DOI: 10.20535/1970.68(2).2024.318208
UDC 53.05: 617.753

OPTIMIZATION METHODS FOR PARAMETER IDENTIFICATION MODEL OF TEST ELECTRORETINOSIGNAL TO ASSESS NEUROTOXICITY RISKS

[1)]Pavlo Tymkiv, [1)]Roman Tkachuk, [2)]Oleksiy Yanenko
[1)]Ternopil Ivan Puluj National Technical University, Ternopil, Ukraine
[2)]National Technical University of Ukraine
"Igor Sikorsky Kyiv Polytechnic Institute", Kyiv, Ukraine
E-mail: t_pavlo_o@ukr.net; romantkachuk48@gmail.com; op291@meta.ua

Introduction. *The development of advanced optimization methods plays a crucial role in the enhancement of diagnostic tools in the biomedical field, particularly in the analysis of complex physiological signals. Electroretinography (ERG) is a widely used diagnostic technique that records electrical responses generated by the retina in response to light stimuli, providing valuable insights into the functional health of retinal cells. ERG is instrumental in diagnosing conditions such as retinitis pigmentosa, diabetic retinopathy, and neurotoxicity. However, the analysis of low-intensity electroretinograms (ERG) presents numerous challenges, particularly due to noise and signal distortion, which complicate accurate signal interpretation.*

Main purpose of this study*.This paper is dedicated to developing an expert system for real-time analysis of electroretinographic signals (ERS), focusing on optimizing the parameters of a mathematical model for ERS analysis in conditions where noise and other distortions are present. The primary aim is to improve the accuracy and efficiency of ERG data processing, enabling early detection of neurotoxicity and other retinal conditions. To achieve this, we applied advanced optimization techniques, such as the Nelder-Mead method, known for its effectiveness in handling non-smooth, noisy functions.*

Conclusions. *1. The application of the Nelder-Mead algorithm for optimizing the complex and noisy ERS model significantly improved the performance of ERG data analysis. The algorithm's adaptability to varying optimization conditions allowed for more accurate model parameter determination, particularly in the context of real-time neurotoxicity detection.*

2. Reduction in Processing Time: The time complexity analysis revealed that the Nelder-Mead method reduced the time required to compute the model coefficients by approximately 15%. This improvement was achieved while maintaining the necessary precision for reproducing the test electroretinosignal, making it suitable for real-time applications.

3. Computational Efficiency: One of the key findings of this study is that the use of the Nelder-Mead algorithm reduced the computational load by up to 30%. This makes the method feasible for use in expert systems designed for real-time ERS analysis, allowing for the monitoring of functional changes in the retina during the early stages of neurotoxicity detection.

Keywords: *electroretinogram, low intensity, neurotoxicity, optimization, parametric identification..*

Introduction. Formulation of the problem

Electroretinography (ERG) is a diagnostic method that measures the electrical responses of various retinal cell types, including photoreceptors (rods and cones), bipolar cells, and ganglion cells, to light stimuli. The procedure involves placing electrodes on the cornea to detect the electrical activity generated by the retina in response to light stimuli. This activity is recorded as a waveform, which reflects the functionality of different layers of retinal cells [1].

ERG is particularly important in ophthalmology because it provides detailed information about the health and functioning of the retina, enabling the diagnosis and monitoring of various visual disorders. For example, in conditions such as retinitis pigmentosa — a group of genetic disorders leading to progressive retinal degeneration – ERG can detect functional decline in photoreceptors at early stages, even before clinical symptoms appear [2]. Similarly, in cases of elevated intraocular pressure, characterized by damage to the optic nerve, ERG can help assess the functional integrity of retinal ganglion cells, which are often affected in the early stages of the disease.

Additionally, ERG is a critical tool in detecting diabetic retinopathy, a common complication of diabetes that affects the retinal blood vessels. Through regular ERG testing, ophthalmologists can monitor retinal function over time, enabling timely interventions to prevent or mitigate vision loss. ERG also plays a crucial role in assessing drug toxicity, particularly for medications known to have potential side effects on the retina, helping to detect neurotoxic effects at an early stage [3-4].

