PROBABILITY FUNCTIONS FOR SOME SOUTH DAKOTA SOIL PROPERTIES

Darrell W. DeBoer
Agricultural and Biosystems Engineering Department
South Dakota State University
Brookings, SD 57007

ABSTRACT

Raw data sets stored in file cabinets or other obscure locations have limited utility even though they may have been formulated at considerable effort and expense. Large data sets of chemical and physical property values for soils in the proposed Lake Andes-Wagner and Oahe irrigation projects in South Dakota have been compiled during previous field investigations. Results of this study identify probability density functions (PDFs) that can be used to mathematically describe the frequency and magnitude of soil chemical property and saturated hydraulic conductivity values associated with the project areas.

Keywords

Electrical conductivity, hydraulic conductivity, Lake Andes-Wagner, Oahe, pH, probability, salinity, selenium, sodium

INTRODUCTION

Large data sets of chemical and physical property values for several soils have been compiled during previous field investigations related to two proposed irrigation projects in South Dakota. The Lake Andes-Wagner Project was proposed for Charles Mix County in south-central South Dakota and the Oahe Project for Brown County and Spink County of north-central South Dakota. Much effort and many dollars were invested in the collection and compilation of the data sets. The purpose of this paper is to summarize the cumbersome data sets in the form of probability density functions so interested parties can use them for future analyses.

METHODS

Lake Andes-Wagner Project Area

Scientists of the SD Agricultural Experiment Station conducted field studies of Lake Andes-Wagner Project area soils in Charles Mix County during 1993 (DeBoer et al., 1995). Soil samples were collected along two orthogonal transects

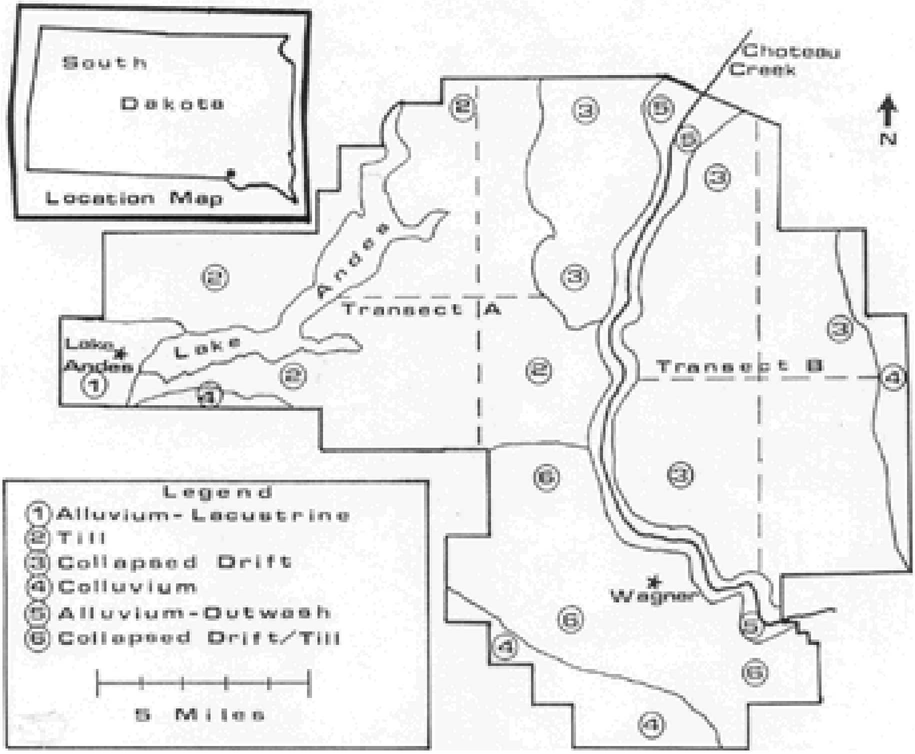

Figure 1. Location of geological materials within the boundaries of the Lake Andes-Wagner Project area and the two soil sample transects used in the field investigations.

as illustrated in Figure 1. Each orthogonal transect consisted of east-west and north-south components. One orthogonal transect (Transect A) was positioned in the till area west of Choteau Creek and a second (Transect B) was located east of Choteau Creek in collapsed drift geologic material. Soil samples were collected at sites about 200 m, (every 1/8 mile), along the transects. Transect A contained 72 sample sites (14.3 km, 8 7/8 miles) in the north-south direction and 59 sample sites (11.8 km, 7 1/4 miles) in the east-west direction, while Transect B had 83 sites (16.7 km, 10 3/8 miles) in the north-south direction and 50 sites (9.9 km, 6 1/8 miles,) in the east-west direction. Each transect contained about 130 sample sites.

Four soil depths at each sample site were used in the investigation: 0.0-0.5 m, 0.5-1.0 m, 1.0-2.0 m and 2.0-3.0 m (0.0-1.6 ft, 1.6-3.3 ft, 3.3-6.6 ft and 6.6-9.9 ft). Soil samples were collected with a Giddings probe and a 76-mm (3 in) diameter core tube to a depth of 1.0 m. A 51-mm (2 in) diameter core tube was used to the 2.0-m depth and a 41-mm (1.6 in) diameter core tube down to the 3.0-m depth. A depth-weighted value was determined for each chemical variable at every sample site and is designated as a composite profile (comp) value.

Selenium (Se) concentration analyses were grouped into three operationally defined fractions (available (F1), conditionally available (F2) and unavailable selenium) (Chao and Sanzolene, 1989). Available Se and conditionally available

Se were extracted from the same soil sample. Available Se was defined as 0.1 M KH_2PO_4 extractable. This includes the H_2O-soluble and nonspecifically adsorbed selenate and the exchangeable, specifically adsorbed selenite. Conditionally available Se was defined as 4M HCL-extractable, which is the Se associated with oxide minerals (Fe, Mn and Al), amorphous minerals, carbonates, acid-volatile sulfides and acid-hydrolyzable organic matter. The Se in this fraction is not readily available; however, it has the potential to become available with changes in pH or redox. Total Se was determined from a separate sample using a mixed acid digestion (Briggs and Crock, 1986). Unavailable Se was calculated as the difference between total Se and the sum of the available and conditionally available Se.

Electrical conductivity (EC) values were determined by a method presented by Rhoades (1982). Likewise, pH and sodium adsorption ratio (SAR) values were determined by standard laboratory methods given by Rhoades (1982).

Saturated hydraulic conductivity data were collected by the Bureau of Reclamation (1985) for the project area. Pump-in, pump-out and piezometer techniques were used to measure *in situ* conductivity values for the soils (Bureau of Reclamation, 1978). Depth weighted conductivity values were reported by the Bureau and are used in this study.

