

Article

Assessing Invasion Risk and Associated Economic Loss of Fish Species Across Global Watersheds Using Machine Learning Algorithms

Sang-Ik Suh ¹, Dahui Kim ¹, Bumseok Lee ², Soheon Lee ² and Seo Jin Ki ^{1,2,3,*}

¹ Department of Energy System Engineering, Gyeongsang National University, 8 Naedong-ro 139beon-gil, Jinju-si, Gyeongsangnam-do 52849, Republic of Korea; sangik.suh@gnu.ac.kr (S.-I.S.)

² Department of Environmental Engineering, Gyeongsang National University, 33 Dongjin-ro, Jinju-si, Gyeongsangnam-do 52725, Republic of Korea

³ Future Convergence Technology Research Institute, Gyeongsang National University, 8 Naedong-ro 139beon-gil, Jinju-si, Gyeongsangnam-do 52849, Republic of Korea

* Correspondence: seojinki@gnu.ac.kr; Tel.: +82-55-772-3341

Abstract

The intention of this study is to compare invasion risk and economic loss for selected fish species in both different watersheds and countries on a global scale. The potential risk of invasion for four different species was assessed by the prediction models developed from machine learning algorithms and joint dataset comprising species occurrence records and relevant environmental information. Combining invasion risk and unit economic costs derived from a comprehensive invasion cost database also enabled us to calculate the economic loss of invaders in three example countries. Results showed that the multi-layer perceptron algorithm, which was selected as the best model out of them, successfully identified areas of very-high-to-no risk for four selected fishes. The potential risk of invasion was relatively high for *Perccottus glenii* and low for *Salmo trutta* in the Republic of Korea, as compared to that of Japan and Australia. In addition, both invasion risk and unit economic costs were found to be largely responsible for the economic loss of selected fish species, specifically the species *Petromyzon marinus* in Australia. We expect that the methodology proposed in this study can be used to address the potential risk of invasion for non-native species along with other similar models such as species distribution models, and can be used to recommend affordable management options for controlling their populations.

Keywords: invasion risk assessment; economic cost estimation; machine learning; global database; non-native fish species; alien species management



Academic Editors: Yulin Kang, Baiyin Liu and Yan Chen

Received: 30 June 2026

Revised: 30 July 2026

Accepted: 31 July 2026

Published: 2 August 2026

Copyright: © 2026 by the authors.

Licensee MDPI, Basel, Switzerland.

This article is an open access article distributed under the terms and

conditions of the [Creative Commons Attribution \(CC BY\)](https://creativecommons.org/licenses/by/4.0/) license.

1. Introduction

Biological invasions by freshwater fish have become one of the major threats to freshwater ecosystems worldwide, leading to biodiversity loss, ecological degradation, a decline in ecosystem services, and increasing socioeconomic costs [1–4]. Invasive freshwater fish can establish self-sustaining populations, alter community structure, compete with native species, modify ecosystem functioning, and threaten both ecological integrity and human wellbeing [5–7]. Consequently, risk assessment for biological invasions has become an essential prerequisite for effective monitoring, prevention, and management of invaders of interest [6,8].

Predicting the potential occurrence of invasive species before establishment is regarded as one of the most effective approaches for supporting early detection and proactive

management of invaders. Species distribution models (SDMs) have been widely used to estimate habitat suitability based on relationships between environmental variables and species occurrence records [9–13]. Recent studies have further demonstrated the applicability of SDMs in predicting the potential distribution of invasive freshwater fish under changing environmental conditions [14,15]. Furthermore, advances in machine learning and deep learning have substantially improved the predictive performance of developed models by successfully identifying complex interactions between environmental variables and occurrence patterns for target species [16–20].

In addition, the rapid development of globally available environmental and biodiversity datasets has further expanded the applicability of species occurrence prediction. In particular, the EarthEnv database allowed us to retrieve standardized high-resolution environmental information for global freshwater resources in terms of climate, topography, land cover, and so on [21], while the global database for freshwater fish species occurrence contained over 100,000 occurrence records for around 15,000 fish species in over 3000 drainage basins to a near global extent [22]. In parallel with recent global studies on freshwater fish invasions [23], those datasets provide a robust foundation for developing transferable prediction models for species occurrence that can be applied across different geographic regions based on standardized environmental and biological information.

Meanwhile, increasing attention has also been directed toward quantifying the economic impacts of biological invasions. Recent studies have demonstrated that biological invasions imposed rapidly increasing economic costs worldwide [24,25], while freshwater invasive fish alone accounted for considerable economic losses across multiple countries and regions [26]. The study of national-scale assessments has further highlighted the substantial economic burden associated with invasive species introductions [27]. Currently, the InvaCost database, a standardized global repository of invasion-related economic losses, serves as the most comprehensive source for estimating the financial consequences of alien species [28]. Furthermore, methodological advances have facilitated standardized analyses of invasion costs [29], and economic assessment has become an increasingly important component of invasion management and policy development [30,31].

To the best of our knowledge, there has been limited research that assessed species occurrence risk and their corresponding economic costs simultaneously. Accordingly, the main objective of this study is to propose an integrated methodology or framework for risk assessment and economic loss estimation of non-native species across various spatial scales such as watersheds or countries. Using machine learning algorithms as well as three global datasets containing freshwater environmental information, species occurrence records, and economic costs of biological invasions, this study would specifically (1) identify the best machine learning model based on various evaluation metrics, as well as (2) screen invasion risk for four selected fish species and (3) evaluate their associated economic loss across three example countries, i.e., Republic of Korea, Japan, and Australia. We believe that this unified approach will allow for direct comparison of invasion risk and costs of target species in different geographic regions. Also, it can be used to prioritize management actions for alien species as well as to allocate conservation resources in a cost-effective way.

