UDC 004.85
DOI https://doi.org/10.24144/2616-7700.2026.48(1).153-161

D. M. Krasii[1], O. O. Larin[2]

[1] National Technical University "Kharkiv Polytechnic Institute",
PhD student of the Department of Mathematical Modeling and Intelligent Computing in Engineering
Danylo.Krasii@infiz.khpi.edu.ua
ORCID: https://orcid.org/0000-0003-1599-2295

[2] National Technical University "Kharkiv Polytechnic Institute",
Professor at the Department of Mathematical Modeling and Intelligent Computing in Engineering,
Doctor of Technical Sciences, Professor
Oleksiy.Larin@khpi.edu.ua
ORCID: https://orcid.org/0000-0002-5721-4400

PHYSICS-GUIDED CHAINED GAUSSIAN PROCESS REGRESSION FOR ATHEROSCLEROTIC PLAQUE PROGRESSION PREDICTION

Predicting atherosclerotic plaque progression from sparse cross-sectional imaging data is critical for cardiovascular risk assessment. Standard regression methods lack physics-based inductive bias, leading to unreliable extrapolation. A Physics-Guided Chained Gaussian Process Regression (PG-CGPR) framework is presented that incorporates Power-Law Damage Accumulation mechanics as a mean function prior with heteroscedastic noise modeling. Cross-sectional observations are transformed into pseudo-longitudinal training pairs using Glagov-tolerant filtering. Validated on cardiovascular imaging data spanning multiple decades, the model maintains predictive accuracy even when extrapolating beyond the training age range, with confidence intervals covering all test observations. The strong physics prior dominates data-driven learning, demonstrating that in small-sample regimes, mechanistic models outperform flexible kernels. The framework enables projection of future vessel geometries for downstream biomechanical analysis.

Keywords: Gaussian process regression, plaque progression, physics-informed machine learning, uncertainty quantification, cardiovascular imaging.

1. Introduction. Atherosclerotic plaque rupture is the primary mechanism underlying acute coronary syndromes and ischemic stroke. The mechanical stability of arterial plaque depends critically on its geometric configuration: fibrous cap thickness, lipid core size, and overall burden determine stress distribution under physiological loading. Predicting how plaque geometry evolves over time is therefore essential for identifying high-risk patients before acute events occur. While current clinical practice relies on cross-sectional imaging, such as intravascular ultrasound (IVUS) or optical coherence tomography (OCT), to provide snapshots of plaque state, the lack of routine longitudinal follow-up creates a fundamental challenge: estimating temporal plaque trajectories from population-level cross-sectional observations.

Standard regression approaches often fail in this setting due to regression toward the prior mean in extrapolation regions, violating physical monotonicity constraints and producing overconfident predictions. To address this, Physics-Informed Machine Learning (PIML) has emerged as a powerful paradigm, embedding physical laws directly into the learning objective to constrain solutions where data is sparse [1, 2].

The problem of learning damage trajectories from sparse data is addressed in this work by proposing a Physics-Guided Chained Gaussian Process Regression

(PG-CGPR) framework. The problem is defined as follows: given cross-sectional observations $\{(t_i, y_i)\}_{i=1}^N$ from N individuals at different ages t_i, the objectives are (1) to learn a population-level damage trajectory $D(t)$ with uncertainty quantification, and (2) to predict patient-specific plaque growth given a baseline measurement. The contributions of this work are:

- A Physics-Guided GPR architecture incorporating Power-Law Damage Accumulation mechanics as a mean function prior, ensuring physically consistent monotonic predictions.
- Heteroscedastic noise modeling that captures increasing biological variability with age.
- Empirical demonstration that strong physics priors can dominate data-driven learning in small-sample regimes.

2. Background and Related Work.

2.1. Gaussian Process Regression and Heteroscedasticity. Gaussian Process Regression (GPR) provides a Bayesian framework for regression with uncertainty quantification [3]. Given observations $\mathcal{D} = \{(t_i, y_i)\}_{i=1}^N$, the model assumes $y_i = f(t_i) + \epsilon_i$, where $f \sim \mathcal{GP}(m(t), k(t, t'))$. While standard GPR assumes constant noise variance, biological processes often exhibit heteroscedasticity, i.e., uncertainty that varies with the state of the system. Chained Gaussian Processes [4] address this by modeling input-dependent noise using two latent functions:

$$y_i = f_1(t_i) + \epsilon_i \cdot g(f_2(t_i))$$

where f_1 models the mean trajectory and f_2 models log-variance. Scaling GPR to larger datasets often requires sparse approximations using inducing points and variational inference [5], which is employed here to handle the non-conjugate likelihood of the Chained GP.

