

Prédiction en ligne par apprentissage et exploitation des caractéristiques géométriques des signaux

On-line prediction with learning using geometrical properties of a signal

par Stéphane LEPERE *, Xianyi ZENG †, Christian VASSEUR *

* Laboratoire I3D, Université des Sciences et Technologies de Lille, Bâtiment P2, F-59655 Villeneuve d'Ascq Cedex, France.

<http://www-i3d.univ-lille1.fr>

† GEMTEX/ENSAIT, 9 rue de l'Ermitage, BP 30329, F-59056 Roubaix Cedex 01, France. <http://www.ensait.fr>

résumé et mots clés

En commande des systèmes, l'anticipation pure apparaît comme la meilleure solution mathématique au problème de l'asservissement des sorties aux consignes, que celles-ci soient connues a priori (cas des systèmes de régulation) ou pas (cas des systèmes de poursuite de trajectoire). D'un point de vue formel, cela est obtenu en mettant le système à asservir en cascade avec son système inverse. Cette solution n'est toutefois pas réalisable physiquement, car le modèle du système n'est que rarement complet. De plus, le système inverse est souvent instable. Enfin, les perturbations observées, soit en sortie, soit en entrée, rendent difficile l'accès à une solution satisfaisante. Dernier aspect : pour être exploitable, l'anticipation doit se faire en temps réel.

La méthode présentée ici tente d'apporter une solution rapide et robuste au problème de prédiction. Elle utilise à la fois les propriétés géométriques du signal à prédire à l'instant considéré (procédure locale), ainsi qu'une base d'apprentissage constituée d'observations passées (procédure globale). Les performances du prédicteur sont évaluées par quelques exemples significatifs et comparées aux performances d'autres prédicteurs.

Prédiction robuste, pente, courbure, cercle osculateur, séries temporelles, temps réel.

abstract and key words

In systems theory, pure anticipation seems the best mathematical solution for the problem of control, whatever the inputs are known a priori (case of the systems of regulation), or not (case of tracking systems). From a formal point of view, this is obtained by putting the system to be controlled in cascade with its opposite system. However, this solution is not realizable physically, because the model of the system is rarely complete. Moreover, the opposite system is often unstable. Lastly, the disturbances observed, either at the output, or at the input, make difficult the access to a satisfactory solution. Last aspect: to be exploitable, anticipation has to be made in real-time.

The method described here tries to bring a fast and robust solution to the problem of prediction. This method uses at the same time the geometrical properties of the signal to be predicted at the considered moment (local procedure), and a learning base constituted by past observations (global procedure). The performances of this predictor are next evaluated on several significative examples and are compared with those of others predictors.

Robust prediction, slope, curvature, osculating circle, time-series, real-time.

1. introduction

La prédiction de séries temporelles est un problème très important, dont la solution trouve de nombreuses applications dans des domaines aussi variés que la prise de décision financière ou économique, la météorologie, la prévision de ventes [BOX 76]. Dans la mise en œuvre des méthodes, on considère que le signal est numérisé et qu'il est représenté par une série temporelle. Ainsi, le problème de la prédiction en temps réel revient à estimer, à chaque instant t , les valeurs futures de la série temporelle, à partir de ses valeurs passées. Le but ultime peut alors être de déterminer un modèle spatio-temporel de comportement de la série.

Il existe de nombreuses méthodes permettant l'apprentissage et la prédiction de séries temporelles. Les méthodes classiques utilisent un modèle d'approximation paramétré : MA (Moving Average), AR (Auto-Regressive), ARMA (combinaison des deux précédents). Ces modèles ont une structure simple et linéaire. Toutefois, ils ne suffisent pas pour réaliser une prédiction correcte de séries temporelles à comportement complexe et non-linéaire [ALI 96].

Des approches plus récentes, basées sur les techniques de « soft computing » (réseaux neuronaux, contrôleurs à logique floue), ont été développées avec succès [LEE 90], [ALI 96], [ISH 96]. Ces outils apparaissent, en effet, comme des approximateurs universels [WAN 92], [HOR 89]. De plus, ils peuvent opérer dans un environnement variable et ont donc d'excellentes capacités d'adaptation. Une difficulté subsiste toutefois : les temps de calculs nécessaires à la recherche de la solution optimale ne permettent pas toujours d'opérer en temps réel.

La méthode de prédiction temps réel proposée dans cet article se décompose en une procédure locale et une procédure globale.

Tout d'abord, la procédure locale repose sur l'analyse des propriétés géométriques du signal. Essentiellement il s'agit de construire, à l'instant considéré, le cercle osculateur à la courbe représentative du signal. Trois informations intrinsèques importantes sont associées à ce cercle :

- le point observé lui-même, qui est le point de contact entre la courbe et le cercle osculateur,
- la courbure, caractérisée par le centre et le rayon du cercle osculateur,

- la tangente, qui est perpendiculaire au rayon vecteur du cercle osculateur, au point observé.

Ensuite, la prédiction est alors simplement effectuée sur la base de ces trois informations complétées par une procédure globale qui utilise les observations passées du signal (base d'apprentissage évolutive).

