Appendix A for:

Understanding drivers of mercury in lake trout (*Salvelinus namaycush*), a top-predator fish in southwest Alaska's parklands

Krista K. Bartz^{1*}, Michael P. Hannam¹, Tammy L. Wilson^{2,1}, Ryan F. Lepak^{3,4}, Jacob M. Ogorek⁵,

Daniel B. Young⁶, Collin A. Eagles-Smith⁷, and David P. Krabbenhoft⁵

¹National Park Service, Southwest Alaska Inventory and Monitoring Network, 240 West 5th Avenue, Anchorage, AK, 99501, USA

²U.S. Geological Survey, Massachusetts Cooperative Fish and Wildlife Research Unit, University of Massachusetts, 160 Holdsworth Way, Amherst, MA 01003, USA (current affiliation)

³Environmental Chemistry and Technology Program, University of Wisconsin-Madison, Madison, WI 53706, USA

⁴U.S. Environmental Protection Agency Office of Research and Development, Center for Computational Toxicology and Exposure, Great Lakes Toxicology and Ecology Division, 6201 Congdon Blvd, Duluth, MN 55804, USA

⁵U.S. Geological Survey, Upper Midwest Water Science Center, Mercury Research Laboratory, 1 Gifford Pinchot Dr, Madison, WI 53726, USA

⁶National Park Service, Lake Clark National Park and Preserve, 240 West 5th Avenue, Anchorage, AK, 99501, USA

⁷U.S. Geological Survey, Forest and Rangeland Ecosystem Science Center, Corvallis, OR 97330, USA

* Denotes corresponding author

Overview of Methods

In this study, we assessed concentrations of mercury (Hg) in lake trout (*Salvelinus namaycush*) from 14 lakes in two National Parks: Katmai National Park and Preserve (KATM) and Lake Clark National Park and Preserve (LACL). Our goals were (1) to evaluate the magnitude of Hg contamination in fish within these lakes, and (2) to determine the relative importance of factors pertaining to loading, methylation, bioaccumulation, and biomagnification in controlling variability of fish Hg concentrations among lakes. To accomplish these goals, we collected water, plankton, and fish in the field, and then analyzed the samples in the lab for Hg and other constituents. In addition to the samples collected in the field and analyzed in the lab, we quantified several factors from pre-existing (often remote sensing) data. Our analyses of water, plankton, and fish samples, and the summaries of pre-existing datasets, produced a large suite of factors, some of which were highly correlated. Therefore, we used principal components analyses to exclude redundant factors. This process resulted in 12 factors selected as covariates for a Bayesian hierarchical model that matched the structure of our conceptual model and data collection.

This appendix provides additional details about the methods used in our study. It also contains tables and figures referenced but not included in the main manuscript, such as Tables A.1, A.2, and A.3 and Figures A.1, A.2, and A.3.

Study Sites

Fourteen lakes were selected as study sites in the two parks. While these lakes vary in area and depth, from 1 to 311 km^2 and from 19 to 266 m, respectively, all are oligotrophic, with low concentrations of nutrients and major ions, and low to moderate acid buffering capacity (Brabets and Ourso, 2006a, 2006b; Chamberlain, 1989; LaPerriere, 1997; Wilkens, 2002). Their stratification patterns are characterized as dimictic or cold polymictic, freezing at least temporarily during winter and stratifying at least weakly during summer. Periodic wind-driven upwelling keeps hypolimnetic water well oxygenated (6-13 mg⋅L⁻¹; Table A.4).

Water Methods

We determined lake profile structure near the midpoint of each lake during the summer of 2016, using a YSI EXO2 multi-parameter sonde that simultaneously recorded water depth, pH, temperature, conductance, dissolved oxygen, and turbidity from the lake surface to the bottom or 50 m depth (whichever was shallower). Using this information, water samples were collected from depths above, within, and below the thermocline via a Teflon-coated, remote-triggered Niskin sampler and new 2-L

2

polyethylene terephthalate glycol-modified (PETG) bottles. Following collection, samples were filtered with pre-ashed quartz fiber filters (0.45 μm). Filter-passing water was poured into 500 mL and 250 mL subsamples and preserved with 1% hydrochloric acid for later total Hg (THg) and methyl-Hg (MeHg) analyses, respectively. Additional filter-passing water was collected in amber glass vials and stored at 4 °C for later measurement of dissolved organic carbon (DOC) and sulfate (SO $_4^2$).

We analyzed THg, MeHg, and DOC concentrations at the United States Geological Survey Mercury Research Laboratory (USGS-MRL) using modified versions of Methods 1631, 1630, and 9060a by the United States Environmental Protection Agency (USEPA, 1998a, 2002, 2004). Concentrations of SO₄²⁻ were analyzed at the University of Wisconsin, Madison using USEPA Method 300 (USEPA, 1997) and a Dionex 2100 ion chromatography system fitted with an AS11 column. All analyses passed established lab quality control criteria.

