Application of Logratios for Compositional Data

Michael Job

Numeric data for earth sciences often represent fractions or percentages of part of a whole, such as the chemical composition of a rock, or oil/water saturations in rock volume for petroleum reservoirs. The individual components can be considered a proportion of the whole composition, which is a constant sum to 100% or 1.0. The preservation of these proportions at unsampled locations after independent estimation or conditional simulation is an appealing concept, but not guaranteed. The variables are therefore not free to vary independently, and the constant sum constraint forces at least one negative correlation, which is known as a spurious correlation. Therefore, a series of transformations using the logarithms of the ratios between the components is used to overcome these two problems. Linear averaging of logarithms results in a geometric rather than arithmetic mean, which will result in a bias. Ordinary kriging and conditional simulation were used on data from the Alberta Oil Sands to assess the performance of the compositional geostatistics approach.

1. Introduction

The compositional data approach in the earth sciences has proven to be successful in non-spatial applications such as mineralogy and petrology (for example Thomas and Atchison, 2005 and Martin-Fernandez et al., 2005), but is yet to find acceptance in geostatistics (Tolosana-Delgado et al., 2008), particularly in industry. There are many situations in geostatistical modeling where the preservation of proportions found in input data is of practical importance in estimates or simulations. While a number of estimation methods have been proposed, there have been various warnings in the literature about the theoretical and practical reliability of these approaches, particularly where the estimate or simulation involves an averaging of log-transformed data.

For example, Lan et al. (2006) discuss the benefits of applying the logratio transform for statistical analysis, but they conclude that using logratios to make estimates using a linear approach such as kriging will result in a bias. This is due to the back-transform of the arithmetic average of logratio values returning the geometric mean of the proportions being studied, which is not the required result for variables that average linearly.

However, Tolosana-Delgado et al. (2008) maintain that the back-transformed results after a 'standard' geostatistical (cokriging) approach is used are linear, unbiased (null expected error) and with minimal variance between the true and estimated value on a relative scale. In addition, the method will yield positive and bounded compositions, but this is a feature of the back-transform.

2. Theory

The field of compositional statistics is largely based on the initial work of Atchison (1986), and these concepts were extended to spatial data by numerous workers in the 1990's and 2000's, notably Pawlowsky-Glahn and Olea (2004). Only the basic concepts needed to understand the methodology are presented here - the interested reader can refer to the above-mentioned books or the CCG Guidebook on Compositional Geostatistics (Manchuk, 2008) for further details.

The basic premise of compositional data analysis is that the data contains information about the relative magnitudes of the components, not just the absolute magnitudes, and therefore these relationships can be expressed as ratios. However, mathematical analysis of these ratios directly is problematic due to the constant sum constraint, meaning that correlations of the raw components and associated forms of standard multivariate statistical analysis designed for unconstrained data are not suitable (Aitchison, 1999). Standard multivariate analysis is applicable for unconstrained vector data from real Euclidean space, but the sample space of compositions is constrained to the simplex, a generalization of a triangle and tetrahedron (Aitchison et al., 2002). Aitchison (1986) also explained that the logarithm of ratios are easier to handle mathematically and interpret than statistically than the ratios themselves, and proposed a number of transforms of the ratios using logarithms – these transforms do not lose any of the information about the composition, as 'there is a one-to-one correspondence between any D-part composition (i.e. consisting of D components, $x_1,\ldots,x_D$) and its logratio vector' (Aitchison, 1999).

In addition, there is the problem of 'subcompositional incoherence' when dealing with the ratios directly (Aitchison, 1986). A subcomposition is a subset extracted from a full composition, and normalized. The covariance

relationships between the variables in the subcomposition are not the same as between the same variables in the full composition, and there may be no relationship between the two covariance structures. Working with logratio methods leads to consistent results whether working with a full composition or a subcomposition (Pawlowsky-Glahn and Egozcue, 2006).

There are four main logratio transforms: the additive logratio (alr), centered logratio (clr), multiplicative logratio (mlr) (Aitchison, 1986), and the isometric logratio (ilr), introduced by Egozcue et al., 2003.
The alr transform is shown in Equation 1,

$$y_i = \log\left(\frac{x_i}{x_D}\right), \qquad i = 1,\ldots,d \tag{1}$$

where the denominator (x_D) can be any of the components but the same component must be used for all data points and must be strictly >0. Interestingly, the choice of denominator does not affect the outcome of the alr transform or back-transform. This transformation results in one less transformed variable than the number of components considered. The alr back-transform is shown in Equation 2.

$$x_i = \frac{\exp(y_i)}{\sum_{i=1}^{d}\exp(y_i)+1}, \qquad i = 1,\ldots,d \tag{2}$$

The clr transform is shown in Equation 3,

$$y_i = \log\left(\frac{x_i}{g(\mathbf{x})}\right), i = 1,\ldots,D \tag{3}$$

where g(x) is the geometric mean of all components. The clr back-transform is shown in Equation 4.

$$x_i = \frac{\exp(y_i)}{\sum_{i=1}^{d}\exp(y_i)}, \qquad i = 1,\ldots,d \tag{4}$$

Note that these equations refer to natural logarithms, but the transforms can also be used with the logarithm of different bases.

There are two other important transformation methods; the mlr is similar to the alr, but uses the 'filler' component (a component introduced to ensure the composition sums to unity) as the denominator. For example, consider assay data for an iron ore deposit with Fe%, P%, SiO_2%, Al_2O_3% and LOI% (Loss on Ignition), and usually CaO%, MgO%, TiO_2%, S%, MnO% and K2O%. Due to other trace compounds, the assayed composites will not sum to 100%. Therefore, a filler component is needed. The oil sands data set used for this exercise sum to 100%; therefore, a filler is not required.

The ilr transform (Egozcue et al., 2003), is shown in Equation 5.

$$ilr\,(x) = V.\,clr(x) \tag{5}$$

where V is a matrix of D rows and (D – 1) columns such that $V \,.\, V^t = I_{D-1}$ (identity matrix of D – 1 elements) and $V \,.\, V^t = I_D + a1$, where a may be any value, and 1 is a matrix full on ones (Tolosana-Delgado, 2008).

