# Applied Stochastic Control: A Forty-Year Long Paper Trail

**Guy A. Dumont**

*This paper is dedicated to Karl Johan Åström, whom I wish to thank for the inspiration he has provided me for the last 25 years.*

***Abstract*** This paper rapidly reviews the development of stochastic control theory and its close relationship with the paper industry. Techniques such as maximum-likelihood identification, minimum-variance control, and the self-tuning regulator, all developed by K. J. Åström and his colleagues were first implemented in the Swedish paper industry. Further work on dual adaptive control, auto-tuning and control loop performance monitoring is also briefly reviewed in that context.

## 1. Introduction

The 1960's and 1970's saw rapid development in stochastic control theory spearheaded by K. J. Åström and his colleagues. That period saw the development of maximum-likelihood identification, minimum-variance control and the self-tuning regulator. A remarkable aspect of this work is that industrial applications motivated most of this work, and generally immediately followed the theoretical development. The paper industry was at the forefront of that work. Numerous extensions of that work took place over the last decades, with development of dual adaptive controllers, PID auto-tuners and control performance monitoring tools, just to name a few.

After giving some background information on the papermaking process, this paper first reviews the development of that period. It then describes three areas those development have impacted upon, namely dual adaptive control, PID auto-tuning and control loop performance monitoring, briefly reviewing the state of the art and mentioning some open problems.

## 2. Paper Machine Control

Åström's interest in paper machine control started in early 1963 while he was a young researcher at the IBM Nordic laboratory in Lindingö, Sweden. There he got involved in a study of the feasibility of applying computer control to the paper-making process that was to have profound repercussions in the fields of system identification, stochastic control and adaptive control. Before we go any further, it is necessary to provide the neophyte with a brief overview of the papermaking process. Figure 1 depicts a cartoon-like representation of the paper machine. After cleaning and screening, thin stock (a mixture consisting of about 99% water and 1% fibre) is fed to the headbox. The headbox is essentially a pressurized tank whose purpose is to tranform a circular flow into a thin linear one of controlled velocity for delivery of the stock on a moving wire. In the first few meters of wire, the bed of stock drains through the wire mesh, leaving a mat of fibres that increases in thickness until there is no more free stock on top of it available to drain. This corresponds to the so-called dry line, clearly visible on single-wire, or fourdrinier machines. At that point, the dry-weight of the paper sheet is pretty much set, as it is not possible to further move the fibres around. The main remaining task is to remove the residual water from the wet web, first through pressing between two rolls, then by contact with a large number of steam-heated dryer rolls. Finally, the thickness and surface proerties are affected by the calender stack. Before it is wound up on a reel, the produced paper sheet is scanned by sensors such as a beta-ray gauge to measure basis weight; an infrared gauge to measure the sheet residual moisture content (typically around 5–6%), and a magnetic reluctance or laser gauge to measure caliper (i.e., thickness). One generally distinguishes between machine-direction (MD) control, i.e. in the direction that the paper moves, and cross-direction (CD) control, i.e. across the width of the paper machine. CD requires an array of actuators distributed across the width of the machine. It will not be discussed in details here. For MD control of basis weight, the actuation point is the thick stock valve, that controls the fibre concentration in the thin stock, while for MD control of moisture, the





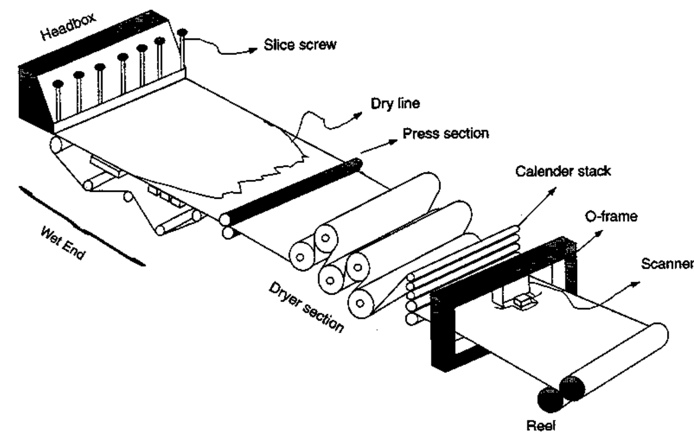


**Figure 1** A paper machine

actuation is usually performed on the steam pressure on one of the last dryer-roll sections. Because the thick stock valve is located before screening and cleaning, there is a much more significant dead time between the actuation and the measurement points for the basis weight control loop than there is for the moisture loop.

In the 1960–70's, Åström was very involved in developing first-principles, dynamic models of paper machines. This work is summarized in a set of lecture notes by Åström (1973). A most interesting study of that time is presented inÅström and Häggman (1974). The authors showed by performing identification experiments on paper machines that the retention of fibres on the wire has to depend on basis weight. While deriving a lumped parameter model, Åström showed that the average retention $r$ has to be characterized by its harmonic mean given by

$$r(w) = w / \int_0^w \frac{1}{r_h(x)} dx \tag{1}$$

where $r_h$ is the instantaneous retention and $w$ is the dry weight of the sheet. This work is relevant to current preoccupations with modelling and controlling wet-end retention, specially on closed-cycle paper machines. There is indeed a significant amount of current work directed at a better understanding of the dynamics of the wet end of the paper machine with the aim of improving the control of the retention of the various components of the thin stock such as fibres, fines and fillers. This renewed interest in the dynamic modelling of the paper machine is also motivated by the realization that increasing the bandwidth of machine-direction can only be accomplished by a clever integration of process and control design.