Cette méthode étant basée sur l'utilisation du rayon de courbure, on fait l'hypothèse que l'évolution du signal est de classe C^2 . La section 2 donne une présentation formelle de la méthode avec ses aspects locaux et globaux, dans laquelle on ne tient pas compte des problèmes de bruits. La section 3 décrit une mise en œuvre robuste, dans laquelle les bruits sont pris en compte. Enfin, la section 4 donne quelques exemples permettant d'évaluer les performances du prédicteur et de les comparer à celles d'autres dispositifs.

2. présentation de la méthode

2.1. notations et principe

On considère (cf. [figure 1]) :

- Un signal $y(t) \in \mathbb{R}$,
- A le point de coordonnées $(t, y(t))$,
- \mathcal{C} le cercle osculateur à la courbe $y(t)$ en A , de centre $O = (t_O, y_O)$ et de rayon $R(t)$.

On désigne alors par $\hat{y}(t+h)$ la prédiction de $y(t)$, faite à l'instant t , h représentant l'horizon de prédiction.

Dans de telles conditions, le problème de prédiction, qui a pour objectif de trouver le point P de coordonnées $(t+h, \hat{y}(t+h))$, peut se ramener, à l'instant t , à la recherche d'un coefficient $\lambda(t) \in \mathbb{R}$ tel que le point P puisse s'écrire $P = \lambda(t)M + (1 - \lambda(t))N$, où le point N appartient à la tangente à la courbe en A , et le point M est l'intersection de \mathcal{C} et de ON .

D'après cette définition, le coefficient $\lambda(t)$ peut être interprété comme étant la pondération optimale entre une prédiction sur le cercle osculateur (point M) et une prédiction le long de la tangente en A (point N).

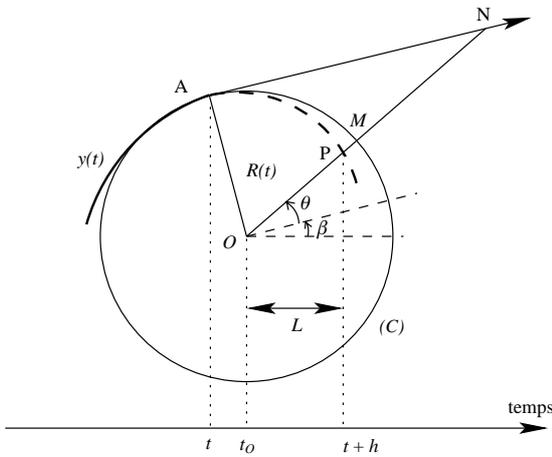


Figure 1. – $y(t)$ et le cercle osculateur

2.2. considérations géométriques (aspect local)

On suppose dans cette section que la valeur $y(t+h)$ est connue, et on calcule $\lambda(t)$ correspondant à cette valeur. Au point A, le rayon de courbure algébrique (dont le signe dépend du sens de concavité) est donné par la formule suivante :

$$R(t) = \frac{(1 + y'(t)^2)^{3/2}}{y''(t)}.$$

On définit ensuite (voir figure 1), $L = t+h - t_0$, $\beta = \arctan y'(t)$ angle entre l'axe des temps et la tangente, et θ donné par l'expression :

$$\theta = \arctan\left(\frac{y(t+h) - y_0}{t+h - t_0}\right) - \beta, \quad (1)$$

qui représente l'angle entre la tangente et le vecteur \overrightarrow{OP} dans le cas où $L > 0$, et l'angle entre la tangente et \overrightarrow{PO} lorsque $L < 0$. On montre qu'avec ces définitions il est possible de déterminer le coefficient de pondération $\lambda(t)$ selon l'expression suivante :

$$\lambda(t) = \frac{\overline{OP} - \overline{ON}}{\overline{OM} - \overline{ON}} = \frac{L \sin(\theta) + R \cos(\theta + \beta)}{R(1 - \sin|\theta|)\cos(\theta + \beta)} \quad (2)$$

qui est valable quels que soient les signes de $y''(t)$ et de L .

2.3. mise en œuvre sur les séries temporelles (aspect global)

Dans cette section, nous nous intéressons à la recherche de $\lambda(t)$ lorsque la valeur de $y(t+h)$ n'est pas connue.

Pour cela, le signal $y(t)$ étant échantillonné à la période T_e , on note $y(iT_e) = y_i$ pour tout indice i . La série temporelle $\{y_i\}$ est alors traitée comme suit [figure 2] :

On considère se trouver à l'instant $t = qT_e$. L'horizon de prédiction est défini par $h = nT_e$. Dans ces conditions, si on note $\lambda_i = \lambda(iT_e)$, la connaissance du passé du signal permet de déterminer la suite $\{\lambda_i\}$, jusqu'à l'indice $i = q - n + 1$, à l'aide de l'équation (2).

On considère ensuite une fenêtre d'observation glissante f_i de longueur fixe N , notée $f_i = \{y_{i-N+1}, \dots, y_i\}$, et on définit la relation $\Lambda : f_i \mapsto \lambda_i$.