To lower detection limits for MeHg, we used dedicated sample bottles (only used for extremely lowlevel waters) and batch-analysis of samples (separated from higher-level MeHg ecosystems), as described by Ogorek et al. (2021). We grouped depths by lake during analysis and increased the vigor and frequency of cleaning procedures for supplies and equipment. These measures improved the MeHg method detection limit from 0.040 ng⋅L⁻¹ to 0.010 ng⋅L⁻¹ (Lepak et al., 2015). However, MeHg concentrations in filtered water from many lakes were below this threshold, so we report these values as approximations. Quality assurance metrics for analytes measured in water samples, including MeHg, are included in Table A.5.

Plankton Methods

We collected bulk plankton from the top 20 m of the water column near the midpoint of each lake using a 1 m diameter, 63 µm mesh net during the summer of 2016. Once collected, plankton were sieved using 20 cm² size-sequential Nitex screens (500, 243, 118, and 63 μ m) and frozen on the date of collection. Frozen plankton captured on the Nitex screens were transported to the USGS-MRL and lyophilized.

At the lab, we measured the mass of dried plankton captured on each screen. Samples were then removed from the screens and prepared for MeHg analysis via a 4.5 M nitric acid extraction, using methods described by Hammerschmidt and Fitzgerald (2006). Subsequently, an aliquot of the extractant was analyzed for MeHg concentration directly by ethylation with sodium tetraethylborate, purge and trap analyte capture, and cold vapor atomic fluorescence spectrometry. Analyses of plankton met USGS-

3

MRL quality control criteria. Recoveries of standard reference material (SRM) IAEA 452 were consistently within 10% of reported values, reagent blanks were negligible, and a secondary standard for verifying instrument calibration was within 10% of the expected concentration.

Plankton MeHg concentrations pertained to a single bulk sample per lake, with each individual size fraction (>500, 243-500, 118-243, and 63-118 μm) normalized relative to its mass contribution to the bulk sample as:

$$
Bulk\; MeHg = \frac{\sum_{i=1}^{4} (MeHg_i \cdot Mass_i)}{\sum_{i=1}^{4} (Mass_i)}
$$

where i = one of the four size fractions and $Mass$ = the weight of that size fraction in grams (g). See Table A.6 for plankton-related analysis results.

Calculating the bioaccumulation factor (BAF) for size-sieved plankton in a given lake required two types of data: (1) the MeHg content of the smallest size fraction (63-118 μm) of plankton (MeHg_{plankton}, ng∙g⁻¹ dw), and (2) the MeHg content of filter-passing water (MeHg_{water}, ng⋅L⁻¹):

$$
BAF = log_{10} \frac{[MeHg_{plantkton}(ng \cdot kg^{-1}dw)]}{[MeHg_{water}(ng \cdot L^{-1}dw)]},
$$

However, many of the MeHg_{water} concentrations were at or below our method detection limit (0.010 ng⋅L⁻¹), making these estimations challenging. Since benthic fluxes of filter-passing THg and MeHg were not observed in the lake profile data, we elected to average the MeHg concentrations for the epilimnion and thermocline layers, including approximations, to provide a more conservative approximation.

Fish Methods

Due to financial and logistical constraints, and a desire to minimize mortality of a long-lived species, we incorporated archived (previously frozen) lake trout samples from some lakes (Table A.7). As a result, a temporal disconnect exists between fish collections (2011 – 2016) and water and plankton collections (2016), which in turn may impact the comparability of analytes sampled in different years. Seasonal and inter-annual variability exists for some analytes but, for the purposes of this study, we assumed the values measured are best estimates for the entire study window.

After fish were caught, axial muscle tissue samples generally were removed and frozen (<-20°C) until laboratory analysis, although fish collected in 2011 and 2012 were frozen whole, and later thawed for tissue sampling. In either case, thawed tissue samples were lyophilized and homogenized and moisture content was determined, prior to analysis.

THg analysis was conducted at the USGS-MRL or USGS Contaminant Ecology Research Lab using direct combustion and atomic absorbance spectroscopy following USEPA Method 7473 (USEPA, 1998b). As with plankton, analyses of fish met established quality control criteria. Recoveries of SRM IAEA 407 were 100 ± 6 %, reagent blanks were negligible, and triplicate analysis performed once every 10 samples achieved a relative standard deviation of <10%.

The homogenized tissue samples were also analyzed for carbon and nitrogen stable isotope ratios ($\delta^{13}C$ and δ^{15} N, respectively), which serve as time-integrated indicators of a consumer's foraging ecology (Perga and Gerdeaux, 2005). Specifically, δ^{13} C can be used to estimate reliance on different carbon sources (e.g., benthic vs. pelagic), while $\delta^{15}N$ can be used to estimate trophic position (Cabana and Rasmussen, 1994; Post, 2002) and, where salmon are present, marine-derived nitrogen (Kline et al., 1990; Naiman et al., 2002). The δ notation indicates the difference between a sample and a standard (Peterson and Fry, 1987), and is defined by the following equation:

$$
\delta^{13}C \text{ or } \delta^{15}N = \left[\left(\frac{R_{sample}}{R_{standard}}\right) - 1\right] \times 1000
$$

where R is the ratio of the heavy to light isotope (13 C: 12 C or 15 N: 14 N) and standards are Vienna Peedee Belemnite for C and air for N. Analysis of δ^{13} C and δ^{15} N was conducted at the University of California Davis Stable Isotope Facility on a continuous flow isotope ratio mass spectrometer. SRMs included Bovine Liver, Nylon 5, Glutamic Acid, and Enriched Alanine, and results met lab standards (± 0.03‰). Triplicate analysis, run on 10% of samples, consistently achieved a relative standard deviation of <1%.

