales. Ces asymptotes sont très utiles car elles permettent de représenter des phénomènes bornés comme la piézométrie (en valeur supérieure par exemple) ou les fuites qui s'annulent en dessous d'une certaine cote.

La puissance du réseau de neurones se trouve également dans la fonction de transfert non linéaire, qui permet d'imaginer des produits complexes entre les paramètres. Par exemple supposons un réseau très compact à deux couches, avec deux neurones sur la première, un sur la sortie. Appelons x et y les variables d'entrée. La première couche applique à x et y la partie droite (bleue sur la figure 3) de la sigmoïde qui approxime un logarithme. Un neurone de la 2ème couche peut en faire la somme et leur appliquer, à l'inverse, l'approximation de l'exponentielle (en rouge). Le tableau 1 récapitule l'ensemble des transformations subies par x et y durant leur passage dans le réseau de neurone.

Neurone concerné	Action dans le neurone	Équation
Neurone de la couche d'entrée	Entrée	x et y
	Pondération	$a \cdot x + b \cdot y + c$
	Transfert	$\ln(a \cdot x + b \cdot y + c)$
Neurone de la couche de sortie	Pondération	$g \cdot \ln(a \cdot x + b \cdot y + c) + h \cdot \ln(d \cdot x + e \cdot y + f) + i$
	Transfert	$\exp(g \cdot \ln(a \cdot x + b \cdot y + c) + h \cdot \ln(d \cdot x + e \cdot y + f) + i)$
	Sortie	$\exp(i) \times (a \cdot x + b \cdot y + c)^g \times (d \cdot x + e \cdot y + f)^h$

Tableau 1 : Puissance d'un réseau de neurones très compact, avec les paramètres $a, b, c, d, e, f, g, h, i$. Si $i=b=c=d=f=0$ le produit $x^g y^h$ est modélisé en sortie.

Le nombre de paramètres pour ce réseau très compact (3 neurones) est de 9. Il est important de remarquer qu'avec très peu de neurones et uniquement 9 paramètres, les possibilités de combinaison sont déjà très variées. Cet exemple n'est qu'un cas particulier en considérant uniquement des portions de courbe de la sigmoïde, et sans prendre en compte les possibilités offertes par les asymptotes.

Ainsi un réseau de neurones peut facilement se construire à partir d'une variable explicative z (creux relatif) et des variables explicatives z^2, z^3 (et suivantes si nécessaire). Il est aussi possible d'imaginer un effet hydrostatique z qui n'est pas le même suivant la saison S , permettant ainsi un degré de liberté supplémentaire par rapport à HST. L'introduction d'une nouvelle variable explicative au modèle se trouve également simplifiée puisqu'il n'est pas nécessaire de connaître *a priori* sa loi d'évolution. Par contre ce type de réseau de neurones ne permet pas de prendre en compte l'effet retardé d'une variable explicative (température, pluie par exemple). Il est nécessaire de retarder cette variable avant de l'intégrer au réseau de neurones.

3. APPLICATION A L'ANALYSE DES SOUS-PRESSIONS

3.1 Problématique et lacunes du modèle HST

Les sous-pressions mesurées dans les barrages voûte au niveau du contact béton-rocher et à l'amont sont en général très influencées par la cote amont. Par application du modèle HST l'effet hydrostatique calculé est un effet moyen qui est identique quelle que soit la saison. En réalité, une analyse plus fine des données montre que cet effet hydrostatique n'est pas le même en été et en hiver. La zone dans laquelle les sous-pressions sont mesurées est une « zone fracturée » et son état de fracturation va dépendre de la position du barrage : en été, avec la dilatation du béton, le barrage prend une position à l'amont et la fissuration en pied d'ouvrage a

tendance à se refermer. Réciproquement en hiver, avec la contraction du béton, le barrage se positionne plus à l'aval et la fissuration en pied amont est plus importante. Cet effet n'est pas intégré par le modèle HST qui considère un effet hydrostatique moyenné sur l'ensemble des saisons. Il y a donc un intérêt évident à utiliser ici les réseaux de neurones qui ont la capacité à créer des effets hydrostatiques différenciés suivant les saisons. Sur le même principe l'effet saisonnier n'est pas forcément identique à RN ou à Rvide. Un réseau de neurones aura la capacité à saisir ce lien croisé entre la saison et la cote.

3.2 Exemple d'application et comparaison HST versus réseau de neurones sur le barrage de Pareloup

La barrage de Pareloup est un barrage voûte de 45 m de hauteur à simple courbure. Dans cet exemple on s'intéresse au piézomètre PPG4 (figure 4). Ce piézomètre est situé en fondation superficielle et ses mesures sont sensibles à la charge hydrostatatique et à l'état de fissuration de la fondation. Le réseau de neurones (figure 5) utilisera 4 variables explicatives en entrée : le temps (t), la saison (cos(S) et sin(S)) ainsi que la cote de retenue (z). L'architecture se compose de 8 neurones : 4 en couche d'entrée, 3 en couche cachée et 1 neurone en sortie. Au total ce réseau comprend 39 paramètres (les poids des branches), à comparer aux 10 paramètres d'HST. Au bout de 200000 itérations le réseau de neurones converge vers une solution qui est numériquement plus performante qu'HST, ce qui est logique étant donné les degrés de liberté utilisés par les neurones. Il est toujours indispensable de s'assurer qu'il n'y pas de surapprentissage (apprentissage d'un bruit sans lien avec les variables explicatives d'entrée) en comparant les erreurs sur les bases d'apprentissage et de test. La figure 6 présente l'évolution de l'erreur de modélisation sur les bases d'apprentissage et de test. L'évolution parallèle et décroissante des deux courbes démontre qu'il n'y a pas eu de sur-apprentissage.

