BAYESIAN METHODS FOR CONTROL LOOP PERFORMANCE ASSESSMENT IN CROSS-DIRECTIONAL CONTROL

A.R. Taylor [*,1] **S.R Duncan** [*,2]

[*] *Department of Engineering, University of Oxford, UK*

Abstract: The minimum variance (MV) lower bound has been applied to many multivariable control systems in order to assess their performance based on routine operating data. However, such analysis often depends on the selection of a suitable dynamic model of the data and for multivariable systems, there can be many candidate models. Also, uncertainty is often not considered, because standard approximations do not exist for the sampling distribution of these multivariable performance indices. This paper addresses these two issues by using the Bayesian approach to vector autoregression (VAR) modelling with Markov Chain Monte Carlo (MCMC) numerical methods. Dynamic model selection is carried out by using Reversible Jump (RJ) MCMC and it is shown that MCMC can be used to the estimate the non-standard distributions that exist in multivariable MV performance indices. The approach is applied to data from an industrial cross-directional (CD) control system by using a more general class of model than has previously been studied for these systems.

Keywords: Performance Analysis, Multivariable Control, Identification

1. INTRODUCTION

One of the most widespread approaches that is used in control loop performance assessment is to apply the minimum variance (MV) lower bound to routine operating data. In this paper, the MV approach is applied to cross-directional (CD) control systems in order to measure the quantity of closed loop variation that could, in principle, be eliminated by using MV control. CD control systems are used to control product uniformity in many web forming processes and Section 4 of this paper studies the performance of a system that is used in a plastic film extrusion process. Typically, as illustrated in Figure 1, CD control systems are based on a finished product gauge that is situated down-stream from an array of cross directional actuators and the control objective is to minimise the cross-web product variation that is observed by the gauge. The MV lower bound has already been applied in CD control (Duncan *et al.*, 2000), but this paper offers a more general model of the process output and gives better insight into the measured system performance. Alternative approaches to CD controller performance assessment include applying the concept of bandwidth to the CD spectrum and the two-dimensional (2D) spectrum of the steady state system in order to report spatial and dynamic performance of the controller. These principles are used with wavelet analysis in a system that partitions variation for performance assessment (Jiao *et al.*, 2000). A drawback of using the bandwidth is the level of process and controller knowledge that is required in order to make the performance assessment.

[1] sponsored by the UK Engineering and Physical Sciences Research Council
[2] corresponding author: stephen.duncan@eng.ox.ac.uk

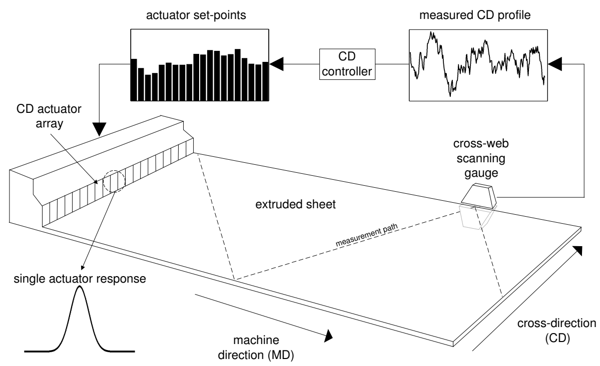

Fig. 1. Schematic diagram of a CD control system.

This paper has three further sections. The CD controller and its MV performance index are defined in Section 2. The Bayesian statistical model and numerical methods that are used to evaluate this model and the performance index are given in Section 3. Finally, in Section 4, a data set from the industrial process is studied.

2. MV ASSESSMENT IN CD CONTROL

Let the CD control system be given by

$$\mathbf{y}_t = q^{-d} g(q^{-1})\mathbf{G}\mathbf{u}_t + \mathbf{C}(q^{-1})\mathbf{e}_t \qquad (1)$$

where $\mathbf{y}_t$ is the $(M \times 1)$ output vector at time t, $\mathbf{e}_t$ is the disturbance vector and $\mathbf{C}(q^{-1})$ is a matrix polynomial giving the disturbance dynamics (Duncan *et al.*, 2000). The actuator set-points are given by the $(N \times 1)$ set-point vector, $\mathbf{u}_t$. It is assumed that the response of the actuator array separates into a scalar dynamic response, which is given by $g(q^{-1})$, and an $(M \times N)$ matrix, $\mathbf{G}$, that describes the steady state spatial response. In this work, the delay structure is assumed to be $q^{-d}\mathbf{I}_N$, where d is the common delay and $\mathbf{I}_N$ is the identity matrix.

