

Lissage à retard fixe robuste aux données aberrantes

Boujemaa AIT EL FQUIH, Cédric GOUY-PAILLER

CEA, LIST, Laboratoire d'Outils pour l'Analyse de Données, 91191 Gif-sur-Yvette, CEDEX, France

boujemaa.aitelfquih@cea.fr, cedric.gouy-pailler@cea.fr

Résumé – Un problème important en traitement du signal consiste en l'estimation d'un processus aléatoire en présence de valeurs aberrantes parmi les observations. Des travaux récents ont été effectués dans le cadre d'un modèle d'état à bruit de mesure suivant une loi de *Student*, qui est en fait une loi à queue lourde, ce qui lui permet de tenir compte des valeurs aberrantes. Le présent travail se focalise sur le problème de lissage à retard fixe, à savoir l'estimation *en ligne*, au fur et à mesure que les observations arrivent, de l'état du système à un instant passé avec un retard fixe. Une solution de ce problème est fournie par un algorithme de calcul approché de type Kalman basé sur l'approche bayésienne variationnelle. La robustesse de cet algorithme est illustrée par des simulations ainsi qu'un exemple de données réelles de pollution de l'air.

Abstract – A recurrent issue in signal processing involves the estimation of a stochastic process in the presence of outliers among the observations. Recent works have been done within the framework of a state space model with a heavy-tailed *Student's t-distribution* for the measurement noise component. This approach therefore takes into account the presence of outliers. This work focuses on the problem of fixed-lag smoothing, namely the online estimation of the past state of the system, after a predetermined fixed delay. We propose an approximate Kalman-type algorithm based on the variational Bayesian approach. The robustness of this algorithm is illustrated by some simulations and an example of air pollution data.

1 Introduction

Le problème d'estimation d'un processus inobservable $\mathbf{x} = \{\mathbf{x}_n\}_{n=0}^N$ à partir d'un processus observé $\mathbf{y} = \{\mathbf{y}_n\}_{n=0}^N$ bénéficie d'un intérêt particulier dans de nombreuses applications comme par exemple la poursuite de cible, la navigation ou les communications numériques [1] [2] [3]. Soit $\mathbf{x}_n \in \mathbb{R}^{n_x}$ et $\mathbf{y}_n \in \mathbb{R}^{n_y}$. Ce problème est classiquement étudié dans le cadre du modèle d'état

$$\begin{cases} \mathbf{x}_{n+1} &= \mathbf{F}_n \mathbf{x}_n + \mathbf{u}_n \\ \mathbf{y}_n &= \mathbf{H}_n \mathbf{x}_n + \mathbf{v}_n \end{cases}, \quad (1)$$

dans lequel le bruit dit de système $\mathbf{u} = \{\mathbf{u}_n\}_n$ et le bruit de mesure $\mathbf{v} = \{\mathbf{v}_n\}_n$ sont supposés indépendants, conjointement indépendants et indépendants de \mathbf{x}_0 . Le problème fondamental associé au modèle (1) est celui du *filtrage*, à savoir, l'estimation à chaque instant n de l'état du système \mathbf{x}_n à partir de l'historique des mesures $\mathbf{y}_{0:n} \stackrel{\text{def}}{=} \{\mathbf{y}_0, \mathbf{y}_1, \dots, \mathbf{y}_n\}$. Dans la plupart des situations les bruits \mathbf{u} et \mathbf{v} ainsi que l'état initial \mathbf{x}_0 sont gaussiens. On note pour la suite $\mathbf{x}_0 \sim \mathcal{N}(\hat{\mathbf{x}}_0, \mathbf{P}_0)$, $\mathbf{u}_n \sim \mathcal{N}(\mathbf{0}, \mathbf{Q}_n)$ et $\mathbf{v}_n \sim \mathcal{N}(\mathbf{0}, \mathbf{R}_n)$. Dans de telles situations, le filtre de Kalman est vu comme un outil incontournable de résolution du problème du filtrage puisqu'il permet un calcul analytique et récursif de l'estimateur de la moyenne *a posteriori*, qui minimise en fait l'erreur quadratique moyenne (EQM) [4] [5].

