A simplified adaptive model predictive controller

Mark S. Ma* and D. C. Williams*
Research assistant and assistant professor, respectively, Dept. of Chemical Engineering, Auburn University, Auburn, Ala. 36849.

ABSTRACT: *We propose an adaptive controller based on predictive control and a simplified identification algorithm. The controller has several advantages for practical operation. It is robust, it obtains initial parameter estimates easily, and it simplifies on-line computation. A simulation study of basis weight control on a paper machine demonstrates its performance.*

KEYWORDS
Computers
Mathematical models
Process control
Process simulation

The advanced control method

The paper machine is difficult to control because it is too complex to be described completely by a mathematical model, it is subject to unknown disturbances, and its operating characteristics may change from time to time. The control method we propose combines two control methods, model predictive control and adaptive control. Model predictive control can be made insensitive to modelling errors. Adaptive control automatically adjusts for changing process characteristics and unknown disturbances. The combined method retains the advantages of each. We have also developed a simplified version that requires fewer calculations and could be implemented in a less expensive control system than the full version.

We tested the proposed method by computer simulations, using a paper machine model that was developed from experimental data and used by other researchers in adaptive control experiments. Both the simplified and full versions performed well in simulation, with little difference in performance. From the simulation results, we conclude that the simplified adaptive model is an attractive candidate for paper machine control and should be further studied experimentally.

Introduction

Adaptive control has great potential for improved operation in the process industries. Paper machine control is a particularly good candidate for adaptive control, as shown by some early examples of the use of minimum-variance control (*1*) and the self-tuning regulator (*2*). Other applications and simulation studies have been reported (*3–5*). Here, we develop an adaptive controller based on model predictive control with a simplified identification algorithm that reduces the requirements for on-line computing speed and memory. The use of the controller is demonstrated through simulation of basis weight control of a paper machine.

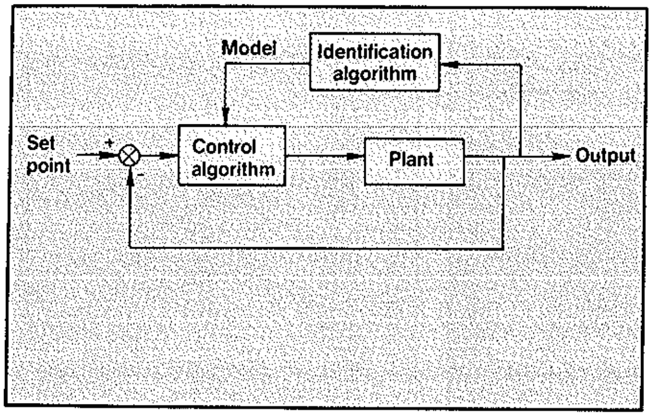
Model predictive control is the most important development in process control within the last decade. Developed independently in France as "model predictive heuristic control" (*6*) and in the United States as "dynamic matrix control" (*7*), model predictive control has been used successfully where conventional control has failed. The major advantage of model predictive control is that the control algorithm can be made robust, or insensitive to modelling errors, in a direct manner.

In theory, any adaptive controller is robust. In practice, the adaptive mechanism, which is usually the identification of a process model, may be switched off, and the control is then operated nonadaptively. In practice, therefore, it is important that the control algorithm be robust under nonadaptive conditions. An adaptive controller based on model predictive control retains the advantages of insensitivity to modelling errors if the adaptative mechanism is switched off.

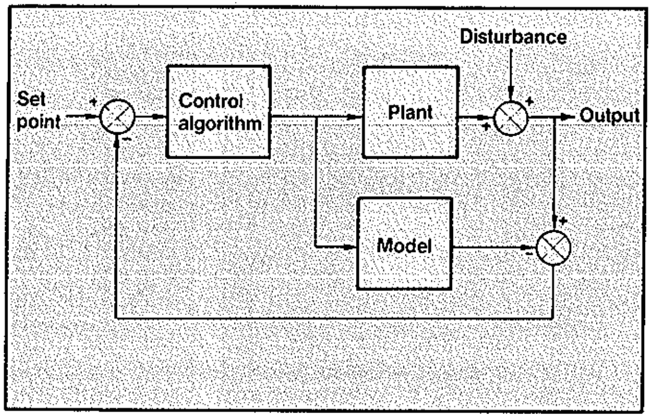
A major portion of the on-line computation of adaptive control comes from the identification of a process model. Although the cost of computing has continued to decrease over the last decade, the computational requirements for adaptive control may be greater than some control computers can achieve in real time. A simplified identification algorithm can reduce the computation requirements and allow adaptive control to be implemented on less expensive computers.

