**Mixed effects model approaches for estimating size-adjusted contaminant concentrations in fish populations.**

Brian W. Kielstra, Calvin Kluke, *“Tom Johnston, Satyendra Bhavsar, Rob Mackereth, Stephanie Melles,”* Gretchen L. Lescord, Erik Emilson

# Key words

Mercury, Arsenic, Fish, Mixed effects models, …

# Abstract (200 words max for ES&T Letters)

Bigger fish are known to have higher concentrations of bioaccumulative contaminants like mercury (Hg). Sampling event regression (SERs) are frequently used to estimate size-adjusted concentrations of these contaminants in fish populations prior to making comparisons or further modeling. However, this approach can be limited by sample size constraints within and across waterbodies. Herein, we describe a mixed effects model approach that borrows strength from all available observations to estimate size-adjusted contaminant concentrations in individual fish populations and then compare amongst inference types and SERs. [Add sentence or two summarizing main results - accurate estimates with n = 2 fish, comparisons among approaches]. We recommend INLA as a reliable means of estimating size-adjusted contaminant concentration in fish populations, particularly when sample sizes are limited.

# Synopsis

TOC Art

# Introduction (Currently 641 words - we should try to cut out ~140 words if possible).

Consumption of wild-caught fish, while abundant with healthy fats and protein, is considered a potentially detrimental source of contaminants to the human body (citation). Bioaccumulative contaminants - which are incorporated into tissues faster than they can be excreted - are of particular concern because they can reach elevated levels in fish, even in remote waterbodies distant from primary sources of pollution. Due to their accumulation over time, the concentrations of these contaminants are often correlated with metrics of fish body size (e.g., weight, length). As a result, many monitoring programs and studies account for fish weight or length in their consumption recommendations (e.g., MECP 2017). This accumulation-size relationship is also accounted for in statistical models used in studies of contaminants (e.g., add example citation) to avoid confounding effects of fish size. A common approach involved in modelling effect-contaminant relationships in fish is to determine the effects of various potential drivers on a size-adjusted estimate of the concentration of the contaminant across lakes (i.e., comparing effects on concentrations of the same size fish from each lake). Sampling event regressions (SERs) are a commonly used approach for estimating size-adjusted means of contaminant concentrations at the population-level; SERs regress contaminant concentrations against metrics of body size within a fish population (usually a unique species-lake combination for a given year) and use the resulting linear equation to estimate the concentration at a chosen size (e.g., 500 cm or 1kg).

The strength of fish contaminant-size relationships vary among different types of contaminants and across fish species or waterbodies, due to complex environmental and metabolic factors. For example, mercury (Hg) is a highly bioaccumulative contaminant that, due to its ubiquitous dispersion, is a global concern that is routinely monitored in fish. Mercury is well known to accumulate in the muscle tissue of fish, due its chemical speciation and inter-organ transport (Peng et al. 2017). Many studies have found strong positive correlations between Hg concentrations ([Hg]) and body size in both freshwater (citations) and marine (citations) environments and, thus, size-adjustments of Hg concentrations prior to making comparisons are common. In contrast, arsenic is less bioaccumulative in fish, showing mixed and weaker relationships with metrics of body size (Lescord et al. 2020; Kluke et al. in prep). Depending on its chemical speciation, arsenic can be a harmful carcinogenic contaminant, but it is less widely distributed and often originates from a localized point source of contamination or natural abundance. It is also less routinely monitored and researched when compared to [Hg] in fish, which results in less data to be available for size-adjustment models. However, not accounting for size in population-level comparisons of arsenic concentrations ([As]) could mean that some variation is unaccounted for, possibly weakening predictions and limiting our understanding of environmental cycling and bioaccumulation patterns.

In general, research and monitoring programs sample waterbodies for fish by collecting a targeted sample size (e.g., 10-20 individuals/species) across as broad of a size range as possible. However, such opportunistic sampling can limit the number of fish caught, the size range represented, or the distribution of fish sizes for a given population – all of which may lead to over/under estimating concentrations when using an SER approach. Alternatively, the deficient population may be excluded altogether, sacrificing valuable data often from underrepresented waterbodies or fish species. Bayesian linear mixed effects models are a potential solution as they utilize all data across sampling structures (e.g., waterbodies) to inform predictions and allow for explicitly modeling variation of random effects variables.

