A Short Note on an Update to the Change of Support Program

Yevgeniy V. Zagayevskiy and Clayton V. Deutsch

Accounting for change of support is important in ore reserve estimation and mine planning. The distribution of the data should be adjusted for the larger scale relevant to mining. The three most common methods of change of support are affine correction (AC), indirect lognormal correction (ILC), and discrete Gaussian model (DGM). The methods adjust the distribution differently depending on expected change of shape of distribution, extreme values, and level of symmetry. Theory is revisited and implemented properly in a modified `histscale` program. Results from a gold case study are shown.

1. Introduction

Volume support of data is an important factor in geomodeling, which should be properly taken into account to avoid over or underestimation of ore reserves and effectively optimize mine development plan. This simple preventive measure leads to economically feasible solution with the highest possible NPV. Therefore, data distribution should be adjusted to support level of a model. Size of selective mining unit (SMU) is usually about volume of truckload, but data used to describe local spatial continuity are sampled at smaller scale of drill hole cores. Estimates of ore reserves derived from core data will probably overestimate actual minable tonnage, because of their higher variability. This is typical situation for frequently encountered highly skewed ore grade data. In order to avoid overestimation of ore reserves support of the data has to be changed from core to truck load. Similar problem arises in modeling of a petroleum reservoir, where porosity or permeability attributes have to be modeled at geological reservoir scale, which is larger than scale of core plugs. Thus, data support impacts on managerial decisions and scaling of data distribution is vital procedure in feasibility study of mine planning.

Generally speaking, support level effects distribution of the data in a way that changes its variance and symmetry (Isaaks and Srivastava, 1989). It is proven that variability of data at block support is lower than variability at point support, since high and low values jointly average out each other. Remarkably, but mean of the distribution remains unchanged. This implies that shape of the distribution is more symmetric for block support data rather than for point support data. Thus, despite size of discretization block, derived estimates describe a field only at support level of data used in modeling.

Distribution of the data can be adjusted properly using variance adjustment methods. Some of them are affine correction, indirect lognormal correction (ILC), and correction based on discrete Gaussian model (DGM). The methods differ in degree of preservation of shape of distribution, minimum and maximum values. More explanation is given in next sections. FORTRAN program `histscale.exe` is slightly modified and extended in order to implement these variance adjustment methods in strict compliance with theory and to reproduce target variance exactly keeping mean constant. Level of change of support can be defined through variance adjustment factor *f*, point and block dispersion variances and gamma bar from semivariogram model (average semivariogram value at block support).

This paper is organized in following manner. Theoretical background of change of data support and dispersion variance is presented first. Then the variance adjustment methods are discussed along with implementation details. Description of program `histscale.exe` and its parameter file is given next. Finally case study is carried out on gold dataset to compare variance adjustment methods and show their unique features. Conclusion is made in the end. Derivation of ILC is shown in the Appendix.

2. Theoretical Background of Change of Support

Data sampled at different support level of same quantity possess distinct variability, but same mean. Larger scale data have smaller variance than smaller scale data have. The relationship between variability of the data at different support levels can be expressed by law of additivity of variance or Krige's relation, which is shown in the Equation (1). Here dispersion variance is the variance that relates data from different support levels, i.e. while values represent one support, mean is coming from another one (Equation (2)). Thus, total variance of data is a sum of variance of the data within blocks and variance between blocks (variance of blocks within whole domain *A*). Point · and block *B* support data may be core and truck load samples respectively.

$$D^2(\cdot, A) = D^2(\cdot, B) + D^2(B, A) \tag{1}$$

$$D^2(V_1, V_2) = \frac{1}{n} \sum_{i=1}^{n} \left[x_i(V_1) - m(V_2) \right]^2, \quad V_1 < V_2 \tag{2}$$

where, $\cdot$, B, and A are the point and block supports and entire domain; $D^2(\cdot,A)$ is the point support variance or total variance of the data; $D^2(\cdot,B)$ is the variance within a block; $D^2(B,A)$ is the variance of blocks; V_1 and V_2 are the arbitrary volume supports; n is the number of the data; $x_i(V_1)$ is the ith data value at support V_1; $m(V_2)$ is the mean of the data at support V_2.

Dispersion variance can also be expressed through average value of semivariogram model for specific block size (gamma bar). Volume variance relation between dispersion variance and gamma bar is shown in the Equations (3) and (4). See derivations in (Isaaks and Srivastava, 1989).

$$D^2(B, A) = \bar{\gamma}(A, A) - \bar{\gamma}(B, B) \tag{3}$$

$$D^2(\cdot, V) = \bar{\gamma}(V, V) \approx \frac{1}{N} \sum_{i=1}^{N} \sum_{j=1}^{N} \gamma(\boldsymbol{h}_{ij}), \quad \bar{\gamma}(\cdot, \cdot) = 0 \tag{4}$$

where, $\bar{\gamma}(V, V)$ is the average semivariogram value for support V level data; N is the number of points for volume V discretization; $\gamma(h_{ij})$ is the semivariogram model value for lag vector h_{ij} computed between discretized locations i and j contained inside the volume V.

So, the task of histogram scaling operation is to adjust distribution of data from one scale to another. Usually values of the data $\{x_i(\cdot), i = 1, ..., n\}$ at point scale are altered to get distribution of the same data at block scale with dispersion variance $D^2(B,A)$. Relationship between these two variances can be expressed through single value called variance adjustment (reduction) factor f: ratio of block to point support dispersion variances (Equation (5)). The factor f can also be derived from semivariogram model and its gamma bar values taking into account Krige's and the Equation (4). Resulting relationship is shown in the Equation (6).

$$f = \frac{D^2(B, A)}{D^2(\cdot, A)} \in [0.1] \tag{5}$$

$$f = 1 - \frac{\bar{\gamma}(B, B)}{\bar{\gamma}(A, A)}, \quad \bar{\gamma}(A, A) = \sigma^2 \tag{6}$$

where, σ^2 is the point scale dispersion variance or global variance of the data.