The advancement of computer technology and digital analysis has significantly expanded the capabilities of ERG. Modern ERG systems utilize photostimulators, such as xenon lamps and LEDs, to deliver controlled light stimuli with specific intensities and wavelengths. This technological advancement allows for standardized and reproducible testing conditions, which are essential for accurate diagnostics and longitudinal studies. Digital signal processing techniques facilitate automated data recording, enabling precise measurements of retinal responses, including key morphological parameters like amplitude and latency of the waveform components.

Overall, the integration of ERG into clinical practice has revolutionized the approach to early-stage retinal disease diagnosis. By providing a comprehensive assessment of retinal function, ERG not only aids in the early detection of retinal pathology but also deepens the understanding of the underlying mechanisms of various retinal disorders. This, in turn, facilitates the development of targeted therapies with the potential to improve patient outcomes through more personalized and timely medical interventions.

Preliminary information

However, the study of electroretinograms (ERGs), particularly at low light intensities, presents several challenges that can affect the accuracy and reliability of the results. Firstly, the signal-to-noise ratio in ERGs can be quite low, especially under conditions of low illumination. This means that the actual signal from the retina may be masked by background electrical activity or noise from other sources, making it difficult to isolate the important characteristics of the ERG. Informative parameters such as amplitude and latency can be distorted by this noise, complicating their analysis and interpretation.

Secondly, artifacts caused by eye movements, blinking, or even slight shifts in the positioning of electrodes can degrade the ERG signal. These artifacts, under certain conditions and circumstances, can distort the signal, leading to inaccurate conclusions if not properly identified and addressed. The presence of such artifacts requires the implementation of sophisticated filtering and signal processing techniques to clean the data before analysis [5].

Moreover, the ERG signal itself is complex and comprises multiple components that reflect the activity of different retinal layers and cell types. Analyzing these components requires advanced mathematical and computational tools to decompose the signal into its constituent parts and accurately attribute them to specific retinal functions. This process can be further complicated in pathological conditions where normal signal patterns are disrupted.

The variability of ERG responses between patients adds yet another layer of complexity. Factors such as age, individual retinal characteristics, and overall health can cause significant differences in ERG readings, making it difficult to establish standard reference values. These inter-individual variations mean that large datasets are necessary to improve modern models and algorithms capable of accounting for these differences and producing reliable diagnostic information when developing an expert system prototype

Figure 1 shows a real electroretinogram without preprocessing.

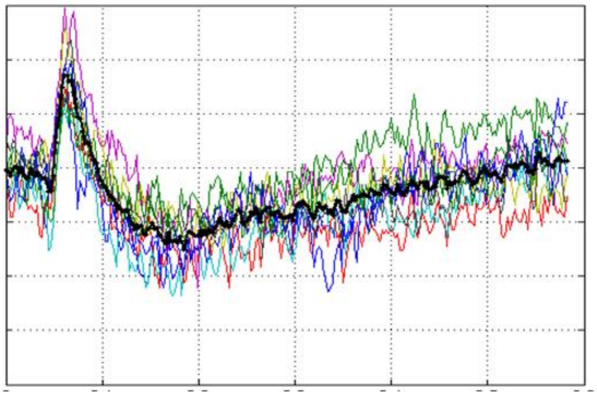

Fig.1. Electroretinogram without proper processing: on the ordinate axis - the price of division is 100 μV, on the abscissa axis - 50 ms.

The volume of data generated during ERG recordings can vary significantly, with a single test producing thousands of data points for each eye. Analyzing this data requires powerful computational resources and efficient processing algorithms. In this context, the use of machine learning and artificial intelligence is becoming increasingly important, offering potential solutions for automating the analysis and visualization of ERG data.

Interpreting ERG results for complex pathological conditions or subtle functional changes, such as those occurring in the early stages of neurotoxicity detection, demands a deep understanding of both retinal physiology and the mechanisms behind the formation of electroretinographic signals. This often involves integrating ERG data with other diagnostic modalities and clinical information to achieve a comprehensive assessment of retinal health.

Main purpose of this study

The goal of this study is to enhance the mathematical model of ERG, analyze existing methods and algorithms for processing low-intensity noisy signals, and develop the foundations for an expert information-measurement system for studying changes in the functional state of the visual analyzer during the early stages of neurotoxicity detection.

