A new conceptual and statistical modelling framework for pan-Canadian mapping of aquatic ecosystem sensitivity to mercury deposition

Murray Richardson[1], Rachel Plewes[1], David Depew[2], Chris Eckley[3], Ashu Dastoor[4], Satyendra Bhavsar[5]

[1]Department of Geography and Environmental Studies, Carleton University, Email: murray_richardson@carleton.ca

[2] Queen's University [3]U.S. Environmental Protection Agency [4]Environment Canada [5]Ontario Ministry of the Environment

INTRODUCTION AND OBJECTIVE

Aquatic ecosystem Hg sensitivity can be thought of as the relative efficiency with which a lake and its watershed transforms an inorganic Hg load to methyl-mercury in fish (Munthe *et al*., 2007). It is generally accepted that some lakes are more sensitive to atmospherically deposited Hg than others, but it remains difficult to statistically or mechanistically predict the impact of atmospheric Hg pollution on aquatic food webs. The objective of this research is to statistically model Hg load-to-uptake response in lakes across Canada using both "state-based" (fish tissue Hg concentrations) and "response-based" (Hg load-to-uptake relationships) statistical models. The ultimate endpoint for this 3-year project is to develop regionally-varying criteria for maximum tolerable Hg deposition loads at the national scale.

DATASETS AND METHODS

Response variable: Fish tissue Hg concentrations standardized to 12 cm yellow perch using the National Descriptive Model for Mercury in Fish (NDMMF, Depew *et al.*, 2013), aggregated to National Ecological Framework (NEF) ecodistrict level (median concentrations from all sampled lakes within an ecodistrict); n = 4571 lakes in 276 of 1025 NEF ecodistricts.

Predictor variables: Annual net mercury deposition across Canada (Dastoor *et al*., 2004); National Ecological Framework climate and ecosystem variables: Mean Annual Temperature, Total Annual Precipitation, Annual Solar Radiation, Percent Organic Soils, Percent Coniferous Vegetation (variables chosen following redundancy analysis).

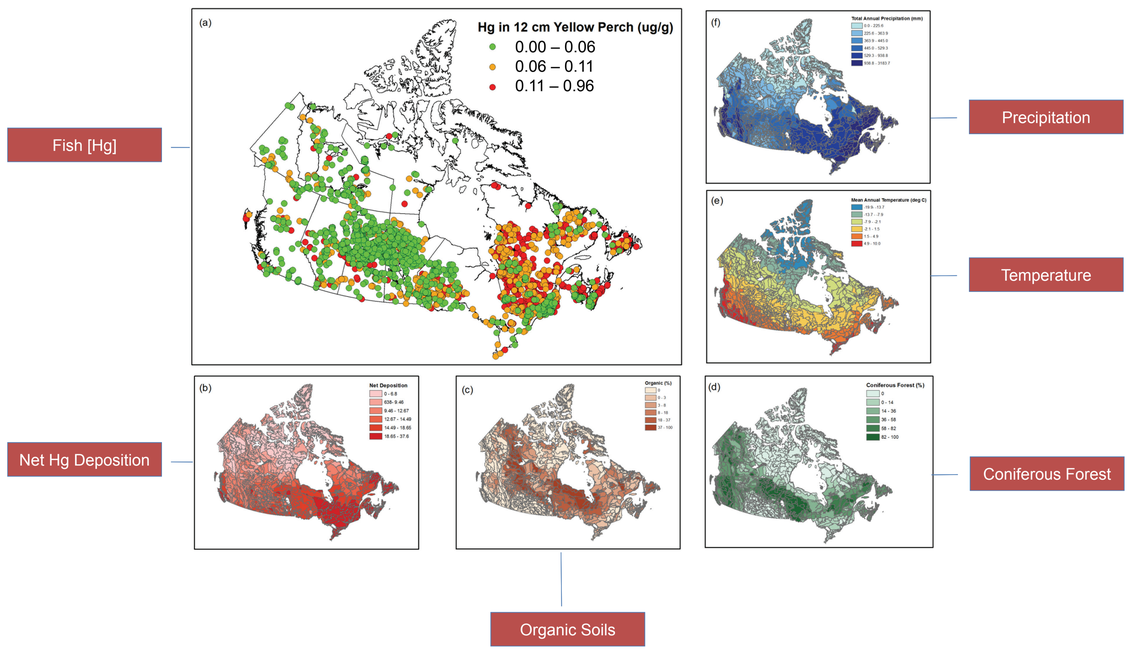

Figure 1: National scale maps showing datasets used in both the state-based and response-based statistical models.

