eDNA provides accurate and precise estimates of abundance by integrating bioenergetics and particle mass-balance

Beaulieu J[1,2], MC Yates[1,2], DJ Fraser[2,3], ME Cristescu[2,4] and AM Derry[1,2]

[1] Département des sciences biologiques, Université du Québec à Montréal, Montréal, QC, Canada

[2] Groupe interuniversitaire en limnologie (GRIL), Montréal, QC, Canada

[3] Department of Biology, Concordia University, Montreal, QC, Canada

[4] Department of Biology, McGill University, Montréal, QC, Canada

ORCID and emails:

Julien Beaulieu: **0000-0003-2597-8921** beaulieu.julien.3@courrier.uqam.ca
Matthew C. Yates**: 0000-0002-9199-1078** matthew.yates@outlook.com
Dylan J. Fraser**: 0000-0002-5686-7338** dylan.fraser@concordia.ca
Melania E. Cristescu: **0000-0002-0854-4040** melania.cristescu@mcgill.ca
Alison M. Derry: **0000-0001-5768-8027** derry.alison@uqam.ca

Data availability statement

Data and codes will be available upon manuscript acceptance

Funding statement

This research was funded by a Fonds de Recherche du Québec—Nature et Technologie (FRQNT) team grant to AMD, MEC, and DJF, a grant from the Natural Sciences and Engineering Research Council of Canada (NSERC Strategic Project grant, AMD, DJF), and a NSERC discovery grant to AMD.

Ethics approval statement

All field work and fish handling protocols were approved in permits issued by Parks Canada, Banff-Yoho and Kootenay National Parks (KooNP- 2017- 24,999), and the Concordia University Animal Ethics Committee.

Conflict of interest disclosure

There are no competing interests to declare

Abstract

Anthropogenic activities have led to an unprecedented crisis in freshwater biodiversity loss. The capacity to monitor the abundance of wild populations is critical to conserving biodiversity, but conventional physical specimen collection methods are invasive, costly, and labour-intensive. Environmental DNA (eDNA) offers a promising alternative, being easy to sample, with studies under controlled laboratory conditions showing consistent correlations between eDNA concentration and abundance. However, applying eDNA to monitor abundance remains contentious, as eDNA particle dynamics and the ecology of eDNA production can decouple this relationship in natural ecosystems. To address this, we integrated bioenergetics and mass-balance frameworks to relate eDNA concentrations to freshwater fish population *abundance* estimated through conventional mark-recapture in Brook Trout (*Salvelinus fontinalis*) across nine Rocky Mountains lakes, five of which underwent size-selectively harvest over two years. Our integrated framework improved the variance explained in eDNA concentrations from 24% to 71%. The integrated model accurately distinguished most abundance estimates across populations and sampling periods, detecting both natural and harvest-induced reductions in abundance within several populations. This study is the first to empirically integrate the DNA production ecology and particle dynamics, demonstrating that properly accounting for 'the ecology of eDNA' enables rapid, accurate abundance quantification. We also discuss future research and applications of eDNA in light of these results.

Significance statement

Addressing the unprecedented freshwater biodiversity crisis requires cost- and time-effective methods for estimating abundance and biomass. Our study empirically demonstrates the potential of eDNA for rapid and accurate assessment of abundance in natural populations by integrating hydrology, chemistry and bioenergetics into a mass-balance modelling framework. Through a replicated whole-ecosystem experiment, we show that eDNA not only detects differences in fish abundance and biomass across populations, but also captures natural and experimentally induced abundance changes within populations. Our modeling approach enables eDNA to generate abundance and biomass estimates comparable in precision and accuracy to high-quality mark-recapture methods.