This work proposes an improved ERG mathematical model and outlines the core principles for building a prototype of an expert information-measurement system to study changes in the functional state of the visual analyzer during the early stages of neurotoxicity detection. These results aim to improve the diagnostic utility of ERG by enhancing the accuracy, reliability, and interpretability of real-world data.

Figure 2 shows the structure of the information and measurement system (IMS) of the expert system for assessing neurotoxicity.

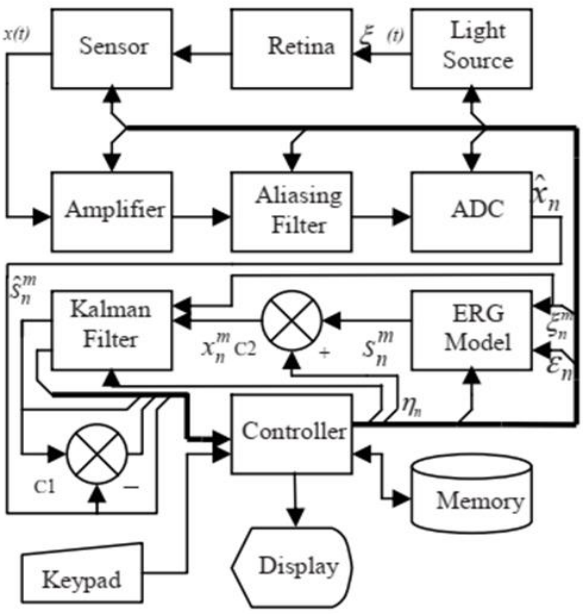

Fig. 2. Structure of the Information and Measurement System (IMS) of the Expert System for Neurotoxicity Assessment

To determine the parameters of the mathematical model (coefficients of the difference equation), the method of direct exhaustive search was employed. This method guarantees a predefined level of accuracy and convergence but requires substantial computational time.

The real-time processing of ERG data by the expert system ensures its use for remote, automated, real-time monitoring of the eye and visual system, particularly in various toxicological contexts [6]. To facilitate its operation, it is necessary to reduce the duration of the research procedures and develop processing algorithms to achieve the set goals.

The real-time processing of ERG data by the In recent studies, advanced algorithms such as the Hooke-Jeeves method and the Conjugate Gradient method have been used for the parametric identification of electroretinographic signal models.

Figure 3 shows an electroretinogram after optimal processing with a higher and lower level of neurotoxicity.

The Hooke-Jeeves method is known for its simplicity and effectiveness, particularly in cases where the function lacks an analytical derivative or is non-differentiable. This method involves exploratory and pattern moves to find the function's minimum by gradually refining the search direction. However, its efficiency and computational complexity significantly depend on the initial approximation conditions [7, 8].

The Conjugate Gradient method is an effective algorithm for minimizing quadratic functions in nonlinear spaces. It iteratively updates the search direction using gradient information to accelerate convergence.

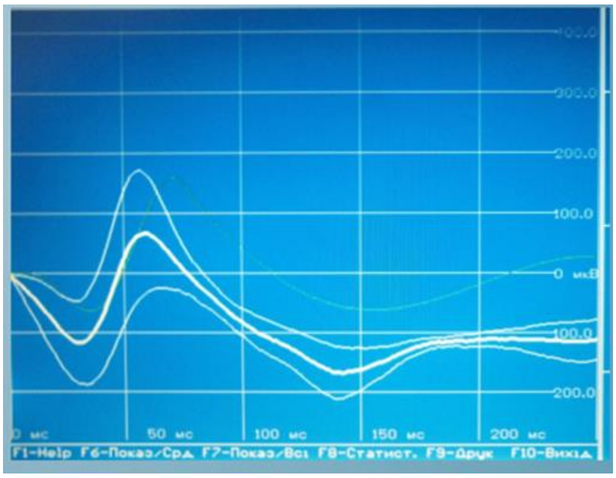

Fig. 3. Electroretinogram for the control dose (thick line) and ERG with a higher and lower level of neurotoxicity (after its optimal processing): on the ordinate axis - the price of division is 100 μV, on the abscissa axis - 50 ms.