The minimum variance lower bound represents the output variation that would be achieved were a minimum variance controller to be used in place of the existing controller. In the case where there are no restrictions on the actuator settings and assuming that both the process and disturbance are stable, the only limitation that is imposed on the performance of a single input single output (SISO) minimum variance controller is that of the process delay. In CD control systems however, the achievable performance is further restricted by uncontrollable spatial variation and this analysis aims to quantify the amount of spatial and dynamic variation that would not exist in the process output, were a minimum variance CD controller employed. In order to distinguish the variation in this way, the uncontrollable spatial variation is not removed prior to derivation of the MV controller.

To solve the minimum variance problem, a d-step ahead prediction is formed in the process output, $\mathbf{y}_t$. Usually, this is achieved by factorising $q^d\mathbf{C}(q^{-1})\mathbf{e}_t$ into predictable and unpredictable dynamic components. However, as mentioned, uncontrollable spatial modes must also be considered in CD control systems and these can be factored out as follows:

$$\mathbf{e}_{c,t} = \mathcal{P}_c\mathbf{e}_t \ ; \ \mathbf{e}_{u,t} = \mathcal{P}_u\mathbf{e}_t \qquad (2)$$

such that $\mathbf{e}_t = \mathbf{e}_{c,t} + \mathbf{e}_{u,t}$. The matrices, $\mathcal{P}_c$ and $\mathcal{P}_u = \mathbf{I} - \mathcal{P}_u$ are M to M projections into the controllable and uncontrollable spatial domains. These are defined from knowledge of the process. For example, the matrix form of the discrete Fourier transform, $\mathbf{F}$ (Stewart, 2000), can be used when the highest controllable spatial frequency is known. This gives $\mathcal{P}_u = \mathbf{F}_u\mathbf{F}_u^T$ and $\mathcal{P}_c = \mathbf{F}_c\mathbf{F}_c^T$, where the columns of $\mathbf{F}$ are divided between $\mathbf{F}_c$ and $\mathbf{F}_u$ according to this frequency. Alternatively, if the spatial interaction matrix, $\mathbf{G}$ is known, the vectors that span its column space give suitable projections (Duncan *et al.*, 2000).

Returning to $q^d\mathbf{C}(q^{-1})$, the d-step ahead prediction of the disturbance, this can be given by

$$q^d\mathcal{P}_c\mathbf{C}_u(q^{-1}) + \mathcal{P}_c\mathbf{C}_c(q^{-1}) + \mathcal{P}_u\mathbf{C}(q^{-1}) \qquad (3)$$

The term, $q^d\mathcal{P}_c\mathbf{C}_u(q^{-1})$ depends on future innovations and its best estimate is $\mathrm{E}[\mathbf{e}_t]$, which is assumed to be zero. The term, $\mathcal{P}_u\mathbf{C}_u(q^{-1})$ is assumed orthogonal to the spatial response, so its d-step ahead prediction is the same expression. This gives the minimum variance controller, $\mathbf{K}_{MV}(q^{-1})$ and when this is applied to the system equation, (1) the process output becomes

$$\mathbf{y}_t(MV) = \left(\mathcal{P}_c\mathbf{C}_u(q^{-1}) + \mathcal{P}_u\mathbf{C}(q^{-1})\right)\mathbf{e}_t \quad (4)$$

such that under minimum variance CD control, the output variation is of the sum of all disturbance variation in uncontrollable spatial modes and the unpredictable dynamic variation in controllable spatial modes. Given a suitable spatial projection matrix, $\mathcal{P}_c$, a closed loop moving average model estimated from the process output, $\mathbf{H}(q^{-1})$, and the error covariance matrix, $\mathbf{\Sigma}_e$, the performance index, $\eta(d)$ is given by

$$\mathbf{\Sigma}_c(MV) = \sum_{j=0}^{d-1} \mathcal{P}_c\mathbf{H}(j)\Sigma_e\mathbf{H}^T(j)\mathcal{P}_c^T \qquad (5)$$

$$\mathbf{\Sigma}_u(MV) = \sum_{t=1}^{T} \mathcal{P}_u\mathbf{y}_t\mathbf{y}_t^T\mathcal{P}_u^T$$

$$\eta(d) = 1 - \frac{\mathrm{trace}(\mathbf{\Sigma}_c(MV) + \mathbf{\Sigma}_u(MV))}{\sum_{t=1}^{T}\mathbf{y}_t\mathbf{y}_t^T}$$

whilst the performance of each CD lane, $\eta(d, m)$ can be extracted from the individual matrix terms.