3. Variance Adjustment Methods

Procedure for adjustment of distribution of data from one support level to another is the quantile to quantile transformation, in which mean of the distribution is kept constant, but variance changes according to variance adjustment factor f (Equation (5)). Quantile values x_q are the same as data values x, but ranked from lowest to largest. Expression for the quantile transformation can be found in the Equation (7), whose graphical representation is shown schematically in the Figure 2 through quantile-quantile (Q-Q) plot. Type of variance adjustment method and value of f determine shape and position of the curve on Q-Q plot, i.e. they establish relationship between quantiles from different distributions.

$$x'_{q,i} = F'^{-1} \left[F(x_{q,i}) \right], \quad i = 1, ..., n \tag{7}$$

where, $x_{q,i}$ and $x'_{q,i}$ are the ith quantiles of the original data and transformed values; F and F' are the cumulative distribution functions (CDF) of original data and transformed values; n is the number of the data.

Three the most common variance adjustment methods for quantities that average arithmetically, like ore grade or reservoir porosity, are affine correction (AC), indirect lognormal correction (ILC), and correction based on discrete Gaussian model (DGM). Even though all of them are employed for change of support of a distribution, they generate slightly different distributions using same factor f. Choice of the method depends on required shape of transformed distribution and based on some prior information. Entropy of original data (spatial connectedness of extreme values) can help to decide on correction method: more chaotic fields with higher entropy average to symmetric distribution faster than less entropic spatial variables.

General features of these distribution correction methods are as follows. Affine correction does not change shape of original distribution (degree of symmetry of distribution remains the same, which is expressed through skewness coefficient k). However, this method arbitrary changes minimum and maximum values that might be unrealistic. Affine correction is good for variance adjustment coefficients larger than 0.7 (Isaaks and Srivastava, 1989). Indirect lognormal correction makes distribution more symmetric when its support changes from point to block, what is usually the case in practice. ILC keeps minimum and maximum (extreme) values relatively unchanged, while variance changes according to f. ILC is successfully implemented for wider range of factor f: from 1.0 to up to 0.5. This method is good for lognormal distributions. If distribution is far from lognormal, further minor corrections are necessary. Discrete Gaussian model is not a trivial transformation technique. It is based on representation of quantiles by Hermite polynomials (Wackernagel, 2003). DGM makes original distribution even more symmetric and Gaussian-like for lower values of f than ILC does. Minimum and maximum values do not change very much. Comparison of the methods and dependence of transformations on value of variance adjustment factor f are presented in case study section of this paper.

Equation for affine correction is shown below:

$$x'_q = \sqrt{f} \cdot \left(x_q - m\right) + m \tag{8}$$

where, m is the mean of original distribution.

Indirect lognormal correction consists of three steps. Equations for first step of ILC can be expressed as follows:

$$x'_q = a \cdot \left(x_q\right)^b \tag{9}$$

$$b = \sqrt{\frac{\ln\left(1 + f \cdot \{\sigma/m\}^2\right)}{\ln\left(1 + \{\sigma/m\}^2\right)}} \tag{10}$$

$$a = \left(\frac{m}{\sqrt{1 + \{\sigma/m\}^2}}\right)^{1-b} \tag{11}$$

where, σ is the standard deviation of original distribution.

If adjusted distribution is far from lognormal, further correction of the distribution is required (Equation (12)). However, even though second adjustment preserves mean of original distribution, ultimately it does not correct variance in accordance to specified factor f. Therefore, iterative correction is recommended for last step of ILC in order to get target mean and variance of corrected distribution. Iterative procedure is shown in the Equation (13).

$$x''_q = x'_q \cdot \frac{m}{m'} \tag{12}$$

where, m' is the mean of adjusted distribution from the Equations (9) – (11).

$$x''_{q,j+1} = \left(\frac{x''_{q,j} - m''_j}{\sigma''_j} \right) \cdot f \cdot \sigma + m, \quad j = 1,...,\sqrt{(m''_{j+1} - m)^2 + (\sigma''_{j+1} - f \cdot \sigma)^2} < \xi \tag{13}$$

where, j is the iteration index; ξ is the criterion for termination of iterative correction, it should be a very small number.

Discrete Gaussian model is based on expansion of data values at point and block support through Hermite polynomials $H_p(y)$, normal scores y, anamorphosis function φ, and point-block correlation coefficient r between Hermite polynomials in normal space. Hermite polynomials form basis for nonlinear uniform conditioning (Neufeld, 2005) and disjunctive kriging (Kumar, 2010) estimation methods. Anamorphosis function φ is a general name of nonlinear function that bijectively (single value in one space corresponds to only single value in another space) relates any random variable X to Gaussian random variable Y (Wackernagel, 2003), see equation below.

$$X = \varphi(Y) \tag{14}$$

Quantiles of original and corrected data values and their variances may be expressed through anamorphosis coefficients φ_p and Hermite polynomials $H_p(y)$:

$$x_q = \sum_{p=0}^{\infty} \varphi_p \cdot H_p(y) \approx \sum_{p=0}^{N_p} \varphi_p \cdot H_p(y) \tag{15}$$

$$x'_q = \sum_{p=0}^{\infty} \varphi_p \cdot r^p \cdot H_p(y) \approx \sum_{p=0}^{N_p} \varphi_p \cdot r^p \cdot H_p(y) \tag{16}$$

$$\sigma^2 = \sum_{p=0}^{\infty} \varphi_p^2 \approx \sum_{p=0}^{N_p} \varphi_p^2 \tag{17}$$

$$(\sigma')^2 = (f \cdot \sigma)^2 = \sum_{p=0}^{\infty} \varphi_p^2 \cdot r^{2p} \approx \sum_{p=0}^{N_p} \varphi_p^2 \cdot r^{2p} \tag{18}$$

where, N_p is the finite number of Hermite polynomials used for expansion of quantiles.

