Conference on Applied Statistics in Agriculture | 1991 - 3rd Annual Conference Proceedings

EVALUATION OF METHANE GAS PRODUCTION IN A SIMULTANEOUS REGRESSION SYSTEM

Mark W. Jenner

Jon Maatta

Dennis M. Sievers

EVALUATION OF METHANE GAS PRODUCTION IN A SIMULTANEOUS REGRESSION SYSTEM

Mark W. Jenner
200 Mumford Hall
University of Missouri-Columbia
Columbia, MO 65211

Jon Maatta
Department of Mathematics
Plymouth State College
Plymouth, NH 03264

Dennis M. Sievers
211 Agricultural Engineering
University of Missouri-Columbia
Columbia, MO 65211

Abstract

Methane gas production is a function of volatile solids activity in anaerobic digesters. Increasing the solids retention time of the swine manure digester system without increasing the hydraulic retention time would theoretically increase the methane gas production efficiency. Coagulation treatments were performed on the effluent of the second digester in a system of two digesters in series.

The objective of this paper is to describe mathematically the relationship of the coagulation treatments in the second digester to biogas production and volatile solids retention. An initial, single equation, ordinary least squares regression produced statistically significant parameter estimates, but failed to accurately describe the treatment activity occurring in the second digester. To assess the treatment activity of the second digester and account for the activity of first digester on the second, data was regressed through a simultaneous equation system. Both two-stage and three-stage least squares regression were examined.

1. Introduction

One of the difficulties associated with anaerobic digestion of livestock wastes is maintaining a high solids retention time (SRT) in the reactor while keeping the hydraulic retention time (HRT) to a minimum. Minimum HRT results in a smaller reactor size and economical savings.

This study was concerned with increasing the retention time of the small organic particles (<0.210 mm in diameter) contained in swine manure inside an anaerobic baffle reactor (BR).

Biochemical analysis of swine manure at the University of Missouri indicated that over 50% of the organic fraction of swine manure was found to be smaller than 0.210 mm in diameter and contained approximately 40% of the potential methane in the waste (Sievers et al, 1980). It was also found that these fine organic particles, composed primarily of protein, lipids and cellulose, were the slowest to degrade (in an anaerobic environment). If these small, slowly degrading organic particles could be retained in the digester system for a longer time without increasing the digester liquid volume, methane gas production efficiency could be increased.

The vertical baffles, in a BR digester, produce an over-and-under liquid flow, which restricts the flow of suspended solids. This allows settling of solids within the BR digester. Previous work with an anaerobic BR treating swine waste Sievers (1988) suggested that the SRT of these small particles could be increased by developing a sludge layer within the BR or increasing the density of the particles.

The objectives of this study are two-fold. First, the objective of the empirical study was to increase the SRT of the small organic particles in a BR by the addition of a chemical coagulant. Coagulation should increase the density of the small organic particles allowing them to remain settled in the BR for longer periods of time resulting in increased digestion and methane recovery. And second, to describe mathematically the influence of the coagulant on biogas production and organic volatile solids (VS) accumulation in the BR.

2. Data Collection

The digestion system consisted of one 10 liter conventionally mixed digester (D1) and a second digester, a 5 liter BR (Figure 1, Jenner, Mark W., 1990). Loading a liquid volume of 1 liter per day, at a rate of 4 g VS per day, the combined two digester system had a 15 day HRT.

This system design was chosen because it allowed examination of the slowly degrading particles in the BR. Siever's work with potential methane rate indicate that 89% of the digestion takes place in the first 10 days of digestion (Figure 2). The next 8% of digestion takes place in the next 5 days. The primary component of the residual VS after 10 days of digestion are the slowly degrading VS particles. Coagulating these particles within the enhanced settling environment of the BR should increase the SRT of the digester system.

Extensive screening of two polymers was conducted on the digester effluent to determine most effective concentration rate, mixing speed and duration. Both digesters were stabilized (four HRT cycles) and reference samples of effluent were taken. The BR digester was loaded with the effluent from the first digester (D1).

The D1 effluent in the D1-BR digestion system was exposed to two separate treatments. Treatment 1 completely stopped BR digestion. An alternative treatment, Treatment 2, was successfully examined without cessation of BR digestion.

<u>Treatment 1</u> The first system coagulation treatment, 300 ppm Percol 763 mixed into digester effluent, resulted in a significant increase in the BR volatile solids concentration. Within 10 days all gas production had stopped. The BR was flushed with effluent without polymer for two weeks with no success in restarting gas production. At that time the BR was opened, all heavy floc was physically removed, the BR resealed and BR digestion restabilized.