Key words: Environmental DNA (eDNA), Quantitative PCR (qPCR), abundance, density, biomass, Allometric scaling, Mass balance

Introduction

Anthropogenic activities have triggered an unparalleled biodiversity crisis in freshwater ecosystems (1). Conserving remaining biodiversity, remediating degraded ecosystems, and halting the decline of endangered populations demand rigorous monitoring and proactive aquatic ecosystem management (2). However, freshwater ecosystems are notoriously difficult to monitor, with the decline in freshwater fish populations, often referred to as an 'invisible collapse' (3, 4). Monitoring the abundance and biomass of species in aquatic ecosystems can be particularly challenging, due to its labour- and resource-intensive nature. These challenges are magnified at larger geographic scale, were obtaining such data for remote or large freshwater waterbodies can be exceedingly difficult if not impossible. As a result, the availability of such critical data for conservation and management is often severely limited.

The analysis of environmental DNA (eDNA), which refers to DNA particles found in environmental mediums like air, soil, and water (5), has revolutionised biomonitoring in natural ecosystems. The analysis of eDNA has proven highly effective for quantifying ecosystem-level biodiversity as well as detecting cryptic, rare, and traditionally difficult-to-detect species in freshwater ecosystems (6, 7). However, numerous studies have demonstrated consistent and significant positive correlations with the concentration of eDNA in an aquatic environment and the abundance or biomass of organisms within that environment, demonstrating its strong potential for additionally providing estimates of population abundance and/or biomass in natural ecosystems (8, 9). Yet, the biological processes involved in the production of eDNA production are complex and, as a physical particle, numerous environmental parameters can impact its degradation and transportation in natural ecosystems (10, 11). Collectively, these processes can decouple the relationships between the concentration of eDNA in natural ecosystems and organism

abundance/biomass (12). This can be readily observed in the discrepancy between the strong correlation eDNA exhibits with abundance under controlled laboratory conditions relative to the moderate correlations observed in studies conducted within natural ecosystems (9).

Our understanding of the behaviour and dynamics of eDNA particles in natural ecosystems is constantly growing and improving through rigorous research occurring in labs across the globe. The decay rate of eDNA, for example, can depend on complex interactions among abiotic conditions (e.g., temperature, pH and UV), biotic conditions (e.g., bacterial activity), and eDNA properties (amplicon and particle size) (13–18). In both lentic and lotic systems, eDNA particles move within water masses at different rates, which can additionally affect their deposition. The concentration measured at one place can thus be influenced by eDNA production upstream and/or its movement within currents or within thermocline layers (19–22).

Metabolic and bioenergetic theory has also recently been integrated as a framework to relate eDNA production to population density and biomass by accounting for the physiological processes involved in its production (23). Metabolic processes, for example, scale allometrically with organism size (24, 25). The strong link between eDNA production and organism metabolism implies that eDNA production likely scales allometrically within and across species (23, 26). Similarly, temperature has long been recognized as having a strong impact upon physiological rates within poikilothermic fishes (27), which similarly extends to eDNA production (16, 28). As a result, eDNA production can be modelled as a function of allometrically scaled organismal mass weighted with a temperature-dependent function to adjust for the impact of the metabolic rate of the organism on eDNA production (12, 26, 29, 30), e.g.:

$$eDNA\ production = I_0 \sum_{i=1}^{N} (M_i^b) \times f(T) \quad (1)$$

Where I_0 is the eDNA production coefficient, N is the number of individuals in the population, M_i is the mass of the ith individual, b is the allometric scaling coefficient, and $f(T)$ is the temperature-dependent function, which considers the maximal metabolic rate of the organism at a specific temperature (12).

A substantial body of work describes the complex dynamics of eDNA production, degradation, and transport, all of which interact to influence eDNA concentrations in natural ecosystems. Integrating across these processes provides an understanding of particle dynamics that affect particle input and output rates; as a result, the application of mass-balance approaches are ideally suited to modelling eDNA steady-state concentrations, in a similar manner as classic models used for estimating phosphorus concentration (16, 31, 32). If we assume the concentration of eDNA to be constant in time, we can state:

$$eDNA_{produced} + eDNA_{imported} = eDNA_{degraded} + eDNA_{sedimented} + eDNA_{exported} \quad (2)$$

