A NEW ROBUST ESTIMATION METHOD FOR ARMA MODELS

Yacine Chakhchoukh, Patrick Panciatici

RTE, DMA
Versailles, France

Pascal Bondon, Lamine Mili

CNRS, Univ. Paris-Sud – Virginia Tech-ECE Dept
Gif-sur-Yvette, France– Falls Church, USA

ABSTRACT

This paper presents a new robust method to estimate the parameters of ARMA models. This method makes use of the autocorrelations estimates based on the ratio of medians together with a robust filter cleaner able to reject a large fraction of outliers, and a Gaussian maximum likelihood estimation which handles missing values. The main advantages of the procedure are its easiness, robustness and fast execution. Its effectiveness is demonstrated on an example of the forecasting of the French daily electricity consumptions.

***Index Terms*—** Robustness, time series, ARMA models, outliers, missing values.

1. INTRODUCTION

Robust estimation is important when estimating a statistical model. When the data contains deviant observations termed outliers, the classical statistical estimators of an ARMA model become unreliable. Thus order selection, parameter estimation, and forecasting can be affected notably. In the robust statistics literature, several methods were proposed mainly for iid data and for regression models. Some methods were proposed in the context of time series such as the filtered–τ, filtered M–, generalized M– and the so-called *Residual Autocovariance* (RA)–estimators [1, 2]. In this paper, we propose a new robust procedure where the autocorrelations estimates are based on the ratio of medians. In practice, our method is comparable to the above mentioned methods in terms of performance of estimation. While being highly robust with a breakdown point of 25 %, it compares favorably to other methods in terms of simplicity and computing time. When applied to the actual French daily electric consumptions, it significantly reduces the running time as compared to the filtered-τ estimators. We analyze the robustness of our method using the standard tools such as the time series influence functionals defined in [1] and the maximum bias curve, which is computed numerically. The effectiveness of our method is demonstrated on an example of the forecasting of the French daily electricity consumptions. The paper is organized as follows. In Section 2, we present the method and analyze its robustness. Section 3 presents the simulation results in the case of electricity load modeling and forecasting. Finally, Section 4 concludes the paper.

2. MEDIAN-BASED FILTERING

Our idea is to robustly estimate the autocorrelation and partial-autocorrelation functions, then to fit a high order AR model and filtering with this AR model to "clean" the data from the outliers, and finally to estimate an ARMA model in the presence of missing values.

First, we show how to estimate the correlation ρ in a Gaussian vector using medians, the latter being known to lead to robust estimates. Consider a zero mean Gaussian vector (X, Y) with density

$$\varphi_{\rho,\sigma^2}(x,y) = \frac{\exp\left[-\frac{1}{2\sigma^2(1-\rho^2)}\left(x^2+y^2-2\rho xy\right)\right]}{2\pi\sigma^2\sqrt{1-\rho^2}}. \quad (1)$$

The density of the product XY is given by

$$\begin{aligned} f_{\rho,\sigma^2}(v) &= \int_{-\infty}^{+\infty} \varphi_{\rho,\sigma^2}\left(x, \frac{v}{x}\right)\frac{dx}{|x|} \\ &= \frac{e^{\frac{\rho v}{\sigma^2(1-\rho^2)}}}{\pi\sigma^2\sqrt{1-\rho^2}} K_0\left(\frac{|v|}{\sigma^2(1-\rho^2)}\right) \end{aligned}$$

where $K_0(\cdot)$ is the modified Bessel function of the second kind [3]. The random variable X^2/σ^2 follows a standard χ^2 distribution. For any distribution function F, the median of F is defined as

$$\xi_F = \inf\{x : F(x) \geq 1/2\}. \quad (2)$$

Corresponding to a sample $\{X_1, \dots, X_n\}$ of observations on F, the sample median $\widehat{\xi}_F$ is defined as the median of the sample distribution function. Let F_{ρ,σ^2}, resp. G_{σ^2} be the distribution function of XY, resp. X^2. The ratio τ of medians satisfies

$$\tau = \xi_{F_{\rho,\sigma^2}}/\xi_{G_{\sigma^2}} = \xi_{F_{\rho,1}}/\xi_{G_1}, \quad (3)$$

where $\xi_{G_1} \simeq 0.45$. An explicit relation between τ and ρ does not seem to exist. Fig. 1 represents ρ as a function of τ, $\rho = r(\tau)$, and is obtained numerically.

