A Preliminary Study on Convolutional Neural Network–Based Classification of *Salvinia biloba* Raddi Growth Stages

Brenda Diniz [1], **Wellynton Diniz** [2], **Roselene M. Schneider** [1], **Leandro D. Battirola** [1], **Ricardo L. T. de Andrade** [3]

[1]Institute of Agricultural and Environmental Sciences
Universidade Federal de Mato Grosso (UFMT)
Sinop – MT – Brazil

[2]Department of Computer Engineering
Faculdade de Tecnologia de Sinop (FASTECH)
Sinop, MT– Brazil

[3]Institute of Natural, Human and Social Sciences
Universidade Federal de Mato Grosso (UFMT)
Sinop, MT – Brazil

brenda.benites@sou.ufmt.br, wellyntondiniz@gmail.com

{roselene.schneider, leandro.battirola, ricardo.andrade}@ufmt.br

Abstract. *Salvinia biloba Raddi (Salviniaceae) is used in phytoremediation systems and requires monitoring to avoid pollutant release during senescence. This paper reports a preliminary computer vision approach to classify three growth stages (young, intermediate, and advanced) from images using CNNs. We evaluated DenseNet, ResNet, MobileNet, and a custom CNN with transfer learning and data augmentation on a dataset of 312 labeled images. All models achieved accuracy above 89%, and MobileNet obtained the best performance (95.74% accuracy; 0.95 macro F1-score). Errors occurred only between adjacent stages, suggesting an ordinal pattern. Results indicate that lightweight CNNs can support automated monitoring in phytoremediation contexts.*

1. Introduction

The presence of organic and inorganic pollutants has significantly compromised the quality of water resources. In this context, concern has intensified regarding the availability of water under suitable conditions for use, especially in light of the increase in the generation and discharge of effluents into water bodies. This process contributes to environmental stress and the degradation of aquatic ecosystems due to the continuous input of pollutant loads [Ansari et al. 2020].

As a sustainable alternative for the treatment of these effluents, phytoremediation techniques stand out, as they use plants as biological agents for contaminant removal. This approach is characterized as a simple, low-cost, and environmentally friendly technology [Pawaiya and Suthar 2025].

Among the plant species employed in this process, aquatic macrophytes exhibit a high capacity for the removal of organic and inorganic contaminants [Wang et al. 2021, Choudhury et al. 2024]. As an example, species of the genus *Salvinia* (Salviniaceae) stand out, such as *Salvinia biloba* Raddi, which has been used in different contexts for pollutant removal [Zevallos et al. 2018, Freitas et al. 2025].

The application of these species integrates nature-based solutions (NbS) strategies, grounded in the use of ecosystem services to promote environmental conservation and the sustainable management of natural resources [Chairat and Gheewala 2024]. [Mancuso et al. 2021] emphasize that the use of NbS to mitigate diffuse contamination problems in water bodies represents a sustainable alternative, particularly with regard to nitrogen removal in impacted areas. However, the authors highlight the need for proper management of these systems to ensure their long-term efficiency.

In this context, to enable the application of aquatic plants in phytoremediation techniques, the adoption of appropriate management practices in effluent treatment systems becomes essential. The maintenance and monitoring of plant biomass are fundamental steps, since, upon reaching the senescence stage, plants may release back into the aquatic environment the pollutants previously assimilated, thereby compromising treatment efficiency [Kröger et al. 2007].

Macrophyte identification technologies based on image analysis can be employed in environmental monitoring, especially in the assessment of water resource quality. These tools enable the detection, classification, and monitoring of vegetation cover in aquatic environments, contributing to the management and conservation of these ecosystems [Levachou and Stonevičius 2025, Palanikkumar et al. 2025].

Considering this context, image-based identification techniques already used in the monitoring of water bodies show potential to be adapted to effluent treatment systems that employ macrophytes. The use of artificial intelligence (AI) tools to identify the degree of maturation and the physiological stage of these plants may contribute to the optimization of management practices and to the strengthening of NbS strategies.

Thus, the objective of this study is to conduct a preliminary investigation aimed at developing a model capable of determining the maturation stage of *Salvinia biloba* based on images.

For this purpose, four Convolutional Neural Network (CNN) models were evaluated: ResNet, DenseNet, MobileNet, and a custom CNN, using a dataset of 312 images distributed among maturation stages 1, 2, and 3. Model performance was analyzed using the metrics of precision, recall, F1-score, and accuracy.