2.2. Continuum Damage Mechanics. The continuum damage mechanics framework describes damage evolution in materials under sustained loading. The Power-Law Damage Accumulation model, which describes delayed failure in soft and fibrous materials [6], defines damage evolution $D(t)$ as:

$$D(t) = \text{scale} \cdot \left[1 - \left(1 - \frac{t}{T_f}\right)^{1/(c+1)}\right]$$

where T_f is the failure time and c is a material exponent. This S-curve captures the slow initiation and rapid acceleration phases characteristic of cumulative damage, which is analogous to the progression of biological tissue degradation and plaque vulnerability [7].

2.3. Arterial Remodeling. Early-stage atherosclerosis exhibits compensatory outward remodeling known as the Glagov phenomenon, where the arterial wall expands to preserve lumen area despite plaque growth [8]. This compensation continues until approximately 40% plaque burden, after which stenotic narrowing begins. This phenomenon implies that plaque burden measurements may not increase monotonically in early stages if referenced against a remodeling vessel wall, a nuance that must be accommodated in data filtering. More recent constrained

mixture models of arterial growth [9] further link hemodynamic stress to wall remodeling, justifying the isotropic growth assumptions used in simplified contour prediction models.

3. Methodology.

3.1. Data Preparation and Image Processing. The foundation of the predictive model is accurate quantification of plaque burden from imaging. Optical Coherence Tomography (OCT) or IVUS images are processed via a semi-automated pipeline [10]. This involves deep learning-based segmentation to extract the lumen and external elastic membrane (EEM) contours, followed by artifact removal.

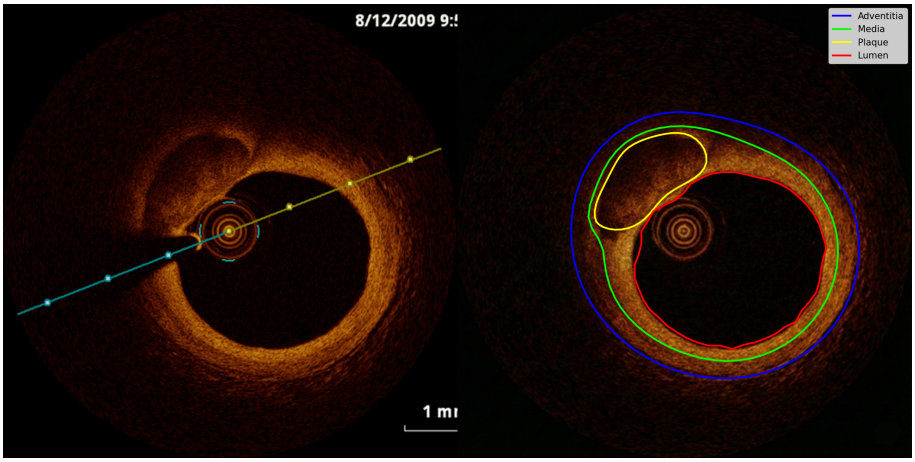

Figure 1. Example of OCT image processing pipeline. (Left) Raw OCT cross-section showing arterial wall and lumen. (Right) Processed image with segmented contours for the lumen (red) and EEM (blue) used to calculate plaque burden.

From these calibrated contours, plaque burden B is computed as the relative cross-sectional area:

$$B = \frac{A_{\text{wall}} - A_{\text{lumen}}}{A_{\text{wall}}}$$

This dimensionless metric serves as the target variable y for the regression model.

3.2. Pseudo-longitudinal Data Generation. Since true longitudinal data is unavailable, synthetic baseline-followup pairs are generated by matching individuals from the cross-sectional population. A rigorous filtering process is employed to ensure physical plausibility. Pairs (t_i, y_i) and (t_j, y_j) are weighted by a Gaussian kernel $w_{\text{time}}(t_i, t_j)$ centered on a target gap Δt_{target}.

Crucially, a Glagov-tolerant filter is applied. Pairs are accepted where $y_j - y_i \geq -\tau_{\text{Glagov}}$, allowing for slight apparent decreases in burden due to compensatory remodeling [8], while pairs that violate maximum biological growth rates ($r_{\max}$) or imply unrealistic jumps in population quantile ($q_{\max}$) are rejected. This ensures the training data reflects valid biological progression rather than random noise.