Where the eDNA production rate is given by the metabolic/bioenergetics function described above (equation 1), and eDNA imported describes the rate eDNA enters the system from up-stream sources. The rate at which eDNA is degraded is typically described by an exponential decay curve as a function of temperature, pH, and other factors that can influence eDNA degradation rate, with 'k' being the degradation constant (e.g. Kagzi et al., 2022). No research has yet examined eDNA sedimentation rates independent of particle decay, but sedimentation of particles in the water are also typically modeled using exponential decay curves (31). The rate at which eDNA is exported (i.e., transported) refers to the amount of eDNA exiting by the outflow of a water body, which can be calculated as the outflow times the concentration at the surface (31). By developing equation 2 and isolating eDNA concentration, the formula can be expressed as:

$$[eDNA]_{water} = \frac{W}{Q + (k_{deg} + k_{sed})V} \quad (3)$$

Where $[eDNA]_{water}$ is the eDNA concentration in the water, W is the average eDNA production (equation 1) plus the average eDNA importation from upstream, V is the volume of the water body, Q is the water flow rate, K_{deg} is the degradation constant at these specific environmental conditions, and K_{sed} is the sedimentation rate.

Applying a modeling technique that integrates both bioenergetics and mass balance approaches to eDNA concentrations to estimate total population abundance and biomass in natural ecosystems remains unexplored but holds considerable potential for improving the capacity to link eDNA concentrations observed in natural ecosystems to wild population abundance. With refined quantitative knowledge on eDNA production, decay, and transport rates, eDNA could become a cost-effective tool for species quantification in aquatic systems (12, 34). To date, however, no study has integrated across all these processes to generate total population abundance/biomass estimates in natural ecosystems.

We investigated the application of an integrated bioenergetics and mass-balance approach to estimate population abundance from eDNA concentrations observed in natural lentic ecosystems, aiming to: i) differentiate total abundance *across* populations; and ii) detect changes in abundance *within* populations. To accomplish these objectives, we quantified Brook Trout (*Salvelinus fontinalis*) eDNA concentrations and total abundance (using mark-recapture techniques) in nine Rocky Mountain lakes (CA) across two seasons for two years, while simultaneously conducting an annual experimental size-selective harvest in five of the nine lakes to manipulate total abundance and population density across the study period. We also investigated the additional power of an integrated bioenergetics and mass-balance approach in explaining

observed eDNA concentrations relative to approaches previously taken in the literature, including i) correlating eDNA concentrations as a linear function of population abundance; and ii) correlating eDNA concentrations as a function of allometrically scaled population biomass. We predicted that taking an integrated bioenergetic and mass-balance approach that accounts for the processes involved in eDNA production, degradation, and transportation would increase the power of eDNA to detect changes in the abundance and biomass of Brook Trout across space and time.

Methods

Study system

Our study included nine headwater mountain lakes in the Canadian Rockies, located within Banff, Yoho, and Kootenay National Parks (Canada; Fig. 1). The nine study lakes contain fish communities largely dominated by Brook Trout (*Salvelinus fontinalis*) that were introduced to the lakes between the 1920s and the 1970s and are now invasive in the region (35). Six of the lakes were solely inhabited by Brook Trout, while the three other had fish communities also containing one to five other species. Five of the nine study lakes were subjected to a size-selective harvest treatment where approximatively 64 % of each population was removed in the Fall of each year from 2017 to 2019, targeting the largest-bodied individuals for harvest. At the beginning of the experiment (2017), the Brook Trout populations differed in density and population size structure between lakes, and the harvesting treatment was successful at inducing changes in population demography over the duration of the three-year experimental manipulation (36). The other four study lakes were monitored as references in which Brook Trout populations were not directly manipulated (Appendix 1). All lakes are small, cold, oligotrophic headwater lakes, but they span a

1000 m elevation gradient over which there are differences between lakes in temperature, pH, and nutrient concentrations (Appendix 1; Trépanier-Leroux et al., 2023).

