A Study on the Optimum Scheme for Determination of Operation Time of Line Feeders in Automatic Combination Weighers

James N. Keraita
Graduate School of Mechanical Engineering, Yeungnam University, Gyeongsan 712-749, Korea
Kyo-Hyoung Kim*
School of Mechanical Engineering, Yeungnam University, Gyeongsan 712-749, Korea

In an automatic combination weigher, the line feeders distribute the product to several weighing hoppers. The ability to supply appropriate amount of product to the weighing hoppers for each combination operation is crucial for the overall performance. Determining the right duration of operating a line feeder to supply a given amount of product becomes very challenging in case of products which are irregular in volume or specific gravity such as granular secondary processed foods. In this research, several schemes were investigated to determine the best way for a line feeder to approximate the next operating time in order to supply a set amount of irregular goods to the corresponding weighing hopper. Results obtained show that a weighted least squares method (WLS) employing 10 data points is the most effective in determining the operating times of line feeders.

Key Words : Automatic Combination Weigher, Accurate Weighments, Vibratory Linear Feeder, Least Squares Method, OLS (Ordinary Least Squares), WLS (Weighted Least Squares)

1. Introduction

Weighing and packaging is a major concern in any production line. An automatic combination weigher is a fully automatic computer-controlled scale whose function is to weigh accurately, efficiently and reliably at high speed. In the radial type, products are fed through a circular feeder located at the centre to a number of weighing channels. Each channel typically consists of a line feeder, a preliminary (pool) hopper, a weighing hopper and a memory hopper (optional). If a certain amount of products are collected in the preliminary hopper, both feeders stop and the products in the preliminary hopper are transferred to the weighing hoppers where the weight is accurately measured by a load cell. The calculation unit then calculates the total weights of many different combinations of weighing (or memory) hopper contents and selects the best combination.

For example in a 10 channel weigher, for four hoppers to accurately provide a total of 200 g (distributed equally), each line feeder should supply close to 50 g to each corresponding weighing hoppers. The combination operation will then pick four hoppers that together best total 200 g. Products from the selected hoppers are released to the packaging machine via discharge chutes. New product is then promptly fed into each of the emptied hoppers, thus continuing the weighing operation. Automatic combination weighers can be found in many factories and packing houses that supply packaged goods to supermarkets and other

* Corresponding Author,
E-mail : khykim@yumail.ac.kr
TEL : +82-53-810-2444; **FAX** : +82-53-813-3703
School of Mechanical Engineering, Yeungnam University, Gyeongsan 712-749, Korea. (Manuscript **Received** December 6, 2005; **Revised** June 13, 2006)

outlets. Almost any type of food product is successfully handled by combination weighers; crisps, nuts, grains, sweets, frozen chips and many more. A wide range of non-food products is also comfortably handled, from detergents and hardware to pharmaceutical, amongst others.

During the last two decades vibratory conveyors have dominated the transport of products prior to weighing to such an extent that it is rarely possible to replace them by any other technology. They are particularly advantageous when individual parts need to be separated from a group and fed at a controlled rate into another piece of equipment. Designing linear vibratory conveyors in industrial applications is based on already established theory on the dynamics of a body on a vibrating plate (Winkler, 1978). The most promising design is the conveyor with inclined motion. Factors that could affect the rigid body's conveying velocity are the angle of vibration, amplitude of vibration, coefficient of friction, inclination of the plate angle to the horizontal and the operating frequency. Because of the complicated interactions between the vibrating trough and the particles, both glide and throw movements frequently appear within one oscillation cycle (Lim, 1997). Hongler and Figour (1989) concluded that the motion of a part on the track can be either of the sliding type (S-regime), of the hopping type (H-regime), or a combination of both (HS-regime).

While many studies are restricted to the sliding regime, it is actually the hopping regime where the conveying rate is highest (Winkler, 1979). The conveying rate in the chaotic region is roughly independent of external variations in parameters. Han and Lee (2002) identified this regime through numerical simulation and experimental analysis. In general the transport rate is difficult to calculate in a purely analytical manner. Conveying the products at the highest velocity will however not provide the optimum performance of automatic combination weighers unless appropriate amount of the product is supplied to the hoppers for the combination operation. Previous models like those manufactured by Ishida Europe use some form of level (volume) sensors on input chute to control the product flow to the line feeders. This means that the line feeders cannot be operated consecutively with the in feed conveyor and product control is not restricted to each individual channel. The level sensor could be installed in or before the preliminary hopper as done in some models by Anritsu Corporation but this means using as many sensors as the number of channels. In other weighers in the market, the line feeders are operated in a preset fixed time though the exact transport rate is difficult to determine.

