

QSAR evaluation & development for predicting acute responses in fish to exposure to pesticides and their degradates:

Introduction to **TMPPL**
(Toxicity Model for Predicting Pesticide Lethality)

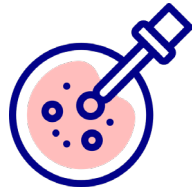
Eco-NAMs Webinar Series

2026/09/17

The views expressed in this presentation are those of the authors and do not necessarily represent the views or policies of the U.S. Environmental Protection Agency

QSARs are a type of New Approach Methodology (NAM)

- **QSAR** = Quantitative Structure-Activity Relationship
- **NAM** = alternative method to reduce or replace traditional toxicity testing on animals



- U.S. Environmental Protection Agency (USEPA) is committed to **reduce, replace or refine (3Rs)** testing of vertebrate animals while maintaining protectiveness of risk assessments
- **QSARs** qualify as a potential replacement for animal testing in certain contexts

How does USEPA use QSARs in pesticide risk assessments?



- Pesticides & their degradates are assessed for potential ecological risk by USEPA during pesticide registration process
 - Registrants submit *in vivo* toxicity studies on pesticide parent chemical but seldomly on all degradates of potential concern; therefore, predictive models are often used to estimate degrade toxicity
- USEPA currently uses ECOSAR to assess toxicity of pesticide degradates to aquatic organisms
 - ECOSAR estimates are not used as regulatory endpoint thresholds but inform degrade inclusion in Residues of Concern (RoC) for risk assessment and possible data needs
- ECOSAR was developed by USEPA for industrial chemicals
 - 38/120 structural classes have pesticides in training data to inform predictions
- RoC assignment often based on non-pesticide chemical classes
 - Uncertain whether results reflect pesticide/degrade mechanism of toxicity
- USEPA is evaluating QSARs that may have higher prediction accuracy for pesticides and degradates
 - May improve confidence in assignment of degradates to RoC

How do USEPA's QSAR tools for acute fish toxicity perform on pesticides and degradates?



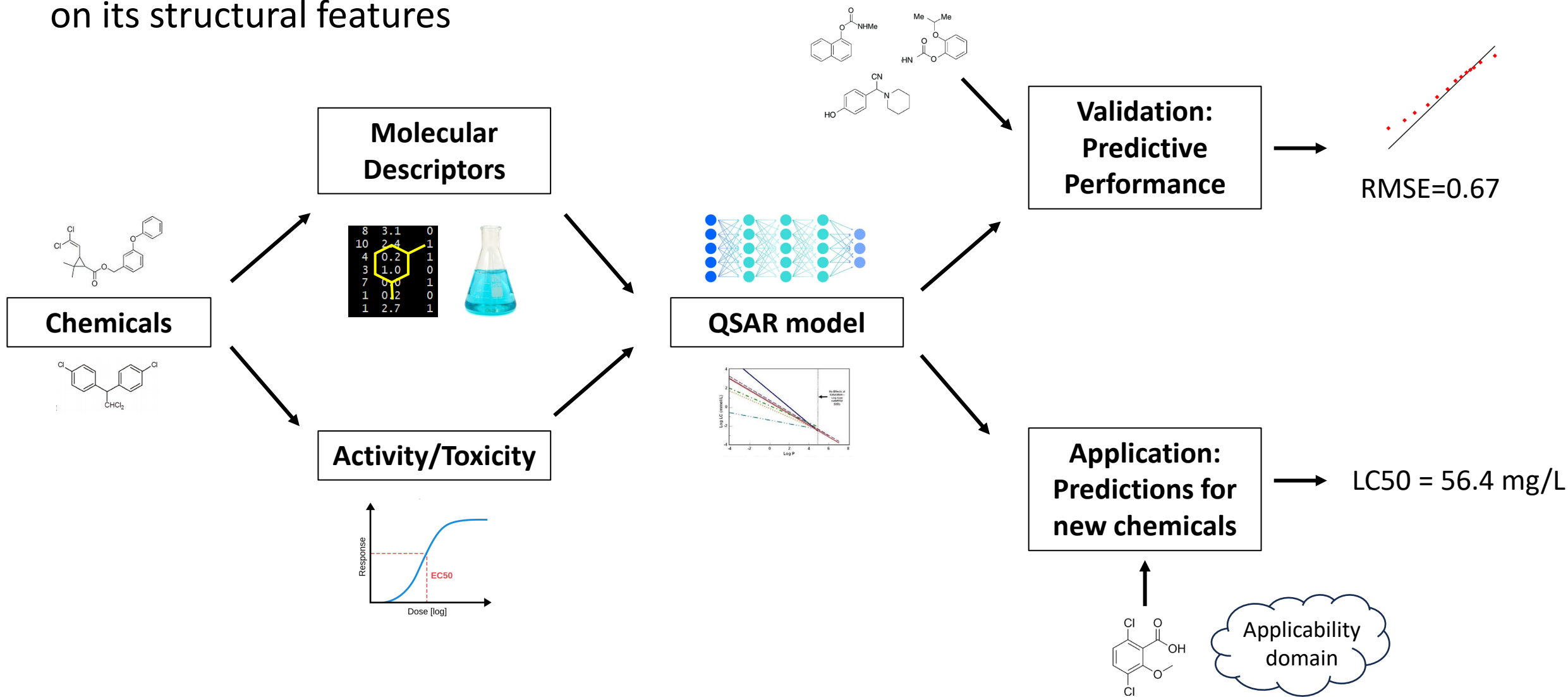
- USEPA evaluated three existing USEPA-developed QSAR models and a newly developed pesticide-specific QSAR model for predictive performance on pesticides and degradates
 - Ecological Structure-Activity Relationship Model (ECOSAR)
 - Toxicity Estimation Software Tool (TEST)
 - FishTox
 - Toxicity Model for Predicting Pesticide Lethality (TMPPL) – *new*
- Evaluated performance using 266 pesticides and 32 degradates with *in vivo* fish acute toxicity data

Outline

- Overview of pesticide dataset and QSAR models
 - ECOSAR, TEST, FishTox, TMPPL
- QSAR model performance assessment for pesticides & degradates
- Residue of Concern (RoC) assessment
 - ECOSAR vs. TMPPL

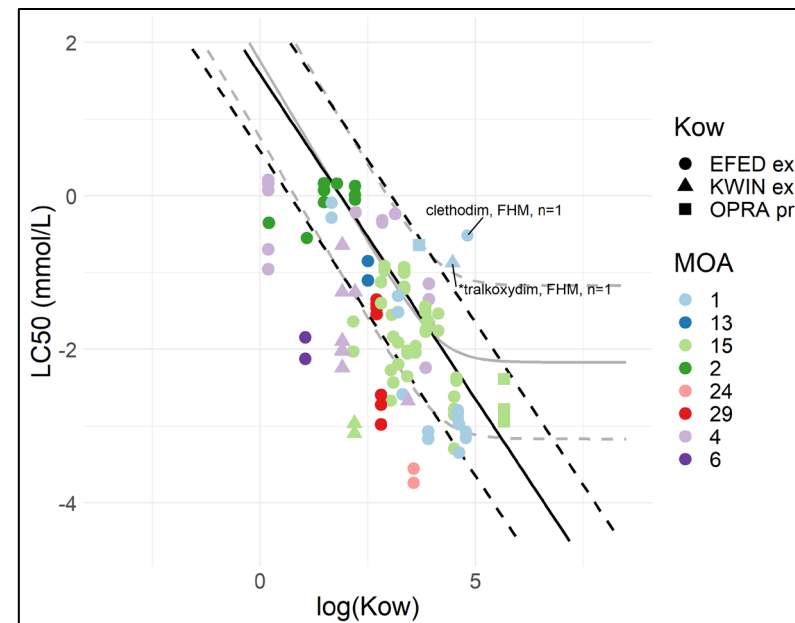
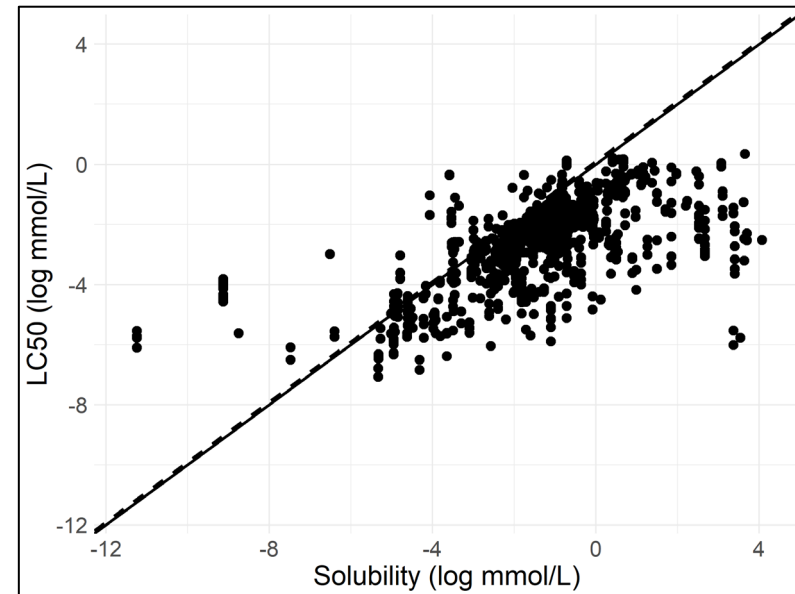
What is a Quantitative Structure-Activity Relationship (QSAR) model?

- Computational approach for predicting a chemical's biological activity or toxicity based on its structural features



Pesticide acute fish tox dataset for model evaluation/development

- *In vivo* pesticide LC50s for 17 fish species from USEPA science documents
- Removed:
 - Metals, mixtures
 - LC50 values $> 1.25 \times$ solubility
 - LC50 values $> 10 \times$ ECOSAR baseline prediction for neutral organics
- Used geometric means for multiple chemical-species LC50s
- Did not aggregate data for isomers
- 266 chemicals, 744 chemical-species observations



ECOSAR model

- USEPA has existing guidance on use of ECOSAR for estimating toxicity of degradates to inform RoC for risk assessment
- Well-established model for aquatic taxa (fish, invertebrates, plants) developed and used by USEPA for many chemicals (primarily industrial)
- Acute (LC50) toxicity predictions for **generic fish**
 - Training using data from bluegill, common carp, fathead minnow, rainbow trout, etc.
- Uses **LC50 vs. LogKow** linear regressions by **chemical structural class**
- Used minimum (i.e., most potent) LC50 prediction across all predicted classes for model assessment

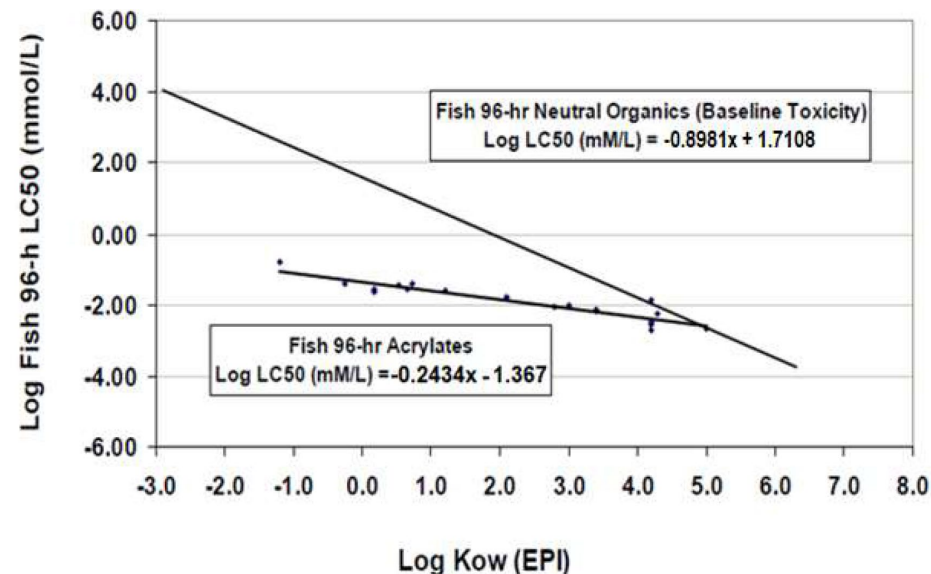
ECOLOGICAL Structure-Activity Relationship Model

(ECOSAR)

Class Program

ESTIMATING TOXICITY OF INDUSTRIAL CHEMICALS
TO AQUATIC ORGANISMS USING THE
ECOSAR (ECOLOGICAL STRUCTURE-ACTIVITY RELATIONSHIP) CLASS
PROGRAM

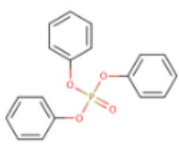
Version 2.2



TEST model

- Todd Martin, USEPA
- Acute aquatic and mammalian toxicity, physical properties, etc.
- Acute (LC50) toxicity predictions for **fathead minnow (FHM)**
- Caveat: Relaxed constraint requiring all molecular fragments of pesticide to be in training set for ~20% of evaluated pesticides