2. Materials and Methods

2.1. Compilation of Datasets for Risk Assessment and Loss Estimation

We compiled three datasets describing environmental variables and fish species occurrence in global freshwater (for invasion risk assessment) as well as economic costs of biological invasions worldwide (for economic loss estimation). The compiled dataset was then processed in the following way: (1) combining two different datasets associated with environmental variables and fish species occurrence for model development, (2) predicting

the probability of species occurrence using machine learning algorithms, (3) developing standard economic costs for all tested fish species according to unit area, and (4) estimating the economic loss across different geographic regions (e.g., watersheds or countries) by merging the results of (2) and (3).

Note that environmental variables used as independent variables in machine learning are retrieved from the EarthEnv database [21], which provides diverse sets of environmental variables for analyses of freshwater ecosystems to a near global extent at a spatial resolution of 1 km. The types of environmental variables provided in the database included climate, topography, land cover, geology, and soil. The total number of environmental variables calculated based on various metrics such as average, minimum, range, sum, and inverse distance-weighted average was 322, excluding two variables filling and removing certain grid cells under the quality control category.

Also, a global database on fish species occurrence is adopted as independent variables to predict the invasion probability of the selected fish species [22]. The database included 14,953 fish species that inhabited 3119 drainage basins accounting for more than 80% of the Earth's surface. In fact, the database, which was created from 1436 literature sources including books, papers, theses, and online sources, was organized into three sub-datasets in terms of the species occurrence records (consisting of six columns, including either the native or exotic status of the occurrence records), geographic information on drainage basins (consisting of nine columns given in table and shapefile formats), and information sources (consisting of three columns, including the type of source).

Figure 1a illustrates the distribution of global watersheds used for spatial representation of freshwater fish occurrence ($n = 3119$). In the figure, gray and black lines indicate the boundaries for individual watersheds and countries, respectively. Also, shaded areas in white ($n = 1928$) and dark green colors ($n = 551$) denote the watersheds without and with occurrence records for 43 target fishes, which are currently designated as invasive or potentially invasive species by the Ministry of Environment in the Republic of Korea. Here, those areas in beige color ($n = 640$) represent the watersheds additionally selected to reduce bias in machine learning models by enhancing their spatial coverage across all watersheds of the world. Figure 1b–d display the watershed boundaries for three example countries (i.e., the Republic of Korea, Japan, and Australia) evaluated in this study. Specifically, we selected two additional countries, Japan and Australia, which have a similarity and dissimilarity to the Republic of Korea in terms of environmental characteristics for comparison of invasion risk among the countries. Figure 1e is inset, exhibiting the annual minimum water temperature in the selected watershed.

The economic loss across different watersheds was computed based on the probability of species occurrence for each watershed (predicted from machine learning algorithms), unit economic costs (adjusted by area from the global loss database caused by biological invasions, InvaCost), and the area of each watershed, which can be written as:

$$Loss_i = P_{occurrence, i} \times C_{unit, i} \times A_{watershed}$$

where $C_{unit, i} = \sum_{j=1}^n Cost_j / \sum_{j=1}^n Area_j$.

Here, the terms $Loss$, $P_{occurrence}$, C_{unit} , and $A_{watershed}$ indicate the total economic cost for each watershed, the probability of species occurrence for each watershed, the normalized economic cost (USD/km²), and the area of each watershed (km²), respectively. The subscripts i and j signify individual fish species and different literature sources (for the same species), respectively.