Thus, in order to obtain corrected quantiles x'_q anamorphosis function coefficients φ_p, Hermite polynomials $H_p(y)$ and correlation coefficient r should be defined. Anamorphosis coefficients are approximated through following equation:

$$\varphi_p = \sum_{i=2}^{n} (x_{q,i-1} - x_{q,i}) \cdot \frac{1}{\sqrt{p}} \cdot H_{p-1}(y_i) \cdot g(y_i), \quad \varphi_0 = E[X] = m \tag{19}$$

$$g(y) = \frac{1}{\sqrt{2\pi}} \cdot \exp(-\frac{y^2}{2}) \tag{20}$$

where, n is the number of data; i is the index of quantiles; $g(y)$ is the probability density function (PDF) of normal scores Y.

General formula for Hermite polynomials, its first two values and recurrence formula are expressed as:

$$H_p(y) = \frac{1}{\sqrt{p! \cdot g(y)}} \cdot \frac{d^p g(y)}{dy^p}$$

$$H_0(y) = 1$$

$$H_1(y) = -y \tag{21}$$

$$H_{p+1}(y) = -\frac{1}{\sqrt{p+1}} \cdot y \cdot H_p(y) - \sqrt{\frac{p}{p+1}} H_{p-1}(y)$$

Once Hermite polynomials are derived, anamorphosis coefficients are estimated from the Equation (19) and correlation coefficient is found recursively from the Equation (18). Quality of anamorphosis estimates may be checked by comparing values at right and left-hand sides of the Equation (17). Also expansion of original values through Hermite polynomials may be checked by using the Equation (15), where CDF's of original and approximated data values are compared to each other. In case if mean and variance are not reproduced close enough to target values, recursive procedure shown in the Equation (13) may be applied.

4. Program Description

This section is devoted for description of parameter file of the program `histscale.exe` that performs adjustment of distribution of data or variable in order to change its support. Sample parameter file of the program is shown in the Figure 1.

```
                    Parameters for HISTSCALE
                    ************************

          START OF PARAMETERS:
1-        data.out                     -input file with data
2-        4         0                  -   columns for variable and weight
3-        -1.0      1.0e21             -   trimming limits
4-        1                            -option for computation of variance adjustment factor f
5-        0.80                         -   1-value of variance adjustment factor f
6-        16.0      3.2                -   2-dispersion variances at point and block supports
7-        10.0   10.0   10.0           -   3-dispersion variances from gammabar: size of block
8-        5      5      5              -      discretization of grid in X, Y, Z directions
9-        2  0.1                       -      standardized variogram model: # str., nugget effect
10-       1  0.7    0.0    0.0    0.0  -      type of str.#1, var. contr., ang1, ang2, ang3
11-              100.0  300.0   10.0   -      semivariogram ranges a_hmax, a_hmin, a_vert
12-       1  0.2    0.0    0.0    0.0  -      type of str.#2, var. contr., ang1, ang2, ang3
13-              500.0 1500.0   20.0   -      semivariogram ranges a_hmax, a_hmin, a_vert
14-       0.000001                     -DGM: acceptable error for disp. variance at block support
15-       100                          -      number of Hermite polynomials to use
16-       histscale.out                -output file for adjusted data
17-       summary.out                  -output file for summary statistics
```

Figure 1: Parameter file of program `histcale.exe` for adjustment of distribution variance according to

support (scale) level

Name of input file with data is specified on first line. Columns for an attribute and data weights are defined on next line. Then trimming limits are entered. Line 4 is reserved for computation option of variance adjustment factor *f*. It can be 1) defined directly, 2) through dispersion variances at point and block support or 3) from gamma bar, where standardized semivariogram model with sill of one, block size and number of

discretization points in X, Y and Z directions should be specified (see Deutsch and Journel, 1998, for proper definition of semivariogram model). All these parameters are found on lines 5 to 13. Quality of discrete Gaussian model is tuned by minimizing difference between actual block variance and block variance approximated through Hermite polynomials $H_p(y)$ and anamorphosis coefficients φ_p and by number of polynomials used in the expansion of data values. Thus, acceptable error for dispersion variance at block support is defined on line 14 and number of Hermite polynomials to use is specified on line 15. Name for output file is specified on line 16. Original data, data corrected by AC, ILC, and DGM are stored in four columns of the output file. Brief summary statistics are reported in second output file, whose name is defined on last line 17. Mean, variance, standard deviation, coefficient of variation, minimum and maximum, lower quartile, mean and upper quartile, skewness coefficient, and obtained variance adjustment factors between original data and its corrected values can be found in the summary statistics file for all three methods. Also mean squared error between original data and approximated values from Hermite polynomials and anamorphosis function (Equation (15)) and difference of their variances (Equation (17)) are reported in the file as well.

5. Case Study

Gold *Au* grade from `red.dat` dataset is used to examine properties of three variance adjustment methods (affine correction, ILC, DGM), which are used for changing data distribution according to change of support. FORTRAN program `histscale.exe` is employed to carry out scaling of the distribution. Remotely similar Visual Basic program is available as well (Oz et al, 2000). Observation locations of gold data are presented in the Figure 3. It is obvious that declustering for this dataset might be omitted. Comparison of the methods for variance adjustment factor *f* = 0.8 is shown in the Figure 4 in the form of histograms. Effect of variance adjustment factor on shape and position of the distribution for every method is shown in the Figure 5 – Figure 7. Concise summary statistics are tabulated in the Table 1. Corrected mean from all methods is exactly the same as original mean, and variance adjusted precisely according to factor *f*. Recursive corrective procedure is applied to ILC and DGM.