Individual Predictions	
Method	Predicted value -Log10(mol/L)
Hierarchical clustering	6.42
Single model	6.05
Group contribution	N/A
Nearest neighbor	5.30

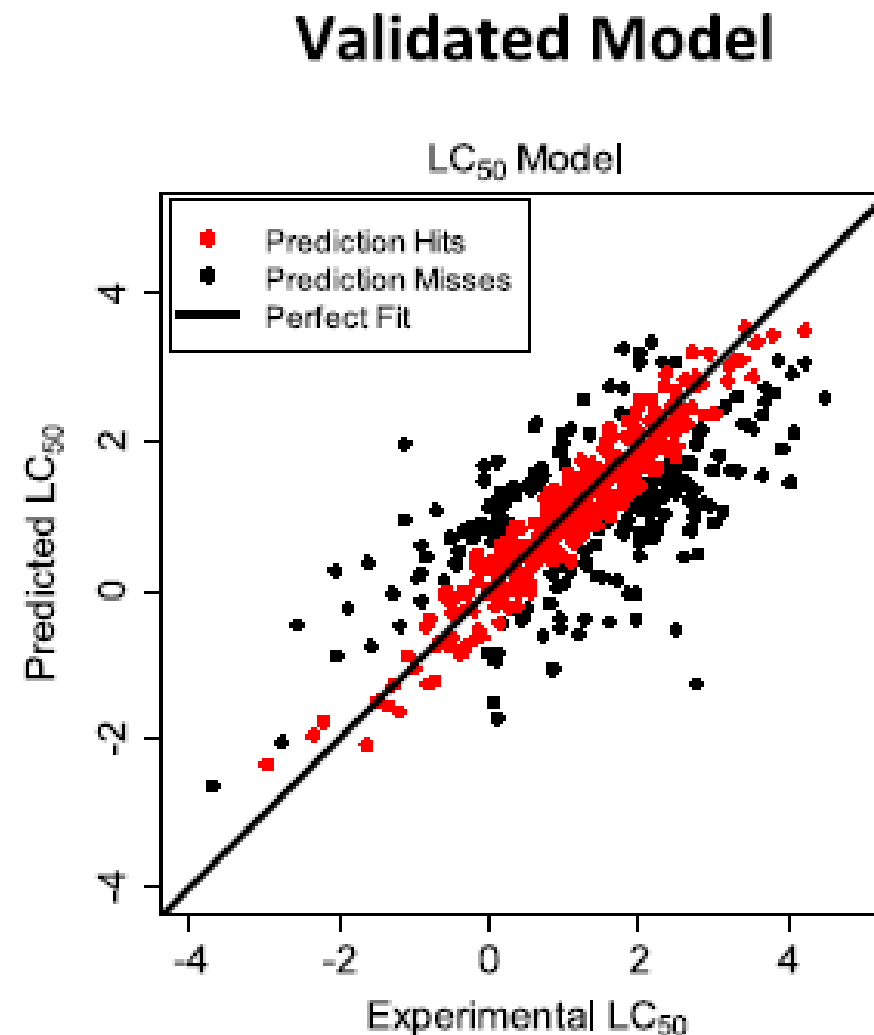


Uses a variety of QSAR methodologies:

- Hierarchical method: Hierarchical clustering into structurally similar clusters with multilinear regression using molecular descriptors within clusters
- Single model method: Multilinear regression using molecular descriptors as independent variables.
- Group contribution method: Multilinear regression model using molecular fragment counts as independent variables.
- Nearest neighbor method: Average of the 3 chemicals in the training set that are most similar to the test chemical.
- **Consensus method: Average of the predicted toxicities from the above QSAR methods** (provided the predictions are within the respective applicability domains [ADs]).

FishTox model

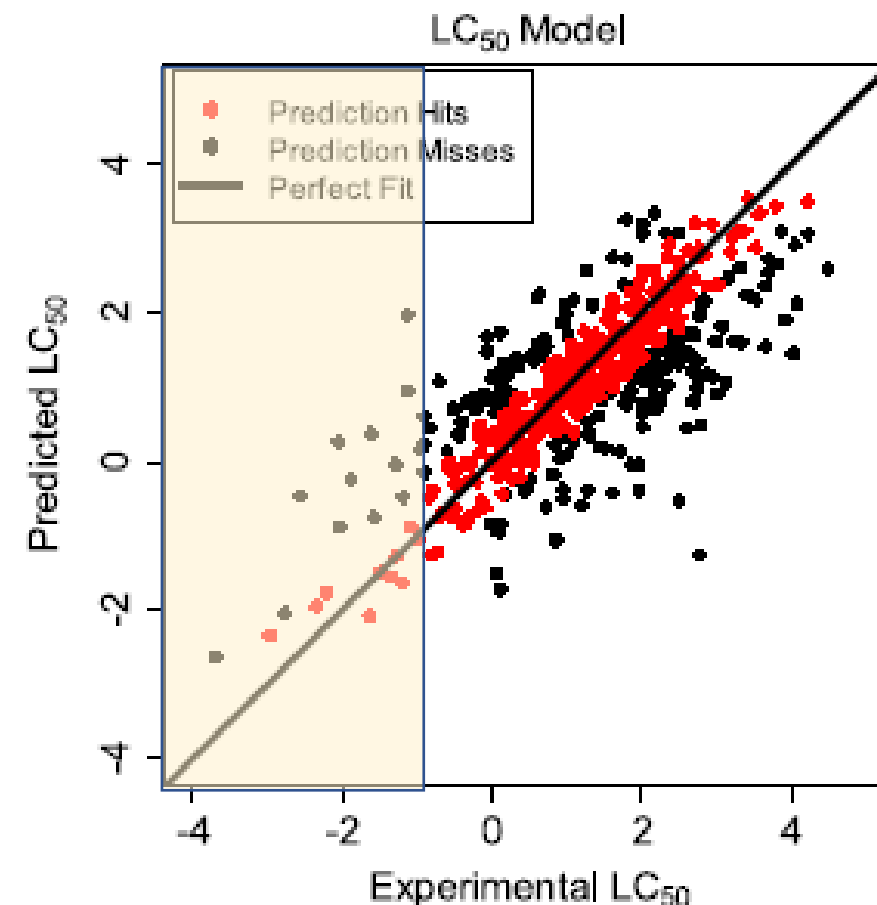
- Sheffield & Judson (2019) publication
- Acute (LC50) predictions for **multiple fish species**
- **Ensemble of 3 machine learning models** (*random forest, gradient boosted trees, support vector regression*) with chemical descriptors from:
 - PaDEL: Pharmaceutical Data Exploration Laboratory
 - Descriptors such as electrotopological states and autocorrelations; atom, ring and bond counts.
 - OPERA: OPEn structure–activity/property Relationship App
 - Predictions of physicochemical properties such as LogP, vapor pressure and water solubility
 - Uses PaDEL descriptors
- Includes fish species and exposure method (e.g., static, renewal) as predictors



FishTox model

- Sheffield & Judson (2019) publication
- Acute (LC50) predictions for **multiple fish species**
- **Ensemble of 3 machine learning models** (*random forest, gradient boosted trees, support vector regression*) with chemical descriptors from:
 - PaDEL: Pharmaceutical Data Exploration Laboratory
 - Descriptors such as electrotopological states and autocorrelations; atom, ring and bond counts.
 - OPERA: OPEn structure–activity/property Relationship App
 - Predictions of physicochemical properties such as LogP, vapor pressure and water solubility
 - Uses PaDEL descriptors
- Includes fish species and exposure method (e.g., static, renewal) as predictors

Validated Model



TMPPL model development for pesticides – candidate predictors

- 266 chemicals, 744 observations
- ClassyFire chemical classes
 - **Superclass, class, subclass**
- **Fish species** (all 17 species)
- Exposure method
 - **Static, renewal, flow-through**
- **RAC** (Pesticide target taxon)
 - Insecticide, Fungicide, Herbicide
- Physicochemical properties
 - **OPERA LogD at pH=7.4 (canned)**
 - **OPERA LogD at pH=7 computed with equation**
 - **OPERA LogD at pH=7 computed from avg. pKa**
 - **Solubility** (Priority order: registrant exp., EPI Suite exp., OPERA pred., EPI Suite pred.)
 - **Vapor Pressure:** (Priority order: registrant exp., EPI Suite exp., OPERA pred., EPI Suite pred.)

ClassyFire examples

superclass	class	subclass
benzenoids	benzene and substituted derivatives	phenylphosphonothioates
benzenoids	benzene and substituted derivatives	phenylpropanes
benzenoids	benzene and substituted derivatives	trifluoromethylbenzenes
benzenoids	benzene and substituted derivatives	xylenes
benzenoids	indanes	indanes
benzenoids	naphthalenes	naphthalenes
benzenoids	phenol ethers	phenol ethers
benzenoids	phenols	nitrophenols
benzenoids	tetralins	tetralins
lipids and lipid-like molecules	fatty acyls	fatty acid esters
lipids and lipid-like molecules	prenol lipids	monoterpenoids
lipids and lipid-like molecules	steroids and steroid derivatives	steroid lactones
organic acids and derivatives	carboximidic acids and derivatives	carboximidic acids and derivatives
organic acids and derivatives	carboxylic acids and derivatives	amino acids, peptides, and analogues

The logD at pH = 7 for a molecule with multiple acidic and basic pKa values is calculated using the formula:

$$\log D = \log P - \log \left(1 + \sum_{i=1}^{n_{\text{acid}}} 10^{\text{pH} - \text{pK}_{\text{a},i}} + \sum_{j=1}^{n_{\text{base}}} 10^{\text{pK}_{\text{a},j} - \text{pH}} \right)$$
, where logP is the partition coefficient of the neutral form, $\text{pK}_{\text{a},i}$ are the acidic pKa values, and $\text{pK}_{\text{a},j}$ are the basic pKa values.

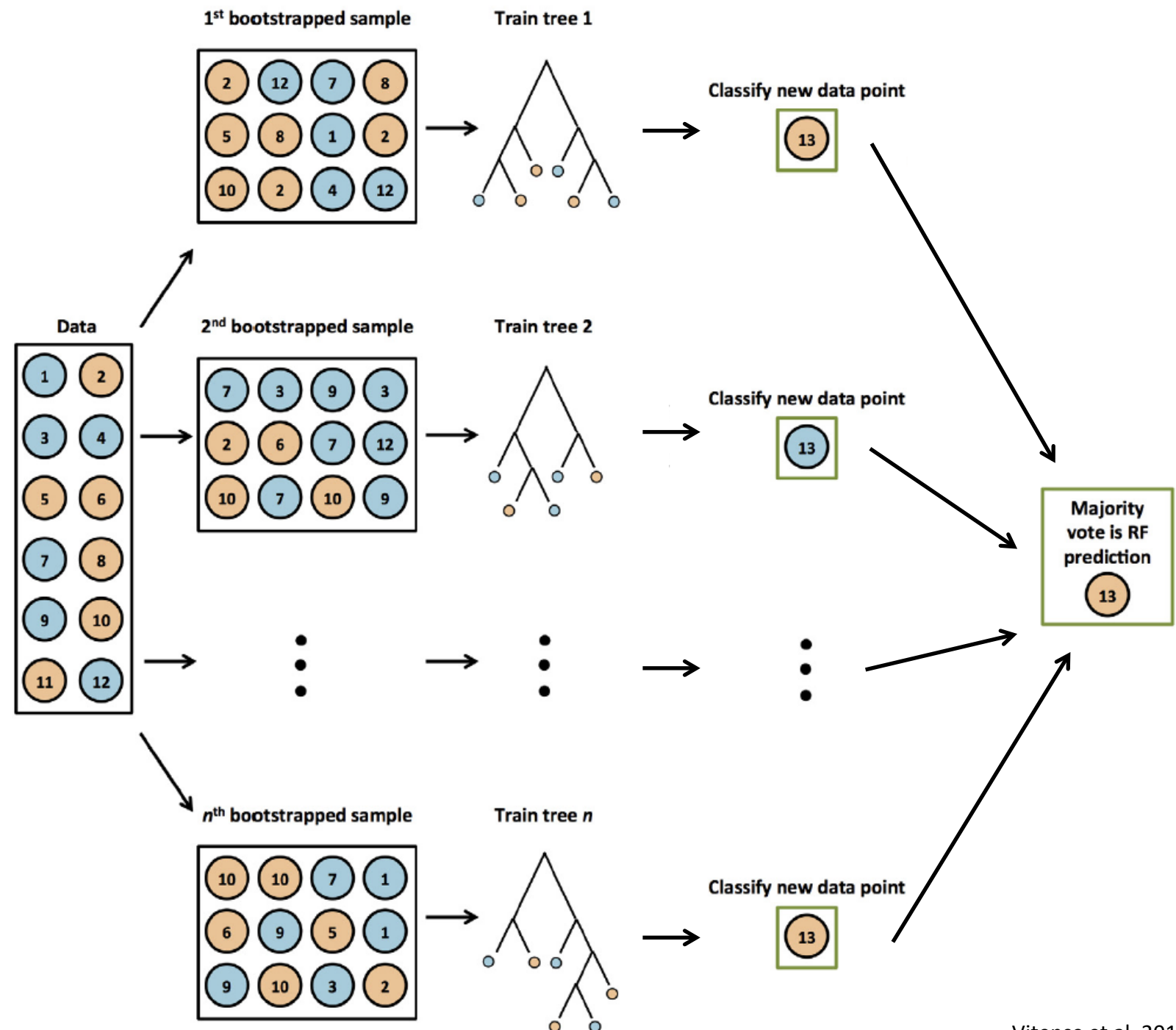
TMPPL is built with a type of random forest algorithm