Clearly these methods cannot deliver accurate weighments in case of products that are irregular in volume and/or specific gravity as commonly encountered. This paper suggests a method of basing the time on analyzing the relationship between previous operating times and the amount of products delivered to the hoppers using the least squares method (LSM). Various schemes are investigated and compared with the fixed time method.

2. Least Squares Method

2.1 Formulation of the least squares problem

In its simplest form the least squares method will be illustrated in this section. Suppose that the relationship between two groups of variables x and y can be best described by the equation of a straight line:

$$y = a_1 x + a_0 \tag{1}$$

One could arbitrarily choose two sets of x and y quantities and solve for the two unknown parameters a_1 and a_0. Yet, the line constructed by the computed a_1 and a_0 parameters might not pass through some sets of x and y (Fig. 1), since the information associated with those points were not used in computing a_1 and a_0. The reasons for all the points not being located on one straight line could be:

(1) Errors in data set.
(2) Inaccurate model for the data set.

The least squares criterion requires that the sum of the squares of the deviations separating

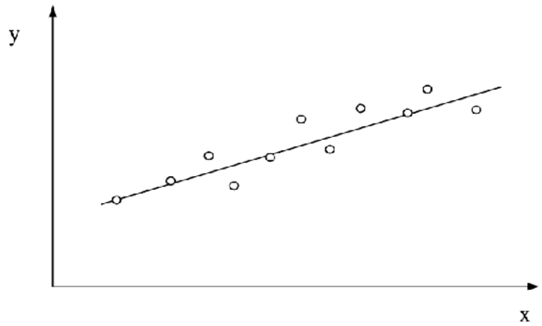

Fig. 1 Straight line fitting of data

the data points from the curve will be minimum. These deviations are simply the difference between the estimated values of y (hereafter denoted by $\hat{y}$) from Eq. (1) and the actual measured values of $y(y_r)$. In other words, the deviations are the errors associated with the value of y predicted by the model and the actual measured data. The sum of the squares of the offsets is used instead of the offset absolute values because this allows the residuals to be treated as a continuous differentiable quantity. The least squares method approach will use information associated with all of the x and y sets, to determine the "best" estimates of a_1 and a_0. In the least squares sense, the error can be expressed as;

$$E=\sum_{r=1}^{N} e_r^2 = \sum_{r=1}^{N}[y_r-\hat{y}_r]^2=\sum_{r=1}^{N}[y_r-(a_1x_r+a_0)]^2 \quad (2)$$

where E is the total error and e_r is the error at a particular data point.

Minimization of the error with respect to a_1 and a_0 results in the following equations which are called the normal equations for the least squares problem.

$$\frac{\partial E}{\partial a_1}=\sum_{r=1}^{N} 2[y_r-(a_1x_r+a_0)][-x_r]=0 \quad (3)$$

$$\frac{\partial E}{\partial a_0}=\sum_{r=1}^{N} 2[y_r-(a_1x_r+a_0)][-1]=0 \quad (4)$$

Eqs. (3) and (4) can be solved for the unknown parameters a_1 and a_2.

$$a_1=\frac{N\sum_{r=1}^{N} x_ry_r-\left(\sum_{r=1}^{N} x_r\right)\left(\sum_{r=1}^{N} y_r\right)}{N\sum_{r=1}^{N} x_r^2-\left(\sum_{r=1}^{N} x_r\right)^2} \quad (5)$$

$$a_0=\frac{\left(\sum_{r=1}^{N} y_r\right)\left(\sum_{r=1}^{N} x_r^2\right)-\left(\sum_{r=1}^{N} x_r\right)\left(\sum_{r=1}^{N} x_ry_r\right)}{N\sum_{r=1}^{N} x_r^2-\left(\sum_{r=1}^{N} x_r\right)^2} \quad (6)$$

The a_1 and a_0 values computed from Eqs. (5) and (6) represent the characteristics of a straight line which would "best" describe the x and y sets of values.