Field sampling of eDNA

Environmental DNA samples were collected each year between 2018 and 2019 for two time periods: i) in the summer before harvesting (June/July); and ii) in the summer/fall following harvest activities. A systematic stratified survey design was used to collect eDNA samples, in which 4 samples each were collected from the pelagic and littoral zones of each study lake. In the littoral zones, eDNA sample sites were distributed equidistantly around the perimeter of the lake. Littoral samples were collected approximatively 1-3 m from the shore at a minimum depth of 30 cm from the surface, but approximately 15 cm above the lake bottom. In the pelagic zones, four samples were collected at equidistant intervals across a transect that followed the longest axis of the lake and went through the lake center (defined as the midpoint of the longest axis). Pelagic samples were collected at a depth of 0.5 to 1 m from the surface. Water samples were collected with 700 ml Whirl-Pak™ bags from a canoe assigned to sample a single specific lake, or from an inflatable kayak, which was decontaminated 48 hours prior with a 2% regular strength household bleach solution for 15 min (including paddle and lifejacket) and left to dry in the sun.

Water samples were immediately filtered on the lakeshore on a water filtering manifold (Wildco, Florida, USA) that was decontaminated with a 20-30% bleach solution 2-12 h prior to sampling. Manifold components were stored individually in plastic zippered bags after bleaching and during transportation. All samples were stored in the shade prior to sampling, and all filtering was conducted in the shade under a tarp. One litre of water from each of the eight water sampling sites per lake was filtered through a 0.7 μm glass fibre filter (GE Healthcare Life Sciences, Ontario, Canada) using a vacuum hand pump (Soil Moisture, California, USA). Vacuum pumps were

decontaminated between lakes by wiping with a 30% bleach solution and resting for 10 minutes, after which they were wiped with distilled water to remove residual bleach. All littoral samples were filtered on one manifold, and all pelagic samples were filtered on another to prevent cross-contamination between samples from different lake zones. As a negative control, 500 mL of distilled water was filtered prior to filtering lake samples. After filtration, filters were folded using metal forceps (decontaminated using a 30% bleach solution for 15 minutes) and stored in a sterile 2 mL microcentrifuge tube containing 700 µL of ATL buffer that was then individually placed in a plastic zippered bag and stored in a cooler bag that was decontaminated by wiping with a 30% bleach solution. If a filter became clogged (i.e. < 1 L of water was filtered), the final volume of water filtered was recorded and the sample was then stored in buffer. Samples were then transported in a cooler bag with bleached frozen gel packs to Kootenay Crossing (BC, Canada), where they were placed in a second zippered plastic bag and stored in a -20 ℃ freezer. At the end of the sampling season, all samples were transported to Montreal on dry ice and placed in a -80 ℃ freezer until extraction. All protocols for contamination prevention in the field, as well as during transport and storage can be found in Appendix 2.

Laboratory analysis of eDNA

DNA was extracted from each filter was extracted using Qiagen DNeasy Blood and Tissue™ kit and Qiashredder™ columns following a modified extraction protocol (see Yates et al. (2021a) for more details). Extracted DNA was eluted into 130 µL of AE buffer and stored in a clean -20 ℃ freezer solely dedicated to the storage of extracted eDNA product (i.e. no post-PCR products or tissue samples). An extraction blank of 700 µL of ATL buffer was included in all batches as a control. Methods for contamination prevention are detailed in Supplementary Information (Appendix 2).

Brook Trout eDNA concentration was quantified using quantitative PCR (qPCR) with the *cyt b* TaqMan minor groove assay (BRK2) described in Wilcox et al. (2013). For details on PCR reaction preparation, component concentrations, standard curve preparation, and cycling conditions, see Yates et al., (2021a). All samples were run in triplicate in a 20 μL reaction volume with 5 μL of template DNA on a Stratagene MX 3000P thermal cycler. Reactions were spiked with an internal positive control (IPC) to test for inhibition, which was defined as a 1 > Ct shift in the IPC relative to a reaction containing 5 μL of nuclease-free ultrapure water. If inhibition was detected, samples were diluted down to 3 μL and re-run. Brook Trout eDNA concentrations at each site were calculated by averaging replicate values, and final copy number data was converted to eDNA concentration per L of water filtered (copies/L). Standard curve R^2 values for the qPCR plates ranged from 0.98-0.99, and reaction efficiency varied between 83.1 and 98.6%. The efficiency of two of the 11 plates was suboptimal (< 85%). However, data generated from these two plates were retained, as their inclusion did not impact results (Appendix 3). Samples associated with negative controls that contained more than 5 copies/rx (on average) were removed from the analysis due to contamination concerns (Cobb Lake samples from fall 2019). The majority of samples collected across lakes per sampling period (26/36 sampling events) showed no contamination, while the remaining contamination in field controls was low enough to not impact the estimated lake eDNA concentrations.

