

Performance Monitoring for Model Predictive Control Maintenance

Per Erik Modén, Michael Lundh

Abstract—An economically motivated performance measure is proposed for use with model predictive control applications. The typical case handled is that you have a trade-off situation between two conflicting goals: (i) the need to be close to a constraint, since this will give good production economy, and (ii) the need to avoid constraint violation, since that would mean producing an inferior quality, exceeding environmental constraints (thereby requiring a fee), or having some other costly drawback. The proposed performance measure uses consumption and production rates in the process and the corresponding cost and benefit factors, combined with the risks of constraint violations and their associated costs. There is a discussion on the hard task of choosing threshold values, and the approach is illustrated by applying it to a couple of examples. One of the examples is the Autoprofit benchmark pulp digester model.

I. INTRODUCTION

IT is a common experience that model based control (model predictive control, MPC, for example) usually delivers good performance when it has been recently commissioned. Its operation leads to economic benefit. But over time things change. What often happens within at most a couple of years, is that a model/plant mismatch has emerged, which affects the performance negatively. To remedy that, you first of all need to observe the performance drop, and then perform additional investigations and tests that in the end may lead to an updated model and a retuned controller. This is part of an overall goal to extend the beneficial life-time of model based control.

In this paper, we focus on monitoring of the performance, and not just the control performance in the sense of how accurately all variables are controlled. One often referred approach to control performance monitoring compares the achieved control accuracy, in terms of variances, with a theoretical limit, the minimum variance control case. This is what you do with Harris' index [8], and its extensions to multivariable control as exemplified by [6] and with interpretations stated in [9]. Some alternatives, like [1], compare with a more realistic case but do require a valid model. Our interest is rather how economically the plant or process is operated. Some control performance drops have

large impact on the economy of operation, while others may have little effect on costs and benefits. Let us say that the relevant aspects of control accuracy are in focus.

It is assumed that benefits and costs are mainly expressed using easily quantifiable entities like the production rate, and the consumption of energy, raw material and additives. We will assume that there is a benefit or cost factor associated with each of these basic variables. That gives a linear expression for the 'unconstrained profit'. When trying to increase this, you will eventually reach a constraint. This is often a constraint on a quality measure, meaning that this quality measure becomes the critical variable. This critical variable is to be kept from violating its constraint, since otherwise the product will have a much lower value (not 'in spec').

If we had perfect control of the critical variable, in the sense that it had zero variance, we could choose to control it exactly to the constraint. In real life, this is not possible, and we will have to back off from the constraint in order to avoid too many violations. This is a trade-off between strive for economic production in terms of the 'unconstrained profit' on one hand and on the other hand the need to avoid costly constraint violations. We will simply use a cost factor for violations in contrast to the elaborated cost functions in for example [7]. Alternatives are given in [2], which provides an overview of economic performance assessment.

A totally different aspect of performance monitoring for model based control is that malfunction of actuators or sensors may be responsible for most of the performance drops. Another common reason is an operators' choice to shut off the automatic control and run manually. The latter can in itself be an indication that the model based control does not work as intended. But these aspects should be handled on another level, and they are not addressed in this paper. The situation considered here is fully automatic operation in steady state. In case the performance monitoring presented here detects a drop that coincides with detection of malfunction in base layer control loops, the cause would be considered as determined by that low level detection, rather than leading to high level tests to diagnose whether the cause could for example be that the model/plant mismatch has increased.

In section II, the paper presents and analyses a performance measure. The choice of threshold values is discussed in section III. Two example cases illustrate the

The research leading to these results has received funding from the European Union's Seventh Framework Programme (FP7/2007-2013) under grant agreement n° 257059, The 'Autoprofit' project (www.fp7-autoprofit.eu)

P. E. Modén and M. Lundh are with ABB AB, Corporate Research, SE 72178 Västerås, Sweden. (e-mails: per-erik.moden@se.abb.com, michael.lundh@se.abb.com)

methods in section IV, and some concluding comments are given in section V.

