SIGNAL MONITORING USING ADAPTIVE THRESHOLD CLASSIFIER IN PULP & PAPER PROCESSES

Hiltunen, J.*, Tervaskanto, M.*, Kivikunnas, S., Pohjanheimo, L.**, Haltamo, J.[+]**

**Systems Engineering Laboratory, Department of Process and Environmental Engineering*
P.O.Box 4300, FIN- 90014, University of Oulu, Finland
Email: firstname.lastname@oulu.fi
***VTT – Technical Research Centre of Finland, Oulu, Finland*
[+]UPM-Kymmene, Kajaani, Finland

Abstract: Typically in pulp and paper processes raw material quality variation due to seasonal deviations as well as measurement drifts cause difficulties in setting the tight alarm thresholds for quality and control measurements. By using adaptive thresholds, more sensitive measurement range and thus reduced quality variation can be achieved. In this paper, new adaptive classification algorithm is proposed and validated using simulated and real mill data.

Keywords: Adaptive systems, Classification, Monitoring, Pulp industry, Paper industry

1. INTRODUCTION

There is a need in the pulp and paper industry for adaptive thresholds due to seasonal deviations in the properties of the raw materials. These properties have a great influence on the quality of end products. Especially in cold climates ice and water cause chip handling problems and difficulties in latter process stages. Also the huge temperature differences between summer and winter times especially in Scandinavia and in northern America cause variation to the wood fiber properties and other process variables. Several process stages needed in paper manufacturing process and different control actions done by several operators cause variation to pulp and paper properties as well.

There are a few reported studies on the effects of the seasonal deviations to thermomechanical pulping (TMP) and final paper quality. According to Cutshall (1990), one key point to reduce variation in the pulp and paper processes is to understand fluctuations in raw material properties caused by different circumstances in soil, wind, temperature etc.

Strand (1987) used factor analysis in the characterisation of the high yield pulps and Browne *et al.* (2003), expanded that study by using factor analysis in studying chip properties linked to the time of the year. They examined several properties from TMP pulp and showed that pulp properties depend on wood properties in two different ways. A portion of the variation in pulp properties is related to the brightness and another portion is due to physical properties of the wood. Seasonal and geographic variations have an effect on both of these wood properties.

Fuhr *et al.* (1998) studied the effects of the seasonal deviation to the TMP pulp, paper machine's wet end properties and final newsprint paper quality. They showed that seasonal deviations in raw material cause systematic deviation at least following variables: TMP fiber length, bark content, white water pH, headbox sulphite content and cross directional tear strength.

Tyrväinen (1997) showed that bonding properties and brightness of the TMP pulp change across the season. Wood (2000) stated that when wood-induced variations of pulp properties are present it is unlikely that refining control systems will be able to bring all the properties within specifications.

Thresholds for warning and alarming are usually kept unchanged in process monitoring. Indeed, they are tuned occasionally by operators typically without any specific rules. In this paper's approach, the thresholds are changed all the time but the relative proportions of normal, warning and alarming observations are set and kept unchanged. With this kind of adaptivity the variations in process raw materials and in process's actual performance are taken into account.

Classifying new observations into existing groups means that any new groups are not going to be formed as in clustering cases (James, 1985). This paper is not either going to proof that any of these assumed groups really exist within some data as in the case of variance analysis. The only target is to make a decision with each observation if it is a member of normal, warning or alarming groups while the bounds (thresholds) of the groups are changing. The classification itself is reduced to simple comparisons with these bounds and the main interest and focus is in the adaptation mechanism.

2. ADAPTIVE CLASSIFIER ALGORITHM

An Adaptive Threshold Classifier (ATC) is a novel procedure, which classifies the measurement values in different subgroups according to their adaptive thresholds. ATC can be used both in the evaluation of the performance of the processes and in the identification of the faults in measurement and control loops. The algorithm consists of superimposed calculation blocks in which the counters and adaptive thresholds are taken place.

