

Identification d'un système non linéaire de vases communicants par réseaux de neurones à temps continu

Laurent SIGNAC, François BENOIT-MARAND

Laboratoire d'Automatique et d'Informatique Industrielle
Bâtiment de Mécanique 40, avenue du recteur Pineau, 86022 Poitiers, France

Laurent.Signac@univ-poitiers.fr, Francois.Benoit.Marand@etu.univ-poitiers.fr
http://laidi.univ-poitiers.fr

Résumé— Nous proposons dans cette communication d'illustrer l'utilisation des réseaux de neurones à temps continu pour l'identification des paramètres physiques d'un système non linéaire réel.

Si la faisabilité théorique de la plupart des éléments présentés ici ont déjà été discutés, une utilisation des réseaux de neurones à temps continu sur un système non linéaire réel dont nous pouvons par ailleurs estimer les caractéristiques faisait jusqu'à présent défaut.

Le dispositif expérimental qui servira de support à la présente communication est un système à vases communicants comportant deux cuves. Nous utiliserons dans un premier temps les réseaux de neurones à temps continu pour obtenir un modèle boîte noire et en tirerons quelques conclusions. Puis, nous montrerons que l'intérêt des réseaux de neurones à temps continu réside dans la possibilité de les utiliser pour retrouver les paramètres physiques qui apparaissent dans le modèle mathématique du dispositif étudié. Dans le cas présent, nous estimerons les caractéristiques des vannes du système.

Mots-clés— Réseaux de neurones, Identification, Systèmes non linéaires

I. INTRODUCTION

Nous avons récemment présenté les réseaux de neurones à temps continu [2], qui permettent d'identifier des systèmes non linéaires avec une grande précision [6], tout en permettant un retour aux paramètres physiques par une phase de réduction de modèle.

Jusqu'à alors, nous avons essentiellement testé notre méthode en simulation, afin d'en mesurer les défauts et d'en présenter les avantages. Nous proposons ici de confronter l'identification par réseaux de neurones à temps continu à un système réel non-linéaire basé sur les vases communicants.

Dans un premier temps, nous présenterons le dispositif expérimental, puis nous détaillerons la procédure d'identification du système par un réseau de neurones à temps continu. Ensuite, après une étape de réduction de modèle, nous déduirons certaines caractéristiques physiques du dispositif expérimental. Nous validerons les résultats obtenus en les confrontant avec ceux obtenus par d'autres moyens.

Enfin, nous concluons et présenterons quelques perspectives.

II. DISPOSITIF À VASES COMMUNICANTS

Le système que nous utilisons est représenté figure 1. Il est formé de deux cuves cylindriques qui communiquent par l'intermédiaire d'une vanne. La cuve 1 se vide dans un bac de vidange, et la cuve 2 est remplie par une pompe dont nous contrôlons le débit. Les débits respectifs de remplissage, de

vidange et de communication sont : d , Q_1 et Q_2 . Les hauteurs de liquide sont h_1 et h_2 . La section des cuves ayant pour surface S , nous pouvons écrire pour les deux cuves :

$$\begin{cases} S \times \frac{dh_1}{dt} = Q_2 - Q_1 \\ S \times \frac{dh_2}{dt} = d - Q_2 \end{cases} \quad (1)$$

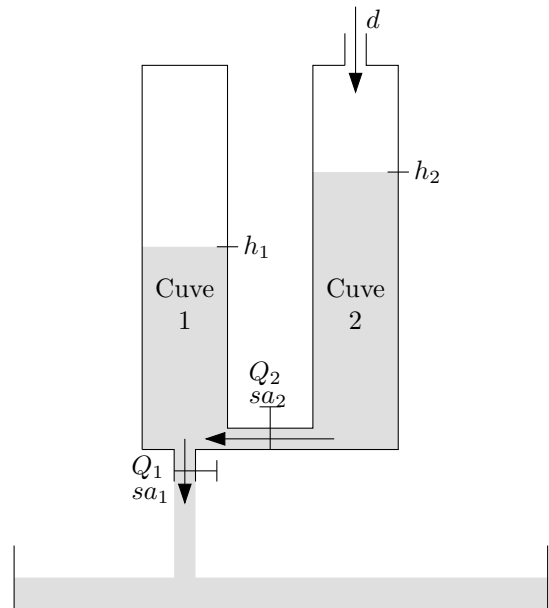


Fig. 1. Dispositif à vases communicants

Les débits Q_1 et Q_2 ne dépendent que des caractéristiques des vannes et des hauteurs de liquide. Dans le système d'équations (2), sa_1 et sa_2 sont des constantes (en faisant l'hypothèse d'un écoulement laminaire) caractéristiques des vannes, et g est l'accélération due à la gravité.

