This is to certify that the

thesis entitled

GEOSTATISTICAL ANALYSIS OF PENNSYLVANIAN SEDIMENTS IN THE EASTERN MICHIGAN BASIN

presented by

Timothy Gerard Monaghan

has been accepted towards fulfillment of the requirements for

~~Masters~~ degree in ~~Geological~~ Sciences

Major professor

Date May 12, 1998

O-7639 *MSU is an Affirmative Action/Equal Opportunity Institution*

GEOSTATISTICAL ANALYSIS OF PENNSYLVANIAN SEDIMENTS IN THE EASTERN MICHIGAN BASIN

By

Timothy Gerard Monaghan

A THESIS

Submitted to
Michigan State University
in partial fulfillment of the requirements
for the degree of

MASTER OF SCIENCE

Department of Geological Sciences

1998

Abstract

GEOSTATISTICAL ANALYSIS OF PENNSYLVANIAN SEDIMENTS IN THE EASTERN MICHIGAN BASIN

By

Timothy Gerard Monaghan

The purpose of this study is to determine how well sandstone distribution within the Pennsylvanian Strata of the Michigan Basin can be predicted. This analysis provides a test of the hypothesis that lithologic variability is predictable in ancient fluvial sediments. A secondary hypothesis is that coal exploration logs, recorded by geologists, are more accurate than water well logs which were primarily performed by non-geologists.

Geostatistics can evaluate large data sets to determine the most likely distribution of variables between known data points. We used the Michigan Computerized Groundwater Resources Information Bank (MCGRIB) to determine any predictable spatial patterns within the Pennsylvanian Strata of the Eastern Michigan Basin.

We discovered that while there are widespread trends within data, there are too many errors within the data to predict sandstone distribution at a local scale.

TABLE OF CONTENTS

List of Tables	iv
List of Figures	v
Introduction	1
Geostatistics	3
Previous Work	12
Methods	26
Results	33
Discussion	43
Conclusions	50
Works Cited	51

LIST OF TABLES

Table 1 27
Township Locations and Designations in Bay County Data Set

Table 2 37
Z-star and standard deviation of
the indicator data sets (thickness ss).

Table 3 38
Z-star and standard deviation of the data sets (thickness ss).

Table 4 39 and 43
Statistical outputs and values for data sets.

Table 5 47
Median sandstone thicknesses for townships in Bay County.

LIST OF FIGURES

Figure 1 3
Typical variogram with a gaussian model.

Figure 2 5
Typical variogram with sinusoidal variations.

Figure 3 7
All types of variogram models superimposed.

Figure 4 12
Simplified stratigraphic column from the area of study.

Figure 5 14
Structural contours of Pennsylvanian Sediments.

Figure 6 15
Isopach of Winn Formation in Central Michigan.

Figure 7 23
County map of Michigan with Bay County highlighted

Figure 8 24
Township and section map of Bay County Michigan.

Figure 9 25
Location of sandstone in Bay County.

Figure 10 29
Cumulative percent plots for thickness of sandstone (thick ss) and log thickness sandstone (log ss).

Figure 11 33
Variogram for area bay6c, Gaussian Model: Nugget = 750, C - Value = 1500, Range = 1.5.

Figure 12 33
Variogram for area bay6c, Linear Model: Nugget = 750,
C - Value = 1500, Range = 1.5.

Figure 13 34
Variogram for area bay5a, Gaussian Model: Nugget = 475,
C - Value = 325, Range = 1.

Figure 14 34
Variogram for area bay7a, Gaussian Model: Nugget = 25,
C - Value = 23, Range = 1.6.

Figure 15 35
Variogram for area bay7e, Gaussian Model: Nugget = 22,
C - Value = 15, Range = 1.5.

Figure 16 36
Error map of area bay7e.

Figure 17 36
Error map of area bay6a.

Figure 18 41
Linear regression plot of bay6c.dat.

Figure 19 41
Linear regression plot of entire Bay County data set.

Figure 20 42
Linear regression plot of entire Bay County data set,
sorted for all rock well logs.

Introduction

Geologic data are commonly displayed as a spatial array; often the most appropriate type of display consists of X, Y, Z data used to construct a contour surface. Geostatistics can be used to characterize data value variation with distance, to estimate data values at locations without observations and to characterize these estimations statistically. We will use geostatistics to determine if a sandstone thickness at a random point can be predicted based on well log data from the surrounding area. The data set used is the Michigan Computerized Groundwater Resources Information Bank (MCGRIB) of computerized well logs, which includes private and municipal water well logs along with coal exploration boring logs (Monaghan and Larson 1985).

The purpose of this study is to determine how well the distribution of sandstones of Pennsylvanian strata in Bay County, Michigan can be predicted. This analysis provides a test of the hypothesis that lithologic variability is predictable in ancient fluvial sediments. A secondary hypothesis is that coal exploration logs, recorded by geologists, are more accurate than water well logs which were recorded mainly by non-geologists.

The MCGRIB data base used in this study encompasses more than 10 counties in the lower peninsula of Michigan. Some of the logs are highly detailed with recorded thickness intervals to the nearest half foot. Bay County was chosen for detailed study because of higher density of coal boring logs. If a suitable geostatistical model can be developed with Bay County data, then it might be applied to other areas in Michigan that contain Pennsylvanian strata to determine if the depositional environment is constant across the basin. If the depositional environment is constant across the basin, the model can then be used to predict the distribution of sandstone.

Geostatistics

Geostatistics refers to statistical methods used to evaluate data variation with distance. Variography is the basic method for evaluating how data that has relatively normal distribution and is continuous varies with

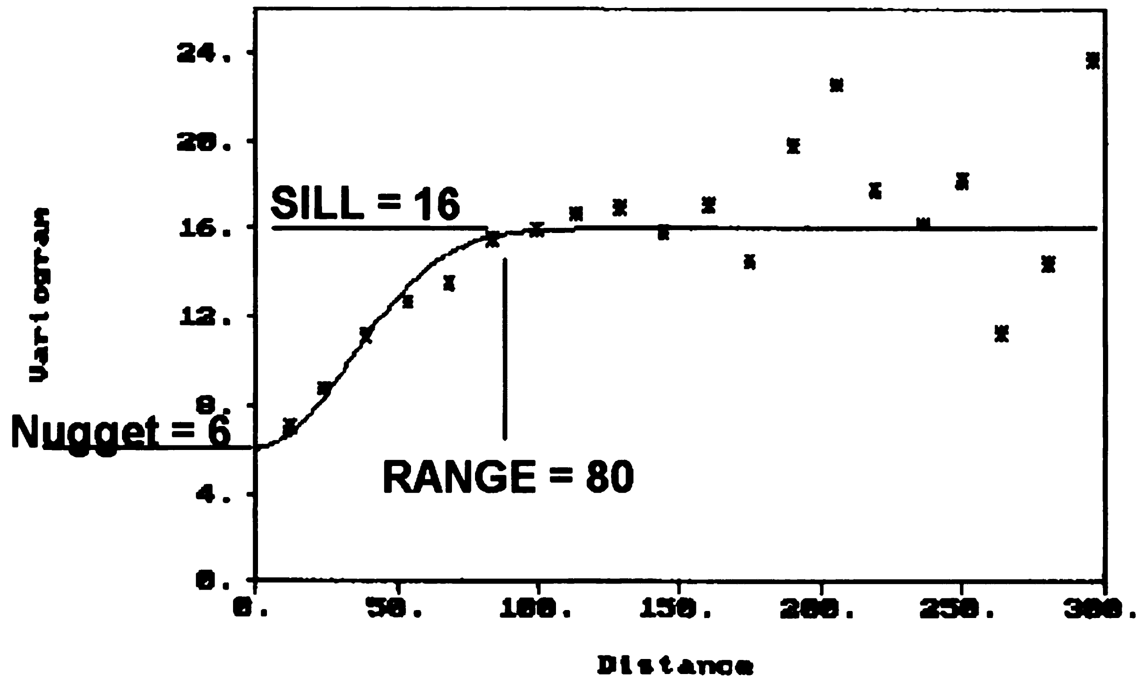

Figure 1
Typical variogram with a gaussian model.
Range, Sill and Nugget are noted.

distance. A variogram (figure 1) is a plot of variance between data points versus distance between those points; generally as distance increases so does

$$\gamma(h) = {}^{1}/_{2N(h)}\, \Sigma_{(i,j)}(v_i - v_j)^2$$

h = radial distance from one point to any other.

$\Sigma_{(i,j)}(v_i - v_j)^2$ = sum of the variance between two distances.

Mathematical formula for the variogram.

the variance plot (Grant et. al. 1994, Isaaks and Srivastava 1989, Olea R A 1996, and Wolf et. al. 1996). The variogram visually displays cumulative variance between all data points at a given distance. The terminology for this distance is lag or lag spacing. Change in cumulative variance may be predictable over a certain distance until the variance of the entire data set is reached. The range of a data set is the lag distance where cumulative variance no longer changes with lag distance and cumulative variance equals variance of the data set. Variogram sill is the cumulative variance of the data and is a point where the variogram levels out or where data variance no longer changes significantly with distance (Grant et. al. 1994, Isaaks and Srivastava 1989, Johnson and Dreiss 1989, Olea R A 1996, and Wolf et. al. 1996). If the variogram shows variance at zero distance, it has a nugget. Nugget may be a measure of the precision of a data set. If data has variance at zero distance, or a nugget value, then some of the data can be assumed to reflect error or a fine scale variability (Grant et. al. 1994, Isaaks and Srivastava 1989, Liu et. al. 1996, Olea R A 1996, and Wolf et. al. 1996). A hole effect is a sinusoidal variogram once the sill has been achieved (figure 2). Time series analysis may be used to determine if periodicity of variance is predictable. Johnson and Davis (1989) and Desbarats and Bachu (1994) suggested that the periodicity of the data could be a measure of the length (horizontal) or thickness (vertical) of a sedimentary body.

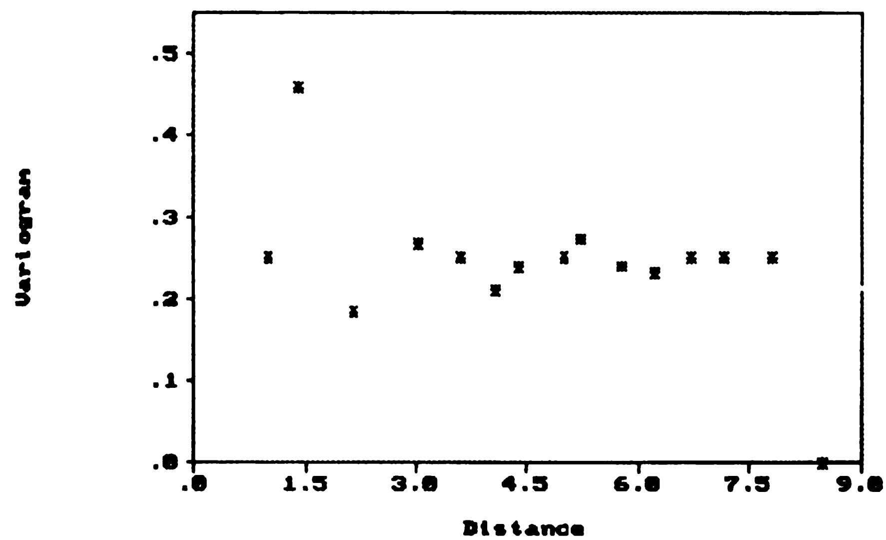

Figure 2
Typical variogram with sinusoidal variations.

