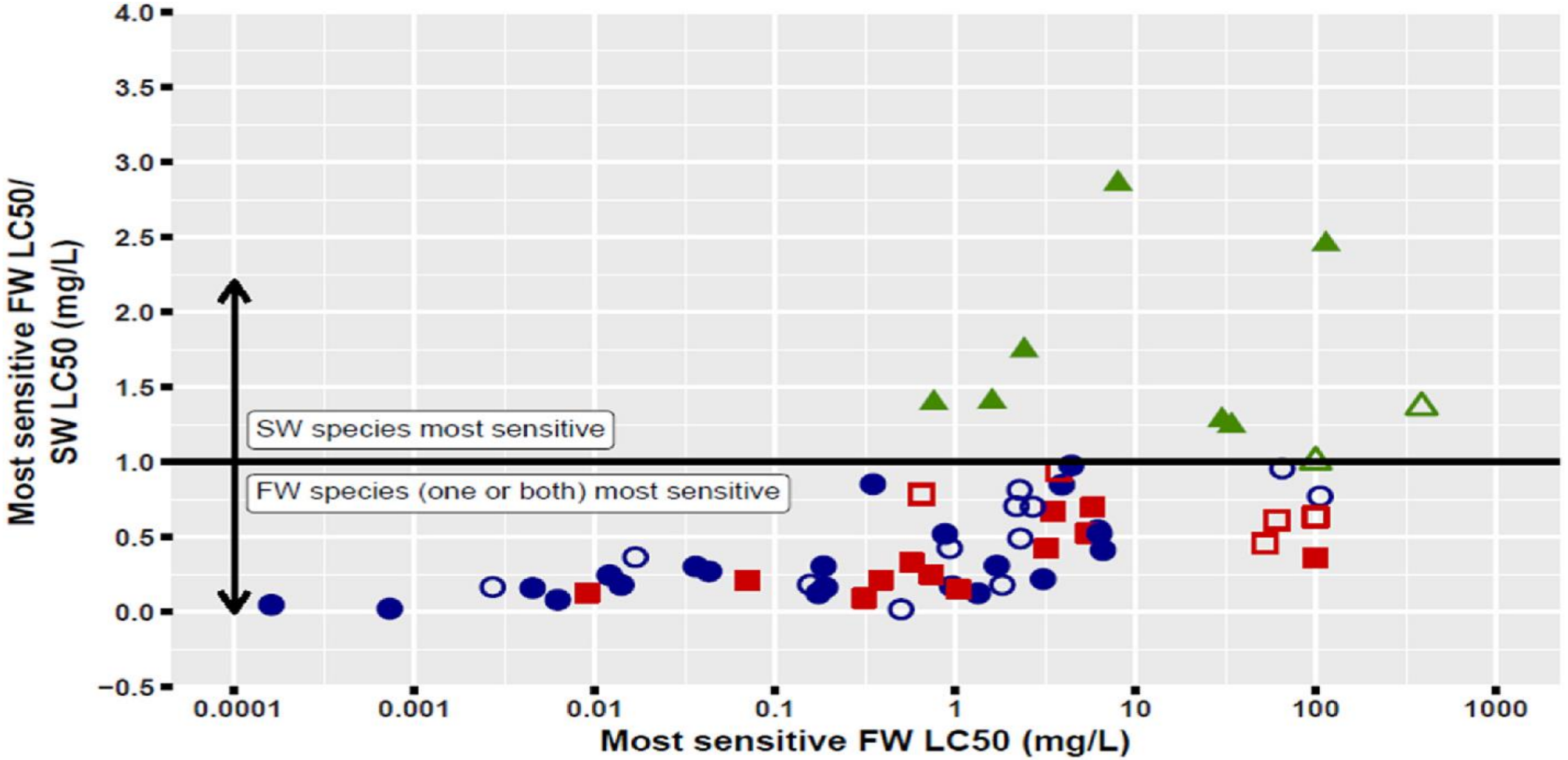
Activities in EPA/OPP to Reduce, Replace or Refine Acute Fish Testing

William P. Eckel, Ph.D.