$$\begin{cases} Q_1 = sa_1 \sqrt{2gh_1} \\ Q_2 = sa_2 \operatorname{sgn}(h_2 - h_1) \sqrt{2g|h_2 - h_1|} \end{cases} \quad (2)$$

Notre dispositif expérimental est donc régi par le système d'équations (3).

$$\begin{cases} S \times \frac{dh_1}{dt} = sa_2 \operatorname{sgn}(h_2 - h_1) \sqrt{2g|h_2 - h_1|} - sa_1 \sqrt{2gh_1} \\ S \times \frac{dh_2}{dt} = d - Q_2 = sa_2 \operatorname{sgn}(h_2 - h_1) \sqrt{2g|h_2 - h_1|} \end{cases} \quad (3)$$

L'entrée de notre système sera le débit imposé d . Les deux sorties de notre système seront les hauteurs de liquide h_1 et h_2 . Les paramètres physiques à identifier seront les caractéristiques des vannes sa_1 et sa_2 .

Entrées et sorties sont mesurées à l'aide de capteurs que nous supposerons parfaits.

Sur la figure 2 est représenté le jeu de données mesuré.

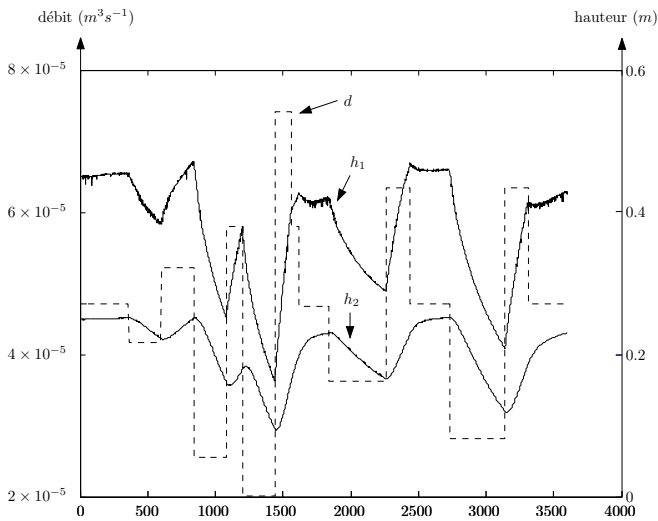


Fig. 2. Jeu de données mesuré

C'est ce jeu de données que nous utiliserons pour modéliser le dispositif par un modèle à base de réseau de neurones à temps continu.

III. IDENTIFICATION PAR RÉSEAUX DE NEURONES À TEMPS CONTINU

A. Principe des RNTC

Dans cette section, nous allons présenter le principe des réseaux de neurones à temps continu, tels qu'ils ont été définis dans [2]. Par souci de simplicité et de clarté, nous considérerons le cas d'un RNTC ayant une seule entrée et une seule sortie. Dans le cas contraire (et dans le cas de notre dispositif expérimental), il suffit d'étendre ce qui suit à des signaux vectoriels.

Un réseau de neurones à temps continu (RNTC) se présente sous la forme d'un réseau de type feed-forward à une couche cachée, suivi d'un ou plusieurs intégrateurs. Dans le cas de notre système expérimental, nous n'aurons qu'un seul intégrateur, comme indiqué figure 3, le système d'équations différentielles (3) qui constitue le modèle mathématique du dispositif étant d'ordre 1.

Le réseau de neurones en lui-même est un réseau feed-forward classique à une couche cachée, et correspond au schéma de la figure 4. Le choix d'un tel réseau se justifie par ses propriétés d'approximateur universel [1], [9].