II. PERFORMANCE EXPRESSION

We introduce a vector of cost or benefit variables, z , and a corresponding vector of cost or benefit factors, f . Typically one benefit variable would be production rate, and the corresponding benefit factor would be value per amount produced (here assuming that it is in spec). A typical cost variable is the consumed power, and its corresponding cost factor is the cost per unit of consumed energy. If there is only one constraint involved, we denote the probability of violating it by p , and the cost per time unit of doing it by c . In the typical case of a quality variable being the critical one, the cost factor c would be the production rate multiplied by the difference in value (per amount produced) between product in spec and out of spec. In case more than one constraint is involved, introduce a vector p of violation probabilities and a corresponding vector of cost factors, c . (We assume that the cost contributions from different constraint violations are independent of each other.) The expression for the performance, or profit, P_t , is

$$P_t = f^T z - c^T p. \quad (1)$$

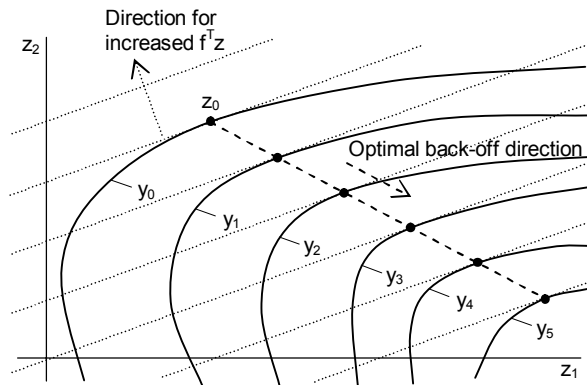


Fig. 1 Steady state operating points for a case with two cost/benefit variables, z_1 and z_2 . Dotted straight lines are level curves for the ‘unconstrained profit’, $f^T z$. Solid lines are level curves for the critical variable, y . Ignoring the risk of violating the constraint, the optimal operating points with the critical variable at values y_0, y_1, \dots, y_5 , would be the bullets along the dashed line. If y_0 is the constraint, and y_1, \dots, y_5 represent increasing back-off from it, then backing off along the dashed line gives the least reduction of the ‘unconstrained profit’, $f^T z$.

An illustration is given in Fig. 1, which shows equidistant level curves for the ‘unconstrained profit’ $f^T z$ in a case with two cost/benefit variables, z_1 and z_2 . It also shows level curves for the critical variable, y , at its constraint y_0 and a few values y_1, \dots, y_5 representing increasing amount of back-off from the constraint. The optimal operating point for a given value of y , when only the ‘unconstrained profit’ is considered, is obtained as the point where a level curve for the unconstrained profit touches (is tangent to) the level curve for this y value. In the figure, z_0 marks that point for $y = y_0$.

Backing off from the constraint is required, since there will always be some spread in the obtained value of the critical variable. The larger its standard deviation, the further you need to back off. The backing off causes a cost, a drop in the unconstrained profit, which is balanced by reduced costs of constraint violations as their frequency of occurrence drops when you increase the back-off.

The optimal back-off direction may vary, so it is not always described by a straight line. However, when backing off along a straight line, as in Fig. 1, we can express the operating point as

$$z = z_0 - c_z b, \quad (2)$$

where c_z gives the back-off direction and b is a scalar, the back-off distance. Introducing a new scalar, $c_b = f^T c_z$, and assuming that only one constraint is involved, we can rewrite (1) as

$$P_t = f^T z_0 - c_b b - cp. \quad (3)$$

Here everything that varies is scalar, so it is easy to visualize in a plot. Fig. 2 shows a principle example, how the profit could depend on the standard deviation of the critical variable. Suppose the critical variable has a symmetric distribution (for example Gaussian). Then operation at the constraint ($y = y_0$ and $z = z_0$) would give the profit

$$P_0 = f^T z_0 - c/2, \quad (4)$$

since the probability of constraint violation would be 0.5.

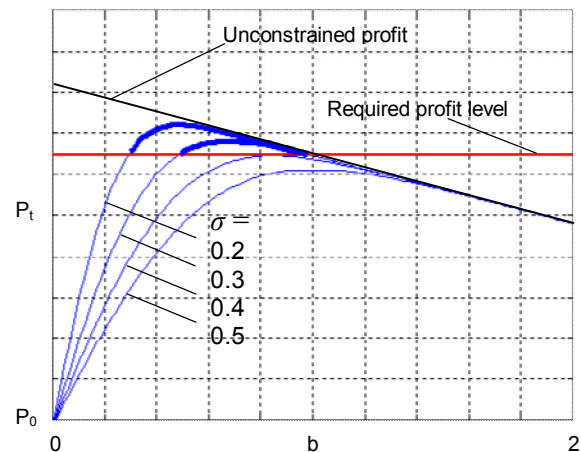


Fig. 2 Profit measure P_t as function of back-off distance b for a few different values of critical variable standard deviation, σ . There is also indicated a suggested required level of the profit measure.