Data are searched by comparing a data point with all others within a given area. This search is performed in a circular or elliptical shape. The search pattern (ellipse) defines an area, around a given point, in which we assume that data values are most alike. This pattern is usually oriented to exploit any trends in data distribution. Because data values in a given area are usually similar, data values may have a preferred orientation. Thus, if there is a preferred orientation that has a diagonal trend through the data, the ellipse will have its long axis oriented along that trend. In an elliptical search, values at greater distances along the preferred orientation are as important as those closer data points perpendicular to the orientation. This utilization of a preferred orientation can be used to effectively evaluate changes in variance, since it evaluates variance trends that are intrinsic to the data. For efficient searches of data values, the search ellipse should be oriented along any preferred orientation, so the evaluation will be performed on data values that are most similar. This search will then allow a preferred orientation within the data to be preserved, by allowing data points that are similar to maintain their similarity.

Spatial variation can also be analyzed with indicator data. Indicator data are raw data converted to a one or zero. For example, data may be converted to a value of one, if the item of interest is present and to zero if the item of interest is absent or vice versa. Indicators can be a powerful method of evaluating data because it allows data to be evaluated by presence or absence of data values, instead of the raw data values. Raw data can vary greatly within a given data set and abnormally high and low values can impede variogram analysis due to high variance values at close lags. Indicator data discounts the significance of abnormally high or low values. This method is used to evaluate variance with distance with only magnitude being altered; any spatial relationships between data points are preserved (Solow 1993).

A map created from indicator data may be used as a probability plot. A value of 0.5 is the separation point of the two values (zero and one) and there is an equal probability of either variable at the 0.5 contour. Values less than 0.5 are more likely to have zero value and vice versa.

Four types of models are commonly fitted to variograms; spherical,

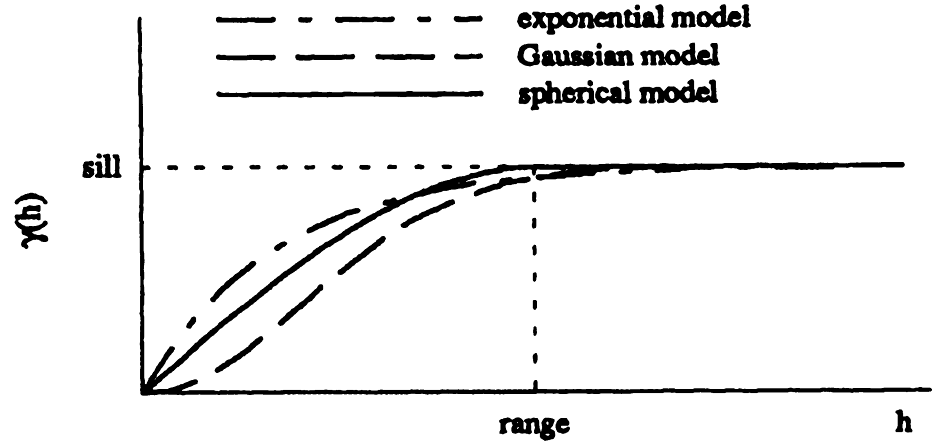

Figure 3
All types of variogram models superimposed.
From Isaaks and Srivastava (1989)

exponential, gaussian and linear. Each model gives a mathematical formula that can be used to predict variance with distance, using range, sill and any nugget value to define the shape of the model (figure 3). Keckler (1994), used a “scale” value, which is the difference between sill and nugget, to describe the model and the search radius to define the range. Similarly, GEOEAS defines the sill value as a difference between nugget and an actual sill. The models are fit to a variogram and used to determine if the data is described with that particular mathematical description. The model that appears to match the variogram best is then used in further evaluations.

The Spherical model is described as follows:

$$\gamma(h) = C\,[1.5\,h - 0.5\,h^3] \text{ if } 0 \le h \le 1$$
$$\text{or } \gamma(h) = C \text{ if } h = 1$$

where C = scale or nugget + sill of the data
where h = relative range of the data points
Modified from Keckler 1994.

This equation describes data varying with distance to a polynomial formula. The model slope changes in a parabolic shape until sill is achieved. Once sill is achieved, the model levels out to a scale value, which is equal to the difference between sill and any nugget value. This sill value should be approximately equal to the total variance of the data whole set.

The exponential model is described as follows:

$$\gamma(h) = C\,[1 - e^{-h}]$$

where C = scale or nugget + sill of the data
where h = relative range of the data points
Modified from Keckler 1994.

This model tends to mimic a parabolic curve and has an apex of the curve greater than the spherical model. As with the spherical model, the model levels out to a scale value.

The Gaussian model is described below. This model, while similar to the exponential model, tends to have a shallower slope at lower ranges and a higher slope at greater ranges until leveling out to a sill value once a range is

$$\gamma(h) = C\,[1 - e^{-h^2}]$$

where C = scale or nugget + sill of the data
where h = relative range of the data points
Modified from Keckler 1994.

reached. The curve is less than the spherical model and has a more complex shape.

The linear is described as follows:

$$\gamma(h) = C\,h$$

where C = scale or nugget + sill of the data
where h = relative range of the data points
Modified from Keckler 1994.

This model is a straight line with no change in slope. This model will not reach a sill value. Data described by this model has constant variation with distance.

Kriging is a mathematical method that uses a weighted average to interpolate data values at both known and unknown points. The weighting factor for each data point is based on inverse distance between that data point and all other data points within the search area. This evaluation is performed on all data points in the data set. The weighting factor is multiplied by the variance between data point values, based on the mathematical model developed with variography. This provides variance information which is weighted by distance. Values are then calculated for all points, usually in continuous x - y coordinate pairs, which will give an estimated surface. Thus, kriging is used to calculate a surface between known data points with a model developed by variograph analysis. An output can be created that describes both quartile values and individual data points in both estimated and actual data values along with standard deviation information. This output can be compared to outputs of other models that have been developed. This statistical evaluation can also be used to compare observed verses calculated points and provides a method of determining if the model gives reasonable estimates for all the data.

Kriging can also be used to evaluate a variogram model, by using the model to estimate data values at all known points. Each point is removed in turn and determined how well the model describes that data point, this process is termed cross-validation. Cross-validation compares the computed value with actual data points and allows for an output. This output can then be statistically evaluated, by any number of methods, to determine the accuracy of a given variogram model. Cross-validation is evaluated by two general methods, an error map or x-y scatter plot. The error map is a measure of relative error in the data estimation, which is a measure of proportional error of an estimated point. This type of visual display does not allow for an absolute value to be placed on a given point. The error map is useful to determine general trends in model error along with problem areas in an estimation. An error map can also be used to compare different models for relative error that exists in each. If a model appears to be poor at only one point, a reasonable explanation could be a few bad points. If, on the other hand, it is poor in multiple areas, then the entire model needs to be re-developed. The error map is generally used to compare relative accuracy of different models in an attempt to determine if a model could be valid for the data set. Once a model is determined to characterize a data set with little relative error, an x-y scatter plot is used. The scatter plot is used as an absolute measure of actual error within the model. Scatter plot information can be output as x-y data for linear regression analysis. Linear regression allows a quantitative measure of model error. This measure of error can be evaluated by statistical methods, such as r^2, p-score or t-score. These statistical methods are then used as an indicator of evaluation significance.

Previous Work

The Pennsylvanian sediments are described in detail by Vugrinovich (1984). We have included the detailed lithologic descriptions to show the variation that is present within the Michigan Basin Pennsylvanian sediments. The formation names used are those proposed by Vugrinovich (1984). The

SYSTEM	SERIES	FORMATION	GROUP
JURASSIC	KIMMERIDGIAN	UNDIFFERENTIATED RED BEDS	
PENNSYLVANIAN	DESMOINESIAN	WINN	SAGINAW
		VERN MEMBER	
	ATOKAN		
	MORROWAN	LAKE GEORGE	
		HEMLOCK LAKE	
		SIX LAKES MEMBER	
		PARMA	
MISSISSIPPIAN	CHESTERIAN	BAYPORT	GRAND RAPIDS
	MERAMECIAN		
	OSAGEAN	MICHIGAN	

Figure 4
Simplified Stratigraphic Column from the Area of Study
Modified from Vugrinovich (1994)

individual units can have sharp lithologic changes, with some of the units grading into others. The units are also a complex conglomeration of aquifers and aquacludes which can be similar to other cyclothem deposits (Westjohn and Weaver 1996). The Parma primarily consists of a clean, well-sorted quartz sandstone with areas of localized siltstone and shaly siltstone. The Saginaw Group consists of the lower Six Lakes Member, the Hemlock Lake Formation, the Lake George Formation, and the Winn Formation with the Verne Member. The Six Lakes Member consists of a light colored micritic limestone with some areas of silty limestone and areas of anhydrite and gypsum. The Hemlock Lake Formation contains two distinct sequences, a lower and an upper. The lower sequence is thinly bedded sandstone, siltstone and shale with minor carbonate. The upper sequence is primarily shale with minor amounts of siltstone and carbonate. Coal is also present in the lower portion of the unit. The Lake George Formation is primarily well-sorted quartz sandstone. There are areas of minor fine grained sediments consisting of shale, siltstone and poorly sorted sandstone. The Winn Formation consists of shale, siltstone and sandstone sequences. The unit is predominantly a dark gray, soft, clayey shale. Some of the shale is dolomitic, hard and massive, with quartz grains and carbonaceous inclusions. The siltstone is also gray and tends to be associated with shale. The sandstone is dirty and poorly sorted, although the individual beds can be well sorted. The Verne Member of the Winn Formation is of limited extent, primarily in the eastern portions of the basin. The unit consists of calcareous black shale and black argillaceous limestone that contain brachiopod and mollusc fossils.

