Совместный выпуск

№ 1
2026

ВЕСТНИК
Национальной инженерной академии Республики Казахстан

ВЫЧИСЛИТЕЛЬНЫЕ ТЕХНОЛОГИИ
Федеральный исследовательский центр информационных и вычислительных технологий

ВЕСТНИК
Национальной инженерной академии
Республики Казахстан

Главный редактор
академик Б. Т. Жумагулов

ВЫЧИСЛИТЕЛЬНЫЕ ТЕХНОЛОГИИ

Федеральный исследовательский центр
информационных и вычислительных технологий

Главный редактор
академик Ю. И. Шокин

по материалам Международной конференции
«Вычислительные и информационные технологии

в науке, технике и образовании»
(CITech-2026)

8-9 января 2026 года

Алматы, 2026

ПРОГРАММНЫЙ КОМИТЕТ

Сопредседатели:

1. **Туймебаев Жансеит Кансеитович,** Председатель Правления - Ректор Казахского национального университета им. аль-Фараби, Казахстан
2. **Жумагулов Бакытжан Турсынович,** академик НИА и НАН РК, Президент Национальной инженерной академий Республики Казахстан, Казахстан

Члены:

1. **Абдибеков Уалихан Сейдильдаевич**, Казахский национальный университет им. аль-Фараби, Казахстан
2. **Арипов Мирсаид**, профессор, Национальный университет Узбекистана им. М. Улугбека, Узбекистан
3. **Ахмед-Заки Дархан**, профессор, ректор ЮКУ им. Ауэзова, Казахстан
4. **Барахнин Владимир Борисович**, д.т.н., профессор, Федеральный исследовательский центр информационных и вычислительных технологий (ФИЦ ИВТ), Россия
5. **Бектемесов Мактагали Абдимажитович**, директор института информационных и вычислительных технологий, Казахстан
6. **Бычков Игорь Вячеславович**, академик РАН, Институт динамики систем и теории управления СО РАН, Россия
7. **Вабищевич Петр Николаевич**, д.ф.-м.н., профессор, Московский государственный университет им. Ломоносова, Россия
8. **Василевский Юрий Викторович**, д.ф.-м.н., профессор, директор Института вычислительной математики им. Г. И. Марчука Российской академии наук
9. **Войцик Вальдемар**, профессор, Люблинский технический университет, Польша
10. **Гусев Олег Игоревич**, к.ф.-м.н., заместитель директора по научной работе ФИЦ ИВТ, Новосибирск, Россия.
11. **Есекеев Куанышбек Бахытбекович**, советник Президента Республики Казахстан, Казахстан
12. **Ешмамматова Дильфуза Бахромовна**, д.ф.-м.н., профессор, Ташкентский государственный транспортный университет, Узбекистан
13. **Жакебаев Даурен Бакытбекулы**, член-корреспондент НИА РК, Казахский национальный университет им. аль-Фараби, Казахстан
14. **Имангалиев Ернар Имангалиевич**, проректор ЮКУ им. Ауэзова, Казахстан
15. **Иманкулов Тимур Сакенович**, Казахский национальный университет им. аль-Фараби, Казахстан
16. **Исахов Алибек Абдиашимович**, Казахстанско-Британский технический университет, Казахстан
17. **Кабанихин Сергей Игоревич**, член-корреспондент РАН, Институт вычислительной математики и математической геофизики СО РАН, Россия
18. **Кальменов Тынысбек Шарипович**, академик НАН РК, Институт математики и математического моделирования, Казахстан
19. **Калимолдаев Максат Нурадилович**, академик НАН РК, Институт Информационных и Вычислительных Технологий, Казахстан
20. **Калтаев Айдархан Жусупбекович**, академик НАН РК, Казахский национальный технический университет им. Сатпаева, Казахстан
21. **Картбаев Тимур Саатдинович**, Казахский женский педагогический университет, Казахстан
22. **Кристоф Жосеран**, Политехнический Институт Парижа, Франция

23. **Кыдырбекулы Алматбек Балгабекулы**, директор Научно-исследовательского института математики и механики
24. **Мамырбаев Оркен Жумажанович**, Институт информационных и вычислительных технологий, Казахстан
25. **Медведев Сергей Борисович**, д.ф.-м.н., директор ФИЦ ИВТ, Новосибирск, Россия
26. **Милошевич Хранислав**, профессор, Университет Приштины, Сербия
27. **Москвичев Владимир Викторович**, д.т.н., профессор, ФИЦ ИВТ, Новосибирск, Россия
28. **Ольшанский Максим Александрович**, д.ф.-м.н., профессор, Хьюстонский университет, США
29. **Отелбаев Мухтарбай Отелбаевич**, академик НАН РК, Институт математики и математического моделирования, Казахстан
30. **Павлов Сергей Владимирович**, Винницкий технический университет, Украйна
31. **Пейман Гиви**, профессор, Университет Питтсбурга, США
32. **Пестунов Игорь Алексеевич**, к.ф.-м.н., доцент, ФИЦ ИВТ, Новосибирск, Россия
33. **Потапов Вадим Петрович,** д.т.н., профессор, ФИЦ ИВТ, Новосибирск, Россия
34. **Потатуркин Олег**, профессор, Институт автоматики и электрометрии СО РАН, Россия
35. **Реш Майкл**, профессор, высокопроизводительный вычислительный центр в Штутгарте, Германия
36. **Рознер Карл**, профессор, Технологический университет Дармштадта, Германия
37. **Рябко Борис Яковлевич**, д.т.н., профессор, ФИЦ ИВТ, Новосибирск, Россия
38. **Садыбеков Махмуд Абдысаметович**, академик НАН РК, Институт математики и математического моделирования, Казахстан
39. **Садовский Владимир**, профессор, Институт вычислительного моделирования СО РАН, Россия
40. **Смагин Сергей**, член-корреспондент РАН, Компьютерный центр ДВО РАН, Россия
41. **Сойфер Виктор**, академик РАН, Самарский государственный аэрокосмический университет им. Королева, Россия
42. **Стемпковский Александр**, член-корреспондент РАН, Институт проблем проектирования в Микроэлектронике РАН, Россия
43. **Чубаров Леонид Борисович**, д.ф.-м.н., профессор, ФИЦ ИВТ, Новосибирск, Россия
44. **Темирбеков Нурлан Муханович**, академик НИА и НАН РК, Казахский национальный университет им. аль-Фараби, Казахстан
45. **Томас Бёниш**, высокопроизводительный вычислительный центр в Штутгарте, Германия
46. **Турицын Сергей**, профессор, Университет Астон, Великобритания
47. **Урмашев Байдаулет Амантаевич**, Казахский национальный университет им. аль-Фараби, Казахстан
48. **Федорук Михаил Петрович**, академик РАН, Новосибирский государственный университет, Россия
49. **Хабаши Вагди Джордж**, профессор, Университет Макгилла, Канада
50. **Шайдуров Владимир**, член-корреспондент РАН, Институт вычислительного моделирования СО РАН, Россия
51. **Шредер Вольфганг**, профессор, RTWH Ахенский университет, Германия
52. **Хужаёров Бахтиёр Хужаёрович**, д.ф.-м.н., профессор, Самаркандский государственный университет им. Шарофа Рашидова, Узбекистан

Investigation of Computational Algorithms Describing Transient Regimes for the Unsteady Navier–Stokes Equations

A.K. MURATOVA[1], N.M. TEMIRBEKOV[2]

[1]Al-Farabi Kazakh National University, Almaty, Kazakhstan
[2]Al-Farabi Kazakh National University, Almaty, Kazakhstan
e-mail: aluamuratova1@gmail.com

The problem of numerical simulation of two-dimensional unsteady flows of a viscous incompressible fluid in a bounded domain is considered. Such problems arise in the investigation of flows in closed and semi-bounded regions, in the analysis of vortex structures, as well as in hydrodynamics and thermophysics problems related to engineering and applied applications.

The mathematical model of these processes is given by the Navier–Stokes equations, which describe the time evolution of the velocity and pressure fields. These equations constitute a nonlinear system of second-order partial differential equations and, in the general case, do not admit analytical solutions. Therefore, numerical methods are widely used for their study.

A significant feature of the incompressible Navier–Stokes equations is the presence of the incompressibility condition, which imposes an additional constraint on the velocity field and leads to the appearance of pressure as an auxiliary unknown. The pressure does not possess its own evolution equation and is determined from the divergence-free condition of the velocity field, which substantially complicates the numerical solution of the problem.

One of the effective approaches to solving this problem is the splitting method based on physical processes [1]. The main idea of such methods consists in the separate treatment of convection, viscosity, and pressure effects at each time step. This allows one to simplify the numerical implementation of the algorithm and to ensure the fulfillment of the incompressibility condition.

In the present work, a numerical algorithm based on the splitting method is considered for solving the two-dimensional unsteady Navier–Stokes equations. The algorithm is implemented on a regular rectangular grid and is aimed at modeling vortex flows in a bounded domain.

1. Problem Statement

The problem of numerical simulation of an unsteady flow of a viscous incompressible fluid in a two-dimensional bounded domain $\Omega \subset \mathbb{R}^2$ with a smooth boundary $\partial\Omega$ over the time interval $t \in (0, T)$ is considered, where $T > 0$ is a given final time.

The fluid motion is described by the incompressible Navier–Stokes equations:

$$\frac{\partial \mathbf{v}}{\partial t} + (\mathbf{v} \cdot \nabla)\mathbf{v} + \nabla p = \frac{1}{Re}\,\Delta \mathbf{v}, \qquad \nabla \cdot \mathbf{v} = 0, \tag{1}$$

where $\mathbf{v}(x, y, t) = (u(x, y, t),\, v(x, y, t))$ is the fluid velocity vector, $p(x, y, t)$ is the pressure, Re is the Reynolds number, ∇ is the gradient operator, $\nabla \cdot (\cdot)$ is the divergence operator, and Δ is the Laplace operator.

The first equation of system (1) represents the momentum conservation law, taking into account the contributions of convective transport, pressure, and viscous diffusion. The second equation is the incompressibility condition, which imposes an additional constraint on the velocity field.

The problem is supplemented with initial conditions specifying the velocity distribution at the initial time:

$$\mathbf{v}(x, y, 0) = \mathbf{v}_0(x, y), \qquad (x, y) \in \Omega, \tag{2}$$

where $\mathbf{v}_0$ is a given vector field.

On the boundary $\partial\Omega$, no-slip boundary conditions corresponding to the absence of fluid slip along solid walls are imposed:

$$\mathbf{v}|_{\partial\Omega} = 0. \tag{3}$$

Such conditions are standard in the modeling of viscous fluid flows in closed and bounded domains.

It should be noted that the pressure $p(x, y, t)$ in the incompressible Navier–Stokes equations does not possess its own evolution equation and is determined indirectly from the incompressibility condition. This fact leads to the necessity of a special numerical approach for pressure computation, which ensures the fulfillment of the condition $\nabla \cdot \mathbf{v} = 0$ at each time step.

In the general case, system (1)–(3) represents a nonlinear system of second-order partial differential equations which, except for certain special cases, does not admit analytical solutions. Therefore, numerical methods are employed for its investigation.

2. Numerical Method

To solve the unsteady Navier–Stokes equations numerically, the splitting method based on physical processes is employed. The main idea of the method consists in the separate treatment, at each time step, of convective and viscous effects as well as pressure, which makes it possible to ensure the fulfillment of the incompressibility condition for the velocity field.

Let the discrete velocity values $\mathbf{v}^n = (u^n, v^n)$ be known at time t^n. The transition to the next time layer $t^{n+1} = t^n + \tau$ is carried out in three stages.

I. Computation of the intermediate velocity field.

At the first stage, the pressure is neglected, and the intermediate velocity $\tilde{\mathbf{v}}$ is determined from the equations

$$\frac{\tilde{\mathbf{v}} - \mathbf{v}^n}{\tau} + (\mathbf{v}^n \cdot \nabla)\mathbf{v}^n = \frac{1}{Re}\,\Delta\mathbf{v}^n. \tag{4}$$

Conservative finite-difference approximations are used for the velocity components. In particular, for the u-component we have

$$\begin{aligned}\frac{\tilde{u}_{i+\frac12,j} - u^n_{i+\frac12,j}}{\tau} - \frac{(u^n_{i,j})^2 - (u^n_{i+1,j})^2}{h_1} + \frac{(uv)^n_{i+\frac12,j-\frac12} - (uv)^n_{i+\frac12,j+\frac12}}{h_2} \\ = \frac{1}{Re}\left(\frac{u^n_{i+\frac32,j} - 2u^n_{i+\frac12,j} + u^n_{i-\frac12,j}}{h_1^2} + \frac{u^n_{i+\frac12,j+1} - 2u^n_{i+\frac12,j} + u^n_{i+\frac12,j-1}}{h_2^2}\right),\end{aligned} \tag{5}$$

and for the component v

$$\begin{aligned}\frac{\tilde{v}_{i,j+\frac12} - v^n_{i,j+\frac12}}{\tau} - \frac{(uv)^n_{i-\frac12,j+\frac12} - (uv)^n_{i+\frac12,j+\frac12}}{h_1} + \frac{(v^n_{i,j})^2 - (v^n_{i,j+1})^2}{h_2} \\ = \frac{1}{Re}\left(\frac{v^n_{i+1,j+\frac12} - 2v^n_{i,j+\frac12} + v^n_{i-1,j+\frac12}}{h_1^2} + \frac{v^n_{i,j+\frac32} - 2v^n_{i,j+\frac12} + v^n_{i,j-\frac12}}{h_2^2}\right).\end{aligned} \tag{6}$$

II. Velocity correction by pressure.

At this stage, an elliptic problem for the pressure, derived earlier, is solved. After the pressure has been computed, the corrected velocity field is determined by the formulas

$$\begin{aligned} u^{n+1}_{i+\frac{1}{2},j} &= \tilde{u}_{i+\frac{1}{2},j} - \tau \frac{p^{n+1}_{i+1,j} - p^{n+1}_{i,j}}{h_1}, \\ v^{n+1}_{i,j+\frac{1}{2}} &= \tilde{v}_{i,j+\frac{1}{2}} - \tau \frac{p^{n+1}_{i,j+1} - p^{n+1}_{i,j}}{h_2}. \end{aligned} \tag{7}$$

III. Incompressibility condition and pressure equation.

The velocity field $\mathbf{v}^{n+1}$ must satisfy the incompressibility condition

$$\nabla_h \cdot \mathbf{v}^{n+1} = 0. \tag{8}$$

Substituting the velocity correction formulas into the incompressibility condition, we obtain the pressure equation, which in discrete form is approximated by the finite-difference scheme

$$\frac{p^{n+1}_{i+1,j} - 2p^{n+1}_{i,j} + p^{n+1}_{i-1,j}}{h_1^2} + \frac{p^{n+1}_{i,j+1} - 2p^{n+1}_{i,j} + p^{n+1}_{i,j-1}}{h_2^2} = \frac{1}{\tau}\left(\frac{\tilde{u}_{i+\frac{1}{2},j} - \tilde{u}_{i-\frac{1}{2},j}}{h_1} + \frac{\tilde{v}_{i,j+\frac{1}{2}} - \tilde{v}_{i,j-\frac{1}{2}}}{h_2} \right). \tag{9}$$

The solution of this equation makes it possible to determine the pressure, after which the velocity correction at the new time level is completed.

3. Boundary conditions for the velocity field.

The splitting scheme is closed by approximating the boundary conditions for the velocity components.

On the lower boundary of the domain, the following conditions are imposed:

$$u^{n+1}_{i+\frac{1}{2},\,\frac{1}{2}} = 0, \qquad v^{n+1}_{i,\,\frac{1}{2}} = 0, \qquad i = 1, \ldots, N_1. \tag{7}$$

On the left boundary, the discrete boundary conditions take the form

$$u^{n+1}_{\frac{1}{2},j} = 0, \qquad v^{n+1}_{\frac{1}{2},j+\frac{1}{2}} = 0, \qquad j = 1, \ldots, N_2. \tag{8}$$

On the upper boundary of the square domain, the following conditions are satisfied:

$$u^{n+1}_{i+\frac{1}{2},\,N_2+\frac{1}{2}} = 0, \qquad v^{n+1}_{i,\,N_2+\frac{1}{2}} = 0, \qquad i = 1, \ldots, N_1. \tag{9}$$

On the right boundary, the boundary conditions are given by

$$u^{n+1}_{N_1+\frac{1}{2},j} = 0, \qquad v^{n+1}_{N_1+\frac{1}{2},j+\frac{1}{2}} = 0, \qquad j = 1, \ldots, N_2. \tag{10}$$

Thus, both the normal and tangential components of the velocity are fully prescribed on the boundary of the domain, which ensures the correct enforcement of the no-slip condition within the finite-difference scheme.

4. Finite-difference elliptic scheme for pressure and reconstruction of boundary equations.

Using the boundary conditions for the velocity components, we derive finite-difference equations for the pressure at the boundary nodes of the computational domain. Since no boundary conditions for the pressure are prescribed on the boundary, it is necessary to construct additional equations to close system (7).

To this end, the boundary values

$$\begin{aligned} &u^{n+1}_{i+\frac{1}{2},1},\ u^{n+1}_{i+\frac{1}{2},N_2}, \quad i=1,\dots,N_1-1,\\ &v^{n+1}_{1,j+\frac{1}{2}},\ v^{n+1}_{N_1,j+\frac{1}{2}}, \quad j=1,\dots,N_2-1 \end{aligned} \tag{11}$$

must be determined.

We demonstrate the approximation procedure for the lower boundary. Expanding the function $u^{n+1}_{i+\frac{1}{2},\frac{1}{2}}$ into a Taylor series in a neighborhood of the nodes $(i+\frac{1}{2},1)$, we obtain a Taylor series in a neighborhood of the nodes $(i+\frac{1}{2},1)$, we obtain

$$u^{n+1}_{i+\frac{1}{2},\frac{1}{2}} = u^{n+1}_{i+\frac{1}{2},1} - \frac{h_2}{2}\left(\frac{\partial u^{n+1}}{\partial y}\right)_{i+\frac{1}{2},1} + \frac{h_2^2}{8}\left(\frac{\partial^2 u^{n+1}}{\partial y^2}\right)_{i+\frac{1}{2},1} + O(h_2^3). \tag{12}$$

Since $u^{n+1}_{i+\frac{1}{2},\frac{1}{2}} = 0$, it follows that

$$u^{n+1}_{i+\frac{1}{2},1} = \frac{h_2}{2}\left(\frac{\partial u^{n+1}}{\partial y}\right)_{i+\frac{1}{2},1} - \frac{h_2^2}{8}\left(\frac{\partial^2 u^{n+1}}{\partial y^2}\right)_{i+\frac{1}{2},1} + O(h_2^3). \tag{13}$$

Projecting the first momentum equation onto the Ox_1 axis, we obtain

$$\left(\frac{\partial p^{n+1}}{\partial x}\right)_{i+\frac{1}{2},1} = \frac{1}{Re}\left(\frac{\partial^2 u^{n+1}}{\partial y^2}\right)_{x,i+\frac{1}{2};y,1}. \tag{17}$$

From relations (16) and (17) we obtain

$$u^{n+1}_{i+\frac{1}{2},1} = \frac{h_2}{2}\left(\frac{\partial u^{n}}{\partial y}\right)_{i+\frac{1}{2},1} - Re\frac{h_2^2}{8}\left(\frac{\partial p^{n+1}}{\partial x}\right)_{i+\frac{1}{2},1}. \tag{14}$$

We approximate the derivative $\left(\frac{\partial u^n}{\partial y}\right)_{i+\frac{1}{2},1}$ using the discrete values of the function u^n at the nodes $(i+\frac{1}{2},\frac{1}{2})$, $(i+\frac{1}{2},1)$, and $(i+\frac{1}{2},2)$. For this purpose, the following Taylor expansions are used:

$$\begin{aligned} 0 = u^{n+1}_{i+\frac{1}{2},\frac{1}{2}} &= u^{n+1}_{i+\frac{1}{2},1} - \frac{h_2}{2}\left(\frac{\partial u^n}{\partial y}\right)_{i+\frac{1}{2},1} + \frac{h_2^2}{8}\left(\frac{\partial^2 u^n}{\partial y^2}\right)_{i+\frac{1}{2},1} + O(h_2^3),\\ u^{n+1}_{i+\frac{1}{2},1} &= u^{n+1}_{i+\frac{1}{2},1},\\ u^{n+1}_{i+\frac{1}{2},2} &= u^{n+1}_{i+\frac{1}{2},1} + h_2\left(\frac{\partial u^n}{\partial y}\right)_{i+\frac{1}{2},1} + \frac{h_2^2}{2}\left(\frac{\partial^2 u^n}{\partial y^2}\right)_{i+\frac{1}{2},1} + O(h_2^3). \end{aligned} \tag{15}$$

Using the method of undetermined coefficients, we obtain a second-order approximation for $\left(\frac{\partial u^n}{\partial y}\right)_{i+\frac{1}{2},1}$ in terms of the function values at three points. The first equation is multiplied by -4, the second by 3, and the third by 1:

$$3h_2\left(\frac{\partial u^n}{\partial y}\right)_{i+\frac{1}{2},1} = -4u^n_{i+\frac{1}{2},\frac{1}{2}} + 3u^n_{i+\frac{1}{2},1} + u^n_{i+\frac{1}{2},2} + O(h_2^3). \tag{16}$$

Hence, we obtain the second-order finite-difference formula for the first derivative

$$\left(\frac{\partial u^n}{\partial y}\right)_{i+\frac{1}{2},1} \approx \frac{3u^n_{i+\frac{1}{2},1} + u^n_{i+\frac{1}{2},2}}{3h_2} + O(h_2^2). \tag{17}$$

Substituting relation (20) into (19), we obtain

$$u^{n+1}_{i+\frac{1}{2},1} = \frac{3u^n_{i+\frac{1}{2},1} + u^n_{i+\frac{1}{2},2}}{6} - Re\frac{h_2^2}{8}\frac{p^{n+1}_{i+1,1} - p^{n+1}_{i,1}}{h_1}. \tag{18}$$

Comparing relation (21) with the first formula in (7) for $j = 1$, we obtain

$$\bar{u}_{i+\frac{1}{2},1} = \frac{3u^n_{i+\frac{1}{2},1} + u^n_{i+\frac{1}{2},2}}{6}, \qquad \tau = \tau_0 = \frac{h_2^2}{8\nu}. \tag{19}$$

To determine the pressure at the interior nodes of the grid domain Ω_h, from relations (7)–(8) we obtain a finite-difference elliptic scheme

$$\left(p^{n+1}_{i,j}\right)_{x\bar{x}} + \left(p^{n+1}_{i,j}\right)_{y\bar{y}} = \Delta_h p^n_{i,j} + \frac{1}{\tau}\,\mathrm{div}_h\,\bar{\mathbf{v}}^{\,n+\frac{1}{2}}, \qquad i = 2,\ldots,N_1-2,\ j = 2,\ldots,N_2-2. \tag{20}$$

Due to the absence of boundary conditions for the pressure at the boundary nodes, it is necessary to construct additional equations based on relations of the form (22). For the lower boundary, we obtain the incompressibility equation at $j = 1$:

$$\frac{u^{n+1}_{i+\frac{1}{2},1} - u^{n+1}_{i-\frac{1}{2},1}}{h_1} + \frac{v^{n+1}_{i,\frac{3}{2}} - v^{n+1}_{i,\frac{1}{2}}}{h_2} = 0, \qquad i = 1,\ldots,N_1. \tag{21}$$

Substituting into (24) the expressions

$$u^{n+1}_{i+\frac{1}{2},1} = \bar{u}_{i+\frac{1}{2},1} - \tau\left(\frac{p^{n+1}_{i+1,1} - p^{n+1}_{i,1}}{h_1} - \frac{p^n_{i+1,1} - p^n_{i,1}}{h_1}\right),$$

$$v^{n+1}_{i,\frac{3}{2}} = \bar{v}_{i,\frac{3}{2}} - \tau\left(\frac{p^{n+1}_{i,2} - p^{n+1}_{i,1}}{h_2} - \frac{p^n_{i,2} - p^n_{i,1}}{h_2}\right), \qquad v^{n+1}_{i,\frac{1}{2}} = 0,$$

we obtain the equation for determining the pressure at the nodes $(i,1)$, $i=2,\ldots,N_1-1$:

$$\frac{\tau_0}{h_1^2}\left(p_{i+1,1}^{n+1}-2p_{i,1}^{n+1}+p_{i-1,1}^{n+1}\right)+\frac{\tau}{h_2^2}\left(p_{i,2}^{n+1}-p_{i,1}^{n+1}\right)=\frac{\tau_0}{h_1^2}\left(p_{i+1,1}^{n}-2p_{i,1}^{n}+p_{i-1,1}^{n}\right)$$
$$+\frac{\tau}{h_2^2}\left(p_{i,2}^{n}-p_{i,1}^{n}\right)+\frac{\bar{u}_{i+\frac{1}{2},1}-\bar{u}_{i-\frac{1}{2},1}}{h_1}+\frac{\bar{v}_{i,\frac{3}{2}}-\bar{v}_{i,\frac{1}{2}}}{h_2}. \quad (22)$$

In an analogous manner, the discrete boundary conditions for the velocity components on the upper, left, and right boundaries of the domain are formulated.

As an illustration of the application of the derived finite-difference scheme, Fig. 1 presents an example of a steady flow of a viscous incompressible fluid in a closed domain.

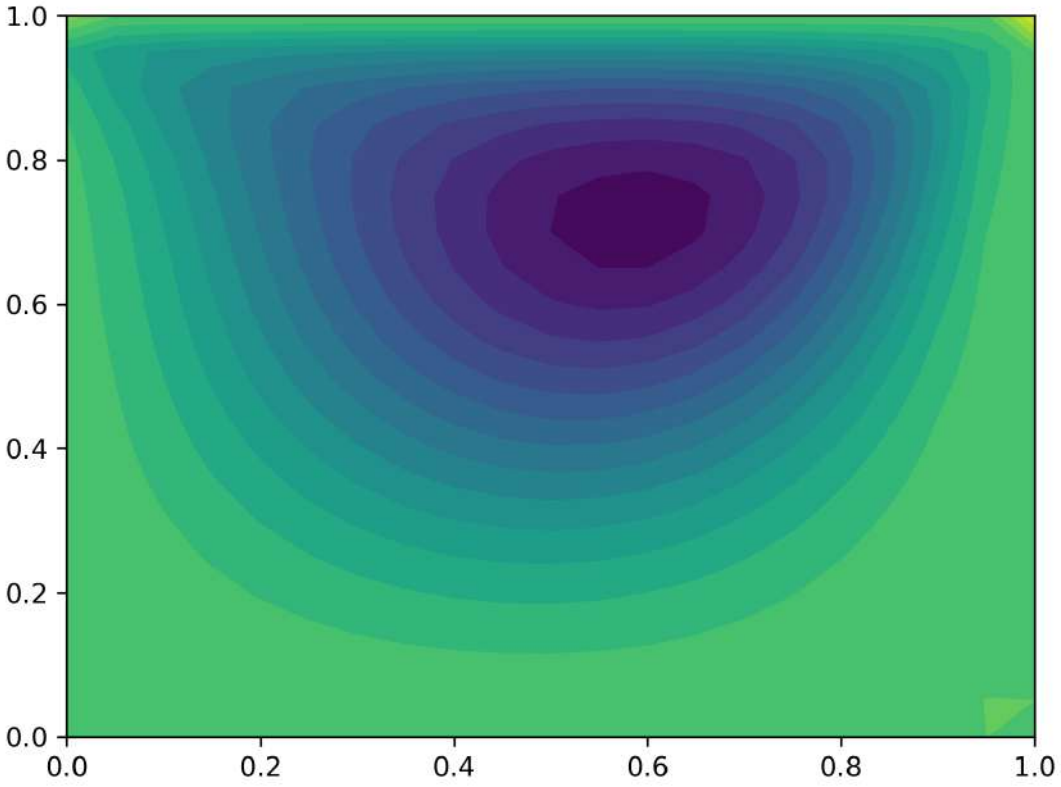


FIGURE 1. Stream function isolines

CONCLUSIONS

The obtained finite-difference elliptic scheme for the pressure and the derived boundary equations ensure the consistency of the pressure field with the velocity boundary conditions. The application of the projection method makes it possible to correctly satisfy the incompressibility condition at each time step. The presented example of a steady flow demonstrates the physical adequacy and the efficiency of the proposed numerical approach.

The developed numerical algorithm is based on a consistent splitting of the Navier–Stokes equations with respect to physical processes and is implemented on a regular staggered grid. The proposed approach provides a stable and accurate approximation of the velocity and pressure fields in bounded domains and allows one to correctly take into account the influence of boundary conditions. The obtained results indicate that the method can be effectively applied

to the numerical simulation of incompressible viscous flows in cavity-type domains and can be extended to more complex geometries, higher Reynolds numbers, and unsteady flow regimes. Various numerical approaches to the incompressible Navier–Stokes equations, including projection methods, pressure correction schemes, and cavity-flow configurations, have been extensively investigated in the literature [2, 3, 6, 5, 7, 4]

Keywords: Navier–Stokes equations, incompressible fluid, splitting method, finite difference scheme, numerical simulation.

AMS Subject Classification: 35Q30, 65M06, 76D05.

References

[1] Belotserkovskii O. M. *Numerical Modeling in Continuum Mechanics*. Moscow: Nauka, 1984 (in Russian).
[2] Chorin A. J. A numerical method for solving incompressible viscous flow problems. *Journal of Computational Physics*, 1967, Vol. 2, No. 1, pp. 12–26.
[3] Temam R. *Navier–Stokes Equations: Theory and Numerical Analysis*. Amsterdam: North-Holland, 1977.
[4] Lomasov D., Vabishchevich P. Iterative solutions of the incompressible fluid flow in a lid-driven cavity at large Reynolds numbers. In: *Current Problems of Applied Mathematics and Computer Systems (CPAMCS 2024)*, Ed. by I. Samoylenko et al., Lecture Notes in Networks and Systems, Vol. 1585. Cham: Springer, 2025. DOI: 10.1007/978-3-032-01831-1_47.
[5] Abouhamza A., Pierre R. A neutral stability curve for incompressible flows in a driven cavity. *Mathematical and Computer Modelling*, 2003, Vol. 38, pp. 141–157.
[6] Erturk E., Corke T. C., Gökçöl C. Numerical solutions of two-dimensional steady incompressible driven cavity flow at high Reynolds numbers. *International Journal for Numerical Methods in Fluids*, 2005, Vol. 48, pp. 747–774. DOI: 10.1002/fld.953.
[7] Erturk E., Allahviranloo T. Bifurcation and multiplicity of solutions of the Navier–Stokes equations in driven semi-elliptical cavity flow. *Mathematics*, 2022, Vol. 10, No. 22, Article 4242. DOI: 10.3390/math10224242.

TRAINING ALGORITHM OF A FULLY CONNECTED NEURAL NETWORK USING BACKPROPAGATION

N.M. TEMIRBEKOV[1,2], Z.I. SEITMURATOVA[1,2]

[1]Al-Farabi Kazakh National University,
[2]National Engineering Academy of the Republic of Kazakhstan

The application of neural networks in geology contributes to accelerating research processes, improving the accuracy and reliability of results, and reducing the costs of geological exploration. At present, several advanced neural network–based systems, such as GigaChat, YaGPT, GPT-4o, and Gemini 1.5 Flash, represent state-of-the-art technologies for modern geology. Across these neural network platforms, geographic information system (GIS) technologies play a central role, followed by 3D modeling of geological structures, artificial intelligence, remote sensing of the Earth, and the analysis of large-scale geological data (Big Data), including multiparametric geostatistics.

Artificial intelligence enables faster processing of geological data, more accurate identification of hidden patterns, and a reduction in decision-making risks. Therefore, this study focuses on a fully connected neural network trained using the backpropagation algorithm. A general overview of the development of machine learning in geological sciences is presented in [1].

A comprehensive review of machine learning applications has been conducted across multiple geoscience disciplines. Remote sensing was among the first fields to adopt machine learning methods, as demonstrated by Lary et al. in 2016 [2]. This was followed by applications in solid Earth geosciences in the work of Bergen et al. (2019) [3], hydrogeophysics in the study by Shen (2018) [4], seismology by Kong et al. (2019) [5], seismic interpretation presented by Wang et al. (2018) [6], and geochemistry discussed in the work of Zuo et al. (2019) [7].

Since the mid-twentieth century, when the first computing machines appeared, significant changes have occurred. Computational resources are now widely available in both hardware and software forms, and high-performance computing has become accessible to a broad audience through cloud computing providers. High-quality machine learning software is extensively available due to the free and open-source software movement, while major technology companies (Google, Facebook, Microsoft) actively compete by promoting their open machine learning frameworks, such as TensorFlow, PyTorch, and CNTK2. In addition, independent developments have gained widespread adoption, including scikit-learn [8] (Pedregosa et al., 2011) and XGBoost [9] (Chen and Guestrin, 2016).

The influence of scikit-learn has been particularly significant in shaping modern machine learning software ecosystems through the introduction of a unified application programming interface (API) [10] (Buitinck et al., 2013). The considered algorithm consists of three stages: forward propagation, backward propagation, and updating of the trainable parameters. As a result of the forward propagation stage, the error $E(\Omega^t)$ is computed. During the backward propagation stage, the error $E(\Omega^t)$ is used to evaluate the gradient vector $\frac{\partial E}{\partial \Omega}(\Omega^t)$. Subsequently, the trainable parameters are updated.

$$\Omega^{t+1} = \Omega^t - \alpha * \frac{\partial E}{\partial \Omega}(\Omega^t) \tag{1}$$

Assume that the dimension of the input vector, denoted as INPUT_DIM, is equal to 4, corresponding to four input features. The number of classes is assumed to be 3; therefore, the dimension of the output layer (OUTPUT_DIM) is equal to 3. A two-layer neural network is

employed in this study. The number of neurons in the first (hidden) layer, denoted as H_DIM, is set to 5. These parameters may be varied in subsequent computational experiments.

We now describe the backpropagation algorithm. For simplicity, let us consider the input vector $x = [x_1, x_2]$ and the output vector $h = [h_1, h_2, h_3]$. The vector h is the output of a nonlinear activation function that depends on the input vector x, the trainable weight parameters w, and the bias parameters.

$$h = F(xW + b), \tag{2}$$

where the parameters are defined as follows.

$$w = \begin{bmatrix} w_{11} & w_{12} & w_{13} \\ w_{21} & w_{22} & w_{23} \end{bmatrix}, \qquad b = [b_1, b_2, b_3]. \tag{3}$$

In expanded form, equation (2) can be written as follows

$$\begin{aligned} t_1 &= F(x_1 w_{11} + x_2 w_{21} + b_1), \\ t_2 &= F(x_1 w_{12} + x_2 w_{22} + b_2), \\ t_3 &= F(x_1 w_{13} + x_2 w_{23} + b_3). \end{aligned} \tag{4}$$

If we introduce the notation $t = xW + b$, then t depends linearly on x. Furthermore, $h = F(t)$ where $F(t)$ is a nonlinear activation function.

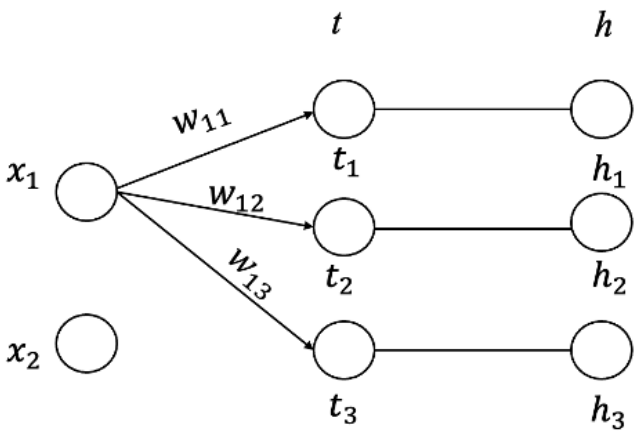


FIGURE 1. Neural network architecture

Let us write the expressions for t and h in expanded form

$$\begin{cases} t_1 = x_1 w_{11} + x_2 w_{21} + b_1, \\ t_2 = x_1 w_{12} + x_2 w_{22} + b_2, \\ t_3 = x_1 w_{13} + x_2 w_{23} + b_3. \end{cases} \tag{5}$$

For the output vector

$$\begin{cases} h_1 = F(t_1), \\ h_2 = F(t_2), \\ h_3 = F(t_3). \end{cases} \tag{6}$$

the computational graph can be represented as follows:

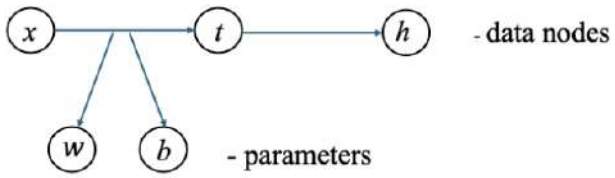


FIGURE 2. Computational graph

To estimate the neural network parameters during the optimization process using stochastic gradient descent, the gradient vector $\frac{\partial E}{\partial \Omega}(\Omega^t)$ is required.

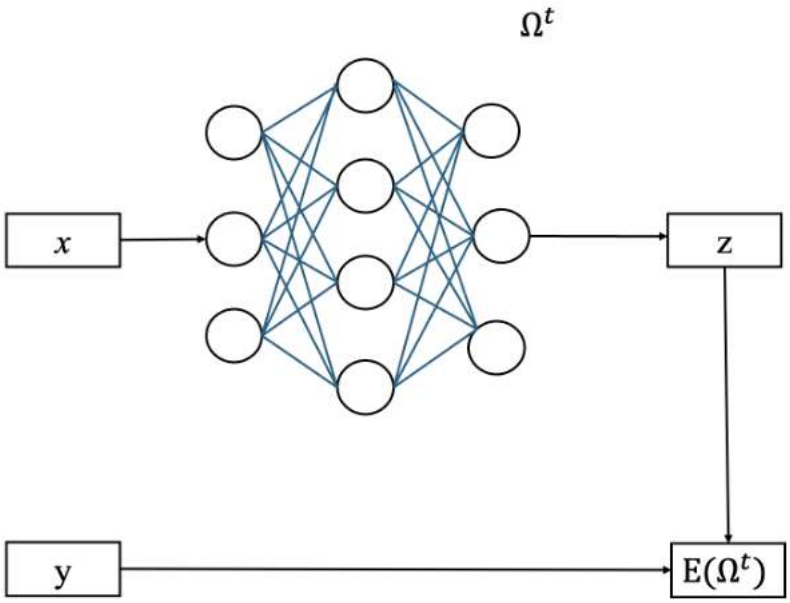


FIGURE 3. Forward propagation implementation scheme

Figure 3 illustrates the forward propagation scheme. Here, x denotes the input vector, z represents the vector of predicted probabilities at the output layer, and y is the known ground-truth label, represented as a vector consisting of zeros with a single unit entry.

To update the trainable parameters $\Omega = (w, b)$, it is necessary to perform backward propagation (BACKWARD).

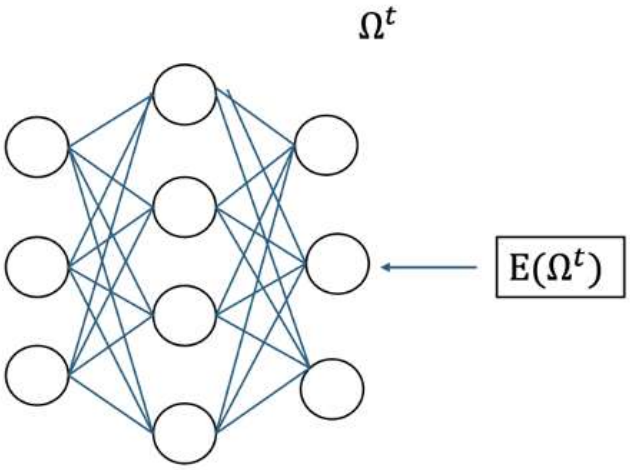


FIGURE 4. Backward propagation along the computational graph

We require $\frac{\partial E}{\partial w}$ and $\frac{\partial E}{\partial b}$.

We will use an important result from mathematical analysis, namely the chain rule, which is the differentiation rule for composite functions.

Let the gradient of the error with respect to the output vector be given by

$$\frac{\partial E}{\partial h} = \left[\frac{\partial E}{\partial h_1}, \frac{\partial E}{\partial h_2}, \frac{\partial E}{\partial h_3}\right],$$

and the components of this vector will be specified later. The rule for computing the error gradients is as follows:

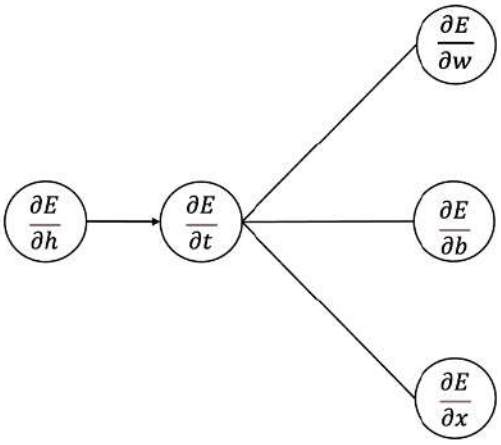


FIGURE 5. Rule for computing the error gradients

Figure 5 illustrates the scheme for the sequential computation of gradients. As a result, the gradient $\frac{\partial E}{\partial w}$ is obtained in the form of a (2×3) matrix, $\frac{\partial E}{\partial b}$ is a row vector of size (1×3), and the input gradient $\frac{\partial E}{\partial x}$ is a row vector of size (1×2). The input gradient $\frac{\partial E}{\partial x}$ is required to propagate the gradient to the previous layer.

