A Short Note on the Proportional Effect and Direct Sequential Simulation

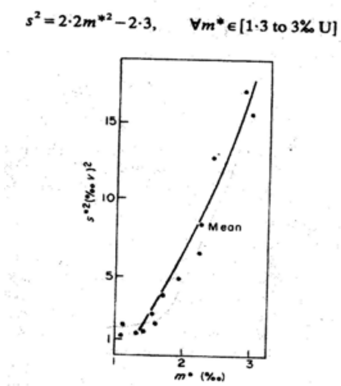
B. Oz (boz@ualberta.ca) and C. V. Deutsch (cdeutsch@ualberta.ca)
University of Alberta, Edmonton, Alberta, CANADA

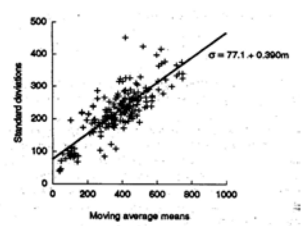
Abstract

Direct sequential simulation (DSS) is receiving attention as a procedure to simultaneously incorporate multiscale data correctly accounting for averaging. An implicit assumption in DSS is that the kriging variance provides the variance of local distributions of uncertainty. The kriging variance, however, depends only on the data configuration and is independent of the data values. This is reasonable for data that follow the congenial Gaussian distribution, but it is completely unreasonable for real earth science data in original units. Real data show a proportional effect, that is, increased variability in high-valued areas. We show this effect with a number of datasets and present some ideas to account for it in DSS.

Introduction

The variability of earth science data depends on the magnitude of the variable. This dependence is called "heteroscedasticity" or the "proportional effect." David, Journel and Huijbregts and others carefully documented the proportional effect in the early days of geostatistics. Here are two representative examples from *Mining Geostatistics* and *An Introduction to Applied Geostatistics*:





These examples and many other published examples show a quadratic relationship between the variance and the mean (a linear relationship between the standard deviation and the mean). In fact, a characteristic of the multivariate lognormal distribution is an analytical quadratic relationship between the variance and mean.

More correctly the proportional effect was described in the context of the variogram being different in different areas. In practice, it was observed that the variograms could be made equal by dividing by a function of the experimental mean (II.35 in *Mining Geostatistics*):

$$\frac{\gamma(\mathbf{h},u_0)}{f(m(u_0))} \approx \frac{\gamma(\mathbf{h},u_0^{'})}{f(m(u_0^{'}))} \tag{1}$$

The two variograms $\gamma(\mathbf{h},u_0)$ and $\gamma(\mathbf{h},u_0^{'})$ are said to differ from each other by a *proportional effect*. The proportional effect function *f* must be determined from the data. As mentioned, it is often quadratic. This assumes that there is a stationary variogram model $\gamma_0(\mathbf{h})$ that is independent of location u_0 such that:

$$\gamma(\mathbf{h},u_0) = f\left(m^*(u_0)\right)\cdot\gamma_0(\mathbf{h}) \tag{2}$$

This common observation led to the use of relative variograms (see GSLIB Chapter 3) where the experimental variogram is divided by the square of the mean.

Of course, kriging weights only depend on the shape of the variogram. Multiplying the variogram locally be a function of the local mean does not change the kriging weights. Only the kriging variance will change. The kriging variance by itself is not particularly useful. Practice evolved to inference of robust variograms such as relative variograms, correlograms, or using logarithms of the data.

Modern geostatistics is more concerned with simulation techniques that allow modeling heterogeneity than with straight kriging that allows optimal estimation. Indicator simulation is insensitive to the proportional effect because of the 0/1 coding. Gaussian techniques are also immune to the proportional effect; the normal transformation effectively removes the proportional effect. Theoretically, lognormal data with a quadratic proportional effect would have none after normal scores transformation.

This discussion is very relevant because of the increasing consideration of direct simulation methods. These methods **must** consider the proportional effect to provide reasonable measures of local uncertainty. Isaaks and Srivastava make a point of this in their *Assessing Uncertainty* Chapter (p 522-3):

> *The improvement in the results [when using the proportional effect] is quite remarkable. If a proportional effect exists, it must be taken into account when assessing local uncertainty. By rescaling the variogram to a sill of one and locally correcting the relative kriging variance, on can build confidence intervals that reflect local conditions.*

We show some examples of the proportional effect and discuss how we are going to account for it in the DSS formalism.