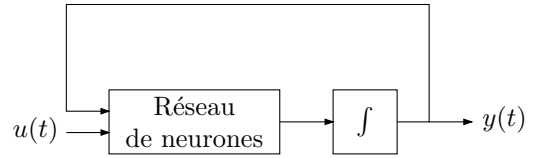


Fig. 3. Réseau de neurones à temps continu

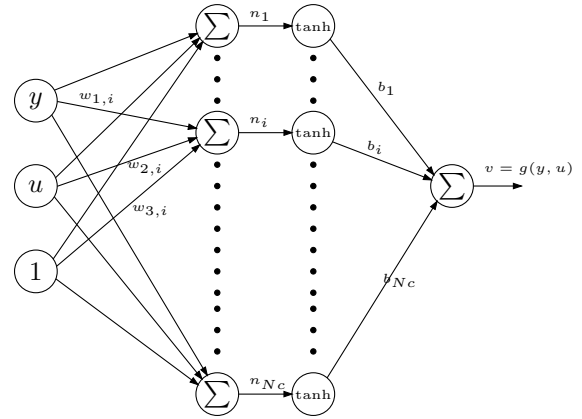


Fig. 4. Réseau de neurones feed-forward

Un modèle par RNTC, avec un seul intégrateur, une entrée, une sortie et N_c neurones cachés correspond donc au système d'équations suivant :

$$\frac{dy}{dt} = g(y, u)$$

$$\text{et } g(y, u) = v = \sum_{i=1}^{N_c} b_i \tanh(w_{1,i}y + w_{2,i}u + w_{3,i}) \quad (4)$$

B. Identification

En raison de son insensibilité au bruit [10], nous utiliserons une méthode d'identification par erreur de sortie, comme indiqué sur la figure 5.

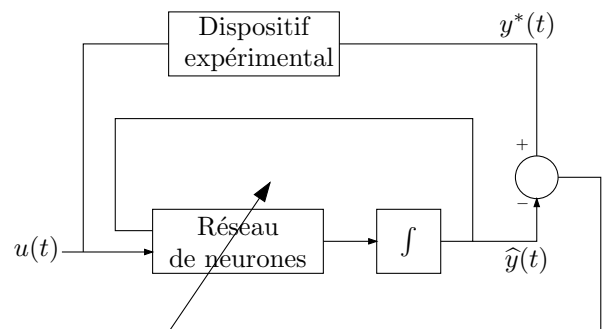


Fig. 5. Identification par erreur de sortie

Nous minimiserons le critère classique suivant :

$$J = \sum_{j=0}^{K-1} (y^*(j\Delta t) - \hat{y}(j\Delta t))^2 = \sum_{j=0}^{K-1} e_j^2$$

où K est le nombre d'échantillons.

Nous utilisons l'algorithme de Levenberg Marquardt [11], [8] pour minimiser le critère J . En appelant $\theta = (w_{11}, \dots, b_1, \dots, b_{N_c})$ les paramètres de la fonction g , nous pouvons donc itérer le vecteur paramètre suivant :

$$\theta_{i+1} = \theta_i + \Delta\theta \text{ avec } \Delta\theta = (\tilde{H}_\theta + \lambda I)^{-1} J'_\theta$$

où J'_θ est le gradient de J par rapport à θ , \tilde{H}_θ est la matrice du pseudo-hessien de J par rapport à θ et λ est un paramètre de contrôle de l'algorithme. J'_θ et \tilde{H}_θ sont calculés en utilisant les relations suivantes :

$$J'_\theta = -2 \sum_{k=0}^{K-1} \left(e_k \cdot \frac{\partial \hat{y}}{\partial \theta}(k\Delta t) \right)$$

$$\left[\tilde{H}_\theta \right]_{j,l} = 2 \sum_{k=1}^K \left(\frac{\partial \hat{y}}{\partial \theta_j}(k\Delta t) \cdot \frac{\partial \hat{y}}{\partial \theta_l}(k\Delta t) \right)$$

Soit $\sigma_p(k\Delta t) = \frac{\partial \hat{y}}{\partial \theta_p}(k\Delta t)$ la sensibilité de \hat{y} par rapport au paramètre θ_p . Nous devons évaluer $\sigma_p(k\Delta t)$ pour tout p et tout k .

Le calcul des fonctions de sensibilité est réalisé par rétropropagation dynamique [12] :

$$\begin{aligned} \frac{d\sigma_p}{dt} &= \frac{d}{dt} \left(\frac{\partial \hat{y}}{\partial \theta_p} \right) = \frac{\partial}{\partial \theta_p} \left(\frac{d\hat{y}}{dt} \right) \\ &= \frac{\partial}{\partial \theta_p} (g(\hat{y}(\theta, t), u(t), \theta)) \\ &= \frac{\partial \hat{v}}{\partial \hat{y}} \times \frac{\partial \hat{y}}{\partial \theta_p} + \frac{\partial \hat{v}}{\partial \theta_p} \\ &= \frac{\partial \hat{v}}{\partial \hat{y}} \sigma_p + \frac{\partial \hat{v}}{\partial \theta_p} \end{aligned}$$

où v est la sortie directe du réseau feed-forward (fig. 4).