Oahe Project Area

Bureau of Reclamation employees conducted field drainage investigations regarding chemical and water flow characteristics of soils within the boundaries of the Oahe Project area in Brown and Spink counties (Bureau of Reclamation, 1974) as shown in Figure 2. Spink County data sets were collected in 1963, 1971 and 1972 while the Brown County data were collected in 1973 and 1974. Soil samples were collected from Lake Plain subsoils at depths ranging from 1.2 to 4.4 m (4 to 14.5 ft) in Spink County and to 5.7 m (18.5 ft) in Brown County. The soil sample depths were arbitrarily subdivided into three layers to discern chemical differences as a function of depth. The three layers were layer 1: 1.2 to 2.1 m (4.0 to 7.0 ft), layer 2: 2.1 to 3.1 m (7.0 to 10.0 ft) and layer 3: 3.1 to 4.4 or 5.7 m (10.0 to 14.5 ft or 10.0 to 18.5 ft). The dimensions of the top two layers are similar to the second and third layers of the Lake Andes-Wagner study. Composite profile values were not determined for Oahe Project soils because soil samples for the three layers were not always collected at the same sample site.

Most of the soil hydraulic conductivity tests were conducted *in situ* where water tables or saturated soil conditions were not present. In this case, the shallow well pump-in test was used to obtain saturated hydraulic conductivities (Bureau of Reclamation, 1978). When water tables were present, single auger pump-out tests were used. However, the data sets used in this study should be considered as pump-in data sets. Five centimeter diameter undisturbed soil cores were collected at each of the test sites. A portion of each core was used for chemical analyses in the laboratory.

Chemical properties evaluated in the investigations were the electrical conductivity (EC), pH and sodium (Na %) or exchangeable sodium percent (ESP) of aqueous extracts of soil pastes. Bureau of Reclamation personnel conducted all

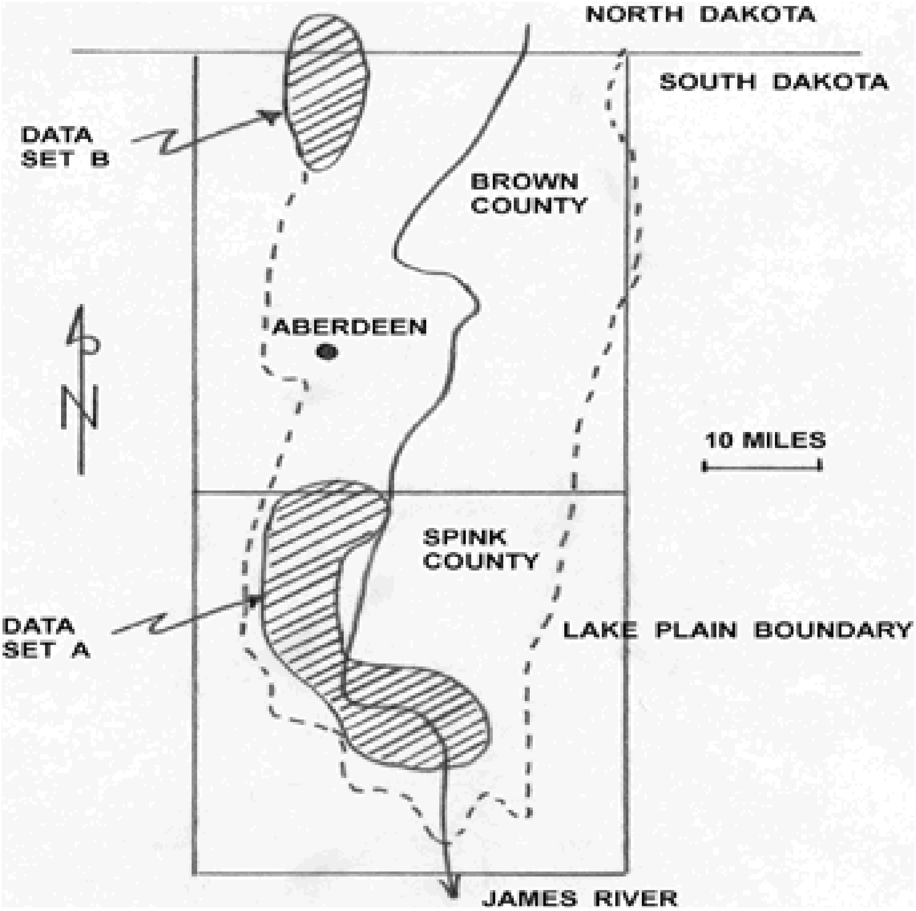

Figure 2. Collection location of Brown and Spink County data sets.

laboratory analyses. Specific laboratory procedures used for the chemical analyses are unknown because personnel associated with the project were not available for consultation.

Statistical Representations

Each data set will be represented mathematically by a Probability Density Function (PDF) which is an equation that can be used to describe the frequency of occurrence and the range of soil parameter data values. Such an equation can be readily adapted to computer computations and simulations. A PDF consists of a dependent parameter (Y), an independent parameter (X) (for example, soil salinity as measured by electrical conductivity) and constants derived from an analysis of the soil data sets (for example, A, B, C and D) and can be represented in the following general form.

$$Y = f(X, A, B, C, D)$$

Four common PDFs (Haan, 1977) were considered for application in this study and are as follows:

The Exponential PDF: $Y = Ae^{-BX}$

The Log-Normal PDF: $Y = (A/C)(1/X)e^{-0.5\{((L-B)/C)^2\}}$
where L = natural logarithm of X = ln (X)

The Normal PDF: $Y = (A/C)e^{-0.5\{((X-B)/C)^2\}}$

The Weibull PDF: $Y = AX^B e^{-(X^C/D)}$

Cumulative distribution equations can also be mathematically expressed as follows:

Exponential: $Z = 1 - e^{-BX}$

Weibull: $Z = 1 - e^{-(X^C/D)}$

Cumulative distributions for the log-normal and normal PDFs are difficult to obtain so it is recommended that one use an applied statistical reference such as Haan (1977) to properly use these two PDFs or to take advantage of the cumulative curve generating capacity of the SAS Institute (1999) computer program.

