A Calculation Method to Evaluate the Control Effect by PLS Regression Analysis

—How Should We Evaluate the Real Control Effect ? —

Yoshitatsu Mori^{*2}

Core Technology Laboratory, Oji Holdings Corporation^{*1}

Abstract

In the pulp and paper manufacturing process, numbers of material properties and end product qualities are changing steadily with time. Therefore it is considerably difficult to judge the variance of the production cost or the improved effect etc. through the operation data properly and fairly, because of their disturbances, when they are so tiny amount.

This paper describes a new method to evaluate the effect of the production cost and control effect etc. more accurately when a new control function was introduced or when process operation methods were changed for some purpose in the production processes.

Here, as an example, a case of a pulp bleaching process was taken up.

The proposed method is follows :

At first, find out the reliable multiple linear regression equation by using statistic analysis PLS (Partial Least Squares) regression method, for example, between the production cost as dependent variable and some primary explaining variables relates to the bleaching process. As it is well known, because the PLS analysis result is superior to the common regression analysis by MLR (Multiple Linear Regression) method in the stability and reliability. Sometimes, the regression coefficients calculated by MLR are tending to be inappropriate by the effects of over fitting or their collinearity.

Next, to get rid of the random variance and disturbance effects, correct and compensate the targeted “production cost” appropriately by “approximate linear correction method” so as to equalize the manufacturing conditions and incoming and outgoing material properties, utilizing the regression coefficients that were obtained by PLS regression model in advance.

As a result, it become possible to judge the control effect fairly and acceptably, even if the variance of the production cost is tiny change or the each explaining variable’s operating condition levels were changing intricately with each other.

Keywords : statistics analysis, PLS, production cost, pulp bleaching process

1. Introduction

When operating process is modified or new operating system is introduced in a mill, it is not easy to evaluate

^{*1}11-10-6 Shinonome, Koto-ku, Tokyo 135-8558, Japan

^{*2}E-mail : mori1yoshitatsu@oji-gr.com

how that change affects the operation, especially when its effect is small.

In this analysis, operations before and after the change are monitored for a long period of run generally. At that time, they are compared with each other.

Let us take kraft pulp bleaching process as an example, here. When you calculate the effect of process change on production cost, strict comparison can be done only when variables like qualities of pulp coming into and going out the process should be kept identical for the two comparing operations.

However, in actual operation, the quality of pulp coming into the process is varying over time, and the quality of bleached pulp is also fluctuating within its acceptable range. Furthermore, as the specification of pulp quality is changing from time to time, so its quality follows.

Then, how could we evaluate the real effect of the process change, using those operational data of long run, with reasonable acceptance. To answer the question, we applied PLS (Partial Least Squares) regression method of statistic analysis.

2. Multi-sequential bleaching process of pulp

Kraft pulp is produced from wood chips by digestion process. The resulting pulp is dark brown and is bleached to high brightness using bleaching chemicals in sequential stages like oxygen delignification followed with multi-sequential bleaching. We took hardwood bleached kraft pulp (LBKP) as our sample. The flow diagram of ECF multi-stage bleaching of our pulp is illustrated in Fig. 1.

This bleaching sequence is abbreviated as D0-Eop-D1. In the first D0 stage, coloring matters in pulp are decomposed by chlorine dioxide solution (ClO_2). Those components are extracted with sodium hydroxide (NaOH), and the pulp is further bleached by the cooperative work of

oxygen gas and hydrogen peroxide in Eop stage. Finally, the pulp is bleached to specified brightness by chlorine dioxide solution (ClO_2) again in D1 stage. The brightness of pulp coming into D0 is about 40% ISO. By passing through the bleaching stages, the brightness increases and reaches to about 85% at the end. At the same time, saving chemical cost in bleaching is an important concern.