Let it be given that

$$\frac{\partial E}{\partial h} = \left[\frac{\partial E}{\partial h_1}, \frac{\partial E}{\partial h_2}, \frac{\partial E}{\partial h_3}\right],$$

$$\frac{\partial E}{\partial t_1} = \frac{\partial E}{\partial h_1} \cdot \frac{\partial h_1}{\partial t_1} = \frac{\partial E}{\partial h_1} \cdot F'(t_1), \tag{7}$$

then analogously we obtain

$$\frac{\partial E}{\partial t_2} = \frac{\partial E}{\partial h_2} \cdot \frac{\partial h_2}{\partial t_2} = \frac{\partial E}{\partial h_2} \cdot F'(t_2), \tag{8}$$

$$\frac{\partial E}{\partial t_3} = \frac{\partial E}{\partial h_3} \cdot \frac{\partial h_3}{\partial t_3} = \frac{\partial E}{\partial h_3} \cdot F'(t_3). \tag{9}$$

Let us rewrite formulas (7)–(9) in a more compact form. For this purpose, we introduce the following notation:

$$\frac{\partial E}{\partial h} = \left[\frac{\partial E}{\partial h_1}, \frac{\partial E}{\partial h_2}, \frac{\partial E}{\partial h_3}\right], \qquad F'(t) = \left[F'(t_1),\, F'(t_2),\, F'(t_3)\right]. \tag{10}$$

Then, formulas (7)–(9) take the following form

$$\frac{\partial E}{\partial t} = \frac{\partial E}{\partial h} \otimes F'(t), \tag{11}$$

Here $\otimes$ denotes the element-wise (Hadamard) product. The Hadamard product is defined as a binary operation between two matrices of the same dimensions.

We now write the formulas for computing $\frac{\partial E}{\partial w}$, using the following notation:

$$\frac{\partial E}{\partial w_{11}} = \frac{\partial E}{\partial t_1} \cdot \frac{\partial t_1}{\partial w_{11}} = \frac{\partial E}{\partial t_1} \cdot x_1.$$

$$\frac{\partial E}{\partial w_{21}} = \frac{\partial E}{\partial t_1} \cdot \frac{\partial t_1}{\partial w_{21}} = \frac{\partial E}{\partial t_1} \cdot x_2, \tag{12}$$

$$\frac{\partial E}{\partial w_{12}} = \frac{\partial E}{\partial t_2} \cdot \frac{\partial t_2}{\partial w_{12}} = \frac{\partial E}{\partial t_2} \cdot x_1, \tag{13}$$

$$\frac{\partial E}{\partial w_{22}} = \frac{\partial E}{\partial t_2} \cdot \frac{\partial t_2}{\partial w_{22}} = \frac{\partial E}{\partial t_2} \cdot x_2, \tag{14}$$

$$\frac{\partial E}{\partial w_{13}} = \frac{\partial E}{\partial t_3} \cdot \frac{\partial t_3}{\partial w_{13}} = \frac{\partial E}{\partial t_3} \cdot x_1, \tag{15}$$

$$\frac{\partial E}{\partial w_{23}} = \frac{\partial E}{\partial t_3} \cdot \frac{\partial t_3}{\partial w_{23}} = \frac{\partial E}{\partial t_3} \cdot x_2. \tag{16}$$

Using the transposed input vector

$$x^T = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$$

and

$$\frac{\partial E}{\partial t} = \left[\frac{\partial E}{\partial t_1}, \frac{\partial E}{\partial t_2}, \frac{\partial E}{\partial t_3}\right],$$

we rewrite equation (12) in a compact matrix form.

$$\frac{\partial E}{\partial w} = x^T \cdot \frac{\partial E}{\partial t}. \tag{17}$$

The gradient $\frac{\partial E}{\partial w}$ is a matrix of size (2×3). Let us derive the formulas for $\frac{\partial E}{\partial b}$.

$$\frac{\partial E}{\partial b_1} = \frac{\partial E}{\partial t_1} \cdot \frac{\partial t_1}{\partial b_1} = \frac{\partial E}{\partial t_1} \cdot 1 = \frac{\partial E}{\partial t_1}. \tag{18}$$

Analogously, we obtain

$$\frac{\partial E}{\partial b_2} = \frac{\partial E}{\partial t_2}, \qquad \frac{\partial E}{\partial b_3} = \frac{\partial E}{\partial t_3}. \tag{19}$$

in a compact form

$$\frac{\partial E}{\partial b} = \frac{\partial E}{\partial t}. \tag{20}$$

The formulas for the input gradient $\frac{\partial E}{\partial x}$ are determined from the following chain rule for differentiating composite functions.

$$\frac{\partial E}{\partial x_1} = \frac{\partial E}{\partial t_1} \cdot \frac{\partial t_1}{\partial x_1} + \frac{\partial E}{\partial t_2} \cdot \frac{\partial t_2}{\partial x_1} + \frac{\partial E}{\partial t_3} \cdot \frac{\partial t_3}{\partial x_1}, \tag{21}$$

From equation (5), we obtain

$$\frac{\partial t_1}{\partial x_1} = w_{11}, \qquad \frac{\partial t_2}{\partial x_1} = w_{12}, \qquad \frac{\partial t_3}{\partial x_1} = w_{13},$$

therefore,

$$\frac{\partial E}{\partial x_1} = \frac{\partial E}{\partial t_1} \cdot w_{11} + \frac{\partial E}{\partial t_2} \cdot w_{12} + \frac{\partial E}{\partial t_3} \cdot w_{13}. \tag{22}$$

By analogous reasoning, we have

$$\frac{\partial E}{\partial x_2} = \frac{\partial E}{\partial t_1} \cdot w_{21} + \frac{\partial E}{\partial t_2} \cdot w_{22} + \frac{\partial E}{\partial t_3} \cdot w_{23}, \tag{23}$$

or, in matrix form

$$\frac{\partial E}{\partial x} = \frac{\partial E}{\partial t} \cdot W^T, \tag{24}$$

Here $\frac{\partial E}{\partial x}$ is a row vector of size (1×2), $\frac{\partial E}{\partial t}$ is a row vector of size (1×3) and W^T is a matrix of size (2×3).

Finally, it remains to answer the question: how and where do we obtain $\frac{\partial E}{\partial h}$?

Let us first describe in more detail the computation of the error. To determine the error, it is necessary to first compute the vector z using the Softmax activation function at the output layer. This activation function is commonly used in classification problems because it normalizes the activations across all output units so that their sum equals one. This is achieved by applying the exponential function to each output t_i and dividing this value by the sum of the exponentials over all outputs:

$$z = \text{softmax}(t) = S(t) = \left\{ \frac{e^{t_i}}{\sum_j e^{t_j}} \right\} \tag{25}$$

Classification problems represent a special class of tasks in which a different type of loss function, known as cross-entropy, is employed. The cross-entropy loss depends on the true label y and the prediction produced at the output layer.

$$E = CrossEntropy(z, y) = CE(z, y) = -\sum_i y_i \ln z_i. \tag{26}$$

Here y is a vector consisting of zeros and a single unit entry (one-hot encoding). Let us transform equation (22) using equation (21).

$$E = -\sum_i y_i \ln \frac{e^{t_i}}{\sum_j e^{t_j}} = -\sum_i y_i \left(t_i - \ln \sum_j e^{t_j} \right) = -\sum_i y_i t_i + \left(\sum_i y_i \right) \ln \sum_j e^{t_j}$$

Since $\sum_i y_i = 1$, we obtain the prediction error formula

$$E = -\sum_i y_i t_i + \ln \sum_j e^{t_j}. \tag{27}$$

From (23), the derivative of E with respect to t_k can be derived

$$\frac{\partial E}{\partial t_k} = -y_k + \frac{e^{t_k}}{\sum_j e^{t_j}} = S(t_k) - y_k, \tag{28}$$

In compact form

$$\frac{\partial E}{\partial t} = S(t) - y = z - y. \tag{29}$$

We now choose the activation function. Let us consider

$$F(t) = \mathrm{ReLU}(t) = \max(0, t). \tag{30}$$

The graph of this function is shown in Figure 6.

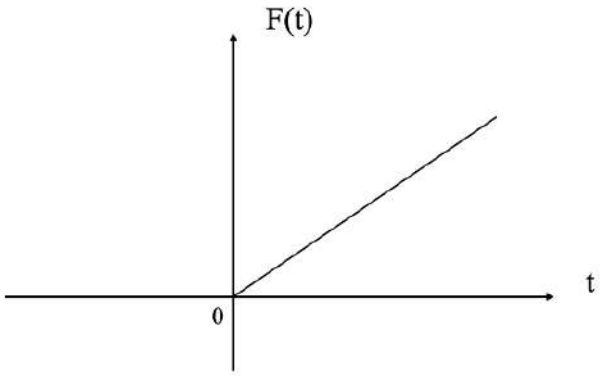


FIGURE 6. Activation function

The derivative of this function is given by

$$F'(t) = \begin{cases} 1, & t \geq 0, \\ 0, & t < 0. \end{cases} \tag{31}$$

Let us consider an example of a specific neural network consisting of two fully connected layers. The computational graph of such a neural network is as follows:

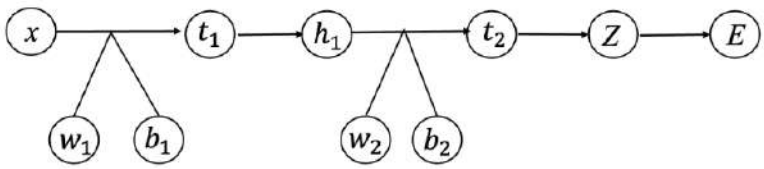


FIGURE 7. Computational graph of the neural network

The neuron computations are performed as follows:

$$t_i = xw_1 + b_1, \quad h_1 = F(t_1), \quad t_2 = h_1 w_2 + b_2, \quad z = S(t_2), \quad E = CE(z, y).$$

As a result, the prediction error E is obtained.

Next, the error backpropagation procedure is performed. The computational graph has the following form:

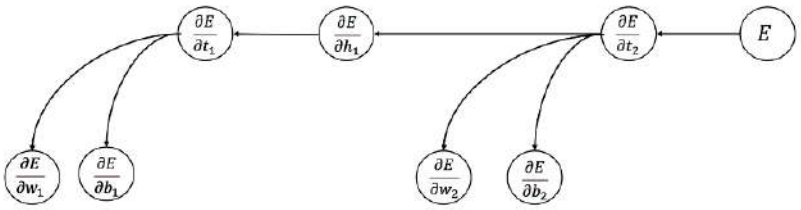


FIGURE 8. Computational graph of error backpropagation

These gradients are computed using the following formulas:

$$\frac{\partial E}{\partial t_2} = S(t_2) - y = z - y,$$

$$\frac{\partial E}{\partial w_2} = h_1^T \cdot \frac{\partial E}{\partial t_2},$$

$$\frac{\partial E}{\partial b_2} = \frac{\partial E}{\partial t_2}.$$

$$\frac{\partial E}{\partial h_1} = \frac{\partial E}{\partial t_2} \cdot W_2^T, \qquad \frac{\partial E}{\partial t_1} = \frac{\partial E}{\partial h_1} \otimes F'(t_1), \tag{32}$$

$$\frac{\partial E}{\partial w_1} = x^T \cdot \frac{\partial E}{\partial t_1}, \qquad \frac{\partial E}{\partial b_1} = \frac{\partial E}{\partial t_1}.$$

Thus, the required gradients $\frac{\partial E}{\partial w_2}$, $\frac{\partial E}{\partial b_2}$, $\frac{\partial E}{\partial w_1}$, $\frac{\partial E}{\partial b_1}$ are obtained, and the trainable parameters are updated according to formula (1).

In this work, a transition is made from mathematical foundations and theoretical concepts to the development of software for model validation. The presentation of results moves from textual descriptions to code examples that complement the theoretical framework and illustrate model verification as well as the explanation of machine learning methods.

REFERENCES

[1] Dramsch J.S. 70 years of machine learning in geoscience in review. *Machine Learning in Geoscience*, pages 1–55. Elsevier, 2020. https://doi.org/10.1016/bs.agph.2020.08.002.

[2] Lary D.J., Alavi A.H., Gandomi A.H., and Walker A.L. Machine learning in geosciences and remote sensing. *Geoscience Frontiers*, 7(1):3–10, 2016. doi:10.1016/j.gsf.2015.07.003.

[3] Bergen K.J., Johnson P.A., Maarten V., and Beroza G.C. Machine learning for data-driven discovery in solid earth geoscience. *Science*, 363(6433):eaau0323, 2019.

[4] Shen C. A transdisciplinary review of deep learning research and its relevance for water resources scientists. *Water Resources Research*, 54(11):8558–8593, 2018.

[5] Kong Q., Trugman D.T., Ross Z.E., Bianco M.J., Meade B.J., and Gerstoft P. Machine learning in seismology: Turning data into insights. *Seismological Research Letters*, 90(1):3–14, 2019.

[6] Wang Z., Di H., Shafiq M.A., Alaudah Y., and AlRegib G. Successful leveraging of image processing and machine learning in seismic structural interpretation: A review. *The Leading Edge*, 37(6):451–461, 2018.

[7] Zuo R., Xiong Y., Wang J., and Carranza E.J.M. Deep learning and its application in geochemical mapping. *Earth-Science Reviews*, 2019.

[8] Pedregosa F., Varoquaux G., Gramfort A., Michel V., Grisel O., Blondel M., Prettenhofer P., Weiss R., Dubourg V., VanderPlas J., Passos A., Cournapeau D., Brucher M., Perrot M., and Duchesnay E. Scikit-learn: Machine learning in Python. *Journal of Machine Learning Research*, 12:2825–2830, 2011.

[9] Chen T. and Guestrin C. XGBoost: A scalable tree boosting system. In *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, pages 785–794, New York, NY, USA, 2016. ACM. doi:10.1145/2939672.2939785.

[10] Buitinck L., Louppe G., Blondel M., Pedregosa F., Mueller A., Grisel O., Niculae V., Prettenhofer P., Gramfort A., Grobler J., Layton R., VanderPlas J., Joly A., Holt B., and Varoquaux G. API design for machine learning software: experiences from the scikit-learn project. In *ECML PKDD Workshop: Languages for Data Mining and Machine Learning*, pages 108–122, 2013.

NUMERICAL MODELING OF SOME PROBLEM OF FINANCIAL MATHEMATICS

ALTYN BAITELIEVA[1], ILYAS SHAKENOV[2], KANAT SHAKENOV[3]

[1]Al-Farabi Kazakh National University, Almaty, Kazakhstan
[2]Sky Students Company, Almaty, Kazakhstan
[3]Al-Farabi Kazakh National University, Almaty, Kazakhstan
e-mail: shakenov2000@mail.ru

This article examines the calculation of the option price $V(t,\, x)$, the stock price $x(t)$, and the optimal stopping (execution) time τ; $(\equiv t)$ over both finite and infinite time horizons. It then delves into determining a fair value for American-style options, leveraging the optimal stopping time within the framework of diffusion processes in stock markets, represented by $(B,\, S)$. Additionally, the article explores the pricing of European-style options, starting with the buyer's perspective and then transitioning to the seller's viewpoint. The problems are solved either analytically, when the optimal stopping time is pre-determined, or numerically using methods like the sweep method and finite element techniques.These methods are applied by reducing the problem to Stefan's problem, where $Y^*(t,\, x)$ represents the rational option value, τ_T^* indicates the rational execution time, and $x^*(t)$ corresponds to the rational stock price.

1. Introduction

Building on the work in [1], this paper investigates various aspects of calculating the option price $V(t, x)$, the stock price $x(t)$, and the optimal stopping (execution) time τ $(\equiv t)$ over finite and infinite time intervals. In then explores the determination of a fair price for American-style options, utilizing the optimal stopping time within diffusion-based $(B,\, S)$-stock market models. The discussion proceeds to address the pricing of European-style options, starting with an analysis on the buyer's perspective, particularly the call option, followed by a focus on the put option. The problems are solved either exactly when the optimal stopping time is predetermined or numerically by reformulating them into the Stefan problem. In mathematical physics, the Stefan problem arises in the study of physical processes associated with the phase transformation of matter and consists in finding a function $u = u(t, x)$ that describes the temperature regime of the phases and the separation boundary $x = x(t),\ t \geq 0$ of these phases.

In the case of standard buyer and seller options, a two-phase situation also takes place – when searching for optimal stopping rules, we can restrict ourselves to considering only two simply connected phases: the area of continuation of observations C^T and the area D^T.

All problems can be solved analytically if the optimal stopping time is known beforehand or numerically if it is not.

The results of numerical modeling of the Stefan's problem by the sweep method and the finite element method for standard call and ask options are presented. As well as a comparative analysis of the numerical results by the sweep method and the finite element method (FEA).

2. Theory of calculations in stochastic financial models. American type options on diffusion $(B,\, S)$ stock markets. Optimal stopping problem

Consider an optimal stopping problem in a discrete and finite setting for a stochastic sequence $f = \big(f_n\big)_{n \leq N}$. This problem is closely related to American-type options, where the option

holder has the right to choose the exercise time, which can be viewed here as an "optimization decision". [2], [3], [4], [5], [6], [7].

Let $f = \Big(f_n,\, F_n\Big)_{0\le n\le N}$ be a stochastic sequence defined on the probability space $\Big(\Omega,\, F,\, (F_n)_{0\le n\le N},\, \mathbf{P}\Big)$, with $F_0 = \Big\{\emptyset,\, \Omega\Big\}$, $F_N = F$. We assume that $\mathbf{E}|f_n| < \infty$ for all $0 \le N < \infty$. TThe task at hand is twofold: 1) To find functions (prices)

$$V_n^N = \sup_{\tau\in M_n^N} \mathbf{E} f_\tau, \tag{1}$$

where the supremum is taken over the class M_n^N of all stopping times τ such that $0 \le \tau \le N$, 2) To find the optimal stopping time, which we assume exists in this setting. A distinctive feature of this finite $N < \infty$ case is that the problem can be solved using backward induction, which is a key method for determining both the prices V_n^N, and the corresponding optimal stopping times. We define the sequence $\gamma^N = \Big(\gamma_n^N\Big)_{0\le n\le N}$ in the following artificial way:

$$\gamma_n^N = f_N,\;\; \gamma_n^N = \max\Big(f_n,\, \mathbf{E}\Big(\gamma_{n+1}^N | F_n\Big)\Big). \tag{2}$$

We then define the optimal stopping time sequence $0 \le \tau \le N$ as:

$$\tau_n^N = \max\Big\{n \le i \le N :\; f_i = \gamma_i^N\Big\}. \tag{3}$$

The following result is one of the central theorems in the theory of optimal stopping problems over a finite time interval $0 \le \tau \le N$.

Theorem 2.1. *The sequence* $\gamma^N = \Big(\gamma_n^N\Big)_{0\le n\le N}$ *defined by recurrence relations (2) and the moments* τ_n^N, $\gamma^N = \Big(\gamma_n^N\Big)_{0\le n\le N}$, *have the following properties:*

(a) $\tau_n^N \in \mathrm{M}_n^N$;

(b) $\mathbf{E}\Big(f_{\tau_n^N} | F_n\Big) = \gamma_n^N$;

(c) $\mathbf{E}\Big(f_\tau | F_n\Big) \le \mathbf{E}\Big(f_{\tau_n^N} | F_n\Big) = \gamma_n^N$ *for anyone* $\tau \in \mathrm{M}_n^N$;

(d) $\gamma_n^N = \operatorname{ess\,sup}_{\tau\in\mathrm{M}_n^N} \mathbf{E}\Big(f_\tau | F_n\Big)$ *and in particular*

$$\gamma_0^N = \sup_{\tau\in\mathrm{M}_n^N} \mathbf{E} f_\tau = \mathbf{E} f_{\tau_0^N};$$

(e) $V_n^N = \mathbf{E}\gamma_n^N$.

Theorem 2.1 without proof.

3. Standard call option

In the context of options and other financial instruments, two scenarios are typically considered: the first, where the time parameter t is confined to a finite interval $[0,\, T]$, and the second, where it extends over an infinite interval $[0,\, \infty)$. The second case is often viewed as an idealized scenario, though it is mathematically simpler than the first, where the timing of decisions at any given moment t is significantly influenced by the remaining time $T - t$ until the contract's expiration. We begin by considering the second case.

Let us assume that on the filtered probability space $\Big(\Omega,\, F,\, (F_t)_{t\ge0},\, \mathbf{P}\Big)$, where $(F_t)_{t\ge0}$ is the Brownian (Wiener) filtration, i.e. the flow of σ-algebras $F_t = \sigma\Big(F_t^0 \cup \mathrm{N}\Big)$, $F_t^0 = \sigma\Big(B_s,\, s \le t\Big)$,

$\mathrm{N} = \Big\{A \in F : \mathbf{P}\big(A\big) = 0\Big\}$, is given a standard Wiener process $W_t = \Big(W_t\Big)_{t\geq 0}$ and the diffusion $\big(B,\, S\big)$-market has the following structure:

$$dB_t = rB_t dt, \;\; B_0 > 0, \tag{4}$$

$$dS_t = S_t\big(\mu dt + \sigma dW_t\big), \;\; S_0 > 0, \tag{5}$$

in this case, r represents the interest rate, while μ and σ are the parameters of the geometric Brownian motion that governs the stock price process $S_t = \Big(S_t\Big)_{t\geq 0}$. S_t The stock price follows the stochastic differential equation (5), and this equation, with the initial condition S_0 independent of the standard Wiener process $W_t = \Big(W_t\Big)_{t\geq 0}$, has the explicit solution $S_t = S_0 e^{\mu t} e^{\sigma W_t - \sigma^2 t/2}$. In this expression, μ represents the drift parameter, indicating the average rate of change in the stock price, while σ^2 denotes the diffusion term, often referred to as the differential dispersion or volatility. These parameters can have different interpretations depending on the specific context of the problem, but in this case, they are assumed to be deterministic. [11], [12], [13], [14], [15], [16].

For a standard discounted call option, the payoff function f_t is defined as follows:

$$f_t = e^{-\lambda t} g\big(S_t\big), \tag{6}$$

where $g(x) = \big(x - K\big)^+$, $x \in E = \big(0,\, \infty\big)$, K is **strike price**.

Suppose that

$$V^*(x) = \sup B_0 \tilde{\mathbf{E}}_x \frac{f\tau}{B_\tau} \tag{7}$$

where the supremum is taken over the class of all finite stopping times

$$\mathrm{M}_0^\infty = \Big\{\tau = \tau(\omega) :\, 0 \leq \tau(\omega) < \infty,\; \omega \in \Omega\Big\} \tag{8}$$

and $\tilde{\mathbf{E}}_x$ denotes the expectation with respect to the martingale measure $\tilde{\mathbf{P}}_x$ under which the process $S_t = \Big(S_t\Big)_{t\geq 0}$ follows a stochastic differential equation.

$$dS_t = S_t\big(rdt + \sigma dW_t\big), \;\; S_0 = x. \tag{9}$$

For convenience, we will assume from the outset that $\mu = r$. With this assumption, the tilde symbols in $\tilde{\mathbf{P}}_x$ and $\tilde{\mathbf{E}}_x$ can be dropped, and we will use "$\sim$" instead.

Thus, let:

$$V^*(x) = \sup_{\tau \in \mathrm{M}_0^\infty} \mathbf{E}_x e^{-(\lambda + r)\tau} \Big(S_\tau - K\Big)^+. \tag{10}$$

For various applications, it is useful to consider, in addition to the class M_0^∞, the class of

$$\bar{\mathrm{M}}_0^\infty = \Big\{\tau = \tau(\omega) :\, 0 \leq \tau(\omega) \leq \infty,\; \omega \in \Omega\Big\}$$

those Markov times that can take the values $+\infty$, and assume

$$\bar{V}^*(x) = \sup_{\tau \in \bar{\mathrm{M}}_0^\infty} \mathbf{E}_x e^{-(\lambda + r)\tau} \Big(S_\tau - K\Big)^+ I\big(\tau < \infty\big). \tag{11}$$

The determination of the functions $V^*(x)$ and $\bar{V}^*(x)$ is closely related to the standard American-style call option being analyzed, as these values directly correspond to the rational values. This assumes that the option buyer has the ability to choose the time of exercise either from the class M_0^∞, alternatively, in the class $\bar{\mathrm{M}}_0^\infty$ with $S_0 = x$, the case where $\tau = \infty$ corresponds to the situation where the option is never exercised. If τ^* and $\bar{\tau}^*$ are the optimal stopping times for solving problems (10) and (11) these will indicate the optimal moments for the buyer to exercise the option within the classes M_0^∞ or $\bar{\mathrm{M}}_0^\infty$.

Now, let us analyze the optimal stopping problems (10) and (11) in the case where $\lambda > 0$. The theorem holds in this context.

Theorem 3.1. *If $\lambda > 0$, then for any $x \in \big(0,\, \infty\big)$*

$$V^*(x) = \bar{V}^*(x) = \left\{ \begin{array}{ll} x - K & \textit{if} \quad x \geq x^*, \\ c^* x^{z_1} & \textit{if} \quad x < x^*, \end{array} \right. \tag{12}$$

where

$$z_1 = \frac{1}{2} \cdot \frac{-2r + \sigma^2 + \sqrt{4r^2 + 4r\sigma^2 + \sigma^4 + 8\sigma^2\lambda}}{\sigma^2}, \tag{13}$$

denote by $\kappa = -2r + \sigma^2 + \sqrt{4r^2 + 4r\sigma^2 + \sigma^4 + 8\sigma^2\lambda}$. Then

$$c^* = \frac{-\big(-x + K\big)}{\exp\Big(-\frac{1}{2}\frac{1}{\sigma^2}\Big(\ln\Big(\frac{2\sigma^2}{\kappa\big(-x+K\big)}\Big)\kappa\Big)\Big)}, \tag{14}$$

$$x^* = -\frac{1}{2}\frac{1}{\sigma^2}\Big(\kappa\big(-x + K\big)\Big). \tag{15}$$

In the class $\bar{\mathrm{M}}_0^{\infty}$ there is an optimal moment, and as such we can take the moment

$$\tau^* = \inf\Big\{t \geq 0 : \quad S_t \geq x^*\Big\}. \tag{16}$$

Where in

$$\mathbf{P}_x\big(\tau^* < \infty\big) = \left\{ \begin{array}{ll} 1 & \textit{if} \quad x \geq \frac{\sigma^2}{2} \quad \textit{or} \quad x \geq x^*, \\ \Big(\frac{x}{x^*}\Big)^{1-\frac{2r}{\sigma^2}} & \textit{if} \quad x < \frac{\sigma^2}{2} \quad \textit{and} \quad x < x^*. \end{array} \right. \tag{17}$$

General Scheme of the Proof of Theorem 3.1. Consider more general optimal stopping problems than problems (10) and (11). Let

$$V^*(x) = \sup_{\tau \in \mathrm{M}_0^{\infty}} \mathbf{E}_x e^{-\beta\tau} g\big(S_\tau\big), \tag{18}$$

$$V^*(x) = \sup_{\tau \in \bar{\mathrm{M}}_0^{\infty}} \mathbf{E}_x e^{-\beta\tau} g\big(S_\tau\big) I\big(\tau < \infty\big) \tag{19}$$

be the prices in optimal stopping problems for the Markov process $S = \Big(S_t,\ F_t,\ \mathbf{P}_x\Big)_{t\geq 0}$, $x \in E = \Big(0,\ \infty\Big)$ where $\mathbf{P}_x$ is the probability distribution of the process S with $S_0 = x$, $\beta > 0$ and $g = g(x)$ is some Borel function.

If $g = g(x)$ is a non-negative function, then, according to the general theory of optimal stopping rules for Markov processes:

$$\textbf{1)} \qquad V^*(x) = \bar{V}^*(x),\ x \in E; \tag{20}$$

2) $V^*(x)$ is the least–excessive majorant of the function $g(x)$, i.e. smallest among functions $V(x)$, such that

$$V(x) \geq g(x),\ V(x) \geq e^{-\beta\tau} T_t V(x), \tag{21}$$

where $T_t V(x) = \mathbf{E}_x V(x)$;

$$\textbf{3)} \qquad V^*(x) = \lim_n \lim_N Q_n^N g(x), \tag{22}$$

where

$$Q_n g(x) = \max\Big(g(x),\ e^{-\beta 2^n} T_{2^{-n}} g(x)\Big); \tag{23}$$

4) if $\mathbf{E}_x\Big(\sup\limits_t e^{-\beta t}g\big(S_t\big)\Big) < \infty$, then for each $\varepsilon > 0$ moment

$$\tau_\varepsilon = \inf\Big\{t:\ V^*\big(S_t\big) \le e^{-\beta t}g\big(S_t\big)\Big\} \tag{24}$$

is the ε-optimal stopping moment in the class M_0^∞, i.e. $\mathbf{P}_x\big(\tau_x < \infty\big) = 1,\ x \in E$ and $V^*(x) - \varepsilon \le \mathbf{E}_x e^{-\beta\tau_\varepsilon}g\big(S_{\tau_\varepsilon}\big)$;

5) if the moment

$$\tau_0 = \inf\Big\{t:\ V^*\big(S_t\big) \le e^{-\beta t}g\big(S_t\big)\Big\} \tag{25}$$

is the stopping point $\Big(\mathbf{P}_x\big(\tau_0 < \infty\big) = 1,\ x \in E\Big)$, then it is optimal in the class M_0^∞:

$$V^*(x) = \mathbf{E}_x e^{-\beta\tau_0}g\big(S_{\tau_0}\big),\ \ x \in E; \tag{26}$$

moreover, if some other stopping time τ_1 is also optimal, then $\mathbf{P}_x\big(\tau_0 \le \tau_1\big) = 1,\ x \in E$, i.e. τ_0 is the smallest optimal stopping time.

Let $C^* = \Big\{x \in E:\ V^* > g(x)\Big\}$ and $D^* = \Big\{x \in E:\ V^* = g(x)\Big\}$. It follows from (22) and (23) that the function $V^* = V^*(x)$ has a simple structure, being a downward convex function that $E = \big(0,\ \infty\big)$ majorizes the function $g = g(x)$. Moreover, there exists x^* such that $C^* = \Big\{x:\ x < x^*\Big\}$ and $D^* = \Big\{x:\ x \ge x^*\Big\}$.

Thus, the solution of problem (10), (11) is reduced to finding the value x^* and function $V^*(x)$ $\Big(V^*(x) = \bar{V}^*(x)\Big)$. It follows from the works:that the required value x^* and $V^*(x)$ is the least $\beta = \big(\lambda + r\big)$-excessive majorant of the function $g(x)$, must be solutions of the following Stefan problem or the problem with a free boundary.

$$\mathrm{L}\tilde{V}(x) = \big(\lambda + r\big)\tilde{V}(x),\ x < \tilde{x}, \qquad \tilde{V}(x) = g(x),\ x \ge \tilde{x}, \qquad \frac{d\tilde{V}(x)}{dx}\bigg|_{x\uparrow\tilde{x}} = \frac{dg(x)}{dx}\bigg|_{x\downarrow\tilde{x}}$$

or

$$\frac{1}{2}\sigma^2x^2\frac{d^2\tilde{V}(x)}{dx^2} + rx\frac{d\tilde{V}(x)}{dx} - \big(\lambda + r\big)\tilde{V}(x) = 0,\ x < \tilde{x}, \tag{27}$$

$$\tilde{V}(x) = g(x),\ x \ge \tilde{x}, \tag{28}$$

$$\frac{d\tilde{V}(x)}{dx}\bigg|_{x\uparrow\tilde{x}} = \frac{dg(x)}{dx}\bigg|_{x\downarrow\tilde{x}}, \tag{29}$$

where

$$\mathrm{L} = rx\frac{d}{dx} + \frac{\sigma^2}{2}x^2\frac{d^2}{dx^2} \tag{30}$$

is an infinitesimal process operator $S = \big(S_t\big)_{t\ge0}$ with a stochastic differential

$$dS_t = S_t\big(rdt + \sigma dW_t\big).$$

The solution of equation (27) in the region unknown $\big(0,\ \tilde{x}\big)$ so far has the form:

$$\tilde{V}(x) = c_1x^{\frac{1}{2}\cdot\frac{-2r+\sigma^2+k}{\sigma^2}} + c_2x^{-\frac{1}{2}\cdot\frac{2r-\sigma^2+k}{\sigma^2}},$$

where $k = \sqrt{4r^2 + 4r\sigma^2 + \sigma^4 + 8\sigma^2\lambda}$.

The characteristic equation has the form:

$$z^2 - \big(1 - \frac{2r}{\sigma^2}\big)z - \frac{2\big(\lambda + r\big)}{\sigma^2},$$

and the roots

$$z_1 = \frac{1}{2} \cdot \frac{-2r + \sigma^2 + k}{\sigma^2}, \qquad z_2 = \frac{1}{2} \cdot \frac{2r - \sigma^2 + k}{\sigma^2}.$$

Because $\lambda > 0$, then $z_1 > 1$. If $\lambda = 0$, then $z_1 = 1$. Root $z_2 < 0$. Then from the general solution of equation (27) $\tilde{V}(x) = c_1 x^{z_1} + c_2 x^{z_2}$, given that $z_2 < 0$ we get $c_2 = 0$. This follows from the fact that, according to the meaning of the problem under consideration, the function $\tilde{V}(x)$ must be a non-negative non-decreasing function. Also, if $c_2 \neq 0$, then $V(x) \to \pm\infty$ for $x \downarrow 0$, which should be excluded according to the meaning of the problem under consideration: $V^* \geq 0$ and $V^* \leq x$.

Thus, $\tilde{V}(x) = c_1 x^{\frac{1}{2} \cdot \frac{-2r+\sigma^2+k}{\sigma^2}}$ for $x < \tilde{x}$, where c_1 and the "free" boundary $\tilde{x}$ are still unknown constants, to determine which we use condition (28) and the "smooth gluing" condition (29). Condition (28) gives the relation

$$c_1 \tilde{x}^{z_1} = x - K. \tag{31}$$

Condition (29) takes the form

$$c_1 z_1 \tilde{x}^{z_1 - 1} = 1. \tag{32}$$

Or we get a system of two equations for the unknowns c_1 and $\tilde{x}$:

$$\begin{cases} c_1 \tilde{x}^{\frac{1}{2} \cdot \frac{\kappa}{\sigma^2}} = x - K, \\ c_1 \dfrac{1}{2} \cdot \dfrac{\kappa}{\sigma^2} \tilde{x}^{\frac{1}{2} \cdot \frac{\kappa}{\sigma^2}} = 1. \end{cases} \tag{33}$$

Solution of system (33):

$$c_1 = \frac{-\big(x + K\big)}{\exp\left(-\dfrac{1}{2}\dfrac{1}{\sigma^2}\left(\ln\left(-\dfrac{2\sigma^2}{\kappa\big(-x + K\big)}\right)\kappa\right)\right)},$$

$$\tilde{x} = -\frac{1}{2}\frac{1}{\sigma^2}\Big(\kappa\big(-x + K\big)\Big). \tag{34}$$

Thus, the solution $\tilde{V}(x)$ of problem (27) – (29) can be represented as follows:

$$\tilde{V}(x) = \begin{cases} x - K, & x \geq \tilde{x}, \\ c_1 x^{z_1}, & x < \tilde{x}, \end{cases} \tag{35}$$

where $\tilde{x}$ and c_1 are determined from (34) and

$$z_1 = \frac{1}{2} \cdot \left(\frac{\kappa}{\sigma^2}\right) = \frac{1}{2} \cdot \left(\frac{-2r + \sigma^2 + \sqrt{4r^2 + 4r\sigma^2 + \sigma^4 + 8\sigma^2\lambda}}{\sigma^2}\right).$$

Theorem 3.1 is proved.

It can be proved that the found function $\tilde{V}(x)$ coincides with the price $V^*(x)$, and the moment $\tilde{\tau} = \inf\Big\{t \geq 0 : \ S_t \geq \tilde{x}\Big\}$ is optimal in the class $\bar{\mathrm{M}}_0^\infty$ and in the class M_0^∞ if $\mathbf{P}_x\big(\tilde{\tau} < \infty\big)$.

4. Standard put option

The analysis of put options with payoff functions $f_t = e^{-\lambda t} g\big(S_t\big)$, where $g(x) = \big(K - x\big)^+$, $x \in E = \big(0, \infty\big)$, is conducted in the same manner as for call options. In this case, the function is bounded. We assume that the diffusion B, S-market is described as (4) and (5) from the **Standard Buy Option** and

$$U_*(x) = \sup_{\tau \in \mathrm{M}_0^\infty} \mathbf{E}_x e^{-\big(\lambda + r\big)\tau} \Big(K - S_\tau\Big)^+, \quad \bar{U}_*(x) = \sup_{\tau \in \bar{\mathrm{M}}_0^\infty} \mathbf{E}_x e^{-\big(\lambda + r\big)\tau} \Big(K - S_\tau\Big)^+ I\big(\tau < \infty\big).$$

Here $U_*(x)$ and $\bar{U}_*(x)$ are rational prices for the seller by class M_0^∞ and $\bar{\mathrm{M}}_0^\infty$ all finite and infinite stopping times.

Theorem 4.1. *Let $\lambda \geq 0$. Then*

$$U_*(x) = \bar{U}_*(x) = \begin{cases} x - K, & x \leq x^*, \\ c_* x^{z_2}, & x > x^*, \end{cases}$$

where

$$z_2 = -\frac{1}{2} \cdot \frac{2r - \sigma^2 + k}{\sigma^2}, \quad c_* = \big|z_2\big|^{|z_2|} \left(\frac{K}{1 + |z_2|}\right)^{1 + |z_2|},$$

$$x_* = K \cdot \frac{|z_2|}{1 + |z_2|}, \quad k = \sqrt{4r^2 + 4r\sigma^2 + \sigma^4 + 8\sigma^2\lambda}.$$

There exists an optimal moment in the class $\bar{\mathrm{M}}_0^\infty$, and the moment $\tilde{\tau} = \inf\Big\{t \geq 0 : \ S_t \geq x_*\Big\}$. Where in

$$\mathbf{P}_x\big(\tau_* < \infty\big) = \begin{cases} 1 & \text{if} \quad r \leq \frac{\sigma^2}{2} \quad \text{or} \quad x \leq x_*, \\ \left(\frac{x_*}{x}\right)^{\frac{2r}{\sigma^2} - 1} & \text{if} \quad x > \frac{\sigma^2}{2} \quad \text{and} \quad x > x_*. \end{cases}$$

Here the function $g(x) = \Big(K - x\Big)^+$ is limited. Theorem 4.1 is proved similarly to Theorem 3.1.

5. Standard European call option

Following [39], using Brownian motion and solutions of stochastic differential equations, one can give a probabilistic representation of solutions $V = V\big(t, x\big)$ of the Cauchy problem for parabolic equations of type $\dfrac{\partial V}{\partial t} = \mathrm{L}(x)V$, where $\mathrm{L}(x)$ is an elliptic operator. See [1] for details. It can also be shown that the function $V = V\big(t, x\big)$, under the assumption that it belongs to the class $\mathrm{C}^{1,2}$, satisfies $t > 0$ and $x \in E$ the equation

$$\frac{\partial V}{\partial t} + \beta V = \mathrm{L}V, \tag{36}$$

where

$$\mathrm{L}V\big(t, x\big) = rx \frac{\partial V\big(t, x\big)}{\partial x} + \frac{1}{2}\sigma^2 x^2 \frac{\partial^2 V\big(t, x\big)}{\partial x^2} \tag{37}$$

with initial condition

$$V\big(t, x\big) = g(x). \tag{38}$$

Let $t \in [0,\, T]$. Let's introduce a function $Y\big(t,\, x\big) = Y\big(T-t,\, x\big)$. Then for $Y\big(t,\, x\big)$ and (36), (38) it follows that the function $Y = Y\big(t,\, x\big)$ for $t < T$ satisfies the equation

$$\frac{\partial Y}{\partial t} + \beta Y = \mathrm{L}Y, \tag{39}$$

with the boundary condition

$$Y\big(T,\, x\big) = g(x). \tag{40}$$

The fundamental equation (39) is the **Feynman-Kac** equation. This equation is used to calculate the rational value $\Big(V\big(T,\, x\big) \,=\, Y\big(0,\, x\big)\Big)$ of a standard European call option with $g(x) = \big(x-K\big)^+$ and $\lambda = 0$.

Finding Rational Cost $V^*\big(T,\, x\big) = Y^*\big(0,\, x\big)$.

Let's define

$$\tau_0^T = \inf\Big\{0 \le t \le T :\; Y^*\big(t,\, S_t\big) = g\big(S_t\big)\Big\} \tag{41}$$

and

$$D_t^T = \Big\{x \in E :\; Y^*(t,\, x) = g(x)\Big\}, \tag{42}$$

$$C_t^T = \Big\{x \in E :\; Y^*(t,\, x) > g(x)\Big\}. \tag{43}$$

For $s \le t$ we have

$$Y^*(s,\, x) = V^*\big(T-s,\, x\big) \ge V\big(T-t,\, x\big) = Y^*(t,\, x).$$

Therefore, for $0 \le s \le t < T$

$$D_0^T \subseteq D_s^T \subseteq D_t^T \quad \text{and} \quad C_0^T \supseteq C_s^T \supseteq C_t^T.$$

In the case $t = T$, obviously, $D_T^T = E$ and $C_T^T = \varnothing$. Areas

$$D^T = \Big\{(t,\, x) :\; t \in \big[0,\, T\big),\; x \in D_t^T\Big\} \tag{44}$$

and

$$C^T = \Big\{(t,\, x) :\; t \in \big[0,\, T\big),\; x \in C_t^T\Big\} \tag{45}$$

in phase space $\big[0,\, T\big) \times E$ are called the regions of **stopping and continuing observations**, respectively. This is due to the fact that in "typical" optimal stopping problems for Markov processes, the moment τ_0^T turns out to be optimal:

$$\mathbf{E}_x e^{-\beta\tau_0^T}\, g\Big(S_{\tau_0^T}\Big) = V^*\big(T,\, x\big). \tag{46}$$

Because the

$$\tau_0^T = \inf\Big\{0 \le t \le T :\; S_t \in D_t^T\Big\}, \tag{47}$$

then the interpretation

$$D^T = \bigcup_{t<T}\Big(\big\{t\big\} \times D_t^T\Big)$$

the concept of the stopping region is clear: if: $\Big(t,\, S_t\Big) \in D^T$, then the observation stops.

