Proceedings of the European Control Conference 2007
Kos, Greece, July 2-5, 2007

Prediction of lacking control power in power plants using statistical methods

P.F. Odgaard & B. Mataji & J. Stoustrup

Abstract— **Prediction of the performance of plants like power plants is of interest, since the plant operator can use these predictions to optimize the plant production. In this paper the focus is addressed on a special case where a combination of high coal moisture content and a high load limits the possible plant load, meaning that the requested plant load cannot be met. The available models are in this case uncertain. Instead statistical methods are used to predict upper and lower uncertainty bounds on the prediction. Two different methods are used. The first relies on statistics of recent prediction errors; the second uses operating point depending statistics of prediction errors. Using these methods on the previous mentioned case, it can be concluded that the second method can be used to predict the power plant performance, while the first method has problems predicting the uncertain performance of the plant.**

I. INTRODUCTION

Monitoring of plant performance is an important task in optimizing power plants production. Since early detection of eventual problems or faults in the plant is essential to achieve the required plant efficiency. A large set of different problems is considered such as: badly tuned controllers, faults etc. Some examples of dealing with these problems are [1] which deals with performance monitoring of power plants. Monitoring of control loops are dealt with in a number of papers, a review of some these can be found in [2]. [3] and [4] deal with fault detection in general. Some examples on fault detection in power plants are [5], [6] and [7].

Another kind of performance problems, which are interesting to deal with, is due problematic combinations of operating conditions (load request, operator set references, known and unknown disturbances). Where each of the conditions individually are in the accepted region, nevertheless, the combination of these conditions decreases the performance of the plant or lead to failure of the plant. An example of such a problem can be found in the coal mills delivering dried and pulverized coal to the furnace in a coal-fired power plant. The problem occurs while high coal flow is requested at the same time, as the moisture content in the raw coal is high. In some cases this will result in not enough energy being available in the mill to dry the coal particles. Consequently the coal particles will be accumulated in the mill. In [8] a solution to this problem is proposed by bounding the achievable load.

Another approach to accommodate this problem is to inform the power plant operators with predictions of the future performance given the known conditions and planned load/reference changes. The predicted performance shall consequently be validated against the required plant performance for this given situation. However, the available models are uncertain which is a problem while predicting the future of the plant. In [9] and [10] statistical methods are developed to predict the uncertainty of the previous mentioned prediction. The first method uses statistics of recent windows of prediction, where a window is formed for each prediction length in question. The second method uses a bank of prediction error statistics depending on the operating point of the plant, as well as on prediction horizon. In this paper these two methods are applied to the mentioned coal-fired power plant, using data containing the problem due to the combination of high coal moisture content and high plant load. Two sets of experimental data are applied to these methods, one before the coal accumulation is occurring and one during the coal accumulation. These examples are used to see how well the future plant performance is predicted. A reliable prediction is necessary if one would like to use these predictions of the future plant behavior and performance.

In Section II the power plant dealt with in the paper is described. This is followed by a description of the used statistically based prediction methods of uncertain systems in Section III. In Section IV the prediction methods are applied to experimental data from the power plant. In this experiment conditions with high moisture content and high load are present. A conclusion is drawn in Section V.

II. PLANT DESCRIPTION

The proposed scheme for predicting the uncertain closed loop system is illustrated using measurements from a power plant. The model of the plant is a combination of a furnace model found in [11] and [12], and extensions consisting of a coal mill model, and controllers. An overview of the model structure can be seen in Fig. 1. The coal mill pulverizes and dries the coal dust, before it is blown into the furnace by the primary air flow. $\dot{m}_{\mathrm{c}}[n]$ denotes the actual coal flow, $\dot{m}_{\mathrm{c,ref}}[n]$ denotes the reference/requested coal flow, $\dot{m}_{\mathrm{pa}}[n]$ denotes the actual primary air flow, $\dot{m}_{\mathrm{pa,ref}}[n]$ denotes the reference/requested primary air flow.