3. CALCULATING THE INDEX

The performance measure defined above can be estimated through the closed loop vector moving average model

$$\mathbf{y}_t = \mathbf{H}(q^{-1})\mathbf{e}_t \tag{6}$$

where $\mathbf{H}(q^{-1})$ denotes the moving average coefficients and $\mathbf{e}_t$ is usually assumed to be multivariate Gaussian with mean μ_e and variance, $\mathbf{\Sigma}_e$. Under a minimum variance controller, the lag order of $\mathbf{H}(q^{-1})$ is $(d-1)$, but the cross lag order of $\mathbf{H}(q^{-1})$ cannot be determined so directly. To address this problem in this paper, the Bayesian approach to vector autoregression (VAR) modelling is used and this employs a model of the form

$$\mathbf{y}_t = \mathbf{\Theta}(q^{-1})\mathbf{y}_t + \mathbf{e}_t \tag{7}$$

where $\mathbf{\Theta}(q^{-1})$ is a matrix polynomial representing the autoregressive coefficients. The entries, $\theta(i, j, k)$, enumerate the VAR relationship from CD measurement cell j to cell i at lag k. The coefficients can be specified in a number of intuitive ways:

- scalar $\mathbf{\Theta}(q^{-1})$ is the univariate autoregression model (Duncan *et al.*, 2000)
- circulant $\mathbf{\Theta}(q^{-1})$ and $\mathbf{S}_e$ is equivalent to modelling the system through its spatial modes using univariate autoregression models (Stewart, 2000; Duncan *et al.*, 2000)
- Toeplitz $\mathbf{\Theta}(q^{-1})$ and $\mathbf{S}_e$ gives a slightly more realistic representation without the wrap around of the circulant form (Stewart, 2000)
- unconstrained $\mathbf{\Theta}(q^{-1})$ and $\mathbf{S}_e$ gives rise to a difficult estimation problem.

The circulant and Toeplitz models give the case where the VAR coefficients are the same across the sheet. However, by the nature some CD processes, the coefficients of $\mathbf{\Theta}(q^{-1})$ may only be similar within localised regions of the web (Mijanovic, 2004), especially when there is a process fault. The algorithm that is now specified can deal with all these cases, but the more general model is very computationally demanding. Methods that reduce the output dimension, such as principal components analysis (PCA) can also be used in conjunction with this approach (Rigopoulos and Arkun, 2003).

Bayesian methods are used to identify the VAR models because they offer control over the spatial structure of the dynamic model through the prior distributions that are required to perform the analysis. For example, these prior distributions can be used to force improbable, high order regression coefficients, which may be influenced by noise, towards zero. The Minnesota prior that is used in econometric modelling has such a structure in which prior belief is that high order coefficients are very small (Lükepohl, 1993). More sophisticated priors are used in spatial statistics, where distributed models have been developed in which coefficients belonging to neighbouring sites are assumed to be similar. These spatial models are intuitively applicable to the problem studied here. However, a major difficulty of VAR modelling is in model selection. This occurs because there are a huge number of model combinations that can be selected from all possible models and this problem is addressed here by the Reversible Jump (RJ) algorithm, which is outlined below, once the basic procedure for estimating the posterior distribution of the performance index has been described. This uses Bayes Factors, which do not depend on the asymptotic assumptions that are required in many forms of penalised likelihood analysis, such as the Akaike Information Criterion (AIC) (Denison *et al.*, 2002).

Closed form solutions are often not available to the integrals that must be solved in order to perform Bayesian inference. The issue applies here to the normalising divisors of the posterior distribution of the performance index and the VAR structure (7). Markov Chain Monte Carlo (MCMC) methods (Denison *et al.*, 2002) are used here to overcome these difficulties. MCMC is a form of simulation in which the aim is to generate random samples from a target distribution in order to make empirical inferences about that distribution. Two forms of MCMC iteration are used in this paper as follows

Gibbs sampling is carried out by factorising the otherwise analytically intractable posterior distribution into full conditional distributions of parameter subsets that individually have standard forms. Each iteration of the Gibbs sampler consists of a random sample being drawn from each full conditional distribution in turn, given the current iteration values of all other parameters.

