Régressions par machines à vecteurs supports pour la prédiction de séries chaotiques

Herwig Wendt, Patrick Flandrin et Patrice Abry

Laboratoire de Physique, UMR 5672, CNRS, Ecole Normale Supérieure de Lyon, 46, allée d'Italie, 69364, Lyon cedex 7, France.
prénom.nom@ens-lyon.fr

Résumé

Les machines à vecteurs supports (SVM) sont utilisées pour réaliser des prévisions de séries temporelles chaotiques. Les performances sont illustrées par une mise en oeuvre, d'une part, sur des données synthétiques produites par le système dynamique de Hénon et, d'autre part, sur des données expérimentales issues de la base de données de Santa Fe. Nous comparons positivement nos résultats à ceux proposés antérieurement dans la littérature.

I. Régressions par machines à vecteurs supports

- Estimateurs non linéaires universels de fonctions $y = f(\mathbf{x})$
- Objectif: trouver $h^* \in \mathcal{H}$, $h^*(\mathbf{x}) = \hat{f}(\mathbf{x})$ à partir d'un nombre limité d'échantillons $S=\{(\mathbf{x}_1, y_1), \cdots, (\mathbf{x}_n, y_n)\} \in \mathbb{R}^d \times \mathbb{R}$.
- Fonctions hypothèses h:
  - des *hyperplans* $(\mathbf{w}, b)$ d'un espace $\mathcal{F}$ de *caractéristiques*
  - $\mathcal{F} \leftrightarrow$ espace original: transformation non linéaire via *noyau* $K(\cdot,\cdot)$ (*astuce du noyau*).
- Principe d'inférence:
  - minimisation d'inégalités sur le *risque* $R[h] = \int L_\varepsilon(y, h(\mathbf{x}))dP(\mathbf{x}, y)$
  - risque: l'espérance mathématique d'une fonction de coût $L_\varepsilon(y, h(\mathbf{x}))$.
- Algorithme:
  - un problème de programmation quadratique; le problème dual de:
  - $h^*(\mathbf{w}^*, b^*, \mathbf{x}) = \arg\min_{(\mathbf{w},b)} C \sum_{(\mathbf{x}_i,y_i)\in S} L_\varepsilon(y_i, h(\mathbf{w}, b, \mathbf{x}_i)) + \frac{1}{2}\|\mathbf{w}\|^2.$ (1)
- Paramètres libres: C, ε (de la fonction de coût), les paramètres du noyau
- Solution:
  - de forme $h^*(\mathbf{x}) = \hat{f}(\mathbf{x}) = \sum_{(\mathbf{x}_i,y_i)\in S} \alpha_i^* K(\mathbf{x}_i, \mathbf{x}) + b^*$ (2)
  - obtention des multiplicateurs α_i^* : via la résolution du problème dual de (1)