Cependant, en présence de valeurs aberrantes (ou extrêmes) parmi les données mesurées, c'est-à-dire des données qui se situent "loin" des autres observations [6], l'hypothèse gaussienne sur le bruit de mesure n'est plus valable et la performance du filtre de Kalman se dégrade. Ce problème provient d'une sous-estimation de la probabilité d'apparition de valeurs extrêmes

dans le cas gaussien par rapport à la réalité des mesures. Or dans le cas gaussien, le filtre de Kalman considère toutes les mesures comme possibles physiquement et ignore la possibilité de présence d'anomalies. Cette hypothèse conduit parfois à une perte de performance du filtre de Kalman auquel cas il serait convenable de modéliser le bruit de mesure par une distribution à queue lourde (probabilité plus grande d'apparition de valeurs extrêmes). Dans cet esprit, un processus auxiliaire indépendant $w = \{w_n\}_n$ distribué selon une loi Gamma a été introduit dans [7] (voir aussi les références de cet article) tel que \mathbf{v} est gaussien conditionnellement à w avec ^{1 2}

$$p(\mathbf{v}_n | w_n) \sim \mathcal{N}\left(\mathbf{0}, \frac{\mathbf{R}_n}{w_n}\right), \quad p(w_n) \sim \Gamma\left(\frac{\alpha_n}{2}, \frac{\beta_n}{2}\right). \quad (2)$$

Les propriétés (2) assurent que le bruit \mathbf{v}_n associé à la mesure \mathbf{y}_n suit une loi de Student, qui notons-le, est une loi à queue lourde, ce qui permet aux valeurs aberrantes des observations d'être mieux représentées. Dans ce modèle robuste le problème de filtrage a été abordé dans [7] en introduisant un algorithme de calcul approché à base de filtre de Kalman. Le problème de filtrage a été ensuite étendu dans [9] au problème de Lissage à Intervalle Fixe (LIF) qui consiste à estimer, pour chaque instant n , l'état \mathbf{x}_n à partir du même intervalle de données $\mathbf{y}_{0:N}$ (voir également [10] [11]). En outre, le filtrage *rétrograde* (dans le sens décroissant du temps n) ainsi que le LIF ont été récemment abordé dans [12] en utilisant une approche

1. $p(\cdot)$ définit la densité de probabilité par rapport à la mesure de Lebesgue.

2. Le modèle dans [7] peut tout-de-même être considéré comme un cas particulier de (1)-(2) en ce sens que (\mathbf{x}, \mathbf{y}) est stationnaire conditionnellement à w avec des matrices de covariances \mathbf{Q} et \mathbf{R} diagonales (homogénéité spatiale).

basée sur les propriétés de markovianité rétrograde du modèle (1)-(2).

Le fait qu'il exploite l'ensemble des données, y compris les futures, le LIF conduirait à une meilleure qualité d'estimation par rapport au filtrage, bien que ce soit au prix d'une augmentation du coût de calcul. Dans le présent travail on s'intéresse au Lissage à Retard Fixe (LRF) qui quant à lui, consiste à remonter à partir des données $\mathbf{y}_{0:n}$, à l'état passé avec un retard fixe L , \mathbf{x}_{n-L} avec $0 \leq L \leq n$. Contrairement au LIF qui est un algorithme *hors-ligne*, le LRF fournit une estimation *en-ligne* de l'état du système au fur et à mesure que les données futures avec un décalage fixe L arrivent ; bien entendu le LRF se réduit au filtrage lorsque le retard L devient nul. Par ailleurs, le choix du retard L définit le compromis entre qualité d'estimation et coût de calcul. Plus L est grand plus l'estimation devient meilleure (puisque l'on utilise plus de données dans l'estimation), mais avec un coût de calcul plus important. C'est ainsi que le LRF est préféré au LIF dans les applications qui nécessitent *un monitoring* fin et quasi-temps réel du processus sous-jacent (contamination, pollution, ...). L'algorithme LRF proposé pour le modèle robuste (1)-(2) est un algorithme de calcul approché de type Kalman basé sur l'Approche Bayésienne Variationnelle (ABV) [13]. Cet algorithme est décrit en détails dans le paragraphe 2. Un exemple d'application sur des données réelles de pollution de l'air est étudié dans le paragraphe 3, suivie de la conclusion dans le paragraphe 4.