Le calcul des fonctions de sensibilité se fait donc par simulation, et les termes classiques : $\frac{dv}{d\theta_p}$ et $\frac{dv}{d\hat{y}}$ sont calculés par rétropropagation du gradient [7], [13]. Dans les équations suivantes, r_i représente l'entrée i du réseau¹ :

$$\begin{aligned} \frac{d\hat{v}}{db_i}(k\Delta t) &= \tanh(n_i(k\Delta t)) \\ \frac{d\hat{v}}{dw_{i,j}}(k\Delta t) &= r_i(k\Delta t) \cdot \tanh'(n_j(k\Delta t)) \cdot b_j \\ \frac{d\hat{v}}{d\hat{y}}(k\Delta t) &= \sum_{i=1}^{N_c} w_{1,i} \cdot \tanh'(n_i(k\Delta t)) \cdot b_i \end{aligned}$$

Nous sommes donc maintenant en mesure d'utiliser l'algorithme de Levenberg-Marquardt pour approcher notre jeu de données (figure 2) par un modèle à base de RNTC.

C. Résultats expérimentaux

Les données représentées figure 2 ont été utilisées pour optimiser le modèle RNTC selon la méthode décrite dans

¹Par exemple, $r_1 = \hat{y}$ et $r_2 = u$.

la section précédente. Sur la figure 6 sont représentés à nouveau les points de mesure en traits pleins, ainsi que la série temporelle produite par le modèle en pointillés. Malgré quelques différences, en partie imputables aux capteurs de mesure, le modèle suit correctement les données mesurées.

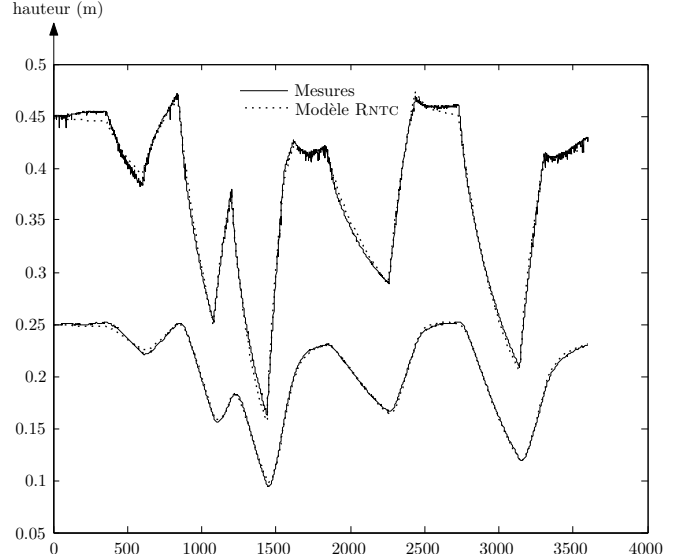


Fig. 6. Résultats de l'identification par RNTC

Ce premier résultat «boîte noire» nous permet déjà de conclure que :

- le processus physique observé est effectivement modélisable par des équations différentielles d'ordre 1 ;
- le nombre de neurones dans la couche cachée, choisi arbitrairement (10 dans notre cas), est suffisant pour modéliser le système ;
- l'approximation n'étant pas parfaite (les données simulées ne suivant pas exactement les données mesurées dans certaines zones), le nombre de neurones dans la couche cachée n'est probablement pas trop élevé.

IV. RÉDUCTION DE MODÈLE

L'utilisation des RNTC présente plusieurs avantages. En premier lieu, ils constituent un modèle boîte noire permettant d'approcher des systèmes non-linéaires avec efficacité, comme nous l'avons vu dans la section précédente. Mais l'intérêt principal réside dans la possibilité de déconnecter le réseau feed-forward de la partie intégration dans le RNTC, afin de disposer d'une approximation des fonctions qui apparaissent dans les équations différentielles qui régissent le processus physique.