As seen in Fig. 2 you can get better economic performance the lower standard deviation you manage to achieve in the critical variable. It also gets less critical how far you back off. But the real benefit is what you can get by operating closer to the constraint, a well-known circumstance motivating ‘constraint pushing’. If you are too close to the constraint, you will gain by increasing the back-off. But if you back off too far, you start losing again. And if

your standard deviation is too high, it may be impossible to reach the profit goal. This is the case in the example of Fig. 2, if the standard deviation is higher than 0.4. With exactly this amount of variation you can just about reach the acceptable economic performance, by backing off 0.85 from the constraint. With standard deviation 0.3, you would be okay with a back off between 0.5 and 1.0, and with standard deviation 0.2 the economic demand would be fulfilled for any back-off value between 0.3 and 1.0.

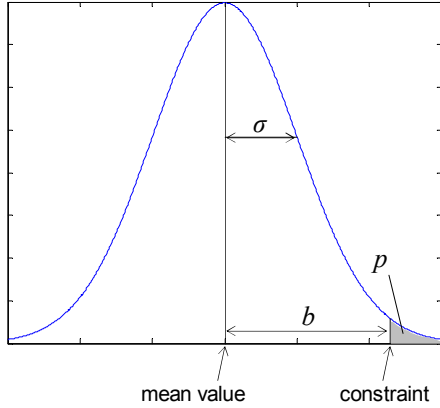


Fig. 3 Illustration of the back-off b from an upper constraint, when the probability distribution is Gaussian with standard deviation σ . The constraint violation probability is p . It gets lower the further you back off.

The case illustrated in Fig. 3 has a single constraint and a Gaussian probability distribution. For that case, we can give an analytical expression for the optimal back-off, b^* , see [10]. With performance depending on the back-off distance of the critical variable according to (3), and Gaussian distribution with standard deviation σ , we have the profit

$$P_t = f^T z_0 - c_b b - c \left(0.5 - \frac{1}{\sigma\sqrt{2\pi}} \int_0^b e^{-\frac{x^2}{2\sigma^2}} dx \right). \quad (5)$$

Taking the derivative of this expression with respect to b and setting it to zero gives the optimal back-off

$$b^* = \sigma \sqrt{-2 \ln \left(\frac{c_b}{c} \sigma \sqrt{2\pi} \right)}. \quad (6)$$

Further investigation on what you could gain from a certain standard deviation reduction can be found in [3].

Up to now, we have presented the theoretical aspect of the performance, using a probability distribution. During operation, this distribution will not necessarily be known. But it is pretty easy to observe the frequency of constraint violations (as estimate of the probability) and also to observe the cost and benefit variables. Thus it is easy to calculate a performance measure according to (1) on-line, present it to the operators, and use it for monitoring and create an event when it drops too much. And to do this, you do not need to restrict the number of variables used, or the

back-off direction, like we have done above in this section for illustration purpose.

III. THRESHOLD VALUES AND MONITORING

We want to use the expression (1) as performance measure for monitoring. The variables in z are assumed to be readily available. You need to have the cost factors in f and c determined. It remains to get the frequencies of constraint violations as estimate of the probabilities in p . Another way would be to estimate the variances of the involved constrained variables and calculate the probabilities based on assumption of Gaussian distributions. Since we want low probabilities, it is obvious that we need a long time window for direct estimation of the frequencies. This, on the other hand, is in good agreement with the typical intended use of the performance measure, which is as indicator that a slowly growing model/plant mismatch has now become so serious that it is time to deal with it. We will of course get a performance drop according to this measure also in some other situations. It may for example be increased disturbance levels or just changed disturbance characteristics that caused the drop rather than actually changed response behavior of the plant. As long as the economy is affected, it may be just as serious and also require some correcting action.