Herein, we present a novel statistical approach to generating size-adjusted contaminant estimates at the population level using Bayesian inference through . More specifically, we developed a mixed model approach that borrows strength from fish across all populations (i.e., pooled across species and waterbodies) to generate predictions for individual lakes . [Additional details - slopes, intercepts, ML, AB, B?]. We compared the results of our INLA models to the traditional SER approach and to real-world measurements at the targeted size metrics (i.e., 0.5 and 1kg). Two contaminant measures - total [Hg] and [As] in fish muscle tissue - were used in this study to assess differences in the modeling approaches based on bioaccumulative potential and data availability.

# Materials and methods

## Datasets

We obtained fish-level Hg and As data for inland waterbodies across Ontario from the Ontario Ministry of Environment, Conservation, and Parks (OMECP). Although individual fish sampling protocols varied, muscle tissue sent to OMECP is analyzed using standardized methods as part of the Fish Contaminants Monitoring Program. More specifically, total [Hg] were measured using cold vapor-flameless atomic absorption spectroscopy (CV-FAAS) following protocol HGBIO-WS057 and total [As] was measured using Inductively Coupled Plasma Mass Spectrometry (ICP-MS) following method BIOTA-E3461. We limited our analyses to inland lakes and three species of interest: *Salvelinus namaycush namaycush* (common Lake Trout, hereafter LT), *Esox lucius* (Northern Pike, hereafter NP), and *Sander vitreus* (Walleye, hereafter WE). These three species represent important food fish in Ontario, sought by subsistence and sport fishers across the province. While they are all predatory fish, LT are generally restricted to profundal zones, while NP and WAL have broader movement patterns. In total, the final dataset we used included [Hg] in 37,923 fish and [As] in 1,001 fish. A full description of the data availability for each fish species and contaminant is in Table 1.

TABLE: number of sampling events with at least four fish, number of sampling events with more than one fish, for each species/contaminant.