3. Dealing with Zeros

It is possible in any given data set that some of the components have a null value. Zeros are problematic as the logarithm of zero is undefined and also cannot be used as a denominator which is needed for some logratio transforms.

In many instances the zero could be due to the component being below the detection limit of the instrument used. In this case, it is common to select a small value to assign the sample. Setting the missing value to half the detection limit of the instrument or analysis method being used is common practice in industry. Adding this small value causes the sum>1.0 and the other components are restandardized to maintain the sum=1.0.

It is also possible that a zero value implies a total absence of a component, such as the complete absence of gold in unmineralised country rock surrounding a gold deposit. One solution to deal with this case is to separate these zeros from the rest of the population by domaining and considering a separate domain with *n*-1 components.

It may be difficult to separate the domain into zones with identically informed components and an amalgamation of similar variables (Martin-Fernandez et al., 2000) can be considered. This involves amalgamating components such that the resulting compositions no longer contains zeros; it can only be used if there are more components measured than are needed for the study and if the amalgamated components do not contain one of the primary variables of interest. It is also possible to combine the amalgamation and domaining approach, so that only selected zones or domains need to be subjected to amalgamation.

4. Case Study: Data

The data used for this case study comes from the Alberta Oil Sands and contains four components, bitumen (B), water (W), coarse solids (C) and fine solids (F). These four components complete a whole composition (i.e. sum to unity). The coarse/fine division is generally accepted to be 44 micrometre particle size (e.g. Romanova et al., 2003), with the proportion of fines being a key driver of the bitumen recovery during processing (e.g. Wik et al., 2008).

The data consists of vertical drill holes, with a maximum depth of 126m. Drill hole data was selected from the main bitumenised horizon, within a 2,000m x 3,000m area that was drilled with a hole spacing of approximately 100m x 100m. Data is collected at 1.5m intervals down the hole.

There are a small number of data points where one or more of the components are zero. Three of the data points had a value of zero for all components, and a further twenty-one data points (from a limited number of drill-holes) had a bitumen value of one, and zero for all the other components. These are clearly database anomalies incorrectly representing null or missing values. These data points were therefore removed from the data set.

However, there are twenty-one instances where bitumen is zero, but the other components are valid (and sum to unity), and three examples of this with fines. It is unclear whether these data points represent 'true' or 'below-detection' zeros; the true absence of fines seems unlikely, and the true absence of bitumen also seems unlikely, given that the intervals above and below these zeros do contain (sometimes significant) bitumen. Amalgamation of variables which seem mutually exclusive is difficult to justify in this particular data set as all components are important.

Local upscaling (compositing with an adjacent sample) is another approach, but for the purposes of this exercise, they were eliminated from the data set. This leaves 6,555 'valid' data points (Figure 1).

5. Transforms

The alr and clr transforms were calculated, with the coarse fraction selected as the denominator for alr because it is the major component of the composition.

Statistical and spatial analysis of the three alr variables (alrB:C, alrF:C and alrW:C) and the four clr variables (clrB, clrC, clrF and clrW) was undertaken, with the basic statistics and histograms for the raw and transformed data shown in Figure 1 to Figure 3, and the variograms for the transformed data in Figure 4 and Figure 5. Scatterplots of the alr and clr variables are shown in Figure 6 and Figure 7.

Grid declustering for the untransformed and transformed data was also undertaken for later comparison to the estimates and simulations.

6. Estimation

A model with size 10x10x1.5m (with 200 cells in E, 300 in N and 44 vertically) over same area as the selected drilling data was constructed. Ordinary Kriging was run independently for all seven logratio transformed components. The mean of the estimates for the logratio data match the declustered input data indicating no bias in the transformed space (see parts of Table 2).

These logratio estimates were then back-transformed into original components, and validation of these estimates against the input data showed that there were some inconsistencies for both logratio transform methods (see Table 1)

- The means for bitumen in the estimates are slightly less than the raw mean of the drilling, but very close to the declustered mean of 0.0858 (from cell declustering).

- The means for coarse in the estimates are above the raw mean of the drilling, and well above the declustered mean of 0.5797.
- The means for fines in the estimates are less than the raw mean of the drilling, and well below the declustered mean of 0.269.
- The means for water in the estimates are below the raw mean of the drilling, and well below the declustered mean of 0.0655.

In addition, the maximum bitumen value from both methods is actually higher than the maximum bitumen value in the drilling, which is a concerning result given that bitumen is the main variable of interest. Of course, kriging does not necessarily honor the bounds of the input data (due to negative weights), but the data in Table 2 shows that the estimated logratio variables do fall within the bounds of the input data. It is therefore possible that there are inconsistencies and distortions in the relative estimates for the logratio variables due to the use of different variograms, and therefore the use of different weights. These inconsistencies may therefore be carried through the back-transform, which by construction, will result in a composition with the appropriate constant sum. However, these high bitumen values represent less than 0.02% of the blocks located at about the centre of the grid.

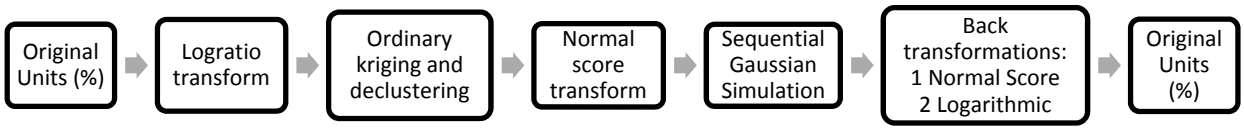
Q-Q plots showing the distribution of the input drilling against the model output for the two different logratio methods are shown in Figure 8. These plots show that the models have not reproduced the input data very well – for bitumen and the coarse fraction, the model grades at quantiles below the mean are higher than the input data, with the model grades lower than the input data above the mean. This holds for both logratio methods, with the water and fine fraction component antithetic to bitumen and the coarse fraction.