As it was mentioned in theoretical section of this paper, affine correction preserves symmetry of original distribution, which is expressed through skewness coefficient. But minimum and maximum values move to median significantly with decrease of factor *f*. Median slowly moves to mean of the distribution. ILC improves symmetry of original distribution (skewness tends to be closer to zero) trying to preserve its shape as well. Minimum value remains almost unchanged for higher *f* and slowly approaches median, when *f* is getting smaller. Maximum value is smaller for ILC than for affine correction, what is probably right thing for examined distribution, since high values average out faster than low values, which are represented by the peak of zero values on histogram of original distribution. DGM does not improve symmetry very much for high *f*, but slowly leans to more symmetric and Gaussian distribution for lower *f* values changing shape of the distribution. Minimum and maximum values do not change as much as for first two methods. Median and mean values slowly converge for this correction method.

To sum up, affine correction preserves shape of the distribution and does not improve symmetry. Minimum and maximum values change significantly and arbitrarily. ILC increases symmetry of the distribution, but tries not to adjust its shape very much. Extreme values do not change as much as in affine corrected distribution. DGM makes distribution more symmetric and Gaussian. Minimum and maximum values do not change very much in comparison to other methods' results. Also median slowly drifts to mean with decrease of variance adjustment factor *f* for all methods, except ILC.

Table 1: Summary statistics

Statistic	Original Data	Affine Corrected Data				Indirect Lognormal Corrected Data				Discrete Gaussian Model Corrected Data			
Variance adjustment factor		0.8	0.6	0.4	0.2	0.8	0.6	0.4	0.2	0.8	0.6	0.4	0.2
Minimum	0.00	0.14	0.30	0.49	0.75	0.09	0.21	0.36	0.56	0.00	0.00	0.06	0.28
Median	1.22	1.23	1.25	1.27	1.29	1.28	1.36	1.43	1.49	1.11	1.18	1.24	1.30
Maximum	4.45	4.12	3.75	3.31	2.74	4.01	3.53	2.99	2.36	4.20	3.87	3.45	2.84
Skewness	0.58	0.58	0.58	0.58	0.58	0.49	0.37	0.21	-0.05	0.59	0.56	0.49	0.36

6. Conclusion

Program `histscale.exe` was slightly modified and expanded in order to implement effect of change of support on data distribution right. Three variance adjustment methods, affine correction, indirect lognormal correction, and discrete Gaussian model, are imbedded in the program with iterative calibration to get corrected distribution with unchanged mean and properly adjusted variance. The methods generate distributions with different change in shape, symmetry and extreme values, which were presented in the case study. AC does not change symmetry of distribution, however significantly changes minimum and maximum values. ILC improves symmetry, but tries to keep shape of distribution similar to original. Maximum and minimum values are not changed very much and in accordance to theory. DGM makes distribution more symmetric and Gaussian. Extreme values change the least. All methods work the best for larger values of variance reduction factor *f* (f > 0.5).

References

Deutsch, C.V. and Journel, A.G., 1998, *GSLIB: Geostatistical Software Library and User's Guide*, Oxford University Press, New York, 2nd Ed., 369 pp.

Isaaks, E.H and Srivastava, R.H., 1989, *Introduction to Applied Geostatistics*, Oxford University Press, New York, 547 pp.

Kumar, A., 2010, *Guide to Recoverable Reserves with Disjunctive Kriging*, Center for Computational Geostatistics, Guidebook Series Vol. 9, 35 pp.

Neufeld, C.T., 2005, *Guide to Recoverable Reserves with Uniform Conditioning*, Center for Computational Geostatistics, Guidebook Series Vol. 4, 40 pp.

Oz, B., Deutsch, C.V. and Frykman, P., 2000, A VisualBasic Program for Histogram and Variogram Scaling, *Center for Computational Geostatistics 2*, 1 - 18

Wackernagel, H., 2003, *Multivariate Geostatistics: An Introduction with Applications*, Springer, New York, 387 pp.

Appendix – Derivation of equations of indirect lognormal correction (ILC)

When affine correction is applied to change support of normally distributed data from point to block scale, shape of the distribution does not change. In general, adjusted distribution tends to be more symmetric and Gaussian, what is already a case for normal distribution. Recall that logarithmic units of log normally distributed data follow normal distribution. Thus, shape of affine corrected distribution of logarithmic values of these data should remain the same. This idea is used to derive equation for ILC. Expression for affine corrected log units of lognormal data is shown below:

$$\ln(x'_q) = \sqrt{f_{\ln(x_q)}} \cdot \left(\ln(x_q) - m_{\ln(x_q)}\right) + m_{\ln(x_q)} \tag{22}$$

$$m_{\ln(x_q)} = \ln\left(\frac{m}{\sqrt{1 + \{\sigma/m\}^2}}\right) \tag{23}$$

$$\sigma^2_{\ln(x_q)} = \ln\left(1 + \{\sigma/m\}^2\right) \tag{24}$$

$$f_{\ln(x_q)} = \frac{(\sigma')^2_{\ln(x_q)}}{\sigma^2_{\ln(x_q)}} \tag{25}$$

$$f = \frac{(\sigma'^2)}{\sigma^2} \tag{26}$$

where, m and σ^2 are the original mean and variance of the data; $(\sigma')^2$ is the corrected variance of the data; $m_{ln(xq)}$ and $\sigma^2{}_{ln(xq)}$ are the original mean and variance of logarithmic values of the data; $(\sigma')^2{}_{ln(xq)}$ is the corrected variance of logarithmic values of the data; f and $f_{ln(xq)}$ are the variance adjustment factors for original and logarithmic units of the data.