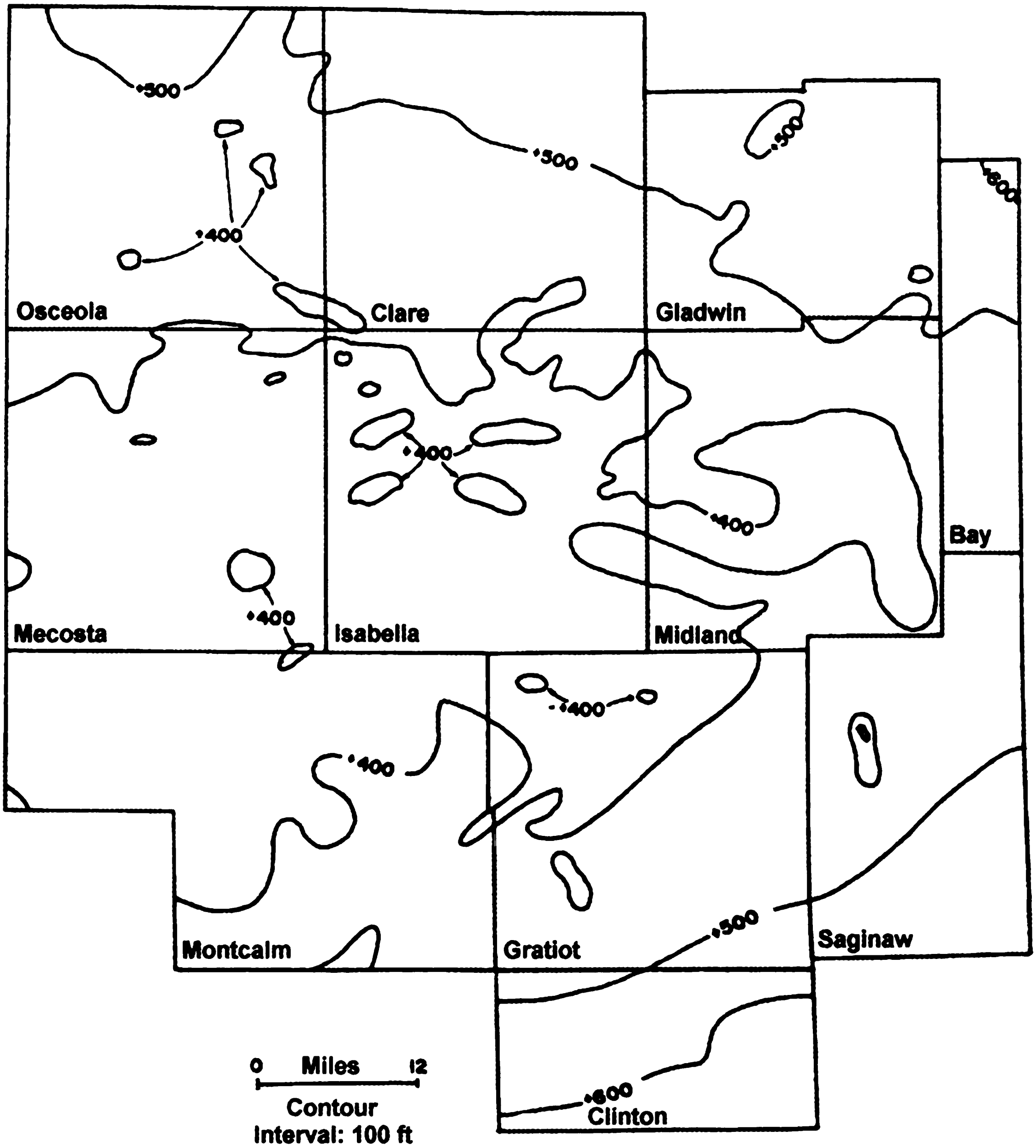

Figure 5
Structural Contours of Pennsylvanian Sediments
from Vugrinovich (1984)

Based on Lilienthal's (1978) interpretations of gamma ray logs of oil and gas wells, the base of the Pennsylvanian in northern Bay County is at 120 feet in elevation (above sea level). Northern Saginaw County logs show the

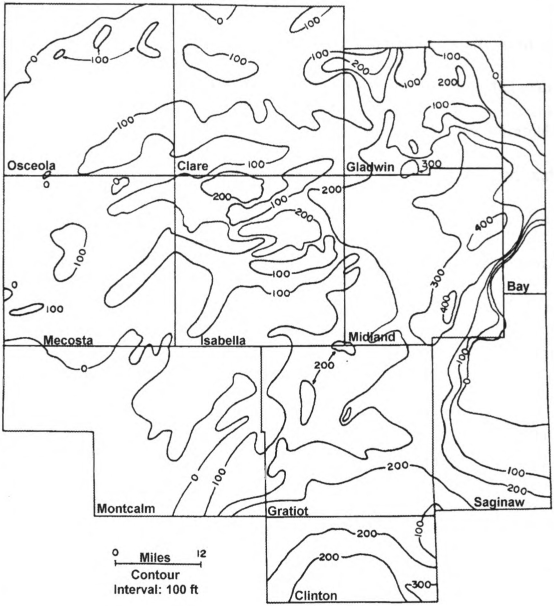

Figure 6
Isopach of Winn Formation in Central Michigan
From Vugrinovich (1984)

base of Pennsylvanian rocks lies at 95 feet above sea level. The contact with glacial drift was not located, but it can be assumed that the glacial drift lies on the Pennsylvanian rocks as noted by Dorr and Eschman (1970).

Pennsylvanian sediments in the Michigan Basin are up to 700 feet thick (Shideler and Wanless 1965). In Mid-Michigan the sediments are up to 600 feet thick (figure 5). The Winn Formation (Pennsylvanian) in Mid-Michigan is up to 200 feet thick (figure 6). Pennsylvanian sediments lie unconformably above the Mississippian Coldwater Shale and Bayport Limestone and under Jurassic Red Beds or Pleistocene glacial drift. The sediments were deposited onto an eroded surface which appears to control the sediment thickness (Shideler 1969). Shideler (1969) and Velbel et. al. (1994) described the sediments as lithologically variable with many channel scours and fills of sand and gravel, along with over-bank sand and associated shale, and coal deposits. Some of these coal beds are economically developable units (Dorr and Eschman 1970). Fine sand and shale units are highly variable and are generally localized within coarser units (Velbel et. al. 1994). Furthermore, these fine textured units tend to be deeper in the units of the Pennsylvanian System (Velbel et. al. 1994). During the Pennsylvanian Period, clastic sediments of the Saginaw Group and Parma Formation began to infill a shallow sea in the Michigan Basin from east to west (Velbel et. al. 1994, Shideler 1969, Shideler and Wanless 1965, and Newcombe 1932). This sea was relatively isolated because of many arches surrounding the basin (Shideler 1969). Marine limestone was deposited in western areas and deltaic and fluvial sand was deposited in eastern areas (Shideler 1969). Fluvial sand has been interpreted as a "coarse grained meandering" river deposit with low to intermediate braiding and fining upward sandy channel fills (Velbel et. al. 1994).

Dorr and Eschman (1970), Lilienthal (1978), Vugrinovich (1984), and Velbel et al (1994) suggested that the Pennsylvanian sediments are fluvial in origin, most likely deposited by a large river system. Vugrinovich (1984), indicated that the Pennsylvanian system started with a transgression of the sea to the west, followed by an associated regression and subsequent transgression. The Lake George Formation, which is near the middle of the system, has been interpreted as a meandering river system, based on Vugrinovich's descriptions of bed forms. Velbel et al (1994), in their work at Grand Ledge, described the bedforms as tabular-planar cross-bedded units with cross-bedding parallel to channel margins, implying lateral accretion.

The depositional environment was further determined to be a meandering river system.

Meandering fluvial environments have been studied by many authors at varying scales. Bridge et. al. (1995) traced meander scrolls in a small river system in Scotland. One of the sedimentary sequences distinguished was the fining upward bar sequence. These units consisted of tens of centimeters thick gravel beds that grade to fine crossbedded sand strata. Peat deposits commonly underlie this material. Miall (1994) distinguished lateral-accretion deposits in the multistory sand bodies that make up the Castlegate Sandstone of eastern Utah. Such deposits were also noted by Velbel et al (1994) within Pennsylvanian Strata at Grand Ledge, Michigan. In the Castlegate Sandstone, Miall determined that some of these units were hundreds of meters in length and even greater in width. Bridge et. al. (1995) determined that the channel of River South Esk, in Scotland, moved laterally over 7 meters in 18 years. This type of channel meandering could lead to the large sediment deposits that have been noted in the Michigan Basin Pennsylvanian system and Miall's description of the Castlegate Sandstone. Bridge et. al. (1995) also distinguished centimeters to decimeters thick beds within their stratigraphic sequence that were apparently deposited by seasonal flood events.

Jordan and Pryor (1992) recognized 6 levels of heterogeneity in the meandering Mississippi River from Cairo, Illinois to Memphis, Tennessee. These levels range from the scale of individual beds (level 6) to the entire river meander channel deposits (level 1). Level 1 heterogeneity is most likely to be determined at a scale represented by the MCGRIB data set. Level 1 heterogeneity is 10 - 15 miles wide and 10's of miles in length with an average thickness of 20 feet (Jordan and Pryor 1992). Sediments in this type of heterogeneity consist within meander scrolls of sand bars formed as the channel migrates laterally across clay and silt bodies that have infilled abandoned channels. This infilling of abandoned channels occurs during the periodic flood events. These are all overlain by silts and clay deposited on levees and flood plains.

Fluvial environments can include predictable and fairly homogenous lithologic units. For example, Desbarats and Bachu, (1994) predicted hydrologic transmissivity in an aquifer consisting of fluvial sandstones and associated shale aquacludes. Since these deposits can be predictable, a random and representative sampling may be used to predict the distribution of lithologies between data points. Davis et al (1993) studied hydrologic variability in the Sierra Ladrones Formation of central New Mexico in an attempt to correlate outcrop observations with core data. The ancestral river systems in the area of study were between 10 and 20 miles in width, with a 20 foot high outcrop being studied. They showed that the maximum variation in lithology of fluvial sediments, for a given lag, was perpendicular to flow-direction. The highest variation is from channel to overbank deposits and the least variation is down-flow because channel fill deposits are primarily sand bars of similar texture.

May and Schmitz (1996) showed that a coarse sand body could be distinguished from other finer textured sediments in a channel meander belt. Their characterization was performed using a sinuosity ratio, channel length to valley length and average sand body widths of the inferred stream type. Braided streams are narrower and have a sinuosity ratio of less than 1.5 (Robinson and McCabe 1997). Channel and facies type were determined with lithologic cores and interpretation of sedimentary structures in those cores. Johnson and Dreiss (1989) applied a hole effect using variography and were able to distinguish sand bodies using only lithologic information from the Santa Clara Valley of California. A horizontal variogram was used to distinguish width and a vertical variogram (down-hole) was used to distinguish thickness of a sand unit. Obviously sampling must be less than the unit width or thickness, if width and thickness are to be distinguished using this method.

Johnson and Dreiss (1989) attempted to distinguish hydrogeological sedimentary facies using an indicator variogram. An assumption in indicator evaluation with two different variables is that a 0.5 value can be used to separate different hydrologic or lithologic units. Indicators are determined from inferred permeability or lithology from borehole data by presence or absence of data. More efficient evaluations can be performed on data that has similar values, such as indicators or data within a facies. Data with a preferred orientation can be searched to exploit this preferred orientation. Matheron and de Marsily (1980), Smith and Schwartz (1980), Gelhar and Axness (1983), Fogg (1986) and Guven et. al. (1986) used data distribution, direction and orientation to determine search parameters for evauation of a particular data set. General orientation of data values, in spatial plots of data values can be used to determine primary search axis. In fluvial deposits this search evaluation is important since the widest range of grain size is from channel deposits to over-bank deposits.

