New approaches to the study of the material composition and geodynamic evolution of the early Precambrian complexes

E V Martynov¹, N E Kozlov¹,², N O Sorokhtin³, T S Marchuk¹

¹Geological Institute of the Kola Science Centre RAS, Apatity, Russia
²Apatity Branch of Murmansk State Technical University, Apatity, Russia
³P.P. Shirshov Institute of Oceanology RAS, Moscow, Russia

mart@geoksc.apatity.ru, kozlov@geoksc.apatity.ru, nsorokhtin@ocean.ru, marchuk@geoksc.apatity.ru

Abstract. Geological interpretation of formation processes of Precambrian assemblages with a complicated structure (reconstruction of geodynamic formation settings of metamorphic rocks protoliths) requires approaches other than those applied in studying Phanerozoic objects. The current paper reports some new approaches to studying the material composition and geodynamic evolution of early Precambrian complexes. They are based on the authors’ techniques of identifying the original nature of metamorphites, statistic modeling of varied features of chemical compositions of Precambrian objects in regard to standards of geodynamic settings in the Phanerozoic and considering them in geodynamic reconstructions.

1. Introduction
The approaches Issues of reconstruction of supracrustal complexes’ protoliths have been studied by numerous Russian and foreign scientists. The following researchers should be named among them: M.R. Bhatia, S.D. Werner, A.F. Grachev, N.L. Dobretsov, N.A. Domaratsky, A.S. Kerr, S.M. Lennon, V.A. Makrygina, A.N. Neyelova, J.A. Pears, A.A. Predovsky, O.M. Rozen, Ch. Roche, A.V. Sidorenko, A. Simonen, S.R. Tailor, V.I. Feldman, Ya.E. Yudovich and many other. This school developed in the 1980-1990s very actively, but now attracts interest of only few scientists, which seems to the authors of this work to be wrong.

As a rule, most petrogeochemical techniques applied to reconstruct settings for the emplacement of the Precambrian complexes’ protoliths have been originally developed for studying settings of Phanerozoic units. These are mainly based on the comparison of selected chemical composition parameters of metamorphic rocks and reference samples of Phanerozoic rocks. A typical drawback of these techniques lies in the overlap of areas occupied by rocks of various groups in the attribute space. For this reason, their mechanical transfer to Precambrian units often leads to an increased uncertainty of results. This needs to be accounted for in further reconstructions of the emplacement settings of protoliths in the Precambrian rock units.

2. Results and Discussion
The results of our studies showed (at a level of statistically significative patterns in geochemical composition of the rocks in geological units) that the Earth’s evolution processes had been evolving nonlinearly. Therefore, the uniformitarianism concept cannot be completely without introducing some corrections applied to the study of the geodynamic settings in the Precambrian. When drawing
parallels between Precambrian and Phanerozoic rock units genetically linked with certain regimes, it is essential to assume the nature of the Precambrian period in the planet evolution on the one hand and some affinity in the style of how the geodynamic settings evolved during the whole geological history, on the other hand. This approach was described as a provision on homological series of geodynamic regimes [1], [2].

It is also important to emphasize that many researchers currently involve data on the content of trace and first of all rare-earth elements, to which the priority is given to solve tasks of geochemical reconstructions. Without denying the possibility of such an approach, the authors of this paper emphasize that the study of the content and distribution of the rock-forming elements in the rocks remains the most important distinctive feature of the methods, which they developed. The information of their content in a rock represents its fundamental property and has to underlie petrogeochemical reconstructions (similarly to the way this is done at the development of various petrogeochemical rock classifications). Besides, applying data on the contents of minor elements when generating reference samples due to often incomparable quality of the analytical data from different laboratories substantially degrades reliability of conclusions derived using these reference samples. In view of the above, the data on the content of trace elements and their isotopes for the solution of the stated tasks in a general case may merely be additional. Although these are of high importance and sometimes have top priority for the solution of other local tasks. It shall be thus noted that, instead of 10 rock-forming elements (SiO₂, TiO₂, Al₂O₃, FeO, Fe₂O₃, MnO, MgO, CaO, Na₂O, and K₂O), we use 8 chemical composition parameters converting parameters (FeO, Fe₂O₃ and MnO) into a new parameter tentatively called ΣFeO. This enables achieving better comparability of the results of chemical analyzes from different laboratories.