The regions D^T and C^T can have highly complex structures depending on the properties of the functions $g = g(x)$ and the characteristics of the process $S = \Big(S_t\Big)_{t \le T}$. For instance, these regions, when viewed as sets $\big[0,\, T\big) \times E$ can be multiply connected, consisting of several "islands" where stopping occurs, and so on.

In the case of standard buyer and seller options (call and put options), where $g(x) = \big(x-K\big)^+$ for the call and $g(x) = \big(K-x\big)^+$ for the put, the regions D^T and C^T are simply connected.

For these options, the boundary of the stopping region ∂D^T can be represented as:

$$\partial D^T = \Big\{\big(t,\, x\big) :\ t \in \big[0,\, T\big),\ x = x^*(t)\Big\},$$

where, in the case of a call option

$$x^*(t) = \inf\Big\{x \in E :\ Y^*(t,\, x) = \big(x - K\big)^+\Big\}$$

and in the case of a put option

$$x^*(t) = \sup\Big\{x \in E :\ Y^*(t,\, x) = \big(K - x\big)^+\Big\}.$$

6. Stefan's Problem for Standard Buy and Ask Options

Here, following [39], we present the precise formulations of the Stefan problems corresponding to the two types of call and put options. We also discuss the qualitative properties of the solutions $Y^* = Y^*(t,\, x)$ and $x^* = x^*(t)$ related to these problems.

A) Buyer's option. Suppose that the $\big(B,\, S\big)$-market is described by relations (4) and (5) with $\mu = r$, $0 \le t \le T$, and the payout function at the moment t has the form $f_t = e^{-\lambda t} g\big(S_t\big)$, where $\lambda \ge 0$ and $g(x) = \big(x - K\big)^+,\ x \in E = (0,\, \infty)$. The main results related to the considered option are as follows.

1) **The rational value** $V^*\big(T,\, x\big),\ x = S_0$, of such a discounted option is determined by the formula

$$V^*\big(T,\, x\big) = \sup_{\tau \in \mathrm{M}_0^T} \mathbf{E}_x e^{-\beta\tau} g\big(S_\tau\big), \tag{48}$$

where $\beta = \lambda + r$ and $\mathbf{E}_x$ is the averaging over the original magningal measure under the assumption $S_0 = x$.

2) Let for $t \in \big[0,\, T\big]$ and $x \in E$

$$Y^*(t,\, x) = \sup_{\tau \in \mathrm{M}_t^T} \mathbf{E}_{t,\,x} e^{-\beta(\tau - t)} g\big(S_t\big), \tag{49}$$

where $\mathbf{E}_{t,\,x}$ denotes the expectation with respect to the martingale measure, assuming $x = S_t$. The function $Y^* = Y^*(t,\, x)$ riepresents the least excessive upper bound of the function $g(x)$.

3) There exists a rational cost

$$V^*\big(T,\, x\big) = Y^*\big(0,\, x\big), \tag{50}$$

aand the rational moment for the buyer to terminate the observation and exercise the option is the time at which

$$\tau_T^* = \inf\Big\{0 \le t \le T :\ Y^*\big(t,\, S_t\big) = g\big(S_t\big)\Big\} \tag{51}$$

or, equivalently,

$$\tau_T^* = \inf\Big\{0 \le t \le T :\ \big(t,\, S_t\big) \in D^T \cup \Big\{\big(T,\, x\big) :\ x \in E\Big\}\Big\}. \tag{52}$$

4) The stopping region D^T nd the continuation region C^T are simply connected and are characterized by the following structure:

$$D^T = \bigcup_{0 \le t < T} \Big\{(t,\, x) :\ Y^*(t,\, x) = g(x)\Big\}, \tag{53}$$

$$C^T = \bigcup_{0 \le t < T} \Big\{(t,\, x) :\ Y^*(t,\, x) > g(x)\Big\}. \tag{54}$$

5) The function $Y^* = Y^*(t,\, x)$ on $\big[0,\, T\big) \times E$ belongs to the class $\mathrm{C}^{1,\,2}$. For each fixed $x \in E$ function $Y^*(\cdot,\, x)$ is non-increasing in t; for each fixed $t \in \big[0,\, T\big)$, the function $Y^*(\cdot,\, x)$ is nondecreasing and convex down in x.

6) **The boundary function** $x^* = x^*(t)$ is non-increasing on $\big[0,\, T\big)$, and the sets C_t^T and D_t^T for $t < T$ have the following form:

$$C_t^T = \Big\{x \in E :\ S_t < x^*(t)\Big\}, \quad D_t^T = \Big\{x \in E :\ S_t \geq x^*(t)\Big\}.$$

For $t = T$ many $C_T^T = \emptyset$ and $D_T^T = E$.

If $\lambda = 0$, then $x^*(t) = \infty$, which corresponds to that for each $t < T$

$$C_t^T = E, \ \ D_t^T = \emptyset.$$

In other words, for all $t < T$, observations should be continue regardless of the values taken by the prices S_t, which follows from the fact that the process $\Big(e^{-rt}\big(S_t - K\big)^+\Big)_{t\geq 0}$ is a submartingale, and by the Doob halting theorem for any $\Big(e^{-rt}\big(S_t - K\big)^+\Big)_{t\geq 0}$

$$\mathbf{E}_x e^{-r\tau}\Big(S_\tau - K\Big)^+ \leq \mathbf{E}_x e^{-rt}\Big(S_T - K\Big)^+.$$

A similar result holds in the discrete time case, which can be interpreted as follows: standard American-style and European-style call options are essentially equivalent.

7) The function $Y^* = Y^*(t,\, x)$, where $t \in \big[0,\, T\big)$, $\ x \in E$, along with the boundary function $x^* = x^*(t)$, where $0 \leq t < T$ is a solution to the "two-phase" Stefan problem, or a problem with a moving (free) boundary. In the region $C^T = \Big\{(t,\, x) :\ x < x^*(t),\ t \in \big[0,\ T\big)\Big\}$, the equation is:

$$-\frac{\partial Y^*(t,\, x)}{\partial t} + \beta Y^*(t,\, x) = \mathrm{L}Y^*(t,\, x); \tag{55}$$

in the region $D^T \cup \Big\{\big(T,\, x\big) :\ x \in E\Big\}$

$$Y^*(t,\, x) = g(x); \tag{56}$$

at the boundary $x^* = x^*(t),\ 0 \leq t < T$, the section of the "two phases" the Dirichlet condition is fulfilled:

$$Y^*(t,\, x^*(t)) = g(x^*(t)); \tag{57}$$

and the Neumann condition:

$$\frac{\partial Y^*(t,\, x)}{\partial x}\Big|_{x\uparrow x^*(t)} = \frac{dg(x)}{dx}\Big|_{x\downarrow x^*(t)}, \tag{58}$$

which is often called the smooth gluing condition.

For the solvability of the Stefan's problem (55) – (58) and properties of the boundary function $x^* = x^*(t)$.

B) Seller option. In this case $g(x) = \big(K - x\big)^+$. Properties 1) – 4) of the call option remain in effect, and the function $Y^* = Y^*(t,\, x)$ again belongs to the class $\mathrm{C}^{1,\,2}$. For each fixed $x \in E$ function $Y^*(\cdot,\, x)$ is non-increasing in t; for each fixed $t \in \big[0,\, T\big)$, the function $Y^*(t,\, \cdot)$ is non-increasing and convex downwards in x.

For any $\lambda \geq 0$ set C_t^T and D_t^T have the following form for $t < T$:

$$C_t^T = \Big\{x :\ S_t > x^*(t)\Big\}, \quad D_t^T = \Big\{x :\ S_t \leq x^*(t)\Big\}.$$

For $t = T$ many $C_T^T = \emptyset$ and $D_T^T = E$.

The boundary function $x^* = x^*(t)$ is a non-decreasing function in t; if $\lambda = 0$, then $\lim\limits_{t\uparrow T} x^*(t) = K$.

The Stefan's problem for $Y^*\big(t,\, x\big)$ and $x^*(t)$ is formulated in a similar way. In this case, equation (55) and conditions (56), (57) are combined, and condition (58) for $0 \le t < T$ takes the following form:

$$\frac{\partial Y^*\big(t,\, x\big)}{\partial x}\bigg|_{x\uparrow x^*(t)} = \frac{dg(x)}{dx}\bigg|_{x\downarrow x^*(t)}, \tag{59}$$

where

$$\frac{dg(x)}{dx}\bigg|_{x\downarrow x^*(t)} = -1, \tag{60}$$

because the $g(x) = \big(K - x\big)^+$ and $x^*(t) < K$.

For more information about the properties of the functions $Y^*(t,\, x)$ and $x^*(t)$ can be found in the work [1], which is specifically devoted to the standard American type put option and contains extensive information related to other options. [2], [3], [4], [5].

Some information about Brownian motion and stochastic processes, controlled diffusion processes, option prices and their models and calculations, American options, optimal stopping in American put options are contained in the following works: [11], [12], [13], [14], [15], [16].

7. Numerical solution of the Stefan's problem (55) – (58) for the call option

In the area $C^T = \Big\{\big(t,\, x\big) :\ x < x^*(t),\ t \in \big[0,\, T\big)\Big\}$ consider the equation

$$-\frac{\partial Y^*(t,\, x)}{\partial t} + \beta Y^*(t,\, x) = \mathrm{L}Y^*(t,\, x), \tag{61}$$

where $\beta = \lambda + r,\ \mathrm{L}Y^*(t,\, x) = rx\dfrac{\partial Y^*(t,\, x)}{\partial x} + \dfrac{1}{2}\sigma^2x^2\dfrac{\partial^2 Y^*(t,\, x)}{\partial x^2}$ and in the area $D^T \cup \Big\{\big(T,\, x\big) :\ x \in E\Big\}$ consider

$$Y^*(t,\, x) = g(x) \tag{62}$$

at the boundary $x^* = x^*(t),\ 0 \le t < T$, the section of "two phases", the Dirichlet condition is fulfilled:

$$Y^*\big(t,\, x^*(t)\big) = g\big(x^*(t)\big); \tag{63}$$

and the Neumann condition:

$$\frac{\partial Y^*(t,\, x)}{\partial x}\bigg|_{x\uparrow x^*(t)} = \frac{dg(x)}{dx}\bigg|_{x\downarrow x^*(t)}. \tag{64}$$

Let us discretize the phase domain $C^T = \Big\{(t,\, x) :\ x < x^*(t),\ t \in \big[0,\, T\big)\Big\}$ with t respect to the step τ, $t_n = n\tau,\ n = 0,\, 1,\, 2,\, \ldots,\, N,\ \tau = \frac{T}{N}$, with x respect to the step h, $x_i = ih,\ i = 0,\, 1,\, \ldots$. We also discretize the area $D^T \cup \Big\{\big(T,\, x\big) :\ \ x \in E\Big\}$. Omit index * above $Y^*(t,\, x)$). We approximate (61) by an implicit scheme, and for the discrete domain C^T_{ni} we obtain the difference equation

$$\alpha_i Y^{n+1}_{i+1} + \beta_i Y^{n+1}_i + \gamma_i Y^{n+1}_{i-1} = -\frac{1}{\tau}Y^n_i, \tag{65}$$

where $\alpha_i = -\Big(\frac{rx_i}{h} + \frac{\sigma^2x_i^2}{2h^2}\Big)$, $\beta_i = \Big(\frac{rx_i}{h} + \frac{\sigma^2x_i^2}{h^2} + \beta - \frac{1}{\tau}\Big)$, $\gamma_i = -\frac{\sigma^2x_i^2}{2h^2}$. In the discrete domain D^T_{ni}, we write the Dirichlet condition (63):

$$Y_i^n = g_i \quad \text{or} \quad Y^n\left(x_i^*\right) = g(x_i^*) \tag{66}$$

and the Neumann condition (64):

$$\left.\frac{Y_{i+1}^n - Y_i^n}{h}\right|_{x_i = x_i^* + 0} = \left.\frac{g_{i+1} - g_i}{h}\right|_{x_i = x_i^* - 0}. \tag{67}$$

If the condition $\beta > \frac{1}{\tau}$, is satisfied, equation (65) can be solved, for instance, using the sweep method. To solve (65) we ensure that at the boundary $\left(x^*\right)_i^n$, the two-phase conditions (66) and (67) are met. At each step, we verify that the "front" $\left(x^*\right)_i^n$ is defined at a grid point. If not, we can adjust the step sizes τ and h.

Next, issues of numerical modeling of the Stefan's problem are considered. [8], [9], [10], [17], [18].

8. Numerical modeling by the method of running the Stefan's problem for standard call options

As a first example for a call option, we consider an American call option with financial variables $K = 10$, $\sigma = 0.6$, $r = 0.25$, $\delta = 0.2$, $x_0 = 10$ and $T = 1$. For these data, proposes a value of $Y^*\left(0,\, K\right) = 2.18728$, which corresponds to the variant shown in following Figure 1

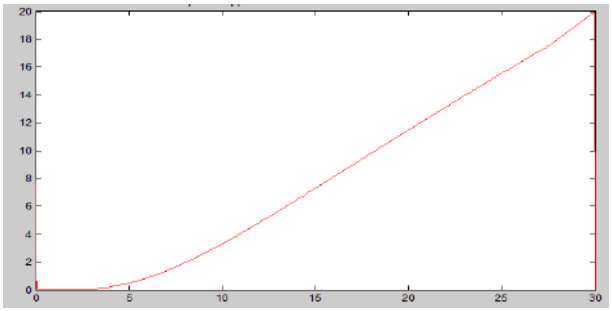


Figure 1. American call option pricing for $K = 10$, $\sigma = 0.6$, $r = 0.25$, $\delta = 0$, $T = 1$ (grid dimension 1600×1600).

The selected values for the test data $\left(K = 10,\, \sigma = 0.6,\, r = 0.25,\, \delta = 0,\, T = 1\right)$ are shown in Table 1 and visualized in Figure 2.

Table 1. Values $Y^*(0, x)$ of the American call option at K=10, σ=0.6, r=0.25, δ=0, T=1.

m (grid dimension), $Y^*_{\max}$	Sweep method	Finite element method (FEM)
100	2.181171	2.186701
200	2.186031	2.187181
400	2.186941	2.187251
800	2.187191	2.187271
1600	2.187261	2.187281

As a second example for call option, we consider an American call option with financial variables $K = 50$, $\sigma = 0.4$, $r = 0.1$, $\delta = 0$, $x_0 = 50$ and $T = 5/12$. For these data proposes a value of $Y^*\left(0,\, K\right) = 21.28638$, which corresponds to the variant shown in Figure 3.

The selected values for the test data $\left(K = 50,\, \sigma = 0.4,\, r = 0.1,\, \delta = 0,\, T = 5/12\right)$ are shown in Table 2 and visualized in Figure 4.

Table 2. Values $Y^*(0, x)$ of the American call option at K=50, σ=0.4, r=0.1, δ=0, T=5/12.

m (grid dimension), $Y^*_{\max}$	Sweep method	Finite element method (FEM),

100	21.280341	21.285981
200	21.285171	21.286281
400	21.286051	21.286351
800	21.286281	21.286371
1600	21.286351	21.286381

9. Numerical modeling by the method of running the Stefan's problem for standard put options

As the first example for a put option, we consider an American put option with financial variables $K = 10$, $\sigma = 0.6$, $r = 0.25$, $\delta = 0.2$, $x_0 = 10$, $T = 1$, which is shown in Figure 5.

This curve, shown in Figure 6, defines the option's early exercise strategy.

As a second example for a put option, we consider an American put option with financial variables $K = 50$, $\sigma = 0.4$, $r = 0.1$, $\delta = 0$, $x_0 = 50$ and $T = 5/12$, which is shown in Figure 7

We also calculate the point $x^*(0)$ for early exercise of the put option. Numerical results are given in Table 3 and illustrated in Figure 8.

Table 3. The boundary $x^*(0)$ of the early exercise of the $K = 50$, $\sigma = 0.4$, $r = 0.1$, $\delta = 0$, $T = 5/12$.

m (grid dimension), $Y^*_{\max}$	Sweep method	Finite element method (FEM)
100	40.93651	37.70211
200	37.04091	37.00011
400	37.04091	36.65201
800	36.58081	36.30571
1600	36.35291	36.30571

10. Conclusion

The advantage of finite difference methods (the sweep method) compared to the finite element method (FEA), in particular, the solution by the sweep method provides knowledge about the development of the option value function for each time step, i.e. the entire term structure of an American put option can easily visualize. See Figure 9. The accuracy of the sweep method is higher than FEA, see Figures 2, 4 and 8. At present $t = 0$, which is the "leading edge" of the surface, the shape of the cost function $Y^*(0, x)$ can be clearly seen, as shown in Figure 9. By selecting the $Y^*(t, x)$ functions for each t during the life of the option, one can obtain a complete term structure, and by the maturity date T approaches the non-smooth payoff function $Y^*\big(K - x\big)^+$. Of the entire surface of options, of particular interest is the development of the high contact point over time, namely $x^*(t)$, which is shown in Figure 10. This curve can be obtained by projecting the upper contact point at each time step onto the $x^*(t)$ – t-plane, and it determines the option's early exercise strategy.

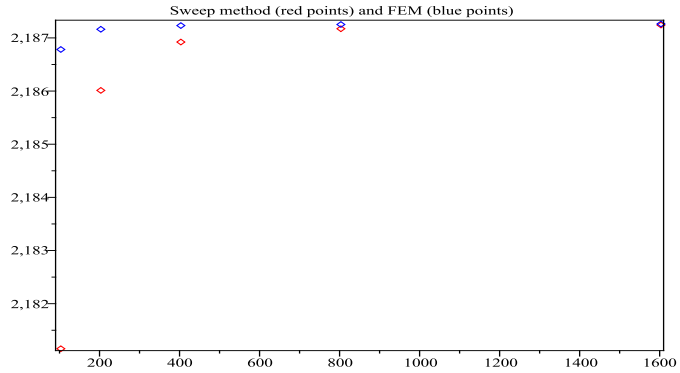


FIGURE 2. Convergence of the price an American call option $Y^*(0, x)$ $\big(K = 10,$ $\sigma = 0.6,\ r = 0.25,\ \delta = 0.2,\ T = 1\big)$ with grid dimension.

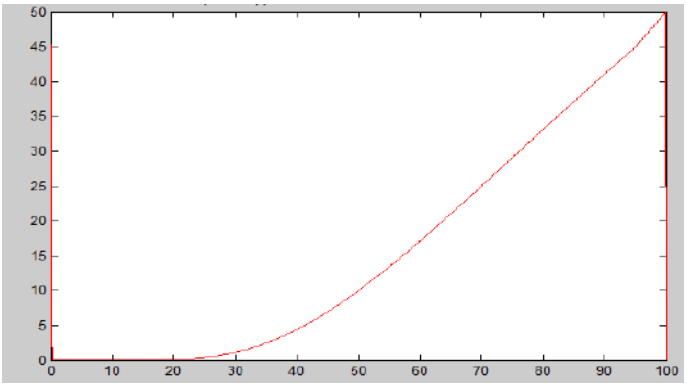


FIGURE 3. American call option pricing for $K = 50$, $\sigma = 0.4$, $r = 0.1$, $\delta = 0$, $T = 5/12$ (grid dimension 1600×1600).

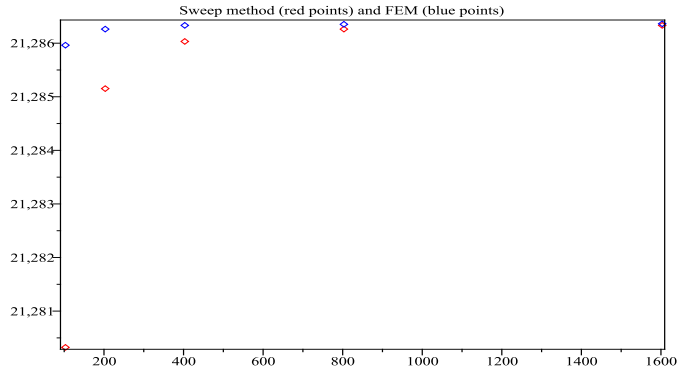


FIGURE 4. Convergence of the price of an American call option $Y^*(0, x)$ $\big(K =$ $50,\ \sigma = 0.4,\ r = 0.1,\ \delta = 0,\ T = 5/12\big)$ with increasing grid dimension.

Keywords: option prices, stock prices, equity diffusion markets, options of American and European types, Stefan's problem, numerical modeling.

AMS Subject Classification: 35R60, 39A50, 60G40, 65M06, 65M08.

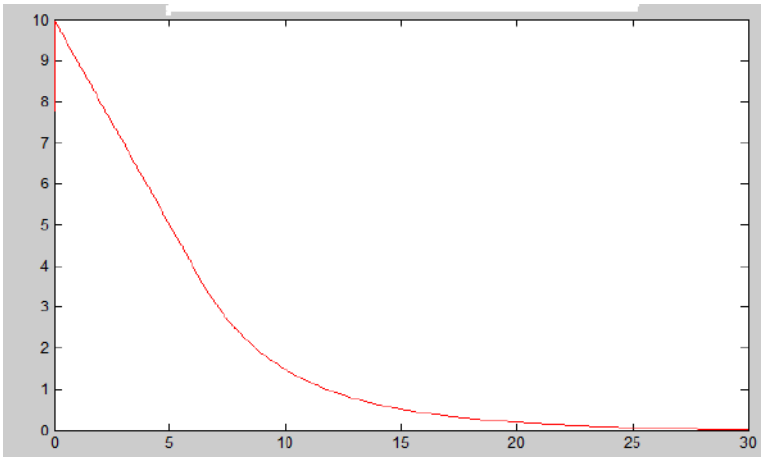


Figure 5. American put option pricing function $Y^*(0, x)$ for $K = 10$, $\sigma = 0.6$, $r = 0.25$, $\delta = 0.2$, $x_0 = 10$, $T = 1$ (grid dimension 1600×1600).

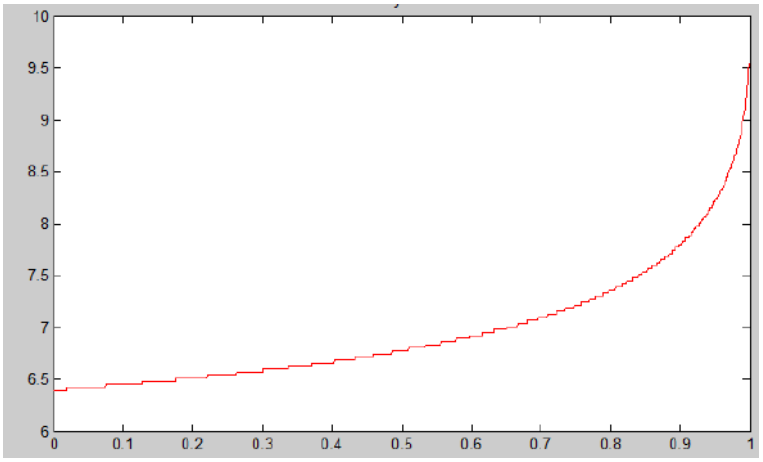


Figure 6. Time structure of the early exercise boundary $x^*(t)$ of a put option $K = 10$, $\sigma = 0.6$, $r = 0.25$, $\delta = 0$, $T = 1$.

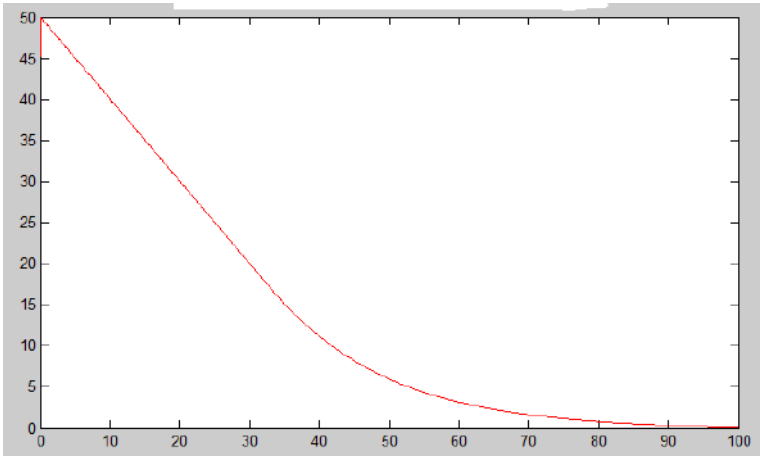


Figure 7. The pricing function $Y^*(0, x)$ of an American put option at $K = 50$, $\sigma = 0.4$, $r = 0.1$, $\delta = 0$, $T = 5/12$ (grid dimension 1600×1600).

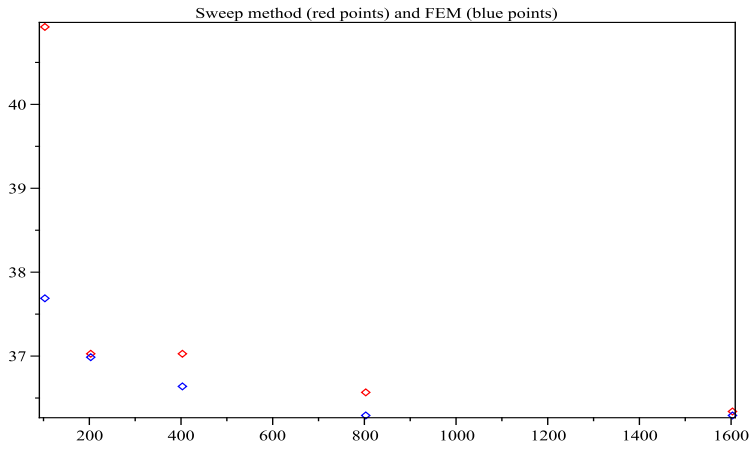


Figure 8. Convergence of the price of an American put option $\left(K = 50,\, \sigma = 0.4,\, r = 0.1,\, \delta = 0,\, T = 5/12\right)$ with increasing grid dimension.

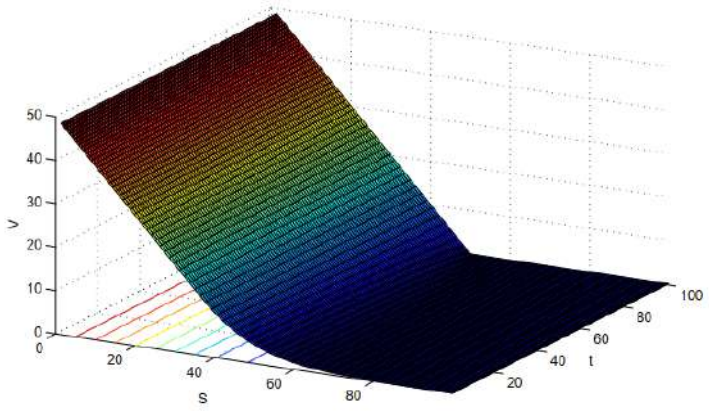


Figure 9. Structure $Y^*(t,\, x)$ of an American put option

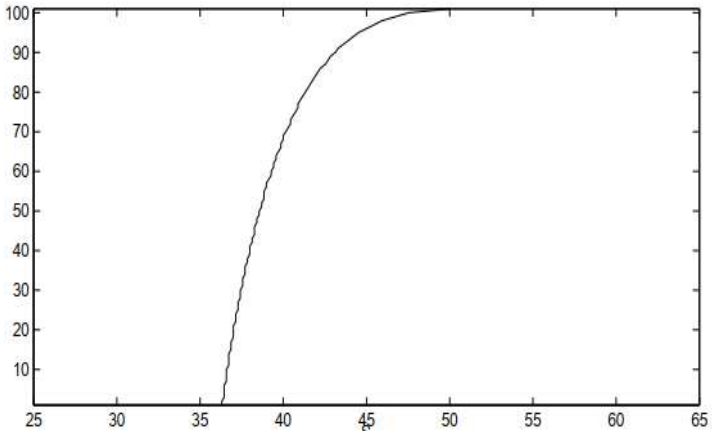


Figure 10. Structure of the free margin $x^*(t)$ of an American put option

References

[1] Shiryaev A. N., *Fundamentals of stochastic financial mathematics*, Vol.2, Theory, FAZIS, 1998, 544 p.

[2] Myneni R., The pricing of the American option, *Annals of Applied Probability*, Vol.2, No.1, 1992, pp.1-23.

[3] Carr P., Jarrow R., Myneni R., Alternative characterizations of American put options, *Mathematical Finance*, Vol.2, No.2, 1992, pp.87-106.

[4] Geske R., Johnson H. E., The American put options valued analytically, *Journal of Finance*, Vol.39, 1984, pp.1511-1524.

[5] Jacka S. D., Optimal stopping and the American put, *Mathematical Finance*, – V.1, No.2, 1991, pp.1-14.

[6] Shiryaev A. N., *Statistical sequential analysis. Optimal stopping rules*, Ed.2. Recycled, Nauka, 1976, 451 p.

[7] Shakenov K., Solution of Equation for Ruin Probability of Company for Some Risk Model by Monte Carlo Methods *Advances in Intelligent Systems and Computing* **441**, Chapter 12, Springer-Heidelberg Proceedings, Vol."Intelligent Mathematics II: Applied Mathematics and Approximation Theory" (Contributions from AMAT 2015), 2015, pp.169-182.

[8] Javierre-Perez E., Literature Study: Numerical methods for solving Stefan problems *Delft: Delft University of Technology, Report 03-16*, 2003.

[9] Rüdiger Seydel, Tools for Computational Finance, *Springer*, 2009.

[10] Shakenov K. K., Dispersion of the estimate for the solution of a system of linearized perturbed difference Navier–Stokes equations, *Computing Technologies*, V.7, No.3, 2002.

[11] Shakenov K. K., Stochastic processes and stochastic differential equations, *Bulletin KazNU, Series mathematics, mechanics, computer science*, No.7(35), 2002, pp.35-42.

[12] Shakenov K. K., Application of weak convergence of probability measures for approximation of diffusion processes,*Bulletin KazNU, Series mathematics, mechanics, computer science*, No. 7(35), 2002, pp.43-50.

[13] Shakenov K. K., Solution of Mixed Problem for Elliptic Equation by Monte Carlo and Probability-Difference Methods, *7th International Summer School, Conference "Let's Face Chaos through Nonlinear Dynamics", at the University of Maribor, 29 June – 13 July 2008, Slovenia*, American Institute of Physics, AIP Conference Proceedings 1076, 2008, pp.213-218.

[14] Kanat Shakenov., The Solution of the Inverse Problem of Stochastic Optimal Control, *Rev. Bull. Cal. Math. Soc.*, Vol.20(1), 2012, pp.43-50.

[15] Shakenov K., The Solution of the Initial Mixed Boundary Value Problem for Hyperbolic Equations by Monte Carlo and Probability Difference methods. *Trends in Mathematics. Fourier Analysis. Pseudo-differential Operators, Time-Frequency Analysis and Partial Differential Equations.* Springer International Publishing Switzerland, Birkhüser, 2014, pp.349-355.

[16] Serovajsky S., Shakenov I., Two forms of Gradient Approximation for an Optimization Problem for the Heat Equation, *Mathematical Modelling of Natural Phenomena*, Vol.12, No.3(2017), 2017, pp.139-145.

[17] Shakenov K., Baitelieva A., Solution of the Same Financial Mathematics Problem by Reducing to the Stefan Problem, *Vestnik KazNRTU*, No.1(137), 2020, pp.589-596.

[18] Shakenov K., Baitelieva A., Numerical Solution to Stefan's Problem for Buyer Option, *Vestnik KazNRTU*, No.6(142), 2020, pp.683-687.

AI AS A GEOPOLITICAL FACTOR: THE GROWING IMPORTANCE OF IMPROVING THE ALGORITHMIC FRAMEWORK

IBRAGIM SULEIMENOV[1], AKHAT BAKIROV[1]

[1]National Engineering Academy of the Republic of Kazakhstan, Almaty, Kazakhstan
e-mail: axatmr@mail.ru

The rapid development of AI is becoming a geopolitical factor, which, among other things, is accompanied by a significant increase in energy costs and associated factors. In this regard, improving the algorithmic foundation of AI is particularly urgent, including by overcoming one of the main drawbacks of the von Neumann architecture: the spatial separation of the memory unit and the computing processor necessitates the continuous movement of data between them, and, consequently, increased energy consumption. This paper examines the possibilities of improving the algorithmic foundation of AI using neuromorphic materials. It is shown that the advantages of neuromorphic materials remain unrealized due to a certain contradiction: existing approaches rely on the use of the same algorithms developed on known neural networks (which are, in fact, nothing more than computer simulations). It is shown that this contradiction can be overcome in the future by creating hybrid systems based on trigger calculators and multi-valued logic (modular arithmetic). It is demonstrated that the next logical step in the development of AI as a factor of social significance is the creation of sociomorphic materials.

Keywords: AI geopolitics, Neuromorphic materials, Post-Neumann architecture, sociomorphic materials.

12-11, 12F99

1. Introduction

Over the past two to three years, "artificial intelligence" in the public and, to a large extent, applied sense has become almost synonymous with large language models (LLM) and the generative systems built around them (chatbots, assistants, copy-pilots, code generators, etc.). Their successes are obvious: scalable training on large data sets has yielded a qualitative leap in interface usability (question-answering) and in a range of tasks—from search and summarization to programming. However, it is precisely the dominance of the LLM paradigm that has exposed a number of systemic limitations that are already impacting the scientific agenda, the economics of implementation, and the long-term prospects for approaching "strong AI."

The first part of this paper presents a brief overview of the key challenges and reasons why further progress in this field cannot be reduced to simply "even more data and computation"—a fundamental shift in the fundamental paradigm of AI development is required. The second part of this paper examines the prerequisites for the formation of a new AI development paradigm, arguing that it is not so much technical solutions that are at the forefront here, but rather the algorithmic foundation. It is also argued that consideration of this issue requires an analysis of the philosophical aspect (in particular, in connection with the feasibility of forming a new AI development paradigm taking into account the achievements of the theory of scientific revolutions).

2. Exhaustion of the Potential of the Existing AI Development Paradigm

2.1. Energy, Infrastructure, and the "Physics" of Scaling

Modern LLMs require large computing clusters on GPUs/TPUs and dense data center infrastructure. This turns progress in AI into not only a mathematical and algorithmic challenge,

but also an energy challenge: the growth in the power of "training" and especially "inference" (servicing user requests) computations increases the load on power grids and accelerates the construction of data centers [1-2]. It's significant that estimates of data center energy consumption are already comparable to those of large countries: a review by MIT News (based on academic and industry estimates) estimates that global data center energy consumption reached approximately 460 TWh in 2022 and could approach 1050 TWh by 2026 [2]. While not all of this load is associated with generative AI, LLMs are becoming one of the drivers of growth in computing power and density. An important detail: the energy cost doesn't end with the training phase. As LLMs become widely used in products, inference could become the dominant source of energy consumption, since each user request triggers a calculation on specialized hardware. This shifts the issue from the realm of isolated "expensive training sessions" to the realm of daily operational load. Water is also a factor: in some architectures, data center cooling relies on water loops. The same MIT News review notes that significant volumes of water can be used for cooling, citing a benchmark of approximately 2 liters of water per 1 kWh of data center energy consumption (an estimate dependent on climate and engineering design). At the corporate reporting level, this is also documented as a "bottleneck": for example, Microsoft explicitly states that this is becoming a serious problem due to the growth and increasing intensity of data center operations [3]. Thus, the "cost" of LLM manifests itself not only in electricity bills, but also in water, building materials, logistics, and so on. The consequence of this set of problems is strategic: if the "next leap" requires disproportionately more hardware and energy, then the development trajectory becomes vulnerable to limitations of the power grid, the pace of data center construction, the price of accelerators, and the geography of supply. This is no longer an abstract discussion, but a factor in the pace of research and competition.

2.2. The Pretraining Plateau and Industry Leaders' Doubts about the Prospects of "Current Scaling"

Model progress in the 2010s and early 2020s was often described by "scaling laws": more data + more parameters + more computation → better quality. However, by 2024, the position that simply increasing pretraining is reaching a plateau and no longer guarantees a qualitative leap has become increasingly clear.

A Reuters article, dedicated to the shift in focus of large labs, notes that the results of pretraining scaling have "reached a plateau," and that the "era of scaling" has given way to "the era of searching for the next algorithm"; the emphasis is shifting to other methods of improvement (including increasing inference-stage computation—test-time computing) [4]. The same publication notes practical obstacles: the high cost of training runs (tens of millions of dollars), the length (months), the risk of hardware failure, the lack of "readily accessible data," and even energy constraints that hinder training [4]. This is important as a symptom: the limitations are discussed not by outside critics, but by the participants who built the paradigm itself.

Yann LeCun shares a similar idea—that "the future is not limited to LLM in its current form." In an interview with IITM Shaastra, he explicitly states that "the future of AI is not LLM," emphasizing the limitations of the current approach [5]. The tone among the leaders may differ significantly, but the common denominator is that the "just scale it" trajectory is no longer considered a reliable path to strong AI. This means that architecture and learning modes require conceptual expansion (world models, causality, planning, integration of symbolic and neural network components, new computational foundations, etc.), otherwise, cost growth will outpace improvements in capabilities.

2.3. Unreliability, "Hallucinations," and the Limits of Trust

In this regard, the issue of response validity and reproducibility comes to the fore. LLMs are capable of generating plausible text, but they do not guarantee truthfulness and can confidently "invent" facts, references, legal precedents, or medical details. In 2024, this issue became the subject of work in leading journals: an article in Nature explicitly states that LLMs "often

hallucinate false conclusions and unsupported answers," which limits their implementation in critical areas [6]. The authors propose a statistically valid approach to detecting some hallucinations (via "semantic entropy"), but the very formulation of the problem demonstrates the need for add-ons that "insure" the model, since a "pure generator" does not provide the required robustness. Importantly, this unreliability is not an isolated bug, but a systemic consequence of LLMs being optimized for the plausibility of text continuations, not for verifiable truth. This leads to practical limitations: external validation is required (retrieval-augmented generation, fact-checking, citations, robustness testing), as well as evaluation procedures that measure not the "beauty of the text," but the probability and cost of errors.

2.4. "Synthetic degradation":

This occurs when models are trained on AI data and begin to "degrade" due to the declining quality of the data amid the exponential growth of AI content. More and more text, code, images, and "training" corpora are becoming generative in origin. This creates the risk of recursive contamination of training data: new-generation models are trained on the outputs of previous-generation models.

In 2024, Nature published a paper on model collapse—the degradation of models when trained on recursively generated data: under certain "self-copying" regimes, the data distribution "collapses," diversity decreases, and rare but important patterns begin to disappear [7]. This formalizes an intuitive danger: synthetic data is useful as a tool (for example, for class balancing or generating training problems), but with the uncontrolled growth of AI content in the information space, an obvious risk arises: new models will increasingly learn from secondary reflections of reality, losing quality and stability. Consequently, the challenge is no longer simply "collecting more data," but ensuring the provenance, diversity, and "purity" of the data, otherwise scaling becomes degradation.

2.5. Dataset Transparency, Licenses, and Legal Barriers

Modern LLMs are trained on huge data collections, the origin of which is often poorly documented. This creates two problems: (1) legal – copyright holders challenge the use of their data; (2) scientific and engineering – unknown origin and quality of the data lead to hidden biases and errors. MIT News, describing an audit of over 1,800 text datasets, indicates that over 70% of the datasets lacked licensing information, and approximately 50% had metadata errors [8]. This means that the typical LLM training pipeline loses context: what data can be used, under what conditions, with what restrictions and risks. Against this backdrop, legal pressure is increasing: a U.S. report The Copyright Office (May 2025) emphasizes that dozens of legal disputes are ongoing around training generative models on copyrighted works, and that the balance between innovation and protecting the creative ecosystem is becoming a key policy issue [9]. Thus, "data" in the LLM era is not simply a resource, but an object of rights and obligations; and the opacity of data origin is gradually becoming a technological risk (limited scalability, rising licensing costs, the need for filtering and documentation).

2.6. Security and Vulnerabilities (Prompt Injection, etc.)

LLM systems are vulnerable to input data manipulation (prompt injection), confidential information leaks, jailbreaking, and the injection of malicious instructions into the content that the model subsequently processes (e.g., in RAG scripts). Significantly, a specialized threat taxonomy has emerged for LLM applications: OWASP has released a Top 10 list of risks for LLM-based applications, with the top items being Prompt Injection, data leaks, authenticity/integrity issues, excessive agent privileges, and so on [10]. This shifts the discussion from "the model is sometimes wrong" to the realm of security engineering: LLMs are becoming a component of information systems and must be designed according to secure-by-design principles, otherwise their natural "text-awareness" becomes a front line of attack.

