Technometrics

ISSN: 0040-1706 (Print) 1537-2723 (Online) Journal homepage: https://www.tandfonline.com/loi/utch20

Two Projection Methods for Use in the Analysis of Multivariate Process Data With an Illustration in Petrochemical Production

Julie Badcock, Trevor C Bailey, Philip Jonathan & Wojtek J Krzanowski

To cite this article: Julie Badcock, Trevor C Bailey, Philip Jonathan & Wojtek J Krzanowski (2004) Two Projection Methods for Use in the Analysis of Multivariate Process Data With an Illustration in Petrochemical Production, Technometrics, 46:4, 392-403, DOI: 10.1198/004017004000000491

To link to this article: https://doi.org/10.1198/004017004000000491

Published online: 01 Jan 2012.

Two Projection Methods for Use in the Analysis of Multivariate Process Data With an Illustration in Petrochemical Production

Julie Badcock

Atass Ltd.
Cullompton, U.K. EX15 2NW

Trevor C. Bailey

Department of Mathematical Sciences
University of Exeter
Exeter, U.K. EX4 4QE
(*T.C.Bailey@exeter.ac.uk*)

Philip Jonathan

Shell Research Ltd.
Chester, U.K. CH1 3SH

Wojtek J. Krzanowski

Department of Mathematical Sciences
University of Exeter
Exeter, U.K. EX4 4QE

Principal components analysis (PCA) is often used in the analysis of multivariate process data to identify important combinations of the original variables on which to focus for more detailed study. However, PCA and other related projection techniques from the standard multivariate repertoire are not explicitly designed to address or to exploit the strong autocorrelation and temporal cross-correlation structures that are often present in multivariate process data. Here we propose two alternative projection techniques that do focus on the temporal structure in such data and that therefore produce components that may have some analytical advantages over those resulting from more conventional multivariate methods. As in PCA, both of our suggested methods linearly transform the original p-variate time series into uncorrelated components; however, unlike PCA, they concentrate on deriving components with particular temporal correlation properties, rather than those with maximal variance. The first technique finds components that exhibit distinctly different autocorrelation structures via modification of a signal-noise decomposition method used in image analysis. The second method draws on ideas from common PCA to produce components that are not only uncorrelated as in PCA, but that also have approximately zero temporally lagged cross-correlations for all time lags. We present the technical details for these two methods, assess their performance through simulation studies, and illustrate their use on multivariate output measures from a fluidized catalytic cracking unit used in petrochemical production, contrasting the results obtained with those from standard PCA.

KEY WORDS: Multivariate process control; Principal components analysis; Temporally structured components; Temporally uncorrelated components.

1. INTRODUCTION

There are numerous well-established, extensively investigated, and widely used methods for the analysis of univariate process data (see, e.g., Wetherill and Brown 1991; Bissel 1994). In contrast, the techniques available for the analysis of multivariate process measurements are less well developed (Champ, Tracy, and Young 1997). This is unfortunate, because multivariate data are commonly encountered in many practical process control situations, and thus there is an important requirement to develop and extend the existing repertoire of methods applicable to such data.

In general, multivariate process data may comprise variables measured on categorical, discrete, or interval-ratio scales, and possibly a mixture of these, but throughout this article we focus on multivariate measurements on an interval-ratio scale. Many datasets arising in industrial process control are of that nature, and we use a typical example from petrochemical manufacturing to illustrate the methods that we propose here. In addition to correlation between variables, a particular feature of such process data is the presence of strong temporal (serial) correlation both within and between variables. Exploitation of this temporal structure forms the basis for the new methods that we suggest.

Various approaches have already been proposed for analyzing multivariate process data depending on whether the objective is description, control, or prediction. Methods include: different forms of multivariate control charts (Alt and Smith 1988; Fuchs and Benjamini 1994; Lowry and Montgomery 1995; Lui 1995; Runger, Alt, and Montgomery 1996; Aparisi, Jabaloyes, and Carrión 1999), the incorporation of time series models in control charts (VanBrackle and Reynolds 1997; Jiang, Tsui, and Woodall 2000), use of principal components analysis (PCA) and modifications thereof (Jackson 1985; Runger and Alt 1996; Wilkström et al. 1998a), and dynamic biplots (Kulkarni and Paranjape 1984; Sparks, Adolphson, and Phatak 1997). Dynamic factor models (Molenaar 1985) are another possible approach in which latent variable modeling is used to systematically extract information from the multivariate time series, although to date, applications of this method have been in econometrics, psychology, or environmetrics rather than in industrial process control.

TECHNOMETRICS, NOVEMBER 2004, VOL. 46, NO. 4
DOI 10.1198/004017004000000491

Of this previous work, the most pertinent to the discussion here is that using PCA or related multivariate projection techniques to decompose the raw data into uncorrelated components for subsequent analysis. This is a common approach in the analysis of multivariate process data, at least at a preliminary stage. However, although undoubtedly useful, the basic forms of such analyses are not oriented specifically to autocorrelated data or to the explicit decomposition of temporal structure in the multivariate process. The need to allow for the strong temporal correlations that are often present in process data has long been recognized in univariate process control; for example, Faltin, Mastrangelo, Runger, and Ryan (1997) and Harris and Ross (1991) have considered how to adapt univariate exponentially weighted moving average (EWMA) and cumulative sum (CUSUM) charts to allow for an autocorrelated process. More recently, modifications to PCA for multivariate temporally correlated processes have been suggested by, for example, Wilkström et al. (1998b), Wachs and Lewin (1999), Ku, Storer, and Georgakis (2001). In this article we propose two additional multivariate methods that may be useful in analyzing temporally correlated multivariate process data. As in PCA, both methods linearly transform the original p-variate time series, $\mathbf{y}(t) = y_1(t), \ldots, y_p(t)$, into uncorrelated components, $\mathbf{u}(t) = u_1(t), \ldots, u_p(t)$. However, rather than focusing on components that explain overall variance in the original data, our methods instead produce components with particular temporal autocorrelation or temporal cross-correlation properties. Such components may provide some analytical advantages over those resulting from either standard PCA or from previously suggested PCA modifications designed to cater for temporal structure. We illustrate that by comparing results from our methods with those from PCA using both simulated data with known structure and real data obtained from a petrochemical production process.

The remainder of the article is organized as follows. In Section 2 we introduce and describe a multivariate and serially correlated dataset obtained from a fluidized catalytic cracking unit (FCC) used in petrochemical production and report the results of standard PCA on these data to aid comparisons in subsequent sections of the article. In Section 3 we develop the first of our suggested alternative methods, reporting on simulations to establish its efficacy and comparing results for the FCC data with those reported using PCA in Section 2. Our idea here is to find "temporally structured" as opposed to "variance maximizing" components, that is, components that exhibit distinct differences in the persistence of autocorrelation. Standard PCA will not necessarily achieve this aim. The advantage of doing this is that components of this type enable the analyst to characterize and separately study short-, intermediate-, or long-term behavior in the multivariate process and capitalize on this knowledge in further analysis. For example, components with autocorrelation extending over only short time lags may be identified with the more transient aspects of process behavior (in the limit this might be considered to be "noise" in the data), whereas those with autocorrelation of a longer duration will relate to long- or medium-term influences, which are perhaps more important in process control. Such components may thus allow one to capture a fuller picture of the process dynamics than might otherwise be achieved.

In Section 4 we outline the second of our transformation methods, again substantiating its performance on simulated data and then comparing results for the FCC data with those from PCA obtained in Section 2. Here our aim is to remove temporal cross-correlation between the components that we derive. Standard PCA produces components, $u_i(t)$, $i = 1, \ldots, p$, that are instantaneously uncorrelated [i.e., $\operatorname{corr}(u_i(t), u_j(t)) = 0$ for $i \neq j$], but pairs of these components are not necessarily "temporally uncorrelated" [i.e., it is not the case that for all time lags δ, $\operatorname{corr}(u_i(t), u_j(t - \delta)) = 0$, $i \neq j$]. The components obtained from our approach approximately satisfy this second condition, which renders such components potentially more appropriate than ordinary principal components if separate univariate process control techniques are subsequently to be applied to each component. For example, if one such component signals an "out-of-control" point, it does not trigger a signal for future points in other components, as would be the case for temporally cross-correlated components, such as those that typically result from PCA. Independent process control analyses of each component thus are not compromised by any temporal dependency between them. Finally, in Section 5 we provide a brief summary and discussion.

2. FLOW COMPOSITION DATA FROM A CONTINUOUS FLUIDIZED CATALYTIC CRACKING UNIT

The oil industry is under constant pressure to improve the performance of its manufacturing facilities. Large-scale petrochemical processes typically involve the separation and purification of mixtures of hydrocarbons and their conversion to chemicals with high economic value. Statistical interrogation of historical plant operating data has an important contribution to make toward improving process understanding and optimizing performance (Gray and Handwerk 1994). A modern refinery or petrochemicals plant consists of a complex network of interlinked processes. Processing is conducted in a sequence of stages, with the output from one processing stage becoming the input for the next. Most major manufacturing units are equipped with supervisory computer systems that provide automatic control and allow continuous accumulation of operational data that characterize process behavior. General features of data samples drawn from continuous manufacturing processes are that they are multivariate and both autocorrelated and cross-correlated in time.