The results demonstrated that all models achieved satisfactory performance, with accuracy above 88%. The best performance was obtained by MobileNet with transfer learning, achieving an accuracy of 95.74%, a macro F1-score of 0.9546, and the lowest loss value (0.1221). These findings highlight the potential of applying deep learning techniques for the automatic identification of the maturation stage of *Salvinia biloba*, contributing to the improvement of monitoring and management of macrophytes in NbS-based systems.

2. Related Work

The study conducted by [Francisco et al. 2024] present an evaluation of the maturation stage of Banana Prata Catarina (*Musa spp.* Musaceae) through the development of Convolutional Neural Network (CNN) models for the classification of different stages, using an image dataset of the cultivar.

In turn, [Boas et al. 2025] propose an AI-based approach for the classification of diseases in coffee leaves, integrating deep learning and computer vision techniques. The study employs CNNs with the application of transfer learning and data augmentation, using architectures such as DenseNet, MobileNet, and ResNet. Additionally, [Machado et al. 2025] investigated the automatic classification of coffee leaf rust, considering both binary and multiclass classification strategies. For this purpose, they evaluated different CNN architectures, including AlexNet, DenseNet, Inception, ResNet, SqueezeNet, and VGG.

For the image-based assessment of aquatic macrophytes, [Palanikkumar et al. 2025] propose a hybrid machine learning approach integrated with IoT, combining temporal analysis and data clustering. The Hybrid Machine Learning System (HMLS) achieved better performance when employing Graph Neural Networks (GNN) and Long Short-Term Memory (LSTM) networks.

In their study,[Bai and Bai 2024] applied CNNs to a dataset constructed through web crawling techniques for the acquisition of aquatic plant images. For modeling, they employed the DenseNet169 architecture to develop the EFL-DenseNet model. After pre-training on the ImageNet dataset, transfer learning was applied to adapt the model to the internal dataset intended for aquatic plant recognition.

In contrast, [Patil et al. 2025] employed pre-trained CNNs (ResNet50, VGG16, and InceptionV3) with transfer learning to more accurately identify the growth stages of the macrophyte *Eichhornia crassipes* (Mart.) Solms (Pontederiaceae). The models were integrated with a customized SimpleCNN through Ensemble Learning, enhancing classification robustness.

It is observed that the use of CNNs and transfer learning strategies for plant classification is widespread, enabling applicability and advancements in the automatic estimation of the maturation stage of aquatic macrophytes applied to phytoremediation.

3. Material and Methods

The methodology adopted in this study was structured into three main stages: (1) data collection, (2) model training, and (3) results evaluation. This organization was defined to ensure clarity in the experimental design and reproducibility of the process. In the first stage, the image dataset was organized and prepared. In the second stage, experiments were conducted using the DenseNet, MobileNet, ResNet, and CNNProper architectures, employing transfer learning and data augmentation techniques. Finally, in the third stage, the models were evaluated using quantitative metrics, enabling a systematic comparison of their performance. Figure 1 presents the overall workflow of the stages that compose the experimental procedure.

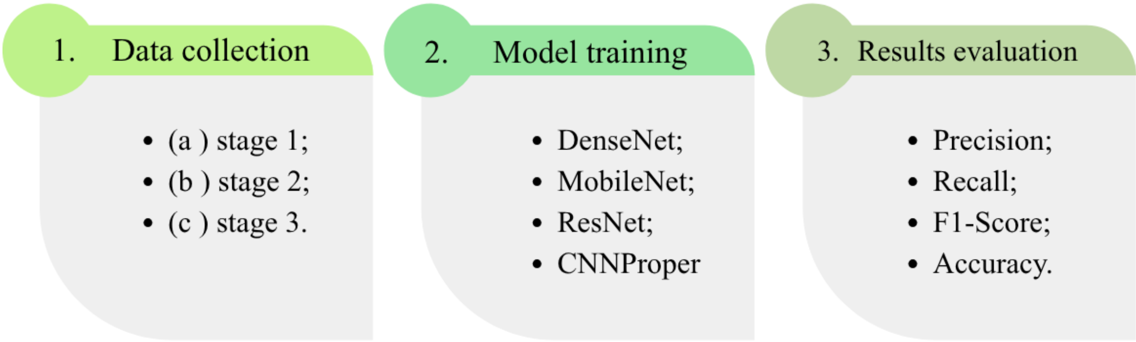

Figure 1. Steps of the methodology.