## 3. Breakthroughs

The ten-year period from 1963 to 1973 would see an amazing chain of breakthroughs achieved by Åström and his colleagues. Early on in the IBM project, it was decided to attempt the development of a systematic procedure to design the controller for the process during normal operation using linear stochastic control theory, that could be adopted for general use.

A linear, continuous dynamics system whose input $u(t)$ and output $y(t)$ are sampled at a time interval $h$ is considered. Because of linearity, the disturbance can be modelled as an additive signal $d(t)$ on the plant output. Also, assume that a time-delay $T_d = (d-1)h$ exists between the input and the output. At the sampling instants, this system can then be described as

$$y(t) = \frac{B_1(q)}{A_1(q)}u(t) + w(t) \tag{2}$$

where $q$ is the forward-shift operator, i.e. $qy(t) = y(t+1)$, and the relative $\deg A - \deg B = d$ corresponds to the delay. Åström showed that if the disturbance $w(t)$ is a stationary random process with a rational spectral density, then it can be represented by

$$w(t) = \frac{C_1(q)}{A_2(q)}e(t) \tag{3}$$

where $e(t)$ is a sequence of independent, equally distributed, zero-mean random variables with variance $\sigma^2$. Åström also showed that $C_1$ and $A_2$ can always be chosen so that they have no roots outside the unit disc. This system is then rewritten as

$$A(q)y(t) = B(q)u(t) + C(q)e(t) \tag{4}$$

with $A = A_1A_2$, $B = B_1A_2$ and $C = C_1A_1$. Before designing a controller, a model of the plant in the above form has to be determined. This is the system identification problem that was first solved by Åström and Bohlin.





### Maximum-Likelihood Identification

Before designing a controller, it is necessary to build an ARMAX model of the plant. For the model 4, the least-squares algorithm gives consistent estimates of $A$ and $B$ if and only if $C = 1$. Building on previous work by Galtieri (Galtieri (1964)), Åström and Bohlin (1965) proposed to use the maximum-likelihood estimate to identify parameters of $A$, $B$ ad $C$, given a sequence of input-output pairs $\{u(t), y(t), t = 1, \cdots, N\}$. In particular, they showed that maximizing the likelihood function $L$ is equivalent to minimizing the loss function

$$J(\theta) = \frac{1}{2}\sum_{t=1}^{N} \varepsilon(t)^2 \tag{5}$$

where $\varepsilon$ are the residuals obtained as

$$C(q)\varepsilon(t) = A(q)y(t) - B(q)u(t) \tag{6}$$

Note that $\varepsilon$ is a one-step ahead prediction error, and can be seen as a reconstruction of the white noise sequence $e(t)$. $J(\theta)$ is convex in the coefficients of $A$ and $B$ but not in those of $C$ and thus it has to be minimized numerically, for instance using the Newton-Raphson gradient method, as proposed by Åström and Bohlin (1965). The interpretation of $\varepsilon$ as a prediction error has led to the development of prediction error methods for identification, see Ljung (1987), now the backbone of modern system identification theory for single-input, single-output systems. Indeed, the maximum-likelihood method has been proven to be both consistent and efficient, i.e. it gives the most accurate, unbiased estimates. In a now classical paper (Åström (1967)), Åström then applied it to identify the dynamics of the basis weight, dry weight and moisture loops on a production paper machine. It is interesting to note that in the same issue of the IBM Journal, a paper by Dahlin (Dahlin (1967)) also dealt with the identification of the paper machine dynamics, using classical techniques based on bump tests.

### Minimum Variance Control

Once an ARMAX model of the plant has been identified, the next step is to proceed to the design of the controller. Concentrating on the regulation problem, Åström went on to develop a design procedure for the minimum-variance controller. The key to the derivation of the minimum-variance controller was to rewrite the ARMAX model of the plant in the so-called predictor form. Omitting the argument $q$ in the various polynomials, the output at time $t+d$ can be written as:

$$y(t+d) = \frac{B}{A}\, u(t+d) + \frac{C}{A}\, e(t+d) \tag{7}$$

The second term in the right-hand side of the above equation consists of noise terms which at time $t$ are future and unknown, and of some past and present terms. Because the noise $e(t)$ is white, those future noise terms cannot be predicted. The impulse response of the noise filter can be used to separate the future unknown terms from the past and present ones

$$\begin{aligned}\frac{C}{A}\, e(t+d) &= \underbrace{e(t+d) + f_1\, e(t+d) + \ldots + f_{d-1}\, e(t+1)}_{\text{future unknown terms}} \\ &+ \underbrace{f_d\, e(t) + \cdots}_{\text{present and past terms}}\end{aligned}$$