3.3. Physics-Guided Chained GPR. A Chained GP is employed where the mean function of the first latent f_1 (trajectory) is explicitly set to the Power-Law

Damage Accumulation formulation:

$$m_1(t) = \text{scale} \cdot \left[1 - \left(1 - \frac{t}{T_f}\right)^{1/(c+1)}\right]$$

To ensure numerical stability during optimization, a soft extension of this function is implemented for $t > 0.95T_f$.

The second latent f_2 (log-variance) utilizes a physics-based mean function proportional to the damage rate, $m_2(t) \propto (dD/dt)^\beta$, reflecting the hypothesis that biological variability increases as the disease state accelerates. The kernel for f_2 combines a Linear and RBF kernel to ensure that uncertainty grows monotonically in extrapolation regions [11], preventing the "variance collapse" often seen in standard GPs.

The model is trained by maximizing the Evidence Lower Bound (ELBO), augmented with penalty terms $\mathcal{L}_{\text{physics}}$ to enforce monotonicity and growth constraints:

$$\mathcal{L} = \mathcal{L}_{\text{ELBO}} - \lambda_1 \mathcal{L}_{\text{mono}} - \lambda_2 \mathcal{L}_{\text{growth}} - \lambda_3 \mathcal{L}_{\text{max}}$$

3.4. Patient-Specific Personalization. For a new patient with baseline measurement (t_b, y_b), exact Gaussian conditioning is performed. The joint distribution of the function values at the baseline age t_b and future age t_f is multivariate normal. Conditioning on the observation y_b yields the posterior predictive distribution:

$$\mu_{\text{post}} = \mu_f + \frac{K_{bf}}{K_{bb} + \sigma_n^2}(y_b - \mu_b)$$

This shifts the population-level curve to pass through the patient's specific data point, effectively "personalizing" the risk trajectory.

4. Numerical Experiments and Results.

4.1. Dataset and Experimental Setup. The methodology was validated on a dataset of $N = 57$ individuals (ages 17–92). Figure 2 shows the distribution of relative plaque burden measurements across the age range. To test extrapolation capabilities, three "Life Fraction" (LF) subsets were created: LF0.5 (training on ages 17–47), LF0.75 (training on ages 17–65), and LF1.0 (full dataset).

4.2. Population Model Performance. The physics-guided mean function successfully prevents the unphysical regression to zero often seen in standard GPs. Figure 3 shows the model fit for each life fraction subset. As shown in Table 1, the model maintains low error rates even when extrapolating decades beyond the training data (LF0.5 scenario).

Table 1.

Population model performance across life fractions.

Metric	**LF50**	**LF75**	**LF100**
Training MAPE	31.8%	27.1%	24.0%
Extrapolation MAPE	15.4%	14.1%	–
95% CI Coverage	100%	100%	100%

The variance behavior is particularly notable. The use of the power-law rate-based mean for f_2 results in uncertainty that naturally expands near the theoretical

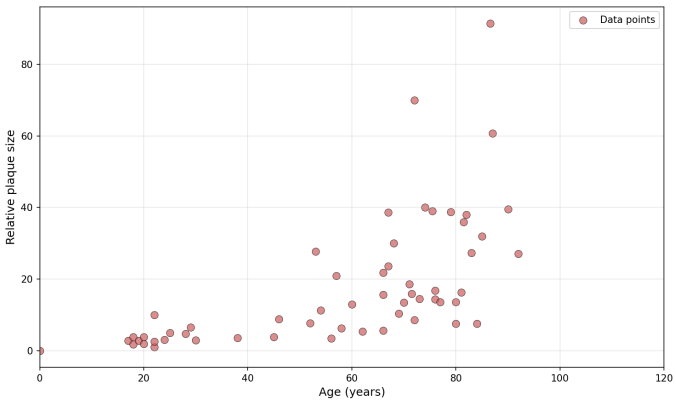

Figure 2. Cross-sectional dataset: relative plaque burden versus patient age (N=57, ages 17–92).