*Ma's current title and address are assistant professor, Dept. of Chemical Engineering, Mississippi State University, Starkville, Miss. 39762. Williams' current title is senior research engineer. Correspondence should be addressed to D. C. Williams, HIMONT Research and Development Center, 800 Greenbank Road, Wilmington, Del. 19808.

1. Adaptive control structure



2. Model predictive control structure



Overview of the adaptive controller

The adaptive controller can be conveniently divided into two parts: a model identification algorithm and a control algorithm (**Fig. 1**). By "control algorithm," we mean only the portion of the overall control system indicated in Fig. 1. We use the phrase "control system" to indicate the overall control system—both the control algorithm and the identification algorithm.

At each sampling time, we (a) measure the output, (b) use the identification algorithm to obtain a new model estimate, (c) calculate new parameters for the control algorithm based on the new model estimate, and (d) calculate the control input. The parameters of the control algorithm are calculated under the assumption that the current model estimate is the true model. This separation of the identification and control functions is called the "certainty equivalence principle" in stochastic control theory and can be shown to be optimal under certain conditions. Here, we use a model predictive control algorithm with a step response model. The identification algorithm is a simplified least squares algorithm.

Process model

Model predictive control requires a process model (**Fig. 2**) for predicting future output values. Here, the process to be controlled is assumed to be described in the absence of any disturbance by

$$y(k) = \sum_{i=1}^{n} a_i y(k-i) + \sum_{i=1}^{n} b_i u(k-d-i) \quad (1)$$

where

y = the controlled output

u = the manipulated input

d = a constant delay.

The z-transform of the model is

$$A(z^{-1})y(z) = B(z^{-1})u(z) \quad (2)$$

where

$$A(z^{-1}) = 1 - a_1 z^{-1} - a_2 z^{-2} - \ldots - a_n z^{-n} \quad (3)$$

$$B(z^{-1}) = 1 - b_1 z^{-1} - b_2 z^{-2} - \ldots - b_n z^{-n} \quad (4)$$

Division of Eq. 2 by $A(z^{-1})$ and inversion of the resulting z-transform gives

$$y(k) = \sum_{i=1}^{\infty} c_i u(k-d-i) \quad (5)$$

If the process is open-loop stable, Eq. 5 may be truncated to yield

$$y(k) = \sum_{i=1}^{m} c_i u(k-d-i) \quad (6)$$

Writing Eq. 6 for $y(k)$ and for $y(k-1)$ and subtracting gives

$$y(k) = y(k-1) + \sum_{i=1}^{m} c_i \Delta u(k-d-i) \quad (7)$$

where

$$\Delta u(k-d-i) = u(k-d-i) - u(k-1-d-i) \quad (8)$$

Equation 7 is the model form to be used in the development of the nonadaptive control algorithm. A major advantage of Eq. 7 over Eq. 1 is that the coefficients c_i are simply the outputs $y(k+1)$ for a unit step input at $k = 0$. Therefore, Eq. 7 is called a "step response model." A second advantage is that Eq. 7 can be obtained without requiring knowledge of the system order, n. The disadvantage of Eq. 7 in comparison with Eq. 1 is that the number of coefficients c_i will be much larger than the number of coefficients a_i and b_i.