Smith et. al. (1993) used multi-variable indicators to evaluate soil quality which were converted into an indicator variogram. Indicator variograms were determined to be a better indication of the soil quality relationships than the standard variogram, due to the complex relationships between soil nutrient values. This study showed that in a complex environment an indicator could be used to suit the environmental variable of interest. Vaughan et al (1995) used geostatistics to estimate the salinization of a soil in the San Joaquin Valley California. Vaughan et al (1995) used geostatistics because regression analysis did not adequately describe known points and another procedure was deemed necessary. The study determined that variography resulted in accurate data prediction, except at study area boundaries.

Figure 7
County Map of Michigan
with Bay County Highlighted

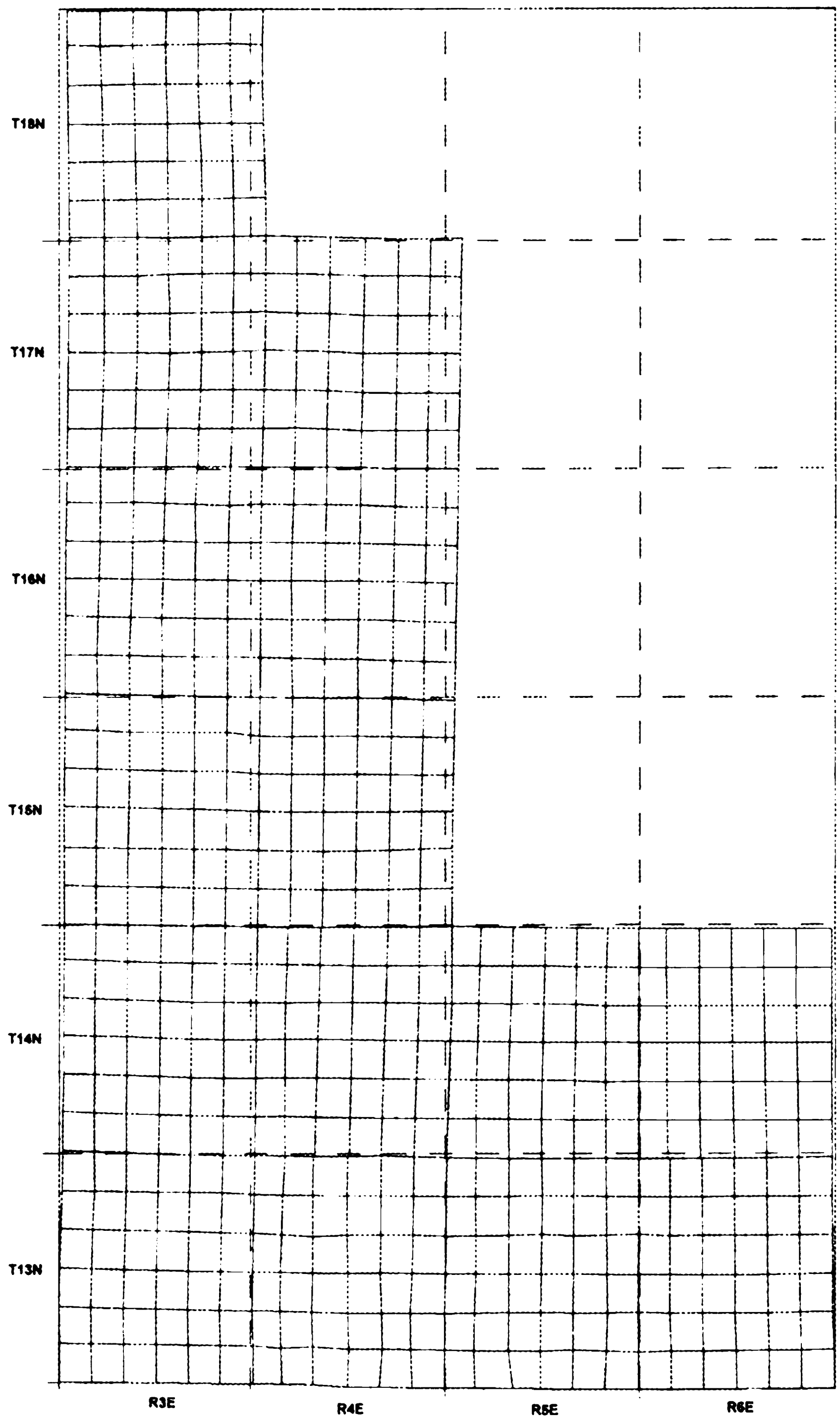

Figure 8
Township and Section Map
Bay County Michigan

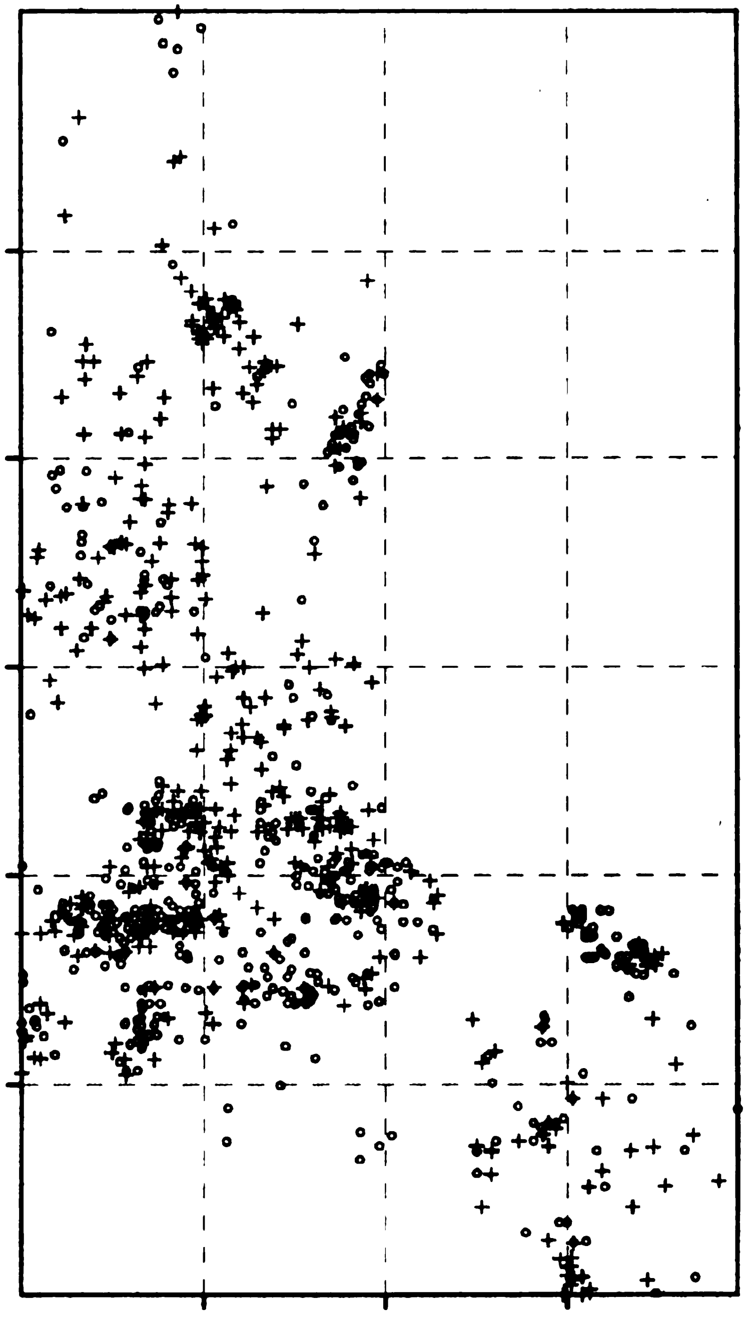

Figure 9
Location of Sandstone in Bay County
+ = Sandstone is Present o = Sandstone is Absent
420 to 400 Foot Interval

Methods

The data for this study is from the MCGRIB data set of well logs in the Michigan Basin. Monaghan and Larson (1985) edited the data by determining if the log made geologic and stratigraphic sense. For example, if a log had lithologies that are not in the region or lithologies that seemed out of sequence (e.g. rock over glacial drift), that well was eliminated. Questionable lithologic units were sometimes given a code for an unknown lithology. Lithologic codes under 50 represent glacial drift, while code values 50 or greater

```
7 13N 4E18501SWNWNW COL  192469999-9999
43.52967187926N84.05152063322W   4823645.294N    738263.425W 16
 3.750  2.230 36.912 32.179
 635.0  158.6  476.4  515.0 9999.0  120.0521920
12   5.0  630.010   45.0  585.019   11.0  574.024    2.0  572.019   47.0  525.0
10   4.0  521.019    6.0  515.052    .5  514.577    6.5  508.052    8.0  500.0
75   1.1  498.977    3.9  495.052    7.0  488.075    1.4  486.652    .6  486.0
77   6.0  480.052    .6  479.472    .3  479.175    2.5  476.672    .2  476.4

7 [county code] 13N 4E 18 [township range section] 501 [well type] SWNWNW [well location]
COL [log type (coal)]  192469999-9999  [well identification]
43.52967187926N84.05152063322W [latitude longitude]  4823645.294N 738263.425W 16 [UTM id]
 3.750  2.230 [county plane projection]  36.912 32.179  [state plane projection]
 635.0  [ground elevation] 158.6 [well depth]  476.4 [bottom elevation]  515.0 [bedrock elevation]
 9999.0 [static water level]  120.0 [drift thickness]
 52  19  20 [drift base lithology, top bedrock lithology, total number of lithologies]
        [followed by lithology, thickness and unit base elevation]
12   5.0  630.010   45.0  585.019   11.0  574.024    2.0  572.019   47.0  525.0
10   4.0  521.019    6.0  515.052    .5  514.577    6.5  508.052    8.0  500.0
75   1.1  498.977    3.9  495.052    7.0  488.075    1.4  486.652    .6  486.0
77   6.0  480.052    .6  479.472    .3  479.175    2.5  476.672    .2  476.4
```

Typical coal log with explanation.
Modified from Monaghan and Larson 1985

correspond to the bedrock lithologies. Codes used in this study are sandstone/shale (51), shale (52) and sandstone (50). Because the code 51 (sandstone/shale) was present in less than 10% of the wells and was such a vague term, it was ignored for this study.

Considerable effort was directed toward determining an area for study within the MCGRIB data set. The data was divided into smaller subsets based on individual legal townships and given short hand descriptions (table 1). Variogram evaluation of the entire Bay County data set was not possible due to software limitations. GEOEAS software can efficiently evaluate no more than one hundred data points using variography and the Bay County

Township location	Township designation	Township location	Township designation
T17N R3E	bay1	T15N R4E	bay6
T17N R4E	bay2	T14N R3E	bay7
T16N R3E	bay3	T14N R4E	bay8
T16N R4E	bay4	T14N R2E	bay9
T15N R3E	bay5	T13N R2E	bay10

Table 1
Township locations and designations in Bay County Data set

data set has well over 500 data points. GEOEAS creates a "pair comparison file" (PCF) to evaluate a data set using a variogram and it allows a maximum of 16,384 pairs for evaluation (50 to 75 points maximum). To achieve good results with GEOEAS, the area of study should have a relatively large amount of well distributed data points. Although many townships within Bay County have relatively good data distribution, there are often gaps in data coverage.