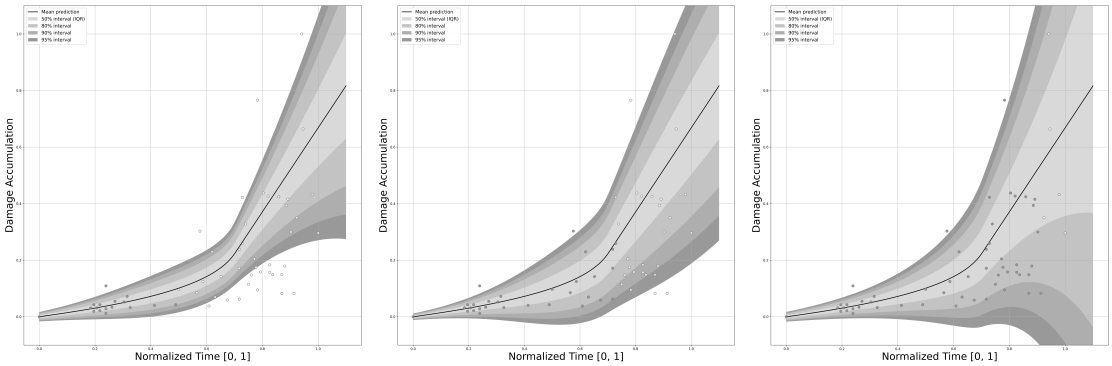

Figure 3. Training data distribution and model fit for different life fractions. (a) LF0.5 (50% of lifecycle), (b) LF0.75, and (c) LF1.0. The gray-filled dots represent the training data (pseudo-longitudinal pairs), and the shaded region indicates the model's uncertainty.

failure time T_f. This aligns with the "Explosion of Uncertainty" concept in risk assessment [12], where the predictability of a system degrades as it approaches a critical transition (rupture).

4.3. Plaque Growth Visualization. Probabilistic forecasts for plaque burden evolution were generated. Figure 4 illustrates the learned trajectories. The model captures the non-linear acceleration of damage in later years. The confidence intervals (gray bands) effectively capture the spread of the data, ensuring that no patient observations fall outside the predicted risk corridors.

4.4. Personalization and Conditioning. The personalization capability was evaluated using the $M = 35$ synthetic follow-up pairs. The exact Gaussian conditioning provided a robust posterior mean but showed minimal error reduction compared to the population prior ($< 0.1\%$). This result indicates that the Physics-Guided Mean Function (m_1) is so effective at capturing the underlying biological trend that the residual epistemic uncertainty (K_{bb}) collapses to near zero during training. The physics prior describes the data so well that the model sees little need for a flexible, data-driven covariance structure to correct it. This confirms the dominance and validity of the Power-Law Damage Accumulation model for this

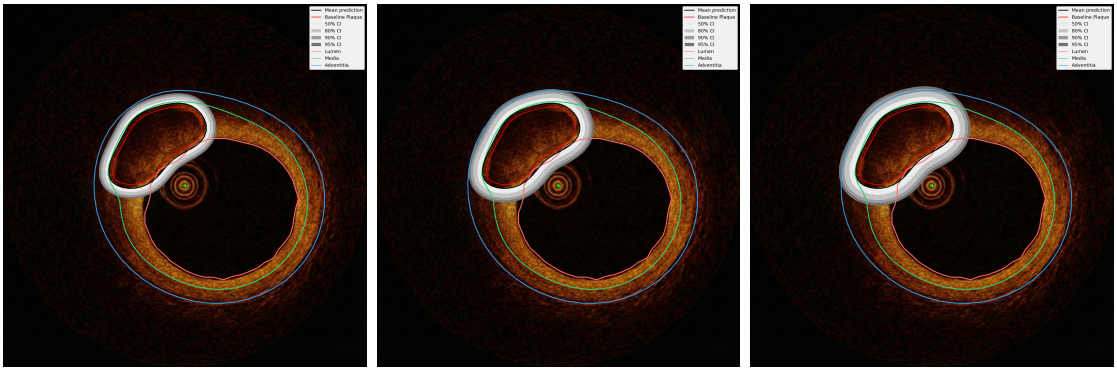

Figure 4. Plaque growth predictions and uncertainty quantification. The solid black line represents the mean trajectory $D(t)$, and the gray bands represent the 95% confidence intervals. Note the monotonic growth and widening uncertainty in extrapolation regions (LF0.5), consistent with damage mechanics theory.

application.

4.5. 5-Year Contour Prediction. To demonstrate clinical utility, the predicted burden B_{t+5} was projected onto the vessel geometry. Using an isotropic growth assumption supported by recent biomechanical studies [9, 13], the lumen radius was contracted according to $r_{\text{lumen}}^{(f)} = r_{\text{wall}}\sqrt{1 - B_f}$. This allows for the visualization of future stenosis and potential occlusion risks, bridging the gap between abstract regression numbers and patient anatomy.