2 Lissage bayésien robuste à retard fixe

Dans le cadre bayésien général, le lissage à retard fixe L consiste à calculer dans un premier temps la loi dite de *lissage* $p(\mathbf{x}_{n-L}|\mathbf{y}_{0:n})$; un estimateur (classiquement la moyenne *a posteriori*) est ensuite déduit via,

$$\mathbb{E}_{p(\mathbf{x}_{n-L}|\mathbf{y}_{0:n})}[\mathbf{x}_{n-L}] = \int \mathbf{x}_{n-L} p(\mathbf{x}_{n-L}|\mathbf{y}_{0:n}) d\mathbf{x}_{n-L}. \quad (3)$$

En pratique, une technique de calcul classique (voir ex. [5]³) consiste à se placer dans le cadre d'un modèle à état augmenté $\{\bar{\mathbf{x}}_n, \mathbf{y}_n\}_n$, avec $\bar{\mathbf{x}}_n = [\mathbf{x}_n^T, \mathbf{x}_{n-1}^T, \dots, \mathbf{x}_{n-L}^T]^T$, dans lequel la loi $p(\bar{\mathbf{x}}_n|\mathbf{y}_{0:n})$ est calculée par un algorithme de filtrage classique ; la loi de lissage d'intérêt est ensuite déduite par marginalisation :

$$p(\mathbf{x}_{n-L}|\mathbf{y}_{0:n}) = \int p(\bar{\mathbf{x}}_n|\mathbf{y}_{0:n}) d\mathbf{x}_n d\mathbf{x}_{n-1} \dots d\mathbf{x}_{n-L+1}. \quad (4)$$

Partant de (1)-(2), le modèle augmenté s'écrit

$$\begin{cases} \bar{\mathbf{x}}_{n+1} &= \bar{\mathbf{F}}_n \bar{\mathbf{x}}_n + \bar{\mathbf{u}}_n \\ \mathbf{y}_n &= \bar{\mathbf{H}}_n \bar{\mathbf{x}}_n + \mathbf{v}_n \end{cases}, \quad (5)$$

3. Dans [5] le modèle est linéaire et gaussien dans lequel les lois $p(\bar{\mathbf{x}}_n|\mathbf{y}_{0:n})$ et $p(\mathbf{x}_{n-L}|\mathbf{y}_{0:n})$ sont gaussiennes. La propagation de ces lois se réduit ainsi à la propagation de leurs paramètres.

$$\bar{\mathbf{F}}_n = \begin{bmatrix} \mathbf{F}_n & \mathbf{0} & \dots & \mathbf{0} \\ \mathbf{I} & \mathbf{0} & \dots & \mathbf{0} \\ \vdots & \ddots & \ddots & \vdots \\ \mathbf{0} & \dots & \mathbf{I} & \mathbf{0} \end{bmatrix}, \bar{\mathbf{u}}_n = \begin{bmatrix} \mathbf{u}_n \\ \mathbf{0} \\ \vdots \\ \mathbf{0} \end{bmatrix} \text{ et } \bar{\mathbf{H}}_n = \begin{bmatrix} \mathbf{H}_n^T \\ \mathbf{0} \\ \vdots \\ \mathbf{0} \end{bmatrix}^T.$$