Township Location	Sections	File Designation	Wells in file
T14N R3E	25, 26, 27, 28	Bay5a.dat	53
T15N R3E	1, 2, 3, 10, 11, 12	Bay7a.dat	93
T15N R3E	10, 11, 12	Bay7e.dat	70
T15N R4E	26, 27, 28, 33, 34, 35	Bay6c.dat	25
All	All	Bay.dat	499

Township Locations, File Designations
Sections used, and Wells in File
For Areas Studied in Bay County.

Such gaps can hamper proper variogram evaluation (Isaaks and Srivastava 1989). Townships T15N R3E and T14N R3E have the largest quantity of data within the Bay County data set, but again there are gaps in data coverage. Plots of data location with respect to individual sections within townships revealed two areas; T15N R3E sections 25, 26, 35, 36 (area bay5a) and T14N R3E sections 1, 2, 3, 10, 11, 12 (area bay7a) with the most evenly distributed data. Sections 10, 11, 12 within T14N R3E (area bay7e) contained the majority of data for a smaller subset of T14N R3E sections 1, 2, 3, 10, 11, 12, so it was used as the basis for all subsequent evaluation. A final area was used for evaluation; T15N R4E sections 26, 27, 28, 33, 34, 35 (area bay6c). This area has more sandstone, but less total well logs than the larger bay7e and bay5a areas.

Probability plots indicate the data has relatively normal distribution and were used because a large quantity of zero values tends to skew results of a histogram. Distribution on a cumulative percent plot tended to be straighter with the data than with a log normal data plot, but neither plot shows a clear normal distribution (figure 10).

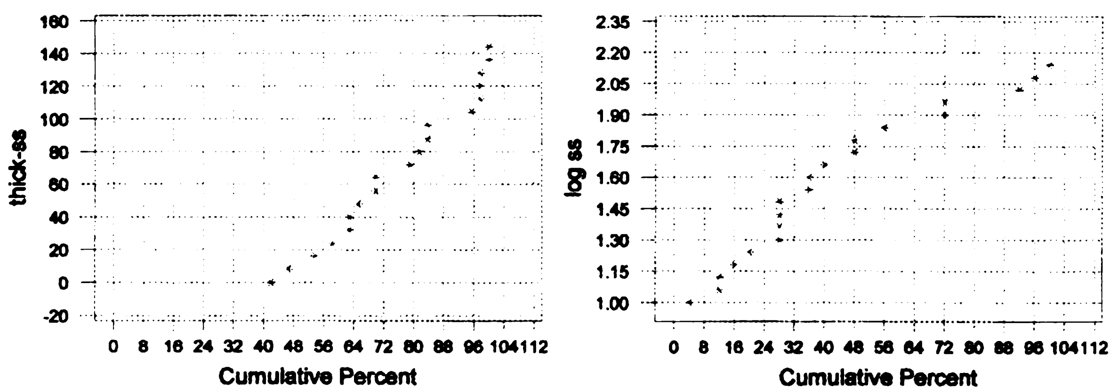

Figure 10
Cumulative percent plots for
thickness of sandstone (thick ss)
and log thickness sandstone (log ss).

Well location was also a problem with this data set. Because wells were only located to the 3rd ¼ section, multiple wells within this 40 acre parcel have the same location. A program was written to reproject raw wells and assign random location within the 40 acre parcel. This simple reprojection eliminated any identical data points.

The data was sorted into smaller, easier to manage data sets based on specific townships with all or a specific set of sections. Specific thickness intervals were chosen, then a data subset created. Wells were also sorted for type: coal boring or all rock logs. Coal logs were used in primary evaluation

Bay County T15N, R3E = bay5.dat	File identification.
12	Number of variables.
northing	Northing value.
easting	Easting value.
top rock	Elevation of bedrock surface.
top ss	Top of the first sandstone unit elevation.
thick ss	Thickness of the first sandstone unit.
ind ss	Indicator value of first sandstone unit.
top sh	Top of the first shale unit elevation.
thick sh	Thickness of the first shale unit.
ind sh	Indicator value of first shale unit.
top all	Top of the first sandstone or sandstone/shale unit elevation.
thick all	Thickness of the first sandstone or sandstone/shale unit.
ind all	Indicator value of first sandstone or sandstone/shale unit.

[last columns give a absolute location identification of the well in question]

186.079	259.378	498.0	420.0	45.0	1	400.0	0.0	0	420.0	45.0	1	15N	3E	2501COL	43.732786	84.075009
184.931	259.123	468.0	420.0	86.0	1	400.0	0.0	0	420.0	86.0	1	15N	3E	2502COL	43.731102	84.087422
180.798	260.249	535.0	433.5	29.5	1	400.0	0.0	0	433.5	29.5	1	15N	3E	5502COL	43.738464	84.132153
179.187	258.481	507.0	425.0	6.0	1	400.0	0.0	0	425.0	6.0	1	15N	3E	6501COL	43.726832	84.149540

Description of the GEO-EAS data format.

Header description top to bottom, is left to right in data fields

because they were assumed to be more reliable. The data were evaluated in four distinct smaller sub-sets mentioned above. These include bay7a, bay5a, bay6c and bay7e. The entire Bay County data set was used only in final kriging due to software limitations. Data sets were sorted for only coal logs in the sections specified, except for bay6c which was sorted for all wells to increase data quantity. The two subsets, bay7a and bay5a, generally have continuous data distribution throughout the entire area of coverage with bay7e being a subset of bay7a.dat. The two subsets, bay7a and bay5a had the best distribution with very few unrepresented areas. These two areas are also contiguous, T15N R3E sections are on the lower tier of the township and T14N R3E sections are in the upper tier of the township. This allowed data from two continuous areas, which should have similar geostatistical results, to be evaluated individually. Using R^2 of estimated thickness verses actual thickness, the results from these areas showed no significance (table #4), so bay6c was determined to be the area for final data evaluation.

Glacial influence in the study area needed to be determined to be certain that sandstone absence was not due to removal by glaciers. Based on the lowest elevation of the top of bedrock, an elevation of 420 feet above sea level was the limit of glacial erosion within Bay County data. Based on raw data evaluation, this low elevation occurred at well location 179.4434 east and 235.5822 north. Due to glacial erosion, an interval of 420 to 400 feet of elevation was chosen. Any interval above this interval would be unreliable because a sandstone unit may be absent due to glacial erosion. Conversely units below this interval would be less reliable due to lack of well quantity, which drops significantly with depth. For example, bay7 has 239 wells in the 420 to 400 foot interval while only 26 wells extend into the 360 to 340 foot interval. Bay5 begins with 74 well in the 420 to 400 foot interval and only 34 wells occur in the 360 to 340 foot interval. For good data quantity and distribution, the interval from 420 to 400 feet was used in this study. In addition this interval range allows evaluation of the level 1 heterogeneity. Jordan and Pryor (1992) described the heterogeneity at this scale to be 20' thick.

The well separation is as little as zero distance with average separation appearing to be 500 to 1000 feet. The data were kriged based on a variogram with a lag spacing of 0.110 or 0.130. A lag spacing of 1.0 has an absolute measurement of about 2000 feet, so variogram lag spacing used in this study would be about 200 feet to 250 feet. The thickness of the first sandstone unit encountered, when sorting well logs, was used as the sandstone thickness for that well location. Because mean thickness of sandstone units in the study area is 30 feet, the 420 to 400 foot interval should place the evaluation of sandstone thicknesses within the interval described as a level 1 heterogeneity. The cross-validation information was output to a data file and a linear regression analysis was performed on actual thickness verses estimated thickness.

Results

The sandstone thickness data were best fit by a gaussian model. This type of model tends to have less variation at close lags which increases as the

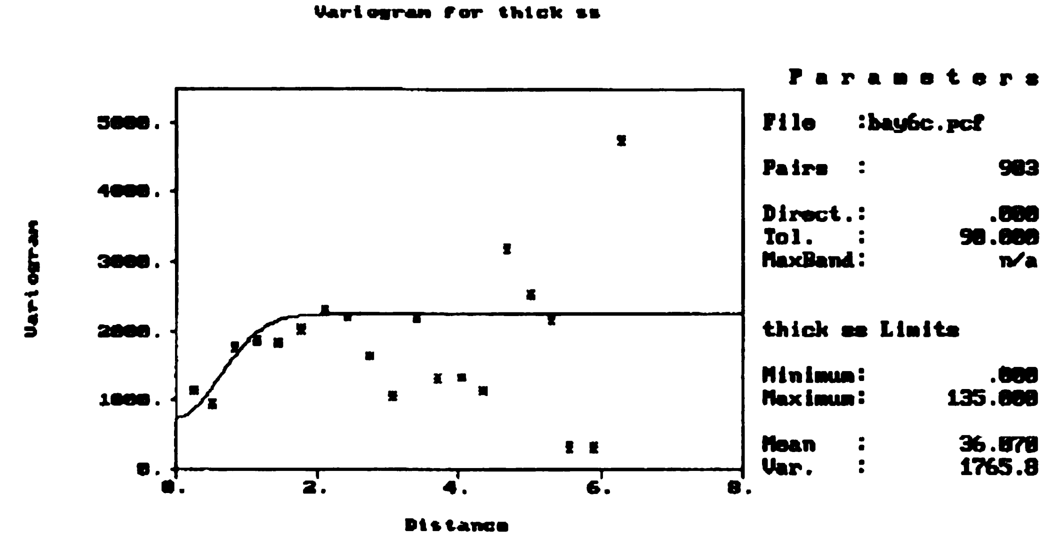

Figure 11
Variogram for area bay6c.
Gaussian Model: Nugget = 750, C - Value = 1500, Range = 1.5