The least squares problem can be formulated in matrix notation as:

$$\{Y\}=[X]\{A\} \quad (7)$$

where;

$$\{Y\}=\begin{Bmatrix} y_1 \\ y_2 \\ \vdots \\ y_N \end{Bmatrix} \quad [X]=\begin{bmatrix} x_1 & 1 \\ x_2 & 1 \\ \vdots & \vdots \\ x_N & 1 \end{bmatrix} \quad \{A\}=\begin{Bmatrix} a_1 \\ a_0 \end{Bmatrix}$$

The equations presented by matrix Eq. (7) are in general a set of inconsistent and overdetermined equations. Inconsistent, since it is not usually possible to find $\{A\}$ that would satisfy all the individual equations of Eq. (7), and overdetermined, since the number of equations is larger than the number of unknowns. The least squares solution of Eq. (7) is:

$$[X]^T\{Y\}=[X]^T[X]\{A\} \quad (8)$$

and solving for unknown vector $\{A\}$ yields;

$$\{A\}=([X]^T[X])^{-1}[X]^T\{Y\} \quad (9)$$

provided that $([X]^T[X])^{-1}$ exists. In case the inverse does not exist, one could then use numerical techniques to solve for the vector $\{A\}$.

In cases where matrices A, X, and Y are complex valued, the Eqs. (8) and (9). The hermitian operator, H, is the complex conjugate transpose. Hence, the unknown vector $\{A\}$ is given by:

$$\{A\}=([X]^H[X])^{-1}[X]^H\{Y\} \quad (10)$$

In general the relationship of Eq. (1) could be in the form of:

$$y=a_N x^N+a_{N-1}x^{N-1}+\cdots+a_1 x+a_0 \qquad (11)$$

A set of equations similar to the formulation above could be written and solved to obtain the least squares estimation of unknown parameters a_N, a_{N-1}, $\cdots$, a_1, a_0. Further, the least squares method stated above could easily be extended to problems involving more than one independent variable. For example, z could be expressed in terms of x and y:

$$z=a_2 y+a_1 x+a_0 \qquad (12)$$

The corresponding normal equations for the least squares problem can be solved for the unknown parameters a_2, a_1, and a_0. The computed values a_2, a_1, and a_0 represent the characteristics of a plane which would "best" describe the x, y, and z sets of values. The above theory and formulation could be expanded to least squares estimation of a surface and eventually to higher order dimensions.

2.2 Weighted least squares method

One of the common assumptions underlying most process modeling methods, including linear and nonlinear least squares regression, is that each data point provides equally precise information about the deterministic part of the total process variation. In other words, the standard deviation of the error term is constant over all values of the predictor or explanatory variables. This assumption, however, clearly does not hold, even approximately, in many modeling application. In situations when it may not be reasonable to assume that every observation should be treated equally, weighted least squares (WLS) can often be used to maximize the efficiency of parameter estimation. This is done by attempting to give each data point its proper amount of influence over the parameter estimates. A procedure that treats all of the data equally would give less precisely measured points more influence than they should have and would give highly precise points too little influence. If a nonconstant variance is seen in a plot of residuals versus fits, then the WLS should be considered.

Individual equations in matrix Eq. (7) could be multiplied by a weighting factor to give that equation more or less weight in the computation. The weighting factors could be presented in form of a diagonal $(N\ X\ N)$ matrix, W. The diagonal element in row i represents the weighting factor corresponding to equation i, and the off diagonal terms are all zero. Matrix W is premultiplied to both sides of Eq. (7).

$$[W]\{Y\}=[W][X]\{A\} \qquad (13)$$

where:

$$[W]=\begin{bmatrix} w_1 & 0 & \cdots & 0 \\ 0 & w_2 & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & \cdots & w_n \end{bmatrix}$$

Solving the vector $\{A\}$ in Eq. (13) yields;

$$\{A\}=([X]^H[W]^H[W][X])^{-1}[X]^H[W]^H[W]\{Y\} \qquad (14)$$

The biggest disadvantage of WLS, which may easily be overlooked, is probably the fact that the theory behind this method is based on the assumption that the weights are known exactly. This is almost never the case in real applications and estimated weights must be used instead. It is clear that only when the weights are estimated with high enough precision, will their use significantly improve the parameter estimation compared to results obtained if all the data points were weighted equally i.e. ordinary least squares (OLS) method.

