# Assessing Anthropogenic Stressors to Global Freshwater Habitats and Inland Fisheries

Gretchen L. Stokes[1], Abigail J. Lynch[2], John V. Flores[1], Jesse P. Wong[3], Connor A. Morang[1], Chelsie Romulo[4], Simon Funge-Smith[5], John Valbo-Jorgensen[5] & Samuel J. Smidt[1]

[1]University of Florida, [2]U.S. Geological Survey, [3]George Mason University, [4]University of Northern Colorado, [5]Food and Agriculture Organization of the United Nations

UF Water Institute Symposium

## 1 INTRODUCTION

Anthropogenic factors like land development and overharvest are **transforming global ecosystems** at unprecedented rates. Addressing these factors with effective management requires information on the **links** between **human activities** and **ecological impacts.**

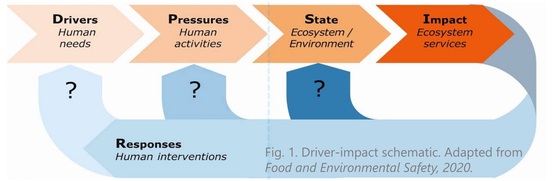



Fig. 1. Driver-impact schematic. Adapted from *Food and Environmental Safety, 2020.*

Systematic literature reviews remain an important approach for synthesizing documented factors. **Literature synthesis** has been dependent on manual efforts; automated processes are increasingly utilized with the explosion of published literature.

This study uses a **coupled review process** to capitalize on the advantages of each method while overcoming their disadvantages.

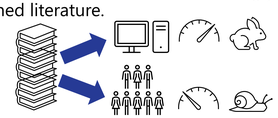


Fig. 2. Literature review by automated (top) and manual (bottom) methods.

**Purpose:**

**Summarize evidence of direct anthropogenic threats to global inland fisheries using coupled manual and automated methods**

## 2 METHODS

### LITERATURE CLASSIFICATION

*Does this abstract contain evidence of a direct, documented threat to fishes/fisheries in the specified hydrological basin?*

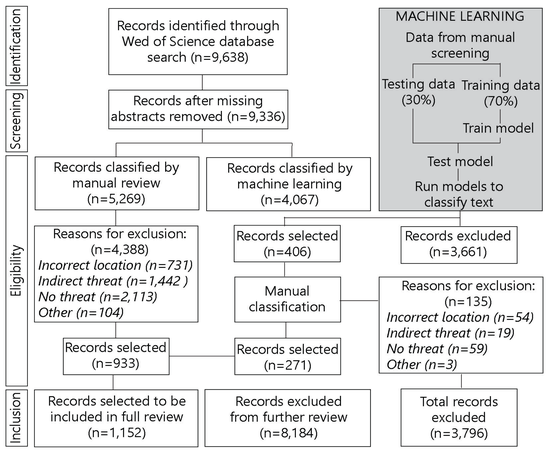



Fig. 3. Bibliographic literature search process: 1) article identification, 2) article screening, 3) eligibility for inclusion as a direct threat, and 4) selection for inclusion in data collection.

## 3 RESULTS

9,336 literature sources screened.

1,152 contained direct threats to fish.

### LITERATURE CLASSIFICATION

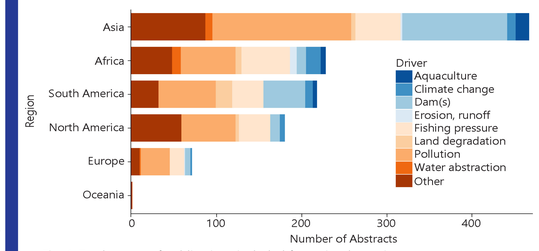



Fig. 4. Total counts of publications included for review by regions.

We screened 9,336 abstracts from **45 major river basins** published in 1,008 distinct peer-reviewed journals from 1990 to 2020 with **340 hours** manual effort and **41 hours** of automated effort.