Dans le nouveau modèle (5)-(2) la loi $p(\bar{\mathbf{x}}_n|\mathbf{y}_{0:n})$ peut se calculer via

$$p(\bar{\mathbf{x}}_n|\mathbf{y}_{0:n}) = \int p(\bar{\mathbf{x}}_n|w_n, \mathbf{y}_{0:n}) p(w_n|\mathbf{y}_{0:n}) dw_n, \quad (6)$$

or une telle expression nécessite de connaître $p(\bar{\mathbf{x}}_n|w_n, \mathbf{y}_{0:n})$ et $p(w_n|\mathbf{y}_{0:n})$. La première loi étant gaussienne de paramètres calculables analytiquement grâce au fait que le processus couple $(\bar{\mathbf{x}}, \mathbf{y})$ soit linéaire et gaussien conditionnellement à w ; la seconde loi quant à elle est inaccessible analytiquement, ce qui rend (entre autres raisons) impossible le calcul analytique de (6). Afin de remédier à ce handicap, nous utilisons une technique de calcul approché basée sur l'ABV.

2.1 Filtrage bayésien variationnel

L'ABV consiste à approcher une loi jointe par un produit de lois marginales séparables au sens de la minimisation de la distance de *Kullback-Leibler* [13]. Dans notre cas, l'objectif est d'approcher la loi jointe *a posteriori* $p(\bar{\mathbf{x}}_n, w_n|\mathbf{y}_{0:n})$ par un produit $q(\bar{\mathbf{x}}_n|\mathbf{y}_{0:n}) \times q(w_n|\mathbf{y}_{0:n})$. Le caractère exponentiel de notre modèle (nous avons des lois gaussiennes et de type Gamma) conduit aux solutions :

$$q(\bar{\mathbf{x}}_n|\mathbf{y}_{0:n}) \propto \exp \left(\int \ln(p(\bar{\mathbf{x}}_n, w_n, \mathbf{y}_{0:n})) q(w_n|\mathbf{y}_{0:n}) dw_n \right) \quad (7)$$

$$q(w_n|\mathbf{y}_{0:n}) \propto \exp \left(\int \ln(p(\bar{\mathbf{x}}_n, w_n, \mathbf{y}_{0:n})) q(\bar{\mathbf{x}}_n|\mathbf{y}_{0:n}) d\bar{\mathbf{x}}_n \right) \quad (8)$$

Dans (7) et (8), le calcul de chacune des lois $q(\bar{\mathbf{x}}_n|\mathbf{y}_{0:n})$ et $q(w_n|\mathbf{y}_{0:n})$ requiert la connaissance de l'autre ce qui rend impossible la résolution exacte de ces équations. Afin de remédier à cette difficulté, une idée consiste à procéder en itérant de telle sorte que chacune de ces deux lois est mise à jour en fixant l'autre. En pratique, le calcul analytique de (7) et (8) à une itération donnée exploite le caractère gaussien (resp. Gamma) des lois du modèle $p(\bar{\mathbf{x}}_n|\bar{\mathbf{x}}_{n-1})$ et $p(\mathbf{y}_n|\bar{\mathbf{x}}_n, w_n)$ (resp. $p(w_n)$). Plus précisément, en supposant connue $q(w_n|\mathbf{y}_{0:n})$, le calcul de (7) conduit à une loi gaussienne $q(\bar{\mathbf{x}}_n|\mathbf{y}_{0:n})$ de paramètres calculables analytiquement via le filtre de Kalman. De même, connaissant $q(\bar{\mathbf{x}}_n|\mathbf{y}_{0:n})$, le calcul de (8) conduit à une loi *a posteriori* $q(w_n|\mathbf{y}_{0:n})$ Gamma (de même famille que l'a priori $p(w_n)$) et de paramètres calculables analytiquement. Considérons les notations suivantes par la suite. $\hat{w}_n|_j = \mathbb{E}_{q(w_n|\mathbf{y}_{0:j})}[w_n]$, $\hat{\bar{\mathbf{x}}}_n|_j = \mathbb{E}_{q(\bar{\mathbf{x}}_n|\mathbf{y}_{0:j})}[\bar{\mathbf{x}}_n]$ et $\bar{\mathbf{P}}_n|_{i:j} = \mathbb{E}_{q(\bar{\mathbf{x}}_n|\mathbf{y}_{0:j})}[(\bar{\mathbf{x}}_n - \hat{\bar{\mathbf{x}}}_n|_j)(\bar{\mathbf{x}}_n - \hat{\bar{\mathbf{x}}}_n|_j)^T]$. Cet algorithme de calcul est décrit dans la proposition suivante. Une version similaire a été introduite récemment dans la littérature (voir ex. [7]) et dans laquelle, contrairement à Prop. 1, les équations de Kalman sont écrites en forme information (propagation des paramètres d'information de la loi gaussienne en question).

