

Kansas State University Libraries

New Prairie Press

Conference on Applied Statistics in Agriculture

1999 - 11th Annual Conference Proceedings

CHARACTERIZING THE STATISTICAL DISTRIBUTION OF ORGANIC CARBON AND EXTRACTABLE PHOSPHORUS AT A REGIONAL SCALE

John J. Brejda

David W. Meek

Douglas L. Karlen

Follow this and additional works at: <https://newprairiepress.org/agstatconference>



Part of the [Agriculture Commons](#), and the [Applied Statistics Commons](#)



This work is licensed under a [Creative Commons Attribution-Noncommercial-No Derivative Works 4.0 License](#).

Recommended Citation

Brejda, John J.; Meek, David W.; and Karlen, Douglas L. (1999). "CHARACTERIZING THE STATISTICAL DISTRIBUTION OF ORGANIC CARBON AND EXTRACTABLE PHOSPHORUS AT A REGIONAL SCALE," *Conference on Applied Statistics in Agriculture*. <https://doi.org/10.4148/2475-7772.1261>

This is brought to you for free and open access by the Conferences at New Prairie Press. It has been accepted for inclusion in Conference on Applied Statistics in Agriculture by an authorized administrator of New Prairie Press. For more information, please contact cads@k-state.edu.

**CHARACTERIZING THE STATISTICAL DISTRIBUTION OF ORGANIC CARBON
AND EXTRACTABLE PHOSPHORUS AT A REGIONAL SCALE**

John J. Brejda, David W. Meek, and Douglas L. Karlen

J. J. Brejda, USDA-ARS Wheat, Sorghum and Forage Research Unit, 344 Keim Hall, University of Nebraska, Lincoln, NE, 68583, D. W. Meek and D. L. Karlen, National Soil Tilth Lab, 2150 Pammel Drive, Ames, IA, 50011.

ABSTRACT

Greater awareness of potential environmental problems has created the need to monitor total organic carbon (TOC) and extractable phosphorus (P) concentrations at a regional scale. The probability distribution of these soil properties can have a significant effect on the power of statistical tests and the quality of inferences applied to these properties. The objectives of this study were to: (1) evaluate the probability distribution of TOC and extractable P at the regional scale in three Major Land Resource Areas (MLRA), and (2) identify appropriate transformations that will result in a normal distribution. Both TOC and extractable P were non-normally distributed in all three MLRAs. Suggested power transformations did not result in normality, but a natural log and negative binomial transformation did produce distributions that met the assumptions of normality in most cases. Statistical analysis of TOC and extractable P data at the regional scale will need to take into account the non-normal distribution of these properties for accurate and precise estimates.

INTRODUCTION

Soil conservation and agricultural resource condition on private lands in the United States are monitored by the USDA-Natural Resources Conservation Service (NRCS) using the National Resource Inventory (NRI) (Kellogg et al., 1994). However, with the NRI no soil samples are collected and in most cases NRI sample points are not visited. Rather, NRCS uses data from the NRI to calculate an estimated rate of soil erosion using the Universal Soil Loss Equation to determine whether the Nation's soils are improving, stable, or degrading (US Congress, Office of Technology Assessment, 1995).

Several current and emerging environmental problems have created the need for NRCS to consider ways to improve their ability to monitor soil and agricultural resource condition using the NRI. One possibility is to visit NRI sample points and collect soil samples which would be analyzed for appropriate

properties. Because of the diversity of soils across the US such an assessment may be most feasible if conducted on a regional basis, provided each region contains similar soil and land-use patterns. Major Land Resource Areas (MLRA) are geographic units of several thousand hectares in extent that contain similar patterns of soils, climate, water resources, and land uses (USDA-SCS, 1981). They are important in agricultural planning at the state, regional, and national levels (USDA-SCS, 1981). Thus, the MLRA offers an appropriate regional scale unit for monitoring soil properties as part of environmental and ecological assessments.

Two soil properties of agronomic and environmental importance are total organic carbon (TOC) and extractable phosphorus (P). Total organic C affects several soil properties important for crop growth and environmental quality (Larson and Pierce, 1991; Seybold et al., 1997). In addition, increasing TOC concentrations in agricultural soils has been proposed as a mechanism to sequester CO₂ from the atmosphere to reduce atmospheric greenhouse gas concentrations (Bruce et al., 1999; Lal et al., 1999).

