ASSESSMENT AND IMPROVEMENT OF PROCESS VARIABLE REPRODUCIBILITY IN COMPOSTING REACTORS

P. D. Schloss, L. P. Walker

ABSTRACT. *The high variability found in experimental measurements of composting state variables is the major factor that limits detection of statistically significant relationships between experimental treatments. The use of statistical techniques, such as analysis of variance (ANOVA) and power analysis, have proven to be powerful methods for assessing the effects of experimental variability on the ability to detect statistically significant differences between treatments. In two previous studies, nested ANOVAs and power analyses were used to test the effects of mixing and microbial inoculation on the biological activity in composting reactors as measured by temperature. The primary conclusion from these studies was that within–treatment variability limits the ability to detect statistically significant differences. In addition, it was shown that by controlling initial microbial populations in composting reactors, it is possible to increase process reproducibility. In this article, the results of the two previous studies are synthesized with the aid of sensitivity analyses using empirical mathematical models to describe the effects of temperature and moisture content on process behavior. Finally, a simulated power analysis was performed using the data from an inoculum study to address the effects of variability on experimental design. Combined with power analysis, the sensitivity analysis further demonstrates the great need for improved process reproducibility in the field of composting at both the research and application levels.*

Keywords. *Composting, Power analysis, Sensitivity analysis, Experimental design, Mixing, Inoculum.*

In two previous composting studies addressing the effects of periodic mixing and inoculation on process dynamics, large variations were observed in temperature and moisture content between designed replicates of a single experimental treatment (Schloss et al., 2000; Schloss and Walker, 2000). The large variation within replicates suggests that uncontrolled variables had substantial effects on the process state variables. These findings are not unique to these experiments. The limited ability to draw reliable conclusions from composting studies has also been experienced by others (Clark et al., 1977; Michel et al., 1996; Walker et al., 1999).

The two previous studies (Schloss et al., 2000; Schloss and Walker, 2000) have yielded several interesting observations regarding the variation found in temperature and moisture content data in composting. First, large variation inhibits the ability to statistically detect physically meaningful differences. For example, a statistical analysis made it clear that differences that appeared physically different (i.e., average differences in temperature greater than 10°C) were in fact not statistically significant. Second, when the contents of a reactor are mixed, the reproducibility of moisture content data is improved. This is possibly due to the redistribution of moisture across the reactor, which prevents the formation of isolated dry zones that may be present in one reactor but not another. Third, the addition of a microorganism–rich inoculum to a composting reactor reduced the variability in temperature data. It cannot be ruled out that small initial differences in a number of variables continually manifest themselves throughout the process. However, it was hypothesized that inoculation provides a homogenous initial microbial population and that the reactors used in non–inoculated studies each have different quantities and types of microbial populations so that initial differences are reduced. Finally, we encountered two unexplainable results: decreased reproducibility of moisture content data with inoculation, and decreased temperature reproducibility with mixing. Beyond unintentional differences in experimental procedure, it is difficult to ascribe meaning to these results.

A natural question to pose is "Why be concerned with process reproducibility in composting?" First, it is important to know how long the process will last. The ability to design a process so that it reaches the same state (i.e., temperature) from replicate to replicate is important for commercial practitioners who need to develop appropriate retention times and loading capacities. Second, at an academic level, if it is not possible to reliably detect large differences statistically, then it becomes difficult to reasonably optimize or understand the effects of potentially salient variables on the process. Third, from a toxicology standpoint, current regulations dictate the period of time that composted material must remain above a specified temperature in order to ensure pathogen destruction. For example, in New York State, a common method of composting potentially pathogenic

Article was submitted for review in July 2000; approved for publication by Biological Engineering Division of ASAE in April 2001. Presented at the 2000 ASAE Annual Meeting as Paper No. 00–8006.

The authors are **Patrick D. Schloss,** ***ASAE Member,*** Graduate Research Assistant, and **Larry P. Walker,** ***ASAE Member,*** Professor, Department of Biological and Environmental Engineering, Cornell University, Ithaca, New York. **Corresponding author:** Larry P. Walker, Riley–Robb Hall, Cornell University, Ithaca, NY 14853; phone: 607–255–2478; fax: 607–255–4080; e–mail: lpw1@cornell.edu.

wastes is in periodically mixed windrows. According to state law, the temperature of the composting waste must be maintained above 55°C for at least 15 consecutive days (New York State Department of Environmental Conservation, 1996). If temperatures vary by more than 15°C between replicates, it becomes more difficult to reproducibly comply with these regulations. Finally, as will be demonstrated in this article, even small differences in temperature and moisture content may manifest themselves as large differences in process performance.

The main goal of the current study was to further analyze the results, specifically the degree of variability, found in earlier studies using a power analysis and mathematical models. This analysis was then used to determine which avenues of research are most appropriate for improving the experimental design of composting studies and field–scale implementation. Finally, suggested tolerance levels for variability are proposed to address the problem of statistical reproducibility.

Materials and Methods

Reactor Description

In the two previous studies, 30 L bench–scale composting reactors made of schedule 40 PVC pipe were used (Schloss et al., 2000; Schloss and Walker, 2000). Big Red puppy food (Pro–Pet, Inc., Syracuse, N.Y.) was mixed with maple wood chips (Coastal Lumber, Cayuta, N.Y.) to obtain a C:N ratio of 18:1 and loaded into the reactors in order to achieve a dry bulk density of approximately 280 kg/m^3. The initial moisture content of the loaded synthetic food waste (SFW) in all experiments was 55% (w.b.). The contents of the reactors were aerated with water–saturated air at a flow rate of 5.25 L min^{-1}. For each study, two experimental replicates were performed sequentially. Within each experimental replicate, each treatment addressed in the study was replicated twice. Therefore, each treatment was replicated four times. The four treatments used in the first study included reactors mixed every 24, 96, and 192 hrs and reactors that were left unmixed. The treatments in the second study included reactors that used tap water as the sole source of moisture, while in the inoculated reactors, 25% of the initial moisture in each reactor was from primary wastewater. Further details of the reactor design, substrate preparation, data acquisition, sampling methodology, and experimental design have been described elsewhere (Hall et al., 1995; VanderGheynst et al., 1997; Schloss et al., 2000; Schloss and Walker, 2000).