Proposition 1 On considère le modèle (5)-(2). Les paramètres $\hat{\mathbf{x}}_{n|n}$ et $\bar{\mathbf{P}}_{n|n}$ peuvent être calculés à partir de $\hat{\mathbf{x}}_{n-1|n-1}$ et $\bar{\mathbf{P}}_{n-1|n-1}$ comme suit.

• Prédiction.

$$\begin{aligned}\hat{\mathbf{x}}_{n|n-1} &= \bar{\mathbf{F}}_{n-1} \hat{\mathbf{x}}_{n-1|n-1} \\ \bar{\mathbf{P}}_{n|n-1} &= \bar{\mathbf{F}}_{n-1} \bar{\mathbf{P}}_{n-1|n-1} \bar{\mathbf{F}}_{n-1}^T + \bar{\mathbf{Q}}_{n-1}\end{aligned}$$

• Mise à jour. On procède par itérations en démarrant par l'initialisation $\hat{w}_{n|n} = \frac{\alpha_n}{\beta_n}$. Pour $i = 1, \dots, I \gg 1$,

$$\begin{aligned}\mathbf{L}_n &= \bar{\mathbf{H}}_n \bar{\mathbf{P}}_{n|n-1} \bar{\mathbf{H}}_n^T + \frac{\mathbf{R}_n}{\hat{w}_{n|n}} \\ \bar{\mathbf{K}}_n &= \bar{\mathbf{P}}_{n|n-1} \bar{\mathbf{H}}_n^T \mathbf{L}_n^{-1} \\ \tilde{\mathbf{y}}_n &= \mathbf{y}_n - \bar{\mathbf{H}}_n \hat{\mathbf{x}}_{n|n-1} \\ \hat{\mathbf{x}}_{n|n} &= \hat{\mathbf{x}}_{n|n-1} + \bar{\mathbf{K}}_n \tilde{\mathbf{y}}_n \\ \bar{\mathbf{P}}_{n|n} &= \bar{\mathbf{P}}_{n|n-1} - \bar{\mathbf{P}}_{n|n-1} \bar{\mathbf{H}}_n^T \bar{\mathbf{K}}_n^T \\ \hat{w}_{n|n} &= \frac{\alpha_n + n_{\mathbf{y}}}{\beta_n + \|\mathbf{y}_n - \bar{\mathbf{H}}_n \hat{\mathbf{x}}_{n|n}\|_{\mathbf{R}_n}^2 + \text{Tr}[\bar{\mathbf{H}}_n^T \mathbf{R}_n^{-1} \bar{\mathbf{H}}_n \bar{\mathbf{P}}_{n|n}]}\end{aligned}$$

où $\|\mathbf{v}\|_{\mathbf{M}}^2 \stackrel{\text{def}}{=} \mathbf{v}^T \mathbf{M} \mathbf{v}$ pour un vecteur \mathbf{v} et une matrice définie positive \mathbf{M} ; $\text{Tr}[\cdot]$ représente la trace d'une matrice.