P.F. Odgaard and J. Stoustrup are at Section of Automation and Control, Department of Electronic Systems, Aalborg University, Fredrik Bajers Vej 7C, DK-9220 Aalborg East, Denmark, {odgaard,jakob}@control.aau.dk
P.F. Odgaard and J. Stoustrup are at Section of Automation and Control, Department of Electronic Systems, Aalborg University, Fredrik Bajers Vej 7C, DK-9220 Aalborg East, Denmark, {odgaard,jakob}@control.aau.dk
B. Mataji is at Dong Energy A/S, Kraftværksvej 53, DK-7000 Fredericia, Denmark, bama@elsam-eng.com

ISBN: 978-960-89028-5-5

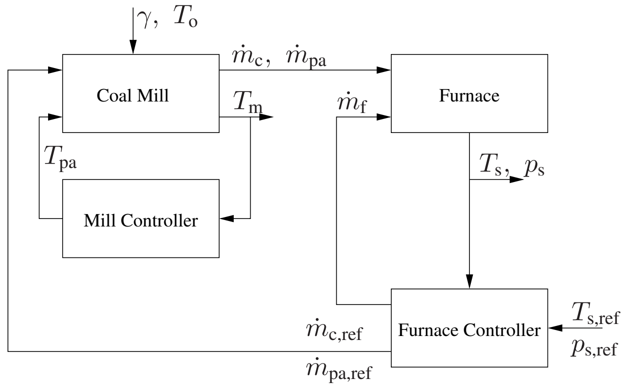

Fig. 1. Illustration of model structure

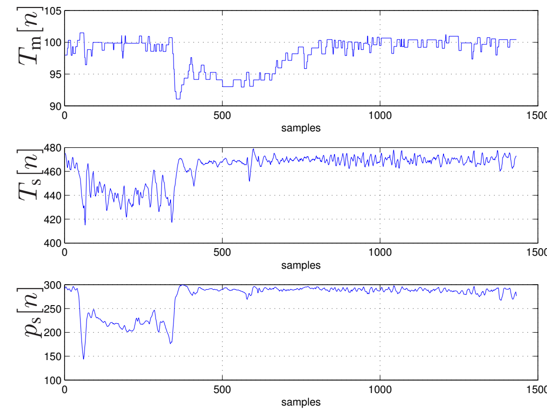

Fig. 2. Plots of the measurements of $T_m[n]$, $T_s[n]$ and $p_s[n]$.

Two disturbances to the coal mill is considered, these are the outside temperature, $T_o[n]$ and the coal moisture content, $\gamma[n]$. The temperature of the primary flow, $T_{pa}[n]$, which is used to dry and lift the coal dust in to the furnace, T_m, are used to manipulate the coal dust temperature to be at 100°C. In the furnace the coal dust is burned and the hot flue gas is used to heat water to pressurized steam. The steam temperature, $T_s[n]$, and pressure, $p_s[n]$, are used to control the plant, since references are given to those. This control results in coal flow and feed water flow, $\dot{m}_f[n]$, requirements.

The non-linear plant model is subsequently linearized and reduced to a 5 state linear model. The outputs of this linear model are coal mill temperature, $T_m[n]$, Steam temperature, $T_s[n]$, and steam pressure, $p_s[n]$. Controlled inputs are reference to steam temperature, $T_{s,ref}[n]$, and reference to steam pressure, $p_{s,ref}[n]$. Two other inputs to the model are not controlled, the first is the coal moisture, $\gamma[n]$, is estimated using the method presented in [13], and the second is the outside temperature which is provided in measurement set.

