ACTA MATHEMATICA VIETNAMICA
Volume 34, Number 1, 2009, pp. 19–26

ON OUTLIER DETECTION IN MULTIVARIATE TIME SERIES

JEAN-MARIE HELBLING AND ROBERT CLÉROUX

Dedicated to Nguyen Van Hien on the occasion of his sixty-fifth birthday

Abstract. This paper deals with the detection of outliers in multivariate time series models. The different detection methods found in the literature are reviewed and a new one is suggested. It is based on the coefficient of vector autocorrelation. We obtain its influence function which is used in a heuristic for detecting outliers. Then the distribution of the influence function is obtained and used for testing the hypothesis of presence of outliers.

1. Introduction

The detection of outliers is an important problem in model building, inference and analysis of multivariate time series. Indeed, the presence of outliers, even in small quantity, can lead to biased estimation of the parameters, to a misspecification of the model and to inappropriate predictions. In the recent literature much importance has been given to this problem in the univariate context by the following authors: Fox [10], Abraham and Box [1], Kitagawa [16], Chernick, Downing and Pike [7], Martin [21], Yatawara [26], Abraham and Chuang [2], Abraham and Yatawara [3], Tsay [24], Lattin [17], Li and Hui [18] and Ljung [20]. Li and Hui [19] considered modeling multivariate time series via robust methods. Tsay, Peña and Pankratz [25] considered this problem in the multivariate framework. Then modern techniques have been applied to this problem like, for instance, projection pursuit by Galeano, Peña and Tsay [11] or independent component analysis by Baragona and Battaglia [5].

In this paper we consider the approach proposed by Tsay, Peña and Pankratz [25] and its advantages. We also consider the usual solutions proposed in the literature and suggest a new method of detecting outliers in multivariate time series.

2. Outliers in the multivariate time series model

Let $\mathbf{X}_t = (X_{1t}, \cdots, X_{rt})'$ be a r-dimensional vector representing a multivariate

Received March 31, 2008; in revised form September 16, 2008.
2000 Mathematics Subject Classification. 62H20, 62P20, 62E15, 62H99.
Key words and phrases. Outlier detection, multivariate time series, vector autocorrelation, influence function.

VARMA time series (vector autoregressive moving average)

$$\mathbf{\Phi}(B)\mathbf{X}_t = \mathbf{C} + \mathbf{\Theta}(B)\mathbf{E}_t, \quad \text{with} \quad t = 1, \cdots, n, \tag{2.1}$$

where B is the $r \times r$ matrix backshift operator such that $B\mathbf{X}_t = \mathbf{X}_{t-1}$ and

$$\mathbf{\Phi}(B) = \mathbf{I} - \mathbf{\Phi}_1 B - \cdots - \mathbf{\Phi}_p B^p$$

$$\mathbf{\Theta}(B) = \mathbf{I} - \mathbf{\Theta}_1 B - \cdots - \mathbf{\Theta}_q B^q$$

are matrix polynomials of orders p and q respectively, $\mathbf{C}$ is a r-dimensional constant vector and $\mathbf{E}_t = (E_{1t}, \cdots, E_{rt})'$ is a sequence of independent white noise vectors of means $\mathbf{0}$ and covariance matrices $\mathbf{\Sigma}$.
If $\mathbf{X}_t$ is invertible, model (2.1) above can be written as

$$\mathbf{\Pi}(B)\mathbf{X}_t = \mathbf{C}_{\mathbf{\Pi}} + \mathbf{E}_t,$$

where $\mathbf{\Pi}(B) = \mathbf{\Theta}(B)^{-1}\mathbf{\Phi}(B) = \mathbf{I} - \sum_{i=1}^{\infty} \mathbf{\Pi}_i B^i$ and $\mathbf{C}_{\mathbf{\Pi}} = \mathbf{\Theta}(1)^{-1}\mathbf{C}$. Similarly (2.1) can be written as:

$$\mathbf{X}_t = \mathbf{C}_{\mathbf{\Psi}} + \mathbf{\Psi}(B)\mathbf{E}_t$$

where $\mathbf{\Phi}(1)\mathbf{C}_{\mathbf{\Psi}} = \mathbf{C}$ and $\mathbf{\Phi}(B)\mathbf{\Psi}(B) = \mathbf{\Theta}(B)$ with $\mathbf{\Psi}(B) = \mathbf{I} + \sum_{i=1}^{\infty} \mathbf{\Psi}_i B^i$.