2.2 Lissage à retard fixe

Nous arrivons maintenant à notre but ultime de calcul *en ligne*, pour un retard fixe $L \geq 0$, d'une approximation $\hat{\mathbf{x}}_{n-L|n}$ de l'estimateur de la moyenne *a posteriori* $\mathbb{E}_{p(\mathbf{x}_{n-L}|\mathbf{y}_{0:n})}[\mathbf{x}_{n-L}]$ et la matrice de covariance associée $\bar{\mathbf{P}}_{n-L|n}$. L'algorithme de calcul de ces paramètres est déduit de Prop. 1 (rappelons que \mathbf{x}_{n-L} est "une composante" de $\bar{\mathbf{x}}_n$). De manière plus générale, les équations de calcul de $\hat{\mathbf{x}}_{n|n}$ et $\bar{\mathbf{P}}_{n|n}$ dans Prop. 1 conduisent au calcul de tous les estimateurs de lissage $\hat{\mathbf{x}}_{n-k|n}$, avec $k = 0, 1, \dots, L$, ainsi qu'aux matrices de covariance $\bar{\mathbf{P}}_{n-k|n}$ (notons que le cas particulier $k = 0$ correspond au filtrage). Soit la notation

$$\bar{\mathbf{P}}_{n|j} = \begin{bmatrix} \mathbf{P}_{n,j} & \mathbf{P}_{n,n-1|j} & \cdots & \mathbf{P}_{n,n-L|j} \\ \mathbf{P}_{n-1,n|j} & \mathbf{P}_{n-1|j} & \cdots & \mathbf{P}_{n-1,n-L|j} \\ \vdots & \cdots & \ddots & \vdots \\ \mathbf{P}_{n-L,n|j} & \mathbf{P}_{n-L,n-1|j} & \cdots & \mathbf{P}_{n-L|j} \end{bmatrix}.$$

L'algorithme de lissage à retard fixe est décrit dans la proposition suivante.

Proposition 2 On considère le modèle (1)-(2). Supposons qu'à l'instant $n = L$ nous avons $\hat{\mathbf{x}}_{n|n}$ et $\bar{\mathbf{P}}_{n|n}$ grâce à Prop. 1. Donc, pour $n = L + 1, \dots, N$,

• Pour $k = 1, 2, \dots, L$,

$$\begin{aligned}\mathbf{P}_{n,n-k|n-1} &= \mathbf{F}_{n-1} \mathbf{P}_{n-1,n-k|n-1} \\ \mathbf{K}_{n,k} &= \mathbf{P}_{n,n-k|n-1} \mathbf{H}_n^T \mathbf{L}_n^{-1} \\ \mathbf{P}_{n,n-k|n} &= \mathbf{P}_{n,n-k|n-1} - \mathbf{K}_{n,k} \mathbf{L}_n \mathbf{K}_{n,k}^T \\ \hat{\mathbf{x}}_{n-k|n} &= \hat{\mathbf{x}}_{n-k|n-1} + \mathbf{K}_{n,k} \tilde{\mathbf{y}}_n \\ \bar{\mathbf{P}}_{n-k|n} &= \bar{\mathbf{P}}_{n-k|n-1} - \mathbf{K}_{n,k} \mathbf{L}_n \mathbf{K}_{n,k}^T.\end{aligned}$$

• Mettre en œuvre l'algorithme de filtrage de Prop. 1 à l'instant suivant dans le cadre du modèle (\mathbf{x}, \mathbf{y}) : en remplaçant dans Prop. 1, $\hat{\mathbf{x}}_{n|j}$, $\bar{\mathbf{P}}_{n|j}$, $\bar{\mathbf{Q}}_n$ et $\bar{\mathbf{H}}_n$ par $\hat{\mathbf{x}}_{n|j}$, $\bar{\mathbf{P}}_{n|j}$, \mathbf{Q}_n et \mathbf{H}_n respectivement avec $j = n - 1, n$.

3 Application

Afin de mettre en évidence la validité de notre algorithme et sa robustesse face aux valeurs aberrantes, un exemple de simulations suivi d'un exemple de données réelles sont étudiés.