Incidentally, various online sensors are installed in this bleaching line like those for measuring brightness and Kappa number. Pulp brightness is monitored both at the inlet and the outlet of the process. Kappa number is measured at the inlet and K number is controlled at the outlet of the process. Every stage is maintained at about 65°C which is its reaction temperature. The total retention time of pulp through the process is relatively long and is from 3 to 5 hours, depending on production rate.

3. Production cost and advanced control

To increase productivity in chemical plants, an advanced control called MPC (Model Prediction Control)¹⁾ has been introduced in many industries. It was reported that the merit of introducing the concept was about 3 % reduction of the production cost²⁾.

In our ECF bleaching process, bleaching chemical cost is calculated everyday on yen (¥)/A. D. kg basis, and is controlled as one of important operating concerns. The value is ¥ 2-3/A. D. kg.

If advanced control is introduced, the saving in bleaching chemical cost, which will be shorted as bleach cost from now on, will be ¥ 0.06-0.09/A. D. kg (= ¥ 2-3/A. D. kg $\times$ 3 %). As it is very small and will be affected by such disturbances as fluctuation of incoming pulp quality as well as finished one, it is very difficult to evaluate real merits by advanced control itself.

So, we studied to evaluate more accurately the effect of

Fig. 1 ECF bleach process

system change by statistical analysis based on data obtained in long run operation, eliminating influences caused by disturbances in variables like pulp quality.

4. Factors affecting to total production cost in a bleaching process

We monitored the bleach cost (average of a day) data in a mill operation for the period of about 400 days, which was divided into two terms. In one term, 230 days, the process was manually operated and in the other term, 170 days, the process was advanced controlled. The distributions of the bleach cost are shown in Fig. 2.

The vertical axis is for bleach cost. Data under the manual operation were marked with ● in the left part and the right part represents data under the controlled operation, marked with □. The horizontal axis is dislocated a bit arbitrary so that many marks could be separated each other to be clearly identified. The data in each part were averaged and the result was shown with a horizontal line. The both sets of data were vertically distributed in a fairly wide range. When the averaged values were compared with each other, the controlled operation gave less bleach cost though the difference was considerably small. As operating variables like pulp quality coming into and going out the process are always varying, simple comparison between the two averages did not represent the real effect of the an advanced control, and was not reliable enough.

We understood that variances in operating variables, which would be described as follows in ① to ⑥, should be taken into account, and influences caused by their time-dependent changes should be corrected to evaluate real bleach cost. They were :

① Variance of pulp brightness coming into D 0 stage

It fairly varies depending on preceding operations. This fluctuation influences following bleaching of pulp and affects bleach cost.

② Variance of brightness of finished pulp after D 1 stage

Its target brightness is changed occasionally according to product instruction. It also deviates from the targeted value in normal operation. The deviation affects to bleach cost.

③ Variance of Kappa number of pulp coming into D 0 stage

It varies, like pulp brightness as mentioned in ①, depending on precedent operations. This fluctuation affects bleaching of pulp that follows and then to bleach cost. Kappa number is an indication of residual lignin in pulp.

④ Variance of K number (permanganate number) of pulp after D 1 stage

Sometimes, K number of finished pulp fluctuates. It is caused when chemicals are not properly added in the process, and bleach cost is affected. Beside, K number is an indication of residual hexenuronic acid that will induce discoloration trouble (color reversion) of finished product. In addition to them,

⑤ Change in production rate

Bleaching towers have their own volumes. If a production rate is increased, retention time in every tower becomes less, and it gets difficult to bleach pulp up to targeted brightness. At the same time, load at washing filter at every stage increases, therefore its efficiency goes down, more chemicals are required and bleach cost gets worse.

⑥ Variance of ratio of wood chips species fed to digester

Here, wood chips of more than 10 wood species are stored separately and available in wood yard. They are selected and mixed as instructed and fed to digesters. Their bleachability is dependent on conditions of wood chips such as wood species, place of origin and state of storage. They make pulp behave differently in bleaching performances and then the wood species and those blending ratio affect to bleach cost.