The "goodness" of any least squares estimation process is measured by the coefficient of correlation parameter which is defined in terms of total variation and explained variation (Strang, 1986). The total variation of y is defined as;

$$\Delta y_T=\sum_{r=1}^{N}(y_r-\bar{y})^2 \qquad (15)$$

where y_r is the measured value of y and $\bar{y}$ is the mean value. The total variation consists of two parts as shown in Fig. 2:

(1) the explained variation, $\sum_{r=1}^{N}(\hat{y}_r-\bar{y})^2$.

(2) the unexplained variation, $\sum_{r=1}^{N}(y_r-\hat{y}_r)^2$

The terms explained variation and unexplained

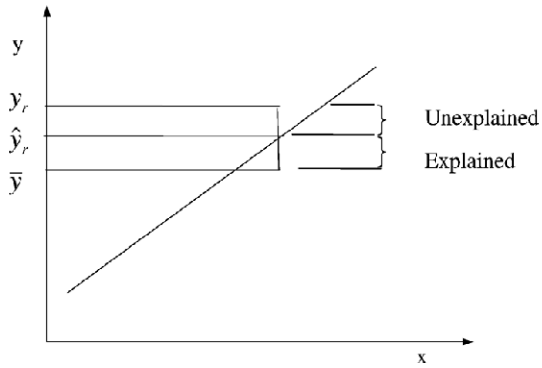

Fig. 2 Explained and unexplained variation

variation are used to denote the fact that the deviations $(\hat{y}_r - \bar{y})$ have a definite pattern, while the deviations $(y_r - \hat{y}_r)$ are random and unpredictable.

The correlation coefficient γ^2, is defined as :

$$\gamma^2 = \left(\frac{\sum_{r=1}^{N} (\hat{y}_r - \bar{y})^2}{\sum_{r=1}^{N} (y_r - \bar{y})^2} \right) \tag{16}$$

The magnitude of γ^2 varies between 0 and 1. A value of 0 indicates no correlation between dependent and independent variable (s), while a value of 1 indicates perfect correlation i.e. the paired values (x_i, y_i) all lie on a straight line.

2.3 Approaches to weight estimation

2.3.1 Replicated data

Optimal results, which minimize the uncertainty in the parameter estimators, are obtained when the weights, w_i, used to estimate the values of the unknown parameters are inversely proportional to the variances at each combination of predictor variable values (Ryan, 1997).

$$w_i \propto \frac{1}{\sigma_i^2} \tag{17}$$

With $w_i = 1/\sigma_1^2$, points with a lower variance are given a greater statistical weight.

If there are replicates in the data, the most obvious way to estimate the weights is to set the weight for each data point equal to the reciprocal of the sample variance obtained from the set of replicate measurements to which the data point belongs. Though this method appears attractive, it requires a very large number of replicates at each combination of predictor variables. A slightly improved strategy for estimating the weights could be to find a function that relates the standard deviation of the response at each combination of predictor variable values to the predictor variables themselves.

2.3.2 Estimating weights using residuals

Most commonly, the pattern of nonconstant variance is that either the standard deviation or the variance of the residuals is linearly related to the mean (Helsel and Hirsch, 1992). This occurs theoretically in most skewed distributions, for instance. Then the absolute residuals essentially are estimates of standard deviation. So if a plot of absolute residuals versus fits looks linear, a regression line (response=absolute residuals, predictor=fits) could be fitted to the pattern. The predicted values from this regression could be viewed as smoothed estimates of the standard deviations of the points. So, the weights in a WLS regression would be ;

$$w_i = \frac{1}{(\hat{s}_i)^2} \tag{18}$$

It is possible to further improve the estimated parameters by iteration. That is by taking the residuals from the weighted least squares estimation and fitting a variance function to them. The new set of weights will be used again to perform WLS estimation. This iteration procedure could be continued until the estimated model parameters seem to stabilize.

2.3.3 Special techniques for estimating weights

There are many other specialized techniques of estimating the weighting matrix (or function) depending on the specific application. For example in filter design, the weighting function should be such as to assign more importance to stop-bands than pass-bands. For applications where the last measurements should be weighted more heavily than preceding measurements (e.g. neural networks), a forgetting factor β whose magnitude is less than one could be used (Lee and Cheng, 1997). The last equation is multiplied by β^0 and preceding equations are progressively multiplied by β^1, $\beta^2 \cdots \beta^N$; where N is the number of equations. In predictive lossless image coding (Meyer and Tischer), the weight of an observed pixel mainly depends on how close the pixel is to the current pixel.