Some Examples

A scatterplot of local means and local standard deviations from moving window calculations is a good way to check for the proportional effect. We explored the proportional effect on a number of different data sets. For all data sets, depending on the available data, the spatial domain was divided into even rectangular or square grid blocks. Then, summary statistics were calculated for non-overlapping windows. The summary statistics were calculated for the original data units and also normal score transforms. The relationship between the local mean and standard deviation values illustrates the proportional effect.

Figure 1 shows the results for the Walker lake data set. The original digital elavation model was processed to produce a data set consisting of three variables measured at each 78,000 points on a 260x300 rectangular grid. The first variable was considered in 5 by 5 non-overlapping grid blocks. The pronounced proportional effect in original units is significantly reduced by normal score transformation.

Figure 2 shows the results for the true data set from GSLIB. Note the characteristic behavior. Figure 3 shows the results for the Dallas (Pb) data set. The size of the field is 11702 m by 10716 m with 180 measurements. 1000 m by 1000 m non-overlapping grid blocks were considered with a minimum of 4 data per grid block. Figure 4 shows the results for some North Sea porosity data – one well is shown at the top. Figure 5 shows the results for the corresponding permeability data. Figure 6 shows the results for some porosity data from a Texas reservoir. Figure 7 shows the results for the corresponding permeability data.

These data were not chosen to illustrate the proportional effect. They just happened to be easily accessible datasets. Other data were also considered. The results were similar in all cases. Even for variables like porosity that are *nearly* normal, there is a significant proportional effect.

Consequences

These examples illustrate the dependence of the variance on the local mean. A linear relationship was systematically observed on the standard deviation versus mean plot, which means that the variance-mean relationship is quadratic. Everything else being equal, the local variance will be four times larger between areas where the mean doubles. The "stationary" kriging variance based on a single variogram is ***not a good measure of local variability***. The kriging variance works okay after Gaussian transformation, but the central idea of DSS is precisely to avoid that transformation. We must account for the proportional effect in DSS.

Proposed Approach

Our idea is to use the same approach proposed by the pioneers of geostatistics: use a standardized variogram, calculate the standardized kriging variance $\sigma_K^2(u)$ and the rescale that variance to a local measure of variability $\sigma_{QS}^2(u)$, that is,

$$\sigma_{QS}^2(u) = f\left(m^*(u_0)\right) \cdot \sigma_K^2(u) \tag{3}$$

This requires two additional steps (1) fitting the proportional effect *f(m)*, and (2) calculating the local mean at each location $m^*(u)$. The proportional effect can be fitted with regression from plots like those shown in the center of figures 1 through 7. The local mean can be calculated by a number of methods: kriging with a large search radius or moving window averages. These calculations and other implementation details will need to be resolved.

Inconsistencies in fitting the proportional effect and local means will result in inaccurate and/or imprecise distributions of local uncertainty and lack of reproduction in the global histogram. Cross validation and checking the local distributions of uncertainty (Deutsch, 1996) will quickly reveal problems and help determine the optimal approach.

Final Comment

Application of DSS without explicit accounting for the proportional effect will certainly lead to incorrect distributions of uncertainty. The fact that DSS will reproduce the covariance is not by itself sufficient to justify using the method. We also require the global distribution to be reproduced within statistical/ergodic fluctuations and that local distributions of uncertainty fairly represent our state of incomplete knowledge.

References

- M. David (1988). Handbook of Applied Advanced Geostatistical Ore Reserve Estimation, Elsevier, Amsterdam
- Deutsch, C.V., "Direct Assessment of Local Accuracy and Precision," Geostatistics Wollongong '96, Vol 1, Baafi and Schofield, editors, Kluwer Academic Publishers, September 1996, pp 115-125.
- Deutsch, C.V. and Journel A.G. (1998). *GSLIB: Geostatistical Software Library: and User's Guide*, Oxford University Press, New York, 2nd Ed.
- Isaaks, E. H., and Srivastava, R. M. (1989). *An Introduction to Applied Geostatistics*, Oxford University Press, New York.
- Journel, A.G. and Huijbregts, Ch.J. (1978). *Mining Geostatistics*, Academic Press, New York.