	Drilling			alr OK estimate			clr OK estimate		
VARIABLE	**Min**	**Max**	**Decl. Mean**	**Min**	**Max**	**Mean**	**Min**	**Max**	**Mean**
Bitumen	0.0001	0.1882	0.0858	0.0004	0.2348	0.0867	0.0005	0.2070	0.0865
Coarse	0.0008	0.9111	0.5797	0.0614	0.8705	0.6310	0.0704	0.8607	0.6282
Fines	0.0008	0.9105	0.2690	0.0066	0.8313	0.2161	0.0076	0.8384	0.2198
Water	0.0030	0.2884	0.0655	0.0065	0.1939	0.0663	0.0066	0.2041	0.0655
Total	0.9999	1.0001	1	1	1	1	1	1	1

Table 1: Comparison of input data and kriged models.

7. Conditional Simulation

To perform simulation, the transformed logratio variables were then transformed to a Gaussian distribution, similar to the methodology of Boisvert et al. (2009). Variograms were modeled for each of the Gaussian transformed variables. Fifty independent realizations were run for each of the seven variables, using sequential Gaussian simulation (SGS). Two sets of simulations were run 1) using a small sequential neighbourhood (20 nodes in E, 20 nodes in N, 5 nodes in RL), and a second using a larger neighbourhood. A summary of the transformations and procedure is shown below (modified after Boisvert et al., 2009).



Checks were performed in normal space to assess the reproduction of the input distribution and variogram and were found to be unsatisfactory, with the variance of the simulations being much lower than the input except in the downhole case, where the variances were higher. Example variograms for the realizations showing this are in Figure 9. The variograms for the other components and for the sequential neighbourhood are nearly identical, even with the larger search.

The Gaussian simulations were back-transformed into the logratio variables, and a comparison of the logratio input data and the simulated logratio values (before the final logratio back-transform to raw components) showed that the basic statistics were very well reproduced (see Table 2), with an acceptable variance reproduction and the histograms of selected realizations match the input data very well. Even though the variance has not been reproduced in Gaussian space, it appears acceptable in logratios.

These simulated logratio variables were then transformed back into their raw components. Comparison of these with the raw input data shows that, similar to the kriging, the mean of individual components has not been very well reproduced, and the maxima of the simulations are well above the input data (see Table 3).

		Count	Minimum	Maximum	Declustered Mean	Std. Dev.	Variance
alrB:C	**Input**	6494	-7.788	3.263	-2.139	0.690	0.476
	OK Estimate	2417495	-6.426	-0.415	-2.137	0.461	0.212
	SGS_standard	2639845	-7.788	3.263	-2.133	0.707	0.500
	SGS_sequential	2640000	-7.788	3.263	-2.132	0.697	0.485
alrF:C	**Input**	6494	-6.941	6.926	-1.352	1.614	2.605
	OK Estimate	2417495	-4.864	2.621	-1.349	1.094	1.196
	SGS_standard	2639845	-6.941	6.926	-1.350	1.647	2.712
	SGS_sequential	2640000	-6.941	6.926	-1.350	1.649	2.720
alrW:C	**Input**	6494	-5.639	5.322	-2.331	1.075	1.156
	OK Estimate	2417495	-4.854	0.509	-2.329	0.740	0.548
	SGS_standard	2639845	-5.639	5.322	-2.332	1.084	1.175
	SGS_sequential	2640000	-5.639	5.322	-2.334	1.078	1.162
clrB	**Input**	6556	-5.876	1.586	-0.683	1.029	1.060
	OK Estimate	2417495	-4.668	1.015	-0.683	0.751	0.564
	SGS_standard	2639845	-5.876	1.586	-0.683	1.062	1.127
	SGS_sequential	2640000	-5.876	1.586	-0.682	1.058	1.120
clrC	**Input**	6556	-3.878	3.403	1.452	0.562	0.315
	OK Estimate	2417495	-0.242	2.671	1.450	0.384	0.147
	SGS_standard	2639845	-3.878	3.403	1.455	0.587	0.345
	SGS_sequential	2640000	-3.878	3.403	1.457	0.583	0.340
clrF	**Input**	6556	-3.719	3.524	0.107	1.067	1.138
	OK Estimate	2417495	-2.163	3.017	0.108	0.764	0.584
	SGS_standard	2639845	-3.719	3.524	0.097	1.104	1.219
	SGS_sequential	2640000	-3.719	3.524	0.101	1.110	1.231
clrW	**Input**	6556	-2.960	1.555	-0.878	0.578	0.334
	OK Estimate	2417495	-2.238	0.783	-0.877	0.381	0.145
	SGS_standard	2639845	-2.960	1.555	-0.884	0.570	0.325
	SGS_sequential	2640000	-2.960	1.555	-0.884	0.574	0.329

Table 2: Logratio Input data, OK, and conditional simulation basic statistics

		Minimum	Maximum	Mean	Std. Dev.	Variance
Bitumen	**Raw Input**	0.000	0.188	0.090	0.053	0.003
	Declustered Input	0.000	0.188	0.086	0.056	0.003
	alr_SGS_standard	0.000	0.377	0.078	0.040	0.002
	alr_SGS_sequential	0.000	0.406	0.078	0.040	0.002
	clr_SGS_standard	0.000	0.201	0.085	0.040	0.002
	clr_SGS_sequential	0.000	0.212	0.085	0.040	0.002
Coarse	**Raw Input**	0.001	0.911	0.604	0.204	0.042
	Declustered Input	0.001	0.911	0.580	0.227	0.052
	alr_SGS_standard	0.005	0.995	0.617	0.235	0.055
	alr_SGS_sequential	0.005	0.995	0.617	0.235	0.055
	clr_SGS_standard	0.073	0.925	0.643	0.137	0.019
	clr_SGS_sequential	0.007	0.929	0.643	0.137	0.019
Fines	**Raw Input**	0.001	0.911	0.244	0.223	0.050
	Declustered Input	0.001	0.911	0.269	0.248	0.062
	alr_SGS_standard	0.001	0.962	0.243	0.222	0.049
	alr_SGS_sequential	0.001	0.965	0.244	0.223	0.050
	clr_SGS_standard	0.003	0.874	0.208	0.148	0.022
	clr_SGS_sequential	0.001	0.907	0.209	0.149	0.022
Water	**Raw Input**	0.003	0.288	0.061	0.036	0.001
	Declustered Input	0.003	0.288	0.066	0.038	0.002
	alr_SGS_standard	0.003	0.532	0.062	0.033	0.001
	alr_SGS_sequential	0.003	0.546	0.062	0.033	0.001
	clr_SGS_standard	0.005	0.363	0.063	0.022	0.000
	clr_SGS_sequential	0.003	0.375	0.064	0.022	0.001

Table 3: Input data and four simulation methods basic statistics.