A. Identification des paramètres physiques

Ainsi, si nous disposons d'un ou plusieurs modèles de systèmes d'équations différentielles, nous pouvons les confronter *très rapidement* à la partie statique du RNTC, afin par exemple d'estimer certains paramètres des modèles mathématiques. Nous proposons ici d'illustrer ce principe en utilisant le modèle présenté dans la section II pour déduire les caractéristiques sa_1 et sa_2 des vannes. Nous pourrions valider ce résultat en estimant ces mêmes paramètres d'une autre manière.

La réduction de modèle est une étape très simple qui consiste à sélectionner une trajectoire temporelle (simulée avec le RNTC, donc sans bruit de mesure) sur laquelle nous supposons que le modèle RNTC est correct (nous prendrons ici les points obtenus avec le RNTC sur la figure 6).

Puis nous réalisons une identification statique très rapide (figure 7) entre une fonction candidate f , issue du modèle physique (3) et la fonction g calculée par le réseau feed-forward.

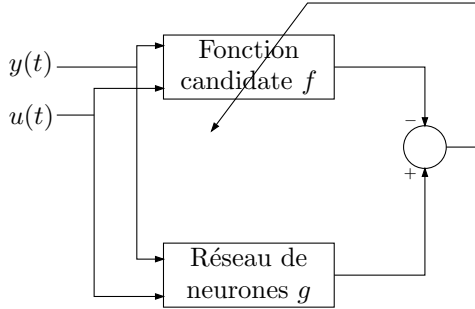


Fig. 7. Identification statique entre le réseau et la fonction candidate

L'identification statique est faite par des méthodes standards (moindres carrés non linéaires) permettant de minimiser :

$$\sum_{i=0}^{i=K-1} (g(i\Delta t) - f(i\Delta t))^2$$

Après l'étape de réduction de modèle, nous obtenons les estimations suivantes des caractéristiques des vannes :

$$\begin{cases} sa_1 = 2.188 \times 10^{-5} m^2 \\ sa_2 = 2.327 \times 10^{-5} m^2 \end{cases} \quad (5)$$

B. Validation des résultats

Les estimations obtenues précédemment peuvent être validées de différentes façons. La méthode de validation est naturellement spécifique au dispositif étudié, alors que l'identification par RNTC est générale. Pour valider nos résultats, nous pouvons :

1. confronter le modèle mathématique (3), complété par les paramètres estimés (5), à des mesures réalisées sur le système physique ;
2. identifier directement les paramètres sa_1 et sa_2 sur des mesures issues d'expérimentations éventuellement plus simples (à débit d'entrée nul par exemple) ;
3. calculer sa_1 en chronométrant simplement des temps de vidange ou en mesurant des hauteurs d'équilibre.

La méthode la plus immédiate est la troisième. L'équation qui régit la vidange de la cuve 1 seule est simplement :

$$S \times \frac{dh_1}{dt} = -sa_1 \sqrt{2gh_1}$$

Si à $t = 0$, $h_1 = h_0$ nous pouvons résoudre cette équation :

$$h_1 = \left(\sqrt{h_0} - \frac{1}{2} B t \right)^2 \quad \text{avec } B = sa_1 \sqrt{2gS}^{-1}$$

En mesurant le temps de passage de la cuve de $h_1 = 0.3 m$ à $h_1 = 0.2 m$ par exemple (ce temps vaut 32.74 s dans notre cas), nous pouvons calculer :

$$sa_1 = 2.135 \times 10^{-5} m^2$$

Si nous utilisons la seconde méthode, nous pouvons estimer à la fois sa_1 et sa_2 . L'expérience a été réalisée avec des cuves initialement pleines (figure 8), et un débit d'entrée nul. Nous avons ainsi obtenu :

$$\begin{cases} sa_1 = 2.119 \times 10^{-5} m^2 \\ sa_2 = 2.570 \times 10^{-5} m^2 \end{cases} \quad (6)$$