Phosphorus is an essential nutrient for plant growth and is often applied to crop fields to ensure levels are adequate for crop production. However, enrichment of P resulting from agricultural, urban, and industrial activities can seriously degrade water quality and aquatic ecosystems (Carpenter et al., 1998; Correll, 1998). Nonpoint sources of P are hard to monitor and regulate because they are dispersed over wide geographic areas. Thus, NRCS may be interested in monitoring the concentrations and distribution of TOC and extractable P in soils on a regional scale using the NRI.

The probability distribution of soil properties can have a significant effect on the power of statistical techniques applied to these properties. Most classical univariate and multivariate statistical analyses require the data to be normally distributed and have equal variances (Sokal and Rohlf, 1981; Johnson and Wichern, 1992; Sharma, 1996). Violation of these assumptions can reduce the power of the statistical tests and lead to incorrect conclusions (Sharma, 1996). Further, the quality of inferences made using these methods depends on how closely the true parent population resembles a normal distribution (Johnson and Wichern, 1992).

At the plot and field scale, the probability distribution exhibited by many soil properties is often non-normally distributed (Edmonds and Lentner, 1987; Edmonds et al., 1988; Parkin et al., 1988; Parkin and Robinson, 1992; 1994; Starr et al., 1992; 1995). In Iowa, Cambardella et al. (1994) reported that in one 36 ha field, 25 of 27 soil properties were non-normally distributed, and in another 96 ha field, 4 of 14 soil properties were non-normally distributed. In Maryland, Starr et al. (1995) reported that denitrification rates and NO₃-N and orthophosphate P concentrations were log-normally distributed.

The distribution and variability of soil properties are also scale dependent (Beckett and Webster, 1971; Starr et al., 1992; Parkin, 1993; Seyfried and Wilcox, 1995). At the regional scale, smaller scale sources of variability such as landform, landscape position, and land-use may become subsumed into the larger scale variability (Beckett and Webster, 1971; Hillel, 1991; Seyfried and Wilcox, 1995). Thus, the probability distribution of soil properties observed at the plot and field scales may not apply at the regional scale. The objectives of this study were to: (1) evaluate the probability distribution of TOC and extractable P at the regional scale in three different geographic regions, and (2) identify appropriate transformations that will meet the assumptions of a normal distribution.

MATERIALS AND METHODS

Three MLRAs, designated 105, 9, and 67, were selected to evaluate the probability distribution of soil properties at the regional scale. MLRA 105 comprises the Northern Mississippi Valley Loess Hills, covers 27,090 km², and is located in southwestern Wisconsin, northwestern Illinois, northeastern Iowa, and southeastern Minnesota (USDA-SCS, 1981). Elevation ranges from 200 m on river valley floors to 400 m on the highest ridges. Average annual precipitation ranges from 750 to 900 mm, two-thirds or more of which falls during the growing season. Average annual temperature ranges from 7^o to 10^oC. About 40% of the agricultural land is cultivated for row crop and small grain production, 20% is in permanent pasture and hayland, and the remainder is forested.

MLRA 9 comprises the Palouse and Nez Perce Prairies, covers 23,140 km², and is located in southeastern Washington, northwestern Idaho, and northeastern Oregon (USDA-SCS, 1981). Elevation ranges from 600 to 1200 m. Average annual precipitation ranges from 375 to 625 mm and is evenly distributed throughout the fall, winter, and spring, with the summers being relatively dry. Average annual temperature ranges from 7^o to 12^oC. About 50% of the agricultural land is cropland, 40% is native rangeland, and the remainder is used for vegetable production.

MLRA 67 comprises the Central High Plains, covers 74,410 km², and is located in eastern Colorado, southeastern Wyoming, and western Nebraska (USDA-SCS, 1981). Elevation ranges from 1100 to 1800 m, increasing gradually from east to west. Average annual precipitation ranges from 325 to 425 mm with maximum precipitation falling in late spring and early autumn. Average annual temperature ranges from 7^o to 10^oC. About 25% of the land is farmed for small grain production and 60% is native rangeland comprised of mid and short grass species. The remainder of the area is irrigated and used to grow row crops and vegetables.