The detection of what has caused the observed performance drop will not be dealt with in this paper. Let us just mention that it may require further investigation, for example using a detection experiment as the one presented in [11] or other approaches as described in [4] and [5]. If there was a simultaneous detection of a problem in any of the base layer control loops that the MPC depends on, then that would naturally be suspected as the main cause of the performance drop. It is natural to assume that performance monitoring is used at control loop level, before you start using it on MPC level.

Having the performance measure, you can of course directly start using it just to record the performance and have it available when you would like to check how profitable the operation seems to be. But you would most probably not want to bother about it until there is reason to act. This means that you have to find suitable threshold values, so that you get a notification when a performance drop exceeds such a threshold. You would perhaps combine the measure according to (1) with other indices, for example as simple as the standard deviation of some controlled or manipulated variables, and apply some heuristic logics before giving the notification. In any case the selection of suitable threshold values may be tricky.

In the case illustrated in Fig. 2 there was a predefined required minimum profit level. Well, predefined or not, a reasonable required level may have been established during commissioning. This is probably the way the threshold for

the profit measure would be found in most cases. If we make a plot as in Fig. 2, we can also draw a conclusion on the required control accuracy for the critical variable. For the case shown there, you will need a standard deviation 0.4 or less. Otherwise you cannot reach that required profit level.

Considering a situation like the one in Fig. 2, you can conclude that increased standard deviation of the critical variable will inevitably lead to a performance drop. A performance drop may also occur if control accuracy is lost for some other variable in such a way that it prevents operation at the desired point. You can also have a drop in this profit measure without any loss of control accuracy, but in that case it is probably due to an operator choice to move away from the optimal operating point (or range). If you want to exclude the latter case, the criterion for notification could be that we have both a performance loss according to the profit measure and a loss in control accuracy for at least one supervised variable. We gave a suggestion how to choose the standard deviation threshold for the critical variable. For other supervised variables the threshold should be chosen with similar margin to the base-line case used in commissioning.

So far, we have not discussed any significance of how the profit measure evolves. But, if there is a steady trend of decreasing profit according to this measure, at which level should you react? If we foresee a gradually growing model/plant mismatch, this is exactly what you would expect to see. And, of course, it is very interesting to know when it would be most beneficial to update the model to agree with the new plant behavior. We can actually state that, if we make a few assumptions. These are of course simplifications, but the exercise may still be of value.

Assume that the model update will restore performance to its original level. Assume also that we have a slow trend of growing model/plant mismatch, and that this will repeat with roughly the same speed also after the model update. Assume also that we know the cost required to get the model updated. Concerning plant operation, this cost could be due to a larger back-off required during excitation, if the model update relies on closed-loop identification.

If we let P_c denote the base-line profit (as established at commissioning), the accumulated cost of performance loss up to time t will be

$$L(t) = \int_0^t P_c - P_t(s) ds. \quad (7)$$

This is represented by the upper shaded area in Fig. 4, where it has been decided to update the model at time t_i , as this loss has reached the value $L(t_i)$. Now let L_i denote the cost of updating the model. This experiment cost is assumed to be independent of when it is done. It turns out that the most beneficial time to make the model update, under the

given assumptions, is when the average total cost per time unit is equal to the present instantaneous loss. In other words, choose the instant t_i such that

$$L(t_i) + L_i = t_i [P_c - P_t(t_i)]. \quad (8)$$

This is easily understood geometrically from Fig. 4, where L_i is the lower shaded area. The average loss per time unit $[L(t)+L_i]/t$ will increase both if you choose to update the model earlier ($t < t_i$) and if you choose to do it later ($t > t_i$).

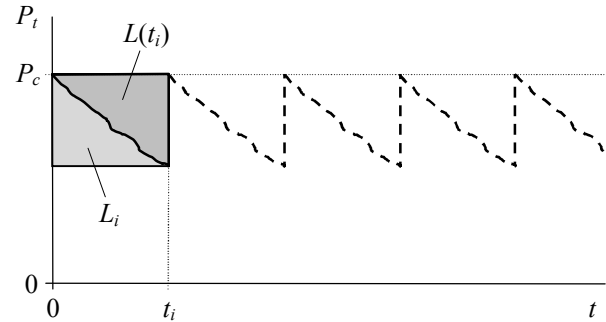


Fig. 4 Profit decreasing as a slow trend until the model is updated, The accumulated loss since last update is $L(t)$ and the cost of updating the model is L_i . The pattern is supposed to repeat itself as indicated dashed. The optimal instant t_i is indicated.

