Appl Biochem Biotechnol (2011) 163:826–835
DOI 10.1007/s12010-010-9086-x

The Slope Method: A Tool for Analyzing Semi-Continuous Data

Aaron Douglas Smith · Mark T. Holtzapple

Received: 9 June 2010 / Accepted: 7 September 2010 /
Published online: 22 September 2010

Abstract The MixAlco process is a biorefinery that converts lignocellulose into useful chemicals and hydrocarbon fuels via mixed-acid fermentation. For a semi-continuous-staged fermentation train, during each transfer, discrete amounts of material are moved between fermentors and data are tabulated. Because of natural day-to-day variations, the data are inherently noisy. To calculate performance parameters (e.g., yield, conversion, selectivity, productivity), the average flowrate of each stream component must be determined. To minimize error associated with noise, three data analysis methods were compared: Average, Accumulation, and Slope. The Average method determines the flowrate by averaging the amounts moved each transfer. The Accumulation method stores the solids and liquids that exit the fermentation train in separate vessels. After an extended time period, the mass in each storage vessel is measured so the average flowrate can be calculated. The Slope method calculates the flowrate of material in each stream from the slope of the moving cumulative sum with respect to time. For all three methods, the measured rates were virtually identical; thus *accuracy* was not affected by the method. However, for the examples presented, the Average method had >40% error and the Slope method <4% error; thus, *precision* was significantly affected by the method. The Accumulation method calculated the flowrate with a single data point so it is not possible to determine the statistical error.

Keywords Slope method · Error reduction · Semi-continuous · Average method · Accumulation method · Data analysis · Staged fermentations · Analytical methods · MixAlco process

A. D. Smith (✉) · M. T. Holtzapple
Department of Chemical Engineering, Texas A&M University, 3122 TAMU, College Station, TX 77843, USA
e-mail: AaronSmith328@gmail.com

M. T. Holtzapple
e-mail: m-holtzapple@tamu.edu

Introduction

The MixAlco process is a biorefinery that converts lignocellulose into useful chemicals and hydrocarbon fuels via mixed-acid fermentation [1–3]. In the laboratory and pilot plant, countercurrent-staged mixed-acid fermentations are operated in a semi-continuous fashion (Fig. 1) [4–7]; every 2 or 3 days, a discrete amount of fermentation material is removed from each fermentor and filtered. Mass balances determine the amounts of filtered solids and liquids that are transferred in opposite directions to the respective adjacent fermentor (Fig. 1a). Between countercurrent transfers, the individual fermentors operate in batch mode (Fig. 1b).

To monitor and measure the performance of the semi-continuous fermentation, the stream masses from each countercurrent transfer are tabulated. These discrete quantities may be raw data (e.g., total stream mass or volume) or calculated component quantities (e.g., liquid, dry solid, non-acid volatile solids (NAVS), ash, or carboxylic acid (product) masses). Fermentation performance values (e.g., yield, conversion, selectivity, productivity) are calculated from steady-state average rates of these stream components. Because the semi-continuous data can be very noisy, minimizing error when calculating the average steady-state component rates is a challenge.

Historically, to determine steady-state component flowrates, the Average and Accumulation methods have been used. The *Average method* determines the flowrate by averaging the amounts moved each transfer, which represent instantaneous rates (i.e., differentiated data). As a consequence, very large errors make statistically meaningful conclusions difficult. The *Accumulation method* physically stores the filtered solids and liquid exiting the fermentor train in a storage vessel. After an extended time period (i.e., 1 or 2 months), the accumulated mass in the storage vessel is measured resulting in a single data point. It does not provide sufficient data to calculate error; thus, statistically meaningful comparisons are impossible. The *Slope method* is a recent innovation that measures the stream rate and error from the data *trend*. In all three methods, the average stream flowrate is theoretically identical; however, the objective is to