Statistical Analysis

Because of the two experimental replicates within each study, a nested analysis of variance (ANOVA) was used to test the differences in temperature and moisture content for each treatment within the same study. This method of statistical analysis will be briefly reviewed because it is described in detail elsewhere (Sokal and Rohlf, 1995; Schloss et al., 2000; Schloss and Walker, 2000). In order to simplify the analysis and determine the effect of each treatment on temperature and moisture, the time profile for each variable was segmented into 12–hr periods. The two hours prior to the end–point of each period were used for data acquisition, and the nested ANOVA was performed for single heights in the reactor. In addition to this first set of data, a second analysis was done for each study in which the effects of each treatment on spatial gradients within a reactor were tested. As in the first analysis, the data were segmented into 12–hr periods, but the second analysis required temperature and moisture data from all height positions in the reactor for a single treatment. If differences between treatments or positions in the reactor were detected, then Tukey's least significant difference technique was used to limit compounding of the Type I error (Sokal and Rohlf, 1995).

The above methods of analysis were used for both studies, and the results are reported in detail in Schloss et al. (2000) and Schloss and Walker (2000). In addition to the ANOVA and Tukey's tests, a power analysis was performed in the second study. Power is the probability of detecting a statistically significant difference that actually exists. Typically, researchers are content with a power of 0.80, also expressed as 80%. This is interpreted as meaning that if a difference exists and an experiment is repeated 10 times, then the difference will be detected 8 of those times. Power is a function of the difference being tested (δ), the variability among replicates (σ^2), the number of replicates (n), and the method of testing the statistical hypothesis. These parameters are used to calculate the non–central parameter (ϕ^2), which describes the difference between the distribution of a randomly occurring population and the distribution representing the data used in the analysis:

$$\phi^2 = \nu_1 \left(\frac{MS_{treat}}{MS_{exper}} - 1 \right) \quad (1)$$

where

ϕ^2 = non–central parameter
ν_1 = degrees of freedom used for MS_{treat}
MS_{treat} = sum of the mean–squared error among treatments
MS_{exper} = sum of the mean–squared error among experiments.

To perform the power analysis, it was necessary to vary δ to determine the ability of the ANOVA to detect different effect sizes between any two treatments. The expression used to calculate MS_{treat} was rewritten so that δ equaled twice the difference between the mean of one treatment and the mean for all the data combined:

$$MS_{treat} = \frac{nb}{2\nu_1} \delta^2 \quad (2)$$

where b = the number of experimental groups (b = 2 for both studies).

Using the values from equations 1 and 2, the power was calculated as demonstrated by Lindman (1991):

$$\text{Power} = \Pr[F^{(\nu', \nu_2)} < F'] \quad (3)$$

where

F' = value of the non–central F distribution
F = ratio of the MS_{treat} and MS_{exper} terms
ν' = first degree of freedom for use with non–central F distribution
ν_2 = degrees of freedom for the MS_{exper} term.

The MS_{exper} terms were calculated in this study using the SPSS computer package (SPSS, Inc., Chicago, Ill.). The

power analyses were performed using an Excel '97 spreadsheet (Microsoft, Redmond, Wash.).

Mathematical Models Used in Analysis

Using SFW, Richard and Walker (1998) proposed a kinetic model describing the relationship between temperature and rate of CO_2 evolution. Their model is based on the Cardinal Temperature Model with Inflection (CTMI) proposed by Rosso et al. (1995):

$$R_{CO_2} = [R_{CO_2,opt}(T-T_{max})(T-T_{min})^2] \div \{(T_{opt}-T_{min})[(T_{opt}-T_{min})(T-T_{opt}) - (T_{opt}-T_{max})(T_{opt}+T_{min}-2T)]\} \quad (4)$$

where

R_{CO2} = rate of CO_2 evolution (g CO_2 × kg initial volatile solids^{-1} × d^{-1})
$R_{CO2,\ opt}$ = maximum rate of CO_2 evolution (g × kg^{-1} × d^{-1})
T_{min} = temperature at minimum rate of CO_2 evolution (°C)
T_{opt} = temperature at optimum rate of CO_2 evolution (°C)
T_{max} = temperature at maximum rate of CO_2 evolution (°C).

Table 1 lists the values and standard deviations of these parameters, as reported by Richard and Walker (1998).

For purposes of a sensitivity analysis, the derivative of equation 4 was taken with respect to temperature (T) and expressed as follows:

$$\frac{dR_{CO_2}}{R_{CO_2}} = \left[\frac{(T-T_{min})+2(T-T_{max})}{(T-T_{max})(T-T_{min})} - \frac{3T_{opt}-2T_{max}-T_{min}}{(T_{opt}-T_{min})(T-T_{opt})-(T_{opt}-T_{max})(T_{opt}+T_{min}-2T)}\right] dT \quad (5)$$

This sensitivity analysis will be used to predict the effect of variation in temperature on predicted process performance. For example, by varying dT, it is possible to determine the percent change in the rate of CO_2 evolution for a given temperature.

According to Haug (1993), the rate of substrate degradation can be modeled as a function of R_{CO2} using various correction factors for sub–optimal levels of moisture, oxygen, and porosity. In this study, only the moisture–based correction is analyzed because this was the only variable of the three that was substantially below optimal levels in our previous two studies. The moisture correction factor (f_{H2O}) was derived by Haug (1993) as a means of expressing the effects of moisture demonstrated experimentally by others. While the model is generally inaccurate above 60% (w.b.), its use to predict process performance at lower moisture contents has been validated in three previous studies (Schulze, 1961; Snell, 1957; Jeris and Regan, 1973):

$$f_{H_2O} = [\exp(-17.684M_{wb} + 7.0622) + 1]^{-1} \quad (6)$$

where

f_{H2O} = fraction of maximum uptake rate
M_{wb} = moisture content (w.b., fraction).

The correction factor (f_{H2O}) is the ratio of the rate of CO_2 evolution at the given moisture content and temperature to the rate of CO_2 evolution at the optimal moisture content and the given temperature.

In order to perform a sensitivity analysis to determine the effect of variation in moisture content on errors in predicting process performance, equation 6 was differentiated with respect to M_{wb}, yielding the following equation:

$$\frac{df_{H_2O}}{f_{H_2O}} = \frac{17.684\exp(-17.684M_{wb}+7.0622)}{\exp(-17.684M_{wb}+7.0622)+1} dM_{wb} \quad (7)$$

Similar to equation 5, by varying the dM_{wb}, it is possible to determine the percent change in process performance, as measured by f_{H2O}, at any given moisture content.

Table 1. Parameter values calculated by Richard and Walker (1998) for equations 4 and 5.