Nice features

- No distributional assumptions
- Do not have to specify form of relationships, interactions of interest
- Discovers complex relationships between predictors and response

Each tree has low bias but high variance (overfit)

Aggregating predictions from many trees built with different samples maintains low bias but decreases variance of predictions (high accuracy)



Summary of model features

Model	Target fish species	Quantitative methodology	Training data
ECOSAR	Generic	Linear regression by chemical class	USEPA ecotoxicity studies for industrial chemicals; Industry-submitted test data; USEPA historical pesticide ecotoxicity database; Peer-reviewed literature
TEST	Fathead minnow	Model averaging (nearest neighbor, multilinear regression, hierarchical clustering)	USEPA's ECOTOX Knowledgebase (fathead minnow)
FishTox	Multiple species	Ensemble modeling (random forest, gradient boosted trees, support vector regression)	USEPA's ECOTOX Knowledgebase; ECHA (European Chemicals Agency) database
TMPPL	Multiple species	Random forest	Registrant-submitted pesticide toxicity studies

Overview of performance assessment across models

- Error metrics: **Mean absolute error** (MAE), **Root mean squared error** (RMSE), **Bias**, **R^2**
- USEPA risk assessment criteria: % **Good** (<5x), % **Fair** (5-10x), % **Poor** (>10x)

Overview of performance assessment across models

- Error metrics: Mean absolute error (MAE), Root mean squared error (RMSE), Bias, R^2
- USEPA risk assessment criteria: % Good (<5x), % Fair (5-10x), % Poor (>10x)
- **3 USEPA QSARs vs. Full TMPPL model (all predictors)**
 - **Cross-validated (CV) errors** = Predict new, unseen **chemicals**
 - **Out-of-bag (OOB) errors** = Predict new **chemical-species** combinations
 - **Compared performance using all chemicals**
 - 1) FHM, 2) All chem-species observations, and 3) Chemical averages

Overview of performance assessment across models

- Error metrics: Mean absolute error (MAE), Root mean squared error (RMSE), Bias, R^2
- USEPA risk assessment criteria: % Good (<5x), % Fair (5-10x), % Poor (>10x)
- 3 USEPA QSARs vs. Full TMPPL model (all predictors)
 - Cross-validated (CV) errors = Predict new, unseen chemicals
 - Out-of-bag (OOB) errors = Predict new chemical-species combinations
 - Compared performance using all chemicals
 - 1) FHM, 2) All chem-species observations, and 3) Chemical averages
- **Reduced TMPPL model (best predictor subset)**
 - Use recursive feature elimination to **select best predictor subset** with training set
 - Compared performance on **hold-out data** (40 chemicals not in training set)

Overview of performance assessment across models

- Error metrics: Mean absolute error (MAE), Root mean squared error (RMSE), Bias, R^2
- USEPA risk assessment criteria: % Good (<5x), % Fair (5-10x), % Poor (>10x)
- 3 USEPA QSARs vs. Full TMPPL model (all predictors)
 - Cross-validated (CV) errors = Predict new, unseen chemicals
 - Out-of-bag (OOB) errors = Predict new chemical-species combinations
 - Compared performance using all chemicals
 - 1) FHM, 2) All chem-species observations, and 3) Chemical averages
- Reduced TMPPL model (best predictor subset)
 - Use recursive feature elimination to select best predictor subset with training set
 - Compared performance on hold-out data (40 chemicals not in training set)
- Reduced TMPPL model performance for **European Food Safety Authority (EFSA) chemicals unregistered in U.S.** (n=17) and **degradates** (n=32)

Overview of performance assessment across models

- Error metrics: Mean absolute error (MAE), Root mean squared error (RMSE), Bias, R^2
- USEPA risk assessment criteria: % Good (<5x), % Fair (5-10x), % Poor (>10x)
- **3 USEPA QSARs vs. Full TMPPL model (all predictors)**
 - Cross-validated (CV) errors = Predict new, unseen chemicals
 - Out-of-bag (OOB) errors = Predict new chemical-species combinations
 - Compared performance using all chemicals
 - 1) FHM, 2) All chem-species observations, and 3) Chemical averages
- Reduced TMPPL model (best predictor subset)
 - Use recursive feature elimination to select best predictor subset with training set
 - Compared performance on hold-out data (40 chemicals not in training set)
- Reduced TMPPL model performance for European Food Safety Authority (EFSA) chemicals unregistered in U.S. (n=17) and degradates (n=32)

Performance for
fathead minnow (FHM)

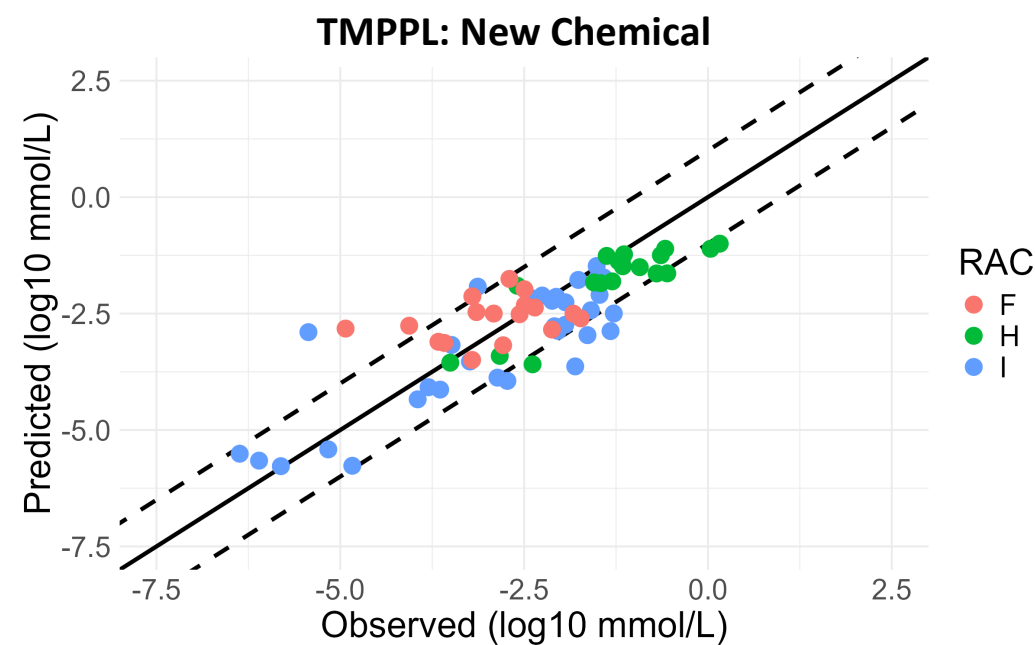
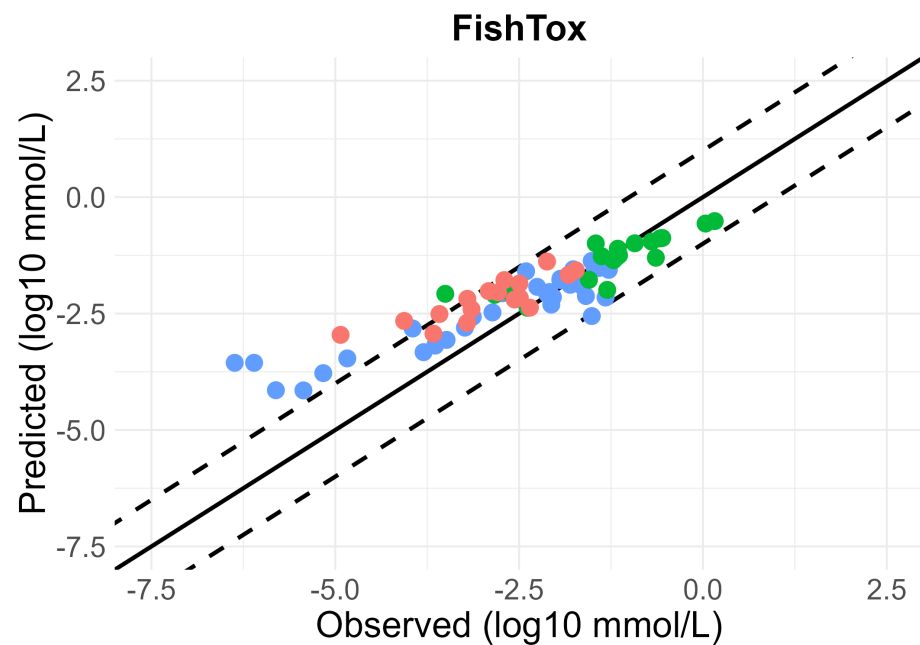
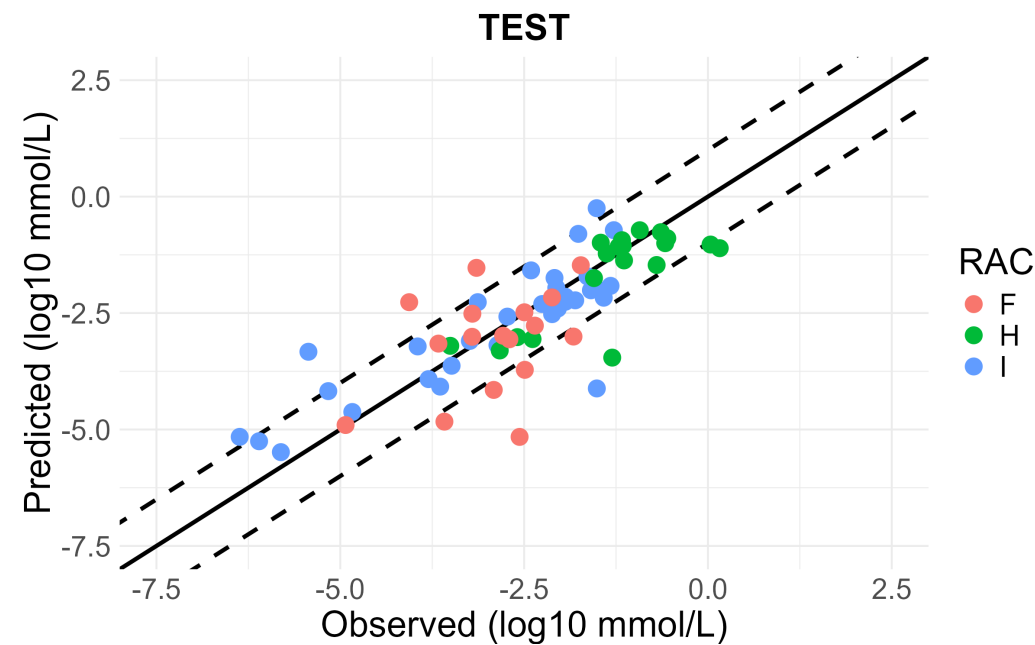
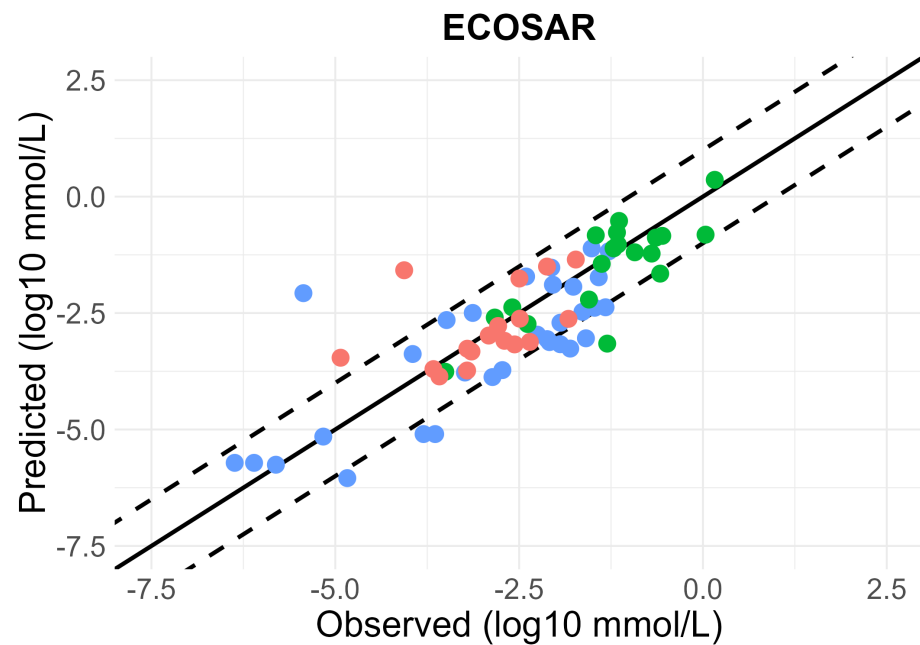
Performance – FHM