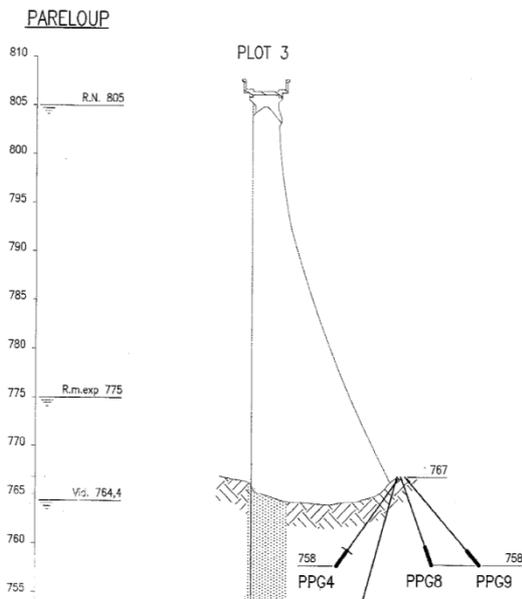


Figure 4 : positionnement du piézomètre PPG4 sur le barrage de Pareloup

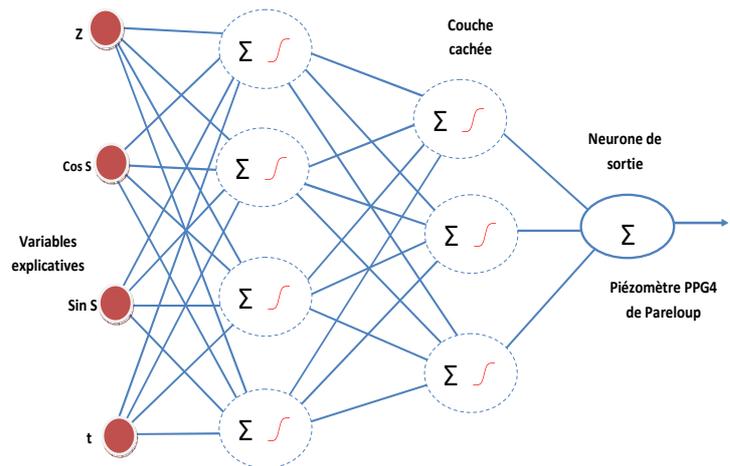


Figure 5 : architecture du réseau de neurones dit "compact" utilisé pour l'analyse du piézomètre PPG4

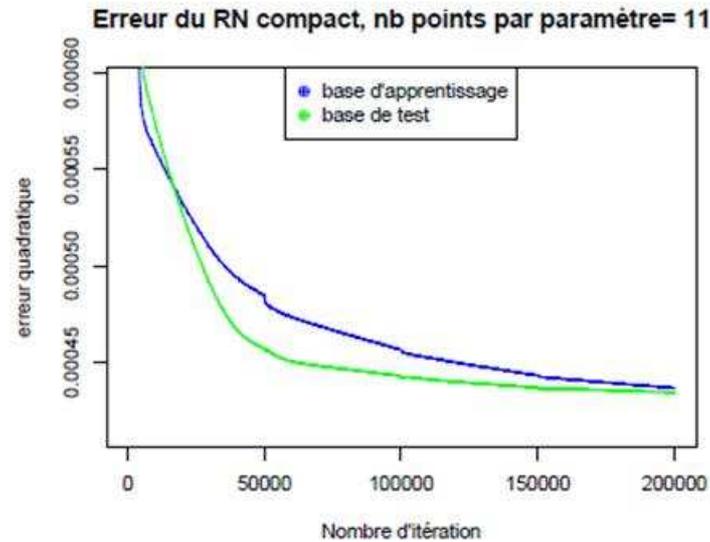


Figure 6 : évolution de l'erreur quadratique sur la base d'apprentissage et la base de test. La sortie du réseau de neurones est normée entre 0 et 1 ce qui explique ici les très faibles valeurs de l'écart quadratique moyen.

L'écart-type des résidus de modélisation vaut 0,30 m pour les neurones, contre 0,56 m avec la correction HST ce qui représente une réduction de dispersion de l'ordre de 40%. Cette performance du réseau de neurones est bien illustrée sur la figure 7 ci-dessous qui présente l'allure des données brutes et corrigées.

Données corrigées

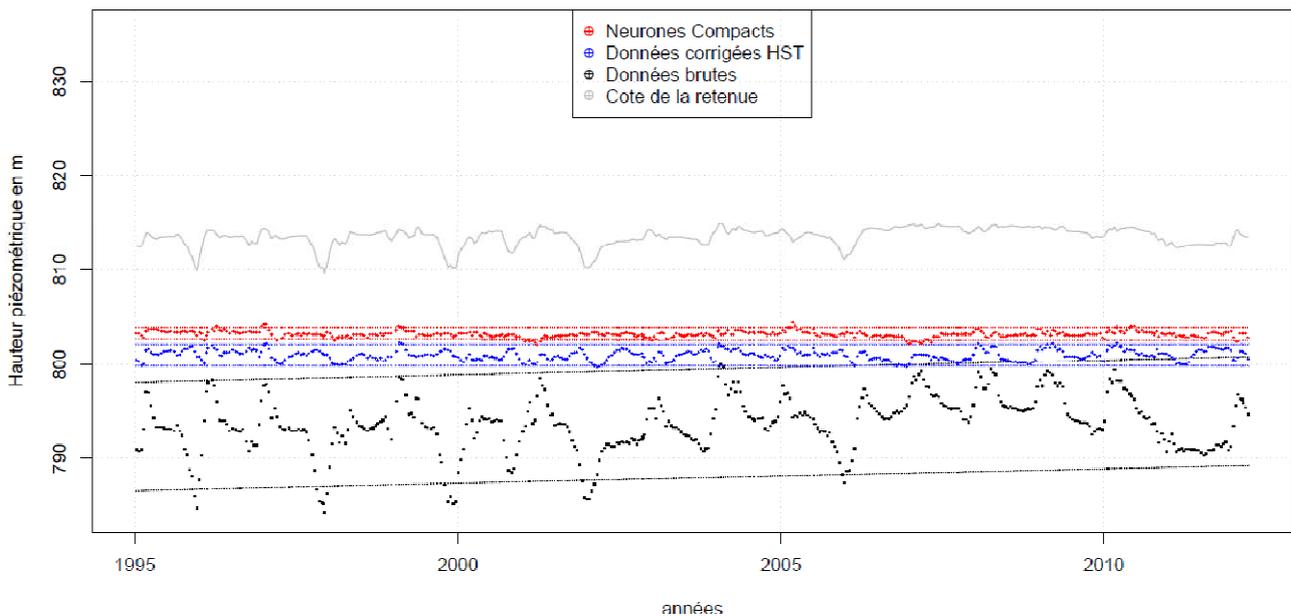


Figure 7 : Evolution comparée des données corrigées HST et neurones ; il est intéressant de remarquer que les données corrigées HST (bleu) présentent encore un aspect cyclique (saisonnier) après correction dû à la non prise en compte de l'effet hydrostatique dépendant de la saison. Cet aspect cyclique a complètement disparu avec les neurones.