3.1. Data Collection

The present study used a proprietary dataset of *Salvinia biloba* images as the data source. The plants were cultivated in a polyethylene (PVC) tank with a capacity of 350 liters. For image acquisition, specimens at different maturation stages were selected, washed, dried, and transported to the laboratory, where they were placed on a white background to standardize the images. Image collection was carried out in January and February 2026, using a mobile phone as the capture device.

The repository is publicly available on the Zenodo platform (https://doi.org/10.5281/zenodo.18825469) and contains labeled images categorized into three maturation stages, classified as “stage 1,” “stage 2,” and “stage 3.” The dataset comprises 312 images, with 100 corresponding to stage 1 (young plants), 90 to stage 2 (early senescence stage), and 122 to stage 3 (advanced senescence stage), all with a resolution of 3024 × 4032 pixels.

The classification of maturation stages 1, 2, and 3 was performed based on leaf coloration, considering stage 1 as plants predominantly green in color; stage 2 as those presenting a mixture of tones ranging from green to brown; and stage 3 as plants predominantly brown, as shown in Figure 2. The selection and definition of the stages were carried out through visual inspection.

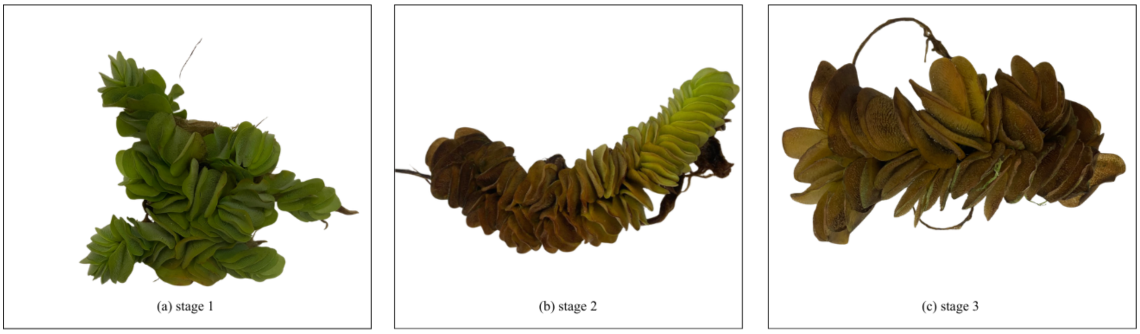

Figure 2. Samples of macrophytes exemplifying the observed maturation stages of the collected specimens: (a) stage 1, (b) stage 2, and (c) stage 3.

3.2. Model Training

The images were organized into separate directories according to their respective stages (1, 2, and 3). The dataset was split using a stratified approach, preserving the original proportion among the stages. A total of 70% of the images were allocated for training, 15% for validation, and 15% for testing. Stratification was applied to prevent statistical bias resulting from potential imbalance among the categories.

Four convolutional neural network architectures widely used in the literature were evaluated (DenseNet, MobileNet, ResNet, and CNNProper). All architectures were initialized with weights pre-trained on the ImageNet dataset, adopting a transfer learning strategy. The final classification layer was replaced with a new fully connected layer with two outputs, corresponding to the problem stages.

Due to the limited size of the dataset, data augmentation techniques were applied exclusively to the training set in order to increase sample variability and reduce the risk of overfitting. The transformations employed included random resized cropping, random rotation up to $\pm 20^{\circ}$, random horizontal flipping, slight variations in brightness, contrast, and saturation, as well as normalization using the statistical parameters of ImageNet.

The transformations were applied dynamically during batch loading (on-the-fly), without physically generating or storing additional images. Therefore, the original dataset size remained unchanged. However, considering the training subset and the 12 epochs, the model was exposed to approximately 2,616 augmented image variations during training. The validation and test sets were subjected only to resizing, center cropping, and normalization, ensuring an unbiased evaluation of performance.

Model training was conducted over 12 epochs using the AdamW optimizer with a learning rate set to 1×10^{-4} and a batch size of 16 images per iteration. All experiments were carried out using the PyTorch library, ensuring standardization in the implementation of the architectures and reproducibility of the results.