8. Discussion and Future Work

With the methodology explored, the kriging and simulation have not performed adequately. Production of maximum values from the simulations higher than that of the input data is a problem. The very different results derived from the alr and clr transform method (especially when compared to the mean of the inputs) is also concerning.

The following comments can be made regarding the means of the simulations versus the expected (i.e. declustered means)

- Bitumen – alr lower, clr equivalent;
- Coarse – alr higher, clr much higher;
- Fines – alr lower, clr much lower; and
- Water – alr and clr reasonable.

It appears that, for as yet undetermined reasons, the major component of the solids (coarse) has been over-stated, and the minor component (fines) has been slightly under-stated. This may be due to the geometric averaging discussed above for logratios as the coarse fraction is the major component. There may also be distortion on the back-transform where the relatively large geometric mean for the coarse fraction is divided by the relatively small sum of the geometric means of all the components.

Tolosana-Delgado et al. (2008) in their worked example use the ilr transform and co-kriging, not the alr and clr transforms and independent kriging. It is unclear if these different approaches are enough to explain the substantially different results from their example, and from the application described here.

Therefore, future work to consider would be the application of full cokriging and cosimulation, and the utilization of the ilr transform in addition to the alr and clr transforms.

9. Conclusions

The methodology presented has the beneficial property that all the variables of interest are positive and sum to unity; however, the results of this study show that a straightforward application of what would be considered 'normal' geostatistical practice leads to erratic reproduction of the input data and bias for the data set analyzed.

The compositional data approach is yet to find acceptance in geostatistics, which is not surprising, given 1) the complexity of the numerous transforms (normalizing, logratio, Gaussian) 2) the difficulty of dealing with zeros (which are very common), and 3) the inconsistent reproduction of basic input statistics when following a standard simulation/kriging approach.

References

Aitchison, J., 1986, The statistical analysis of compositional data, The Blackburn Press, 416p.

Aitchison, J., 1999, Logratios and Natural Laws in Compositional Data Analysis: Mathematical Geology, Vol. 31, No. 5, pp. 563-580.

Aitchison, J., Barcelo-Vidal, C. and Pawlowsky-Glahn, V., 2002, Some comments on compositional data analysis in archaeometry, in particular the fallacies in Tangri and Wright's dismissal of logratio analysis, Archaeometry, Vol. 44, No. 2, pp. 295-304.

Boisvert, J.B., Rossi, M.E. and Deutsch, C.V, 2009, Multivariate geostatistical simulation of proportions and nonadditive geometallurgical variables, Eleventh Annual Report of the Centre for Computational Geostatistics, University of Alberta, pp. 303-1 – 303-8.

Egozcue, J. J., Pawlowsky-Glahn, V., Mateu-Figueras, G., and Barcelo-Vidal, C., 2003, Isometric logratio transformations for compositional data analysis: Mathematical Geology, Vol. 35, No. 3, pp. 279-300.

Lan, Z., Leuangthong, O. and Deutsch, C.V., 2006. Why Logratios are a Bad Idea for Multiscale Facies Modeling, Eighth Annual Report of the Centre for Computational Geostatistics, University of Alberta, pp. 211-1 – 211-11.

Manchuk, J. G., 2008, Guide to geostatistics with compositional data, Centre for Computational Geostatistics (CCG) Guidebook Series, University of Alberta, Vol. 7, 34p.

Martin-Fernandez, J.A., Barcelo-Vidal, C. and Pawlowsky-Glahn, V., 2000, Zero replacement in compositional data sets, Studies in Classification, Data Analysis, and Knowledge Organization (eds Kiers, H., Rasson, J., Groenen, P. and Shader, M.), Springer-Verlag, Berlin, pp. 155-160.

Martin-Fernandez, J.A., Barcelo-Vidal, C., Pawlowsky-Glahn, V., Kovacs, L.O. and Kovacs G.P., 2005, Subcompositional patterns in Cenozoic volcanic rocks of Hungary: Mathematical Geology, Vol. 37, No. 7, pp.729-752.

Pawlowsky-Glahn, V. and Egozcue, J.J., 2006, Compositional data and their analysis: an introduction, in Compositional data analysis in the geosciences: from theory to practice (eds. Buccianti, A., Mateu_Figueras, G. and Pawlowsky-Glahn, V.), Geological Society, London, Special Publications, 264, pp. 1-10.

Pawlowsky-Glahn, V. and Olea, R., 2004, Geostatistical analysis of compositional data, Oxford University Press, 304p.

Romanova, U.G., Yarranton, H.W. and Schramm, L.L., 2003. "Towards the Improvement of the Efficiency of Oil Sands Froth Treatment". *Canadian International Petroleum Conference, 2003.* Paper 2003-010.

Thomas, C.W. and Aitchison, J., 2005, Compositional data analysis of geological variability and process: A case study: Mathematical Geology, Vol. 37, No. 7, pp. 753-772.

Tolosana-Delgado, R., 2008, Compositional data analysis in a nutshell, University of Gottingen on-line reference, http://www.sediment.uni-goettingen.de/staff/tolosana/extra/CoDaNutshell.pdf

Tolosana-Delgado, R., Egozcue, J.J. and Pawlowsky-Glahn, V., 2008, Cokriging of Compositions: Log-ratios and Unbiasedness, Proceedings of the Eighth International Geostatistics Congress (eds. Ortiz, J.M and Emery, X.), pp. 299-308.

