MULTIVARIATE CONTROL LOOP PERFORMANCE ASSESSMENT FOR CROSS-DIRECTIONAL CONTROL

A.R. Taylor and S.R. Duncan*
Department of Engineering Science,
University of Oxford, Oxford, UK
* author for correspondence stephen.duncan@eng.ox.ac.uk

ABSTRACT

Control loop performance assessment analyses routine operating data to identify poorly performing control systems. Such systems can then be optimised in order to improve the productivity of the plant. This paper uses vector autoregression (VAR) models and the principle of the minimum variance lower bound to assess the performance of a cross-directional (CD) control system on a plastic film process.

INTRODUCTION

With the data that is now available from most control systems and the constant drive towards increased productivity in industry, there is considerable interest in methods that assess closed loop performance based on routine operating data. The goal of such analysis is to identify systems that are not performing well in order that corrective measures can be taken to improve the situation. In practice however, there are many factors that can influence the measured performance of a control system and the performance assessment often serves as a filter to eliminate control loops that do not require attention.

One of the most widespread approaches to control loop performance assessment is to apply the minimum variance (MV) lower bound to routine operating data. A minimum variance approach is used in this paper and this gives a statistic that measures the quantity of closed loop variation that could, in principle, be eliminated by a minimum variance controller. The minimum variance lower bound has already been applied to cross-directional (CD) control systems in [1], but this paper takes a more general approach to modelling the process output for the performance assessment and gives more insight into the measured system performance.

Alternative approaches to CD controller performance assessment include applying the spatial bandwidth to the cross-directional spectrum. Further to this, the two dimensional bandwidth can be used to report spatial and dynamic performance of the system. These principles are used in a system that uses wavelet analysis to partition variation for performance assessment [2]. A drawback of this approach is the level of process and controller knowledge that is required in order to make the performance assessment.

The rest of this paper is split into three sections. In the first instance, the performance index to be used is defined. The index is then discussed and a statistical model is introduced in order to facilitate its calculation. This is followed by an example in which two data sets from a plastic film process are studied. Finally, the main results of the paper are summarised.

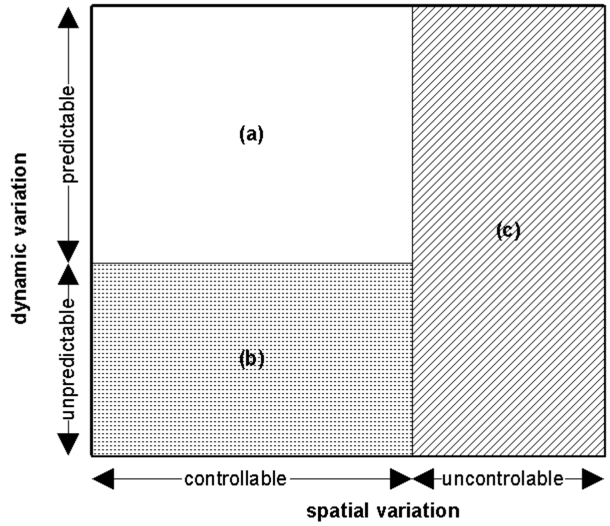

Figure 1 – partitioning the observed closed loop CD profile variation into "controllable" and "uncontrollable" constituents.

CD CONTROL MINIMUM VARIANCE ASSESSMENT

The minimum variance lower bound represents the output variation that would be achieved were a minimum variance controller used in place of the existing controller. In the case when there are no restrictions on the actuator settings, assuming that the process and disturbance transfer functions 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.

This analysis aims to identify components of variation that exist in the process output of a CD controller with respect to the minimum variance lower bound. These components are illustrated in Figure 1. In order to distinguish the variation in this way, the uncontrollable spatial variation is not removed prior to derivation of the MV controller. 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 $(M\times1)$ output vector, $\mathbf{e}_t$ is the disturbance and $\mathbf{C}(q^{-1})$ is a matrix polynomial that describes the disturbance dynamics. The actuator set-points are given by the $(N\times1)$ set-point vector, $\mathbf{u}_t$. It is assumed that the response of the actuator array to these set-points separates into a scalar dynamic response, which is given by $g(q^{-1})$, and a steady state spatial response, $\mathbf{G}$ $(M\times N)$. The delay structure of this system is assumed to be $q^{-d}\mathbf{I}$, where d is the common delay and $\mathbf{I}$ is the $N\times N$ identity matrix.