Bold*	= Best	Good: <5x Fair: 5-10x Poor: >10x
Bold	= Second Best	
Text	= Third Best	

Model	N	MAE	RMSE	Bias	% Good	% Poor	R ²
ECOSAR	68	4.61	7.66	0.67	62	21	0.61
TEST	68	4.26	7.39	0.81*	68*	21	0.62
FishTox	68	4.00*	6.75	2.37	68*	19*	0.65
TMPPL: new chemical (CV)	68	4.33	6.56*	0.63	65	22	0.66*
TMPPL: new chem-spec. (OOB)	68	3.05	3.89	0.52	75	10	0.82

- All models perform well for FHM (MAE<5; Good>60%; R²>0.60)
- FishTox has substantial bias for more toxic compounds
- TEST most competitive here for FHM, least biased
- TMPPL has good performance for FHM when predicting new chemicals (CV)
- TMPPL has excellent performance if allowed to be trained on different species of same chemical (OOB)

Predicted vs. Observed – FHM



Performance for All Observed
Chemical-Species Combinations

Performance – All Chem-Species Observations

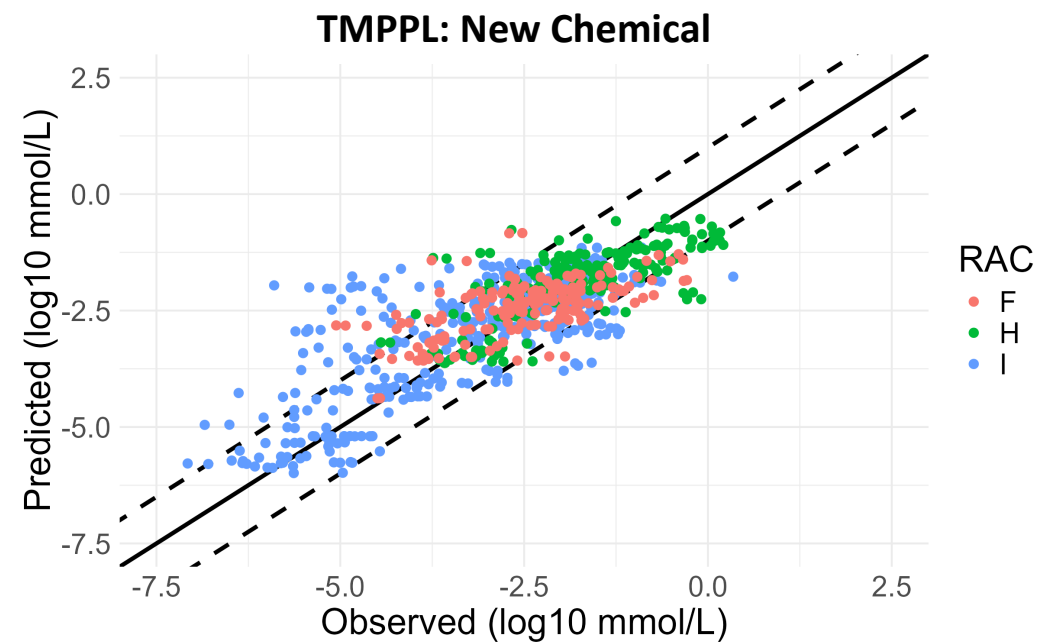
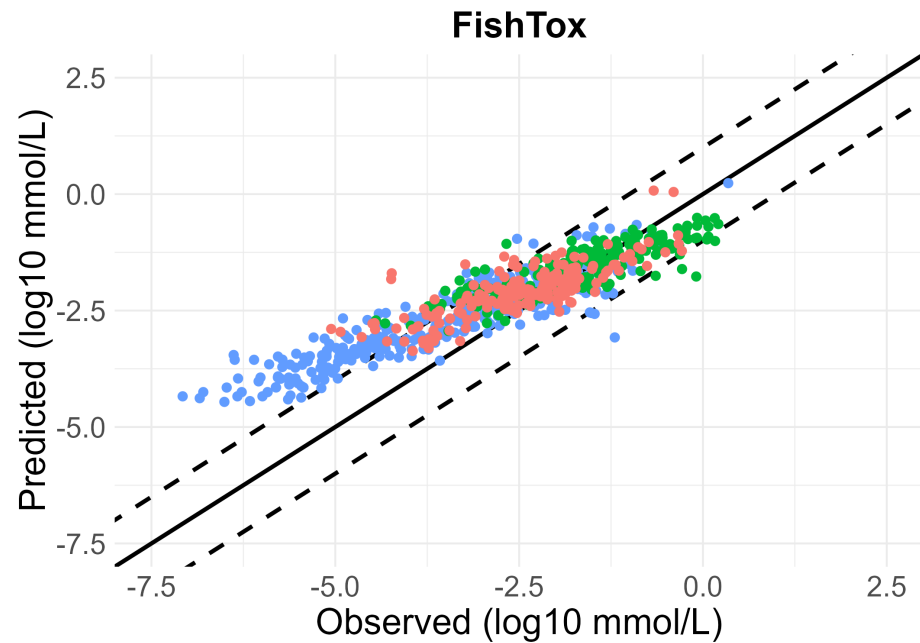
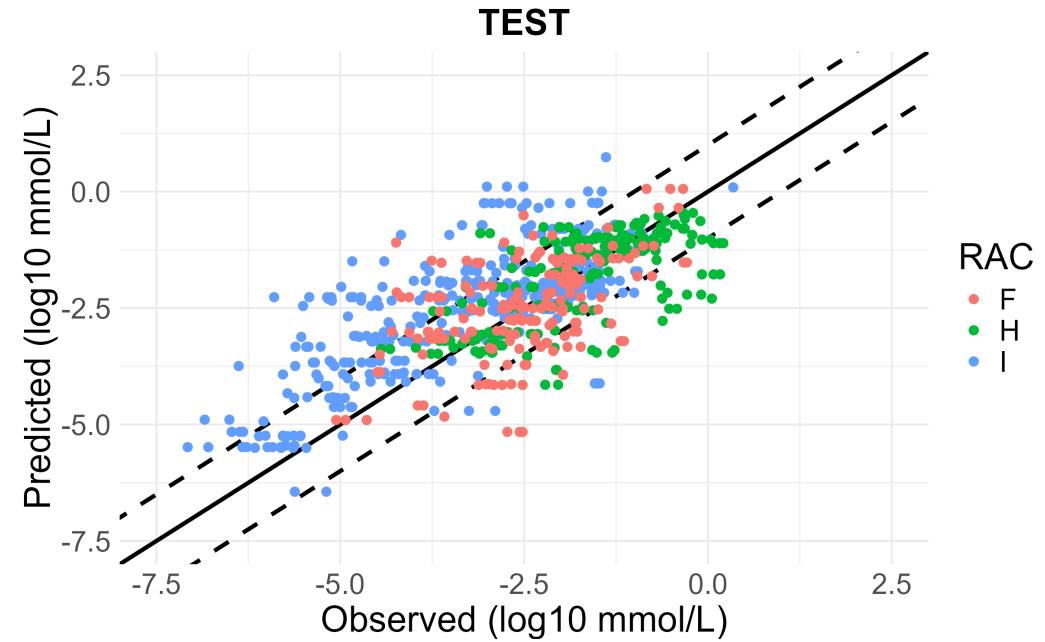
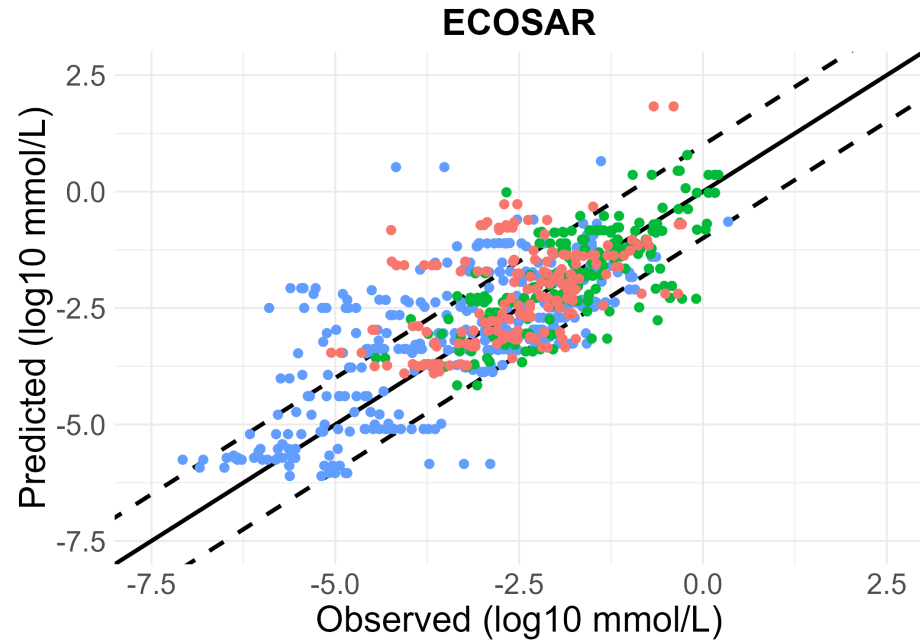
Bold* = Best
Bold = Second Best
Text = Third Best

Good: <5x
Fair: 5-10x
Poor: >10x

Model	N	MAE	RMSE	Bias	% Good	% Poor	R ²
ECOSAR	744	5.98	11.34	1.50	56	26	0.46
TEST	744	6.77	12.57	2.14	55	34	0.41
FishTox	<u>741</u>	4.90	7.84	3.14	61	27	0.61
TMPPL: new chemical (CV)	744	4.37*	7.53*	1.25*	65*	20*	0.62*
TMPPL: new chem-spec. (OOB)	744	3.02	4.33	1.04	78	12	0.80

- TMPPL has top performance across all metrics
- FishTox also performs well except for bias
- Results match expectations:
 - TMPPL and FishTox perform best for multi-species data
 - Produce species-specific predictions; both machine learning approaches
 - ECOSAR performs better than TEST
 - Multiple fish species used to build ECOSAR models

Predicted vs. Observed – All Chem-Species Observations



Performance for Chemical Averages

Performance – Chemical averages

Bold* = Best
Bold = Second Best
Text = Third Best

Good: <5x
Fair: 5-10x
Poor: >10x

Model	N	MAE	RMSE	Bias	% Good	% Poor	R ²
ECOSAR	266	5.21	10.23	1.38	60	24	0.51
TEST	266	6.22	10.72	1.56	57	31	0.49
FishTox	<u>264</u>	4.50	7.19	2.83	64	23	0.65
TMPPL: new chemical (CV)	266	3.96*	6.29*	1.11*	66*	16*	0.70*
TMPPL: new chem-spec. (OOB)	266	2.57	3.67	0.99	83	8	0.85

- TMPPL has top performance across all metrics
- FishTox also performs well except for bias
- Results match expectations:
 - TMPPL and FishTox perform best for multi-species data
 - Produce species-specific predictions; both machine learning approaches
 - ECOSAR performs better than TEST
 - Multiple fish species used to build ECOSAR models