To understand their effects, one factor, ⑤ production rate, was taken as an example. Bleach costs were plotted

Fig. 2 Operation mode vs. Bleach cost (direct analysis with original data)

Fig. 3 Production rate vs. Bleach cost (● : Manual data, □ : Control data)

against production rates in Fig. 3. This kind of study is often done in mill operation.

Here, they were grouped and marked differently by shapes with regard to the modes of operation as in Fig. 2 (● was for manual operation and □ was controlled operation).

Then, regression lines by least-squares method for the two groups were drawn in Fig. 3, one for the manual operation and the other for the controlled operation. Though data were scattering, the bleach cost tended to increase along with the production rate as indicated by the regression lines both of which had positive gradients. The fact that the data of the controlled operation is distributed below those of the manual operation data in whole. And the regression line of the former also located below that of the latter suggested that the controlled operation had definite improvement though it was small.

The production rate data in Fig. 3 varied from 400 to 750 tons/day. From that, it was understood that there was some relationship between bleach cost and production rate. However, bleach cost is affected by other factors described in Chapter 4 from ① to ⑥ as well. Therefore rational correction should be made with these influential factors to evaluate the real effect of the controlled operation.

5. Correction of production cost by regression analysis

When new control system is introduced or operating process is modified, production costs, for example, one of which is bleach cost, are affected. To estimate their effects with high reliability, we tried to use multi-regression analysis as a method of cancelling influences caused by time-dependent changes of operational variables like those of raw material qualities and final product qualities. From this standpoint we devised a method to correct the production costs by multiple linear regression analysis.

As the first step, bleach cost was taken as an objective variable, and factors which were supposed to be influential to it were selected as explanatory variables. The relation-

ship among those variables would be represented with a multiple linear regression equation in more reliable form.

After their coefficients of regression were calculated, bleach cost was corrected as the second step adequately, with a linear approximation model, using those coefficients, to get a more accurate value.

In our case, an objective variable was bleach cost. Also, explanatory variables were : ① brightness of incoming pulp, ② increase in pulp brightness (Δ brightness) which was difference in brightness between final bleached pulp and incoming pulp, ③ Kappa number of incoming pulp, ④ K number of final pulp, ⑤ production rate, and ⑥ blending ratio among wood species. Here, as ten kinds of wood were blended in operation, a blending ratio of wood species had ten variables. So, the number of explanatory variables was 15 in total. All data of variables used for analysis were corresponded to data in Fig. 2 in the same course of time. They were highly reliable data as every value was calculated as average of one-day operation as routine control.

As data were collected in a long term, prices of chemicals might vary during monitoring. So, we fixed a representative unit-price for every chemical in our analysis to exclude the effect of price changes.

6. Multiple linear regression (MLR) and PLS regression

Let us briefly explain statistical analysis method. Multiple linear regression (MLR), which utilized frequently, uses a least squares method. Its regression equation is expressed in the equation (1).

Y is an objective variable. X_j ($j = 1, 2, \dots, n$) are explanatory variables (input variables). $a_1, a_2, \dots, a_n$ are regression coefficients. a_0 is a constant, and n is the number of explanatory variables.

• Multiple linear regression equation :

$$Y = a_0 + a_1 \cdot X_1 + a_2 \cdot X_2 + \dots + a_n \cdot X_n \quad \dots\dots(1)$$

In case where strong interrelations exist among explanatory variables, some troubles like emergence of instability of estimate values caused by multi-collinearity or

over-fitting occurs as a number of variables increases. Then, some of regression coefficients in the regression equation become unreasonable, which is a disadvantage of MLR.

PLS (Partial Least Squares) regression method was developed to overcome those defects by Prof. Wold (father and son) of Umea University in Sweden in the 1980 s, and was used in a field of analytical chemistry as "Chemometrics". PLS regression method^{(3), (4)} is an intermediate concept between "Principal component analysis" and "MLR". By introducing "Latent variables", it reduces a number of apparent explanatory variables by contraction operation and uses latent variables as explanatory variables (input variables). So, it can be free from a problem of non-independency among input variables. It makes the variance of latent variables maximum and, at the same time, makes the correlation between an objective variable and latent variables greatest possible. It helps to find a more stable regression equation than MLR by stepwise procedure.