Brook Trout mark-recapture and population size structure estimation

Mark-recapture studies of Brook Trout were conducted in all lakes during May and June in both 2018 and 2019. The exception was Cobb Lake where, in addition to mark-recapture sampling in May/June, a series of isolated marking events continued until September for both years. Fish were captured using a combination of fyke nets, angling, and backpack electrofishing. Large (1 m

hoop diameter, 2 cm mesh) and small (0.7 m hoop diameter and 0.8 cm mesh) fyke nets were distributed around the perimeter of lakes with the lead attached to shore and the end of the trap facing the center of the lake. Nets were checked daily. Angling was used to supplement Brook Trout capture efforts at sites where fyke catchability was low (predominantly Cobb Lake). Marks were also assigned to Brook Trout captured by electrofishing the shore and inlets/outlets of lakes.

Captured Brook Trout were anesthetized using clove oil and measured for fork length (± 1 mm) and mass (± 0.1 g). Any unmarked Brook Trout were gastrically tagged with a BioMark HPT8 pre-loaded Passive Integrated Transponder (PIT) tag (Boise, Idaho, USA). Only Brook Trout greater than or equal to 80 mm were tagged to reduce tagging mortality. Other species, when present, were not tagged. The tag number of any recaptured Brook Trout was recorded. Recovered fish were released in the center of the lake to standardize release location and promote mixing. Marking ceased once recapture ratios approached twenty five percent for several consecutive days to standardize marking efforts across all populations and to ensure that enough Brook Trout were tagged to facilitate estimation of census size (Nc) with credible intervals within 10% to 25% of true values, following general methodologies reviewed in Krebs (2009).

To obtain representative snapshots of the size structure of each Brook Trout population before and following harvest, population size structure estimates were conducted in August 2018 and 2019. The exception was Cobb Lake where size structure assessments continued to October 12th. During mark-recapture, fyke-nets were deployed in littoral zone areas extending to the centre of the lake and, as a result, size-structure assessments obtained from fyke nets may be more biased towards small-medium bodied individuals that tend to prefer littoral habitats (41). To obtain a relatively unbiased estimate of population size structure, Brook Trout were captured in large and small sinking mixed mesh gillnets with clear monofilament distributed throughout the lake. Large

mixed-mesh gillnets were 15.6 m long, 1.8 m deep and had an equal area of 64-51-89-38-76 mm mesh panels. Small mixed-mesh gillnets were 12.5 meters long, 1.8 meters deep, and consisted of an equal area of 32-19-38-13-25 mm mesh panels. Such index nets are widely used in North America for size structure assessments (42–46), as these tend to capture a representative size/age structure of populations (47). Nets were checked daily and moved to different locations across the lake, to capture a representative sample of Brook Trout in each lake. Sampling ceased when approximately five to ten percent of the population was captured, apart from Cobb Lake where the harvested fish were also included for size structure assessment such that approximately 71% of individuals were used. Captured Brook Trout were euthanized with clove oil, PIT tags were recorded, and length/mass data were collected as described for the marking period.

Schnabel population size estimates were used to estimate the abundance of Brook Trout in a lake (48). All size structure assessment removals were pooled together into one final sampling event for the population estimates which controlled for the removal of marks at large. Note that population estimates only account for Brook Trout greater than the minimum tagging size (80 mm fork length), which correspond to all individuals age 1 and older (36). All population estimates were conducted in R (49) with the mrClosed function from the Fisheries Stock Assessment package FSA (50). Credible intervals for Schnabel population estimates followed recommendations from Seber (2002) as implemented in the FSA package.