L'ensemble des couples $(f_i; \lambda_i)$ constitue le graphe de Λ . Le sous-graphe correspondant aux valeurs de $i \in J_s(q)$ avec :

$$J_s^{(q)} = \{q - \max(n, N) - s + 1, \dots, q - \max(n, N)\},$$

constitue une base d'apprentissage glissante de taille s , caractérisant le passé du signal lorsque l'on se trouve en qT_e (cf. [figure 2]).

La recherche du coefficient λ_q qui fournit la prédiction \hat{y}_{q+n} est basée sur une extrapolation de Λ qui utilise les fenêtres passées f_i et la fenêtre actuelle f_q . Cette extrapolation nécessite un calcul de distance entre chaque f_i et f_q . Afin de mettre en évidence des équations de récurrence permettant d'accélérer les calculs et donc de faciliter une mise en œuvre en temps réel, nous avons adopté comme distance la distance quadratique notée $d_i^{(q)}$:

$$d_i^{(q)} = d(f_i, f_q) = \sum_{k=0}^{N-1} (y_{i-k} - y_{q-k})^2, \text{ où } i \in J_s^{(q)}. \quad (3)$$

De cette manière, on obtient :

$$d_{i+1}^{(q+1)} = \sum_{k=0}^{N-1} (y_{i+1-k} - y_{q+1-k})^2 = d_i^{(q)} + (y_{i+1} - y_{q+1})^2 - (y_{i-N+2} - y_{q-N+2})^2, \quad \forall i \in J_s^{(q)},$$

Le passage de qT_e à $(q+1)T_e$ ne nécessite donc que 4 additions/soustractions et 2 produits par distance calculée.

Dans ces conditions, la prédiction est opérée selon la procédure suivante :

Procédure d'estimation de y_{q+n}

1. – pour $i \in J_s^{(q)}$
 - calculer les distances $d_i^{(q)} = d(f_i, f_q)$
 - définir les poids $\mu_i = \left(\frac{1}{d_i^{(q)}}\right)^p$ où $p \in \mathbb{N}^*$.
2. – estimer λ_q par :

$$\hat{\lambda}_q = \sum_{i \in J_s^{(q)}} \mu_i \lambda_i / \sum_{i \in J_s^{(q)}} \mu_i \quad (4)$$

3. – à partir du $\hat{\lambda}_q$, estimer θ_q à l'aide de l'équation (2)
4. – estimer y_{q+n} , à partir de l'équation (1).

Commentaires :

– $\hat{\lambda}_q$ est défini comme étant la somme pondérée des λ_i passés, un λ_i donné ayant d'autant plus de poids que la fenêtre f_i qui lui est associée ressemble à (est proche de) la fenêtre actuelle f_q .