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Sampling event regressions For each contaminant-species-waterbody-year combination (i.e., a sampling event), we developed log-contaminant (ug/g) by log-weight (g) regression models for those combinations with at least 5 sampled individuals (cite XX, XX, XX, …). Each model used the following formula:  where *i* is an individual fish for a species-waterbody-year combination; these models were run until all combinations were exhausted. We used these models to generate contaminant predictions and their 95% confidence intervals for a representative 1000 g fish per sampling event. Mixed effects regression models using maximum likelihood, approximate Bayesian, and Bayesian inference For each contaminant-species combination, we developed log-contaminant (ug/g) by log-weight (g) mixed effects regression models. We allowed for random variation in the slope and intercept per waterbody and random variation in the intercept by sampling event (i.e., waterbody-year combination). This identical model structure was fit using maximum likelihood inference in *lme4* (**ML** models; cite XX), using approximate Bayesian inference using integrated nested Laplace approximation in *INLA* through *R-INLA* (**AB** models; cite XX), and using Bayesian inference with Markov Chain Monte Carlo in *Stan* through *rstanarm* (**B** models; cite XX) in *R* (cite XX). We used default priors for both *R-INLA* and *rstanarm’s* *Stan* implementation; we used XX chains, a XX burn-in, and a XX sampling rate for the *Stan* model. We undertook posterior sampling checks for the *R-INLA* and *Stan* model. Conceptually, the set of models looked like:  *log(contaminantijk) = ꞵ1 + ꞵ2 × log(weightijk) + b b1j + b*1*jk εi*  *εi ~ N(0, σ2)* ,  Log(contaminantxyz) = Intercept + Log(Weightxyz) + (Log(Weightiy)|Lakeiy) + (1|Lake:Yeariyz) for species *i*.  Again, we used these models to generate contaminant predictions and their 95% confidence intervals for a representative 1000 g fish for each species-waterbody-year combination. For ML models, we generated 95% bootstrapped confidence intervals for the model coefficients and predictions. In ML models, predictions typically do not incorporate uncertainty of the random effects; we partially overcame this by using *lme4’s* “bootMer” parametric bootstrapping with 2000 simulations and *use.u = TRUE* settings - this assumes random effect estimates as constant but resamples at the observation-level (cite bootMer documation). For the AB and B models, we generated analogous 0.5 quantile and 95% credible intervals (0.025 and 0.975 quantiles) from the posterior distributions. Comparing fit and accuracy of predictions across model types We calculated four measures of fit to compare across model types. We calculated the root mean squared error (RMSE) at the individual lake level, RMSEevent, and pooling all observations, RMSEglobal. Similarly, we took the R2 from a linear model of observed versus fitted values at the sampling event level, R2event, and pooling all observations R2global. This gave us a means of comparing SER with the mixed effects models.    In addition to evaluating model fit, we used 1000 g +/- 50 g fish from the Hg dataset and As dataset for comparing predictions against observations. We used RMSEglobal and R2global for each contaminant-species-model type (SER, ML, AB, and B) combination for comparison. Visually comparing precision of predictions when increasing sampling event observations For a subset of contaminant-species-waterbody-year combinations, we compared the precision of 1000 g fish predictions when increasing sampling event observations. To do this, we first took the waterbody-year combination with the most samples per contaminant-species combination. We then randomly sampled one observation and excluded all others from the sampling event, re-ran the models, and generated the 1000 g fish predictions and associated 95% confidence/credible interval. In subsequent runs, we added a new randomly sampled observation to the previously selected observation(s) while excluding others. We did not track the effect of increasing sampling event observations with the Bayesian inference model since each model took approximately XX hours to run and XX x XX iterations made this investigation unfeasible. For display purposes, we found this to be reasonable.  *Other comparisons or interest?* Results and discussion General results here  *Hg- and As-weight relationships*  In general, Model summaries Format - descriptor format and up to two follow up paragraphs  The mixed effects models produced more predictions than the SER models. For Hg, we could use seven more lakes for LT, 53 more for NP, and 23 more for WE than the SER approach (i.e., sampling events with <4 fish were discarded). For As, we could use XX more for LT, XX more for NP, and XX more for WE than the SER approach.  The models were a good fit for the data based on the R2 of linear models relating fitted values to the observed values. Comparing accuracy of predictions across model types The 1000 g predictions were a good fit to the 1000 g +/- 50 g observations across all  contaminant-species and model types based on the R2 of linear models between observed and  predicted. In general, R2 was higher for Hg than As.  *Visually comparing precision of predictions when increasing sampling event observations* Comparing accuracy and precision of predictions when increasing sampling event observationsAcknowledgementsReferencesFigures   Figure 1 - Distribution of sampling event regression R2 values (histogram and SER vertical line) and mixed effects models R2 obtained from linear models of observed versus fitted values (vertical lines) by fish species. For mixed effects models, ML is maximum likelihood, AB is approximate Bayesian, and MB is Monte Carlo Bayesian (see main text for distinction).  Probably need a graph of accuracy vs fish size for each model to compare that  Results of prediction accuracy of each fish from each approach with the total # of fish included in the sampling event that it was collected    Figure 3 - Tables Table 1: Sampling characteristics of Hg and As datasets.   |  |  |  |  |  |  | | --- | --- | --- | --- | --- | --- | | Contaminant | Species | nobservation | nlake | nobs/lake in year; median  (min, max) | Fish weight  5% (2.5%, 97.5%) quantiles | | Hg | LT | 6122 | 411 | 10 (1-24) | 1300 (155, 5800) | |  | NP | 13870 | 1001 | 9 (1-31) | 1150 (275, 5262) | |  | WE | 17931 | 955 | 10 (1-55) | 754 (120, 3018) | |  |  |  |  |  |  | | As | LT | 120 | 22 | 4 (1-9) | 984 (105, 6502) | |  | NP | 465 | 90 | 3 (1-22) | 1260 (204, 6479) | |  | WE | 416 | 83 | 3 (1-30) | 955 (100, 4026) | |
|  |

Table 2: Model summaries for Hg and As models for Lake Trout (LT), Northern Pike (NP), and Walleye (WE) based on Sampling Event Regression (SER), maximum likelihood mixed effects models (ML), approximate Bayesian mixed effects models (AB), and MCMC Bayesian mixed effects models (MB). Root mean squared error compares fitted versus observed values. R2 is the coefficient of determination from a linear model of fitted versus observed values. RMSE (1000 +/- 50 g; n) is the root mean squared error for all fish 1000 g +/- 50 g (Hg) or 100 g (As) and *n* is the number of observations. Columns denoted 2 represent the standard deviation of the estimated intercepts and slopes in the case of the waterbody-year-species sampling event regressions