The calculation proceeds in the following sequence:

- initialisation of the algorithm
- signal status validation and scaling
- classification
- counter updating
- counter scaling
- error signal calculation
- step length calculation
- threshold updating
- threshold overlap checking
- safety threshold checking

2.1 Initialisation of Algorithm

Initial thresholds should be placed as near the adapted values as possible and expert knowledge can be used in determining the preliminary alarming and warning threshold values.

If the initial values of the thresholds are not close enough to the adapted values, the thresholds fluctuates and the settling time grows excessively. The initial values could be calculated by using the standard deviation of the measurement values. A good starting point and a good practise (in the case of five subgroups) is as follows:

$$T_{au}(0)=m+3s \qquad (1)$$

$$T_{al}(0)=m-3s \qquad (2)$$

$$T_{wu}(0)=m+2s \qquad (3)$$

$$T_{wl}(0)=m-2s \qquad (4)$$

where $T_{cj}(0)$ are the initial values of the thresholds and the subscript c indicates warning (w) or alarming (a) and subscript j upper (u) or lower (l) areas, respectively. In the case of the measurements containing only three classes, only upper limits for the alarms and warnings exist. The symbols m and s denote the mean value and the standard deviation of the classified signal, respectively.

It is supposed that the classified signals are Gaussian distributed, containing white noise. The initial ratios of the specific area (IR_{ij}) could be determined by the probabilities for an observation belonging to the certain subgroup:

$$IR_{ru}=P(s(k)\geq m+3s) \qquad (5)$$

$$IR_{rl}=P(s(k)\leq m-3s) \qquad (6)$$

$$IR_{yu}=P(m+2s\leq s(k)<m+3s) \qquad (7)$$

$$IR_{yl}=P(m-3s<s(k)\leq m-2s) \qquad (8)$$

$$IR_{g}=P(m-2s<s(k)<m+2s) \qquad (9)$$

where $P(.)$ is a probability function for the observations of the signal $s(k)$ (k denotes discrete instants of time) located in the respective subgroups. The subscripts r, y and g denote red, yellow and green areas, respectively.

2.2 Signal Status Validation and Scaling

Due to the adaptive nature of the ATC algorithm, it is necessary to validate the condition of the signals prior to classification. The minimum requirement is to separate the "dead" signals, where there is no variation at all, and on the other hand the signals that have unstable oscillation. It is also important to check the operation points of the signals to interrupt calculations under known special situations (start up, maintenance etc.).

A practical way to identify "dead" measurements is to limit the minimum acceptable standard deviations of the signal. Limiting the maximum acceptable standard deviations identifies unstable signals. Furthermore, it is useful to limit the rate of change of the standard deviations. The absolute limit values indicating the signal to noise ratio (*SNR*) are also needed. The following equations may be used for all of the presented purposes

$$a_{min} * s(k-1) < s(k) < a_{max} * s(k-1) \tag{10}$$

$$b_{min} < SNR(k) < b_{max}, \tag{11}$$

where *SNR* is calculated simply by the following equation

$$SNR(k) = \frac{\mu(k)}{s(k)} \tag{12}$$

and a_{min}, a_{max} ,b_{min} and b_{max} are fixed parameters.

All the signals to be classified are scaled to zero mean and unit variance by using the common scaling equation

$$s(k) = \frac{r(k) - m(k)}{s(k)} \tag{13}$$

where $r(.)$ is the original signal and $s(.)$ is the scaled one. The number of observations in the equations (*10-13*) depends on the system's dynamics. Before the scaling the calculations of the mean and deviation values are based on the observations already validated from the reasonability standpoint, excluding abnormal situations of the system. Thus the scaling of the initial values becomes equal for the individual measurements.