2.7. Supply Chains, Materials, and "Hidden" Environmental Effects

Finally, there are issues that are rarely covered in brief overviews but are crucial for long-term prospects: material intensity and supply chain sustainability. Growing demand for GPUs/TPUs

entails material extraction and processing, chemically complex manufacturing, logistics, and subsequent disposal. In the research literature, this is reflected in the need to evaluate LLM not only by text quality, but also by sustainability, privacy, the "digital divide," and ethical implications. The environmental footprint of raw materials and accelerator production is emphasized separately [11]. Taken together, this means that even if algorithmic progress continues, its cost is increasingly dependent on "non-algorithmic" factors—energy, water, materials, regulation, and safety.

2.8. Conclusions for Section 2

The issues listed above set the following framework. Modern AI in its LLM form is a powerful, but resource-intensive and conceptually limited tool. The resource component (electricity, water, equipment) is becoming a factor in the speed of progress [1–3]; "pure scaling" faces a plateau and data shortage, as acknowledged by industry leaders [4–5]; the lack of reliability and hallucinations require separate scientific and engineering "insurance" [6]; Recursive learning within the existing paradigm creates a risk of degradation in the future [7]; the opacity of datasets and the legal status of data become a key barrier to scaling [8–9]; the security of LLM applications requires independent standards and practices [10]. In this logic, further development of AI cannot be viewed as a linear continuation of the current architecture/paradigm: a new algorithmic foundation and new engineering principles are required, otherwise the increase in costs will outpace the growth of utility.

3. A New Paradigm for AI Development: From Neuromorphic to Sociomorphic Materials

3.1. Neuromorphic Materials and Computing Technology with a Quasi-Biological Basis

The concept of "strong" AI, in one way or another, presupposes a comparison of AI with a biological prototype, i.e., with human intelligence. Therefore, it seems quite logical that in recent years, increased attention has been paid by researchers to the development of computing systems based on physicochemical structures, which can be called quasi-biological. In particular, the literature reflects attempts to implement computing devices based on biological polymers, which are informational macromolecules (primarily DNA [12-14]). Such attempts correspond to a general trend associated, in particular, with the development of various types of organic electronics [15-17], including organic transistors [18-20], organic optoelectronics [21-23], etc. The need to modernize the elemental base of computing technology (which, in turn, presupposes the use of a non-trivial physical and chemical basis) is beyond doubt. The potential for further development of classical semiconductor technology built on the von Neumann architecture has already been largely exhausted [24-26]. The authors of works such as [27,28] share this opinion. At a minimum, the von Neumann architecture is currently considered to lead to increased energy consumption, which is associated with the constant exchange of data between the computing nodes themselves and the storage elements. It is clear that the crisis trends in the development of AI (Section 2) are exacerbated by the crisis trends in the development of computing technology itself, in particular, this directly concerns energy consumption issues.

From this perspective, neuromorphic materials, which de facto represent various options for the physical implementation of neural networks [29-31], are of considerable interest. Such materials, including those based on organic electronics, have also been actively developed in recent years [32-34]. A key advantage of such materials is their ability to implement parallel computing. This significantly improves performance and saves energy. Furthermore, it is possible to implement neuromorphic materials, initially oriented toward parallel computing, specifically modular arithmetic and algebraic ring theory [35]. In this regard, it is worth emphasizing that existing processors de facto perform calculations modulo a number representing some power of two. Consequently, it is permissible to consider creating computers that utilize modulo operations for other numbers, especially since this approach is fully justified, for example, in terms of improving convolutional neural networks [36].

Neuromorphic materials have already proven themselves quite successful in practice. For example, work [37] demonstrated that they can successfully address the problem of sensory

multitask learning aimed at creating systems analogous to the human retina and designed for pattern recognition. In [38,39], it was also shown that organic transistors used as the basis for the construction of neural networks (or their analogues) provide decentralized training on the corresponding carrier (crystal). From the perspective of the goals of this work, it is also appropriate to note the advances in the field of interfacing AI with neuromorphic materials [40-42].

However, if we proceed from the basic tenets of the modernized theory of scientific revolutions [43], then attention should be focused on the existence of a very specific contradiction, noted, in particular, in [26]. On the one hand, existing varieties of neuromorphic materials are oriented towards the use of learning algorithms developed for classical neural networks [37-39]. On the other hand, the use of such materials in computing must necessarily presuppose the utmost transparency of the arithmetic/logical operations they perform. Consequently, a contradiction arises due to the fact that existing types of neural networks generally remain logically opaque. The algorithms for training them are well-known and well-developed, but the training results are often difficult not only to predict but even to interpret. Difficulties of this kind, as is well known, led to the need to develop explainable neural networks [44-46] (and, correspondingly, explainable AI [47-49]). A clear illustration of this conclusion is provided by works such as [37-39]: the algorithm for the functioning of neuromorphic materials used in them was essentially a combination of neural network algorithms and the algorithms on which classical logical circuits are built. Therefore, the development of specific algorithms specifically designed to ensure the functioning of neuromorphic materials is of interest. Moreover, based on the provisions of the modernized theory of scientific revolutions [43], it can be concluded that different types of neuromorphic materials will ultimately be oriented toward different algorithmic foundations, including those using nontrivial algebraic structures (finite algebraic rings, Galois fields), as well as the closely related multi-valued logic. This point of view is largely consistent with the opinion of the authors of [50], who substantiate the importance of forming a new paradigm for biological artificial intelligence. A definite step in this direction was made in [35], where it was shown that there exists an algorithm structurally based on neural networks and allowing the implementation of an analogue of a binary adder. As current practice in computer development shows, this is sufficient for implementing an analogue of existing processors. The advantage of this algorithm for its application to neuromorphic materials is that it is initially oriented towards parallel computing, i.e., it automatically ensures the implementation of one of the main advantages of the physical implementation of neural networks. Algorithm [35] is based on the use of residue number systems (RNS), which are increasingly being used [51-53]. Residue number systems can be put in correspondence with a certain type of finite algebraic rings (rings of subtraction classes of an integer by some modulus). We also note that electronic adders and modulo-integer multipliers [54, 55], which find various practical applications [56, 57], and which can also be used to implement computational techniques based on residue number systems, have been actively developed recently. Investigations into the field of modular adders and multipliers are also reflected in the patent literature [58-61].

3.2. Sociomorphic Materials: The Question of the Complementarity of Society and AI

There is no need to prove that the development of AI is acquiring pronounced socioeconomic significance, which is why AI is becoming a geopolitical factor. It follows that the creation of AI that would be maximally complementary to society is of current interest [62]. Based on this premise, the cited work proposed an interpretation of the term "sociomorphic materials" that, among other things, reflects philosophical concepts about the nature of complex systems, based on modern achievements in this field [63]. Specifically, the current level of research in the field of complex systems theory makes it possible to discover general patterns inherent in such systems, regardless of their nature. The simplest application of sociomorphic materials is the simulation of processes occurring in society, which is also inherently a complex system. The concept of sociomorphic materials also allows for a slightly different approach to the problem of "strong"

AI. Specifically, alongside the mundane (personal) level of information processing, there is also a suprapersonal level [64].

The existence of this level is determined by the fact that human intellect (as well as consciousness) has a dual nature—it simultaneously contains both individual and collective "components" [65]. This can be proven without even resorting to mathematical models, one of which is proposed in [63]. Indeed, intelligence is an object of a purely informational nature; human intellect (as well as consciousness) does not physically exist. Physically, only electrochemical impulses, exchanged between neurons in the brain, exist. Similarly, the exchange of information between individuals is also reduced to the exchange of signals between neurons, but only belonging to different brains. As a result of interpersonal communications, therefore, a common neural network is formed, which (on a planetary scale) can be identified with the noosphere.

Human consciousness is an information object formed by the neural network of the brain. Similarly, the global neural network is also capable of forming non-trivial information objects of various types, which are only indirectly related to the memory and consciousness of individuals. The qualitative differences between the above-mentioned levels of information processing can be illustrated by the following argument. As is well known, the ability of neural networks to process information nonlinearly depends on the number of network elements. Otherwise, it would be pointless to implement neural networks containing an ever-increasing number of elements.

Consequently, the global neural network—the noosphere—possesses certain additional properties and capabilities that cannot be reduced to the properties of local neural networks formed by individual brains. It is this fact that forms the basis for the conclusion about the existence of a suprapersonal level of information processing [65]. At this level of information processing, a wide variety of information objects are formed [66], including, in particular, any highly developed scientific theories, natural languages, etc. The collective unconscious is also formed at this level [67] (this concept is one of the key concepts in analytical psychology). Furthermore, everything that constitutes human intelligence cannot be unrelated to the information objects of the type mentioned above. In particular, the intelligence of a particular individual is inherently dependent on the sum of their accumulated knowledge, their skills in using it in practice, etc. The source of such knowledge are information objects formed at the transpersonal level of information processing. Consequently, human intelligence can be considered a projection of a set of transpersonal information objects (their totality can be interpreted as the collective consciousness) onto a relatively independent fragment of the noosphere, localized in the individual brain. The individuality of intelligence is largely determined by which specific components of the above-mentioned information objects are "projected" onto the individual brain [64]. Therefore, from the perspective of the neural network theory of the noosphere, intelligence is that component of the personality structure that is closest to the collective consciousness.

This conclusion allows for a significant rethinking of the concept of "strong" AI. Indeed, "strong" AI was previously considered in terms of its gradual approximation of individual intelligence. The neural network theory of the noosphere suggests the existence of at least one more "object of comparison"—the collective consciousness. The concept of sociomorphic materials [62] is aimed at implementing "strong" AI precisely in this vein. As noted in the cited work, the first step in this regard is the synthesis of materials that model the behavior of complex systems, aimed at the most complete interpretation of the category of complexity. The advances in high-molecular chemistry discussed above, aimed at creating neuromorphic materials, make this approach quite sensible. Sociomorphic materials can be created precisely on this basis, but with due consideration for the need to modernize the algorithmic foundation.

4. Conclusions

The existing paradigm for the development of AI has largely exhausted its potential for further development, which is becoming increasingly evident as AI-based tools are increasingly implemented and made accessible to a wider range of consumers. Based on the fundamental concepts of the theory of scientific revolutions, it is reasonable to conclude that the further

development of AI will necessarily be linked to the modernization of its algorithmic foundation. Such modernization is inevitably linked to the modernization of the hardware components of computing technology, for which the quasi-biological component is becoming increasingly important. The conclusion regarding the dual nature of human intelligence and consciousness, as well as the fact that AI is acquiring increasingly pronounced geopolitical significance, forces us to significantly rethink the concept of "strong" AI. This concept does not necessarily need to focus on gradually approaching individual intelligence. It is more appropriate to interpret it in terms of approaching collective consciousness, which determines the relevance of the development of sociomorphic materials, which can be realized based on modern advances in the synthesis and research of neuromorphic materials.

Funding: This research is funded by the Committee of Science of the Ministry of Science and Higher Education of the Republic of Kazakhstan: AP23490107

References

[1] International Energy Agency. Energy and AI. Paris: IEA, 2025, *URL: https://www.iea.org/reports/energy-and-ai*, Access date: 25.12.2025.

[2] MIT News. Explained: Generative AI's environmental impact, 2025, *URL: https://news.mit.edu/2025/explained-generative-ai-environmental-impact-0117*, Access date: 25.12.2025.

[3] Smith B., Nakagawa M. Our 2024 Environmental Sustainability Report. Microsoft On the Issues. 2024, *URL: https://blogs.microsoft.com/on-the-issues/2024/05/15/microsoft-environmental-sustainability-report-2024/*, Access date: 25.12.2025.

[4] Hu K., Tong A. OpenAI and others seek new path to smarter AI as current methods hit limitations // Reuters, *URL: https://www.reuters.com/technology/artificial-intelligence/openai-rivals-seek-new-path-smarter-ai-current-methods-hit-limitations-2024-11-11/*, Access date: 25.12.2025.

[5] IITM Shaastra. Future of AI is not with LLMs: Yann LeCun. 2024, *URL: https://shaastra.org/articles/future-of-ai-is-not-with-llms-yann-lecun/*, Access date: 25.12.2025.

[6] Farquhar S., Kossen J., Kuhn L., Gal Y. et al., Detecting hallucinations in large language models using semantic entropy, *Nature*, Vol.630, No.8017, 2024, pp.625-630.

[7] Shumailov I., et al., AI models collapse when trained on recursively generated data, *Nature*, Vol.631, No.8022, 2024, pp.755-759.

[8] Study: Transparency is often lacking in datasets used to train large language models // MIT News, *URL: https://news.mit.edu/2024/study-large-language-models-datasets-lack-transparency-0830*, Access date: 25.12.2025.

[9] United States Copyright Office. Copyright and Artificial Intelligence. Part 3: Generative AI Training (Pre-Publication Version). A report of the Register of Copyrights. May 2025, *URL: https://www.copyright.gov/ai/Copyright-and-Artificial-Intelligence-Part-3-Generative-AI-Training-Report-Pre-Publication-Version.pdf*, Access date: 25.12.2025.

[10] OWASP. OWASP Top 10 for LLM Applications (v1.1). 2024, *URL: https://owasp.org/www-project-top-10-for-large-language-model-applications/*, Access date: 25.12.2025.

[11] Mishra P., et al., Sustainability in large language model supply chains—insights and recommendations using analysis of utility for affecting factors, *Scientific Reports*, Vol.15, No.1, 2025, pp.33524.

[12] Demir, B., Akin Gultakti, C., Koker, Z., Anantram, M. P., Oren, E. E., Electronic Properties of DNA Origami Nanostructures Revealed by In Silico Calculations, *J. Phys. Chem.*, Vol.128, No.19, 2024, pp.4646–4654.

[13] Katz, E., *DNA- and RNA-Based Computing Systems, 1st ed.*, WILEY-VCH: Weinheim, 2021.

[14] Fan, D.; Wang, J.; Wang, E.; Dong, S., Propelling DNA Computing with Materials' Power: Recent Advancements in Innovative DNA Logic Computing Systems and Smart Bio-Applications, *Advanced Science*, Vol.7, No.24, 2020, pp.2001766.

[15] Krauhausen, I.; Coen, C.; Spolaor, S.; Gkoupidenis, P.; Van De Burgt, Y., Brain-Inspired Organic Electronics: Merging Neuromorphic Computing and Bioelectronics Using Conductive Polymers, *Adv Funct Materials*, Vol.34, No.15, 2024, pp.2307729.

[16] Friederich, P., Fediai, A., Kaiser, S., Konrad, M., Jung, N., & Wenzel, W., Toward design of novel materials for organic electronics, *Advanced Materials*, Vol.31, No.26, 2019, pp.1808256.

[17] Casalini, S., Bortolotti, C. A., Leonardi, F., & Biscarini, F., Self-assembled monolayers in organic electronics, *Chemical Society Reviews*, Vol.46, No.1, 2017, pp.40-71.

[18] Kukhta, N. A., Marks, A., & Luscombe, C. K., Molecular design strategies toward improvement of charge injection and ionic conduction in organic mixed ionic–electronic conductors for organic electrochemical transistors, *Chemical Reviews*, Vol.122, No.4, 2021, pp.4325-4355.

[19] Rivnay, J., Inal, S., Salleo, A., Owens, R. M., Berggren, M., & Malliaras, G. G., Organic electrochemical transistors, *Nature Reviews Materials*, Vol.3, No.2, 2018, pp.1-14.
[20] Huang, W., Chen, J., Yao, Y., Zheng, D., Ji, X., Feng, L. W., ... & Facchetti, A., Vertical organic electrochemical transistors for complementary circuits, *Nature*, Vol.613, No.7944, 2023, pp.496-502.
[21] Bryce, M. R., A review of functional linear carbon chains (oligoynes, polyynes, cumulenes) and their applications as molecular wires in molecular electronics and optoelectronics, *Journal of Materials Chemistry C*, Vol.9, No.33, 2021, pp.10524-10546.
[22] Yu, P., Zhen, Y., Dong, H., & Hu, W., Crystal engineering of organic optoelectronic materials, *Chem*, Vol.5, No.11, 2019, pp.2814-2853.
[23] Ullbrich, S., Benduhn, J., Jia, X., Nikolis, V. C., Tvingstedt, K., Piersimoni, F., ... & Vandewal, K., Emissive and charge-generating donor–acceptor interfaces for organic optoelectronics with low voltage losses, *Nature materials*, Vol.18, No.5, 2019, pp.459-464.
[24] Shalf J., The future of computing beyond Moore's Law, *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences*, Vol.378, No.2166, 2020.
[25] Leiserson, C. E., Thompson, N. C., Emer, J. S., Kuszmaul, B. C., Lampson, B. W., Sanchez, D., & Schardl, T. B., There's plenty of room at the Top: What will drive computer performance after Moore's law?, *Science*, Vol.368, No.6495, 2020, pp.eaam9744.
[26] Suleimenov, I.; Gabrielyan, O.; Kopishev, E.; Kadyrzhan, A.; Bakirov, A.; Vitulyova, Y., Advanced Applications of Polymer Hydrogels in Electronics and Signal Processing, *Gels*, Vol.10, No.715, 2024.
[27] Niu, X.; Tian, B.; Zhu, Q.; Dkhil, B.; Duan, C., Ferroelectric Polymers for Neuromorphic Computing, *Applied Physics Reviews*, Vol.9, No.2, 2022, pp.021309.
[28] Zhang, B.; Chen, W.; Zeng, J.; Fan, F.; Gu, J.; Chen, X.; Yan, L.; Xie, G.; Liu, S.; Yan, Q.; Baik, S. J.; Zhang, Z.-G.; Chen, W.; Hou, J.; El-Khouly, M. E.; Zhang, Z.; Liu, G.; Chen, Y., 90% Yield Production of Polymer Nano-Memristor for in-Memory Computing, *Nat Commun*, Vol.12, No.1, 2021, pp.1984.
[29] Gumyusenge, A.; Melianas, A.; Keene, S. T.; Salleo, A., Materials Strategies for Organic Neuromorphic Devices, *Annu. Rev. Mater. Res*, Vol.51, No.1, 2021, pp.47–71.
[30] Melianas, A.; Quill, T. J.; LeCroy, G.; Tuchman, Y.; Loo, H. V.; Keene, S. T.; Giovannitti, A.; Lee, H. R.; Maria, I. P.; McCulloch, I.; Salleo, A., Temperature-Resilient Solid-State Organic Artificial Synapses for Neuromorphic Computing, *Sci. Adv.*, Vol.6, No.27, 2020, pp.eabb2958.
[31] Park, M. H., Kim, Y., Choi, M. J., Kim, Y. B., Yun, J. M., Jeong, J. H., ... & Kang, S. J., Enhanced In-Sensor Computing with Spike Number-Dependent Plasticity Characteristics in an InGaSnO Optical Neuromorphic Device for Accelerating Machine Vision, *ACS nano*, Vol.19, No.13, 2025, pp.13107-13117.
[32] Hoch, F. L., Wang, Q., Lim, K. G., & Loke, D. K., Multifunctional Organic Materials, Devices, and Mechanisms for Neuroscience, Neuromorphic Computing, and Bioelectronics, *Nano-Micro Letters*, Vol.17, No.1, 2025, pp.1-26.
[33] Kim, K. N., Sung, M. J., Park, H. L., & Lee, T. W., Organic synaptic transistors for bio-hybrid neuromorphic electronics, *Advanced Electronic Materials*, Vol.8, No.1, 2022, pp.2100935.
[34] Tuchman, Y., Mangoma, T. N., Gkoupidenis, P., Van De Burgt, Y., John, R. A., Mathews, N., ... & Salleo, A., Organic neuromorphic devices: Past, present, and future challenges, *MRS Bulletin*, Vol.45, No.8, 2020, pp.619-630.
[35] Shaltykova, D.; Kadyrzhan, K.; Caiko, J.; Vitulyova, Y.; Suleimenov, I., Trigger-Based Systems as a Promising Foundation for the Development of Computing Architectures Based on Neuromorphic Materials, *Technologies*, Vol.13, 2025, pp.326.
[36] Suleimenov, I.; Kadyrzhan, A.; Matrassulova, D.; Vitulyova, Y., Peculiarities of Applying Partial Convolutions to the Computation of Reduced Numerical Convolutions, *Appl. Sci.*, Vol.14, 2024, pp.6388.
[37] Wu, X.; Wang, S.; Huang, W.; Dong, Y.; Wang, Z.; Huang, W., Wearable In-Sensor Reservoir Computing Using Optoelectronic Polymers with through-Space Charge-Transport Characteristics for Multi-Task Learning, *Nat Commun*, Vol.14, No.1, 2023, pp.468.
[38] Krauhausen, I.; Koutsouras, D. A.; Melianas, A.; Keene, S. T.; Lieberth, K.; Ledanseur, H.; Sheelamanthula, R.; Giovannitti, A.; Torricelli, F.; Mcculloch, I.; Blom, P. W. M.; Salleo, A.; Van De Burgt, Y.; Gkoupidenis, P., Organic Neuromorphic Electronics for Sensorimotor Integration and Learning in Robotics, *Sci. Adv.*, Vol.7, No.50, 2021, pp.eabl5068.
[39] Wang, W.; Jiang, Y.; Zhong, D.; Zhang, Z.; Choudhury, S.; Lai, J.-C.; Gong, H.; Niu, S.; Yan, X.; Zheng, Y.; Shih, C.-C.; Ning, R.; Lin, Q.; Li, D.; Kim, Y.-H.; Kim, J.; Wang, Y.-X.; Zhao, C.; Xu, C.; Ji, X.; Nishio, Y.; Lyu, H.; Tok, J. B.-H.; Bao, Z., Neuromorphic Sensorimotor Loop Embodied by Monolithically Integrated, Low-Voltage, Soft e-Skin, *Science*, Vol.380, No.6646, 2023, pp.735–742.
[40] Choi, Y., Jeong, S., Jeong, H., Han, S., Ko, J., Yu, S. E., ... & Bae, S. H., Advanced AI computing enabled by 2D material-based neuromorphic devices, *npj Unconventional Computing*, Vol.2, No.1, 2025, pp.8.
[41] Gandolfi, D., Mapelli, J., & Puglisi, F. M., Brain-inspired computing: from neuroscience to neuromorphic electronics for new forms of artificial intelligence, *Frontiers in Neuroscience*, Vol.19, 2025, pp.1565811.

[42] Li, Y., Ding, G., Zhai, Y., Lv, Z., Yan, Y., Xue, S., ... & Han, S. T., MXene-Based Flexible Memory and Neuromorphic Devices, *Small*, 2025, pp.2410914.
[43] Suleimenov, I.; Gabrielyan, O.; Vitulyova, Y., Dialectics of Scientific Revolutions from the Point of View of Innovations Theory, *wisdom*, Vol.24, No.4, 2022, pp.25–35.
[44] Ibrahim, R., & Shafiq, M. O., Explainable convolutional neural networks: a taxonomy, review, and future directions, *ACM Computing Surveys*, Vol.55, No.10, 2023, pp.1-37.
[45] Damilola, S., A Review of Unsupervised Artificial Neural Networks with Applications, *IJCA*, Vol.181, No.40, 2019, pp.22–26.
[46] Dike, H. U.; Zhou, Y.; Deveerasetty, K. K.; Wu, Q., Unsupervised Learning Based On Artificial Neural Network: A Review, *In 2018 IEEE International Conference on Cyborg and Bionic Systems (CBS); IEEE: Shenzhen*, 2018, pp.322–327.
[47] Adadi, A.; Berrada, M., Peeking Inside the Black-Box: A Survey on Explainable Artificial Intelligence (XAI), *IEEE Access*, Vol.6, 2018, pp.52138–52160.
[48] Carloni, G., Berti, A., & Colantonio, S., The role of causality in explainable artificial intelligence, *Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery*, Vol.15, No.2, 2025, pp.e70015.
[49] Mersha, M., Lam, K., Wood, J., AlShami, A. K., & Kalita, J., Explainable artificial intelligence: A survey of needs, techniques, applications, and future direction, *Neurocomputing*, Vol.599, 2024, pp.128111.
[50] Perera, J., Balasubramaniam, S., Somathilaka, S., Wen, Q., Li, X., Kasthurirathna, D., ... & Nelson, T., Wet-neuromorphic computing: A new paradigm for biological artificial intelligence, *IEEE Intelligent Systems*, 2025.
[51] Oke, A. A., Nathaniel, B. A., Bukola, B. F., & Ayopo, O. A., Residue number system based applications: a literature review, *Annals. Computer Science Series*, Vol.19, No.1, 2021.
[52] Chang, C. H., Molahosseini, A. S., Zarandi, A. A. E., & Tay, T. F., Residue number systems: A new paradigm to datapath optimization for low-power and high-performance digital signal processing applications, *IEEE circuits and systems magazine*, Vol.15, No.4, 2015, pp.26-44.
[53] Younes, D., & Steffan, P., Efficient image processing application using residue number system, *In Proceedings of the 20th International Conference Mixed Design of Integrated Circuits and Systems-MIXDES 2013*, pp.468-472.
[54] J. -L. Beuchat,, Some modular adders and multipliers for field programmable gate arrays, *Proceedings International Parallel and Distributed Processing Symposium, Nice, France, 2003*, pp.8.
[55] Kuo C-T, Wu Y-C., FPGA Implementation of a Novel Multifunction Modulo (2n ± 1) Multiplier Using Radix-4 Booth Encoding Scheme, *Applied Sciences*, Vol.13, No.18, 2023, pp.10407.
[56] Ahmed A.H. Abd-Elkader, Mostafa Rashdan, El-Sayed A.M. Hasaneen, Hesham F.A. Hamed, Efficient implementation of Montgomery modular multiplier on FPGA, *Computers & Electrical Engineering*, Vol.97, 2022, pp.107585.
[57] B. K. Patel and J. Kanungo, Efficient Tree Multiplier Design by using Modulo 2n + 1 Adder, *Emerging Trends in Industry 4.0*, 2021, pp.1-6.
[58] Irkhin V.P., Obukhov A.N., Gul'bin S.S., Patent RU 2 145 112 C1. G06F 7/49, Device for modulo addition and subtraction of numbers,
[59] Hidenori Ebihara, Kiyoto c/o Kawasaki Denki Co. Ltd. Kawasaki, Patent EP0801345B1. Circuit for modulo multiplication and exponentiation arithmetic.
[60] Valeriy Pavlovich Suprun, Patent EA030205B1. Modulo four adder, *2018-07-31*.
[61] Petrenko Viacheslav Ivanovich, Stepanian Nerses Ernestovich, Nelidin Iurii Romanovich, Patent RU2724597C1. Russian Federation, G06F 7/72. Multi-digit parallel adder modulo with serial transfer: 2019144521, *27.12.2019*.
[62] Shaltykova, D.; Sedláková, Z.; Kopishev, E.; Suleimenov, I., From Neuromorphic to Sociomorphic Materials: Perspectives and Prognoses, *Symmetry*, Vol.17, 2025, pp.2110.
[63] Suleimenov, I. E., Matrassulova, D. K., Moldakhan, I., Vitulyova, Y. S., Kabdushev, S. B., & Bakirov, A. S., Distributed memory of neural networks and the problem of the intelligences essence, *Bulletin of Electrical Engineering and Informatics*, Vol.11, No.1, 2022, pp.510-520.
[64] Suleimenov, I., Gabrielyan, O., & Matrassulova, D., Philosophical Foundations of Sciences and Prospects of Multivalued Logic in Describing Thinking, *Science & Education*, 2025, pp.1-19.
[65] Suleimenov, I. E.; Gabrielyan, O. A.; Bakirov, A. S.; Vitulyova, Y. S., Dialectical Understanding of Information in the Context of the Artificial Intelligence Problems, *IOP Conf. Ser.: Mater. Sci. Eng.*, Vol.630, No.1, 2019, pp.012007.
[66] Suleimenov, I. E., Gabrielyan, O. A., Massalimova, A. R., & Vitulyova, Y. S., WORLD SPIRIT FROM THE STANDPOINT OF MODERN INFORMATION THEORY, *European Journal of Science and Theology*, Vol.20, No.1, 2024, pp.19-31.
[67] Ibragim, S., Oleg, G., & Vitulyova, Y., Problems of many-valued logic from the point of view of the theory of socio-cultural code, *Journal of Ecohumanism*, Vol.3, No.4, 2024, pp.236-248.

EXPERIENCE IN DEVELOPING A DATABASE FRAMEWORK FOR THE IMPLEMENTATION OF AN INTEGRATED MINERALOGICAL INDICATOR MAPPING METHODOLOGY AIMED AT IDENTIFYING PROSPECTIVE ORE LOCALIZATION ZONES

GALINA LEBEDEVA[1]

[1]Ecoservice-S LLP, East-Kazakhstan Branch, Ust-Kamenogorsk, Kazakhstan
e-mail: lebedeva.gv@ecoservice.kz

Ecoservice LLP carries out scientific research within the framework of the Project "Methods and Technologies for Mineral Deposit Prospecting and Evaluation Using Artificial Intelligence."

The objective of the project is the implementation of a set of measures aimed at developing a methodology for integrated mapping of indicators to identify prospective localizations of ore objects within the study area. The study area includes the following ore subzones: the Aleisk–Ashalinskaya polymetallic subzone, the Priirtysh and Kalbinskaya rare-metal subzones, and the West Kalbinskaya gold-bearing subzone. The total area of investigation is approximately 45 thousand square kilometers. The analytical coverage includes gold, rare-metal, and polymetallic deposits, ore occurrences, and mineralization points. In 2025, materials related to gold-bearing objects were prepared for analytical purposes.

The main task of the first stage of the work is the collection and systematization of archival materials, data obtained from internet sources, monographs, and dissertations. The results of this stage include completed standardized memos prepared in Word format and a consolidated table compiled in Excel format.

Preparatory Stage

At the preparatory stage, the data structure and data format were established. The main authors of the work are Academician of the International Academy of Sciences of Ecology, Safety, and Nature Management (MANEB) P. G. Kayukov and Leading Researcher, Candidate of Geological and Mineralogical Sciences O. D. Gavrilenko.

Work with Standardized Memos in Word Format.

- Using cartograms of geological and geophysical exploration coverage and the catalogue of the Vostkaznedra Department archive, reports containing data on prospecting and geological exploration works for the target objects were selected. In parallel, a search for alternative data sources available on the Internet was conducted.
- The textual and graphical contents of the reports were analyzed, the most informative materials were selected, and applications for the acquisition of geological information were prepared and submitted to the Regional Center for Geological Information 'Kazgeoinform" LLP.

The research is funded by the Science Committee of the Ministry of Science and Higher Education of the Republic of Kazakhstan (Grant No. BR27100483) "Development of predictive exploration technologies for identifying ore-prospective areas based on data analysis from the unified subsurface user platform 'Minerals.gov.kz' using artificial intelligence and remote sensing methods".

- Standardized memos were completed for each ore object. The work with the memos was carried out directly in the archive collections, using information extracted from the accessible textual parts of the reports, typically Volume I. The memos are required for subsequent work with the main Excel-format table, as they allow assessment of the completeness and quality of information for each selected study object.

Work with the Table of Prospecting Prerequisites and Indicators in Excel Format.

- The main table of prospecting prerequisites and indicators in Excel format was compiled after obtaining materials from the Kazgeoinform archive collections and completing the standardized memos.

The final Excel table contains coordinates for 243 gold-bearing mineralization points, gold deposits, and gold ore occurrences. Among them, 70 deposits and ore occurrences are described in detail, in accordance with the completeness of data presented in archival reports and other published sources. In addition to the textual information, graphical materials containing geological–tectonic, geochemical, and geophysical data were selected. The table and graphical materials were transferred for processing into a relational geographic information system (ArcGIS Pro). This software was used for georeferencing and vectorization of graphical files, as well as for classification and integration of information into the database structure. A schematic map showing the spatial distribution of gold deposits, ore occurrences, and mineralization points included in the final Excel table of prospecting prerequisites and indicators is shown in Figure 1.

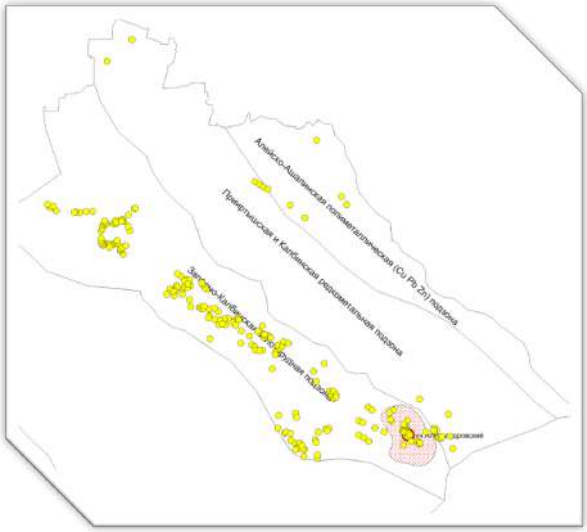


FIGURE 1. Spatial distribution scheme of gold deposits, ore occurrences, and mineralization points included in the final Excel table of prospecting prerequisites and indicators; the Aleksandrovskiy site is shown in red

EXAMPLE OF COMPLETING THE STANDARDIZED MEMO AND THE EXCEL TABLE

Compilation of the Standardized Memo. As an example of material systematization, the Aleksandrovskiy gold-bearing site is considered (see Appendix A, Table A1).

- The first step is to determine the location of the site. The location of the Aleksandrovskiy site indicates that it lies within the ore subzones considered in this study (Figure 1).
- Using the reference book *"East Kazakhstan Region. Gold"* [1], data on the Kuludzhun ore field were extracted and entered into the standardized memo, including the name of the ore belt, ore district, and general characteristics of the ore bodies.

- Based on the keywords "West Kalbinskiy gold-bearing belt," "Kuludzhun," "Kalba," and "gold deposits of East Kazakhstan," archival reports were identified in the Vostkaznedra Department collections [2, 3, 4, 7, 8, 9, 10].
- Directly within the archive collections, available information on prospecting prerequisites and indicators was extracted from the reports and recorded in the standardized memo. In some cases, such information may be absent for various reasons, including the exclusion of necessary exploration objectives from the technical assignment or the low technical level of exploration methods applied in earlier periods.

The standardized memo compiled by the authors represents the initial stage in the creation of the consolidated table. The memo reflects information available in open-access archival collections of the Vostkaznedra Department and is therefore incomplete. Nevertheless, even this limited dataset demonstrates the prospectivity of the considered site and motivates its further detailed analysis.

Conclusion for the Aleksandrovskiy Site Based on the Standardized Memo. The Aleksandrovskiy site is considered prospective based on the following factors:

- proximity to granitoids of the Kunushskiy complex, which are associated with ore formation within the Kuludzhun ore district (Dyachkov B. A.);
- identified ore zones with a total strike length of 170 m and ore intervals 1–3 m thick, with gold grades of 4–10 g/t.

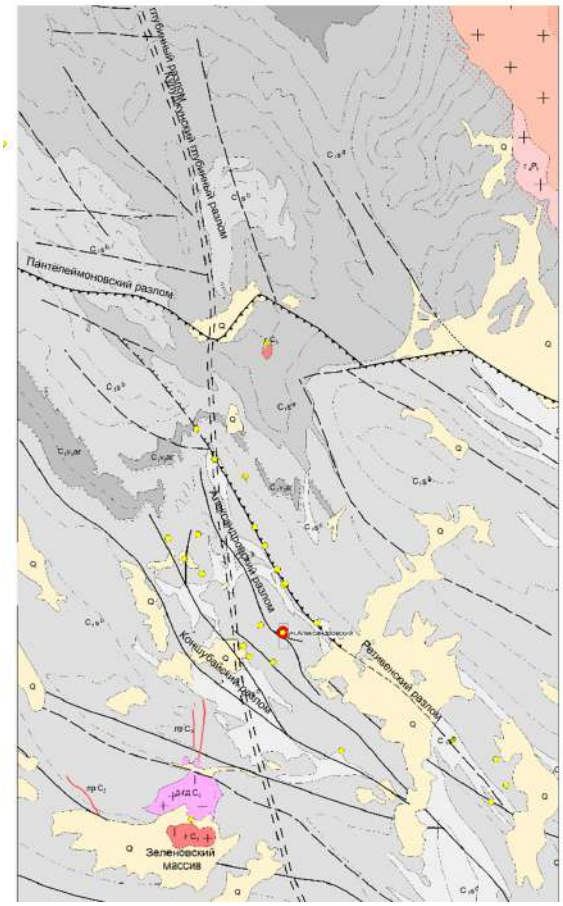
Work with the Table of Prospecting Prerequisites and Indicators in Excel Format. After receiving archival reports from the Regional Center for Geological Information "Kazgeoinform" LLP, the table of prospecting prerequisites and indicators was completed in Excel format. Graphical materials were reviewed separately, and the most informative datasets suitable for further processing in a geographic information system (ArcGIS Pro) were selected. The selected materials were numbered, and references to raster files were recorded in the table.

The table summarizes concise descriptions of gold-bearing sites, with emphasis on characteristics relevant to forecasting and identifying new prospective gold-bearing zones. The table incorporates information from archival reports [2, 3, 4, 7, 8, 9, 10] and published geological studies [1, 5, 6, 11], both for the Aleksandrovskiy site and for other gold occurrences within the test area. This dataset forms the basis for subsequent geological studies aimed at compiling a mineragenic map of the study area and developing a prototype methodology for prospecting typical industrial ore deposits using artificial intelligence.

Analysis of Graphical Appendices. The geological effectiveness of graphical materials included in the table of prospecting prerequisites and indicators is illustrated using the Aleksandrovskiy site as an example. Based on the analysis, the most informative graphical files were selected from the report by Verentsov [3] published in 1977. These materials allow assessment of prospecting prerequisites, indicators, and the overall prospectivity of the site. The selected files were georeferenced and vectorized by specialists in the WGS 84 coordinate system.

In other reports, graphical materials for the Aleksandrovskiy site are either absent (Vvedenskiy [2], Okunev [9], Minazov [8], Stepanov [10]), lack coordinate information required for georeferencing (Derevtsov [4], 1987), or largely duplicate the information presented in the 1977 report (Titov [7], 1978).

Figure 2 illustrates the geological and structural prospecting criteria of the Aleksandrovskiy site: (1) the presence of widely developed fault systems, including major faults, faults of local significance, and minor local fractures with predominantly sublatitudinal and northwestern trends; (2) localization of the site within the Aleksandrovskiy fault zone; (3) proximity to granitoid intrusions of the Kunushskiy complex (C_3kn), which are closely associated with gold deposits of the Kuludzhun ore district.



(a)

(b)

FIGURE 2. Geological map of the Aleksandrovskiy site showing its position relative to tectonic faults and granitoid intrusions of the Kunushskiy complex

Figure 3 illustrates indirect prospecting indicators: (1) a complex of geochemical dispersion halos of gold, arsenic, antimony, tungsten, and silver; (2) development of fractured and hydrothermally altered rocks; (3) numerous quartz veins, many of which were identified as gold-bearing as early as the nineteenth century.

Tectonic faults, quartz veins, and metasomatic alteration were identified both visually, as a result of geological investigations, and by geophysical methods, particularly induced polarization (IP).

The Aleksandrovskiy site is clearly expressed on this correlation scheme.

(a)

(b)

FIGURE 3. Correlation scheme of geophysical and geochemical fields and halos of the Kuludzhun site

For the Aleksandrovskiy site, a geological map with cross-sections along profiles was also compiled, incorporating the results of sampling from trenches and boreholes completed within zones of hydrothermally altered rocks. These rocks host gold–sulfide mineralization and are characterized by intense schistosity, iron oxidation, silicification, sericitization, graphitization, and kaolinization. The results of sampling from mine workings and boreholes confirm the prospectivity of the site. Sampling represents a direct prospecting indicator for the presence of mineralization.

In addition, the same report [3] presents the results of geophysical investigations using the IP method, including areas recommended for detailed prospecting and prospecting–evaluation works. The Aleksandrovskiy site is included among the recommended areas.

Results of Investigations for the Aleksandrovskiy Site. Analysis of the textual and graphical materials from archival reports makes it possible to conclude that, based on integrated investigation of the area, there is a high probability of discovering new gold ore bodies.

Prospecting Criteria. Prospecting criteria:

- All gold mineralization points within the Kuludzhun ore field, including the Aleksandrovskiy site, are developed along northwestern-trending fault zones in proximity to

granitoids of the Kunushskiy complex, which, according to Dyachkov B. A., are genetically associated with gold deposits of the studied area (Figure 2).

- Economic concentrations of gold are spatially associated with nodes of fault intersections and zones of hydrothermal alteration of rocks (Figure 3).

Indirect prospecting indicators:

- The central and northwestern parts of the Aleksandrovskiy zone are marked by secondary dispersion halos of gold (0.005–0.01 g/t), arsenic (0.005–0.01%), and tungsten (0.001–0.005%). Geophysical investigations delineated tectonic zones, silicification zones, zones of pyrrhotite mineralization, and zones of hydrothermally altered rocks.

Direct prospecting indicators:

- The Aleksandrovskiy area was explored by trenches spaced at intervals of 100–300 m and traced along strike for approximately 2000 m. The studied area represents the southeastern branch of the Aleksandrovskiy schistosity zone system and, in its northwestern part, includes the well-known gold-bearing Aleksandrovskaya quartz vein, which has been known since 1871. The vein was exploited by prospectors and a mine since 1907. In its central part, it was mined to a depth of 97 m and, on average, has a thickness of 5–10 m, being discontinuously traced over a length of 250 m.
- During 1975–1977, trenches, pits with crosscuts, and borehole No. 5 with a depth of 212 m were completed within the Aleksandrovskiy zone. Trenches intersected ore intervals with gold grades of 4–10 g/t, while gold content in the host rocks does not exceed 0.3–0.4 g/t.