January 31, 2024

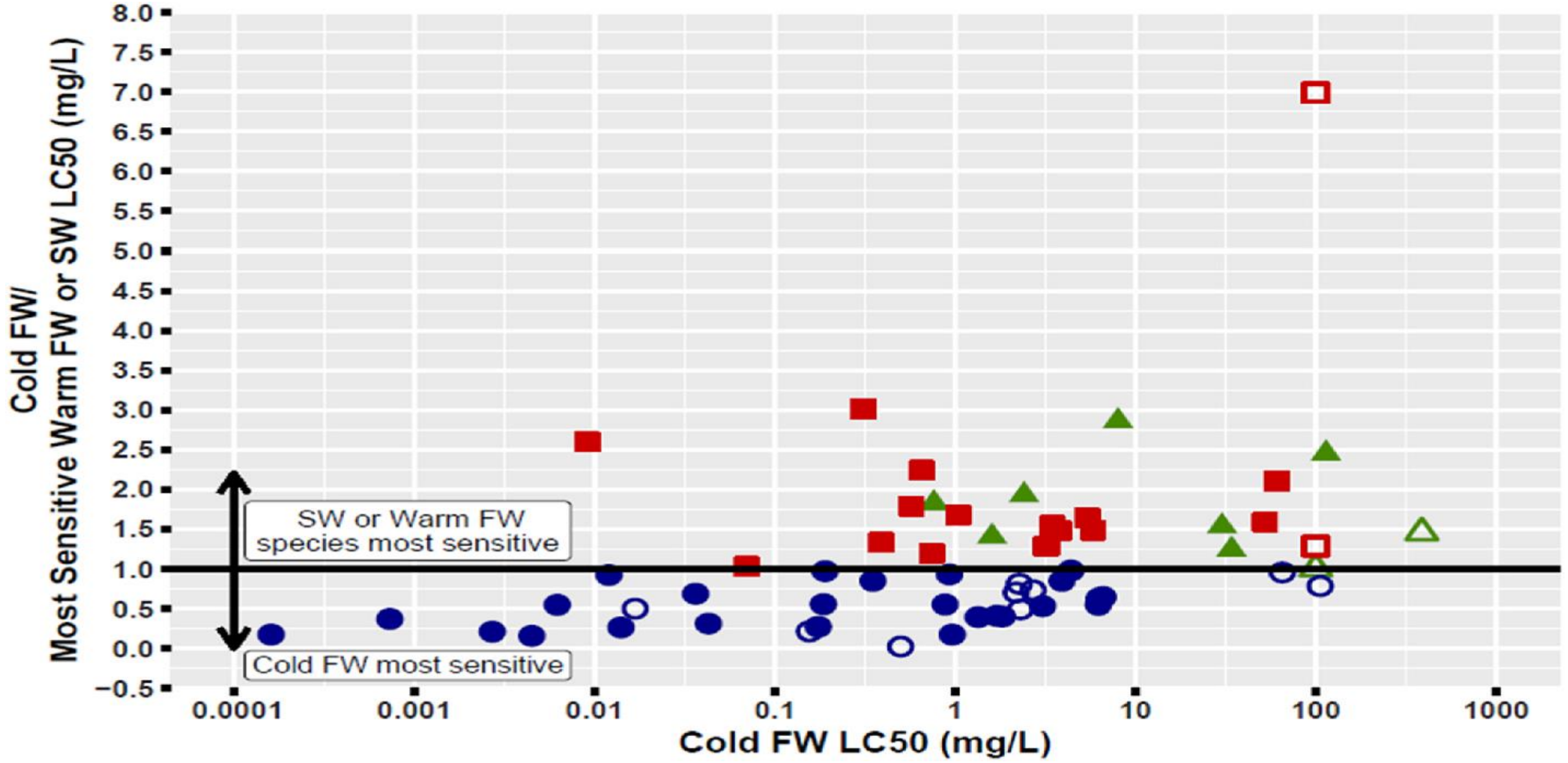
Relative Sensitivity of FW versus SW



Retrospective Analysis of 3 Test Species

- 181 cases of pesticides tested on cold freshwater, warm freshwater, and estuarine species
- With few exceptions, the estuarine species was the least sensitive and could be dropped without affecting the protectiveness of the risk assessment
- Regulatory Toxicology and Pharmacology 2023, 139 (March):105340

Relative Sensitivity of Cold FW vs Warm FW and SW



QSAR Model for Fish LC50

- OPP currently uses the OPPT ECOSAR model to estimate toxicity of metabolites when data are lacking
- OPP worked with ORD-Duluth to develop a new QSAR model based on acute fish data submitted to EPA ("Random Forest" -RF- model)
- RF model had best predictive ability overall (vs ECOTOX, FishTox, TEST) because it
  - is trained on pesticide data
  - algorithm accounts for predictor interactions and non-linear relationships
  - Works best on structural classes in training set

QSAR Evaluation and Development for the Prediction of Acute Responses in Fish to Exposure to Pesticides and their Degradates

Kelsey Vitense[1], Richard Kolanczyk[1], William Eckel[2], Greg Elonen[3], Tim Dawson[3], Dale Hoff[1]

[1]U.S. Environmental Protection Agency, Office of Research and Development, [2]U.S. Environmental Protection Agency, Office of Chemical Safety and Pollution Prevention, [3]SpecPro Professional Services

Introduction

- When little or no observed toxicity data are available for ecotoxicological hazard assessments of pesticides or their degradates, Quantitative Structure Activity Relationship (QSAR) models may be used to assess the potential acute toxicity of chemicals in fish.