Size-selective harvesting of Brook Trout

Populations at Cobb, Mud, Olive, Ross, and Temple lakes were harvested annually with size- selective gillnets (25– 89 mm mesh) targeting the top 75% (but effectively harvesting the top 64% on average) of individual body sizes according to the size distributions estimated during stock assessment (i.e. depleting the largest size classes sequentially until the only remaining individuals

were in the lowest 25% quartile of the original size structure). Details can be found in Matte et al. (2023).

Lake temperature and its effect on the physiology of eDNA production

To adjust for maximal metabolic rate of fish at a given temperature, we used a temperature-dependent function specific to Brook Trout published from previous empirical studies (29, 52). Temperature data were constituted of both high frequency (15 min) surface temperature from data loggers (Onset Hobo®; UA-002-64, Onset Computer Corporation, Cape Cod, Massachusetts) and vertical profiles from a YSI multiparameter sonde (model 10,102,030; YSI Inc.). The dataloggers were deployed at 0.5 m from the surface in June and retrieved in September. The YSI vertical profiles were taken at the deepest point of the lake at the same time as eDNA sampling, except six lakes at Fall 2019 where a YSI malfunction necessitated us to use data from the same time period in 2018 as a proxy. There was approximately a +/- 2°C difference between fall 2018 and 2019 for the lakes that had YSI data for both seasons, and as a result this likely had very little impact on model estimates. When possible, mean surface temperature over the interval from two weeks before to two weeks after each eDNA sampling event per lake was used as input in the temperature-dependent function in the eDNA mass balance model. However, due to logistic constraints and material failures, some lakes and sampling periods had temperature data for only a subset of the time intervals or only the epilimnion mean temperature from the YSI vertical profiles. A table with the lakes, dates, and temperature data can be found in the supplementary materials (Appendix 4).

Lake hydrology

To calculate the rate of lake outflow, lake watershed area was estimated using the flow direction and flow accumulation hydrologic models from ArcGIS pro V.3.2. The input data were 15 m resolution Digital Elevation Models (DEM) for 7 out of the 9 study lakes. However, the 15

m resolution was too low to effectively determine the watershed for two of the lakes (Mud and McNair). Fifteen m resolution DEM are publicly available form the Government of Canada and under Open Government Licence (https://open.canada.ca/en/open-government-licence-canada). The Mud and McNair lakes required high resolution data, but none were available for their area. The Mud Lake watershed area was therefore estimated by process of elimination by estimating the watershed area of the surrounding larger water courses. The remaining area that was not part of any other watershed was assumed to be Mud Lake's watershed. We could not obtain watershed area for McNair Lake. The lake outflow rate was estimated by multiplying each lake watershed area with the flux per area of watershed from hydrological stations from the same area (Kootenay: Kootenay River: 50.886944, -116.045833; Banff: Pipestone River: 51.433056, -116.174722; Environment and Natural Resources Canada, 2024). The rate of outflow for each lake at each sampling period was calculated as the mean outflow from two weeks before the sampling date to two weeks after.

Lake volume and epilimnion volume

Detailed bathymetric maps and lake volume produced by high frequency sonar were available for only four of our lakes (Temple, Olive, Helen, and Margaret lakes). For the remaining five study lakes, we used older bathymetric maps from Parks Canada that were georeferenced and digitised in arcGIS pro®. The depth between isobaths were interpolated using Bayesian Kriging method. The epilimnion volume was taken as the input in the mass balance model (equation 5 and 8). Epilimnetic volume was chosen because eDNA samples were collected from the epilimnion of each lake; this was done because most study lakes undergo thermal stratification during summer, and because Brook Trout spend most of their time feeding and resting in surface waters (41). Additionally, research has demonstrated that the transfer of eDNA across the thermocline is limited

during lake stratification (54). The epilimnion volume was calculated as the mean thermocline depth times the lake surface area from the bathymetric maps. The mean thermocline depth was calculated from two YSI® vertical temperature profiles per summer. For the study lakes that were not thermally stratified (assessed from the YSI vertical temperature profile) at the time of eDNA sampling period, total lake volume (estimated from the bathymetric maps) was used as input in mass-balance models for that time point because eDNA signal is expected to become homogenous throughout the water column during mixing (54).Unfortunately, the extent of thermal stratification during the fall of 2019 could not be quantified due to the malfunction of the YSI. Lakes that were mixed in the fall of 2018 were therefore also assumed to be mixing in fall 2019 except for Margaret Lake, where stratification was very weak in fall 2018 and it was not fully mixed. As fall 2019 was colder than fall 2018, we assumed that Margaret Lake was mixing in fall 2019; this increased model fit for this lake.

