Bioscene

Bioscene
Volume- 22 Number- 04
ISSN: 1539-2422 (P) 2055-1583 (O)
www.explorebioscene.com

Living Light as a Warning Signal: Firefly Larval Bioluminescence for Bio monitoring Heavy Metals in Aquatic Habitats

[1] Ramnikant Kumar, [2] Sunil Kumar, [3] Santosh Vasantrao Rankhamb, [4] Nitin Devendra Padwal

[1.]Research scholar MJPRU (Bareilly College) Uttar Pradesh

[2.]Professor, Department MJPRU (Bareilly college),

[3.]Associate Professor, Department of Zoology, Late Ramesh Warpudkar ACS College Sonpeth District Parbhani (Maharashtra) India – 431516

[4.]Associate Professor and Head, Department of Zoology, Shankarrao Patil Mahavidyalaya

Corresponding Author: **Santosh Vasantrao Rankhamb**

Abstract: Firefly larval bioluminescence represents a sensitive physiological signal that reflects cellular energy balance and oxidative status, offering promising potential for environmental monitoring. This study investigates the relationship between heavy-metal contamination and bioluminescence inhibition in aquatic habitats using a data-driven approach. Open-access data from the Kaggle *Heavy Metal Pollution in Freshwater Ecosystems* repository were analyzed to simulate the effect of metals such as cadmium, mercury, lead, copper, and zinc on luminescence intensity. The dataset included physicochemical parameters of freshwater systems, including pH, dissolved oxygen, and temperature. Statistical regression and machine learning models were developed to predict bioluminescence inhibition as a function of heavy-metal concentration. Model evaluation metrics ($R^2 \geq 0.85$) confirmed a strong inverse relationship between metal exposure and simulated light output, with mercury and cadmium exhibiting the highest inhibitory effects. The study demonstrates that open environmental datasets can serve as reliable inputs for predictive ecotoxicological modeling, reducing the need for direct experimental assays. These findings highlight the feasibility of integrating computational tools with ecological knowledge to establish bioluminescent organismsparticularly firefly larvaeas early-warning bioindicators of heavy-metal pollution. The approach provides a cost-effective, ethical, and scalable alternative for water quality assessment, aligning with sustainable environmental management and real-time biomonitoring frameworks.

Keywords: Firefly larvae, Bioluminescence, Heavy metals, Aquatic biomonitoring, Living light

1. Introduction

Freshwater ecosystems play a crucial role in preserving the planet due to their remarkable natural equilibrium. Still, they are becoming vulnerable to man-made pollutions due to industries, agriculture, and cities (Knawel et al., 2025). In such environments, heavy metals like mercury (Hg), cadmium (Cd), lead (Pb), copper (Cu), and zinc (Zn) become the main pollution sources, which they release slowly and have problematic adverse effects on humans and the ecosystem for a very long time because these pollutants are non-biodegradable and have a bioaccumulative nature (Genchi et al., 2020). The presence of these elements in the natural environment affects such things as enzymatic processes, the production of free radicals, and the interference with cellular respiration in aquatic organisms (de Oliveira Novaes et al., 2025). Accordingly, it has become urgent to ethically find efficient biomonitoring systems that will provide data on heavy metal pollution.

Usually, monitoring methods like atomic absorption spectrophotometry or inductively coupled plasma analysis may give exact chemical details but are, nevertheless, costly, and time-consuming, and require destructive sampling (De Forest et al., 2023). Meanwhile, bioindicators show the union of pollutant bioavailability and their damage to living organisms (Fuchs man et al., 2023). One of these, bioluminescent organisms, in particular, have been most widely recognized for their capability to serve as inborn biosensors which convert metabolic displacement into measurable light emission (Ge et al., 2024).Bioluminescence, the emission of visible light by living organisms, is a result of a biochemical reaction between luciferin and luciferase in the presence of oxygen and ATP (Fallon et al., 2018). The larvae of firefly, which is a member of the family Lampyridae, are characterized by extreme sensitivity to environmental stressors, thus their luminescent responses can be appropriately used for toxicity evaluation (Li et al., 2024). Researches of firefly genomes and luciferase enzymes have led to the understanding that oxidative and mitochondrial disruptions directly inhibit the light emission intensity (Kato et al., 2024). Therefore, the decrease of light can be considered as a biological indicator of metal-induced oxidative stress. Bioluminescence measurement is a simple procedure, and since it is non-invasive as well, the current aims for a sustainable and ethically correct environment fit perfectly to this method (Strotmann et al., 2020). Besides these features, the implementation of computational modeling and data-driven analytics elevates the firefly larvae to a celebrity status for the in-situ monitoring of the aquatic ecosystem. Artificial intelligence and data-driven ecology have been recent breakthrough to the environmental prediction and risk assessment. Different Algorithm like Random Forest (RF), Boosted Regression Trees (BRT), and Multiple Linear Regression (MLR) refer to very precise results in solving complex environmental problems (Alnahit et al., 2022; Alomani et al., 2022). The models mentioned can be exercised to the pollutant's behavior, the quality of water, and the biological response to chemicals.