3.1 Exemple de simulations

On considère $\mathbf{F} = \begin{bmatrix} 0.85 & 0.01 \\ 0.13 & 0.7 \end{bmatrix}$, $\mathbf{H} = [0.37 \ 0.55]$, $\mathbf{Q} = 0.1\mathbf{I}_2$ et $R = 0.1$. On génère $N = 500$ mesures contenant des valeurs atypiques. Les mesures atypiques sont générées de telle sorte que $w_n = 1$ avec une probabilité de 85% et $w_n = 0.02$ avec une probabilité de 15%. On s'intéresse à la reconstruction des mesures $\hat{y}_{n-L|n}$ calculées après avoir estimé l'état $\hat{\mathbf{x}}_{n-L|n}$ ($\hat{y}_{n-L|n} = \mathbf{H} \hat{\mathbf{x}}_{n-L|n}$). L'algorithme utilise $\alpha_n = \beta_n = 2$ et un nombre d'itérations $I = 10$. Le résultat de reconstruction à l'aide du lissage avec $L = 10$ ainsi que le filtrage ($L = 0$) est indiqué dans la Fig. 1. Cette figure montre que les algorithmes sont bien robustes aux données aberrantes. Les Fig. 2 et 3 représentent l'EQM moyennée sur 10 réalisations indépendantes. Nous constatons bien que le lissage fourni "une erreur" inférieure à celle du filtrage.

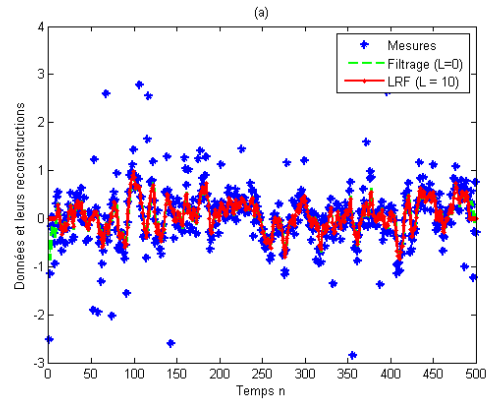


FIGURE 1 – Données observées y_n et leurs reconstructions.

3.2 Exemple de données réelles

Cet exemple s'inscrit dans la surveillance de la qualité de l'air dans une station de métro parisienne (données RATP – Régie Autonome des Transports Parisiens, disponibles sur le site <http://data.ratp.fr>). Nous avons considéré plus particulièrement une série de mesures de concentration d'oxyde d'azote (NO) (en $\mu\text{g}/\text{m}^3$) prises sur une période approximative d'une heure à la station de Auber, en fonction de certains paramètres climatiques que sont la température (en $^\circ\text{C}$) et l'humidité relative

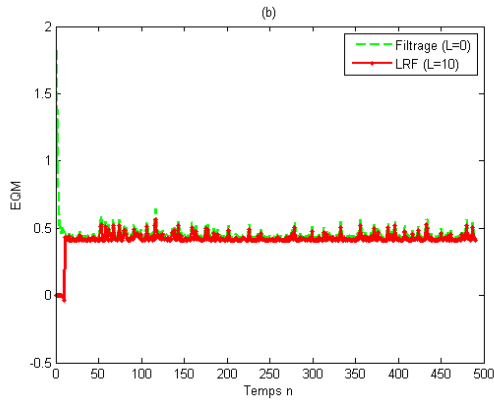


FIGURE 2 – EQM de l'estimation de l'état x_n .

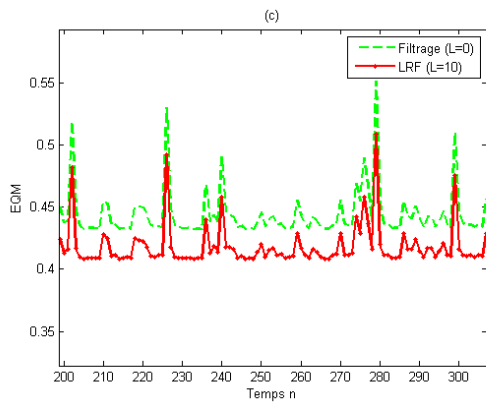


FIGURE 3 – Zoom de Fig. 2.