Metropolis Hastings (MH) sampling copes with posterior distributions that cannot be factorised as above. Each MH sampler iteration, candidate samples are generated from a proposal distribution and these are accepted as true random samples from the target distribution according to an acceptance probability that depends on the current state of the MCMC sample sequence. The RJ algorithm is a special case of the MH sampler in which the Markov chain transitions can involve parameter vector dimension changes.

Both the above samplers result in a Markov chain whose steady state distribution is the target distribution. In this analysis, a Gibbs sampler is used to sample the parameters, $(\mathbf{\Theta}, \mathbf{\Sigma})$ and the RJ MH sampler is used to sample the VAR lag structure. The performance index, $\eta(d)$ can then be calculated each iteration, based on the current state of the MCMC Markov chain. As already mentioned, this gives rise to a further advantage of the Bayesian approach in estimating the posterior distribution of $\eta(d)$, which can be used to measure uncertainty in the performance. This overcomes an outstanding problem in multivariate control loop performance assessment. However, it is noted that, in a similar manner to the above, the bootstrap method could be used to provide confidence intervals around a least squares, or maximum likelihood estimate of the $\eta(d)$.

The likelihood of the time series model (7) is

$$l(\mathbf{\Sigma}, \mathbf{\Theta}|\mathbf{Y}) = (2\pi)^{-T/2}|\mathbf{\Sigma}|^{-0.5} \qquad (8)$$
$$\times \exp(-0.5\sum_{t=1}^{T} (\ \mathbf{y}_t - \mathbf{X}_t\mathbf{\Theta})^T\mathbf{\Sigma}^{-1}(\mathbf{y}_t - \mathbf{X}_t\mathbf{\Theta}))$$

where $\mathbf{X}_t$ represents the matrix of lagged and cross-lagged observations that forms the basis of the given VAR model at time t (Lükepohl, 1993).

In order to form the Bayesian model, prior distributions must be assigned to the covariance matrix, $\mathbf{\Sigma}$, the VAR parameters, $\mathbf{\Theta}$ and the model order. In this paper, a conjugate, Wishart, prior distribution is assigned to precision, $\mathbf{P}_e$, which is the inverse of covariance matrix. This has parameters q and $\mathbf{R}$ that combine to give the expectation of this distribution. The VAR parameters are assigned a multivariate Gaussian prior with zero mean and variance covariance matrix, $\mathbf{W}$. The entries in $\mathbf{W}$ are used here to represent prior belief that the associated VAR coefficients are similar and it is these weights that determine the effective structure of the resulting VAR model. It is noted that the model equations can be re-expressed in a more compact form for the scalar, circulant and Toeplitz forms of (7). The resulting posterior distribution, $f(\mathbf{P}_e, \mathbf{\Theta}|\mathbf{Y})$, is proportional to

$$|\mathbf{P}_e|^{0.5(T+q)-1}|\Omega|^{0.5}\exp(-0.5\mathrm{trace}\mathbf{P}_e\mathbf{R})$$
$$\times \exp(-0.5\sum_{t=1}^{T}(\mathbf{y}_t - \mathbf{X}_t\mathbf{\Theta})^T\mathbf{P}_e(\mathbf{y}_t - \mathbf{X}_t\mathbf{\Theta}))$$
$$\times \exp(-0.5\mathbf{\Theta}^T\Omega\mathbf{\Theta}) \qquad (9)$$

Although marginal posterior distributions can be derived from this expression for the parameters, these do not have standard forms (Tiao and Zellner, 1964). However, the full conditional distributions for $\mathbf{\Theta}$ and $\mathbf{P}_e$ do have standard forms, which can be used to form a Gibbs sampler (Denison *et al.*, 2002). Let $\mathbf{Y}$ represent the process output stacked into a vector, length $(T \times M)$ and $\mathbf{X}$ denote the corresponding VAR regressor matrix. By manipulating (9), mainly to complete the square in $\mathbf{\Theta}$, the full conditional distribution of $\mathbf{\Theta}$ is $\mathcal{MVN}(\hat{\mathbf{\Theta}}, P_\Theta^{-1})$ where

$$\mathbf{P}_\Theta = (\Omega + \mathbf{X}^T\mathbf{P}_E\mathbf{X})$$
$$\hat{\mathbf{\Theta}} = \mathbf{P}_\Theta^{-1}\mathbf{X}^T\mathbf{P}_E\mathbf{Y} \qquad (10)$$