Parameter	Estimate	Std. Dev.
$R_{CO2,\ opt}$ (g kg^{-1} d^{-1})	178	37
T_{min} (°C)	5	N.A.
T_{opt} (°C)	58.6	5.6
T_{max} (°C)	71.6	5.7

Results and Discussion

Empirical Description of Variability

Figures 1 and 2 reveal large differences in temperature and moisture content between mixed and unmixed reactors that were not statistically significant. The vertical bars in each figure denote points in the experiment where statistically significant differences were detected ($p < 0.05$). Statistically significant temperature differences were only detected between treatments in the mixing study. These bars do not appear at some of the largest gaps between treatments for temperature and moisture content. The reason for this inability to detect statistically significant differences when they actually exist lies in the high levels of variation between reactors within treatments. Consequently, only those between–treatment differences that were located in regions in which there was small within–treatment variability were actually statistically significant.

In addition, important conclusions were drawn from the comparison of data for a single treatment within and between experimental groupings. While most composting studies fail to use multiple replications or statistical procedures to assess differences between treatments, the replication applied to this study demonstrated the large variability between a set of composting reactors operated under the same treatment conditions. For example, the temperature data shown in figure 3 were recorded at 30 cm above the reactor floor in the unmixed reactors. These profiles represent the tremendous variation typically found in composting studies. Similarly, figure 4 demonstrates the high levels of variability found in moisture content data. This figure represents the moisture content measured at 10 cm above the floor of each reactor in four replicate unmixed reactors from the first study. While

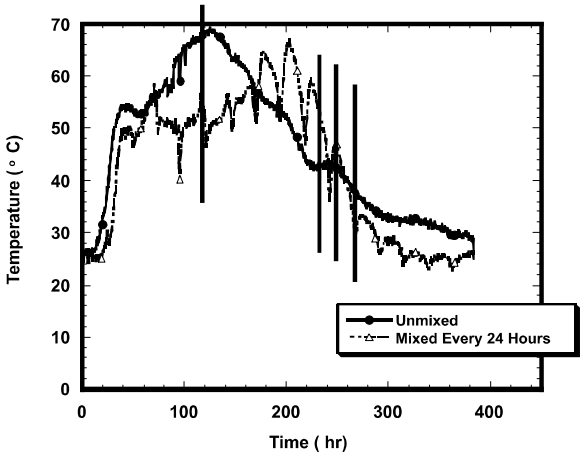

Figure 1. Comparison of average temperatures taken at 30 cm in reactors mixed every 24 hrs and those left unmixed. Vertical bars denote times at which statistically significant differences were detected between any of the four treatments.

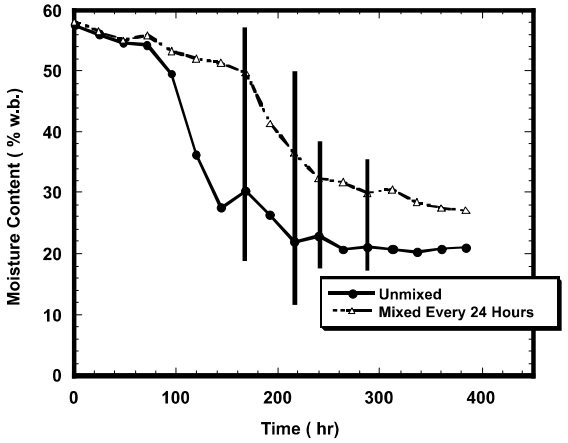

Figure 2. Comparison of average moisture contents taken at 10 cm in reactors mixed every 24 hrs and those left unmixed. Vertical bars denote times at which statistically significant differences were detected between any of the four treatments.

there is high variability throughout the process, the variability appears to increase once drying commences.

Failure to detect differences between two treatments can be attributed to two scenarios: either there was no "true" difference, or there was not enough statistical power to detect differences that actually existed. In order to determine which of the two scenarios holds for these studies, the power to detect differences of 5, 10, and 15°C, as well as differences of 5, 10, and 15 percentage points (w.b.) between treatments, was calculated for each height and time point. As shown in table 2, there were median power values of 0.06, 0.13, and 0.36 to detect differences of 5, 10, and 15°C, respectively, at 30 cm above the reactor floor in the mixing study. Median power values to detect moisture content differences of 5, 10, and 15 percentage points (w.b.) were 0.07, 0.28, and 0.63, respectively (table 3, power values for both variables from the mixing study have not been previously published). Temperature results are presented from 30 cm because this elevation was the hottest point in the reactor, while moisture

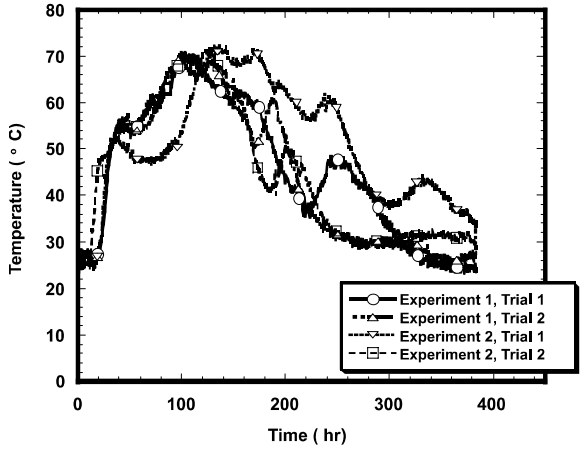

Figure 3. Decomposition of the curve representing the average temperature experienced in an unmixed reactor at 30 cm (in Figure 1) into the four replicate profiles.

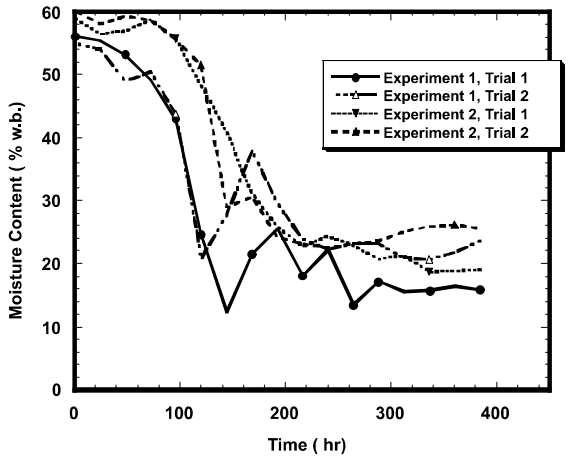

Figure 4. Decomposition of the curve representing the average moisture content experienced in an unmixed reactor at 10 cm (in figure 2) into the four replicate profiles.

content results are presented because this region dried the fastest.