Thus, from Equations (22) – (26):

$$\ln(x'_q) = \sqrt{\frac{\ln\left(1 + f \cdot \{\sigma/m\}^2\right)}{\ln\left(1 + \{\sigma/m\}^2\right)}} \cdot \left(\ln(x_q) - \ln\left(\frac{m}{\sqrt{1+\{\sigma/m\}^2}}\right)\right) + \ln\left(\frac{m}{\sqrt{1+\{\sigma/m\}^2}}\right) \qquad (27)$$

$$\ln(x'_q) = \sqrt{\frac{\ln\left(1 + f \cdot \{\sigma/m\}^2\right)}{\ln\left(1 + \{\sigma/m\}^2\right)}} \cdot \ln(x_q) + \ln\left(\frac{m}{\sqrt{1+\{\sigma/m\}^2}}\right) - \sqrt{\frac{\ln\left(1 + f \cdot \{\sigma/m\}^2\right)}{\ln\left(1 + \{\sigma/m\}^2\right)}} \cdot \ln\left(\frac{m}{\sqrt{1+\{\sigma/m\}^2}}\right) \qquad (28)$$

$$\ln(x'_q) = \cdot\ln(x_q^{\sqrt{\frac{\ln\left(1+f\cdot\{\sigma/m\}^2\right)}{\ln\left(1+\{\sigma/m\}^2\right)}}}) + \ln\left(\left[\frac{m}{\sqrt{1+\{\sigma/m\}^2}}\right]^{1-\sqrt{\frac{\ln\left(1+f\cdot\{\sigma/m\}^2\right)}{\ln\left(1+\{\sigma/m\}^2\right)}}}\right) \qquad (29)$$

$$x'_q = \left[\frac{m}{\sqrt{1+\{\sigma/m\}^2}}\right]^{1-\sqrt{\frac{\ln\left(1+f\cdot\{\sigma/m\}^2\right)}{\ln\left(1+\{\sigma/m\}^2\right)}}} \cdot x_q^{\sqrt{\frac{\ln\left(1+f\cdot\{\sigma/m\}^2\right)}{\ln\left(1+\{\sigma/m\}^2\right)}}} \qquad (30)$$

$$x'_q = a \cdot x_q^b \qquad (31)$$

$$b = \sqrt{\frac{\ln\left(1 + f \cdot \{\sigma/m\}^2\right)}{\ln\left(1 + \{\sigma/m\}^2\right)}}, \quad a = \left[\frac{m}{\sqrt{1+\{\sigma/m\}^2}}\right]^{1-b} \qquad (32)$$
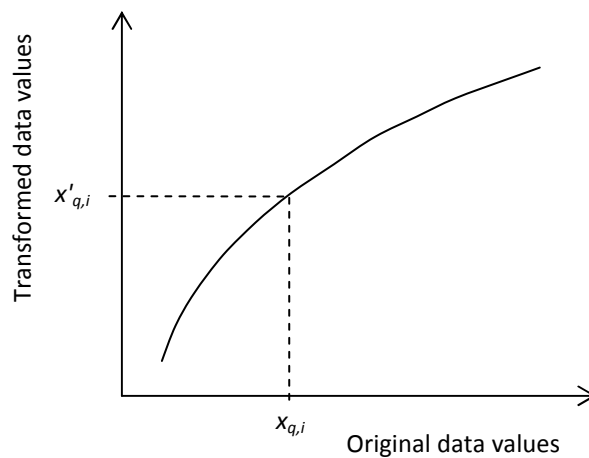

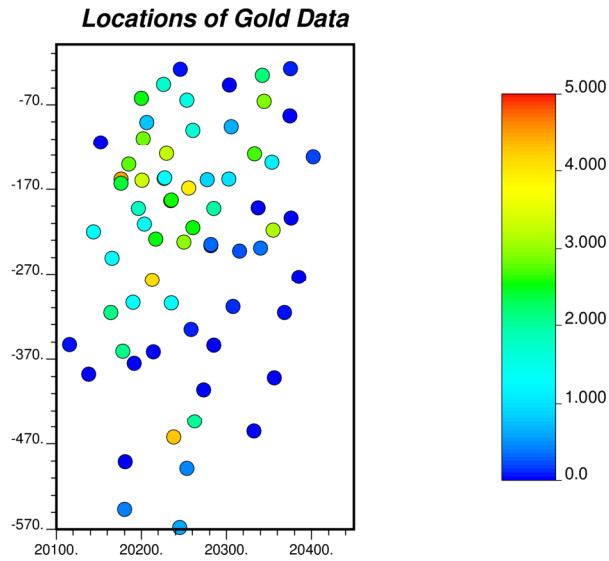**Figure 2**: Schematic Q-Q plot of change of support level of a distribution