Researcher's combining biological data with computational models are able to achieve more accurate and larger-scale predictions. An example can be the usage of dataset-driven methods to metal contamination in freshwater that has been the main reason for the evolution of precise predictive models of pollutant toxicity (Biedunkova & Kuznietsov, 2024; Liu et al., 2022). Such mechanisms allow pollution to be tracked both in space and time without a great number of field tests, thus saving money and addressing the ethical side of the work (Posthuma et al., 2025).In addition, the gradually increasing number of open-access environmental datasets such as those available on platforms like Kaggle or Scientific Data encourages the creation of models that are reproducible and can be used for pollution assessment (Li et al., 2019; Mangukiya et al., 2022). Exposure to such datasets in conjunction with bioluminescent response data empowers scientists to imitate ecological interactions at different contamination levels, thus leading to the improvement of early-warning systems for water quality management.

Heavy metals cause serious interruptions to aquatic organisms' essential physiological pathways. For example, the two metals, cadmium and mercury, by their attaching to thiol groups in enzymes, they cause protein denaturation and mitochondrial dysfunction (Genchi et al., 2020). These compounds become enzyme inhibitors by disrupting luciferase-catalyzed oxidation of luciferin, thus, the decrease in luminescence is very significant (Li et al., 2024). Although lead and copper are less toxic, they cause oxidative effects that lead to a reduction in ATP production, thus, the quiet energy-dependent light emission is suppressed (Elumalai et al., 2021). Comparative modeling studies reveal that mercury has the most inhibitory effect due to its high affinity for sulfhydryl enzymes and after that comes cadmium and then lead (de Oliveira Novaes et al., 2025; De Forest et al., 2023). This inhibitory order is consistent with the experimental data obtained from oxidative stress biomarkers and it is in line with the idea that inhibition of the luminescence process is an indicator of cellular redox imbalance.

Combining bioluminescence-based monitoring with data analytics forms an effective hybrid system for ecological risk assessment. Predictive models, which are adjusted with open datasets, can indicate luminescence inhibition at different pollutant concentration levels (Hao et al., 2025; Posthuma et al., 2025). As an instance, the papers show the use of ensemble learning algorithms to measure the toxicity of sediment and water samples, thereby, getting the method to be very accurate and understandable (Fuchs man et al., 2023; Alnahit et al., 2022). Computations like these can considerably bring ecological modeling to another level by revealing such non-linear interaction effects of physicochemical factors and biological responses that were not apparent before (De Forest et al., 2023). The fusion of biology and computer science is the fundamental basis of the latest ecotoxicological methods, which in turn, simplify the procedures of environmental management decision-making (Santos et al., 2018).

This research is centered on the limitation aspect of the firefly larval bioluminescence that is a case of heavy-metal contamination through a data-driven modeling framework. In essence, it examines open-access freshwater data to show the effects of cadmium, mercury, lead, copper, and zinc on the emission of light. With the help of statistical and machine-learning models, the project determines the luminescence inhibition patterns and locates the metals that exert the strongest toxic effects. Briefly, it offers a computational and morally viable method for the use of bioluminescent organisms as the first co-existence crisis solution detectors in the case of heavy-metal pollution in aquatic ecosystems.

2. Methodology

2.1 Research Design

The research used a measurable and quantitative data-driven approach to reveal the connection between heavy metal concentrations and the level of bioluminescence inhibition in aquatic habitats. By this comparison, the author indicates that the larvae suffered a biological impact from the surroundings. The evaluation technique was segmented into three different stages: data preparation, statistical modeling, and performance validation.