Conclusions. The scientific research conducted within the framework of the Project "Methods and Technologies for Mineral Deposit Prospecting and Evaluation Using Artificial Intelligence" has not yet been completed; however, even at the intermediate stage and based on the example of a single gold-bearing site, it can be concluded that indirect prospecting indicators in the form of integrated geochemical and geophysical halos can be effectively applied in practice to identify prospective mineralized areas. Exploration works carried out at the Aleksandrovskiy site, including trenching and drilling, confirm this conclusion.

Based on professional experience in geological exploration, this study is considered to be of significant interest, as it integrates a large volume of heterogeneous geological information, forming a solid basis for subsequent analytical investigations aimed at forecasting areas potentially prospective for near-surface mineral deposits. With the availability of geological, geophysical, and geochemical information from deeper levels, the approach may also be applied to the identification of deposits at greater depths.

References

[1] East Kazakhstan Region. Gold. *Reference Book*. Ust-Kamenogorsk, Kazakhstan.
[2] Vvedenskiy A. V. Report on regional prospecting for gold deposits at the Kuludzhun site in 1982–1984. Vostkaznedra Department Archive, Ust-Kamenogorsk, Kazakhstan, 1984.
[3] Verentsov Yu. I. Report on areal gold prospecting within the Kuludzhun ore node conducted by the Kuludzhun exploration party in 1975–1977. Vostkaznedra Department Archive, Ust-Kamenogorsk, Kazakhstan, 1977.
[4] Derevtsov G. V. Report on detailed prospecting works at the Kuludzhun site in 1979–1987. Vostkaznedra Department Archive, Ust-Kamenogorsk, Kazakhstan.
[5] Dyachkov B. A., Zimanovskaya N. A. On the geological position and age of gold deposits in Eastern Kazakhstan. Almaty, 2013.
[6] Gold and Rare-Metal Deposits. East Kazakhstan Technical University named after D. Serikbayev. Ust-Kamenogorsk, Kazakhstan.
[7] Titov V. I. Report on prospecting and evaluation works at the Aleksandrovskiy site conducted in 1977–1978. Vostkaznedra Department Archive, Ust-Kamenogorsk, Kazakhstan, 1978.
[8] Minazov R. S. Report on prospecting and evaluation works at the Kuludzhun ore field in 1993–1996. Vostkaznedra Department Archive, Ust-Kamenogorsk, Kazakhstan, 1996.
[9] Okunev E. V. Prospects for development of a number of conserved quartz-vein gold deposits of Southern Kalba. Almaty, 1972.
[10] Stepanov A. E. Geological structure and mineral resources of map sheets M-44-106-B and 107-A. Final report of the Mirolyubovskaya prospecting and survey party for geological mapping at a scale of 1:50 000. Vostkaznedra Department Archive, Kazakhstan, 1968.
[11] Shcherba G. N. *The Great Altai*. Almaty, 1998.

TABLE 1. Standardized memo for the Aleksandrovskiy gold-bearing site

Type of information	**Information content**
Deposit	Aleksandrovskiy site
Report references / indicators	Reference book *"East Kazakhstan Region. Gold"*; archival reports [2–4, 7–10]
List of economic components	Primary component: gold
Ore-forming genesis	Hydrothermal
Morphological type of deposit	Quartz-vein and veinlet-disseminated
Formation	Gold-bearing formation
Structural position	Inner part of the Kuludzhun flexure, along a northwestern-trending tectonic fault; the northwestern segment includes gold-bearing quartz veins
Prospecting indicators at discovery	The Aleksandrovskaya vein has been known since 1871 and was discovered by a prospector based on secondary minerals exposed at the surface
Physical–geological models of ore bodies	En echelon quartz veins with variable thickness (0.1–2.0 m) and length (tens to hundreds of meters), locally transitioning into veinlet zones; delineated by drilling
Structural control	Gold-bearing quartz veins are confined to tectonic faults developed within the inner part of the Kuludzhun flexure
Physical properties of rocks and ores	Not reported in archival sources
Absolute geochronology	Not available
Quantitative description of geophysical fields	Presented in the appendix to the table of prospecting prerequisites and indicators
Interpretation of geophysical anomalies	Presented in the appendix to the table of prospecting prerequisites and indicators
Characteristics of prospecting surveys	Presented in the appendix to the table of prospecting prerequisites and indicators
Prospecting methods at different exploration stages	Reconnaissance surveys, geochemical and geophysical investigations, trenching, drilling, and sampling of workings and boreholes
Geological efficiency of applied methods	Not evaluated
Current status of the deposit	Under development
Deposit scale	Small
Lithology of ores and host rocks	Host sand–shale sequences and molasse-type sandy–siltstone deposits of the Kalba Formation (Lower Carboniferous)
Petrochemical and chemical properties of rocks and ores	Not reported in archival sources
Mineralogical features of ores	Average sulfide content 0.2–0.5%; in individual veins, scheelite content reaches up to 10% and stibnite up to 10%
Mineralogical zoning	Hydrothermal alteration of rocks; no information available on oxidation zones or weathering crust
Geochemistry	Graphical appendices: correlation scheme of geophysical and geochemical datasets at a scale of 1:10 000
Landscape conditions	Smoothed mountainous terrain with intensive turf cover in river valleys and on slopes
Loose cover characteristics	Proluvial deposits in river valleys
Exploration data (GIS format)	Geological logs of borehole No. 5 with logging data at a scale of 1:200
Prospectivity	Prospective; ore intersections traced along strike for 170 m

ARCGIS PRO–BASED MULTILAYER FEATURE ENGINEERING FOR MACHINE LEARNING PREDICTION OF GOLD MINERALIZATION

BEKDAULET KHUDAIBERGEN[1], ALMAS TEMIRBEKOV[2]

[1]Department of Computational and Statistical Sciences, Al-Farabi Kazakh National University, Almaty, Kazakhstan
[2]Department of Computational and Statistical Sciences, Al-Farabi Kazakh National University, Almaty, Kazakhstan
e-mail: secdet29@gmail.com

Abstract

This study presents a geospatial–machine learning framework for predicting gold mineralization using multilayer feature engineering developed in ArcGIS Pro. The study area was divided into four mineragenic subzones, and a unified database of verified mineral occurrences and background samples was constructed.

A high-resolution digital elevation model (DEM) was reprojected into WGS 84 / UTM Zone 45N, enabling extraction of key geomorphological attributes such as slope, curvature, and structural edge indicators [1]. Spatial predictors were integrated into a Python-based workflow using `pandas`, `numpy`, and `scikit-learn` [2].

Several supervised machine learning models were evaluated using cross-validation procedures. Future work planned for 2025–2026 includes the incorporation of regional geochemical datasets, interpolated geochemical surfaces, and advanced DEM-derived metrics such as topographic position index (TPI), terrain ruggedness index (TRI), and surface roughness.

Benchmarking of modern machine learning architectures [3] will support the development of a reproducible and scalable mineral prospectivity mapping pipeline suitable for regional gold exploration.

Keywords: gold mineralization, ArcGIS Pro, geospatial feature engineering, machine learning, DEM analysis.

AMS Subject Classification: 62H30, 68T09, 86A60.

1. Introduction: Relevance of the topic

Exploration of prospective mineralization zones in geologically complex regions remains a pressing task in modern applied geology. The intricate tectonic structure, heterogeneous terrain, and uneven distribution of known deposits significantly hinder the application of traditional forecasting methods, which are predominantly based on qualitative expert interpretation of data.

In the context of large volumes of heterogeneous spatial information, the use of Geographic Information Systems (GIS) becomes particularly valuable, enabling comprehensive analysis, visualization, and integration of diverse data types. In this study, the ArcGIS Pro software suite serves as the primary platform for multi-layer geospatial analysis, encompassing data preprocessing, derivation of secondary characteristics, and compilation of a unified feature set for subsequent modeling.

Augmenting the GIS approach with machine learning methods allows for the identification of hidden patterns in the spatial distribution of mineral deposits, shifting the focus from qualitative

This work was carried out within the framework of academic research at Al-Farabi Kazakh National University.

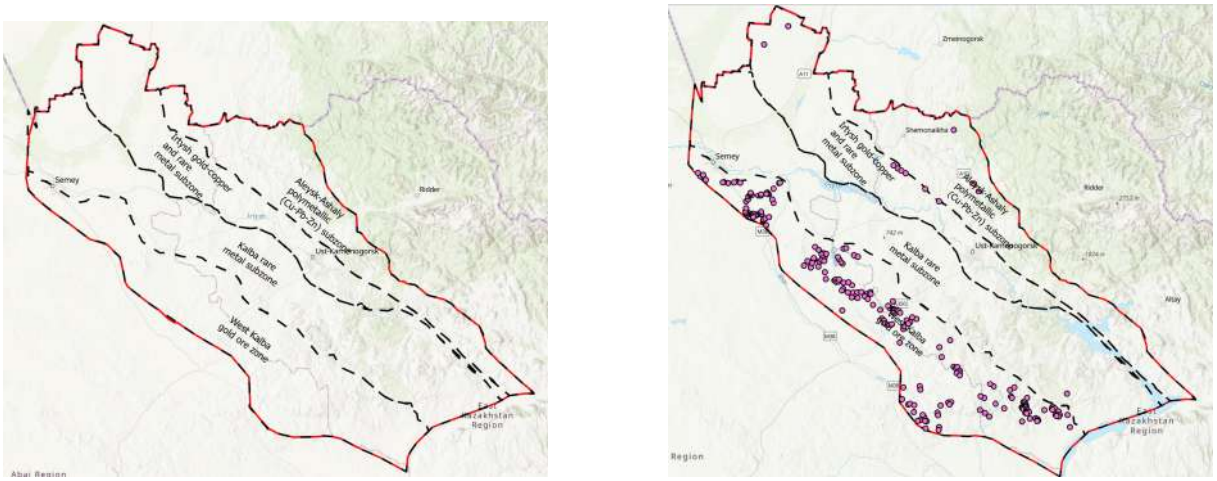
analysis to quantitative estimation of mineralization probability. This approach provides a more objective, reproducible, and interpretable assessment of an area's mineral potential.

Thus, integrating the capabilities of ArcGIS Pro with machine learning algorithms represents a relevant and effective strategy for improving the reliability of predictive modeling and supporting decision-making during the early stages of geological exploration.

2. Study Area

The study area is located in Eastern Kazakhstan, within a fold-mountain zone characterized by a complex tectonic structure and the presence of known mineral occurrences. The region is divided into four metallogenic subzones:

- **Irtysh gold-copper and rare metal subzone** – known for orogenic gold deposits, porphyry copper-gold systems, and associated rare-metal (REE, W, Mo) mineralization, hosted in late Paleozoic volcanic-plutonic complexes.
- **Alei-Ashaly polymetallic subzone (Cu–Pb–Zn)** – dominated by volcanogenic-hydrothermal and skarn-type polymetallic deposits within Middle to Upper Paleozoic carbonate-terrigenous sequences.
- **Kalba rare metal subzone** – characterized by Li-Be-Ta-Nb pegmatites and greisen-type Sn-W mineralization, related to Permian-Triassic granitoid magmatism.
- **West Kalba gold ore zone** – hosts mesothermal gold-quartz veins and stockwork mineralization, structurally controlled by regional shear zones and fault intersections.



(A) Geological overview of the study region (B) Deposit and subzone distribution

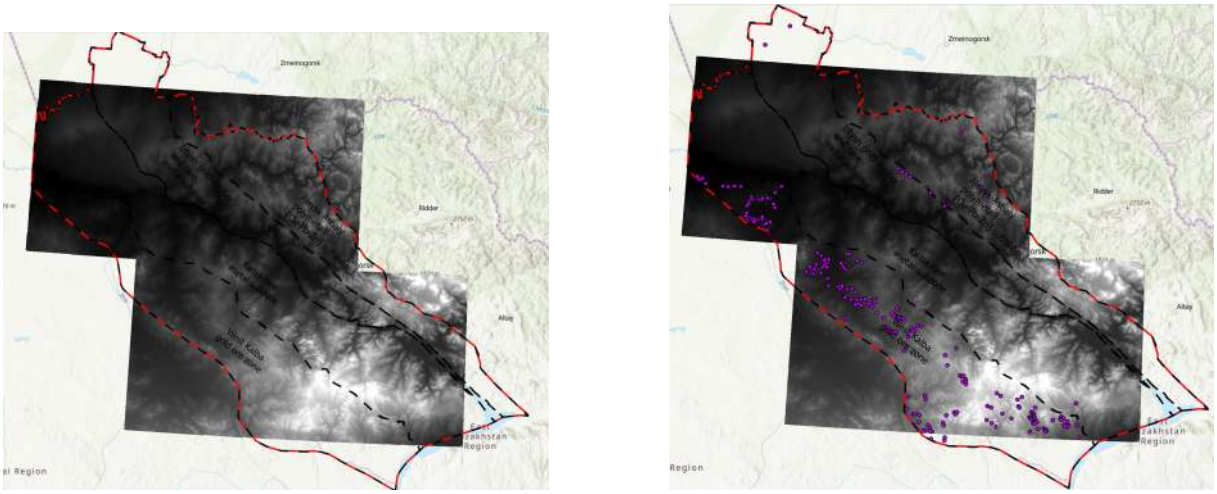
Figure 1. Spatial and geological setting of the study area in Eastern Kazakhstan

For spatial analysis, the WGS 84 / UTM Zone 45N projection (EPSG:32645) was employed, providing metric accuracy essential for geospatial modeling and integration with remote-sensing data. The study area lies within longitudes 72°–78°E, where this projection ensures minimal distortion.

Regional mineralization is structurally controlled by the Irtysh shear zone and the Kalba-Narym fault system, accompanied by NW-SE and NE-SW secondary faults. Mineralization is spatially linked to Late Paleozoic – Early Mesozoic intrusives and to specific lithostratigraphic units, notably Devonian-Carboniferous volcanic-sedimentary sequences and Ordovician-Silurian metamorphic basement.

3. Data for Machine Learning

The task is formulated as a binary classification problem. The feature set comprises point coordinates (Latitude, Longitude), terrain elevation (ELEVATION), and spectral indices derived from satellite imagery. Elevation values, extracted from a 30-meter resolution digital elevation model (SRTM), characterize the geomorphological profile of the area, including slope, aspect, and local relief.



(A) Satellite imagery of mountainous terrain (B) Digital elevation model (hillshade)

Figure 2. Input data sources for the study

The following table presents a subset of the point dataset used for modeling, including coordinates and corresponding elevation values.

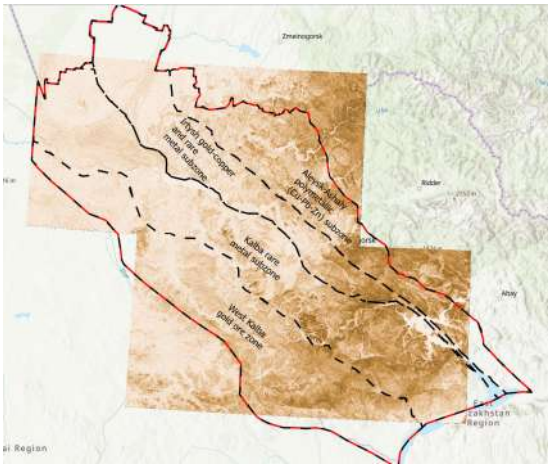
Table 1. Sampling point coordinates and elevation values

№	SHAPE	Latitude	Longitude	ELEVATION (m)
1	Point Z	51.09121	80.59377	0
2	Point Z	51.206	80.8016	0
3	Point Z	50.45553	81.83131	415
4	Point Z	50.4393	81.86952	444
⋮	⋮	⋮	⋮	⋮
224	Point Z	49.92259	81.56034	459

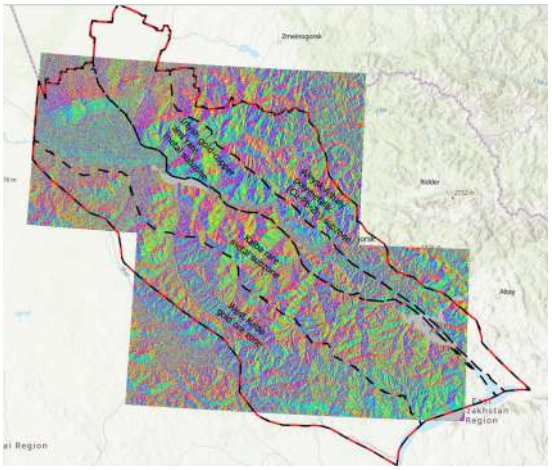
Data specifications:

- **Coordinate system:** WGS 84 / UTM Zone 45N (EPSG:32645)
- **Elevation source:** SRTM DEM (30 m resolution)
- **Satellite data:** Sentinel-2 multispectral imagery (10–20 m resolution)
- **Total sample size:** 224 points (112 mineralized, 112 background)
- **Spectral indices:** NDVI, NDWI, and iron-oxide ratios were calculated for each point

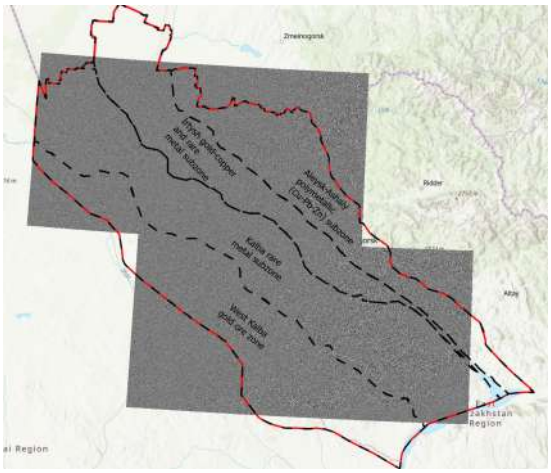
Additional figures used to obtain data that will help improve the quality of the model:



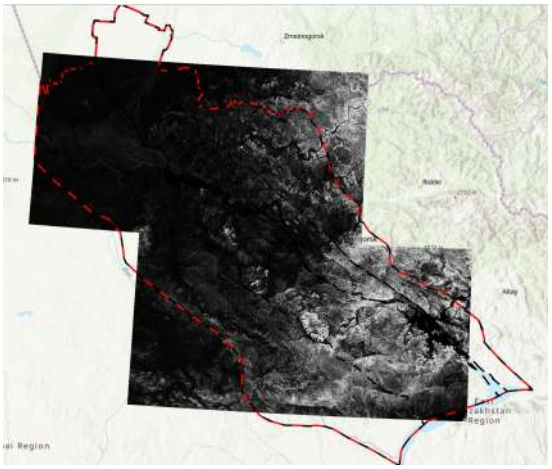
(A) Slope steepness



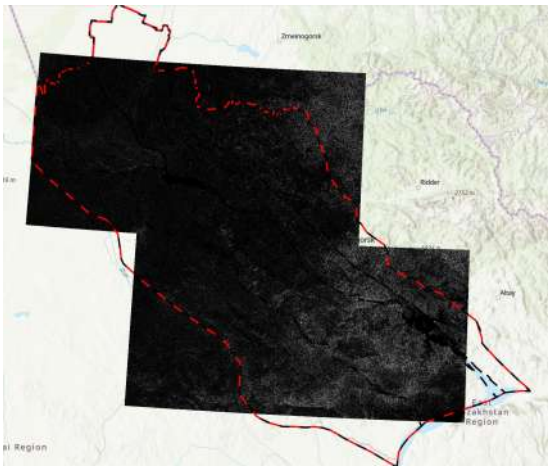
(B) Slope aspect



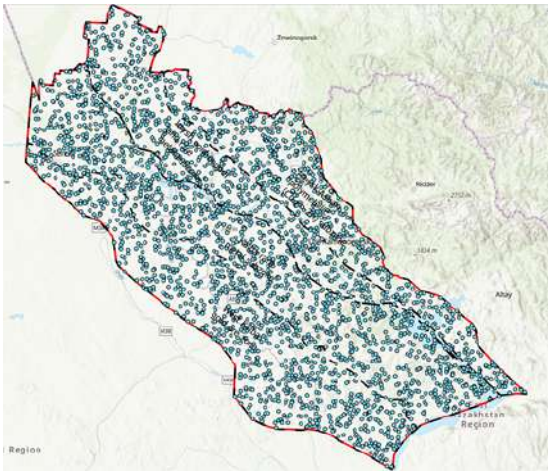
(C) General curvature



(D) Smoothed slope



(E) Locally averaged slope



(F) Background layer

FIGURE 3. Derived geomorphometric features from the DEM using ArcGIS Pro tools.

4. Core Concept of the Model

4.1. **Problem Statement.** Each point i is represented by a feature vector

$$\mathbf{x}_i = (x_{i1}, \ldots, x_{ip}) \in \mathbb{R}^p,$$

containing geomorphometric parameters (elevation, slope, aspect, curvature, etc.) derived in ArcGIS Pro.

The target variable is binary:

$$y_i = \begin{cases} 1 & \text{(gold ore deposit)}, \\ 0 & \text{(background area)}. \end{cases}$$

Thus, we solve $\mathbf{x}_i \longrightarrow y_i \in \{0, 1\}$.

4.2. **Probabilistic Model.** We estimate the conditional probability

$$P(y = 1 \mid \mathbf{x}),$$

interpreted as the probability of gold mineralization given geomorphological conditions.

4.3. **Random Forest Method.** Random Forest is an ensemble of N decision trees

$$\{T_1, \ldots, T_N\},$$

each trained on a bootstrap sample with random feature subsets.

4.4. **Decision Tree.** Each tree T_k provides a piecewise-constant approximation:

$$T_k(\mathbf{x}) = p_{k,j}, \quad \mathbf{x} \in R_{k,j},$$

where $R_{k,j}$ is a region and $p_{k,j}$ is the proportion of class 1 in that leaf.

4.5. **Splitting Criterion.** The Gini impurity $G = 1 - \sum_{c \in \{0,1\}} p_c^2$ is minimized by selecting splits that maximize

$$\Delta G = G_{\text{parent}} - \left(\frac{n_L}{n} G_L + \frac{n_R}{n} G_R \right),$$

with n_L, n_R denoting sample sizes in left/right child nodes.

4.6. **Probability Estimation.** The final gold-mineralization probability is obtained by averaging over all trees:

$$\hat{P}(y = 1 \mid \mathbf{x}) = \frac{1}{N} \sum_{k=1}^{N} T_k(\mathbf{x}).$$

4.7. **Classification Rule.** A binary decision is obtained via a threshold τ (usually $\tau = 0.5$):

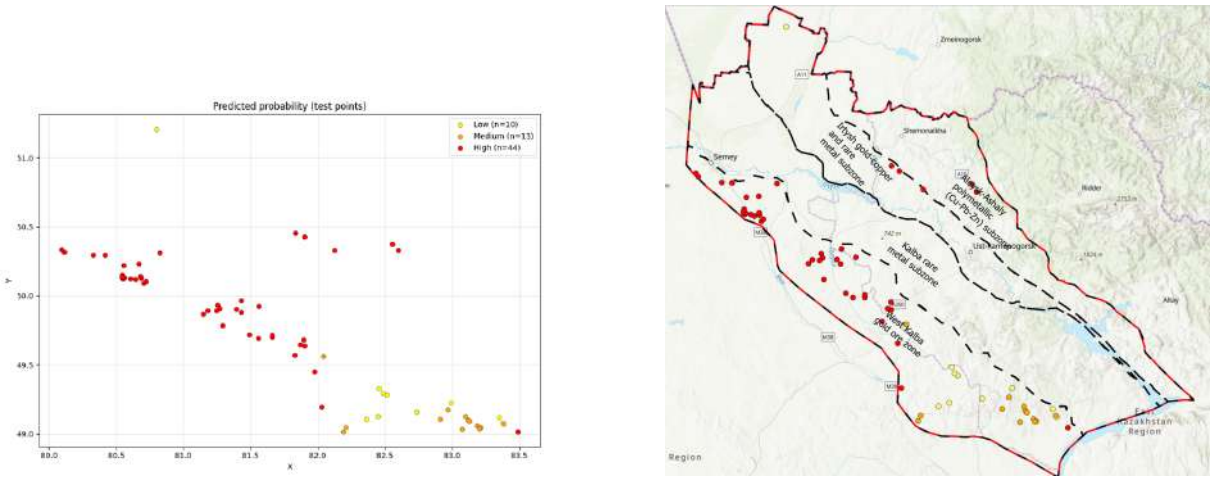
$$\hat{y} = \begin{cases} 1 & \text{if } \hat{P}(y = 1 \mid \mathbf{x}) \geq \tau, \\ 0 & \text{otherwise}. \end{cases}$$

In practice, multi-level classification (low/medium/high probability) is also used.

4.8. **Spatial Interpretation.** The model defines a mapping $f : \mathbb{R}^p \to [0, 1]$, which, after georeferencing, produces a spatial probability field used to delineate prospective areas for geological exploration.

5. Results: Probability of gold mineralization

The presented results were obtained using the Random Forest model and reflect the spatial distribution of gold mineralization probability based on morphometric terrain characteristics.



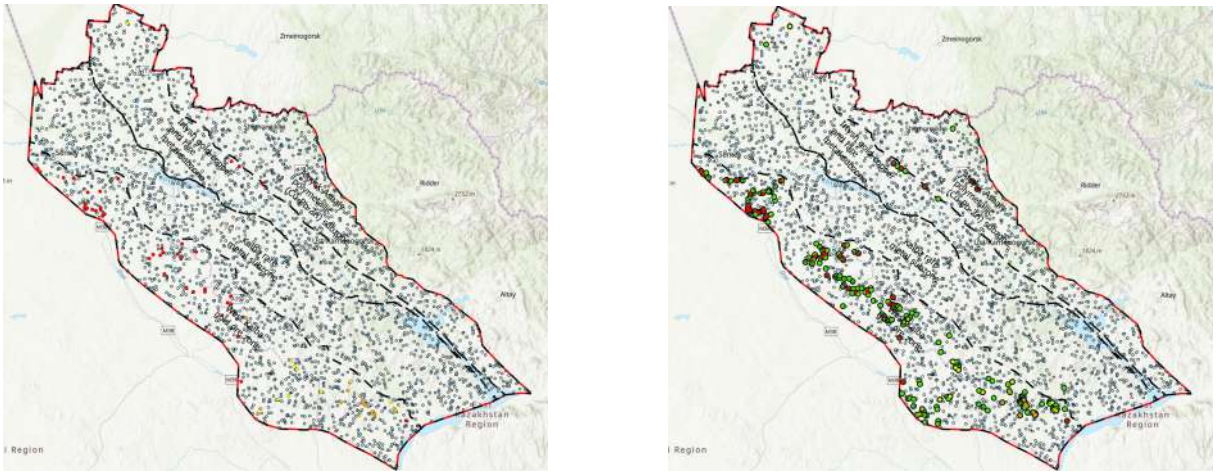
(A) Predicted probability map

(B) Model validation results

Figure 4. Output of the Random Forest classification model.

The Random Forest model demonstrates robust classification performance with an accuracy of **84%** and a ROC-AUC of ≈**0.86**. High precision and F1-score values for the gold mineralization class indicate reliable predictions with a moderate loss in recall. Following the established recommendations, three types of areas were delineated:

(1) Zones with existing gold deposits,
(2) Zones without identified deposits,
(3) Prospective zones predicted by the machine learning model.



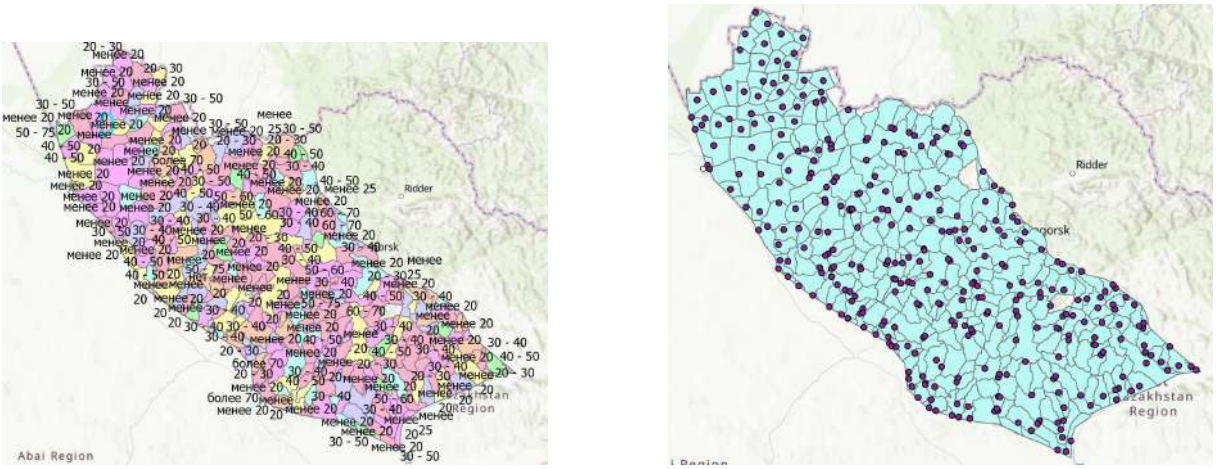
(A) Spatial overlay of predictions and known deposits

(B) Delineated prospective zones for exploration

Figure 5. Spatial interpretation of model predictions.

6. Future Work for 2026

In the forthcoming work of 2026, the expansion and refinement of the model are planned through the integration of radioactivity data and geochemical indicators. A joint spatial analysis of radionuclide distribution, geochemical anomalies, and morphometric terrain characteristics using machine learning methods is anticipated. This approach will enable a shift from point-based prediction to the delineation of prospective zones and will enhance the reliability of forecasting gold-mineralized areas.



(a) Spatial distribution of uranium (U) and thorium (Th) anomalies

(b) Integrated radioactive and geochemical layer with known deposit locations

Figure 6. Planned additional data layers for model enhancement.

Key enhancements planned for 2026:

- **Radioactivity data:** Incorporation of aerial gamma-ray spectrometry measurements (K, U, Th channels) to identify hydrothermal alteration zones associated with mineralization.
- **Geochemical sampling:** Systematic soil and rock chip sampling for multi-element analysis (Au, As, Sb, Cu, Pb, Zn) across the study area.
- **Integrated modeling:** Fusion of radiometric, geochemical, and morphometric data layers within a machine learning framework to improve spatial targeting accuracy.
- **Validation:** Field verification of high-probability zones identified by the enhanced model through geological mapping and sampling.

This multi-parameter approach is expected to reduce exploration uncertainty by providing complementary evidence from independent data sources, potentially revealing mineralization patterns not detectable through single-data-type analysis.

Conclusion

This study successfully demonstrated the integration of ArcGIS Pro for advanced geospatial feature engineering with machine learning for gold mineralization prediction. The developed framework, utilizing a Random Forest classifier on DEM-derived features, achieved an accuracy of 84% and identified prospective zones that align with known geological controls. The workflow provides a reproducible and scalable approach for mineral prospectivity mapping. Future integration of geochemical and radiometric data promises to further enhance model robustness and exploration targeting capabilities in complex geological terrains.

References

[1] Burrough, P. A., McDonnell, R. A., *Principles of Geographical Information Systems*, Oxford University Press, 1998.

[2] Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., et al., *Scikit-learn: Machine Learning in Python*, Journal of Machine Learning Research, **12**, 2011, 2825–2830.

[3] Cracknell, M. J., Reading, A. M., *Geological mapping using remote sensing data and machine learning for mineral exploration*, Computers & Geosciences, Vol. 63, 2014, pp. 22–34.

PREDICTIVE MODELING OF SUBSURFACE MINERALIZATION BASED ON GEOSPATIAL DATA AND ARTIFICIAL INTELLIGENCE METHODS

LAURA TEMIRBEKOVA [1], BAKYTZHAN OMIRZHANOVA [2], ELSHAT TOKTARBEK [3], BEKARYS OTEBAK [1]

[1]Abai Kazakh National Pedagogical University, Almaty, Kazakhstan,
[2]National Engineering Academy of the Republic of Kazakhstan,Almaty, Kazakhstan
[3]Eurasian Bank, Almaty, Kazakhstan

e-mail: laura.tem.2020@gmail.com

This study presents an integrated workflow for predictive modeling of subsurface characteristics based on multicomponent geospatial data and artificial intelligence methods. In the modern context of declining fixed exploration budgets, increasingly complex targets, and rapidly growing volumes of multiparametric datasets, effective management and integration of existing data are essential for mineral exploration operations. Machine learning algorithms such as Random Forest (RF) and Gaussian Process Regression (GPR) were applied. The results indicate that the GPR model provides the most optimal performance for mapping lithium (Li) potential in the study area.

Keywords: Geospatial data, Subsurface modeling, Machine learning, ArcGIS, Remote sensing, Mineralization mapping.

AMS Subject Classification: 62M30, 68T01, 68T05, 86A60, 35R30.

Introduction. Mineral prospectivity mapping (MPM) or modeling is a method for identifying and ranking areas where undiscovered mineral deposits of a specific type are likely to occur. This is achieved by establishing correlations between geological features (input variables) and the occurrence of target mineral deposits (output variables) [1]. Geospatial information is often collected through costly and labor-intensive field surveys, which are limited by difficult accessibility of territories and a small number of samples. At the same time, satellite remote sensing has proven to be effective in mineral deposit exploration and mineral association mapping by detecting spectral anomalies [2–3]. Over recent decades, data-driven modeling approaches have significantly advanced the development of MPM. Alongside probabilistic methods such as weights of evidence and logistic regression, increasing attention has been paid to machine learning algorithms that effectively address multidimensional classification problems and reveal complex nonlinear relationships associated with mineralization [2,4–5]. In recent years, deep learning has been actively developed and applied to various geoscience problems, including geochemical mapping and mineral deposit prediction [4,6]. Nevertheless, no universal algorithm exists for MPM, which necessitates a comparative analysis of multiple predictive methods.

1. Random Forest (RF)

Breiman (2001) developed the Random Forest (RF) machine learning algorithm, which performs repeated predictions of the same phenomenon by combining multiple decision trees. Typically, the RF method generates a large number of decision trees constructed from several subsets of the original training dataset using the bootstrap aggregating (bagging) technique [7]. Subsequently, at each node of a tree, informative layers (features) are selected and applied to increase diversity and promote the growth of the "forest." In this study, the root node of each decision tree was split into several leaf nodes, after which the optimal split providing the highest accuracy of the resulting trees was selected. Various parameters can be used to evaluate the accuracy of trees in the RF method, the most common of which are the Gini index (GI), information gain ratio, and the chi-square criterion [1]. In

The research is funded by Science Committee of the Ministry of Science and Higher Education of the Republic of Kazakhstan (Grant No. BR27100483 "Development of predictive exploration technologies for identifying ore-prospective areas based on data analysis from the unified subsurface user platform "Minerals.gov.kz" using artificial intelligence and remote sensing methods").

the present study, the Gini index (GI) was employed to compute the purity of leaf nodes relative to root nodes, as expressed by the following equation (1).

$$I_G f = \sum_{i=1}^{n} f_i(1 - f_i) \tag{1}$$

where f_i can be described as the probability of belonging to a class i at the node n and can be calculated using equation (2):

$$f_i = \frac{m_j}{m} \tag{2}$$

where m_j— the number of samples belonging to a class j, m— the total number of samples at the node.

2. Gaussian Process Regression

Gaussian Process Regression was applied to model the spatial distribution of lithium concentration. Unlike the Random Forest method, Gaussian Process Regression provides a smooth estimation of the background geochemical field and allows for a quantitative assessment of prediction uncertainty. The model exhibits a tendency to revert to the mean, emphasizing regional trends rather than local anomalies.

Gaussian Process Regression (GPR) [8] is a probabilistic nonlinear regression method based on a Bayesian framework, which treats the target relationship as a realization of a Gaussian random process. Unlike deterministic machine learning models, GPR enables not only the estimation of predicted values of the studied parameter but also the quantitative characterization of prediction uncertainty. The method is widely used in spatial analysis, geostatistics, and interpolation of geophysical and geochemical data, and is mathematically closely related to kriging techniques.

Let a training dataset be given:

$$D = \{(x_i, y_i)\}_{i=1}^{n} \tag{3}$$

where $x_i \in R^d$- the input feature vector (in this study, spatial coordinates), y_i- the observed value of the geochemical parameter (lithium concentration).

In GPR, it is assumed that the function values are distributed according to a Gaussian process:

$$y(x) \sim \mathcal{GP}(m(x), k(x, x^{'})) \tag{4}$$

Where $m(x)$- the mean function (usually assumed to be zero), $k(x, x^{'})$ - the covariance function (kernel), which defines the spatial correlation between points.

In this study, a composite covariance function is used:

$$k(x, x^{'}) = \sigma_f^2 exp(-\frac{\|x - x^{'}\|^2}{2l^2}) + \sigma_n^2 \delta(x, x^{'}), \tag{5}$$

where:

l - the correlation length parameter (length scale),

σ_f^2 - the signal variance,

σ_n^2- the noise level (white noise),

$\delta(x, x^{'})$- the Kronecker delta function.

The exponential–quadratic kernel (RBF) ensures smoothness of the predicted surface and reflects the assumption of spatial correlation of geochemical values. Training of the GPR model involves optimization of the kernel hyperparameters by maximizing the log marginal likelihood of the observed data. To improve model stability, the input coordinates are pre-scaled using standard normalization.

3. Study Area

The Kalba–Narym ore district, located in the eastern part of the Republic of Kazakhstan, is considered in this study. It is confined to the Kalba–Narym fold belt, which represents one of the key metallogenic structures of Eastern Kazakhstan. The district is characterized by a complex tectonic framework, intensive granitoid magmatism, and widespread rare-metal mineralization. From a geodynamic perspective, the area was formed under Late Paleozoic collisional and post-collisional tectonic conditions, which led to the concentration of rare and non-ferrous metals. Previous expeditionary studies conducted geochemical mapping of unconsolidated deposits within the Rudny Altai, Kalba–Narym, and West Kalba structural–metallogenic zones at a scale of 1:500,000, including the collection of lithochemical samples [9-11]. During field investigations, soil samples were collected from pits to depths of up to 20 cm using an approximate 5 × 5 km sampling grid. Soil and sediment sampling was carried out in accordance with GOST 28168 standards. The locations of sampling points were determined using a GARMIN GPS navigator. In total, 777 soil samples were collected over the surveyed area exceeding 40,000 square kilometers. Figure 1 presents a map of the study area.

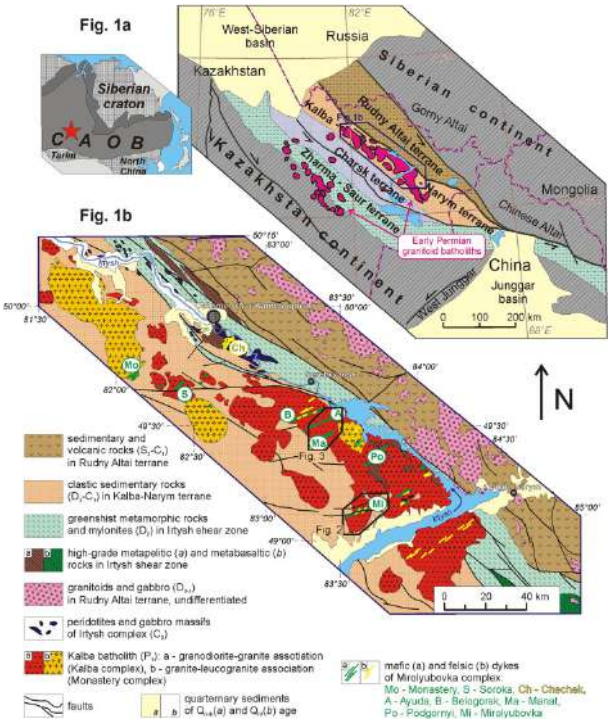


FIGURE 1. Location map of the Kalba, Irtysh region, and Rudny Altai.

4. Prediction Results

In this study, two artificial intelligence methods were implemented to predict the content of the chemical element Li in the subsurface environment and to analyze its spatial distribution. Figure 2 illustrates the spatial distribution of lithium concentration at the Earth's surface based on the described field investigations.

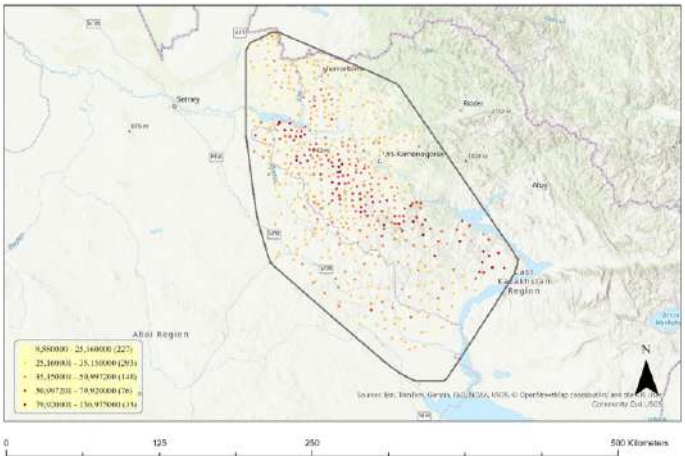


FIGURE 2. Spatial distribution of lithium concentration at the Earth's surface

Figures 3 and 4 present comparative maps of the predicted lithium concentration values obtained using two artificial intelligence methods—Random Forest (RF) and Gaussian Process Regression (GPR). Both approaches were applied to the same set of initial geochemical data and the same predictive feature layers, which allows for a consistent comparison of the results and an analysis of differences in spatial patterns.