- However, the data used to develop QSARs often do not represent the broad range of pesticidal structures and modes-of-action that contribute to pesticide toxicity.

Research questions:

1. How well can existing QSARs predict acute toxicity (LC50s) in fish?
2. Can we develop a new model with better prediction accuracy for acute pesticide toxicity in fish?

Methods

- Compiled LC50s for 266 pesticides and 17 fish species (n=744 chemical-species LC50s) from data evaluation records (DERs),17 unregistered EFSA pesticides (n=26 LC50s), and 32 degradates (n=39 LC50s).
- Evaluated parent chemical predictive accuracy for 3 existing QSARs (ECOSAR, TEST, FishTox) and a new random forest (RF) model trained using strictly pesticide data (Table 1).
  - RF predictors: 1) solubility, 2) n-octanol/water distribution coef. (LogD, pH=7), 3) vapor pressure, 4) ClassyFire superclass, 5) ClassyFire class, 6) ClassyFire subclass, 7) pesticide target (Fungicide/Herbicide/Insecticide Resistance Action Committee [RAC]), 8) test fish species, and 9) test exposure dynamics.
- Evaluated performance of full RF model (all predictors) using cross validation (CV) to assess on full parent dataset (Fig. 1).
- After initial assessment, identified reduced RF model (smallest predictor subset) using recursive feature elimination (RFE; Fig. 2), retaining fish species *a priori*.
- Assessed reduced RF model on 1) holdout set of 40 chemicals in well-represented ClassyFire subclasses (roughly in model domain of applicability [DOA]), 2) EFSA parent chemicals, and 3) degradate dataset.