Suppose the required profit level is P_r . Then, if the measure drops to P_r before the time t_i is reached, the average profit including the cost for model updating, cannot reach the required level. If the gradually decreased performance is a perfectly linear trend, you will pick the t_i giving $L(t_i) = L_i$. In this case, with required profit level P_r , the maximum allowed cost for the updating is given by

$$L_i \leq (P_c - P_r) t_i / 2, \quad (9)$$

where t_i is the time it takes for the performance to drop from P_c to P_r . Otherwise your average profit will stay below P_r .

The conclusion from the discussion above is the following. You should get a notification on performance drop based on the present value of the profit measure (1), possibly in combination with increased standard deviation of at least one supervised variable. You should also get it in a situation with gradually decreased performance, when the condition (8) gets fulfilled.

IV. EXAMPLES

Let us shortly return to the example used in Fig. 2 and present two recordings, one with good performance directly after commissioning, Fig. 5, and one after some model/plant mismatch has occurred, Fig. 6. The critical variable is shown, and it has an upper constraint equal to 1.

We have the production rate as benefit variable z_1 with corresponding benefit factor $f_1 = 3.25$. The cost variables have been recast into one, z_2 , which equals the back-off distance, with corresponding cost factor $f_2 = 0.0325 * z_1$. The cost factor for constraint violation is $c = 3.25 * z_1$. The

required performance is 2.90. In Fig. 5, we have the production rate 1, and the recorded performance measure is 3.21. In Fig. 6, the control has deteriorated, and a larger back-off has been applied. Still, the constraint violation probability has increased. The production rate has also dropped to 0.91, and the recorded performance measure is now 2.85. That is below the required profit, so a notification of performance drop is due.

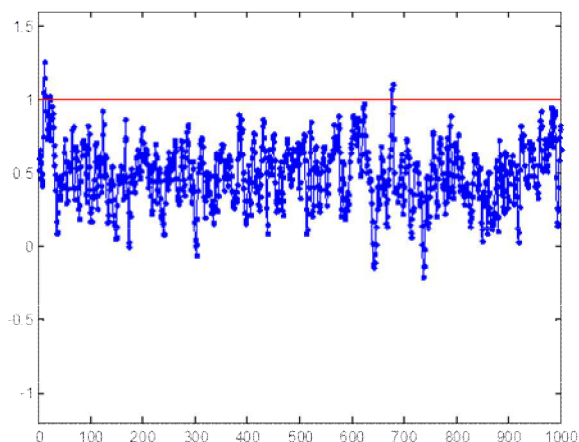


Fig. 5 The standard deviation of the critical variable is 0.2, and it is operated with a back-off of 0.49. The recorded frequency of violations is 0.8%.

The example above was constructed just to illustrate the principle and does not represent any known real process. In the next example we use a pulp digester simulator ([12], [13]) to demonstrate how the profit function and some related properties are influenced by a process change.

The pulp digester converts wood chips to pulp to be used for making paper. In this process, a mixture of wood chips and white liquor is heated under pressure to dissolve the lignin in the wood chips from the cellulose fiber.

The MPC was designed for hardwood for production of pulp with kappa 90. A main objective for control is to keep the kappa number at its set-point to produce pulp of the correct quality. Pulp with a kappa number above 93 has to be sold for a lower price, which is undesired. Hence, the MPC has an upper constraint defined for kappa=93. Five other PVs are only used in so called predict mode, meaning that they have neither set-points nor constraints. They are only used to improve the state estimation. In addition to these six PVs, four additional measurements have been added to the simulator for calculation of key performance indices. They indicate raw material flow rate, y_7 , production rate y_8 , consumed power y_9 , and rate of added chemicals y_{10} . They are not used in the MPC. Five MVs are also available. The sampling interval is 600 s.