Locations of Gold Data

Figure 3: Locations of gold *Au* data from `red.dat` dataset
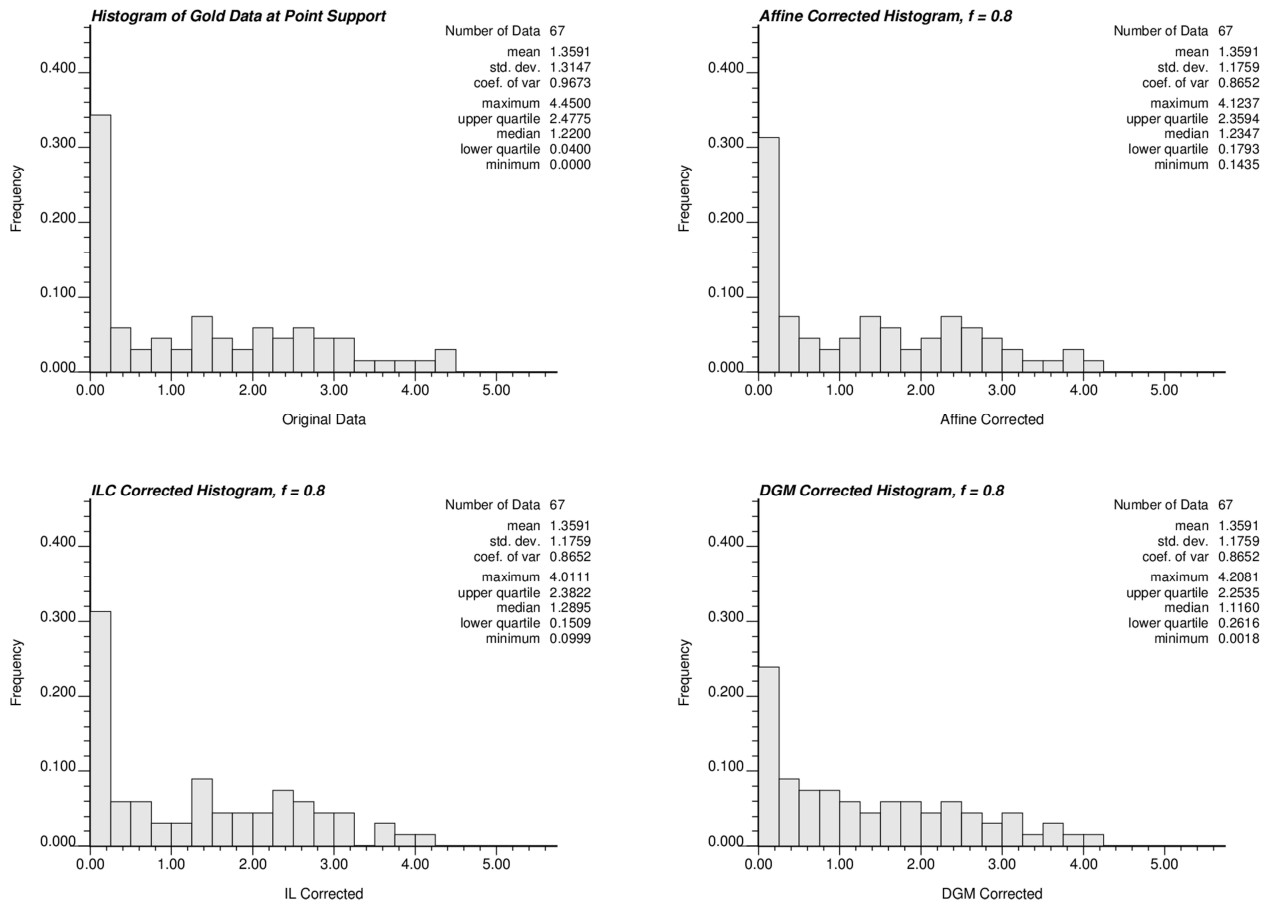