2.3 Classification

The scaled values are fed to the threshold comparison block. The ATC compares the scaled signal value with the determined thresholds and adds one increment unit to the counter of the corresponding subgroup area. At the same time the sum counter is also incremented. The subgroup areas could be defined as follows:

The upper red area,

$$R_{ru}(k): s(k) = T_{au}(k) \tag{14}$$

The upper yellow area,

$$R_{yu}(k): T_{wu}(k) \le s(k) < T_{au}(k) \tag{15}$$

The green area,

$$R_g(k): T_{wl} < s(k) < T_{wu}(k) \tag{16}$$

The lower yellow area,

$$R_{yl}(k): T_{wl}(k) \le s(k) < T_{al}(k) \tag{17}$$

The lower red area,

$$R_{rl}(k): s(k) \le T_{al}(k) \tag{18}$$

It is also possible to define only three subgroup areas by excluding either the upper or the lower red and yellow subgroup areas. In the case of five subgroup areas the counter increments are made as follows:

$$C_{ij}(k) = C_{ij}(k-1) + inc \tag{19}$$

where $C_{ij}(.)$ is the subgroup area counter corresponding the latest signal value. The subscript *i* denotes red (r), yellow (y) or green (g) colours and *j* denotes upper (u) and lower (l) areas. In the case of green subgroup, the upper and lower areas are not separated. The increment unit *inc* is stated to be one. After increments all the counters are filtered to maintain their sum value constant. Indeed, the increment is filtered out so that only the mutual ratios of the counters vary.

$$C_{ij}(k) = aC_{ij}(k-1) \tag{20}$$

The filtering factor α is maintained constant during on-line calculations, but it is calculated off-line as follows:

$$a = \frac{C_{max}}{C_{max} + inc}, \; a < 1 \tag{21}$$

where C_{max} is a tuning parameter describing the maximum value of counters. It affects to the sensitivity of the threshold changes defining their frequency band.

After the counter increments and filtering the new subgroup areas are calculated as follows:

$$R_{ij}(k) = C_{ij}(k) / SC \tag{22}$$

where *SC* is the sum of subgroup area counters, which defines the adaptation speed. The slower the changes of the process are, the bigger the *SC* value should be.

A new subgroup area is then compared to the initial ratios of the subgroup areas producing an error signal $e_{ij}(.)$

$$e_{ij}(k) = R_{ij}(k) - IR_{ij} \tag{23}$$

Next, a step length for a threshold update are calculated based on the error signal

$$h_{cj}(k) = b * e_{ij}(k) \tag{24}$$

where b is a tuning factor. The value of the factor is depending on the filtering parameter C_{max}: high values of b presume high values of C_{max} and vice versa.

Finally, the thresholds are adapted as follows:

$$T_{cj}(k+1)= T_{cj}(k)+h_{cj}(k), \text{ when } j=u, \tag{25}$$

$$T_{cj}(k+1)= T_{cj}(k)- h_{cj}(k), \text{ when } j=l \tag{26}$$

When the ratios of the observation numbers differ from their initial values, the thresholds are moved so that the mutual ratios of observation numbers approach the preset ones.

2.4 Overlap and safety checking

It is possible that the lower threshold for some reason gains greater value than the upper threshold. These kind of faulty situations are checked and blocked out by using the following comparison:

$$T_{cl}(k)<T_{cu}(k), \forall\, k=0,1,2,3,...n \tag{27}$$

$$T_{wu}(k)<T_{au}(k), \forall\, k=0,1,2,3,...n \tag{28}$$

$$T_{al}(k)<T_{wl}(k), \forall\, k=0,1,2,3,...n \tag{29}$$

To ensure a general reasonability of the thresholds, the following comparisons are also necessary:

$$ST_{cul}<T_{cu}(k)<ST_{cuu}, \forall\, k=0,1,2,3,...n \tag{30}$$

$$ST_{cll}<T_{cl}(k)<ST_{clu}, \forall\, k=0,1,2,3,...n \tag{31}$$

where ST_{cjj} denotes the so called safety thresholds and they can be set by using expert knowledge of the acceptable signal ranges.

3. RESULTS & DISCUSSION

The ATC algorithm is based on the assumption that the data is Gaussian distributed containing white noise. That is an adequate assumption because most of the measurements in the pulp and paper manufacturing processes have a normal distribution (Cutshall, 1990). In the following results, both simulated Gaussian type data and experimental data from TMP -process and paper mill is used.