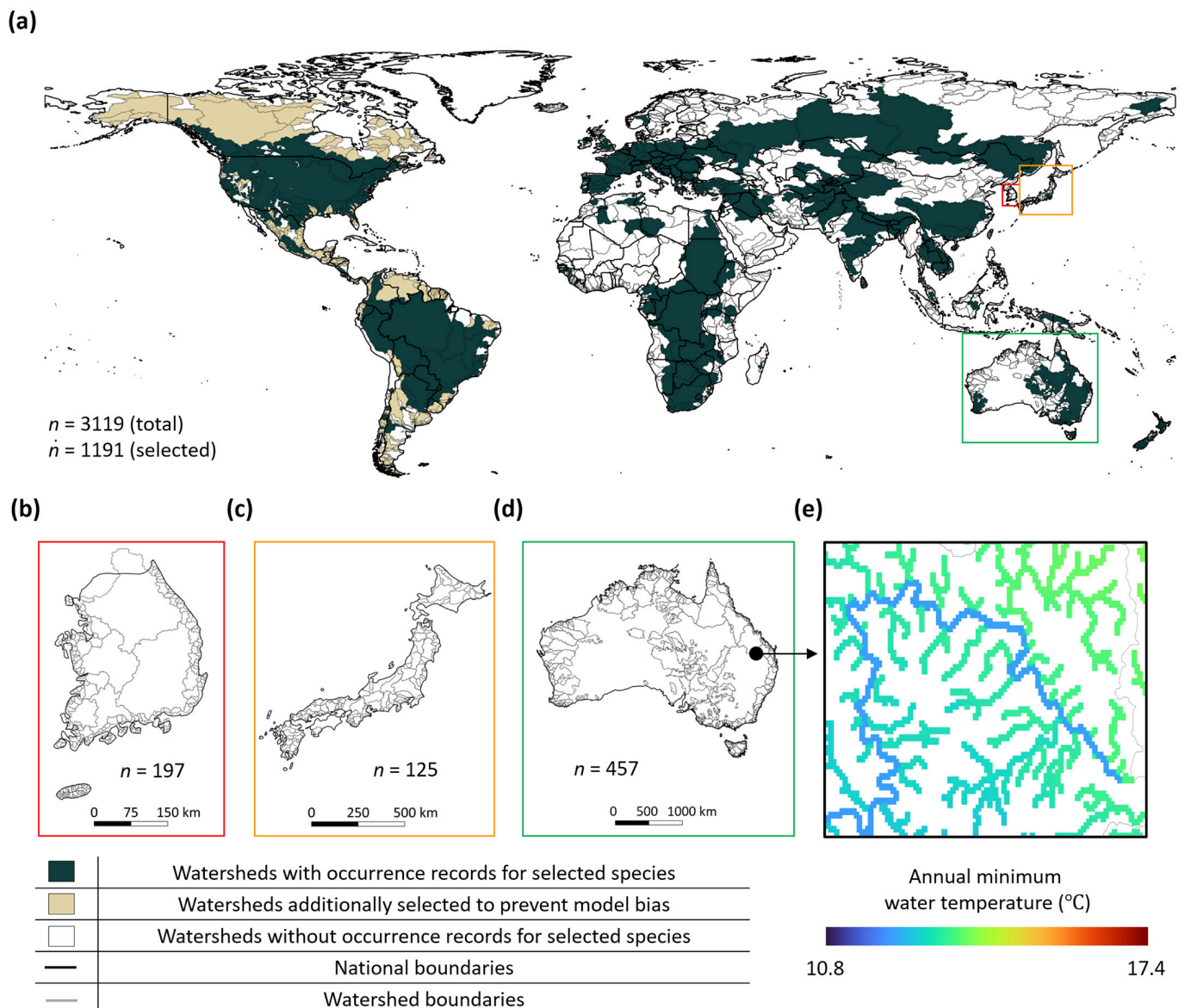


Figure 1. Maps showing the boundaries of watersheds in (a) the world, (b) Republic of Korea, (c) Japan, and (d) Australia as well as (e) a particular environmental variable in the selected watershed. Note that the databases on freshwater fish species occurrence, including global watersheds in shapefile, and environmental variables are available at https://doi.org/10.6084/m9.figshare.c.3739145_D5 (accessed on 15 June 2026) and <https://www.earthenv.org/streams> (accessed on 15 June 2026), respectively.

Briefly, InvaCost contained 2419 cost estimates for 343 invasive species, which belonged to various taxonomic kingdoms such as Animalia, Bacteria, and Chromista, across 90 countries [28]. All recorded costs in the database were converted to 2017 US Dollars based on two conversion methods, market exchange rates and purchasing power parity, to allow for global comparisons. Also, as costs in the database were provided at various spatial scales such as continent, country, basin, and river, we then transformed those original cost estimates to cost estimates per unit area (USD/km²). Note that when information on the spatial coverage for a particular cost was not available, we used official area data for target geographic regions readily available online. Table 1 exhibits unit economic costs for four target fish species, which are normalized by the entire area in different regional studies. As shown in the table, the unit economic costs vary substantially across selected species due to the area of occurrence and associated economic loss for each species. More

information on the selected fish species in terms of biological characteristics and occurrence records in three different example countries is also given in Table 2.

Table 1. Unit economic costs calculated for four target fish species ¹.

Fish Species	The Number of Regional Studies Compiled	Areas (Units: km ²)	Cost Estimates Converted to USD for the Year 2017 (Units: USD)	Unit Economic Costs (Units: USD/km ²)
<i>Morone chrysops</i>	1	47,000	3,393,637.15	72.21
<i>Perccottus glenii</i>	1	65,300	7337.83	0.11
<i>Petromyzon marinus</i>	10	244,106	848,409,287.57	3475.58
<i>Salmo trutta</i>	2	340	188,535.40	554.52

Note: ¹ The original dataset can be downloaded at <https://doi.org/10.6084/m9.figshare.12668570> (accessed on 15 June 2026).

Table 2. Biological characteristics and occurrence records of four selected fish species in three example countries, Republic of Korea, Japan, and Australia ¹.

Scientific Names	Common Names	Max Total Length (cm)	Max Weight (kg)	Temperature Range (°C)	Occurrence Records		
					Korea	Japan	Australia
<i>Morone chrysops</i>	White bass	45	3.1	12–32	–	–	–
<i>Perccottus glenii</i>	Amur sleeper	25	0.25	15–30	✓	–	–
<i>Petromyzon marinus</i>	Sea lamprey	120	2.5	1–20	–	–	–
<i>Salmo trutta</i>	Brown trout	140	50.0	18–24	–	✓	✓

Note: ¹ All data presented is obtained from <https://www.fishbase.se/search.php> (accessed on 15 June 2026).

2.2. Machine Learning Algorithms for Invasion Risk Prediction

Invasion risk for selected fish species (i.e., the probability of species occurrence) was predicted by four different machine learning algorithms, extreme gradient boosting (XGBoost), gradient boosted decision tree (GBDT), multi-layer perceptron (MLP), and random forest (RF), which are widely adopted in ecology. In particular, XGBoost was found to show excellent performance on tabular datasets, whereas GBDT and MLP, which worked well for both regression and classification tasks, were successfully capable of capturing complex, non-linear patterns in highly dimensional datasets [32]. RF, which required relatively little parameter tuning, was often claimed to be an ensemble model with strong generalization performance and robustness to noise [33].