The authors of this work have rather long been engaged with this issue [1], [3], [6].

The geochemical approach to the solution of reconstruction tasks applies a databank containing information about chemical composition of the rocks in the Precambrian D = {Dᵢ} and Phanerozoic A = {Aᵢ} metamorphic units, the combinations of which are used as references of the settings of the rock associations.

A convenient mathematical model for describing chemical composition of any geological unit represented by a complex of rocks characterized by n-parameters lies in an n-dimensional random value. For the description of Precambrian and Phanerozoic units, a set of n-dimensional random values {ηᵢ} and {ξⱼ} is respectively used. Each random value ηᵢ is represented by a selection of n-dimensional vectors Yᵢ = {Yᵢk | Yᵢk ∈ Rⁿ} and ξⱼ by a selection Xⱼ = {Xⱼk | Xⱼk ∈ Rⁿ}.  

2.1. Task of modeling distinctions of Precambrian and Phanerozoic units

Suppose a population of heterochronous rock masses consists of two populations, which are represented by the Phanerozoic reference rock associations A = {Aᵢ} and Precambrian metamorphic units D = {Dᵢ}. Their chemical compositions are shown as selections of n-dimensional random values X = {Xᵢ} and Y = {Yⱼ}. For identifying the nature of distinctions in chemical compositions of the Precambrian and Phanerozoic rock masses (targets), it is suggested to apply a method of building a minimal (quantitatively) system of linear functions {Pᵢ}, which optimally, i.e., statistically significantly describes these distinctions.

The search for such distinctions starts with an attempt to describe these using a single linear function P(x) = (c,x), where c ∈ Rⁿ, x ∈ Rⁿ and c are the vector of unit length (|c| = 1). For this purpose it is necessary to solve the following optimization task. The functional J(c) = minₙₙ₀(λ{(c,Xₙ₀)},{(c,Yₙ₀)}), where λ(·) is a Puri-Sen-Tamura statistic [7-8] shall be selected as a quality criterion. Then the optimization task may be formulated as follows:

\[ \max_c J(c) \]

with the following limitations:

\[ M((c,Xₙ₀)) < M((c,Yₙ₀)) \text{ for any pair } (i,j); \]
\[ |c| = 1; \]
\[ J(c) > \chi^2(δ), \]
where $\chi^2(\delta)$ means the $\chi^2$ quantile value of the distribution for a selected significance level $\delta$, $M(.)$ denotes an average of the set $c \in \mathbb{R}^n$. It is suggested to implement the verification of the inequality $M\{(c,X_{ik})\} < M\{(c,Y_{jk})\}$ using the Puri-Sen-Tamura criterion. To solve it, it is possible to use the Nelder-Mead method [9].

If the task has no solution, a conclusion is made of the impossibility to describe statistically significant distinctions between the parameters of the chemical composition for the Precambrian and Phanerozoic targets using one linear function. In this case, an attempt should be undertaken to describe the required distinctions using several linear functions.

In case when the constructed system is made of several linear functions, it should be optimized. This step is defined by a requirement to select most stable, out of all possible, description of the distinctions. For the optimization of the system $\{P_l\}$, it is suggested to use the following method.

The stability level of the solution shall be chosen as a quality criterion in the form of the functional $J = \min\{(c_i,c_j)\}$ for all pairs $(i,j)$ such that $i \neq j$ takes the value in the interval $[-1, 1]$.

Based on the foregoing, the selection of a stable description may be implemented as a result of solution the following optimization task:

$$\max J$$

with the following limitations:

$$M\{(c,X_{ik})\} < M\{(c,Y_{jk})\} \text{ for any pair } (i,j);$$

$$\|c_l\| = 1 \text{ for any } l;$$

$$\Lambda\{(c,X_{ik}),(c,Y_{jk})\} > \chi^2(\delta) \text{ for any pair } (i,j)$$

at the selected significance level $\delta$.