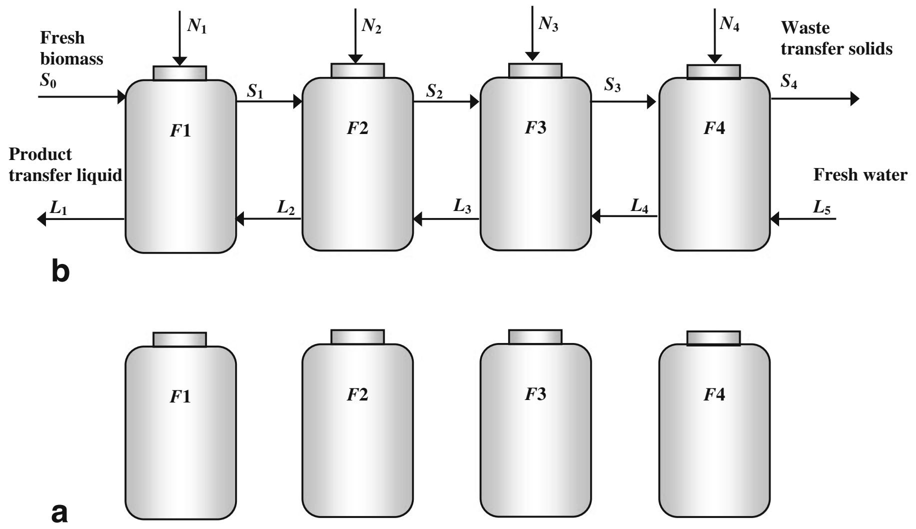

Fig. 1 **a** Countercurrent movement of material during a transfer. S_i, L_i, and N_i represent the transfer solid, transfer liquid, and nutrient streams, respectively. *Fi* represents the fermentations mass in fermentor *i*. (**b**) Batch operation between transfers

dramatically reduce the error so that statistically meaningful comparisons can be made between two different fermentation trains.

Slope-Based Methods

When analyzing data, slopes are commonly used to determine rates; however, the correct method is not always apparent and has been the subject of research. For example, Michaelis–Menten parameters are determined from the slope and intercept of a Lineweaver–Burk plot, which was a refinement of previous technique [8–13]. A few literature articles explicitly develop a "Slope method" to determine a *rate*. To simplify what was previously a laborious calculation, Thomas [12] developed a Slope method to determine the reaction rate constant of polluted water; however, his Slope method was understood to be less accurate than the traditional method. Hurwitz [9] showed that phosphorus uptake in baby chickens, which is a semi-continuous process, is linear and the slope determines the rate. To determine the rate yeast uptake sugar, Trevelyn [13] applied a material balance around a continuously fed fermentor and measured the rate of change of the extracellular sugar concentration from the slope of an appropriate plot.

In the above-referenced articles, the focus of the "Slope methods" was the *value* of interest. In contrast, the motivation for the Slope method presented this paper is *error reduction*. The Slope method is further differentiated because of its application to semi-continuous fermentations and its use of a moving cumulative sum.

Methods

To demonstrate and compare the Average, Accumulation, and Slope methods, the example data (Smith et al., unpublished data) shown in Table 1 lists the mass removed from Fermentor 4 (S_4) during each transfer. For data in Table 1, all stream rate calculations were made over the steady-state region (days 76–162). In practice, aspects of these methods have been combined (e.g., Average method used to determine total stream flowrate; Accumulation method used to determine a component rate of the same stream); however, in this paper each will be discussed and evaluated as a separate method so that the advantages and disadvantages are apparent. Unless otherwise stated, "error" refers to two standard deviations.

Average Method

In the Average method [4, 5, 14–16], samples are taken from the inlet and outlet streams to measure properties (e.g., moisture content, acid concentration, ash content) so that the mass of each stream component may be determined. Implementing the Average method requires that data be collected and samples analyzed upon each transfer, which can be a logistical challenge.