Comparison of the power values from the inoculum study, the mixing study, and those calculated from Michel et al. (1996; Schloss and Walker, 2000), revealed that the greatest observed statistical power to detect differences in temperature and moisture content was in the inoculum study (tables 2 and 3). However, these values are still not ideal for sound experimental design. The inoculum study data presented in figure 5 show that, while there are a few physically significant differences, they remain undetectable due to the limited power to detect differences. Because of the obvious limits in detecting differences, it is not possible to conclude, with confidence, that the statistically significant differences detected are the only true differences present in these studies.

Effects of Mixing and Inoculation on Statistical Power

The issue of the effects of mixing and inoculation on experimental design was further analyzed through another

Table 2. Results of power analysis using temperature data taken at 30 cm and showing the number of tests that had an observed statistical power for a given range to detect three effect sizes using data from the mixing (Schloss et al., 2000), inoculum (Schloss and Walker, 2000), and Michel (Michel et al., 1996) studies (e.g., 2 of 33 ANOVAs had a statistical power between 0.60 and 0.69 when attempting to detect a 5°C difference at 30 cm in the mixing study).

Statistical Power	Small Effect Size 5°C Difference			Medium Effect Size 10°C Difference			Large Effect Size 15°C Difference		
	Mixing	Inoculum	Michel	Mixing	Inoculum	Michel	Mixing	Inoculum	Michel
0.99–				2	1		2	3	
0.95–0.98					1			5	
0.90–0.94							1	1	
0.80–0.89		1			4		1	2	3
0.70–0.79					2			3	1
0.60–0.69	2	1			1		3	2	6
0.50–0.59				2	3		4	2	5
0.40–0.49					3	3	4	3	4
0.30–0.39		5		1	1	7	4	3	3
0.20–0.29		3		7	5	9	3	7	
0.10–0.19	2	6	3	9	10	3	7	2	
0.05–0.09	29	19	19	12	4		4	2	
N	33	35	22	33	35	22	33	35	22
Mean	0.099	0.177	0.081	0.214	0.395	0.307	0.497	0.568	0.600
Median	0.06	0.08	0.08	0.13	0.27	0.30	0.43	0.51	0.60
St. Dev.	0.14	0.18	0.01	0.24	0.30	0.10	0.38	0.317	0.16
% > 0.80	0.0	2.9	0.0	6.1	17.1	0.0	12.1	31.4	13.6

Table 3. Results of power analysis using moisture content data taken at 10 cm and showing the number of tests that had an observed statistical power for a given range to detect three effect sizes in the mixing (Schloss et al., 2000) and inoculum (Schloss and Walker, 2000) studies (e.g., 13 of 17 ANOVAs had a statistical power between 0.05 and 0.09 when attempting to detect a 5% (w.b.) difference at 10 cm in the mixing study).

Statistical Power	Small Effect Size 5% w.b. Difference		Medium Effect Size 10% w.b. Difference		Large Effect Size 15% w.b. Difference	
	Mixing	Inoculum	Mixing	Inoculum	Mixing	Inoculum
0.99–		1	1	3	2	8
0.95–0.98				1	2	2
0.90–0.94				1		3
0.80–0.89			1	4	3	1
0.70–0.79		2	1	2	1	1
0.60–0.69	1		1	3	1	1
0.50–0.59		2			1	
0.40–0.49		1	3	1		1
0.30–0.39		3	1	1		
0.20–0.29	1	5	2	1	1	1
0.10–0.19	2	2		1	3	
0.05–0.09	13	2	7		3	
N	17	18	17	18	17	18
Mean	0.166	0.372	0.341	0.713	0.540	0.870
Median	0.07	0.31	0.28	0.78	0.63	0.96
St. Dev.	0.14	0.26	0.30	0.26	0.38	0.20
% > 0.80	0.0	5.6	11.8	50.0	41.2	77.8

power analysis. In this analysis, the ability to detect spatial differences within a given treatment was determined. Using the data from the mixing study (Schloss et al., 2000), the median power values to detect spatial differences of 5, 10, and 15°C was 0.08, 0.31, and 0.63, respectively, in unmixed reactors, while in reactors mixed every 24 hrs the values were 0.07, 0.20, and 0.45, respectively. The power values for unmixed reactors and those mixed every 24 hrs are presented in table 4. When considering the other intervals between mixing in this analysis, there was no discernable relationship between frequency of mixing and power (data not presented).

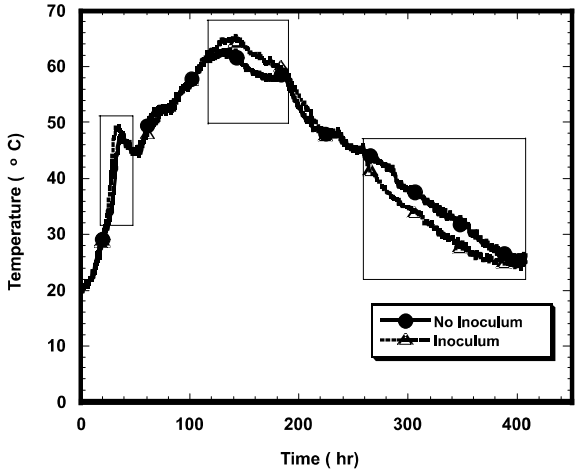

Figure 5. Comparison of average temperatures taken at 30 cm in inoculated and non–inoculated reactors. No statistically significant differences were detected between the two treatments. Potential physically significant differences are contained within the rectangular regions.

While mixing may have a beneficial effect on the process, considering the intervals between mixing tested, mixing should not be considered as a method of improving experimental design. However, comparison of power values from inoculated and non–inoculated reactors showed that inoculation resulted in improved statistical power (table 4). This analysis provided evidence that the addition of an inoculum increased the ability to detect differences in a hypothesis test (Schloss and Walker, 2000). In addition, the power analysis suggested that the lower–than–desired statistical power found in the previous hypothesis tests may have been reduced because non–inoculated reactors were used. Future experiments using an inoculum in all reactors to test the effect of a separate treatment would probably obtain greater statistical power because the variability introduced by non–inoculated reactors would be absent.

Table 4. Results of power analysis using temperature data for selected treatments and showing the number of tests that had an observed statistical power for a given range to detect three effect sizes in the mixing (Schloss et al., 2000) and inoculum (Schloss and Walker, 2000) studies (e.g., 1 of 35 ANOVAs had a statistical power greater than 0.99 when attempting to detect a 5°C difference across the bed of an inoculated reactor).