In particular, when studying rocks with basic compositions in the Precambrian and Phanerozoic targets, we managed to describe their distinctions using two linear functions. It is possible to attempt describing the offset of the Precambrian targets relative to the Phanerozoic reference rock units in a multi-dimensional attribute space (Precambrian targets and Phanerozoic reference rock units are represented by sets of figurative points) using a second-order surface separating the Precambrian and Phanerozoic targets [2]. In this case, the criterion defining that this surface is a separating one is the statistical significance of the distinction between the sets of figurative points in each target and the sets of projections for these points on the above surface while the set of figurative points for the Precambrian and Phanerozoic targets are positioned on the opposite sides of this surface (See Fig. 1). We use an optimal separating surface, for which the minimum proximity of the entire set of targets to the surface is maximal.

2.2. Task of accounting for chemical composition features in Precambrian units at the reconstruction of the emplacement settings for their protoliths

The features of the chemical rock compositions for the Precambrian targets are represented by a set of samples of the chemical rock composition parameters $Y = \{Y_j\}$ relative to the Phanerozoic targets represented by a set of samples $X = \{X_i\}$ are defined by a system of linear functions $\{P_l \mid P_l(x) = (c_l,x)\}$ and $\|c_l\| = 1$ derived during the modeling of the nature for the distinctions in the Precambrian and Phanerozoic targets. To account for these features, this research proposes a procedure of multi-variant (variant $V_l$ corresponds to each function $P_l$) reconstruction.

The reconstruction method is based on the technique developed by us that enables reconstructing the emplacement settings for the protoliths of the Precambrian targets through the consideration of the preliminarily established nature of their offsetting chemical composition relative to the Phanerozoic targets. The reliability of the reconstruction results is achieved by reducing the dimension of the attribute space as a consequence of transition from the study of n-dimensional random values to (n–1)-dimensional ones through projecting n-dimensional vectors for the samples of sets $X$ and $Y$ on a hyperplane orthogonal to $c_l$. In this case, a procedure of multi-variant reconstruction is used.

Another method for the reconstruction of emplacement settings for the Precambrian rock protoliths relies on the description of distinctions in chemical compositions of the Precambrian and Phanerozoic...
targets using the second-order surfaces. The simplest way to describe such surface is using the following equation:

$$\frac{2}{1} - \ln(2)\left( S_{11}^{1/2} \right) = - \ln(2)\left( S_{22}^{1/2} \right),$$

where $S_{11}$ and $S_{22}$ are estimates of the covariance matrices, $\bar{X}$ and $\bar{Y}$ are average values of the samples $X$ and $Y$.

By projecting the sample points from $X$ on this surface, the reference settings generate. Similarly, by projecting points of the studied Precambrian target on the given surface, its image is created. The comparison of this image with the constructed references allows estimating the extent of its similarity with each reference. Therefore, it is possible to reconstruct the protolith emplacement setting for the studied Precambrian target (Fig. 1).

**Figure 1.** Illustration on a 2D plane to the method for the reconstruction of emplacement settings for the Precambrian rock protoliths on the basis of the corresponding chemical rock compositions by constructing a second-order surface. In the Figure, the sets $\{a\}$, $\{\ast\}$ and $\{\circ\}$ represent Phanerozoic targets with different emplacement settings (the arcs show reference settings) and $\{\bullet\}$ are the Precambrian targets.

2.3. **Task of modeling linear variability trends in the chemical rock compositions relative to the assigned partial order**

Suppose the set $\{\xi_i\}$ of n-dimensional random values is represented by a combination of samples $Z = \{Z_i\}$, where $Z_i = \{Z_{ik} \mid Z_{ik} \in \mathbb{R}^n\}$. The partial order «<»: $U = \{<Z_i,Z_j> \mid Z_i < Z_j\}$ is assigned for the set $Z^*Z$. If $c \in \mathbb{R}^n$ is an n-dimensional vector, the scalar product $(c, \xi_i)$, where $c \in \mathbb{R}^n$ is a unit length vector ($c = \|1\|$), represents a one-dimensional random value. It is suggested to search for a linear trend relative to an assigned partial order in the form of a solution of the following optimization task [2].