Consider now a Gaussian stationary time series $\{X_t\}$ with variance σ^2 and autocorrelation function $\rho(\cdot)$. For each $k \in \mathbb{N}$, we define $\widehat{\tau}(k)$ by

$$\widehat{\tau}(k) = \widehat{\xi}_{F_{\rho(k),\sigma^2}}/\widehat{\xi}_{G_{\sigma^2}}, \quad (4)$$

where $F_{\rho(k),\sigma^2}$, resp. G_{σ^2} are the monovariate distribution function of $\{X_t X_{t-k}\}$, resp. $\{X_t^2\}$. The sample medians $\widehat{\xi}_{F_{\rho(k),\sigma^2}}$ and $\widehat{\xi}_{G_{\sigma^2}}$ are calculated using the single ergodic and stationary time series $\{X_t\}$. Then a robust estimate $\widehat{\rho}(k)$ of $\rho(k)$ is obtained by the relation $\widehat{\rho}(k) = r(\widehat{\tau}(k))$. Call the obtained estimator ratio-of-medians-based estimator (RME).

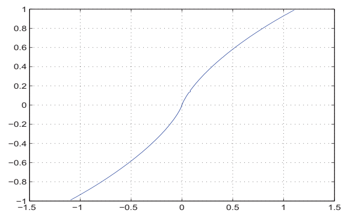

Fig. 1. Correlation coefficient of X and Y versus the ratio of medians, τ, given by (3).

2.1. Deriving the influence function of the new estimator

Consider the estimation of ϕ in the simple case of the first-order autoregressive model AR(1)

$$Y_t = \phi Y_{t-1} + \varepsilon_t, \;\; \varepsilon_t \sim N(0, \sigma_\varepsilon^2) \tag{5}$$

Our estimator is $\widehat{\phi}_r = \widehat{\rho}(1) = r(\widehat{\tau}(1))$. $\widehat{\phi}_{r,\infty}$, resp. $\widehat{\tau}_\infty$ is the asymptotic value of $\widehat{\phi}_r$, resp. $\widehat{\tau}$ when the length of the series tends to infinity. The mapping: $\tau \rightarrow \phi$, $|\phi| < 1$ is one to one. robustness means that outliers cause only small deviations to an estimator. Fig.1 represents a numerical approximation of $r(\cdot)$. If $\widehat{\tau}_\infty$ is robust then $\widehat{\phi}_{r,\infty}$ is robust since $r(\cdot)$ is continuous (Fig.1).

To analyze the infinitesimal effect of outliers, we calculate the influence function. In the literature, we encountered two definitions of influence functions for time series. The first definition was given by Künsch [1] and the second one, which we use in this work, was introduced by Martin and Yohai [1, 4], who define the influence function by[1, page 302]

$$\text{IF}\left(\{F^\epsilon_{X,Z,W}\}; \widehat{\lambda}\right) = \lim_{\substack{\epsilon \to 0 \\ >}} \frac{1}{\epsilon}\left(\widehat{\lambda}_\infty(F_Y^\epsilon) - \widehat{\lambda}_\infty(F_X)\right) = \frac{\partial}{\partial \epsilon}\widehat{\lambda}_\infty(F_Y^\epsilon)|_{\epsilon \to 0}$$

$\widehat{\lambda}_\infty(F)$ is the asymptotic value of the estimate when the underlying distribution is F; $F^\epsilon_{X,Z,W}$ is the joint distribution of the processes $\{X_t\}$, $\{Z_t^\epsilon\}$ and $\{W_t\}$; F_Y^ϵ is the joint distribution of the process $\{Y_t^\epsilon\}$; $Y_t^\epsilon = (1 - Z_t^\epsilon)X_t + Z_t^\epsilon W_t$; $\{X_t\}$ is a stationary Gaussian process with a joint distribution F_X. W_t is an outlier generating process and $\{Z_t^\epsilon\}$ is a 0-1 process with $\text{P}(Z_t^\epsilon = 1) = \epsilon$. The time series influence function is a functional on a probability distribution space.