Equations 1 and 2 mathematically describe the Average method. For a given transfer period i, dM_i/dt is the stream rate of material M.

$$\frac{dM_i}{dt} = \frac{M_i}{t_i - t_{i-1}} = \frac{M_i}{T_i} \tag{1}$$

where,

M_i	amount of material M transferred
t_i	time when transfer i was executed.
$T_i=t_i-t_{i-1}$	time since the previous transfer

Table 1 Example countercurrent transferred stream data (Smith et al., unpublished—train 2, stream S_4)

A Row #	B Date	C Time of day	D Nominal time (day)	E Exact time (day)	F Material transferred (g)	G Σ Material transferred (g)
1	10/23	1617	0	0.00	26.3	26.3
2	10/25	1056	2	1.78	79.1	105.4
3	10/27	1345	4	3.89	129.1	234.5
4	10/29	1618	6	6.00	196.7	431.2
5	10/31	1645	8	8.02	166.0	597.2
65	3/23	1017	151	150.75	168.5	12013.6
66	3/25	1652	153	153.70	196.4	12210.0
67	3/27	1709	155	155.71	96.0	12306.0
68	3/30	0934	158	158.40	138.8	12444.8
69	4/1	1802	160	160.75	152.8	12597.6
70	4/3	1744	162	162.74	204.5	12802.1
Average					159.5 g/transfer	
2 standard deviations					79.9 g/transfer	

The Average method assumes the transfer period T is constant. Thus, the average steady-state stream rate $\left(d\overline{M}/dt\right)$ is determined according to Eq. 2.

$$\frac{d\overline{M}}{dt} = \frac{\sum_{i=1}^{N} M_i}{NT_i} \tag{2}$$

where,

N number of transfers (i.e., data points) during steady state

Figure 2a, b and f illustrates the Average method, where the horizontal solid line represents the average mass removed per transfer. The steady-state stream or stream component flowrate is a simple arithmetic average of the material per transfer (i.e., steady-state average of column F in Table 1), multiplied by the appropriate conversion factor (e.g., average transfer frequency; transfers/day) so that the flowrate is computed in the desired time units (typically per day). When comparing the Average method's values in Figs. 2a, b and f and 3, the apparent discrepancy results from a difference in units. To determine the error of the average, the standard deviation (or double standard deviation; 95% confidence interval) of the data set (Table 1, column F) is calculated.

Accumulation Method

In the Accumulation method [4, 5, 14–16], solids and fluids exiting the fermentation train are *accumulated* in a container once steady state begins. The accumulated material is frozen and analyzed at the end of the fermentation. Because only *one* measurement is used to determine the average flowrate and stream component masses, it is impossible to determine the statistical error. An additional limitation is that the Accumulation method can only be applied to outlet streams (i.e., product transfer liquid (L_1), and waste transfer solids (S_4)).