PLS regression is expressed in the equations (2) and (3). Y in the equation (2) is an objective variable. T_i ($i=1, 2, \dots, r$) is a latent variable. b_i is a regression coefficient of T_i and b_0 is a constant.

A latent variable T_i in the equation (2) is expressed by the equation (3). X_j ($j=1, 2, \dots, n$) is a explanatory variable (input variable) in the original. T_i consists of those input variables and their weight w_{ij} ($i=1, 2, \dots, r, j=1, 2, \dots, n$) and its constants w_{i0} ($i=1, 2, \dots, r, j=0$). Please note a following relation : $r \leq n$.

• PLS regression equation :

$$Y = b_0 + b_1 \cdot T_1 + b_2 \cdot T_2 + \dots + b_r \cdot T_r \quad \dots\dots(2)$$

The first latent variable :

$$T_1 = w_{10} + w_{11} \cdot X_1 + w_{12} \cdot X_2 + \dots + w_{1n} \cdot X_n$$

The second latent variable :

$$T_2 = w_{20} + w_{21} \cdot X_1 + w_{22} \cdot X_2 + \dots + w_{2n} \cdot X_n \quad \dots\dots(3)$$

.....

The r-th latent variable :

$$T_r = w_{r0} + w_{r1} \cdot X_1 + w_{r2} \cdot X_2 + \dots + w_{rn} \cdot X_n$$

$$\begin{aligned} Y &= (b_0 + b_1 \cdot w_{10} + b_2 \cdot w_{20} + \dots + b_r \cdot w_{r0}) + \\ &\quad (b_1 \cdot w_{11} + b_2 \cdot w_{21} + \dots + b_r \cdot w_{r1}) \cdot X_1 + \\ &\quad (b_1 \cdot w_{12} + b_2 \cdot w_{22} + \dots + b_r \cdot w_{r2}) \cdot X_2 + \dots \\ &\quad \dots + (b_1 \cdot w_{1n} + b_2 \cdot w_{2n} + \dots + b_r \cdot w_{rn}) \cdot X_n \\ &= c_0 + c_1 \cdot X_1 + c_2 \cdot X_2 + \dots + c_n \cdot X_n \quad \dots\dots(4) \end{aligned}$$

Latent variables $T_1, T_2, \dots, T_r$ in the right hand side of the equation (2) are substituted with the corresponding $T_1, T_2, \dots, T_r$ in the equations (3). Then, the equation (4) is obtained which consists of original input variables and their weights. It is similar to the equation (1) which consists of original input variables and helps us intuitively understand the meaning of regression equation. $c_1, c_2, \dots, c_n$ are coefficients of regression obtained by unwinding latent variables and c_0 is a constant. Whether they are good

enough can be evaluated in the same manner as in MLR.

The number of latent variables r used in the algorithm of PLS regression can be as large as n which is the number of the original input variables. A latent variable which is firstly calculated is the first latent variable. A latent variable which is calculated for the residual component obtained by PLS regression equation using the first latent variable is the second latent variable. So, there are the third latent variable, the forth latent variable and so on. PLS regression equation using as many a number of latent variables as that of explanatory variables is theoretically the same to MLR, though some difference may occur due to accumulation of calculation errors.