Despite its efficiency, this method has certain limitations: its effectiveness and time complexity depend on the initial approximation conditions, and additional processing is required for effectively handling initial constraints.

Research objects and research methodology

Given the limitations of the Hooke-Jeeves and Conjugate Gradient methods, more efficient optimization approaches are proposed. Among these are the Broyden-Fletcher-Goldfarb-Shanno (BFGS) method and the Nelder-Mead method.

The BFGS method is an iterative algorithm designed for minimizing functions with nonlinear constraints. The core idea of this method is to approximate the Hessian matrix (which estimates the second derivatives of the function) using a quasi-Newton approach. This approximation significantly reduces computational load compared to exact Hessian calculations. However, the BFGS method also requires storing the Hessian matrix approximation at each iteration, which can increase memory demands, and the choice and determination of the initial approximation can significantly affect the method's speed and convergence [9].

The BFGS update rule for the Hessian matrix Hk at iteration k is mathematically expressed as:

$$H_{k+1} = H_k + \frac{y_k y_k^T}{y_k^T S_k} - \frac{H_k s_k s_k^T H_h}{s_k^T H_k s_k}, \qquad (1)$$

where:

$$s_k = x_{k+1} - x_k \qquad (2)$$

and

$$y_k = \nabla f(x_{k+1}) - \nabla f(x_k). \quad (3)$$

The well-known Nelder-Mead method is one of the most widely used derivative-free optimization techniques. It operates through an iterative process where a set of points, called a simplex, is modified at each iteration to search for the minimum or maximum of a function. The core idea of the Nelder-Mead method lies in gradually expanding or contracting the simplex in the direction of optimization based on function comparisons at different points.

The Nelder-Mead algorithm involves several key steps: reflection, expansion, contraction, and shrinkage. These steps are used to transform the simplex in the function's landscape to converge toward an optimal solution. For example, the reflection step is mathematically represented as:

$$x_r = x_c + \alpha(x_c - x_h), \quad (4)$$

where, x_c – is the centroid of the simplex, excluding the worst point x_h, and α is the reflection coefficient.

The convergence of optimization algorithms, such as the Nelder-Mead and BFGS methods, can be determined based on several criteria:

- gradient norm – for derivative-based methods like BFGS, convergence can be assessed using the norm of the gradient. If $\|\nabla f(x_k)\|$ falls below a predefined threshold, the algorithm is considered to have converged.
- objective function value – convergence can also be evaluated by monitoring changes in the objective function's value. If the difference $|f(x_{k+1}) - f(x_k)|$ is below a certain acceptable deviation, the algorithm is deemed convergent.
- parameter change – in derivative-free methods like Nelder-Mead, convergence can be indicated by changes in parameter values. If the simplex becomes sufficiently small, such that $\|x_{(i+1)} - x_i\|$ falls below a certain threshold for all points in the simplex, convergence is achieved.

The convergence of these algorithms depends on the nature of the objective function, initial conditions, and specific problem constraints. Selecting appropriate convergence criteria and considering the task's context allows for maximizing the optimization process's efficiency.

The BFGS method is an iterative algorithm for minimizing functions subject to nonlinear constraints. Its core idea lies in approximating the quasi-Newton Hessian matrix, which estimates second derivatives of functions. On the other hand, the Nelder-Mead method is one of the most popular derivative-free optimization methods. It utilizes an iterative process, where the simplex (a geometric shape in the parameter space) is gradually expanded or contracted in the direction of optimization based on function comparisons at different points. Evaluating the time and hardware complexity helps determine the suitability of the BFGS method for large-scale tasks and real-time processing.

The evaluation of the time and hardware complexity of optimization algorithms like the Nelder-Mead method involves several aspects:

a) time complexity – the time complexity of the Nelder-Mead method is characterized by its adaptive nature and lack of explicit gradients. However, it is often described as having a worst-case time complexity of $O(n^2)$, where n – is the number of parameters or the dimensionality of the parameter space. The method iteratively adjusts the simplex to minimize or maximize the objective function, and each iteration requires evaluating the function at several points in the simplex. This contributes to the overall time complexity.

b) hardware complexity – similar to the BFGS method, the Nelder-Mead method does not require specialized hardware and can be executed on standard computing devices. However, memory requirements may vary depending on the parameter space's size and the number of iterations needed for convergence. The method primarily uses memory to store the vertices of the simplex and function evaluations, which can increase with higher dimensional parameter spaces.

c) memory requirements for the Nelder-Mead method primarily depend on the dimensionality of the parameter space and the number of iterations. Although the method does not explicitly compute gradients or Hessian matrices, it still needs to store function evaluations and simplex vertices. As the number of parameters increases, memory usage may rise, potentially becoming a limiting factor for large-scale optimization problems.