The map generated using the Random Forest method is characterized by higher contrast and a wider range of predicted values. The RF model effectively highlights localized areas of elevated concentrations, resulting in a more

"patchy" spatial structure. This pattern reflects the algorithm's ability to capture nonlinear relationships and local data features; however, RF does not explicitly account for spatial autocorrelation and does not provide estimates of prediction uncertainty. As a result, in areas with low density of input observations, zones with potentially overestimated predictions may appear, which require additional verification and geological justification.

In contrast, the GPR results exhibit a smoother spatial field of predicted values and a more stable delineation of regional trends. The GPR model accounts for spatial correlation between sampling points and produces predictions in the form of a mean value accompanied by an uncertainty estimate. Consequently, the resulting surface is generally less sensitive to isolated extreme values and better represents the background component of the geochemical field. This property is particularly important for interpreting results in peripheral areas with sparse observations, where GPR typically reduces the likelihood of artificial anomalies and provides a more conservative assessment.

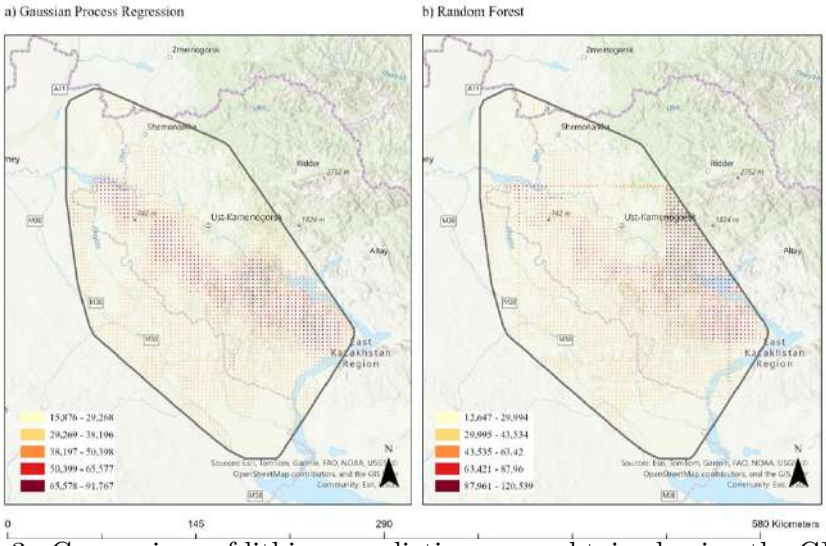


FIGURE 3. Comparison of lithium predictive maps obtained using the GPR and RF methods

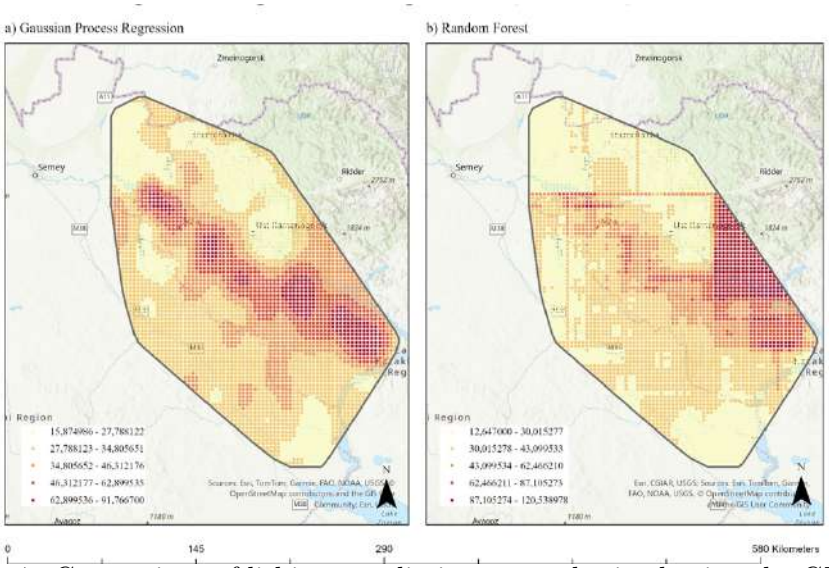


FIGURE 4. Comparison of lithium predictive maps obtained using the GPR and RF methods.

The probabilistic maps constructed using the Gaussian Process Regression model represent the spatial distribution of the probability of exceeding specified threshold values of lithium concentration. The map $P(Li > 50$

ppm) delineates zones of potential lithium enrichment and can be used for regional-scale screening. The map $P(Li > 65$ ppm) highlights areas with a high probability of economically significant lithium concentrations and is therefore of greatest interest for detailed geological exploration. A comparison of these maps indicates that the most prospective zones are localized within areas exhibiting a high probability of exceeding both threshold values.

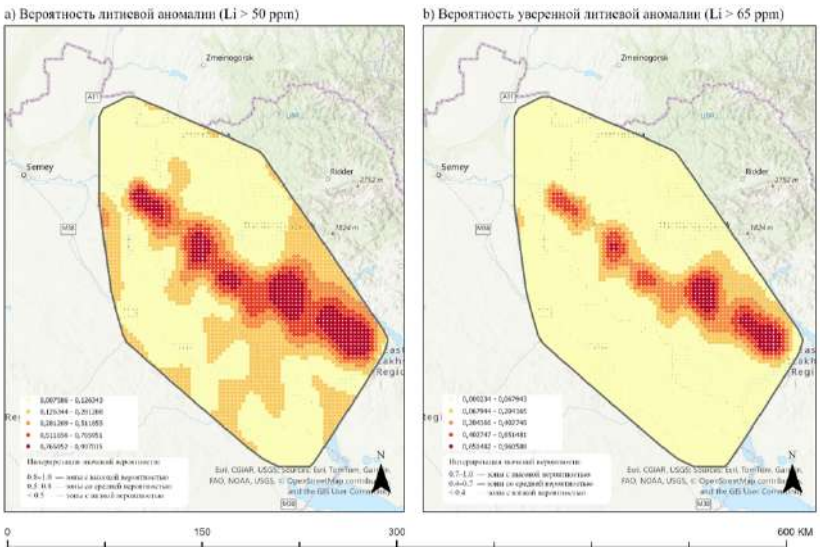


FIGURE 5. Probabilistic lithium prediction maps generated using the GPR method.

Figure 6 presents a comparison of the predicted lithium concentration maps obtained using the GPR and RF methods. Figure 7 shows the spatial distribution of uncertainties in lithium concentration prediction.

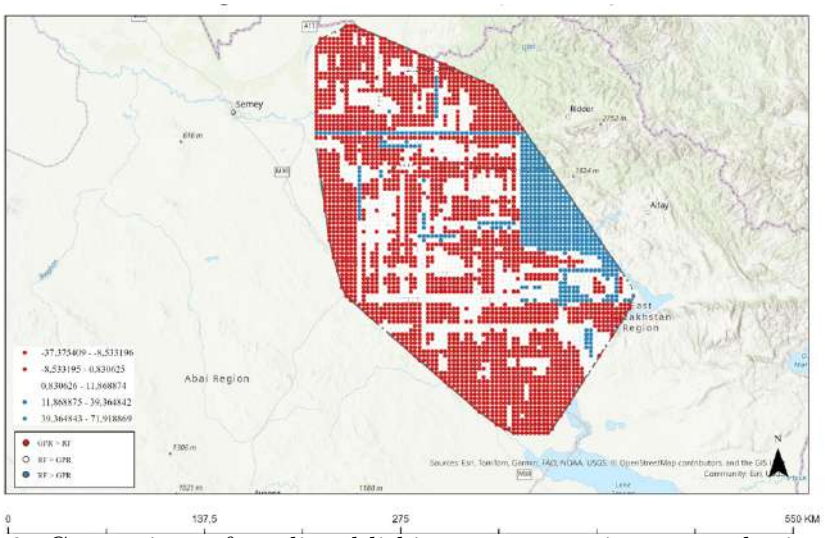


FIGURE 6. Comparison of predicted lithium concentration maps obtained using the GPR and RF methods

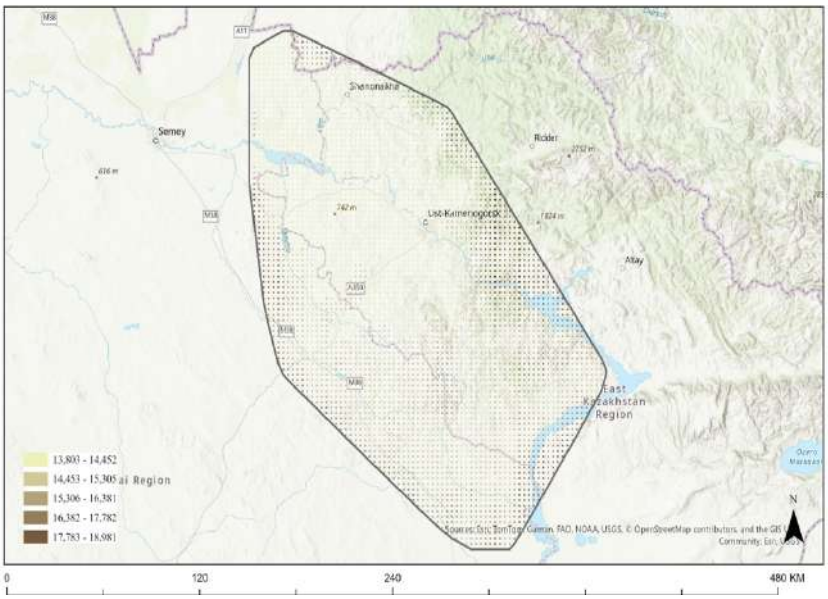


FIGURE 7. Map of lithium concentration prediction uncertainty

Conclusion The comparison of the maps indicates that areas with elevated predicted values identified by both methods partially overlap, confirming the presence of a stable spatial signal in the input data. At the same time, differences in the intensity and boundaries of anomalous zones reflect the differing "sensitivities" of the methods: RF places greater emphasis on local extrema, whereas GPR primarily highlights persistent regional structures. From a practical perspective, this implies that RF results are well suited for the initial identification of potential anomalies, while GPR results are more appropriate for refining spatial trends and assessing prediction reliability. Thus, the combined application of RF and GPR provides a more comprehensive representation of the predicted lithium distribution. RF offers high sensitivity to localized anomalous features, whereas GPR enhances interpretability and reliability by accounting for spatial structure and enabling quantitative uncertainty assessment. The resulting comparative maps form a basis for the further delineation of prospective areas and for developing recommendations for detailed geological exploration.

REFERENCES

[1] Mahboob, M.A., Celik, T., Genc, B., Predictive modelling of mineral prospectivity using satellite remote sensing and machine learning algorithms, *Remote Sensing Applications: Society and Environment*, Vol.36, No.101316, 2024.

[2] Shirmard, H., Farahbakhsh, E., Muller, R.D., Chandra, R., A review of machine learning in processing remote sensing data for mineral exploration, *Remote Sensing of Environment*, Vol.268, No.112750, 2022.

[3] Fu, Y., Cheng, Q., Jing, L., Ye, B., Fu, H. Mineral prospectivity mapping of porphyry copper deposits based on remote sensing imagery and geochemical data in the Duolong ore district, *Remote Sensing, MDPI*, Vol. 15, No.439, 2023.

[4] Mahboob, M. , Celik, T., Genc, B. Review of machine learning-based mineral resource estimation, *Journal of the Southern African Institute of Mining and Metallurgy*, Vol. 122, 2023, pp.655–664.

[5] Jung, D., Choi, Y. Systematic review of machine learning applications in mining: exploration, exploitation, and reclamation. *Minerals*, Vol. 11, 2021, 148.

[6] Kong, Y., Chen, G., Liu, B., Xie, M., Yu, Z., Li, C., Wu, Y., Gao, Y., Zha, S., Zhang, H. 3D mineral prospectivity mapping of the Zaozigou gold deposit, West Qinling, China: machine learning-based mineral prediction. *Minerals*, Vol. 12, 2022, 1361.

[7] Wang, X., Yuan, P., Mao, Z., You, M. Molten steel temperature prediction model based on bootstrap feature subsets ensemble regression trees. *Knowledge-Based Systems*, Vol. 101, 2016, pp.48-59.

[8] Murphy, K.P., *Machine Learning: A Probabilistic Perspective*, MIT Press, Cambridge, MA, USA, 2012, 1098 p.

[9] Temirbekov, N.M., Los, V.L., Imangaliev, E.I., Baigereev, D.R., Temirbekova, L.N., Nurmangalieva, M.B. A geoinformation system module based on numerical modeling of inverse geochemical problems using regularization algorithms. *Bulletin of Abai Kazakh National Pedagogical University. Series "Physical and Mathematical Sciences",* , Vol. 3(75), 2021, pp.15-28.

[10] N. Temirbekov, Ye. Imangaliyev, D. Baigereyev, M. Nurmangaliyeva, L.Temirbekova, D.T. Pham. Numerical simulation of inverse geochemistry problem by regularizing algorithms, *Cogent Engineering 2022-12-31 — Journal article*, 9:2003522, 2022, 21 p. https://doi.org/10.1080/23311916.2021.2003522

[11] Syrym Kasenov, Aigerim Tleulesova, Almas Temirbekov, Zholaman Bektemessov, Rysbike Asanova, *Numerical Solution of the Inverse Thermoacoustics Problem Using QFT and Gradient Method* , Fractal and Fractional, MDPI, Vol.9, No.6, https://doi.org/10.3390/fractalfract9060370

NUMERICAL METHODS FOR SOLVING A GEOCHEMICAL INVERSE PROBLEM BASED ON A FREDHOLM INTEGRAL EQUATION

DINARA TAMABAY[1,2], ARUZHAN SEITOVA[1], AYAUZHAN SEITOVA[1]

[1]Faculty of Mechanics and Mathematics, Al-Farabi Kazakh National University, Almaty, Kazakhstan
[2]National Engineering Academy of the Republic of Kazakhstan, Almaty, Kazakhstan

e-mail: dtamabay@gmail.com, aruzhanseitova55@gmail.com, aiauzhanseitova77@gmail.com

This paper addresses a geochemical inverse problem aimed at reconstructing an unknown subsurface distribution from surface measurement data. The problem is formulated as a Fredholm integral equation of the first kind, which belongs to the class of ill-posed problems and is highly sensitive to measurement noise.

Three numerical approaches are employed to solve the problem: the classical Galerkin method, the Galerkin method combined with Lavrentiev regularization, and a method based on Singular Value Decomposition (SVD). The discretization of the integral equation leads to a poorly conditioned system of linear algebraic equations. To ensure stability, the influence of regularization techniques and singular value truncation is investigated. The performance of the proposed methods is evaluated through numerical experiments with artificially perturbed data, allowing a comparative analysis of their stability and reconstruction accuracy.

The results demonstrate that the classical Galerkin method is highly sensitive to noise, whereas the Lavrentiev-regularized Galerkin method and the SVD-based approach significantly improve the stability of the solution. The accuracy of the reconstructed solutions is assessed and their physical meaning is interpreted within a geochemical context. The proposed methodology proves to be an effective numerical tool for solving geochemical inverse problems governed by the Fredholm integral equations.

Keywords: Fredholm integral equation, inverse problems, Galerkin method, Lavrentiev regularization, singular value decomposition, geochemical modeling

AMS Subject Classification: 45B05, 65R30, 65F22, 86A32.

References

[1] Temirbekova L. N., *Numerical Methods and Geographic Information System for Geochemical and Geophysical Problems: Monograph*, Almaty, 2022.

[2] Tikhonov A. N., Arsenin V. Y., *Solutions of Ill-Posed Problems*, Winston, Washington, 1977.

This work was partially supported by National Engineering Academy of the Republic of Kazakhstan.

1. Introduction

Predicting the subsurface distribution of mineral deposits based on surface measurements is a key challenge in modern geochemistry and mineral exploration. Direct investigation methods, such as drilling, require significant time and financial resources, making the development of mathematical methods for interpreting indirect measurements highly relevant.

In many geochemical tasks, the relationship between surface measurements and subsurface properties is of integral nature. These problems naturally lead to Fredholm integral equations of the first kind. A distinctive feature of such equations is their ill-posedness, manifested as a high sensitivity of the solution to errors in the input data.

Given the presence of noise and incomplete experimental information, classical solution methods often fail without additional stabilization. Therefore, regularized numerical methods play a critical role. The aim of this study is to investigate and compare numerical methods for solving geochemical inverse problems based on Fredholm integral equations of the first kind.

2. Mathematical Formulation of the Geochemical Inverse Problem

Let

$$\Omega \subset \mathbb{R}^n$$

denote the investigated geological domain. Let f(x) represent the measured surface geochemical data, and u(s) denote the unknown function describing the subsurface distribution of the chemical element.

The relationship between these quantities is modeled by the Fredholm integral equation of the first kind:

$$\int_{\Omega} K(x,s)\,u(s)\,ds = f(x), \quad x \in \Omega \tag{1}$$

where K(x,s) is the kernel of the integral operator, reflecting the physical-chemical processes of material migration in the geological medium.

The integral operator K is compact, which leads to the ill-posedness of the problem in the sense of Hadamard. Small perturbations in f(x), which are inevitable in experimental measurements, can result in large deviations in the solution u(s).

3. Discretization of the Integral Equation

For numerical implementation, the integral equation is discretized. Let the domain Ω be divided into nodes $\{s_j\}_{j=1}^{N}$ with corresponding quadrature weights w_j. Then, the integral can be approximated as:

$$\int_{\Omega} K(x_i,s)\,u(s)\,ds \approx \sum_{j=1}^{N} K(x_i,s_j)\,u(s_j)\,w_j, \quad i = 1,\ldots,M \tag{2}$$

As a result, we obtain a system of linear algebraic equations:

$$\mathbf{K}\,\mathbf{u} = \mathbf{f} \tag{3}$$

where

$\mathbf{K} \in \mathbb{R}^{M\times N}$ — discretized integral operator,

$\mathbf{u} \in \mathbb{R}^{N}$ — vector of unknown concentrations,

$\mathbf{f} \in \mathbb{R}^{M}$ — vector of observed data.

The resulting system is ill-conditioned, reflecting the ill-posedness of the original geochemical problem.

4. Lavrentiev Regularization

To stabilize the solution, Lavrentiev regularization is applied by minimizing the functional:

$$J(\mathbf{u}) = \|\mathbf{K}\mathbf{u} - \mathbf{f}\|_2^2 + \alpha\|\mathbf{u}\|_2^2, \tag{4}$$

where $\alpha > 0$ is the regularization parameter. Unlike Tikhonov regularization, this method does not use the operator $\mathbf{L}$ and directly smooths the solution.

Minimizing this functional leads to the normal equations:

$$(\mathbf{K}^T\mathbf{K} + \alpha\mathbf{I})\mathbf{u} = \mathbf{K}^T\mathbf{f}. \tag{5}$$

where I is the identity matrix. The regularization term suppresses the influence of noisy components, ensuring solution stability. Geochemically, this corresponds to assuming smooth variations in subsurface concentration.

5. Choice of Optimal Regularization Parameter (α)

The quality of the Lavrentiev regularized solution heavily depends on the choice of the parameter $\alpha > 0$. A common and effective method for selecting α is the **L-curve criterion**.

The L-curve plots the norm of the regularized solution $\|\mathbf{L}\mathbf{u}_\alpha\|$ versus the residual norm $\|\mathbf{K}\mathbf{u}_\alpha - \mathbf{f}\|$ in log-log scale:

$$(\|\mathbf{K}\mathbf{u}_\alpha - \mathbf{f}\|_2, \|\mathbf{u}_\alpha\|_2), \quad \alpha > 0 \tag{6}$$

The curve typically has an "L" shape. The **corner** of the L-curve corresponds to a balanced choice of α, providing a compromise between:

- **Fidelity to the data** — small residual norm, and
- **Stability/smoothness** — small solution norm.

Mathematically, the corner can be found by maximizing the curvature $\kappa(\alpha)$ of the L-curve:

$$\kappa(\alpha) = \frac{|\rho'(\alpha)\eta''(\alpha) - \rho''(\alpha)\eta'(\alpha)|}{(\rho'(\alpha)^2 + \eta'(\alpha)^2)^{3/2}}, \tag{7}$$

where $\rho(\alpha) = \|\mathbf{K}\mathbf{u}_\alpha - \mathbf{f}\|_2$ and $\eta(\alpha) = \|\mathbf{I}\mathbf{u}_\alpha\|_2$.

This method is widely used in geophysical and geochemical inverse problems to select an optimal regularization parameter automatically.

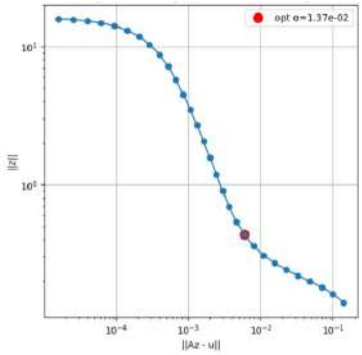


FIGURE 1. L-curve example

6. Galerkin Method

The Galerkin method approximates the unknown function in a finite-dimensional basis:

$$u(s) \approx \sum_{k=1}^{n} c_k \varphi_k(s), \tag{8}$$

where $\{\varphi_k\}_{k=1}^n$ is a chosen system of basis functions.

Substituting this approximation into the integral equation and requiring orthogonality of the residual yields the system:

$$\sum_{k=1}^{n} c_k \int_\Omega \int_\Omega K(x,s)\,\varphi_k(s)\,\varphi_i(x)\,ds\,dx = \int_\Omega f(x)\,\varphi_i(x)\,dx, \quad i = 1, \ldots, n \tag{9}$$

The Galerkin method incorporates a priori information about the solution's structure and provides good interpretability, which is important for analyzing geochemical data.

7. Singular Value Decomposition (SVD)

After discretization, the problem reduces to solving

$$\mathbf{K}\mathbf{u} = \mathbf{f}, \tag{10}$$

where $\mathbf{K}$ is ill-conditioned. Small errors in $\mathbf{f}$ can significantly distort $\mathbf{u}$.

To analyze operator properties and construct a stable solution, Singular Value Decomposition (SVD) is applied:

$$\mathbf{K} = \mathbf{U}\,\mathbf{\Sigma}\,\mathbf{V}^T, \tag{11}$$

where:

- $\mathbf{U}$ contains orthonormal left singular vectors, associated with observed data;
- $\mathbf{V}$ contains orthonormal right singular vectors, associated with the subsurface solution;
- $\mathbf{\Sigma} = \mathrm{diag}(\sigma_1, \ldots, \sigma_r)$ contains singular values ordered as $\sigma_1 \geq \sigma_2 \geq \cdots \geq \sigma_r > 0$.

Large singular values correspond to stable solution components, whereas small ones correspond to noise-sensitive components.

The formal SVD solution is

$$\mathbf{u} = \sum_{i=1}^{r} \frac{\mathbf{u}_i^T \mathbf{f}}{\sigma_i}\,\mathbf{v}_i. \tag{12}$$

which is unstable for small σ_i. Stability is improved by truncated SVD:

$$\mathbf{u}_{r_0} = \sum_{i=1}^{r_0} \frac{\mathbf{u}_i^T \mathbf{f}}{\sigma_i}\,\mathbf{v}_i, \tag{13}$$

where $r_0 \leq r$ is the truncation index chosen to remove noise-sensitive components corresponding to small singular values.

8. Numerical Experiments

Numerical experiments were conducted on geochemical models predicting gold concentrations at depths of 100, 200, 300 meters, and 400 meters, based on surface data.

Findings:

Lavrentiev regularization produces the most stable and smooth solution;

The Galerkin method yields an interpretable subsurface structure;

SVD allows control over spectral influence of noise and evaluates the informative content of solution components.

All methods proved effective in the presence of noisy data typical of field geochemical measurements.

9. Data and Results Visualization

The geochemical data used in the study include surface measurements of gold concentrations at different locations. The following figures illustrate the raw data and the reconstructed subsurface distributions obtained using the numerical methods described above.

	s	y	z_depth100
0	307500	5.4635e+06	0.0405427
1	307550	5.4635e+06	-0.00425739
2	307600	5.4635e+06	0.00953923
3	307650	5.4635e+06	0.00274169
4	307700	5.4635e+06	0.00573197
5	307750	5.4635e+06	0.00427825
6	307800	5.4635e+06	0.00486882
7	307850	5.4635e+06	0.00454624
8	307900	5.4635e+06	0.00464886
9	307950	5.4635e+06	0.00456601

	s	y	z_depth100
6451	310550	5.468e+06	0.00678059
6452	310600	5.468e+06	0.00662791
6453	310650	5.468e+06	0.00918708
6454	310700	5.468e+06	0.00540721
6455	310750	5.468e+06	0.013628
6456	310800	5.468e+06	0.00712232
6457	310850	5.468e+06	0.00669029
6458	310900	5.468e+06	0.0165885
6459	310950	5.468e+06	-0.0073657
6460	311000	5.468e+06	0.0732627

FIGURE 2. numerical values of the calculated data (100m)

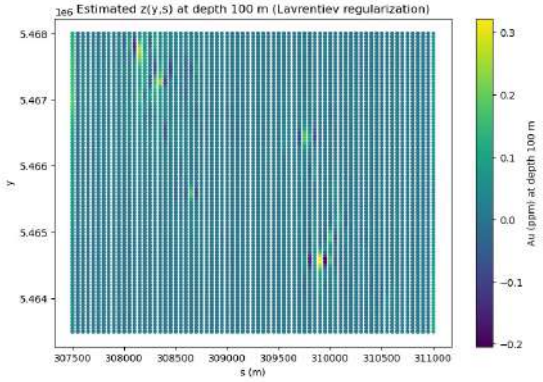


FIGURE 3. Reconstructed subsurface distribution using Lavrentiev regularization (100 m).

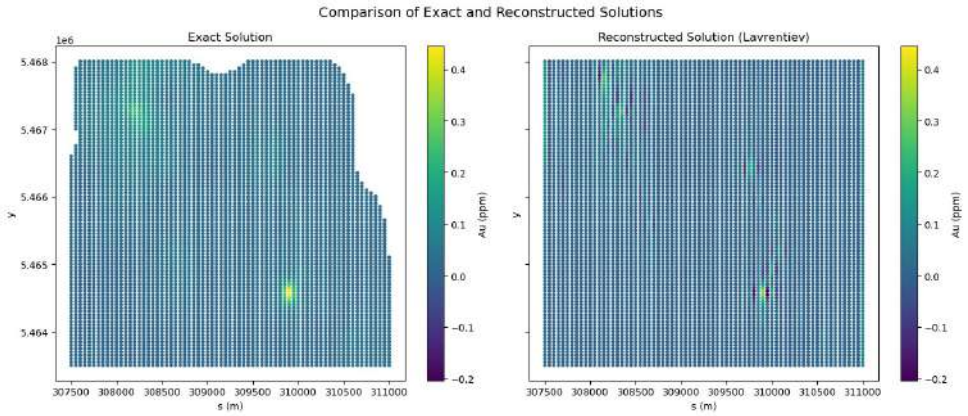


FIGURE 4. Comparison of Exact and Reconstructed Solutions (100m)

10.Comparative analysis In this study, the distribution of gold at various depths was investigated based on geochemical data obtained from the surface. The problem was formulated as an inverse problem, with the aim of reconstructing the subsurface distribution of gold using surface measurements.

During the study, for example, the gold distributions at depths of 100 m and 200 m were determined and comparatively analyzed. In addition, the data at a depth of 100 m were conditionally treated as surface data and used to recompute the gold distribution at a depth of 200 m. The differences between the obtained results were analyzed to assess depth-dependent variations. This approach helps to identify whether certain regular patterns exist in the distribution of gold.

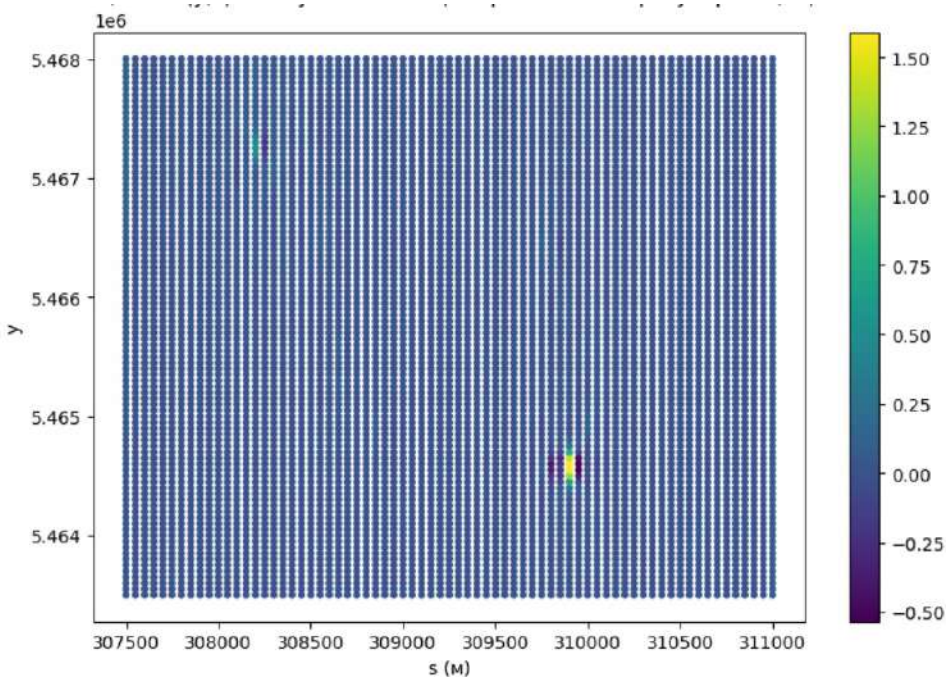


FIGURE 5. Lavrentiev 200m ($h_0 = 100$m)

11. Conclusion

This study investigated numerical methods for solving geochemical inverse problems based on Fredholm integral equations of the first kind. The methods considered include:

Lavrentiev regularization;

Galerkin method;

Singular value decomposition (SVD).

It was shown that their combined use yields physically meaningful, stable solutions that adequately reflect subsurface mineral distributions. The proposed approaches are applicable for predicting mineral concentrations based on indirect surface measurements.

ROBUSTNESS OF FBCSP AGAINST VISUAL ALPHA MASKING: A RELIABLE BASELINE WITH NESTED CROSS-VALIDATION

AIGERIM AITIM[1], ZHULDYZ BAGYBEK[1]

[1]International Information Technology University, Almaty, Kazakhstan
e-mail: bagybekzuldyz@gmail.com, a.aitim@iitu.edu.kz

Motor imagery (MI) based brain–computer interfaces (BCIs) rely on the reliable interpretation of electroencephalography (EEG) signals, which is particularly important in rehabilitation and assistive applications. In practice, however, MI decoding performance is often affected by pronounced inter-subject variability and by differences in experimental and evaluation protocols. Filter Bank Common Spatial Patterns (FBCSP) remains one of the most commonly used baseline methods in this domain, yet its reported performance varies considerably across studies due to inconsistent preprocessing, feature selection strategies, and validation procedures.

In this study, an adjusted FBCSP-based workflow is considered with the goal of reducing the influence of common methodological biases. Particular attention is paid to the organization of the evaluation procedure. Nested stratified cross-validation is used to avoid information leakage and to preserve class balance across folds. In addition, Independent Component Analysis (ICA) is applied to identify and remove components associated with ocular and muscle artifacts, using their spatial distribution and temporal characteristics. When tested on the training session of the BCI Competition IV-2a dataset, the proposed pipeline achieves an average within-session accuracy of 81.8% with a Cohen's Kappa value of $\kappa = 0.757$.

In addition to overall performance measures, the behavior of the classifier is examined in subjects with comparatively low decoding accuracy. The analysis suggests that in some cases motor imagery–related activity is present but masked by strong background rhythms originating from the occipital region. This effect, referred to as Visual Alpha Masking, indicates that dominant visual alpha activity can obscure sensorimotor signals at the scalp level. The observations support the use of spatial filtering techniques, which make it possible to separate weaker motor-related patterns from more prominent background activity.

Keywords: Brain–Computer Interface, Motor Imagery, Electroencephalography, Filter Bank Common Spatial Patterns, Nested Stratified Cross-Validation, Independent Component Analysis

AMS Subject Classification: 92C55

1. Introduction

Brain–computer interfaces (BCIs) are designed to translate patterns of brain activity into control commands for external devices. One of the most common approaches in non-invasive BCI research is motor imagery (MI), which is based on changes in sensorimotor rhythms during imagined movements. Owing to its physiological relevance and absence of external stimulation, MI has been widely investigated in rehabilitation-oriented applications, particularly for patients with motor impairments [1]. At the same time, MI-based systems are known to be sensitive to variability in EEG signals and to experimental conditions.

A large number of studies report high classification accuracy for MI decoding; however, comparable results are often difficult to reproduce. Even when the same datasets and algorithms are used, performance estimates may differ substantially across publications. This inconsistency is frequently related to the organization of data processing and evaluation rather than to the

choice of the classifier itself. In particular, information leakage may occur if feature selection is performed before data partitioning. In addition, insufficient suppression of ocular and muscular artifacts can lead to classifiers that rely on non-neural activity. Validation schemes that do not control class balance or trial structure further contribute to optimistic performance estimates that cannot be generalized [2].

An additional challenge is the so-called phenomenon of BCI illiteracy, where a considerable proportion of users fail to achieve reliable control of MI-based systems. It is commonly reported that approximately 15–30% of participants do not exceed the accuracy threshold required for effective communication [3]. The underlying reasons for this limitation are still debated and are frequently described in terms of poor signal quality or weak sensorimotor modulation, without a detailed analysis of the contributing factors.

In this study, we focus on establishing a transparent and reproducible baseline for MI classification using the Filter Bank Common Spatial Patterns (FBCSP) method. A nested stratified cross-validation scheme is employed to prevent information leakage and to ensure fair performance evaluation. Furthermore, artifact suppression is performed using automated Independent Component Analysis. The analysis is conducted in a within-session setting in order to estimate the upper bound of decoding performance under controlled conditions. Beyond reporting classification accuracy, we investigate the mechanisms behind performance degradation in low-performing subjects and introduce the concept of *Visual Alpha Masking*, which describes the spatial dominance of occipital alpha activity over weaker sensorimotor patterns.

The main contribution of this study is threefold. First, a leakage-free FBCSP pipeline with nested stratified cross-validation is established as a reliable baseline for within-session motor imagery decoding. Second, a systematic stability analysis demonstrates that the proposed framework is robust to feature dimensionality. Third, a detailed case study introduces the concept of *Visual Alpha Masking*, providing a neurophysiological explanation for successful decoding in subjects traditionally considered BCI-illiterate.

2. Methodological Framework

The experiments were carried out using the BCI Competition IV-2a dataset [4], which is widely employed as a reference benchmark in motor imagery studies. The dataset consists of EEG recordings from nine healthy subjects (A01–A09) performing four motor imagery tasks: left hand, right hand, feet, and tongue. Signals were acquired from 22 Ag/AgCl electrodes positioned according to the international 10–20 system and sampled at 250 Hz. To reduce the influence of session-to-session non-stationarity and to focus on the intrinsic properties of the decoding pipeline, only the training session was considered in this study. Each subject contributed 288 trials in total.

Prior to feature extraction, EEG data were subjected to a sequence of preprocessing steps aimed at improving the signal-to-noise ratio while preserving task-related temporal information. First, the signals were filtered using a zero-phase finite impulse response (FIR) band-pass filter in the range of 8–28 Hz, covering the sensorimotor μ and β rhythms. A linear-phase FIR design was selected to avoid temporal distortions of event-related activity. The filter coefficients were obtained using a Hamming window to reduce spectral leakage.

After temporal filtering, the signals were re-referenced using the common average reference (CAR). This operation suppresses spatially global activity by subtracting the mean signal across all channels and enhances local cortical sources associated with motor imagery. To further reduce the influence of non-neural artifacts, Independent Component Analysis (ICA) was applied to separate neural activity from physiological contamination. The FastICA algorithm [5] was used to decompose the EEG into statistically independent components. Components were excluded when their scalp topographies exhibited pronounced frontal dominance, indicative of ocular artifacts, or when their temporal activations showed high kurtosis, consistent with transient muscle-related activity.

Feature extraction was performed using the Filter Bank Common Spatial Patterns (FBCSP) method. The EEG signals were decomposed into eight overlapping frequency bands (4–8 Hz, 8–12 Hz, ..., 28–32 Hz). For each sub-band, class-specific covariance matrices were estimated and spatial filters were obtained by solving a generalized eigenvalue problem that maximizes the variance ratio between classes,

$$J(w) = \frac{w^T \Sigma_1 w}{w^T (\Sigma_1 + \Sigma_2) w}, \tag{1}$$

where Σ_1 and Σ_2 denote the covariance matrices of the two motor imagery classes. The resulting spatial filters were applied to the EEG signals, and trial-wise variance values were computed from the projected data. These values were then transformed using a logarithmic scale to stabilize their distribution and facilitate subsequent linear classification.

Classification performance was assessed using a nested stratified cross-validation strategy. The outer cross-validation loop consisted of 10 folds and served exclusively for the final evaluation on held-out data. Within each outer training split, an additional 5-fold stratified cross-validation was employed to rank and select features based on mutual information. This strict separation between feature selection, model optimization, and evaluation avoids information leakage and provides a realistic estimate of within-session generalization performance.

3. Results

The effect of preprocessing was first assessed to verify that artifact suppression did not interfere with neural activity related to motor imagery. Particular attention was paid to the behavior of the ICA decomposition, as improper component rejection may introduce non-neural bias into subsequent analysis.

Figure 1 illustrates representative ICA component topographies. A component with pronounced frontal dominance was identified as ocular activity and removed, whereas a retained component exhibited a more central distribution consistent with sensorimotor sources: (a) excluded component with frontal dominance (ocular artifact), (b) retained component with a more central distribution. These examples indicate that the adopted ICA rejection criteria primarily target non-neural activity while preserving task-relevant spatial structure.

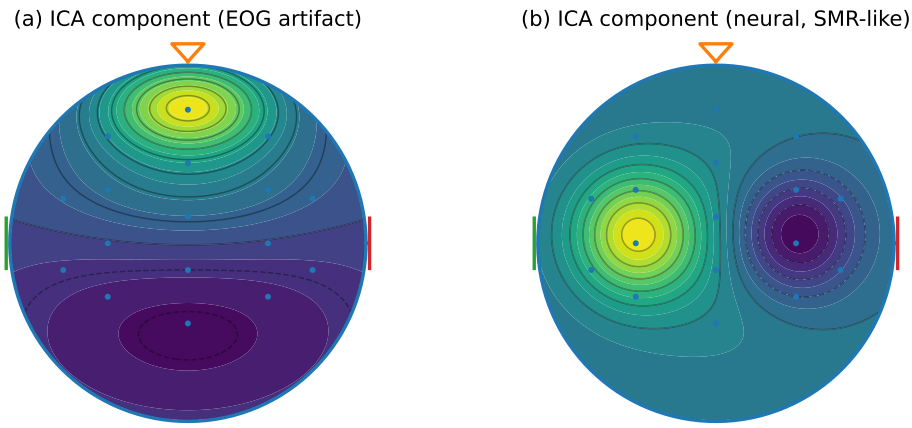


Figure 1. ICA-based artifact rejection examples

The stability of the extracted FBCSP representation was examined by varying the number of selected features k from 16 to 128. This analysis evaluates whether classification performance is driven by a narrow feature subset or remains stable under feature-count changes.

As shown in Fig. 2, classification accuracy remains stable across the tested range of k values, suggesting that discriminative information is distributed across multiple spatial–spectral features rather than concentrated in a single peak.

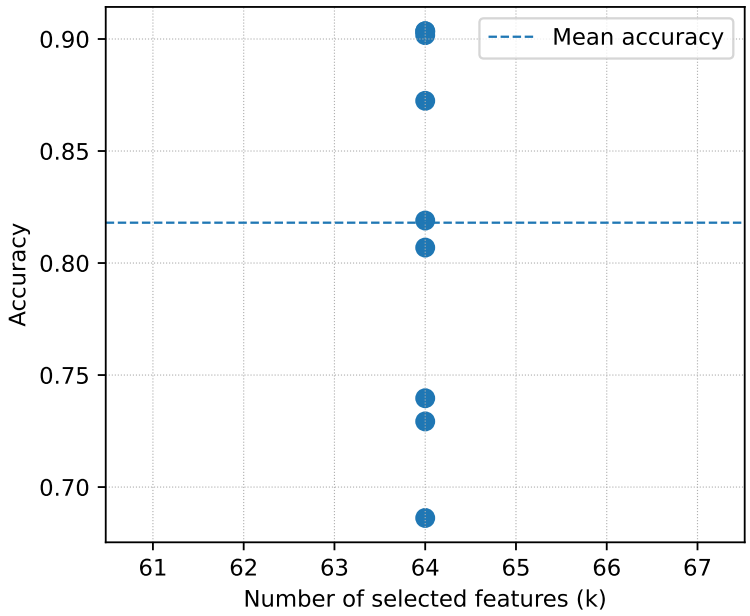


FIGURE 2. Sensitivity to the number of selected features k.

Following confirmation of feature stability, the complete FBCSP pipeline was evaluated using the training session of the BCI Competition IV-2a dataset. The mean classification accuracy across subjects reached **81.8%** with a standard deviation of ±0.9%, accompanied by a Cohen's Kappa value of **0.757**. Individual results for all subjects are reported in Table 1.

TABLE 1. Within-subject performance (training session).