Examples of observed values, a representative Weibull PDF and a companion cumulative distribution for the saturated hydraulic conductivity data set of Spink County in the Oahe Project area are presented in Figure 3. The agreement

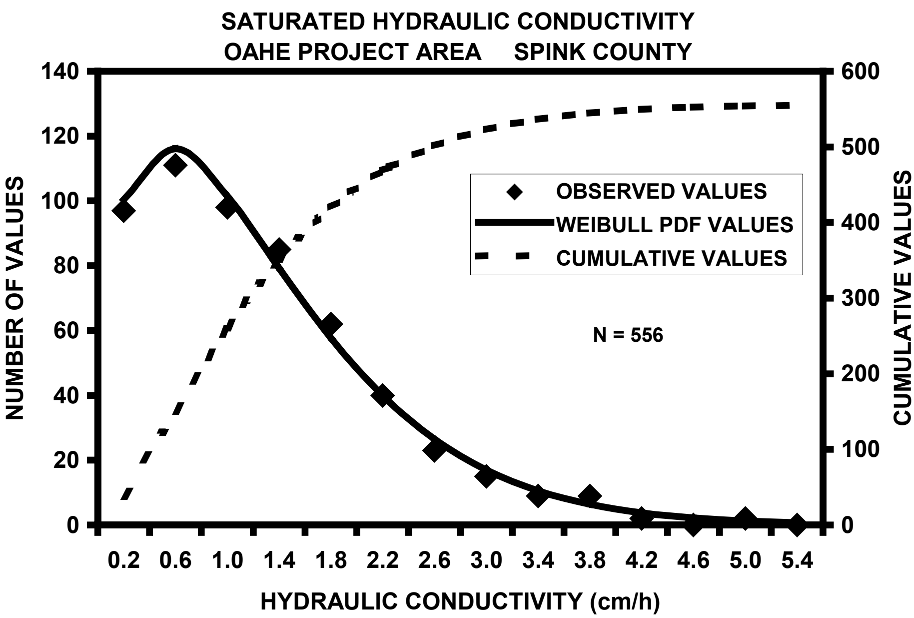

Figure 3. Observed, Weibull PDF and cumulative values for the Spink County saturated hydraulic conductivity data set.

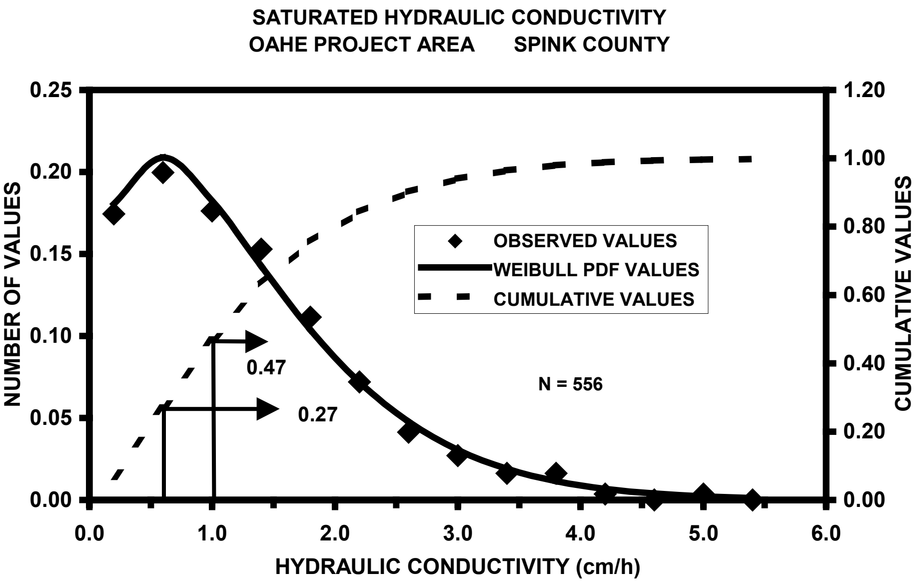

Figure 4. Decimal representations of the Figure 3 data set.

between observed and PDF values is one of the best for all data sets analyzed in this reported study. The cumulative value curve represents a sum of the area under the PDF curve moving from left to right and increasing to a maximum value of 556, the total number of data values in the data set. Normally, decimal equivalent values are used to describe such statistical distributions. Thus, values on the vertical axes of Figure 3 need to be converted to decimal equivalents by dividing each value by the total number of values in the data set (N = 556) as shown in Figure 4. The decimal form of the cumulative curve is used to estimate probabilities of occurrence.

The Kolmogorov-Smirnov and Cramer-von Mises goodness-of-fit statistics were employed to ascertain whether or not a given PDF could be used to describe a field data set (SAS Institute, 1999). In addition, the chi-square statistic, a less discerning statistic, was also used to evaluate data sets that appeared to have reasonable PDF fits but were not justified by the Kolmogorov-Smirnov and Cramer-von Mises statistics (Ostle, 1963).

Chemical parameter values of collapsed drift and till geological materials were represented in the Lake Andes-Wagner data sets. Visual inspections indicated that the shapes of the frequency distributions for a specific chemical parameter were similar for the two materials. So a statistical "t" test was used to determine if the means of the two data sets in question were statistically different from each other. If they were not, the data sets were combined into one combined data set.

The range of each data set was subdivided into 10 to15 classes as suggested by Haan (1977). Then data histograms were determined using the predetermined class widths for PDF analyses. Data histograms were used to identify "outlier" data values that were substantially larger than most of the data set values and caused standard deviation values to be non-representative of the data sets. The "outlier" data values were removed from data sets used for subsequent analyses. A maximum of about 6 % of the data values were removed from two available selenium data sets with an average of about 2 % for all the Lake Andes-Wagner Project chemical variable data sets. No "outliers" were removed from the Oahe Project chemical variable data sets. Three data values out of 355 were removed from the Brown County, layers 1 and 3 hydraulic conductivity data set. No other data were removed from the remaining hydraulic conductivity data sets.

RESULTS

Chemical Parameters

Lake Andes-Wagner Project Area. Means, standard deviations and the number of data values in each final data set are summarized in Table 1. Probability density function (PDF) parameter values were determined for the readily available selenium (F1), conditionally available selenium (F2), total selenium (TOT), electrical conductivity (EC), pH and sodium adsorption ratio (SAR) data sets as presented in Table 2. The Exponential, Log-Normal, Normal and Weibull PDFs can be used to describe 2, 23, 3 and 10 data sets, respectively. Three data sets could not be described with a PDF. Most of the selenium and EC PDFs are valid for both the drift and till geologic materials while the pH and SAR PDFs tend to be specific for each material. Examples showing excellent and good agreement between field data sets and PDFs are illustrated in Figures 5 and 6. Likewise, an example of a data set which could not be described by a PDF is presented in Figure 7. Figure 8 shows a data set-PDF scenario where the chi-square "goodness of fit" criterion was used to justify use of the Log-Normal PDF.

Oahe Project Area. Means, standard deviations and the number of data values in each final data set are summarized in Table 3. About one half of the data sets could not be described by a PDF (Table 4). The electrical conductivity (EC) data sets for Spink County followed Log-Normal PDFs and three of the possible six pH PDFs followed a Normal distribution. A Weibull PDF can be used to describe the exchangeable sodium percentage (ESP) data set for the second soil layer (2.1 to 3.1 m) in Brown County. The ESP data set for the third layer (3.1 to 5.7 m) is an example of a data set that could not be described by a PDF (Figure 9).