Further to this, in order to complete the factorisation of the posterior that is required to form a Gibbs sampler, by inspection, the full conditional distribution of $\mathbf{P}_e$ is Wishart with parameters $(T+q)$ and $\mathbf{R}+(\mathbf{Y}-\mathbf{X\Theta})^T\otimes(\mathbf{Y}-\mathbf{X\Theta})$. However, this depends on the VAR model being specified and to overcome this, the problem of model selection is now addressed by forming a MCMC sampler that estimates the posterior distribution of the lag structure in addition to that of the parameters of the model. This important part of the analysis is carried out by using RJ sampling. As already mentioned above, the RJ MH algorithm can deal with parameter vector dimension changes and it can be used to marginalise the distribution of non-trivial random variables, such as the performance index, $\eta(d)$. This is achieved by sampling from the model order posterior distribution in addition to that of the parameters. However, one of the difficulties this method poses is that dummy variables must be set up in order to handle MCMC steps that result in parameter vector dimension changes, and these must have a meaningful deterministic relation to the current state of the Markov chain. In a procedure similar to that given in (Denison *et al.*, 2002), this difficulty is handled by treating the linear parameter vector, $\mathbf{\Theta}$ as a nuisance variable and this is integrated out before the dimension changing step.

Changes to the model structure give rise to changes in the regressor matrix, $\mathbf{X}$, and a prior distribution is now defined over the model space in order to facilitate Bayesian model selection. Let n_ϕ^+ and m_ϕ^+ denote the maximal dynamic and spatial lags of the VAR models to be considered. This two-dimensional lag structure can represented by a $(m_\phi^+ \times n_\phi^+)$ matrix of indicator variables, $\mathbf{Q}$ in which $Q(i,j) = 1$ determines that the coefficient, $\phi(i,j,k)$, is present in the model. By definition, this structure is constrained to be spatially symmetric, where if $\Theta(i,j)$ is present in the model at lag k, then so is $\Phi(-i,j)$. Without any further structural constraints on $\mathbf{Q}$ (constraints on the structure are discussed later) the priors for each entry are assumed to be Bernoulli with $p(Q(i,j,k) = 1) = p_Q\exp(-0.5\sigma_Q^{-2}(i+j)^2)$. The parameters p_Q and σ_Q must be chosen to reflect prior belief, for example higher order models may be considered to be less probable. This gives the overall prior, $p(\mathbf{Q})$

$$\prod_{i=1}^{m_\phi^+}\prod_{j=1}^{n_\phi^+}(p_Q \exp(-0.5\sigma_Q^{-2}(i+j)^2))^{Q(i,j)} \quad (11)$$

and the un-normalised posterior is given by the product (11) $\times$ (8). Terms that do not change with model order can be dropped and it is assumed that the maximal design matrix, $\mathbf{X}$, is formed from the data at the onset, such that T remains the same irrespective of the model order.

For a given model order by (8)-(9), the terms in the posterior that involve the linear parameters have a Gaussian form and by the law of total probability, these can be integrated out to leave the posterior distribution of just $\mathbf{P}_e$ and $\mathbf{Q}$

$$\begin{aligned} f(\mathbf{P}_e, \mathbf{Q}|.) &\propto \int f(\mathbf{\Sigma}, \mathbf{\Theta}, \mathbf{Q}|.)d\mathbf{\Theta} \quad (12)\\ &\propto p(\mathbf{Q})|\Omega|^{\frac{1}{2}}|\mathbf{P}_e|^{\frac{(T+q)}{2}-1}|P_\Theta|^{-0.5}|\Omega|^{0.5}\\ &\times \exp\left\{-0.5(\text{trace}(\mathbf{R}\mathbf{P}_e))\right\}\\ &\times \exp\left\{-0.5(\mathbf{Y}^T\mathbf{P}_E\mathbf{Y} - \hat{\mathbf{\Theta}}^T\mathbf{P}_\Theta\hat{\mathbf{\Theta}}))\right\} \end{aligned}$$

This distribution does not have a standard form and it cannot readily be used in a Gibbs sampler. However, given its form up to the normalising constant, a Metropolis step can be used sample the VAR model order from this distribution.