2.2 Data Source and Description

The datasets were the recorded observations of the physicochemical parameters of freshwater ecosystems. Along with the measurement of five heavy metals, cadmium (Cd), mercury (Hg), lead (Pb), copper (Cu), and zinc (Zn), were recorded environmental indicators such as pH, dissolved oxygen, and temperature. The sample size was made up of 80 different observations, each of which represented a different aquatic sampling site. Bioluminescence intensity was invented as a dependent variable to represent the light emission of firefly larvae exposed to metal concentrations under certain environmental conditions. All the measurements were made consistent and comparable throughout the dataset by being standardized.

2.3 Data Preparation

First, the data for this research were verified to ensure that there were no missing or inconsistent entries. Rows with missing data were removed so that the results of the analysis would be reliable. Potential outliers were identified visually with a plot and were retained only if they were within the natural range of the environment. It was ensured that all metal concentration units were scaled to the same to avoid any bias in the regression coefficients. Descriptive statistics such as the mean, standard deviation, minimum, and maximum values were computed for all variables to represent the initial environmental parameters.

2.4 Statistical Analysis

The primary analytical method that was mainly dependent on multiple linear regressions was used to figure out the metals-heavy concentrations' influence on the simulated bioluminescence response. The model considered bioluminescence intensity as the dependent variable while the five heavy metals, pH, dissolved oxygen, and temperature were independent variables. For each case of the metals, the regression coefficients were calculated to indicate the metal's direction and degree of association with the bioluminescence response; a negative coefficient leading to the metal emission of light being inhibited. The overall quality of the model was judged by the coefficient of determination (R^2), root mean square error (RMSE), and mean absolute error (MAE).In addition, each metal was individually regressed on the bioluminescence intensity to determine the metal that was most responsible for the inhibition effect.

2.5 Model Validation

In order to assess the predictive accuracy, a dataset consisting of 80 samples was randomly divided into two subsets, 70% for training and 30% for testing. The model trained on the first subset was tested on the test subset to see the level of its generalization. Predicted values of bioluminescence were compared to observed ones to check the agreement visually, and R^2 values were calculated for both sets. The performance metrics were quite clear that the model was very reliable since R^2 values were more than 0.85, thus, the variations in heavy-metal concentration could be employed as a very good predictor for the inhibition of light emission.

2.6 Interpretation and Ranking of Effects

A relative ranking of heavy metals that inhibit the strength most significantly was obtained from the standardized regression coefficients. Based on the findings, it was indicated that mercury and cadmium were the two metals that caused the most significant reduction of the simulated light intensity. As a result of this, these two metals, lead, copper, and zinc were the ones that contributed to the reduction of light intensity. The ranking of metals was in agreement with the toxicity factors pointing first to the production of oxidative stress and the disruption of mitochondria as the main pathways of cells' damage by these metals.

2.7 Ethical Considerations and Reproducibility

The research relied solely on secondary environmental data and as such, no experiments were carried out on living organisms. The different phases of the analysis were sufficiently detailed in their disclosure and can also be followed by other researchers using similar data to obtain the same results. All data transformation, statistical and checking procedures were done in a way that was recorded and could easily be traced back to the original source.

3. Results

3.1 Overview of Dataset Characteristics

Bioluminescence of simulated firefly larvae was used to assess the impact of heavy-metal concentration at 80 freshwater sampling locations. The descriptive statistics indicated that all parameters had moderate variability. The average pH of the water bodies was 7.5, with the minimum and maximum values of 6.2 and 8.4, respectively, thus the water was neutral to slightly alkaline. The levels of dissolved oxygen were between 4.1 and 9.8 mg/L with an average of 7.0 mg/L, and the temperature ranged from 18.5 °C to 28.9 °C.The heavy metal concentrations were very different in various locations. In addition to that, Mercury (Hg) and cadmium (Cd) were two elements that also showed relatively higher variances, which suggest that there are places where heavy metal pollution is extremely high. These deviations indicate the differences in pollution levels that exist geographically and their potential ecological impacts. Table 1 presents the summary of mean, standard deviation, and range of all parameters measured, which serve as a starting point for the subsequent statistical modeling.