Statistical	Small Effect Size 5°C Difference				Medium Effect Size 10°C Difference				Large Effect Size 15°C Difference			
Power	Unmixed	24 h	Inoculum	No Inoculum	Unmixed	24 h	Inoculum	No Inoculum	Unmixed	24 h	Inoculum	No Inoculum
0.99–			1	1	6	2	2	3	11	3	7	9
0.95–0.98	1	1			2		1	4		1	1	
0.90–0.94			1			2	2	2		2	9	1
0.80–0.89	1	1			3		2		3	2		1
0.70–0.79	2			1		1	1		1	5	2	1
0.60–0.69	2			1			7	1	2	1	4	2
0.50–0.59			1	1	2	4	2	1	2	1	3	5
0.40–0.49	2		1	4	1	3	2		1	2	5	2
0.30–0.39	1	2	1	1	3	3	4	3	4	3	2	7
0.20–0.29	2	1	3		3	2	6	7	6	4		3
0.10–0.19	3	6	10	2	7	6	4	10	3	5	1	3
0.05–0.09	19	22	17	24	6	10	2	4		3	1	1
N	33	33	35	35	33	33	35	35	33	33	35	35
Mean	0.239	0.150	0.194	0.202	0.455	0.335	0.490	0.407	0.617	0.522	0.721	0.580
Median	0.08	0.07	0.10	0.08	0.31	0.20	0.43	0.24	0.63	0.45	0.78	0.52
St. Dev.	0.27	0.21	0.22	0.24	0.38	0.30	0.29	0.36	0.33	0.34	0.27	0.31
% > 0.80	6.1	6.1	5.7	0.3	33.3	12.1	20.0	25.7	42.4	24.2	48.6	31.4

While there appear to be relationships that may be drawn from these two studies using temperature data, the situation is more confusing when one analyzes the power values resulting from moisture content data. Table 5 summarizes the results presented in Schloss and Walker (2000), as well as the power values calculated from the data of Schloss et al. (2000). Based on the results from the mixing study, mixing appears to improve the ability to detect differences in moisture content across the bed of the reactor. However, the difference between each of the mixed treatments is small, and there is not enough information to determine if there is a trend relating the interval between mixing to statistical power. At a gross level, mixing improves the power to detect differences in moisture content. According to the power values reported in the inoculum study (Schloss and Walker, 2000), inoculation decreases the ability to detect differences in moisture content. There is not a clear explanation for why inoculation improves temperature reproducibility but increases variability in moisture content.

Meaning of Physically Significant Differences

This article and our previous studies have focused on the ability to detect temperature differences of 5, 10, and 15°C, as well as differences in moisture content of 5, 10, and 15 percentage points. Previous discussion of these values has centered on the biological meaning of the differences. For example, a difference of 15°C would differentiate between mesophilic and thermophilic temperatures. In addition,

Table 5. Results of power analysis using moisture content data for selected treatments and showing the number of tests that had an observed statistical power for a given range to detect three effect sizes in the mixing (Schloss et al., 2000) and inoculum (Schloss and Walker, 2000) studies (e.g., 3 of 18 ANOVAs had a statistical power between 0.70 and 0.79 when attempting to detect a 15% (w.b.) difference across the bed of an inoculated reactor).

Statistical	Small Effect Size 5% w.b. Difference				Medium Effect Size 10% w.b. Difference				Large Effect Size 15% w.b. Difference			
Power	Unmixed	24 h	Inoculum	No Inoculum	Unmixed	24 h	Inoculum	No Inoculum	Unmixed	24 h	Inoculum	No Inoculum
0.99–			1			3	1	2		4	4	6
0.95–0.98		1					2	2	1		1	5
0.90–0.94						1		2			1	1
0.80–0.89		1		1			2		2		1	2
0.70–0.79		1			1			3	1		3	
0.60–0.69				1				3		1	2	1
0.50–0.59			1		1		2	2				1
0.40–0.49			1	2	1		1		1	1		
0.30–0.39		1		2	1	1	4	1	1	6	2	
0.20–0.29	1		2	3		1		1	2	2	3	1
0.10–0.19	2		2	5	3	7	4		4	3	1	1
0.05–0.09	14	13	11	4	10	4	2	2	5			
N	17	17	18	18	17	17	18	18	17	17	18	18
Mean	0.077	0.224	0.202	0.270	0.188	0.336	0.456	0.650	0.325	0.476	0.657	0.833
Median	0.06	0.07	0.09	0.20	0.08	0.14	0.39	0.70	0.18	0.34	0.73	0.96
St. Dev.	0.04	0.31	0.25	0.21	0.20	0.37	0.33	0.30	0.31	0.32	0.32	0.27
% > 0.80	0.0	11.8	5.6	5.6	0.0	23.5	27.8	33.3	17.6	23.5	38.9	77.8

according to the discussion provided by Haug (1993), a moisture difference of 15 percentage points would cause the moister material to degrade faster than the drier material. This justification provides an empirical method of determining when physically significant differences exist. However, application of a sensitivity analysis using mathematical models would strengthen the rationale for the differences that are tested for in the power analyses.

Using the CTMI model (eq. 4) implemented by Richard and Walker (1998) and equation 5, figure 6 was constructed by evaluating equation 5 when dT equaled 1, 5, 10, and 15°C. Based on figure 6, if two sets of reactors are 1, 5, 10, or 15°C different at a given time, then one could determine what effect this difference had on the rate of CO_2 evolution, a parameter directly related to the rate of substrate degradation. For example, if two reactors were statistically 60°C, but actually differed by 10°C, then the rate of CO_2 evolution in the two reactors would differ by 41.4%.

Several important trends are suggested in figure 6. First, the percent difference in the rate of CO_2 evolution increases as the temperature difference increases. Second, the change in the percent difference in the rate of CO_2 evolution increases with respect to temperature. Finally, even small differences of 5°C can result in large differences in the rate of CO_2 evolution (33.1%).

Based on the limitations of instrumentation and the sensitivity analysis, it is suggested that future power analyses focus on reporting the power to detect differences of 1, 5, and 10°C. It is apparent that differences of 10°C must be reliably detected because they can result in large variations (66.3%) in the rate of substrate degradation. The difference of 1°C is suggested because it is the lowest temperature difference that can be reliably detected with a traditional data acquisition system (Hall et al., 1995). At this point in the development of methods in the field of composting research, an acceptable goal would be to pursue methods for obtaining power values above 0.80 for detection of 10°C temperature differences in 75% of the tests. Another goal would be to obtain power values above 0.80 to detect temperature differences of 5°C for 40% of the statistical tests.