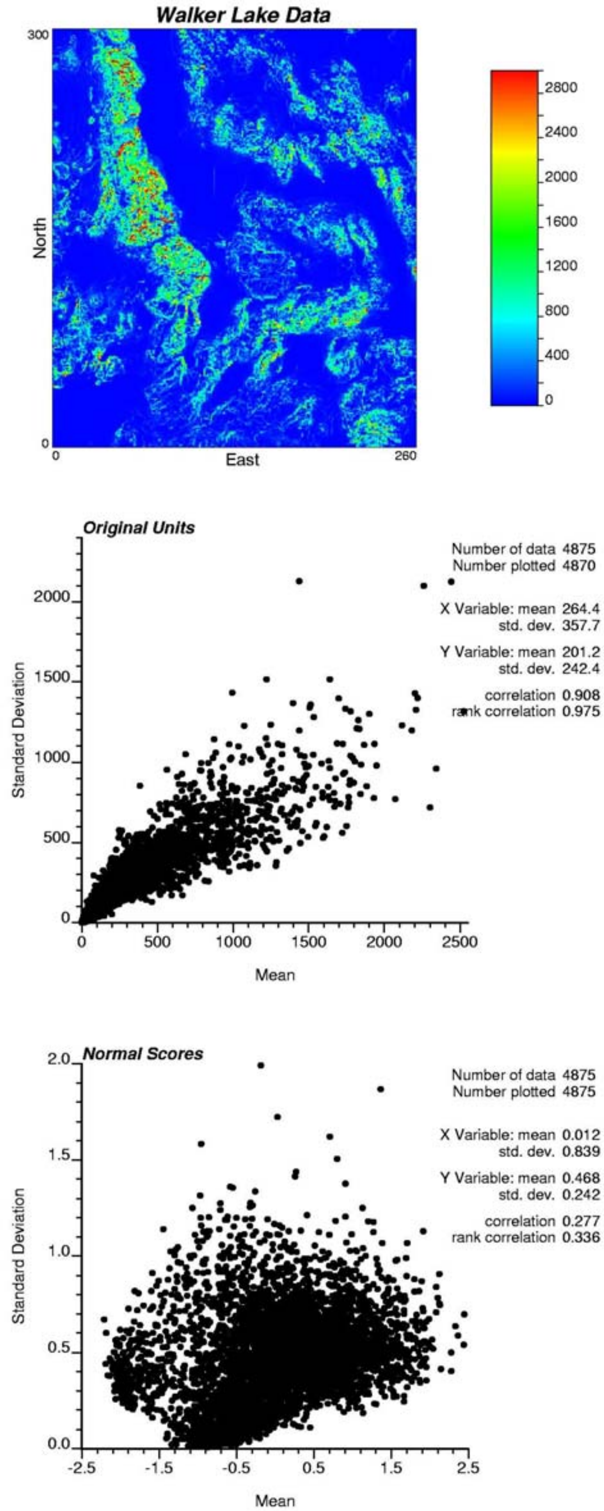


Figure 1: Check for the proportional effect with the Walker Lake data. The color scale plot shows the data distribution, the central plot shows the relationship between the mean and standard deviation in original data units, and the bottom plot shows the relationship after normal score transformation.