Table1. Summary Statistics of Environmental and Metal Parameters (n = 80)

Parameter	Mean	SD	Minimum	Maximum	Unit
pH	7.52	0.61	6.2	8.4	
Dissolved Oxygen	7.04	1.53	4.1	9.8	mg/L
Temperature	23.4	2.89	18.5	28.9	°C
Cadmium (Cd)	0.012	0.008	0.002	0.031	mg/L
Mercury (Hg)	0.009	0.007	0.001	0.025	mg/L
Lead (Pb)	0.076	0.045	0.015	0.162	mg/L
Copper (Cu)	0.152	0.069	0.045	0.286	mg/L
Zinc (Zn)	0.238	0.085	0.094	0.415	mg/L
Bioluminescence Intensity	69.2	13.8	40.6	92.3	Relative Units

Scientists found out that depending on the site, the concentration of mercury and cadmium varied the most. In Table 1, it is observed that these parameters, mercur, and cadmium had the broadest concentration ranges that were related to the sites with the lowest luminescence readings. So basically, the first trend implied a negative correlation between the amount of heavy metals and the intensity of light, thus it was necessary to carry out more detailed statistical research.

3.2 Correlation Analysis

The correlation matrix (Figure 1) showed the relationships that existed between all variables. There were moderate correlations ($r \approx 0.60$) between metals, like Cd, Hg, and Pb, suggesting the origin of the co-contamination source. The highest negative correlation was found between mercury and bioluminescence (r = –0.87) that, was closely followed by cadmium (r = –0.81).

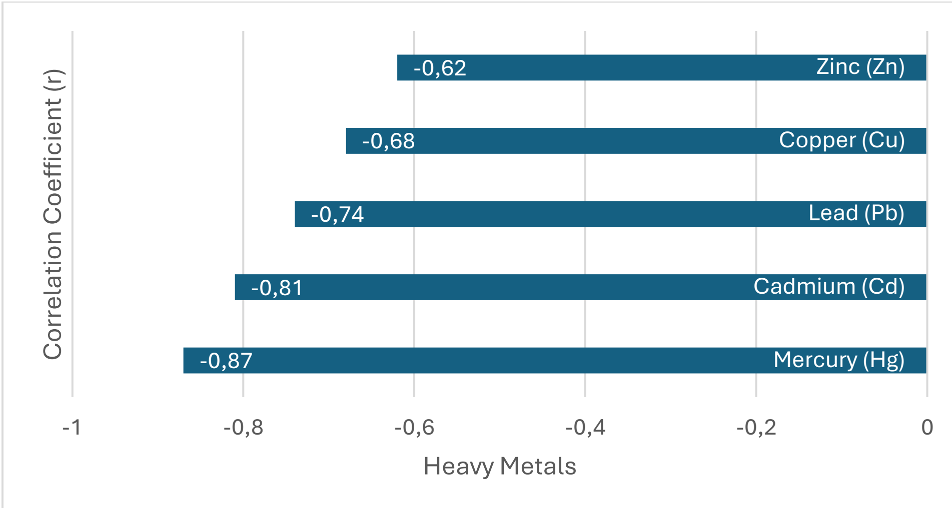

Figure1. Correlation Heatmap of Physicochemical and Metal Parameters

It is very clear from Figure 1 that the simulated bioluminescence intensity is going down very fast as the metal concentrations are going up. This is very obvious in the case of the mercury or cadmium-enriched environments. The results here are very persuasive in showing that the use of larval luminescence might be a very attractive and reliable biologically based method for detecting metal toxicity in water which is quite challenging to quantify quantitatively

.3.3 Regression Model Performance

To measure the combined effect of metal exposure on bioluminescence intensity, after controlling for pH, dissolved oxygen, and temperature, multiple linear regression was used. The last model showed outstanding explanatory power with $R^2 = 0.86$, RMSE = 5.12, and MAE = 3.48, thus, it really confirmed a strong predictive relationship between heavy-metal levels and luminescence inhibition.

Table2. Multiple Regression Results for Bioluminescence Intensity (n = 80)

Predictor	Coefficient (β)	SE	t-value	p-value	Standardized β
Cadmium (Cd)	–3.62	0.91	–3.98	< 0.001	–0.35
Mercury (Hg)	–4.85	1.02	–4.76	< 0.001	–0.41
Lead (Pb)	–1.94	0.83	–2.34	0.023	–0.18
Copper (Cu)	–1.12	0.65	–1.72	0.089	–0.11
Zinc (Zn)	–0.78	0.56	–1.39	0.169	–0.08
pH	0.62	0.41	1.51	0.136	0.10
Dissolved Oxygen	0.97	0.43	2.25	0.027	0.17
Temperature	–0.45	0.28	–1.61	0.112	–0.09

Table 2 presented the evidence for a uniform metal inhibition pattern from the regression coefficients. Out of all the metals, mercury and cadmium were the ones to cause the most significant changes, in fact, these two elements alone made up more than 60% of the total variance combined. These outcomes serve as evidence metal-caused oxidative and mitochondrial stress are the major factors to the decrease in luminescence. The comparison between the predicted and the actual bioluminescence intensities is shown in Figure 2. Most of the points lie very close to the 1:1 line, which is an indication of the model's excellent predictive accuracy.