Summary of model performance

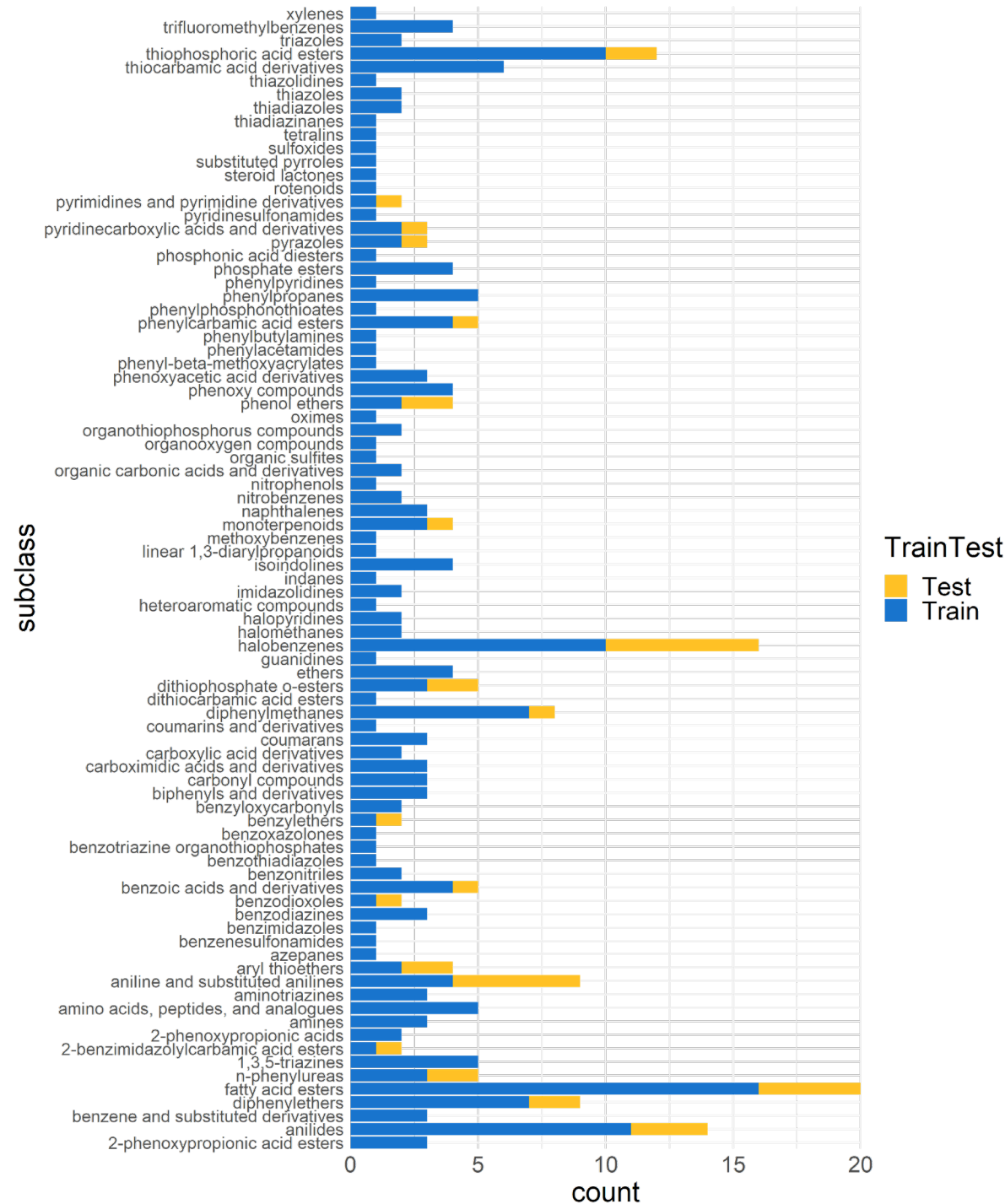
Model	Target fish species	Quantitative methodology	Fathead minnow prediction good on average (MAE $\leq$ 5)	Chemical-species prediction good on average (MAE $\leq$ 5)	Chemical mean prediction good on average (MAE $\leq$ 5)	Low bias
ECOSAR	Generic	Linear regression by chemical class	Y	N	N	Y
TEST	Fathead minnow	Model averaging (nearest neighbor, multilinear regression, hierarchical clustering)	Y	N	N	Y
FishTox	Multiple species	Ensemble modeling (random forest, gradient boosted trees, and support vector regression)	Y	Y	Y	N
TMPPL	Multiple species	Random forest	Y	Y	Y	Y

Overview of performance assessment across models

- Error metrics: Mean absolute error (MAE), Root mean squared error (RMSE), Bias, R^2
- USEPA risk assessment criteria: % Good (<5x), % Fair (5-10x), % Poor (>10x)
- 3 USEPA QSARs vs. Full TMPPL model (all predictors)
 - Cross-validated (CV) errors = Predict new, unseen chemicals
 - Out-of-bag (OOB) errors = Predict new chemical-species combinations
 - Compared performance using all chemicals
 - 1) FHM, 2) All chem-species observations, and 3) Chemical averages
- **Reduced TMPPL model (best predictor subset)**
 - Use recursive feature elimination to select best predictor subset with training set
 - Assessed performance on hold-out data (40 chemicals not in training set)
- Reduced TMPPL model performance for European Food Safety Authority (EFSA) chemicals unregistered in U.S. (n=17) and degradates (n=32)

Set aside holdout test set for validation of reduced TMPPL model

- Randomly sampled 40 chemicals (15%) with selection probability proportional to number of chemicals in ClassyFire subclass
- Subclasses in test set are all represented in training set
- This test set comprises chemicals roughly in applicability domain (AD)
- Perform recursive feature elimination on remaining 226 chemicals to identify optimal, minimal predictor subset



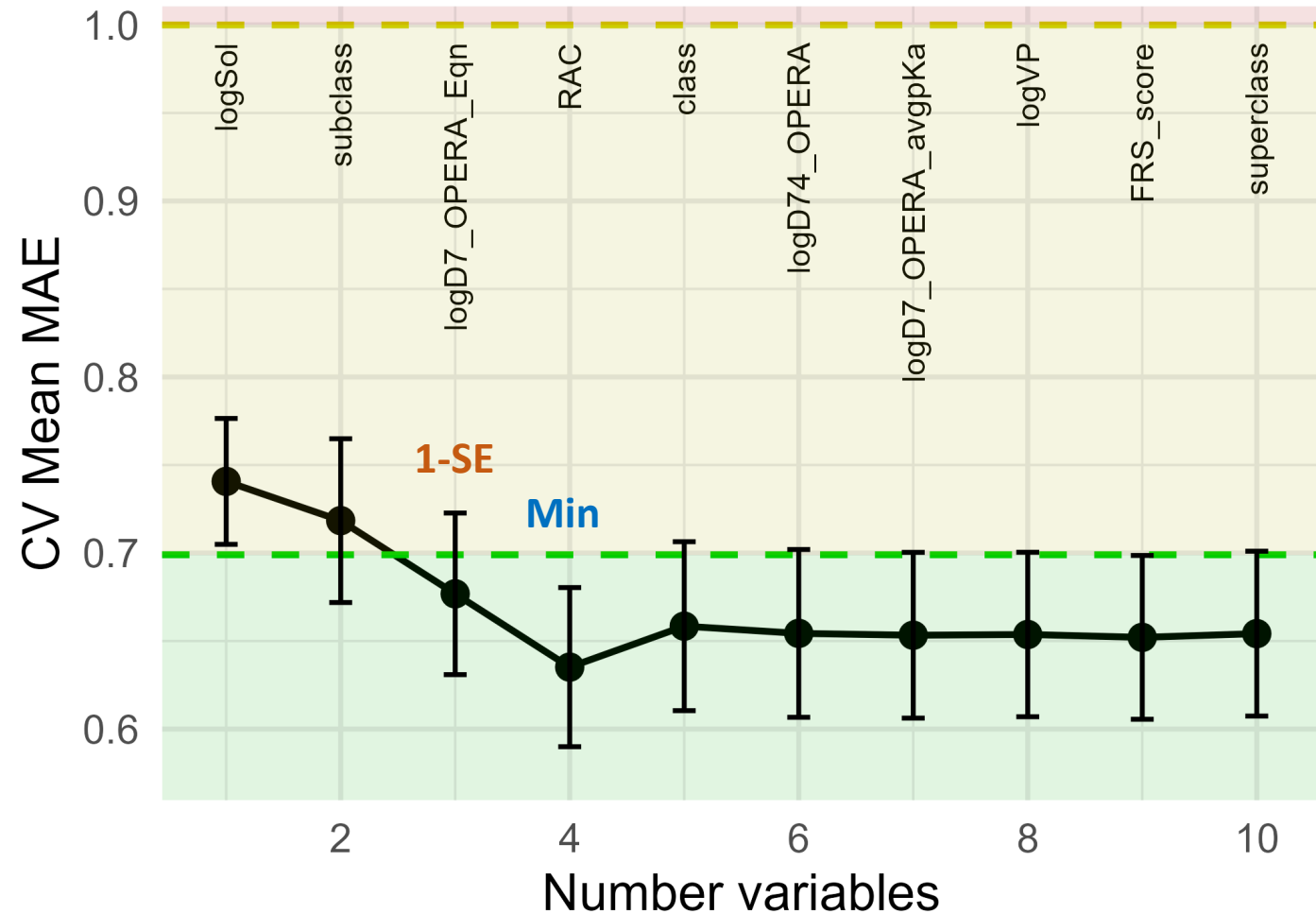
Recursive feature elimination using CV-estimated Mean Absolute Error

RFE order

1. Solubility
 2. Subclass
 3. logD7_OPERA_Eqn
 4. RAC (target taxon)
 5. Class
 6. logD74_OPERA
 7. logD7_OPERA_avgpKa
 8. Vapor pressure
 9. FRS_score
 10. Superclass
- Min error
- 1-SE rule

Final models:

- **TMPPL:** Fish species, Solubility, Subclass, LogD, RAC
- **TMPPLgen:** Fish species, Solubility, Subclass, LogD



*Retain test fish species at each iteration
(+1 to x-axis)

Reduced TMPPL models: Performance for Chemical Averages in Holdout Test Set

Performance – Chemical averages

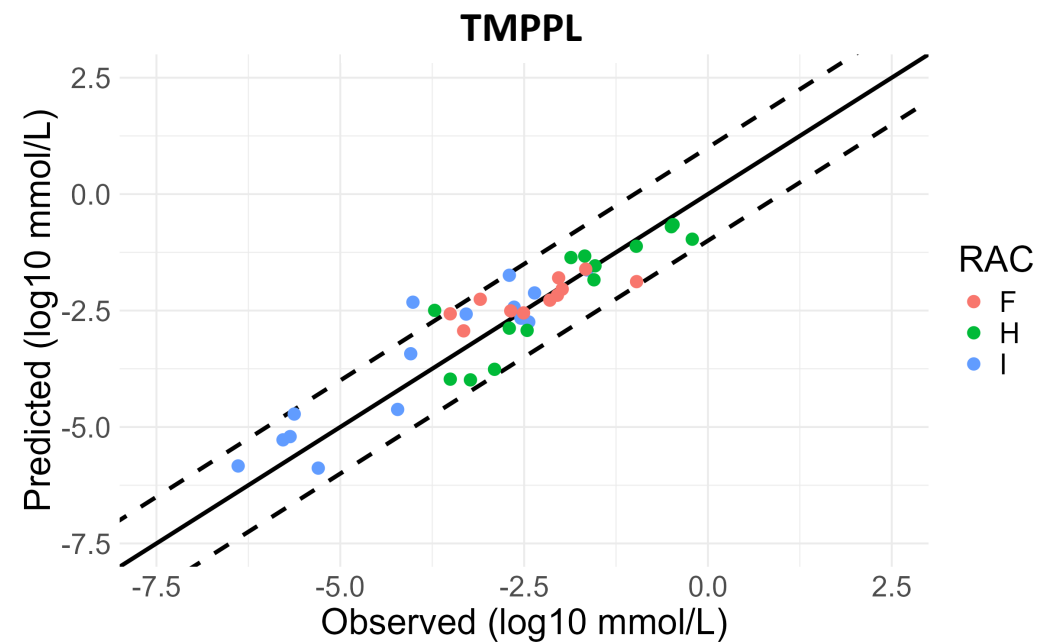
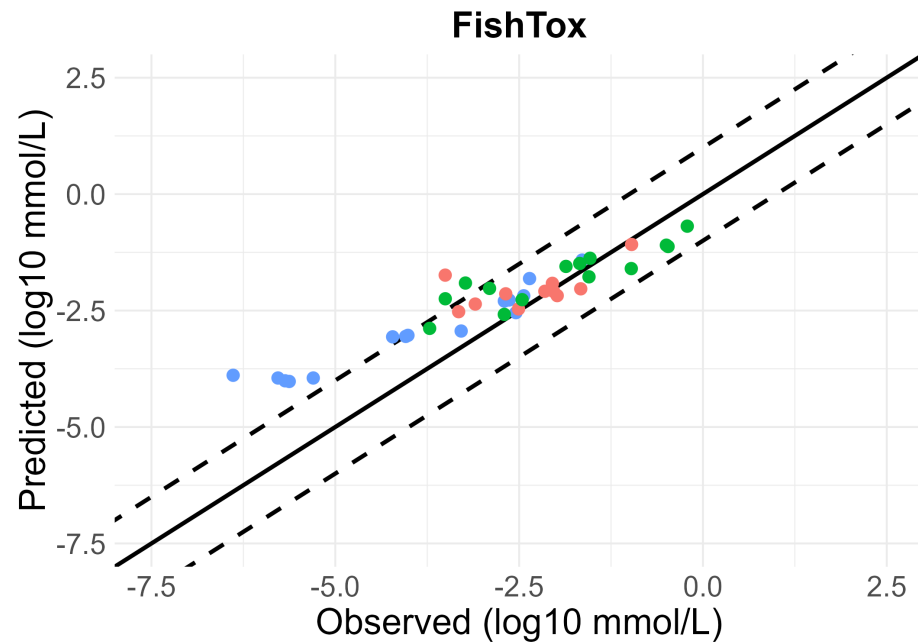
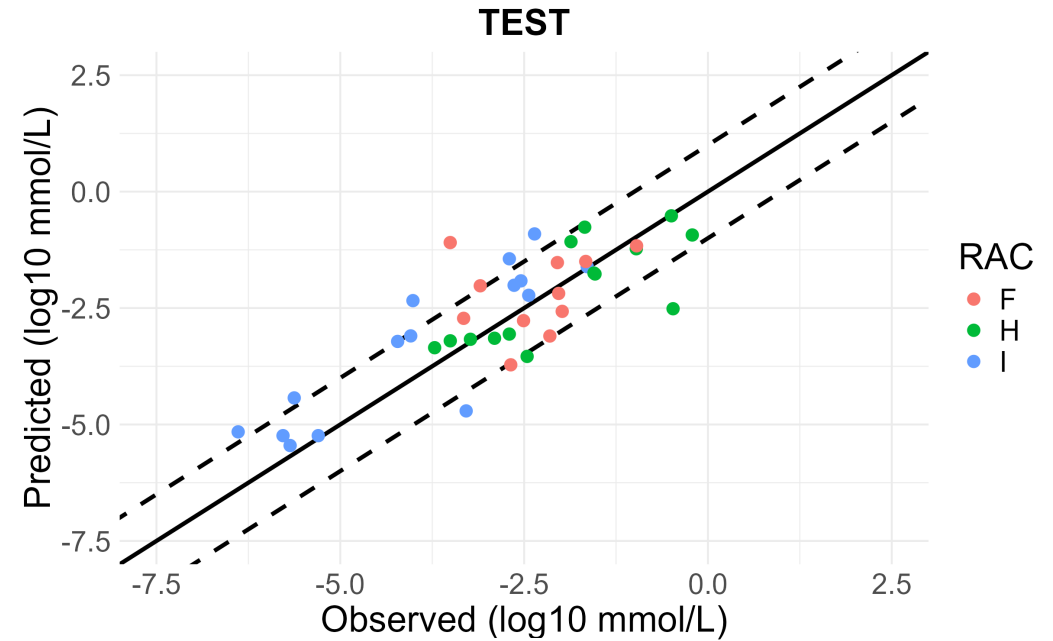
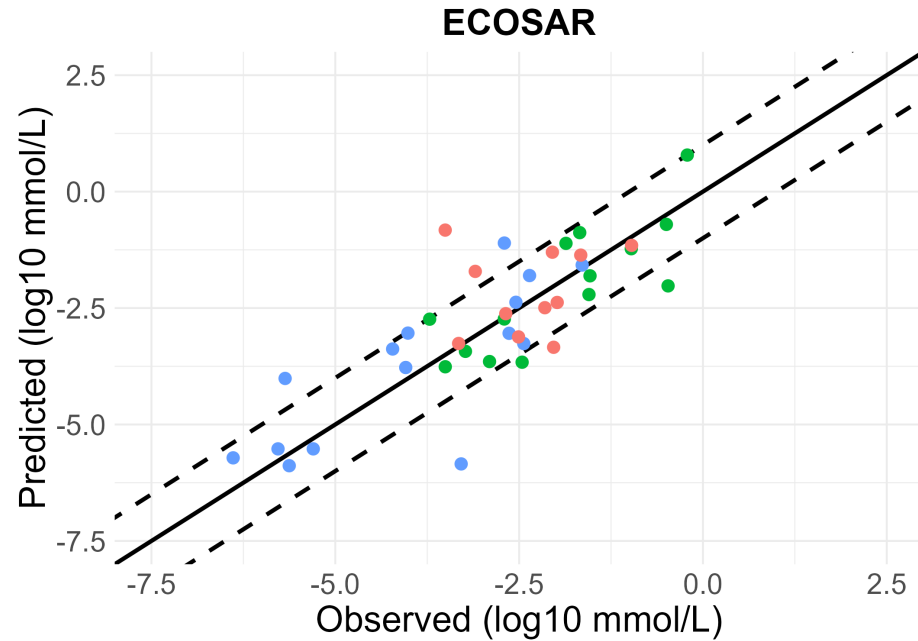
Bold* = Best
Bold = Second Best
Text = Third Best