Wik, S., Sparks, B.D., Ng, S., Tu, Y., Li, Z., Chung, K.H., and Kotlyar, L.S., 2008. "Effect of bitumen composition and process water chemistry on model oilsands separation using a warm slurry extraction process simulation". Fuel, Vol. 87, Issue 7, pp. 1413-1421.

Figures

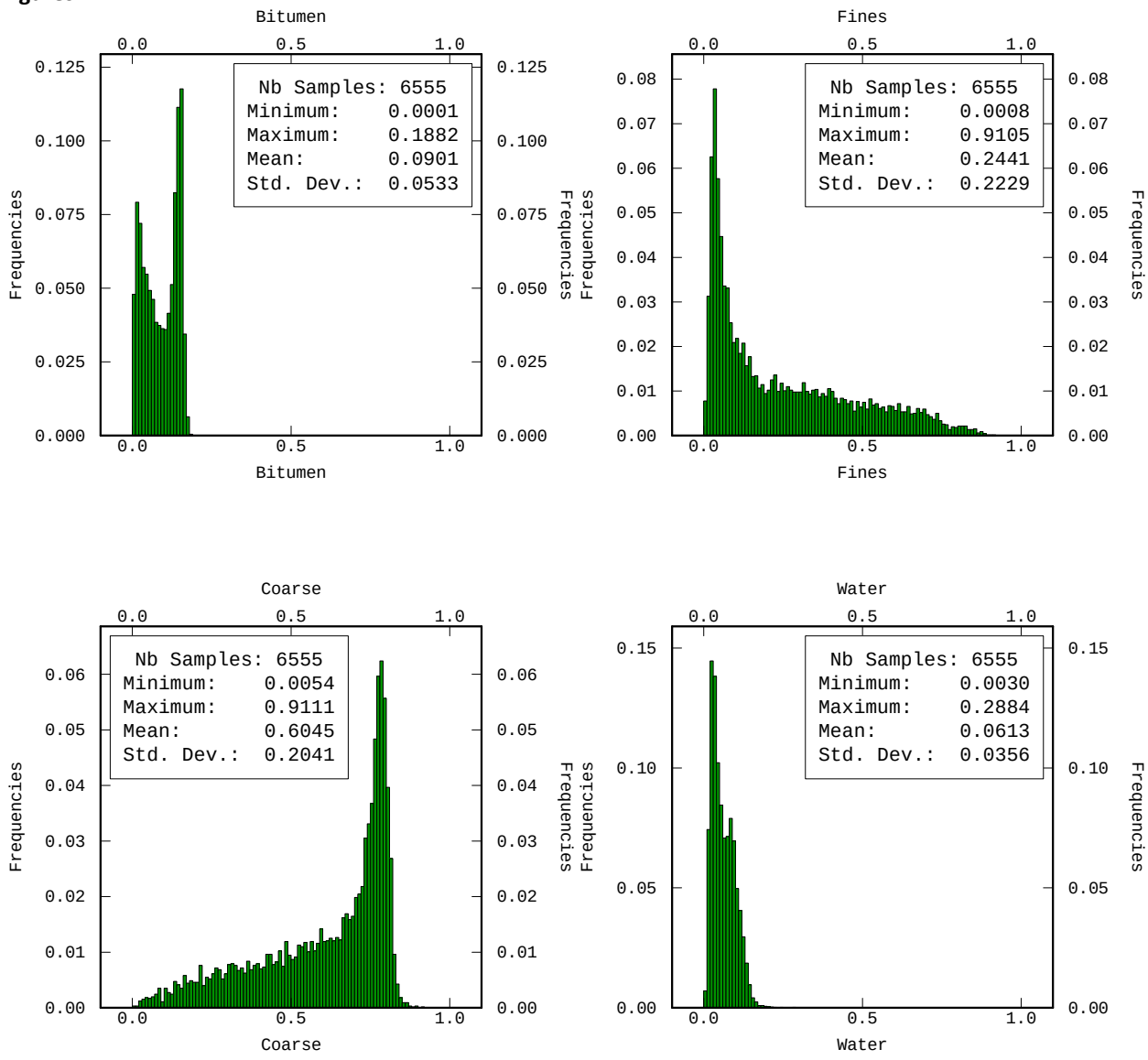


Figure 1: Histograms and statistics of raw data.

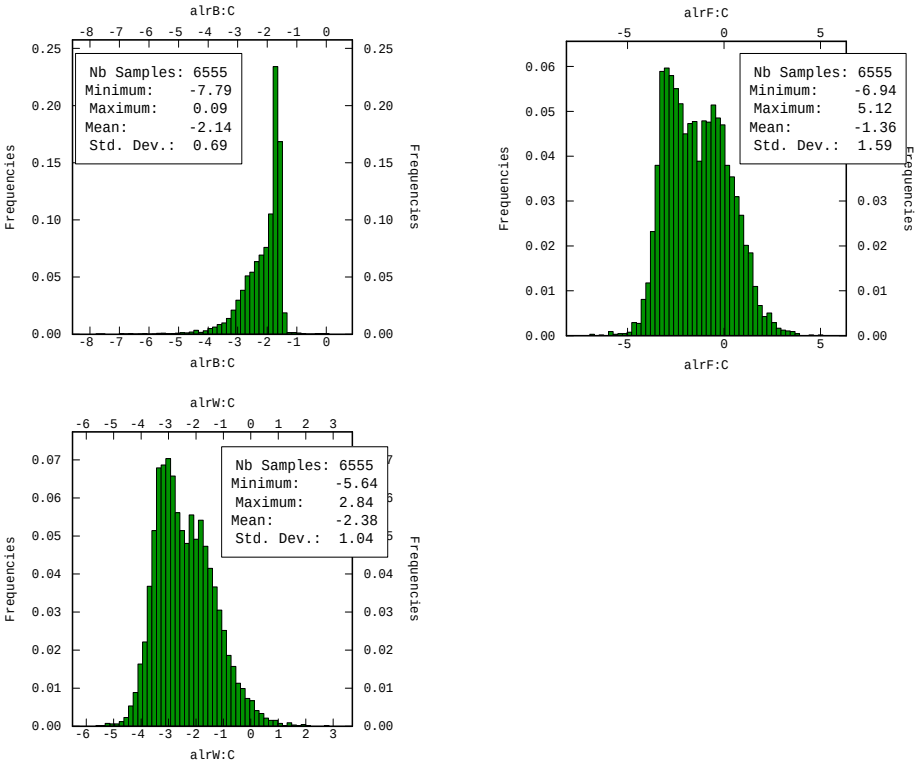


Figure 2: Histograms and statistics of alr-transformed data.