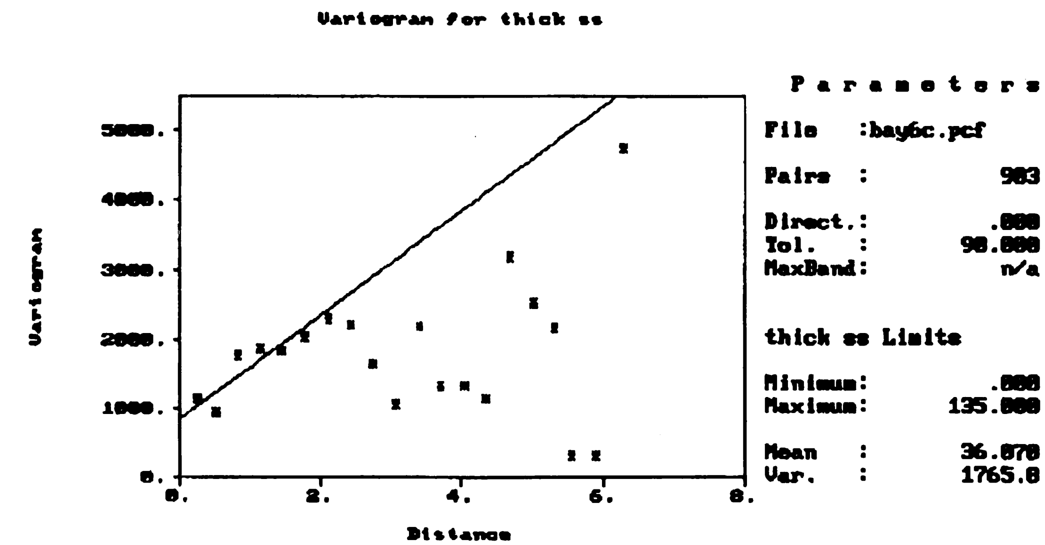

Figure 12
Variogram for area bay6c.
Linear Model: Nugget = 750, C - Value = 1500, Range = 1.5

range is approached. Although some of the data fit a linear model for short distances, a gaussian model tended to better characterize the entire data set

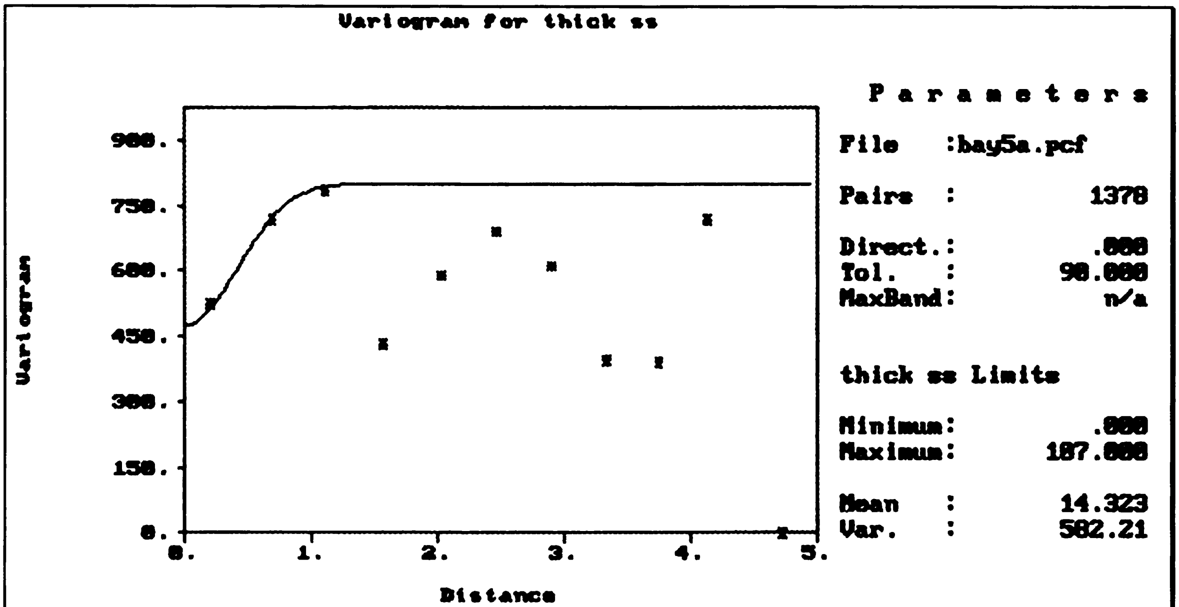

Figure 13
Variogram for area bay5a.
Gaussian Model: Nugget = 475, C - Value = 325, Range = 1

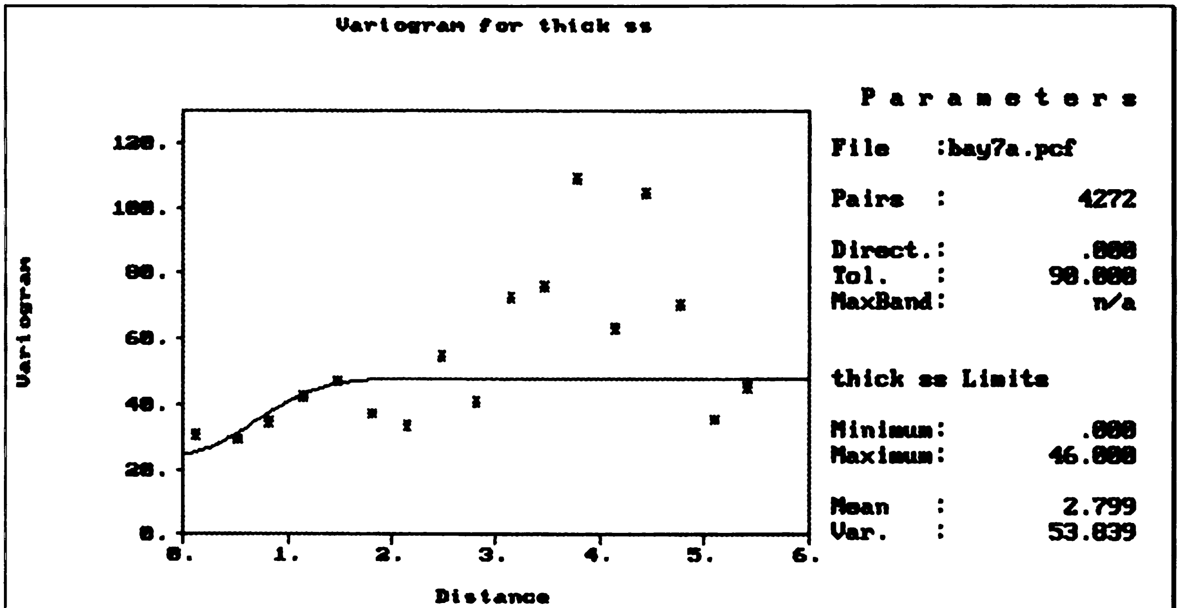

Figure 14
Variogram for area bay7a.
Gaussian Model: Nugget = 25, C - Value = 23, Range = 1.6

(figures 11 & 12). The gaussian model fits areas bay5a (figure 13),bay7a (figure 14), bay7e (figure 15), bay6c (figure 11) as well as the entire Bay County data set.

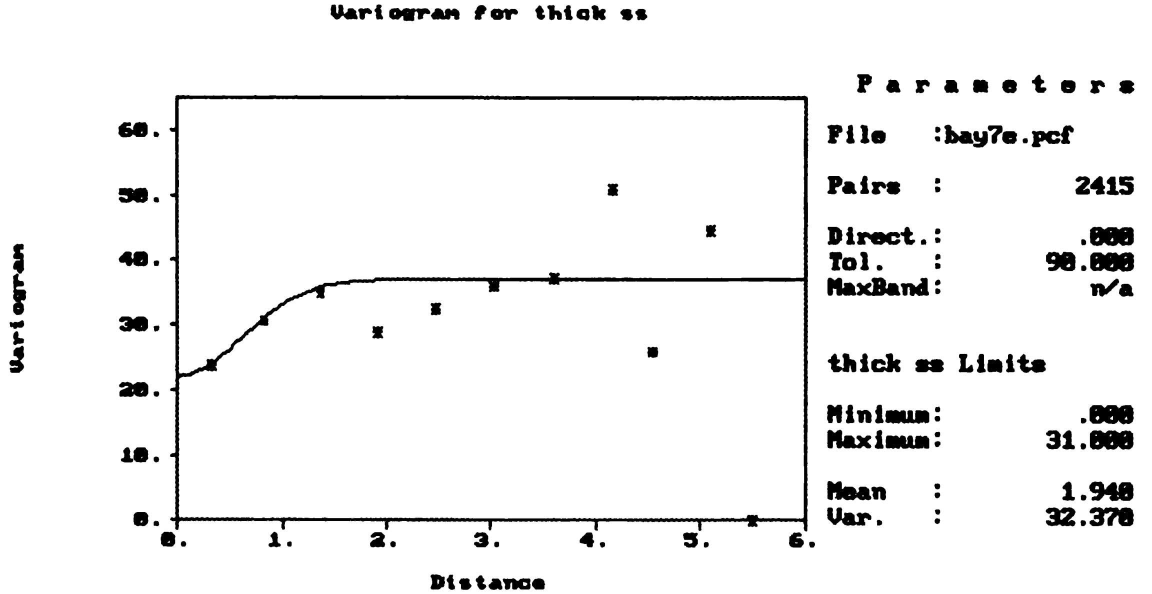

Figure 15
Variogram for area bay7e.
Gaussian Model: Nugget = 22, C - Value = 15, Range = 1.5

Error maps show many areas with relatively small differences between actual and predicted values and few areas of large relative error (figures 16 and 17). The areas with larger error are confined to small groups of highly variable data. These groups of data apparently skew the results, because they are closely spaced but also have a wide range of values. Two attempts were made to eliminate this phenomena, reprojection of data points and the use of indicator data. Change variance with respect to distance is assumed to be near immediate, with some of the data in these areas having widely variant data

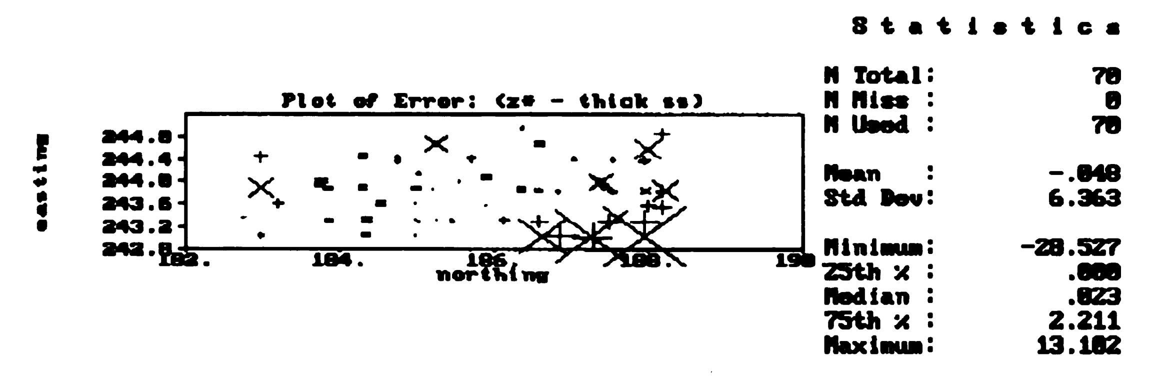

Figure 16
Error map of area bay7e
Error is proportional to symbol size

values. A gradual change in variance is needed to predict change in variance with distance. The reprojection moved data location a maximum of 100 feet. Reprojection, however, did not disperse data location enough to allow for gradual change in variance with distance.