The FCC process is one of the most important in a refinery, converting heavy crude oil fractions into more valuable gasoline and lighter products via catalytic cracking reactions (Sadeghbeigi 1995). FCC throughput is measured in thousands of (metric) tons per day, and hence very small improvements in performance yield substantial monetary benefits. We focus here on a subset of output measures collected from the Shell U.K. Stanlow FCC unit. Data collected from this process have been described in more detail by Jonathan and Kaskavelis (1998) and are broadly typical of the kind of multivariate process data obtained from many industrial manufacturing operations. The particular subset that we consider consists of daily measurements of the percentage concentrations of just two chemicals in one of the major FCC output streams. These chemicals, denoted here simply as A and B, were selected because they represent two of the more valuable components in the process. Each chemical was recorded daily at three different points of the process over a total of 665 days, giving rise to a multivariate time series on six variables denoted A1, A2, A3, B1, B2, and B3.

The measured concentration of chemicals A and B can be expected to vary over time for many reasons. First, there is variation in chemical composition of feed to the FCC due to processing of crude oil from different sources or from variation in operation of the crude distiller process upstream of the FCC; there is also the potential to reroute hydrocarbon streams from other parts of the refinery to the FCC. Second, the FCC is a complex processing unit, the routine operation of which is largely regulated by automatic process control schemes. These schemes, based on such ideas as feedback adjustment (see, e.g., Box and Luceno 1997), are essential for efficient operation but can introduce oscillations in process operation. Third, FCC operation is also subject to numerical optimization to achieve good economic performance. The optimizer adjusts operating parameters on a daily basis. Fourth, FCC operation is further manually adjusted on a periodic basis to adjust the rates at which different products are produced. For example, regulatory requirements for motor gasoline in summer are different to the winter requirements. Finally, as for any industrial process, there are other essentially uncontrollable sources of variation in production, including such effects as operator preference (which tends to vary from shift to shift), causing cyclic trends in operation.

The chemical bases corresponding to A and B are closely related, and the operation of the FCC unit can be modified to produce more of one of A or B at the expense of the other. We therefore expect variables measuring A to be negatively correlated with those measuring B. In contrast, we would expect positive correlation between the measures on each chemical taken at the three different points of the process.

As is typical of process measures encountered in practice, the raw FCC data contained missing values where instruments failed to record measurements on one or more of the variables. Some of these were at isolated time points, whereas others involved short contiguous runs of values. Here we used linear interpolation to impute the short runs, whereas we imputed single missing values by the overall mean of the variable. In addition, there were some outlying observations that clearly corresponded to spurious readings from faulty equipment (because they were not physically possible in the corresponding processes), and these were also replaced by mean imputation. There are arguments against the treatment of missing values by mean imputation and linear interpolation when analyzing covariance and autocovariance matrices (as we do in later sections of this article) because this practice can lead to bias in their estimation. However, the number of missing values here is not large relative to the total length of the observed series, and the dataset that results is believed to be of a quality that is generally acceptable for the purposes of illustrating the methods developed in the subsequent sections of this article, where presentation and exposition is simplified by having a complete set of data. However, it should be noted that the methods that we propose could be used without modification in conjunction with more sophisticated methods for handling missing values if this were felt to appropriate in a particular application, for example, robust estimation of covariance matrices (Little 1988). An alternative strategy, when the proportion of missing values is relatively low, is simply to ignore the missing values and base estimates of covariance and autocovariance matrices on reduced sample sizes, because our methods do not necessarily require that the estimates of the various matrices involved be based on equal sample sizes.

In all of the analyses that we report here, the raw data were standardized so that each variable was mean-centered with unit variance. The standardized base dataset is depicted in Figure 1. As seen from the estimated autocorrelation functions for each of the six process variables, shown in Figure 2, and from their sample correlation matrix, shown in Table 1, the six responses are both correlated and autocorrelated.

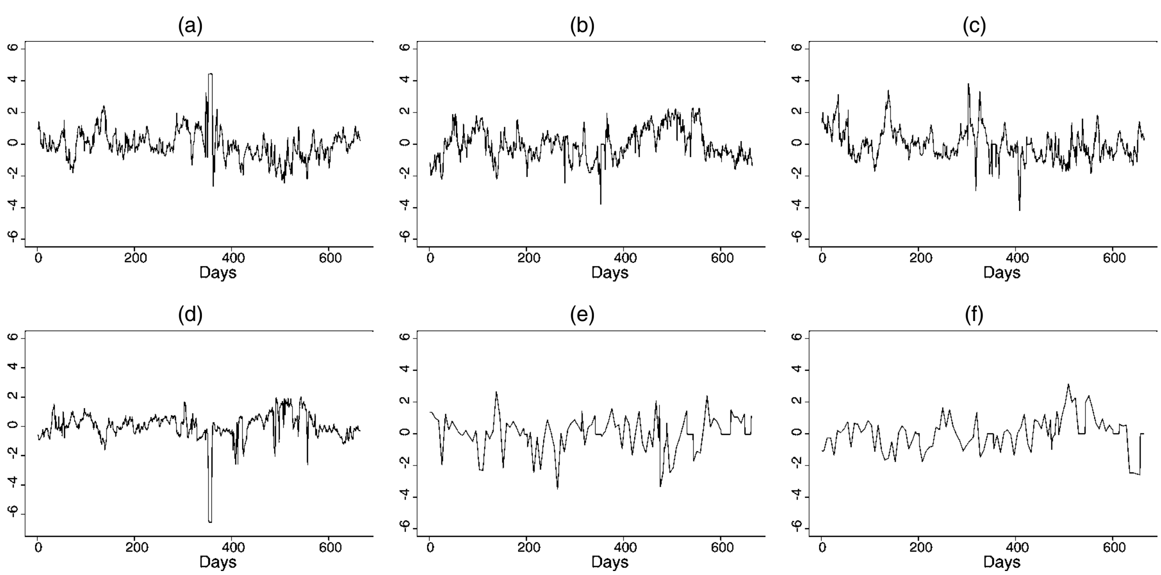

Figure 1. Time Series Plots of Original Standardized Variables for FCC Data (a) A1; (b) B1; (c) A2; (d) B2; (e) A3; (f) B3.

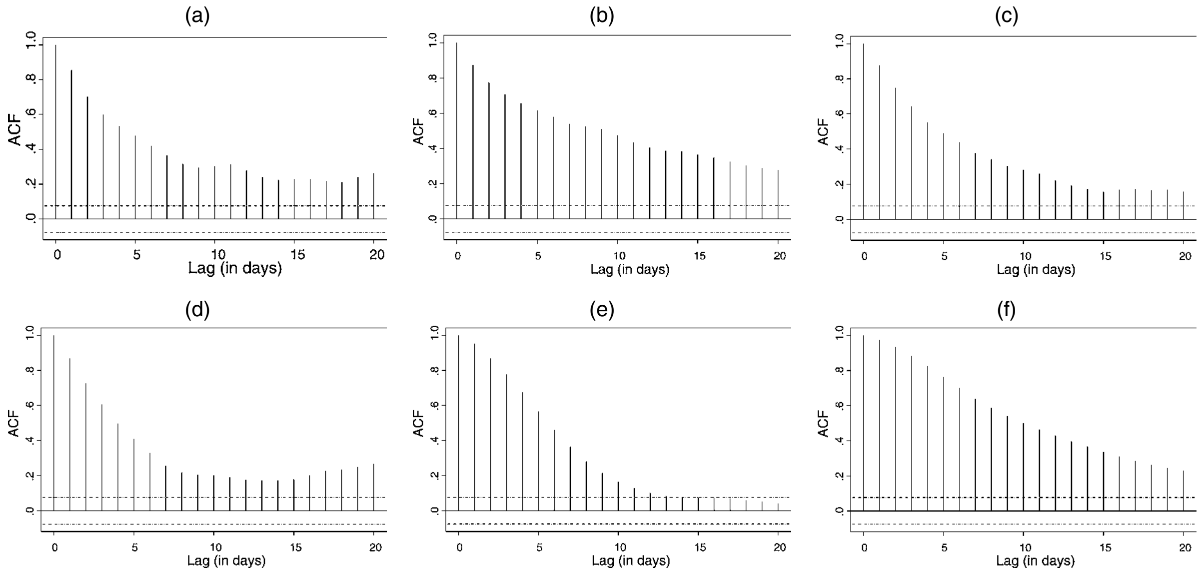

Figure 2. Estimated Autocorrelation Functions of FCC Original Standardized Variables, With 95% Significance Bands Overlaid. (a) A1; (b) B1; (c) A2; (d) B2; (e) A3; (f) B3.

For the purposes of comparison in later sections, we report here the results of a conventional PCA of these data, focusing particularly on how the autocorrelations and lagged cross-correlations between variables in the raw data are affected by the analysis. Figure 3 shows the estimated autocorrelations over 20 daily lags of the 6 principal components with 95% significance bands overlaid in each case. Like the original variables, all of the principal components are quite strongly autocorrelated even over lags of several days. Component 1 exhibits a somewhat more persistent autocorrelation than the others, whereas the autocorrelation structures of components 2–6 are very similar. Therefore, these principal components do not convincingly characterize constituents in the process of a long-, medium-, or short-term nature.