Given an observed multivariate time series $\mathbf{Y} = (\mathbf{Y}_1', \cdots, \mathbf{Y}_n')'$ with $\mathbf{Y}_t = (Y_{1t}, \cdots, Y_{rt})'$, the presence of outliers can be modelized as:

$$\mathbf{Y}_t = \mathbf{X}_t + \boldsymbol{\alpha}(B)\mathbf{w}I_t^{(h)}, \tag{2.2}$$

where $I_t^{(h)}$ is an indicator variable characterizing the outlier at time h, that is $I_h^{(h)} = 1$ and $I_t^{(h)} = 0$ if $t \neq h$, and $\mathbf{w} = (w_1, \cdots, w_r)'$ is its impact on the series and $\mathbf{X}_t$ follows (2.1).

The outliers are then classified in four categories:

- The multivariate additive outliers (MAO) affect only one observation of the series and not the future values.
  In terms of polynomials the MAO are modelized by letting $\boldsymbol{\alpha}(B) = \mathbf{I}$ in (2.2).
- The multivariate innovational outliers (MIO) have a temporary impact on the series like an innovation.
  Then we let $\boldsymbol{\alpha}(B) = \mathbf{\Psi}(B)$ in (2.2).
- The multivariate level shift (MLS) increase or decrease all the observations at a given point of the series by a constant.
  In this case we have $\boldsymbol{\alpha}(B) = (1 - B)^{-1}\mathbf{I}$ in (2.2).
- The multivariate temporary change (MTC) increase or decrease drastically the level of the series which rapidly returns to its initial level exponentially.
  Then we have $\boldsymbol{\alpha}(B) = (\mathbf{I} - \delta\mathbf{I}B)^{-1}$ with $0 < \delta < 1$ in (2.2).

In general, the MAO and the MIO are considered as non-typical observations whereas the MTC and the MLS as structural changes.
The problem of outlier detection in multivariate time series is a complex problem because the different components of $\mathbf{X}_t$ can be affected by different types of outliers.

3. Outliers with respect to vector autocorrelation

Let $\mathbf{X}_t = (X_{1t}, X_{2t}, \cdots, X_{rt})'$ be a multivariate stationary time series of dimension r with

$$\begin{aligned} E(\mathbf{X}_t) &= 0, \\ E(\mathbf{X}_t\mathbf{X}'_{t+k}) &= \mathbf{\Gamma}(k) = (\gamma_{ij}(k)) \text{ a } r \times r \text{ matrix and} \\ E(\mathbf{X}_t\mathbf{X}'_t) &= E(\mathbf{X}_{t+k}\mathbf{X}'_{t+k}) = \mathbf{\Gamma}(0) = (\gamma_{ij}(0)) \text{ another } r \times r \text{ matrix.} \end{aligned}$$

We know that $\mathbf{\Gamma}(0)$ is symmetric but not $\mathbf{\Gamma}(k)$ for $k \neq 0$.

Consider the vector of dimension $2r$, $\mathbf{X} = \begin{pmatrix} \mathbf{X}_t \\ \mathbf{X}_{t+k} \end{pmatrix}$ with the $2r \times 2r$ covariance matrix $\mathbf{\Gamma} = \begin{pmatrix} \mathbf{\Gamma}(0) & \mathbf{\Gamma}(k) \\ \mathbf{\Gamma}'(k) & \mathbf{\Gamma}(0) \end{pmatrix}$.

Then the coefficient of vector autocorrelation is defined by (see Roy and Cléroux [23] for example)

$$\lambda(k) = \frac{tr\mathbf{\Gamma}(k)\mathbf{\Gamma}'(k)}{tr\mathbf{\Gamma}^2(0)}, \tag{3.1}$$

where $tr(.)$ is the trace operator.