The corrective operation to bleach cost described in Chapter 5 as the second step is done by applying $c_1, c_2, \dots, c_n$ which are coefficients in the equation (4) to the equation (5), which is a linear approximation model. Y_m is a value of the objective variable to be corrected. X_{jm} is an observed value of explanatory variable X_j at the time of a sequential number of m and it has to be adjusted to a value under the same operational condition, which is $X_{j(\text{Base})}$. Note : $j=1, 2, \dots, n, n$ is a number of explanatory variables and m is a sequential number of time.

$$Y_m \text{ (after corrected)} = Y_m \text{ (before corrected)} -$$

$$\sum_{j=1}^n \{c_j \cdot (X_{jm}) - X_{j(\text{Base})}\} \quad \dots\dots(5)$$

7. Analysis with MLR

A set of data of variables specified in Chapter 5 as from ① to ⑥ was prepared for every value presented in Fig. 2 and Fig. 3. They were analyzed with MLR and the result was in Table 1.

The second column from the right in the table is for the manual operation analysis results in which the control system was not used. The most right column is for the controlled operation analysis results. N. A. (not available) in the table means that a number of data for the corresponding factor was too small compared to that of whole data

Table 1 An example of MLR analysis result

(Intercept)			expect ed sign	(Manual)	(Control)
①	In Pulp Brightness	V1	-	-0.098964	-0.119203
②	ΔBrightness (Out-In)	V2	+	-0.004683	0.005880
③	In Pulp Kappa No.	V3	+	0.021771	-0.048993
④	Out Pulp K No.	V4	-	-0.692537	-0.363426
⑤	Production Rate	V5	+	-0.000215	0.000225
⑥	Wood Chip species	V6 C1		0.002993	-0.000049
		V7 C2		0.004627	0.000198
		V8 C3		-0.010194	NA
		V9 C4		0.005146	0.008564
		V10 C5		0.000365	-0.002346
		V11 C6		-0.001989	NA
		V12 C7		0.007084	0.002215
		V13 C8		0.002909	0.002019
		V14 C9		0.007204	0.000190
		V15 C10		NA	0.003390

and they were not used in calculation by statistical analysis software “R”⁵⁾.

At the same time, every explanatory variable was marked with a theoretical sign, which was expected in normal bleaching reaction. In the manual operation, the sign of a explanatory variable, ⑤ production rate, was minus, contrary to the slope observed in Fig. 3, and it was not reasonable.

A variable, ② Δ brightness, in the manual operation and a variable, ③ incoming Kappa number, in the controlled operation also had signs reverse to those reasonably expected. They suggested that the linear regression equation was not reliable enough and moving forward to the next step of the equation (5) with the result of MLR was not practicable. As large data which consisted of one objective variable and 15 explanatory variables, which seemed to have fairly large multi-collinearity among them, were forced to form one regression equation by MLR, there would be excess over-fitting.

8. Analysis by PLS regression and correction of bleach cost

This problem in MLR regression can be removed in PLS regression which uses latent variables. In PLS regression, a most probable latent variable was taken as the first latent variable, and then the second latent variable and so on were taken into the regression stepwise. By doing so, the regression equation is improved to a more probable form.

Whether the equation was improved or not was judged

by several viewpoints. They are : 1) signs of coefficients $c_1, c_2, \dots, c_n$ in the equation (4), which are calculated as equivalents to regression coefficients of original explanatory variables (input variables), 2) values of those coefficients, 3) declining behavior of root mean squared error of prediction (RMSEP) by increasing the number of latent variables.

The result of PLS regression analysis of the same data used for MLR analysis, in which bleach cost was a objective variable and 15 explanatory variables were taken, are presented in Table 2. The upper half was for the manual operation and the lower half was for the controlled operation analysis results. As the number of explanatory variables was fifteen, the number of latent variables in PLS analysis could be increased up to fifteen in the calculation.

Fig. 4 depicts the transition curve of root mean squared error of prediction (RMSEP) against the number of latent variables taken into the regression analysis. The vertical axis was for RMSEP and the horizontal axis was for the number of latent variables. The behavior of the transition curve suggests how properly the regression equation by PLS analysis represents the data. Finally, taking into account the results in Table 2 and Fig. 4, the number of latent variables was decided to get a most probable PLS equation.