Figure 4 shows the estimated cross-correlation functions over 20 daily lags between various pairs of the first three principal components, which cumulatively account for 78% of the variance in the original data. The 95% significance bands are overlaid in each case, and the cross-correlations are significantly different from 0 for the longer lags in both cases. These principal components are therefore temporally cross-correlated, even though the sample correlation between them at any particular time point is 0. Cross-correlation functions for other pairs of principal components show a similar situation, and at various time lags there are significant lagged associations between all pairs of principal components. Separate univariate process control analyses of these components thus would be confounded rather than independent, and conclusions drawn from such analyses may be compromised by the lagged temporal correlations. For example, knowledge of the occurrence of an "alarm" in a univariate control chart based on one particular principal component modifies our future expectations concerning "alarms" for control charts based on the other principal components.

It is of course not surprising that PCA fails to isolate a distinctive autocorrelation structure convincingly, or that it does not remove temporal cross-correlations in the data, because these do not form part of the objectives of the method. PCA simply concentrates on finding an orthogonal rotation of the variables that maximizes variance under the assumption that the observations for a variable are independent (rather than potentially autocorrelated and temporally cross-correlated). However, the results given earlier for this dataset highlight the point that autocorrelation and cross-correlation structures in the principal components may remain strong. Alternative projection methods are therefore needed if the objective of the analysis is to produce components with particular and possibly more desirable temporal structure. In the next two sections we suggest two possible methods relating to that aim.

Table 1. Sample Correlations of Original Standardized Variables for FCC Data

A1	1.00					
B1	−.47	1.00				
A2	.54	−.48	1.00			
B2	−.59	.49	−.11	1.00		
A3	.31	−.44	.36	−.31	1.00	
B3	−.46	.55	−.36	.40	−.26	1.00

3. "TEMPORALLY STRUCTURED" COMPONENTS

Following the ideas outlined in Section 1, the objective of the first of our alternative projection ideas is to derive a set of "temporally structured" components, that is, uncorrelated components with distinctly different autocorrelation structures. One way to achieve this objective is to identify components with minimum and maximum autocorrelation at some appropriately chosen temporal lag δ. More specifically, we attempt to determine linear combinations of the original data, $u_i(t) = \mathbf{a}_i'\mathbf{y}(t)$, $i = 1, \ldots, p$, such that $\mathrm{corr}(u_i(t), u_j(t)) = 0$ for $i \neq j$ and so that these components successively minimize autocorrelation at the predetermined temporal lag δ, that is, $\mathrm{corr}(u_i(t), u_i(t-\delta)) \leq \mathrm{corr}(u_j(t), u_j(t-\delta))$ for $i < j$.

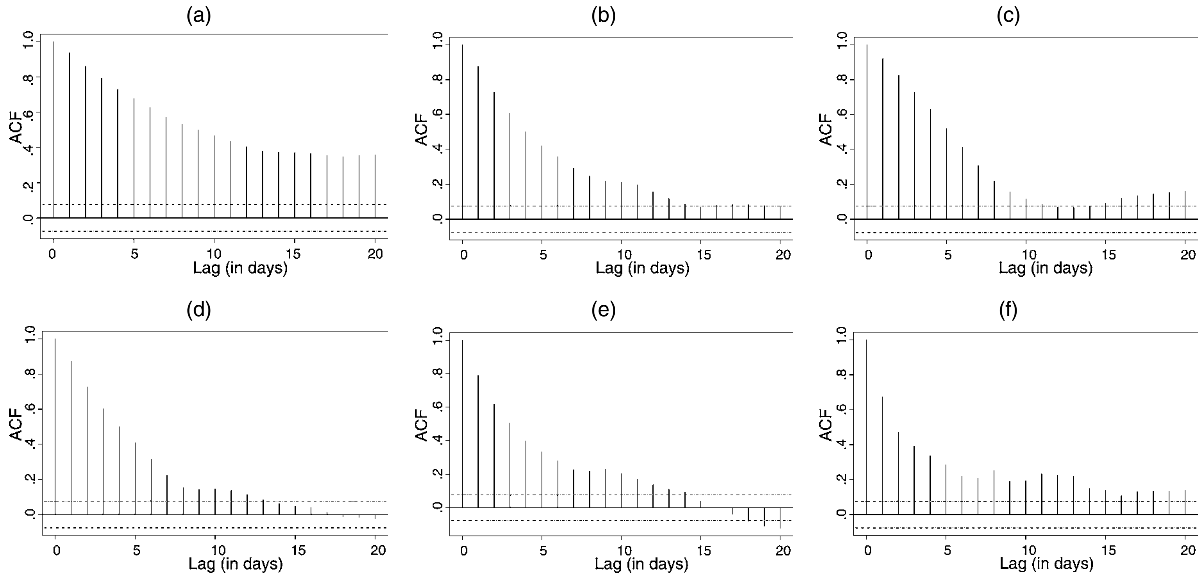

Figure 3. Estimated Autocorrelation Functions of FCC Principal Component Scores, With 95% Significance Bands Overlaid. (a) Principal component 1; (b) 2; (c) 3; (d) 4; (e) 5; (f) 6.

This idea is familiar from the image processing literature, being originally studied in that context by Switzer (1985). It has also recently been used in a more general spatial context by Bailey and Krzanowski (2000). Switzer was interested in removing noise from observed multivariate satellite imagery and argued that if a spatial lag δ was chosen judiciously (typically small in magnitude), then a linear component of the data with minimal spatial autocorrelation at that lag could be associated with "noise" in the image, on the basis that any noise component would tend to have only short-range spatial dependence in contrast with signal components where spatial autocorrelations would persist over longer distances. Once identified, the noise component could then be filtered from the data to produce a cleaner image.

Adapting the mathematical derivation of Switzer's "min/max spatial factors" to our purposes here, and to the time as opposed to the spatial domain, is straightforward and leads to the result that the coefficients, $\mathbf{a}_i$, $i = 1, \dots, p$, that we seek for our "temporally structured" components are solutions to the eigenvalue problem $(\mathbf{C}(\delta) - \lambda(\delta)\boldsymbol{\Sigma})\mathbf{a} = \mathbf{0}$ (see Switzer 1985 for mathematical details), where $\boldsymbol{\Sigma} = \mathrm{cov}(\mathbf{y}(t))$ denotes the covariance matrix of the original observed process, and $\mathbf{C}(\delta) = \mathrm{cov}(\mathbf{y}(t), \mathbf{y}(t-\delta))$ is its autocovariance matrix at time lag δ (i.e., that of its auto- and cross- covariances at lag δ). So, put more simply, our required coefficients, $\mathbf{a}_i$, for each component, $u_i(t) = \mathbf{a}_i'\mathbf{y}(t)$, are the eigenvectors of $\boldsymbol{\Sigma}^{-1}\mathbf{C}(\delta)$, these being arranged in ascending order of their corresponding eigenvalues, $\lambda_i(\delta)$. By convention, the eigenvectors are normalized so that $u_i(t)$ have unit variance, and then it follows that the ith eigenvalue is simply equal to the autocorrelation of the ith component at lag δ, that is, $\lambda_i(\delta) = \mathrm{corr}(u_i(t), u_i(t-\delta))$.

Practical application of the foregoing method depends on the assumption that the multivariate series under investigation is at least second-order stationary, or otherwise the autocovariance is not defined. For most multivariate process datasets encountered in practice, stationarity of both variance and mean are common assumptions and we adopt that stronger stationarity condition here and in subsequent sections of the article. A more difficult practical issue is that the process covariance matrix, $\boldsymbol{\Sigma}$, and autocovariance matrix at lag δ, $\mathbf{C}(\delta)$, would be generally unknown for an observed series and need to be estimated to apply our suggested technique. It is well known (e.g., Percival 1993) that the classical sample estimators of covariance and autocovariance are biased for an observed time series with unknown mean, even when the series is stationary, and

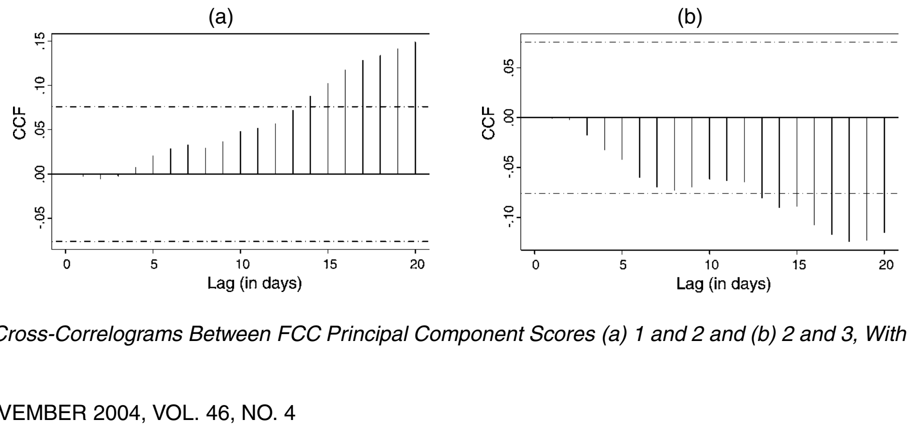

Figure 4. Estimated Cross-Correlograms Between FCC Principal Component Scores (a) 1 and 2 and (b) 2 and 3, With 95% Significance Bands Overlayed.