Hydraulic Conductivity

Lake Andes-Wagner Project Area. Means, standard deviations and the number of data values in each final data set are summarized in Table 5. The drift and till data sets are statistically different from each other and are defined by individual Log-Normal PDFs (Table 6). Mean conductivity values for the drift (N = 65) and till (N = 77) data sets are 3.49 and 1.12 cm/h, respectively. Figure 10 shows a good agreement between observed and PDF values for the drift data set.

Table 1. Mean and standard deviation values for the chemical property data sets of soils in the Lake Andes-Wagner Project area. "Outlier" data were removed from original data sets.

CHEMICAL PARAMETER	SOIL LAYER*	SOIL MATERIAL	NUMBER OF DATA VALUES	MEAN	STANDARD DEVIATION
F1 - Se	1	DRIFT&TILL	254	53.8 ppb	29.1 ppb
F1 - Se	2	DRIFT&TILL	249	147.0 ppb	180.4 ppb
F1 - Se	3	DRIFT&TILL	250	427.1 ppb	557.5 ppb
F1 - Se	4	DRIFT&TILL	253	500.6 ppb	636.1 ppb
F1 - Se	Comp	DRIFT&TILL	259	415.7 ppb	524.4 ppb
F2 - Se	1	DRIFT	131	255.5 ppb	92.9 ppb
F2 - Se	2	DRIFT	131	418.8 ppb	211.6 ppb
F2 - Se	1	TILL	125	199.8 ppb	93.4 ppb
F2 - Se	2	TILL	127	361.4 ppb	207.3 ppb
F2 - Se	3	DRIFT&TILL	255	521.8 ppb	290.7 ppb
F2 - Se	4	DRIFT&TILL	258	689.1 ppb	378.4 ppb
F2 - Se	Comp	DRIFT&TILL	258	522.1 ppb	250.2 ppb
TOTAL - Se	1	DRIFT&TILL	259	892.1 ppb	282.1 ppb
TOTAL - Se	2	DRIFT&TILL	255	881.8 ppb	467.9 ppb
TOTAL - Se	3	DRIFT&TILL	253	1219.0 ppb	883.0 ppb
TOTAL - Se	4	DRIFT&TILL	253	1487.0 ppb	984.4 ppb
TOTAL - Se	Comp	DRIFT&TILL	259	1277.0 ppb	760.6 ppb
EC	1	DRIFT&TILL	256	1.24 dS/m	1.23 dS/m
EC	2	DRIFT&TILL	236	3.97 dS/m	2.86 dS/m
EC	3	DRIFT&TILL	263	6.09 dS/m	3.03 dS/m
EC	4	DRIFT&TILL	264	6.11 dS/m	3.00 dS/m
EC	Comp	DRIFT&TILL	264	4.98 dS/m	2.19 dS/m
pH	1	DRIFT&TILL	264	7.90	0.363
pH	2	DRIFT	131	7.80	0.282
pH	3	DRIFT	130	7.63	0.244
pH	4	DRIFT	132	7.62	0.314
pH	Comp	DRIFT	132	7.69	0.179
pH	2	TILL	131	7.89	0.286
pH	3	TILL	130	7.72	0.311
pH	4	TILL	128	7.72	0.311
pH	Comp	TILL	131	7.79	0.158
SAR	1	DRIFT	133	2.39	2.90
SAR	2	DRIFT	133	3.29	3.05
SAR	3	DRIFT	133	4.85	4.14
SAR	4	DRIFT	133	4.91	4.18
SAR	Comp	DRIFT	133	4.20	2.89
SAR	1	TILL	127	0.63	0.84
SAR	2	TILL	131	2.16	2.18
SAR	3	TILL	128	3.27	2.76
SAR	4	TILL	130	3.84	2.79
SAR	Comp	TILL	131	3.00	2.18

*LAYER 1 (0.0-0.5 m); LAYER 2 (0.5-1.0 m); LAYER 3 (1.0-2.0 m); LAYER 4 (2.0-3.0 m); LAYER Comp (0.0-3.0 m, COMPOSITE)

Table 2. Probability density function (PDF) parameters for the chemical properties of soils in the Lake Andes-Wagner Project area.

PDF	CHEMICAL PARAMETER	SOIL LAYER	SOIL MATERIAL	PDF PARAMETER VALUES				APPLICABLE DATA RANGE
				A	B	C	D	
Exponential	SAR	2	TILL	0.370	0.463			0 – 8.8
	SAR	3	TILL	0.306	0.306			0 – 11.0
Log-Normal	F1 - Se	1	DRIFT&TILL	3.99	3.86	0.507		0 – 160 ppb
	F1 - Se	2	DRIFT&TILL	23.9	4.34	1.21		0 – 900 ppb
	F1 - Se	3	DRIFT&TILL	79.8	5.09	1.62		0 – 2400 ppb
	F1 - Se	4 *	DRIFT&TILL	99.7	5.20	1.72		0 – 3000 ppb
	F1 - Se	Comp *	DRIFT&TILL	79.8	5.23	1.34		0 – 2400 ppb
	F2 - Se	1	DRIFT	16.0	5.47	0.394		0 – 520 ppb
	F2 - Se	2	TILL	32.9	5.73	0.578		0 – 1120 ppb
	TOTAL - Se	1 *	DRIFT&TILL	39.9	6.74	0.329		200 – 1800 ppb
	TOTAL - Se	2	DRIFT&TILL	79.8	6.66	0.490		200 – 3000 ppb
	TOTAL - Se	3	DRIFT&TILL	139.0	6.86	0.705		0 – 4200 ppb
	TOTAL - Se	4	DRIFT&TILL	139.0	7.08	0.702		0 – 5250 ppb
	TOTAL - Se	Comp	DRIFT&TILL	120.0	6.99	0.567		0 – 3900 ppb
	pH	1	DRIFT&TILL	0.0798	2.07	0.0458		6.7 – 8.9
	pH	2*	DRIFT	0.0399	2.05	0.0358		7.2 – 8.6
	pH	2*	TILL	0.0399	2.06	0.0360		7.2 – 8.6
	pH	3*	DRIFT	0.0399	2.03	0.0318		7.0 – 8.3
	pH	3	TILL	0.0399	2.05	0.0335		7.0 – 8.5
	pH	4*	DRIFT	0.0399	2.03	0.0411		6.8 – 8.6
	pH	4	TILL	0.0399	2.05	0.0302		6.9 – 8.5
	pH	Comp	DRIFT	0.0399	2.04	0.0231		7.3 – 8.4
	pH	Comp	TILL	0.0399	2.05	0.0203		7.2 – 8.3
	SAR	1	DRIFT	0.399	-0.0472	1.55		0.0 – 14.0
	SAR	1	TILL	0.200	-1.07	0.841		0.0 – 6.0
Normal	EC	3	DRIFT&TILL	0.399	6.09	3.03		0 – 15.0 dS/m
	EC	4	DRIFT&TILL	0.399	6.11	3.00		0 – 14.0 dS/m
	EC	Comp	DRIFT&TILL	0.319	4.98	2.19		0 – 12.0 dS/m
Weibull	F2-Se	1	TILL	0.000427	1.26	2.26	212,000	0 – 520 ppb
	F2-Se	2	DRIFT	0.000704	1.01	2.01	228,000	0 – 1040 ppb
	F2-Se	3	DRIFT&TILL	0.00103	0.900	1.90	184,000	0 – 1500 ppb
	F2-Se	4	DRIFT&TILL	0.000857	0.911	1.91	335,000	0 – 1950 ppb
	F2-Se	Comp	DRIFT&TILL	0.000151	1.23	2.23	1,480,000	100 – 1400 ppb
	SAR	2	DRIFT	0.290	0.0782	1.08	3.73	0 – 15.0
	SAR	3	DRIFT	0.275	0.177	1.18	6.86	0 – 20.8
	SAR	4	DRIFT	0.272	0.173	1.17	6.90	0 – 19.2
	SAR	Comp	DRIFT	0.174	0.524	1.52	10.5	0 – 14.4
	SAR	Comp	TILL	0.216	0.366	1.37	5.06	0 – 9.6
NONE	EC	1	DRIFT&TILL					
	EC	2	DRIFT&TILL					
	SAR	4	TILL					