Figure 4: Comparison of variance adjustment methods on gold data from `red.dat` dataset: original histogram, affine corrected histogram, histogram is corrected according to indirect lognormal correction, and histogram is corrected according to discrete Gaussian model. Variance reduction factor *f* is 0.8.
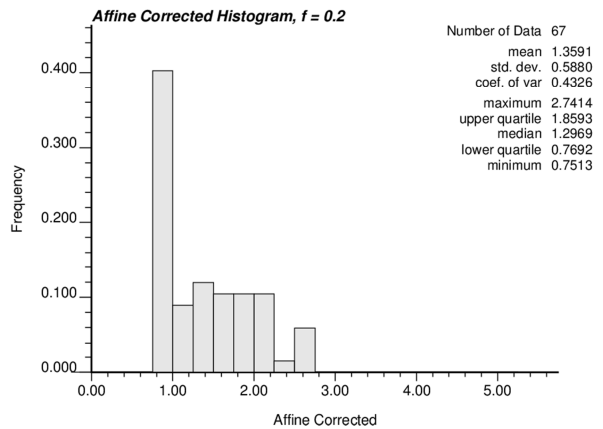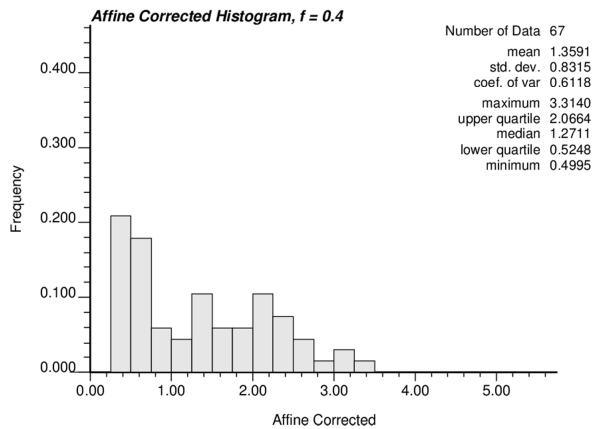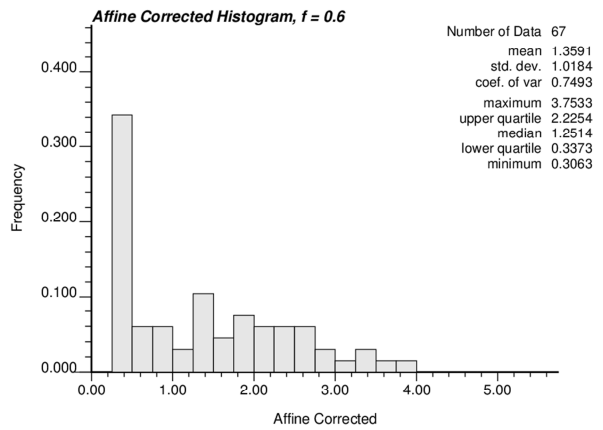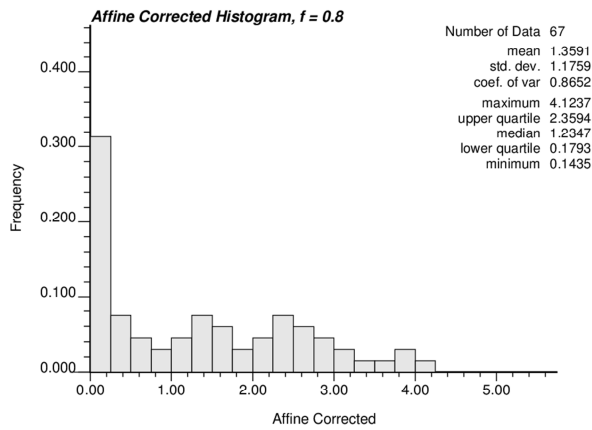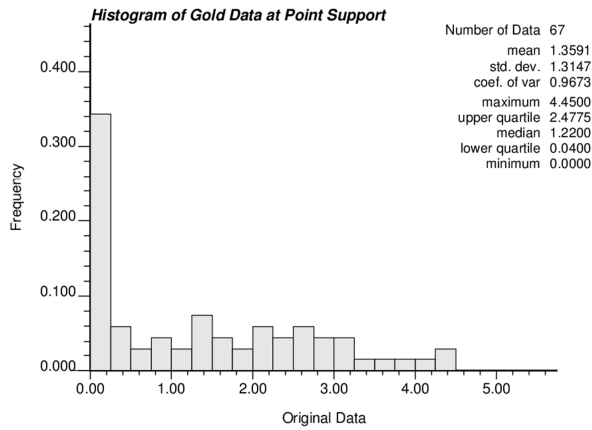
Figure 5: Influence of variance reduction factor *f* on change of shape of histogram for affine correction: original histogram, *f* = 0.8, *f* = 0.6, *f* = 0.4 and *f* = 0.2.
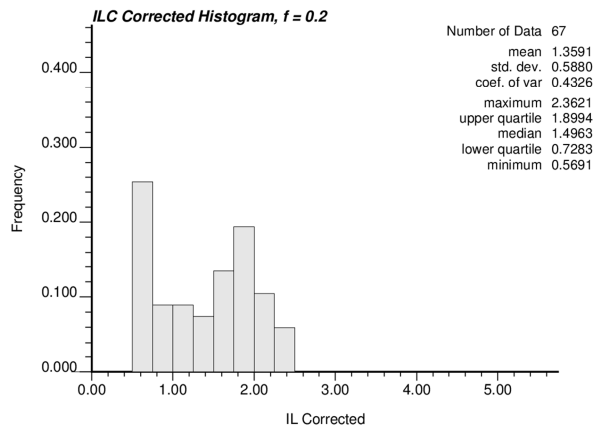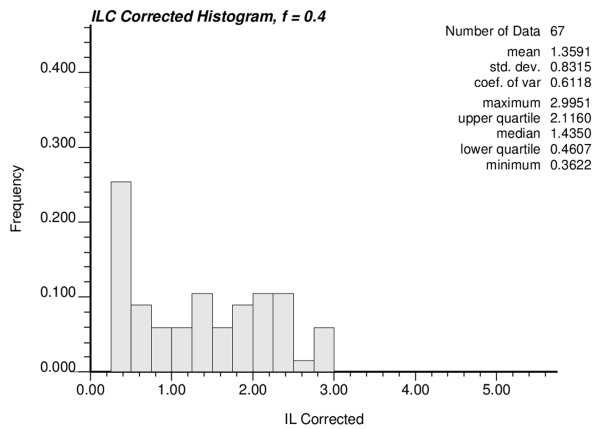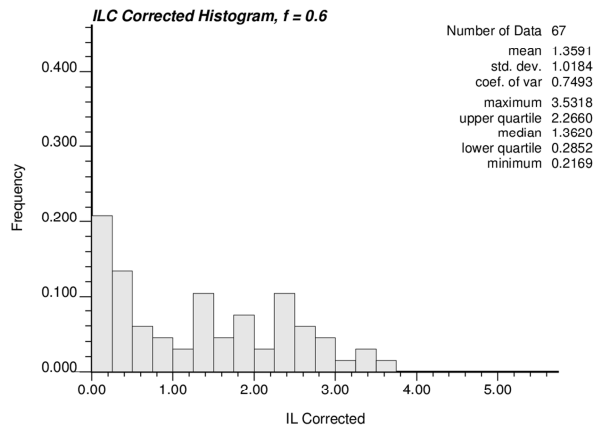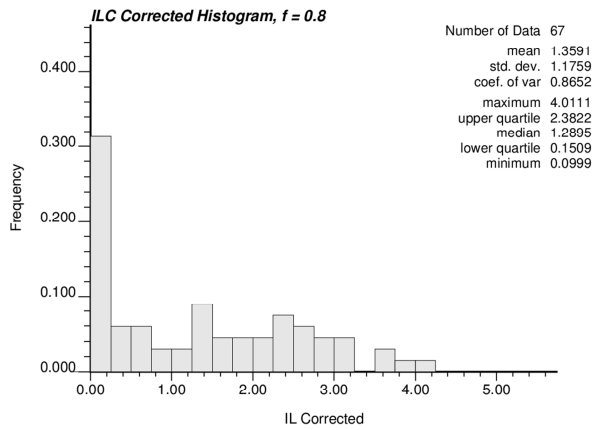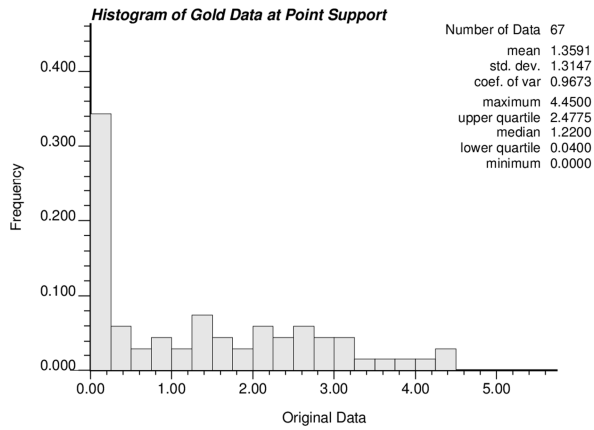
Figure 6: Influence of variance reduction factor *f* on change of shape of histogram for indirect lognormal correction (ILC): original histogram, *f* = 0.8, *f* = 0.6, *f* = 0.4 and *f* = 0.2.
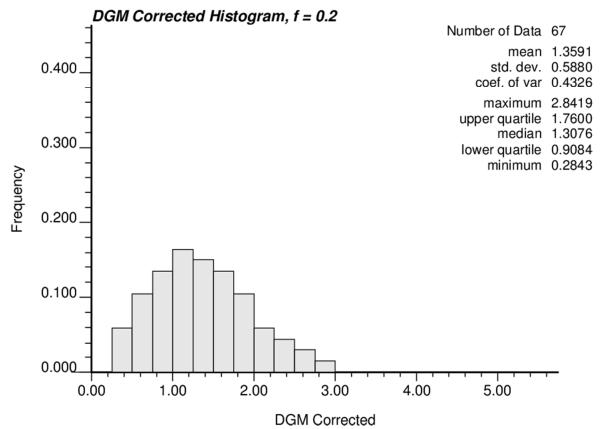
Histogram of Gold Data at Point Support