Au-delà de la réduction attendue de la dispersion, il faut analyser quels facteurs explicatifs ont été déterminés par les neurones pour améliorer le modèle. Pour afficher un effet hydrostatique avec les neurones il suffit de faire varier z de 1 (barrage vide) à 0 (barrage plein). Il faut également préciser une saison ($\cos(S)$ et $\sin(S)$) ainsi qu'un instant (variable t). La figure 7 présente l'effet hydrostatique (en 2003) à différentes saisons :

- Janvier pour la période décembre-février correspondant à l'hiver ;
- Avril pour la période mars-mai correspondant au printemps ;
- Juillet pour la période juin-août correspondant à l'été ;
- Octobre pour la période septembre-novembre correspondant à l'automne.

Afin d'afficher des effets calculés proches des mesures, le creux relatif z est pris parmi les cotes de l'ouvrage rencontrées à chaque saison. Ainsi lorsqu'il y a une grande densité de points (par exemple les points rouges

entre la cote 798 m et 804 m en avril figure 8) on considèrera que le réseau de neurones est bien calé dans cette zone car il s'appuie sur un nombre important de mesures (courbe avril : mesures du mois de mars à juin). A l'inverse les deux points rouges isolés à 797 m et 795,5 m doivent être interprétés avec précaution : sur la période mars-mai seules deux mesures sont présentes à une cote inférieure à 798 m. Bien que le réseau de neurones soit capable de tracer un effet hydrostatique à cote basse (sous 796 m) cet effet extrapolé ne sera pas représenté graphiquement.

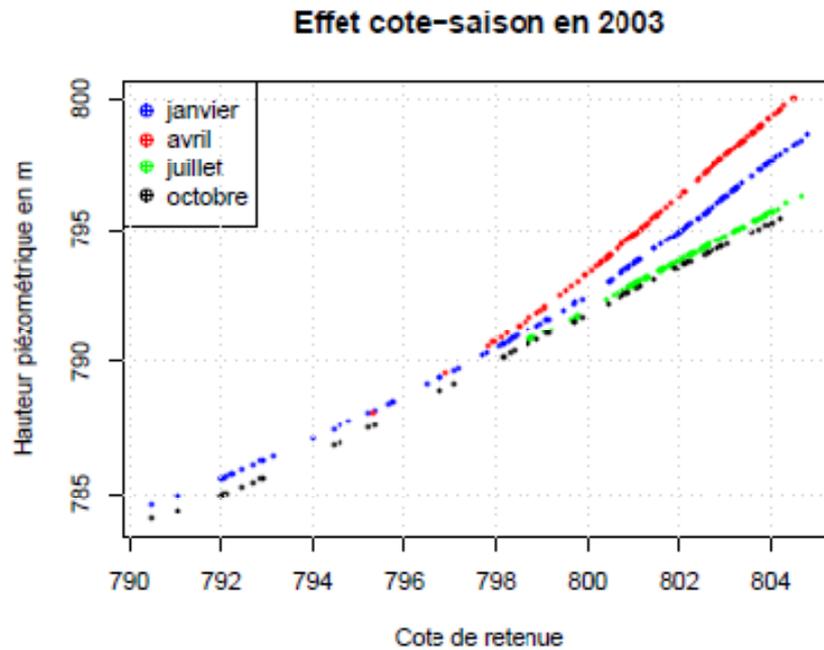


Figure 8 : Evolution de la hauteur piezométrique en fonction de la saison calculé par réseau de neurones. A cote pleine, il existe une différence d'environ 5 m suivant l'état thermique (saison) du barrage.

L'effet hydrostatique calculé par les neurones met clairement en évidence un couplage hydrostatique : en avril lorsque le barrage est le plus froid la fissuration en fondation atteint son maximum et la charge d'eau atteint plus facilement le piézomètre PPG4 : la figure 8 montre un effet hydrostatique avril (en rouge) qui présente la plus forte pente. La pente est également importante en janvier (courbe bleue) mais moins qu'en avril à cause de l'inertie thermique car le barrage n'a pas encore eu le temps de subir son refroidissement hivernal maximum. A l'inverse à la fin de l'été le barrage a sa dilatation maximum et l'état de fissuration en fondation est donc minimum. Les pentes des effets hydrostatiques en juillet et octobre sont les plus faibles. Bien qu'il y ait peu de points en-dessous de la cote 798 m il est intéressant de noter que l'effet hydrostatique semble indépendant de la saison sous cette cote. Il est toutefois difficile de confirmer rigoureusement cette constatation étant donné le faible nombre de situations où la cote est inférieure à 798 m.

Cet effet hydrostatique différencié suivant les saisons explique la meilleure capacité prédictive des neurones par rapport à HST. L'effet hydrostatique calculé par HST est un effet moyenné entre les saisons, il correspond sommairement à l'effet hydrostatique calculé en janvier.

Le même effet peut être observé sur la fonction saisonnière : la figure 9a présente l'évolution de l'effet saisonnier en fonction du remplissage. L'effet saisonnier apparaît plus important lorsque le réservoir est plein ce qui est logique car le chargement hydrostatique a tendance à maximiser l'état de fracturation de la fondation et permet ainsi plus facilement la transmission de la charge. Lorsque l'ouvrage est peu rempli (dernier quartile) la fissuration de la fondation est quasiment refermée et l'effet saisonnier est insignifiant. Il est rassurant de remarquer que l'effet saisonnier calculé par le modèle historique HST (Figure 9b) a la même allure. Le saisonnier d'HST correspond environ au saisonnier des neurones pris à cote médiane.

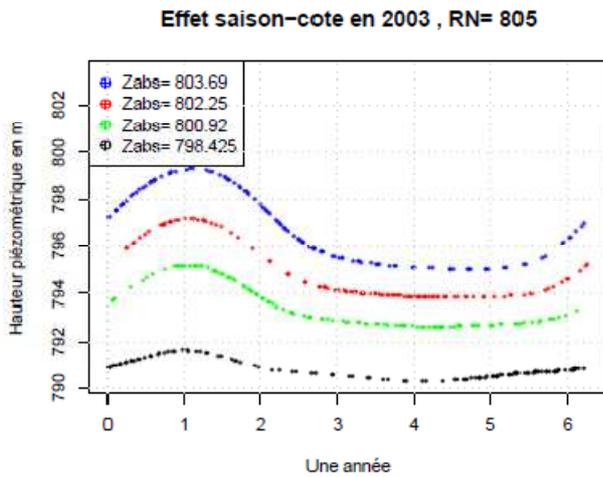


Figure 9a : évolution de l'effet saisonnier en fonction du remplissage. Les cotes indiquées dans la légende correspondent aux valeurs médianes des 4 quartiles.