The experimental data used in this work is sampled with an interval of 60s at the power plant. The data contains load change from 85% load down to 65% load, at sample 65, and up again to 85% load at sample 340. The measured outputs can be seen in Fig. 2. The moisture content, on the other hand, is increasing during the experiment from 14% to 15.5% at the time of the second load change, consequently not enough energy is available to heat and evaporate the moisture from the pulverized coal. This can be seen by the plot of $T_m[n]$ which decreases below the evaporation point of the moisture. This drop in $T_m[n]$ is an example of a non-acceptable performance. A consequence is that these wet coal particles are too heavy to be lifted up to the furnace by the primary air flow. Therefore the coal particles are accumulated inside the coal mill, as a result a higher coal flow is requested by the furnace controller, however, this does lead to even more coal being accumulated in the coal mill instead of being blown into the furnace. In this case the moisture content drops again, resulting in more coal being blown into the furnace than requested. Such a situation could result in an overheating of the plant. A safety stop is consequently necessary. Stops of the power plant are highly costly, so these should be avoided if possible.

III. Statistical uncertainty prediction methods

These two statistically based predictors used in this paper are presented in [9] and [10]. These are based on the same general structure of the predictor, which is illustrated by Fig. 3. Where the system inputs and outputs are used to estimate the present state values, these are fed to the predictor together with system inputs and outputs in order to predict the expected values as well as the uncertainty bounds. $\hat{\mathbf{x}}[n]$ and $\hat{\mathbf{d}}[n]$ denote the estimated state and disturbance vectors for the time instance n. $\hat{\mathbf{y}}[n]$ is the vector of the predicted system output for the time n, $\epsilon_u[n]$ and $\epsilon_l[n]$ denote respectively upper and lower bounds on system prediction for the time n.

The observer and predictor (prediction model) will subsequently be described in more details.

The close loop model is uncertain with respect to the real system. Consequently an observer is introduced in order to estimate the value of the states at the sample time n.

$$\hat{\mathbf{x}}[n] = \mathbf{\Gamma}\left(\hat{\mathbf{x}}[n-1], \mathbf{u}[n], \mathbf{y}[n]\right), \tag{1}$$

where $\mathbf{\Gamma}$ is an operator representing the observer, and $\hat{\mathbf{x}}[n]$ is the estimated state vector at time n, $\mathbf{u}[n]$ is a vector of plant inputs and $\mathbf{y}[n]$ is a vector of plant outputs. The used type of observer depends on the application, which this scheme is applied to. In specific example and optimal unknown input observer is used, see [3].

The estimated states can be used to predict the state and the output vectors a number of samples/steps into the future. In these specific cases the reference is partly known in the future due prediction of the required plant production, such as power plants since the general power production is known one day ahead. The disturbance might be known

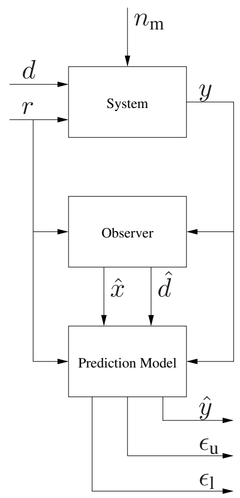

Fig. 3. Overview of the predictor structure, where the observer estimates the present states and in some cases the disturbances as well. The prediction model predicts the system behavior and uncertainties k steps into the future.

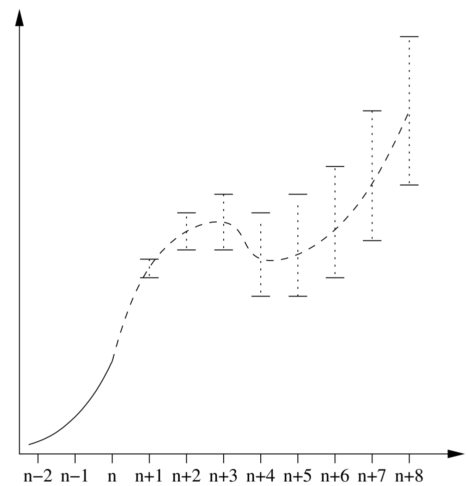

Fig. 4. Illustration of the uncertainty in the prediction. The uncertainty at each prediction step is increased as the number of prediction steps increases.