As mentioned in Chapter 4, bleach cost tends to increase with increasing production rate. So, the sign of the regression coefficient of the production rate was expected as plus. Moreover, from the line's gradient tendency, its value would be between +0.0001 to +0.0010, judging from Fig. 3 in which the approximate regression lines

Table 2 Analyzed results by PLS (Upper/Manual operation mode, Lower/Control mode)

Manual

		expected sign	Y.1.comp	Y.2.comp	Y.3.comp	Y.4.comp	Y.5.comp	Y.6.comp	Y.7.comp	Y.8.comp	Y.9.comp	Y.10.comp	Y.11.comp	Y.12.comp	Y.13.comp	Y.14.comp	Y.15.comp	
(Intercept)			2.384003	2.492003	2.441583	2.479962	2.414729	2.555069	2.832014	2.881189	2.810221	3.013221	7.400856	12.045939	10.346088	7.297383	20844984	
①	In Pulp Brightness V1	-	-0.000020	-0.000307	-0.000860	-0.001667	-0.003244	-0.016890	-0.039183	-0.049605	-0.083981	-0.093260	-0.121428	-0.162059	-0.131968	-0.038984	-0.038987	
②	ΔBrightness (Out-In) V2	+	0.000017	0.000253	0.000703	0.001544	0.002788	0.013826	0.032100	0.040889	0.054057	0.054443	0.016355	-0.071368	-0.041228	-0.004683	-0.004683	
③	In Pulp Kappa No. V3	+	0.000006	0.000078	0.000222	0.000520	0.000937	0.004829	0.011188	0.014101	0.018593	0.017292	-0.054686	0.018787	0.022054	0.021771	0.021771	
④	Out Pulp K No. V4	-	-0.000002	-0.000037	-0.000103	-0.000219	-0.000389	-0.002078	-0.004971	-0.006419	-0.009277	-0.014680	-0.197085	-0.541712	-0.655405	-0.852527	-0.892362	
⑤	Production Rate V5	+	0.000002	0.000359	0.000445	0.000382	0.000451	0.000224	0.000130	0.000204	-0.000180	-0.000160	-0.000135	-0.000206	-0.000219	-0.000215	-0.000215	
⑥	Wood Chip species	V6 C1	-0.000039	-0.000642	-0.001775	-0.003502	-0.005582	-0.007529	-0.002460	0.000956	0.001046	0.000345	0.000304	0.006197	-0.000582	0.001179	-208450	
		V7 C2	-0.000122	-0.001110	-0.000603	-0.000647	-0.000013	-0.000741	-0.002803	-0.002285	0.001375	-0.000009	0.003148	0.008559	0.000990	0.002812	-208450	
		V8 C3	0.000000	-0.000007	-0.000020	-0.000039	-0.000005	-0.000348	-0.000800	-0.001058	-0.001548	-0.002412	-0.002978	-0.009705	0.004687	-0.012009	-208450	
		V9 C4	-0.000113	-0.000105	-0.000247	-0.000041	0.000296	0.000503	0.007722	0.004426	0.004517	0.003639	0.006349	0.008861	0.001913	0.003232	-208450	
		V10 C5	0.000040	0.000698	0.001742	0.003422	0.005427	0.002401	0.008277	0.005401	0.001272	0.001487	0.002620	0.004379	-0.002353	-0.001450	-208450	
		V11 C6	-0.000012	-0.000151	-0.000388	-0.000772	-0.001289	-0.005539	-0.012087	-0.014438	-0.013979	-0.008533	-0.009935	-0.009597	0.010689	0.003604	0.005270	-208450
		V12 C7	0.000007	-0.000408	-0.001008	0.000096	-0.000177	-0.000758	-0.002859	-0.002202	0.002202	0.002202	0.002202	0.002202	0.002202	0.002202	0.002202	
		V13 C8	0.000025	0.000105	-0.001287	-0.001732	-0.001058	-0.001856	-0.004698	-0.004388	-0.000988	-0.002090	0.001243	0.006490	-0.000988	0.001095	-208450	
		V14 C9	0.000122	0.001844	0.004068	0.004103	0.003785	0.003146	0.001412	0.001454	0.000587	0.003871	0.008573	0.010726	0.003587	0.005380	-208450	
		V15 C10	-0.000007	-0.000225	-0.000483	-0.000887	-0.001305	0.006021	0.008275	0.007988	0.000768	0.000715	0.000787	0.001983	0.002848	-0.004128	-0.001814	-208450