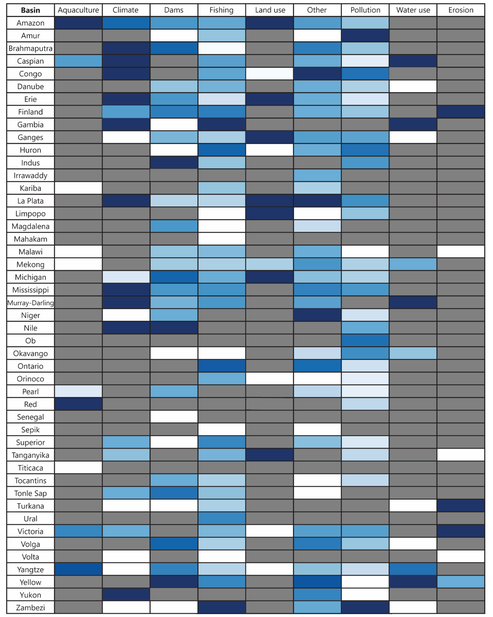
### MACHINE LEARNING PERFORMANCE

Table 1. Machine learning outputs from models where training data was applied to unclassified articles.

| | Naïve Bayes | | | Logistic Regression | | | SVM | | | kNN | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Category | Prec. | Recall | F1 | Prec. | Recall | F1 | Prec. | Recall | F1 | Prec. | Recall | F1 |
| Direct | 0.56 | 0.45 | 0.50 | 0.64 | 0.40 | 0.50 | 1.00 | 0.03 | 0.06 | 0.49 | 0.36 | 0.41 |
| Indirect | 0.90 | 0.93 | 0.92 | 0.89 | 0.96 | 0.92 | 0.84 | 1.00 | 0.92 | 0.88 | 0.93 | 0.91 |
| Average Accuracy | | | 0.86 | | | 0.87 | | | 0.84 | | | 0.84 |

Natural language processing was **90% accurate** classifying irrelevant papers. A recall rate of 64-67% suggests if only machine learning was used, 2/3 of papers with direct drivers would be selected. Manual checks aligned with model metrics with **false positives** in **33.3%** of abstracts and **false negatives** in **8.4%.**

### THREATS TO FISHERIES




Climate change, pollution, land use, and fishing pressure were associated with **higher overall threat** scores across basins.

**Important gaps remain** for documented direct evidence of some types of threats, seen here in gray boxes.

Fig. 5. Mean threat scores (1-4) by river basins important to inland fisheries and drivers types, where darker colors represent higher threat and lighter colors represent lower threat.