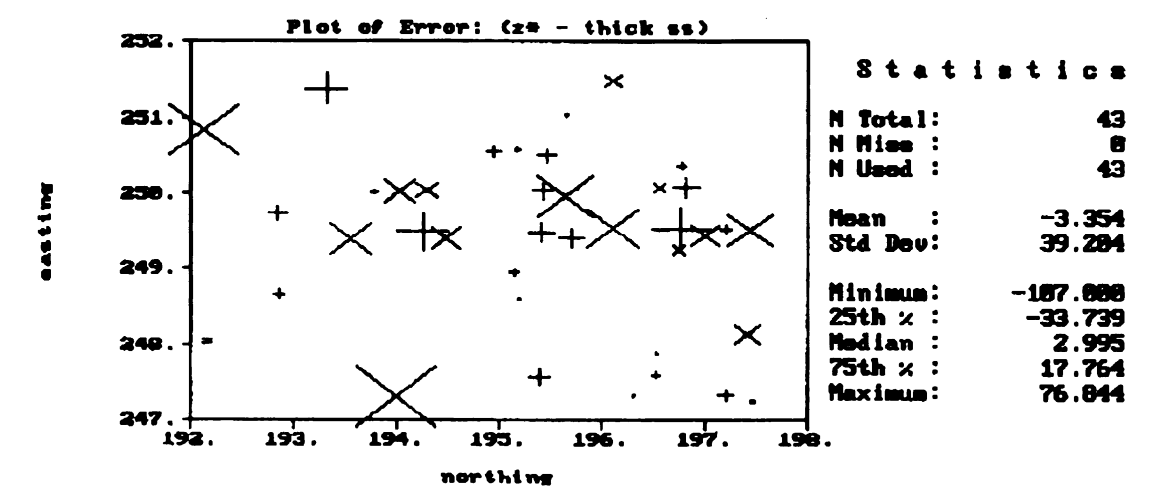

Figure 17
Error map of Area BAY6A

The use of indicator data to eliminate data variability at short lags was also not successful. Area bay7a and bay5a were evaluated using indicator data. The standard deviation for bay7a is 0.484 and for bay5a is 0.5. These

	bay5a.dat	bay7a.dat
minimum	-3.818	-3.926
25th percentile	-1.236	-1.535
50th percentile	0.000	0.000
75th percentile	0.975	1.074
maximum	4.593	4.050
standard deviation	1.842	1.819

Table 2
z-star and standard deviation of the indicator data sets (thickness ss)

standard deviation values are what would normally be derived from data that is either one or zero. This indicator evaluation is assumed not to be any better than what can be expected of randomly chosen data.

To determine if the closely spaced, highly variable, data groups were representative of the entire data set, they were evaluated individually. Three groups of data were evaluated singly and totally as a group. Each group had only 7 or 8 well logs and because of small sample size, they were difficult to evaluate. The error maps had small relative error, but regression and statistical analysis revealed poor model performance. The best linear regression evaluation had a t-test value of 0.057 and a p-score of 0.875; these values indicate no significance.

The following statistical results are for the thicknesses of sandstone units (50) in coal logs of the individual subsets. The search radius used was 1.5 lag units which is the range on the model developed from variogram analysis. The evaluation of the smaller data sets was poor using the error map information (figures 16 and 17). The raw statistics results for models used to

	bay5a.dat	bay6c.dat	bay7a.dat	bay7e.dat
data σ	24.36	42.52	7.377	5.731
indicator σ	0.50	0.50	0.484	0.463
minimum	-95.309	-107.000	-45.199	-28.527
25^{th} percentile	-5.362	-33.739	0.000	0.000
50^{th} percentile	3.718	2.995	0.102	0.023
75^{th} percentile	11.407	17.764	2.998	2.211
maximum	40.109	76.844	24.775	13.102
z-star σ	10.34	26.43	3.956	2.763

Table 3

σ = standard deviation

z-star and standard deviation of the data sets (thickness all)

evaluate area bay7a and bay7e are similar, while bay5a performed poorly (table 3), using a standard deviation test. The z-star is the estimation of data value at a given point.

Area bay6c was the only small data set to have a statistically significant correlation between calculated and observed sandstone thickness. Since the majority of estimation error does not fall within two standard deviations about the mean, statistically the evaluation is not valid (table 2). The model developed for area bay6c.dat was the only model that performed well enough in preliminary statistical information to be applied to the other data sets (table 3).

To determine if coal logs and water well rock logs are equally valid in precision, an evaluation was performed on separately sorted well logs. Two subsets of the entire county were formed, one sorted for coal logs and another sorted for all rock logs. Since only coal logs were encountered in the 420 to 400 foot interval of the smaller data sets, the coal logs to rock log accuracy could not be evaluated in the smaller data sets. The model developed for bay6c.dat was also applied to the entire Bay County data set, both coal and water well log sorts. Except for abnormally high and low maximum thickness values, the model performed similar to the small data sets (table 4).

Linear regression analysis and significance tests were performed on the kriged output from the models (figures 18, 19 and 20). This was an observed thickness (THK SS) verses calculated thickness (Z-star) analysis. The model used was that determined for area bay6c. This model was applied to all data sets, with entire Bay County data set sorted for coal logs (bay.dat) and entire Bay County Data set sorted for all rock wells (bay-w.dat). Since the average sandstone unit thickness was less that 5 feet, large estimated thickness could be suspect. In a clipped data set, values of sandstone thickness that were greater than 100 feet, both original thickness and estimated thickness, were removed. This clipping was performed in an attempt to eliminate abnormally high and low data values, since these tail values could skew results significantly. This clipped data was evaluated using the same statistical methods as the other data sets.

	R^2	N	p-score	t-score target	t-score
bay5a.dat	0.043	52	0.136	1.671	1.463
bay6c.dat	0.186	42	0.004	1.684	2.720
bay7a.dat	0.003	92	0.601	1.658	0.522
bay7e.dat	0.000	69	0.998	1.698	*
bay.dat	0.184	477	*	1.645	10.219
bay-x.dat	0.23	548	0.000	1.645	11.224
bay-w.dat	0.308	970	*	1.645	17.109
bay-w-x.dat	0.177	878	0.000	1.645	12.439

Table 4
Statistical outputs and values for data sets
* - Indicates uncalculatable values
Bay-x and Bay-w-x - calculated and thickness values over 100 eliminated
T-Score over target indicates significance
P-Score less than 0.05 indicates significance

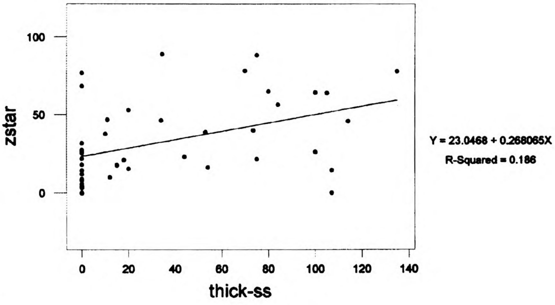

Figure 18
Linear regression plot of bay6c.dat

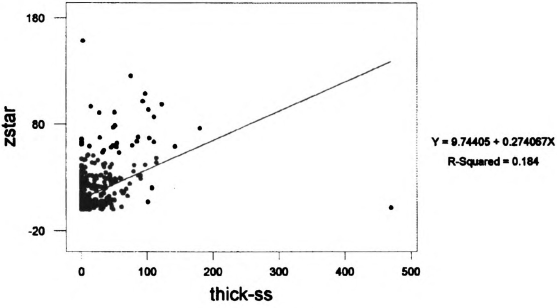

Figure 19
Linear regression plot of entire Bay County data set
Coal Logs Only

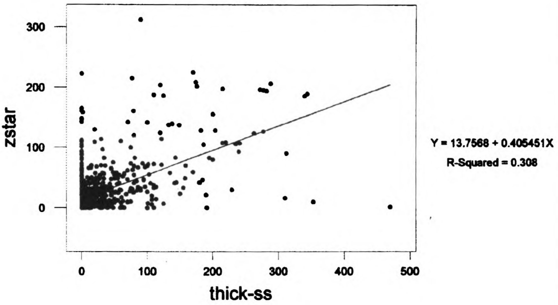

Figure 20
Regression plot of entire Bay County data set
Sorted for all rock well logs

Discussion

With respect to error maps, our geostatistical estimation was relatively accurate. Errors were spread throughout the study area, with high error concentrated around areas with close well spacing. However, this performance was not confirmed in linear regression analysis of model output.

	R^2	N	p-score	t-score target	t-score
bay5a.dat	0.043	52	0.136	1.671	1.463
bay6c.dat	0.186	42	0.004	1.684	2.720
bay7a.dat	0.003	92	0.601	1.658	0.522
bay7e.dat	0.000	69	0.998	1.698	*
bay.dat	0.184	477	*	1.645	10.219
bay-x.dat	0.23	548	0.000	1.645	11.224
bay-w.dat	0.308	970	*	1.645	17.109
bay-w-x.dat	0.177	878	0.000	1.645	12.439

Table 4
Statistical outputs and values for data sets
* - Indicates Uncalculatable values
Bay-x and Bay-w-x - calculated and thickness values over 100 eliminated
T-Score over target indicates significance
P-Score less than 0.05 indicates significance

This discrepancy can be expected, since error maps only indicate relative error of observed thickness compared to estimated thickness. The statistical evaluations (table 4) for bay6c.dat indicates some data correlation along with significance of some degree. The model, as applied to bay5a, bay7a and bay7e showed no significance, while the model applied to the entire Bay County data was significant. The larger data sets have somewhat improved performance, which is attributed to a predominance of widespread data location throughout the study area. Models developed should more closely resemble widely dispersed data, which has a change in variance relatively small with respect to distance. The closely spaced data has a relatively high change in variance with respect to distance. These different data variance types could distort the variogram and assign greater importance to data variance that are closer to average thickness value.

Data point locations could explain poor model estimation. For example, some groups of data points had zero distance separation and some of these data points had widely variable thickness values. Reprojection of data points, in an attempt to move wells at zero distance in order to allow for more predictable change in variance with distance, was insufficient to allow thickness changes to be predictable. Even though well locations were displaced, much of the widespread data had less variation with distance than that of zero distance groups. This fact was indicated by the large nugget value in the model. Geostatistical methods can not efficiently evaluate relatively high variation with respect to distance and relatively low variation with respect to distance. To achieve a proper estimation, the change in variance must be predictable with respect to distance. If one of these data variance types does not dominate, it will be difficult to model any change in variation with distance because mathematical models will predict higher variation than is true for the data set.

Well location estimation could explain some model variance. Wells are spaced at irregular intervals, with an average distance of 500 to 1000 feet, with some of the wells located at zero distance. This spacing should be sufficient to allow evaluation of the presumed channel width of 10's of miles. Reprojection of well locations should not alter the evaluation significantly, due to accuracy of original well location. Wells were located within a 10 acre parcel of land, which has a width and length of 660 feet. Location information as determined for each well within a given parcel was the center of that parcel. If a well was near a corner of a given parcel and its location was noted as the center of the parcel, an error of up to 450 feet could have been introduced to the evaluation. Well reprojection moved well location up to 100 feet, which was enough to eliminate wells located at identical locations, but not enough to alter error introduced in the original location.