To account for the δ13C fractionation associated with lipid formation in fish, a mathematical lipid correction was applied to the lake trout raw δ^{13} C values and labeled as δ^{13} C_{lipid-free} (Table A.7). The correction followed Eq. 4 from Hoffman et al., 2015:

$$
\delta^{13}C_{lipid-free} = \delta^{13}C_{raw} + \frac{(-6.5*(3.8 - C:N_{raw}))}{C:N_{raw}}
$$

where -6.5 and 3.8 were Hoffman et al.'s derived values for $\Delta\delta^{13}C_{lipid}$ and C : N_{raw} , respectively. Quality assurance data for analysis of δ^{13} C and δ^{15} N in lake trout samples are included in Table A.8.

GIS Methods

Several factors quantified for this study were based on pre-existing, rather than original, data. For example, volcano proximity, glacier cover, and wetland cover were derived from remote sensing data using Geographic Information Systems (GIS).

Volcano proximity was quantified for each lake as a unitless index, similar to that developed by Gustafson and Parker (1992). The index used the distance (d) between all volcano (v) and lake (l) pairs, filtered to include only distances <100 km, then summed by lake:

$$
Volcano\ proximity = \sum_{l=1}^{14} \frac{1}{d_{vl}^2}
$$

so higher proximity index values signified more volcanoes nearby. Distances were calculated via the Near tool in ArcMap (v 10.7.1) from the point representing the center of the volcano (AKDNR, 2011) to the closest point on the lake edge (NPS, 2006). We defined the radius as 100 km to approximate the distance gaseous elemental Hg (Hg⁰) from volcanoes might travel. We focused on Hg⁰ because most naturally emitted Hg is Hg⁰ (UNEP, 2013), including that from volcanoes (Bagnato et al., 2007) and geothermal features (Hall et al., 2006). The exact travel distance will likely depend on the volcano altitude, wind strength, and plume size, in addition to the Hg form emitted (i.e., gaseous elemental, gaseous oxidized, or particle-bound).

Glacier cover was calculated for each lake basin using shapefiles of glacial area within the parks at two time intervals: current (2000's) and historical (1950's) (Arendt and Rich, 2013). Specifically, glacial area was summed across glacier types within a basin, and the sums were divided by the total basin area (USGS, 2019) to calculate percent cover for each time interval. Glacier loss was calculated as the difference between current and historical glacier cover in a lake's basin.

Wetland cover was calculated for each lake basin using shapefiles of wetland area from the National Wetlands Inventory (NWI; USFWS, 2011), where available. Like glacier area, wetland area was summed across wetland types within a basin and then divided by basin area. For basins where NWI data were not available, wetland cover was estimated from mean basin slope (NPS, 2009), via a relationship developed in 21 nearby basins where NWI and slope data existed (R^2 = 0.92, p < 0.01; Figure A.4).

GIS methods for glacier and wetland cover relied on lake basin boundaries. Existing boundaries from the USGS Watershed Boundary Dataset (WBD; USGS, 2019) were sufficient for 6 of the 14 lakes (i.e.,

6

Telaquana, Kontrashibuna, Clark, Kukaklek, Nonvianuk, and Hammersly). For the other eight lakes, the existing WBD perimeter needed to be revised, typically because the downstream-most point was well below the lake outlet. For those eight lakes, we used two alternative data sources as guides for revising the downstream perimeter of the WBD boundary. Specifically, we used an older basin boundary geodatabase where available (NPS, 2006); where the geodatabase was unavailable, we used the contour lines from a raster image of a scanned USGS topological map (1:63,360 scale; USGS, 2007).

Covariate Selection

Methods thus far produced 13 fish-level and 21 lake-level factors for potential use as covariates in the Bayesian hierarchical model (Table A.9). Given the sample size – particularly at the lake level – we needed to reduce the pool of covariates to a manageable number.

The number of fish-level covariates was reduced by omitting those with a large number of undetermined values (e.g., sex) or no clear link to a particular process (e.g., latitude, longitude). Age, length, and weight were highly correlated, so we selected the covariate mostly strongly related to fish total Hg (i.e., age). To reduce the number of lake-level covariates, we performed separate principal component analyses (PCA) for each of three processes: loading, methylation, and biomagnification. We then selected non-redundant covariates that best explained each of the two main axes in the PCA biplots. Redundancy was identified by visually identifying covariates with similar vectors in each biplot. From the groups of redundant covariates, we chose the covariate with the highest loading and/or the clearest mechanistic justification. In the loading biplot, two covariates with clear and distinct mechanistic justifications (volcano proximity and spawning density) presented similar vectors. We verified that these covariates were not highly correlated before including both. We also examined variance inflation factors in the model.