Disturbance models

To this point we have assumed that the process is not subject to disturbances, which is generally unrealistic. If we add a disturbance to Eq. 7, we obtain

$$y(k) = y(k-1) + \sum_{i=1}^{m} c_i \Delta u(k-d-i) + d(k) - d(k-1) \quad (9)$$

Several disturbance models are commonly used for $d(k)$. In one model, which often models drifting process conditions, $d(k)$ is assumed to be constant. In another model, $d(k)$ is assumed to be an independent random variable with zero mean, which often models measurement noise. In another model, it is assumed that

$$d(k) = \sum_{i=0}^{n} e_i v(k-i) \quad (10)$$

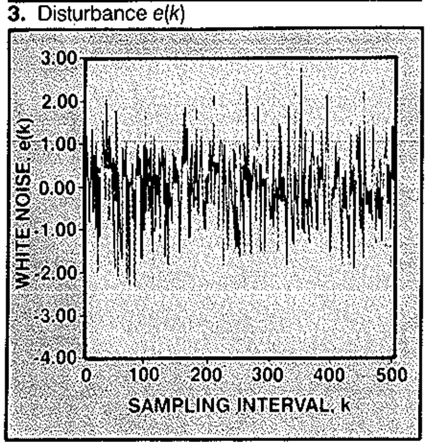
where $v(k)$ is an independent random variable with zero mean. This model is generally more accurate than the previous two, but it is more difficult to use if the disturbance is unmeasured, which is commonly the case. In another common model, $d(k)$ is the sum of a constant term and a random term. Other disturbance models can be used, but these are the most common. Here, we assume that $d(k)$ is an independent random variable with zero mean. We have examined control schemes and other disturbance models (*8*), but we have not reported on these here because the results are essentially the same for all disturbance models tested.

Control algorithm

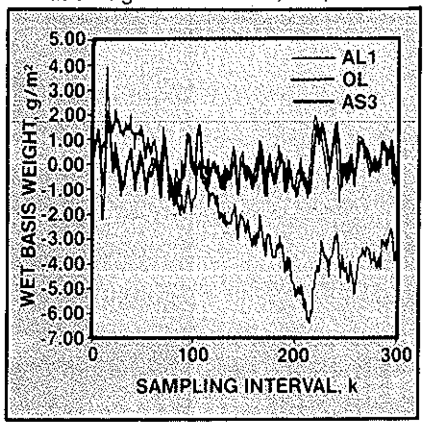
The formulation of the control algorithm is essentially taken from Tung (*9*).

This formulation is the same as

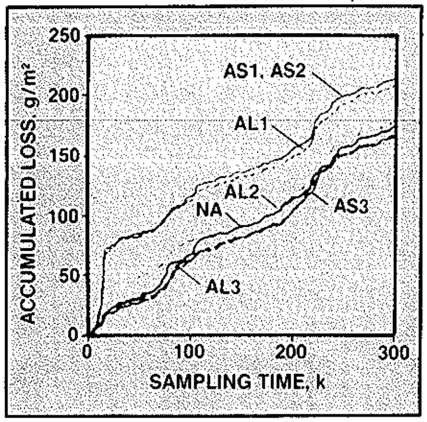
3. Disturbance $e(k)$



4. Basis weight for Cases OL, AL1, and AS3



5. Accumulated loss for all closed-loop cases



dynamic matrix control (7). The control is chosen as the solution of the following optimization problem.

At each sampling time, choose the sequence of inputs $\Delta u(k), \Delta u(k+1), \ldots, \Delta u(k+N-1)$, which minimizes the objective function:

$$J=(S-Y)^T W(S-Y)+\Delta U^T Q \Delta U \quad (11)$$

where Y is the vector of predicted outputs:

$$Y^T=[y(k+1+d), y(k+2+d), \ldots, y(k+N+d) \quad (12)$$

S is the vector of set points:

$$S^T=[s(k+1+d), s(k+2+d), \ldots, s(k+N+d) \quad (13)$$

ΔU is the vector of future control inputs:

$$\Delta U^T=[\Delta u(k), \Delta u(k+1), \ldots, \Delta u(k+N-1)] \quad (14)$$

and W and Q are diagonal nonnegative weighting matrices:

$$W=\mathrm{diag}[w_1, w_2, \ldots, w_N] \quad (15)$$

$$Q=\mathrm{diag}[q_1, q_2, \ldots, q_N] \quad (16)$$

The predicted outputs are determined from recursive application of Eq. 9. Since we assume $d(k)$ to be an independent random variable with zero mean, the best estimate of $d(k)$ is zero. We can write the predicted outputs in a convenient vector-matrix notation as follows:

$$Y=G\Delta U+\Xi \quad (17)$$

where G is an $N \times N$ matrix with elements

$$G_{ij}=h_{1+i-j} \text{ for } i \leq j \quad (18A)$$

$$G_{ij}=0 \text{ for } i>j \quad (18B)$$

and

$$\Xi^T=[\xi_1, \xi_2, \ldots, \xi_N] \quad (19)$$

The h_i are defined recursively by these equations:

$$h_i=0 \text{ for } i \leq 0 \quad (20)$$

$$h_1=c_1 \quad (21)$$

$$h_i=h_{i-1}+c_i \quad (22)$$

The ξ_i are defined recursively by

$$\xi_0=y(k+d) \quad (23)$$

$$\xi_i=\xi_{i-1}+\sum_{j=i+1}^{m} c_i \Delta u(k+i-j) \quad (24)$$

Setting the derivative of Eq. 9 with respect to Eq. 12 to zero yields the control algorithm

$$\Delta U=(G^T WG+Q)^{-1} G^T W(S-\Xi) \quad (25)$$

Since ΔU is calculated at each sampling time, it is necessary to calculate explicitly only $\Delta u(k)$.

The matrix inversion in Eq. 25 can be performed recursively as a result of the special structure:

$$1/q_2, 1/q_3, \ldots, 1/q_N \quad (26)$$

$$D_{k+1}=D_k-(w_{k+1} D_k g_{k+1} g^T_{k+1} D_k)/(1+w_{k+1} g^T_{k+1} D_k g_{k+1}) \quad (27)$$

$$D_N=(G^T WG+Q)^{-1} \quad (28)$$

where g_i^T is the i^{th} row of G.

Identification algorithm

Least squares identification (LSI) is often used for several reasons. LSI converges more quickly than gradient methods. It does not require a knowledge of statistical properties as do other more powerful identification methods such as the maximum likelihood method. Perhaps most importantly, it is based on a simple idea, which makes it acceptable to the engineer who may not have rigorous training in statistics. The standard LSI algorithm in recursive form is:

$$\epsilon(k)=y(k)-y(k-1)-\Psi^T(k)\Theta(k-1) \quad (29)$$

$$\gamma(k)=P(k-1)\Psi(k)[1+\Psi^T(k) P(k-1)\Psi(k)]^{-1} \quad (30)$$

$$P(k)=[I-\gamma(k)\Psi^T(k)]P(k-1) \quad (31)$$

$$\Theta(k)=\Theta(k-1)+\gamma(k)\epsilon(k) \quad (32)$$

where

$$\Psi^T(k)=[\Delta u(k-1-d)\,\Delta u(k-2-d) \ldots, \Delta u(k-m-d) \quad (33)$$

$$\Theta^T(k)=[c_1(k), c_2(k), \ldots, c_m(k)] \quad (34)$$

The matrix inversion in Eq. 30 consumes most of the on-line computational time. A simplification of the standard algorithm (10) eliminates the need for this matrix inversion. The simplification is that $P(k)$ is forced to be a diagonal matrix. The basis for this approximation is that most of the information in the matrix P is contained in the diagonal elements. If P is forced to be diagonal, the matrix inversion is replaced with scalar division. We call this algorithm "simplified recursive identification" (SRI). The SRI algorithm is:

$$\epsilon(k)=y(k)-y(k-1)-\Psi^T(k)\Theta(k-1) \quad (29)$$

$$\gamma_i(k)=p_i(k-1)\psi_i(k)/[1+\Psi^T(k) P(k-1)\Psi(k)]^{-1} \quad (35)$$

$$p_i(k)=[1-\gamma_i(k)\psi_i(k)]\,p_i(k-1) \quad (36)$$

$$\Theta(k)=\Theta(k-1)+\gamma(k)\epsilon(k) \quad (32)$$

where

$$\Psi^T(k)=[\Delta u(k-1-d)\,\Delta u(k-2-d) \ldots) \Delta u(k-m-d)] \quad (33)$$

$$\Theta^T(k)=[c_1(k), c_2(k), \ldots, c_m(k)] \quad (34)$$

$$\psi_i(k)=\Delta u\,(k-i-d) \quad (37)$$

$$\partial(k)=[\partial_1(k), \partial_2(k), \ldots, \partial_m(k)] \quad (38)$$

SRI is the same as LSI except that the matrix calculations of Eqs. 30 and 31 are replaced with the scaler calculations of Eqs. 35 and 36. This reduces the on-line computational time.