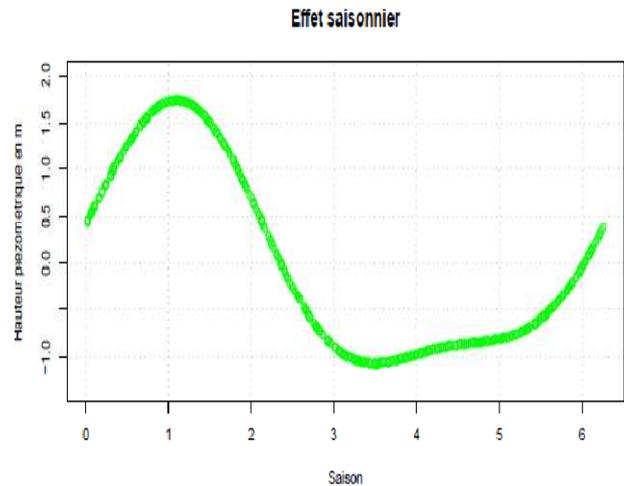


Figure 9b : Effet saisonnier calculé par HST.

Les neurones ont la capacité de croiser l'ensemble des variables explicatives. Il est important de vérifier qu'il n'existe pas d'autres liens croisés entre variables explicatives comme le temps par exemple. La figure 10 démontre que les effets hydrostatiques et saisonniers sont indépendants de l'évolution temporelle. Les effets calculés sont identiques en 1997, 2003 et 2009.

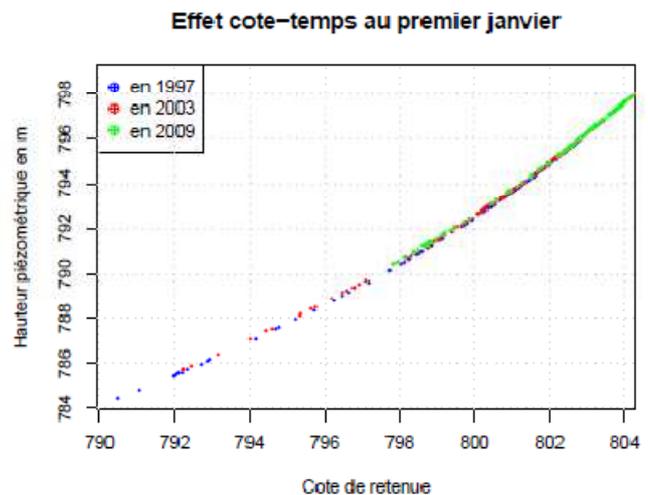
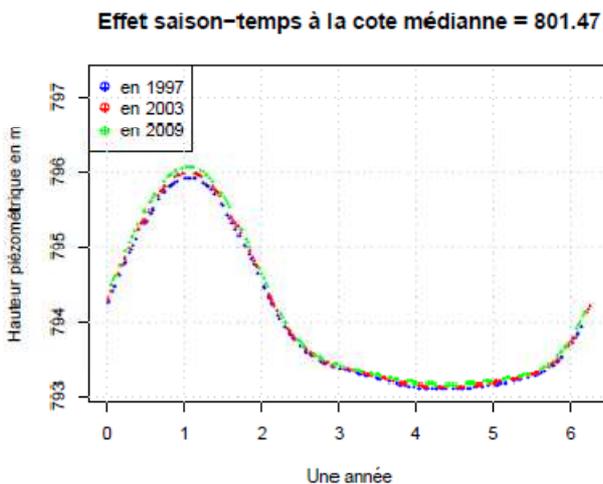


Figure 10 : Effets saisonniers et hydrostatiques calculés par les neurones à 3 instants : 1997 (début de calage), 2003 et 2009 (fin de calage). On notera que ces effets restent constants dans le temps.

3.3 Modèle à termes croisés imposés

Des modèles physico-statistiques peuvent être conçus spécifiquement pour l'analyse des sous-pressions. Un modèle intéressant qui intègre explicitement le couplage cote-saison est décrit ci-après.

La sous-pression mesurée dépend du produit de la charge d'eau et d'une longueur fissurée :

$$NP_{brut} = NP_{aval} + \frac{(RN - R_{vide}) \cdot (1 - z) \cdot (d + LF)}{Eb} \quad (7)$$

Où :

NP_{brut} : niveau piézométrique brut sous la voûte au point considéré

NP_{aval} : niveau piézométrique à l'aplomb du parement aval

d : distance entre le piézomètre et le parement aval (ici d = 5,8 m pour l'analyse de PPG4)

LF : longueur qui représente l'état de fissuration sous l'ouvrage

E_b : épaisseur du barrage à la base du plot considéré (ici E_b = 15,4 m pour l'analyse de PPG4)

Les autres variables (z, RN, R_{vide}) ont la même signification qu'au paragraphe décrivant HST (§2.1)

Cette longueur fissurée (LF) dépend elle-même d'un effet hydrostatique, d'une variation saisonnière et d'une évolution irréversible dans le temps :

$$LF = L0 + \xi(z) + \xi(S) + \xi(t) \quad (8)$$

Où :

$L0$: longueur fissurée à l'instant initial, barrage plein et saison nulle

$\xi(z)$: effet hydrostatique expliquant l'ouverture de la fissure

$\xi(S)$: effet saisonnier expliquant l'ouverture de la fissure

$\xi(t)$: effet irréversible expliquant l'ouverture de la fissure

L'expression algébrique des 3 fonctions ξ est identique à HST (§2.1).

Le calage de ce modèle s'effectue en minimisant les carrés des écarts modèle-mesure. Le couplage entre la cote de retenue et la saison est ici explicite par l'intermédiaire du terme $(1-z)$ en facteur de la longueur fissurée. L'effet hydrostatique avec ce modèle à termes croisés est présenté sur la figure 11. L'effet hydrostatique maximum en hiver est retrouvé et l'effet saisonnier est également maximum à cote pleine. L'écart-type des résidus avec le modèle croisé vaut 0,33 m, valeur très proche du modèle neuronal (0,30 m). Les deux modèles développés sont donc bien cohérents pour l'analyse des sous-pressions.