Bayesian Hierarchical Model

Observations of total Hg concentration (y_i) in lake trout ($i = 1...N$ fish) were modeled as a lognormal variable.

$$
y_i = e^{z_i}
$$

$$
z_i \sim \mathcal{N}(\hat{z}_i, \sigma_1^2)
$$

The expected value (\hat{z}_i) was a function of a random intercept ($\alpha_{1,j}$, $j=1...L$ lakes) and individual-level charactersitics, where **X** is the $i \times k$ matrix of fish-level covariate values and α is the $k \times j$ matrix of lake-varying random intercepts and regression coefficients to be estimated.

$$
\hat{z}_i = X_i \alpha_{j[i]}
$$

Lake-varying intercept and slope random effects were modeled as draws from a multivariate normal distribution centered on a lake-specific predicted intercept $(\beta_{1,j})$ and regional average slopes ($\beta_{2:5}$).

$$
\begin{pmatrix} \alpha_{1,j} \\ \alpha_{2,j} \\ \alpha_{3,j} \\ \alpha_{4,j} \\ \alpha_{5,j} \end{pmatrix} \sim \mathcal{MVN} \begin{pmatrix} \beta_{1,j} \\ \beta_2 \\ \beta_3 \\ \beta_4 \\ \beta_5 \end{pmatrix}
$$

The lake-varying intercept was, in turn, predicted by an intercept (λ_1) and two lake-level covariates: fish species richness (λ_2) and plankton MeHg concentration (λ_3), where W is a $j \times 3$ matrix containing observations of fish species richness and plankton MeHg at each lake.

$$
\beta_{1,j}=W_j\lambda
$$

We expect that any lake level-variability in loading or methylation would influence plankton MeHg and subsequently lake trout Hg, so $m-1$ different lake-level variables related to loading, methylation, and bioaccumulation were used to predict plankton MeHg (W_2) which was measured in *j* lakes. Let U be a $j \times m$ matrix of observations of m covariates at *j* lakes and an intercept dummy variable, and θ be an m -length vector of regression coefficients to be estimated.

$$
W_{2j} \sim \mathcal{N}(\widehat{W}_{2j}, \sigma_2^2)
$$

$$
\widehat{W}_{2j} = U_j \Theta
$$

Indicator Variable Selection

Lake-level covariates were selected using indicator variable selection (Hooten and Hobbs, 2015; Kuo and Mallick, 1998). Specifically, we used a random effect variant of the Kuo and Mallick model, as described by O'Hara and Sillanpää (2009). To implement this, the covariate effects (λ, θ) in the above model are the combination of the estimated covariate effect (v, η) and an indicator variable (I_2, I_2) .

$$
\lambda_l = I_{1,l} v_l
$$

\n
$$
I_{1,l} \sim \text{Bernoulli}(p_1)
$$

\n
$$
\theta_m = I_{2,m} \eta_m
$$

\n
$$
I_{2,m} \sim \text{Bernoulli}(p_2)
$$

The indicator variables were Bernoulli variables with probabilities (p_2, p_3) that were estimated as hyperparameters. The indicator variable for each covariate could take a value of 0 or 1 in each Markov chain Monte Carlo (MCMC) sample, effectively excluding or including the covariate. In this way, the mean of the indicator (I) for each covariate represents an inclusion probability for that covariate, informed by the data. To implement the random effects approach, which improves mixing, we placed a level-specific (lake-level, plankton-MeHg-level) hyperprior on the variance of each covariate effect (Cunningham et al., 2018). This constrained the covariate effect estimates to remain in a reasonable range when $I = 0$ and that covariate is decoupled from the likelihood.

$$
v_l \sim \mathcal{N}(0, \zeta_v^2)
$$

\n
$$
\zeta_v \sim Uniform(0, 20)
$$

\n
$$
\eta_m \sim \mathcal{N}\left(0, \zeta_\eta^2\right)
$$

\n
$$
\zeta_\eta \sim Uniform(0, 20)
$$

Prior probabilities for all parameters were diffuse or weakly informative (Table A.10). We fit our model with JAGS (Plummer, 2017), via the JagsUI package (Kellner and Meredith, 2021) in R Statistical Software (R Core Team, 2019). Three MCMC chains were burned in for 5000 iterations, or until the Gelman-Rubin statistic was less than 1.1, indicating convergence.

Model Adequacy

We assessed the adequacy of model fit with posterior predictive checks (Gelman and Hill, 2007) of the mean and standard deviation of lake trout THg and plankton MeHg concentrations (Table A.11), and by graphically examining realizations of posterior predictive distributions alongside observed data (Figure A.5). At each step of the MCMC, we simulated a full dataset (y_{rev} , W_{2rev}) from our model and the parameter values at that step. We examined discrepancies (i.e., how frequently summaries of the simulated data exceeded summaries of the observed data; Table A.11), and graphically examined simulated datasets compared to the observed dataset. Model fit was generally good but did not capture the extreme peak at the mode of the observed data (Figure A.5), leading to larger standard deviations in most simulated data sets compared to the observed data.

Disclaimer

Any use of trade, firm, or product names is for descriptive purposes only and does not imply endorsement by the U.S. Government. Data collected for this study are available at https://doi.org.10.5066/P9UEP9C5.

9

Literature Cited

[dataset] Alaska Department of Natural Resources (AKDNR), 2011. Alaska volcanoes. https://data-soadnr.opendata.arcgis.com/maps/56a6b26082304fdbb78d5a6e49f18894/about.