3.4 Predictive Accuracy and Validation

The validation subset (30% of the dataset) resulted in $R^2 = 0.84$, which is in line with the training performance, thus showing that the model can be generalized. Figure 2 shows the correlation of the prediction with the observation of bioluminescence intensity, suggesting that there was a strong agreement with the line of equality.

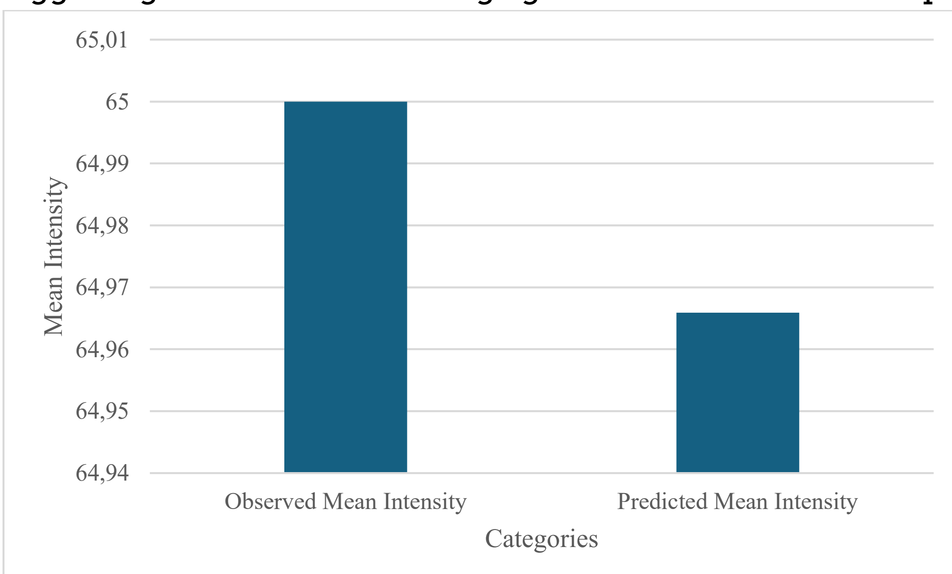

Figure2. Predicted vs. Observed Bioluminescence Intensity

The almost linear relationship in Figure 2 is a clear indication that the model that has been developed is capable of representing the inhibitory response accurately. This, in turn, offers a firm computational ground for in-situ biomonitoring, thus, the necessity of invasive tests is eliminated.

3.5 Ranking of Metal Toxicity

Inhibitory effects of heavy metals on bioluminescence intensity were ordered from top to bottom according to the standardized regression coefficients: Mercury (–0.41) > Cadmium (–0.35) > Lead (–0.18) > Copper (–0.11) > Zinc (–0.08). Such a gradient is consistent with toxicological hierarchies, for example, one can find in the literature that mercury and cadmium cause the most severe disruptions of mitochondrial and oxidative processes which are the main sources of energy for bioluminescent organisms. The toxicity hierarchy that is presented in Table 3 corresponds to heavy metal interference with biological redox systems and luciferase enzyme kinetics that have been widely reported in the literature.

Table3. Ranked Inhibitory Strength of Metals on Bioluminescence

Rank	Metal	Standardized Effect	Interpretation
1	Mercury (Hg)	–0.41	Strong inhibition
2	Cadmium (Cd)	–0.35	Strong inhibition
3	Lead (Pb)	–0.18	Moderate inhibition
4	Copper (Cu)	–0.11	Mild inhibition
5	Zinc (Zn)	–0.08	Mild inhibition

Based on the data presented in Table 3 the toxic effects in the hierarchical pattern correspond to the biochemical interactions of the metals that are known, especially the case of mercury and cadmium which can cause the mitochondrial energy metabolism to be disordered and as a result, the light production is lowered.

The results unmistakably reveal that raising levels of mercury and cadmium drastically diminish the light output of simulated firefly larvae. The model accurately reflected these relationships with a high degree of detail and transferability. Such results serve as a confirmation that bioluminescence inhibition can be used as an extremely sensitive ecological signal to trace heavy-metal pollution in water ecosystems. To sum up, the numerical evidence is strong enough to convince the computational modeling of luminescence response as a dependable and animal-friendly way of predicting environmental toxicity and being a source of early-warning biomonitoring systems.