All machine learning models adopted for this study were implemented using scikit-learn (version 1.0.2), a machine learning package in Python (version 3.7.10). Note that dependent and independent variables were normalized using the MinMaxScaler in scikit-learn so that all variables had a similar influence on the model. We randomly allocated 80% of the entire dataset for training and the remaining 20% for testing. The developed models were evaluated using three important metrics typically applied for classification problems in machine learning, i.e., accuracy, F1 Score, and Receiver Operating Characteristic–Area Under the Curve (ROC-AUC). Five-fold cross-validation was also used to fairly assess the performance of all tested models. We adopted Latin hypercube sampling, which explored the search space more efficiently than random and grid search, to tune hyperparameters of all machine learning models using the pyDOE2 package (version 1.3.0). A list of hyperparameters and their search space defined for hyperparameter optimization in different machine learning algorithms are given in Table 3.

Table 3. The search range and best values adopted for hyperparameter optimization in four different machine learning algorithms.

Algorithms	Hyperparameters	Data Types	Search Range	Best Values
XGBoost	n_estimators	integer	15–40	40
	learning_rate	log continuous	0.01–0.20	5.19×10^{-2}
	max_depth	integer	2–3	2
	min_child_weight	log continuous	0.5–15	6.50×10^{-1}
	subsample	continuous	0.60–1.00	7.78×10^{-1}
	colsample_bytree	continuous	0.50–1.00	7.36×10^{-1}
	gamma	continuous	0–3	1.46×10^{-1}
	reg_alpha	log continuous	0.0001–3	5.39×10^{-4}
GBDT	reg_lambda	log continuous	0.01–30	2.53×10^{-2}
	max_iter	integer	20–60	53
	learning_rate	log continuous	0.01–0.20	1.37×10^{-2}
	max_leaf_nodes	integer	7–31	25
	min_samples_leaf	integer	10–40	21
	l2_regularization	log continuous	0.0001–3	4.06×10^{-3}
MLP	max_bins	integer	64–255	180
	learning_rate_init	log continuous	0.00003–0.01	2.85×10^{-3}
	alpha	log continuous	0.000001–0.01	3.46×10^{-4}
	batch_size	categorical	16, 32, 64, 128	16
	max_iter	fixed	30	30
RF	early_stopping	fixed	True	True
	n_estimators	integer	150–350	173
	max_depth	integer	5–20	14
	min_samples_split	integer	2–20	4
	min_samples_leaf	integer	1–10	3
	max_features	categorical	sqrt, log2, 0.25, 0.50, 0.75	log2
	class_weight	categorical	balanced, balanced_subsample	balanced_subsample

3. Results and Discussion

3.1. Performance Assessment of Machine Learning Models

Figure 2 compares the performance of four machine learning models, XGBoost, GBDT, MLP, and RF, for test datasets in predicting the probability of four target fish species occurrence, based on three evaluation metrics. Note that the scores of all three metrics range from 0 to 1, where higher (close to 1) and lower values (close to 0) indicate better- and worse-performing models, respectively, regardless of evaluation metrics. The box plot in the figure also shows the results of five-fold cross-validation for different machine learning models. If the interquartile range, the difference or the range between the upper quartile (or Q3) and the lower quartile (or Q1) in the box plot is narrow, then there is low variability among the results. It was found from the figure that MLP outperformed other machine learning models, specifically in terms of accuracy (i.e., the percentage of correct predictions out of all predictions) and F1 Score (i.e., a single metric balancing precision and recall). Also, the performance of GBDT was comparable to that of MLP only with respect to ROC-AUC, which was known as a more robust measure than accuracy on imbalanced datasets. Interestingly, the predictive performance of RF was not as good as that of MLP, specifically in terms of both F1 Score and ROC-AUC. Overall, MLP was found to achieve the highest prediction performance out of four tested models. Therefore, we only adopted the prediction model developed from MLP in assessing the occurrence probability (i.e., invasion risk) of four target fish species. Table 3 also displays a series of optimal

hyperparameter values for different machine learning models, which are determined prior to their performance assessment. Note that no obvious differences are observed between two models which include all independent variables and the top 10 variables (selected based on feature importance), regardless of four tested algorithms, in terms of accuracy (-0.0064 – 0.0072), F1 Score (-0.0033 – 0.0023), and ROC-AUC (-0.0024 – 0.0013).

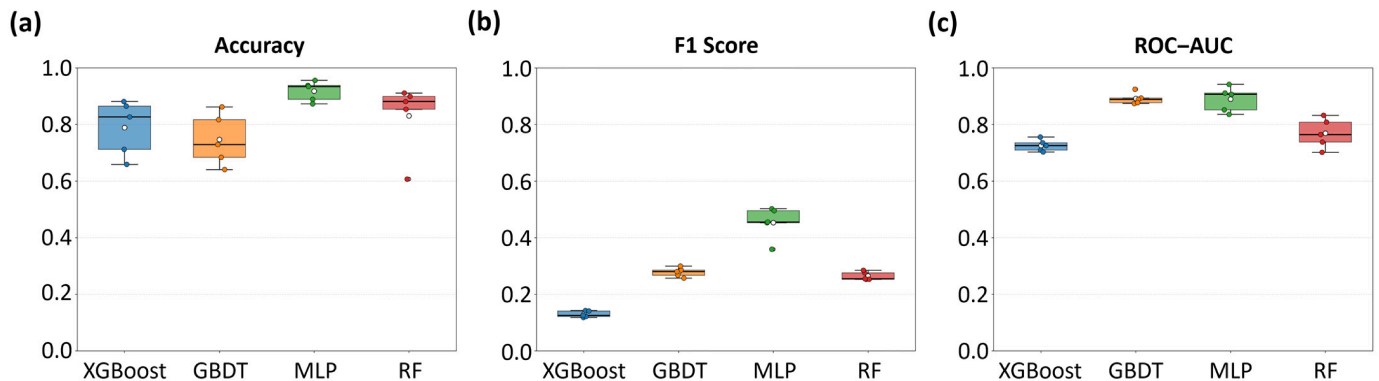


Figure 2. Performance comparison of four machine learning algorithms for test datasets with respect to (a) Accuracy, (b) F1 Score, and (c) ROC-AUC.