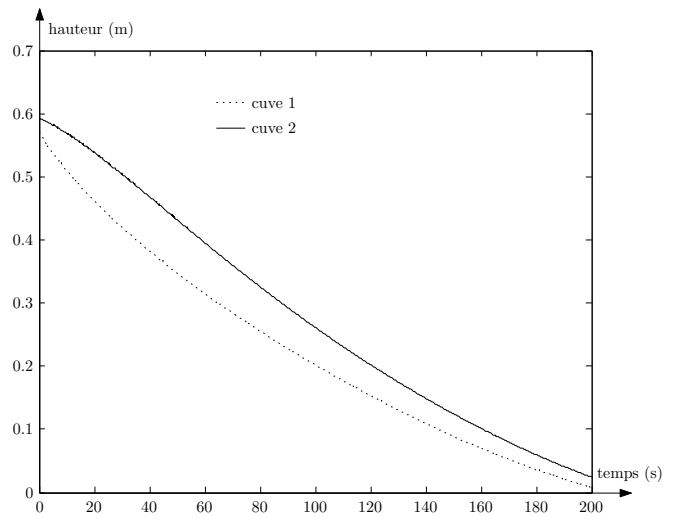


Fig. 8. Mesure des temps de vidange des deux cuves pour estimer les paramètres des vannes

Chaque expérimentation donnera des valeurs sensiblement différentes. Ces erreurs sont dues d'une part aux mesures approximatives (erreurs de calibration des capteurs par exemple), et naturellement au fait que le modèle mathématique n'est pas parfait, les «caractéristiques» des vannes dépendant en réalité du débit de l'écoulement.

Néanmoins, les valeurs estimées par réduction de modèle sur un RNTC sont relativement proches des valeurs obtenues autrement puisque l'erreur relative sur les paramètres ne dépasse pas 10%.

Enfin, nous pouvons confronter notre fonction candidate et les estimations de sa_1 et sa_2 avec les mesures expérimentales, comme indiqué sur la figure 9.

Les résultats ne sont pas parfaits (comparer les figures 9 et 6) en ce sens que l'étape de réduction de modèle a dégradé les résultats que nous obtenions avec le modèle boîte noire, mais il est important de noter qu'*a priori*, aucun autre couple de paramètres ne donnera de résultats franchement meilleurs, l'erreur commise étant due à une approximation dans le modèle mathématique (sa_1 et sa_2 ne sont en fait pas des constantes). Pour illustrer ce dernier point, nous avons représenté les mesures, et la simulation obtenue en utilisant le modèle mathématique (3) et les paramètres (6) sur la figure 10. Cette figure est à comparer avec la figure 9.

La réduction de modèle donne donc des résultats corrects et permet de plus de pointer du doigt l'insuffisance du modèle mathématique (3) choisi (bien que ce ne soit pas le but ici).

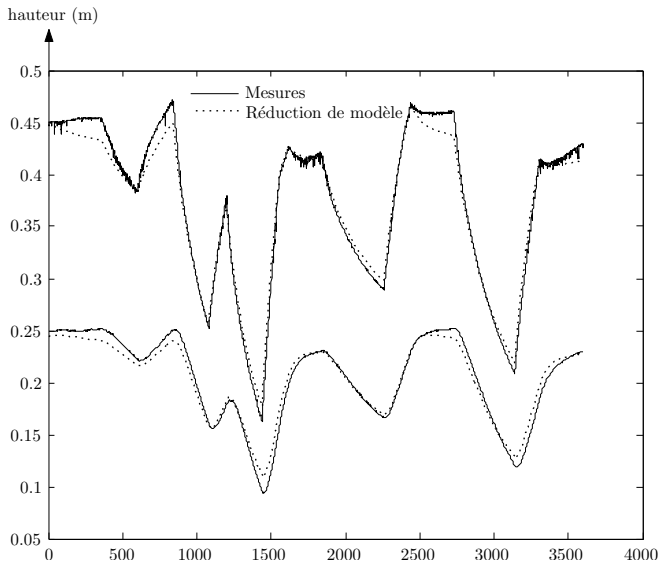


Fig. 9. Comparaison des mesures avec le modèle (3) utilisant les paramètres estimés (5)

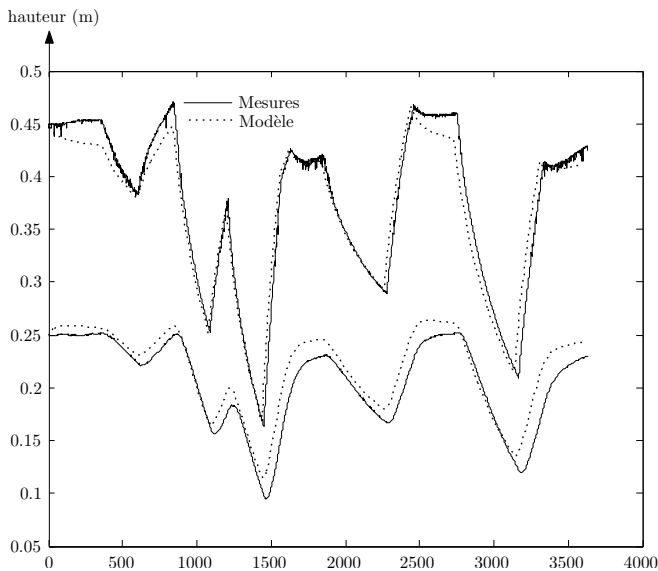


Fig. 10. Comparaison des mesures avec le modèle (3) utilisant les paramètres (6)