up to time n, I.e. by estimation. Subsequently these are denoted: $\hat{\mathbf{r}}[n]$ and $\hat{\mathbf{d}}[n]$. The k-step predictor of the output, $\mathbf{y}[n+k|n]$, and states, $\mathbf{x}[n+k|n]$, are computed by

$$\mathbf{x}[n+1|n] = \mathbf{f}_{\mathrm{m}}\left(\hat{\mathbf{x}}[n], \hat{\mathbf{r}}[n], \hat{\mathbf{d}}[n]\right), \tag{2}$$

where

$$\mathbf{x}[n+2|n] = \mathbf{f}_{\mathrm{m}}\left(\mathbf{x}[n+1|n], \hat{\mathbf{r}}[n], \hat{\mathbf{d}}[n]\right), \tag{3}$$

continue this recursive process until $\mathbf{x}[n+k|n]$ is computed. Then compute

$$\mathbf{y}[n+k|n] = \mathbf{g}_{\mathrm{m}}\left(\mathbf{x}[n+k|n], \hat{\mathbf{r}}[n], \hat{\mathbf{d}}[n]\right). \tag{4}$$

Now where the k-step predictor is defined, it is possible to define a k-step prediction error residual.

$$\psi[n+k|n] = \mathbf{y}[n+k] - \mathbf{y}_{\mathrm{m}}[n+k|n]. \tag{5}$$

This prediction residual defined in (5) can of course only be computed earliest at sample n.

As previously stated the model is assumed to be uncertain in relation to the real system. It also means that $y[n+N+1|n]$ is more certain than $y[n+N+2|n]$ validated in terms of the variance of $\psi[n+N+1|n]$ is smaller than the variance of $\psi[n+N+2|n]$. In other words the prediction is expected to be more uncertain as the prediction horizon increases. This is illustrated by Fig. 4. The predicted system value is drawn with the dashed line, (from sample $n+1$ to sample $n+8$), the measured system output value is drawn with solid line (sample $n-2$ to n). The uncertainty of the predicted values are marked by the vertical markings, with the small horizontal lines at the ends, the distance between these end markings represent the uncertainty for the specific predicted system value.

The question is how this uncertainty shall be dealt with. The mentioned predictor does predict the future expected output values of the dynamical systems.

In the context mentioned in the introduction (see Section I), it is as well interesting to predict a region in which output can be expected to be in, e.g. for the operator to see how the plant will perform given the prespecified conditions and references. These uncertainty predictions can be used to compute the upper and lower bounds of the prediction as:

$$\epsilon_{\mathrm{u}}[n+k|n] = \hat{\mathbf{y}}[n+k|n] + \epsilon[n+k|n], \tag{6}$$
$$\epsilon_{\mathrm{l}}[n+k|n] = \hat{\mathbf{y}}[n+k|n] - \epsilon[n+k|n]. \tag{7}$$

In this paper two different statistically based methods for predicting the uncertainty are compared. The first method computes mean and variance of a given number of the most resent prediction errors at the prediction steps in question, see [9]. The advantage of this method is to be found if the model uncertainty is strongly time varying but independent of operating points. The second method uses a bank of precomputed mean and variance of prediction errors at different operating points (references, disturbances etc.), see [10]. This method is preferable if model uncertainties are depending on the operating point. It is as well possible to adapt the uncertainty model bank to present prediction errors, as some time dependency can be included in the uncertainty models.

A. *The uncertainty predictor - method I*

In this approach the uncertainty of the prediction can be represented by a distribution depending on the references and eventual disturbances. For simplifications the prediction errors are assumed to be a normal random process, with a specific variance and mean depending on the prediction horizon, depending on the vector of references, $\mathbf{r}$, and the vector of disturbances, $\mathbf{d}$. $\epsilon[n+k|n]$ is the predicted uncertainty at sample $n+k$ given estimate at n.

$$\epsilon[n+k|n] = \Phi\left(\sigma_{k,n}, \mu_{k,n}\right), \qquad (8)$$

where Φ is the normal distributed random process, $\sigma_{k,n}$ is the variance of the k-step prediction error at sample n, $\mu_{k,n}$ is the mean of the k-step prediction error at sample n. These statistics are computed based on recorded prediction errors, where only the M most recent time samples are considered.