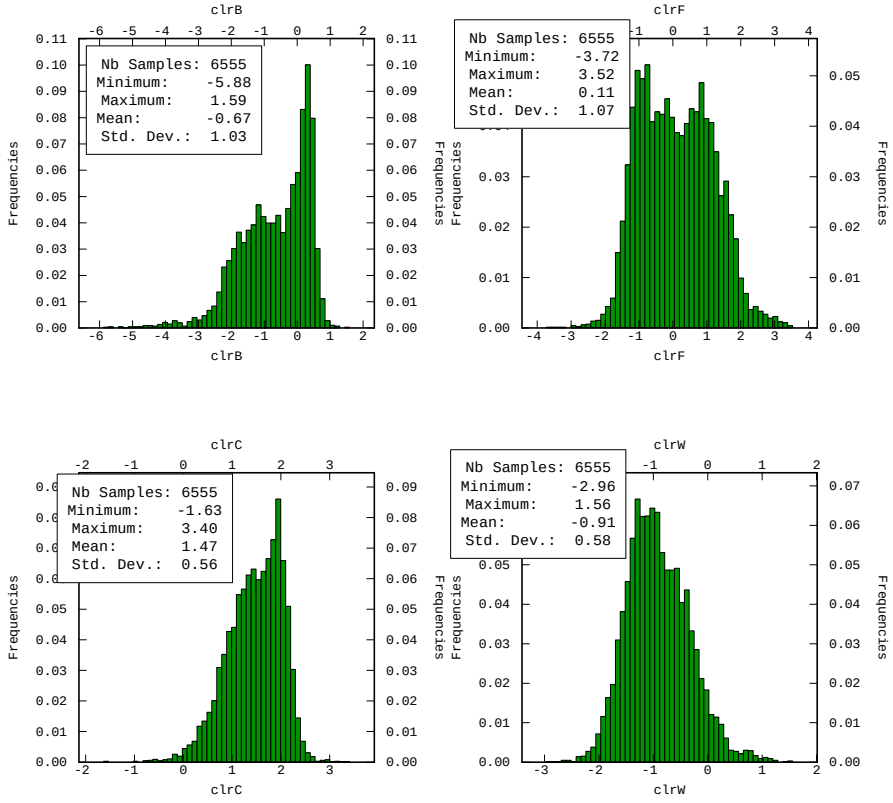


Figure 3: Histograms and statistics of clr-transformed data.

Variogram Models									
	Mathematical	Nugget	Nugget	Range			Sill		
Domain	Rotation *	(C0)	(as %)	Major	Semi	Minor	Sill	(as %)	Structure
alrB:C	0,0,0	0.11	23.1%	200	150	20	0.273	57.4%	1
				450	450	25	0.076	16.0%	2
				10000	10000	30	0.017	3.6%	3
alrF:C	0,0,0	0.46	18.2%	150	60	15	1	39.5%	1
				250	170	25	0.77	30.4%	2
				650	1200	40	0.3	11.9%	3
alrW:C	0,0,0	0.30	27.2%	180	125	25	0.666	60.5%	1
				550	920	30	0.135	12.3%	2
clrB	45,0,0	0.3	28.3%	200	140	20	0.63	59.4%	1
				840	840	30	0.131	12.3%	2
clrC	45,0,0	0.115	36.4%	180	180	20	0.148	46.8%	1
				780	700	25	0.053	16.8%	2
clrF	45,0,0	0.324	28.5%	150	150	10	0.618	54.4%	1
				860	860	25	0.193	17.0%	2
clrW	45,0,0	0.118	35.3%	150	100	20	0.177	53.0%	1
				710	650	25	0.039	11.7%	2

* Rotation 0,0,0 means major direction towards 090

* Rotation 45,0,0 means major direction towards 045

Table 4: Variogram Models. All structures are spherical.

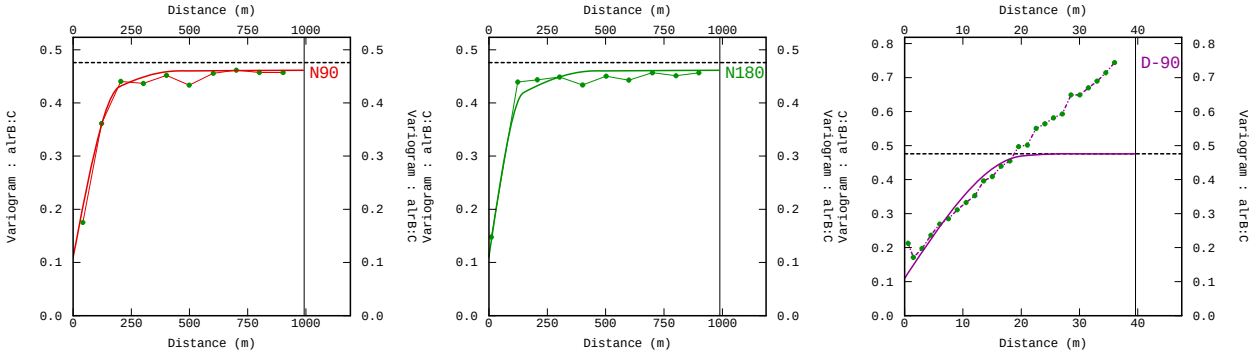


Figure 4: Example alr variogram models (alrB:C). (red=major direction, green=semi major direction, purple = minor direction)