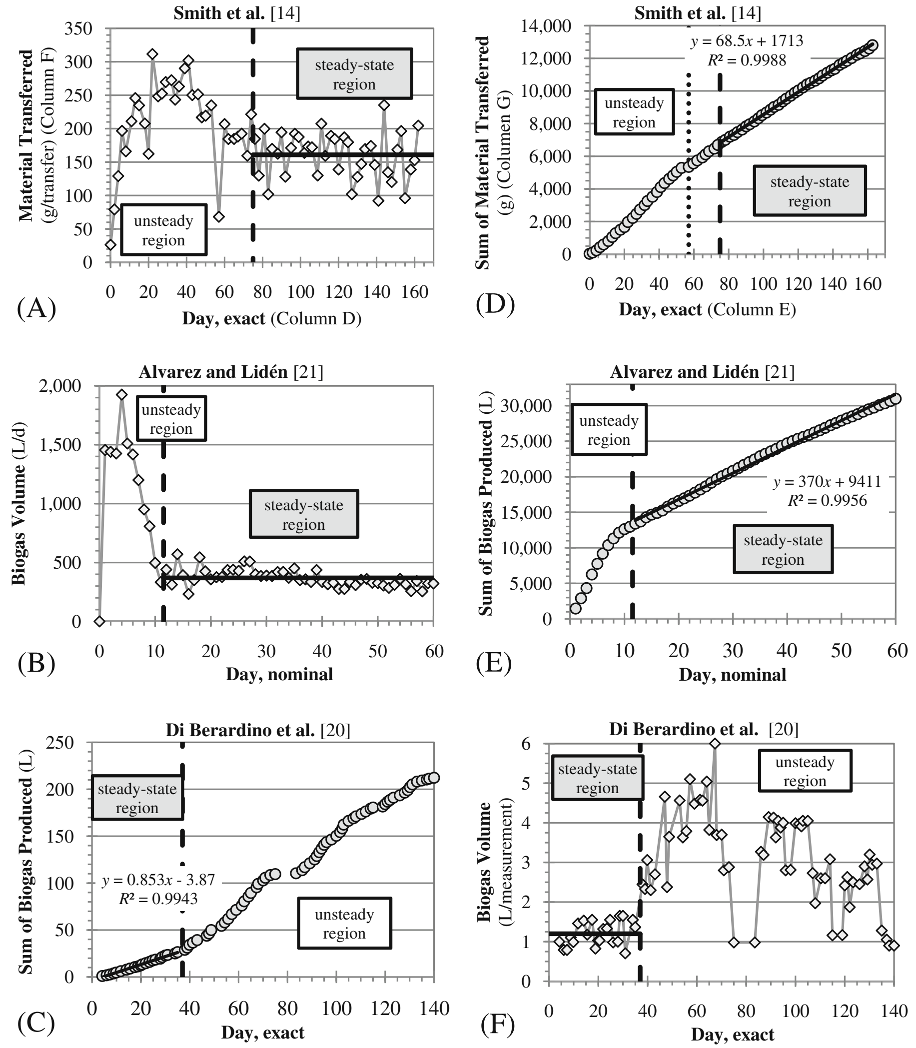

Fig. 2 *Dashed line* demarcates the steady-state and unsteady regions. *Solid line* represents steady-state trend/average. **a**, **b**, **f** For each respective study, raw semi-continuous data (quantity per time increment) are plotted with respect to time (Average method); **c**, **d**, **e** the moving cumulative sum of the respective material/gas with respect to time (Slope method). *Dotted line* (**d**) represents the beginning of an observable steady-state region using the Slope method

Although it can be used to determine the *overall* performance of the fermentation train, it cannot be used to evaluate the performance of a *single* fermentor.

Mathematically, the Accumulation method (Eq. 3) is similar to Eq. 1. The numerator is the sum of material M transfer, but the individual masses are unknown. The denominator is the length of the steady-state (SS) period.

$$\frac{d\overline{M}}{dt} = \frac{\sum M_i}{t_{\mathrm{SS\ end}} - t_{\mathrm{SS\ start}}} \tag{3}$$

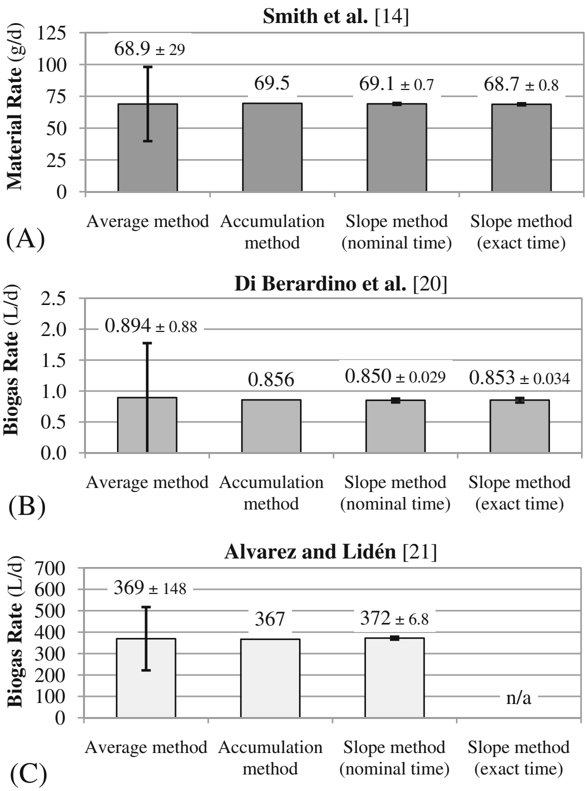

Fig. 3 Comparison of steady-state material flowrate calculation methods for the three semi-continuous data sets. *Error bars* represent two standard deviations (95% confidence interval). For comparison, all flowrates and errors are evaluated over the same time period. Errors were calculated according to methods described in the "Methods" section