that this bias can be surprisingly large when strong autocorrelations are present, particularly for estimates of autocovariance at long lags. Furthermore, the classical sample estimate of $\mathbf{C}(\delta)$ does not even guarantee a symmetric and positive definite matrix (e.g., Wackernagel 1995). In contrast, Haslett (1997) suggested a natural sample estimator for the variogram that is unbiased for a mean and variance stationary process and has the added advantage of having only a small bias even when strict stationarity does not apply. For that reason, it is convenient to reformulate our technique in terms of the matrix of semivariances, $\boldsymbol{\Gamma}(\delta) = (1/2)\operatorname{cov}(\mathbf{y}(t) - \mathbf{y}(t-\delta))$, using the relationship between semivariance and covariance for a stationary process, that is, $\boldsymbol{\Gamma}(\delta) = \boldsymbol{\Sigma} - \mathbf{C}(\delta)$. We may then express $(\mathbf{C}(\delta) - \lambda\boldsymbol{\Sigma})\mathbf{a} = \mathbf{0}$ equivalently as $(\boldsymbol{\Gamma}(\delta) - \nu(\delta)\boldsymbol{\Sigma})\mathbf{a} = \mathbf{0}$, where $\nu(\delta) = (1 - \lambda(\delta))$, and estimate $\boldsymbol{\Gamma}(\delta)$ using Haslett's natural estimator, that is, $\hat{\boldsymbol{\Gamma}}(\delta) = \{\hat{\gamma}_{ij}(\delta)\}$, where

$$\hat{\gamma}_{ij}(\delta) = \frac{1}{2(T-\delta-1)} \sum_{t=\delta+1}^{T} \big(d_i(t) - \bar{d}_i\big)\big(d_j(t) - \bar{d}_j\big),$$

with $d_i(t) = y_i(t) - y_i(t-\delta)$ and T denoting the length of the observed series.

The problem of estimating the covariance matrix, $\boldsymbol{\Sigma}$, remains, but the bias in the classical sample covariance estimator is less than that experienced in autocovariance estimation, and the classical covariance estimator is asymptotically unbiased for a stationary series. In the applications reported in this article we are dealing with stationary series observed at several hundreds of time points, and so we simply adopt the classical estimator throughout; that is, we use the sample covariance matrix, $\hat{\boldsymbol{\Sigma}}$, to estimate $\boldsymbol{\Sigma}$. For shorter series, a bias reduction technique such as the jackknife could be used in estimating $\boldsymbol{\Sigma}$, a simple example being the method suggested by Quenouille (1957) in which sample covariance matrices $\hat{\boldsymbol{\Sigma}}_1$, $\hat{\boldsymbol{\Sigma}}_2$, and $\hat{\boldsymbol{\Sigma}}$ are formed from the first half, second half, and whole of the observed series, and then an overall estimate of $\boldsymbol{\Sigma}$ is derived as $\tilde{\boldsymbol{\Sigma}} = 2\hat{\boldsymbol{\Sigma}} - (\hat{\boldsymbol{\Sigma}}_1 + \hat{\boldsymbol{\Sigma}}_2)/2$. Alternatively, one might use the sample semivariance estimate given earlier, but computed at a lag guaranteed to be large enough to exceed the range of the autocovariance structure in the process.

Using the estimates discussed earlier, the required coefficients for our suggested components are therefore obtained in practice as normalized eigenvectors of $\hat{\boldsymbol{\Sigma}}^{-1}\hat{\boldsymbol{\Gamma}}(\delta)$, where these are arranged in descending order of corresponding eigenvalues $\nu_i(\delta)$, because these are now equal to $1 - \operatorname{corr}(u_i(t), u_i(t-\delta))$. Note, however, that $\hat{\boldsymbol{\Sigma}}^{-1}\hat{\boldsymbol{\Gamma}}(\delta)$ is not a symmetric matrix, so it is useful to make a final modification for computational convenience. If $\hat{\boldsymbol{\Sigma}} = \mathbf{F}\mathbf{F}'$ is the spectral decomposition of the estimated covariance matrix $\hat{\boldsymbol{\Sigma}}$, then $(\hat{\boldsymbol{\Gamma}}(\delta) - \nu(\delta)\hat{\boldsymbol{\Sigma}})\mathbf{a} = \mathbf{0}$ reduces to

$$\big(\mathbf{F}^{-1}\hat{\boldsymbol{\Gamma}}(\delta)(\mathbf{F}')^{-1} - \nu(\delta)\mathbf{I}\big)\mathbf{e} = \mathbf{0}, \qquad (1)$$

where $\mathbf{e} = \mathbf{F}'\mathbf{a}$. Thus, finally, our required "temporally structured" components $u_i(t) = \mathbf{a}_i'\mathbf{y}(t)$ are given by $\mathbf{a}_i = (\mathbf{F}')^{-1}\mathbf{e}_i$, where $\mathbf{e}_i$ are the normalized eigenvectors of the symmetric matrix $\mathbf{F}^{-1}\hat{\boldsymbol{\Gamma}}(\delta)(\mathbf{F}')^{-1}$ in descending order of corresponding eigenvalues $\nu_i(\delta) = 1 - \operatorname{corr}(u_i(t), u_i(t-\delta))$.

Given a judicious choice of the prespecified time lag δ, this method provides a straightforward way to transform multivariate process data into a set of constituent components that are uncorrelated and may be interpretable in terms of physically meaningful long-, medium-, or short-range features of the process. The "success" of an analysis of this nature will depend on a suitable choice of δ and will be one that results in a distinct separation between components with short-range and long-range autocorrelation structures. This in turn corresponds to component autocorrelations at lag δ being close in absolute value to either end of the $[0, 1]$ interval. Because $\nu_i(\delta) = 1 - \operatorname{corr}(u_i(t), u_i(t-\delta))$, this suggests that we need to choose δ so as to optimize a function that characterizes the range or spread of the eigenvalues obtained at lag δ. Various possible functions could be adopted; here we suggest simply taking the variance of these eigenvalues,

$$G(\delta) = \frac{1}{p-1} \sum_{i=1}^{p} \big(\nu_i(\delta) - \bar{\nu}(\delta)\big)^2. \qquad (2)$$

Maximizing (2) over a sensible range of lags δ will then indicate a value of δ that produces components from (1) that strongly separate long- and short-term components. $G(\delta)$ is not, however, a simple monotonic function, and so there could be several local maxima and possibly no clear global maximum. If this is the case, then it is a matter of trying each of the local maxima in contention to judge the best suitable lag.

We can use simulation to see how our suggested method performs in general in uncovering the underlying structure of a multivariate series constructed from a convolution of components with distinctly different temporal structures. To do this, we simulated 100 datasets, each composed of 500 observations from three independent ARMA(1, 1) time series. The latter were chosen so as to have short-, intermediate-, and long-range autocorrelation functions ($\phi = .64$ and $\theta = -.20$ for series 1, $\phi = .83$ and $\theta = -.60$ for series 2, and $\phi = .92$ and $\theta = -.15$ for series 3). Random errors were drawn in each case from an N(0, 4) distribution for 75% of the time and from an N(0, 25) distribution for the remaining time. Once generated, the three series in each of the 100 datasets were linearly combined using randomly selected nonsingular matrices to produce 100 realizations of a trivariate, correlated multivariate process of length 500. Our suggested method for deriving temporally structured components, implemented via (1), with δ chosen via optimization of (2), was then applied in each case (where a choice of $\delta = 5$ was found appropriate). For comparison, a standard PCA was also applied for each dataset, and then the autocorrelation functions for both the temporally structured components and the principal component scores were estimated for each realization. These functions were then averaged over the 100 datasets. The resulting average estimated autocorrelation functions for each of the three temporally structured components and for each of the three principal components are depicted in Figure 5, along with those of the three original constituent ARMA series. The maximum and minimum autocorrelation values obtained over the 100 datasets for each lag for each of the various components are also overlaid in Figure 5.