This can be rewritten as

$$\frac{C}{A} q^{d-1}\, e(t+1) = Fe(t+1) + \frac{qG}{A}\, e(t) \tag{8}$$

i.e., $F$ and $G$ satisfy the following Diophantine equation with deg $F = d-1$ and deg $G = n-1$:

$$q^{d-1} C(q) = A(q)F(q) + G(q) \tag{9}$$

Reconstructing the noise sequence $e(t)$ as

$$e(t) = [Ay(t) - Bu(t)]/C) \tag{10}$$

the $d$-steps ahead predictor becomes:

$$y(t+d) = \frac{qG}{C} y(t) + \frac{qBF}{C} u(t) + F\, e(t+1) \tag{11}$$

The predictor can be written as

$$y(t+d) = \hat{y}(t+d|t) + Fe(t+1)$$

If $y_r$ denotes the setpoint, then the minimum variance controller is obtained by setting

$$\hat{y}(t+d|t) = y_r$$





thus giving the controller:

$$u(t) = \frac{C(q)}{B(q)F(q)}\, y_r - \frac{G(q)}{B(q)F(q)}\, y(t) \tag{12}$$

Note that when minimum-variance control is used, then

$$\begin{aligned} y(t+d) &= Fe(t+1) \\ &= e(t+d) + f_1 e(t+d-1) + \ldots + f_{d-1} e(t+1)\end{aligned}$$

*i.e.,* the output is a moving-average process of order $d-1$. A characteristic of such a process is that its autocorrelation function vanishes after $d$ lags. This is an important feature of minimum-variance control, which will prove very useful to test the optimality of a stochastic control scheme.

Åström then went on to apply minimum-variance controllers to the dry-weight and moisture control loops on the previously identified paper machine. Results indicated that minimum variance was indeed achieved, roughly halving the dry-weight and moisture standard deviations compared with the situation prior to the installation of those controllers. This, I believe was the first succesful attempt at computer control of a paper machine, before commercial systems became widely available. It is also worthy to note that even today those commercial systems are actually based on more conservative design techniques, such as detuned Dahlin controller or Smith predictor, and that none of those systems is actually even attempting to solve the regulation problem. In that sense, Åström's work on dry weight and moisture control remains very contemporary and can still be considered a benchmark in the paper industry.

Åström summarized the findings of this project in a now classic book, Åström (1970). Of course, Åström was very much aware of the deficiencies of the minimum-variance controller, namely its hyperactivity and its sensitivity to parameter variations. The latter realization in particular was to lead to the development of the first practical adaptive control scheme, the self-tuning regulator.

## Self-Tuning Control

Åström and Wittenmark's 1973 (Åström and Wittenmark (1973)) seminal paper on self-tuning regulator is a model of elegance and simplicity, while it managed to revolutionize the field of adaptive control. The scheme builds on the idea of combining a recursive least-squares estimator and a controller first proposed by Kalman (1958). The key idea is to reparameterize the plant model in terms of the controller parameters using the predictor form derived above, but rewritten for $y(t)$ and in terms of polynomials in the backward shift operator $q^{-1}$ as

$$y(t) = \frac{1}{C}\left(Gy(t-d) + BFu(t-d)\right) + Fe(t) \tag{13}$$

or

$$y(t) = S(q^{-1})y_f(t-d) + R(q^{-1})u_f(t-d) + F(q^{-1})e(t) \tag{14}$$

with $\deg R = n+d-1$ and $\deg S = n-1$ and

$$y_f(t) = \frac{y(t)}{C(q^{-1})} \qquad u_f(t) = \frac{u(t)}{C(q^{-1})}$$

In this predictor model, the controller polynomials $R$ and $S$ appear directly since the control law

$$R(q^{-1})u(t) = -S(q^{-1})y_f(t) \tag{15}$$

gives $y(t) = Fe(t)$, i.e. minimum-variance control.

Then if $1/\hat{C}(q^{-1})$ is a stable filter, it is possible to get estimates $\hat{R}$ and $\hat{S}$ from

$$y(t+d) = S(q^{-1})y_f(t) + R(q^{-1})u_f(t) + \varepsilon(t+d)$$

where

$$y_f(t) = \frac{y(t)}{\hat{C}(q^{-1})} \qquad u_f(t) = \frac{u(t)}{\hat{C}(q^{-1})}$$

using a simple recursive least-squares algorithm with the regressor $\varphi$ and the parameter estimate

$$\varphi^T(t) = \frac{1}{\hat{C}}\left[u(t)\dots u(t-n-d+1)y(t)\dots y(t-n+1)\right]$$
$$\theta^T = [r_0 \dots r_{n+d-1} s_0 \dots s_{n-1}]$$

Because the residual $\varepsilon = Fe$ is uncorrelated with the regressor $\varphi$, simple least-squares can converge to the true parameters. To complete the scheme, it then suffices to apply the control law.

$$\hat{R}(q^{-1})u(t) = -\hat{S}(q^{-1})y(t)$$





Although this scheme seemingly requires knowledge of the $C$-polynomial, the remarkable fact is that the controller will converge to the minimum-variance controller even when $\hat{C} \neq C$ is used to generate $y_f$ and $u_f$. This is known as the self-tuning property. It is easy to add feedforward and command signals

$$y(t+d) = S(q^{-1})y_f(t) + R(q^{-1})u_f(t) + S_{ff}(q^{-1})v_f(t) + \varepsilon(t+d)$$

where $v_f$ is the filtered measured disturbance.

The controller is then

$$\hat{R}(q^{-1})u(t) = -\hat{S}(q^{-1})y(t) - \hat{S}_{ff}(q^{-1})v(t)$$

By replacing $d$ by $h > d$, we obtain a self-tuning moving-average controller. By sufficiently large $h$, no zeros are cancelled, and non-minimum phase systems can be controlled.