Slope Method

The Slope method, illustrated in Fig. 2c, d and e, determines the average flowrate based on the data *trend* rather than averaging discrete transfer quantities. The amount transferred M_i and time are treated as independent variables. To determine the average flowrate, the average mass of M transferred and average transfer time period are calculated separately and then divided, as shown in Eq. 4.

$$\frac{d\overline{M}}{dt} = \frac{\frac{\sum_{i=1}^{N} M_i}{N}}{\frac{\sum_{i=1}^{N} (t_i - t_{i-1})}{N}} \tag{4}$$

Adjusting time such that steady state starts at t_0=0, Eq. 4 may be simplified and rearranged into a "y=m·x" form.

$$\sum_{i=1}^{N} M_i = \frac{d\overline{M}}{dt} \cdot t \tag{5}$$

Equation 5 shows that the average flowrate of material M may be determined from the slope (not the y-axis value; Fig. 2c, d and e) of the moving cumulative sum of M_i with respect to time (e.g., Table 1; column G).

Equations 6–11 [17] outline the regression formulas used to calculate the slope m and the standard deviation D_m of the slope.

$$m = \frac{S_{xy}}{S_{xx}} \tag{6}$$

$$D_m = \sqrt{\frac{D_r^2}{S_{xx}}} \tag{7}$$

where,

x_i time data
y_i moving-cumulative-sum data
N number of (x_i, y_i) data points

$$S_{xy} = \sum x_i y_i - \frac{\sum x_i \sum y_i}{N} \tag{8}$$

$$S_{xx} = \sum x_i^2 - \frac{\left(\sum x_i\right)^2}{N} \tag{9}$$

$$S_{yy} = \sum y_i^2 - \frac{\left(\sum y_i\right)^2}{N} \tag{10}$$

$$D_r = \sqrt{\frac{S_{yy} - m^2 S_{xx}}{N-2}} \tag{11}$$

Zellmer [18] states that S looks like a variance; however, it is just the sums of squares of the deviations of x (or y) from the mean of x (or y).

In Microsoft Excel, the slope m and standard deviation of the slope D_m may be calculated using Eqs. 12 and 13.

$$m = \mathbf{SLOPE}(Y_i, X_i) \tag{12}$$

$$D_m = \mathbf{SQRT}\left(\mathbf{STEYX}(Yi, Xi)^2 / \left(\left(\mathbf{COUNT}(Yi) - 1\right)^* \mathbf{STDEV}(Xi)^2\right)\right) \tag{13}$$

where, X_i and Y_i are the x_i and y_i data arrays, respectively, and the Microsoft Excel functions are in bold.

The Slope and Accumulation methods are similar except for the time period over which the accumulation is done. In the Slope method, the accumulation period is the shortest possible (i.e., each transfer). In the Accumulation method, the accumulation period is the longest possible (i.e., the entire steady-state period). If logistics (e.g., limited labor relative to the large number of fermentation trains) prohibit the analysis of exit streams after *each* transfer, then one could compromise and accumulate over a set time period such as one week (e.g., three transfers).

Order of Calculations

In the Slope method, when using a spreadsheet to calculate performance parameters, the following order of operations should be followed:

1. From the raw data (e.g., time, total stream mass, moisture content, acid concentration), calculate the mass of each stream component (e.g., acid, NAVS, dry solid) for each transfer.
2. For each stream component, add a column in which to calculate the moving cumulative sum.
3. Using the steady-state moving cumulative sum, time, and Eqs. 6–13, calculate the component flowrate and error.
4. Calculate the desired performance parameters (e.g., yield, conversion, selectivity, productivity) from these steady-state rates.
5. Use sum-of-squared errors techniques to determine the propagation of errors for the desired performance parameters.