Good: <5x
Fair: 5-10x
Poor: >10x

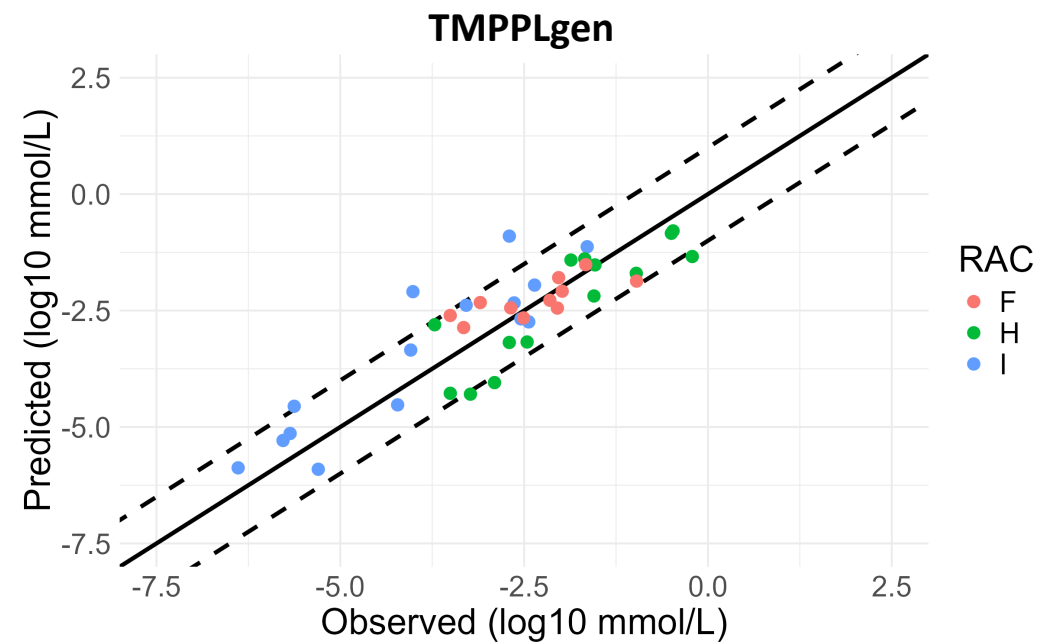
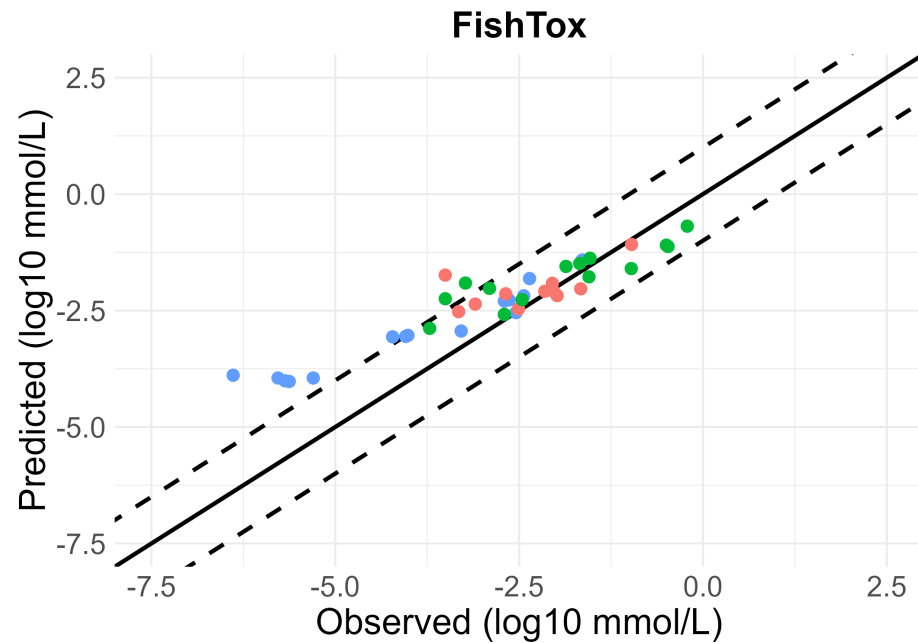
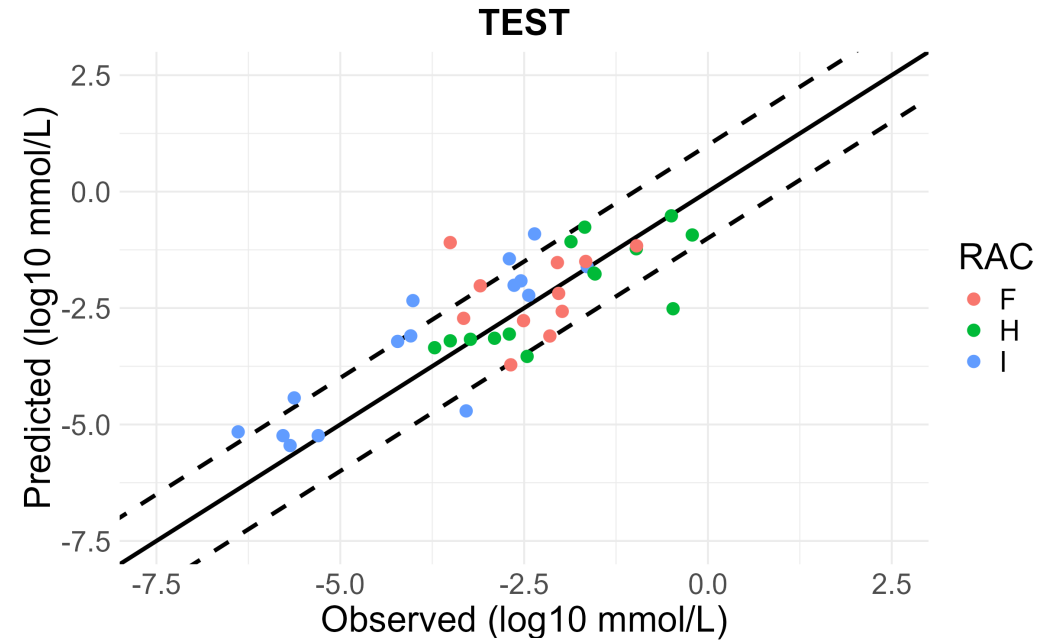
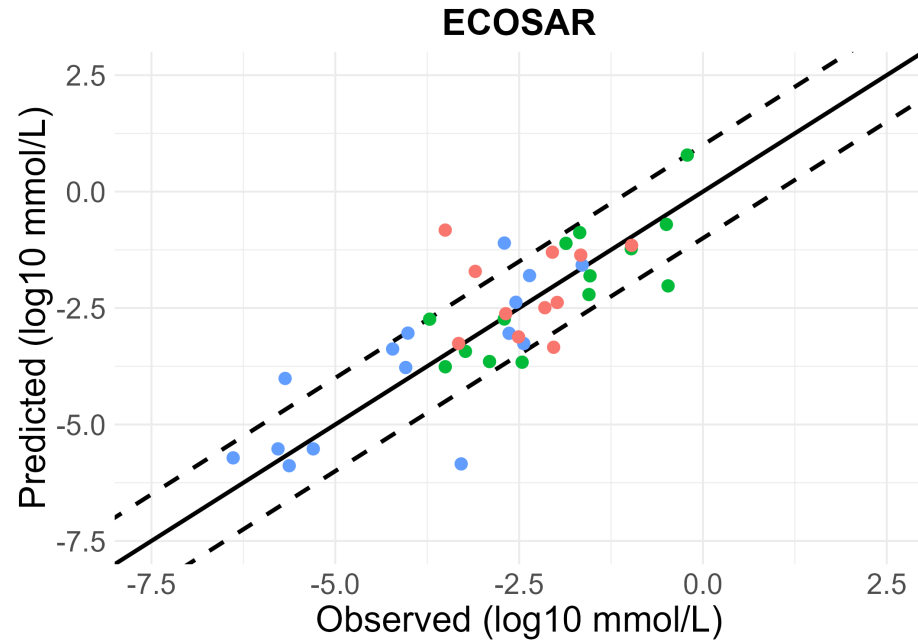
Model	N	MAE	RMSE	Bias	% Good	% Poor	R ²
ECOSAR	40	5.1	8.8	1.2	58	20	0.59
TEST	40	5.0	8.0	1.6	58	30	0.63
FishTox	40	4.7	7.9	3.2	63	23	0.63
TMPPL	40	2.9*	3.9*	1.3	73*	5*	0.84*
TMPPLgen	40	4.0	5.4	1.2	65	15	0.75

- TMPPL outperforms other QSARs on test set
 - TMPPL (with RAC) has better performance than TMPPLgen (without RAC – pesticide target taxon)

Predicted vs. Observed – Chemical Averages



Predicted vs. Observed – Chemical Averages



Overview of performance assessment across models

- Error metrics: Mean absolute error (MAE), Root mean squared error (RMSE), Bias, R^2
- USEPA risk assessment criteria: % Good (<5x), % Fair (5-10x), % Poor (>10x)
- 3 USEPA QSARs vs. Full TMPPL model (all predictors)
 - Cross-validated (CV) errors = Predict new, unseen chemicals
 - Out-of-bag (OOB) errors = Predict new chemical-species combinations
 - Compared performance using all chemicals
 - 1) FHM, 2) All chem-species observations, and 3) Chemical averages
- Reduced TMPPL model (best predictor subset)
 - Use recursive feature elimination to select best predictor subset with training set
 - Compared performance on hold-out data (40 chemicals not in training set)
- Reduced TMPPL model performance for **European Food Safety Authority (EFSA) chemicals unregistered in U.S. (n=17)** and **degradates (n=32)**

Reduced TMPPL model:
Performance for Chemical
Averages in EFSA Test Set and
Degradate Dataset