In some situations it may be convenient to let $\mathbf{Y}_t = \mathbf{D}^{-1/2}\mathbf{X}_t$ for each t where $\mathbf{D} = \text{diag}(\gamma_{ii}(0))$. Then the covariance matrix of $\mathbf{Y}_t$ is the correlation matrix of $\mathbf{X}_t$ and the corresponding $\lambda(k)$ is obtained from correlation matrices. The influence function of a point $\mathbf{Y}$ or of a point $\mathbf{X}$ would be obtained in a similar way.

Let F be a distribution function and $\tilde{F} = (1-\epsilon)F + \epsilon\delta_{\mathbf{X}}$ be a perturbation of F by $\delta_{\mathbf{X}}$, the distribution function which assigns unit probability to the point $\mathbf{X}$. Let $\theta = T(F)$ be any parameter expressed as a functional of the distribution function F and let $\tilde{\theta} = T(\tilde{F})$. Hampel [12] defined the theoretical influence function $I(\mathbf{X};\theta)$ at $\mathbf{X}$ as

$$I(\mathbf{X};\theta) = \lim_{\epsilon \to 0} \left(\frac{\tilde{\theta} - \theta}{\epsilon}\right). \tag{3.2}$$

The theoretical influence function can be used to determine the influence of the point $\mathbf{X}$ on the parameter θ. Influence functions are important tools for outlier detection. They have been used by Devlin, Gnanadesikan and Kettenring [9] in the context of bivariate correlation, by Campbell [6] for outlier detection in discriminant analysis, by Cléroux, Helbling and Ranger [8] for outlier detection in multivariate data and in the multivariate linear regression model, and by Chernick, Downing and Pike [7] and Li and Hui [18] for outlier detection in univariate time series data. The concept of influential functional, similar in spirit to Hampel's influence function, has been introduced by Martin and Yohai [22] for univariate time series.

Using an approach similar to that in Cléroux, Helbling and Ranger [8] we obtain, after lengthy algebraic manipulations, the theoretical influence function of $\lambda(k)$:

$$I(\mathbf{X};\lambda(k)) = \lambda(k)\Big(\frac{2\mathbf{X}_t'\mathbf{\Gamma}(k)\mathbf{X}_{t+k}}{tr\mathbf{\Gamma}(k)\mathbf{\Gamma}'(k)} - \frac{\mathbf{X}_t'\mathbf{\Gamma}(0)\mathbf{X}_t + \mathbf{X}_{t+k}'\mathbf{\Gamma}(0)\mathbf{X}_{t+k}}{tr\mathbf{\Gamma}^2(0)}\Big). \tag{3.3}$$

It can be seen that $E[I(\mathbf{X};\lambda(k)] = 0$ and if $\mathbf{X}$ has the multivariate normal distribution, then $Var[I(\mathbf{X};\lambda(k)] = 4\lambda^2(k)\sigma^2$, where

$$\begin{aligned}\sigma^2 &= 1 + \frac{tr\mathbf{\Gamma}^4(0)}{(tr\mathbf{\Gamma}^2(0))^2} + \frac{tr\mathbf{\Gamma}(0)\mathbf{\Gamma}(k)\mathbf{\Gamma}(0)\mathbf{\Gamma}'(k)}{(tr\mathbf{\Gamma}^2(0))^2} + \frac{tr\mathbf{\Gamma}(0)\mathbf{\Gamma}(k)\mathbf{\Gamma}(0)\mathbf{\Gamma}'(k)}{(tr\mathbf{\Gamma}(k)\mathbf{\Gamma}'(k))^2}\\ &\quad - \frac{2tr\mathbf{\Gamma}^2(0)\mathbf{\Gamma}(k)\mathbf{\Gamma}'(k)}{tr\mathbf{\Gamma}^2(0)tr\mathbf{\Gamma}(k)\mathbf{\Gamma}'(k)} - \frac{2tr\mathbf{\Gamma}^2(0)\mathbf{\Gamma}'(k)\mathbf{\Gamma}(k)}{tr\mathbf{\Gamma}^2(0)tr\mathbf{\Gamma}(k)\mathbf{\Gamma}'(k)}.\end{aligned} \tag{3.4}$$