We choose a contamination model that simplifies the calculation and better represents the actual French daily load consumption series. We consider the case of replacement isolated outliers where $\{W_t\}$ is independent of $\{X_t\}$. $\{X_t\}$ follows a Gaussian process AR(1) with correlation ϕ and variance σ^2. $\{W_t\}$ is an iid Gaussian process with variance κ^2, $N(0, \kappa^2)$. One should notice that even if the Gaussian contamination is centered, it causes a bias to a classical estimator. The estimator of lag-1 autocorrelation coefficient, for example, converges to $\widehat{\phi}_\infty(F_Y^\epsilon) = \rho(1) = \text{E}(Y_t^\epsilon Y_{t-1}^\epsilon)/\text{E}[(Y_t^\epsilon)^2]$. In the previously contaminated case, $\widehat{\phi}_\infty(F_Y^\epsilon)$ is equal to

$$\frac{(1-\epsilon)^2\phi\sigma^2 + 2\epsilon(1-\epsilon)\text{E}(X_t W_{t-1}) + \epsilon^2 \text{E}(W_t W_{t-1})}{(1-\epsilon)\sigma^2 + \epsilon\kappa^2}$$

The bias of $\widehat{\phi}_\infty(F_Y^\epsilon)$ becomes

$$b = \phi - \frac{\text{E}(Y_t^\epsilon Y_{t-1}^\epsilon)}{\text{E}[(Y_t^\epsilon)^2]} = \phi\frac{\epsilon(\kappa^2 + \sigma^2) - \epsilon^2\sigma^2}{(1-\epsilon)\sigma^2 + \epsilon\kappa^2}$$

Thus, the centered Gaussian contamination introduces a bias that depends on κ. The definition of the influence function in the case of time series is more general than in the case of iid. In this paper, a Gaussian form of outlier-generating process $\{W_t\}$ is chosen. $\widehat{\phi}_{r,\infty}$ depends only on the joint distribution $F_{(Y_t^\epsilon, Y_{t-1}^\epsilon)}$. The joint density $f_{(Y_t^\epsilon, Y_{t-1}^\epsilon)}(x, y)$ is

$$(1-\epsilon)^2\varphi_{(\phi,\sigma^2)}(x,y) + \epsilon(1-\epsilon)\varphi_{\kappa^2}(x)\varphi_{\sigma^2}(y) + \epsilon(1-\epsilon)\varphi_{\sigma^2}(x)\varphi_{\kappa^2}(y) + \epsilon^2\varphi_{(0,\kappa^2)}(x,y)$$

where $\varphi_{\kappa^2}(x) = \frac{1}{\sqrt{2\pi\kappa^2}}\exp\left(-\frac{x^2}{2\kappa^2}\right)$.

First, we evaluate $\text{IF}\left(\{F^\epsilon_{X,Z,W}\}; \widehat{\xi}_{H_{\phi,\kappa,\sigma}}\right)$ denoted by IF_1 evaluated using the monovariate probability distribution function $H_{\phi,\kappa,\sigma}$ of the contaminated product, $\{Y_t^\epsilon Y_{t-1}^\epsilon\}$ with a probability density function $h_{\phi,\kappa,\sigma}(v)$

$$(1-\epsilon)^2 f_{\phi,\sigma^2}(v) + \frac{2\epsilon(1-\epsilon)}{\pi\sigma\kappa}K_0\left(\frac{|v|}{\sigma\kappa}\right) + \epsilon^2 f_{(0,\kappa^2)}(v)$$