The impact of this work on the adaptive control field cannot be overstated. It revitalized the field, which by then was largely concerned with model-reference adaptive control. Of course, the STR has severe shortcomings, such as lack of tuning parameters (!?), hyperactivity, lack of robustness, etc... However, those led several researchers to propose various improvements, making adaptive control the subject of intense research for the following 12-15 years. However, Åström and his colleagues moved rapidly to demonstrate the applicability of the proposed STR to industrial control problems, for control of ore crushers in the cement industry, Borisson and Syding (1973), and once again for control of basis-weight and moisture on paper machines, Borisson and Wittenmark (1974), Cegrell and Hedqvist (1973). An application of the self-tuning regulator to titanium dioxide kilns, Dumont and Bélanger (1978) would stay in continuous use from 1977 to 1991, to be replaced by an adaptive controller of a newer generation based on a representation of the plant dynamics by a truncated Laguerre series, Dumont (1992). In fact, those particular kilns have been under adaptive control continuously since 1977. A very successful commercial application of adaptive is for control of ultrafiltration of blood in dialysis machines, Sternby (1996), with several thousands units in operation. Companies like 3M have developed their own adaptive control schemes and implemented them in large number on programmable logic controllers (PLC), Alam and Burhardt (1979). Various commercial adaptive controllers have also been marketed, such as the Novatune from ABB, see Bengtsson and Egardt (1984), the Firstloop controller from First Control Systems, and more recently the Brainwave controller from Universal Dynamics Technologies, Dumont (1998).


**Figure 2** An artistic rendition of a mid-eighties adaptive control session

Yet, it would take many more years before the stability and convergence of those schemes could be analyzed, their stochastic, nonlinear and time-varying nature making the analysis a formidable task. In 1982, the soundness of the adaptive control approach was seriously challenged when robustness issues were raised. This was to lead to an increase in activity in the field, not to mention to a series of very lively adaptive control sessions at major control conferences, see Figure 2. Eventually, this resulted in a better understanding of adaptive control and the design of more robust schemes. For an overview of adaptive control theory, see Åström and Wittenmark (1995).

Multivariable self-tuning schemes were also rapidly proposed. A notable example is the work of Borisson (1979) who developed a multivariable self-tuning regulator for a restricted class of multivariable systems described by a linear vector difference equation

$$\mathbf{A}(q^{-1})\mathbf{y}(t) = \mathbf{B}(q^{-1})\mathbf{u}(t) + \mathbf{C}(q^{-1})\mathbf{e}(t) \tag{16}$$

where **y**, **u** and **e** are all vectors of dimension $p$. The noise $\mathbf{e}(t)$ is now a sequence of independent, equally distributed random vectors with zero mean and covariance **R**. **A**, **B** and **c** are $p \quad p$ polynomial matrices

$$\mathbf{A}(q) = \mathbf{I} + \mathbf{A}_1 q + \cdots + \mathbf{A}_n q^n$$
$$\mathbf{B}(q) = \mathbf{B}_0 + \mathbf{B}_1 q + \cdots + \mathbf{B}_{n-1} q^{n-1}$$





where $\mathbf{B}_0$ is nonsingular

$$\mathbf{C}(q) = \mathbf{I} + \mathbf{C}_1 q + \cdots + \mathbf{C}_n q^n$$

Nonsingularity of $\mathbf{B}_0$ implies that all outputs in the system have the same delay. Under those conditions, the scheme proposed by Borisson has the self-tuning property, i.e. it converges to minimum-variance control. Borisson tested his proposed scheme on a simulated paper machine headbox, a process which would then subsequently be used as a benchmark process in several studies of multivariable adaptive control.

Work on multivariable adaptive control on systems with multiple delays was first proposed by Goodwin *et al.* (1980). Although there has since been a significant amount of work on multivariable adaptive control, it has led to very few industrial applications of true multivariable adaptive control.

## 4. Let's Get Active!

In the self-tuning controller, learning is incidental, i.e. the STC does not attempt to improve learning by properly exciting the plant. Rather, the estimator passively counts on the input, generated by the controller to keep the plant output at setpoint, to sufficently excite the plant in order to estimate its dynamics.

Feldbaum was the first to realize that the control of an unknown plant can be formulated as an optimal stochastic control problem, see Feldbaum (1965). The resulting dual controller then generates a control signal that continuously finds the optimal compromise between regulation and excitation of the plant. Unfortunately, the problem is numerically intractable except for very simple, academic examples, see for instance Åström and Wittenmark (1971) and Åström and Helmersson (1986). Although it might be dangerous to draw conclusions from results obtained on two simple examples, those studies show that the control signal can roughly be thought as consisting of three components, a control term, a caution term and a probing term. Typically, the control term dominates when the uncertainty about the plant is low. When uncertainty is large and the control error is large, caution prevail. When uncertainty is large and the control error is small, then probing dominates. Those observations confirm what would intuitively be expected from a dual controller. Since in practice dual control cannot be computed, ways have to be found to approximate it. For instance, it was shown in Allison *et al.* (1995b) that the generalized minimum-variance controller with the proper choice of control weight will behave on the average like the cautious controller, but without the so-called turn-off phenomenon. The latter is characterized by control inactivity during periods of large uncertainty. The cautious control law for the integrator with unknown gain is given by

$$u(t) = -\frac{-\hat{b}}{\hat{b}^2 + p_b}\, y(t)$$

where $\hat{b}$ is the gain estimate and $p_b$ its covariance. It is clear that in periods of large uncertainty, i.e. when $p_b >> \hat{b}$, the controller gain is very small. Since the only excitation to the plant is provided by the controller, the plant is no more excited, and the uncertainty cannot be reduced. The system will stay like this until there is a perturbation sufficiently large to produce some control activity. In contrast, the control law that minimizes

$$J = E(y^2 + \rho u^2)$$

is given by

$$u(t) = -\frac{-\hat{b}}{\hat{b}^2 + \rho}\, y(t)$$

By choosing $\rho$ equal to the average covariance, the two control laws will, on the average display the same caution, but in the latter case without exhibiting turn-off periods. On the other hand, the latter controller will be overly cautious in periods of low uncertainty.