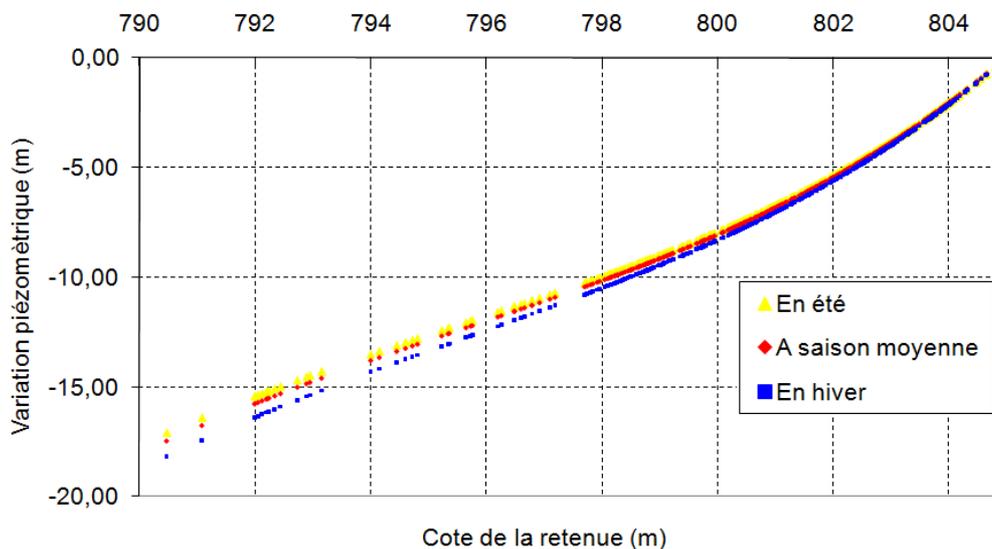


Figure 11 : effet hydrostatique calculé par le modèle à terme croisés. L'effet hydrostatique est bien différencié suivant les saisons, avec un effet plus marqué en hiver (idem réseaux de neurones). L'amplitude été-hiver est ici moindre qu'avec les neurones car le modèle croisé impose des courbes homothétiques.

3.4 Conclusions sur l'analyse des sous-pressions

Les sous-pressions présentent sur certains ouvrages un effet hydrostatique qui dépend de l'état saisonnier (thermique) de l'ouvrage. Cet effet a été bien vu et modélisé par les neurones et le modèle à termes croisés. La justesse des données corrigées se trouve améliorée et l'appréciation des effets est affinée suivant la saison. Ces résultats ne sont pas en désaccord avec HST puisqu'il est possible de retrouver les mêmes effets hydrostatiques en avril en calant un modèle HST uniquement sur les mesures d'avril. De même l'effet hydrostatique d'HST est un effet moyenné entre toutes les saisons ; en moyennant les effets hydrostatiques des neurones ou du modèle à terme croisés sur les différentes saisons c'est bien l'effet hydrostatique d'HST qui est retrouvé. Ainsi les réseaux de neurones et le modèle à termes croisés viennent enrichir l'analyse HST sans entrer en contradiction avec le modèle historique.

4. APPLICATION A L'ANALYSE DES FUITES

4.1 Problématique de l'analyse des fuites

Les débits de fuites sont quelquefois mal corrigés par HST car il est possible qu'il existe dans certains cas des effets à seuil (déversement à partir d'une certaine cote) qui sont mal modélisés par des polynômes. L'influence de la pluie ou de la fonte des neiges n'est pas modélisée par HST, ce qui contribue également à

une dispersion importante des données corrigées. L'influence de la pluie ou de de la neige n'est pas connue *a priori*. Le réseau de neurones peut aider à déterminer la loi de l'effet correspondant (polynomial, à seuil, ...) et à détecter d'éventuels effets croisés (avec la saison par exemple).

4.2 Analyse des fuites du barrage de Bissorte

4.2.1 Travail préliminaire sur les données de pluie et de neige

Un réseau de neurones est capable de modéliser une grande variété de fonctions mais le type de réseau de neurones présenté en 2.2 n'est pas capable de modéliser un retard. L'effet de la pluie ou de la neige sur les débits de fuite n'est pas immédiat à cause de phénomènes lents d'infiltration et d'alimentation de nappes de versant. Par conséquent il faut retarder les mesures journalières de pluie ou de fonte des neiges. La neige est mesurée grâce à un NRC² situé à Bissorte à une altitude équivalente au barrage. La fonte a été construite comme la différence de teneur en eau entre deux jours consécutifs. Il est rapidement apparu que les données de fonte comme de pluie sous leur forme brute n'étaient pas assez pertinentes de part leur caractère assez irrégulier et très variable dans le temps. Cette donnée de fonte doit servir plutôt à définir un ordre de grandeur de la période et de l'intensité de la fonte printanière. Un lissage a été réalisé à l'aide d'un filtre passe-bas en transformée de Fourier rapide (FFT). La fréquence de filtrage a été optimisée par itérations sur le coefficient de corrélation obtenu. Un retard de même type que celui utilisé pour l'analyse des cellules de pression interstitielle dans les remblais a été testé. Ce retard transforme une variable instantanée (pluie ou neige) en une variable retardée via le produit de convolution suivant :

$$Xr(t) = \int_0^{+\infty} X(t) \cdot P(t - \tau) d\tau \quad (9)$$

Où X(t) désigne la variable instantanée (pluie ou neige) et Xr(t) désigne la variable retardée.

La fonction P(t) dite de réponse impulsionnelle du milieu s'inspire de l'équation de diffusion de la chaleur en milieu semi-fini et s'écrit :

$$P(t) = \sqrt{\frac{T_0}{\pi}} \cdot \frac{e^{-\frac{t}{T_0}}}{t^{\frac{3}{2}}} \quad (10)$$

Avec T₀ un temps caractéristique de diffusion dans le milieu.

La figure 12 présente les résultats du lissage et du retard sur la fonte de Bissorte.