Suppose $J(c) = \min_U \Lambda(\{(c,Z_{ik})\}, \{(c,Z_{jk})\})$, where $\Lambda(\{(c,Z_{ik})\}, \{(c,Z_{jk})\})$ is the Puri-Sen-Tamura statistic value. It is required to find such a unit length vector ($c = \|1\|$) that:

max. $J(c)$

at the following limitations:

- $M(\{(c,Z_{ik})\}) < M(\{(c,Z_{jk})\})$ on the set $U$;
- $c = \|1\|$;
- $\Lambda(\{(c,Z_{ik})\}, \{(c,Z_{jk})\}) > \chi^2(\delta)$ on the set $U$. 

Here, $\chi^2(\delta)$ means the quantile value $\chi^2$, which is the distribution for selected significance level $\delta$ for all pairs $<Z_i,Z_j>$ such that $Z_i<Z_j$. To solve it, it is possible to use the aforementioned Nelder-Mead method [9]. A linear function $F = (c,x)$ represents the variability trend or the chemical rock compositions relative to the assigned partial order.

To illustrate the above stated, we will give the following hypothetical example (Fig. 2). Let us assume that in a two-dimensional attribute space some combination of rock units $Z = \{A,B,C,J,H,D,E,K,M\}$ is assigned, and each of them is represented by a figurative point (for instance, point of averages). The partial order $U = \{B<F, C<D, D<E, H<K\}$ (in the Figure «<», it is shown as «→») for this combination is represented as a graph. It is required to find the linear function $P(c) = (c,x)$ such that the functional $J(c)$ assumes a maximum value.

![Figure 2](image)

**Figure 2.** Illustration on a 2D plane to the task of modeling linear variability trends for the chemical rock compositions relative to the assigned partial order. The points $\{A,B,C,J,H,D,E,K,M\}$ represent projections of points $\{A,B,C,J,H,D,E,K,M\}$ on the P axis assigned by a linear function $P(c) = (c,x)$.

2.4. **Task of estimating the degree of similarity for the geological sites**

As a rule, these chemical rock compositions in the geological sites are represented by heterogeneous samples (polymodal composition distribution, presence of anomalous values, etc.). For this reason, we face a problem of selecting the optimum similarity measure between geological sites. In scientific publications, the following basic similarity measures for the samples $X = \{X_i\}_{i=1,\ldots,n_1}$ and $Y = \{Y_j\}_{j=1,\ldots,n_2}$ are discussed, and namely:

1. minimum local distance measured under the nearest neighbor principle $d(X,Y) = \min_{i,j} d(X_i,Y_j)$, where $d$ is the Euclidean distance between $X_i$ and $Y_j$;

2. maximum local distance measured under the distant neighbor principle $d(X,Y) = \max_{i,j} d(X_i,Y_j)$;

3. centroid distance $d(X,Y) = d(\bar{X},\bar{Y})$;

4. average distance (average relationship) $d(X,Y) = \frac{1}{n_1n_2} \sum_{i=1}^{n_1} \sum_{j=1}^{n_2} d(X_i,Y_j)$;

5. Hausdorff distance $d(X,Y) = \max_i \min_j d(X_i,Y_j), \max_j \min_i d(X_i,Y_j)$;
6. Mahalanobis distance $d^2(X,Y) = (\bar{X} - \bar{Y})^T S^{-1} (\bar{X} - \bar{Y})$, where $S$ is the estimate of generalized covariance matrix ($S = \frac{S_1}{n_1} + \frac{S_2}{n_2}$, where $S_1$ and $S_2$ are covariance matrices for $X$ and $Y$).

The practice of our studies demonstrates that the application of any of these measures due to heterogeneity of samples characterizing the studied geological sites usually does not allow obtaining a stable and statistically substantiated result. Thus, we were forced to build another similarity measure for the geological sites different from the aforementioned ones. This measure, as the practice of our studies showed, turned out to be more applicable for the solution of the issues under study. For the construction of this similarity, one of the objects is previously divided into a set of homogeneous groups [10].