3.2. Invasion Risk for Target Fish Species

Figure 3 exhibits the invasion risk of two target fish species (i.e., *Perccottus glenii* and *Salmo trutta*) out of four species for various watersheds in three different geographic regions (i.e., the Republic of Korea, Japan, and Australia) using the best prediction model which is chosen in the previous section. In the figure, the potential risk of invasion for different species is indicated by a color index: 0 (no risk), 0–20% (very low), 20–40% (low), 40–60% (medium), 60–80% (high), and 80–100% (very high). A gray line indicates the boundaries for individual watersheds distributed in each country. Out of two fish species tested, *Perccottus glenii* presented slightly higher invasion risk than *Salmo trutta* in three example countries. However, the potential risk of invasion for *Perccottus glenii* was much higher in the Republic of Korea than in Japan and Australia. In addition, while high invasion risk for *Perccottus glenii* was observed in southern and southwestern watersheds in Japan, its invasion risk was only high in the upper parts of Australia. In contrast, the potential risk of invasion for *Salmo trutta* was estimated to be very high only in a few watersheds in Australia and Japan, as well as very low to negligible in the Republic of Korea. All these results revealed that invasion risk considerably varied according to various spatial scales (i.e., watersheds and countries) and target fish species, which was likely attributed to different environmental characteristics (in regions of interest), habitat suitability (in species of interest), and a combination of both.

3.3. Economic Loss for Target Fish Species

Table 4 presents the total economic costs likely caused by the invasion risk of four different fish species across different watersheds in three example countries. Note that the average invasion risk for each fish species calculated by country is provided in the table in order to give a better understanding of the differences in invasion risk among countries and species, so it is not used for estimation of the total economic costs. It should also be noted that invasion risk for target species differs significantly across individual watersheds (see Figure 3). In the table, the relative rank by loss is also indicated by a color index: 0 (no loss), 0– 10^2 (very low), 10^2 – 10^4 (low), 10^4 – 10^5 (medium), 10^5 – 10^6 (high), and $\geq 10^6$ (very high). As shown in the table, the economic loss for *Petromyzon marinus* was very high in Australia, which was followed by *Salmo trutta* in Australia (classified as “high”), *Salmo trutta* in Japan

(classified as “medium”), *Morone chrysops* in Japan (classified as “medium”), and so on. Also, the species *Perccottus glenii* was shown to yield very-low-to-negligible economic loss in three tested countries. These results implied that the economic loss for target fish species was greatly affected by both invasion risk in different regions and unit economic costs. From these results, we were also able to identify areas of concern (e.g., watersheds or countries) which were highly vulnerable to invaders and to prioritize management options for either their control or eradication.