Neige : Lissage FFT et le retard d = 1

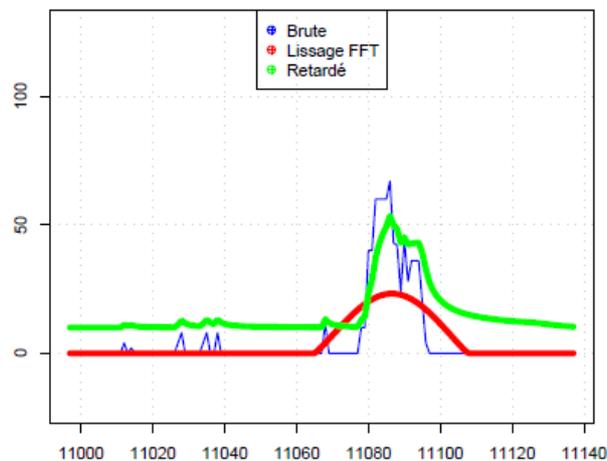


Figure 12 : lissage et retard appliqués à la fonction de fonte à Bissorte. L'abscisse représente le temps (une année de mesure) et l'ordonnée représente la fonte (baisse de teneur en eau du sol entre deux jours consécutifs, valeur nulle si accumulation ou stagnation du manteau neigeux)

² Nivomètre à Rayonnement Cosmique : permet la mesure de la masse en eau d'un manteau neigeux posé sur le sol en comptant l'atténuation du rayonnement cosmique

4.2.2 Premiers résultats obtenus grâce aux réseaux de neurones

Les premiers réseaux de neurones testés utilisent les variables d'HST en entrée auxquelles on rajoute la pluie retardée, la fonte (lissée) et éventuellement la température de l'air. Le coefficient de corrélation augmente significativement pour la fuite aval rive gauche puisqu'il passe de 0,45 (HST) à 0,60 (neurones).

Il est intéressant de noter que les effets calculés par les neurones pour la pluie et la neige sont linéaires. Leur ordre de grandeur est au moins aussi important que l'effet saisonnier, ce qui indique que ces variables ont un caractère explicatif certain. L'effet de la température semble négligeable. Partant du constat que les effets neige et fonte sont linéaires, il apparaît opportun de se demander si un modèle de régression linéaire plus simple utilisant simplement la neige et la fonte retardées pourrait également convenir. C'est l'objet du paragraphe suivant.

4.2.3 Application d'un modèle de régression linéaire inspiré des résultats des réseaux de neurones

Il est aisé de compléter HST avec des informations de fonte et pluie de la manière suivante :

$$X_i = (t_i) + f(Z_i) + f(S_i) + b_{13} \cdot \text{FonteL} + b_{14} \cdot \text{PJR} + b_0 + \varepsilon_i \quad (11)$$

Avec :

FonteL : fonte lissée

PJR : pluie journalière retardée

b_{13}, b_{14} : coefficient de régression linéaire

Les autres variables sont définies au §2.1.

Ce modèle « HST PJR Fonte » a été testé sur l'ensemble des fuites du barrage de Bissorte. Les résultats sont synthétisés dans le tableau 2.

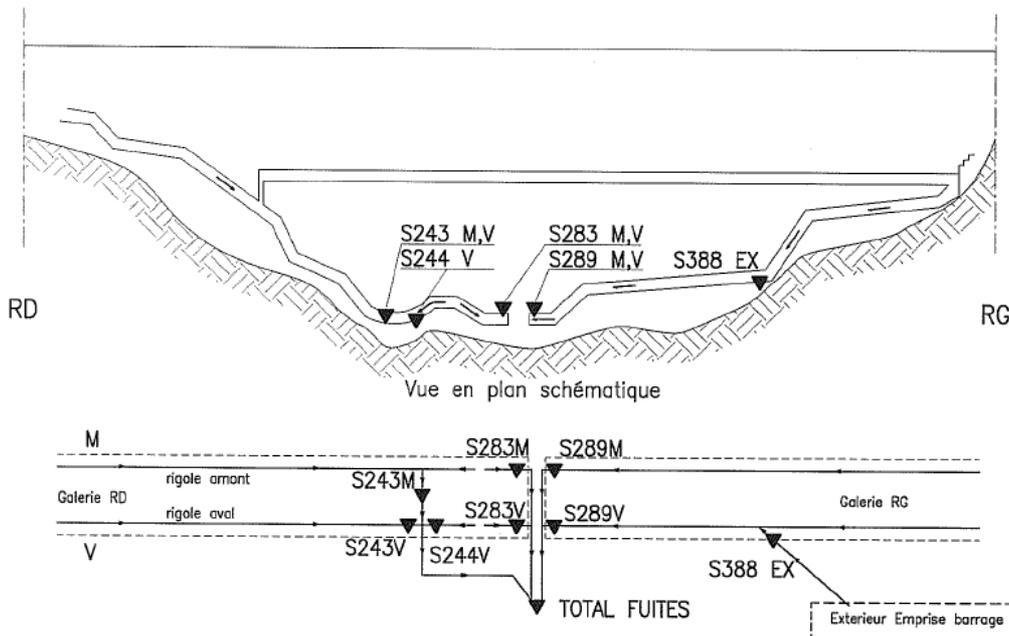


Figure 13 : Résumé de la partie fuites du dispositif d'auscultation du barrage de Bissorte

	Fonte	PJ	R ² ini	ΔS ²	ΔR ²
Amont Rive Droite	–	–	0.55	5 %	0.02
Amont Rive Gauche	–	–	0.7	5 %	0.02
Aval Rive Droite	fort	moyen	0.45	29 %	0.16
Aval Rive Gauche	fort	léger	0.45	32 %	0.17
Somme des quatre	moyen	léger	0.43	17 %	0.10

Tableau 2 : classement des fuites suivant leurs variables explicatives.

Le modèle classe clairement les fuites en deux groupes :

- les fuites de la rigole amont (drainage d'élévation dans le corps de l'ouvrage) qui ne sont pas influencées par la neige ou la pluie. (pas d'amélioration du coefficient de corrélation) ;
- les fuites de la rigole aval (drainage des rives) qui sont clairement influencées par la pluie et la neige (nette amélioration du coefficient de corrélation).

4.3 Conclusions sur l'analyse des fuites

Le modèle neuronal a permis ici de discriminer les variables explicatives d'entrée. La température extérieure ne présente pas de caractère explicatif pour les débits de fuite. Par ailleurs les formes très simples (linéaires) pour les effets de la neige et de la pluie ont permis de mettre au point un modèle simple de régression linéaire. Ce modèle présente des performances en terme de coefficient de corrélation qui sont tout à fait similaires au modèle neuronal. Le modèle neuronal a ainsi permis la mise au point d'un modèle simple et performant : « HST PJR Fonte ». Une sophistication d'« HST PJR Fonte » avec des effets plus complexes n'améliorerait pas les résultats puisque ce dernier possède déjà un coefficient de corrélation très proche du modèle neuronal. Les avantages du modèle simple « HST PJR Fonte » sont de plusieurs ordres :

- temps de calcul très rapide (quasi-instantané) ;
- intégration possible directement dans les outils industriels d'EDF d'analyse des données d'auscultation ;
- compréhension aisée du modèle et de ses effets qui permet d'envisager une utilisation en dépouillement ;