Experimental Design and Sample Analysis

In MLRA 105 and 9 a statistically representative sample of 200 points were selected using the NRI sampling framework. The design of the NRI is a stratified two-stage area sample with primary sampling units (PSU) as the first stage and sample points within PSU as the second stage. Detailed descriptions of sample point selection within the NRI framework are presented elsewhere (Goebel and Baker, 1982; Nusser and Goebel, 1997; Nusser et al., 1998). Some points were inaccessible or fell on homesteads, urban areas, road pavement, or rock outcrops. These points were not sampled. As a result, only 186 points were actually sampled in the MLRA 105, and 149 points were sampled in MLRA 9. Sample points were selected without regard to soil series or land-use. In MLRA 67 a statistically representative sample of 100 points was selected using the NRI sampling framework, with restriction that sampling was confined to a single soil series. If the soil present at the sampling site was not the designated series the point was not sampled. As a result, only 64 points were sampled in MLRA 67.

At each point a 1000 cm³ soil sample was collected from the 0-10 cm depth. Each sample was analyzed for TOC by dry combustion using a Leco SC-444 analyzer (Leco Corp., St. Joseph, MI) and extractable P using the procedure of Mehlich (1984) measured by inductively-coupled plasma emission spectroscopy.

Statistical Analysis

A U-test (e.g., Cressie, 1993) was used as a first step in evaluating non-normality in the data. This test standardizes the mean-median difference. Values of $|U| \geq 3$ indicate possible non-normality or the presence of outliers. Normality of the data was tested using the D'Agostino-Pearson K^2 test. This test is an omnibus χ^2 test for detecting deviation from normality caused by either skewness or kurtosis, and has been shown to have good power over a broad range of non-normal distributions (D'Agostino et al., 1990). Skewness was tested using the third sample moment test ($\sqrt{b_1}$) and kurtosis was tested using the fourth sample moment test (b_2). The mean, median, and coefficient of variation for each soil attribute was calculated using PROC UNIVARIATE in SAS, and the K^2 , $\sqrt{b_1}$ and b_2 tests were calculated using a SAS macro provided by D'Agostino et al. (1990). Normality was tested at the $\alpha=0.05$ probability level.

If either of the soil properties were non-normally distributed in any of the MLRAs a power transformation was selected using the procedure of Emerson and Stoto (1983). This procedure calculates the slope of a line between the midspread and centered midrange where the midspread (horizontal axis) is calculated as:

$$\frac{(\underline{X}_U - M)^2 + (M - \underline{X}_L)^2}{4M} \quad [1]$$

and centered midrange (vertical axis) is calculated as:

$$\frac{\underline{X}_U + \underline{X}_L}{2} - M \quad [2]$$

where \underline{X}_U and \underline{X}_L are upper and lower midsummaries and M is the sample median. One minus the slope is the indicated power of transformation. However, this procedure calculates an empirical power transformation. Therefore, we used a set of rules to select the actual transformation evaluated (Table 1). In addition to the power transformations of $X^{0.75}$, $X^{0.5}$, $X^{0.25}$, and natural log ($\ln X$), we examined two other transformation using $X^{-0.25}$ and a negative binomial-normalizing transformation (NB) (Beall, 1942) where:

$$NB = \tau^{-0.5} [\sinh^{-1}((X + 0.5)\tau)^{0.5}] \quad [3]$$

$$\tau = (\sigma^2 - \mu) / \mu^2 \quad [4]$$

The $X^{-0.25}$ transformation was selected for evaluation after examining box plots for each soil variable in each MLRA using the procedure of Emerson and Strenio (1983). The NB transformation was evaluated because the $X^{0.5}$ transformation was recommended in several cases and the $X^{0.5}$ transformation is the limiting case of the negative binomial. Following transformation the data were again tested for normality using the procedure of D'Agostino et al. (1990).

RESULTS AND DISCUSSION

Neither TOC or extractable P were normally distributed in any of the MLRAs (Table 2). U-test values for TOC ranged between 3.5 and 7.7, and for extractable P they ranged between 4.9 and 11.2. MLRA 9 had the highest U-test values for both TOC and extractable P. Both soil properties exhibited significant kurtosis and skewness to the right ($\sqrt{b_1} > 0$). Right-tailed skewness resulted from the presence of a few exceptionally large values for each soil property, lengthening the right-tail. We considered all observations as valid and therefore did not consider it appropriate to discard any of the exceptionally large values in the right-tail as outliers. Right-tailed skewness resulted in the mean being larger than the median (Table 2) and the large values in the right-tail tended to inflate the variance.