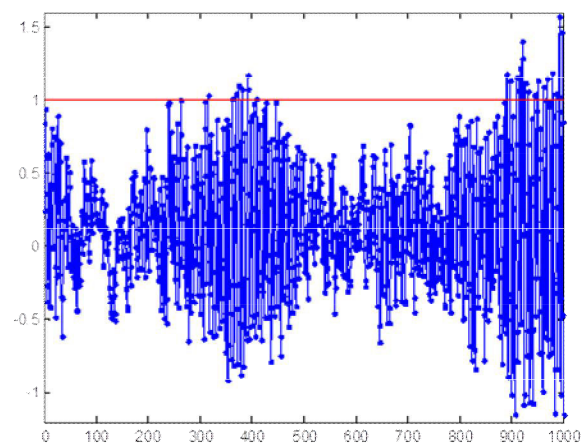


Fig. 6 The standard deviation of the critical variable has increased to 0.45, and the back-off has been increased to 0.9 (which is near the optimal for this standard deviation). The frequency of violations has increased to 2.6% in spite of that increased back-off.

Here the profit function P_t is a measure of the benefit of the production. It is obtained by multiplying the price for pulp with the pulp production rate. This value is then reduced by costs for used energy, used wood and added chemicals. Further it is reduced according to the lower value of the pulp that is outside the specification. This gives the following performance function

$$P_t = 2.5y_8 - y_7 - 0.0001y_9 - 100y_{10} - (1.0y_8)P_{viol}, \quad (10)$$

where P_{viol} is the relative frequency of kappa exceeding the value 93, which causes a decreased profitability due to the lower sales price.

Fig. 7 shows a simulation over 333 hours (2000 samples). Half way through, the wood feed is changed from hardwood to softwood. At the same time the feed rate is adjusted to keep the pulp production at the desired level. The set-point for kappa is 90, and an upper constraint is defined to keep kappa ≤ 93 .

Fig. 8 shows the performance indices during the simulation. The value at a specific sample is the average over a window, N_{win} samples backwards in time. Here the window length is chosen to have a length of 100 samples. The upper figure shows the profit function P_t , the middle figure shows the standard deviation for the control error, and the bottom figure shows the relative frequency of violation of the constraint for kappa.

V. CONCLUSION

It is not an easy task to monitor how well an industrial process performs when controlled by a model predictive controller. Some aspects of this have been discussed in this paper.

In addition to commonly used performance measures, as standard deviations for selected critical variables, we have proposed a performance objective in terms of a profit function. The profit function is based on economic measures which are derived from the value of the production and the cost for the production. Such a performance measure is intended to be more informative than the commonly used standard deviations. A threshold value for the monitoring function has been discussed. Preferably it should be based on what could be recorded directly after the commissioning. Two examples illustrate the use of the proposed profit function for performance monitoring.

REFERENCES

- [1] C. Zhao, H. Su, Y. Gu, and J. Chu, "A pragmatic approach for assessing the economic performance of model predictive control systems and its industrial applications," *Chinese J. Chem. Eng.*, vol. 17, no. 2, pp. 241-250, April 2009.
- [2] M. Bauer, I. K. Craig, "Economic assessment of advanced process control – A survey and framework," *J. Process Contr.*, vol. 18, no. 1, pp. 2-18, January 2008.
- [3] M. Bauer, I. K. Craig, E. Tolsma, and H. de Beer, "A profit index for assessing the benefits of process control," *Ind. Eng. Chem. Res.*, vol. 46, 2007, pp. 5614-5623.
- [4] R. S. Patwardhan and S. L. Shah, "Issues in performance diagnostics of model-based controllers," *J. Process Contr.*, vol. 12, pp. 413-427, 2002.
- [5] A. S. Badwe, R. S. Patwardhan, S. L. Shah, S. C. Patwardhan, and R. D. Gudi, "Quantifying the impact of model-plant mismatch on controller performance," *J. Process Contr.*, vol. 20, pp. 408-425, 2010.
- [6] B. Huang, S. Lu, and F. Xu, "Towards optimal MPC performance: Industrial tools for multivariate control monitoring and diagnosis," presented at International Workshop on Solving Industrial Control and Optimization Problems, Cramado, Brazil, April 6-7, 2006.
- [7] D. Wei, I. K. Craig, and M. Bauer, "Multivariate economic performance assessment of an MPC controlled electric arc furnace," *ISA Transactions*, vol. 46, pp. 429-436, 2007.
- [8] T. J. Harris, "Assessment of control loop performance," *Can. J. Chem. Eng.*, vol. 67, pp. 856-861, 1989.
- [9] T. J. Harris, "Interpretations of multivariate performance assessment indices," *J. Process Contr.*, vol. 19, no. 4, pp. 701-710, April 2009.
- [10] M. Molander, "Short note on economic performance measure," unpublished Autoprofit/ABB document, 9ADB003887-008, 2011.
- [11] A. Mesbah, X. Bombois, J. H.A. Ludlage, and P. M.J. Van den Hof, "Closed-loop performance diagnosis using prediction error identification," in *Proc. 50th IEEE Conference on Decision and Control and European Control Conference*, Orlando, Florida, 2011, pp. 2969 - 2974.
- [12] P.A. Wisniewski, F.J. Doyle, and F. Kayihan, "Fundamental continuous-pulp-digester model for simulation and control", *AIChE J.*, vol. 43, pp. 3175-3192, 1997.
- [13] P.E. Modén, "Continuous pulp digester simulator model", unpublished Autoprofit/ABB document, SECRC/AT/TR-2011/035, 2011.