To solve the minimum variance problem, a d-step ahead prediction is formed in the process output. Usually, $q^d\mathbf{C}(q^{-1})\mathbf{e}_t$ is factorised into predictable and unpredictable dynamic components. However, for CD control, in addition to the unpredictable dynamic variation, there are uncontrollable spatial modes, which can be factored out as follows

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

where $\mathbf{e}_t$=$\mathbf{e}_{c,t}$+$\mathbf{e}_{u,t}$. The matrices, $\mathbf{P}_c$ and $\mathbf{P}_u$=$\mathbf{I}$-$\mathbf{P}_c$ 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, $\mathbf{X}$ [3], can be used when the highest controllable spatial frequency is known. This gives $\mathbf{P}_u$= $\mathbf{X}_u\mathbf{X}_u$' and $\mathbf{P}_c$=$\mathbf{X}_c\mathbf{X}_c$', where the columns of $\mathbf{X}$ are divided between $\mathbf{X}_c$ and $\mathbf{X}_u$ according to this frequency. Alternatively, if the spatial interaction matrix $\mathbf{G}$ is known, the vectors that span its column space give suitable projections [1].

Returning to the d-step ahead prediction of the disturbance

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

The term $q^d\mathbf{P}_c\mathbf{C}_u(q^{-1})$ depends on future innovations and its best estimate is E[$\mathbf{e}_t$], which is assumed to be zero. The term $\mathbf{P}_u\mathbf{C}_u(q^{-1})$ is assumed orthogonal to the spatial response, so its d-step ahead prediction is the same expression to give the controller $\mathbf{K}_{MV}(q^{-1})$ as follows

$$-\mathbf{B}^{-1}(q^{-1})\left[\mathbf{I} - q^{-d}\mathbf{P}_c\mathbf{C}_c(q^{-1})\mathbf{C}^{-1}(q^{-1})\right]^{-1}\mathbf{P}_c\mathbf{C}_c(q^{-1})\mathbf{C}^{-1}(q^{-1}) \qquad (4)$$

Replacing this expression in the system equation gives

$$\mathbf{y}_t(\mathrm{MV}) = (\mathbf{P}_c\mathbf{C}_u(q^{-1}) + \mathbf{P}_u\mathbf{C}(q^{-1}))\mathbf{e}_t \qquad (5)$$

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. This partitioning of variance is illustrated in Figure 2. No variation exists in area (a) of the figure for the output of a minimum variance system.

Given a suitable spatial projection matrix, $\mathbf{P}_c$, a closed loop moving average model estimated from the process output, $\mathbf{H}(q^{-1})$, and the error covariance matrix, Σ_e, the performance index η(d) is given by

$$\mathbf{\Sigma}_c(\mathrm{MV}) = \sum_{j=0}^{d-1}\mathbf{P}_c\mathbf{H}(j)\mathbf{\Sigma}_e\mathbf{H}^{\mathrm{T}}(j)\mathbf{P}_c^{\mathrm{T}} \qquad (6)$$
$$\mathbf{\Sigma}_u(\mathrm{MV}) = \sum_{t=1}^{T}\mathbf{P}_u\mathbf{y}_t^{\mathrm{T}}\mathbf{y}_t\mathbf{P}_u^{\mathrm{T}}$$
$$\eta(d) = 1 - \frac{\mathrm{trace}(\mathbf{\Sigma}_c(\mathrm{MV}) + \mathbf{\Sigma}_u(\mathrm{MV}))}{\sum_{t=1}^{T}\mathbf{y}_t\mathbf{y}_t^{\mathrm{T}}}$$

whilst the performance of each CD lane, $\eta(d,m)$ can be extracted from the individual matrix terms. This is in keeping with the original concept of minimum variance assessment, because the performance is assessed subject to a minimal number of assumptions.

From the above, measuring the performance of the system reduces to the problem of estimating the closed loop filter coefficients, $\mathbf{H}(q^{-1})$ and the error covariance matrix, Σ_e .