The variance $\sigma_{k,n}$ can be computed as

$$\sigma_{k,n} = var\left(\begin{bmatrix}\mathbf{r}[n-M+1+k|n-M+1] \\ \vdots \\ \mathbf{r}[n+k|n]\end{bmatrix}\right). \qquad (9)$$

The mean $\eta_{k,n}$ can be computed in a similar way

$$\mu_{k,n} = mean\left(\begin{bmatrix}\mathbf{r}[n-M+1+k|n-M+1] \\ \vdots \\ \mathbf{r}[n+k|n]\end{bmatrix}\right). \qquad (10)$$

These normal distributions can be used to compute uncertainty bounds on the prediction given a η confidence interval. I.e. the uncertainty bounds bound the possible system output values given probability of η.

B. *The uncertainty predictor - method II*

The second proposed scheme computes the statistics depending on the operational points. $\epsilon[n+k|n]$ is the predicted uncertainty at sample $n+k$ given estimate at n.

$$\epsilon[n+k|n] = \Phi\left(\sigma^k_{\mathbf{r}[n+k],\mathbf{d}[n+k]}, \mu^k_{\mathbf{r}[n+k],\mathbf{d}[n+k]}\right), \qquad (11)$$

where Φ is the normal distributed random process, $\sigma^k_{\mathbf{r}[n+k],\mathbf{d}[n+k]}$ is the variance of the k-step prediction error for reference at time $n+k$ and disturbance at time $n+k$, $\mu^k_{\mathbf{r}[n+k],\mathbf{d}[n+k]}$ is the mean of the k-step prediction error for reference at time $n+k$ and disturbance at time $n+k$.

a) The uncertainty model parameter bank: The statistics of prediction errors depend on the operating points, which can be considered as uncertainty model parameters are stored in a data base, ordered accordingly to the depending variables, e.g. $\mathbf{r}[n+k]$ and $\mathbf{d}[n+k]$ for k-step prediction. For each element in this model parameter bank, model parameters are attached. The simplest way to use these model parameters is to use the instance in the model bank, which is closest to specified depending variables. However, linear interpolation between model elements are obvious to use, see [10].

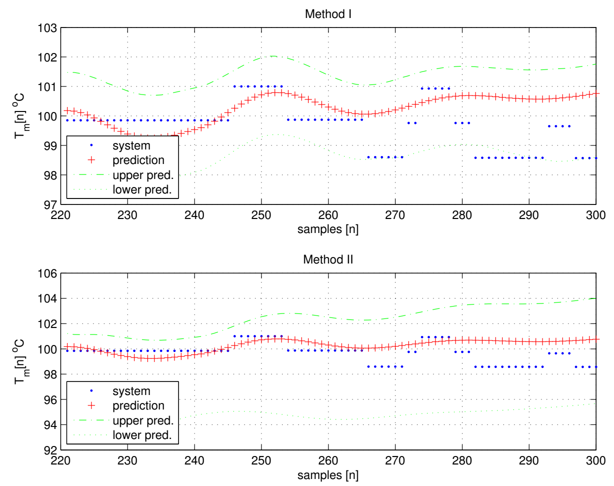

Fig. 5. Plot of prediction using the two statistical based methods for $T_{\mathrm{m}}[n]$, predicted from sample 220 and 80 samples into the future. The upper plot shows method I, and he lower plot shows method II.

IV. Experimental examples

In these experiments the linear model described in Section II and an observer are used to estimate the state values. In addition an optimal unknown input observer is used, see [3]. For the used uncertainty prediction methods a confidence interval at 90% is used.