Let us note in passing that

i) if we let $\hat{\lambda}(k) = \dfrac{tr\mathbf{C}(k)\mathbf{C}'(k)}{tr\mathbf{C}^2(0)}$ where $\mathbf{C}(0)$ and $\mathbf{C}(k)$ are the usual estimates, from a sample of size n, of $\mathbf{\Gamma}(0)$ and $\mathbf{\Gamma}(k)$ respectively, then the asymptotic distribution, as $n \longrightarrow \infty$, of $\sqrt{n}[\hat{\lambda}(k) - \lambda(k)]$ is $\mathrm{N}(0, 4\lambda^2(k)\sigma^2)$ where σ^2 is given by (3.4), when $0 < \lambda(k) < 1$ (see for example, Andrews et al. [4], pp. 29-30).

ii) if $r = 1$ then $\lambda(k) = \rho^2(k)$ where $\rho(k)$ is the autocorrelation coefficient of lag k in a univariate time series. From (3.3) we obtain

$$I(X;\rho(k)) = Z_t Z_{t+k} - \frac{\rho(k)}{2}(Z_t^2 + Z_{t+k}^2), \tag{3.5}$$

where $Z_t = \dfrac{X_t}{\sqrt{\gamma(0)}}$ and $Z_{t+k} = \dfrac{X_{t+k}}{\sqrt{\gamma(0)}}$. Formula (3.5) is identical to that obtained by Chernick, Downing and Pike [7].

iii) formulas (3.3) and (3.4) correspond to similar formulas obtained in Cléroux, Helbling and Ranger [8] in a different context.

4. Methods of outlier detection

From the model of Section 2 we can consider a first method of detecting outliers. This method has been developed by Tsay, Peña and Pankratz [25] and reformulated by Galeano, Peña and Tsay [11]. Here we mention the latter. It is based on the following ideas. If all the parameters of model (2.1) for $\mathbf{X}_t$ are known we obtain, from the observed series $\mathbf{Y}_t$, for $t < h$, the innovations

$$\mathbf{A}_t = \mathbf{\Pi}(B)\mathbf{Y}_t - \mathbf{C}_{\mathbf{\Pi}}. \tag{4.1}$$

Then the relation between the pure white noise $\mathbf{E}_t$ and the computed innovations $\mathbf{A}_t$ is

$$\mathbf{A}_t = \mathbf{E}_t + \boldsymbol{\Gamma}(B)\mathbf{w}I_t^{(h)}, \tag{4.2}$$

where $\boldsymbol{\Gamma}(B) = \boldsymbol{\Pi}(B)\boldsymbol{\alpha}(B) = \mathbf{I} - \sum_{i=1}^{\infty} \boldsymbol{\Gamma}_i B^i$.
It is shown by Tsay, Peña and Pankratz [25] that the estimation of the impact $\mathbf{w}$ of the outlier of type i (i = MAO, MIO, MLS, MTC) is the following when the parameters of the model are known:

$$\mathbf{w}_{i,h} = -(\sum_{j=0}^{n-h} \boldsymbol{\Gamma}'_j \boldsymbol{\Sigma}^{-1} \boldsymbol{\Gamma}_j)^{-1} (\sum_{j=0}^{n-h} \boldsymbol{\Gamma}'_j \boldsymbol{\Sigma}^{-1} \mathbf{A}_{h+j}), \tag{4.3}$$

where $\boldsymbol{\Gamma}_0 = -\mathbf{I}$.

The covariance matrix of this estimator is $\boldsymbol{\Sigma}_{i,h} = (\sum_{j=0}^{n-h} \boldsymbol{\Gamma}'_j \boldsymbol{\Sigma}^{-1} \boldsymbol{\Gamma}_j)^{-1}$. Under the hypothesis $H_0 : \mathbf{w} = 0$ the statistics

$$J_{i,h} = \mathbf{w}'_{i,h} \boldsymbol{\Sigma}_{i,h}^{-1} \mathbf{w}_{i,h}, \quad i = \text{MAO, MIO, MLS, MTC} \tag{4.4}$$

has the chi-squared distribution with r degrees of freedom. Another statistics proposed by the same authors is

$$C_{i,h} = \max\{|w_{j,i,h}|/\sqrt{\sigma_{j,i,h}} : 1 \leq j \leq r\}, \quad i = \text{MAO, MIO, MLS, MTC}$$

where $w_{j,i,h}$ is the j^{th} element of $\mathbf{w}_{i,h}$ and $\sigma_{j,i,h}$ is the j^{th} element of the main diagonal of $\boldsymbol{\Sigma}_{i,h}$.
Clearly, in practical situations, neither the parameters nor the time h, where the outliers occur, are known. Thus the parameters are replaced by their estimators and the time h to be used is obtained by taking a maximum:

$$J_{\max}(i, h_i) = \max_{1 \leq h \leq n} J_{i,h} \quad \text{and} \quad C_{\max}(i, h_i^*) = \max_{1 \leq h \leq n} C_{i,h}$$

for each type of outliers i = MAO, MIO, MLS, MTC.