CALCULATING THE PERFORMANCE INDEX

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

$$\mathbf{y}_t = \mathbf{H}(q^{-1})\mathbf{e}_t \qquad (7)$$

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, Σ_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 from this analysis. To address this problem, the Bayesian approach to vector autoregression (VAR) modelling is used in this paper to calculate the performance index. The form of the model used is

$$\mathbf{y}_t = \mathbf{\Phi}(q^{-1})\mathbf{y}_t + \mathbf{e}_t \qquad (8)$$

where $\Phi(q^{-1})$ is a matrix polynomial representing the autoregressive coefficients. This is a general model form in which each entry, $\Phi(i,j,k)$, gives the VAR relationship from CD cell j to CD cell i at lag k. In CD control, the coefficients can be specified in a number of intuitive ways:

- scalar $\Phi(q^{-1})$ is the univariate autoregression model [1]
- circulant $\Phi(q^{-1})$ and Σ_e is equivalent to modelling the system through its spatial modes using univariate autoregression models [1,3]
- Toeplitz $\Phi(q^{-1})$ and Σ_e gives a slightly more realistic representation without the wrap around of the circulant form [3]
- unconstrained $\Phi(q^{-1})$ and Σ_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 of the process studied here, the values of $\Phi(q^{-1})$ may only be similar within localised regions on the web when there is a process fault. The algorithm that is now specified is able to deal with all these cases, but the more general model is very computationally demanding. Further to this, the approach can be extended by using methods that reduce the output dimension in CD control, such as orthogonal polynomials and principal components analysis (PCA).

The approach used here was chosen because it offers control over the spatial structure of the dynamic model through the prior distributions that are required to perform the analysis. The 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 [4]. 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. This problem is addressed here by employing 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.

MCMC steps to sample the posterior distribution of $\eta(d)$	
Initialisation	Initialise model order to a random value from the *a priori* feasible range. Initialise the covariance, Σ_e, and model parameter values, Φ, with the least squares estimates.
Step1	Sample the value of Φ from its full conditional distribution given the current covariance and model order.
Step2	Sample the value of Σ_e from its full conditional distribution given the linear parameters, Φ, and the model order.
Step3	Sample the model order conditional on the covariance using a MH step.
Step4	Calculate the performance index, $\eta(d)$, given Φ and Σ_e from above. Also calculate any other statistics of interest for posterior analysis. For example, the power spectrum is obtained from impulse response analysis.
Repeat steps 1 to 4 until the values converge in distribution. After this point, the sequences of sampled values can be shown to be random draws from their respective posterior distributions and as a result, these can be used in statistical inference.	

Algorithm 1 – MCMC sampler for the VAR model and performance index posterior distribution.

An advantage of the Bayesian approach is that it estimates the posterior distribution of the performance metric. This measures uncertainty in the performance given the output data and overcomes an outstanding problem in multivariate control loop performance assessment. However, the approach taken is not necessarily unique to Bayesian inference, because the bootstrap could be used to provide confidence intervals through least squares or maximum likelihood estimation of the performance index.

In order to form the Bayesian model, prior distributions must be assigned to the covariance matrix, the VAR parameters 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, Ω. The entries in Ω can be used to represent the prior belief that the corresponding VAR coefficients are similar. It is these weights that determine the effective structure of the resulting VAR model, although the model equations can be re-expressed in a more compact form for the scalar, circulant and Toeplitz forms of (8).

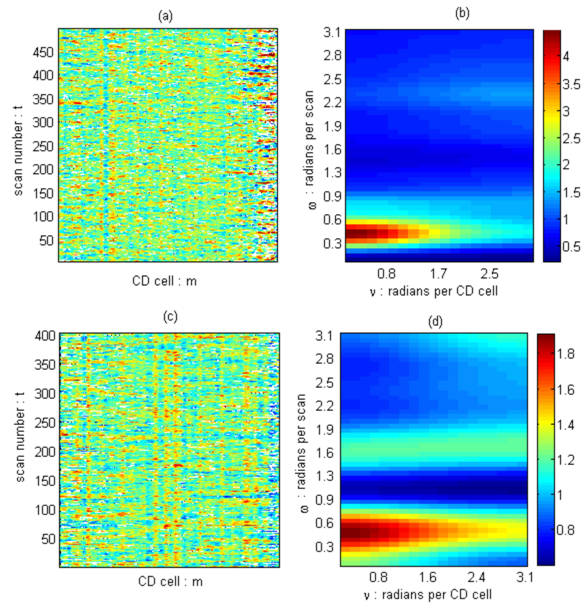

Figure 2 – colourmap plots of raw CD profile data and estimated 2D power spectra. Data set A is given by (a) and (b) and data set B by (c) and (d).

In this paper, Markov Chain Monte Carlo (MCMC) and the Reversible Jump (RJ) algorithm are used to estimate the posterior distributions of the model parameters conditional on these priors [5]. These are numerical integration methods that are used in Bayesian statistics to overcome the problem of intractable integration. A full explanation of MCMC and the algorithm that is used here is out of the scope of this brief paper. As a result, interested readers are referred to [5] and the references therein.