The expression of IF_1 is derived as

$$IF_1 = \sigma^2\sqrt{1-\phi^2}\left[\frac{\pi - 2\int_{-\infty}^{\xi_{F_{\phi,1}}} \frac{\sigma}{\kappa}K_0\left(|v|\frac{\sigma}{\kappa}\right)dv}{e^{\frac{\phi\xi_{F_{\phi,1}}}{(1-\phi^2)}}K_0\left(\frac{|\xi_{F_{\phi,1}}|}{1-\phi^2}\right)}\right]$$

$\eta = \frac{\kappa}{\sigma} > 0 \Rightarrow \eta^{-1}K_0\left(\eta^{-1}|v|\right) > 0$ then $\int_{-\infty}^{\xi_{F_{\phi,1}}}\eta^{-1}K_0\left(\eta^{-1}|v|\right)dv$ is bounded by $\int_{-\infty}^{\infty}\eta^{-1}K_0\left(\eta^{-1}|v|\right)dv = \pi$. This means that the influence function IF_1 is bounded for any $\kappa > 0$ ($\eta > 0$). η is the ratio of variances.

Next, $\text{IF}(\{F^\epsilon_{X,Z,W}\}; \widehat{\xi}_{I_{\sigma,\kappa}})$ is evaluated and denoted by IF_2, which is expressed as

$$IF_2 = \sigma^2\sqrt{2\pi\xi_{G_1}}\left(\text{erf}\left(\frac{\sqrt{\xi_{G_1}}}{\sqrt{2}}\right) - \text{erf}\left(\frac{\sqrt{\xi_{G_1}}}{\sqrt{2}}\frac{\sigma}{\kappa}\right)\right)e^{\frac{\xi_{G_1}}{2}},$$

where $I_{\sigma,\kappa}$ is the monovariate distribution function of $\{(Y_t^\epsilon)^2\}$, $I_{\sigma,\kappa} = (1-\epsilon)G_{\sigma^2} + \epsilon G_{\kappa^2}$ and $\text{erf}(z) = \frac{2}{\sqrt{\pi}}\int_0^z \exp(-x^2)dx$. Finally, it follows that the time series influence function is given by

$$\text{IF}(\{F^\epsilon_{X,Z,W}\}; \widehat{\phi}_r) = r'(\phi)\text{IF}\left(\{F^\epsilon_{X,Z,W}\}; \widehat{\tau}\right) = r'(\phi)\left(\frac{IF_1}{\xi_{G_{\sigma^2}}} - \frac{\xi_{F_{\phi,\sigma^2}}IF_2}{\left(\xi_{G_{\sigma^2}}\right)^2}\right)$$

It is observed that IF_1 and IF_2 are bounded for $\kappa > 0$, ($\eta > 0$). $r'(0)$, $\xi_{F_{\phi,\sigma^2}}$ and $\xi_{G_{\sigma^2}}$ do not depend on the contamination or κ. This implies that $\text{IF}(\{F^\epsilon_{X,Z,W}\}; \widehat{\phi}_r)$ is bounded for $\kappa > 0$. For a given value of ϕ, this influence function is evaluated numerically for $\kappa > 0$. For $\phi = 0.5$, $r'(0.5) \simeq 1.1208$, where $r'(\cdot)$ is the derivative of the $r(\cdot)$ and is computed numerically.

The influence function for the least-squares estimator, which is a non robust estimator, is given by

$$\text{IF}\left(\{F^{\epsilon}_{Z,X,W}\}; \sum Y_t Y_{t-1} / \sum Y_t^2\right) = -\phi\left(1+\frac{\kappa^2}{\sigma^2}\right)$$

It is clear that the influence function of this estimator is not bounded for $\kappa > 0$. This is substantiated in Fig.2, which displays the IF of the classical least squares estimator (solid line) and that of our RME (dashed line). It is observed that the former is not bounded while the latter is.