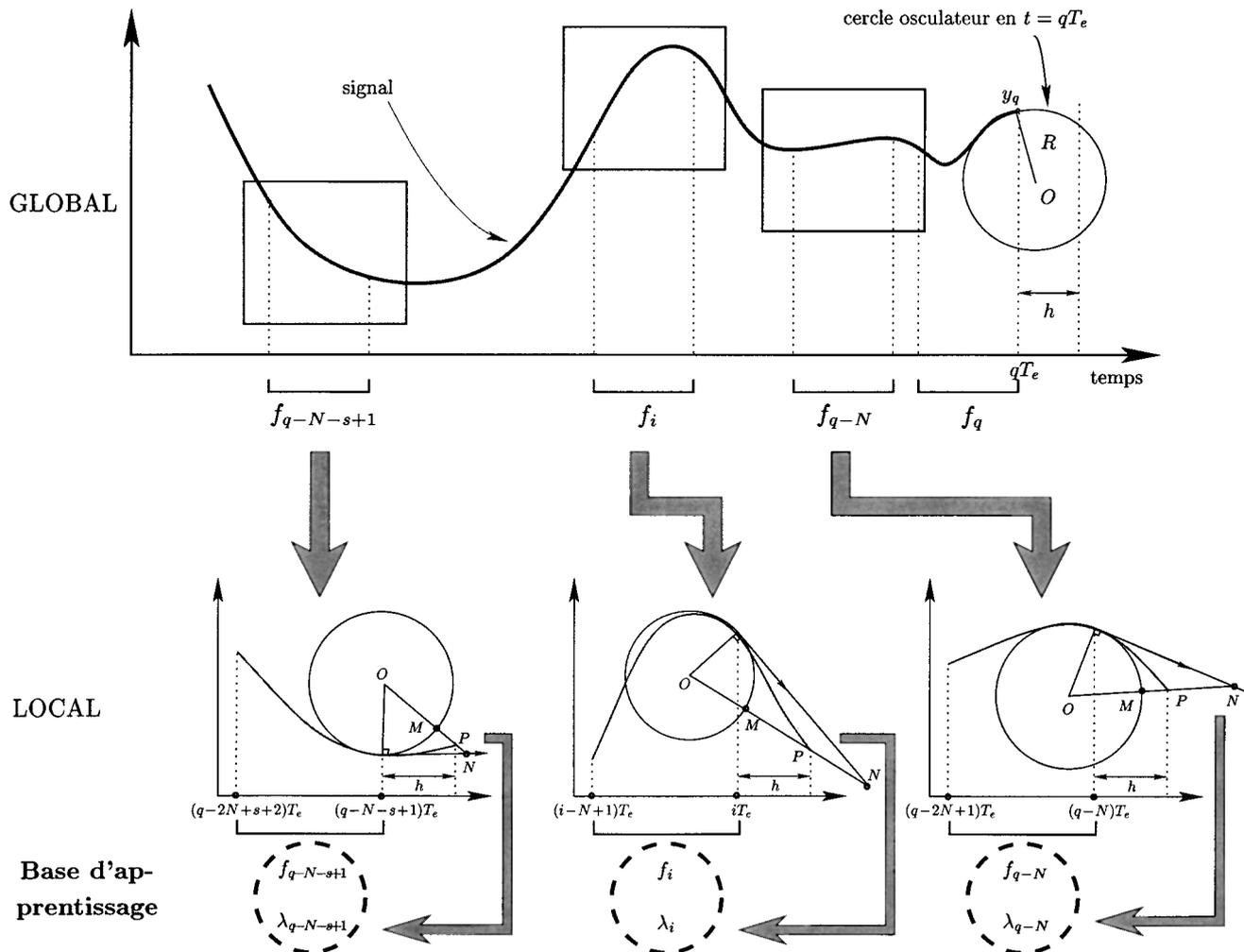


Figure 2. – Principe sur une série temporelle.

– Pour le cas particulier où il existe un $i \in J_s^{(q)}$ pour lequel $d_i^{(q)} = 0$, c'est-à-dire pour lequel la fenêtre f_i est identique à la fenêtre actuelle f_q , on prend comme prédiction l'image de la fenêtre f_i : $y_{q+n} = y_{i+n}$.

– Le paramètre p permet de renforcer le poids donné aux fenêtres proches par rapport aux fenêtres éloignées (contraste). En pratique, il apparaît que ce paramètre influence peu la précision du prédicteur.

3. prédiction robuste

La méthode présentée ci-dessus ne prend pas en compte les bruits du signal qui proviennent, en pratique, de perturbations ou

de dysfonctionnements des capteurs. En particulier, les bruits ne permettent pas de calculer correctement $y'(t)$, $y''(t)$, et donc le rayon $R(t)$ sur lequel s'appuie le modèle utilisé ici.

Afin de pallier cette difficulté, il convient d'opérer sur un signal filtré. La procédure utilisée est classique. Elle est décrite ci-dessous (cf. [figure 3]).

A l'instant $t = qT_e$, on recherche la courbe des moindres carrés définie sur la fenêtre $\{y_{q-M+1}, \dots, y_q\}$, à partir d'un modèle polynomial. Cela revient à minimiser l'erreur :

$$E_q^2 = \sum_{i=0}^{M-1} w_i \left(y_{q-i} - \left(a_0 + a_1 \frac{(-iT_e)}{1!} + a_2 \frac{(-iT_e)^2}{2!} + \dots + a_d \frac{(-iT_e)^d}{d!} \right) \right)^2,$$

avec $w_i = (1 - \gamma)^i$, où $\gamma \in [0, 1[$ est un facteur d'oubli fixé.

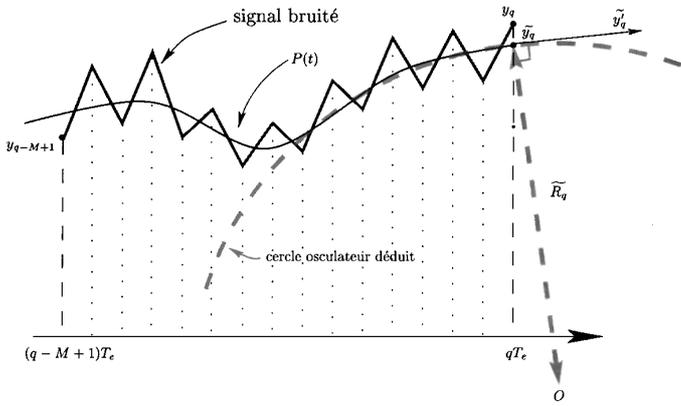


Figure 3. – Principe du filtrage

En posant ensuite $\frac{\partial E_q^2}{\partial a_k} = 0$ pour $k = 0, \dots, d$, on obtient le système :

$$\Phi (a_0, a_1, \dots, a_d)^t = W_q,$$

avec

$$\Phi = \begin{pmatrix} \sum w_i & \sum i w_i & \dots & \sum i^d w_i \\ \sum i w_i & \sum i^2 w_i & & \vdots \\ \vdots & & \ddots & \vdots \\ \sum i^d w_i & \dots & \dots & \sum i^{2d} w_i \end{pmatrix}$$

$$\begin{pmatrix} 1 & 0 & \dots & 0 \\ 0 & -T_e & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \dots & 0 & \frac{(-T_e)^d}{d!} \end{pmatrix},$$

et

$$W_q = \begin{pmatrix} \sum w_i y_{q-i} \\ \sum i w_i y_{q-i} \\ \vdots \\ \sum i^d w_i y_{q-i} \end{pmatrix}$$

La matrice Φ est inversible si est seulement si $M > d$. Dans ce cas, on peut déduire les évaluations des données filtrées $\tilde{y}(t)$, $\frac{d\tilde{y}}{dt}(t)$, et $\frac{d^2\tilde{y}}{dt^2}(t)$ par :

$$\begin{pmatrix} \tilde{y}_q \\ \tilde{y}'_q \\ \tilde{y}''_q \end{pmatrix} = \begin{pmatrix} a_0 \\ a_1 \\ a_2 \end{pmatrix} = \widetilde{\Phi}^{-1} W_q, \quad (5)$$

où $\widetilde{\Phi}^{-1}$ représente la matrice constituée par les 3 premières lignes de Φ^{-1} .