The emulation of the dual controller probing term is more difficult. Many suboptimal ways of introducing probing have been proposed over the years, such as adding a perturbation signal to the control signal, constraining the estimator covariance matrix, or using serial expansion of the dual control performance index to facilitate the computation of the solution. However, the technique that so far seems most promising involves modifications to the loss function to be minimized, see Wittenmark and Elevitch (1985). The basic idea is to introuce in the control loss function a measure of future parameter estimate uncertainty. For instance, $f$ or a delay-free system with unknown gain $b$ consider

$$J = E\{[y(t+1) - y_r]^2\} + \lambda f(p_b(t+2))$$

where $y_r$ is the setpoint, and $\lambda$ is a user-defined weight. The covariance $p_b(t{+}2)$ is the first covariance to be affected by $u(t)$. Thus, $u(t)$ will also attempt to minimize the future parameter-estimate covariance. When delay





is present in the plant, the covariance beyond the delay has to be considered, see Allison *et al.* (1995a). This complicates matters, as this term is no longer deterministic but now depends on future unknown outputs and residuals. Those quantities will have to be replaced by estimates. Allison *et al.* (1995a) used this technique to design and implement what may be the first application of dual control in the process industries, controlling the motor load on a wood chip refiner. Wood chip refiners are used to produce wood pulp by separating fibres mechanically rather than chemically. One of their characteristics is that the static gain between the input (plate gap) and the output (motor load) is nonlinear and time-varying, with occasional sign reversals. The active adaptive controller significantly improved upon an earlier passsive one, Dumont (1982), although it still required heuristics to avoid entering a limit cycle when the setpoint is greater than the maximum achievable load. An added benefit of active learning is the reduced start-up period, during which most of the control action is due to probing, with the result that parameter estimates reach satisfactory values much more rapidly. In Dumont and Åström (1988), nonlinear terms were added to the loss function to ensure that probing would only occur in a safe manner, i.e. in the direction corresponding to a static gain of the right sign. This approach does not require additonal heuristics to handle the case of an unreachable setpoint. Despite those few successful studies, all limited to fairly simple systems, much remains to be done toward the development of easy-to-use, general-purpose probing strategies.

## 5. Back to Basics

Despite the rapid progress in control theory that had taken place during the twenty or so preceeding years, in the early 1980's, the belief prevailed in industry that much of that theory was too esoteric to be used practically and addressed problems of not necessarily great industrial relevance. Citing inspiration from Abraham Lincoln's concern for the common people[1], Åström turned his attention to the common industrial controller, the PID controller. Up until then, most of the attempts at developing automatic PID tuners were using model-based techniques, with the notable exception of Foxboro's EXACT controller, released in 1984, Bristol and Kraus (1984) and based on heuristic logic developed using extensive computer simulation studies. The relay tuning method proposed by Åström

[1]"God must love common people, he made so many of them"

and Hägglund (1984) is based on the classical method for tuning PID controllers of Ziegler and Nichols (1943). That method computes the tuning of PID controllers using knowledge of only one point on the open-loop system's Nyquist curve. The point of intersection of the Nyquist curve with the negative real axis is determined by controlling the process by a simple proportional controller and increasing its gain until a limit cycle is obtained. A drawback of the Ziegler-Nichols method is that the user has no control on the amplitude of the limit cycle. By observing that the same point can be determined by replacing the proportional controller by a relay. This will generally induce a limit cycle in the closed loop with the advantage that its amplitude can be controlled by changing the amplitude of the relay. By automatic determination of the amplitude and period of the oscillation, tuning constants for a simple controller can be computed. This method can be modified by introducing hysteresis in the relay, allowing the determination of a point on the Nyquist away from the negative real axis. Gain and phase margins auto-tuners can then be easily developed. This technique was patented by Åström and Hägglund and has been implemented on several commercial PID controllers, augmented with the automatic building of a gain scheduling look-up table if a gain scheduling variable is available. Auto-tuning has now become a standard feature of commercial PID controllers. For more on PID control and auto-tuning, see respectively Åström and Hägglund (1995) and Åström and Hägglund (1988). The same technique can also be used to derive a rough estimate of the plant dynamics to initalize an adaptive controller.

## 6. What's Up Doc?

In the last ten years, the issue of control loop performance monitoring has become a hot topic both in the research community and the applied control community, particularly in the process industries. This is not a coincidence, considering that the typical control loop represents a $25,000 asset. typical process plant with about 2000 loops, about half of those are wasted assets, as many loops simply do not function properly! The half-life of a control loop tuning is about six months. It typically takes about two hours to manually audit the performance of a control loop. A typical process plant has between 2000 and 4000 loops, and rarely has the personnel with the skills required to perform such an audit. There is thus a need for tools to automatically and continuously monitor the performance of control loops.





For regulatory loops, minimum variance provides a benchmark. Harris showed that it is possible to estimate the minimum variance from closed-loop operating data without knowledge of the plant dynamics, see Harris (1989). Harris' technique is a direct consequence of Åström derivation of the minimum variance controller.