3.3. Results Evaluation

Model performance evaluation was based on the metrics *Accuracy* and *F1-score (macro)*. Considering the classification among stages *1*, *2*, and *3*, the following terms are defined: True Positive (TP), True Negative (TN), False Positive (FP), and False Negative (FN).

Accuracy Accuracy corresponds to the proportion of correct predictions relative to the total number of samples, as shown in Eq. 1.

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \tag{1}$$

Precision, Recall, and F1-score Precision and Recall for a given stage k are defined by Eqs. 2 and 3, respectively. The F1-score for stage k is defined by Eq. 4.

$$Precision_k = \frac{TP_k}{TP_k + FP_k} \tag{2}$$

$$Recall_k = \frac{TP_k}{TP_k + FN_k} \tag{3}$$

$$F1_k = \frac{2 \cdot Precision_k \cdot Recall_k}{Precision_k + Recall_k} \tag{4}$$

F1-score (macro) The macro F1-score is calculated as the arithmetic mean of the F1-score across stages, assigning equal weight to each category, as defined in Eq. 5. For this study, with three stages (*1*, *2*, and *3*), we have $K = 3$.

$$F1_{macro} = \frac{1}{K} \sum_{k=1}^{K} F1_k \tag{5}$$

The choice of the macro F1-score is justified by the imbalance among the stages, allowing for a balanced evaluation of performance across all categories. The results obtained from the four architectures (DenseNet, MobileNet, ResNet, and CNNProper) were quantitatively compared, enabling the identification of the architecture with the best generalization capability for the proposed problem.

4. Results and Discussion

4.1. Overall Performance of the Models

Table 1 presents the performance of the evaluated models in the scenario with three stages (1, 2, and 3). It is observed that MobileNet achieved the best overall performance, reaching an accuracy of 95.74% and a macro F1-score of 0.9546. DenseNet and ResNet showed intermediate performance, while the CNN trained from scratch obtained lower, yet competitive, results.

Table 1. Comparison between models on the test set.

Model	Type	Accuracy	F1 (Macro)	Loss
CNNProper	Without TL	0.8936	0.8889	0.2556
ResNet	TL	0.9149	0.9089	0.1438
DenseNet	TL	0.9362	0.9292	0.2220
MobileNet	**TL**	**0.9574**	**0.9546**	**0.1221**

The obtained results indicate that the image classification problem presents high visual separability, especially between the extreme stages (1 and 3). All evaluated models achieved accuracy above 89%, demonstrating that the visual characteristics associated with the different plant stages are sufficiently discriminative for supervised learning. MobileNet achieved the best overall performance, suggesting that lightweight architectures may be suitable for this type of task, including scenarios with computational constraints.

4.2. Analysis by Stage

Table 2 details the precision, recall, and F1-score metrics for each stage. It is observed that stage *3* achieved the highest recall values in nearly all models, indicating a strong ability to detect the more advanced stages of deterioration. Stage *2* showed greater variability among the models, highlighting increased difficulty in distinguishing this stage.

Table 2. Results by stage on the test set for each model.

Model	Stage	Precision	Recall	F1-score	Support
	1	0.93	0.87	0.90	15
ResNet	2	0.80	0.92	0.86	13
	3	1.00	0.95	0.97	19
	1	0.93	0.93	0.93	15
DenseNet	2	0.92	0.85	0.88	13
	3	0.95	1.00	0.97	19
	1	0.94	1.00	0.97	15
MobileNet	2	0.92	0.92	0.92	13
	3	1.00	0.95	0.97	19
	1	0.88	1.00	0.94	15
CNNProper	2	0.79	0.85	0.81	13
	3	1.00	0.84	0.91	19

The obtained results indicate that the image classification problem exhibits high visual separability, particularly between the extreme stages (1 and 3). All evaluated models achieved accuracy above 89%, demonstrating that the visual characteristics associated with the different plant stages are sufficiently discriminative for supervised learning. MobileNet achieved the best overall performance, suggesting that lightweight architectures may be well suited for this type of task, including scenarios with computational constraints.

4.3. Analysis of Confusion Matrices

Figure 3 presents the confusion matrices for the four evaluated models. It was observed that errors occurred exclusively between adjacent stages ($1 \leftrightarrow 2$ and $2 \leftrightarrow 3$), with no direct confusion between the extreme stages 1 and 3. This pattern suggests that the problem has an ordinal nature, with a gradual transition between the plant stages.