A basic assumption of this study was that data contained in the logs is accurate, and geological logs (oil and petroleum), being highly detailed, were assumed to be the most accurate. Evaluations performed on coal logs only and compared to those for all logs, using the same model on both data sets, were nearly identical. Since identical models performed similarly on each data set, the geologic and water well logs can be assumed to be equal in their accuracy. Monaghan and Larson (1985) reported that many logs were eliminated due to lithology data that was vague or inaccurate. These include bedrock units over glacial sediments, rock descriptions that were uncertain, and wells described by persons that tended to use the aforementioned errors. Vertical control on wells was not always accurate. In many cases vertical elevation used was estimated from a USGS quadrangle map for the area, this estimation could have introduced ±10 feet of vertical error (Monaghan and Larson 1985). Depending on sandstone thickness, this vertical error, along with any error in the depth to unit, could miss a sandstone body in a given interval.

Error in lithologic information could have also been introduced to a log. For example, some drillers noted a sandstone / shale lithology. It is uncertain what this lithologic description is; interbedded sandstone and shale is likely but it could also be a "dirty" sandstone. Bridge et al (1995) were able to discern thin, fine-textured beds within thicker sandy units, which were interpreted as flood events. These could be the origin of the sandstone / shale units. Some logs contain massive thickness of sand, up to 450 feet. Based on

Robinson and McCabe (1997) work, on channel width to sandstone thickness associations, a sandstone thickness or 450 feet will not correspond to a channel width. An explanation for these unusual thicknesses could be a "lumping" factor, which would be ignoring thin units which are either destroyed or mixed with sandstone cuttings in a given log. This could explain the presence of abnormally thick units near thinner units, units greater than 100 feet thick. These abnormally thick units are assumed to be errors and should be ignored.

Poor model accuracy could also be explained by assumptions about depositional environment. Depositional environment throughout the study

TOWNSHIP LOCATION	420-400 ft INTERVAL MEDIAN THICKNESS	400-380 ft INTERVAL MEDIAN THICKNESS
17N 3E	85 feet	106.5 feet
17N 4E	35.5 feet	20.5 feet
16N 3E	10.5 feet	9.5 feet
16N 4E	0 feet	0 feet
15N 3E	4.5 feet	1.9 feet
15N 4E	20 feet	20 feet
14N 3E	0 feet	0 feet
14N 4E	0 feet	0 feet
14N 5E	0 feet	2.5 feet
13N 4E	0 feet	0 feet
14N 5E	0 feet	0 feet

Table 5

Median sandstone thicknesses for townships in Bay County

area was assumed to be a broad meandering river valley, that should have one or more wide channels. Based on Robinson and McCabe (1997), in which a channel depth is proportional to depth, there was not enough sandstone thickness in southern Bay County to correspond to wide channels. Studying median thicknesses and a plot of indicator data shows no apparent patterning to sandstone thicknesses (figures 9 and table 5). Robinson and McCabe (1997) showed that a meandering river with a depth as shallow as 1 meter (3 feet) should have a width of at least 70 meters (230 feet). This relationship is linear on a log thickness vs. log width plot. With sandstone thickness averaging 20 to 30 feet, a river channel should have a width of 10 to 15 miles. There are no areas, in southern Bay County, that have continuous sands of these dimensions.

These are all based on the assumption of continuous channeling and a meandering river system deposit similar to those studied by Robinson and McCabe (1997) and Jordan and Pryor (1992). With a wide channel that was abandoned quickly, it is possible that a majority of the former channel could have been filled with fine textured over-bank deposits. This would be similar to the level three heterogeneity from Jordan and Pryor (1992). This level three heterogeneity consists of individual channel point-bar and channel-splay sand bodies. There are also associated thin sheets and lenses of low-permeability muds. This deposit is contained within the river channel itself and can be up to one mile in width, 2 miles in length and 100 feet in thickness. There appears to be more sand indicators in the northern area of Bay County than in southern portions (figure 9). It is possible that a wide river channel moved north, isolating wide channels, which were then filled with fine textured sediments. These channels later became the thick shales we see today.

Conclusion

When the MCGRIB data base was used to estimate Pennsylvanian sandstone distribution, we were able to show that coal logs are as imprecise as those that describe water wells. This was shown through model performance in wells sorted for coal logs and wells sorted for all rock logs. Geostatistical estimation performance on average was poor, with some correlation between observed and estimated values. The Pennsylvanian sediments in the interval 420 - 400 feet of elevation appear to be part of a meandering river channel. This is similar to what has been reported in previous work, this area being a broad meandering river valley. Due to poor coverage throughout Bay County, data being extensive but not as continuous as is generally required for geostatistical evaluation, sandstone data could not be evaluated with a high degree of certainty. Although, we were able to develop a model that could be used to determine depth to sandstone aquifer with some level of confidence, there are too many areas of error in the data set to allow for this data to be used at a localized level. The MCGRIB data set, although extensive and somewhat detailed, is useful in determining general, county wide trends. The data set was found to be only marginally useful for specific areas.

Works Cited

Bridge J S, Alexander J, Collier R E, Gawthorps R L and Jarvis J, (1995) Ground-Penetration Radar and Coring Used to Study the Large-Scale Structure of Point-Bar Deposits in Three Dimensions, Sedimentology, 42, 839 - 852.

Davis J M, Lohmann R C, Phillips F M, Wilson J L and Love D W, (1993) Architecture of the Sierra Ladrones Formation, Central New Mexico: Depositional Controls of the Permeability Correlation Structure, Geological Society of America Bulletin, 105, 998 - 1007.

Desbarats A J and Bachu S, (1994) Geostatistical Analysis of Aquifer Heterogeneity for the Core Scale to the Basin Scale: a Case Study, Water Resources Research, 30 (3), 673 - 684.

Dorr J A and Eschman D F, (1970) Geology of Michigan: University of Michigan Press, 476.

Englund E and Sparks A, (1991) GEO-EAS, Computer Oriented Geological Society, Computer Software.

Grant C W, Goggin D J, and Harris P M, (1994) Outcrop Analog for Cyclic-Shelf Reservoirs, San Andres Formation of Permian Basin: Stratigraphic Framework, Permeability Distribution, Geostatistics, and Fluid-Flow Modeling: AAPG Bulletin, 78 (1), 23-54.

Isaaks E H and Srivastava R M, (1989) An Introduction to Applied Geostatistics: Oxford University Press, 561.

Johnson N M and Dreiss S J, (1989) Hydrostratigraphic Interpretation Using Indicator Geostatistics, Water Resources Research, 25 (12), 2501 - 2510.

Jordan D W and Pryor W A, (1992) Hierarchical Levels of Heterogeneity in a Mississippi River Meander Belt and Application to Reservoir Systems: The American Association of Petroleum Geologists Bulletin, 76 (10), 1601 - 1624.

Keckler D, (1994) Surfer® for Windows, Golden Software, Inc., Golden Colorado, USA.

Lilienthal R T, (1978) Stratigraphic Cross-Sections of the Michigan Basin, Michigan Department of Natural Resources, Geologic Survey Division, Report of Investigation #19.

Liu K, Boult P, Painter S and Paterson L, (1996) Outcrop Analog for Sandy Braided Stream Reservoirs: Permeability Patterns in the Triassic Hawkesbury Sandstone, Sidney Basin, Australia, AAPG Bulletin, 80 (12), 1850 - 1866.

May J H and Schmitz D W, (1996) Development of a predictive model for defining subsurface sand bodies: Engineering Geology, 42, 175 - 186.

Miall A D, (1994) Reconstructing Fluvial Macroform Architecture from Two-Dimensional Outcrops: Examples from the Castlegate Sandstone, Book Cliffs, Utah, Journal of Sedimentary Research, B64 (2), 146-158.

Miall A D, (1996) The Geology of Fluvial Deposits, Sedimentary Facies, Basin Analysis, and Petroleum Geology, Springer and Company, New York, New York.

Monaghan G W and Larson G L, (1985) A Computerized Ground-Water Resources Information System, Ground Water, 23 (2), 233 - 239.

Monaghan G W (1996), SET4TIM, MUDLAB software, Computer Software.

Newcombe R B, (1932) Oil and Gas Fields of Michigan, A discussion of depositional and structural features of the Michigan Basin, State of Michigan Department of Conservation, Geological Survey Division, Publication #38.

Olea R A, (1996) Fundamentals of Semivariogram Estimation, Modeling, and Usage, *in* Stochastic Modeling and Geostatistics, Principles, Methods, and Case studies, edited by J M Yarns and R L Chambers, AAPG Computer Applications in Geology #3, American Association of Petroleum Geologists, Tulsa OK, USA, 27-35.

Pan G, (1994) Restricted Kriging: A Link Between Sample Value and Sample Configuration, Mathematical Geology, 26 (1), 135 - 155.

Robinson J W, and McCabe P J, (1997) Sandstone-Body and Shale-Body Dimensions in a Braided Fluvial System: Salt Wash Sandstone Member (Morrison Formation), Garfield County, Utah, AAPG Bulletin, 81 (8), 1267-1291.

Shideler G L, (1969) Dispersal Patterns of Pennsylvanian Sandstones in the Michigan Basin, Journal of Sedimentary Petrology, 39 (3), 1229-1237.

Shideler G L, and Wanless H R, (1965) Pennsylvanian sediments of the Michigan Coal Basin, Geological Society of America, Special Paper, 153-154.

Smith J L, Halvorson J J and Papendick R I, (1993) Using Multiple-Variable Indicator Kriging for Evaluating Soil Quality, Soil Science Society of America Journal, 57, 743 - 749.

Solow A R, (1993) On the Efficiency of the Indicator Approach in Geostatistics, Mathematical Geology, 25 (1), 53 - 57.

Vaughan P J, Lesch S M, Corwin D L, and Cone D G, (1995) Water content effect on soil salinity prediction: A geostatistical study using co-kriging: Soil Science Society of America Journal, 59, 1146 - 1156.

Velbel M A, Price J R, and Brandt D S, (1994) Sedimentology, Paleogeography, and Geochemical Weathering of the Pennsylvanian Strata of Grand Ledge, Michigan, Eastern Section AAPG Annual Meeting Field Trip and Great Lakes Section - SEPM 24th annual Fall Field Conference.

Vugrinovich R, (1984) Lithostratigraphy and Depositional Environments of the Pennsylvanian rocks and the Bayport Formation of the Michigan Basin, Michigan Department of Natural Resources, Geologic Survey Division, Report of Investigation #27.

Westjohn D B and Weaver T L, (1996) Hydrogeologic Framework of Pennsylvanian and Late Mississippian Rocks in the Central Lower Peninsula of Michigan, U. S. Geological Survey, Water-Resources Investigation Report 94-4107.

Wolf D J, Withers K D, and Burnaman M D, (1996) Integration of Well and Seismic Data Using Geostatistics, *in* Stochastic Modeling and Geostatistics, Principles, Methods, and Case studies, edited by J M Yarns and R L Chambers, AAPG Computer Applications in Geology #3, American Association of Petroleum Geologists, Tulsa OK, USA, 177-199.