La valeur de W_q se déduit quant à elle facilement de W_{q-1} par la récurrence suivante :

$$\begin{aligned} \sum_{i=0}^{M-1} i^p w_i y_{q-i} &= \sum_{i=-1}^{i=M-2} (i+1)^p w_{i+1} y_{q-1-i} \\ &= (1-\gamma) \sum_{i=0}^{M-1} (i+1)^p w_i y_{q-1-i} - (1-\gamma)^M \\ &\quad M^p y_{q-M} \quad (+y_q \text{ si } p = 0), \end{aligned}$$

ce qui conduit à :

$$W_q = (1-\gamma) \underbrace{\begin{pmatrix} 1 & 0 & \dots & \dots & 0 \\ 1 & 1 & 0 & & \vdots \\ 1 & 2 & 1 & \ddots & \vdots \\ \vdots & & & \ddots & 0 \\ 1 & & & & 1 \end{pmatrix}}_T W_{q-1} + \begin{pmatrix} 1 \\ 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix} y_q - (1-\gamma)^M \begin{pmatrix} 1 \\ M \\ M^2 \\ \vdots \\ M^d \end{pmatrix} y_{q-M},$$

où T est une matrice triangulaire inférieure ayant pour terme général $t_{ij} = C_{i-1}^{j-1} = \frac{(i-1)!}{(i-j)!(j-1)!}$, avec $i, j = 1, \dots, d+1$.

En combinant la relation de récurrence (6) avec l'égalité (5), il est possible d'estimer les valeurs filtrées de y , y' et y'' en $t = qT_e$ en temps réel, et ceci, indépendamment de la taille M de la fenêtre de filtrage. Le facteur d'oubli γ sert quant à lui à régler le poids affecté aux données.

Remarque 1 : Plutôt que d'opérer le filtrage suivant un modèle polynomial, il aurait été plus logique d'utiliser un modèle circulaire qui nous aurait fourni directement le cercle osculateur filtré. Malheureusement, un tel modèle ne fournit pas de relations de récurrence, et aurait donc été plus difficile à appliquer en temps réel.

Remarque 2 : Cette technique très peu gourmande en calculs fournit des évaluations de y' et de y'' au bout de la fenêtre de filtrage, c'est-à-dire exactement en t . Par contre les expressions habituellement utilisées $\frac{y(t)-y(t-T_e)}{T_e}$ et $\frac{y(t)-2y(t-T_e)+y(t-2T_e)}{T_e^2}$, donnent les dérivées première et seconde respectivement en $t - \frac{T_e}{2}$ et en $t - T_e$.

De ce fait, même en l'absence de bruit, nous utilisons cette méthode, et la prédiction n'en sera que meilleure.

Remarque 3 : Dans la mesure où la procédure de filtrage fait apparaître un polynôme, donc une fonction de classe supérieure à C^2 , l'hypothèse de la classe C^2 est automatiquement vérifiée. Nous donnons, dans la section suivante, des exemples montrant la qualité, la rapidité et la robustesse de cette méthode.

4. exemples d'application

4.1. sinusoïde [figure 4]

Nous appliquons dans un premier temps cette méthode à un signal simple : une fonction sinusoïdale de période $T_1 = 50$ pour les 250 premiers points, et de période $T_2 = 100$ pour les 250 points suivants, de manière à illustrer la capacité d'adaptation de l'algorithme. La figure [Figure 4] montre les résultats pour un horizon de prédiction $h = 10$ avec $N = 5$, $M = 3$, $d = 2$, $p = 1$, $\gamma = 0$, et avec une base d'apprentissage évolutive constituée de $s = 100$ points. Sur la figure, le signal prédit est retardé de h . Ceci permet de mieux apprécier l'erreur de prédiction.

Lorsqu'on se trouve entre $t = 250$ et $t = 350$, la base d'apprentissage ne contient pas toutes les données nécessaires à la prédiction de la deuxième partie de la série. L'algorithme utilise donc à la fois les informations relatives à la première sinusoïde et les nouvelles informations concernant la deuxième série. Ensuite, après une période de la deuxième série, la base d'apprentissage contient toutes les informations nécessaires, ce qui conduit à une bonne prédiction.