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| Model | RMSEevent | RMSEglobal | R2overall | 2intercept (waterbody) | 2slope (waterbody) | 2intercept (waterbody, year) | 2residual |
| Hg - LT - SER | 0.27  (0.05, 1.04) |  | 0.60 (0.00, 1.00) | 2.88 | 0.37 | – | – |
| Hg - LT - ML | 0.33 |  | 0.83 | 1.46 (1.35, 1.58) | 0.16  (0.15, 0.18) | 0.12  (0.10, 0.14) | 0.35  (0.35, 0.36) |
| Hg - LT - AB | 0.33 |  | 0.83 | 1.68  (1.51, 1.89) | 0.21  (0.19, 0.24) | 0.14  (0.12, 0.16) | 0.36  (0.35, 0.36) |
| Hg - LT - MB | 0.33 |  | 0.83 | 1.67 (1.49, 1.85) | 0.20  (0.17, 0.22) | 0.15 (0.12, 0.17) | 0.59  (0.59, 0.60) |
|  |  |  |  |  |  |  |  |
| Hg - NP - SER | 0.28  (0.03, 0.70) |  | 0.60  (0.00, 0.99) | 2.54 | 0.35 | – | – |
| Hg - NP - ML | 0.33 |  | 0.81 | 0.99  (0.91, 1.08) | 0.12  (0.11, 0.14) | 0.11 (0.09, 0.12) | 0.35  (0.35. 0.36) |
| Hg - NP - AB | 0.33 |  | 0.81 | 1.24  (1.13, 1.37) | 0.17  (0.16, 0.19) | 0.14  (0.12, 0.15) | 0.36  (0.35, 0.36) |
| Hg - NP - MB | 0.33 |  | 0.81 | 1.23  (1.12, 1.35) | 0.16  (0.15, 0.18) | 0.13  (0.12, 0.16) | 0.60  (0.59, 0.60) |
|  |  |  |  |  |  |  |  |
| Hg - WE - SER | 0.25  (0.03, 0.82) |  | 0.70 (0.00, 0.98) | 2.43 | 0.33 | – | – |
| Hg - WE - ML | 0.30 |  | 0.84 | 1.27  (1.21, 1.32) | 0.14  (0.13, 0.15) | 0.14  (0.13, 0.15) | 0.31  (0.31, 0.32) |
| Hg - WE - AB | 0.30 |  | 0.84 | 1.45  (1.37, 1.57) | 0.19  (0.17, 0.20) | 0.16  (0.15, 0.18) | 0.31  (0.31, 0.32) |
| Hg - WE - MB | 0.30 |  | 0.84 | 1.45  (1.36, 1.56) | 0.17  (0.15, 0.18) | 0.17  (0.16, 0.19) | 0.56  (0.56, 0.56) |
|  |  |  |  |  |  |  |  |
| As - LT - SER | 0.54  (0.10, 1.11) |  | 0.22  (0.00, 0.92) | 2.63 | 0.41 | – | – |
| As - LT - ML | 0.60 |  | 0.79 | 1.05  (0.00, 2.61) | 0.16  (0.01, 0.40) | 0.00  (0.00, 0.36) | 0.66  (0.57, 0.77) |
| As - LT - AB | 0.63 |  | 0.77 | 0.63  (0.33, 1.60) | 0.26  (0.19, 0.35) | 0.01  (0.00, 0.03) | 0.71  (0.62, 0.82) |
| As - LT - B | 0.66 |  | 0.77 | 0.20  (0.03, 1.11) | 0.14  (0.07, 0.23) | 0.20  (0.01, 0.72) | 0.84  (0.78, 0.91) |
|  |  |  |  |  |  |  |  |
| As - NP - SER | 0.29  (0.06, 0.78) |  | 0.34  (0.01, 0.87) | 2.71 | 0.37 | – | – |
| As - NP - ML | 0.38 |  | 0.86 | 1.22  (0.54, 1.80) | 0.26  (0.17, 0.34) | 0.20  (0.00, 0.36) | 0.44  (0.41, 0.47) |
| As - NP - AB | 0.41 |  | 0.84 | 0.79  (0.50, 1.23) | 0.21  (0.17, 0.25) | 0.09  (0.01, 0.23) | 0.46  (0.44, 0.49) |
| As - NP - B | 0.42 |  | 0.85 | 0.52  (0.04. 1.65) | 0.17  (0.10, 0.31) | 0.25  (0.04, 0.48) | 0.68  (0.65, 0.71) |
|  |  |  |  |  |  |  |  |
| As - WE - SER | 0.28  (0.04, 0.75) |  | 0.27  (0.00, 0.98) | 2.37 | 0.36 | – | – |
| As - WE - ML | 0.32 |  | 0.77 | 0.95  (0.43, 1.36) | 0.15  (0.08, 0.21) | 0.08  (0.00, 0.21) | 0.37  (0.34, 0.40) |
| As - WE - AB | 0.32 |  | 0.76 | 1.04  (0.62, 1.70) | 0.21  (0.15, 0.29) | 0.01  (0.00, 0.02) | 0.38  (0.35, 0.41) |
| As - WE - B | 0.37 |  | 0.75 | 0.21  (0.02, 1.27) | 0.08  (0.05, 0.21) | 0.17  (0.01, 0.35) | 0.63  (0.60, 0.65) |