A clear general picture emerges from these simulation results in that the autocorrelation structures of our temporally structured components look very similar to those of the three original constituent series, even given the min/max variations experienced over the 100 realizations. On average, our method

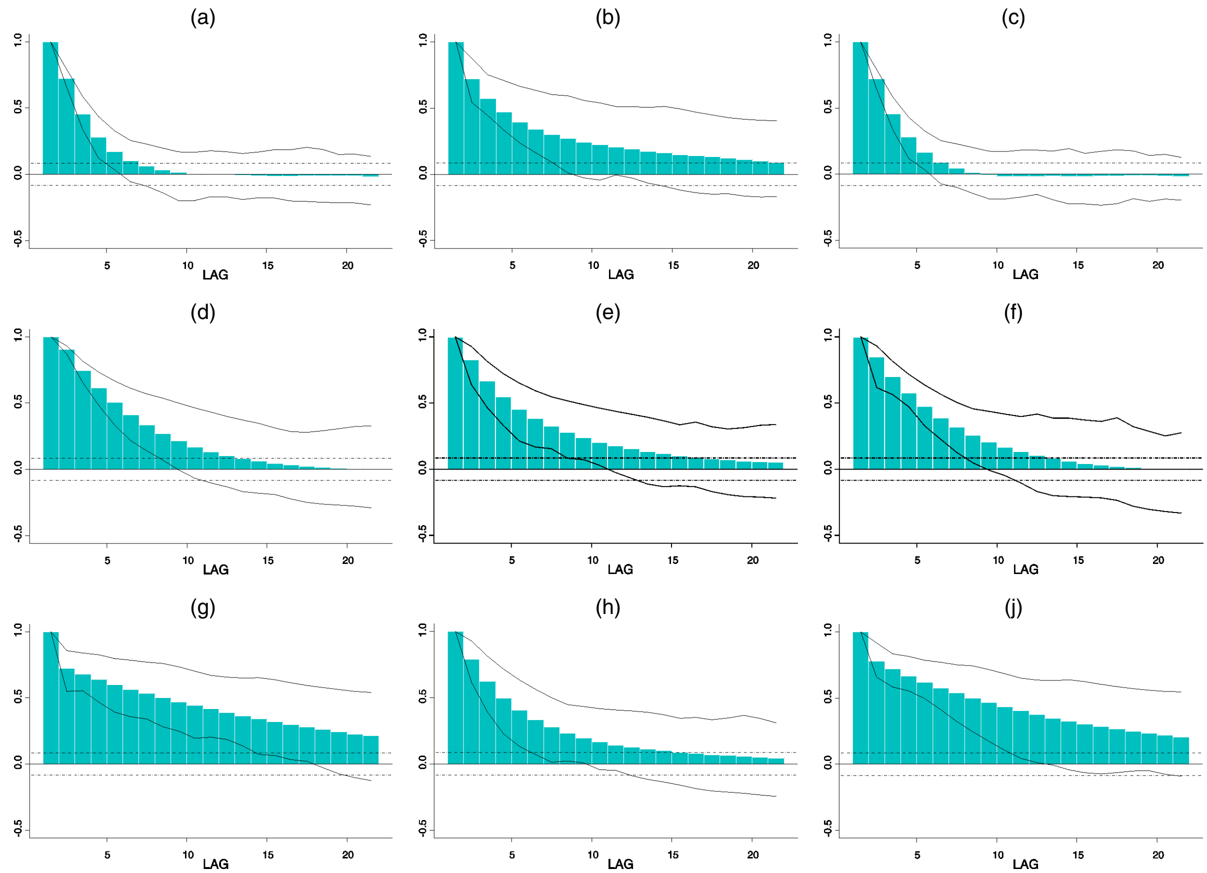

Figure 5. Average Estimated Autocorrelation Functions Over 100 Realizations for Three Constituent ARMA Series [(a) 1, (d) 2, (g) 3] and for the Scores From the "Temporally Structured" [(c) 1, (f) 2, (i) 3] and Principal Components [(b) 1, (e) 2, (h) 3] Derived From Random Linear Combinations of These Series, With Overlaid Envelopes of the Min/Max Values Obtained in the Realizations.

does successfully decompose the convoluted three variable series into its constituent short-, intermediate-, and long-range components. In contrast, the three principal components fail to recapture this underlying structure in the data. On average, the estimated autocorrelation functions of all components are very similar and do not correspond well to the three distinct structures underlying the simulated multivariate series. These results imply that our suggested method performs well (and certainly better than PCA) at uncovering underlying temporally structured components in a multivariate series, and more extensive simulation results, not reported here, confirm this when various other constituent numbers and types of autocorrelation structures are combined to make up the multivariate series analyzed.

Turning to the real FCC dataset introduced in Section 2, values of $G(\delta)$, as defined in expression (2), for δ in the range 0–15 days for these data are shown in Figure 6 (with lags of longer than 2 weeks being of little practical interest here). A lag of 0 days is equivalent to carrying out a standard PCA and is included for comparison. The maximum value of $G(\delta)$ occurs here at $\delta = 7$. Figure 7 shows the autocorrelations of the six temporally structured components derived from (1) using this choice of δ. The components demonstrate markedly different autocorrelation structures. Components 1 and 2, with the shortest-range autocorrelations, are consistent with more transient effects, possibly "noise," in the process, whereas component 6 exhibits autocorrelation extending over the full 20-day period and corresponds to global behavior. The other components display intermediate behavior. This picture contrasts with the autocorrelations of the principal components given previously in Figure 3, where there is less of an obvious difference in the autocorrelation structures of the components, making it more difficult to interpret them in terms of long- or short-term effects.

The coefficients for each of the derived temporally structured components are given in Table 2. In general, when standard-

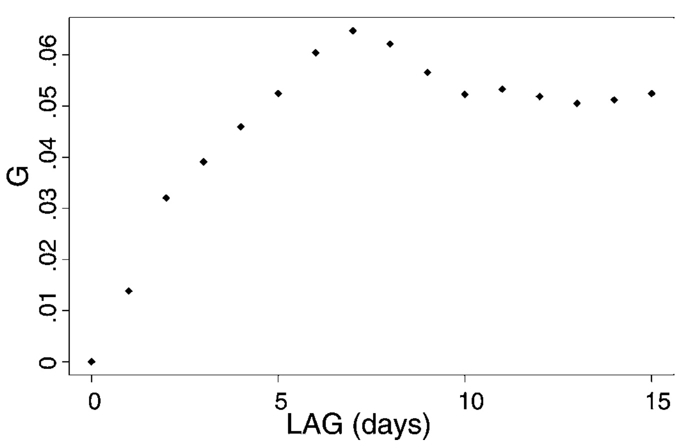

Figure 6. Values of $G(\delta)$ for FCC "Temporally Structured" Components Obtained for δ Ranging Between 1 and 15 Days.

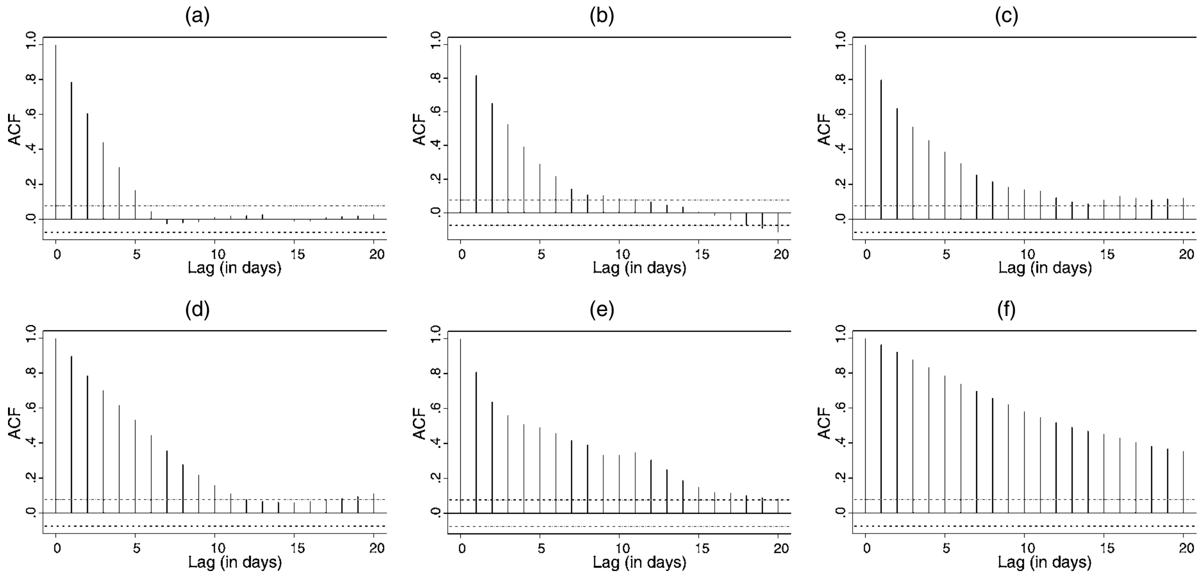

Figure 7. Estimated Autocorrelation Functions of FCC "Temporally Structured" Component Scores [(a) 1, (b) 2, (c) 3, (d) 4, (e) 5, (f) 6] Obtained for $\delta = 7$ Days, With 95% Significance Bands Overlaid.

ized by the variances of the relevant original variables, such coefficients allow the components to be interpreted in terms of contributions from each of the original variables. Here standardization is unnecessary, because the original data were standardized before analysis and so raw coefficients can be used directly. From this it is clear, for example, that the long-range component 6 is essentially a mixture of B1 and B3, whereas the shortest-range component is contrasting B2 with the sum of B1 and A2.