4.2. série de Mackey-Glass

La série chaotique de Mackey-Glass est souvent utilisée pour évaluer la performance des prédicteurs [CHO 96]. Elle est générée à partir de l'équation différentielle :

$$\frac{dy(t)}{dt} = \frac{Ay(t - \tau)}{1 + y^{10}(t - \tau)} - By(t), \quad (7)$$

où $A = 0.2$, $B = 0.1$ et $\tau = 17$ par la méthode de Runge-Kutta d'ordre 4. Pour pouvoir comparer avec d'autres prédicteurs,

nous utilisons la méthodologie habituelle en utilisant les 500 premiers points de la série pour l'apprentissage, et les 500 suivants pour la validation (et de ce fait dans cet exemple, notre base d'apprentissage n'est plus évolutive). La figure [figure 5] montre les résultats pour un horizon de prédiction $h = 6$ et avec $N = 50$, $M = 5$, $d = 4$, $p = 4$, et $\gamma = 0$.

Pour faire la comparaison avec d'autres prédicteurs, nous donnons dans la table [Tableau 1], pour chaque méthode, les temps de calculs et la NRMSE (normalized root mean square error) donnée par la formule :

$$\text{NRMSE} = \left(\frac{\sum_{i=501}^{1000} (y_i - \hat{y}_i)^2}{\sum_{i=501}^{1000} (y_i - \bar{y})^2} \right)^{\frac{1}{2}},$$

où y_i et \hat{y}_i représentent respectivement les vraies valeurs, et les valeurs prédites de la série, alors que \bar{y} dénote la moyenne des données réelles.

Les techniques classiques de « tracking » comme le filtre de Kalman généralisé ou d'autres modèles locaux, bien qu'étant très rapides, ont l'inconvénient de ne pas prendre en compte les observations passées du signal. Les méthodes utilisant une approche globale comme les réseaux de neurones fournissent de

Tableau 1. – Comparaison entre différents prédicteurs

Prédicteur	NRMSE	temps de calcul (sec.)
Modèle AR	0.19	2
Réseau de neurones avec rétropropagation	0.02	987
Réseau de neurones avec cascade corrélation	0.06	151
ANFIS (Adaptative-Network-Based Fuzzy Inference System)	0.007	4620
Prédicteur géométrique	0.04	11

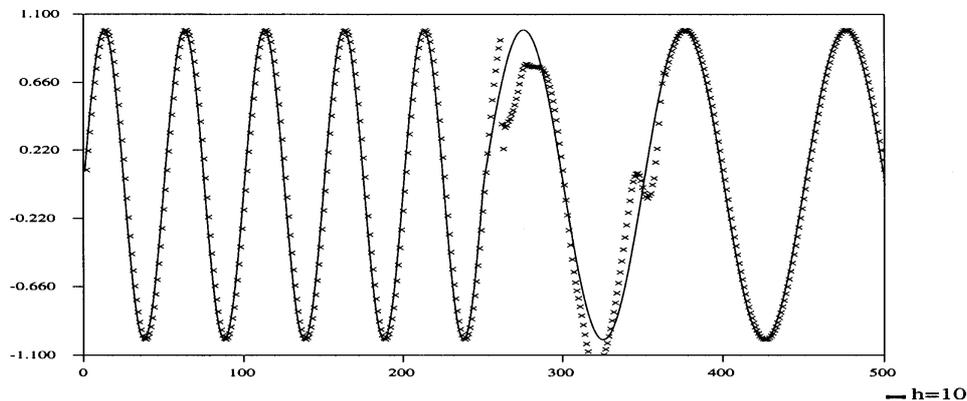


Figure 4. – Prédiction de la sinusoïde avec $h = 10$.

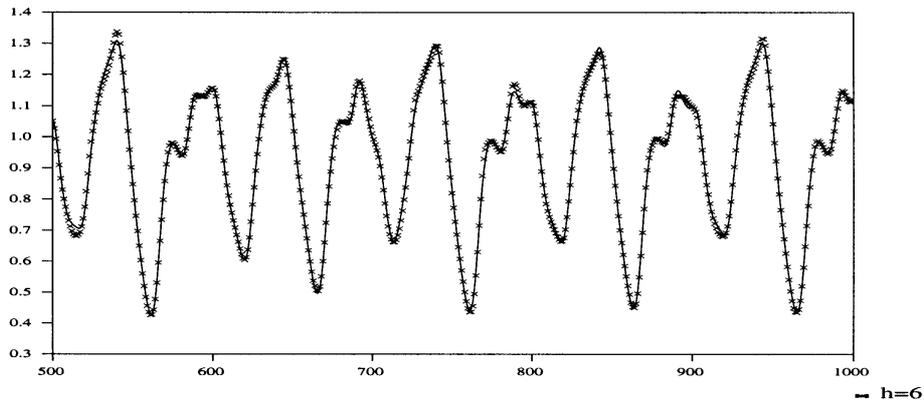


Figure 5. – Prédiction de la série de Mackey-Glass avec $h = 6$.

bons résultats mais demandent plus de calculs. La méthode présentée ici tente d'apporter un compromis entre ces deux aspects, autant en matière de précision qu'en matière de rapidité.