|  |  |  |
| --- | --- | --- |
| Model | R2fixed | R2random |
| Hg - LT - SER |  |  |
| Hg - LT - ML |  |  |
| Hg - LT - AB |  |  |
| Hg - LT - B |  |  |
|  |  |  |
| Hg - NP - SER |  |  |
| Hg - NP - ML |  |  |
| Hg - NP - AB |  |  |
| Hg - NP - MB |  |  |
|  |  |  |
| Hg - WE - SER |  |  |
| Hg - WE - ML |  |  |
| Hg - WE - AB |  |  |
| Hg - WE - MB |  |  |
|  |  |  |
| As - LT - SER |  |  |
| As - LT - ML |  |  |
| As - LT - AB |  |  |
| As - LT - B |  |  |
|  |  |  |
| As - NP - SER |  |  |
| As - NP - ML |  |  |
| As - NP - AB |  |  |
| As - NP - B |  |  |
|  |  |  |
| As - WE - SER |  |  |
| As - WE - ML |  |  |
| As - WE - AB |  |  |
| As - WE - B |  |  |

|  |
| --- |
| R2  (1000 +/- 50 g; *n*) |
| 0.75  (236) |
| 0.75  (243) |
| 0.75 (243) |
| 0.75 (243) |
|  |
| 0.74  (720) |
| 0.72 (773) |
| 0.72 (773) |
| 0.72 (773) |
|  |
| 0.74  (859) |
| 0.76  (882) |
| 0.76  (882) |
| 0.76  (882) |
|  |
|  |
|  |
|  |
|  |
|  |
|  |
|  |
|  |
|  |
|  |
|  |
|  |
|  |
|  |

|  |
| --- |
|  |

Figure 1:

Map of sampling points for Hg and As in Ontario.

Figure 1:

Histogram of SER R2 and ablines for glmmTMB, INLA, and RSTAN

Figure 2:

Histogram of R2 vs. INLA and RSTAN

Overall fit

Predictions and their uncertainty

3 species - NP, WE, NP

We compared sampling event regressions (i.e., 1 regression per sampling event), maximum likelihood mixed effects models, Bayesian mixed effects models using INLA, and Bayesian mixed effects models using Stan.

In sampling event regressions, log-log linear regressions are used. A prediction is made based on this regression at the. To standardize across lakes, fish may be . By borrowing strength of other observations in the dataset, … .

Table:

Confidence intervals

Early e-mails

**Gretchen mentioned this in the past, but I think we have a strong analysis to produce a short paper to CJFAS.** I could lead the paper with Calvin as second author and Gretchen and Erik as senior authors (possibly Stephanie Melles, Rob Mackereth, Tom Johnston, and/or John Gunn since the work builds on things they’ve done). The idea would be to contrast predictions at the whole-lake level using “sampling event regression” (i.e., lake-level regressions) and the mixed model combined approach (glmmTMB [fast, maximum likelihood], INLA [fast – Bayesian], and possibly STAN [slow, like 16 hours slow – Bayesian]. We could compare overall fit and predictions but also use Calvin’s analysis to demonstrate how predictions compare with small sampling sizes and when sites deviate strongly in terms of absolute concentration. We could also consider comparing a strong regression relationship (Hg ~ mass) vs. a weak regression (As ~ mass). The code is written for all of this already. Just needs to be brought together in one paper. Simple and potentially impactful paper (3 studies are currently using this approach).