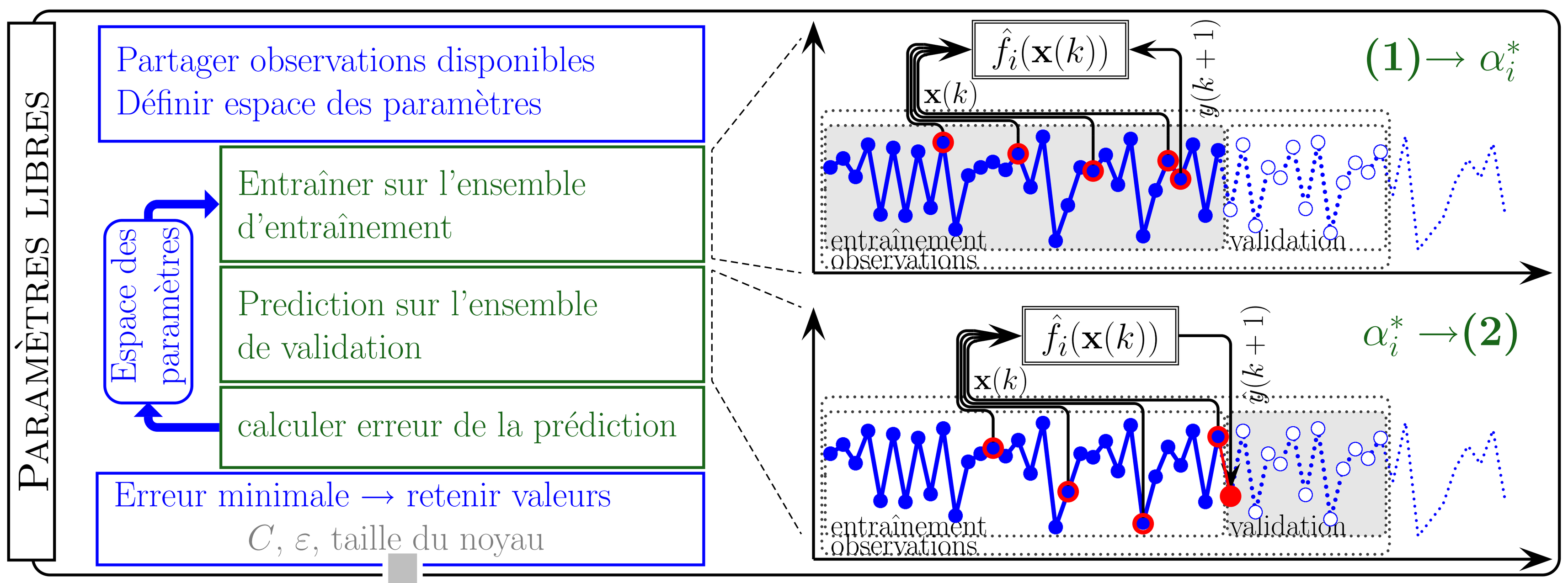
II. Séries chaotiques

- Observations scalaires y de systèmes dynamiques non linéaires
- Systèmes dynamiques non linéaires:
  - étude via *espace de phases*
  - modèle: $y(k) = f(y(k-\kappa_1), y(k-\kappa_2), \cdots, y(k-\kappa_m))$, $\kappa_i \in \mathbb{N}^+$
- Espace de phases reconstruit:
  - les *vecteurs de retard* $\mathbf{x}(k) = [y(k), y(k-\tau), \cdots, y(k-(m-1)\tau)] \in \mathbb{R}^m$
  - la dimension de plongement $m \rightarrow$ des arguments théoriques ou paramètre libre
  - le retard $\tau \rightarrow$ des arguments d'information mutuelle
- Prédiction
  - estimer la fonction $f : \mathbf{x}(k) \rightarrow y(k+1)$ à partir des observations y