3.1 Simulation case using Gaussian data

In the Fig. 1, Gaussian type data with 60000 points is considered and divided into five different regions using the ATC algorithm. The middle region denotes the normal operation of the measured variable. The next regions upward and downward denote

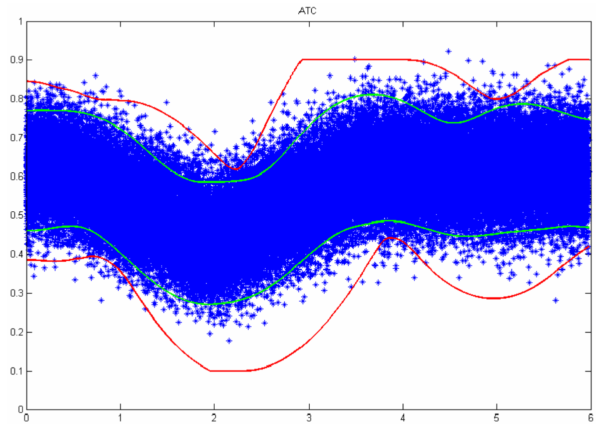

Fig. 1. Classified Gaussian type data.

'uncertain area'. If the point lies in the highest or in the lowest region, it can be considered as an alarm signal.

It can be seen from the Fig. 1, that there has been a downward slope between the points 500 and 2000. Respectively, between the points 2000 and 4000 upward slope has been simulated. The purpose of the slopes has been to test the effects of the drifts in the measurements and to test how effectively the ATC can be adapted to that situation. It can be seen, that the warning limits respond smoothly to the varying situations but the alarming limits are too sensitive to the drifts. It can be also seen, that the lower alarming has reached its safety threshold (0.1) approximately between the points 2000 and 2300. Respectively, the upper alarming limit has reached its safety threshold (0.9) approximately between the points 2900 and 4000.

3.2 TMP –plant case

In the next Figs. 2–5, the TMP reject refiner (RR) data with 4000 points is classified using ATC – algorithm. The data has skewed distribution and therefore it is not exactly Gaussian type. However, the following results show that the ATC is also useful if the data is not exactly normally distributed.

In the Fig. 2, the classification result is shown. The *SC* value used in TMP case and with Gaussian data has been 100000. In the Fig. 3, the percentual proportions of the green subgroup is shown. Initial counter value for the green area is 95000 (95 % of 100000 = 2s). In the ideal case, the green area counter should get values near the initial value. Due to two tails normal distribution curve upper and lower yellow initial counter values are 2350 in the Fig. 4 (2*2350=4700; 3s-2s=4,7 % of 100000). Respectively upper and lower red initial counter values are 150 (2*150=300; >3s =0.3 % of 100000) in the Fig. 5.

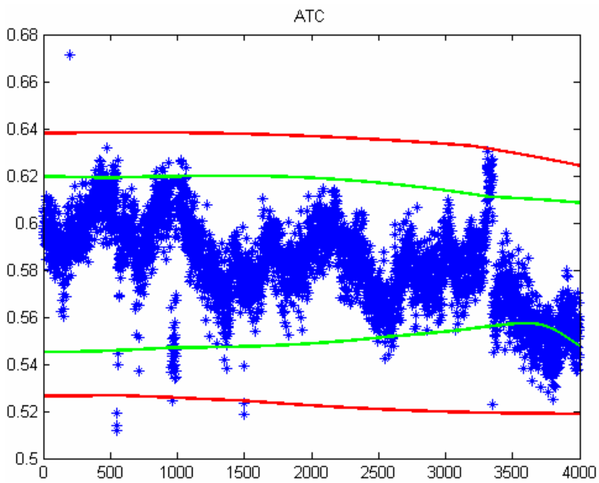

Fig. 2. The classification results of the ATC with RR data.