Statistical analyses

Constructing and evaluating an integrated bioenergetics/mass-balance model

To obtain a representative mean eDNA concentration for each lake and sampling period, the littoral and pelagic samples were weighted according to the littoral and pelagic zone area in each lake. The littoral zone was defined as the area where the depth is bellow or equal to 2 m. To parameterise the mass balance model (quation 3), a non-linear Bayesian model with informative priors was used. The parameters estimated were I_0 (eDNA generation coefficient), b (allometric scaling coefficient), and k (a combined degradation and sedimentation constant). Although equation 3 has both the degradation constant k_{deg} and the sedimentation constant k_{sed}, the current state of knowledge on eDNA degradation and sedimentation does not allow for a precise estimate

of the values for k_{deg} and k_{sed}, and so one "disappearance constant" k was parameterised in the model as the sum of both (equation 4).

$$[eDNA]_{Lake} \sim \frac{I_0 \sum_{i=1}^{N}(M_i^b) \times f(T)}{Q + kV} \tag{4}$$

The model was fitted with a gaussian distribution, four chains, 10 000 iterations, and *adapt_delta = 0.99*. A lake specific k parameter was estimated to account for potential differences in degradation rates between lakes, and a sampling period specific eDNA generation coefficient (I_0) to account for season and sampling period effect on mean Brook Trout eDNA concentrations observed across lakes (Gaudet-Boulay et al., 2023), either due to ecological differences in the eDNA production coefficient (I_0) across seasons or potential (unintentional) methodological differences in sampling. For all analysis, $\sum_{i=1}^{N}(M_i^b)$ was approximated using the Taylor expansion for the first moment (55) such that equation 4 becomes:

$$[eDNA]_{Lake} \sim \frac{I_0 f(T) \times N(\bar{M} + \frac{\sigma_M^2}{2} b(b-1)\bar{M}^{b-2})}{Q + kV} \tag{5}$$

Where $\bar{M}$ is the mean mass of the fish and σ_M^2 is the variance of the mass of the fish. The priors used were:

$$\begin{gathered}
\ln(I_0) \sim normal(10, 2) \\
b \sim beta(2.8, 1.2) \\
\ln(k) \sim normal(-0.7, 0.1) \\
Standard\ deviation_{K\ lake} \sim exponential(3) \\
\sigma \sim \theta + N^{\alpha} \\
\alpha \sim normal(0{,}0.5),\ lb = 0 \\
\theta \sim normal(0{,}1),\ lb = 0
\end{gathered}$$

The prior on I_0 ensured posterior predictions were the correct order of magnitude, the prior on b ensured values were between 0 and 1 with a higher probability density around 0.7, which is

approximately the *b* value estimated from Brook Trout bioenergetic models and expected based on bioenergetic theory (29). The prior for *k* ensured values corresponding roughly to the expected values based on Lamb et al. (2022). eDNA data were scaled to Kilocopies m^{-3} to avoid working with very large numbers. The prior on sigma accounted for heteroscedasticity and increases variance with total abundance (N), and α and θ have a lower bound (*lb*) at zero.