Equations 6 and 7 are plotted in figure 7 as a function of the percent moisture content (w.b.) using dM_{wb} values of 1, 5, 10, and 15 percentage points. It is important to note that, while equation 4 was evaluated by Richard and Walker (1998) using the same substrate as in the mixing and inoculum studies, equation 6 was developed based on studies that used municipal solid wastes and biosolids (Haug, 1993). While the values of f_{H2O} predicted by equation 6 are probably not the same as would be predicted using SFW, the shape and functional dependence of the curve with respect to moisture content is most likely accurate for moisture contents below 60%. An additional problem with equation 6 is that maximum f_{H2O} is predicted when the moisture content of the substrate is 100%. This is not valid, as it has been observed that excessive moisture limits mass transfer of O_2 through the composting matrix and limits rates of degradation (Miller, 1989). However, the predicted f_{H2O} values for the actual moisture content range experienced by the substrate in the two previous studies, as denoted by the shaded box in figure 7, are within the boundaries of the values used by Haug (1993) to develop the model. Despite the limitations of this model in describing the functional relationship between moisture content and f_{H2O}, the model is reasonable for the illustrative purposes of a sensitivity analysis.

It is clear from figure 7 that small differences in moisture content result in large differences in the variation of f_{H2O}. Variation in moisture content has the smallest effect on the variation of the rate of CO_2 evolution at high moisture contents, while at lower moisture contents the variation in CO_2 evolution is much more pronounced. For purposes of discussion, if a composting reactor is considered inactive when the rate of CO_2 evolution is 20% of its optimal rate for the given temperature, then the moisture content range of interest can be further limited to those values between 30% and 55% (w.b.). Even within this limited range, the error in predicting the rate of CO_2 evolution is large compared to relatively small differences in moisture content. In order to limit the error in predicting f_{H2O} to 20% across all moisture contents, the difference between two moisture contents must be below 1 percentage point. Figure 4 demonstrated that the variation in moisture content increases as time progresses

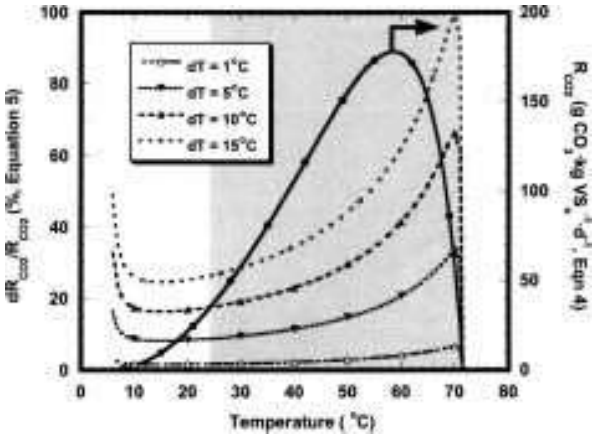

Figure 6. Graphical representation of equations 4 and 5 as functions of temperature and dT. Shaded region represents those temperatures experienced in the reactors used in the mixing (Schloss et al., 2000) and inoculum (Schloss and Walker, 2000) studies.

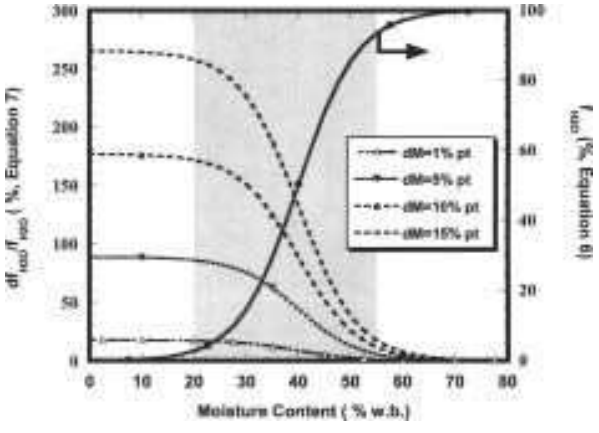

Figure 7. Graphical representation of equations 6 and 7 as functions of moisture content and dM. Shaded region represents those moisture contents experienced in the reactors used in the mixing (Schloss et al., 2000) and inoculum (Schloss and Walker, 2000) studies.

through the process. Because of this, variation has the largest effect on process performance when the bed is driest. This means that when the reactor bed is approaching the inactive phase, differences as small as 5 percentage points can cause errors in predicting the rate of CO_2 evolution by about 75%. In other words, if it is not possible to differentiate between two reactors that have moisture contents that differ by 5 percentage points, then one reactor may be inactive while the other is highly active.

An additional difficulty in the analysis of moisture content is the method of data acquisition. It has previously been demonstrated that in a mixture of dog food and wood chips, moisture is primarily accumulated in the wood chips (Baker et al., 1997). Therefore, when comparing multiple samples, it is necessary to ensure that each sample is representative of other samples and of the location in the reactor from which the sample was taken. This is impossible given the sensitivity the process shows to changes in moisture content. Because it is realistic to expect the sampling procedure to be accurate to within 2.5 percentage points, it is difficult to interpret the meaning of moisture content data in light of this sensitivity analysis. It is suggested that future power analyses attempt to detect differences of 2.5, 5, and 10 percentage points. Based on the limitation of the analytical methods, it is also suggested that new methods be developed that allow for precise *in situ* measurements of moisture content.

Effect on Experimental Design

From the data presented in this study, it is clear that even small levels of variation have a large effect on the ability to detect physically significant differences and the degradation activity throughout the process. Possible factors causing reduced power include too many treatments, limited replication, small number of observations, or the lack of control for salient variables. Each power–limiting factor will be evaluated in an attempt to suggest further avenues of research for improved experimental design. Particular attention will be paid to attempts to detect differences of 10°C with a power of 0.80 in 75% of the statistical tests.

Table 6 contains the results of several simulations that used MS_{exper} values reported by Schloss and Walker (2000, see their fig. 5) to determine the power to detect differences of 10°C. The data used were from ANOVAs attempting to detect differences across the bed of inoculated reactors. These data represent the least–variable experimental condition thus far reported in composting research.

This type of simulation is important because if researchers know what effect size is physically meaningful and the variability associated with the data, then they can implement an experimental design to obtain a specific power. For example, if one wanted to reliably detect 10°C differences with a power of 80%, then the shaded cases in table 6 would be acceptable. If only one reactor is available for each treatment, then the experiment grouping would need to be repeated at least ten times to have sufficient power in a two– or four–treatment study. If the number of reactors representing each treatment were increased to two, then the number of experimental groupings required would reduced. However, more than four groupings would be necessary.