Performance – Chemical Averages

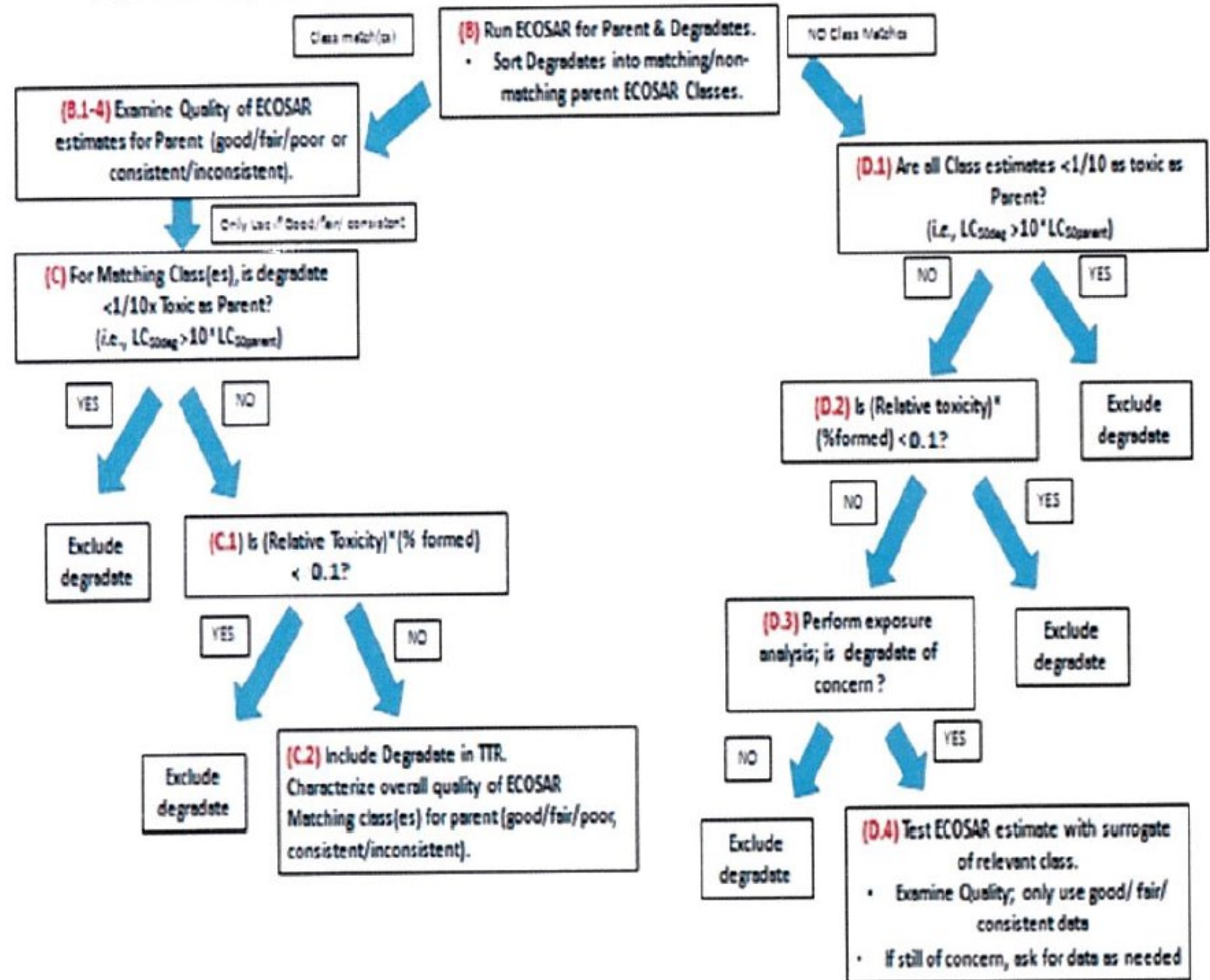
Good: <5x
Fair: 5-10x
Poor: >10x

Model – Test set	N	MAE	RMSE	Bias	% Good	% Poor	R ²
TMPPL – holdout set	40	2.9	3.9	1.3	73	5	0.84
TMPPL – EFSA	17	5.4	8.5	1.9	59	24	0.71
TMPPLgen– EFSA	17	7.7	12.1	1.4	47	29	0.61
TMPPL – Degradates	32	6.4	9.3	1.1	50	34	0.47
TMPPLgen – Degradates	32	8.5	12.3	1.3	47	34	0.34

- Performance on EFSA test set and degradates not as good as holdout set (which is approximately in applicability domain [AD])
- Half of EFSA chemical ClassyFire subclasses not in training data
- One quarter degradate subclasses not in training set

USEPA RoC assessment with TMPPL vs. ECOSAR

Figure 1. Graphical Representation of ECOSAR Guidance

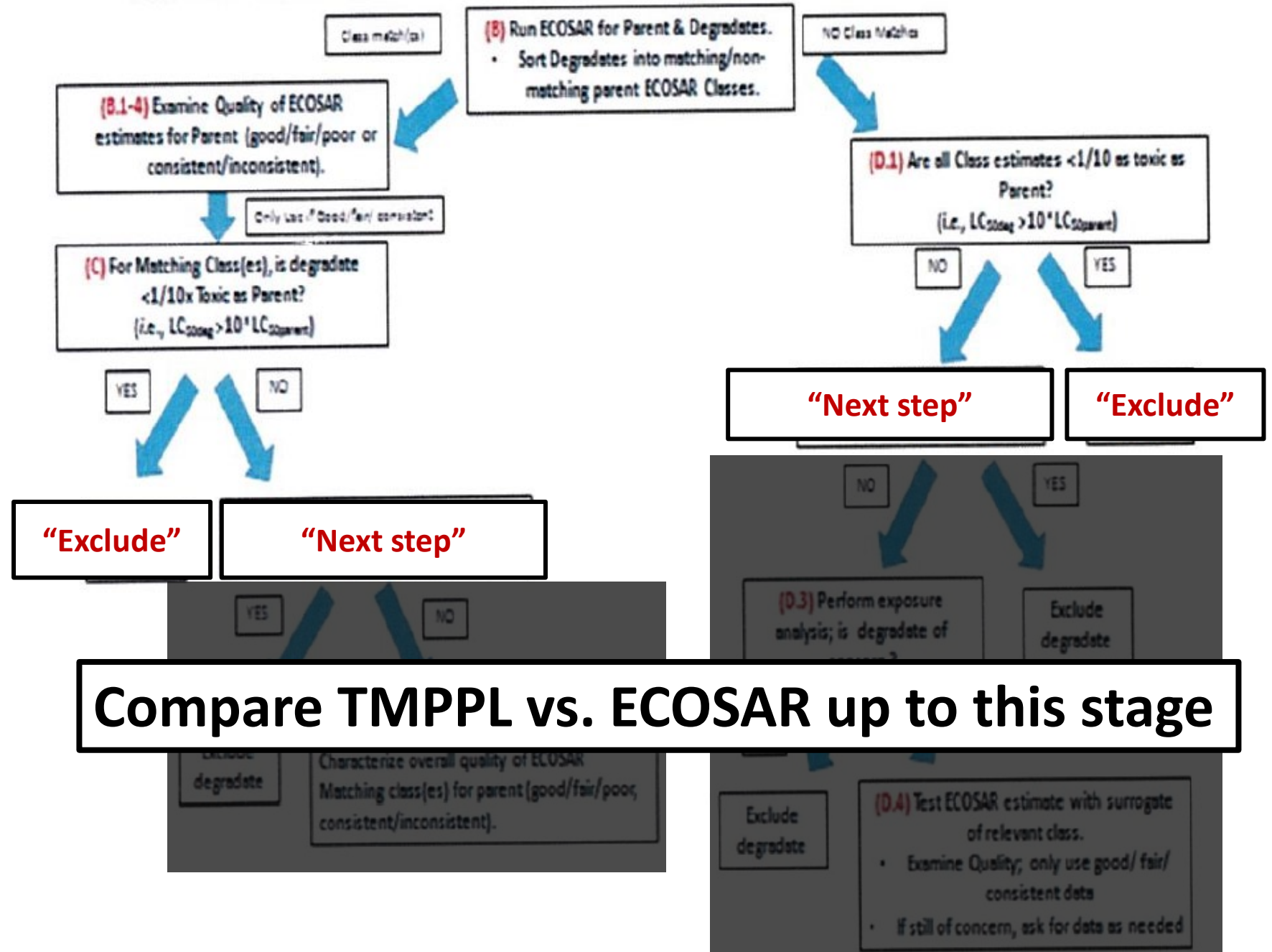


Key Considerations:

- Quality of ECOSAR estimate
- Relative toxicity of pesticide (parent compound) and degrades
- Fraction of degrades formed

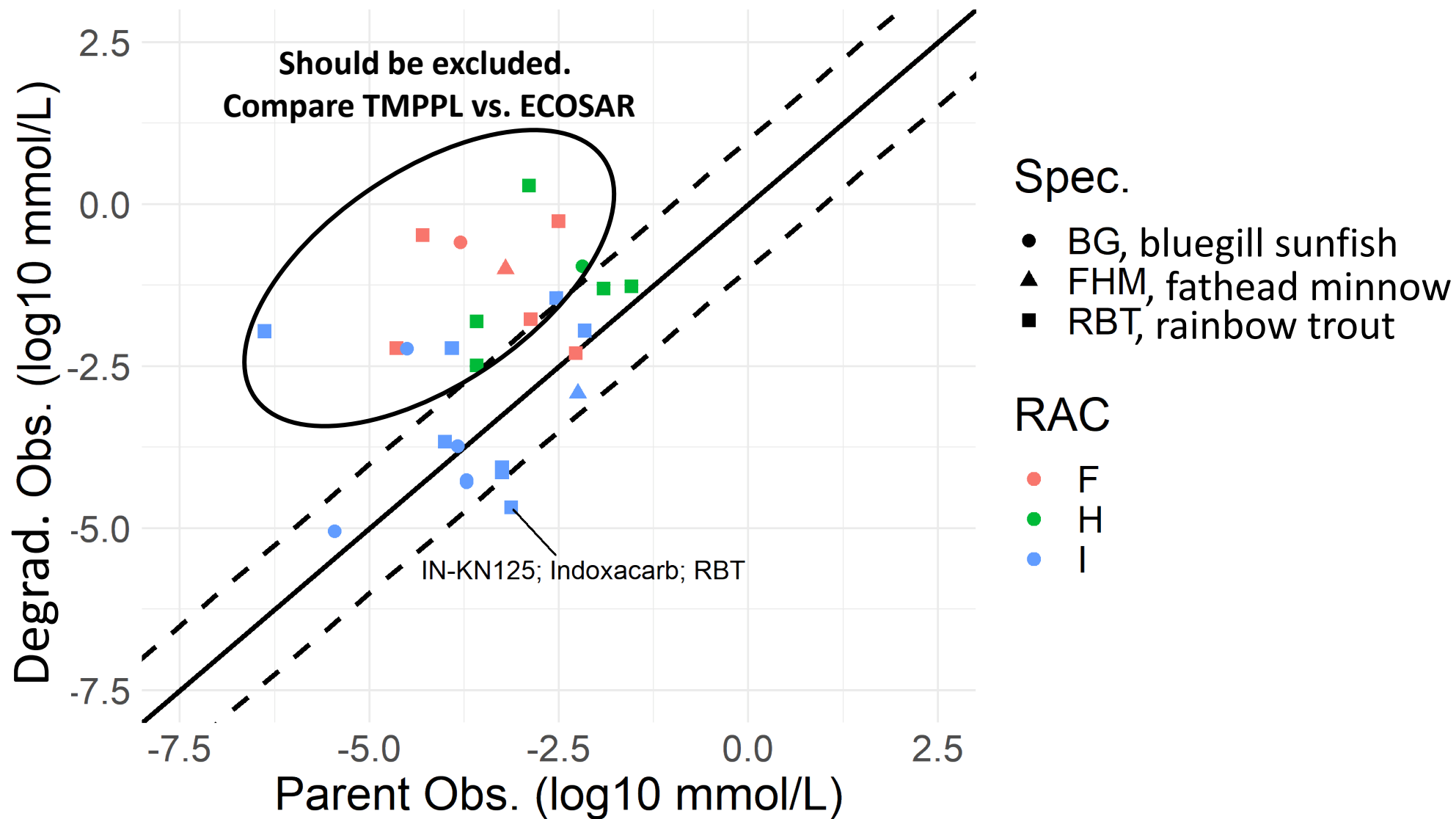
Key Considerations:

- Quality of ECOSAR estimate
- Relative toxicity of pesticide (parent compound) and degradate
- Fraction of degradate formed