Pour souligner l'apport qu'offre l'évaluation de la courbure du signal pour la prédiction, nous avons comparé le prédicteur géométrique avec le prédicteur obtenu en remplaçant dans la procédure d'évaluation de y_{q+n} (voir section 2.3) les étapes (2), (3) et (4) par l'étape (2') :

2' – estimer y_{q+n} par :

$$\hat{y}_{q+n} = \frac{\sum_{i \in J_s^{(q)}} \mu_i y_{i+n}}{\sum_{i \in J_s^{(q)}} \mu_i} \quad (8)$$

Celui-ci est donc basé uniquement sur une comparaison avec les fenêtres les plus proches et ne prend pas en compte la courbure du signal.

Les erreurs de prédiction de la série de Mackey-Glass pour différentes valeurs de h de ces deux prédicteurs sont représentées sur la figure [Figure 6], ainsi que celles du prédicteur uniquement basé sur la courbure (pour lequel la prédiction est prise sur le cercle osculateur). On constate que pour de petites valeurs de h (jusque $h = 9$), le prédicteur géométrique est meilleur que les deux autres, et qu'ensuite le prédicteur utilisant uniquement les

fenêtres devient plus performant. Cela peut s'expliquer de la manière suivante :

Le prédicteur géométrique combine une procédure locale (utilisation du cercle osculateur) et une procédure globale (utilisation de la base d'apprentissage). La procédure locale lui confère une meilleure précision pour les petites valeurs de h . Par contre, pour des valeurs de h plus élevées, une faible erreur sur l'évaluation de l'angle θ_q (étape (3) de la procédure d'estimation de y_{q+n}) entraîne une grande erreur de prédiction. Dans ce cas, la procédure locale pénalise l'ensemble.

On constate par ailleurs que lorsque la procédure locale n'est pas utilisée (prédiction uniquement basée sur l'utilisation des fenêtres), l'erreur de prédiction varie peu avec h .

La méthode a été ensuite appliquée sur le même signal, mais cette fois-ci bruité (bruit uniformément distribué, centré en 0 et d'amplitude égale à 10 % de celle de la série de Mackey-Glass), montrant ainsi la robustesse de la méthode [Figure 7]. La nouvelle RMSE est ici de 0.24 (avec $N = 50$, $M = 30$, $d = 2$, $p = 4$, et $\gamma = 0.1$).

Remarque : Les seuls paramètres inhérents à la méthode sont N et p (les paramètres M , d et γ permettent le réglage du filtrage et ne sont donc pas spécifiquement liés à la méthode). Nous avons constaté que le réglage de ces paramètres influe peu sur les résultats.

5. conclusion

Cette méthode de prédiction basée sur des considérations géométriques et un apprentissage évolutif fournit d'excellents résultats et soutient largement la comparaison avec les méthodes habituelles, que ce soit au niveau de la précision qu'au niveau des temps de calculs. Les performances obtenues proviennent probablement du compromis réalisé entre la procédure locale (approche géométrique) qui améliore la précision pour des pré-

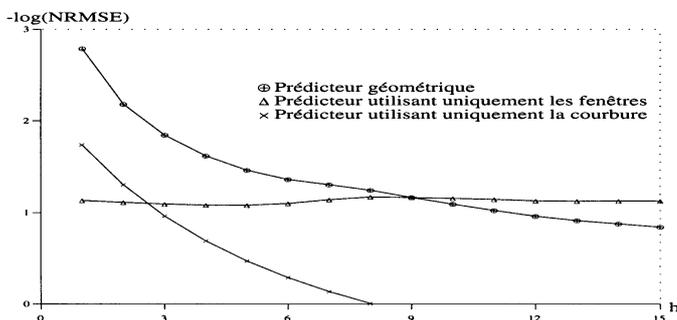


Figure 6. – Erreurs de prédiction pour différentes valeurs de h .

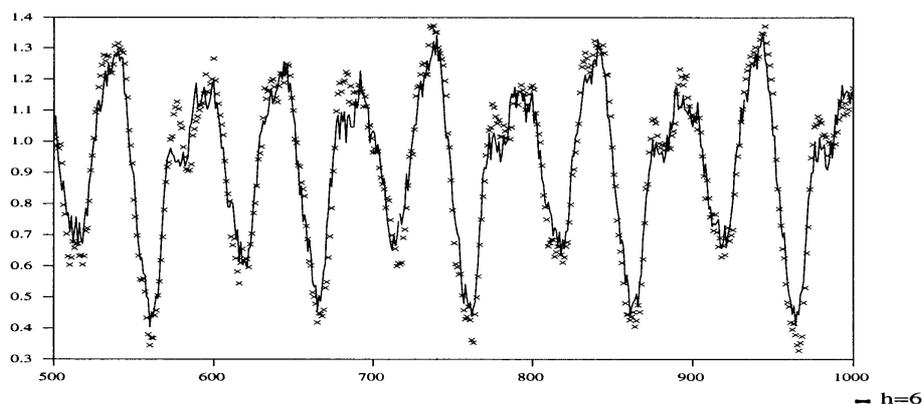


Figure 7. – Prédiction de la série de Mackey-Glass bruitée avec $h = 6$.

dictions à court terme, et la procédure globale (apprentissage évolutif) qui maintient une précision acceptable pour des prédictions à plus long terme. Par ailleurs, la procédure de filtrage intégrée à l'algorithme temps réel confère au prédicteur une bonne immunité au bruit. Enfin, la mise en œuvre d'un ensemble de relations de récurrence permet au prédicteur d'opérer en temps réel, avec une période d'échantillonnage de l'ordre de la milliseconde.