<u>Treatment 2</u> During the BR flushing period a portion of the digester effluent was kept for further testing of polymer concentrations and mixing speeds. Based on supernatent turbidity, we found that we still achieved a significant measure of coagulation using 150 ppm Percol 763, mixed at a higher speed for a longer duration. Loading of this treatment continued for six weeks (eight HRT cycles) with no apparent change in gas production or SRT in the BR.

3. Mathematical Models

The objective of this statistical analysis is to develop a relationship between the coagulation treatments and BR gas production and VS. Initially, recorded variables were transformed and modelled using all available theoretical and empirical information. Later the regression analysis was expanded to a simultaneous system of equations to obtain more efficient parameter estimates.

Initial Ordinary Least Squares Regression Model

Initial model specification was based on the direct relationship between methane gas production and quantity of VS loaded. VS provide the carbon source for methane (CH_4). Data for explanatory variables from both D1 and BR digesters were pooled and regressed in a single equation model (ordinary least squares, OLS). This was based on the assumption that the CH_4:VS relationship was the same for both the D1 digester and the BR (as in the ultimate methane function describe in Figure 2). This OLS model would be more appropriate if no treatments had been imposed on the BR.

A single OLS regression equation was specified and tested. Indicator variables were used to flag BR observations as they entered the model. Indicator variables were also used for the two coagulation treatments and the zero polymer loading periods. In this initial model, the dependent variable, CH4YHAT, was a transformation of biogas production.

$$E(CH4YHAT) = \beta_0 + \beta_1(VSHR)_1 + \beta_2(DGADCH4)_2 + \beta_3(WKNM)_3 + \beta_4(DPLA)_4 + \beta_5(DPLB)_5 + \beta_6(DPLD)_6 \quad (1).$$

CH4YHAT (dependent variable): Methane production efficiency (liters of CH_4/g VS loaded).

VSHR (independent variable): Composite variable of VS retained and time per loading period.
DIGESTER (independent variable): Indicator variable that tracked individual digester performance. Multiplied by all BR specific variables.
*DGADCH4** (independent variable): Measure of D_1 activity in BR observation.
*DPLA** (treatment indicator variable): Initial polymer loading (300 ppm).
*DPLB** (treatment indicator variable): Flushing of initial treatment.
*DPLD** (treatment indicator variable): Second polymer loading (150 ppm).
WKNM - Week number, a time series reference.

* Indicates parameters enter model only when BR observations are active.

The fit of the simple OLS model (Equation 1) was statistically significant (Prob.>F=0.0001), explaining 96% of the variability (Table 1). The parameter estimates were all significant except the two variables pertaining to the first coagulation treatment (DPLA and DPLB). Although these two variables had the greatest impact on the BR (gas production went to zero), they were not significant in the model. This OLS specification of the model tested the data against a single continuous function (Figure 2). The OLS model did not model multiple functions occurring simultaneously within the D1-BR digestion system.

The Simultaneous Regression System

The OLS model specification did not describe the relationship between the coagulation treatment and BR VS accumulation or biogas production. For this to occur at least two equations would be necessary. System regression methods were reviewed and two-stage least-squares (2SLS) and three-stage least squares (3SLS) regression techniques were examined.

These regression system techniques allow the dependent variable in one equation to be an independent variable in another. The "active" dependent variable in an equation is referred to as the *endogenous* variable and the explanatory variables in the equation are the *exogenous* variables. These exogenous variables are composed of independent variables, dependent variables from other equations in the system, lagged variables and indicator variables.

There are four assumptions that must be met in order to get appropriate estimates with 3SLS regression (Koutsayiannis, A., 1978). First, the complete specification for the entire model is known. Second, the random term of each equation is serially independent (non-autocorrelation). The random variables of the various relations of the system are contemporaneously dependent. And finally, the system is overidentified.

Model Specification It is imperative that the model be correctly specified. The OLS model was statistically significant but the specification was incomplete. A measure of gas production was the only endogenous (dependent) variable estimated. VS was included in the initial OLS model, but was not specified as an endogenous variable. Using a system of correctly specified endogenous relationships to model treatment behavior produces more appropriate and efficient parameter estimates.

Within the bounds of the data collected, there were two relationships (biological and physical) occurring in the BR simultaneously. The endogenous variable of the biological relationship is biogas production. This represents the yield of gas produced by the biology of the digester given the exogenous (independent) inputs, which include VS. The endogenous variable of the physical system, VS, is estimated from explanatory variables that influenced level of VS in the BR. From these two endogenous relationships within the BR, there are two equations specified.

A third equation is included in the system to model the biological output of D1, biogas production. This D1 activity equation offers a measure of contribution to the two BR equations. This equation is important, because the BR received only residual materials from the operation of D1. A separate physical relationship in D1 is not included, because the exogenous variables did not differ between the D1 biological and physical systems. Inclusion of two D1 system equations creates multicollinearity problems due to the high correlation between D1 gas production and VS activity.