Results

Table 1. Summary of model characteristics and performance on full parent chemical dataset. 'Good' prediction = ≤5-fold difference between predicted and observed. (Exponentiated mean absolute error)

Model	Target fish species	Quantitative methodology	Fathead minnow prediction good on average	Chemical-species prediction good on average	Chemical mean prediction good on average	Low bias
ECOSAR	Generic	Linear regression by chemical class	Y (4.6x)	N (6.0x)	N (5.2x)	Y
TEST	Fathead minnow	Model averaging (nearest neighbor, multilinear regression, hierarchical clustering)	Y (4.3x)	N (6.8x)	N (6.2x)	Y
FishTox	Multiple species	Ensemble modeling (random forest, gradient boosted trees, and support vector regression)	Y (4.0x)	Y (4.9x)	Y (4.5x)	N
RF	Multiple species	Random forest	Y (4.2x)	Y (4.3x)	Y (3.9x)	Y

- RF model outperformed other models across predictive metrics (mean absolute error [MAE], mean squared error, bias) on the full parent chemical dataset, whether summarized at the chemical-species level or chemical-mean level (Table 1).
- FishTox had second-best performance but exhibited substantial bias (predicted>observed on average; Fig 1).
- ECOSAR outperformed TEST when assessed on the full multi-species fish dataset.
- TEST had its best performance when the data were restricted to fathead minnow.

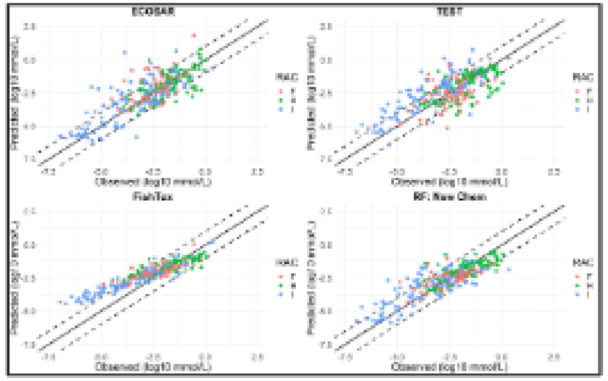


Fig 1. Predicted vs. observed LC50s (chemical means). RAC=Resource Action Committee, F=Fungicide, H=Herbicide, I=Insecticide. RF predictions are CV predictions (chemical data do not overlap between folds).

Results, continued

- Reduced RF model (Fig. 2) had good performance on holdout parent chemicals in well-represented ClassyFire subclasses (Table 2), good/fair performance on degradate set (~1/4 subclasses were novel), and fair performance on EFSA parent chemicals (~1/2 subclasses were novel).

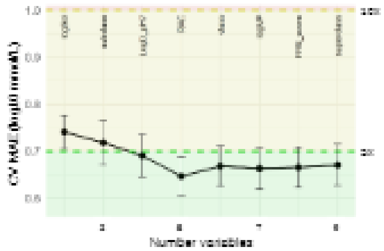


Fig 2. CV error as a function of remaining variables in RF model during RFE. Five predictors in reduced RF are: solubility, ClassyFire subclass, LogD, pesticide target (RAC), and fish species. Least informative predictor listed at the top of each iteration.

Table 2. Performance of reduced RF on test data. Errors computed on $\log_{10}$-scale and exponentiated for interpretation as fold-difference. Bias>1 means predicted>observed on average.

Dataset	N	MAE	RMSE	Bias
Holdout set (~in DOA)	40	3.0	3.8	1.3
EFSA test set	17	6.4	11.3	2.3
Degradates	32	5.0	7.4	1.1

Conclusion

- Relative model performance matches expectations given model complexity and target species (Table 1).
- RF had best predictive ability overall, likely because:
  - It is trained using only pesticide data and a targeted predictor set
  - Algorithm accounts for predictor interactions and non-linear relationships
- RF performs best on chemicals in ClassyFire subclasses represented in training set (i.e., roughly in DOA).

Inter-Agency Effort to Replace or Reduce Fish LC50 Testing

- ICCVAM – Interagency Coordinating Committee on the Validation of Alternative Methods
- Sponsored by National Toxicology Program
- Developing a review paper of the state of the science on alternatives to acute testing of live fish
- Fish embryo tests (FET), rainbow trout gill assay, and other methods