Consider the system

$$y(t) = \frac{B(q^{-1})}{A_1(q^{-1})} q^{-k} u(t) + \frac{C(q^{-1})}{A_2(q^{-1})\Delta^d} e(t) \tag{17}$$

If the process is minimum phase, the minimum variance controller is

$$u(t) = -\frac{A_1(q^{-1})G(q^{-1})}{B_1(q^{-1})F(q^{-1})A_2(q^{-1})\Delta^d} y(t) \tag{18}$$

with

$$C(q^{-1}) = F(q^{-1})A_2(q^{-1})\Delta^d + q^{-k}G(q^{-1}) \tag{19}$$

The minimum variance is then

$$\begin{aligned} \sigma^2_{\text{mv}} &= E[y^2(t+k)] = E\{[F(q^{-1})e(t+k)]^2\} \\ &= \sigma_e^2(1 + f_1^2 + f_2^2 + \ldots + f_{k-1}^2) \end{aligned}$$

We want to estimate the minimum variance when the system is under closed-loop control with the controller

$$u(t) = \frac{N(q^{-1})}{D(q^{-1})} y(t) \tag{20}$$

The closed-loop system is then described by

$$y(t) = \frac{CA_1D}{\Delta^d[A_1A_2D - BNA_2q^{-k}]} e(t) \triangleq H(q^{-1})e(t) \tag{21}$$

One could estimate $H$ and then knowing $N$, $D$, $k$ and $d$, attempt to solve for $A_1$, $A_2$, $B$ and $C$. However, as first shown by Harris, there is a simpler way.

Using the Diophantine equation, one can write

$$y(t) = Fe(t) + q^{-k} \frac{BNA_2\Delta^d F + GA_1D}{(A_1A_2D - BNA_2q^{-k})\Delta^d} e(t) \tag{22}$$

Note that the first $k$ terms of $H$, represented by $F$ are unaffected by the controller, i.e., are feedback invariant.

Thus the procedure is as follows:

- Estimate $H$ in $y(t) = He(t)$ and $\sigma_e$ from the closed-loop data
- Assuming knowledge of the delay $k$, compute the first $k$ terms $h_i$ of the impulse response of $H$
- The minimum variance $\sigma^2_{\mathrm{mvc}}$ is then $(1 + h_1^2 + \cdots + h_{k-1}^2)\sigma_e^2$
- The performance can then be measured by the ratio of the actual variance to the minimum one:

$$P = \frac{\sigma^2_{\mathrm{mv}}}{\sigma^2_y}$$

This technique has led to the development of several control performance monitoring toolboxes, with many of the early systems developed in Canada for applications to the pulp and paper industry, Owen *et al.* (1996), Lynch and Dumont (1996). Several commercial systems are now available. Most of those systems also compute additional measures such as a loop oscillation index or a valve friction index. Although the use of the minimum variance as benchmark performance makes for an index that is relatively easy to compute and requires little knowledge about the plant, it has several shortcomings. Minimum variance control is not necessarily a realistic yardstick in many situations. An index allowing the user to assess the performance with respect to realistic expectations given control energy, actuator constraints and nonlinearities would prove more useful in practice. Several researchers have recently proposed such extensions, see for example Horch and Isaksson (1998). The problem of assessing the performance of a multivariable control system has also attracted attention, see Harris *et al.* (1996), Huang *et al.* (1997), Ettaleb *et al.* (1998). In the multivariable case, the difficulty is compounded by the fact that the interactor matrix (the multivariable equivalent of time delay) is not unique and that to different ordering of outputs, correspond different minimum variances, unless the interactor is diagonal.

Whereas all systems developed so far give some indication of performance, very few propose diagnostics of poor performing loops. For those systems to be useful, they have to propose ways of improving performance, whether it be retuning the controller or servicing a control valve. The latter has received the most attention as stiction in valves has been found to be the major cause of loop oscillation. Because bypass lines are rarely installed in pulp and paper mills, servicing or changing a valve requires a shutdown. Consequently it is important to propose ways of improving the





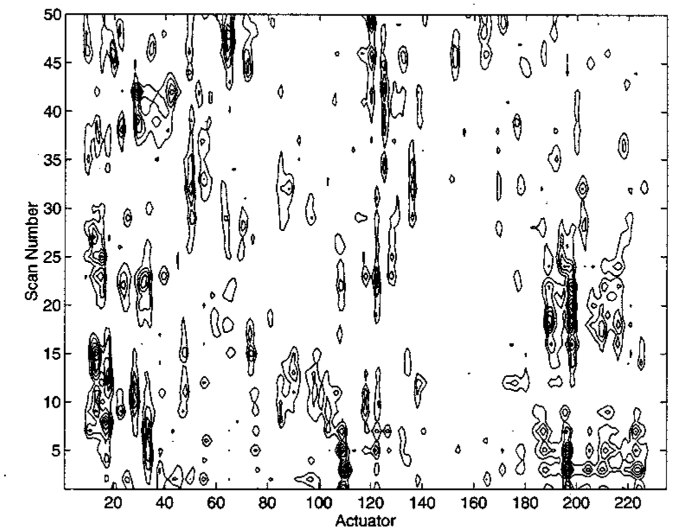


**Figure 3** Contour plot of the squared deviation between achieved and optimal variations on a paper machine, Duncan *et al.* (1999)

performance of the loop whose control valve has been detected as faulty, since it might be weeks or months before it is changed. Proposed methods range from the cleverly simple valve knocker of Hägglund (1995) to the more sophisticated adaptive stiction compensator of Bergstrom (1998).