* Statistical significance based on chi-square criterion.

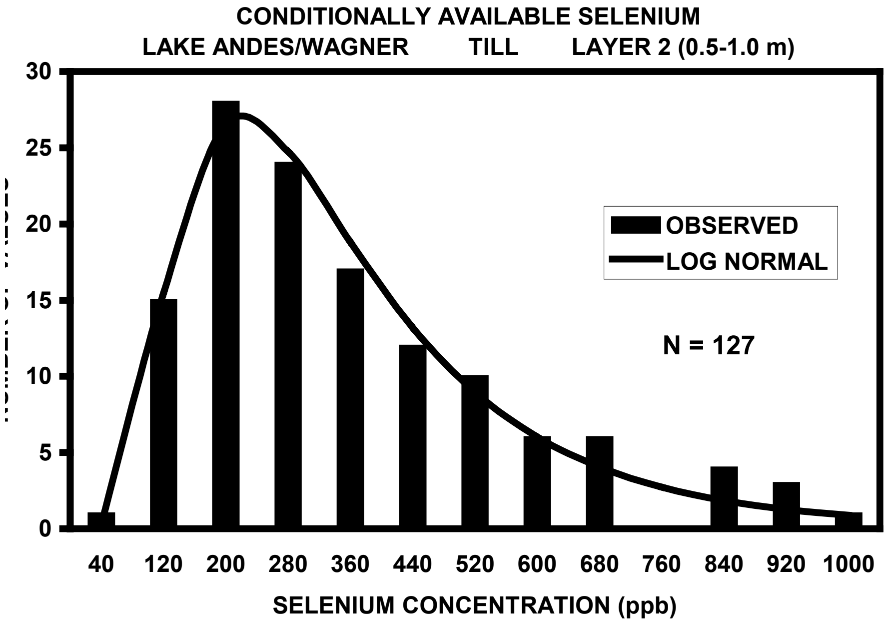

Figure 5. Excellent agreement between observed and Log-Normal PDF values for conditionally available selenium at the 0.5 to 1.0 m depth in till material for the Lake Andes-Wagner Project area.

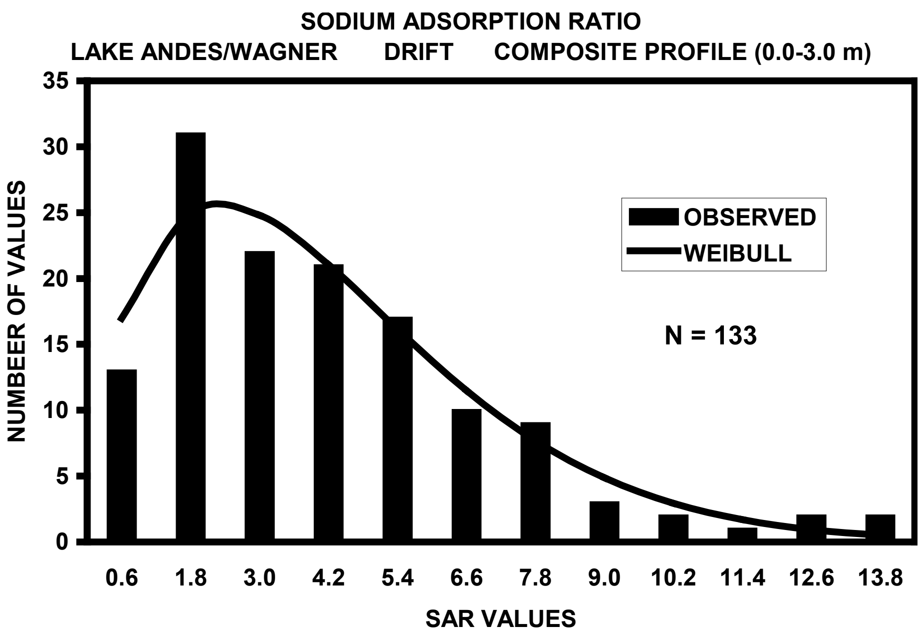

Figure 6. Good agreement between observed and Weibull PDF values for sodium adsorption ratios for a composite profile in collapsed drift material for the Lake Andes-Wagner Project area.

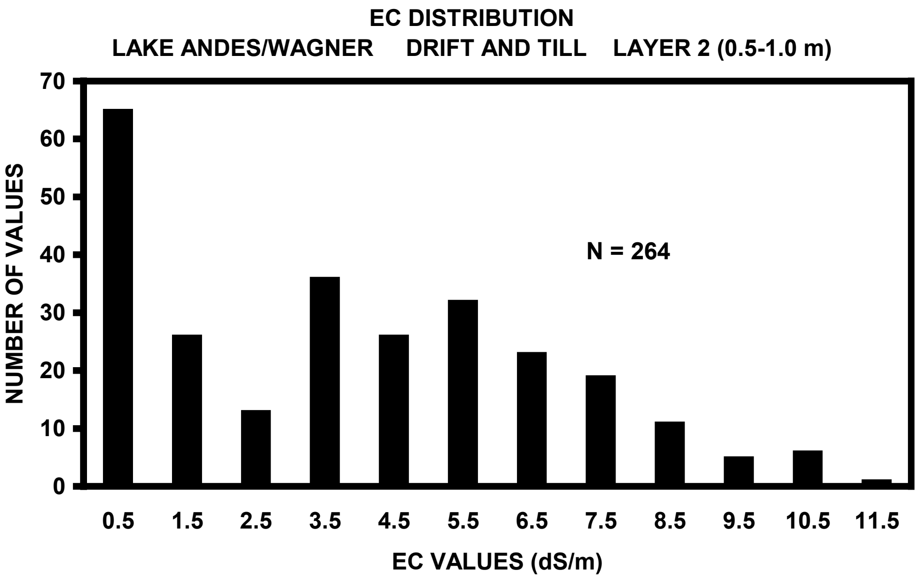

Figure 7. A data set in the Lake Andes-Wagner Project area that could not be described by a PDF.