5. Discussion. The integration of Power-Law Damage Accumulation mechanics into the Gaussian Process framework addresses the critical limitation of standard ML in medicine: reliable extrapolation. Standard RBF kernels regress to a zero mean outside the training data, which is physically impossible for cumulative damage [11]. By enforcing the power-law structure, it is ensured that predictions obey the second law of thermodynamics (entropy/damage accumulation) even in the absence of data.

The results highlight a trade-off between physics-based robustness and data-driven flexibility. The "collapse" of the epistemic covariance suggests that for small datasets ($N = 57$), a strong physical prior is far more valuable than a flexible kernel. This aligns with findings in the broader PIML literature [1, 14], where inductive biases are shown to reduce the sample complexity of learning algorithms.

The Glagov-tolerant filtering proved essential. Without it, the "noise" introduced by compensatory remodeling would force the GP to learn a high noise variance σ_n^2, obscuring the true progression signal. By explicitly modeling this physiological phenomenon, the training data was refined to better reflect the underlying pathology.

Future work will focus on integrating these geometric predictions with Finite Element Analysis (FEA). The predicted contours can serve as geometries for stress analysis, allowing for the calculation of Peak Circumferential Stress (PCS), a direct biomechanical predictor of plaque rupture [15].

6. Conclusions. A Physics-Guided Chained Gaussian Process Regression framework for predicting atherosclerotic plaque progression was presented. By embedding continuum damage mechanics into the probabilistic learning model, the following was achieved:

- **Robust Extrapolation:** The model predicts monotonic damage growth even when trained on only 50% of the lifecycle (MAPE 15.4%).
- **Calibrated Uncertainty:** The heteroscedastic noise model captures the increasing variability of disease presentation with age.
- **Clinical Relevance:** The pipeline transforms raw OCT images into future risk predictions, providing a tangible tool for long-term cardiovascular prognosis.

The proposed method offers a generalized approach for modeling progressive degenerative diseases where data is sparse, but the underlying physics of decay is well-understood.

Conflict of Interest

The authors declare that they have no conflicts of interest in relation to the current study, including financial, personal, authorship, or any other, that could affect the study, as well as the results reported in this paper.

Funding

The research was conducted without financial support.

Data Availability

All data are available, either in numerical or graphical form, in the main text of the manuscript.

Use of artificial intelligence

During the preparation of this manuscript, the authors used Gemini (Google) to improve language, grammar, and check formatting compliance. After using this tool, the authors reviewed and edited the text and take full responsibility for the content of the published manuscript.

Contributions of authors

Krasii D. M.: Conceptualization, Methodology, Writing — original draft. Larin O. O.: Supervision, Formal analysis, Methodology — review & editing.

References

1. Raissi, M., Perdikaris, P., & Karniadakis, G. E. (2019). Physics-informed neural networks: A deep learning framework for solving forward and inverse problems involving