The formulation of these three system equations was based on both anaerobic digestion theory and on empirical observations. The kinds of data and the number of observations collected introduced restrictions into the analysis. A range of physical, chemical, and biological data was collected as a part of this study, but much of this was taken for reference at several stages and did not represent continuous digester activity. These restrictions do not compromise the model, but limit the interpretation of the results to the establishment of a significant treatment/yield relationship. More extensive sampling of all data would have provided more informative results.

The three relationships described in the system are:

$$E(BIOGAS_{D1}) = f((\text{biological parameters})_{D1}) \tag{2a}$$

$$E(BIOGAS_{BR}) = f((\text{biological parameters})_{BR}, (VS_{BR}) \text{ and } BIOGAS_{D1}) \tag{2b}$$

$$E(VS_{BR}) = f((\text{physical parameters})_{BR} \text{ and } VS_{D1}) \tag{2c}$$

Serial Independence Another important assumption is serial independence. The variance of the random error term may be seriously underestimated if the error terms are autocorrelated. Autocorrelation causes inefficient predictions in forecasting models. The model described here is a cross-sectional model, which implies that autocorrelation in this circumstance may not be as influential.

Contemporaneous Dependence The contemporaneous dependence assumption implies that a regression "system" is necessary. Without this simultaneous component, efficient relationships can be estimated in the single equation OLS model. There should be one equation describing each endogenous variable.

Overidentification The fourth assumption assumes overidentification. Identification implies that each equation in the system has a unique statistical form. Overidentification implies that the number of variables excluded from a particular equation is greater than the number of equations less 1.

The conceptual framework of 3SLS regression is as follows. The first stage represents an OLS regression of each equation. In the second stage (2SLS), the predicted endogenous (dependent) variables, from the OLS regression, replace original endogenous variables. Parameter estimates in third stage (3SLS) incorporate the error term estimates generated by a matrix of variances and covariances from the second stage.

4. Results of the Analysis

The three system equations in the specified model are described in Equation 3. The equations are all significant, and fit within the assumption framework of 3SLS. The model is correctly specified. First order autocorrelation is not influential. The explanatory variables, particularly in the BR biogas equation (3b), are contemporaneously dependent. And the system is overidentified.

Model Specification

The Full Model

$$D1GAS_1 = \gamma_{11}BOD10_1 + \gamma_{12}MXD_1 + \gamma_{13}WEEK_1 + \gamma_{14}DGMVS_1 + u_1 \qquad (3a)$$

$$BRGAS_2 = \beta_{21}D1GAS_1 + \beta_{22}BGMVS_3 + \gamma_{21}DGMVS_2 + \gamma_{22}NTRTC_2 + \gamma_{23}NTRTD_2 + u_2 \qquad (3b)$$

$$BGMVS_3 = \gamma_{31}DGMVS_3 + \gamma_{32}BFVS_3 + \gamma_{33}NTRTC_3 + \gamma_{34}NTRTD_3 + \gamma_{35}PVOLIN_3 + u_3 \qquad (3c)$$

where β's represent the coefficients for endogenous variables from other equations and the γ's represent the coefficients for the remaining exogenous variables, where,

D1GAS (endogenous): Digester D1 measured biogas production per loading period (liters/period).
BRGAS (endogenous): Digester BR measured biogas production per loading period (liters/period).
BGMVS (endogenous): VS retained BR, in grams. A transformation of measured VS effluent concentrations and effluent volumes.

BOD10 (exogenous): (D1) A transformation of the Ultimate Methane equation $(1-e^{-kt})$, where k = a constant, 0.217, and t is an estimation of the 10 day HRT, depending on the duration of the loading period.
WEEK (exogenous): week number.
MXD (exogenous): An indicator variable where D_1 digester mixer remained on for a full 24 hours instead of the regular 5 minutes within a 24 hour period.
DGMVS (exogenous): VS retained in D1, in grams. A transformation of measured VS effluent concentrations and effluent volumes.
NTRTC, *NTRTD* (exogenous): Respective 300 ppm and 150 ppm polymer coagulation treatments.
PVOLIN (exogenous): The % volume of D1 effluent that was loaded into BR as influent. This fluctuated depending on D1 effluent volumes needed for chemical samples, and mixing tests.
BFVS (exogenous): VS concentration of BR effluent, a measured value.