The power transformation procedure of Emerson and Stoto (1983) resulted in the selection of transformations of $X^{0.75}$ for TOC in MLRA 105 and extractable P in MLRA 105 and 9, and $X^{0.5}$ for TOC in MLRA 9 and 67 and extractable P in MLRA 67 (Table 3). However, these transformations resulted in a distribution that met the assumption of normality only for TOC in MLRA 67 (Table 2). Concentrations of TOC in MLRA 105 and 9 and extractable P in all three MLRAs required stronger transformations than suggested using the procedure of Emerson and Stoto (1983) to achieve

normality. Although no single transformation worked for both soil properties in all three MLRA, two transformations, log and NB gave the most consistent results. These two transformations resulted in distributions that met the assumption of normality for TOC in MLRA 9 and 67 and for extractable P in all three MLRAs (Table 2). For TOC in MLRA 105, only the $X^{-0.25}$ transformation resulted in a distribution that met the assumption of normality.

The transformation procedure of Emerson and Stoto (1983) is actually a transformation for symmetry. However, it usually also results in normality and the elimination of outliers. Reasons for the suggested transformation not resulting in an adequate normalizing transformation may be: (1) The slope of the line between the points on the horizontal and vertical axes (Eq. 1 and 2) is a first term approximation of a series and is valid only when the relationship is approximately linear. In most cases a departure from linearity in the transformation plot was visually apparent (Fig. 1); (2) For all the suggested transformations symmetry was at best poorly achieved based on visual inspection.

SUMMARY

Results from this study indicate that TOC and extractable P are non-normally distributed at the regional scale. This is consistent with the observed distribution of many soil properties at the plot and field scale (Edmonds and Lentner, 1987; Edmonds et al., 1988; Parkin et al., 1988; Parkin and Robinson, 1992; 1994; Starr et al., 1992; 1995). Both log and NB transformations resulted in distributions that reasonably met the assumption of normality in five out of six cases. In terms of data analysis, however, applying and taking the inverse of the log is much simpler than the NB transformation and would probably be the preferred transformation by most soil scientists. It is recognized that the log-normal distribution is a good approximation for a wide range of probability distributions (Patel and Read, 1996). Researchers willing to do the analyses presented in this paper may sometimes find a better transformation than the log (as with the TOC results for MLRA 105). A next step might be to estimate the nature of the distribution for each of these soil properties using maximum likelihood methods or procedures given in Patel and Read (1996). However, such techniques may be beyond the capability and interest of most soil scientists and therefore are not likely to be used. The best option is to ensure the sampling framework used has sufficient sample size and geographic coverage to adequately represent the heterogeneity within each MLRA so that accurate and precise estimates of the soil properties can be made at a regional scale.

LITERATURE CITED

- Beall, G. 1942. The transformation from entomological field experiments so that the analysis of variance becomes applicable. *Biometrika* 32:243-262.
- Beckett, P.H.T., and R. Webster. 1971. Soil variability: a review. *Soils and Fert.* 34:1-15.
- Bruce, J.P., M. Frome, E. Haites, H. Janzen, R. Lal, and K. Paustian. 1999. Carbon sequestration in soils. *J. Soil Water Conserv.* 54:382-389.
- Cambardella, C.A., T.B. Moorman, J.M. Novak, T.B. Parkin, D.L. Karlen, R.F. Turco, and A.E. Konopka. 1994. Field-scale variability of soil properties in central Iowa soils. *Soil Sci. Soc. Am. J.* 58:1501-1511.
- Carpenter, S.R., N.F. Caraco, D.L. Correll, R.W. Howarth, A.N. Sharpley, and V.H. Smith. 1998. Nonpoint pollution of surface waters with phosphorus and nitrogen. *Ecol. Applic.* 8:559-568.
- Correll, D.L. 1998. The role of phosphorus in the eutrophication of receiving waters: a review. *J. Environ. Qual.* 27:261-266.
- Cressie, N.A.C. 1993. *Statistics for spatial data.* John Wiley and Sons, Inc., NY.
- D'Agostino, R.B., A. Belanger, and R.B. D'Agostino Jr. 1990. A suggestion for using powerful and informative tests of normality. *Am. Stat.* 44:316-321.
- Edmonds, W.J., and M. Lentner. 1987. Soil series differentiae selected by discriminant analysis based on ranks. *Soil Sci. Soc. Am. J.* 51:716-721.
- Edmonds, W.J., D.D. Rector, N.O. Wilson, and T.L. Arnold. 1988. Evaluation of relationships between oak site indices and properties of selected dystrochrepts. *Soil Sci. Soc. Am. J.* 52:204-209.
- Emerson, J.D., and M.A. Stoto. 1983. Transforming data, p. 97-128. In D.C. Hoaglin et al., (ed.) *Understanding Robust and Exploratory Data Analysis.* John Wiley and Sons, Inc., NY.
- Emerson, J.D., and J. Strenio. 1983. Boxplots and batch comparisons, p. 58-96. In D.C. Hoaglin et al., (ed.) *Understanding Robust and Exploratory Data Analysis.* John Wiley and Sons, Inc., NY.
- Goebel, J.J., and H.D. Baker. 1982. The 1982 National Resources Inventory sample design and estimation procedures. Statistical Laboratory, Iowa State University, Ames, IA.
- Hillel, D. 1991. Research in soil physics: A re-review. *Soil Sci.* 151:30-34.
- Johnson, R.A., and D.W. Wichern. 1992. *Applied multivariate statistical analysis.* Third Ed. Prentice Hall, Inc., Upper Saddle River, NJ.
- Kellogg, R.L., G.W. TeSelle, and J.J. Goebel. 1994. Highlights from the 1992 National Resources Inventory. *J. Soil Water Conserv.* 49:521-527.