nonlinear partial differential equations. *Journal of Computational Physics,* 378, 686–707. https://doi.org/10.1016/j.jcp.2018.10.045
2. Karniadakis, G. E., Kevrekidis, I. G., Lu, L., Perdikaris, P., Wang, S., & Yang, L. (2021). Physics-informed machine learning. *Nature Reviews Physics,* 3(6), 422–440. https://doi.org/10.1038/s42254-021-00314-5
3. Rasmussen, C. E., & Williams, C. K. I. (2005). *Gaussian Processes for Machine Learning.* The MIT Press. https://doi.org/10.7551/mitpress/3206.001.0001
4. Saul, A. D., Hensman, J., Vehtari, A., & Lawrence, N. D. (2016). Chained Gaussian Processes (Version 1). arXiv. https://doi.org/10.48550/ARXIV.1604.05263
5. Hensman, J., Fusi, N., & Lawrence, N. D. (2013). Gaussian Processes for Big Data (Version 1). arXiv. https://doi.org/10.48550/ARXIV.1309.6835
6. Lockwood, H. A., Agar, M. H., & Fielding, S. M. (2024). Power law creep and delayed failure of gels and fibrous materials under stress. *Soft Matter,* 20(11), 2474–2479. https://doi.org/10.1039/d3sm01608k
7. Larin, O., & Vodka, O. (2014). A probability approach to the estimation of the process of accumulation of the high-cycle fatigue damage considering the natural aging of a material. *International Journal of Damage Mechanics,* 24(2), 294–310. https://doi.org/10.1177/1056789514536067
8. Glagov, S., Weisenberg, E., Zarins, C. K., Stankunavicius, R., & Kolettis, G. J. (1987). Compensatory Enlargement of Human Atherosclerotic Coronary Arteries. *New England Journal of Medicine,* 316(22), 1371–1375. https://doi.org/10.1056/nejm198705283162204
9. Humphrey, J. D., & Rajagopal, K. R. (2002). A Constrained Mixture Model for Growth and Remodeling of Soft Tissues. *Mathematical Models and Methods in Applied Sciences,* 12(03), 407–430. https://doi.org/10.1142/s0218202502001714
10. Krasii, D., Larin, O., & Gorovyi, I. (2025). Analysis of Optical Coherence Tomography Images of Human Arteriosclerotic Vessels Using Computer Vision and Deep Learning Algorithms. In *Lecture Notes in Networks and Systems* (pp. 107–117). Springer Nature Switzerland. https://doi.org/10.1007/978-3-031-94845-9_9
11. Krasii, D., & Larin, O. (2023). ML-surrogate modeling for the estimation of random system performance parameter progress by the Chained Gaussian Process Regression method. In *2023 IEEE 4th KhPI Week on Advanced Technology (KhPIWeek)* (pp. 1–5). IEEE. https://doi.org/10.1109/khpiweek61412.2023.10312806
12. Sornette, D. (2000). *Critical Phenomena in Natural Sciences.* In Springer Series in Synergetics. Springer Berlin Heidelberg. https://doi.org/10.1007/978-3-662-04174-1
13. Ambrosi, D., Ateshian, G. A., Arruda, E. M., Cowin, S. C., Dumais, J., Goriely, A., Holzapfel, G. A., Humphrey, J. D., Kemkemer, R., Kuhl, E., Olberding, J. E., Taber, L. A., & Garikipati, K. (2011). Perspectives on biological growth and remodeling. *Journal of the Mechanics and Physics of Solids,* 59(4), 863–883. https://doi.org/10.1016/j.jmps.2010.12.011
14. Wang, S., Teng, Y., & Perdikaris, P. (2020). Understanding and mitigating gradient pathologies in physics-informed neural networks (Version 1). arXiv. https://doi.org/10.48550/ARXIV.2001.04536
15. Cheng, G. C., Loree, H. M., Kamm, R. D., Fishbein, M. C., & Lee, R. T. (1993). Distribution of circumferential stress in ruptured and stable atherosclerotic lesions. A structural analysis with histopathological correlation. *Circulation,* 87(4), 1179–1187. https://doi.org/10.1161/01.cir.87.4.1179

Красій Д. М., Ларін О. О. Фізично-керована ланцюгова гаусівська регресія процесу для прогнозування прогресії атеросклеротичної бляшки.

Прогнозування прогресії атеросклеротичної бляшки за розрідженими даними поперечних зображень є критичним для оцінки серцево-судинного ризику. Стандартні методи регресії не мають фізично обґрунтованого індуктивного зміщення, що призводить до ненадійної екстраполяції. Представлено фреймворк фізично-керованого ланцюгового гаусівського процесного регресування (PG-CGPR), який інтегрує механіку степеневого накопичення пошкоджень як апріорну функцію середнього з гетероске-

дастичним моделюванням шуму. Спостереження поперечних перерізів трансформуються у псевдо-лонгітюдні навчальні пари з використанням фільтрації, толерантної до явища Глагова. Валідований на даних серцево-судинних зображень, що охоплюють кілька десятиліть, модель зберігає точність прогнозування навіть при екстраполяції за межі навчального вікового діапазону, з довірчими інтервалами, що покривають усі тестові спостереження. Сильний фізичний апріор домінує над навчанням на основі даних, демонструючи, що в режимах малих вибірок механістичні моделі перевершують гнучкі ядра. Фреймворк дозволяє проектувати майбутні геометрії судин для подальшого біомеханічного аналізу.

Ключові слова: гаусівські процеси, прогресія бляшки, фізично-інформоване машинне навчання, механіка пошкоджень, кількісна оцінка невизначеності, серцево-судинна візуалізація.

Recived: 05.12.2025 Accepted: 22.12.2025 Published: 29.01.2026