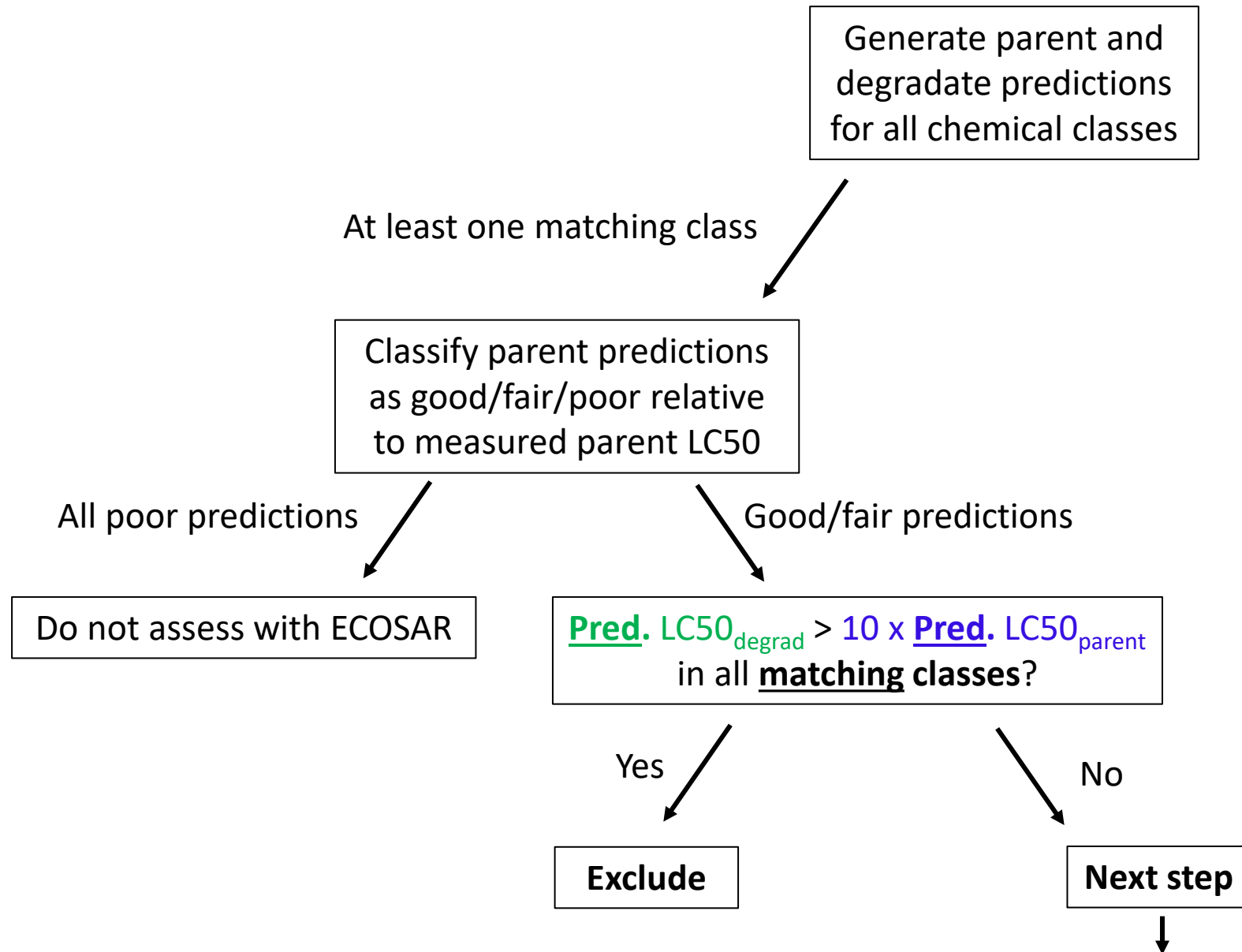


Observed LC50: Degradate vs. Parent

22 degradates with parent data matched by species

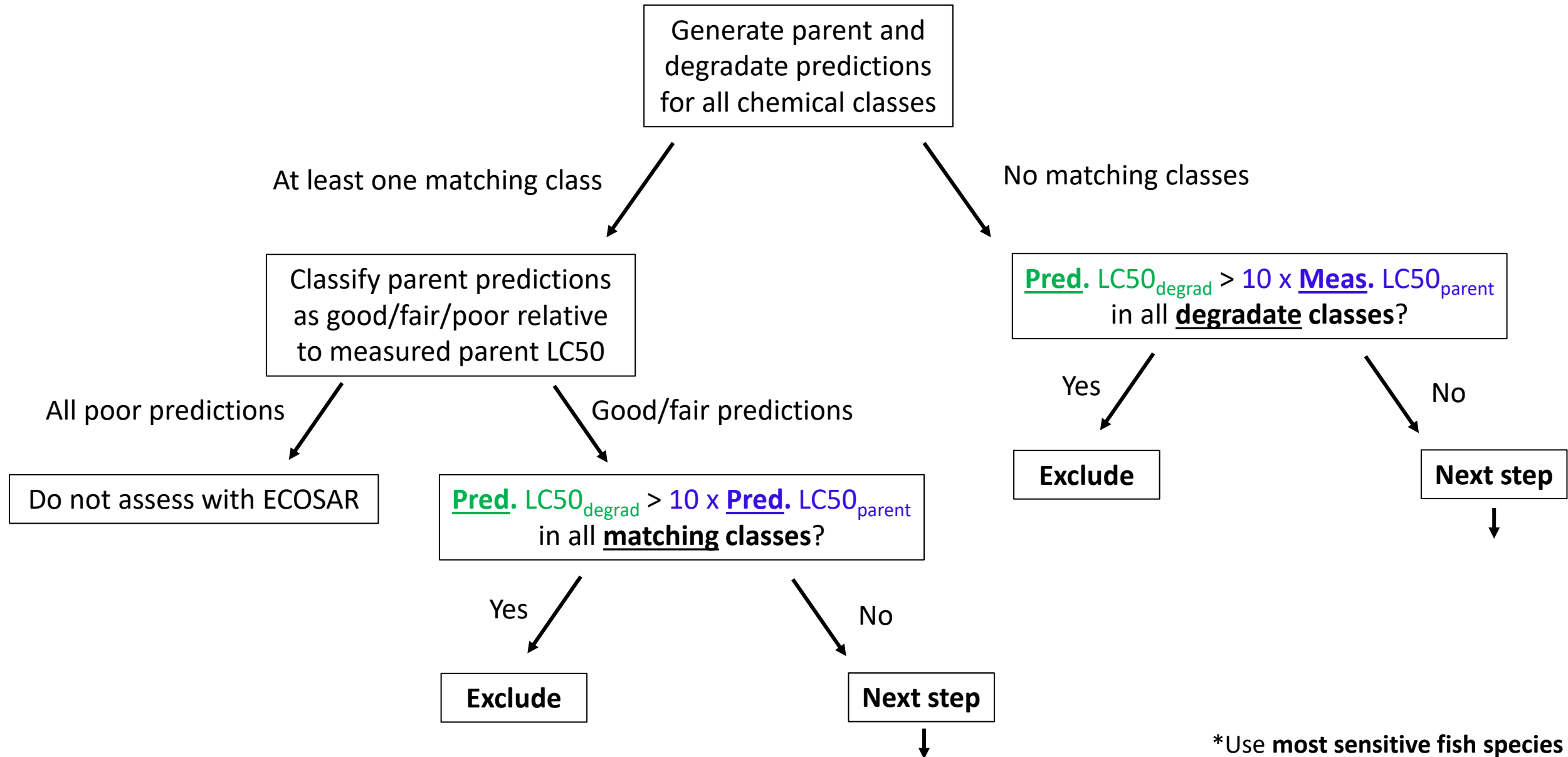


RoC assessment using ECOSAR (multiple chemical classes per compound)

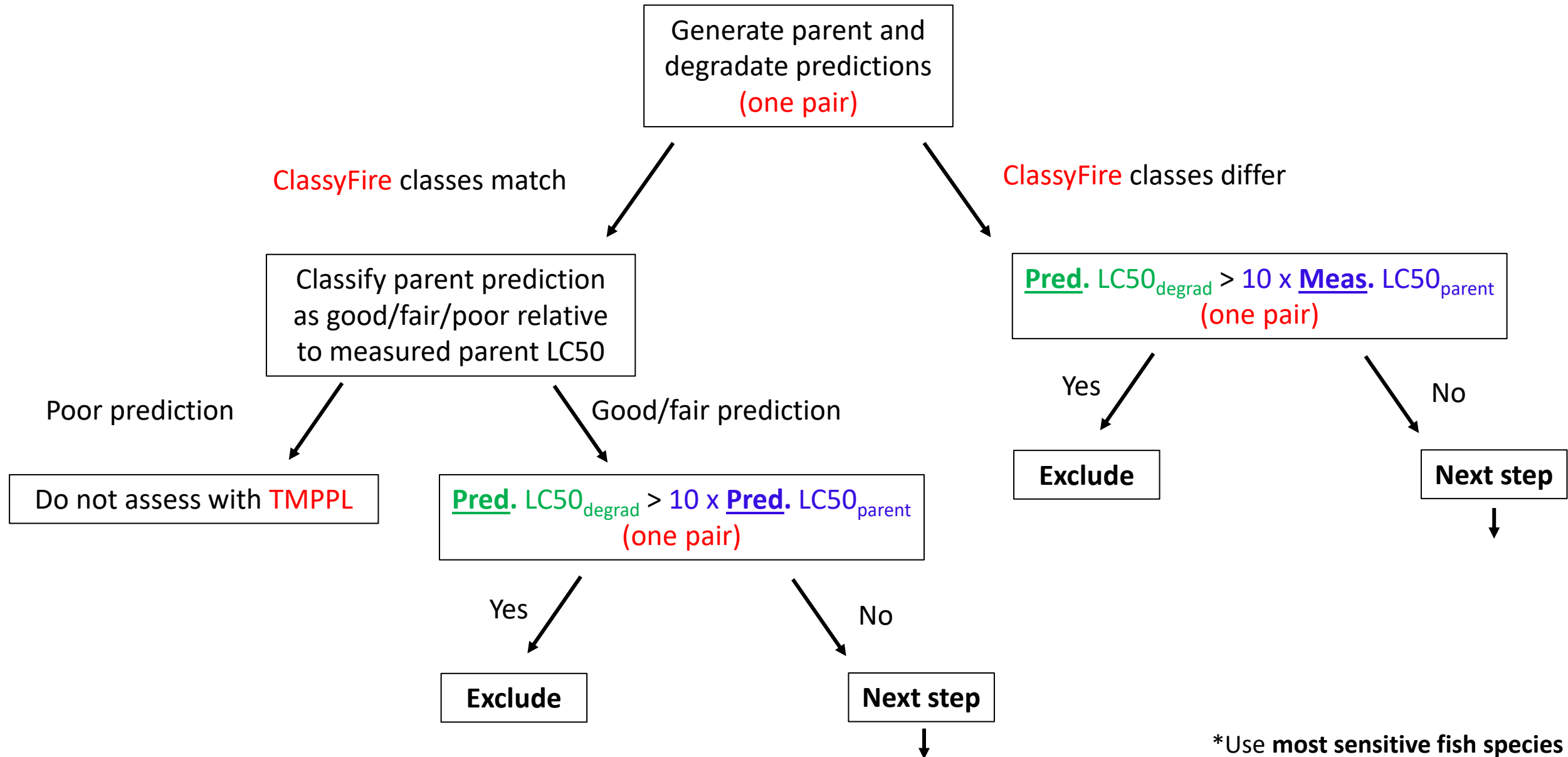


*Use most sensitive fish species

RoC assessment using ECOSAR (multiple chemical classes per compound)



RoC assessment using TMPPL (single chemical class per compound)



Degradate RoC identification accuracy

Model	'Exclude' Accuracy (%)	'Next Step' Accuracy (%)	Overall Accuracy (%)	Degradates Assessed (out of 22)	Degradates Assessed (%)
ECOSAR	67	71	69	16	73
TMPPL	67	78	71	21	95

- TMPPL has comparable or slightly higher RoC accuracy than ECOSAR
 - TMPPL had higher accuracy for 'Next Step' degradates than ECOSAR
- TMPPL tends to err on the conservative side (passing 'Exclude' degradates onto next step)
- Almost all degradates could be assessed with TMPPL (95%) because fewer poor parent predictions
- Small sample size (n=22) limits assessment

SUMMARY

- **Relative model performance matches expectations, esp. given modeled fish species**
 - TMPPL and FishTox best for species-specific predictions, as they are built to do
 - ECOSAR trained on multiple species but produces generic fish prediction
 - TEST trained using only FHM data
- **TMPPL had best predictive ability overall**
 - Trained using data specific to pesticides and targeted predictor set
 - Accounts for predictor interactions and non-linear relationships
- **TMPPL performs better on chemicals in ClassyFire subclasses represented in training set**
- **Assessment of RoCs was improved using TMPPL**
 - Both TMPPL and ECOSAR can be used as lines of evidence in RoC assessment
- **USEPA will continue to evaluate TMPPL for potential use in pesticide assessments**

Acknowledgements

USEPA/OCSP

Dale Hoff

Dan Chang

Todd Martin

Tony Williams

Richard Kolanczyk

USEPA/OPP/EFED

Bill Eckel

Kelsey Vitense

Keith Sappington

Kristin Jones

Katrina White

SpecPro Prof. Services

Tim Dawson

Greg Elonen

Thanks for listening!

Questions?

eckel.william@epa.gov
vitense.kelsey@epa.gov