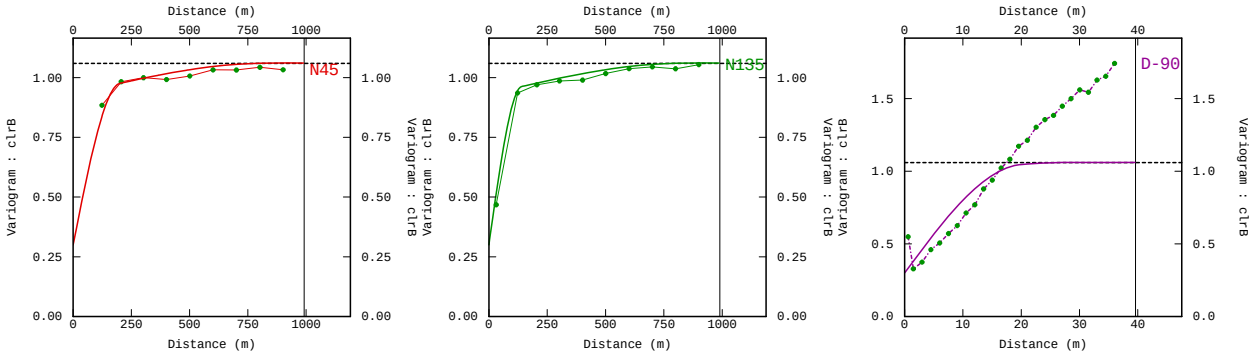


Figure 5: Example clr variogram model (clrB). (red=major direction, green=semi major direction, purple = minor direction).

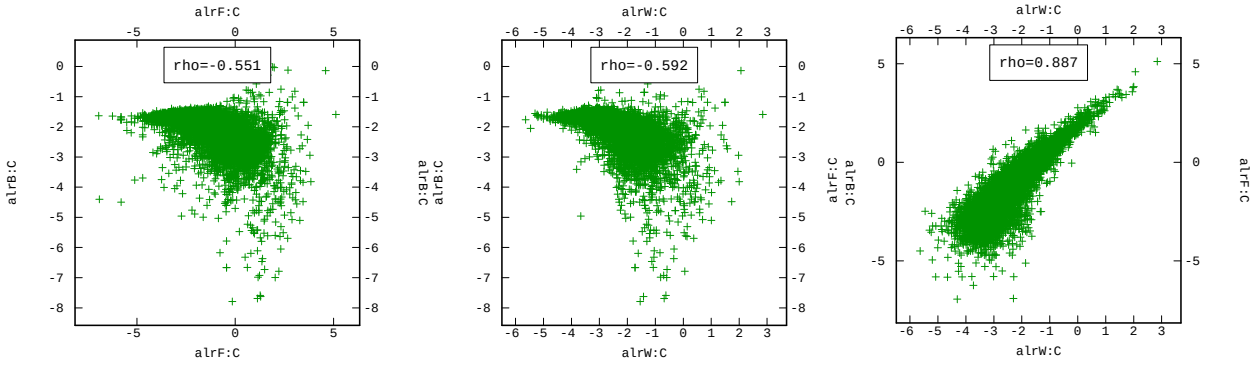


Figure 6: alr scatterplots (alrF:C/B:C left, alr W:C/B:C middle, alr W:C/F:C right)

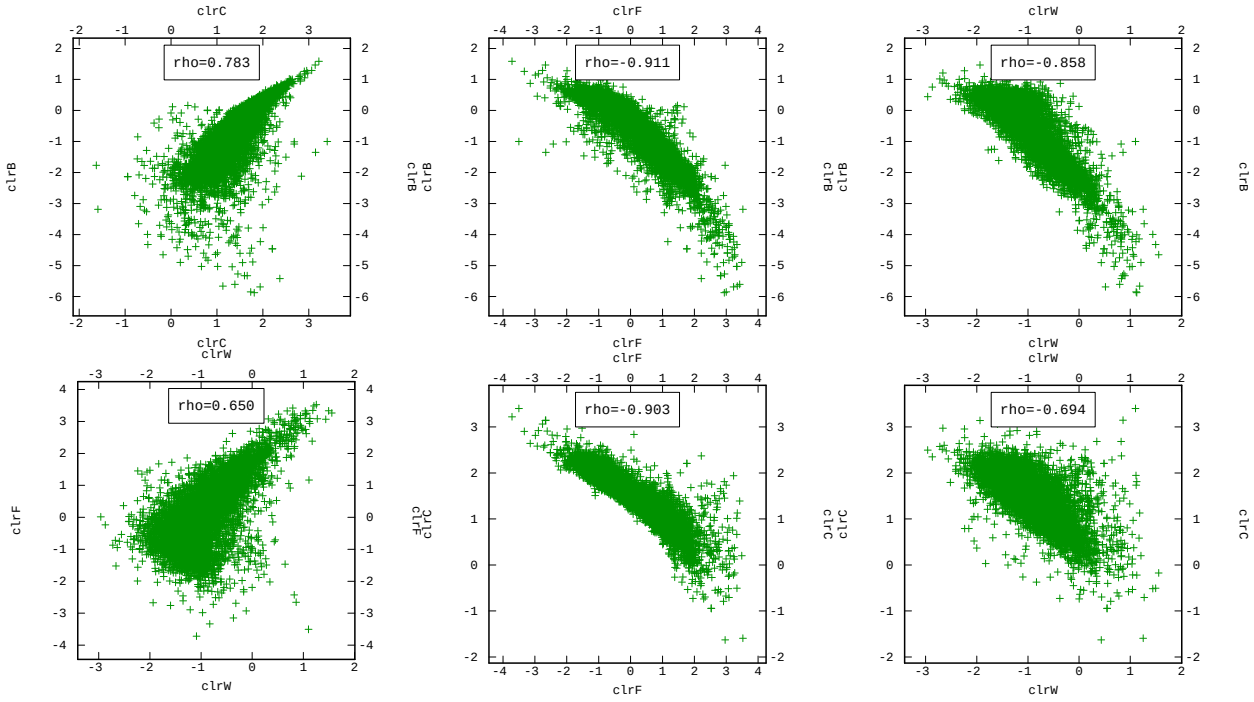


Figure 7: clr scatterplots (clrC/B upper left, clrF/B upper middle, clrW/B upper right, clrW/F lower left, clrF/C lower middle, clrW/C lower right).