V. CONCLUSION ET PERSPECTIVES

Dans cette communication, après avoir rappelé le principe des réseaux de neurones à temps continu (RNTC), nous avons montré qu'ils pouvaient être utilisés sur des données issues d'expérimentations réelles, et non de simulations. Sur ces données (système non linéaire de vases communicants), nous avons utilisé un RNTC comme modèle boîte noire, puis nous avons tiré parti de la possibilité d'isoler la partie statique des RNTC afin d'estimer les paramètres physiques du modèle mathématique à partir du modèle RNTC.

Dans un avenir proche, nous pourrions confronter expérimentalement notre méthode d'identification et de réduction de modèle pour des systèmes d'ordre 2. Ceci permettra entre autres de montrer de quelle façon l'étape boîte noire peut être utilisée pour confirmer que le modèle a un ordre correct.

Enfin, nous avons récemment présenté des résultats théoriques concernant les systèmes d'ordre fractionnaire [4], [5], [3]. Toujours dans un souci de confronter notre méthodologie à l'expérience, nous espérons pouvoir illustrer l'utilisation des RNTC d'ordre non-entiers à des données réelles.

RÉFÉRENCES

- [1] A. Barron. Universal approximation bounds for superpositions of a sigmoidal function. *IEEE Trans. on Information Theory*, 39 :930–945, 1993.
- [2] François Benoit-Marand, Laurent Signac, Smail Bachir, and Jean-Claude Trigeassou. Non linear identification of continuous time systems using neural networks. In *International Workshop on Electronics and System Analysis*, Bilbao, Spain, 2004.
- [3] François Benoit-Marand, Laurent Signac, Amel Benchellal, Thierry Pointot, and Jean-Claude Trigeassou. Identification of diffusive interfaces using a simplified fractional integrator. part ii : non linear case. In *Conference of the IEEE Industrial Electronics Society*, 2006.
- [4] François Benoit-Marand, Laurent Signac, Thierry Pointot, and Jean-Claude Trigeassou. Identification de systèmes fractionnaires non linéaires par réseaux neuronaux à temps continu. In *Conférence Internationale Francophone d'Automatique*, 2006.
- [5] François Benoit-Marand, Laurent Signac, Thierry Pointot, and Jean-Claude Trigeassou. Identification of non linear fractional systems using continuous time neural networks. In *IFAC Workshop on Fractional Differentiation and its Applications*, 2006.
- [6] François Benoit-Marand, Laurent Signac, and Jean-Claude Trigeassou. Neural identification of physical processes : Analysis of disruptive factors. In *Engineering Applications of Neural Networks*, 2005.
- [7] Y. Le Cun. Une procédure d'apprentissage pour réseau à seuil asymétrique. *Cognitiva*, pages 599–604, 1985.
- [8] Martin T. Hagan and Mohammad B. Menhaj. Training feedforward networks with the marquardt algorithm. *IEEE Transactions on neural networks*, 5(6), november 1994.
- [9] Kurt Hornik. Some new results on neural network approximation. *Neural Networks*, 6 :1069–1072, 1993.
- [10] Lennart Ljung. *System Identification - Theory For the User*. PTR Prentice Hall, Upper Saddle River, N.J., 1999.
- [11] D. W. Marquardt. An algorithm for least-squares estimation of non-linear parameters. *Soc. Indust. Appl. Math*, 11(2), 1963.
- [12] K.S. Narendra and K. Parthasarathy. Gradient methods for the optimization of dynamical systems containing neural networks. *IEEE Transactions on Neural Networks*, 2 :252–262, March 1991.
- [13] D. E. Rumelhart, G.E. Hinton, and R.J. Williams. Learning representations by back-propagating errors. *Nature*, 323 :533–536, 1986.