If the studied object $W$ is represented by a selection of $n$-dimensional vectors $X = \{x_i \in \mathbb{R}^n\}$ and there is a population of reference targets $V = \{V_t\}$ represented by samples $\{Y_t\}$, where $Y_t = \{y_{tj} \in \mathbb{R}^n\}$, a task emerges to select a target closest to the target $W$ from $V$ (let us call it a homolog of the target $W$). To solve this task, every sample $Y_t$ is subdivided into homogenous groups $Y_t = \{Y_{tk}\}$ and the distances $R_t = \{r_{it}\}$ from points of the set $X$ and to the set $Y_t$. Here, $r_{it} = \min_k \rho_{itk}$, and $\rho_{itk}$ are minimum distances from the vector $x_i$ and the convex shell of the homogenous group $Y_{tk}$. The median $\text{Me}(R_t)$ as an estimate of the average set $R_t$ is accepted as the estimate of the degree of similarity between the targets $W$ and $V_t$. But the selection of the homolog $V_t \in V$ for target $W$ does not mean that the solution derived on the basis of $\min_t \text{Me}(R_t)$ is a unique one. The point is that there is a possibility of presence of not one but the entire group of targets from the set $V$ whose degrees of similarity with the target $W$ is statistically insignificantly different from each other at the significance level selected by us. In this case, this group of targets should be considered as homologs of the target $W$. For the verification of a hypothesis on the presence of not one but several homologs it is suggested to use the already mentioned Puri-Sen-Tamura criterion.

3. Conclusions

The approaches proposed by the authors allow considerably increasing reliability of reconstructed geodynamic formation settings of metamorphic rocks protoliths. It is defined by the following “factors”: first, by the reliability degree of conducted identification of the metamorphite protonature; second, by the selected method of modeling the varied features of chemical compositions of Precambrian and Phanerozoic objects; third, by the technique of considering this variability and selected similarity degree of the objects.

Acknowledgements

The study has been carried out in the framework of the state orders № 0226-2019-0052 and № 0149-2019-0005.

References

[1] Kozlov N E 1995 Material composition of metamorphic complexes high-baric granulite belts and the formation problem of their protoliths (as exemplified by Lapland granulites). Abstract of doctoral diss. p 36
[2] Kozlov N E, Sorokhtin N O, Martynov E V 2018 Geodynamic Evolution and Metallogeny of Archean Structural and Compositional Complexes in the Northwestern Russian Arctic Minerals 8(12) p 573
[3] Kozlov N E, Martynov E V 1992 On the issue of the primary nature of the Lapland granulite belt metamorphites (Baltic Shield) Geochemistry 1 128-133
[4] Kozlov N E, Martynov E V 1995 Granulite belts in Precambrian and some aspects of petro-geochemical evolution of proto-subduction zones during Earth evolution Geology and geophysics 36(12) 105-112
[5] Kozlov N E, Ivanov A A, Martynov E V, Peredovsky A A 2002 Modeling of formation and
evolution conditions of the Precambrian complexes based on a study of geology of the material composition in their composing rock associations Geology and economic minerals of the Kola Peninsula 3 96-110

[6] Kozlov N E, Martynov E V, Ivanov A A 1999 Features of petro-geochemical distinctions in basic rocks of ensialic and ensimatic complexes (comparative analysis of the Phanerozoic and Precambrian Geochemistry 6 582-588

[7] Purî M L, Sen P K 1971 Nonparametric methods in multivariate analysis (New York: John Wiley & Sons) p 677

[8] Tamura R 1966 Multivariate nonparametric several–sample tests The Annals of Mathematical Statistics 37(3) 611-618

[9] Nelder J A, Mead R A 1964 Simplex Method for Function Minimisation Computer J. 7 308-313

[10] Martynov E V 2008 Math methods of modeling geological processes and phenomena parameters Teaching guide for direction 130100 “Geology and exploration of economic minerals” (Murmansk) p 136