[dataset] Arendt, A., Rich, J.L., 2013. Randolf Glacier Inventory. https://irma.nps.gov/DataStore/Reference/Profile/2221653.

Bagnato, E., Aiuppa, A., Parello, F., Calabrese, S., D'Alessandro, W., Mather, T.A., McGonigle, A.J.S., Pyle, D.M., Wängberg, I., 2007. Degassing of gaseous (elemental and reactive) and particulate mercury from Mount Etna volcano (Southern Italy). Atmos. Environ. 41, 7377–7388. https://doi.org/https://doi.org/10.1016/j.atmosenv.2007.05.060.

Brabets, T.P., Ourso, R.T., 2006a. Water quality, physical habitat, and biology of the Kijik River Basin, Lake Clark National Park and Preserve, Alaska, 2004-2005. U.S. Geological Survey, p. 60. Scientific Investigations Report 2006-5123. https://doi.org/10.3133/sir20065123.

Brabets, T.P., Ourso, R.T., 2006b. Water quality of the Crescent River Basin, Lake Clark National Park and Preserve, Alaska, 2003-2004. U.S. Geological Survey, p. 48. Scientific Investigations Report 2006-5151. [https://doi.org/10.3133/sir20065151.](https://doi.org/10.3133/sir20065151)

Cabana, G., Rasmussen, J.B., 1994. Modelling food chain structure and contaminant bioaccumulation using stable nitrogen isotopes. Nature 372, 255–257. https://doi.org/10.1038/372255a0.

Cameron, C.E., Schaefer, J.R., Muliken, K.M., 2018. Historically active volcanoes of Alaska. Alaska Volcano Observatory, p.1. Miscellaneous Publication 133; version 3.0. https://avo.alaska.edu/images/dbimages/display/1546466677.jpg.

Chamberlain, D.M., 1989. Physical, and biological characteristics, and nutrients limiting primary productivity, Lake Clark, Alaska. M.S. Thesis, p. 135. Michigan Technological University.

Cunningham, C.J., Westley, P.A.H., Adkison, M.D., 2018. Signals of large scale climate drivers, hatchery enhancement, and marine factors in Yukon River Chinook salmon survival revealed with a Bayesian life history model. Glob. Chang. Biol. 24, 4399–4416. https://doi.org/10.1111/gcb.14315.

Gelman, A., Hill, J., 2007. Data Analysis Using Regression and Multilevel/Hierarchical Models. Cambridge University Press, Cambridge.

Gustafson, E.J., Parker, G.R., 1992. Relationships between landcover proportion and indices of landscape spatial pattern. Landsc. Ecol. 7, 101–110. https://doi.org/10.1007/BF02418941.

Hall, B.D., Olson, M.L., Rutter, A.P., Frontiera, R.R., Krabbenhoft, D.P., Gross, D.S., Yuen, M., Rudolph, T.M., Schauer, J.J., 2006. Atmospheric mercury speciation in Yellowstone National Park. Sci. Total Environ. 367, 354–366. https://doi.org/10.1016/j.scitotenv.2005.12.007.

Hammerschmidt, C.R., Fitzgerald, W.F., 2006. Photodecomposition of methylmercury in an arctic Alaskan lake. Environ. Sci. Technol. 40, 1212–1216. https://doi.org/10.1021/es0513234.

Hoffman, J.C., Sierszen, M.E., Cotter, A.M., 2015. Fish tissue lipid‐C:N relationships for correcting δ13C values and estimating lipid content in aquatic food-web studies. Rapid Commun. Mass Spectrom. 29, 2069–2077. https://doi.org/10.1002/rcm.7367.

Hooten, M.B., Hobbs, N.T., 2015. A guide to Bayesian model selection for ecologists. Ecol. Monogr. 85, 3–28. https://doi.org/https://doi.org/10.1890/14-0661.1

Huang, A., Wand, M.P., 2013. Simple marginally noninformative prior distributions for covariance matrices. Bayesian Anal. 8, 439–452. https://doi.org/10.1214/13-BA815.

Kellner, K., Meredith, M., 2021. Package 'jagsUI': A wrapper around "rjags" to streamline "JAGS" analyses. https://cran.r-project.org/web/packages/jagsUI/jagsUI.pdf.

Kline, T.C., Goering, J.J., Mathisen, O.A., Poe, P.H., Parker, P.L., 1990. Recycling of elements transported upstream by runs of Pacific salmon: I. δ15N and δ13C evidence in Sashin Creek, Southeastern Alaska. Can. J. Fish. Aquat. Sci. 47, 136–144. https://doi.org/10.1139/f90-014.

Kuo, L., Mallick, B., 1998. Variable selection for regression models. Sankhyā Indian J. Stat. Ser. B 60, 65– 81. [https://doi.org/10.1038/157869b0.](https://doi.org/10.1038/157869b0)

LaPerriere, J.D., 1997. Limnology of two lake systems of Katmai National Park and Preserve, Alaska: Physical and chemical profiles, major ions, and trace elements. Hydrobiologia 354, 89–99. https://doi.org/10.1023/a:1003067525783.