This work imposes the constraint that only contiguous lag structures exist, thus all possible regression subsets are not considered. This means that if $\theta(i_1, j_1)$ is present in the model, then $\theta(i_2, j_2)$ must also be present for lags within the region $i_2 \leq i_1$ and $j_2 \leq j_1$. This leads to the incremental model order transition kernel defined in Algorithm 1, which adds, or deletes VAR regression coefficients from the model, whilst preserving its contiguous structure. Although this forms an RJ step, the Jacobian for this transformation is unity (Denison *et al.*, 2002).

Algorithm 1 describes the MCMC steps that are used in order to give the required posterior distributions for the performance analysis. Step 1 of this algorithm draws samples from the posterior distribution of the linear parameters. This is a Gibbs sampler step, which samples from the full conditional distribution of the VAR parameters, conditional on the value of the model order and the covariance matrix obtained in the previous iteration of the algorithm. It can be shown that when combined with Step 2, which is a Gibbs step for the covariance, the resulting sequence of samples are random draws from the posterior distribution of the model. In step 4, the performance index and any other statistics of that of concern to the performance analysis are calculated from the current MCMC values of the model parameters. Step 5 is a RJ Metropolis step for the model order. The target distribution for this step was obtained in (10) by integrating the linear parameters from the full model posterior distribution and the proposal distribution used is a random walk within the a priori feasible model space.

Algorithm 1. Posterior distribution of VAR model and $\eta(d)$

Step 1 Initialise the system by setting $\mathbf{P}_e = cov(\mathbf{y} - \mathbf{x}\hat{\theta}_{LS})^{-1}$ where $\hat{\theta}_{LS}$ is the least squares estimate of the parameter vector, $(\mathbf{X}^T\mathbf{X})^{-1}\mathbf{X}^T\mathbf{Y}$.

Step 2 Sample the linear parameters, θ, from their full conditional distribution, $\mathcal{MVN}(\hat{\theta}, P_\theta^{-1})$.

Step 3 Sample $\mathbf{P}_e$ from its Wishart full conditional posterior with parameters $(q+T)$ and $\mathbf{R} + (\mathbf{Y} - \mathbf{X}\theta)^T \otimes (\mathbf{Y} - \mathbf{X}\theta)$.

Step 4 Invert the VAR model to VMA(∞) form, giving $\mathbf{H}(q^{-1})$. Calculate $\eta(d)$ using (5); the VMA model; the projection matrices, $\mathcal{P}_u$ and $\mathcal{P}_c$; and the error covariance $\mathbf{\Sigma}_e = \mathbf{P}_e^{-1}$. Further statistics can be collected for analysis as required, for example the 2D spectrum of the process output can be estimated through the VMA model.

Step 5 Metropolis Hastings sampler step for the posterior distribution of the lag structure, $\mathbf{X}$. This is based on (10).

Step 6 Once samples have converged to steady state, collect for analysis.

Step 7 Repeat steps 2 to 6 until the estimated posterior distributions, $f(\eta(d)|.)$ and $f(\mathbf{X}|.)$ converges.

4. RESULTS

The performance index is now used to assess the CD control system that is used in an industrial plastic film process. This process extrudes polymer feedstock to form a continuous amorphous sheet, which is then drawn in a two-stage process, firstly along its machine direction and then in the cross-direction. The bi-axially drawn finished product is gauged just before it is wound onto rolls at the end of the process. The system delay for the data set considered is 4 scans and the output data is plotted as a topological map in Figure 2(a). The results below have been obtained by using the Toeplitz VAR form of (7) with an unconstrained covariance matrix. The performance analysis is summarised in Table 1. From this, about 11% of observed variation could be eliminated by a MV controller. This suggests that the performance of this system is reasonable. By contrast, a separate analysis based on the known controller settings

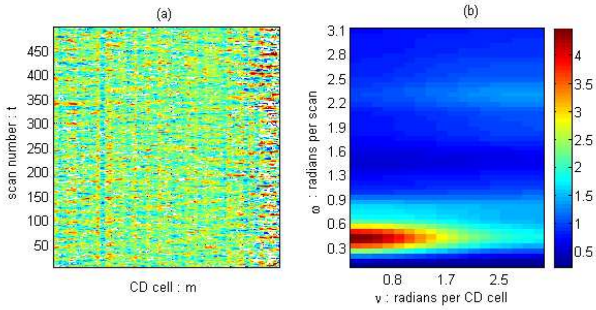

Fig. 2. CD profile data: (a), plotted as a colourmap and (b) the posterior mean 2D power spectrum.