III. Schéma de prédiction

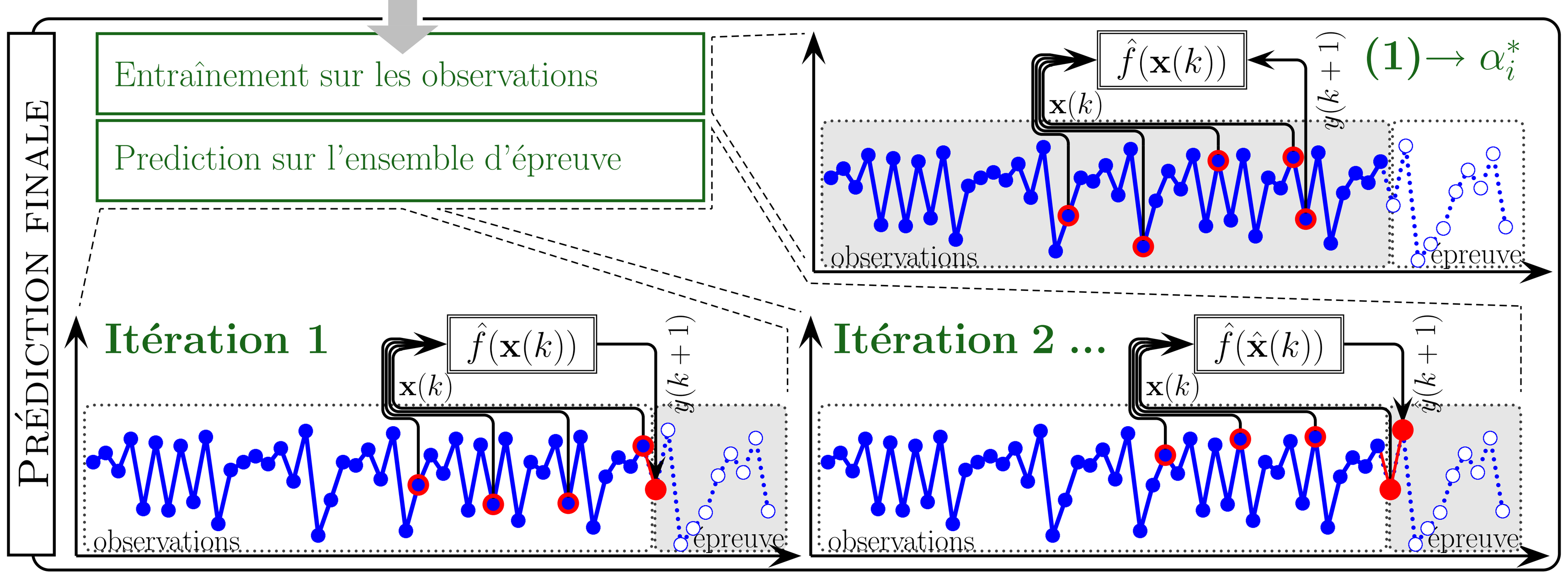
Choix des paramètres libres

- Recherche exhaustive dans l'*espace des paramètres*
  - partage des observations disponibles $\rightarrow$ ensembles d'entraînement et de validation
  - optimisation des performances de la prédiction à un pas sur l'ensemble de *validation*
- C, ϵ, la taille du noyau (fixé gaussien), éventuellement la dimension de plongement m



Réalisation de la prédiction finale

- Valeurs des paramètres $\rightarrow$ erreur de prédiction minimale sur l'ensemble de validation
- Utilisation de la *totalité des observations* disponibles
- Prédictions à plusieurs pas $\rightarrow$ réalisées par *itération* de la prédiciton à un pas



IV. Expérimentations

Evaluation de la performance

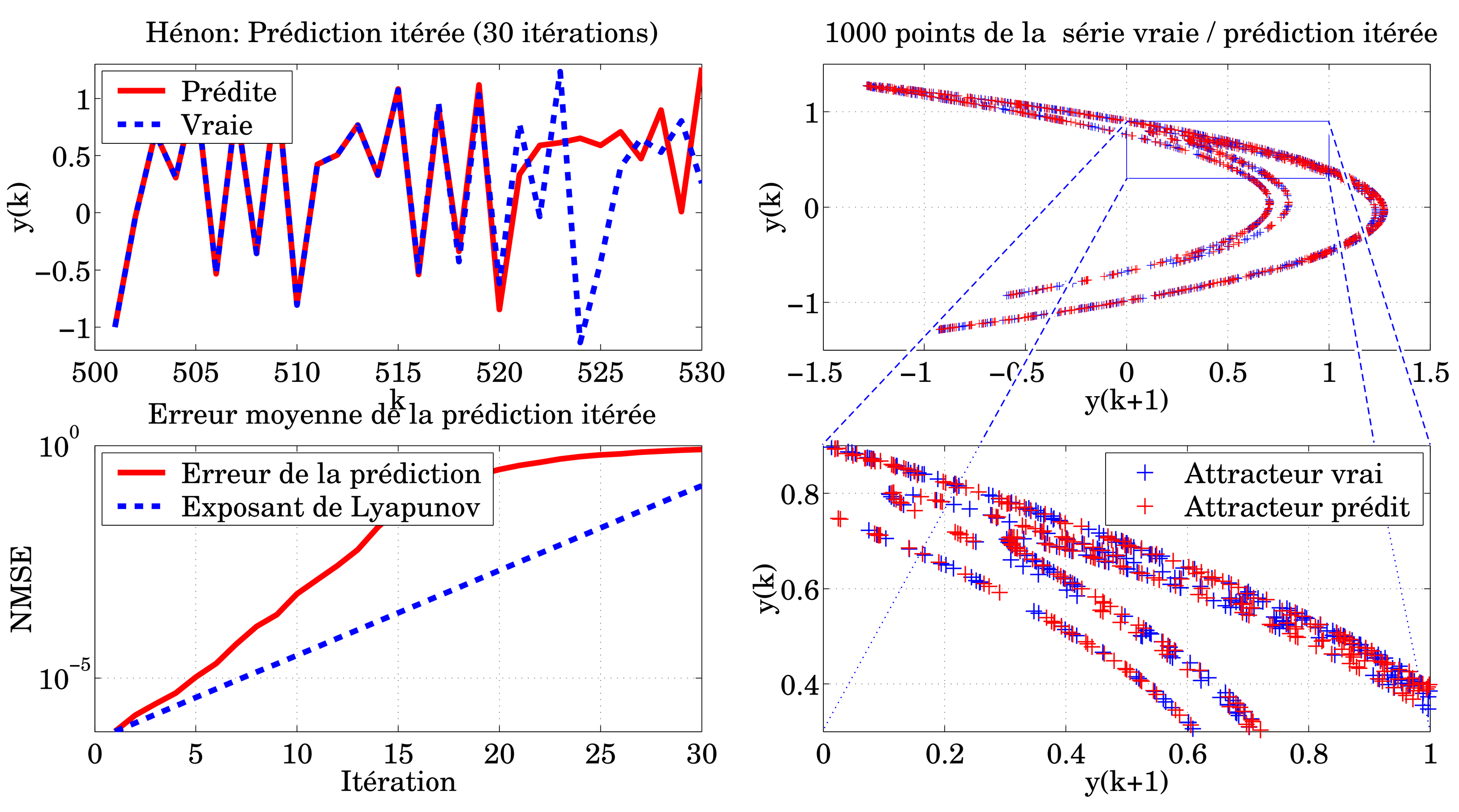
- L'erreur quadratique moyenne normalisée (*NMSE)*:
  - $\text{NMSE} = \frac{\sum_{k\in\mathcal{T}}(y(k)-\hat{y}(k))^2}{\sum_{k\in\mathcal{T}}(y(k)-\bar{y}_\mathcal{T})^2} \approx \frac{1}{\hat{\sigma}_\mathcal{T}^2 N}\sum_{k\in\mathcal{T}}(y(k)-\hat{y}(k))^2$ ($\mathcal{T}$ - l'ensemble d'épreuve)