Whereas the minimum-variance benchmark is applicable to the regulation problem, it is not suitable for the servo problem, i.e. when the main objective is good response to setpoint change and good rejection of load disturbances. Åström (1991) proposed various measures (both qualitative and quantitative) of performance, many of which rely on minimal knowledge of the process. In fact, methods that use knowledge derived from a simple relay experiment can provide a reasonably accurate estimate of the achievable performance when using PI or PID control. Such methods can be incorporated in a PID controller to give it self-diagnosis capabilities, a step toward autonomous PID control. Performance assessment has also been used in adaptive control to decide when to re-initialize the estimator, see Dumont (1992). Performance assessment would also prove useful for iterative control design in order to decide whether to proceed to the next iteration or to stop the process at the current iteration.

Finally, although performance assessment for control of distributed parameter systems is still in its infancy, several inroads have been made in the context of cross-directional control on paper machines, see Nesic *et al.* (1998), Fu *et al.* (1998) and Duncan *et al.* (1999). In the latter reference, the performance of a generalized minimum-variance controller is used as benchmark, assuming knowledge of the temporal and dynamic responses of the system. By adjusting the control weighting term in the GMV loss function, it is possible to obtain a realistic benchmark performance that accounts for the actuator constraints. Figure 3 shows a contour plot of the squared deviation between the actual basis weight variation and the optimal one for 50 scans, i.e. about 25 minutes worth of paper production. Such a display can be used by the operator to assess the tuning of the cross-directional control system, to detect machine malfunctions, and in particular to detect faulty CD actuators. In constrast, Nesic *et al.* (1998) assume little knowledge of the system's responses and estimate the optimal performance through a wavelet-based multi-resolution analysis of the data.

## 7. Conclusions

Process control theory has come a long way since 1962. Although minimum-variance control is rarely found in the pulp and paper industry, or in any industry for that matter, it has had a profound impact on control by leading to the development of closely related techniques such as constrained controllers and predictive controllers, both of which have been applied widely across industries. The current work on performance monitoring of control loops has been largely inspired by the derivation and properties of minimum-variance control. The minimum-variance based STR can be credited for spearheading the rapid development of the adaptive control field in the 1970's and 1980's. This has also had a strong impact on the practice of industrial control, particularly with the development of auto-tuners for PID controllers. The pulp and paper industry, traditionally perceived as a very conservative industry played an important role in demonstrating the potential benefit of those advanced control techniques. Finally, as has been demonstrated here, great theoretical work does not have to be disconnected from real industrial problems. The work reviewed here is an example where the synergy between the very practical needs of an industry and the power of mathematical abstraction has permitted major breakthroughs in both the theory and the practice of control.





## 8. References


Alam, M. and K. Burhardt (1979): "Further work on self-tuning regulators." In *IEEE Conference on Decision and Control*, pp. 616–620. Ft. Lauderdale, FLA.

Allison, B., J. Ciarniello, P.-C. Tessier, and G. Dumont (1995a): "Dual adaptive control of chip refiner motor load." *Automatica*, **31**, pp. 1169–1184.

Allison, B., P.-C. Tessier, and G. Dumont (1995b): "Comparison of suboptimal dual adaptive controllers." In *1995 European Control Conference*. Rome, Italy.

Åström, K. (1967): "Computer control of a paper machine – an application of linear stochastic control theory." *IBM Journal*, **11**, July, pp. 389–405.

Åström, K. (1970): *Introduction to Stochastic Control Theory*. Academic Press, New-York.

Åström, K. (1973): "Lectures on paper machine control – simple papermachine models." Technical Report 7305 C. Division of Automatic Control, Lund Intitute of Technology, Lund, Sweden.

Åström, K. (1991): "Assessment of achievable performance of simple feedback loops." *Int. J. of Adaptive Control and Signal Processing*, **5**, pp. 3–19.

Åström, K. and T. Bohlin (1965): "Numerical identification of linear dynamic systems from normmal operating records." In *2nd IFAC Symposium on the Theory of Self-Adaptive Control Systems*, pp. 96–111. NPL Teddington, England.

Åström, K. and T. Hägglund (1984): "Automatic tuning of simple regulators with specifications on phase and amplitude margins." *Automatica*, **20**, pp. 645–651.

Åström, K. and T. Hägglund (1988): *Automatic Tuning and Adaptation for PID Controllers*. Instrument Society of America, Research Triangle Park, N.C.

Åström, K. and T. Hägglund (1995): *PID controllers: Theory, Design and Tuning*. Instrument Society of America, Research Triangle Park, N.C.

Åström, K. and B. Häggman (1974): "Paper machine dynamics." In *4th IFAC/IFIP International Conference on DIgital Computer Applications to Process Control*, pp. 463–476. Zürich, Switzerland.

Åström, K. and A. Helmersson (1986): "Dual control of an integrator with unknown gain." *Comp. and Math. with Applications*, **12A**, pp. 653–662.

Åström, K. and B. Wittenmark (1971): "Problems of identification and control." *J. of Mathematical Analysis and Applications*, **Vol. 34, No. 1**, April, pp. 90–113.

Åström, K. and B. Wittenmark (1973): "On self-tuning regulators." *Automatica*, **Vol. 9, No. 2**, March, pp. 185–199.