3. Experimental Device and Method

In this research only one line feeder was used in the experiment to deliver a target weight of 50 g. Fig. 3 shows a photograph of the line feeder. The vibrating line feeder consists of a base member below a spring-supported horizontal pan. The drive is an electric motor with a fixed eccentric shaft. In operation, the drive transmits vibration through the support springs to the pan base member. The pan's vibration continuously throws the material upward and forward, thus moving the material in short hops along the conveyor. Any vibrating linear conveyor's operation is typically based on the natural frequency principle. At the natural frequency, the conveyor will vibrate indefinitely with only a small energy input. Once the drive initiates the conveyor's vibration, the supporting springs, by alternately storing and releasing most of the required energy, help maintain constant motion under the conveyed load.

To test whether the method based on least squares method is more accurate in determining the time of operation of the line feeders to supply the right amount of products compared to the fixed time method, the most appropriate fixed them was first established by averaging 20 durations that most accurately delivered 50 g. Then using this average time, 100 trials were carried out.

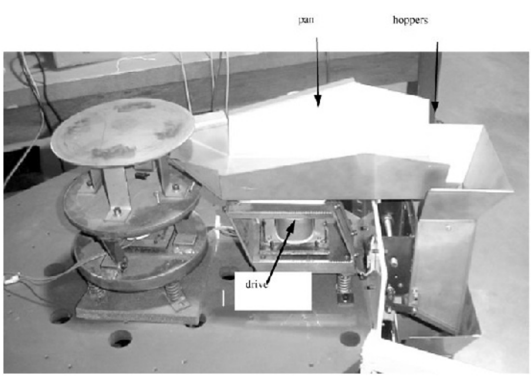

Fig. 3 A photograph of the line feeder

To implement the OLS method, the 20 sets of weight delivered verses time (or part of them depending on the data points preferred), were used as the starting point. In the WLS case, the weights were approximated using residuals. The procedure for implementation of WLS method was as follows:

(1) Fit the regression model by unweighted least squares (i.e. OLS) and analyze the residuals.

(2) Estimate the variance function by regressing the squares residuals on the appropriate predictor (s).

(3) Use the fitted values (i.e. the predicted values) from the estimated variance function to obtain the weights w_i.

(4) Estimate the regression coefficients using these weights.

4. Results and Analysis

Table 1 shows the experimental schemes investigated in this research. The average time to deliver 50 g was found to be 0.761 s. Fig. 4 shows the results obtained from 100 trials by fixing the line feeder operation time at 0.761 s. Though the average value (49.59 g) is very close to the expected value of 50 g, the dispersion as indicated by the standard deviation is quite high. This would give rise to high and unacceptable errors in precision weighing.

Figure 5 (and corresponding Table 2) shows

Table 1 Line feeder schemes

Scheme	Algorithm	Number of Data
1	Fixed Time	Not Applicable
2	Ordinary LSM	5
3	Ordinary LSM	10
4	WLS	5
5	WLS	10
6	WLS	20