- Lal, R., R.F. Follett, J. Kimble, and C.V. Cole. 1999. Managing U.S. cropland to sequester carbon in soil. *J. Soil Water Conserv.* 54:374-381.
- Larson, W.E., and F.J. Pierce. 1991. Conservation and enhancement of soil quality. p. 175-203. *In* Evaluation for Sustainable Land Management in the Developing World. Vol. 2. IBSRAM Proc. 12 (2). Bangkok, Thailand. Int. Board Soil Res. Manage.
- Mehlich, A. 1984. Mehlich 3 soil test extractant: A modification of Mehlich 2 extractant. *Comm. Soil Sci. Plant Anal.* 15:1409-1416.
- Nusser, S.M., and J.J. Goebel. 1997. The National Resource Inventory: A long-term multi-resource monitoring programme. *Environ. Ecol. Stat.* 4:181-204.
- Nusser, S.M., F.J. Breidt, and W.A. Fuller. 1998. Design and estimation for investigating the dynamics of natural resources. *Ecol. Appl.* 8:234-245.
- Parkin, T.B. 1993. Spatial variability of microbial processes in soil - a review. *J. Environ. Qual.* 22:409-417.
- Parkin, T.B., and J.A. Robinson. 1992. Analysis of lognormal data. *Adv. Soil Sci.* 20:193-235.
- Parkin, T.B., and J.A. Robinson. 1994. Statistical treatment of microbial data. p. 15-39. *In* R.W. Weaver et al. (ed.) *Methods of soil analysis. Part 2. Microbial and biochemical properties.* SSSA Book Series 5. Madison, WI.
- Parkin, T.B., J.J. Meisinger, S.T. Chester, J.L. Starr, and J.A. Robinson. 1988. Evaluation of statistical methods for lognormally distributed variables. *Soil Sci. Soc. Am. J.* 52:323-329.
- Patel, J.K., and C.B. Read. 1996. *Handbook of the normal distribution.* 2nd Ed. Marcel Dekker, Inc., NY.
- Seybold, C.A., M.J. Mausbach, D.L. Karlen, and H.H. Rogers. 1997. Quantification of soil quality. p. 387-404. *In* R. Lal et al., (eds.) *Soil Processes and the Carbon Cycle.* CRC Press, Boca Raton, FL.
- Seyfried, M.S., and B.P. Wilcox. 1995. Scale and the nature of spatial variability: field examples having implications for hydrologic models. *Water Resources Res.* 31:173-184.
- Sharma, S. 1996. *Applied multivariate techniques.* John Wiley and Sons, Inc., New York.
- Sokal, R.R., and F.J. Rohlf. 1981. *Biometry.* Second Ed. W.H. Freeman and Co., San Francisco, CA.
- Starr, J.L., T.B. Parkin, and J.J. Meisinger. 1992. Sample size considerations in the determination of soil nitrate. *Soil Sci. Soc. Am. J.* 56:1824-1830.
- Starr, J.L., T.B. Parkin, and J.J. Meisinger. 1995. Influence of sample size on chemical and physical soil measurements. *Soil Sci. Soc. Am. J.* 59:713-719.