Åström, K. and B. Wittenmark (1995): *Adaptive Control, Second Edition*. Addison-Wesley Publishing Co., Reading, Mass.

Bengtsson, G. and B. Egardt (1984): "Experiences with self-tuning control in the process industry." In *9th IFAC World Congress*, vol. XI, pp. 132–140. Budapest, Hungary.

Bergstrom, J. (1998): "Simulation and adaptive control for valve friction problems in a paper mill.". Master's thesis, Dept. of Electrical and Computer Eng., University of British Columbia, Vancouver, Canada.

Borisson, U. (1979): "Self-tuning regulators for a class of multivariable systems." *Automatica*, **15**, pp. 209–215.

Borisson, U. and R. Syding (1973): "Control of an ore crusher – a feasibility study of self-tuning control and remote processing." Technical Report 7318. Division of Automatic Control, Lund Intitute of Technology, Lund, Sweden.

Borisson, U. and B. Wittenmark (1974): "An industrial application of a self-tuning regulator." In *4th IFAC/IFIP International Conference on DIgital Computer Applications to Process Control*, pp. 76–87. Zürich, Switzerland.

Bristol, E. and T. Kraus (1984): "Life with pattern adaptation." In *1984 American Control Conference*. San Diego, Calif.

Cegrell, T. and T. Hedqvist (1973): "Successful adaptive control of paper machines." In *3rd IFAC Symposium on Identification and System Parameter Estimation*. The Hague, Netherlands.

Dahlin, E. (1967): "On-line identification of process dynamics." *IBM Journal*, **11**, jul, pp. 406–426.

Dumont, G. (1982): "Self-tuning control of a chip refiner motor load." *Automatica*, **18**, pp. 307–314.





Dumont, G. (1992): "Fifteen years in the life of an adaptive controller." In *4th IFAC Symp. on Adaptive Systems in Control and Signal Processing*, pp. 109–120. Grenoble, France.

Dumont, G. (1998): "From LUST to Brainwave: Fifteen years of persistent excitation." In *6th IFAC Symp. on Adaptive Systems in Control and Signal Processing*. Glasgow, UK.

Dumont, G. and K. Åström (1988): "Wood chip refiner control." *IEEE Control Systems Magazine*, April, pp. 38–43.

Dumont, G. and P. Bélanger (1978): "Self-tuning control of a titanium dioxide kiln." *IEEE Transactions on Automatic Control*, **23**, pp. 532–538.

Duncan, S., G. Dumont, and D. Gorinevsky (1999): "Evaluating the performance of cross-directional systems." In *American Control Conference*. San Diego, Calif.

Ettaleb, L., G. Dumont, and M. Davies (1998): "Performance assessment of single and multivariable feedback controllers: A unified approach." *submitted to Automatica*.

Feldbaum, A. (1965): *Optimal Control Theory*. Academic Press, New York, N.Y.

Fu, C., S. Neyhan, and S. Bale (1998): "CD control performance assessment." In *Control Systems'98*, pp. 289–297. Porvoo, Finland.

Galtieri, C. (1964): "Problems of estimation in discrete-time processes." Technical Report RI 315. IBM Research Laboratory, San Jose, Calif.

Goodwin, G., R. P., and P. Caines (1980): "Discrete-time multivariable adaptive control." *IEEE Transactions on Automatic Control*, **AC-25**, pp. 449–456.

Hägglund, T. (1995): "The knocker–a compensator for stiction in control valves." Technical Report TFRT -7539. Dept. of Automatic Control, Lund Intitute of Technology, Lund, Sweden.

Harris, T. (1989): "Assessment of control loop performance." *The Canadian Journal of Chemical Engineering*, **67**, pp. 856–861.

Harris, T., F. Boudreau, and J. McGregor (1996): "Performance assessment of multivariable feedback controllers." *Automatica*, **32**, pp. 1505–1518.

Horch, A. and A. Isaksson (1998): "A modified index control performance assessment." In *American Control Conference*, pp. 3430–3434. Philadelphia, PA.

Huang, B., S. Shah, and E. Kwok (1997): "Good, bad or optimal? performance assessment of mimo processes." *Automatica*, **33**, pp. 1175–1183.

Kalman, R. (1958): "Design of a self-optimizing control system." *Trans. ASME*, **80**, pp. 468–478.

Ljung, L. (1987): *System Identification: Theory for the User.* Prentice-Hall, Englewood Cliffs, N.J.

Lynch, C. and G. Dumont (1996): "Control loop performance monitoring." *IEEE Transactions on Control Systems Technology*, March, pp. 185–192.

Nesic, Z., M. Davies, G. Dumont, S. McLeod, and I. Shaw (1998): "Paper machine operator reel/roll display for improved paper quality." In *Control Systems'98*, pp. 41–47. Porvoo, Finland.

Owen, J., D. Read, H. Blekkenhorst, and A.A.Roche (1996): "A mill prototype for automatic monitoring of control loop performance." In *Control Systems'96*, pp. 171–178. Halifax, Canada.

Sternby, J. (1996): "Adaptive control of ultrafiltration." *IEEE Transactions on Control Systems Technology*, **6**, pp. 11–17.

Wittenmark, B. and C. Elevitch (1985): "An adaptive control algorithm with dual features." In *7th IFAC/IFORS Symposium on Identification and System Parameter Estimation.* York, UK.

Ziegler, J. and N. Nichols (1943): "Optimum settings for automatic controllers." *Trans. ASME*, **65**, pp. 433–444.