Control

		expected sign	Y.1.comp	Y.2.comp	Y.3.comp	Y.4.comp	Y.5.comp	Y.6.comp	Y.7.comp	Y.8.comp	Y.9.comp	Y.10.comp	Y.11.comp	Y.12.comp	Y.13.comp	Y.14.comp	Y.15.comp
(Intercept)			1.806111	1.991127	1.950945	1.935269	1.915198	1.740968	1.799173	1.997750	2.022898	2.008870	7.517793	11.095508	7.834082	7.834083	7.834083
①	In Pulp Brightness V1	-	-0.000042	-0.000563	-0.001201	-0.008113	-0.018462	-0.029652	-0.047839	-0.052121	-0.061868	-0.064618	-0.124424	-0.159219	-0.119203	-0.119203	-0.119203
②	ΔBrightness (Out-In) V2	+	0.000032	0.000441	0.000939	0.002220	0.014183	0.022700	0.037187	0.040290	0.045392	0.045409	0.018755	-0.035629	0.005880	0.005880	0.005880
③	In Pulp Kappa No. V3	+	0.000016	0.000178	0.000372	0.002282	0.004962	0.007467	0.011217	0.011781	0.010071	0.007459	-0.003704	-0.002480	-0.048993	-0.048993	-0.048993
④	Out Pulp K No. V4	-	-0.000002	-0.000040	-0.000086	-0.000550	-0.001284	-0.002131	-0.003731	-0.004268	-0.006896	-0.008902	-0.008196	-0.188499	-0.363426	-0.363426	-0.363426
⑤	Production Rate V5	+	0.001238	0.001058	0.001018	0.001029	0.000977	0.000836	0.000539	0.000333	0.000338	0.000338	0.000371	0.000217	0.000225	0.000225	0.000225
⑥	Wood Chip species	V6 C1	-0.000011	-0.000261	-0.000395	-0.001578	-0.004115	-0.008123	0.001971	0.002699	-0.004532	-0.004298	0.000480	0.002495	-0.000049	-0.000049	-0.000049
		V7 C2	-0.000162	-0.002213	-0.001608	-0.001837	-0.001375	0.003259	0.005777	0.004657	0.002128	0.001698	0.001145	0.000775	0.000198	0.000198	0.000198
		V8 C3	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
		V9 C4	-0.000025	-0.000059	0.000440	0.011346	0.022745	0.024799	0.020624	0.018878	0.013134	0.012688	0.012365	0.009906	0.008584	0.008584	0.008584
		V10 C5	-0.000038	-0.000648	-0.001599	-0.003007	-0.002049	0.002633	0.002185	0.004597	-0.001386	-0.001757	-0.001757	-0.002172	-0.002346	-0.002346	-0.002346
		V11 C6	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
		V12 C7	0.000010	0.000159	0.000188	0.000605	0.002864	0.003848	0.008320	0.009136	0.008378	0.004318	0.004426	0.003100	0.002215	0.002215	0.002215
		V13 C8	0.000040	0.000680	0.000377	0.003896	0.006553	0.005177	0.008321	0.007117	0.004247	0.003571	0.002918	0.002363	0.002019	0.002019	0.002019
		V14 C9	0.000000	-0.000009	-0.000094	-0.005108	-0.000537	0.004129	0.008954	0.004880	0.002007	0.002057	0.000882	0.001052	0.000190	0.000190	0.000190
		V15 C10	0.000196	0.002858	0.004678	0.006529	0.009634	0.013082	0.012368	0.011820	0.006311	0.005952	0.006468	0.004973	0.003390	0.003390	0.003390

Fig. 4 RMSEP (Root Mean Squared Error of Prediction) curve transition

Fig. 5 Operation mode vs. Bleach cost (corrected by PLS regression result)

were plotted against the production rate. In Table 2, expected signs of variables were included as in Table 1, and fields of which sign was the same to that of the expected were painted by hatchings.