The Reduced Model

$$D1GAS_1 = \gamma_{11}BOD10_1 + \gamma_{12}MXD_1 + \gamma_{13}WEEK_1 + u_1 \qquad (3a)$$

$$BRGAS_2 = \beta_{21}D1GAS_1 + \beta_{22}BGMVS_3 + \gamma_{21}DGMVS_2 + \gamma_{22}NTRTC_2 + u_2 \qquad (3b)$$

$$BGMVS_3 = \gamma_{31}DGMVS_3 + \gamma_{32}BFVS_3 + \gamma_{33}NTRTC_3 + \gamma_{34}PVOLIN_3 + u_3 \qquad (3c)$$

In the D1 gas production equation (D1GAS), digester volatile solids (DGMVS) was not significant and dropped from the model. NTRTD, the treatment 2 variable was not significant in either BR gas production (BRGAS) or BR VS retained (BGMVS). The 2SLS and 3SLS t-test probabilities for reduced model, parameter estimate significance are listed in Table 2.

First Order Autocorrelation

Serial independence is inconclusive for the two BR equations (Neter, J., W. Wasserman and M. H. Kutner, 1989). The Durban-Watson first order autocorrelation statistics (1.586 and 1.191 for BRGAS and BGMVS, respectively) are between the upper and lower bounds for the 31 observations used in the calculation of the BR gas production and VS retention. In the reduced model the Durban-Watson number for the D1GAS equation is just below the lower bounds of the test. Therefore first order autocorrelation is present and could be improved by correction. Again the autocorrelation assumption is more influential in econometric forecasting models than in cross-sectional models as presented here.

Contemporaneous Dependence

Although endogenous variables are present exogenously only in the BRGAS equation (contemporaneous dependence), the covariances generated are regressed in the third stage to provide efficient parameter estimates. The cross model correlation is presented in Table 3. Care must be exercised at interpreting these values. The correlation values reflect the activity of this study only and should not be used for forecasting estimates. These correlation values incorporate the influence of three coagulation treatments (two treatments and the control) in their formulation. The correlations between the endogenous variables are not as meaningful examined separately as is the model's ability to incorporate the variances and covariances between endogenous variables into the analysis to provide more efficient parameter estimates.

Overidentification

The system is overidentified. Results of the Order Identification Test are described in Equation 4 and Table 5.

$$\begin{bmatrix} \text{Total number of} \\ \text{excluded variables} \end{bmatrix} \geq \begin{bmatrix} \text{number of} \\ \text{equations less 1} \end{bmatrix} \qquad (4)$$

System predicted and actual biogas values for D1 and BR digesters are plotted over time in Figure 3. The duration of the first coagulation treatment, NTRTC, was week 6-9. The duration of the second coagulation treatment, NTRTD, was week 12-17.

BR predicted and BR and D1 actual VS values are plotted over time in Figure 4. Predicted D1 VS values were not available since D1 VS was not estimated in 3SLS. One curious phenomenon occurring during NTRTD treatment was negative values (actual and predicted) for VS retained in the BR. This means that more VS were leaving the BR than remaining. No explanation is available.

5. Summary

The influence of coagulation treatments within a baffle reactor (BR) on biogas production and swine manure volatile solids (VS) retention was described mathematically. The BR was connected in series to a conventionally-mixed anaerobic digester and the BR design was intended to increase the retention time within the digestion system of the slowly degrading VS particles. Initially, a single equation, ordinary least squares regression produced statistically significant parameter estimates, but failed to accurately describe the treatment activity occurring in the second digester.

To assess the treatment activity of the second digester and account for the activity of first digester on the second, data was regressed through a simultaneous equation system. Both two-stage (2SLS) and three-stage least squares (3SLS) regression were examined.

In the reduced 3SLS model, was the only significant coagulation treatment parameter estimate. NTRTC was the treatment that stopped BR digestion and represents the influence of the high concentration of polymer on BR biogas production and on BR VS retention. The highly significant parameter estimate was negative, implying treatment 1 was inversely proportional to BR biogas production. The other coagulation treatment, NTRTD, was not significant in the full model and was removed in the reduced model.

The 3SLS was appropriate and effective at establishing relationship significance between coagulation treatments and BR biogas and VS retention. 3SLS provided more efficient parameter estimates than 2SLS, because it incorporated the variances and covariances into the final regression stage.

6. Acknowledgements

The authors gratefully acknowledge the theoretical guidance from colleagues in the Agricultural Economics Department at the University of Missouri-Columbia and to the referees for their constructive and useful comments.

7. References

________. Proc Syslin. SAS/ETS User's Guide. Version 6, 1st Ed. SAS Institute, NC 1988.

Jenner, Mark W. 1990. Methane Production from Anaerobic Digestion of Swine Manure by Chemical Coagulation in a Baffle Reactor. University of Missouri, MS Thesis.

Koutsoyiannis, A. 1978. Theory of Econometrics. 2nd Ed. Barnes and Noble.