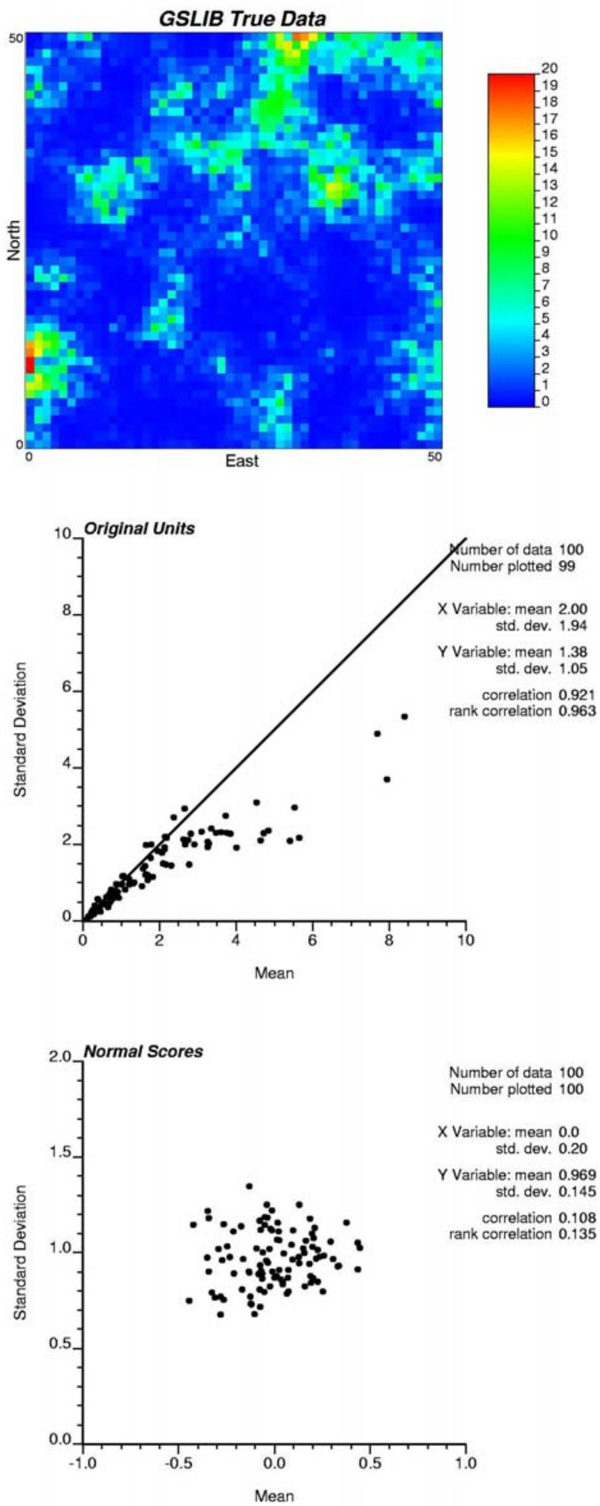


Figure 2: Check for the proportional effect with the GSLIB True data. The color scale plot shows the data distribution, the central plot shows the relationship between the mean and standard deviation in original data units, and the bottom plot shows the relationship after normal score transformation.

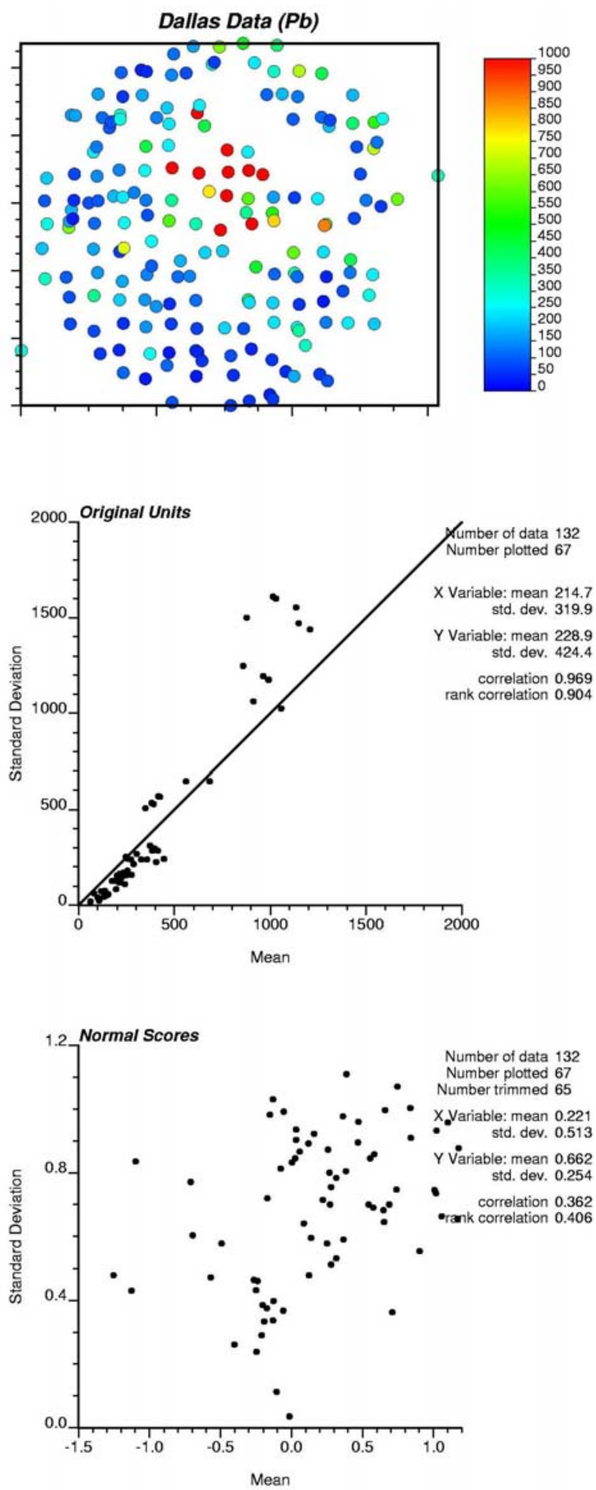


Figure 3: Check for the proportional effect with the Dallas Pb data. The color scale plot shows the data distribution, the central plot shows the relationship between the mean and standard deviation in original data units, and the bottom plot shows the relationship after normal score transformation.