US Congress, Office of Technology Assessment, 1995. Targeting environmental priorities in agriculture: reforming program strategies. Washington, D.C. Gov. Printing Off. OTA-ENV-640. p. 27.

USDA, Soil Conservation Service. 1981. Land Resource Regions and Major Land Resource Areas of the United States. USDA-SCS, Agric. Handb. 296, U.S. Gov. Print. Office, Washington, D.C.

Table 1. Transformation rules used to select the transformation to apply to non-normally distributed total organic carbon and extractable phosphorus data.

Empirical power estimate (p)	Transformation evaluated
$p \leq 0.125$	$\ln(x)$
$0.125 < p \leq 0.375$	$x^{0.25}$
$0.375 < p \leq 0.625$	$x^{0.5}$
$0.625 < p \leq 0.875$	$x^{0.75}$
$0.875 < p \leq 1.125$	none
$1.125 < p \leq 1.375$	$x^{-0.25}$

Table 2. Descriptive and D'Agostino statistics for original and transformed total organic C and extractable P from three Major Land Resource Areas.

Trans-formation	Mean	Median	U-test	CV	D'Agostino Statistics				
					K^2 †	P>F	$\sqrt{b_1}$ ‡	P>F	b_2 §
Total organic C (g kg⁻¹)									
MLRA 105 (Northern Mississippi Valley Loess Hills)									
None	27.6	22.7	3.5	65	132.2 (<0.001)	2.7 (<0.001)	13.4 (<0.001)		
X ^{0.75} ¶	11.7	10.4	6.0	46	97.4 (<0.001)	2.1 (<0.001)	9.3 (<0.001)		
X ^{0.5}	5.1	4.8	4.4	29	62.5 (<0.001)	1.5 (<0.001)	6.4 (<0.001)		
X ^{0.25}	2.2	2.2	3.0	14	31.1 (<0.001)	0.9 (<0.001)	4.5 (0.003)		
Ln X	3.2	3.1	1.7	17	8.8 (0.012)	0.5 (0.011)	3.6 (0.11)		
NB	4.7	4.6	1.9	17	10.6 (0.005)	0.5 (0.005)	3.6 (0.102)		
X ^{-0.25}	0.5	0.5	-0.5	13	1.1 (0.59)	0.0 (0.99)	3.3 (0.302)		
MLRA 9 (Palouse and Nez Perce Prairies)									
None	25.0	22.4	7.7	48	31.4 (<0.001)	1.1 (<0.001)	4.5 (0.006)		
X ^{0.75}	11.0	10.3	2.7	36	16.8 (<0.001)	0.8 (<0.001)	3.6 (0.11)		
Table 2. Continued.									
X ^{0.5} ¶	4.9	4.7	1.8	24	6.4 (0.04)	0.5 (0.012)	3.1 (0.68)		
X ^{0.25}	2.2	2.2	1.0	12	1.2 (0.54)	0.2 (0.30)	2.8 (0.70)		
Ln X	3.1	3.1	0.1	15	0.5 (0.80)	-0.1 (0.57)	2.8 (0.72)		
NB#	5.0	5.0	0.4	20	0.5 (0.79)	0.0 (0.96)	2.7 (0.49)		
X ^{-0.25}	0.5	0.5	0.8	12	5.0 (0.082)	0.4 (.03)	3.1 (0.60)		
MLRA 67 (Central High Plains)									
None	10.5	8.8	3.0	59	11.1 (0.004)	1.1 (0.001)	3.2 (0.48)		
X ^{0.75}	5.7	5.1	2.3	44	7.1 (0.028)	0.8 (0.008)	2.8 (0.94)		
X ^{0.5} ¶	3.1	3.0	1.6	29	3.8 (0.15)	0.6 (0.056)	2.6 (0.66)		
X ^{0.25}	1.7	1.7	0.9	15	0.9 (0.63)	0.2 (0.39)	2.6 (0.67)		
Ln X	2.2	2.2	0.2	27	0.5 (0.77)	-0.2 (0.59)	3.0 (0.64)		
NB	3.3	3.2	0.9	30	1.6 (0.45)	0.2 (0.39)	2.4 (0.36)		
X ^{-0.25}	0.6	0.6	0.6	15	8.5 (0.015)	0.7 (0.028)	4.2 (0.056)		
Extractable P (g kg⁻¹)									
MLRA 105 (Northern Mississippi Valley Loess Hills)									
None	53.5	34.8	7.3	102	120.4 (<0.001)	2.6 (<0.001)	10.7 (<0.001)		

Table 2. Continued.