(en %). $N = 500$ mesures sont considérées en commençant par celle prise le 03/03/2012 à 12h00. Ces mesures contiennent des valeurs aberrantes comme le montre Fig. 4. Cette même Fig. indique la reconstruction des mesures par notre algorithme de lissage avec un retard fixe L de 6 h. Pour mettre en œuvre l'algorithme, les paramètres du modèle sont d'abord estimés par un algorithme de type EM (Expectation-Maximization). Le résultat de reconstruction est concluant dans la mesure où moins d'importance est donnée aux valeurs aberrantes.

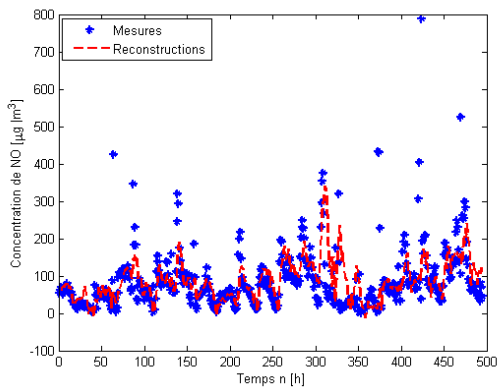


FIGURE 4 – Concentrations mesurées de NO et leurs reconstructions.

4 Conclusion

Cet article a été consacré au problème de lissage à retard fixe dans les situations où les observations sont corrompues par des valeurs aberrantes. Dans un modèle d'état linéaire pour lequel le bruit de mesure suit une loi de Student (à queue lourde), nous avons proposé en utilisant l'approche bayésienne variationnelle un algorithme robuste de calcul approché de type Kalman. L'algorithme a été validé via des simulations ainsi qu'un exemple de données réelles de pollution de l'air d'un environnement souterrain.

Références

- [1] Y. Bar-Shalom, X. Li et T. Kirubarajan, *Estimation with Applications to Tracking and Navigation*. New York : Wiley, 2001.
- [2] H. Durrant-Whyte et T. Bailey, *Simultaneous Localisation and Mapping (SLAM) : Part I The Essential Algorithms*. IEEE Robot. Autom. Mag., 13 : 99-110, 2006.
- [3] P. M. Djuric, J. Zhang, T. Ghirmai, Y. Huang et J. H. Kotecha, *Applications of Particle Filtering to Communications : A Review*. Proceedings of the EUSIPCO, 2002.
- [4] R. E. Kalman, *A New Approach to Linear Filtering and Prediction Problems*. Journal of Basic Eng., Trans. ASME, Series D, 82 : 35-45, 1960.
- [5] B. D. O. Anderson et J. B. Moore, *Optimal Filtering*. Prentice Hall, Englewood Cliffs, New Jersey, 1979.
- [6] D.S. Moore et G. P. McCabe, *Introduction to the Practice of Statistics*. W.H. Freeman & Company, March, 1999.
- [7] J. Ting, E. Theodorou et S. Schaal, *A Kalman filter for robust outlier detection*. Proceedings of the IROS, 2007.
- [8] J. Ting, E. Theodorou et S. Schaal, *Learning an outlier-robust Kalman filter*. Proceedings of the ECML, 2007.
- [9] G. Agamennoni et J. Nieto, *An outlier-robust Kalman filter*. Proceedings of the IEEE ICRA, 2011.
- [10] G. Agamennoni, J. Nieto and E. Nebot, *Approximate Inference in State-Space Models With Heavy-Tailed Noise*. IEEE Trans. on Signal Processing, vol. 60, no. 10, pp.5024–37, Oct. 2012.
- [11] R. Peché, S. Särkkä et J. Hartikainen, *Recursive Outlier-Robust Filtering and Smoothing for Nonlinear Systems using the Multivariate Student-t Distribution*. Proceedings of the MLSP, 2012.
- [12] B. Ait-El-Fquih et C. Gouy-Pailler, *Backward Hidden Markov Chain for Outlier-Robust Filtering and Fixed-Interval Smoothing*. Proceedings of the the IEEE ICASSP, Vancouver, Canada, May 26-31, 2013.
- [13] V. Smidl et A. Quinn, *The Variational Bayes Method in Signal Processing*. Springer, 2006.