The predictor of the uncertain system is applied to two sets of measured data from the plant, presented in Section II, and notice the sample time is 60s. The first set is sampled during the low load, where the system performance is predicted from sample 220 and 80 samples into the future. At this time accumulation of coal is not occurring. The second data set is sampled during the problem of coal accumulation, where the system performance is predicted from sample 621 and 80 samples into the future. Notice that in the plotted prediction the uncertainty bounds do not start together due to the introduction of the mean of the prediction error in the uncertainty models.

The first set of prediction can be seen in Figs. 5-7, in which the upper plot shows the prediction using method I and the lower plot prediction using method II. Fig. 5 illustrates the prediction of $T_{\mathrm{m}}[n]$. From this it can be seen that uncertainty bounds of method II covers the measured system behavior while method I does cover all measurement points except a few points between sample 280 and 290. This means that the uncertainty bounds are computed too narrow.

Fig. 6 illustrates the prediction of $T_{\mathrm{s}}[n]$. From this it can again be seen that uncertainty bounds of method II covers the measured system behavior while method I does not cover all measurement points, (230-246, 246-247, 278-284, 287-300).

Fig. 7 illustrates the prediction of $p_{\mathrm{s}}[n]$. From this it can again be seen that uncertainty bounds of method II

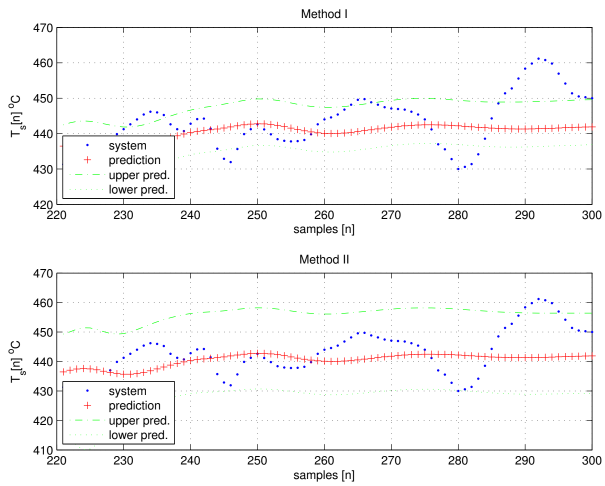

Fig. 6. Plot of prediction using the two statistical based methods for $T_s[n]$, predicted from sample 220 and 80 samples into the future. The upper plot shows method I, and the lower plot shows method II.

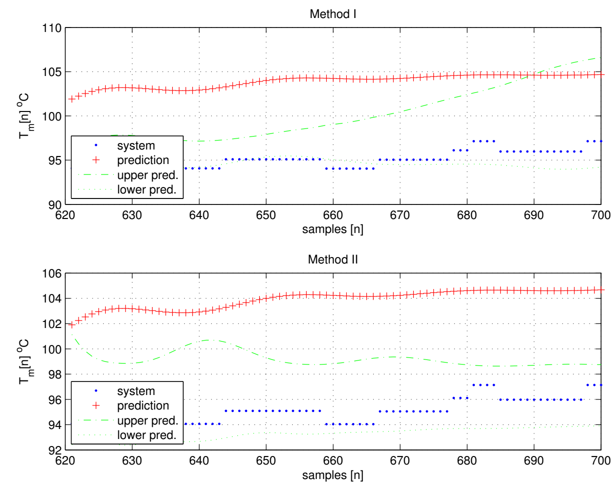

Fig. 8. Plot of prediction using the two statistical based methods for $T_m[n]$, predicted from sample 620 and 80 samples into the future. The upper plot shows method I, and the lower plot shows method II.