The detection of outliers can also be made using projection pursuit. Galeano, Peña and Tsay [11] use this technique to find a linear function of the multivariate time series which maximizes the coefficient of kurtosis in order to find the best univariate representation of the multivariate signal. Afterwards the detection of the times where outliers occur and the estimation of their amplitudes are made using univariate methods.

The detection of outliers in the same framework can also be made by independent component analysis. This method consists of extracting hidden components (latent variables) of multivariate data under the sole hypothesis that the unknown components are mutually independent. Clearly the distribution of these independent components is assumed non gaussian in order to be in interest. It is the case if outliers are present. The presence of an outlier and the time when it occurs is revealed by observing the first few independent components. The search of the independent components is made using algorithms associated with neural networks (Hyvärinen and Oja [13]) and the application to outlier detection in multivariate time series is found in Baragona and Battaglia [5].

We now propose a new method of detecting outliers in multivariate time series based on the coefficient of vector autocorrelation and its influence function obtained in Section 3.
In practice we can replace $\mathbf{\Gamma}(0)$ and $\mathbf{\Gamma}(k)$ in (3.3) by their usual estimators $\mathbf{C}(0)$ and $\mathbf{C}(k)$ respectively to obtain $\hat{I}(\mathbf{x};\lambda(k))$ and then compute the influence of each data point to pinpoint possible outliers. A simple heuristic would be to consider $\mathbf{x}_i$ as being an outlier if its estimated influence is, in absolute value, greater than $6\hat{\lambda}(k)\hat{\sigma}$, which is three times the estimated standard deviation of $I(\mathbf{X};\lambda(k))$, where $\hat{\sigma}$ is obtained from (3.4) by replacing $\mathbf{\Gamma}(0)$ and $\mathbf{\Gamma}(k)$ by $\mathbf{C}(0)$ and $\mathbf{C}(k)$ respectively.
We can also obtain the exact distribution of the theoretical influence function and propose an approximate test for outliers.

Formula (3.3) can be written as a quadratic form $I(\mathbf{X};\lambda(k)) = \mathbf{X}'\mathbf{A}\mathbf{X}$, where $\mathbf{X}' = (\mathbf{X}'_t, \mathbf{X}'_{t+k})$ and

$$\mathbf{A} = \lambda(k) \begin{pmatrix} \dfrac{-\mathbf{\Gamma}(0)}{tr\mathbf{\Gamma}^2(0)} & \dfrac{\mathbf{\Gamma}(k)}{tr\mathbf{\Gamma}(k)\mathbf{\Gamma}'(k)} \\ \dfrac{\mathbf{\Gamma}'(k)}{tr\mathbf{\Gamma}(k)\mathbf{\Gamma}'(k)} & \dfrac{-\mathbf{\Gamma}(0)}{tr\mathbf{\Gamma}^2(0)} \end{pmatrix}. \tag{4.5}$$

Now if $\mathbf{X}$ has the multivariate normal distribution, the distribution of $I(\mathbf{X};\lambda(k))$ is that of $\sum_{i=1}^{2r}\delta_i W_i^2$ where the W_i's are independent identically distributed $\mathrm{N}(0,1)$ random variables and where the δ_i's are the eigenvalues of