Control system

The control system combines the control and identification algorithms. At each sampling time we (a) measure the output, (b) estimate a new model by either LSI (Eqs. 29–34) or SRI (Eqs. 29, 33–38), and (c) calculate the control input by the control algorithm (Eqs. 18–28). Computational time is saved by calculating the control input directly, without explicitly obtaining the parameters of the control algorithm.

Plant model

For our simulation study of basis weight control, we simulate the plant with a model taken from Astrom and Wittenmark (*2*). This model was developed from experimental operating data (*1*) and is as follows:

$$y(z)= u(z)\,(2.307z^{-3} - 2.025z^{-4})/(1-1.283z^{-1}+ 0.495z^{-2}) + e(z)\;\; (0.3821 - 1.438z^{-1} +0.550z^{-2})/[(1- 1.283z^{-1} + 0.495z^{-2})\,(1 - z^{-1})] \quad (39)$$

where

y = basis weight

u = thick stock flow

e = a disturbance.

Although this is not the most recent model, we chose it because it was developed from experimental data and was used in adaptive control experiments.

The disturbance of the plant is modelled as Eq. 10. Furthermore, the disturbance model contains an integrator so that the disturbance model may be loosely viewed as consisting of step inputs of random sign and magnitude. In the process model used by the control algorithm in the simulations presented here, we have assumed that the disturbance is an independent random variable with zero mean. Despite the difference between the "true" distrubance (as modelled by the plant model) and the "assumed" disturbance (as modelled by the process model used in the control), the control algorithm performs well.

Simulations

The simulation has two major goals. The first is simply to determine whether adaptive control based on model predictive control is a candidate for paper machine control. The second is to determine whether the use of the more simple SRI algorithm instead of the more accurate LSI algorithm causes a significant loss in control performance.

All quantities are reported as deviation from steady state. The initial conditions are

$$y=\Delta u = d = s= 0$$

At time k= 0, a random disturbance $e(k)$ is introduced, where $e(k)$ is simulated by pseudo-Gaussian distributed random numbers generated from uniformly distributed random numbers through the central limit theorem. The random number generator used is FORTRAN Uniform (0,1) from Marse and Roberts (*11*). This scheme has been tested by Riggs (*12*). The disturbance $e\,(k)$ is shown in **Fig. 3**. In each simulation, the weighting matrices in equation 11 are chosen as W= 20I and Q = I, where I is the identity matrix.

We compare the results with accumulated loss:

$$\sum_{i=0}^{k} w_i[s(i)-y(i)]^2 \quad (40)$$

In the simulation, we compare the adaptive control with open-loop operation (no control) and with nonadaptive model predictive control using a "perfect" model. We compare the effect on the control of different initial estimates for the model parameters. We also compare control with LSI and SRI.

Case OL is open-loop operation. No control action is taken.

Case NA is nonadaptive MPC with a "perfect" model. The process model used by the control algorithm to predict future outputs is

$$y(k)= 0.93488\Delta u\,(k-4) + 0.05748\Delta u\,(k-5) - 0.38901\Delta u\,(k-6) - 0.52756\Delta u\,(k-7) - 0.48429\Delta u\,(k-8) - 0.36021\Delta u\,(k-9) - 0.22242\Delta u\,(k-10) - 0.10707\Delta u\,(k-11) - 0.27265\Delta u\,(k-12) \quad (41)$$

This model is obtained from the application of the procedure (Eqs. 2–7) to the model given by Astrom (*2*). This model is "perfect" in that it matches the true process (Eq. 39) except that the disturbance terms are omitted.

Cases A(L,S) (Eqs. 1-3) is adaptive control with either LSI (L) or SRI (S) as the identification algorithm. We consider three sets of initial parameter estimates: (1) $c_1(O)$ = 0.015, all others zero, (2) all initial estimates = 0.015, and (3) all initial estimates are given by the perfect model (Eq. 39). In other words,

$$c_1(O)= 0.93488, c_2(O)= 0.057487,$$

and so on.