Another situation to consider is when the number of experimental grouping replications is limited. If one is limited to two groupings, then it would be necessary to run 10 replicate reactors per treatment and experimental grouping. Again, as the number of reactors used decreases, the number of groupings required increases, and vice versa. In the mixing study, four treatments were evaluated. This analysis shows that the statistical tests in that study would have possessed more power to detect differences if only two treatments were addressed. In hindsight, it would have been

Table 6. Results of power analysis simulation using inoculated reactor data used to construct figure 5 of Schloss and Walker (2000). The shaded blocks represent cases in which at least 75% of the statistical tests obtained a power greater than 0.80 when attempting to detect differences of 10°C or less.

	Reactors[a]:	1			2			4			10		
Expt[b]	Treatments[c]:	2	4	6	2	4	6	2	4	6	2	4	6
2	Mean	0.36	0.28	0.23	0.53	0.46	0.41	0.71	0.67	0.64	0.89	0.89	0.87
	Median	0.29	0.17	0.10	0.50	0.39	0.30	0.75	0.72	0.67	0.97	0.98	0.99
	% > 0.80	8.6	8.6	5.7	20.0	20.0	17.1	48.6	48.6	48.6	80.0	80.0	77.1
3	Mean	0.63	0.48	0.40	0.81	0.70	0.63	0.92	0.88	0.83	0.97	0.96	0.96
	Median	0.63	0.41	0.26	0.91	0.80	0.69	1.00	0.99	0.98	>0.99	>0.99	>0.99
	% > 0.80	42.9	20.0	20.0	54.7	51.4	48.6	91.4	77.1	68.6	97.1	94.3	94.3
4	Mean	0.76	0.62	0.52	0.90	0.82	0.75	0.96	0.94	0.91	0.98	0.97	0.97
	Median	0.85	0.64	0.47	0.99	0.96	0.91	>0.99	>0.99	>0.99	>0.99	>0.99	>0.99
	% > 0.80	54.3	48.6	37.1	92.9	65.7	54.3	94.3	94.3	88.6	97.1	97.1	94.3
10	Mean	0.85	0.90	0.85	0.97	0.96	0.95	0.98	0.97	0.97	>0.99	0.99	0.98
	Median	>0.99	>0.99	>0.99	>0.99	>0.99	>0.99	>0.99	>0.99	>0.99	>0.99	>0.99	>0.99
	% > 0.80	94.3	82.9	71.4	97.1	94.3	94.3	97.1	97.1	97.1	100.0	97.1	97.1
15	Mean	0.97	0.95	0.93	0.98	0.97	0.96	0.99	0.98	0.97	>0.99	>0.99	0.99
	Median	>0.99	>0.99	>0.99	>0.99	>0.99	>0.99	>0.99	>0.99	>0.99	>0.99	>0.99	>0.99
	% > 0.80	94.3	94.3	91.4	97.1	97.1	94.3	97.1	97.1	97.1	100.0	100.0	97.1

[a] Number of reactors used in hypothetical experiment, assuming five observations taken from each reactor.
[b] Number of experimental groupings used in hypothetical experiment, assuming MS_{exper} does not vary.
[c] Number of treatments used in hypothetical experiment.

useful to evaluate the shortest possible interval between mixings (1 day) to unmixed reactors.

At some point, the benefit gained in obtaining greater power through increased replication is outweighed by several factors. First, in this analysis it was assumed that the MS_{exper} term would be the same regardless of the number of reactors and experimental groupings used. If it is necessary to perform 10 experimental groupings and each takes approximately three weeks to run from set–up to clean–up, then the study would take more than half a year. Given the large amount of variability found between replicates performed within two months, it is not reasonable to expect constant variability across six months. This is due to factors such as altered ambient air temperature or differences in the abundance of microbial populations from sources not being controlled. Second, the cost of labor and materials to repeat the same experimental grouping ten times would become overwhelming. Third, on a related subject, if it were necessary to use 10 reactors per treatment and experimental grouping, then limitations of labor and cost would again be introduced. For example, if 10 reactors were used in the mixing study for each treatment, then 30 reactors would be mixed on certain days. In order to make reasonable comparisons between treatments, the time of the mixing event would need to be approximately the same for every reactor. This would not be possible. Statistical, practical, and economic reasons currently limit the ability to design experiments with reasonable power. While the reactor system used in these experiments has been engineered to limit obvious sources of variability, additional work may be necessary to refine and standardize experimental techniques used by all researchers. Standardization of procedures will enable engineers to better cooperate to address concerns of variability and improve the process.

If the variability between experimental groupings could be reduced, then the difficulty of obtaining a robust experimental design would be reduced. Table 7 presents data from a simulation in which the MS_{exper} values used to create table 6 were reduced by 50%. Comparison of the two tables demonstrates the effect that the reduction in variability would have on the structure of the experimental design. For a future hypothetical study addressing the effects of two treatments on the process, the data in table 6 suggest that each treatment be replicated with four reactors and three experimental groupings, or with two reactors and four experimental groupings. However, the data in table 7 suggest using two treatments with two reactors and three experimental groupings. In addition to the practical reasons for reducing variability in compost data outlined at the beginning of this article, this analysis demonstrates the effect that methods of improving process reproducibility would have on experimental design in composting studies.

CONCLUSION

The sensitivity analysis represented in figures 6 and 7 demonstrates the previously described effect of variability on process performance. It was shown that a difference as small as 5°C between two reactors can account for a 33% difference in process performance. In addition, a difference of 1 percentage point in moisture content, which is smaller than the sensitivity of current methods, results in a difference of 15% at a moisture content of 30% (w.b.).

The analysis of experimental data in this study demonstrates that it is necessary to question the predictions of kinetic and process models. If it is not possible to detect differences of 15°C because of large variability in an experiment, then the accuracy of the parameters derived from

Table 7. Results of power analysis simulation using inoculated reactor data used to construct figure 5 of Schloss and Walker (2000) when the MS_{exper} used in the power analysis simulation is assumed to be 50% of the value reported. The shaded blocks represent cases in which at least 75% of the statistical tests obtained a power greater than 0.80 when attempting to detect differences of 10°C or less.