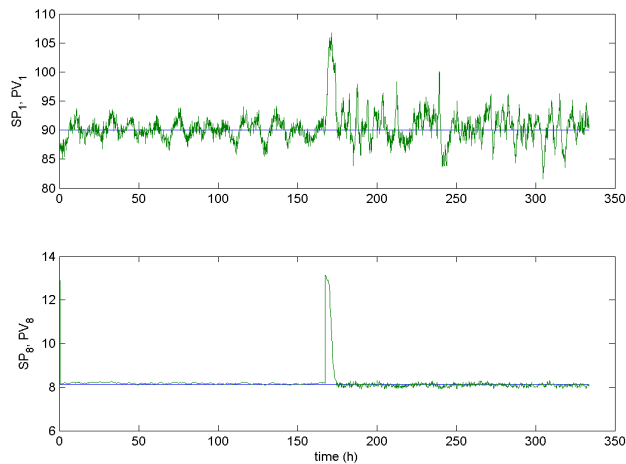


Fig. 7 Pulp digester simulation over 2000 samples. Halfway through, the wood feed is changed from hardwood to softwood. At the same time the feed rate is adjusted to keep the pulp production at the desired level. The upper figure shows the set-point and the measured values for the kappa number. The lower figure shows the pulp production rate.

This example shows that the performance of the pulp digester decreases when it uses a wood type that the MPC was not tuned for. This is particularly seen in the profit function. It can also be spotted in the increased standard deviation for the control error in kappa, and in the increase in relative frequency in the violation of the constraint for $\text{kappa} \leq 93$.

The profit function is here constructed from purely economic quantities, the value of the production reduced by the cost for the production. Therefore it may be fairly easy to find a threshold where the operation of the pulp digester is not profitable enough.

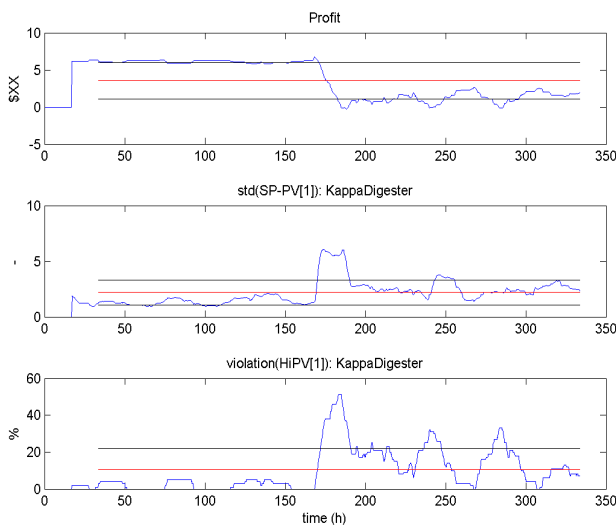


Fig. 8 Performance indices during the simulation. The upper figure shows the profit function P , the middle figure shows the standard deviation for the control error, and the bottom figure shows the relative frequency of violation of the constraint for kappa.