For example, in the upper table in Table 2, the value of RMSEP reduced fairly well at the 9th latent variable. On the other hand, the sign of the coefficient of regression of the production rate changed from plus to minus there which was unreasonable. It was suggested that over-fitting already occurred in the equation. In the lower table of Table 2, all variables for the production rate had the sign of plus. So, we decided to take latent variables down to the 6th in PLS regression equation for both manual and controlled operations.

Regarding the blending ratio of ten different wood species, there were many unknown factors affecting bleach cost. For example, besides original bleachability dependent on wood species, growth condition in forest, part of location in log material (whether it was in growing zone or core zone), and chip storage conditions were influential. As it was difficult to use signs of regression coefficients of wood species for decision making, they were left to the course of calculation.

So, the number of latent variables to take into regression analysis was decided. Then, using regression coefficients obtained in PLS regression equation in Table 2, bleach cost data were corrected by the equation (5), which was a linear approximation, and in which every explanatory variable was adjusted to its own base condition. Regarding $X_{ji}(\text{Base})$, which was an evaluation base condition for a variable X_j , its average was used.

Bleach cost data were corrected by this procedure and presented in Fig. 5 in the same manner and scale as in Fig. 2. When the distribution of data in the vertical axis of the corrected bleach cost was compared with that in Fig. 2, it was recognized to be compressed in a smaller area. Their average values were marked in horizontal lines in Fig. 5 as in Fig. 2. The difference between them was larger than that in Fig. 2. Finally, cost saving by the controlled operation was clearly recognized in the corrected data group.

9. Conclusion

Until now, few reasonable methods have been proposed for evaluating changes in operations like pulp bleaching, and if the effect of change was not big enough, its evaluation was finished with uncertainty in most cases. Our

method, however, which used PLS regression described in this paper⁶⁾⁷⁾, made it possible to evaluate reasonably the effect of changes in operation like introducing new control system or modifying production process, even though they were very small.

We used free software “R” for statistical analysis⁵⁾ in this report. We received much support from staff in charge in our mill operation and members of Yokogawa Electric Corporation. They were Mr. Takahiro Yamamoto of the Third Solution Technology Section, Mr. Masahiro Watanabe of APC Engineering of Technology Section and Mr. Jun Nishimura of Technical Division of Yokogawa & Co., Ltd.

References

- 1) *for example*, I. Hashimoto, S. Hasebe, M. Kano : Process Control Engineering, Asakura Publishing Co., Ltd., 2002, M. Ohshima : Process Control System, Corona Publishing Co., Ltd., 2003 *etc.*
- 2) M. Ogawa, M. Kano : Practice and challenges in chemical process control Application in Japan, System/Control/Information, ISCIE Journal, Vol 52-8, 2008.
- 3) Y. Miyashita, S. Sasaki : Chemo-metrics/Chemical pattern recognition and multivariate analysis, Kyoritsu Publishing Co., Ltd., 1995.
- 4) Y. Ozaki, A. Uda, T. Akai : Multivariate analysis for chemical engineers, Koudansha Scientific Ltd., Publishers, 2007.
- 5) R software for statistical processing : <http://cran.r-project.org/>.
- 6) Y. Mori *et al.* : Japan Patent applied for 2011-248615, 2011.
- 7) Y. Mori : A calculation method to evaluate the control effect by PLS regression analysis, 79 th Pulp and Paper Research Conference, 2012.

(Manuscript accepted 24 March, 2013)