	Reactors[a]:	1			2			4			10		
Expt[b]	Treatments[c]:	2	4	6	2	4	6	2	4	6	2	4	6
2	Mean	0.53	0.43	0.41	0.71	0.67	0.635	0.85	0.85	0.83	0.95	0.95	0.95
	Median	0.50	0.34	0.30	0.75	0.72	0.67	0.94	0.96	0.96	>0.99	>0.99	>0.99
	% > 0.80	20.0	20.0	17.1	48.6	48.6	48.6	68.6	68.6	65.7	94.3	94.3	94.3
3	Mean	0.81	0.67	0.63	92.10	0.88	0.83	0.97	0.95	0.94	0.98	0.98	0.97
	Median	0.91	0.74	0.69	>0.99	0.99	0.98	>0.99	>0.99	>0.99	>0.99	>0.99	>0.99
	% > 0.80	65.7	48.6	48.6	91.4	77.1	68.6	94.3	94.3	94.3	97.1	97.1	97.1
4	Mean	0.90	0.78	0.75	0.96	0.94	0.91	0.98	0.97	0.96	0.99	0.98	0.98
	Median	0.99	0.94	0.91	>0.99	>0.99	>0.99	>0.99	>0.99	>0.99	>0.99	>0.99	>0.99
	% > 0.80	82.9	65.7	54.3	94.3	94.3	88.6	97.1	94.3	94.3	97.1	97.1	97.1
10	Mean	0.97	0.95	0.95	0.98	0.97	0.97	0.99	0.98	0.98	>0.99	>0.99	>0.99
	Median	>0.99	>0.99	>0.99	>0.99	>0.99	>0.99	>0.99	>0.99	>0.99	>0.99	>0.99	>0.99
	% > 0.80	97.1	94.3	94.3	97.1	97.1	97.1	97.1	100.0	100.0	100.0	100.0	100.0
15	Mean	0.98	0.97	0.96	0.99	0.98	0.97	>0.99	>0.99	0.99	>0.99	>0.99	>0.99
	Median	>0.99	>0.99	>0.99	>0.99	>0.99	>0.99	>0.99	>0.99	>0.99	>0.99	>0.99	>0.99
	% > 0.80	97.1	94.3	94.3	97.1	97.1	97.1	100.0	100.0	97.1	100.0	100.0	100.0

[a] Number of reactors used in hypothetical experiment, assuming five observations taken from each reactor.
[b] Number of experimental groupings used in hypothetical experiment, assuming MS_{exper} does not vary.
[c] Number of treatments used in hypothetical experiment.

those experiments for use in models predicting process behavior becomes dubious. For example, table 1 lists the parameter estimates and standard deviations for the parameters used in the CTMI model used by Richard and Walker (1998), and the error in these parameters varies from 8% to 20%. As shown in figure 6, a 7.1% temperature difference (equivalent to 5°C at 70°C) results in a 33.1% difference in the predicted rate of CO_2 evolution.

While it is not possible at this time to obtain consistently insightful information for statistical analysis with moisture content data, future research will focus on limiting the variability in temperature data. The variability in temperature data was reduced by inoculation, but a large amount of variability still exists. Because of the inoculum study, it is expected that other sources of microorganisms are serving as indirect sources of inoculation. These sources may include water, air, substrate, and reactor walls. By limiting these sources of variability, we hypothesize that it will be possible to further reduce experimental variability and to achieve the goals outlined earlier. While the variation in physical parameters, such as initial moisture content, air temperature, substrate density, and recipe mix, are also important in obtaining a reproducible process, variation of these variables will continue to be limited. In addition, experimental designs should be implemented that provide maximum statistical power with minimal resources.

Acknowledgement

This project was supported in part by the U.S. Department of Education under agreement number P200A8045.

References

Baker, C. S., J. S. VanderGheynst, and L. P. Walker. 1997. Equilibrium moisture isotherms for synthetic food waste and biosolids composts. *Compost Science and Util.* 7(1): 6–13.

Clark, C. S., C. O. Buckingham, D. H. Bone, and R H. Clark. 1977. Laboratory scale composting: Techniques. *J. Environ. Eng. Div. of ASCE* 103(EE5): 47–59.

Hall, S. G., D. Aneshansley, and L. P. Walker. 1995. Instrumentation in experimental compost modeling. In *1995 Annual International Summer Meeting*, 14. Chicago, Ill. St. Joseph, Mich.: ASAE.

Haug, R. T. 1993. *The Practical Handbook of Compost Engineering.* Boca Raton, Fla.: Lewis Publishers.

Jeris, J. S., and R. W. Regan. 1973. Controlling environmental parameters for optimum composting: II. Moisture, free air space, and recycle. *Compost Science* 14(2): 8–15.

Lindman, H. R. 1991. *Analysis of Variance in Experimental Design.* New York, N.Y.: Springer–Verlag.

Michel Jr., F. C., L. J. Forney, A. J.–F. Huang, S. Drew, M. Czuprenski, J. D. Linderberg, and C. A. Reddy. 1996. Effects of turning frequency, leaves to grass mix ratio, and windrow vs. pile configuration on composting of yard trimmings. *Compost Science and Util.* 4(1): 26–43.

Miller, F. C. 1989. Matric water potential as an ecological determinant in compost: A substrate dense system. *Microbial Ecology* 18(1): 59–71.

New York State Department of Environmental Conservation. 1996. Section 360–5.3. Operational requirements: sludge and solid waste. In Title 6 of the *Official Compilation of Codes*. Available at http://www.dec.state.ny.us/website/regs/360l.htm#360–5.3. Accessed on 9 July 2000.

Richard, T. L., and L. P. Walker. 1998. Temperature kinetics of aerobic solid–state degradation. In *Proceedings of the Institute of Biological Engineering* 1: A22–A39. Cortland, N.Y.: Institute of Biological Engineering.

Rosso, L., J. R. Lobry, S. Bajard, and J. P. Flandrois. 1995. Convenient model to describe the combined effects of temperature and pH on microbial growth. *Appl. and Environ. Microbiology* 61(2): 610–616.

Schloss, P. D., B. Chaves, and L. P. Walker. 2000. The use of the analysis of variance to assess the influence of mixing during composting. *Process Biochemistry* 35(7): 675–687.

Schloss, P. D., and L. P. Walker. 2000. Measurement of the process performance and variability in inoculated composting reactors using ANOVA and power analysis. *Process Biochemistry* 35(9): 931–942.

Schulze, K. L. 1961. Relationship between moisture content and activity of finished compost. *Compost Science* 2(1): 32–34.

Snell, J. R. 1957. Some engineering aspects of high rate composting. *J. Sanitary Eng. Div., Proceedings of ASCE* 83(SA1): 1178.1–1178.36.

Sokal, R. R., and F. J. Rohlf. 1995. *Biometry.* 3rd ed. New York, N.Y.: W.H. Freeman and Co.

VanderGheynst, J. S., J. M. Gossett, and L. P. Walker. 1997. High–solids aerobic decomposition: Pilot–scale reactor development and experimentation. *Process Biochemistry* 32(5): 361–375.

Walker, L. P., T. D. Nock, J. M. Gossett, and J. S. VanderGheynst. 1999. Managing moisture limitations on microbial activity in high–solids aerobic decomposition: Pilot–scale experimentation. *Process Biochemistry* 34(6–7): 601–612.