Algorithm 1 describes the MCMC steps that are used in order to give the required posterior distributions for the performance analysis. These steps are now briefly explained. Step 1 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. Step 3 is a RJ Metropolis step for the model order. The target distribution for this step is obtained by intergrating 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. MCMC schemes for similar models are discussed in more detail in [5].

RESULTS

The performance index is now evaluated for two data sets taken from a plastic film line. The process has been described in an earlier paper in this volume by these authors. Let these data sets be referred to as data sets A and B respectively. The system delay for data set A is 4 scans and for data set B, the delay was 3 scans. The data is plotted in Figure 2.

	Data A	Data B
Total variation - *MSE(y)*	0.46	0.42
Uncontrollable spatial variation	32%	24%
Unpredictable dynamic variation	60%	66%
Predictable dynamic variation	6%	1%
Controllable steady state profile	2%	8%
$\eta(d)$	0.11	0.11

Table 1 – summary of CD control system performance over the three data sets.

The performance analysis is summarised in Table 1 and the posterior distribution of η(d) is given in Figure 3. This has been obtained using on a Toeplitz VAR model structure. 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 shows that just under 3% of the variation in data set A lies inside the 2D bandwidth of this system [3]. 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 for data set A. The main peak in this chart is outside the very low dynamic bandwidth of the controller.

As well as reporting the posterior mean value of the performance index, $\eta(d)$, Table 1 partitions variation into predictable / unpredictable dynamic and controllable / uncontrollable spatial components and some differences between the two data sets are observed:

- the controller may be behaving more aggressively in data set A. It is suspected that there is more “controllable” dynamic variation present, because the controller is amplifying naturally occurring process variation, possibly due to an unknown disturbance.
- in data set B, it is known that the controller gain was lower and this may explain the reduced levels of “controllable” dynamic variation and the lower VAR model orders in Figure 3(c). However, the controllable DC profile variation is higher.

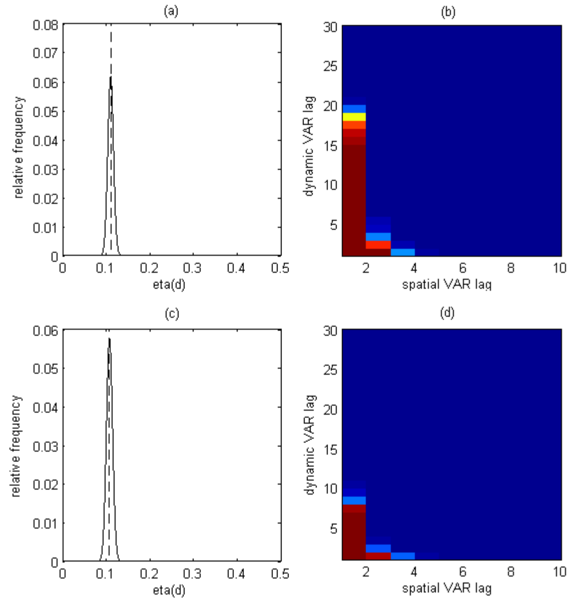

Figure 3 – posterior distributions of the performance statistic, η(d), and the VAR model lag structures.

CONCLUSIONS

This paper presents a method for analysing the performance of CD control systems. Similar principles are used to the methods considered in [1]. However, this paper presents 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. It is interesting to note that although the performance of the system is very similar across these data sets, significant differences can be seen in the detailed components of variation analysis. Further analysis of the system has suggested some plausible explanations for these differences, but further work is required to fully understand the variation observed. A by-product of the approach used is that the posterior distribution of the performance index is evaluated.

REFERENCES

[1] S.R. Duncan, G.A. Dumont and D.M. Gorinevsky. Performance monitoring for cross-directional control systems. Proc. Control Systems Conference 2000, Pulp and Paper Technical Association of Cananda, Victoria, BC, Canada.

[2] X. Jiao, M. Davies and G. Dumont. Wavelet packet analysis of paper machine data for control assessment and trim loss optimisation. Proc. Control Systems Conference 2000, Pulp and Paper Technical Association of Cananda, Victoria, BC, Canada.

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

[4] H. Lukepohl. Introduction to Multiple Time Series Analysis. Springer-Verlag, Berlin, 2^{nd} edition, 1993.

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