Note that the temporally structured components proposed here are uncorrelated, but they are not an orthogonal transformation of the original data; that is, the total variance in the original data is not preserved by the transformation as it would be in standard PCA. For the FCC data analyzed here, the six components taken together account for approximately 64% of the original variance in the data. This is comparable to the 66% that is cumulatively explained by the first two principal components considered in Section 2. Of course, the objective here is not necessarily one of reducing the dimensionality of the data, but rather of identifying long- or short-term constituents of the process behavior. Nevertheless, it is interesting to note that the temporally structured components do retain most of the variance in the original data. Individual variances for each may be calculated directly from the component scores, and, if desired, these may be used to provide a basis for ordering the relative importance of each component and perhaps a justification for retaining only a subset of these in subsequent analyses.

Table 2. Coefficients of FCC "Temporally Structured" Components Obtained for $\delta = 7$ Days

	Component					
Variable	*1*	*2*	*3*	*4*	*5*	*6*
A1	−.14	−.50	.60	−.14	−.65	−.19
B1	.39	−.56	−.39	−.55	.15	.49
A2	.42	.14	.23	−.27	.70	.15
B2	−.78	−.23	.38	−.24	−.16	−.05
A3	−.18	−.60	−.23	.40	.18	−.20
B3	.12	.04	.49	.63	−.03	.81

4. "TEMPORALLY UNCORRELATED" COMPONENTS

A byproduct of the projection technique discussed in the previous section is that it produces uncorrelated components that also have zero cross-correlation at the single specific lag δ on which the analysis is based. Although this property does not necessarily guarantee small cross-correlation at other lags, our experience in practice is that it does tend to result in components with generally smaller lagged cross-correlations than those between the principal components. However, rather than simply relying on this general flattening in cross-correlations induced by enforcing zero cross-correlation at one specific lag, there may also be value (for the reasons outlined in Sec. 1) in components that are more directly designed to minimize cross-correlations over all lags. In this section we therefore focus on the second of the alternative projection ideas mentioned in Section 1 concerning the derivation of a set of "temporally uncorrelated" components, that is, linear combinations of the original data, $u_i(t) = \mathbf{a}_i'\mathbf{y}(t)$, $i = 1, \dots, p$, with the property that $\text{corr}(u_i(t), u_j(t)) = \text{corr}(u_i(t), u_j(t-\delta)) = 0$ for $i \neq j$ and for all time lags δ. The key motivation is that such components may be used in separate univariate analyses with more confidence than would be the case with conventional principal components because, although the latter are uncorrelated at any point in time, they may, as demonstrated in Section 2, remain cross-correlated over a range of time lags.

Our aforementioned objective may be achieved by adapting techniques from those used in common PCA (Flury 1988). To see this, suppose that $\delta_1, \dots, \delta_K$ are a set of K temporal lags chosen so as to reasonably span the full temporal range of lags over which cross-correlation might be expected to persist between the original variables. Then, with notation similar to that used in Section 3, the problem of finding coefficients for "temporally uncorrelated" components involves finding solutions, $\mathbf{a}$, to the set of simultaneous equations

$$\big(\mathbf{C}(\delta_k) - \lambda(\delta_k)\boldsymbol{\Sigma}\big)\mathbf{a} = \mathbf{0} \quad \text{for all } k = 1, \dots, K, \tag{3}$$

where, as before, $\mathbf{C}(\delta_k) = \text{cov}(\mathbf{y}(t), \mathbf{y}(t - \delta_k))$ are the autocovariance matrices of the original process at each of the time lags δ_k and $\boldsymbol{\Sigma} = \text{cov}(\mathbf{y}(t))$ is its covariance matrix.

As in Section 3, issues of practical estimation and computational convenience would arise in directly applying the foregoing formulation to an observed series. So, as before, we reexpress the problem in (3) in terms of sample estimates of semivariance matrices, $\hat{\boldsymbol{\Gamma}}(\delta_k)$, and use the spectral decomposition of the sample estimate of the covariance matrix, $\hat{\boldsymbol{\Sigma}} = \mathbf{FF}'$, to obtain the series of equations

$$\big(\mathbf{F}^{-1}\hat{\boldsymbol{\Gamma}}(\delta_k)(\mathbf{F}')^{-1} - \nu(\delta_k)\mathbf{I}\big)\mathbf{e} = \mathbf{0} \quad \text{for all } k = 1, \dots, K, \tag{4}$$

where $\mathbf{a} = (\mathbf{F}')^{-1}\mathbf{e}$ and $\nu(\delta_k) = 1 - \lambda(\delta_k)$.

Therefore, in practice, the problem of identifying components with zero cross-correlations at each of a set of lags $\delta_1, \dots, \delta_k$ reduces to that of finding common eigenvectors, $\mathbf{e}_i$, $i = 1, \dots, p$, of the set of symmetric matrices $\mathbf{F}^{-1}\hat{\boldsymbol{\Gamma}}(\delta_k)(\mathbf{F}')^{-1}$, $k = 1, \dots, K$. This is essentially an identical problem to that encountered in the well-established method of common PCA, and although there is no exact mathematical solution in the general case, there are maximum likelihood or least squares algorithms readily available to obtain a set of eigenvectors that best approximate the simultaneous diagonalization of these K matrices (Flury and Constantine 1985; Clarkson 1988). Once the best approximating common eigenvectors, $\mathbf{e}_i$, are found via one or other of these algorithms, then the required coefficients, $\mathbf{a}_i$, $i = 1, \dots, p$, for our set of "temporally uncorrelated" components, $u_i(t) = \mathbf{a}_i'\mathbf{y}(t)$, $i = 1, \dots, p$, are simply formed from $\mathbf{a}_i = (\mathbf{F}')^{-1}\mathbf{e}_i$. The K sets of eigenvalues, $\nu_i(\delta_k)$, $i = 1, \dots, p$, $k = 1, \dots, K$, corresponding to the p common eigenvectors $\mathbf{e}_i$, then have analogous properties to the single set encountered in Section 3, that is, $\nu_i(\delta_k) = 1 - \text{corr}(u_i(t), u_i(t - \delta_k))$.

Thus formulated and given a sensible choice of the time lags δ_k, our method provides a practical way to transform multivariate process data into a set of "temporally uncorrelated" constituent components $u_i(t)$, $i = 1, \dots, p$, that is, components for which $\text{corr}(u_i(t), u_j(t)) = 0$ and $\text{corr}(u_i(t), u_j(t - \delta)) \approx 0$ for all $i \neq j$ and all "reasonable" time lags δ. As in Section 3, the proposed method depends on sensible choices of time lags. However, here the lags δ_k are simply required to be representative of all "reasonable" time lags. The choice thus presents little practical difficulty, because it is relatively straightforward to decide a priori on a realistic range over which lagged cross-correlation in the original process might extend. One then chooses K such that equally spaced δ_k, $k = 1, \dots, K$, sufficiently cover that range.

To check the performance of our suggested method in producing components with low cross-correlation, we simulated 100 datasets each composed of 500 observations from a trivariate VARMA (vector ARMA) model chosen to have strong autocorrelation and cross-correlation structures. More specifically, we used a modification of the artificial three-variable VARMA(1, 0) series given by Quenouille (1957), taking his autoregressive parameter matrix

$$\boldsymbol{\Phi} = \begin{pmatrix} 1.0 & -.1 & 0 \\ .2 & 1.0 & -.3 \\ 0 & .9 & 0 \end{pmatrix}$$

but adjusting the error structure to reflect a similar mixture to that used in the simulations in Section 3. Random errors were drawn from a multivariate $\text{N}(\mathbf{0}, \boldsymbol{\Sigma}_1)$ distribution 75% of the time and from an $\text{N}(\mathbf{0}, \boldsymbol{\Sigma}_2)$ distribution the remaining time, with

$$\boldsymbol{\Sigma}_1 = \begin{pmatrix} 4.0 & & \\ 1.6 & 4.0 & \\ 1.6 & 1.6 & 4.0 \end{pmatrix}, \qquad \boldsymbol{\Sigma}_2 = \begin{pmatrix} 25.0 & & \\ 10.0 & 25.0 & \\ 10.0 & 10.0 & 25.0 \end{pmatrix}.$$

Table 3 gives the results of applying our method to the simulated series from this model and includes corresponding comparative results for the raw series. The cells of the table report the averages for the 100 realizations (with standard deviations in parentheses) of the mean absolute values of the cross-correlations over lags 1–10 between the three raw variables and between the "temporally uncorrelated" components. On average, our "temporally uncorrelated" components exhibit markedly less overall cross-correlation than the original variables. Our suggested method is therefore effective in decomposing a correlated multivariate series into components that at least have small, if not zero, temporal cross-correlations, something that cannot be guaranteed for PCA. Space prevents us from reporting the details here, but this picture is confirmed in more extensive simulations involving a variety of other types and higher-dimensional multivariate time series models.

We now return to the FCC example used in Sections 2 and 3 and examine the results of applying our method in that six-variable case. Components were derived via (4), selecting lags, δ_k, $k = 1, \dots, 7$, to correspond to 1, 4, 7, 10, 13, 16, and 19 days. Figure 8 shows the estimated cross-correlation functions over 20 daily lags between the "temporally uncorrelated" component scores 1 and 2 and 2 and 3, with 95% significance bands overlaid in each case. All cross-correlations between the first and second components lie wholly within the 95% significance band, showing that the temporal correlation between components is not significantly different from zero over all lags. The cross-correlations between the second and third components show only a few marginally significant lags. Hence we can conclude that the method has been successful in substantially removing cross-correlations between components. These results may be contrasted with the corresponding cross-correlation functions for the principal components given previously in Figure 4.