State-based model: Classification and Regression Tree approach via Random Forest to predict fish tissue Hg concentrations at the ecodistrict level and to assess relative importances of predictor variables.

Response-based model: Bayesian tree modelling to predict slope and intercept of regression lines relating fish tissue Hg concentrations to annual Hg deposition load for different clusters of ecodistricts. Ecodistrict clusters that are well-described by a unique regression line can be considered to exhibit similar sensitivity to Hg deposition; the slope and intercept parameters can potentially be used to predict maximum tolerable deposition loads (see example – Figure 2).

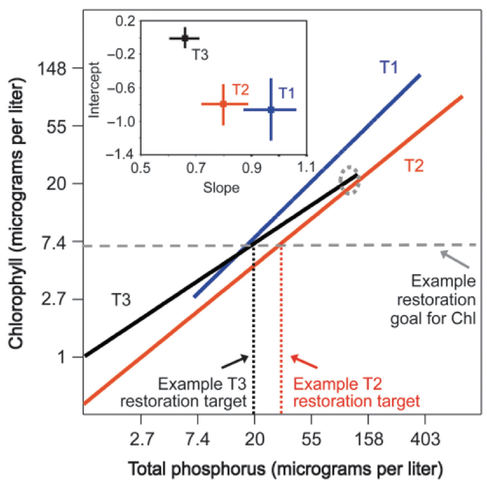

Figure 2: Bayesian trees have been used to set restoration targets for total phosphorous (TP) loading in lakes to maintain acceptable levels of chlorophyll *a*. Each regression line represents a separate group of lakes exhibiting unique (to that group) sensitivity to TP and can be used to determine a maximum tolerable TP concentration . From Soranno *et al.*, (2010).

RESULTS, DISCUSSION AND CONCLUSIONS

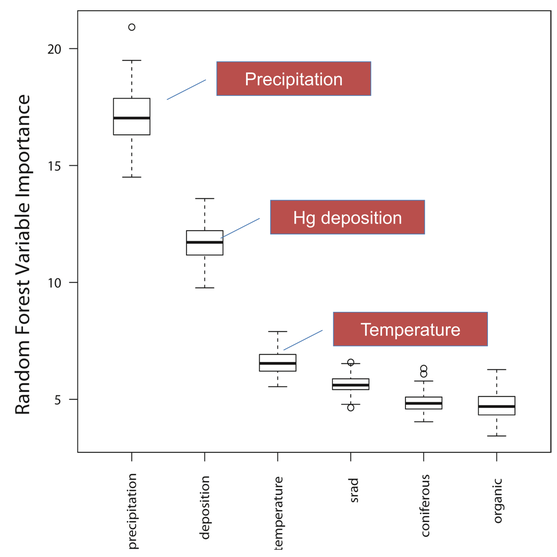

Figure 3: Variable importance plots resulting from state-based Bayesian-tree model.

Variable importance plots resulting from the state-based, random forest model are show in in Figure 3. Total annual precipitation is the most important predictor variable, followed by total Hg deposition and temperature. Hg deposition and precipitation are highly correlated, however, making it difficult to isolate their relative individual importances as predictors of fish Hg concentrations. The response-based approach was developed to address this problem.

Results from the response-based Bayesian tree model are provided in Figures 4-7. The model produced 5 terminal nodes, each representing a cluster of ecodistricts that fall along a unique regression line relating median fish Hg to annual Hg deposition. These ecodistricts were geographically clustered across the country (Figure 4). All regressions were statistically significant with varying slope and intercept terms (Figure 5).

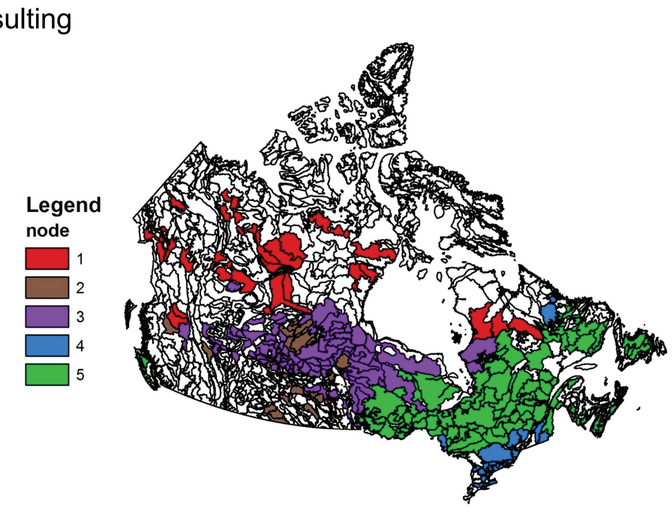

Figure 4: Mapped sensitivity clusters for ecodistricts with available fish Hg data. A similar, predictive map will be produced for all ecodistricts once the response-based model has been finalized.