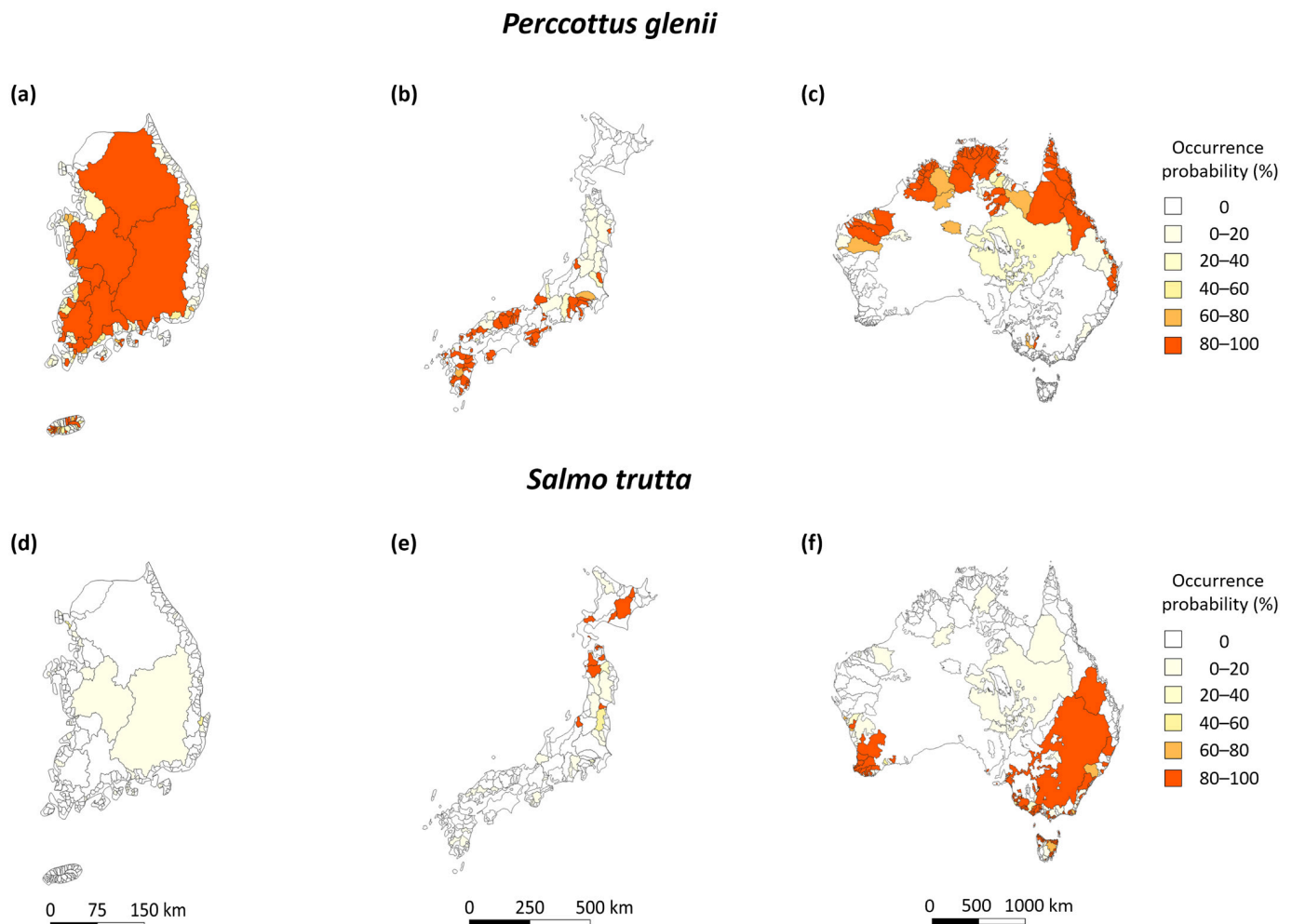


Figure 3. The probability of species occurrence for two example fish species, (a–c) *Perccottus glenii* and (d–f) *Salmo trutta*, in various watersheds in three different countries (i.e., Republic of Korea, Japan, and Australia) which are predicted by MLP.

Table 4. Economic loss and its relative rank for four fish species in three example countries estimated based on the invasion risk for individual watersheds, unit economic costs, and entire watershed areas with invasion risk.

Species	Countries	Invasion Risk on Average (by Country)	Unit Economic Costs (USD/km ²)	Entire Watershed Areas with Invasion Risk (km ²)	Total Economic Costs (USD)	Relative Rank by Loss ¹
		$P_{occurrence}$	C_{unit}	$A_{watershed}$	$Loss$	
<i>Morone chrysops</i>	Korea	0.021	72.21	201.80	304.55	Low
	Japan	0.894	72.21	375.00	24,169.07	Medium
	Australia	0.000	72.21	0	0	No loss

Table 4. Cont.

Species	Countries	Invasion Risk on Average (by Country)	Unit Economic Costs (USD/km ²)	Entire Watershed Areas with Invasion Risk (km ²)	Total Economic Costs (USD)	Relative Rank by Loss ¹
		$P_{occurrence}$	C_{unit}	$A_{watershed}$	$Loss$	
<i>Percottus glenii</i>	Korea	0.004	0.11	181.80	0.08	Very low
	Japan	0.714	0.11	374.00	29.38	Very low
	Australia	0.000	0.11	0	0	No loss
<i>Petromyzon marinus</i>	Korea	0.306	3475.58	0.20	214.50	Low
	Japan	0.000	3475.58	0	0	No loss
	Australia	0.813	3475.58	10,026.00	28,300,807.74	Very high
<i>Salmo trutta</i>	Korea	0.006	554.52	202.00	673.38	Low
	Japan	0.209	554.52	374.00	43,403.54	Medium
	Australia	0.523	554.52	1367.00	396,192.80	High

Note: ¹ 0 = No loss, 0–10² = very low, 10²–10⁴ = low, 10⁴–10⁵ = medium, 10⁵–10⁶ = high, and $\geq 10^6$ = very high.

4. Conclusions

This study was done to introduce a harmonized approach of invasion risk assessment and corresponding economic loss estimation, which could be implemented consistently at various spatial scales such as watershed, city, country, and even continent. Two datasets containing species occurrence records and environmental variables on inland water systems at the global scale were used to develop prediction models for the invasion risk of selected fish species with different machine learning algorithms. The economic loss at the watershed level was computed based on the invasion risk of selected species (predicted by the best machine learning model) and unit economic costs (which are normalized by the accumulated area of different studies from the global loss dataset of biological invasions). The main findings of this study are as follows.

- The predictive performance of MLP was superior to its competitors (i.e., XGBoost, GBDT, and RF) in terms of three evaluation metrics (i.e., accuracy, F1 Score, and ROC-AUC). MLP also showed quite stable performance across five-fold cross-validation.
- The potential risk of invasion (predicted by MLP) was much higher for *Percottus glenii* than for *Salmo trutta*, in three examined countries (i.e., the Republic of Korea, Japan, and Australia). However, we also observed various levels of invasion according to different regions and species.
- The economic loss for target fish species was highly sensitive to invasion risk in individual watersheds and unit economic costs. The species *Petromyzon marinus* was found to result in the largest economic loss in Australia, compared to those of other tested species in three different countries.
- Additional research is still required to explore the benefits and drawbacks of our proposed approach compared to other similar models such as SDMs as well as datasets provided at different spatial resolutions.

Author Contributions: Conceptualization, S.-I.S.; Methodology, S.-I.S.; Software, S.-I.S.; Validation, S.-I.S.; Investigation, D.K., B.L. and S.L.; Resources, S.-I.S.; Data curation, S.-I.S., D.K., B.L. and S.L.; Writing—original draft preparation, S.-I.S. and S.J.K.; Writing—review and editing, S.-I.S. and S.J.K.; Supervision, S.J.K.; Project administration, S.-I.S. and S.J.K.; Funding acquisition, S.J.K. All authors have read and agreed to the published version of the manuscript.