Lee, G.K., Yager, D.B., Mauk, J.L., Granitto, M., Denning, P., Wang, B., Werdon, M.B., 2016. The geochemical atlas of Alaska, 2016, Data Series. Reston, VA. https://doi.org/10.3133/ds908.

Lepak, R.F., Ogorek, J.M., Bartz, K.K., Janssen, S.E., Tate, M.T., Runsheng, Y., Hurley, J.P., Young, D.B., Eagles-Smith, C.A., Krabbenhoft, D.P., 2022. Using carbon, nitrogen, and mercury isotope values to distinguish mercury sources to Alaskan lake trout. Environ. Sci. Technol. Lett. 9, 312–319. https://doi.org/10.1021/acs.estlett.2c00096.

Lepak, R.F., Yin, R., Krabbenhoft, D.P., Ogorek, J.M., Dewild, J.F., Holsen, T.M., Hurley, J.P., 2015. Use of stable isotope signatures to determine mercury sources in the Great Lakes. Environ. Sci. Technol. Lett. 2, 335–341. https://doi.org/10.1021/acs.estlett.5b00277.

Naiman, R.J., Bilby, R.E., Schindler, D.E., Helfield, J.M., 2002. Pacific salmon, nutrients, and the dynamics of freshwater and riparian ecosystems. Ecosystems 5, 399–417. https://doi.org/10.1007/s10021-001- 0083-3.

[dataset] National Park Service (NPS), 2009. Slope angle in degrees of the 30-meter National Elevation Dataset (resampled using cubic convolution) for Alaska including St. Lawrence Is. https://www.usgs.gov/publications/national-elevation-dataset.

[dataset] National Park Service (NPS), 2006a. Southwest Alaska Network lake basins. https://irma.nps.gov/DataStore/Reference/Profile/2216046.

[dataset] National Park Service (NPS), 2006b. Southwest Alaska Network lake basins. https://irma.nps.gov/DataStore/Reference/Profile/2216046.

O'Hara, R.B., Sillanpää, M.J., 2009. A review of bayesian variable selection methods: What, how and which. Bayesian Anal. 4, 85–118. https://doi.org/10.1214/09-BA403.

Ogorek, J.M., Lepak, R.F., Hoffman, J.C., DeWild, J.F., Rosera, T.J., Tate, M.T., Hurley, J.P., Krabbenhoft, D.P., 2021. Enhanced susceptibility of methylmercury bioaccumulation into seston of the Laurentian Great Lakes. Environ. Sci. Technol. 55, 12714–12723. https://doi.org/10.1021/acs.est.1c02319.

Perga, M.E., Gerdeaux, D., 2005. "Are fish what they eat" all year round? Oecologia 144, 598–606. https://doi.org/10.1007/s00442-005-0069-5.

Peterson, B.J., Fry, B., 1987. Stable Isotopes in ecosystem studies. Annu. Rev. Ecol. Syst. 18, 293–320. https://doi.org/10.1146/annurev.es.18.110187.001453.

Plummer, M., 2017. JAGS Version 4.3.0 user manual. [https://people.stat.sc.edu/hansont/stat740/jags_user_manual.pdf.](https://people.stat.sc.edu/hansont/stat740/jags_user_manual.pdf)

Post, D.M., 2002. Using stable isotopes to estimate trophic position: Models, methods, and assumptions. Ecology 83, 703–718. https://doi.org/10.1890/0012-9658(2002)083[0703:USITET]2.0.CO;2.

R Core Team, 2019. R: A language and environment for statistical computing. R Foundation for Statistical Computing. Vienna, Austria. https://www.r-project.org/.

United Nations Environment Programme (UNEP), 2013. Global mercury assessment 2013: Sources, emissions, releases and environmental transport. UNEP Chemicals Branch, p.44. https://wedocs.unep.org/20.500.11822/7984 .

United States Environmental Protection Agency (USEPA), 2004. Method 9060A: Total organic carbon. U.S. Environmental Protection Agency, p. 5. https://www.epa.gov/hw-sw846/sw-846-test-method-9060a-total-organic-carbon.

United States Environmental Protection Agency (USEPA), 2002. Method 1631, Revision E: Mercury in water by oxidation, purge and trap, and cold vapor atomic fluorescence spectrometry. U.S. Environmental Protection Agency, p. 45. https://www.nemi.gov/methods/method_summary/9628/.

United States Environmental Protection Agency (USEPA), 1998a. Method 1630: Methyl mercury in water by distillation, aqueous ethylation, purge and trap, and cold vapor atomic fluorescence spectrometry. U.S. Environmental Protection Agency, p. 55. https://www.epa.gov/sites/default/files/2015- 08/documents/method_1630_1998.pdf.

United States Environmental Protection Agency (USEPA), 1998b. Method 7473: Mercury in solids and solutions by thermal decomposition, amalgamation, and atomic absorption spectrophotometry. U.S. Environmental Protection Agency, p. 17. https://www.epa.gov/sites/default/files/2015- 12/documents/7473.pdf.

United States Environmental Protection Agency (USEPA), 1997. Method 300.1: Determination of inorganic anions in drinking water by ion chromatography. U.S. Environmental Protection Agency, p. 40. https://www.epa.gov/sites/default/files/2015-06/documents/epa-300.1.pdf.