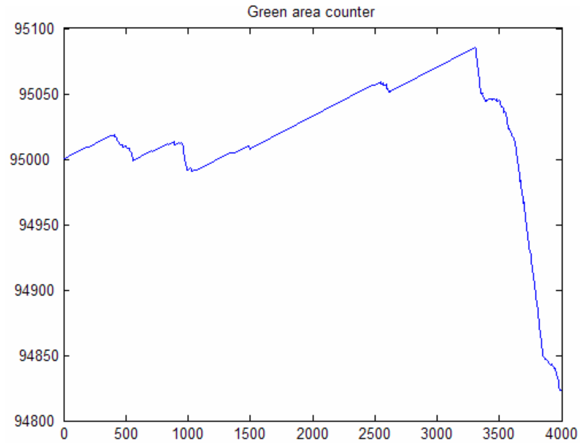

Fig. 3. The green area's percentual proportion.

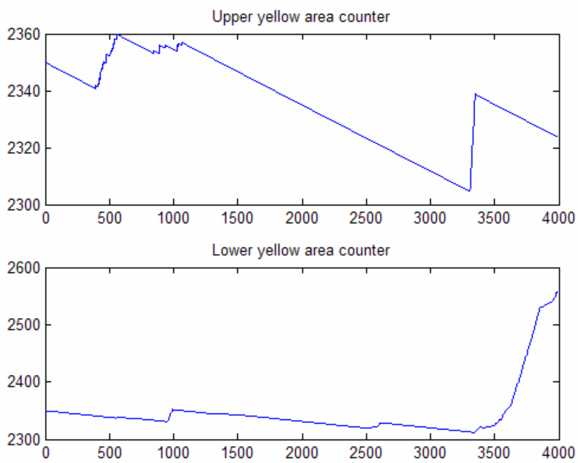

Fig. 4. The yellow area's percentual proportions.

In the Fig. 2, the trend is moving downward as well as the thresholds. Thus the system remains sensible to process faults.

It can be seen from the Figs. 3-5, that the counter values stay fairly close to the initial values. After the point 3300 the lower yellow area counter is

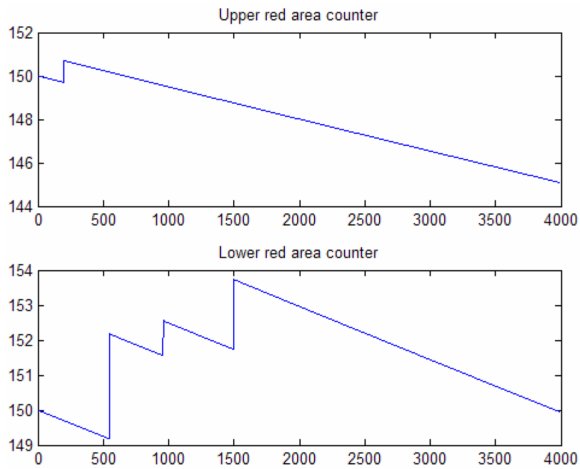

Fig. 5. The red area's percentual proportions.

increasing and the green area counter is decreasing strongly due to downward shift in the trend values. The same readjustment can be seen from the lower warning threshold in the Fig. 2.

3.2 *Paper machine performance monitoring*

This section describes experiences of using ATC with hierarchical monitoring methods that are capable to refine subprocess performance data into overall process status information. ATC was applied to implement adaptive limits for the index calculation and scaling.

The basis of the methodology is in hierarchical division of the plant into controllable subparts: process, subprocess, function and unit. This division follows the practical needs in process monitoring and control environment. Process disturbances are easier to diagnose when there are distinct areas to explore and as a consequence, recovery actions can be reached fast. In the unit level, several calculation and scaling methods were applied. They contain, for example, operations to input control, output standard deviation, rate of change, absolute and relative deviations to the calculation machinery following the hierarchical structure. Unit level indices are aggregated to their corresponding function node by taking the maximum of their absolute value. Function level values are averaged to their top level subprocesses. The same principle is adopted in the process level calculation. Top level process node can serve as input to other higher level monitoring systems. All calculations, beginning from unit level are performed in a normalised scale.