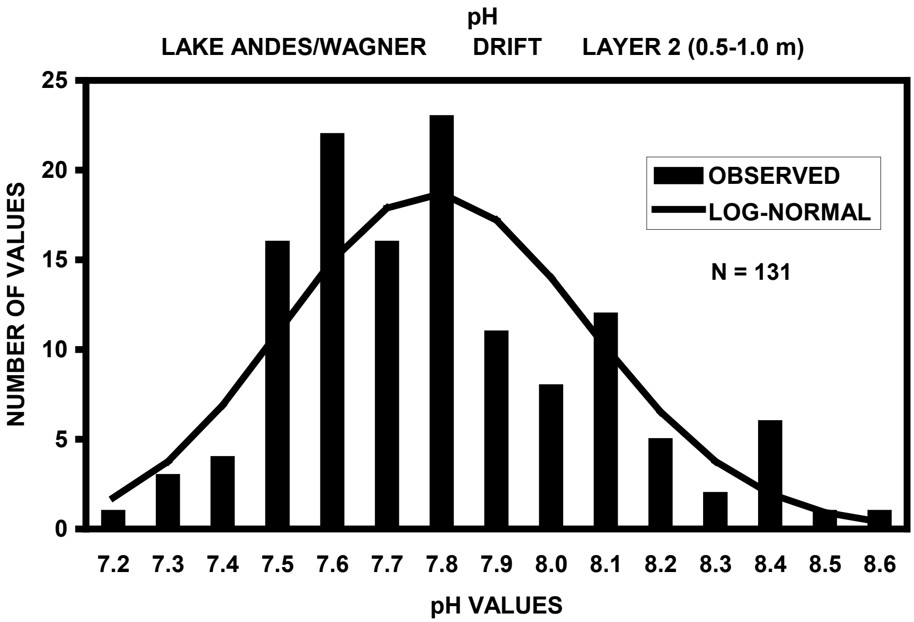

Figure 8. A data set where the chi-square "goodness-of-fit" was used to justify the use of a Log-Normal PDF to describe the data set.

Table 3. Mean and standard deviation values for the chemical property data sets of soils in the Oahe Project area.

CHEMICAL PARAMETER	SOIL LAYER	SOIL MATERIAL	NUMBER OF DATA VALUES	MEAN	STANDARD DEVIATION
EC	1	BROWN CO	251	8.59 mmhos/cm	4.92 mmhos/cm
EC	2	BROWN CO	188	5.92 mmhos/cm	4.19 mmhos/cm
EC	3	BROWN CO	101	5.27 mmhos/cm	3.09 mmhos/cm
EC	1	SPINK CO	227	3.07 mmhos/cm	3.71 mmhos/cm
EC	2	SPINK CO	202	2.31 mmhos/cm	2.82 mmhos/cm
EC	3	SPINK CO	127	1.79 mmhos/cm	2.13 mmhos/cm
ESP	1	BROWN CO	251	8.76 %	5.28 %
ESP	2	BROWN CO	188	6.46 %	5.02 %
ESP	3	BROWN CO	101	5.70 %	3.73 %
Na (%)	1	SPINK CO	194	6.04	4.33
Na (%)	2	SPINK CO	156	5.13	3.64
Na (%)	3	SPINK CO	86	2.67	2.74
pH	1	BROWN CO	251	7.64	0.262
pH	2	BROWN CO	188	7.50	0.252
pH	3	BROWN CO	101	7.38	0.200
pH	1	SPINK CO	227	7.85	0.355
pH	2	SPINK CO	202	7.78	0.408
pH	3	SPINK CO	127	7.75	0.436

* LAYER 1 (1.2-2.1 m); LAYER 2 (2.1-3.1 m); LAYER 3 (3.1-4.4 m or 3.1-5.7 m)

Table 4. Probability density function (PDF) parameter values for chemical properties of Oahe Project soils.

PDF	CHEMICAL PARAMETER	SOIL LAYER	SOIL MATERIAL	PDF PARAMETER VALUES				APPLICABLE DATA RANGE
				A	B	C	D	
Log-Normal	EC	1	SPINK CO	0.399	0.525	1.11		0.0 – 16.0 mmhos/cm
	EC	2	SPINK CO	0.399	0.172	1.23		0.0 – 15.0 mmhos/cm
	EC	3	SPINK CO	0.399	-0.0820	1.21		0.0 – 11.0 mmhos/cm
Normal	pH*	1	BROWN CO	0.0399	5.56	3.48		7.0 – 8.4
	pH*	2	BROWN CO	0.0399	7.50	0.252		7.0 – 8.4
	pH*	1	SPINK CO	0.0798	7.85	0.355		7.0 – 9.0
Weibull	ESP*	2	BROWN CO	0.283	0.0711	1.07	7.56	0.0 – 22.0 %
NONE	EC	1	BROWN CO					
	EC	2	BROWN CO					
	EC	3	BROWN CO					
	ESP	1	BROWN CO					
	ESP	3	BROWN CO					
	pH	3	BROWN CO					
	pH	2	SPINK CO					
	pH	3	SPINK CO					
	Na (%)	1	SPINK CO					
	Na (%)	2	SPINK CO					
	Na (%)	3	SPINK CO					

* Statistical significance based on chi-square criterion.

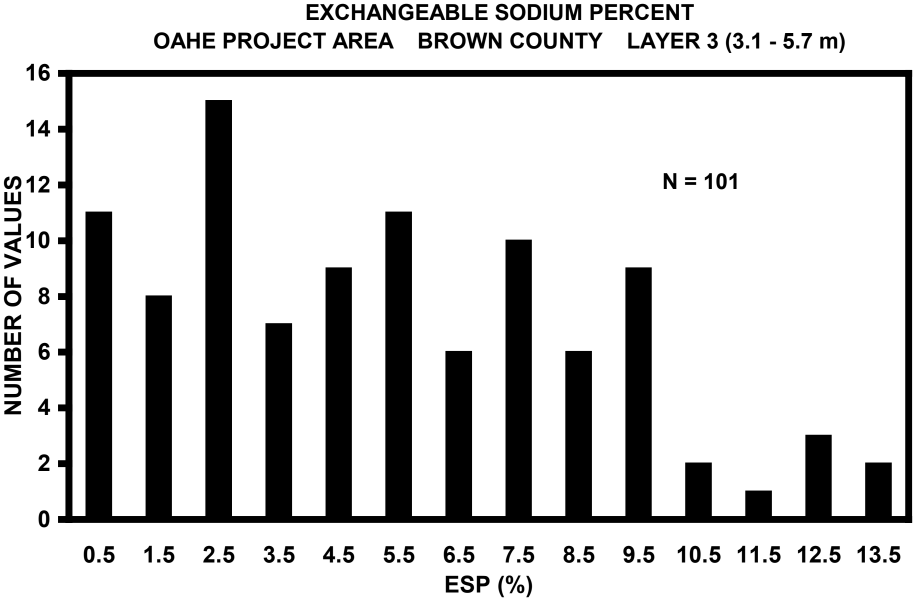

Figure 9. A data set in the Oahe Project area that could not be described by a PDF.