Results

The results are summarized in **Figs. 4 and 5.** Figure 4 shows the response of the basis weight for Cases OL, AL1, and AS3. The open-loop response is poorer than either closed-loop response. The responses for AL1 and AS3 are representative of all the closed-loop responses. (The others are not included for readability.)

The difference in the closed-loop responses are attributable to the initial parameter estimates. Case AL1 shows a larger deviation from steady state than does Case AS3. This reflects the better initial parameter estimates used in Case AS3. The differences between Cases AL1 and AS1, between Cases AL2 and AS2, and between AL3 and AS3 are not shown because they are too small to be readable. This illustrates that the control with the more simple SRI performed as well as control with the more accurate LSI.

Figure 5 shows the accumulated losses for all cases except Case OL. Again, there is an advantage in using better initial parameter estimates. This advantage becomes relatively constant as k increases, showing that the effect of the initial parameter estimates appears only in the initial stages of the control.

The control with SRI performs as well as control with LSI. If there are slight differences in performance, these occur during the initial period. After a few sampling times, the performance is essentially the same.

For some times, the accumulated loss for Case AS3 is less than the loss for Case NA. In other words, the adaptive control with initial parameter estimates from the "perfect" model performs better than the nonadaptive control with the "perfect" model. In both the nonadaptive and adaptive control, we assume the disturbance is an independent random variable with zero mean, which is incorrect. The adaptive control adjusts its model parameter estimates away from the "perfect" model in such a way that future predictions of the output are better and the resulting control is better.

The best performance is achieved with "perfect" initial parameter estimates. "Perfect" estimates are unrealistic in practice, but good estimates are not an unreasonable assumption. This is one advantage of the step response model. The initial parameter estimates can be obtained from a simple, open-loop, step response. The choice of a step response model rather than an impulse response model gives an advantage in that the signal-to-noise ratio for the measurements will generally be better for a step response.

Summary

We have developed an adaptive controller based on model predictive control with a simplified identification algorithm. Simulation testing has shown the controller to be a candidate for paper machine control.

In this proposed adaptive control, the model predictive control algorithm is robust. Consequently, if the adaptive mechanism is switched off, the advantages of model predictive control are retained.

The step response model allows good initial parameter estimates to be obtained from a simple experiment. It also allows the matrix inversion required in the control algorithm to be performed efficiently through recursive equations.

Finally, a simplified recursive identifier eliminates the matrix inversion from the identification algorithm and greatly reduces computational time.

Literature cited

1. Astrom, K. J., IBM J. Res. Dev. 11:389(1967).
2. Astrom, K. J. and Wittenmark, B., Automatica 9: 185(1973).
3. Cegrell, T., and Hedqvist, T., Automatica 11:53(1975).
4. Sikora, R. F., Pulp Paper Can. 85(8): T194(1984).
5. D'Hulster, F. M., DeKeyser, R. M. C., and Cauwenberghe, V., Automatica 19: 407(1983).
6. Richalet, J. A., Rault, A., Testud, J. D., and Papon, J., Automatica 14: 413(1978).
7. Cutler, C. L. and Ramaker, B. L., Joint Automatic Control Conference, American Automatic Control Council, Green Valley Ariz., 1980, Paper WP5-B.
8. Ma, M. S., Ph.D. dissertation, Dept. of Chemical Engineering, Auburn University, 1987.
9. Tung, L., 1983 American Control Conference, American Automatic Control Council, Green Valley, Ariz., p. 349.
10. Farsi, M., Karam, K. Z., and Warwick, K., Electronics Letters 20(22): 913(1984).
11. Marse, K., and Roberts, S. D., Simulation 135 (Oct., 1983).
12. Riggs, T., Ph.D. dissertation, Dept. of Electrical Engineering, Auburn University, 1986.

This research has been supported by the Auburn University Pulp and Paper Research and Education Center and the Auburn University Advanced Manufacturing Technology Center. The assistance of C. L. Phillips and T. Riggs of the Electrical Engineering Dept. is deeply appreciated.

Received for review July 30, 1987.

Accepted Dec. 17, 1987.