Number of Data	67
mean	1.3591
std. dev.	1.3147
coef. of var	0.9673
maximum	4.4500
upper quartile	2.4775
median	1.2200
lower quartile	0.0400
minimum	0.0000

DGM Corrected Histogram, f = 0.8

Number of Data	67
mean	1.3591
std. dev.	1.1759
coef. of var	0.8652
maximum	4.2081
upper quartile	2.2535
median	1.1160
lower quartile	0.2616
minimum	0.0018

DGM Corrected Histogram, f = 0.6

Number of Data	67
mean	1.3591
std. dev.	1.0184
coef. of var	0.7493
maximum	3.8775
upper quartile	2.0901
median	1.1848
lower quartile	0.4757
minimum	0.0023

DGM Corrected Histogram, f = 0.4

Number of Data	67
mean	1.3591
std. dev.	0.8315
coef. of var	0.6118
maximum	3.4530
upper quartile	1.9353
median	1.2494
lower quartile	0.6820
minimum	0.0634

DGM Corrected Histogram, f = 0.2

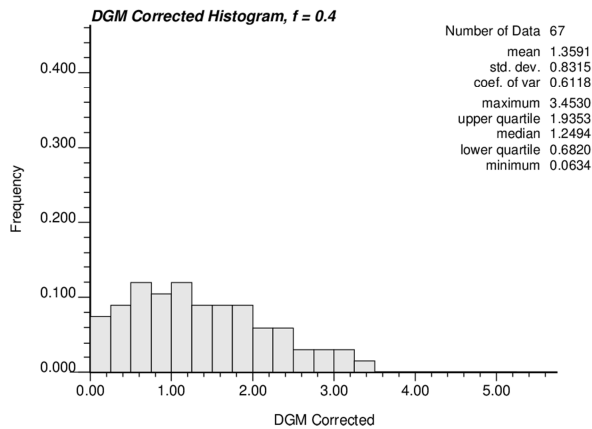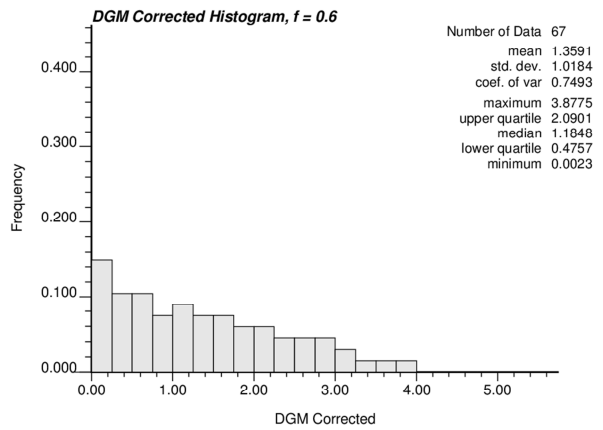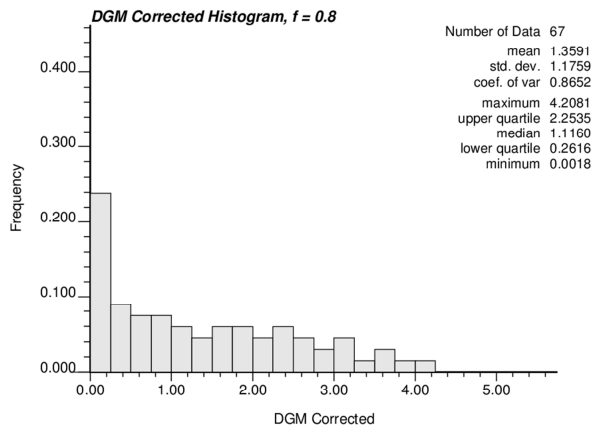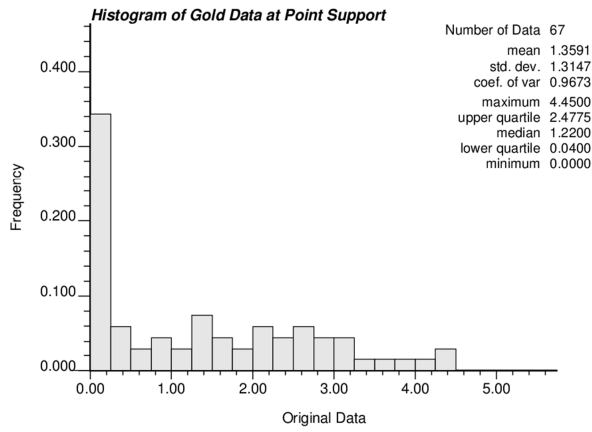|	
Number of Data	67
mean	1.3591
std. dev.	0.5880
coef. of var	0.4326
maximum	2.8419
upper quartile	1.7600
median	1.3076
lower quartile	0.9084
minimum	0.2843

Figure 7: Influence of variance reduction factor f on change of shape of histogram for discrete Gaussian model (DGM): original histogram, f = 0.8, f = 0.6, f = 0.4 and f = 0.2.