In this case, a prototype monitoring system for a SC paper machine at UPM-Kymmene Kajaani Mills in Finland was developed. The prototype application was designed and implemented by following client/server architechtual basis. The system contains a backend database and a Graphic User Interface. A more detailed description of the system can be found in Kivikunnas *et al*. (2002).

During the testing of the basic performance monitoring system, a need for adaptability to slow changes in the process was identified. Add-on functionality was specified to the performance monitoring tool. Scaling limits should be calculated so that the majority of values fall into the specified range without loosing the ability to detect significant process performance changes. This somewhat vague requirement was met by using ATC and strongly tuning the system parameters by following the guidelines described in section 2. The actual implementation was done with the same design principles and tools as the basic performance monitoring system.

Two parallel performance monitoring calculation systems were implemented: one without adaptivity and another one with ATC-properties. These were run for a six month test and gradual development period. Tests indicated that the intended functionality, adaptability to slow changes, was mainly obtained but the price for it at this stage of development was an extensive amount of tuning work. Main findings of this case can be summarised as follows:

- In a large-scale system, the tuning of the parameters of the classifier is a time-consuming exercise. Especially the settling time is difficult to control and needs deeper analytical approach before practical implementations. A tuning and configuration tool with tuning related knowledge management capabilities could boost its exploitation.
- During the project ATC proved to be very useful in defining fixed limits for performance value calculations. It rapidly indicated bad initial limits and errors in the calculation parameters. This finding generated an idea of using ATC with batch history data for preliminary threshold setting. The time schedule of the project however did not make possible the implementation of a test system and validation of the idea.

4. CONCLUSIONS

The ATC algorithm for threshold calculation and adaptation were developed. At first, the algorithm was tested using simulated Gaussian data to ensure its statistical properties. Secondly, two full-scale industrial applications were presented and discussed.

According to the applications, the ATC algorithm was proofed to be both practical and useful. It can be used as an on-line or off-line tuner of the signal thresholds. In the first case, the applicable rate of adaptation asks more workload, but because of the clear nature of the parameters of the algorithm it is however straightforward and easy to understand. For defining only the best constant values of thresholds (the case of off-line tuning), the presented ATC algorithm was even more simple and effortless to use.

REFERENCES

Browne, T., K. Miles, D. McDonald, and J. Wood (2003). *Multivariate analysis of seasonal pulp quality variations in a TMP mill*. 89th Annual meeting, Montreal, Quebec, Canada, 27-30 Jan. 2003, session 5C-4, 9 p.

Cutshall, K. (1990). The Nature of Paper Variation. *Tappi Journal*, **73**(6), pp. 81-90.

Fuhr, B.J, D. Henry, G. Leary, and G. Smith (1998). Seasonal variations at a mechanical newsprint mill. *Pulp & Paper Canada*, **99**(2), pp. 45-49.

James, M. (1985). *Classification algorithms*. John Wiley & Sons, New York. 210 p.

Kivikunnas, S., J. Koskinen, R. Lähdemäki and P. Oksanen (2002). *Hierarchical navigation and visualisation system for paper machine performance monitoring*. European Symposium on Intelligent Technologies, Hybrid Systems and Their Implementation on Smart Adaptive Systems (eunite 2002). Albufeira, Portugal, 19–21 September, 2002. Aachen: Verlag Mainz. P. 80 + Paper on CD-ROM.

Strand, B.C. (1987). *Factor analysis as applied to the characterization of high yield pulps*. Tappi Pulping Conference, Washington D.C., November 1-5, Book 1, pp. 61-66.

Tyrväinen, J. (1997). *Influence of Seasonal Wood Variation on TMP Properties, Bleach Consumption and Newsprint Runnability*. Preprints, 1997 International Mechanical Pulping Conference, Stockholm, Sweden, 9-13 June 1997, pp. 213-225.

Wood, J. (2000). *Wood-Induced Variations on TMP Quality – Their Origins and Control*. TAPPI 2000 Pulping/Process & Product Quality Conference. 5-9 November 2000, Boston, MA, USA.