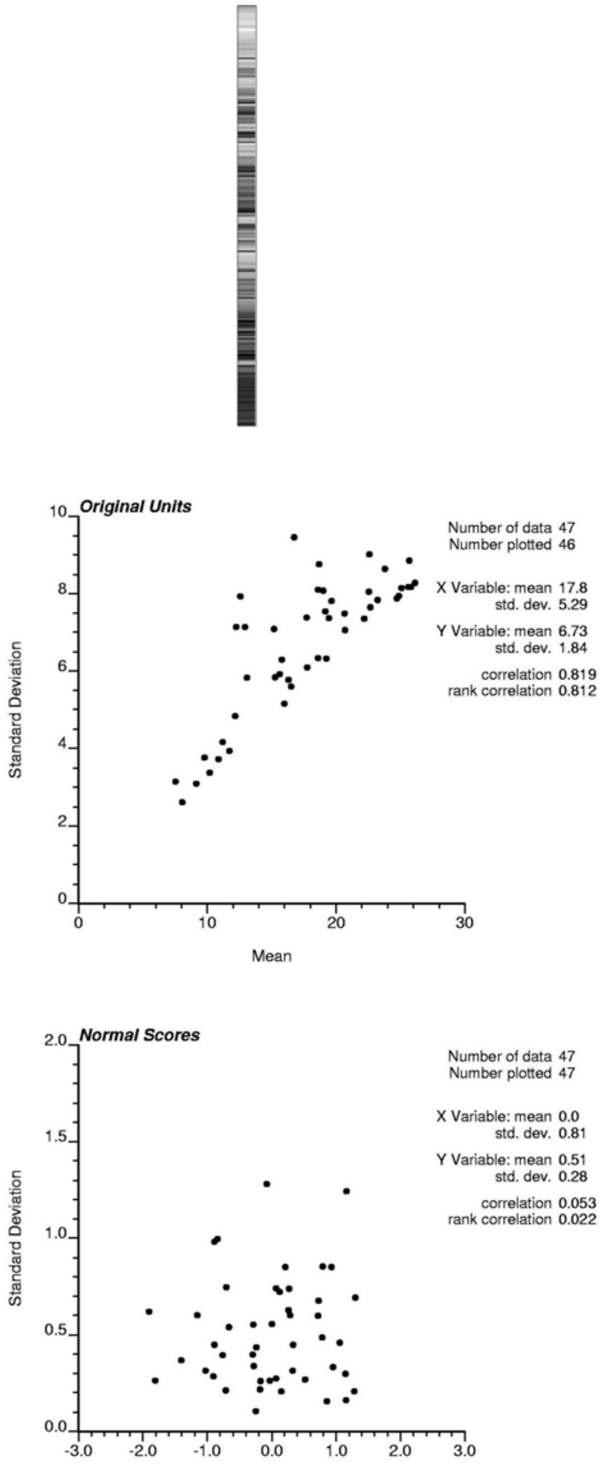


Figure 4: Check for the proportional effect with some North Sea porosity data. The top plot shows one well, the central plot shows the relationship between the mean and standard deviation in original data units, and the bottom plot shows the relationship after normal score transformation.