Table 5. Mean and standard deviation values for the saturated hydraulic conductivity data sets of Lake Andes-Wagner and Oahe Project area soils.

PROJECT AREA	SOIL MATERIAL	NUMBER OF DATA VALUES	MEAN	STANDARD DEVIATION
LAKE ANDES-WAGNER	DRIFT	65	3.49 cm/h	3.49 cm/h
LAKE ANDES-WAGNER	TILL	77	1.12 cm/h	1.05 cm/h
OAHE	SPINK COUNTY	556	1.28 cm/h	0.98 cm/h
OAHE	BROWN COUNTY LAYERS 1 & 3	352	0.99 cm/h	1.22 cm/h
OAHE	BROWN COUNTY LAYER 2	188	1.38 cm/h	1.54 cm/h

Table 6. Probability density function (PDF) values for saturated hydraulic conductivity properties of Lake Andes-Wagner and Oahe Project area soils.

PDF	PROJECT AREA	SOIL MATERIAL	PDF PARAMETER VALUES				APPLICABLE DATA RANGE
			A	B	C	D	
Log-Normal	LAKE ANDES WAGNER	DRIFT	0.319	0.701	1.16		0.0 – 15.0 cm/h
	LAKE ANDES WAGNER	TILL*	0.0399	2.05	0.0203		0.0 – 5.0 cm/h
Weibull	OAHE	SPINK COUNTY	0.344	0.356	1.36	1.58	0.0 – 7.4 cm/h
	OAHE	BROWN COUNTY* LAYERS 1 & 3	0.410	0.0321	1.03	1.01	0.0 – 7.0 cm/h
NONE	OAHE	BROWN COUNTY LAYER 2					

*Statistical significance based on chi-square criterion.

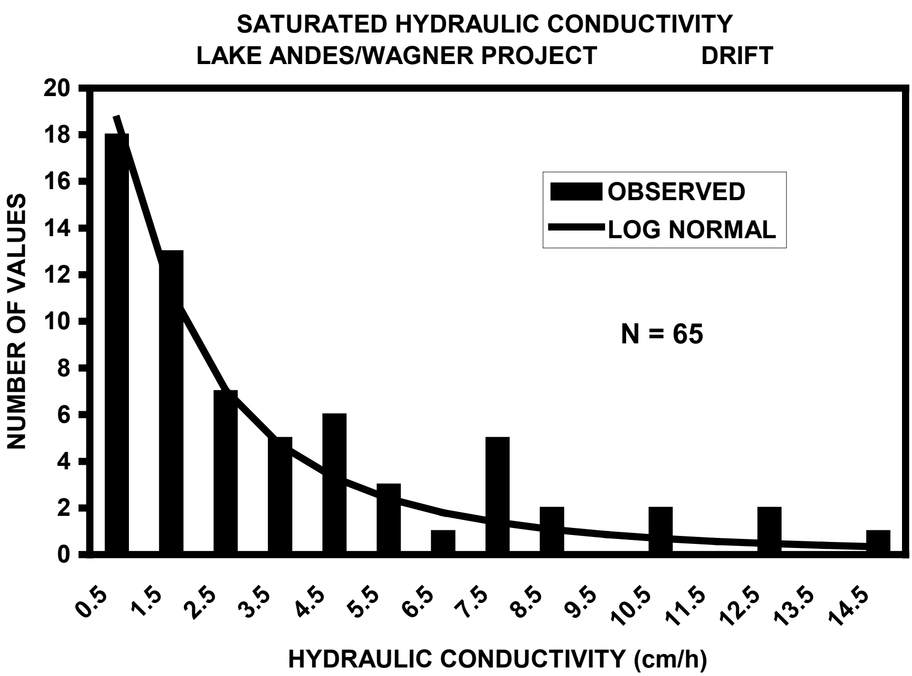

Figure 10. Good agreement between observed and Log-Normal PDF values for saturated hydraulic conductivity in collapsed drift material of the Lake Andes-Wagner Project area.

Oahe Project Area. A statistical test indicated that the conductivity data sets for the three soil layers in Spink County were not different from each other and could be represented by a composite data set (N = 556) and a Weibull PDF (Table 6 and Figure 3). The Weibull PDF can also be used to depict the similar and combined layer 1 and 3 data sets (N = 352) for Brown County while the layer 2 data set (N = 184) can not be describe by a PDF (Figure 11).

APPLICATION OF FINDINGS

Two examples will be used to acquaint the reader with the utility of the study results.

Example 1: What is the probability of having a saturated hydraulic conductivity value of 1.0 cm/h or less in Spink County? What is the probability of a value between 0.6 and 1.0 cm/h?

Using the graphical presentation of the Spink County data set as found in Figure 4, we see that a cumulative value of 0.47 is associated with a 1.0 cm/h conductivity value which means that 47 % of the values can be expected to be less than or equal to 1.0 cm/h. Likewise, a value of 0.27 (27 %) is associated with a conductivity value of 0.6 cm/h. The difference between the 27 and 47 % values equals 20 % and indicates that 20 % of all values can be expected to be between conductivity values of 0.6 and 1.0 cm/h.