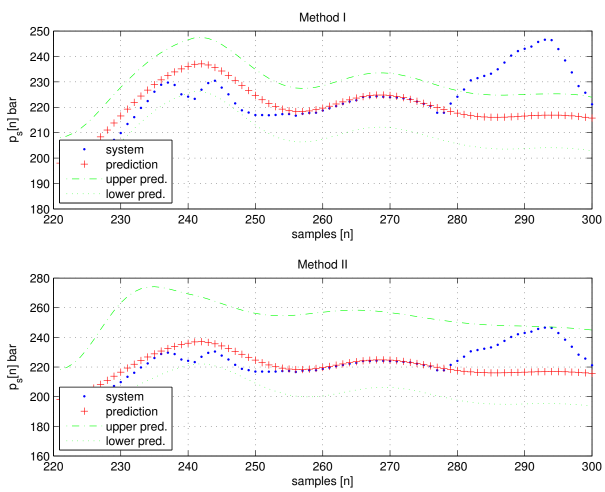

Fig. 7. Plot of prediction using the two statistical based methods for $p_s[n]$, predicted from sample 220 and 80 samples into the future. The upper plot shows method I, and the lower plot shows method II.

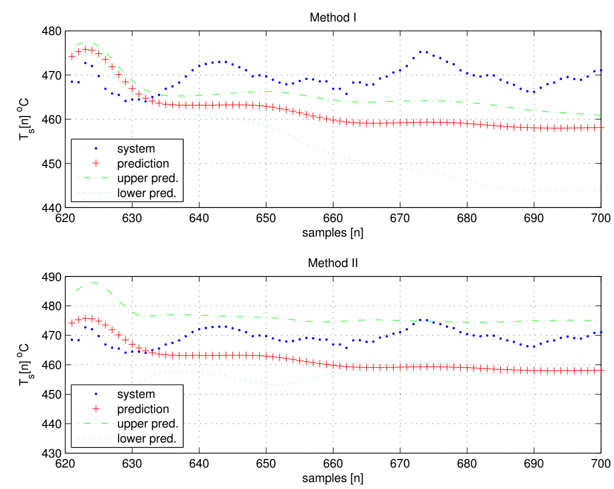

Fig. 9. Plot of prediction using the two statistical based methods for $T_s[n]$, predicted from sample 620 and 80 samples into the future. The upper plot shows method I, and the lower plot shows method II.

covers the measured system behavior while method I does approximately not the last 20 measurement samples.

From this set of experimental data and the prediction of system performance it can be seen that method I using statistics of recent prediction errors has problems covering the system behavior. On the other hand method II using statistics depending on operating points covers the system performance. This is a consequence of the model uncertainty depending more on the point of operation than on time.

The attention is now turned to the second set of measurement and predictions, from sample 620. These plots can be seen in Figs. 8-10. Fig. 8 illustrates the prediction of $T_m[n]$. From this it can be seen that uncertainty bounds of method II covers the measured system behavior while method I does not cover all measurement points.

Fig. 9 illustrates the prediction of $T_s[n]$. From this it can be seen that uncertainty bounds of both method I and method II cover the measured system behavior.

Fig. 10 illustrates the prediction of $p_s[n]$. From this it can be seen that uncertainty bounds of method I and method II covers the measured system behavior.

This set also illustrates the problem of method I and that method II predict uncertainties such that they cover the actual measured outputs. This means that these uncertainty predictions can be used by the operator to see how the future performance of the system given the specified operating conditions. This can be used in other cases than the

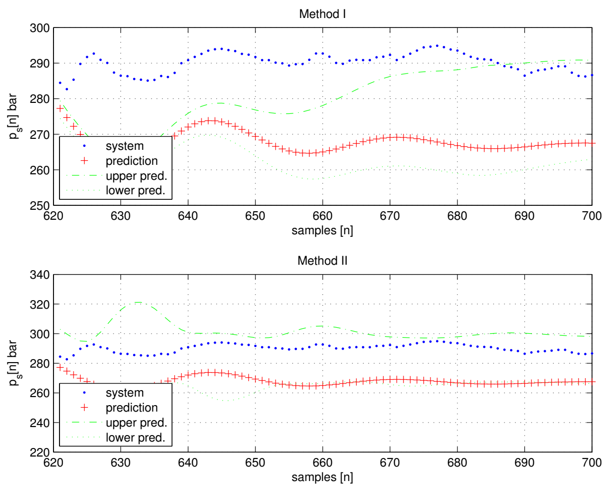

Fig. 10. Plot of prediction using the two statistical based methods for $p_s[n]$, predicted from sample 620 and 80 samples into the future. The upper plot shows method I, and the lower plot shows method II.

one described in this paper, e.g. start-up of the plant where specific state values shall be reached in a given time, in this case the operator can used these uncertainty prediction to see how the start-up process is proceeding and eventually take action if required.