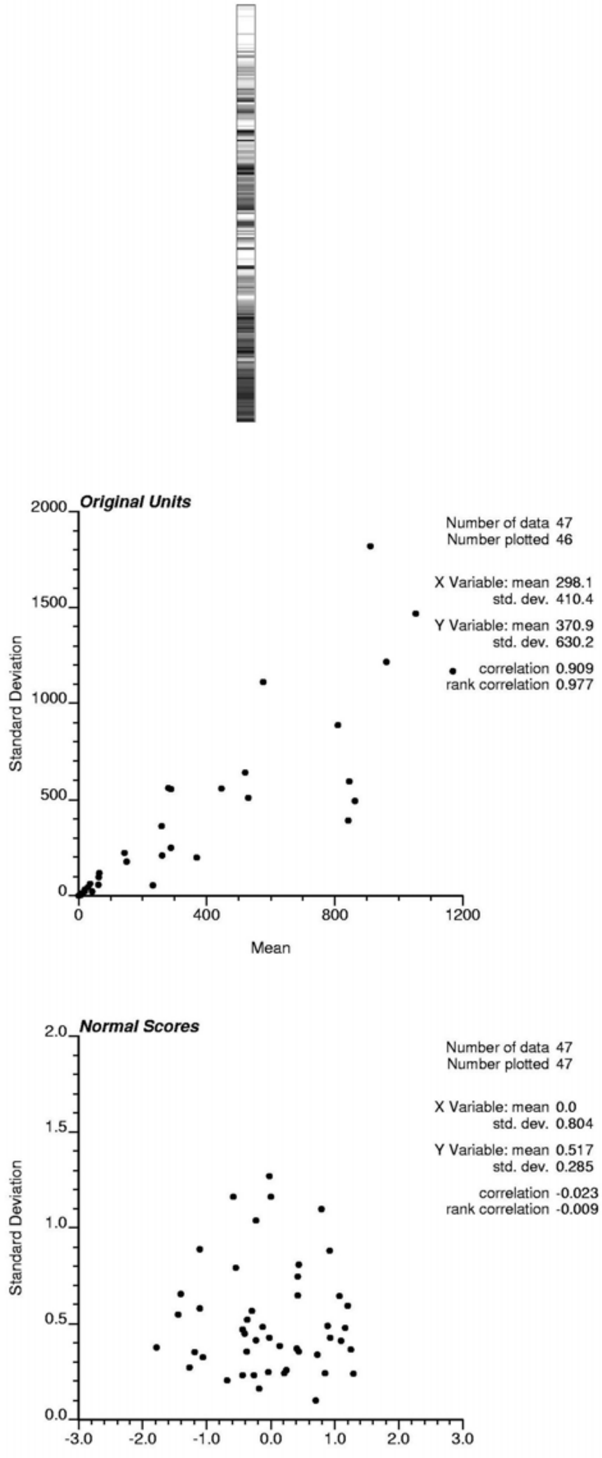


Figure 5: Check for the proportional effect with some North Sea permeability data. The top plot shows one well, the central plot shows the relationship between the mean and standard deviation in original data units, and the bottom plot shows the relationship after normal score transformation.