Funding: This study was supported by the Korea Environment Industry & Technology Institute (KEITI) through the Exotic Invasive Species Management Program, funded by the Korea Ministry of Environment (MOE) [RE201807019].

Data Availability Statement: The original contributions presented in this study are included in the article. Further inquiries can be directed to the corresponding author.

Conflicts of Interest: The authors declare no conflicts of interest.

References

1. Dudgeon, D.; Arthington, A.H.; Gessner, M.O.; Kawabata, Z.-I.; Knowler, D.J.; Lévêque, C.; Naiman, R.J.; Prieur-Richard, A.-H.; Soto, D.; Stiassny, M.L.J.; et al. Freshwater biodiversity: Importance, threats, status and conservation challenges. *Biol. Rev.* **2006**, *81*, 163–182. [\[CrossRef\]](#) [\[PubMed\]](#)
2. Ricciardi, A.; Rasmussen, J.B. Extinction rates of North American freshwater fauna. *Conserv. Biol.* **1999**, *13*, 1220–1222. [\[CrossRef\]](#)
3. Reid, A.J.; Carlson, A.K.; Creed, I.F.; Eliason, E.J.; Gell, P.A.; Johnson, P.T.J.; Kidd, K.A.; MacCormack, T.J.; Olden, J.D.; Ormerod, S.J.; et al. Emerging threats and persistent conservation challenges for freshwater biodiversity. *Biol. Rev.* **2019**, *94*, 849–873. [\[CrossRef\]](#) [\[PubMed\]](#)
4. Tickner, D.; Opperman, J.J.; Abell, R.; Acreman, M.; Arthington, A.H.; Bunn, S.E.; Cooke, S.J.; Dalton, J.; Darwall, W.; Edwards, G.; et al. Bending the curve of global freshwater biodiversity loss: An emergency recovery plan. *BioScience* **2020**, *70*, 330–342. [\[CrossRef\]](#) [\[PubMed\]](#)
5. Pyšek, P.; Hulme, P.E.; Simberloff, D.; Bacher, S.; Blackburn, T.M.; Carlton, J.T.; Dawson, W.; Essl, F.; Foxcroft, L.C.; Genovesi, P.; et al. Scientists’ warning on invasive alien species. *Biol. Rev.* **2020**, *95*, 1511–1534. [\[CrossRef\]](#) [\[PubMed\]](#)
6. Bernery, C.; Bellard, C.; Courchamp, F.; Brosse, S.; Gozlan, R.E.; Jarić, I.; Teletchea, F.; Leroy, B. Freshwater fish invasions: A comprehensive review. *Annu. Rev. Ecol. Evol. Syst.* **2022**, *53*, 427–456. [\[CrossRef\]](#)
7. Strayer, D.L. Alien species in fresh waters: Ecological effects, interactions with other stressors, and prospects for the future. *Freshw. Biol.* **2010**, *55*, 152–174. [\[CrossRef\]](#)
8. Gallardo, B.; Clavero, M.; Sánchez, M.I.; Vilà, M. Global ecological impacts of invasive species in aquatic ecosystems. *Glob. Change Biol.* **2016**, *22*, 151–163. [\[CrossRef\]](#) [\[PubMed\]](#)
9. Elith, J.; Leathwick, J.R. Species distribution models: Ecological explanation and prediction across space and time. *Annu. Rev. Ecol. Evol. Syst.* **2009**, *40*, 677–697. [\[CrossRef\]](#)
10. Franklin, J. *Mapping Species Distributions: Spatial Inference and Prediction*; Cambridge University Press: Cambridge, UK, 2010. [\[CrossRef\]](#)
11. Phillips, S.J.; Anderson, R.P.; Schapire, R.E. Maximum entropy modeling of species geographic distributions. *Ecol. Model.* **2006**, *190*, 231–259. [\[CrossRef\]](#)
12. Thuiller, W. BIOMOD—Optimizing predictions of species distributions and projecting potential future shifts under global change. *Glob. Change Biol.* **2003**, *9*, 1353–1362. [\[CrossRef\]](#)
13. Srivastava, V.; Lafond, V.; Griess, V.C. Species distribution models (SDM): Applications, benefits and challenges in invasive species management. *CAB Rev.* **2019**, *14*, 020. [\[CrossRef\]](#)
14. Kim, Z.; Shim, T.; Koo, Y.-M.; Seo, D.; Kim, Y.-O.; Hwang, S.-J.; Jung, J. Predicting the impact of climate change on freshwater fish distribution by incorporating water flow rate and quality variables. *Sustainability* **2020**, *12*, 10001. [\[CrossRef\]](#)
15. Hong, S.; Jang, I.; Kim, D.; Kim, S.; Park, H.S.; Lee, K. Predicting potential habitat changes of two invasive alien fish species with climate change at a regional scale. *Sustainability* **2022**, *14*, 6093. [\[CrossRef\]](#)
16. Nekrasova, O.; Tytar, V.; Pupins, M.; Čeirāns, A.; Marushchak, O.; Skute, A. A GIS modeling study of the distribution of viviparous invasive alien fish species in Eastern Europe in terms of global climate change, as exemplified by *Poecilia reticulata* Peters, 1859 and *Gambusia holbrooki* Girard, 1859. *Diversity* **2021**, *13*, 385. [\[CrossRef\]](#)
17. Deneu, B.; Servajean, M.; Bonnet, P.; Botella, C.; Munoz, F.; Joly, A. Convolutional neural networks improve species distribution modelling by capturing the spatial structure of the environment. *PLoS Comput. Biol.* **2021**, *17*, e1008856. [\[CrossRef\]](#) [\[PubMed\]](#)
18. Norberg, A.; Abrego, N.; Blanchet, F.G.; Adler, F.R.; Anderson, B.J.; Anttila, J.; Araújo, M.B.; Dallas, T.; Dunson, D.; Elith, J.; et al. A comprehensive evaluation of predictive performance of 33 species distribution models at species and community levels. *Ecol. Monogr.* **2019**, *89*, e01370. [\[CrossRef\]](#)
19. Høye, T.T.; Årje, J.; Bjerger, K.; Hansen, O.L.P.; Iosifidis, A.; Leese, F.; Mann, H.M.R.; Meissner, K.; Melvad, C.; Raitoharju, J. Deep learning and computer vision will transform entomology. *Proc. Natl. Acad. Sci. USA* **2021**, *118*, e2002545117. [\[CrossRef\]](#) [\[PubMed\]](#)
20. Pichler, M.; Hartig, F. Machine learning and deep learning—A review for ecologists. *Methods Ecol. Evol.* **2023**, *14*, 994–1016. [\[CrossRef\]](#)
21. Domisch, S.; Amatulli, G.; Jetz, W. Near-global freshwater-specific environmental variables for biodiversity analyses in 1 km resolution. *Sci. Data* **2015**, *2*, 150073. [\[CrossRef\]](#) [\[PubMed\]](#)