5. APPLICATION A L'ANALYSE DES DEPLACEMENTS

Les déplacements amont-aval des barrages sont également soumis à d'importants effets croisés entre l'effet thermique et la cote. Les mesures brutes de déplacement des barrages poids et voûtes sont actuellement corrigées grâce au modèle HSTThermique qui est une amélioration d'HST grâce à l'intégration d'une information réelle de température. L'expression algébrique du modèle est la suivante :

$$\mathbf{X}_j = (\mathbf{t}_j) + \mathbf{f}(\mathbf{Z}_j) + \mathbf{f}(\mathbf{S}_j) + \mathbf{b}_0 + \mathbf{b}_{13} \cdot \Delta\Theta_R^{TO} + \boldsymbol{\varepsilon}_j \quad (12)$$

Les notations sont identiques à la description du modèle HST (§2.1). $\Delta\Theta_R^{TO}$ désigne l'écart retardé par rapport à la normale saisonnière de la température de l'air. Le retard T_0 est choisi comme étant celui qui optimise la régression linéaire. Ce retard représente un temps de diffusion moyen de la chaleur à travers l'ouvrage. HSTThermique est décrit en détail dans [2].

Un réseau de neurones qui utilise les variables d'HSTThermique en entrée permet de réduire la dispersion des données corrigées de l'ordre de 80% par rapport à HST (HSTThermique permettait déjà une réduction de 50%). La figure 14 présente les données corrigées par les neurones du pendule en clé au barrage de Monteynard. La dispersion est très faible ce qui permet de détecter de manière précoce tout changement de comportement ou anomalie de mesure.

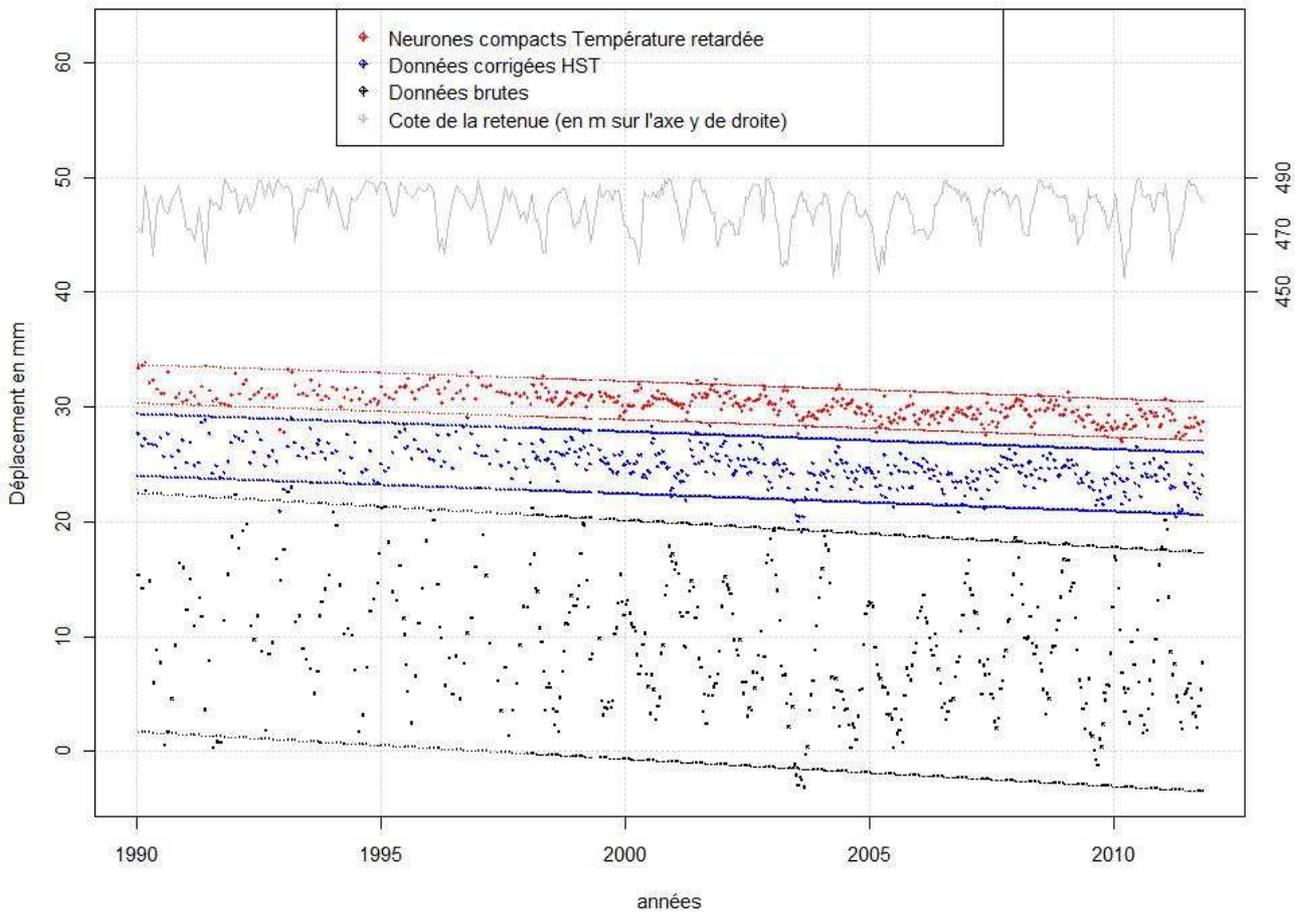


Figure 14 : déplacement amont-aval au plot central du barrage de Monteynard. Evolution dans le temps des données corrigées (neurones et HST).

Il est intéressant de rechercher quels sont les liens croisés que les neurones ont trouvé pour mieux expliquer la mesure. L'analyse systématique de tous les effets croisés (cote-saison, cote-temps, cote- $\Delta\Theta_R^{TO}$, $\Delta\Theta_R^{TO}$ -cote, ...) doit permettre de repérer celui qui est prépondérant et permet une meilleure explication.