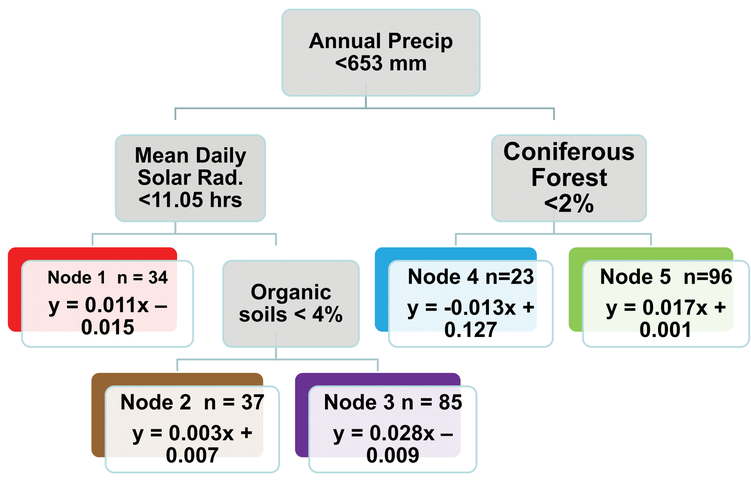

Figure 5: Final Bayesian tree model showing predictor variables and terminal node regression parameters.

Arctic ecodistricts had the largest slope term (most sensitive) and best goodness-of-fit (lowest predictive error), followed by Central Boreal Shield ecodistricts. Lower goodness-of-fit in nodes 2-5 may reflect stronger role of ecosystem factors in modulating Hg sensitivity.

The final response-based model also provides ecodistrict level predictions of median fish Hg concentrations (Figure 6, R^2 = 0.52, RMSE=0.038 μg g^{-1}).

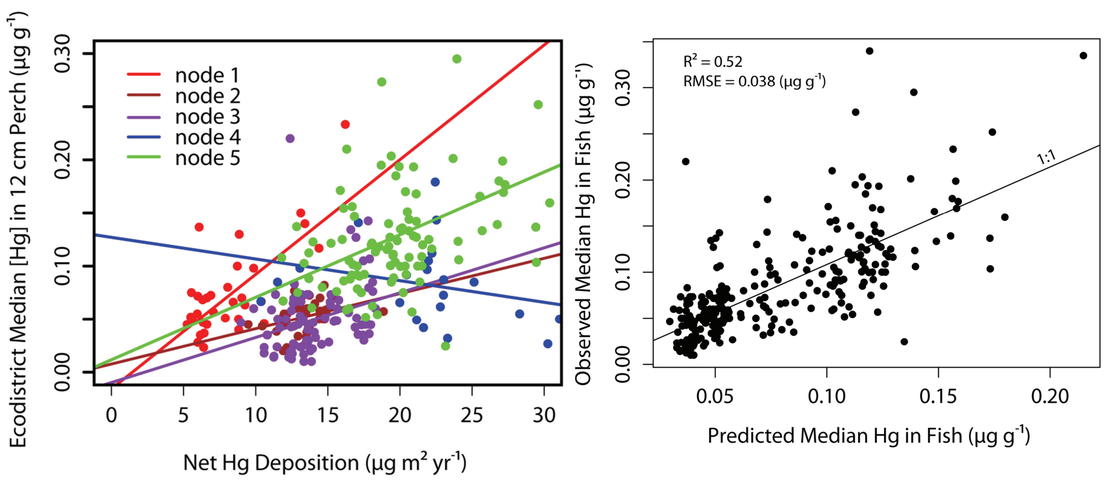

Figure 6: Median fish Hg concentration *vs.* net Hg deposition and fitted regression lines by terminal node (left) and final cross-validated observed *vs.* predicted plot for all ecodistricts (right).

The model explains ~50% of the variability in median, standardized fish Hg concentrations at the ecodistrict level and captures some region-specific variability in aquatic Hg sensitivity at the nationals scale. The final phase of the project will incorporate watershed information at the scale of individual lakes to further refine the response-based Bayesian-tree model.

References: Dastoor *et al*., 2004, *Atmospheric Environment*, Vol. 38(1); Depew et al., 2013, *Canadian Journal of Fisheries and Aquatic Sciences*, Vol. 70; Munthe *et al*., 2007, *Ambio*, Vol. 36(1); Soranno *et al.*, 2010, *Bioscience*. Vol. 60(6). **Funding:** This project is being funded through the Clean Air Mercury Agenda Mercury Science Program, Environment Canada