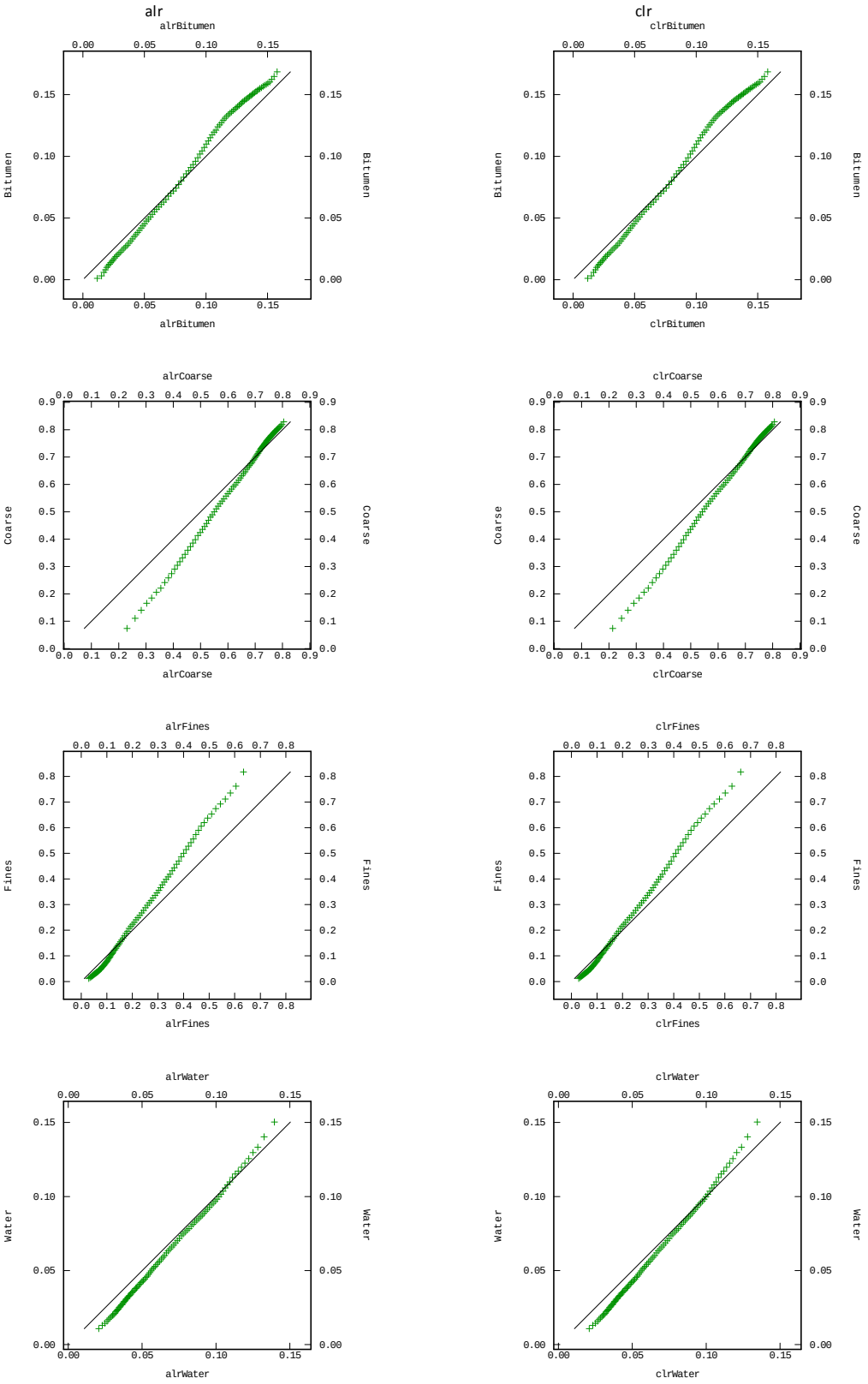


Figure 8: Q-Q plots, drilling v. models, alr left, clr right. Model data on x-axis, drilling on y-axis.

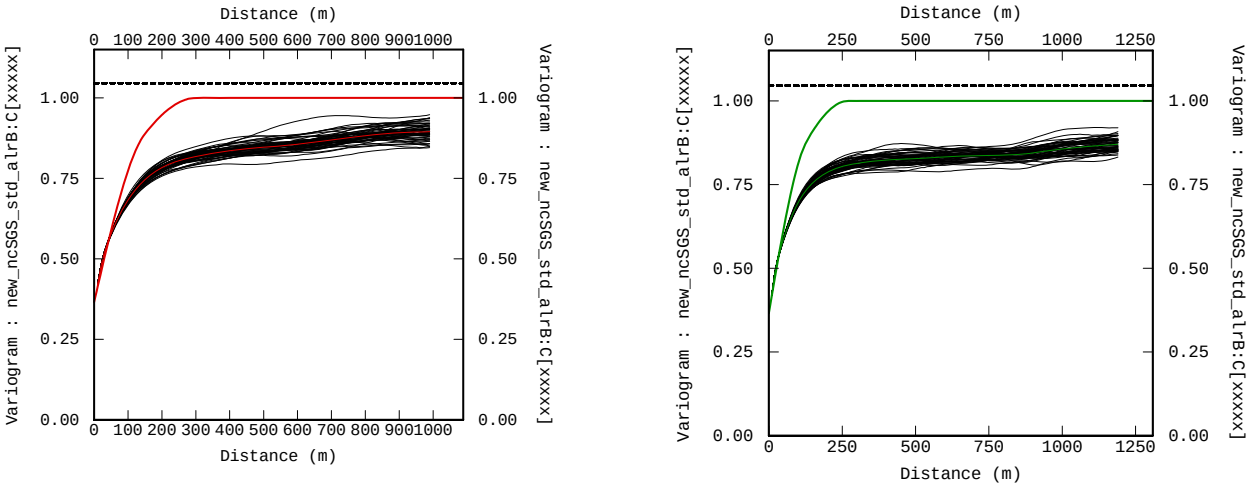


Figure 9: Grid variograms for alrB:C (before Gaussian back-transform).