Hi all,

Hope all is well with you. I’m writing this since my approach to predicting Hg (or, possible As) in fish at the waterbody-level is being used in several projects. Tom, Gretchen, and Calvin were interested in hearing more about my approach. I thought I would include those working on the other projects. I’d rather hear any comments now than stumble into a roadblock down the line. This is exploratory at this point.

Ultimately, we’re taking a Hg~length/mass regression as a backbone model to predict Hg in a fish of standardized size across these large regions. Rob, Tom, and Gretchen (I believe) have used some form of “sampling event regressions” where, for each waterbody:year:species combination, a regression is fit for Hg~length/mass. The model is then used to predict Hg. One disadvantage I saw was needing to discard waterbodies:year:species combinations that had less than a set amount of individuals (e.g., 5 or 10).

My idea was to use mixed effects models as they powerfully borrow strength from other observations when the appropriate hierarchical structures are imposed. They can also be used to estimate variability in regression relationships (i.e., random intercepts and random slopes). By also using species-level fixed effects (i.e., a different base level intercept per species), I developed a “full model” using Bayesian approximations in which waterbody:year:species combinations borrow strength from all other observations in the model. Importantly, and although it is possible, I didn’t predict Hg for fish that were missing from individual lakes. One final advantage of the INLA-Bayesian approach is that you get an estimate of uncertainty for the prediction (less data = less certainty; although I haven’t fully investigated what this looks like in low data lakes). Note that for the As project and Regression Kriging project, we run individual species models.

To summarize the two modelling approaches:

Classic: For each waterbody:year:species, run log(Hg) ~ log(length/mass); predict 1 kg fish for that waterbody:year:species

INLA Bayesian full mixed effects model: Using the full dataset, run log(Hg) ~ log(length/mass) + species + (log(length/mass) | waterbody) + (1|waterbody:year); predict 1 kg fish for each waterbody:year:species used to fit the model. Intercepts and slopes for the regression are allowed to vary by waterbody (<https://peerj.com/articles/4794/> for more info).

The table below shows the results of sampling event regressions broken up by species. Below those species, it shows the results of sampling event regressions if all aggregated together. Below that aggregation, it shows the results of an INLA model using only those species above. The final row shows the results of the full mixed effects INLA model using all species that were measured for Hg (n = 27).

Highlights:

- Under n (waterbody:year), you can see that more than ~200 lakes of data can be added if using the Bayesian approach.

- Under n (fish; median [range]), you can see that between ~10-27 individuals were used to estimate a waterbody-level prediction for the sampling even regressions but ~1-29 individuals were used for the INLA approach (in truth, it’s using all ~22000 observations considering that they borrow strength from other observations).

- Under sd intercept and sd slope, the higher sd intercept than sd slope across all species and models suggests that there is more variability captured in the intercepts than slopes (i.e., there is more variability due to absolute Hg differences than the Hg~length/mass relationships between lakes). The sd intercept and sd slope values are also very similar across species. This same pattern and roughly the same values is captured using INLA.

- Under R2, R2 of observed vs. fitted for INLA is ~0.80 whereas the R2 for the regressions is quite variable (0-0.97 in some cases but the median tends to be around 0.5-0.6).

**Finally, how well to the predictions line up? There is a very strong correlation between the two methods (see figure below separated by species, correlations are given in the figure).**

**So overall, the methods are very comparable but you get to use more data using the mixed effects model approach.**

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Species | n (waterbody:year) | n (fish; median [range]) | sd intercept | sd slope | R2 (median [range]) |
| 80 - Brook Trout | 29 | 15 [10-20] | 2.13 | 0.33 | 0.24 [0.00-0.81] |
| 81 - Lake Trout | 169 | 13 [10-24] | 2.22 | 0.30 | 0.58 [0.00-0.97] |
| 131 - Northern Pike | 274 | 10 [10-23] | 2.27 | 0.32 | 0.52 [0.00-0.94] |
| 316 - Smallmouth Bass | 129 | 11 [10-22] | 2.49 | 0.35 | 0.66 [0.00-0.97] |
| 334 - Walleye | 394 | 15 [10-27] | 2.11 | 0.29 | 0.67 [0.00-0.97] |
| Aggregating all above species | 622 | 12 [10-27] | 2.28 | 0.32 | 0.60 [0.00-0.97] |
| INLA (all above species) | 848 | 10 (1-29) | 2.66 | 0.43 | 0.80 |
| INLA (all species; n = 27) | 849 | 9 (1-29) | 2.34 | 0.38 | 0.82 |