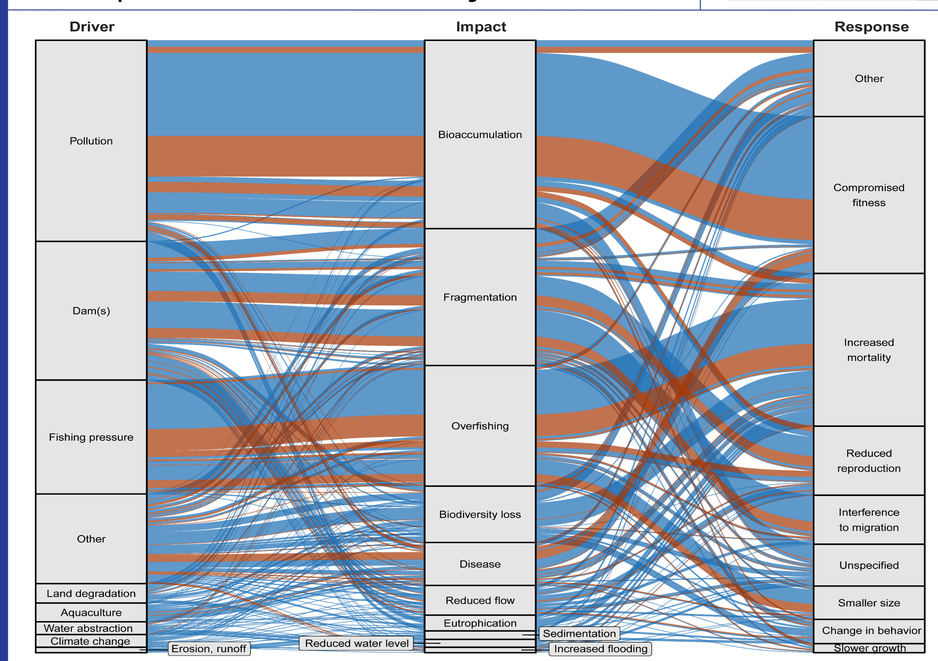




Fig. 6. Alluvial diagram depicting driver – impact – response relationships of documented, direct threats to fisheries.

Single species  Multiple species

### THREAT CHARACTERISTICS

*Drivers*

The most common documented, direct drivers of threats to fisheries were pollution (33%), dams (17%), and fishing pressure (17%). Climate change was documented in a relatively small count of articles but a high proportion of basins (n=21).

*Impacts*

The distribution of impacts strongly corresponded with top drivers, where pollution was linked to 87% of all bioaccumulation impacts, dams were attributed to 56% of fragmentation impacts, and fishing pressure resulted in 93% of overfishing impacts. All driver categories were linked to biodiversity loss.

*Responses*

The most frequently documented drivers were also reflected in the strongest links to fish response categories: mortality (37% linked to fishing pressure), changes in fitness (77% linked to pollution), and reproduction impacts (47% linked to dams).

## 4 DISCUSSION

This study advances the understanding of driver-response-impact links for inland fisheries.

### KEY TAKEAWAYS

1. Eight major drivers recorded at least once per basin → All types of stressors present in major fisheries
2. Documented threats skewed toward pollution → Bias toward acute threats with more visible impacts
3. Literature is incomplete as a standalone weighting method → Threat-impacts not well-studied or documented in published literature
4. Models performed best classifying irrelevant papers → Coupled approaches can improve efficiency
5. Both humans and machine learning are needed for reviews → Manual review provides necessary training data for models to run

### APPLICATIONS

Results will inform the relative importance scores for a global inland fisheries threat assessment, which is necessary **for improving fisheries conservation** and sustainable management.

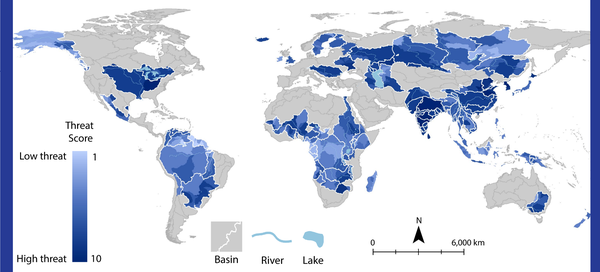



Fig. 7. Pilot threat assessment for major inland fisheries. *Source: FAO 2021*

## 5 CONCLUSIONS

**Literature synthesis contributes valuable threat information for fisheries assessments.**

**+**

**Both humans and machine learning are necessary for effective and efficient literature review.**

## REFERENCES & ACKNOWLEDGEMENTS


- FAO Division of Fisheries
- USGS NCASC
- UF Land & Water Lab
- Siddhesh Ambokar, GMU
- Jeremy Diaz, USGS
- Daniel Wieferich, USGS
- Lynch et al. 2016
- Kim et al. 2019
- Hazlett et al. 2020


This material is based upon work supported by the National Science Foundation Graduate Research Fellowship under Grant No. DGE-1842473.

UF IFAS UNIVERSITY of FLORIDA

Contact Information:
gstokes@ufl.edu

Land and Water Lab
University of Florida
landandwaterlab.org