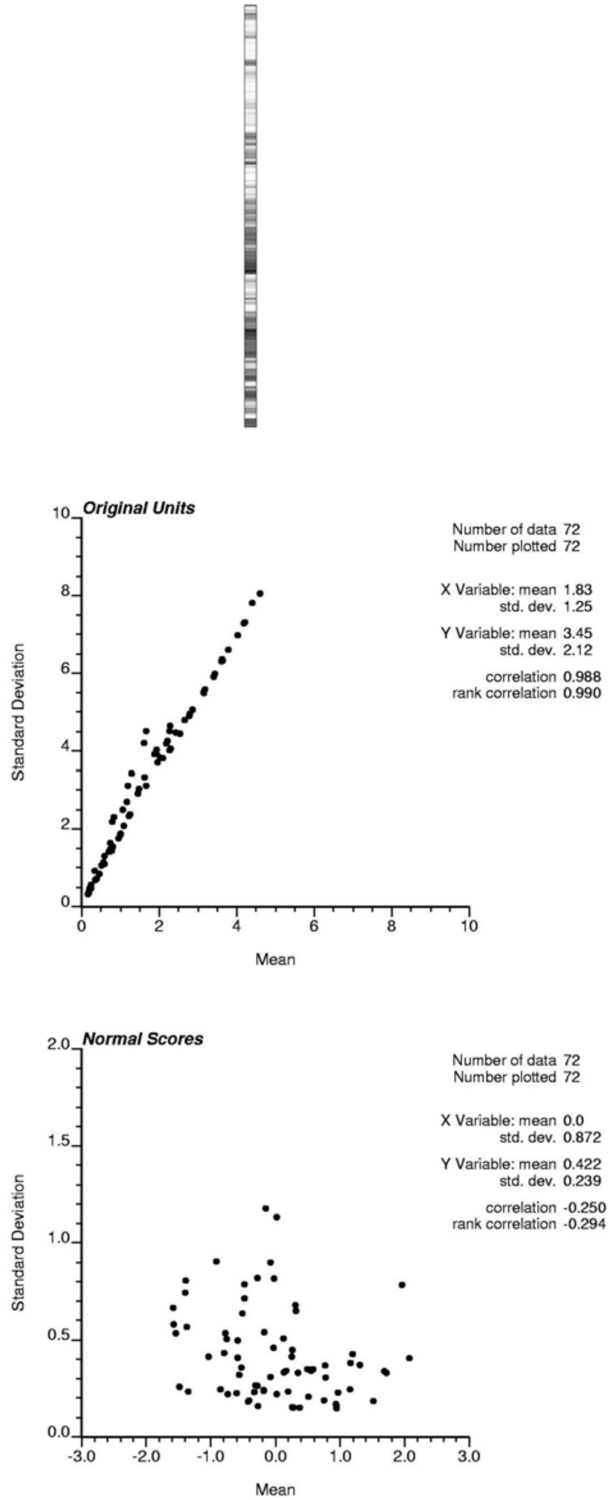


Figure 6: Check for the proportional effect with some Texas porosity data. The top plot shows one well, the central plot shows the relationship between the mean and standard deviation in original data units, and the bottom plot shows the relationship after normal score transformation.

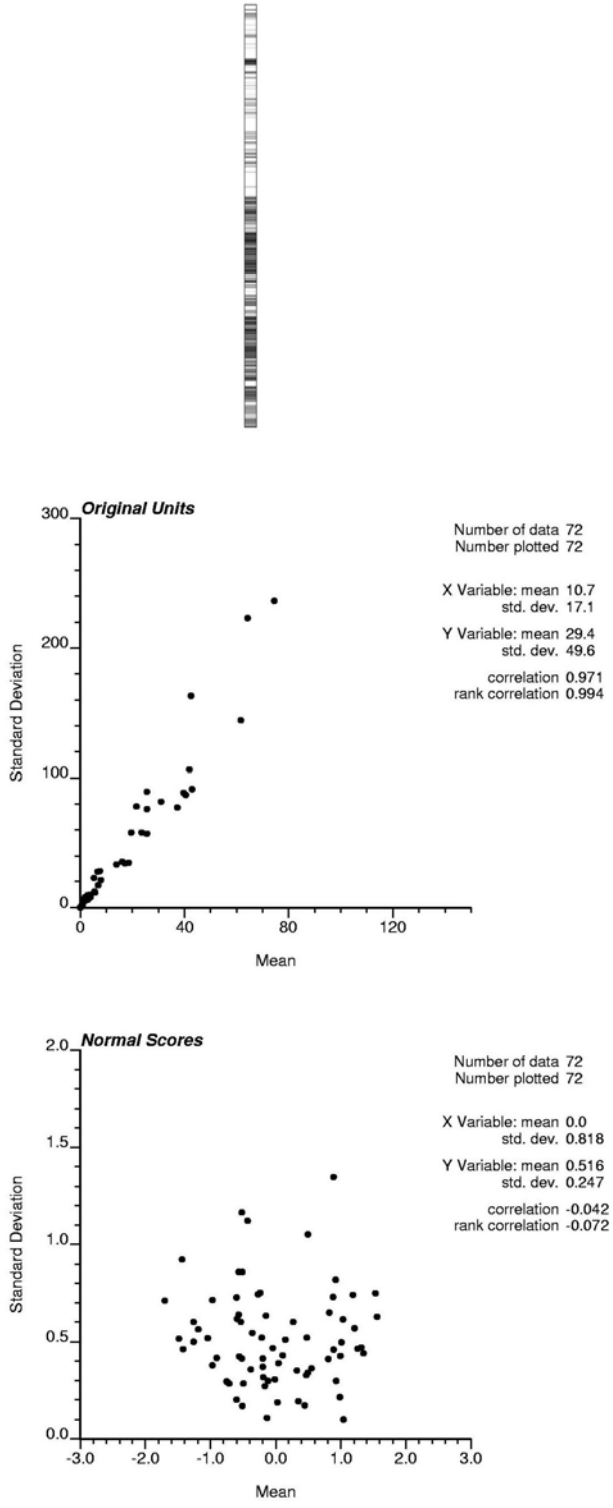


Figure 7: Check for the proportional effect with some Texas permeability data. The top plot shows one well, the central plot shows the relationship between the mean and standard deviation in original data units, and the bottom plot shows the relationship after normal score transformation.