Figure 3. Confusion matrices obtained for the evaluated models on the test set.

4.4. Impact of Transfer Learning

The comparison between the CNN trained from scratch and the models using transfer learning highlights consistent performance gains. MobileNet showed an increase of approximately 6 percentage points in accuracy compared to the custom CNN. This result suggests that visual features previously learned from large-scale datasets contribute to better discrimination of stage 2 of the plant.

Stage 2 exhibited greater variability in recall and F1-score metrics across the models, indicating increased difficulty in separation. This behavior is consistent with the definition of this stage, which is characterized by partial yellowing, visually positioning it between the extreme stages. This result reinforces the hypothesis of morphological continuity among the plant stages.

Despite the numerical superiority of MobileNet, the 95% confidence intervals showed overlap among the models, indicating that the sample size still limits definitive statistical conclusions. Future studies with a larger volume of data may enable a more robust statistical comparison between architectures.

4.5. Implications for Macrophyte Management

The ability of the models to automatically classify the maturation stages of *Salvinia biloba* with high accuracy demonstrates the potential of computer vision as a decision-support tool in effluent treatment.

The developmental stage of aquatic macrophytes is directly related to their efficiency in nutrient uptake, growth rate, and pollutant removal capacity. Young plants (stage 1) tend to exhibit higher metabolic activity and greater nutrient assimilation potential, whereas more advanced stages of senescence (stage 3) may indicate reduced system efficiency and the need for management actions, such as harvesting or biomass replacement [Kröger et al. 2007, Choudhury et al. 2024].

In this context, the use of lightweight models such as MobileNet enables the implementation of automated monitoring systems through mobile devices or fixed cameras installed in treatment ponds. Such an approach would allow continuous monitoring of the maturation stages of macrophytes, reducing the subjectivity inherent to conventional visual assessment and optimizing the frequency of system interventions. Similar strategies have been explored for monitoring fruit maturation stages, as in the study by [Francisco et al. 2024] applied to bananas, highlighting the potential of computer vision as a decision-support tool in management practices.

Furthermore, the ordinal nature observed in the stage classification suggests that the transition process between phases can be progressively monitored, enabling the early identification of the onset of senescence. This information may contribute to more efficient management strategies, such as scheduled biomass removal and maintenance of vegetation cover rates within the systems. Thus, the results of this study not only validate the technical feasibility of applying CNNs to macrophyte classification but also indicate their potential as a tool to support intelligent monitoring of NbS-based systems.

5. Conclusion

This study investigated the performance of different CNN architectures in classifying plant images into three distinct stages: 1, 2, and 3. The results demonstrated that all evaluated architectures achieved satisfactory performance, highlighting the high visual separability among the analyzed stages.

Among the evaluated models, MobileNet achieved the best overall performance, followed by DenseNet and ResNet. The comparison with the CNN trained from scratch demonstrated that the use of transfer learning contributes to improved discrimination between adjacent stages, particularly for stage 2, which proved to be the most challenging.

The analysis of the confusion matrices revealed that errors occurred exclusively between neighboring stages, with no direct confusion between the extreme stages. This behavior suggests that the problem has an ordinal nature, characterized by a gradual transition between the levels of plant deterioration. This finding represents a relevant conceptual contribution, indicating that future approaches may explore ordinal classification models to further improve performance.

From an applied perspective, the results demonstrate the feasibility of using deep learning models for the automatic identification of the maturation stage of *Salvinia biloba*, representing an advancement in the application of computer vision to the monitoring of aquatic macrophytes. However, the study was conducted using individually isolated plants, not encompassing scenarios with continuous coverage, as commonly observed in natural environments or treatment ponds. In such contexts, the greater visual complexity may require additional approaches, such as segmentation techniques and large-scale analysis.

Thus, despite the promising results, the small size of the test set and the relative homogeneity of the image acquisition conditions constitute limitations of this study. Future work should consider larger datasets, greater environmental variability, and evaluation in real-world application scenarios.

The findings indicate that lightweight models, such as MobileNet, show strong potential for practical application in automated monitoring systems, including mobile devices or embedded platforms, contributing to rapid and objective assessment of the plant stage.

6. Acknowledgements

This study was supported by the Research Support Foundation of the State of Mato Grosso (FAPEMAT).