Neter, John, William Wasserman, Michael H. Kutner. 1989. Applied Linear Regression Models. 2nd Ed. IRWIN. Homewood, IL. 1989.

Sievers, D. M., J. R. Fischer, and C. D. Fulhage. 1980. "Design of Sedimentation Tanks for Swine Waste." Presented At 1980 Mid-Central Region Meetings of ASAE, St. Joseph, MI. March 21-22, 1980.

Sievers, Dennis M. 1988. "Particle Trapping in an Anaerobic Baffle Reactor," Paper No. 88-6606. Presented at International Winter Meetings of the American Society of Agricultural Engineers. Hyatt Regency, Chicago, December 13-16, 1988.

Sievers, Dennis M. 1991. Unpublished research on ultimate methane. Personal communication.

Table 1. Results of OLS Model.

Dependent Variable: CH4YHAT

MSE	0.00163	Adj. R^2	0.9610
F value	234.95	Prob.>F	0.0001

Independent Variables:

	t for H_o	Prob>|t|
INTERCEPT	-0.093	0.9265
VSHR	22.596	0.0001
DGADCH4	-5.888	0.0001
WKNM	4.963	0.0001
DPLD	-2.408	0.0197
DPLB	-0.948	0.3474
DPLA	-0.299	0.7661

Table 2. T-test Probabilities for Reduced Model Parameter Estimates in 2SLS and 3SLS Models.

		2SLS	3SLS
D1GAS	INTERCEPT	0.0001	0.0001
	BOD10	0.0001	0.0001
	MXD	0.0001	0.0001
	WEEK	0.0044	0.0001
BRGAS	INTERCEPT	0.1052	0.0158
	NTRTC	0.0001	0.0001
	D1GAS	0.0002	0.0001
	DGMVS	0.0442	0.0019
	BGMVS	0.5738	0.0337
BGMVS	INTERCEPT	0.0001	0.0001
	NTRTC	0.0384	0.0824 *
	BFVS	0.0001	0.0001
	DGMVS	0.0001	0.0001
	PVOLIN	0.0001	0.0001

Shaded cells represent probabilities that are not significant at $\alpha = 0.05$.

* This value was not significant at $\alpha = 0.05$, but it was significant at $\alpha = 0.10$ and kept in the model.