Determination of Steady State

Determination of steady state requires consideration of many factors (e.g., period length, noise tolerance, stability of other parameters) all of which must be evaluated in context to the process. In this paper, the steady-state period was determined by visually inspecting the data (Fig. 2), and the most reasonable period was selected. The purpose of this paper is to compare analytical techniques to determine the error of an average material flow rate, not determine when steady state begins. Although different judgements as to when a period begins and ends do influence the average rate and error, it does not significantly affect the magnitude of the error of each method (i.e., Average method has much greater error than Slope method). Thus, it is more important that the same data set be used compare each method rather than when steady state begins.

Validation

To validate the Slope method, two independent semi-continuous data sets (Fig. 1a of Di Berardino et al. [19], and Fig. 3 data set M2 of Alvarez and Lidén [20]) were obtained from the literature. The data from both papers are biogas production (volume, L) from methane digestors (Fig. 2b and f). For each of these data sets, the average rate and error is determined by the Average, Accumulation, and Slope methods (Fig. 3). The moving cumulative sum of the raw data is plotted in Fig. 2c and e.

Results and Discussion

For all three data sets, Fig. 3 shows that the Average, Accumulation, and Slope methods result in virtually identical flowrates; thus, *accuracy* is not significantly affected by the method. Therefore performance parameters based on rates determined using any of these methods will be similar. However, the error calculated by each method is dramatically different. For all three data sets, the Average and Slope methods had an error of >40% and <4%, respectively. The Accumulation method cannot determine error. The "Average Method" and

"Accumulation Method" sections evaluate the mechanics, advantages, and disadvantages of each method.

Average Method

The Average method has large error (Fig. 3) because it averages instantaneous data (i.e., discrete transfer amounts) that are noisy. Ideally, transfers would be conducted precisely on a constant time interval; however, in practice, this is not always logistically possible. The Average method falsely assumes that the time associated with each transfer is constant (Table 1; Column E); thus, each transfer amount is weighted equally with respect to time. Because the fermentation is active while in batch mode, an increase in the time interval between transfers causes the amount of each stream component to increase (e.g., carboxylic acid; product) or decrease (e.g., NAVS; reactant), thus resulting in noisy instantaneous data. Additionally, operator error, equipment, and feedstock fluctuations contribute variability.

Accumulation Method

The Accumulation method is logistically the simplest, and greatly reduces the volume of data and analysis required. However, it has several key disadvantages that make it the least rigorous method. Because the accumulated material is quantified and characterized at the end of a steady-state period, only one data point is taken; thus, error cannot be calculated nor can statistical comparisons be made.

Because component stream masses are not quantified upon each transfer, acid concentration is the only data from which to judge the onset of steady-state operation (i.e., when sample accumulation begins). Although acid concentration is a good indicator of steady state, using it as the *sole* parameter fails to evaluate the system as a whole. The transition from unsteady to steady (or vice versa) operation is not readily apparent until well after steady state has been established (see Fig. 2a, b and f). If accumulation of material is initiated prematurely, the presumed steady-state measurement will be corrupted. Unlike the Average and Slope methods, the Accumulation method does not provide sufficient data to adjust the steady-state time region during post-experiment analysis so that the most stable data set is evaluated.

Slope Method

Figure 2 clearly demonstrate that the Slope method effectively smoothes semi-continuous data and reduces error because it calculated flowrates based on the data *trend* rather than averaging instantaneous rates. Because the data are smoother, the transition from unsteady to steady-state operation is more apparent. For the Smith et al. (unpublished data) data, the Slope method allows a longer steady-state period to be observed (Fig. 2d; day 58 vs. day 76), which help demonstrate control on a large-scale fermentation train, or satisfy steady-state operation time requirement (e.g., two retention times).

To determine if error in the Slope method can be reduced, the component flow rates were calculated using the exact transfer time and nominal transfer time. For the Smith et al. (unpublished data) and Di Berardino et al. data [19] (Fig. 3), the error slightly increased when the exact time was used; however, both errors are very similar. Exact time data was unavailable for Alvarez and Liden data [20]. It is the authors' opinion that the flowrate determined using the exact time is more accurate and should be used if the data are available.