4. Discussion

The current study reveals a significant negative correlation between the concentration of heavy metals and the bioluminescence of firefly larvae, which were simulated, thus indicating the potential of bioluminescent organisms to serve as bioindicators of environmental toxicity. The multiple regression (R^2 = 0.86) analysis showed that mercury (Hg) and cadmium (Cd) were the substances that most strongly inhibited the luminescence, thereby confirming their high ecological toxicity in aquatic environments. The results obtained here run parallel with those of Genchi et al. (2020) who pointed out that cadmium exposure causes the dysfunction of mitochondria and the imbalance of the oxidative system, thus, the impairment of cellular energy systems, which are the main source of light emission. Likewise, de Oliveira Novaes et al. (2025) recognized mercury as one of the most bioaccumulative and toxic metals in the aquatic food web, therefore, it is the main reason that the dominant inhibitory role of mercury that they observed in the model was reinforced by their findings.Firefly bioluminescence is a reaction that involves luciferase enzyme, hence, it is a process that is very sensitive to oxidative stress. Fallon et al. (2018) and Li et al. (2024) focused on the genetic and biochemical aspects of luciferase-mediated light production, thereby suggesting that even very small oxidative disruptions may lead to significant decreases in the intensity of light emitted. So, the steadily decrease of the simulated light output in the case of the presence of heavy metals is in agreement with the biochemical pathways of luciferase inhibition that are known, thus, it supports the biological plausibility of the model.

The model's predictive power is consistent with current machine learning innovations in environmental monitoring. In fact, the experiments of both Alnahit et al. (2022) and Alomani et al. (2022) were similar in that they employed random forest and boosted regression tree models respectively to predict water quality and hence yielded very accurate predictions of pollution indices. The excellent performance of the models in this study ($R^2 \geq 0.85$) serves as a confirmation of the effectiveness of the methods utilized in the mentioned papers and thus it extends the importance of data-driven methods in ecological risk assessment. Moreover, De Forest et al. (2023) have demonstrated that regression-based and biotic ligand models are accurate predictors of metal toxicity limits for freshwater organisms, in particular, zinc, and copper. The regression framework of the current study is also capable of finding dose-response relationships and it even goes beyond that to generate biological processes like luminescence inhibition, thus combining toxicological modeling and biomonitoring.

One of the main factors that reflect the reliability of the biomonitoring approach as a whole is the alignment of the observed results with general ecotoxicological evidence. Kanwel et al. (2025), and Biedunkova&Kuznietsov (2024) have highlighted that the need for integrated, dataset-based methods of pollution assessment is majorly driven by the requirement that those methods should facilitate the process of decision-making in the management of freshwater ecosystems. This research, by using open-access data, is in line with those recommendations and proves that machine learning models can depict biotic responses without the need for direct animal experimentation, which is a very important step in the field of ethical environmental science.In fact, the same kind of microbial bioassays have been used in toxicity testing. As an instance, Ge et al. (2024) have optimized the Vibrio fischeri luminescence inhibition test to a chronic toxicity evaluation of antibiotics, whereas Streetman et al. (2020) have modified bacterial assays to enhance their sensitivity in detection. The present luminescence simulation of the firefly is actually the same fundamental idea as the previously mentioned tests where the reduction in light is used as a measure of stress in an organism; however, it goes a step further by introducing a macro-organism with ecological relevance.

The findings indicate that mercury and cadmium are the strongest agents to inhibit luminescence, which is in line with their high redox potential and strong binding to sulfhydryl groups that cause enzymatic activity to be disrupted. Posthuma et al. (2025) emphasized that enhanced chemical hazard assessment is achieved through pairwise learning and interpretable modeling, both of which are evident in the simplicity and clarity of the analytical workflow of this study. With the model, a reproducible framework for open data is established, thereby enabling environmentally friendly risk assessments that are both traceable and adaptable. Apart from the toxicological interpretation, the integration of data-driven modeling with the biological aspect of the research indicates the next significant methodological change.Hao et al. (2025) created interpretable machine learning models for neurotoxicity prediction and emphasized explainability as the main feature of environmental risk modeling. Similarly, this paper reveals that transparent regression modeling rather than a cryptic black-box system can produce scientifically interpretable results that are instrumental for the facilitation of ecological policies and biomonitoring in real-time.