Subject	**Accuracy (%)**	**Kappa**	**Artifacts Removed**
A01T	90.3 ± 2.7	0.871	2
A02T	74.0 ± 4.7	0.653	3
A03T	90.2 ± 3.2	0.869	1
A04T	72.9 ± 6.2	0.639	2
A05T	80.7 ± 4.7	0.742	1
A06T	68.6 ± 7.2	0.581	4
A07T	90.3 ± 2.8	0.871	1
A08T	87.2 ± 2.7	0.830	1
A09T	81.9 ± 3.3	0.759	2
Mean	**81.8 ± 0.9**	**0.757**	**1.9**

To assess whether high accuracy values were influenced by class imbalance, accuracy was compared with Cohen's Kappa across subjects. As shown in Fig. 3, both metrics exhibit a strong linear relationship (Pearson correlation $r = 0.98$) , indicating consistent and reliable performance estimates.

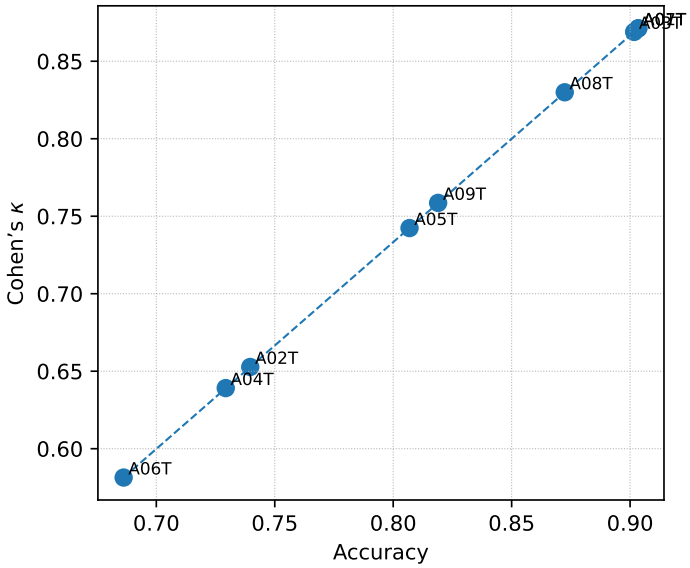


FIGURE 3. Accuracy versus Cohen's Kappa.

Despite high overall performance, classification errors were observed. The confusion matrix for a representative high-performing subject (A01T) is shown in Fig. 4. Misclassifications predominantly occur between left- and right-hand motor imagery, whereas imagery of feet and tongue is classified with higher accuracy.

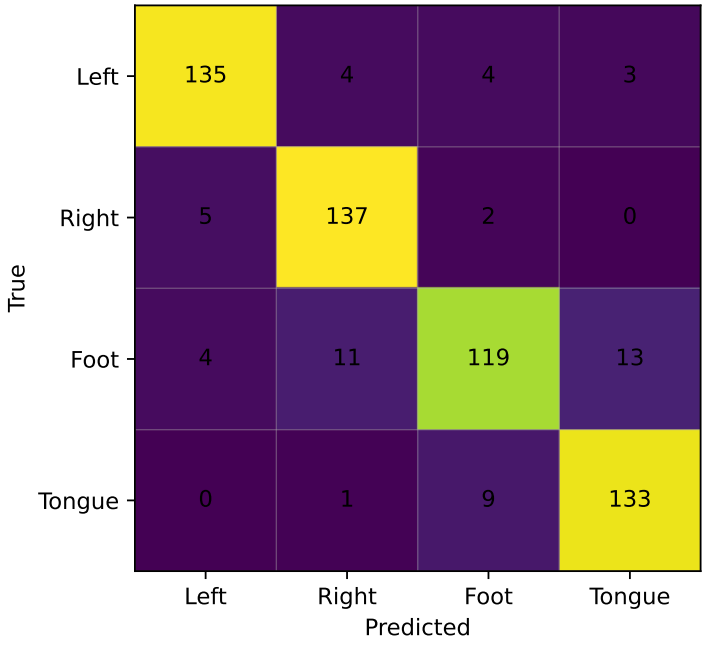


FIGURE 4. Confusion matrix for Subject A01T.

The robustness of the proposed pipeline was further examined on Subject A09, who is often reported as BCI-illiterate due to unstable cross-session performance. In the present within-session evaluation, an accuracy of **81.9%** was achieved, suggesting that the motor imagery signal is present but obscured.

To explore this effect, global pre-trial alpha power was compared between correctly and incorrectly classified trials. As shown in Fig. 5, the distributions largely overlap and no statistically significant difference was detected ($p > 0.05$).

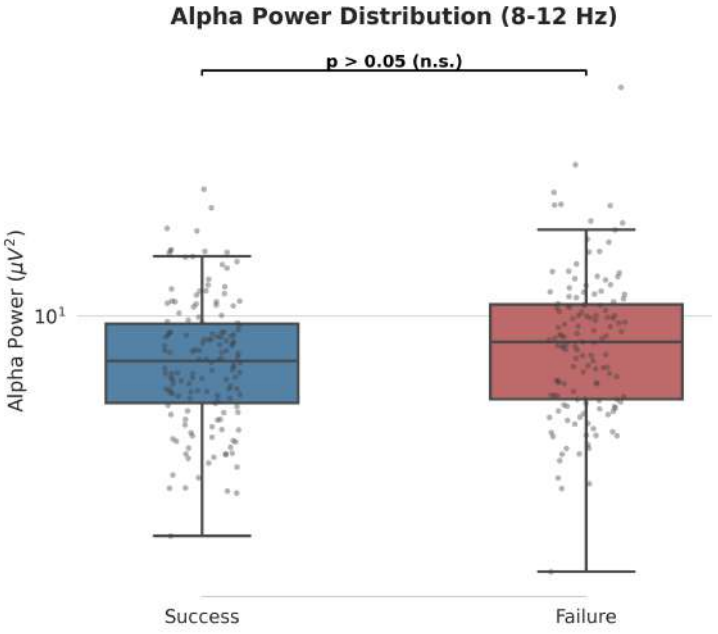


FIGURE 5. Pre-trial alpha power for Subject A09.

Decoding performance across frequency bands for Subject A09 is summarized in Fig. 6. The highest accuracy is observed within the conventional 8–12 Hz alpha band, while neighboring frequency ranges yield lower performance.

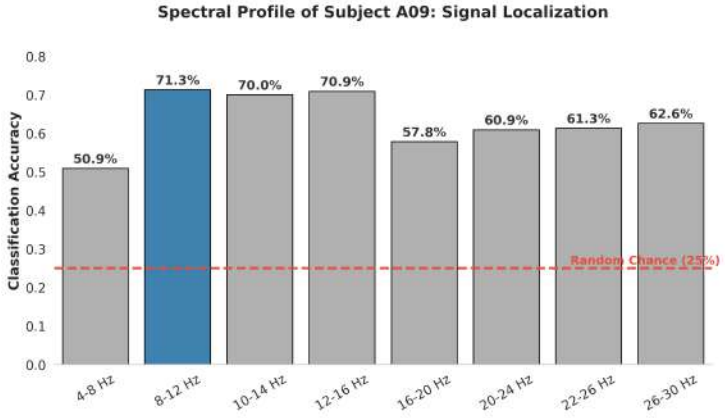


FIGURE 6. Spectral decoding profile for Subject A09.

The spatial structure of the extracted CSP components provides further insight. As illustrated in Fig. 7, dominant components are influenced by occipital alpha activity, while a lower-variance component reveals a localized sensorimotor pattern over C3/C4. This effect, referred to as *Visual Alpha Masking*, indicates that motor-related activity can be spatially overshadowed by strong visual alpha rhythms. Despite reduced signal-to-noise ratio in terms of power, spatial filtering enables recovery of the latent motor command.

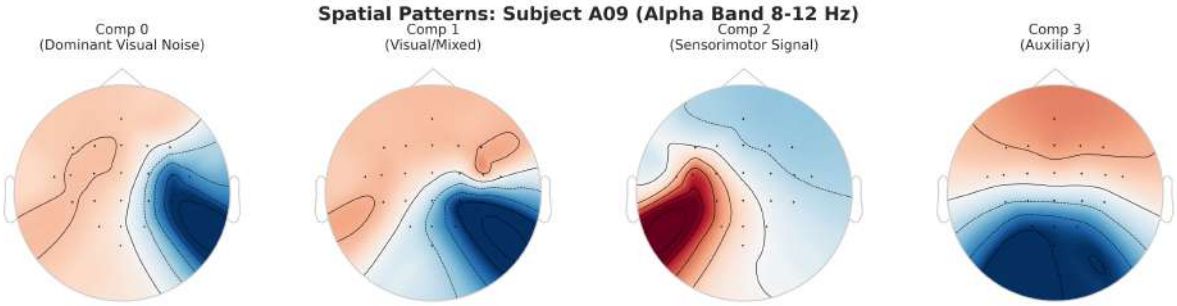


FIGURE 7. Spatial manifestation of Visual Alpha Masking.

4. Discussion

The present within-session results indicate that comparatively low decoding accuracy in motor imagery (MI) BCIs does not necessarily imply an absence of task-related neural information. In some subjects, MI-related patterns may be weakly expressed at the sensor level and partially obscured by competing rhythms with stronger scalp-level amplitude. This interpretation is supported by the analysis of Subject A09, who is frequently reported as unstable in cross-session settings but achieved high within-session accuracy in the present evaluation.

For Subject A09, strong occipital alpha activity can dominate global power measures while the sensorimotor structure remains recoverable through spatial filtering. The comparison of pre-trial alpha power between correctly and incorrectly classified trials did not reveal a clear separation, suggesting that global energy measures are not sufficient to explain trial-level decoding outcomes in this case. In addition, the spectral decoding profile remained centered in the conventional alpha band, which does not support a substantial shift in the informative frequency range as the primary limiting factor.

Further evidence is provided by the spatial domain. While the dominant CSP components for Subject A09 exhibit patterns consistent with visually related alpha activity, lower-variance components preserve a physiologically plausible sensorimotor topology over central regions. This observation motivates the term *Visual Alpha Masking*, describing a scenario in which MI-related activity is present but spatially overshadowed by stronger non-motor rhythms. Importantly, this phenomenon is not readily captured by temporal or power-based summaries alone and becomes apparent when spatial filtering explicitly exploits the topographical organization of neural sources.

From a methodological standpoint, these results highlight the value of spatially adaptive baselines in MI decoding. Spatial filters such as CSP and FBCSP can be interpreted as data-driven spatial projectors that separate overlapping sources even when their relative power differs substantially. Consequently, increased model complexity is not necessarily the most effective first response to low performance; a careful evaluation of whether task-related structure can already be recovered through principled spatial filtering may be informative.

The reliability of the reported performance is further supported by the close correspondence between classification accuracy and Cohen's κ across subjects. The alignment between these metrics suggests that performance differences are unlikely to be primarily driven by class imbalance effects. More generally, the results underline the importance of leakage-free evaluation protocols, particularly in within-subject settings where optimistic bias can be introduced by improper separation of model selection and testing stages.

Several limitations should be noted. First, the analysis is restricted to a within-session setting using the training session of a single benchmark dataset, and the observations may not directly translate to cross-session generalization. Second, the ICA-based artifact rejection relies on automated heuristics and may not fully capture all non-neural sources. Third, *Visual Alpha*

Masking is introduced as an interpretation supported by spatial patterns in the present data; further validation using additional datasets and experimental conditions is needed to assess its generality.

5. Conclusion

This study evaluated an FBCSP-based motor imagery decoding pipeline under leakage-free, nested stratified cross-validation with automated ICA-based artifact rejection in a within-session setting on the BCI Competition IV-2a dataset. The resulting baseline achieved a mean accuracy of 81.8% with a Cohen's κ of 0.757 and exhibited stable performance under changes in the number of selected features.

A focused analysis of Subject A09 provided evidence consistent with the *Visual Alpha Masking* hypothesis: although occipital alpha rhythms can dominate scalp-level power measures, MI-related spatial structure can remain recoverable via spatial filtering. These findings suggest that decoding difficulties may, in some cases, reflect insufficient separation of overlapping sources rather than an absence of task-related information. Overall, the results support the continued use of carefully validated spatial filtering pipelines as interpretable and reliable baselines for motor imagery BCI research, especially when methodological transparency and evaluation rigor are prioritized.

References

[1] Ang K. K., Chin Z. V., Zhang H., Guan C., Filter bank common spatial pattern (FBCSP) in brain–computer interface, *Proceedings of the IEEE International Joint Conference on Neural Networks*, Vol. 1, 2008, pp. 2390–2397.

[2] Schirrmeister R. T., Springenberg J. T., Fiederer L. D. J., Glasstetter M., Eggensperger K., Tangermann M., Hutter F., Burgard W., Ball T., Deep learning with convolutional neural networks for EEG decoding and visualization, *Human Brain Mapping*, Vol. 38, No. 11, 2017, pp. 5391–5420.

[3] Lawhern V. J., Solon A. J., Waytowich N. R., Gordon S. M., Touryan C. J., Oley M. E., EEGNet: a compact convolutional neural network for EEG-based brain–computer interfaces, *Journal of Neural Engineering*, Vol. 15, No. 5, 2018, p. 056013.

[4] Tangermann M., Roberts S. J., Scherer R., Neuper C., Müller-Putz G., Pfurtscheller G., Review of the BCI Competition IV, *Frontiers in Neuroscience*, Vol. 6, 2012, p. 55.

[5] Hyvärinen A., Fast and robust fixed-point algorithms for independent component analysis, *IEEE Transactions on Neural Networks*, Vol. 10, No. 3, 1999, pp. 626–634.

[6] Cohen J., A coefficient of agreement for nominal scales, *Educational and Psychological Measurement*, Vol. 20, No. 1, 1960, pp. 37–46.

COMPREHENSIVE SCIENTIFIC ASSESSMENT OF THE ENVIRONMENTAL CONDITION OF TERRITORIES AS A BASIS FOR GOVERNMENTAL DECISION-MAKING (CASE STUDY OF KENTAU CITY)

S. A. ASKAROV[1], G. D. BERKINBAYEV[1], N. A. YAKOVLEVA[1], E. K. SADVAKASOV[1], A. M. AYAZHANOV[1]

[1]Ecoservice-S LLP, Almaty, Kazakhstan
e-mail: info@ecoservice.kz

Keywords: Kentau city, environmental assessment criteria, integrated environmental study, environmental emergency, pre-crisis ecological zone

Introduction

The conservation of ecosystems and the creation of favorable living conditions for the population are among the key components of the Sustainable Development Goals adopted by the United Nations [1]. In the Republic of Kazakhstan, substantial efforts are being undertaken in this direction. In particular, the Environmental Code adopted in 2021 establishes a legal framework that allows, based on the results of comprehensive studies of environmental conditions and public health, decisions to be made on classifying a territory as a zone of ecological disaster, a zone of emergency ecological situation, or a territory with relatively satisfactory environmental conditions [2], followed by the implementation of a system of measures depending on the assigned environmental status.

It should be noted that, in accordance with the Environmental Code of the Republic of Kazakhstan, comprehensive studies of the ecological condition of a territory must be conducted in compliance with the requirements of the *Criteria for Assessing the Ecological Situation of Territories* [3].

According to Article 405 of the Environmental Code, the results of a comprehensive scientific assessment of the ecological condition of a territory are submitted for expert review to the authorized body in the field of environmental protection, as well as to the authorized state bodies in the fields of healthcare, education, and science.

If environmental problems meeting the established Criteria are identified, the studied territory, based on the conclusions of the expert assessments, is declared by a resolution of the Government of the Republic of Kazakhstan as a zone of emergency ecological situation, or, by a law of the Republic of Kazakhstan, as a zone of ecological disaster [2].

Regulatory legal acts specify the boundaries of a zone of emergency ecological situation or a zone of ecological disaster; the duration of such status; the legal regime governing these zones; measures aimed at stabilizing and reducing adverse environmental conditions within the relevant territory or provisions indicating the need for their development; as well as the procedure for classifying citizens as affected by an emergency ecological situation or ecological disaster and the measures for their social protection [2].

The study was carried out by Ecoservice-S LLP under commission from the Akimat of Kentau within the project “Comprehensive assessment of the environmental condition and public health of Kentau city and adjacent settlements”.

At the same time, measures for the restoration (reproduction) of natural resources, improvement of environmental quality, and provision of medical assistance to the population are developed and implemented in a differentiated manner [2].

Thus, Kazakhstan has the institutional and legal capacity to substantiate the environmental status of a territory on the basis of a comprehensive scientific assessment of environmental conditions and to adopt appropriate state-level decisions.

Below, using the city of Kentau and the adjacent settlements as an example, the practical implementation of this approach is demonstrated.

Main Part

The study was carried out by the well-known Kazakhstani environmental company Ecoservice-S LLP under commission from the Akimat of the city of Kentau within the framework of the project *"Comprehensive Assessment of the Environmental Condition and Public Health of the City of Kentau and Adjacent Settlements"*. The period of investigation covered September 2019 to December 2020.

The relevance of conducting a comprehensive environmental assessment in this region was determined by several factors:

- changes in the geological environment resulting from past mining activities, including heavy-metal contamination of the environment, flooding of building foundations, and the formation of subsidence features within the urban area;
- the fact that the city of Kentau and adjacent settlements (Achisai, Kantagi, Bayaldyr, Karnak, Kushata, Shashtobe, and Burgem) are surrounded by territories classified as zones of ecological pre-crisis condition related to the Aral Sea problem, while not formally belonging to this zone;
- the absence of objective and comprehensive information on the environmental condition and public health in the city of Kentau and the adjacent settlements.

The objective of the study was to provide an integrated characterization of the environmental condition and the level of public health in the city of Kentau and the adjacent settlements (Achisai, Kantagi, Bayaldyr, Karnak, Kushata, Shashtobe, and Burgem) in order to obtain a scientifically substantiated conclusion regarding the possibility of classifying the investigated territories as a zone of emergency ecological situation or a zone of ecological disaster in accordance with the *Criteria for Assessing the Ecological Situation of Territories*, or as a zone of ecological pre-crisis condition according to the classification of environmental disturbance of territories in the Kazakhstani part of the Aral Sea region [4].

The overall design of the study was defined by the requirements specified in the above-mentioned regulatory documents [3, 4]. Specialists from leading research institutes and organizations of the Republic of Kazakhstan and neighboring countries were involved in the study, including the D. A. Kunayev Institute of Mining, the U. U. Uspanov Kazakh Research Institute of Soil Science and Agrochemistry, the Kazakh Research Institute of Oncology and Radiology, the Institute of Biology and Plant Biotechnology, the Institute of General Genetics and Cytology, Khoja Akhmet Yassawi International Kazakh–Turkish University, the Medical Center Hospital of the Presidential Administration, Perm State National Research University, the Ural State Mining University (Yekaterinburg), and the South Ural Federal Scientific Center for Mineralogy and Geoecology of the Russian Academy of Sciences (Miass), among others. Laboratory analyses were conducted exclusively in accredited laboratories.

Field investigations were carried out within the territories of the following settlements: the city of Kentau and the villages of Achisai, Kantagi, Bayaldyr, Karnak, Kushata, Shashtobe, and Burgem. The settlement of Lenger (Tolebi District, Turkestan Region) was selected as the background area. The city of Lenger has socio-economic conditions similar to those of Kentau but is located in an environmentally more favorable area. As comparison territories, the nearest

urban (Turkestan) and rural (Orangai) settlements classified as zones of ecological pre-crisis condition were selected.

The work was conducted along five main directions: assessment of the ecological situation; analysis of geological and hydrogeological conditions; evaluation of public health and the socio-economic situation; and development of a mitigation and action plan.

A comprehensive set of methods was applied during the implementation of the project, including field investigations; laboratory analysis of samples from environmental components (soil, water, air, and agricultural products) and biological media (blood and hair); methods for detecting genetic mutations; assessment of phytotoxicity; cartographic analysis using GIS software and remote sensing data; geophysical methods (vertical electrical sounding (VES), ground-penetrating radar surveys, and others); hydrogeological methods; and methods for assessing public health status (immunological, genetic, anthropometric analyses, medical examinations, etc.).

To investigate natural environments and collect samples, field teams were formed and equipped with vehicles, sampling equipment, and georeferencing tools for precise localization of sampling points.

Within the scope of the environmental assessment block, the following aspects were examined: ambient air quality; drinking and surface water quality; soil contamination levels; radiological conditions; total mutagenic activity of drinking water and residential soil samples; depletion of water resources; ecological condition of soils; quality of agricultural products; degradation of terrestrial ecosystems; condition of vegetation; soil phytotoxicity; and the status of fauna.

Ambient air quality assessment was conducted through field measurements. In total, more than 14 000 air quality measurements were performed. As a result, 195 exceedances of maximum permissible concentrations (MPCs) for single-time measurements and 166 exceedances for average daily values were identified. In most cases, the exceedance levels did not exceed three times the MPC. Elevated concentrations were primarily observed for sulfur dioxide, nitrogen dioxide, suspended particulate matter, and carbon monoxide. Data from the Kazhydromet monitoring system and other organizations were also taken into account.

In accordance with the technical specifications, the composition of atmospheric deposition was analyzed. The results indicated the influence of both regional processes, including dust transport from the Aral Sea region, and local pollution sources. Samples of dry atmospheric deposition revealed elevated concentrations of heavy metals and arsenic, which are characteristic pollutants of the Kentau urban agglomeration. In particular, maximum lead concentrations of 780 mg/kg (24.4 MPC) and 560 mg/kg (17.5 MPC) were recorded in samples collected in the village of Achisai and the city of Kentau, respectively. These results indicate dust emission from lead-contaminated areas and long-range transport of lead-bearing dust.

Analysis of satellite imagery from the *TERRA* and *AQUA* platforms for the period from 2002 to 2020 confirmed the possibility of salt–dust transport from the Aral Sea region. It was noted that under global trends such as increasing air temperatures, more frequent droughts, and intensified storm activity associated with climate change, the removal of dust and salt particles from the dried seabed of the Aral Sea is likely to intensify, followed by their deposition in the region. This process may lead to the desertification and aridization of soils.

The quality of drinking water from centralized and decentralized water supply systems was assessed based on quarterly sampling and laboratory analysis. In total, more than 175 drinking water samples were collected and analyzed from 26 sampling points, covering 11 settlements. In certain water samples, exceedances of existing standards for hardness, total mineralization, and iron content (ranging from 1.1 to 4 MPC) were identified. These exceedances are associated with the natural geochemical background of the region.

Surface water quality assessment was carried out through the analysis of water samples collected from the Bayaldyr, Biresik, Achisai, Khantagi, Karashyk, and Ermekozen rivers, as well

as the Kushkorgan and Shashtobe reservoirs and a drainage adit channel. In a number of samples, exceedances of established regulatory limits were detected for dissolved oxygen, sulfates, lead, zinc, and nitrites.

To assess soil contamination within the Kentau urban agglomeration, extensive field investigations were conducted, including more than 21 000 elemental determinations. The concentrations of heavy metals were determined, including total content, mobile and water-soluble forms, and bioavailability, as well as bacteriological indicators, pesticide residues, and radiological parameters. Areal contamination of soils by heavy metals, particularly lead, was identified. The primary source of contamination is associated with past industrial activities. According to territorial zoning based on the cumulative soil pollution index (Z_c), areas with moderately hazardous and hazardous levels of soil contamination were identified within the city of Kentau and the settlements of Bayaldyr, Kantagi, and Achisai.

The assessment of the radiation situation in the region included pedestrian gamma surveys (more than 12 000 gamma radiation measurements), measurements of indoor radon concentrations using both instantaneous and integrated methods, and determination of radon concentrations in water. The recorded gamma background ranged from 0.09 to 0.15 μSv/h.

Evaluation of water resource depletion within the Kentau urban agglomeration involved the study of five rivers located directly within the study area and reflecting potential changes in local water resources. The assessment of filtration losses in the runoff of the Karashyk, Biresik, and Bayaldyr rivers demonstrated the role of these rivers in the formation of groundwater flows entering the territory of Kentau through near-surface deposits, as well as their influence on the redistribution of groundwater pressures across all aquifers and hydrogeological complexes underlying the city.

The study of vegetation condition revealed that, in certain areas, forage plant species are being replaced by weedy species of low nutritional value; the distribution ranges of endemic plant species are decreasing; and vegetation cover is degrading. In addition, damage to and deterioration of the vitality of woody plant species were observed as a result of habitat changes, reaching up to 80% in some areas.

Investigation of the fauna within the Kentau urban agglomeration showed a significant decline in biodiversity compared to 2017. This trend was particularly evident among insects, the most abundant class of animals, for which species diversity decreased by approximately 30%. As insects are a key component of trophic networks in biocenoses, a reduction in their diversity and abundance inevitably leads to a decline in overall faunal diversity and population size, indicating unfavorable changes in the ecological situation. The proportion of harmful and sanitary–epidemiologically significant species was high at all sampling sites, exceeding 50–60% in most locations and reaching up to 80% in some cases. Such values indicate a substantially disturbed state of the biocenosis (values approaching 30% are generally characteristic of only slightly disturbed ecosystems, whereas unstable ecosystems tend to favor harmful species that are better adapted to adverse environmental conditions).

The results of the conducted studies, characterized by relatively low faunal diversity, low abundance (with the exception of harmful and synanthropic species), and infrequent occurrence, indicate that the biodiversity of the animal world within the Kentau urban agglomeration is currently disturbed.

A forecast of further deterioration of the ecological situation was developed, initially for specific locations (the city of Kentau and the settlements of Karnak, Shashtobe, Kushata, and Orangai), and subsequently for the Kentau urban agglomeration as a whole. It was noted that this forecast is driven not only by the existing unfavorable environmental conditions but also by climate change associated with global warming. As a result, animal species adapted to hot and arid conditions are expected to gain further advantage, as are harmful synanthropic species capable of persisting in close association with human activity.

Assessment of the geological and hydrogeological situation included investigation of groundwater contamination levels, analysis of processes associated with geological deformation and surface subsidence, as well as the study of indicators of induced seismicity and karst formation.

Flooding represents one of the major problems in Kentau, as pressurized groundwater inundation leads to the leaching of gypsum from clay strata, resulting in the formation of cavities and surface subsidence. These processes, in turn, cause damage to buildings founded on clayey deposits.

Observations of surface subsidence and ground displacement within the city of Kentau were conducted using two complementary approaches: satellite (radar) interferometry and instrumental measurements at established monitoring points. Interferometric data revealed both subsidence and uplift processes. It was established that subsidence occurs at average rates of up to 12 mm per year, particularly in the western and southwestern parts of the territory, while uplift reaches rates of up to 7 mm per year. Given that the risk of increasing ground deformation within Kentau persists, it was recommended to continue areal geomonitoring across residential areas in combination with instrumental observations.

Seismological monitoring was carried out within the city of Kentau to assess indicators of potential induced seismicity and to delineate areas susceptible to disturbance from external impacts. The study of karst formation processes in Kentau involved the collection, analysis, and synthesis of results from all previous geophysical investigations, as well as the application of geophysical methods under field conditions, including electrical resistivity surveys and ground-penetrating radar profiling.

Based on electrical survey data, tectonic structures and conductive zones potentially associated with karst manifestations were identified. Integrating all available data, a karst hazard map was compiled, distinguishing zones of very high hazard, high hazard, potential hazard, and low hazard. The low-hazard zone occupies most of the southern part of the city, whereas the very high and high hazard zones are confined to the northeastern and northern parts of Kentau, including the industrial area.

The degree of terrestrial ecosystem degradation was assessed using satellite image interpretation combined with landscape–ecological investigations. The presence of suffosion and collapse phenomena in the northeastern part of Kentau, waste rock dumps, sinkholes associated with karst processes, open pits, and degraded lands, collectively exceeding 5% of the total area, allowed the territory adjacent to the northern part of Kentau and covering approximately 900 ha to be classified as an area with an emergency ecological situation.

A comprehensive assessment of population health was conducted, including analysis of the dynamics and structure of child, perinatal, and infant mortality; evaluation of health status based on screening studies; assessment of morbidity trends; analysis of the incidence of congenital malformations in newborns; in-depth evaluation of morbidity and mortality from oncological diseases; analysis of reproductive health disorders in women; characterization of the frequency of pregnancy complications, adverse pregnancy outcomes, and childbirth complications; assessment of the condition of full-term newborns; investigation of the mental development of school-aged children; determination of the frequency of genetic abnormalities in population cells; immunological studies; detailed medical examinations of selected population groups; assessment of heavy metal accumulation in biological media of children; and evaluation of health risks to the population. The study revealed generally low indicators of population health.

The integrated assessment of environmental conditions and population health provided the basis for classifying the territories of the city of Kentau, as well as the settlements of Karnak, Shashtobe, Kushata, Achisai, Khantagi, Bayaldyr, and Burgem, as a zone of pre-crisis ecological condition. This classification was based on evidence of excessive soil contamination, vegetation degradation, reduction in the abundance and species diversity of fauna, and specific health indicators of the population (including immunological status, results of in-depth medical examinations, and morbidity rates), in accordance with the Law of the Republic of Kazakhstan dated

30 June 1992 No. 1468-XII "On Social Protection of Citizens Affected by Ecological Disaster in the Aral Sea Region".

The influence of the Aral Sea region was confirmed. Interpretation of satellite imagery demonstrated that, beginning in the early 2000s, the transport of sand–silt particles of saline aerosol from the dried bottom of the Aral Sea has reached distances of 700 km and more, covering the territory of the Kentau urban agglomeration. In addition, the results of analyses of dry atmospheric deposition composition indicate the impact of regional processes, including dust transport originating from the Aral Sea region.

Based on the results of investigations conducted in 2019–2020, sufficient grounds were identified for classifying part of the Kentau urban agglomeration as a zone of emergency ecological situation, in accordance with the Criteria for Environmental Situation Assessment [3].

Conclusion

Following the completion of all stages of expert review and coordination procedures stipulated by the Environmental Code of the Republic of Kazakhstan, amendments were introduced to the Law of the Republic of Kazakhstan "On Social Protection of Citizens Affected by Ecological Disaster in the Aral Sea Region". As a result, the territory of the city of Kentau and adjacent settlements was classified as a zone of pre-crisis ecological condition. These amendments entered into force on 1 July 2025.

In addition, in accordance with Order No. 106-p of the Ministry of Ecology and Natural Resources of the Republic of Kazakhstan dated 23 April 2025, the industrial zone of the city of Kentau was officially declared a zone of emergency ecological situation.

It should be emphasized that such a decision to amend the legislative framework, based on the results of a comprehensive environmental and health assessment, was implemented in the Republic of Kazakhstan for the first time.

The adoption of these legislative acts ensures the implementation of a comprehensive set of measures aimed at protecting the population from adverse environmental factors and improving social conditions.

Thus, the comprehensive study conducted in the city of Kentau and adjacent settlements not only provided an objective characterization of environmental conditions and population health, but also served as a scientifically substantiated basis for governmental decision-making with high social significance.

The presented study may serve as a reference example for conducting similar comprehensive assessments in other regions experiencing environmental degradation.

References

[1] United Nations. *Transforming Our World: The 2030 Agenda for Sustainable Development*. Resolution adopted by the General Assembly on 25 September 2015. Available at: `https://docs.un.org/en/A/RES/70/1`.

[2] Republic of Kazakhstan. *Environmental Code of the Republic of Kazakhstan*. Law No. 400-VI ZRK dated 2 January 2021. Available at: `https://adilet.zan.kz/rus/docs/K2100000400`.

[3] Republic of Kazakhstan. *On Approval of Criteria for Assessing the Environmental Situation of Territories*. Order of the Acting Minister of Ecology, Geology and Natural Resources of the Republic of Kazakhstan No. 327 dated 13 August 2021. Registered with the Ministry of Justice of the Republic of Kazakhstan on 16 August 2021, No. 23994. Available at: `https://adilet.zan.kz/rus/docs/V2100023994`.

[4] Republic of Kazakhstan. *On Social Protection of Citizens Affected by Ecological Disaster in the Aral Sea Region*. Law No. 1468-XII dated 30 June 1992. Available at: `https://adilet.zan.kz/rus/docs/Z920002600/`*history*.

Application of the RBF Interpolation Method for Identifying Geochemical Anomalies

[1]Al-Farabi Kazakh National University,

[2]National Engineering Academy of the Republic of Kazakhstan

1. Introduction

In recent years, interpolation methods have been widely applied in geoinformation and geochemical studies to reconstruct continuous concentration fields from a limited number of measured data points. This approach not only enables visualization of the spatial distribution of elements but also facilitates the identification of mineralization patterns and geochemical anomalies.

The present study is devoted to the construction of spatial models of geochemical element concentrations—gold (Au), silver (Ag), aluminum (Al), antimony (Sb), and mercury (Hg)—based on radial basis functions (RBFs). The RBF method provides a smooth approximation and allows for the prediction of concentrations at intermediate locations, which is particularly relevant for geological objects characterized by complex structural features.

The study utilizes geochemical data comprising approximately 800 initial measurement points. Based on these data, 50,000 coordinate points were generated using the RBF method, and the concentration distributions of each element were determined at these locations. For analytical and visualization purposes, two-dimensional maps were constructed to illustrate zones of local accumulation, and ten points with the highest concentration values were identified for each element.

The developed models make it possible not only to trace spatial variations in element concentrations but also to identify areas of potential mineralization, which is of practical significance for geological exploration activities.

Object of the study: Geochemical data represented as a coordinate-based dataset of concentrations of the elements Au, Ag, Al, Sb, and Hg.

Purpose of the study: To determine the spatial distribution of geochemical elements and to visualize their concentrations in two-dimensional coordinate systems using the radial basis function (RBF) method, as well as to identify patterns of their joint distribution, delineate zones of local accumulation, and construct predictive models for assessing potentially mineralized areas within the study region.

2. Source Data and Methodology

2.1. **Source Data.** The initial data were obtained from the file *"Geochemical Data.xlsx"*, which contains spatial coordinates and concentration values of five chemical elements: Au, Ag, Al, Sb, and Hg.

Each row of the dataset represents the result of a geochemical sampling point, including its spatial coordinates and the measured concentrations of the selected elements.

2

№	X (КООРД.)	Y (КООРД.)	AU (PPM)	AG (PPM)	AL (%)	SB (PPM)	HG (PPM)
1	82.315	49.627	0.35	0.42	7.25	1.20	0.08
2	82.317	49.630	0.28	0.31	7.18	1.11	0.06
3	82.320	49.634	0.47	0.55	7.41	1.36	0.09
4	82.324	49.638	0.39	0.44	7.29	1.24	0.07
5	82.328	49.642	0.51	0.60	7.53	1.48	0.10

FIGURE 1. Example of the structure of the initial geochemical dataset

2.2. **RBF Interpolation Method.** In this study, an interpolation method based on radial basis functions (RBFs) was employed. This approach makes it possible to construct a continuous approximating surface of geochemical element concentration distributions from discrete experimental data.

The general interpolation formula can be expressed as follows:

$$f(x, y) = \sum_{i=1}^{N} \omega_i \, \varphi(\|x - x_i\|) \,, \tag{1}$$

where $f(x, y)$ is the computed value of the interpolated function at an arbitrary point in space, $x_i = (x_i, y_i)$ are the coordinates of the known observation points, ω_i are the weighting coefficients, $\varphi(r)$ is the radial basis function, $r = \|x - x_i\|$ denotes the Euclidean distance between the interpolation point and the observation point.

2.3. **Multiquadric (MQ) Radial Basis Function.** In this study, a Multiquadric (MQ) type radial basis function was employed, which is defined as follows:

$$\varphi(r) = \sqrt{1 + (\varepsilon r)^2}, \tag{2}$$

where $\varepsilon > 0$ is the scale (or smoothing) parameter that controls the degree of smoothness and the sensitivity of the interpolation.

Physical and Mathematical Interpretation. The Multiquadric function is characterized by a high degree of smoothness and continuity, which makes it particularly suitable for describing gradual spatial variations in the concentrations of geochemical elements. It effectively models large-scale trends, such as the distribution of mineralization zones, and demonstrates robustness with respect to random noise in the input data.

Due to these properties, the MQ function is widely applied in geological and geophysical problems, especially in modeling the spatial distribution of elements such as gold (Au), silver (Ag), antimony (Sb), and mercury (Hg).
Software: Python (NumPy, SciPy, Matplotlib, Plotly).

3. Results and Analysis

3.1. **Two-Dimensional Interpolation (2D).** Within the framework of the present study, interpolation was performed at 50,000 coordinate points using the radial basis function (RBF) method. For each element (Au, Ag, Al, Sb, and Hg), individual concentration distribution

models were constructed. Based on the obtained results, ten points with the highest predicted concentration values were identified and presented for each element.

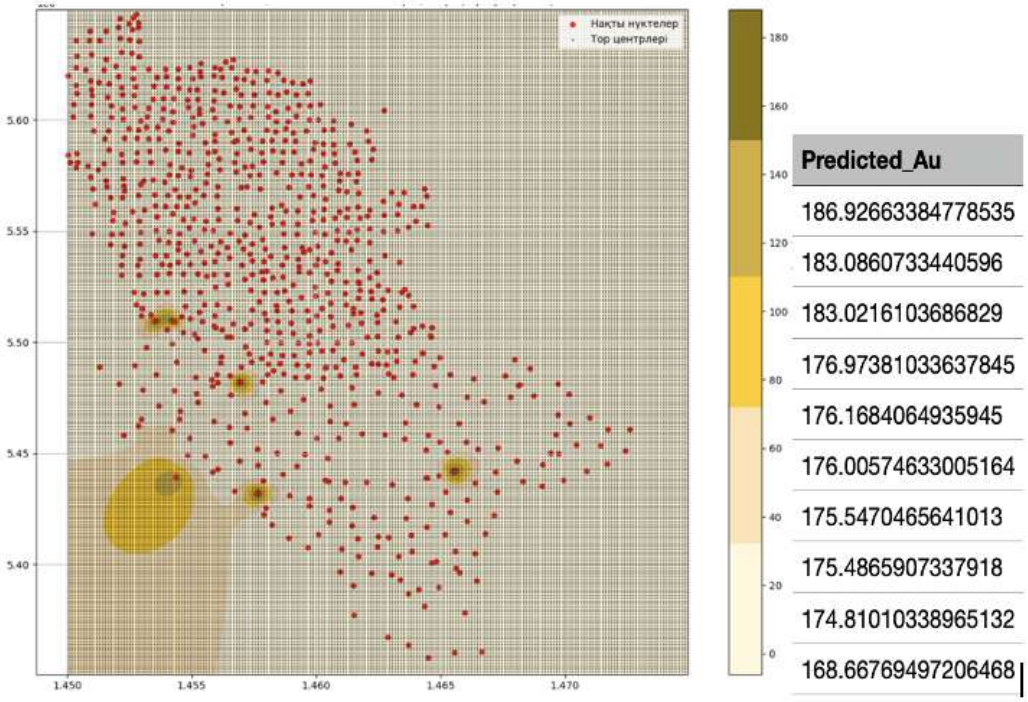


FIGURE 2. Example of the structure of the initial geochemical dataset

Figure 3.1 – Two-dimensional map of the spatial distribution of gold (Au). The interpolation was performed at 50,000 coordinate points using the radial basis function (RBF) method. Gold concentrations are represented in yellow on the map. Ten points with the highest concentration values are highlighted, corresponding to zones of local accumulation and potential gold mineralization.

4

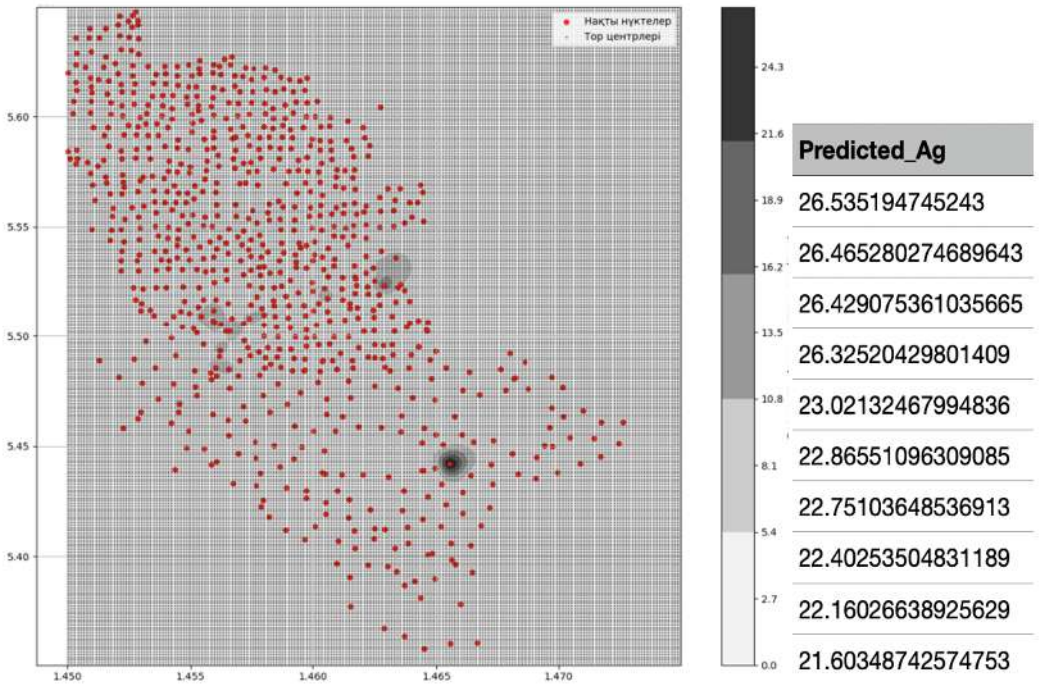


FIGURE 3. Example of the structure of the initial geochemical dataset

The map was constructed based on the interpolation of 50,000 coordinate points. Silver concentrations are shown in gray, and ten maximum points characterizing areas of elevated element content are highlighted. The anomalous zones partially overlap with the gold distribution, indicating a possible Au–Ag association.