Table 1. Posterior mean statistics.

	parameter estimate
Total variation - MSE(y)	0.46
Uncontrollable spatial variation	32%
Unpredictable dynamic variation	60%
Predictable dynamic variation	6%
Controllable steady state profile	2%
$\eta(d)$	0.11

of this system shows that just under 3% of the observed variation lies inside the 2D bandwidth of this system (Stewart, 2000). The variation that is responsible for this difference in performance is mainly dynamic variation and this can be seen in Figure 2(b), which plots the posterior mean 2D power spectrum of the closed loop output. The main peak in this chart is outside the very low dynamic bandwidth of the controller. Although the performance is good, the relatively high level of "controllable" dynamic variation that can also be seen in the entries of Table 1 could be due to the controller behaving more aggressively than optimal and further investigation of the controller settings has been recommended.

The posterior distribution of $\eta(d)$ is given in Figure 3(a). Based on these MCMC results, this distribution is concentrated over a relatively small range of values indicating a relatively high level of certainty in the value of $\eta(d)$. However, under different conditions, for example when the performance is not so good, more variation is observed in these values. Figure 3(b) gives the posterior distribution of the VAR lag structure and it can be seen that in this case, the dynamic variation is concentrated at spatial lag zero.

5. CONCLUSIONS

This paper presents a method for analysing the performance of CD control systems. Similar principles are used to the methods considered previously in (Duncan *et al.*, 2000). However, this paper offers a more general approach to estimating the closed loop filter that is required in order to make the performance assessment and further insight is gained with respect to the reported performance. This is illustrated using data from a plastic film process and further analysis of the system has been recommended in order to explain the underlying variation that was identified. The Bayesian numerical approach of MCMC overcomes the problem of VAR model order selection by integrating over the posterior distribution of the model space and a by-product of this is the posterior distribution of the performance index.

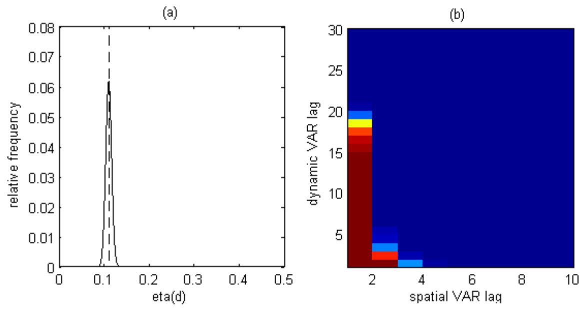

Fig. 3. Posterior distributions of the performance statistic, $\eta(d)$ and the VAR lag structures..

REFERENCES

Denison, D.G.T, C.C. Holmes, B.K. Mallick and A.F.M. Smith (2002). *Bayesian Methods for Nonlinear Regression and Classification*. Wiley. Chichester, UK.

Duncan, S.R., G. Dumont and D. Gorinevsky (2000). Performance monitoring for cross-directional control systems. *Proc. PAPTAC Control Systems Conf., Victoria, BC, Canada* pp. 173–179.

Jiao, X., M. Davies and G. Dumont (2000). Wavelet packet analysis of paper machine data for control assessment and trim loss optimization. *Proc. PAPTAC Control Systems Conf., Victoria, BC, Canada* pp. 179–182.

Lükepohl, H. (1993). *Introduction to Multiple Time Series Analysis*. 2nd ed.. Springer-Verlag. New York, NY.

Mijanovic, S. (2004). Paper Machine Cross-Directional Control Near Spatial Domain Boundaries. PhD Thesis. University of British Columbia. Vancouver, Canada.

Rigopoulos, A. and Y. Arkun (2003). Reduced order cross-directional controller design for sheet forming processes. *IEEE Trans. Contr. Syst. Technol.* **11**(5), 746–756.

Stewart, G.E. (2000). Two Dimensional Loop Shaping Controller Design for Paper Machine Cross-Directional Properties. PhD Thesis. University of British Columbia. Vancouver, Canada.

Tiao, G.C. and A. Zellner (1964). On the Bayesian estimation of multivariate regression. *J. R. Statist. Soc. B* **26**(2), 277–285.