[dataset] United States Fish and Wildlife Service (USFWS), 2011. National Wetlands Inventory. http://www.fws.gov/wetlands/.

[dataset] United States Geological Survey (USGS), 2019. National Map of the Watershed Boundary Dataset. https://apps.nationalmap.gov/downloader/#/.

[dataset] United States Geological Survey (USGS), 2007. Alaska topological map: 1:63,360 scale. https://www.usgs.gov/programs/national-geospatial-program/us-topo-maps-america.

Wells, A., Macander, M., Jorgenson, T., Christopherson, T., Baird, B., Trainor, E., 2013. Ecological land survey and soil landscapes map for Lake Clark National Park and Preserve, Alaska, 2011. National Park Service, p. 376. Natural Resource Technical Report NPS/LACL/NRTR—2013/693. https://irma.nps.gov/DataStore/Reference/Profile/2193315.

Wells, A.F., Frost, G. V., Macander, M.J., Ives, S.L., McNown, R.W., 2021. Ecological land survey and soils inventory for Katmai National Park and Preserve, 2016–2017. National Park Service, p. 416. Natural Resource Report NPS/KATM/NRR—2021/2298. https://doi.org/10.36967/nrr-2287466.

Wilkens, A.X., 2002. The limnology of Lake Clark, Alaska. M.S. Thesis, p. 148. University of Alaska Fairbanks. https://irma.nps.gov/DataStore/Reference/Profile/552342.

Tables

Table A.1. Names of the 24 volcanoes that are within 100 km of at least one of the 14 lakes in this study. General locations and periods of activity are also listed, based primarily on AVO 2016.

^A LACL = Lake Clark National Park and Preserve; KATM = Katmai National Park and Preserve.

^B Volcanoes defined as "historically active" by Cameron et al. (2018) meet at least one of the following criteria since 1700 CE: (1) a documented unquestioned eruption, (2) a strongly suspected eruption, often documented in a historical account with very little information, but never contradicted by current geologic knowledge, (3) persistent (usually on the order of decades but certainly longer than several months) fumaroles, with temperatures (where measured) within ~10 °C of the boiling point, (4) significant, measured, volcanic-related, non-eruptive deformation, or (5) documented earthquake swarm with strongly suspected volcanic cause.

 \textdegree Active within the last 2,000,000 years, but not within the last 10,000 years.

D Active in the Holocene.

Table A.2. Characteristics of the 14 study lakes and their surrounding basins, listed along a north-to-south gradient.

^A LACL = Lake Clark National Park and Preserve; KATM = Katmai National Park and Preserve.

^B Latitude and longitude specify the lake centroid location in decimal degrees.

^C Basin elevation represents a mean value, averaged within the basin boundary.

Table A.3. Analytes measured in filter-passing water collected from various depths at 14 study lakes in 2016. Analytes include total mercury (THg), methylmercury (MeHg), dissolved organic carbon (DOC), and dissolved sulfate (SO₄²⁻). MeHg values are generally below the method detection limit (0.010 ng⋅L⁻¹) and thus should be considered as approximates.

^A LACL = Lake Clark National Park and Preserve; KATM = Katmai National Park and Preserve.

^B Field duplicates were collected at this depth so results in this row are mean values. See Table A.5 for details. \textdegree No data, due to sample loss (Kukaklek) or lack of sample bottles (Clark).

Table A.4. Ranges in dissolved oxygen (DO) concentrations (mg⋅L⁻¹) and saturation (%) along vertical profiles at 14 lakes in the year 2016. Profile dates and maximum depths are also provided. Profile data for additional years and parameters are available at https://irma.nps.gov/aqwebp/.

^A LACL = Lake Clark National Park and Preserve; KATM = Katmai National Park and Preserve.

^B DO concentrations >10 mg∙L⁻¹ at depths of 190 m have been recorded in Lake Clark (Wilkens, 2002).

Park ^A	Lake name	Metric ^B	Depth (m)	THg $(ng \cdot L^{-1})$	MeHg $(ng·L-1)$	DOC $(mg·L-1)$	SO ₄ ² $(mg·L-1)$
LACL	Kijik	Mean	12	0.171	0.012	0.558	9.741
LACL	Kijik	Difference (abs.)	12	0.012	0.010	0.002	0.059
LACL	Clark	Mean	10	0.206	0.003	C	C
LACL	Clark	Difference (abs.)	10	0.032	0.006	C	C
KATM	Hammersly	Mean	7.5	0.239	0.004	0.618	2.913
KATM	Hammersly	Difference (abs.)	7.5	0.033	0.003	0.004	0.965
D	D	Mean blank	D	0.063	0.004	0.161	0.015
D	D	Detection limit ^E	D	0.040	0.010	0.200	0.017

Table A.5. Quality assurance metrics for field duplicates of filter-passing water collected at a subset of lakes and depths.

^A LACL = Lake Clark National Park and Preserve; KATM = Katmai National Park and Preserve.

B Metrics included mean, absolute difference, and percent difference of two duplicate samples.

 \textdegree Field duplicates were not collected for these analytes.

^D Does not apply.

 E Method detection limits (MDL) are reported except for SO₄², which had no established MDL and was therefore calculated as 3 x standard deviation of the mean blank value.