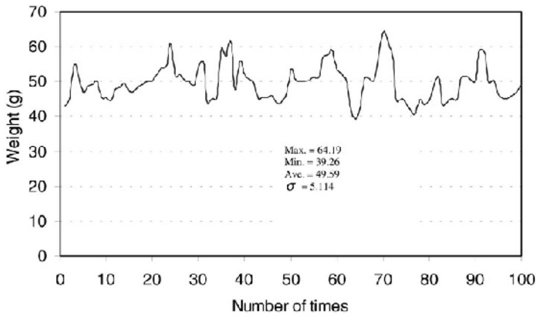

Fig. 4 Results for the fixed time scheme

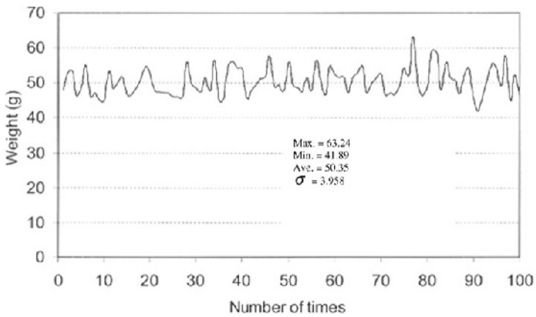

Fig. 5 Results scheme 2's 5th experiment

Table 2 Results of statistical analysis for scheme 2

Parameter	1st	2nd	3rd	4th	5th	Ave.
Max	64.55	66.53	62.18	60.59	63.24	63.42
Min	42.87	44.28	40.41	38.75	41.89	41.64
Ave.	50.62	50.72	49.38	50.41	50.35	50.30
σ	4.216	4.022	4.126	3.862	3.958	4.037

the results of OLS method using 5 previous sets of operating time and weight of products delivered i.e. 5 data points. To increase the validity of the results, each experiment was repeated 5 times to give a total of 500 trials. The results show an improvement over the fixed time scheme. The correlation coefficient for the initial set of data was 0.892. This means that the fit of the linear model to the data is reasonably good and OLS method is therefore justified. The results were only slightly improved by increasing the number of data points to 10 as shown in Fig. 6 and Table 3.

Figure 7 and Table 4 shows the results of WLS scheme implemented with 5 data points. The results show a significant reduction of precision errors over the corresponding OLS. The correlation coefficient for the same initial set of data was 0.924. The increase in the correlation coefficient (from 0.892 for OLS) indicates an improvement in the degree to which the assumed model describes the relationship for the set of data obtained.

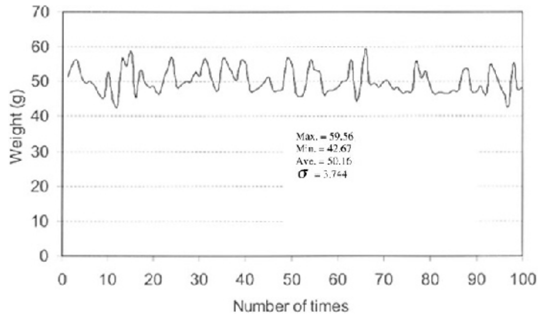

Fig. 6 Results scheme 3's 1st experiment

Table 3 Results of statistical analysis for scheme 3

Parameter	1st	2nd	3rd	4th	5th	Ave.
Max	59.56	64.58	57.02	59.44	62.45	60.61
Min	42.67	41.36	39.48	43.85	44.87	42.45
Ave.	50.16	49.76	50.25	50.42	50.36	50.19
σ	3.744	3.782	3.673	3.823	3.611	3.727

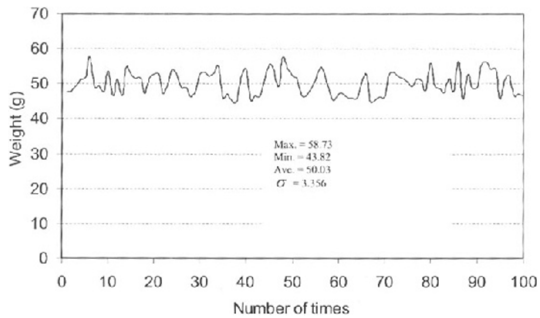

Fig. 7 Results scheme 4's 5th experiment

Table 4 Results of statistical analysis for scheme 4

Parameter	1st	2nd	3rd	4th	5th	Ave.
Max	62.45	56.79	56.34	60.26	57.80	58.73
Min	43.28	41.25	44.56	45.24	44.78	43.82
Ave.	49.59	50.22	49.78	50.43	50.13	50.03
σ	3.406	3.278	3.361	3.386	3.350	3.356

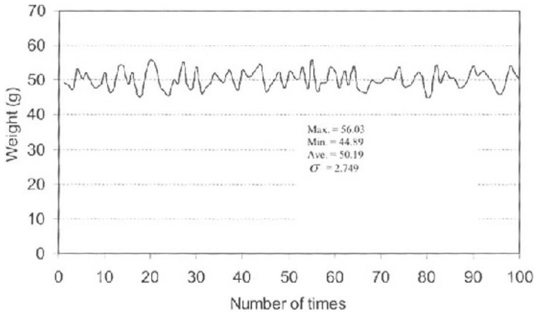

Fig. 8 Results scheme 5's 2nd experiment

Table 5 Results of statistical analysis for scheme 5

Parameter	1st	2nd	3rd	4th	5th	Ave.
Max	55.47	56.03	59.67	58.91	58.18	57.65
Min	46.21	44.89	41.24	46.23	42.34	44.18
Ave.	49.73	50.19	49.83	50.17	49.94	49.97
σ	2.912	2.749	2.872	2.698	2.706	2.787