$X^{0.75}$ ¶	18.4	14.3	8.1	72	86.6 (<0.001)	2.0 (<0.001)	7.5 (<0.001)
$X^{0.5}$ ¶	6.7	5.9	5.6	46	49.3 (<0.001)	1.3 (<0.001)	5.1 (<0.001)
$X^{0.25}$	2.5	2.4	3.3	22	15.9 (<0.001)	0.7 (<0.001)	3.5 (0.145)
Ln X	3.6	3.6	1.2	24	0.3 (0.88)	0.1 (0.62)	2.9 (0.99)
NB	4.3	4.2	1.3	20	0.6 (0.74)	0.1 (0.44)	2.9 (0.93)
$X^{-0.25}$	0.4	0.4	0.9	22	9.6 (0.008)	0.5 (0.004)	3.3 (0.29)
MLRA 9 (Palouse and Nez Perce Prairies)							
None	46.4	35.6	11.2	93	159.9 (<0.001)	3.9 (<0.001)	26.6 (<0.001)
$X^{0.75}$ ¶	16.9	14.6	5.0	63	112.5 (<0.001)	2.7 (<0.001)	15.0 (<0.001)
$X^{0.5}$ ¶	6.3	6.0	3.1	39	62.8 (<0.001)	1.6 (<0.001)	8.0 (<0.001)
$X^{0.25}$	2.5	2.4	1.3	19	20.1 (<0.001)	0.7 (<0.001)	4.5 (0.005)
Ln X	3.6	3.6	-0.3	21	1.2 (0.55)	0.1 (0.26)	3.3 (0.29)
NB	4.0	4.1	-0.2	19	1.3 (0.52)	0.1 (0.54)	3.3 (0.32)
$X^{-0.25}$	0.4	0.4	1.9	19	10.5 (0.005)	0.6 (0.005)	3.7 (0.102)
MLRA 67 (Central High Plains)							
None	39.6	31.8	4.9	65	31.4 (<0.001)	1.7 (<0.001)	6.1 (0.002)

Table 2. Continued.

$X^{0.75}$	15.3	13.4	3.8	47	21.5 (<0.001)	1.4 (<0.001)	4.8 (0.016)
$X^{0.5}$ ¶	6.0	5.6	2.8	31	12.2 (0.002)	1.0 (0.0017)	3.8 (0.12)
$X^{0.25}$	2.4	2.4	1.9	15	5.0 (0.083)	0.6 (0.033)	3.2 (0.52)
Ln X	3.5	3.5	1.0	17	0.9 (0.64)	0.3 (0.34)	2.8 (0.94)
NB	5.2	5.1	1.1	17	1.1 (0.57)	0.3 (0.29)	2.8 (0.95)
$X^{-0.25}$	0.4	0.4	-0.1	14	0.2 (0.91)	0.1 (0.74)	2.7 (0.80)

†The K^2 statistic is approximately chi-square distributed with 2 df when the population is normally distributed.

‡ $\sqrt{b_1} = 0$ when the population is normally distributed. $\sqrt{b_1} > 0$ indicates skewness to the right and $\sqrt{b_1} < 0$ indicates skewness to the left.

§ $b_2 = 3$ when the population is normally distributed. $b_2 > 3$ indicates a leptokurtic (peaked) distribution and $b_2 < 3$ indicates a platykurtic (flat) distribution.

¶Recommended power transformation.

#NB is a negative binomial normalizing transformation.

Table 3. Estimated and actual power transformations evaluated using the transformation rules in Table 1.

MLRA†	Total organic carbon		Extractable phosphorus	
	Estimated power	Actual transformation	Estimated power	Actual transformation
105	0.65	$X^{0.75}$	0.77	$X^{0.75}$
9	0.43	$X^{0.50}$	0.81	$X^{0.75}$
67	0.49	$X^{0.50}$	0.53	$X^{0.50}$

†MLRA 105 = Northern Mississippi Valley Loess Hills; MLRA 9 = Palouse and Nez Perce Prairies; MLRA 67 = Central High Plains.

Figure 1. Transformation plot for symmetry using the Emerson and Stoto procedure on soil total organic carbon concentrations in MLRA 105. Note the significant deviation from linearity which may have caused the procedure to suggest a transformation that failed to produce a normal distribution.