Table A.6. Total mercury (THg) and methylmercury (MeHg) concentrations measured as ng∙g⁻¹ dry weight in size-sieved plankton. The percentage of THg that is MeHg is reported for each size fraction. Bulk MeHg concentrations and bioaccumulation factors (BAFs) are calculated once per lake.

^A LACL = Lake Clark National Park and Preserve; KATM = Katmai National Park and Preserve.

^B Bulk MeHg concentrations pertain to a single bulk sample per lake, including all size fractions (63 - >500 μm).

 c Like bulk MeHg concentrations, BAF was calculated once per lake, but only using the 63 - 118 μm size fraction.

 P The amount collected for this size fraction was too small to enable Hg analysis, so a value was estimated from a linear regression using data from the five other KATM lakes: $y = 0.0762x + 5.364$ (R² = 0.723, F = 46.970, df = 1, p < 0.001), where y = MeHg content in ng g^{-1} and x = the lower value in the size fraction range (i.e., 63 μm).

Table A.7. Field and lab measurements for 158 lake trout sampled from 14 lakes from 2011 to 2016. Field measurements include the coordinates where fish were sampled in decimal degrees, as well as their total length, weight, age, and sex. Lab measurements include total mercury (HgT ng g⁻¹ dry weight), percent moisture, nitrogen and carbon stable isotope ratios (δ¹⁵N, δ¹³C, and δ¹³C_{lipid-free}), C:N ratio, and percent lipid.

^A LACL = Lake Clark National Park and Preserve; KATM = Katmai National Park and Preserve.

 $^{\text{\tiny{\textregistered}}}$ For Snipe Lake, the sample date spanned 2 days (2015-06-15 – 2015-06-16); for Turquoise Lake, it spanned 3 days (2015-09-06 – 2015-09-08).

^C Coordinates for 4 lake trout at Grosvenor Lake and 6 lake trout at Crescent Lake are averages representing multiple locations.

^D Weight (W) was estimated from length (L) for 5 Lachbuna lake trout using a relationship based on 116 other fish from 12 lakes: W = 2E-05L² - 0.0129L + 2.6311 (R² = 0.94).

 E Age (A) was estimated from weight (W) for 2 Kontrashibuna lake trout using a relationship based on 8 other fish in the same lake: A = 20.845W - 1.8304 (R² = 0.75).

 $F M$ = male; F = female; U = unknown.

^G $\delta^{13}C_{l-f} = \delta^{13}C_{lipid-free}$.

^H% lipid (L) was estimated from the molar C:N ratio (R) as: L = 93/(1 + ((0.246*R - 0.775)⁻¹) following Hoffman et al. (2015).

Table A.8 Quality assurance data from triplicate analysis of carbon and nitrogen stable isotope ratios in a subset of lake trout samples. Data include mean values for each triplicate, followed by twice the standard deviation (SD), representing an estimate of relative precision.

^A LACL = Lake Clark National Park and Preserve; KATM = Katmai National Park and Preserve.

^B Sample IDs match those in Table A.7.

Table A.9. List of fish- and lake-level factors considered for inclusion as covariates in the Bayesian hierarchical model, and an accounting of which factors were ultimately included.

Table A.10. Prior probability distributions and their characteristics for parameters in the Bayesian hierarchical model.

^A From Huang and Wand (2013).

Table A.11. Posterior predictive tests of the mean and standard deviation (SD) of lake trout THg (y) and plankton MeHg (W_2), based on observed and simulated replicate (rep) datasets. A perfect model fit would yield probability values of 0.5.

Figure A.1. Total mercury content in ash (n = 21) and non-ash (n = 11) layers of surface soil samples located in or near (within 5 km of) the basin boundaries of our study lakes. Samples were collected for park soil inventories in 2011, 2016, and 2017 (Wells et al., 2013, 2021; see Table S8 in Lepak et al., 2022). Distributions of total mercury do not differ among ash and non-ash soil layers (*W* = 98 and *p* = 0.50, based on a non-parametric Wilcoxon signed-rank test).

Figure A.2. Elevation vs. total mercury content in 73 rock samples (top; *p* = 0.12) and 450 sediment samples (bottom; $p = 0.67$) collected and analyzed previously by the USGS (Lee et al., 2016). Here, the USGS data were trimmed to include only samples collected within study basins that had glacier cover. These data support the idea that mercury concentrations of recently exposed bedrock and sediment near glacier margins upslope are likely comparable to those of surrounding areas downslope.

Figure A.3. Volcano proximity index vs. water total mercury content in our 14 study lakes (*p* = 0.20). The volcano proximity index is unitless; here it is multiplied by 1000 for ease of viewing.

Figure A.4. Relationship between mean slope (°) and total wetland cover (%) in 21 lake basins (*p* <0.01), all but two of which were located within the boundaries of Lake Clark National Park and Preserve. The other two basins (Delight and Desire lakes) were located in another park in southwest Alaska: Kenai Fjords National Park.

Figure A.5. Histogram of observed total mercury (THg) concentrations in lake trout (black bars), overlaid with density plots of 120 model-simulated THg datasets (blue curves). Note that the histogram is scaled to density rather than frequency to allow for comparison with simulated data.