Assessing the time and hardware complexity of the Nelder-Mead method helps evaluate its suitability for optimization tasks and provides insights into resource requirements for various computational problems.

The convergence of optimization algorithms, such as the Nelder-Mead and BFGS methods, typically involves monitoring specific criteria throughout the algorithm's iterations [10]. Here's how convergence can be evaluated for each method:

I. objective function value – track the objective function's value at each iteration. If the function value gradually decreases or stabilizes within a predefined tolerance, it suggests convergence.

II. simplex contraction – check if the simplex contracts toward a minimum or maximum point. Convergence occurs when the simplex contracts to a small size, indicating that the optimum is approached.

III. parameter change – monitor changes in the parameter values from one iteration to the next. Convergence is achieved when the parameter values stabilize into a solution within the specified tolerance.

These criteria provide information on how effectively the algorithms converge toward the optimal solution during the optimization process.

The Nelder-Mead algorithm is a promising derivative-free optimization method widely applied in various fields of science and bioengineering. Its primary advantage lies in its ability to handle functions that may be irregular, non-smooth, or noisy.

Below is a detailed justification of the benefits of this algorithm and an analysis of its convergence rate.

Figure 4 shows the block diagram of the Nelder-Mead simplex algorithm.

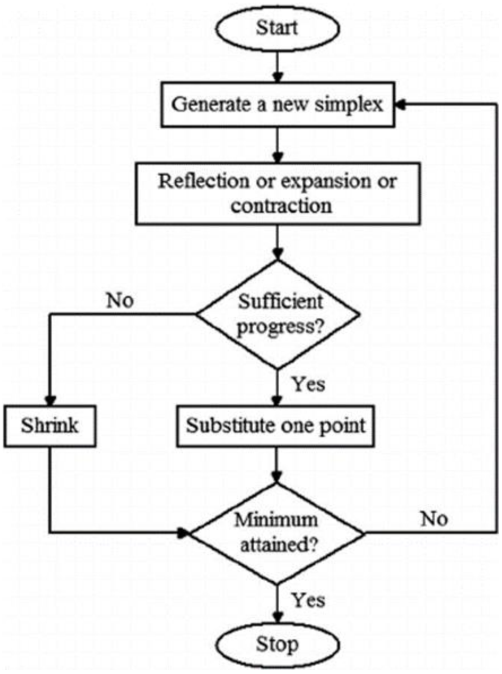

Fig. 4. Flowchart of Nelder-Mead simplex algorithm

The Nelder-Mead algorithm does not require the calculation of function derivatives, making it suitable for optimizing functions where derivative computation is complex or impossible.

An example of determining convergence in the MATLAB environment is shown in Fig. 5.

This is a significant advantage in problems where the function has a complicated or unknown analytical form. The algorithm uses only function values to modify the simplex (a polytope in n-dimensional space), making it appealing for practical applications, including those with multiple local minima. Its flexibility allows the algorithm to adapt to various types of functions and problem geometries. The algorithm efficiently operates in high-dimensional spaces since it does not require the construction of global function models or complex matrix operations. As the method does not use derivatives, it is less sensitive to noise in function values, making it useful for problems where function measurements or estimates contain noise (in our case, during the registration of low-intensity electroretinograms).