22. Tedesco, P.A.; Beauchard, O.; Bigorne, R.; Blanchet, S.; Buisson, L.; Conti, L.; Cornu, J.-F.; Dias, M.S.; Grenouillet, G.; Hugueny, B.; et al. A global database on freshwater fish species occurrence in drainage basins. *Sci. Data* **2017**, *4*, 170141. [[CrossRef](#)] [[PubMed](#)]
23. Xu, M.; Li, S.-P.; Liu, C.; Tedesco, P.A.; Dick, J.T.A.; Fang, M.; Wei, H.; Yu, F.; Shu, L.; Wang, X.; et al. Global freshwater fish invasion linked to the presence of closely related species. *Nat. Commun.* **2024**, *15*, 1411. [[CrossRef](#)] [[PubMed](#)]
24. Diagne, C.; Leroy, B.; Vaissière, A.-C.; Gozlan, R.E.; Roiz, D.; Jarić, I.; Salles, J.-M.; Bradshaw, C.J.A.; Courchamp, F. High and rising economic costs of biological invasions worldwide. *Nature* **2021**, *592*, 571–576. [[CrossRef](#)] [[PubMed](#)]
25. Cuthbert, R.N.; Pattison, Z.; Taylor, N.G.; Verbrugge, L.; Diagne, C.; Ahmed, D.A.; Leroy, B.; Angulo, E.; Briski, E.; Capinha, C.; et al. Global economic costs of aquatic invasive alien species. *Sci. Total Environ.* **2021**, *775*, 145238. [[CrossRef](#)] [[PubMed](#)]
26. Haubrock, P.J.; Bernery, C.; Cuthbert, R.N.; Liu, C.; Kourantidou, M.; Leroy, B.; Turbelin, A.J.; Kramer, A.M.; Verbrugge, L.N.; Diagne, C.; et al. Knowledge gaps in economic costs of invasive alien fish worldwide. *Sci. Total Environ.* **2022**, *803*, 149875. [[CrossRef](#)] [[PubMed](#)]
27. Watari, Y.; Komine, H.; Angulo, E.; Diagne, C.; Ballesteros-Mejia, L.; Courchamp, F. First synthesis of the economic costs of biological invasions in Japan. *NeoBiota* **2021**, *67*, 79–101. [[CrossRef](#)]
28. Diagne, C.; Leroy, B.; Gozlan, R.E.; Vaissière, A.-C.; Assailly, C.; Nuninger, L.; Roiz, D.; Jourdain, F.; Jarić, I.; Courchamp, F. InvaCost, a public database of the economic costs of biological invasions worldwide. *Sci. Data* **2020**, *7*, 277. [[CrossRef](#)] [[PubMed](#)]
29. Leroy, B.; Kramer, A.M.; Vaissière, A.-C.; Kourantidou, M.; Courchamp, F.; Diagne, C. Analysing economic costs of invasive alien species with the invacost R package. *Methods Ecol. Evol.* **2022**, *13*, 1930–1937. [[CrossRef](#)]
30. Epanchin-Niell, R.S. Economics of invasive species policy and management. *Biol. Invasions* **2017**, *19*, 3333–3354. [[CrossRef](#)]
31. Lodge, D.M.; Simonin, P.W.; Burgiel, S.W.; Keller, R.P.; Bossenbroek, J.M.; Jerde, C.L.; Kramer, A.M.; Rutherford, E.S.; Barnes, M.A.; Wittmann, M.E.; et al. Risk analysis and bioeconomics of invasive species to inform policy and management. *Annu. Rev. Environ. Resour.* **2016**, *41*, 453–488. [[CrossRef](#)]
32. Yu, Z.; Bu, C.; Li, Y. Machine learning for ecological analysis. *Chem. Eng. J.* **2025**, *507*, 160780. [[CrossRef](#)]
33. Allen Akselrud, C.I. Random forest regression models in ecology: Accounting for messy biological data and producing predictions with uncertainty. *Fish. Res.* **2024**, *280*, 107161. [[CrossRef](#)]

Disclaimer/Publisher’s Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.