A - La série de Hénon

- Système dynamique: $y(k+1) = 1 - 1.4y(k)^2 + 0.3y(k-1)$
- Expérimentations:
  - les observations disponibles: $N = 500$, dont l'ensemble de validation de $N_V = 50$
  - les paramètres de plongement: $m = 2$, $\tau = 1$
  - des prédictions sur ensembles d'épreuve de 50 échantillons
  - la moyenne sur 50 ensembles d'épreuve différents

Hénon - Résultat exprimé en NMSE (meilleur en gras)

Référence	Méthode	Observations	Préd. un pas
Wendt et al.	SVM Gauss	500	$\mathbf{7.7 \cdot 10^{-7}}$
Omidvar 04	RBF Net	4000	$7.4 \cdot 10^{-5}$
Wah 01	Neural Net	5000	$3.4 \cdot 10^{-5}$



B - Santa Fe Data Set A

- Données expérimentales réelles: fluctuations chaotiques de l'intensité d'un laser
- Expérimentations:
  - observations disponibles: $N = 1000$, dont ensemble de validation de $N_V = 100$
  - paramètres de plongement: $m = 18$, $\tau = 1$
  - prédictions itérées sur ensembles d'épreuve de 100 échantillons

Santa Fe data set A - Résultats exprimés en NMSE (meilleurs en gras)

Référence	Méthode	1001-1100	2181-2280	3871-3970	4001-4100	5181-5280
Wendt et al.	SVM Gauss	$4.46 \cdot 10^{-2}$	$3.93 \cdot 10^{-1}$	$4.31 \cdot 10^{-1}$	$\mathbf{8.89 \cdot 10^{-4}}$	$4.81 \cdot 10^{-2}$
Sauer 94	Local Linear	$7.7 \cdot 10^{-2}$	$1.74 \cdot 10^{-1}$	$1.83 \cdot 10^{-1}$	$6.0 \cdot 10^{-3}$	$1.11 \cdot 10^{-1}$
Small 02	Neural Net	$6.6 \cdot 10^{-2}$	$\mathbf{6.1 \cdot 10^{-2}}$	$\mathbf{8.6 \cdot 10^{-2}}$	$4.79 \cdot 10^{-1}$	$\mathbf{3.8 \cdot 10^{-2}}$
Wah 01	Neural Net	$\mathbf{1.94 \cdot 10^{-2}}$	-	-	-	-
Wan 94	Neural Net	$2.73 \cdot 10^{-2}$	$6.5 \cdot 10^{-2}$	$4.87 \cdot 10^{-1}$	$2.3 \cdot 10^{-2}$	$1.6 \cdot 10^{-1}$