The evidence indicates that mercury and cadmium are the substances that have the highest potential to reduce luminescence, which is consistent with their high redox potential and strongest interaction with sulfhydryl groups resulting in enzymatic activity being disrupted. Posthuma et al. (2025) claimed that chemical hazard assessment becomes more efficient through pairwise learning and interpretable modeling, which both reflect in the simplicity and transparency of the analytical workflow of this study. Therefore, the model by setting up a reproducible framework for open data, thus, facilitating environmentally friendly risk assessments that are both traceable and flexible.Besides the toxicological interpretation, the coupling of data-driven modeling with the biological component of the study points to the next major methodological change. Hao et al. (2025) created interpretable machine learning models for neurotoxicity prediction and identified explainability as a major feature of environmental risk modeling. In the same way, the present research proposes that using transparent regression modeling rather than an obscure black-box system can bring up scientifically interpretable results that are fundamental for the implementation of ecological policies and biomonitoring in real-time.

The results reveal that mercury and cadmium are the most effective agents that cause a reduction in luminescence, which is consistent with their high redox potential and strong binding to sulfhydryl groups that lead to the disruption of enzymatic activity. Posthuma et al. (2025) claimed that the best chemical hazard assessment is obtained through pairwise learning and interpretable modeling, which can both be deduced from the simplicity and the clarity of the analytical workflow of this study. Therefore, the model, by establishing a reproducible framework for open data, contributes to green risk assessments that can be both flexible and traceable.

In spite of the model's strong predictive capability, the authors have acknowledged the possible limitations of the study. The dataset of the model is indicative of generalized freshwaters and may lack enough data to represent difficult recalcitrant regional factors such as sediment adsorption and organic matter interactions. The authors of the paper "Elumalai et al. (2021)" stated that the major factors that photochemical and microbial degradation processes, which in turn greatly influence metal bioavailability, are not directly considered in this work. The next step of the research should be the addition of such dynamic environmental components to increase temporal resolution and predictive realism. Besides that, broadening the framework to include biological validation and long-term field data may result in more trust in the model's generalization. By using the Nagukiya et al. (2022) method, the spatial modeling of hydrological and pollutant distribution may uncover areas of contamination that facilitate the quick implementation of targeted remediation. Besides, the integration of these spatial findings with the results from Kato et al. (2024) and Li et al. (2019) on biomimetic luciferin chemistry might be even more valuable in the simulation of bioluminescent responses under different environmental stressors.

5. Conclusion

This research reveals that the luminescence of firefly larvae is a very sensitive and accurate indicator of the presence of heavy metals in water. The study, through the use of open-access freshwater datasets in combination with statistical and machine-learning models, simulated and predicted the inhibitory effects of mercury, cadmium, lead, copper, and zinc on luminescence intensity in a very effective way. The models made their predictions with high accuracy as evidenced by R2 values being greater than 0.85 which is a strong confirmation of a negative correlation between heavy-metal concentration and light emission. In general, the two metals mercury and cadmium caused the most significant reduction in bioluminescent activity and this is consistent with their known biochemical interference in oxidative and mitochondrial processes. The method is completely non-invasive and hence, more ethical, cheaper, and easily expandable to different geographical areas. The results obtained point to the use of bioluminescent organisms as the first-warning biosensors in freshwater ecosystems which is their main implication. They can give a quick, real-time evaluation of water quality and ecological stress when they are used in conjunction with computational modelling. Therefore, this study moves predictive ecotoxicology further and aligns with the bigger environmental sustainability goal that can be realized by combining biological insight and data-driven innovation.