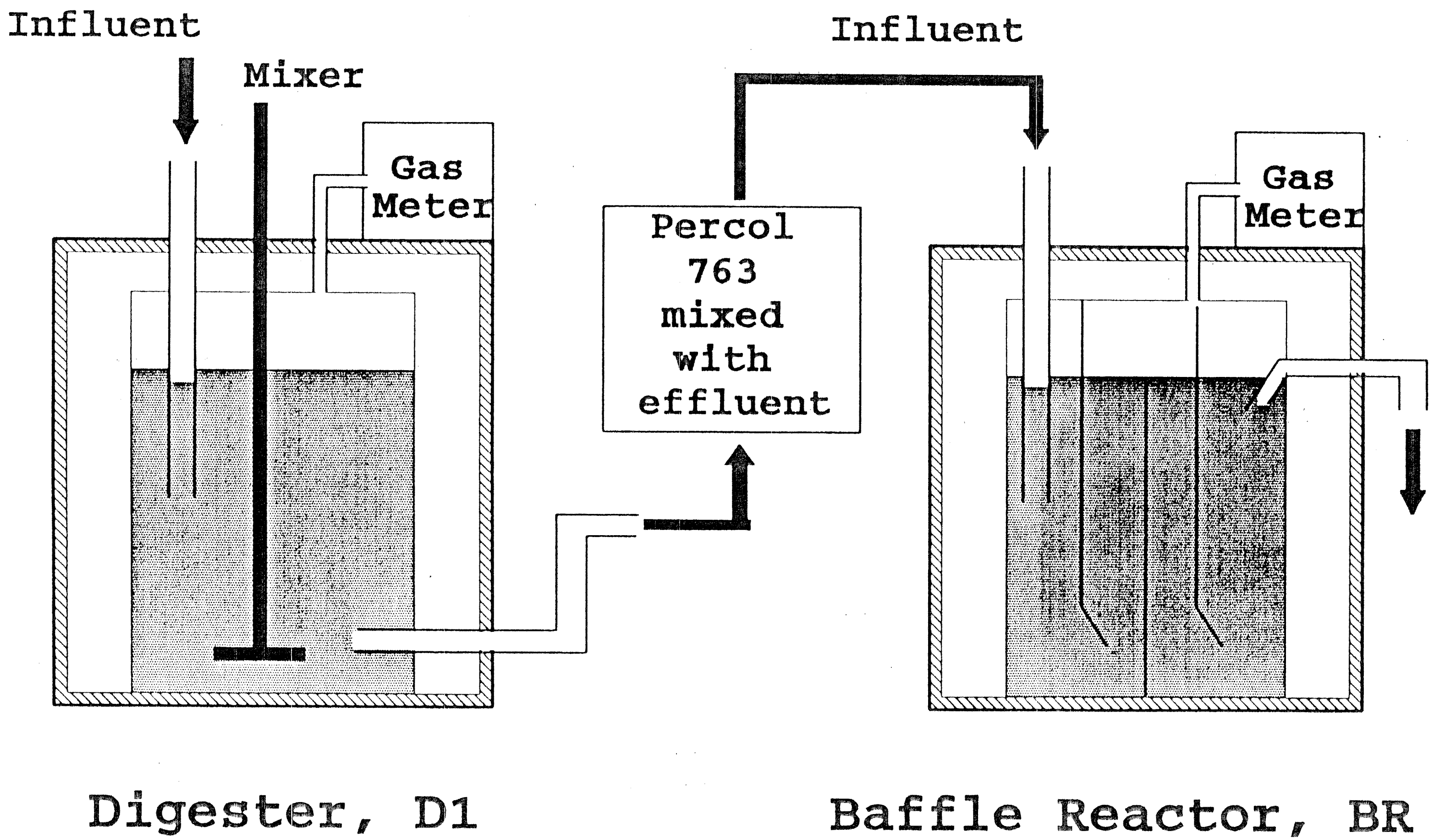

Figure 1. The Digester (D1) - Baffle Reactor (BR) System.