$$\mathbf{\Gamma}\mathbf{A} = \lambda(k) \begin{pmatrix} \dfrac{-\mathbf{\Gamma}^2(0)}{tr\mathbf{\Gamma}^2(0)} + \dfrac{\mathbf{\Gamma}(k)\mathbf{\Gamma}'(k)}{tr\mathbf{\Gamma}(k)\mathbf{\Gamma}'(k)} & \dfrac{\mathbf{\Gamma}(0)\mathbf{\Gamma}(k)}{tr\mathbf{\Gamma}(k)\mathbf{\Gamma}'(k)} - \dfrac{\mathbf{\Gamma}(k)\mathbf{\Gamma}(0)}{tr\mathbf{\Gamma}^2(0)} \\ \dfrac{\mathbf{\Gamma}'(k)\mathbf{\Gamma}(0)}{tr\mathbf{\Gamma}^2(0)} + \dfrac{tr\mathbf{\Gamma}(0)\mathbf{\Gamma}'(k)}{tr\mathbf{\Gamma}(k)\mathbf{\Gamma}'(k)} & \dfrac{\mathbf{\Gamma}'(k)\mathbf{\Gamma}(k)}{tr\mathbf{\Gamma}(k)\mathbf{\Gamma}'(k)} - \dfrac{-\mathbf{\Gamma}^2(0)}{tr\mathbf{\Gamma}^2(0)} \end{pmatrix} \tag{4.6}$$

(see for example Johnson and Kotz [15], p. 150). The percentiles of the distribution of $I(\mathbf{X};\lambda(k))$ can be computed using the Imhof [14] algorithm.

However, the criterion for testing that an extreme observation is an outlier is based on the extreme values of the influence function. The distribution of

$$I_{(1)} = \min_{1\le i\le n} I(\mathbf{X}_i;\lambda(k)) \quad \text{and of} \quad I_{(n)} = \max_{1\le i\le n} I(\mathbf{X}_i;\lambda(k))$$

are thus needed. $I_{(1)}$ will usually be negative and $I_{(n)}$ will usually be positive. Since $I(\mathbf{X}_i;\lambda(k))$, $i = 1, 2, \cdots, n$ are independent it follows that the distribution of $I_{(n)}$ is $G(x) = [H(x)]^n$ and that of $I_{(1)}$ is $L(x) = 1 - [1 - H(x)]^n$ where $H(x)$ is the distribution of $I(\mathbf{X};\lambda(k))$. A procedure for testing extreme observations as significant outliers is the following:

i) compute the eigenvalues $\hat{\delta}_i, i = 1, 2, \cdots, 2r$ of $\hat{\mathbf{\Gamma}}\hat{\mathbf{A}}$ where $\hat{\mathbf{\Gamma}}$ and $\hat{\mathbf{A}}$ are obtained from $\mathbf{\Gamma}$ and $\mathbf{A}$ by replacing $\mathbf{\Gamma}(0)$ and $\mathbf{\Gamma}(k)$ by $\mathbf{C}(0)$ and $\mathbf{C}(k)$ respectively.

ii) obtain the data point $\mathbf{x}_j$ such that $\hat{I}(\mathbf{x}_j; \lambda(k))$ has the largest positive value. Use the distribution $G(x)$ above (with $\hat{\delta}_i$ in place of δ_i for all i) to find the probability p_1 of exceeding this value. If $p_1 < \alpha_1$ then $\mathbf{x}_j$ can be considered as being an outlier at level α_1.

iii) obtain the data point $\mathbf{x}_l$ such that $\hat{I}(\mathbf{x}_l; \lambda(k))$ has the smallest negative value. Use the distribution $L(x)$ above (with $\hat{\delta}_i$ in place of δ_i for all i) to find the probability p_2 of not exceeding this value. If $p_2 < \alpha_2$ then $\mathbf{x}_l$ can be considered as being an outlier at level α_2.

5. Conclusion

In this paper we considered various methods for detecting outliers in multivariate time series. Some methods are based on tests of hypotheses and others are based on projection pursuit and independent component analysis. We introduced the coefficient of vector autocorrelation, obtained its influence function together with its distribution. We also proposed new methods of detecting outliers in the multivariate time series model, a heuristic method based on the graph of the influence function and another consisting of testing for the presence of outliers. All the methods considered in this paper have been seen from a theoretical point of view. Numerical comparisons would be interesting and remain to be done.

Acknowledgement

The authors wish to thank the referee and the editor for their helpful comments.

References

[1] B. Abraham and G. E. P. Box, Bayesian analysis of some outlier problems in time series, *Biometrika* **66** (1979), 229–236.