V. Conclusion

Prediction of the performance of plants like power plants are of interest, since the operator can use these predictions to optimize the plant settings. In this paper the focus is put on a special case where a combination of high coal moisture content and a high load limits the possible plant performance, meaning that the requested load cannot be met. The available models are uncertain especially in this case. Instead statistical methods are used to predict uncertainty bounds on the prediction. Two different methods are used. The first rely on statistics of moving windows of recent prediction errors, the second one uses operating point depending statistics of prediction errors. Using these methods on the experimental data, it can be concluded that the second method can be used to predict the power plant performance given the set of operating conditions with high moisture content and high load, while the first method has problems doing so. In addition to the described problem, this uncertain plant performance prediction can also be of use during plant start-up.

VI. Acknowledgment

The authors acknowledge the Danish Ministry of Science Technology and Innovation, for support to the research program CMBC (Center for Model Based Control), grant no 2002-603/4001-93.

VII. References

[1] J. Ritchie and D. Flynn, "Partial least squares for power plant performance monitoring," in *IFAC control of power plants and power systems 2003*, Seoul, South Korea, September 2003.

[2] T. J. Harris, C. T. Seppala, and L. D. Desborough, "A review of performance monitoring and assessment techniques for univariate and multivariate control systems," *Journal of Process Control*, vol. 9, no. 1, pp. 1–17, February 1999.

[3] J. Chen and R. J. Patton, *Robust model-based fault diagnosis for dynamic systems*, 1st ed. Kluwer academic publishers, 1999.

[4] J. J. Gertler, *Fault Detection and Diagnosis in Engineering System*. Marcel Dekker, Inc., 1998.

[5] P. Odgaard and B. Mataji, "Fault detection in coal mills used in power plant," in *Proceedings of IFAC Symposium on Power Plants and Power Systems*, Kananskis, Alberta, Canada, June 2006.

[6] P. Odgaard, B. Lin, and S. Joergensen, "Observer-based and regression model-based detection of emerging faults in coal mills," in *Proceedings of 6th IFAC Synposium on Fault Detection, Supervision and Safety of Technical Processes, SAFEPROCESS 2006*, Bejing, P.R.China, August 2006.

[7] P. Odgaard and B. Mataji, "Observer based fault detection and moisture estimating in coal mill," 2006, submitted for journal publication.

[8] P. Odgaard, J. Stoustrup, and B. Mataji, "Preventing performance drops of coal mills due to high moisture content," 2007, to appear in proceedings of ECC 2007.

[9] ——, "Performance prediction of uncertain plants with disturbances using statistics of uncertainties: applied to a power plant," 2006, submitted for publication.

[10] ——, "Using reference trajectories to predicted uncertain systems: exemplified on a power plant," 2007, to appear in proceedings of ACC 2007.

[11] J. Bendtsen, J. Stoustrup, and K. Trangbaek, "Multi-dimensional gain scheduling with application to power plant control," in *Proceedings. 42nd IEEE Conference on Decision and Control*, vol. 6, Maui, USA, December 2003, pp. 6553 – 6558.

[12] ——, "Bumpless transfer between advanced controllers with applications to power plant control," in *Proceedings. 42nd IEEE Conference on Decision and Control*, vol. 3, Maui, USA, December 2003, pp. 2059–2064.

[13] P. Odgaard and B. Mataji, "Estimation of moisture content in coal in coal mills," in *Proceedings of IFAC Symposium on Power Plants and Power Systems*, Kananskis, Alberta, Canada, June 2006.