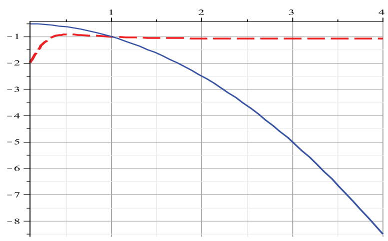

Fig. 2. Influence curves of $\sum Y_t Y_{t-1} / \sum Y_t^2$ (solid line) and RME $\widehat{\phi}_r$ estimator (dashed line) versus $\eta > 0$ for an AR(1) at $\phi = 0.5$ with the previous Gaussian contamination.

2.2. Maximum bias curves in the case of AR(1)

The maximum bias curves of the RME are calculated following the Monte Carlo procedure described in [1, page 305]. For AR(1), Fig.3 depicts the maximum bias curve of our RME together with that of another robust estimator, namely the GM estimator. It is observed from these plots that RME is robust and has a breakdown point of about 25%, that is, it can handle up to 25% of outliers among the data samples. While the GM estimator seems to perform better for small contamination, its robustness degrades with increasing dimensions, signifying a decreasing breakdown point with increasing number of parameters to be estimated, a well-known result in robust statistics literature (i.e. [1]). On the other hand, our RME exhibits a constant breakdown point regardless of the order of the AR model.

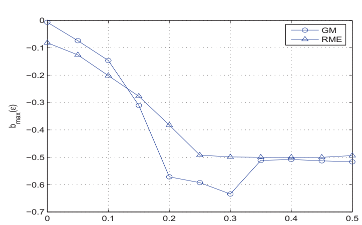

Fig. 3. Maximum bias curves of two robust estimators of an AR(1) , ϕ=0.5

2.3. Breakdown point of the RME

The breakdown point of 25% observed in simulations can be explained as follows. Because in our model, a single outlier y_p affects two terms of the product $y_{p-1}y_p$ and $y_p y_{p+1}$, the breakdown point of the RME is half that of the sample median, yielding BP=1/4.

2.4. Robust estimation of an ARMA using the RME

An ARMA(p, q) model are defined by $\phi_p(B)Y_t = \theta_q(B)\varepsilon_t$, where $\{\varepsilon_t\}$ is a sequence of iid Gaussian variables with variance σ_ε^2 and ϕ_p and θ_q are two polynomials of order p,q given by $\phi_p(B) = 1 - \phi_1 B - \ldots - \phi_p B^p$, $\theta_q(B) = 1 - \theta_1 B - \ldots - \theta_q B^q$, where B is the lag operator defined by $B^l Y_t = Y_{t-l}, l \in \mathbb{N}$.

To improve the efficiency of our method of estimation, the parameters of an ARMA(p,q) model are estimated via the following steps:

- Fit a high order AR(p^*) using the RME, where p^* is selected by a robust order selection criterion subject to being larger than the order of the autoregressive part p.
- Detect the outliers by filtering with the high order AR(p^*), reject them and use a classical maximum likelihood based estimation method of ARMA models with missing values [5].

The robust filtering is based on the state representation of an AR(p^*). The filter used is defined in [1] and based on the robust filter of Masreliez [6], which is termed the filter cleaner.

This filter adapts the outliers with their expected values from the other observations and the structure of the model. While at this stage, we can apply a maximum-likelihood estimator on the 'cleaned' series, we prefer to delete the outliers and apply a classical estimator with missing values [5]. An AR(p^*) of high order is described by the following state space representation

$$\begin{cases} X_t = \Phi X_{t-1} + D\varepsilon_t \\ y_t = G X_t \end{cases}, \quad \Phi = \begin{pmatrix} \phi_1 & \\ \vdots & I_{p^*-1} \\ \phi_{p^*} & 0'_{p^*-1} \end{pmatrix} \quad (6)$$

Here, Φ is the transition matrix, $D = (1, 0, \ldots, 0)'$, $G = (1, 0, \ldots, 0)$, I_k is the $k \times k$ identity matrix and 0_k the zero vector in $\mathbb{R}^k$; dim(Φ)=$k \times k$.