[2] B. Abraham and A. Chuang, Outlier detection and time series modeling, *Technometrics* **31** (1989), 241–248.

[3] B. Abraham and N. Yatawara, A score test for detection of time series outliers, *J. Time Series Anal.* **9** (1988), 109–119.

[4] D. Andrews, P. J. Bickel, F. R. Hampel, P. J. Huber, W. H. Rogers and J. W. Tukey, *Robust Estimates of Location*, Princeton University Press, 1972.

[5] R. Baragona and F. Battaglia, Outliers detection in multivariate time series by independent component analysis, *Neural Computation* **19** (2007), 1962–1984.

[6] N. A. Campbell, The influence function as an aid in outlier detection in discriminant analysis, *Applied Statistics* **27** (1978), 251–258.

[7] M. R. Chernick, D. J. Downing, and D. H. Pike, Detecting outliers in time series data, *J. Ameri. Statist. Ass.* **77** (1982), 743–747.

[8] R. Cléroux, J.-M. Helbling, and N. Ranger, Some methods of detecting multivariate outliers, *Comput. Statistics Quarterly* **3** (1986), 177–195.

[9] S. J. Devlin, R. Gnanadesikan and J. R. Kettenring, Robust estimation and outlier detection with correlation coefficients, *Biometrika* **62** (1975), 531–545.

[10] A. J. Fox, Outliers in time series, *J. Royal Statist. Soc. Ser. B* **34** (1972), 350–363.

[11] P. Galeano, D. Peña, and R. S. Tsay, Outlier detection in multivariate time series by projection pursuit, *J. Amer. Statist. Ass.* **101** (2006), 654–669.

[12] F. R. Hampel, The influence curve and its role in robust estimation, *J. Amer. Statist. Ass.* **69** (1974), 383–393.

[13] A. Hyvärinen and E. Oja, Independent component analysis: algorithms and applications, *Neural Networks* **13** (2000), 411-430.

[14] J. P. Imhof, Computing the distribution of quadratic forms in normal variables, *Biometrika* **48** (1961), 419–426.

[15] N. L. Johnson and S. Kotz, *Continuous Univariate Distributions*, Vol. 2, Houghton Mifflin Company, Boston, 1970.

[16] G. Kitagawa, On the use of aic for the detection of outliers, *Technometrics* **21** (1979), 193-199.

[17] J. M. Lattin, Identifying influential observations in time series data, in: *Computer Science and Statistics*, Academic Press, New York, 1987, pp.21-53.

[18] W. K. Li and Y. V. Hui, On the empirical influence functions of residual autocorrelations in time series, in: *Proc. Bus. Econ. Statist. Sect., American Statistical Association* 1987, pp. 465–468.

[19] W. K. Li and Y. V. Hui, Robust multiple time series modelling, *Biometrika* **76** (1989), 309–315.

[20] G. M. Ljung, On outlier detection in time series, *J. Royal Statist. Soc. Ser. B* **55** (1993), 559–567.

[21] R. D. Martin, *Time Series Outlier Detection and Model Identification with Robust Filter and Smoother Cleaners*, Academic Press, London, 1983.

[22] R. D. Martin and V. J. Yohai, Influence functionals for time series, *Ann. Statist.* **14** (1986), 781-818.

[23] R. Roy and R. Cléroux, Vector cross-correlation in time series and applications, *Internat. Statist. Rev.* **61** (1993), 447–464.

[24] R. S. Tsay, Time series model specification in the presence of outliers, *J. Amer. Statist. Ass.* **81** (1986), 132–141.

[25] R. S. Tsay, D. Peña, and A.E. Pankratz, Outliers in multivariate time series, *Biometrika* **87** (2000), 789–804.

[26] N. Yatawara, *Detection of abrupt system changes with application to outlier analysis in time series*, Unpublished manuscript, Department of Statistics and Actuarial Science, University of Waterloo, 1985.

Institute of Mathematics, Station 8
Swiss Federal Institute of Technology
1015 Lausanne, Switzerland

E-mail address: jean-marie.helbling@epfl.ch

Département de Mathématiques et de Statistique
Université de Montréal, CP 6128, Succ. Centre-Ville
Montréal, Canada H3C3J7

E-mail address: cleroux@dms.umontreal.ca