References

Ansari, A. A., Naeem, M., Gill, S. S., and AlZuaibr, F. M. (2020). Phytoremediation of contaminated waters: An eco-friendly technology based on aquatic macrophytes application. *Egyptian Journal of Aquatic Research*, 46(4):371–376.

Bai, Y. and Bai, X. (2024). Deep learning-based aquatic plant recognition technique and natural ecological aesthetics conservation. *Crop Protection*, 184:106765.

Boas, L. V., Rodrigues, L., e Silva, R. M., Lima, D., and Moreira, L. (2025). Ai-driven approach for digital agriculture: A case study on coffee leaf disease. In *Anais do XVI Workshop de Computação Aplicada à Gestão do Meio Ambiente e Recursos Naturais*, pages 1–10, Porto Alegre, RS, Brasil. SBC.

Chairat, S. and Gheewala, S. H. (2024). The conceptual quantitative assessment framework for nature-based solutions (nbs). *Nature-Based Solutions*, 6:100152.

Choudhury, M. I., Nilsson, J. E., Hylander, S., Hauber, M., Ehde, P. M., Weisner, S. E., and Liess, A. (2024). Enhancing nitrogen removal through macrophyte harvest and installation of woodchips-based floating beds in surface-flow constructed wetlands. *Chemosphere*, 359:142284.

Francisco, R., Pedroso, G., and Ventura, T. (2024). Aplicação de redes neurais convolucionais para classificação de imagens de estágios de maturação da banana prata catarina. In *Anais do XV Workshop de Computação Aplicada à Gestão do Meio Ambiente e Recursos Naturais*, pages 141–148, Porto Alegre, RS, Brasil. SBC.

Freitas, F., Solera, K., Ferreira, L. F., Sheng, L. Y., Moreno, M. I. C., Jacinto, M. J., Battirola, L. D., and de Andrade, R. L. T. (2025). Use of salvinia biloba raddi biomass in the remediation of solutions contaminated by nanoparticles and silver ions. *Brazilian Journal of Biology*, 85:e283123. Open Access.

Kröger, R., Holland, M., Moore, M., and Cooper, C. (2007). Plant senescence: A mechanism for nutrient release in temperate agricultural wetlands. *Environmental Pollution*, 146(1):114–119.

Levachou, Y. and Stonevičius, E. (2025). Near-infrared reflectance thresholding for macrophyte identification in temperate lakes using sentinel-2. *Environmental Conservation*, 52(4):239–244.

Machado, T., Rodrigues, L., Travençolo, B., Costa, C., Lima, D., and Moreira, L. (2025). Evaluation of convolutional neural networks for coffee leaf rust classification. In *Anais do XVI Workshop de Computação Aplicada à Gestão do Meio Ambiente e Recursos Naturais*, pages 11–20, Porto Alegre, RS, Brasil. SBC.

Mancuso, G., Bencresciuto, G. F., Lavrnić, S., and Toscano, A. (2021). Diffuse water pollution from agriculture: A review of nature-based solutions for nitrogen removal and recovery. *Water*, 13(14).

Palanikkumar, D., Anuradha, T., Ramalingam, J., and Sivaraju, S. (2025). A hybrid machine learning strategy for aquatic plant surveillance in sustainable aqua-ecosystems using iot attributes. *Aquaculture*, 609:742779.

Patil, M. L., Mehta, V., Sigtia, S., Pandey, G., and Kumari, S. (2025). Decoding water hyacinth growth stages with deep learning. In *2025 International Conference on Next Generation Communication Information Processing (INCIP)*, pages 692–696.

Pawaiya, A. and Suthar, S. (2025). Dual application of azolla- and wolffia-based floating constructed wetland in sewage treatment and phytomass production under the circular economy model. *Cleaner Water*, 4:100186.

Wang, J., Wang, W., Xiong, J., Li, L., Zhao, B., Sohail, I., and He, Z. (2021). A constructed wetland system with aquatic macrophytes for cleaning contaminated runoff/storm water from urban area in florida. *Journal of Environmental Management*, 280:111794.

Zevallos, W. T., Salvatierra, L. M., Loureiro, D. B., Morató, J., and Pérez, L. M. (2018). Evaluation of the autochthonous free-floating macrophyte salvinia biloba raddi for use in the phytoremediation of water contaminated with lead. *Desalination and Water Treatment*, 103:282–289.