2.4.1. Robust filtering for outlier suppression

A filter can be defined as a mapping from $\mathbb{R}^n$ to $\mathbb{R}^{k\times n}$: $\widehat{X}^n = g^n(Y^n)$ where $\widehat{X}^n = (\widehat{X}_1, \ldots, \widehat{X}_n)$, $Y^n = (Y_1, \ldots, Y_n)$, $Y_1, \ldots, Y_n$ are the observed random variables and $\widehat{X}_1, \ldots, \widehat{X}_n$ are the filtered variables.

This filter will be called resistant if g^n is bounded and continuous. The Kalman filter is continuous and thereby robust against small rounding errors, but it is not bounded and consequently is not robust against gross errors. The proposed procedure, which makes use of the robust filter cleaner [1], consists of the following two steps.

Prediction :
$\widehat{X}_{t|t-1} = \Phi \widehat{X}_{t-1|t-1}$; $\widehat{\varepsilon}_t = Y_t - G\Phi\widehat{X}_{t-1|t-1}$
$P_{t|t-1} = \Phi P_{t-1|t-1}\Phi' + \sigma_\varepsilon^2 DD'$
Correction :
$\widehat{X}_{t|t} = \widehat{X}_{t|t-1} + \frac{1}{s_t} P_{t|t-1} G' \psi\left(\frac{\widehat{\varepsilon}_t}{s_t}\right)$

$P_{t|t} = P_{t|t-1} - \frac{1}{s_t^2} P_{t|t-1} G' w \left(\frac{\widehat{\varepsilon}_t}{s_t}\right) G P_{t|t-1}$
where ψ is a bounded continuous function. A popular choice consists in a redescending ψ function such as the Hampel or the bisquare psi function [1], where the weight function is defined as $w(x) = \frac{\psi(x)}{x}$ and s_t is a robust estimator of scale of the residual $\widehat{\varepsilon}_t$.

3. APPLICATION TO LOAD TIME SERIES FORECASTING

Forecasting load time series using SARIMA models is extensively used in the literature [7, 8]. This work is initiated by RTE, the transmission operator that manages and operates the French electric power transmission system, which is confronted to the presence of outliers in the French daily electric consumptions. RTE uses a SARIMA model in its daily forecasting. The resulting adjusted series exhibit a trend and several major cycles (daily, weekly, seasonal, yearly, etc.). Fig. (4) illustrates the load demand from Saturday July 2^{nd}, 2005 to Saturday July 23^{rd}, 2005. We notice that there is a break appearing on July 14^{th} and lasts until July 17^{th}, 2005 (approximately from observation 600 to 800 on figure (4). July 14^{th} is a public holiday in France. The load time series is first

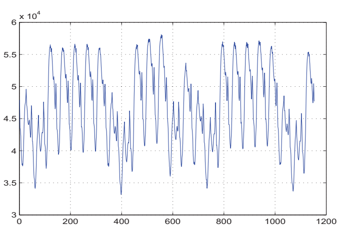

Fig. 4. Half-hourly electricity consumption on July 2-23,2005, France.

corrected from the influence of the weather by using a regression model where the exploratory variables are the temperature and the nebulosity. Then an ARIMA model is fitted to the residuals. A seasonal ARIMA model, SARIMA$(p, d, q) \times (p_1, d_1, q_1)_{s_1}$ follow the equation

$$\phi_p(B)\Phi_{p_1}(B^{s_1})\nabla^d \nabla_{s_1}^{d_1} Y_t = \theta_q(B)\Theta_{q_1}(B^{s_1})\varepsilon_t,$$