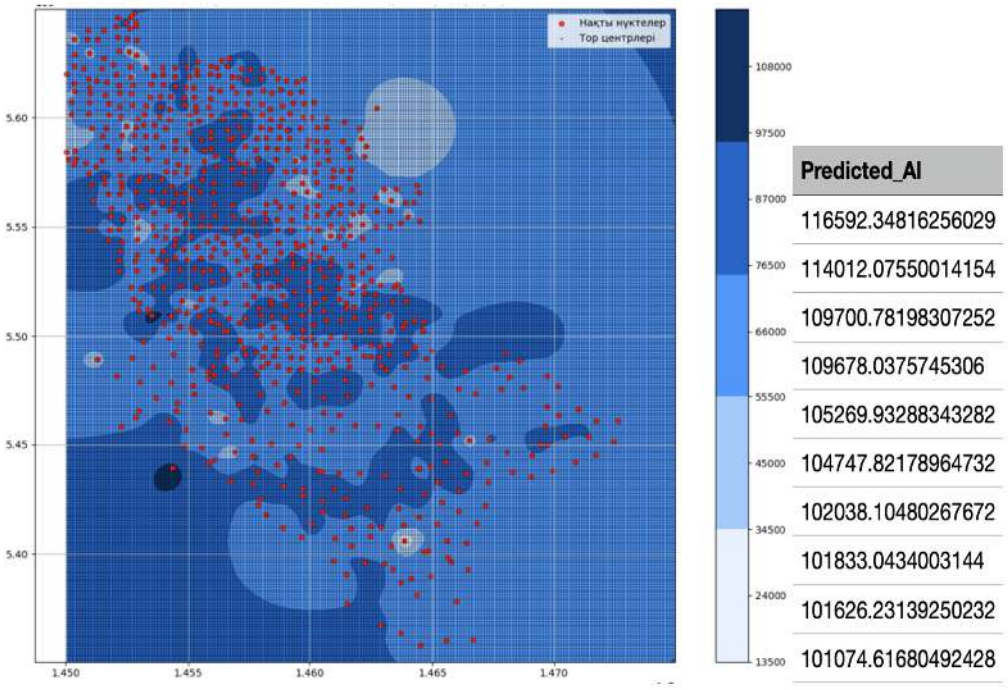


FIGURE 4. Example of the structure of the initial geochemical dataset

The interpolation was performed using 50,000 measurement points. Aluminum is shown in blue and is characterized by a relatively uniform background distribution over the entire study area. The ten highlighted points indicate minimal concentration variations and confirm the lithological homogeneity of the background.

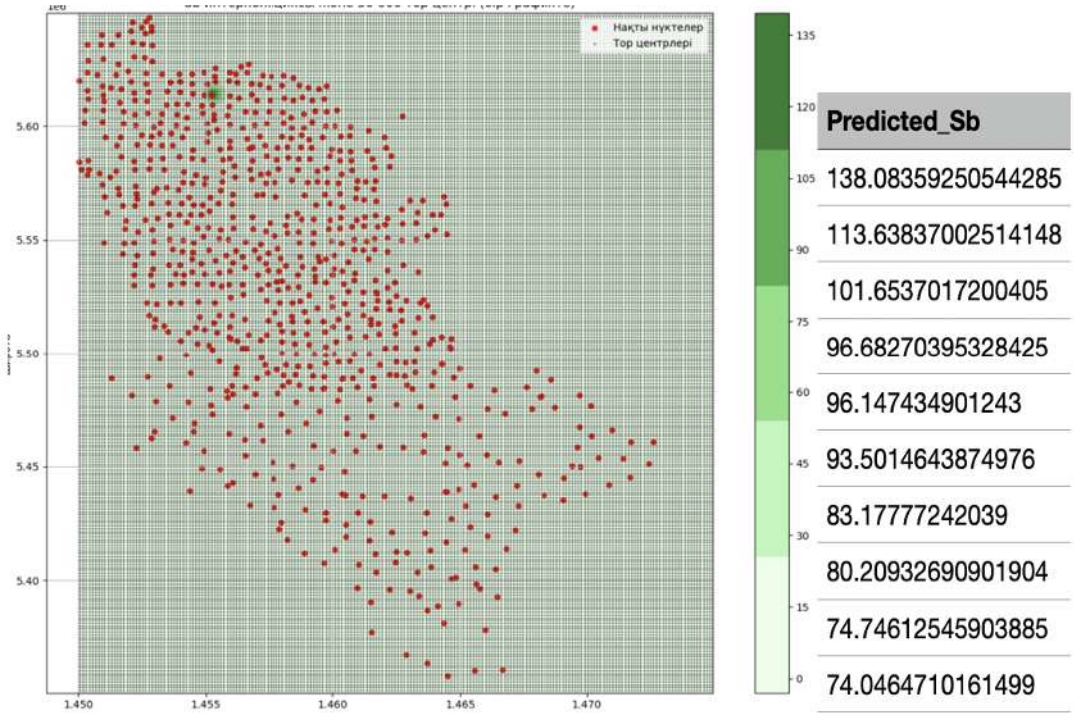


FIGURE 5. Example of the structure of the initial geochemical dataset

The distribution was calculated based on 50,000 coordinate points. Antimony is represented in green. Ten points with maximum concentration values are highlighted on the map, corresponding to tectonic structures and zones of increased hydrothermal activity.

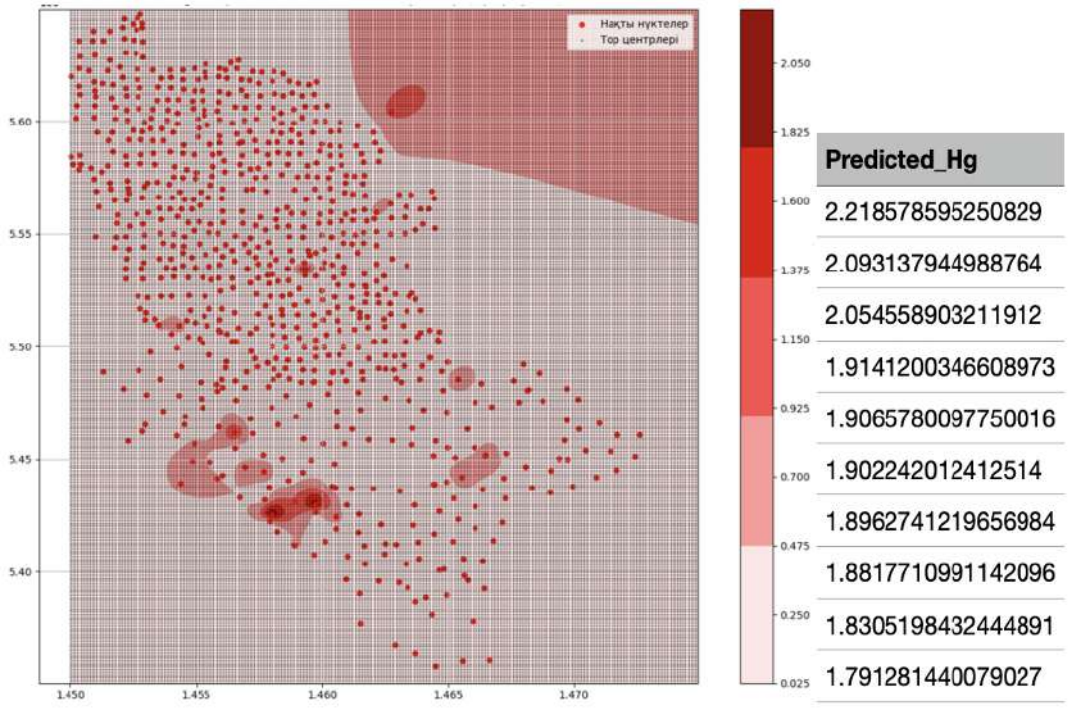


FIGURE 6. Example of the structure of the initial geochemical dataset

The interpolation was performed using 50,000 points. Mercury is shown in red, and ten areas with the highest concentration values are highlighted, predominantly located in the southeastern part of the study area. Zones of elevated Hg content spatially coincide with Sb anomalies, indicating a common hydrothermal origin. Each map illustrates characteristic patterns of the spatial distribution of individual elements:

- **Gold (Au):** distinct local zones of elevated concentrations are observed, primarily located in the central part of the study area.
- **Silver (Ag):** exhibits distribution patterns similar to those of Au.
- **Aluminum (Al):** shows a uniform background distribution across the entire area.
- **Antimony (Sb) and Mercury (Hg):** are characterized by localized anomalous zones that tend to be associated with tectonic structures.

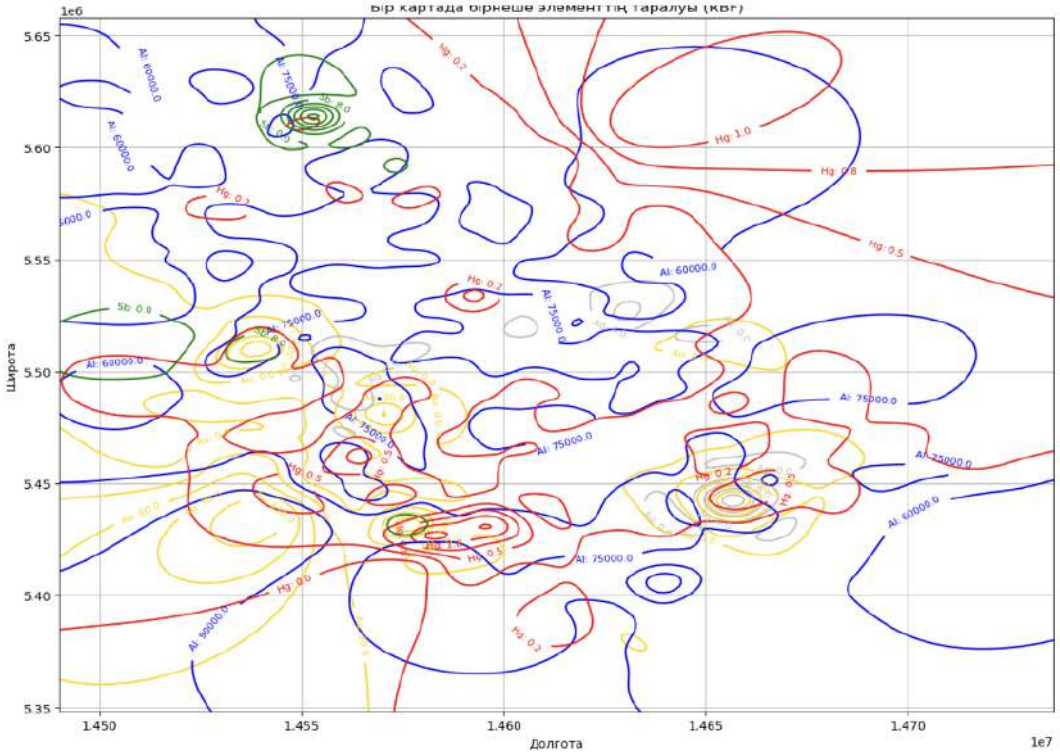


FIGURE 7. Example of the structure of the initial geochemical dataset

At the next stage, all obtained results were visualized on a single combined map, which made it possible to clearly represent the spatial distribution of the predicted values and the relative positions of anomalous zones.

In the combined map, each element is represented by a specific color: Au—yellow, Ag—gray, Al—blue, Hg—red, and Sb—green.

The integrated visualization enabled the identification of spatial relationships between anomalous zones. The overlap of Au and Ag anomalies indicates the presence of potentially polymetallic areas, while the coincidence of Sb and Hg zones suggests the influence of hydrothermal processes.

Gold (Au): as a result of interpolation, several high-concentration zones were identified, which may be considered potential ore clusters.

Silver (Ag): zones overlapping with Au were detected, indicating a possible occurrence of polymetallic mineralization, particularly where spatial superposition of anomalies is observed.

Mercury (Hg) and Antimony (Sb): in certain areas, these elements overlap with Au and Ag anomalies, suggesting a hydrothermal or magmatic origin of the observed anomalies. Aluminum

(Al): exhibits a uniform distribution reflecting the geological background and can be used as a normalizing element in anomaly detection.

3.2. **Conclusions Based on the Interpolation Results.**

- The RBF interpolation method provided stable and physically consistent models of the spatial distribution of the studied elements.
- The spatial overlap of Au and Ag anomalies indicates their possible association within a single mineralized zone.
- Local Sb and Hg anomalies may be related to hydrothermal processes or deep-seated magmatic sources.
- Aluminum (Al) can be used as a background marker in the comparative analysis of other geochemical anomalies.
- The combined map (Figure 3.2) enables the identification of complex geochemical anomalies that are promising for further prospecting and geological exploration activities.

4. Conclusion

In this study, a methodology for spatial modeling of geochemical element concentrations based on the radial basis function (RBF) interpolation method was implemented. This approach made it possible to construct continuous approximating surfaces of element distributions and to identify anomalous concentration zones with a high degree of reliability.

Based on the initial dataset, interpolation was performed at 50,000 coordinate points for five key elements: gold (Au), silver (Ag), aluminum (Al), antimony (Sb), and mercury (Hg). For each element, individual two-dimensional maps were generated, illustrating the distribution of predicted values and concentration isolines.

In addition, ten points with the highest predicted concentration values were identified for each element, enabling the localization of areas with the greatest geochemical potential. The combined spatial distribution map provided a comprehensive visualization of the relationships between mineralization zones of different elements.

Particular attention was given to the overlapping anomalies of Au and Ag, which may indicate the presence of potential polymetallic deposits. Sb and Hg anomalies detected near tectonic faults suggest a probable hydrothermal or magmatic origin of the geochemical processes. Aluminum (Al) exhibited a uniform background distribution and was therefore used as an indicator of lithological homogeneity within the study area. The interpolation results demonstrated high stability and physical consistency of the developed models. The RBF method proved to be effective for processing large datasets (50,000 points) and for visualizing complex spatial relationships. The algorithm not only enabled the reconstruction of continuous concentration fields but also revealed spatial distribution patterns that are difficult to detect using classical statistical methods.

Thus, the application of the radial basis function method made it possible to:

- construct continuous models of geochemical fields;
- identify anomalous zones of elevated element concentrations;
- reveal relationships among different geochemical indicators;
- visualize spatial structures and potential mineralization areas;
- establish a foundation for the further application of machine learning algorithms and geographic information systems (GIS) in the prediction of prospective mineral zones.

The obtained results confirm that the RBF interpolation method is a reliable and efficient tool for the analysis and prediction of geochemical data and can be effectively applied to geological exploration and spatial modeling tasks.

5. References

(1) Buhmann, M. D. *Radial Basis Functions: Theory and Implementations.* Cambridge University Press, Cambridge, 2003.

(2) Fasshauer, G. E. *Meshfree Approximation Methods with MATLAB.* World Scientific, Singapore, 2007.

(3) Mukhamediev, R., Popova, E., et al. Review of artificial intelligence and machine learning technologies. *Mathematics*, 2022.

(4) Shvets, O., Bugubayeva, A., Rakhmetullina, S., Wojcik, W. Development of an intelligent information system for assessing the negative impact of industrial emissions on public health. *Journal of Theoretical and Applied Information Technology*, 2021.

(5) Hardy, R. L. Multiquadric equations of topography and other irregular surfaces. *Journal of Geophysical Research*, 1971, 76(8), 1905–1915.

(6) Franke, R. Scattered data interpolation: Tests of some methods. *Mathematics of Computation*, 1982, 38(157), 181–200.

(7) Kansa, E. J. Multiquadrics—A scattered data approximation scheme with applications to computational fluid dynamics. I: Surface approximations and partial derivative estimates. *Computers & Mathematics with Applications*, 1990, 19(8–9), 127–145.

(8) Lazzaro, D., Montefusco, L. B. Radial basis functions for the multivariate interpolation of large scattered data sets. *Journal of Computational and Applied Mathematics*, 2002, 140(1–2), 521–536.

(9) Tomczak, M., Wróbel, L. C. *The Finite Difference and Meshfree Methods for Solving Partial Differential Equations.* Springer, 2018.

(10) Li, X., Liu, Z., Wu, G. Application of RBF neural networks in geochemical anomaly recognition. *Geoscience Frontiers*, 2019, 10(3), 843–853.

(11) Goovaerts, P. *Geostatistics for Natural Resources Evaluation.* Oxford University Press, Oxford, 1997.

(12) Bishop, C. M. *Pattern Recognition and Machine Learning.* Springer, New York, 2006.

(13) Hastie, T., Tibshirani, R., Friedman, J. *The Elements of Statistical Learning.* Springer, New York, 2009.

(14) Chilès, J.-P., Delfiner, P. *Geostatistics: Modeling Spatial Uncertainty.* Wiley, New York, 2012.

(15) Kikumoto, M., Ooka, R. Numerical study on pollutant dispersion using the RNG k–ε model. *Atmospheric Environment*, 2018, 180, 50–63.

PREPROCESSING OF DATA IN PREDICTING GEOCHEMICAL ANOMALIES USING NEURAL NETWORKS

MOLDIR TANASHOVA[1]

[1]Al-Farabi Kazakh National University, Almaty, Kazakhstan
e-mail: moldir.tanashova@bk.ru

Abstract. Accurate prediction of geochemical anomalies strongly depends on the quality of data preprocessing, since geographic coordinates alone are weakly informative and poorly correlated with element concentrations. In neural-network-based geochemical forecasting, this often leads to overly simplified and compressed predictions that do not reflect the true range of geochemical values.

This study proposes an anisotropy-based data preprocessing approach that accounts for directional spatial variability and enhances the relationship between input features and the target variable. The transformed features are used to train a multilayer perceptron for predicting gold concentrations at unsampled locations.

Numerical experiments demonstrate that anisotropic preprocessing significantly improves the accuracy and stability of geochemical anomaly prediction, confirming the decisive role of data preprocessing in neural-network-based geochemical forecasting.

Keywords: data preprocessing; geochemical anomalies; neural networks; multilayer perceptron; anisotropy.

AMS Subject Classification: The author(s) should provide AMS Subject Classification numbers using the link https://mathscinet.ams.org/msnhtml/msc2020.pdf.

1. Introduction

Modern exploration tasks require not only an increase in data volumes but also the development of advanced data analysis methods capable of accounting for the complex spatial structure of geochemical fields. In this context, the advancement of computational and intelligent technologies aimed at improving the accuracy of geochemical anomaly forecasting is of particular importance.

A number of studies demonstrate the high potential of multilayer perceptrons (MLP) in solving geological forecasting problems. Thus, study [1] shows that the integration of geological, geochemical, and geophysical data within a digital platform enables the identification of spatial patterns of ore bodies and the construction of three-dimensional models of element distribution. At the same time, high forecasting accuracy is achieved under conditions of a pronounced correlation between input features and target characteristics, which emphasizes the critical role of proper data preprocessing and feature structuring.

In study [2], a multilayer perceptron was applied to reconstruct fracture parameters of carbonate reservoirs based on geophysical data. The obtained results confirm the ability of MLP models to efficiently approximate complex nonlinear relationships; however, the performance of the model is largely determined by the use of informative features obtained through data preprocessing and careful input parameter selection.

This research was funded by the Science Committee of the Ministry of Science and Higher Education of the Republic of Kazakhstan (BR) No. 27100483 “Development of predictive exploration technologies for identifying ore-prospective areas based on data analysis from the unified subsurface user platform ”Minerals.gov.kz” using artificial intelligence and remote sensing methods”.

In study [3], a neural network was employed for three-dimensional geological modeling using a large drilling database. Data preprocessing stages, including coordinate normalization, handling of missing values, and integration of heterogeneous data sources, played a decisive role in achieving reliable results. Such preprocessing enabled the identification of spatial structures that are difficult to capture using traditional geostatistical approaches.

Of particular interest is study [4], in which a multilayer perceptron was used to predict the spatial distribution of Cu, Pb, and Zn concentrations in the Hokuroku ore district (Japan). The authors demonstrated that the neural network can reveal hidden spatial patterns even in cases where classical variogram analysis fails to detect significant correlations. This result highlights the potential of neural network models when the spatial structure of the data is adequately represented.

In study [5], neural network methods were compared with geostatistical approaches for assessing the quality of a shale deposit. The results showed that accounting for anisotropy and spatial heterogeneity significantly affects the characteristics of predicted distributions. Moreover, different neural network architectures produce fundamentally different spatial representations, emphasizing the necessity of a well-founded choice of data preprocessing techniques depending on the geological structure of the studied object.

Despite the substantial number of publications devoted to the application of neural networks in geology, many studies still perform spatial prediction by directly mapping geographic coordinates to element concentrations without sufficiently considering the anisotropy of geochemical fields and the spatial organization of the data. Such an approach reduces the informativeness of input features and limits the ability of neural network models to identify latent spatial relationships.

In this regard, an important scientific challenge is the development and analysis of geochemical data preprocessing methods aimed at strengthening the relationship between input and output variables and improving the stability of neural network predictions. This paper focuses on investigating the impact of anisotropic data preprocessing on the accuracy of geochemical anomaly forecasting using multilayer perceptrons.

2. Problem Statement

For the studied territory, the results of geochemical sampling are available for a single mineral commodity, namely gold. Each sampling location is characterized by known geographical coordinates (X_i, Y_i) and an associated gold concentration C_i. The numerical values of the sampling data are summarized in Table 1.

Table 1. Geochemical sampling data and corresponding coordinates

No.	Longitude	Latitude	Grade, g/tonne
1	82.77855	49.96831	0.3639
2	81.80126	49.47704	0.1140
3	81.98683	49.68197	0.3069
⋮	⋮	⋮	⋮
580	82.95850	50.10450	0.3639

The spatial arrangement of the geochemical sampling points is illustrated in Figure 1 by red markers, with darker shading indicating higher values of gold concentration.

In the northwestern part of the study area, an additional grid consisting of 820 points was generated (Figure 1, green markers). A portion of these generated points falls within zones covered by thick unconsolidated deposits, where geochemical sampling data are absent. Since the gold concentrations at these locations are unknown, all generated points are displayed with uniform color intensity.

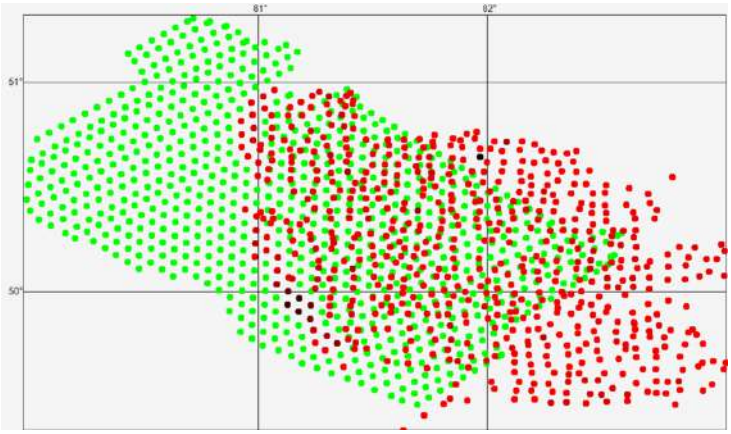


FIGURE 1. Spatial distribution of geochemical sampling locations (red) and generated grid points (green) used to construct prediction maps for the study area.

The objective of this research is to estimate the gold content at the generated grid points by utilizing information from the sampled locations, based on a multilayer perceptron trained on the dataset presented in Table 1.

From a formal perspective, the problem can be described as follows. Let there be a set of sampling locations with known coordinates (X_i, Y_i) and corresponding gold concentrations C_i. A multilayer perceptron (MLP) is used as a nonlinear operator that establishes a functional relationship between the spatial coordinates of a point and the associated gold concentration:

$$\text{MLP} : (X, Y) \to C. \tag{1}$$

The dataset given in Table 1 serves as the training set for the neural network. Consequently, the input layer of the MLP contains two neurons representing the spatial coordinates (X, Y), while the output layer consists of a single neuron corresponding to the predicted gold concentration C. To completely specify the neural network model, it is necessary to define the number of hidden layers and the number of neurons within each hidden layer.

It is important to note that a straightforward mapping between spatial coordinates and geochemical concentrations is typically characterized by weak correlation between the input and output variables. This property complicates the learning process and often results in overly smoothed predictions that fail to capture the true spatial variability of geochemical fields. Therefore, enhancing the informativeness of the input features through appropriate data preprocessing is a key requirement for effectively addressing the stated prediction problem.

3. Feature Transformation Based on an Anisotropic Information Model and Formation of Training and Prediction Sets

Anisotropy in geological and geochemical data is manifested as a dependence of spatial variability on direction. In geochemical studies, this implies that mineral concentrations may vary with different intensities in different spatial directions, which is typically controlled by the geological structure of the area. The physical origin of anisotropy is commonly associated with lithological layering, tectonic dislocations, fracture systems, metamorphic processes, and other structural and textural characteristics of rocks.

In this study, anisotropy is considered in relation to the spatial variability of gold concentration, that is, the dependence of concentration values on the direction of prediction. The input data consist of geochemical sampling results presented in Table 1 and visualized on the spatial distribution map. The color intensity of the sampling points reflects the magnitude of gold concentration, whereas the generated points, for which geochemical measurements are unavailable, are displayed with uniform color.

3.1. **Spatial Problem Statement for Prediction.** The prediction task is formulated for generated points located predominantly in the northwestern direction relative to the area of geochemical sampling. To formally define the geometric boundaries of the study area, convex hulls are constructed for the set of geochemical sampling points and for the set of generated points. This representation makes it possible to clearly identify the overlapping domain and the regions extending beyond the area covered by direct geochemical observations.

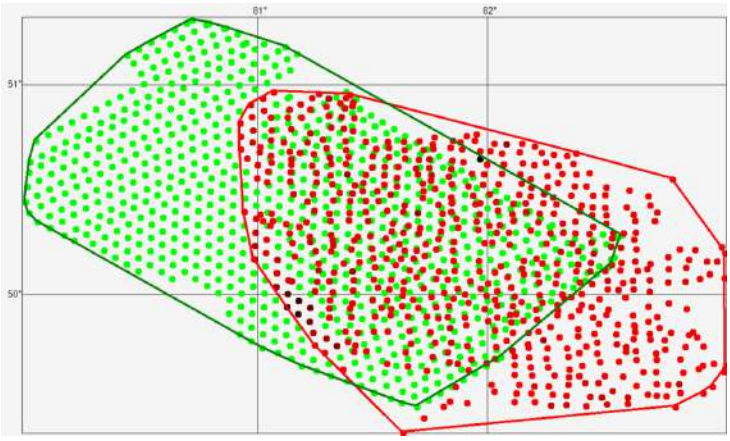


FIGURE 2. Geochemical sampling points (red) and generated points (green) together with their corresponding convex hulls.

Let the following sets be defined:

- R — the set of points with known geochemical sampling results;
- G — the set of generated points;
- $C = R \cup G$ — the union of the two sets;
- $S = R \cap G$ — their intersection;
- $GR = G \setminus R$ — the set of generated points located outside the geochemical sampling domain.

Under these definitions, the problem of forecasting gold concentration at points belonging to the set G can be naturally decomposed into two subproblems. The first subproblem is *interpolation*, which consists in estimating gold concentrations at points located within the spatial domain covered by geochemical sampling data. The second subproblem is *extrapolation*, which involves predicting gold concentrations at points situated outside the geochemical sampling area.

In this study, a unified feature construction and prediction framework is employed, which is applicable to both interpolation and extrapolation tasks.

3.2. **Anisotropy Direction and Selection of the Prediction Point.** To identify the dominant direction of spatial variability, correlation lines are constructed for the set of geochemical sampling points and for the set of generated prediction points. These lines are found to be nearly parallel and therefore define the principal direction of spatial continuity in the data. In the present case, the dominant direction corresponds to a northwestern trend.

For each prediction point, a prediction line is defined as a line perpendicular to the correlation direction and passing through the considered point. This line partitions the plane into two half-spaces. The southeastern half-space contains points with known geochemical sampling results, while the northwestern half-space contains the prediction points.

A prediction point is selected such that a sufficient number of sampling points are located in the southeastern half-space within a distance not exceeding the maximum spacing between neighboring generated points. This condition ensures the availability of representative geochemical information required for constructing a reliable anisotropic model.

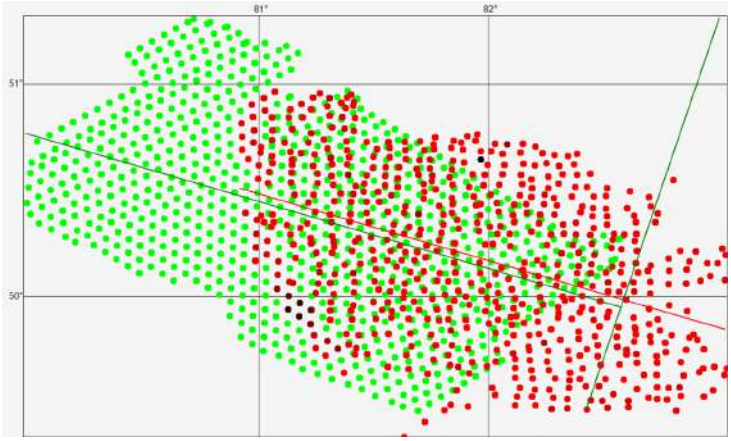


FIGURE 3. Geochemical sampling points and generated prediction points with their corresponding correlation lines.

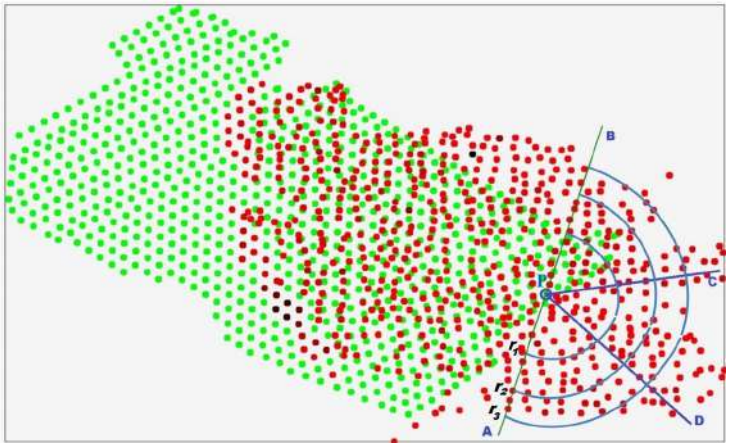


FIGURE 4. Geometric construction for the interpretation of the anisotropic information model converging to the prediction point p.

3.3. **Sector–Radial Decomposition of Space.** For a selected prediction point, the south-eastern half-space is partitioned into N_S angular sectors. Each sector is further subdivided into N_P radial subsectors bounded by arcs of increasing radius. This sector–radial decomposition provides a structured representation of the spatial neighborhood surrounding the prediction point.

The number of sectors and subsectors is selected according to the density of geochemical sampling points, the spatial scale of variability, and the type of the investigated mineral commodity.

Let M_{ij} denote the set of geochemical sampling points that fall into the j-th subsector of the i-th sector. For each subsector, the following quantities are computed:

- the mean gold concentration;
- the weighted mean coordinates of the sampling points.

The mean gold concentration within a subsector is defined as the arithmetic average of concentration values of all sampling points belonging to the corresponding subsector. The weighted mean coordinates are calculated as the average longitude and latitude of the points contained in the subsector.

3.4. **Anisotropic Information Model.** The anisotropic information model converging to a prediction point represents an ordered set of parameters that describes the directional variability of gold concentration as the spatial distance to the prediction point decreases.

The anisotropic information model includes the following components:

(1) the geographic coordinates of the prediction point;
(2) the mean gold concentrations computed for the subsectors;
(3) the weighted mean coordinates of the subsectors.

An example of the anisotropic information model for a single sector is presented in Table 2.

TABLE 2. Anisotropic information model for a single sector

No.	Model element	Symbol	Description
1	Longitude of the prediction point	lon	Geographic longitude
2	Latitude of the prediction point	lat	Geographic latitude
3	Gold concentration at the prediction point	c	Gold concentration value
4	Mean concentration in subsector 1	C_{11}	Averaged gold concentration
5	Mean longitude of subsector 1	Lon_{11}	Geographic longitude
6	Mean latitude of subsector 1	Lat_{11}	Geographic latitude
7	Mean concentration in subsector 2	C_{12}	Averaged gold concentration
8	Mean longitude of subsector 2	Lon_{12}	Geographic longitude
9	Mean latitude of subsector 2	Lat_{12}	Geographic latitude
10	Mean concentration in subsector 3	C_{13}	Averaged gold concentration
11	Mean longitude of subsector 3	Lon_{13}	Geographic longitude
12	Mean latitude of subsector 3	Lat_{13}	Geographic latitude

In contrast to classical anisotropy descriptions based on variogram models, the proposed approach represents anisotropy as an ordered sequence of averaged concentration values and associated spatial coordinates. This sequence captures the directional structure of the geochemical field and provides a compact and informative representation suitable for subsequent neural network modeling.

3.5. **Formation of the Training Dataset for the MLP.** To employ the anisotropic information model in neural network modeling, it is transformed into a set of input features suitable for a multilayer perceptron (MLP). For each sector associated with a prediction point, a single row of the training dataset is constructed.

An example of the format of one training data row is presented in Table 3.

TABLE 3. Format of a single training dataset row

lon	lat	C_{11}	C_{12}	C_{13}	Lon_{11}	Lat_{11}	Lon_{12}	Lat_{12}	Lon_{13}	Lat_{13}	c
81.801	49.477	0.247	0.169	0.417	81.904	49.492	82.109	49.550	82.279	49.600	0.114

In this example, the space surrounding each prediction point is divided into three sectors, and each sector is subdivided into three subsectors. The first eleven parameters in the table constitute the input features of the MLP, while the last parameter represents the output value corresponding to the gold concentration at the prediction point.

For each prediction point, N_S rows of the training dataset are generated, one for each sector. When forming a training dataset for multiple prediction points, the total number of rows is equal to the product of the number of prediction points and the number of sectors. This representation allows the neural network to learn directional patterns of spatial variability encoded in the anisotropic feature structure.

4. Results: Effect of Data Preprocessing

This section illustrates the impact of anisotropic data preprocessing on the quality of geochemical anomaly prediction using multilayer perceptrons. The comparison is carried out on the same spatial domain using two modeling strategies: (i) a baseline MLP trained only on geographic coordinates and (ii) a proposed approach based on anisotropic feature transformation combined with a two-stage MLP architecture.

4.1. **Baseline Model: MLP without Preprocessing.** In the baseline approach, a multilayer perceptron is trained using only geographic coordinates (longitude and latitude) as input features, while gold concentration is used as the target variable. No anisotropic transformation or spatial feature engineering is applied. Consequently, the network attempts to learn the mapping $(lon, lat) \mapsto C$ directly from raw coordinates.

The baseline prediction results are presented in Figure 5. In this case, the predicted field becomes overly simplified and strongly smoothed. The spatial contrast is reduced, and the predicted values tend to be "compressed", which indicates that coordinate-only inputs are weakly informative for representing complex spatial patterns in geochemical data.

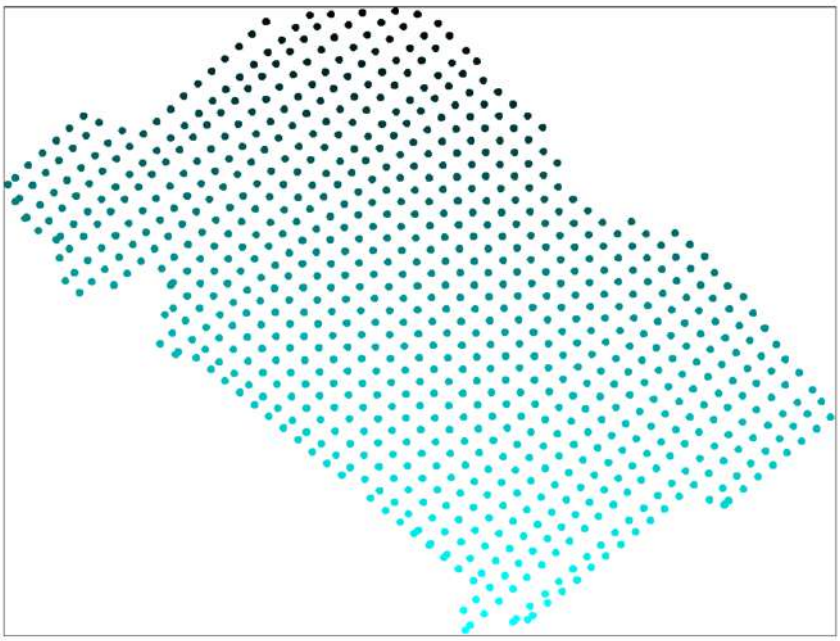


Figure 5. Prediction results obtained using a baseline MLP trained only on geographic coordinates (longitude and latitude) without preprocessing. The output is overly smoothed and exhibits reduced spatial contrast.

4.2. **Proposed Approach: Anisotropic Preprocessing and Two-Stage MLP.** In the proposed method, the original sampling data are first transformed using the anisotropic information model described in Section 3. For each prediction point, the southeastern half-space is decomposed into angular sectors and radial subsectors, and aggregated directional features (sector–subsector means and weighted coordinates) are computed. These features enhance the relationship between the inputs and the target variable by explicitly encoding spatial anisotropy.

The transformed features are then used in a two-stage MLP scheme. At the first stage, the network produces sector-level estimates of gold concentration, capturing directional variability. At the second stage, sector-wise outputs are aggregated to obtain a single final concentration estimate for each prediction point.

The resulting prediction pattern is shown in Figure 6. The map demonstrates increased spatial contrast and better preservation of local heterogeneity, which is consistent with the expected structure of geochemical anomaly fields.

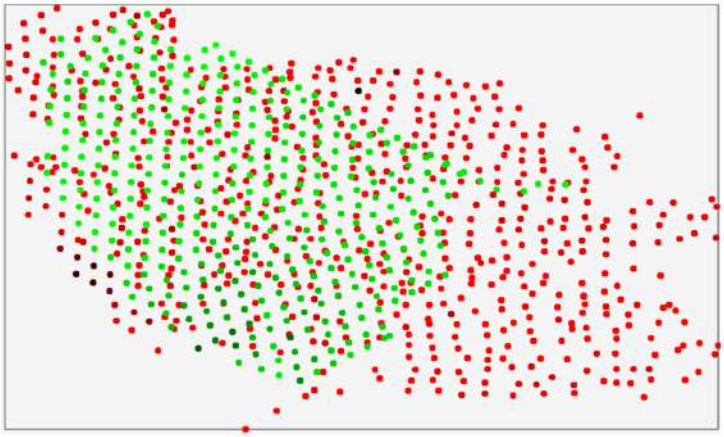


FIGURE 6. Prediction results obtained using anisotropic data preprocessing and a two-stage MLP. The model preserves spatial heterogeneity and highlights local anomaly-like variations.

A qualitative comparison of Figures 6 and 5 demonstrates the decisive role of preprocessing in neural-network-based geochemical forecasting. The anisotropic transformation increases input informativeness and enables the neural network to capture directional spatial variability, resulting in more structured and geologically meaningful prediction patterns. In contrast, the coordinate-only baseline model is prone to excessive smoothing and loss of local variations.

5. CONCLUSIONS

This study demonstrates the importance of data preprocessing in neural-network-based geochemical anomaly prediction. The results show that when multilayer perceptrons are trained on raw spatial coordinates alone, the predictive capability of the model is significantly limited, leading to overly simplified and weakly informative results.

To overcome this issue, an anisotropy-based data preprocessing strategy was proposed. By introducing the concept of anisotropy converging to a prediction point and transforming the spatial neighborhood into structured directional features, the relationship between input variables and predicted outputs is significantly strengthened. This transformation increases the expressive power of the training dataset and enables the neural network to learn spatial dependencies that are not accessible through raw coordinates alone.

Overall, the study highlights the decisive role of data preprocessing in neural-network-based geochemical forecasting and shows that appropriate feature transformation is a prerequisite for achieving reliable and informative predictions of geochemical anomalies.

6. ACKNOWLEDGMENTS

This research was funded by the Science Committee of the Ministry of Science and Higher Education of the Republic of Kazakhstan (No. BR27100483 “Development of predictive exploration technologies for identifying ore-prospective areas based on data analysis from the unified subsurface user platform ”Minerals.gov.kz” using artificial intelligence and remote sensing methods”).

References

[1] Los V.L., Legonkin V.S., Development of mathematical and software tools for forecasting ore mineral deposits, *Geology and Subsoil Protection*, 2023, 1(86), 43–54.

[2] Pei J., Zhang Y., Prediction of reservoir fracture parameters based on the multilayer perceptron machine-learning method: A case study of Ordovician and Cambrian carbonate rocks in the Nanpu Sag, Bohai Bay Basin, China, *Processes*, 2022, 10(11), 2445.

[3] Ji Y., Kim H.-S., Lee M.-G., Cho H.-I., Sang C.-G., 3D mapping of geotechnical layers based on MLP using a borehole database in Seoul, South Korea, *Journal of the Korean Geotechnical Society*, 2021, 37(5), 1–18.

[4] Iwashita T., Ohnishi Y., Takahashi Y., 3D prediction of ore grade distribution using multilayer perceptron neural networks: A case study in the Hokuroku Basin, Japan, *Natural Resources Research*, 2021, 30(2), 1863–1880.

[5] Torres J., Lamelas M.T., Carranza E.J.M., Comparison of neural networks and geostatistical methods for predicting shale quality parameters in northern Spain, *Journal of Petroleum Science and Engineering*, 2020, 192, 107320.

СОДЕРЖАНИЕ