Table 3. Averages for 100 Realizations of Simulated VARMA Series (with standard deviations in parentheses) of the Mean Absolute Values of the Cross-Correlations Over Lags 1 to 10 for Raw Variables and for "Temporally Uncorrelated" Components

	Raw variables		Temporally uncorrelated components	
	1	2	1	2
2	.4735		.1825	
	(.080)		(.060)	
3	.4448	.4311	.0471	.1378
	(.080)	(.074)	(.018)	(.019)

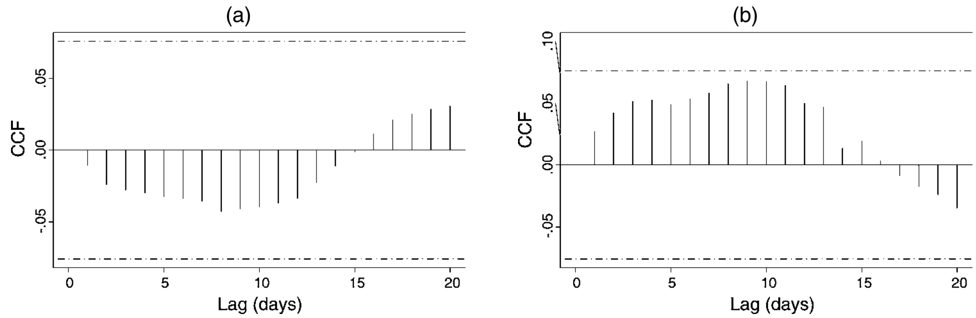

Figure 8. Estimated Cross-Correlograms Between FCC "Temporally Uncorrelated" Component Scores (a) 1 and 2 and (b) 2 and 3, With 95% Significance Bands Overlaid.

Cross-correlations between all possible pairs of the six derived "temporally uncorrelated" components are usefully summarized in Table 4, which reports the mean absolute value over 20 daily lags of the estimated cross-correlations between the "temporally uncorrelated" component scores and between the original FCC variables. Values in this table for the "temporally uncorrelated" components are generally much smaller than the corresponding figures for the original variables, confirming that the minimization of lagged cross-correlations has been consistent over each of the pairs of components. That situation is not guaranteed for the principal components scores derived in Section 2.

Note that, as in Section 3, our "temporally uncorrelated" components are not an orthogonal transformation of the original data and thus, in contrast to the principal components, do not preserve the total variance in the original data. In this case the six "temporally uncorrelated" components account for approximately 72% of the original variance in the data, comparable to the 78% cumulatively explained by the first three principal components in Section 2. The "temporally uncorrelated" components thus retain most of the variance in the original data. Individual variances for each may be calculated directly from the component scores, and, if desired, these may be used to provide a basis for ordering the relative importance of each component and perhaps a justification for retaining only a subset of these in subsequent analyses. An alternative basis for an ordering would be provided by considering the eigenvalue sums, $\sum_{k=1}^{K} \nu_i(\delta_k)$. Arranging the components in descending order of these quantities would correspond to attempting to match the components to short-range as opposed to long-range process behavior. This results in components with similar properties to those obtained in Section 3, but with the added advantage of a lack of lagged cross-correlation. Clearly, however, there is no guarantee that the components here would be as successful in exhibiting distinctly different autocorrelation structures as those obtained by the methods discussed in Section 3. In the case of the FCC data, it is components 4 and 1 that correspond to the shortest- and longest-range components.

The primary motivation for considering a method for deriving "temporally uncorrelated" components was that their lack of cross-correlation would make them well suited to subsequent separate univariate analyses without any need to consider potential difficulties that might arise from variables in these analyses having strong temporal cross-correlation. Figure 9 provides an illustration of this by comparing univariate Shewhart charts for principal components 1, 2, and 3 with similar charts for the three "most important" "temporally uncorrelated" components, that is, those with the three highest individual variances. The 95% upper and lower control limits are also indicated in Figure 9 (adjusted to allow for multiple testing over the six variables).

The general behavior of the process indicated in each of the principal components can be matched to that for a corresponding "temporally uncorrelated" component (1 matches 1, 2 matches 2 and 3 matches minus 3). However, although the overall chart picture is similar for each principal component and for its "temporally uncorrelated" counterpart, a closer inspection reveals some differences in the detail. This is particularly interesting where it concerns inconsistencies in the values that exceed the control limits, and it serves to illustrate that different judgments can arise from "temporally uncorrelated," as opposed to principal, components. The general point to be stressed is that separate significance tests carried out on individual principal components are not independent (due to potential temporal cross-correlations between components), whereas corresponding tests performed on individual "temporally uncorrelated" components are less confounded, because

Table 4. Mean Absolute Auto- and Cross-Correlations Over Daily Lags 1–20 Between FCC "Temporally Uncorrelated" Component Scores and Between Original Standardized Variables

	Temporally uncorrelated components						*FCC original variables*					
	1	*2*	*3*	*4*	*5*	*6*	*1*	*2*	*3*	*4*	*5*	*6*
1	.633						.416					
2	.064	.372					.315	.539				
3	.087	.059	.307				.223	.246	.396			
4	.035	.037	.055	.196			.247	.311	.075	.371		
5	.091	.066	.040	.079	.285		.149	.261	.165	.208	.355	
6	.057	.048	.099	.061	.058	.294	.347	.420	.231	.299	.210	.584

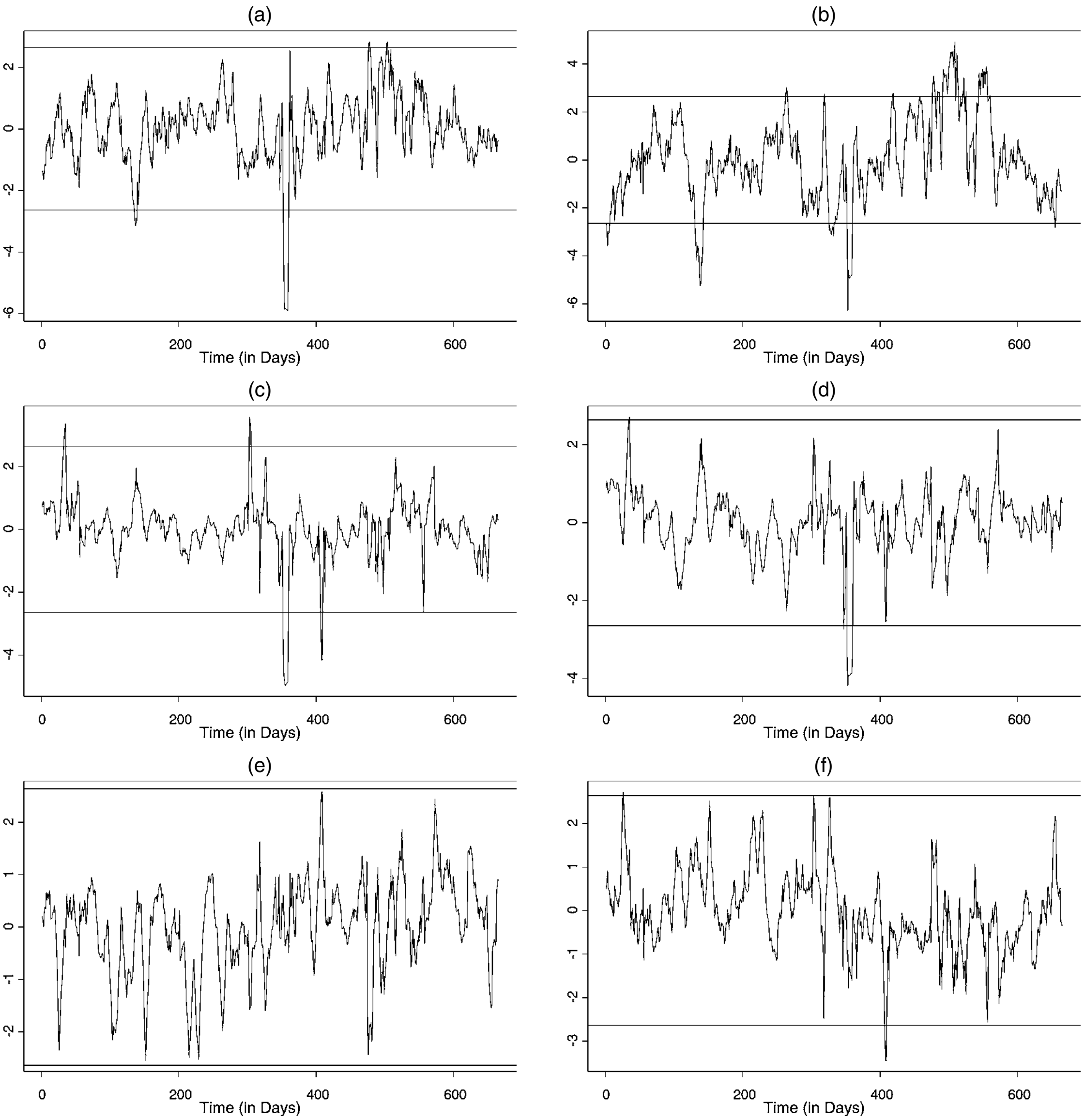

Figure 9. Standard Shewhart Control Charts for FCC Principal Component Scores (a) 1, (c) 2, and (e) 3 and for the Three FCC Temporally Uncorrelated Components With the Largest Variance [(b) 1, (d) 2, (f) 3], With Upper and Lower Control Limits Overlaid.

the lagged cross-correlations between these are generally reduced. We thus may have more confidence in any control warnings given by the "temporally uncorrelated" components.