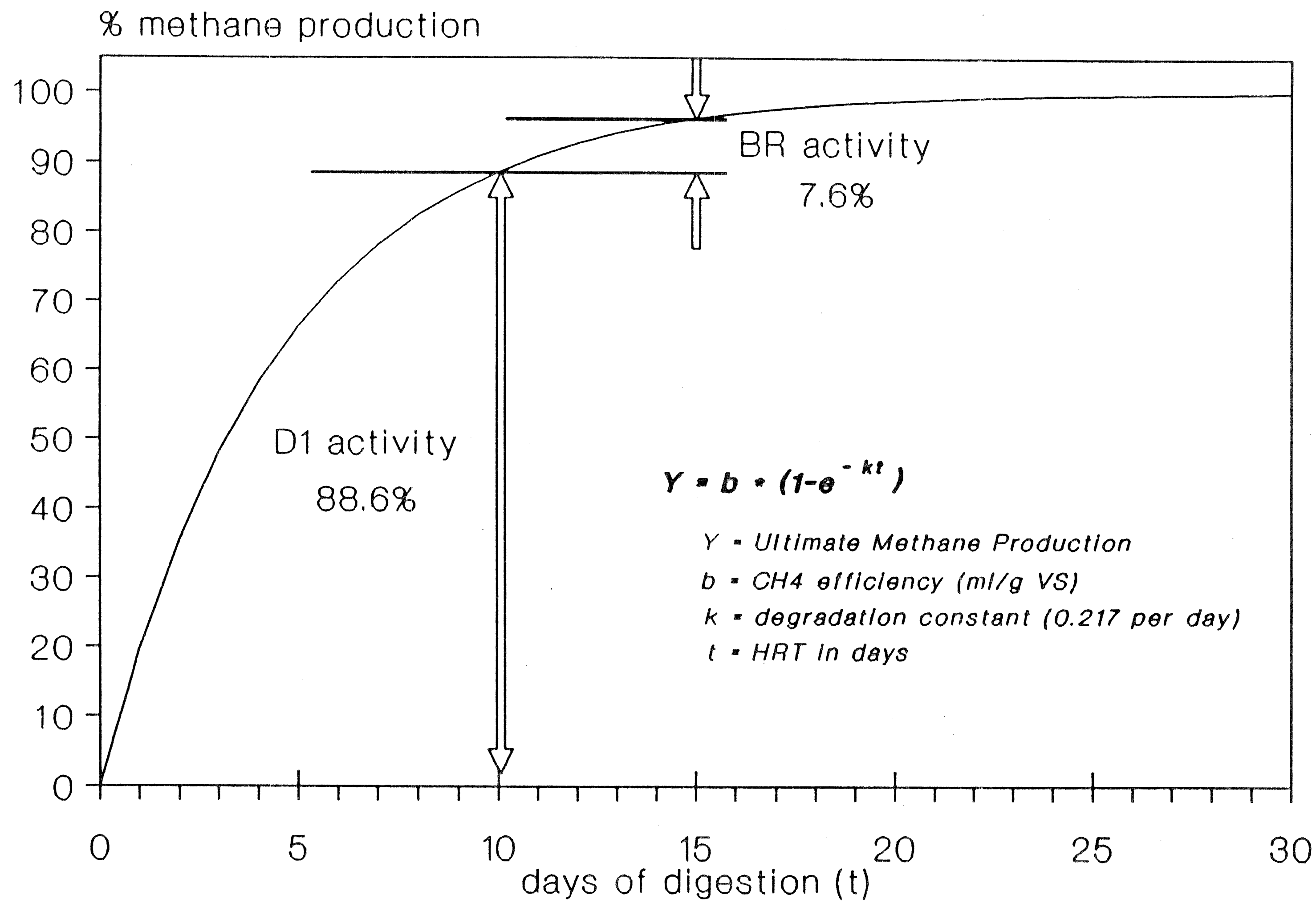

Figure 2. Cummulative methane production over days of digestion from one gram of swine manure VS.

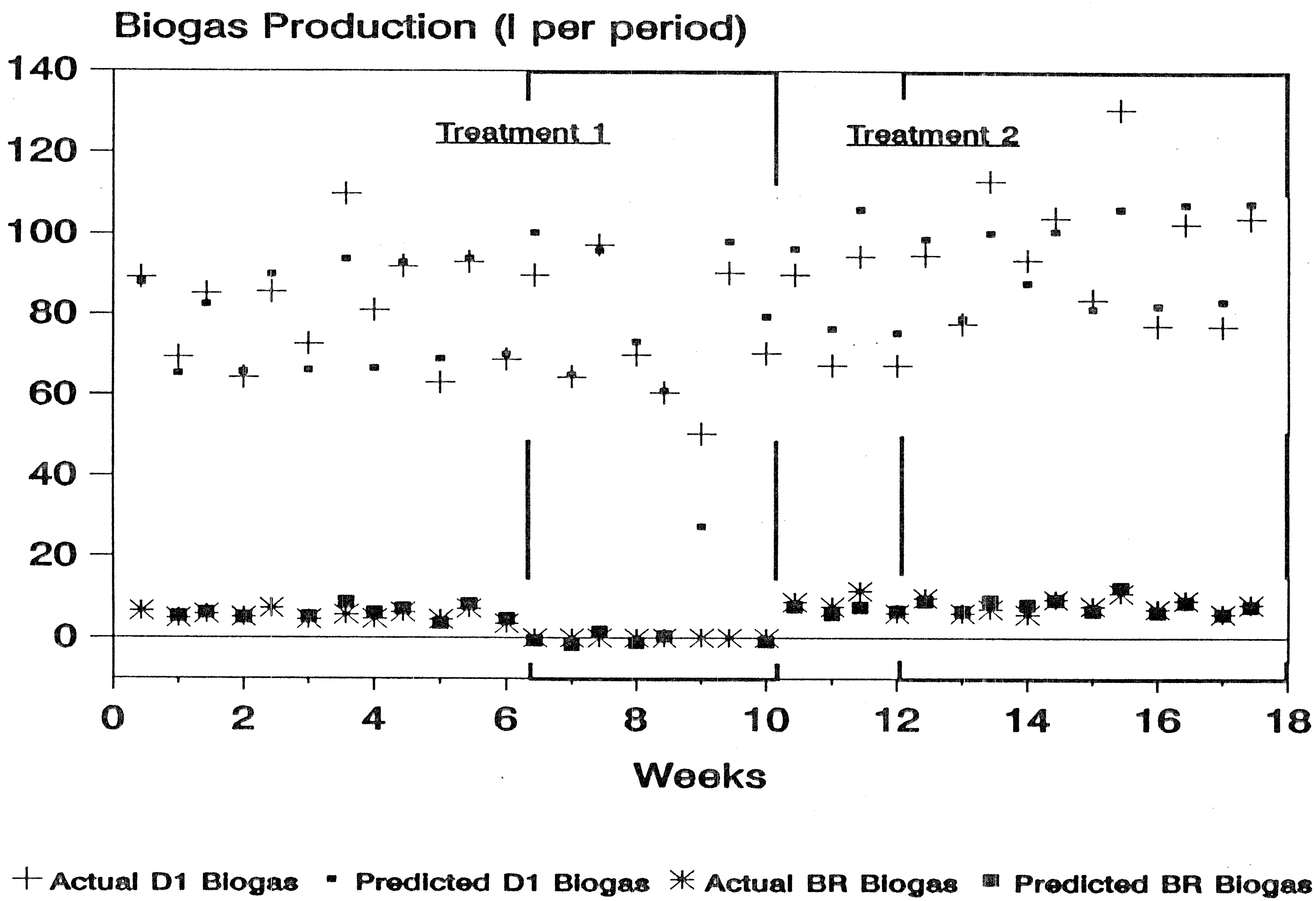

Figure 3. Actual and predicted digester 1 and baffle reactor gas production values over weeks.