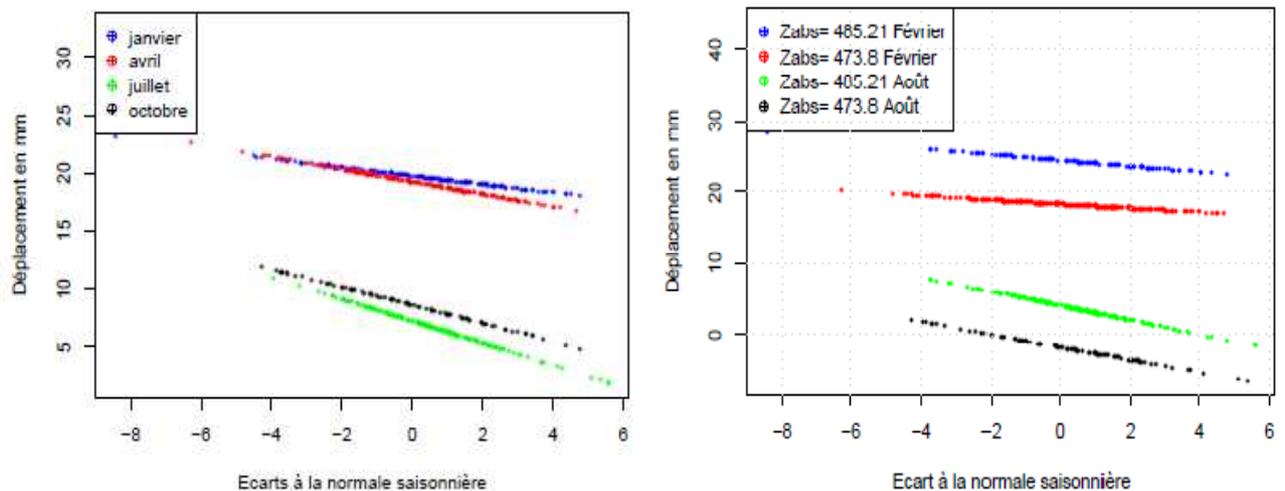


Figure 15 : effet croisés $\Delta\Theta_R^{TO}$ -saison et $\Delta\Theta_R^{TO}$ -cote. La figure de gauche permet d'affirmer que la sensibilité à la température de l'air retardée dépend de la saison. La figure de droite confirme ce lien tout en indiquant que la cote de retenue n'influe pas sur la sensibilité à la température de l'air retardée.

La figure 15 présente les effets croisés qui sont les plus significatifs. Là où HSTThermique prévoyait une sensibilité constante à la température de l'air retardée, le réseau de neurones trouve une sensibilité différente suivant la période de l'année : la sensibilité est plus importante en été (juillet et octobre) qu'en

hiver (janvier et avril). On peut être tenté de penser que cette sensibilité est en partie due à la position du plan d'eau, mais la figure 15 démontre également que le remplissage n'influe quasiment pas sur la sensibilité à la température de l'air retardée : en hiver, à la cote 485 m (médiane des valeurs hautes), la sensibilité est quasiment identique qu'à la cote 474 m (médiane des valeurs basses). Durant l'été la même observation peut être faite. Une hypothèse avancée pour expliquer cet effet est l'influence du rayonnement solaire qui n'est pas la même en été ou en hiver et qui viendrait s'ajouter à l'effet convectif de la température de l'air.

Ainsi les réseaux de neurones détectent des liens croisés entre variables explicatives et proposent des relations de dépendance. Ces relations de dépendance, après analyse approfondie, peuvent être utilisées pour mieux appréhender le comportement de l'ouvrage et proposer une amélioration à base physique des modèles existants.

6. CONCLUSIONS ET PERSPECTIVES

Les exemples d'application aux données d'auscultation des réseaux de neurones ont montré une mise en perspective nouvelle. Au-delà de la performance numérique qui était attendue il est intéressant de noter que des effets croisés entre variables explicatives sont nettement mis en évidence. Cette observation peut expliquer quelquefois les difficultés d'HST à corriger les données. Les neurones doivent être utilisés avec certaines précautions : les effets calculés sont à tracer à proximité des mesures ayant servi à les caler et l'exploration de tous les effets croisés possibles doit être examinée afin de comprendre comment les neurones ont amélioré la modélisation. Le problème du sur-apprentissage (avec du bruit de mesure par exemple) peut être aisément maîtrisé en surveillant l'erreur sur une base de test qui n'a pas vu d'apprentissage. En cas de sur-apprentissage, une baisse du nombre d'itérations ou une simplification de l'architecture du réseau de neurones (diminution des degrés de liberté) sont en général suffisants pour résoudre le problème. Les neurones ne sont pas une finalité en tant que telle puisqu'ils peuvent servir de base à l'élaboration de modèles physico-statistiques s'appuyant sur les effets calculés et liens croisés. Deux exemples de modèles simples ayant des performances identiques aux neurones ont été présentés pour les débits de fuite et les sous-pressions. Le caractère explicatif ou non d'une variable peut également être facilement déterminé sans postuler *a priori* sa loi d'évolution. Les neurones constituent ainsi un outil puissant d'exploration des données qui permet de faire une analyse exhaustive des effets. Il est également envisageable d'utiliser l'outil, une fois bien calé, pour analyser les données d'auscultation en temps réel et détecter de manière précoce une éventuelle anomalie. La politique d'utilisation des neurones à EDF reste encore à définir : dans un contexte industriel d'auscultation, la complexité des neurones rend leur utilisation systématique trop chronophage et coûteuse. A l'inverse, les neurones représentent un atout pour l'appui à l'analyse de certaines mesures complexes. Les résultats obtenus peuvent soit être utilisés et interprétés tels quels, soit servir à la mise au point de modèles à base physique simplifiés.

RÉFÉRENCES ET CITATIONS

- [1] Willm, G., Beaujoint, N. (1967). Les méthodes de surveillance des barrages au service de la production hydraulique d'Électricité de France. Problèmes anciens et solutions nouvelles. In : Proc. of the Ninth International Congress on Large Dams, Istanbul, 1967, pp. 529-550.
- [2] Penot, I., Daumas, B., Fabre, J.P. (2005). Monitoring behaviour. In : International Water Power & Dam Construction. Décembre 2005, pp 24-27.
- [3] Hu, W., Zhang, F., Song, L. (2010). Merging Model of Dam Deformation Analysis Based on Neural Network. In : Joint International Symposium on Deformation Monitoring. 2-4 November 2011. Hong Kong.
- [4] Mata, J., (2011). Interpretation of Concrete Dam Behavior with Artificial Neural Network and Multiple Linear Regression Models. In : Engineering Structures 33 (2011), pp. 903-910.
- [5] Lino, M., (2000). Les Réseaux de Neurones en Auscultation des Barrages – Artificial Neural Network in Dam Monitoring, XXe congrès de la CIGB, Q 78 minutes of the session, Sep. 2000, Beijing, Volume V.