The convergence speed of the Nelder-Mead algorithm depends on several factors, including the initial simplex configuration, the nature of the optimized function, and the stopping criteria [11].

```
% Define the objective function for Nelder-Mead optimization
fun = @(X) calculateMSD(X, s, x);

% Initial guess for optimization
X0 = [b1o; b2o];

% Run Nelder-Mead optimization
options = optimset('Display','iter');
tic;
[X, fval, exitflag, output] = fminsearch(fun, X0, options);
elapsedTime = toc;

% Check convergence
if exitflag > 0
    disp('Nelder-Mead method converged successfully.');
else
    disp('Nelder-Mead method did not converge.');
end

% Display optimal coefficients and search time
b1o_nelder = X(1);
b2o_nelder = X(2);
disp('Optimal coefficients using Nelder-Mead method:');
disp(['b1o: ', num2str(b1o_nelder)]);
disp(['b2o: ', num2str(b2o_nelder)]);
disp(['Elapsed time: ', num2str(elapsedTime), ' seconds.']);
```

Fig. 5. Convergence determination in MATLAB environment

While it is quite effective for many practical problems, it does have limitations.

The Nelder-Mead algorithm exhibits linear convergence, meaning that the speed of approaching the optimal solution increases linearly with the number of iterations. This makes the method less effective for tasks that require high precision. The initial shape and size of the simplex significantly impact the convergence rate. A poorly chosen simplex can lead to slow convergence or even getting trapped in local minima. The time complexity of this algorithm depends on the dimensionality of the parameter space n. At each iteration, the method requires function evaluation at n+1 points, resulting in an O(n) complexity per iteration. However, the total number of iterations can be substantial for problems with high dimensionality or complex topology.

Modern models, algorithms, and novel approaches for processing ERG data have been applied, and the fundamental principles of constructing a prototype expert information-measurement system have been developed to study changes in the functional state of the human visual system at the early stages of detecting neurotoxicity.

Conclusions

1. The Nelder-Mead method, known for its versatility and reliability, was applied for optimizing

complex, non-uniform, or noisy objective functions (such as the electroretinosignal (ERS) model in low-intensity electroretinography). Its adaptability to various optimization conditions allows it to be used in the development of expert systems for neurotoxicity risk identification tasks.

2. The evaluation of the time complexity (computation time comparison) of the Nelder-Mead algorithm showed that the time required to determine the model coefficients was reduced by an average of 15%, considering the required accuracy for reproducing the test electroretinosignal.

3. The use of the Nelder-Mead algorithm enabled an average reduction of 30 % in computational power consumption, making it feasible for application in expert information-measurement systems for real-time ERS analysis and the registration of functional changes that occur during the early stages of detecting neurotoxicity.

References

[1] E. E. Cornish, A. Vaze, R. V. Jamieson, and J. R. Grigg, "The electroretinogram in the genomics era: outer retinal disorders" *Eye*, vol. 35(12), pp. 2406-2418, 2021. DOI: 10.1038/s41433-021-01659-y.

[2] D. L. McCulloch, M. F. Marmor, M. G. Brigell, R. Hamilton, G. E. Holder, R. Tzekov, and International Society for Clinical Electrophysiology of Vision, 2015. ISCEV Standard for full-field clinical electroretinography (2015 update). *Documenta Ophthalmologica*, vol. 130, no. 1, pp.1-12. DOI: 10.1007/s10633-014-9473-7

[3] A. J. Tatham, and F. A. Medeiros, "Detecting Structural Progression in Glaucoma with Optical Coherence Tomography", *Ophthalmology*, vol. 124(4S), pp. S57-S65, 2017. DOI: 10.1016/j.ophtha.2017.07.015

[4] A. G. Robson, J. Nilsson, S. Li, S. Jalali, A. B. Fulton, A. P. Tormene, G. E. Holder, and S. E. Brodie, "ISCEV guide to visual electrodiagnostic procedures", *Documenta Ophthalmologica*, 136(1), pp.1-26, 2018. DOI: 10.1007/s10633-017-9621-y

[5] R. A. Tkachuk, B. I. Yavorsky, and O. P. Yanenko, "Problems of neurotoxicity assessment with using of electroretinography", *Visnyk NTUU KPI Seriia - Radiotekhnika Radioaparatobuduvannia*, 0(61), pp. 108-115, 2015. DOI: 10.20535/RADAP.2015.61.108-115.