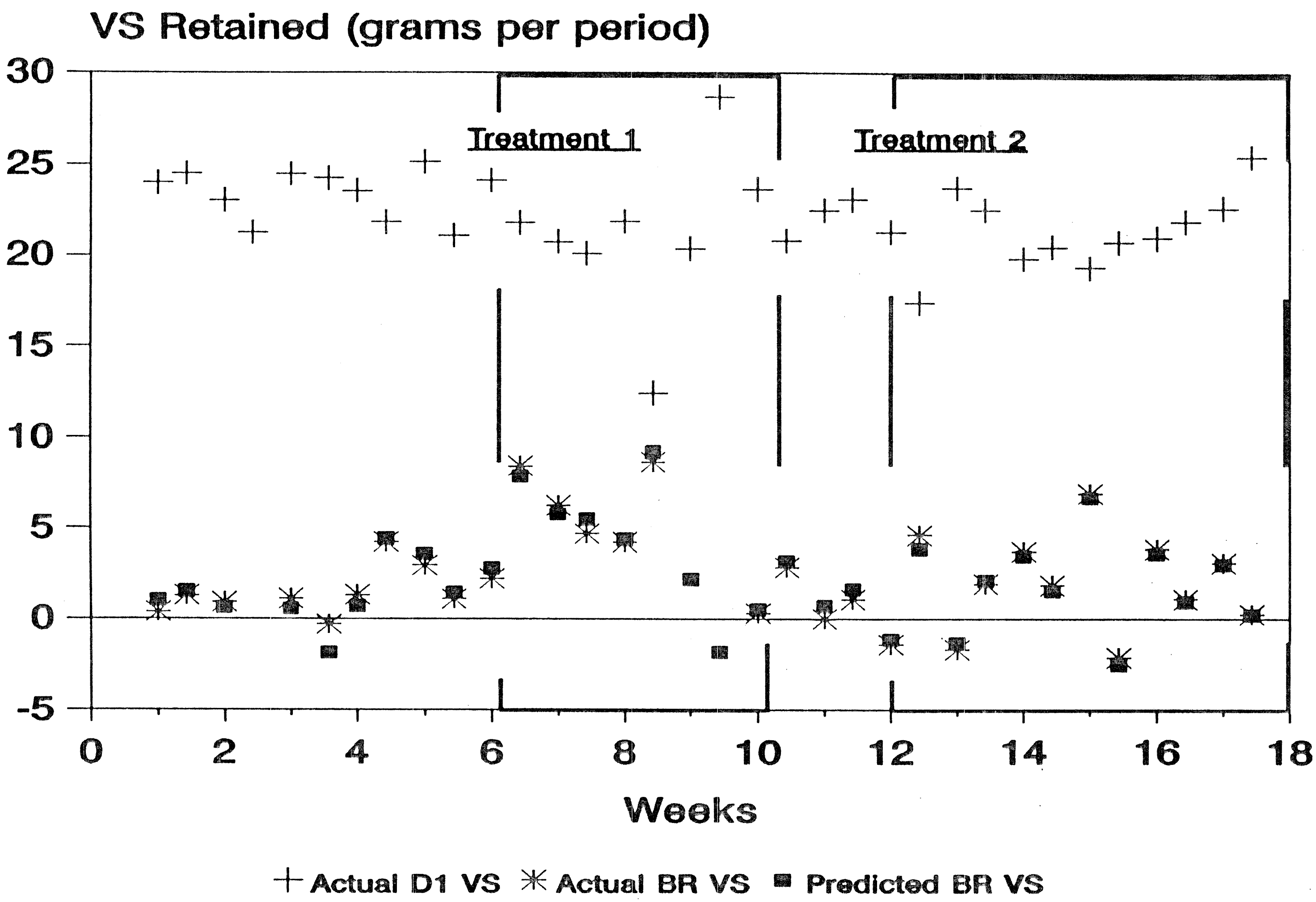

Figure 4. Actual D1 and actual and predicted BR volatile solid activity over time.

Table 3. Cross Model Correlation Matrix from Iterative Three-Stage Least Squares Estimation, SAS version 6.03.

Corr	DGBGAS	BFBGAS	BGMVS
DGBGAS	1	-0.720002807	0.4333485585
BFBGAS	-0.720002807	1	-0.254979208
BGMVS	0.4333485585	-0.254979208	1

System Weighted MSE: 1.13 with 79 degrees of freedom.
System Weighted R-Square: 0.9496

Table 4. The Order Identification Test.

Endogenous Variable	Total Variables		Equation Variables		Equations - 1
D1GAS	10	-	4	>	2
BRGAS	10	-	5	>	2
BGMVS	10	-	5	>	2