Example 2: What is the probability of finding a conditionally available selenium value between 800 and 900 ppb at a depth of 1.5 m (layer 3) in a till soil

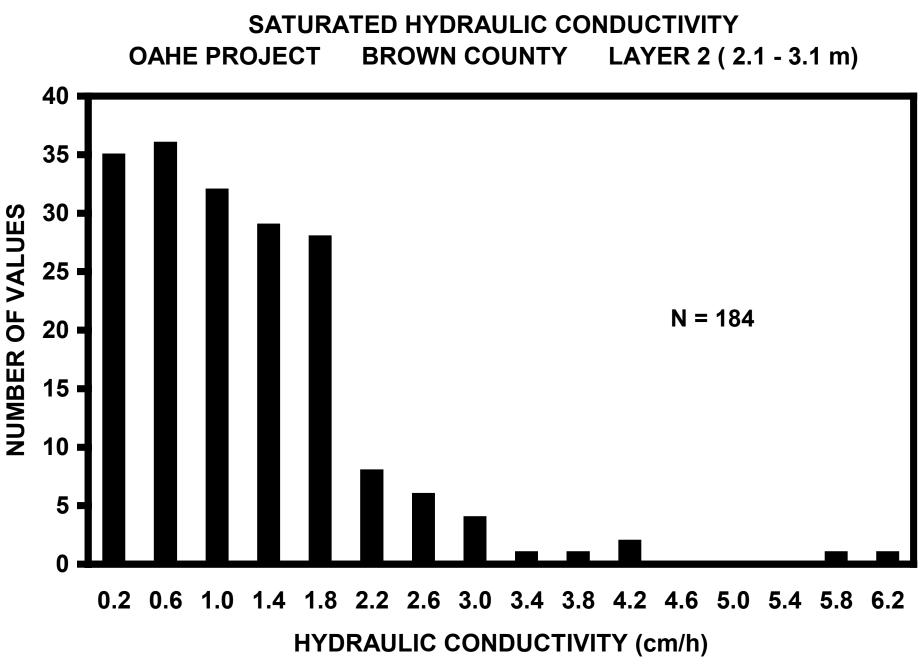

Figure 11. A data set in the Oahe Project area that could not be described by a PDF.

in the Lake Andes Wagner Project area?

We know from Table 2 that the PDF for conditionally available selenium in the soil follows a Weibull mathematical equation and that A = 0.00103, B = 0.900, C = 1.90 and D = 184,000. We also know that we are dealing with X values of 800 and 900 ppb.

The cumulative Weibull equation becomes $Z = 1 - e^{-(X^{1.90}/184{,}000)}$ and $Z = 1 - e^{-(800^{1.90}/184{,}000)}$ for an X value of 800 ppb. This becomes $1.0 - e^{-1.78}$ which in turn becomes 1.0 - (0.168) and results in a value of 0.832. Likewise, $Z = 1 - e^{-2.23}$ for an X value of 900 and we get 1.0 - (0.108) or 0.892. So the difference between the two Z values (0.892 - 0.832 = 0.060) gives us a 6.0 % probability of finding a conditionally available selenium value between 800 and 900 ppb at a depth of 1.5 m in a till soil in the Lake Andes Wagner Project area. That means if we measured conditionally available selenium values for 100 soil samples at a 1.5 depth, we would expect to find 6 (100 times 0.060) samples with conditionally available selenium values between 800 and 900 ppb.

DISCUSSION

Probability density functions are often used for computer modeling purposes with the Log-Normal PDF being one of the most popular. Since the Log-Normal is relatively simple to use, is bounded by zero on the left and is positively skewed, it has received wide acceptance in hydrologic studies (Haan, 1977). Parkin and Robinson (1992) presented a list of abiotic and biological variables that have been observed to be Log-Normally distributed including soil hydraulic conductivity, soil nitrate, soil denitrification, bacterial populations on leaves and species numbers. Doolittle et al. (1995) performed an analysis of the Lake Andes-Wagner selenium data set assuming all data distributions followed a Log-Normal PDF. The results of this reported study show that the Weibull PDF does a better job of describing the conditionally available selenium (F2) data sets.

The Weibull PDF also can be used to describe several of the data sets considered in this study. It, like the Log-Normal, is bounded by zero on the left but is a four parameter relationship in contrast to the two parameter Log-Normal PDF. However, its cumulative relationship requires numerical values for only two parameters and can be expressed in a simple mathematical equation for analytical purposes.

The Normal and Exponential PDFs were seldom the best PDFs for use with the data sets considered in this study. Fifteen of the 64 possible data sets investigated in this study could not be described by a PDF with 11 of them being chemical properties associated with the Oahe Project area.

One other saturated hydraulic conductivity data set was collected in association with the proposed CENDAK Irrigation Project east of the Missouri River in central South Dakota (Bureau of Reclamation, 1988). However, it contains an insufficient number of data values for the development of reliable PDFs.

ACKNOWLEDGEMENTS

The author gratefully acknowledges the efforts of Anthony Bly, Xinjiang Huang, Robert Kohl and James Doolittle in the establishment of the Lake Andes-Wagner soil chemical data sets and the financial support of the S.D. Agricultural Experiment Station and the U.S. Bureau of Reclamation. The reported work is published with the approval of the Director of the S.D. Agricultural Experiment Station as Journal Article No. 3501.

LITERATURE CITED

Briggs, P.H. and J.G. Crock. 1986. Automated determination of total selenium in rocks, soils and plants. US Geological Survey Open-File Report 86-40. Washington, DC.

Bureau of Reclamation. 1974. Unpublished Oahe Unit hydraulic conductivity data. US Department of the Interior. Denver, CO.

Bureau of Reclamation. 1978. Drainage Manual. U.S. Department of the Interior. US Govt. Printing Office, Washington, DC.

Bureau of Reclamation. 1985. Lake Andes-Wagner Unit planning report/draft environmental statement: Appendix G - Drainage. U.S. Department of the Interior, Upper Missouri Region, Billings, MT.

Bureau of Reclamation. 1988. Concluding report of CENDAK drainage investigations. U.S. Department of the Interior, Upper Missouri Region, Billings, MT.

Chao, T.T. and R.F. Sanzolene. 1989. Fractionation of soil selenium by sequential partial dissolution. Soil Sci. Soc. Am. J. 53:385-392.

DeBoer, D.W., J. J. Doolittle and R. A. Kohl. 1995. Spatial variability of selenium, salinity and water retention in a potentially irrigable soil (Lake Andes-Wagner/Marty II Irrigation Project). Completion report. US Bureau of Reclamation Research Grant 2-FG-60-02750. Department of Agricultural and Biosystems Engineering, SD State University. Brookings, SD.

Doolittle, J.J., X. Huang and A.G. Bly. 1995. The spatial variability of selenium in selected soils of Charles Mix County, South Dakota. Proc. of SD Acad. Sci. 74:113-122.

Haan, C. T. 1977. Statistical Methods in Hydrology. The Iowa State University Press. Ames, IA. 378 pp.

Ostle, B. 1963. Statistics in Research. The Iowa State University Press. Ames, IA. 585 pp.

Parkin, T. B. and J. A. Robinson. 1992. Analysis of lognormal data. p 193-235. In: B. A. Stewart (ed.). Advances in Soil Science Vol 20. Springer-Verlag. New York.

Rhoades, J.D. 1982. Soluble salts. p 167-179. In: A.L. Page et al. (ed.) Methods of soil analysis, Part 2. Chemical and microbial properties. Agronomy Monograph No. 9. Am. Soc. of Agron., Madison, WI.

SAS Institute. 1999. The SAS System for Windows, Version 8. SAS Institute, Inc. Cary, NC