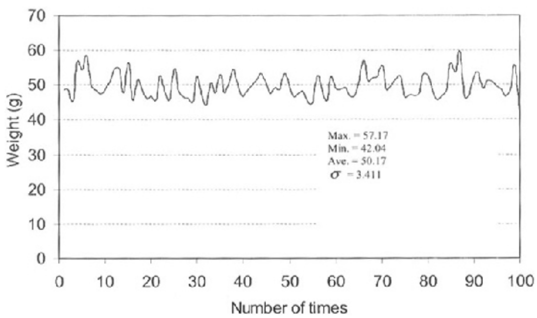

Fig. 9 Results scheme 6's 4th experiment

Table 6 Results of statistical analysis for scheme 6

Parameter	1st	2nd	3rd	4th	5th	Ave.
Max	61.34	62.67	57.28	59.67	57.17	59.63
Min	45.24	44.34	37.34	42.03	42.04	42.20
Average	49.67	49.54	50.27	49.77	50.17	49.88
σ	3.345	3.674	3.563	3.495	3.411	3.498

The results were further improved by using WLS with 10 data points as shown in Fig. 8 and Table 5. However when the number of data points was increased from 10 to 20, the precision errors increased. This is shown in Fig. 9 and Table 6. It can be seen that the standard deviation for instance, is 3.356 compared to 2.787 obtained for the case of WLS with 10 data points. It appears that going too deeper into history does not improve the results. The best results are those of scheme 5 i.e. WLS implemented with 10 data points.

5. Conclusions

In automatic combination weighers the ability for line feeders to accurately deliver target weights to the hoppers for the combination operation is crucial for the overall performance. Determining the right duration of operating a line feeder to supply a given amount of product is very challenging in case of products which are irregular in volume and/or specific gravity. In this work, determining the duration by analyzing the relationship between previous operating times and the amount of products delivered using least squares method was investigated.

Results obtained show that the least squares implementation reduces precision errors compared to using a fixed time. The best results were obtained using a weighted least squares (WLS) method with 10 data points. The weights were approximated by using the residuals from a corresponding ordinary least squares (OLS) model. For a target weight of 50 g, all 500 trials lied between 41.24 g and 59.67 g with an average standard deviation of 2.787. Though the weights are not the final output from the combination weigher, they do indicate the accuracy that can be achieved after the combination operation.

References

Han, I. and Lee, Y., 2002, "Chaotic Dynamics of Repeated Impacts in Vibratory Bowl Feeders," *Journal of Sound and Vibration*, Vol. 249, No. 3, pp. 529～541.

Helsel, D. R. and Hirsch R. M., 1992, *Statistical Methods in Water Resources*, Elsevier, Amsterdam.

Hongler, M. O. and Figour, J., 1989, "Periodic Versus Chaotic Dynamics in Vibratory Feeders," *Helvetica Physica Acta*, Vol. 62, pp. 68～81.

Lee, T. T. and Jeng, J. T., 1997, "Chebyshev

Polynomials-Based (CPB) Unified Model Neural Networks for Function Approximation," *International Society for Optical Engineering*, pp. 372~381.

Lim, G. H., 1997, "On the Conveying Velocity of a Vibratory Feeder," *Computers and Structures*, Vol. 62, pp. 197~203.

Meyer, B. and Tischer, P., 2001, "Glicbawls-Grey Level Image Compression by Adaptive Weighted Least Squares," in *Proc. IEEE Data Compression Conference*, Snowbird, Utah, pp. 503.

Ryan, T.P., 1997, *Modern Regression Methods*, Wiley, New York.

Strang, G., 1986, *Introduction to Applied Mathematics*, Wellesley-Cambridge Press.

Winkler, G., 1978, "Analysing the Vibrating Conveyor," *International Journal of Mechanical Sciences*, Vol. 20, pp. 561~570.

Winkler, G., 1979, "Analysing the Hopping Conveyor," *International Journal of Mechanical Sciences*, Vol. 21, pp. 651~658.