BIBLIOGRAPHIE

[ALI 96] C. Alippi, V. Piuri, « Experimental neural networks for prediction and identification », *IEEE Trans. on Instrumentation and Measurement*, Vol. 45, #2, 1996, p. 670-676.
[BOX 76] G.E.P. Box, G.M. Jenkins, « Time series analysis: forecasting and control », Oakland C.A.: Holden-Day, 1976.

[CHO 96] K.B. Cho, B.H. Wang, « Radial basis function based adaptive fuzzy systems and their applications to system identification and prediction », *Fuzzy Sets and Systems*, Vol. 83, 1996, p. 325-339.
[HOR 89] K. Hornik, M. Stinchcombe, H. White, « Multilayer feedforward networks are universal approximators », *Neural Networks*, Vol. 2, 1989, p. 359-366.
[ISH 96] M. Ishikawa, T. Moriyama, « Prediction of time series by a structural learning of neural networks », *Fuzzy Sets and Systems*, Vol. 82, 1996, p. 167-176.
[JAN 93] J.S.R. Jang, « ANFIS: Adaptive-Neural-Network-Based Fuzzy Inference Systems », *IEEE Trans. on SMC*, Vol. 23, #3, 1993, p. 665-685.
[LEE 90] C.C. Lee, « Fuzzy logic in control systems: fuzzy logic controller - Part I and II », *IEEE Trans. on SMC*, Vol. 20, #2, 1990, p. 404-435.
[MAC 77] M.C. Mackey, L. Glass, « Oscillation and chaos in physiological control systems », *Science*, Vol. 197, 1977, p. 287-289.
[WAN 92] L.X. Wang, « Fuzzy systems are universal approximators », *Proceedings of the IEEE International Conference on Fuzzy Systems*, San Diego, California, U.S.A., 1992, p. 1163-1170.
[ZAR 95] A. Zardecky, « Fuzzy controllers in nuclear material accounting », *Fuzzy Sets and Systems*, Vol. 74, 1995, p. 73-79.

Manuscrit reçu le 23 juillet 1999.

LES AUTEURS

Stéphane LEPÈRE



Stéphane Lepère est né à St-Omer en 1974. Il a obtenu son Diplôme d'Etudes Approfondies en mathématiques appliquées à l'Université des Sciences et Technologies de Lille, en 1997. Il termine actuellement son doctorat au Laboratoire d'Automatique I3D (Interaction, Image & Ingénierie de la Décision). Ses intérêts de recherche actuels comprennent la modélisation de systèmes complexes à l'aide de la logique floue et son application à la prédiction de séries temporelles.

Christian VASSEUR



Né à Cambrai en 1947, Christian Vasseur est Ingénieur (1970) diplômé de l'Institut Industriel du Nord (Ecole Centrale de Lille), Docteur Ingénieur (1972) et Docteur ès Sciences Physiques (1982) diplômé de l'Université des Sciences et Technologies de Lille.

Après un séjour de deux ans au Canada, pendant lequel il a travaillé comme Assistant de Recherche à l'Université de Sherbrooke (Québec), il a intégré l'Université des Sciences et Technologies de Lille, en

1974. Au début des années 80, il a créé, au sein de cette université, une équipe de recherche dédiée à la classification automatique en ligne et au traitement du signal en temps réel.

De 1988 à 1997, il a assuré la direction de l'Ecole Nationale Supérieure des Arts et Industries Textiles (ENSAIT) et, depuis 1997, il a réintégré l'Université des Sciences et Technologies de Lille, où il assure la direction du Laboratoire d'Automatique I3D (Interaction, Image & Ingénierie de la Décision).

Christian Vasseur est le coauteur de plus de 140 publications scientifiques. Ses centres d'intérêt se situent essentiellement dans le domaine de la classification automatique en ligne appliquée au traitement en temps réel du signal et de l'image. S'agissant de l'image, il s'est spécialisé dans l'utilisation de l'imagerie médicale (IRM, scanner, etc.) en stéréotaxie : préparation de protocoles opératoires, modélisation des volumes tumoraux, optimisation des dosages en radiothérapie et plus généralement Thérapie Assistée par l'Image.

Pour plus de détails, voir : <http://www-i3d.univ-lille1.fr/~crva/index.htm>

Xianyi ZENG



Xianyi Zeng est né à Tianjin, R.P. Chine, en 1963. Il a obtenu son diplôme de fin d'études universitaires en informatique de l'Université de Tsinghua, Beijing, en 1986, ses diplômes de doctorat et d'habilitation à diriger des recherches en automatique de l'Université des Sciences et Technologies de Lille, respectivement en 1992 et en 2000. Depuis 1993, il travaille en tant que maître de conférences à l'Ecole Nationale Supérieure des Arts et Industries Textiles de Roubaix. Ses intérêts

de recherche comprennent l'analyse de données, la reconnaissance de formes, la modélisation des systèmes complexes et ses applications dans l'industrie textile.