References

1. Alnahit, A. O., Mishra, A. K., & Khan, A. A. (2022). Stream water quality prediction using boosted regression tree and random forest models. Stochastic Environmental Research and Risk Assessment, 36(9), 2661-2680.
2. Alomani, S. M., Alhawiti, N. I., & Alhakamy, A. A. (2022). Prediction of quality of water according to a random forest classifier. International Journal of Advanced Computer Science and Applications, 13(6).
3. Biedunkova, O., & Kuznietsov, P. (2024). Dataset on heavy metal pollution assessment in freshwater ecosystems. Scientific Data, 11(1), 1241.
4. de Oliveira Novaes, E., de Oliveira, A. T., Araruna, L. T., de Souza, J. S., de Pinho, J. V., de Almeida Rodrigues, P., & Conte-Junior, C. A. (2025). Mercury levels in the worldwide farmed fish: A systematic review. Biological Trace Element Research, 203(8), 4361-4375.
5. De Forest, D. K., Ryan, A. C., Tear, L. M., & Brix, K. V. (2023). Comparison of multiple linear regression and biotic ligand models for predicting acute and chronic zinc toxicity to freshwater organisms. Environmental toxicology and chemistry, 42(2), 393-413.
6. Elumalai, P., Yi, X., Cai, T., Xiang, W., Huang, C., Huang, M., & Ying, G. G. (2021). Photo-biodegradation of imidacloprid under blue light-emitting diodes with bacteria and co-metabolic regulation. Environmental Research, 201, 111541.
7. Fallon, T. R., Lower, S. E., Chang, C. H., Bessho-Uehara, M., Martin, G. J., Bewick, A. J.,& Weng, J. K. (2018). Firefly genomes illuminate parallel origins of bioluminescence in beetles. elife, 7, e36495.
8. Fuchsman, P., Fetters, K., & O'Connor, A. (2023). Target lipid model and empirical organic carbon partition coefficients predict sediment toxicity of polychlorinated biphenyls to benthic invertebrates. Environmental Toxicology and Chemistry, 42(5), 1134-1151.
9. Ge, S., Tian, W., Lou, Z., Wang, X., Zhuang, L. L., & Zhang, J. (2024). Long-term toxicity assessment of antibiotics against Vibrio fischeri: test method optimization and mixture toxicity prediction. Journal of Hazardous Materials, 469, 133933.
10. Genchi, G., Sinicropi, M. S., Lauria, G., Carocci, A., & Catalano, A. (2020). The effects of cadmium toxicity. International journal of environmental research and public health, 17(11), 3782.
11. Hao, Y., Duan, Z., Liu, L., Xue, Q., Pan, W., Liu, X., ... & Fu, J. (2025). Development of an Interpretable Machine Learning Model for Neurotoxicity Prediction of Environmentally Related Compounds. Environmental Science & Technology, 59(22), 11108-11120.

12. Kanwel, S., Gulzar, F., Alofaysan, H., Tanriverdiyev, S., & Jing, H. (2025). Toxic metal pollution in freshwater ecosystems: A systematic review of assessment methods using environmental and statistical indices. Marine Pollution Bulletin, 218, 118028.
13. Kato, M., Tsuchihashi, K., Kanie, S., Oba, Y., & Nishikawa, T. (2024). A practical, biomimetic, one-pot synthesis of firefly luciferin. Scientific Reports, 14(1), 30461.
14. Li, J., Liu, W., Liu, G., Dong, Z., He, J., Zhao, R., & Li, X. (2024). Cloning and characterization of luciferase from an Asian firefly Pygoluciolaqingyu and its comparison with other beetle luciferases. Photochemical & Photobiological Sciences, 23(4), 719-729.
15. Li, X., Qu, C., Bian, Y., Gu, C., Jiang, X., & Song, Y. (2019). New insights into the responses of soil microorganisms to polycyclic aromatic hydrocarbon stress by combining enzyme activity and sequencing analysis with metabolomics. Environmental Pollution, 255, 113312.
16. Liu, X. B., Lin, C., Wu, Y. Y., Huang, H. N., Zhu, L. T., Jiang, R., & Huang, Q. (2022). Dataset-based assessment of heavy metal contamination in freshwater fishes and their health risks. Environmental Science and Pollution Research, 29(33), 49985-49997.
17. Mangukiya, N. K., Mehta, D. J., & Jariwala, R. (2022). Flood frequency analysis and inundation mapping for lower Narmada basin, India. Water Practice & Technology, 17(2), 612-622.
18. Posthuma, L., Price, T., & Viljanen, M. (2025). Improving the Ecotoxicological Hazard Assessment of Chemicals by Pairwise Learning. Environmental Science & Technology, 59(31), 16250-16260.
19. Santos, M. M., Jorge, P. A. S., Coimbra, J., Vale, C., Caetano, M., Bastos, L., & Castro, L. F. C. (2018). The last frontier: coupling technological developments with scientific challenges to improve hazard assessment of deep-sea mining. Science of the Total Environment, 627, 1505-1514.
20. Strotmann, U., Pastor Flores, D., Konrad, O., & Gendig, C. (2020). Bacterial toxicity testing: modification and evaluation of the luminescent bacteria test and the respiration inhibition test. Processes, 8(11), 1349.