Validation

For three different data sets, Fig. 2 shows that the Slope method effectively smoothes noisy semi-continuous data. For the same data sets, Fig. 3 shows that each method resulted in a similar average rate and the error trend among the methods was consistent; the Average method had the greatest error, the Accumulation had insufficient data to determine error, and the Slope method had the least error. The Slope method dramatically reduced error by a factor of 20–40 compared with the Average method. Based on the data shown in Figs. 2 and 3, the Slope method is an effective and consistent technique for determining average rates of semi-continuous material flows with minimal error.

Conclusions

For semi-continuous fermentations, the Average, Accumulation, and Slope methods produce the same average flowrates. The Accumulation method produces only a single measurement, so it is not possible to determine the statistical error. The Slope method has minimal error (<4%) compared with the Average method, which has much larger error (>40%). The Slope method is useful for analyzing data to determine fermentation performance parameters (e.g., yield, conversion, selectivity, productivity).

References

1. Holtzapple, M., & Granda, C. (2009). *Applied Biochemistry and Biotechnology, 156*(1), 95–106.
2. Holtzapple, M. T., Davison, R. R., Ross, M. K., Aldrett-Lee, S., Nagwani, M., Lee, C.-M., et al. (1999). *Applied Biochemistry and Biotechnology, 77–79*, 23.
3. Smith, A. D., Landoll, M., Falls, M., & Holtzapple, M. T. (2010). Survey of chemical production from lignocellulosic biomass. In K. W. Waldron (Ed.), *Bioalcohol production: Biochemical conversion of lignocellulosic biomass* (p. 560). Cambridge, UK: Woodhead Publishing Ltd.
4. Aiello-Mazzarri, C., Agbogbo, F. K., & Holtzapple, M. T. (2006). *Bioresource Technology, 97*(1), 47–56.
5. Domke, S. B., Aiello-Mazzarri, C., & Holtzapple, M. T. (2004). *Bioresource Technology, 91*(1), 41–51.
6. Moody, A.G. (2006). M.S. thesis, Texas A&M University, College Station, TX.
7. Smith, A.D. (2011). Ph.D. thesis, Texas A&M University, College Station, TX.
8. Donefer, E., Llyod, L. E., & Crampton, E. W. (1960). *Journal of Animal Science, 19*(2), 545–552.
9. Hurwitz, S. (1964). *The Journal of Nutrition, 84*(1), 83–92.
10. Kunz, G. J., & Leeuw, Gd. (1993). *Applied Optics, 32*(18), 3249–3256.
11. Sue, D. Y., Wasserman, K., Moricca, R. B., & Casaburi, R. (1988). *Chest, 94*(5), 931–938.
12. Thomas, H. A., Jr. (1937). *Sewage Works Journal, 9*(3), 425–430.
13. Trevelyan, W. E. (1966). *The Biochemical Journal, 100*, 556–563.
14. Agbogbo, F. K., & Holtzapple, M. T. (2006). *Applied Biochemistry and Biotechnology, 129–132*, 997–1014.
15. Chan, W. N., & Holtzapple, M. T. (2003). *Applied Biochemistry and Biotechnology, 111*, 93–112.
16. Thanakoses, P., Black, A. S., & Holtzapple, M. T. (2003). *Biotechnology and Bioengineering, 83*(2), 191–200.
17. Skoog, D. A., West, D. M., & Holler, F. J. (1996). *Fundamentals of Analytical Chemistry* (7th ed.). Philadelphia, PA: Saunders.
18. Spreadsheet Equivalents of Least Squares Formulas (1997). Available from: http://zimmer.csufresno.edu/~davidz/Stat/LLSTutorial/Formulas/LLSFormulas.html#SWH7th. Accessed 31 Mar 2010
19. Di Berardino, S., Costa, S., & Converti, A. (2000). *Bioresource Technology, 71*(3), 261–266.
20. Alvarez, R., & Lidén, G. (2008). *Renewable Energy, 33*(4), 726–734.