[6] P. Tymkiv, and M. Bachynskiy, "Assessing neurotoxicity risk through electroretinography with reduced light irritation intensity", *Scientific Journal of TNTU*, 111(3), pp. 58-66, 2023. DOI: 10.33108/visnyk_tntu2023.03.058

[7] P. Tymkiv, "Analysis of the Complexity of Algorithms for Finding the Coefficients of the Mathematical Model of Low-Intensity Electroretinosignal," in *Advanced Applied Energy and Information Technologies 2021. Proc. of the International Conference*, Ternopil, 15-17 December 2021, pp.145-150.

[8] Tymkiv, P., Kłos-Witkowska, A. and Andrushchak, I., "Optimization Methods for Determining Coefficients of Mathematical Model of Electroretinosignal for Detection of Neurotoxicity Risks," in *Proc. of the 1st International Workshop on Computer Information Technologies in Industry 4.0* (CITI 2023). Ternopil, Ukraine, June 14-16, 2023, pp.109-116, 2023.

[9] R. Byrd, G. M. Chin, W. Neveitt, J. Nocedal, "On the use of stochastic Hessian information in unconstrained optimization," *SIAM Journal on Optimization*, vol. 21, is. 3, pp. 977-995, 2011. https://optimization-online.org/2010/06/

[10] Reza Barati, "Parameter Estimation of Nonlinear Muskingum Models Using Nelder-Mead Simplex Algorithm," *Journal of Hydrologic Engineering*, vol. 16, no 11, pp. 946-954, 2011. DOI: 10.1061/(ASCE)HE.1943-5584.0000379.

[11] S Takenaga, Y Ozaki, M Onishi, "Practical initialization of the Nelder–Mead method for computationally expensive optimization problems," *Optimization Letters*, Springer, 17:283–297, 2023. https://doi.org/10.1007/s11590-022-01953-y.

УДК 53.05: 617.753

[1)]Павло Тимків, [1)]Роман Ткачук, [2)]Олексій Яненко
[1)]Тернопільський національни технічний університет імені Івана Пулюя, Тернопіль, Україна
[2)]Національний технічний університет України «Київський політехнічний інститут імені Ігоря Сікорського», Київ, Україна

МЕТОДИ ОПТИМІЗАЦІЇ ІДЕНТИФІКАЦІЇ ПАРАМЕТРІВ МОДЕЛІ ТЕСТОВОГО ЕЛЕКТРОРЕТИНОСИГНАЛУ ДЛЯ ОЦІНЮВАННЯ РИЗИКІВ НЕЙРОТОКСИКАЦІЇ

Стаття присвячена дослідженню та розробці оптимізаційних методів для моделювання електроретинографічних сигналів (ЕРС) у контексті виявлення ранніх стадій нейротоксикації. Електроретинографія (ЕРГ) є важливим інструментом для оцінки функціонального стану сітківки ока, що дозволяє діагностувати різноманітні зорові патології, включно з нейротоксикацією на різних етапах. Однак, низька інтенсивність стимуляції та шум у сигналах можуть значно ускладнювати процес обробки даних та їх інтерпретації.

У статті запропоновано використання методу Нелдера-Міда для оптимізації параметрів математичної моделі ЕРС. Цей метод оптимізації без похідних є ефективним для складних та нерівномірних цільових функцій, що робить його перспективним для застосування в задачах, пов'язаних з аналізом низькоінтенсивних сигналів. Було проведено оцінку часової складності алгоритму, що показала скорочення часу визначення коефіцієнтів моделі на 15% при забезпеченні необхідної точності відтворення тестового електроретиносигналу. Крім того, використання алгоритму Нелдера-Міда дозволило зменшити обчислювальні витрати на 30%, що робить його придатним для реального часу.
У дослідженні також описано структуру прототипу експертної інформаційно-вимірювальної системи для моніторингу змін у функціональному стані зорової системи під впливом нейротоксичних чинників. Результати можуть бути застосовані у медичній практиці для діагностики нейротоксикації та інших патологічних станів на ранніх етапах, що дозволить поліпшити якість медичних втручань та своєчасно запобігати серйозним наслідкам для здоров'я пацієнтів.
Ключові слова: електроретиносигнал, низька інтенсивність, нейротоксикація, оптимізація, параметрична ідентифікація.

*Надійшла до редакції
20 вересня 2024 року*

*Рецензовано
10 жовтня 2024 року*