where Y_t is the electricity demand at time t, s_1 is the number of periods in the different seasonal cycles. B is the lag operator. ∇ is the difference operator, ∇_{s_1} is the seasonal difference operator $(B^l Y_t = Y_{t-l}, \nabla = 1 - B, \nabla_{s_1} = 1 - B^{s_1})$. $\phi_p,\ \Phi_{p_1},\ \theta_q,\ \Theta_{q_1}$ are polynomials of order p, p_1, q, q_1. ε_t is a Gaussian white noise from $N(0, \sigma_\varepsilon^2)$. On the daily series of a certain time (12:00), s_1 is equal to 7 to model the within-week seasonal cycle. In figure (5), we show the evolution of the mean absolute percentage error MAPE=$\frac{100}{h}\sum_{t=1}^{h}\left|\frac{Y_t - \widehat{Y}_t}{Y_t}\right|$, where h is the length of prediction. The MAPE was calculated for the classical ARIMA method and the previously defined approach. The mean absolute percentage error of our robust approach and the classical approach are denoted by MAPER and MAPE respectively in figure (5). The forecasting was done on "normal" days, which is natural since the goal of the method

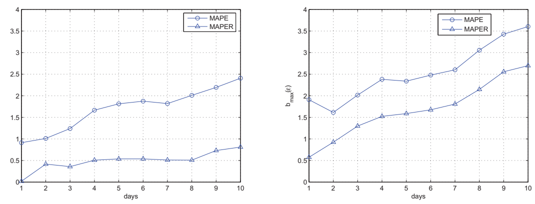

Fig. 5. MAPE forecast accuracy versus lead time for the series at 12:00, 20:00

is to forecast "normal" days with better precision. We remark that for all leading times MAPER $<$ MAPE. This means that the forecasting quality is improved with the new approach for these two hours of the day.

4. CONCLUSION

In this paper, a new robust RME estimation method for ARMA models is proposed and its influence function and maximum bias curve are derived. We compare the performance of our RME method to that of the classical approach based on maximum likelihood estimation in the French daily load forecasting. It is found that our RME method outperforms the current methods. Ongoing effort has been concentrating on computing the asymptotic distribution and variance of the sample median when it is applied to a correlated series. This analysis will allows us to derive confidence intervals for the autocorrelations and possibly for the parameters as well. Furthermore, a comparison of the RME with the median of slopes estimator, which has bias-optimality properties [1, Chapter 5], is worth investigating. Another research work will be to compare the performance of the RME to that of the τ-estimator and the MM-estimator in load forecasting.

5. REFERENCES

[1] Ricardo A. Maronna, R. Douglas Martin, and Victor J. Yohai, *Robust statistics*, Wiley Series in Probability and Statistics. John Wiley & Sons Ltd., Chichester, 2006, Theory and methods.

[2] Oscar H. Bustos and Víctor J. Yohai, "Robust estimates for ARMA models," *J. Amer. Statist. Assoc.*, vol. 81, no. 393, pp. 155–168, 1986.

[3] Milton Abramowitz and Irene A. Stegun, Eds., *Handbook of mathematical functions with formulas, graphs, and mathematical tables*, Dover Publications Inc., New York, 1992, Reprint of the 1972 edition.

[4] R. Douglas Martin and Víctor J. Yohai, "Influence functionals for time series," *Ann. Statist.*, vol. 14, no. 3, pp. 781–855, 1986, With discussion.

[5] Richard H. Jones, "Maximum likelihood fitting of ARMA models to time series with missing observations," *Technometrics*, vol. 22, no. 3, pp. 389–395, 1980.

[6] C. Johan Masreliez and R. Douglas Martin, "Robust Bayesian estimation for the linear model and robustifying the Kalman filter," *IEEE Trans. Automatic Control*, vol. AC-22, no. 3, pp. 361–371, 1977.

[7] Marcelo C. Medeiros and Lacir J. Soares, "Robust statistical methods for electricity load forecasting," in *RTE-VT workshop*, PARIS, May 2006.

[8] L. Mili, P. Panciatici, and Y. Chakhchoukh, "Robust short-term load forecasting using projection statistics," in *RTE-VT workshop*, PARIS, May 2006.