5. CONCLUSIONS

We have proposed two multivariate projection methods that may be useful in the analysis of multivariate process data exhibiting strong temporal autocorrelations and cross-correlations. The first technique derives components that exhibit distinctly different autocorrelation structures and may be used to characterize and separately analyze short-, intermediate-, or long-term behavior in the multivariate process. The second method produces components that are approximately temporally uncorrelated and allow application of separate univariate techniques without such analyses being compromised by temporal dependency. Both methods are relatively straightforward to implement in practice given access to the standard matrix algebra routines that exist in most statistical software. For example, all of the analyses in this article used functions developed within S–PLUS, which are available from the authors on request.

Both methods were demonstrated to be effective, both in simulated studies and when applied to real data from a fluidized catalytic cracking unit typical of the kind of serially correlated multivariate data arising from many industrial manufacturing processes. In the latter case, the resulting components retained most of the variability in the original data while also providing distinctly different, and possibly more advantageous, properties from those that arise from standard PCA applied to the same data.

At this stage, we present these methods primarily as useful exploratory techniques. Further simulation studies are currently being conducted on inferential aspects and such issues as their robustness to the distributional properties of the original data. The possibilities for updating these methods in real time are also being investigated, so as to be able to apply them dynamically in "on-line" process control. However, one benefit of the methods that we suggest is that they depend on few assumptions about the original data, whereas other multivariate approaches assume specific process models. The techniques thus can potentially be applied to a broad range of data, provided that the processes considered are stationary.

The analysis of multivariate process data is a difficult practical problem that is probably best approached using a multiplicity of methods. Ultimately, the justification for any particular methodology is the extent to which it aids in a meaningful and useful analysis of data in practice. We believe that the methods discussed here do represent a useful addition to the growing range of statistical tools in this area, particularly when the data have strong temporal dependencies.

ACKNOWLEDGMENTS

This research was supported through a Cooperative Award in Science and Engineering (CASE) studentship funded jointly by the U.K. Engineering and Physical Sciences Research Council (EPSRC) and Shell Research Ltd. (No. 99316137).

[Received February 2002. Revised September 2002.]

REFERENCES

Alt, F. B., and Smith, N. D. (1988), *Handbook of Statistics, Vol. 7: Quality Control and Reliability*, Amsterdam: Elsevier.

Aparisi, F., Jabaloyes, J., and Carrión, A. (1999), "Statistical Properties of the $|S|$ Multivariate Control Chart," *Communications in Statistics, Part A—Theory and Methods*, 28, 2671–2686.

Bailey, T. C., and Krzanowski, W. J. (2000), "Extensions to Spatial Factor Methods With an Illustration in Geochemistry," *Mathematical Geology*, 32, 657–682.

Bissel, D. (1994), *Statistical Methods for SPC and TQM*, Padstow, U.K.: T. J. Press.

Box, G., and Luceno, A. (1997), *Statistical Control by Monitoring and Feedback Adjustment*, New York: Wiley.

Champ, C. W., Tracy, N. D., and Young, J. C. (1997), "Assessment of Multivariate Process Control Techniques," *Journal of Quality Technology*, 29, 140–143.

Clarkson, D. B. (1988), "ASR71, A Remark on Algorithm AS211, the F–G Diagonalisation Algorithm," *Applied Statistics*, 37, 147–151.

Faltin, F. W., Mastrangelo, C. M., Runger, G. C., and Ryan, T. P. (1997), "Considerations in the Monitoring of Autocorrelated and Independent Data," *Journal of Quality Technology*, 29, 131–133.

Flury, B. N. (1988), *Common Principal Components and Related Multivariate Models*, New York: Wiley.

Flury, B. N., and Constantine, G. (1985), "AS211, the F–G Diagonalisation Algorithm," *Applied Statistics*, 35, 177–183.

Fuchs, C., and Benjamini, Y. (1994), "Multivariate Profile Charts for Statistical Process Control," *Technometrics*, 36, 182–195.

Gray, J. H., and Handwerk, G. E. (1994), *Petroleum Refining: Technology and Economics*, New York: Marcel Dekker.

Harris, T. J., and Ross, W. H. (1991), "Statistical Process Control Procedures for Correlated Observations," *The Canadian Journal of Chemical Engineering*, 69, 48–57.

Haslett, J. (1997), "On the Sample Variogram and the Sample Autocovariance for Non-Stationary Time Series," *The Statistician*, 46, 475–485.

Jackson, J. E. (1985), "Multivariate Quality Control," *Communications in Statistics, Part A—Theory and Methods*, 14, 2657–2688.

Jiang, W., Tsui, K. L., and Woodall, W. H. (2000), "A New SPC Monitoring Method: The ARMA Chart," *Technometrics*, 42, 399–410.

Jonathan, P., and Kaskavelis, E. A. (1998), "Prediction of Flow Composition From a Continuous Fluidised Catalytic Cracking Unit," unpublished technical report, Shell Global Solutions, Chester, U.K.

Ku, W., Storer, R. H., and Georgakis, C. (2001), "Disturbance Detection and Isolation by Dynamic Principal Component Analysis," available at *http://www. emsl.pnl.gov:2080/docs/incinc/proc_shifts/WKdoc.html*.

Kulkarni, S. R., and Paranjape, S. R. (1984), "Use of Andrews' Function Plot Technique to Construct Control Curves for Multivariate Processes," *Communications in Statistics, Part A—Theory and Methods*, 13, 2511–2533.

Little, R. J. A. (1988), "Robust Estimation of the Mean and Covariance Matrix From Data With Missing Values," *Applied Statistician*, 37, 23–38.

Lowry, C. A., and Montgomery, D. C. (1995), "A Review of Multivariate Control Charts," *IIE Transactions Industrial Engineering Research And Development*, 27, 800–810.

Lui, R. Y. (1995), "Control Charts for Multivariate Processes," *Journal of the American Statistical Association*, 90, 1380–1387.

Molenaar, P. C. M. (1985), "A Dynamic Factor Model for the Analysis of Multivariate Time Series," *Psychometrika*, 50, 181–202.

Percival, D. B. (1993), "Three Curious Properties of the Sample Variance and Autocovariance for Stationary Processes With Unknown Mean," *The American Statistician*, 47, 274–276.

Quenouille, M. H. (1957), *The Analysis of Multiple Time Series*, London, U.K.: Griffin.

Runger, G. C., and Alt, F. B. (1996), "Choosing Principal Components for Multivariate Statistical Process Control," *Communications in Statistics, Part A—Theory and Methods*, 25, 909–922.

Runger, G. C., Alt, F. B., and Montgomery, D. C. (1996), "Contributors to a Multivariate Statistical Process Control Chart Signal," *Communications in Statistics, Part A—Theory and Methods*, 25, 2203–2213.

Sadeghbeigi, R. (1995), *Fluid Catalytic Cracking Handbook*, Houston, TX: Gulf.

Sparks, R., Adolphson, A., and Phatak, A. (1997), "Multivariate Process Monitoring Using Dynamic Biplots," *International Statistical Review*, 65, 325–349.

Switzer, P. (1985), "Min/Max Autocorrelation Factors for Multivariate Spatial Imagery," in *Computer Science and Statistics: Proceedings of the 16th Symposium on the Interface*, pp. 13–16.

VanBrackle, L. N., and Reynolds, M. R., Jr. (1997), "EWMA and CUSUM Control Charts in the Presence of Correlation," *Communications in Statistics, Part B—Simulation and Computation*, 26, 979–1008.

Wachs, A., and Lewin, D. R. (1999), "Improved PCA Methods for Process Disturbance and Failure Identification," *AiChE. Journal*, 45, 1688–1700.

Wackernagel, H. (1995), *Multivariate Geostatistics, an Introduction With Applications*, Berlin: Springer-Verlag.

Wetherill, G. B., and Brown, D. W. (1991), *Statistical Process Control—Theory and Practice*, London: Chapman & Hall.

Wilkström, C., Albano, C., Eriksson, L., Fridén, H., Johansson, E., Nordahl, Å., Rännar, S., Sandberg, M., Kettaneh-Wold, N., and Wold, S. (1998a), "Multivariate Process and Quality Monitoring Applied to an Electrolysis Process, Part I. Process Supervision With Multivariate Control Charts," *Chemometrics and Intelligent Laboratory Systems*, 42, 221–231.

——— (1998b), "Multivariate Process and Quality Monitoring Applied to an Electrolysis Process, Part II. Multivariate Time-Series Analysis of Lagged Latent Variables," *Chemometrics and Intelligent Laboratory Systems*, 42, 233–240.