Comparing the mass balance eDNA approach to other existing models

Most previous studies examining relationships between eDNA concentrations and organism abundance/biomass typically directly regress observed eDNA concentrations against estimates of abundance and/or biomass (38). To evaluate the relative importance of accounting for both bioenergetics and particle dynamics (i.e., mass balance modelling), we also fitted models in which eDNA concentrations were directly related to both population biomass (equation 6) and allometrically scaled biomass (equation 7, Yates et al. 2021a). Both models were fitted assuming a Gaussian distribution:

$$[eDNA]_{Lake} \sim intercept + \beta \frac{N \times \bar{M}}{lake\ area\ (ha)} \tag{6}$$

$$[eDNA]_{Lake} \sim \frac{I_0 \sum_{i=1}^{N} (M_i^b)}{lake\ area\ (ha)} \approx \frac{I_0 \times N(\bar{M} + \frac{\sigma_M^2}{2}\, b(b-1)\bar{M}^{b-2})}{lake\ area\ (ha)} \tag{7}$$

With the priors:

$$\begin{gathered} \ln(I_0) \sim normal(0, 5) \\ b \sim beta(2.8, 1.2) \\ \sigma \sim cauchy(0, 5) \end{gathered}$$

The performance of the different models was compared using the bayes R^2 and the value of the slope comparing the predicted and measured eDNA concentrations (the closer to one the better).

Evaluating the efficacy of eDNA to estimate population abundance

We used our integrated bioenergetics/mass-balance approach to derive estimates of population abundance to compare their precision and accuracy relative to estimates of abundance derived from mark-recapture studies. Specifically, we are interested in evaluating the efficacy of eDNA to both i) differentiate abundance *across* study populations; and ii) detect changes in abundance *within* a population over time (i.e., population reductions due to harvesting). To produce posterior estimates of total abundance for each of the populations during each sampling period using the mass balance approach, *N* was isolated from equation 5 and the model was refitted with the same priors. Additionally, season and sampling period can have a strong effect on mean Brook Trout eDNA concentrations observed across lakes (Gaudet-Boulay et al., 2023), either due to ecological differences in the eDNA production coefficient (I_0) across seasons or potential (unintentional) methodological differences in sampling. A sampling-period specific I_0 was therefore also fitted in the model. Equation 5 then becomes:

$$N \sim \frac{[eDNA]_{Lake} \times (Q + kV)}{I_0 f(T) \times \left(\bar{M} + \frac{\sigma_M^2}{2} b(b-1)\bar{M}^{b-2}\right)} \tag{8}$$

With the priors:

$$\begin{gathered}
\ln(I_0) \sim normal(10, 2) \\
b \sim beta(2.8, 1.2) \\
\ln(k) \sim normal(-0.7, 0.1) \\
Standard\ deviation_{K\ lake} \sim exponential(3) \\
Standard\ deviation_{\log(I_0\ Sampling\ period)} \sim exponential(3) \\
\sigma \sim \theta + N^{\alpha} \\
\alpha \sim normal(0{,}0.5),\ lb = 0 \\
\theta \sim normal(0{,}1),\ lb = 0
\end{gathered}$$

In this model, we also accounted for the measurement error on the mark-recapture total abundance estimates (56). We also examined the efficacy of the integrated eDNA model to differentiate population biomass, which can be estimated from the integrated eDNA model (equation 8) for each population by multiplying the posterior distribution of N values by the mean mass of each population. The performance of the model was assessed with the Bayesian R^2 and the number of abundance estimates from eDNA that overlapped with mark-recapture (and their respective 95% credible intervals).

Examining potential sources of bias

To assess the presence of systematic bias of the model, a Bayesian mixed model with non-informative priors was fitted with the z-score of each eDNA prediction as the independent variable. The fixed effects were mark-recapture population estimates, elevation, lake thermal mixing status (yes = stratified or no = mixed), season (summer or fall), lake area and experimental treatment (harvested or reference lake). We included interactions between thermal mixing status and elevation, between season and treatment, and between season and elevation. Lake was included as a random intercept to account for non-independence. This bias testing model was performed on the outputs of the eDNA mass-balance models, including sampling period specific I_0.

Can our nine lakes calibrate a model to predict abundance in novel systems?

Due to repeated sampling within and across years, the model described in equation 8 can estimate lake-specific degradation constants and sampling period specific I_0, and is thus well-suited to monitoring abundance within and across this specific set of nine study lakes. However, this model cannot be extended to estimate population abundance in a lake without pre-existing data (e.g., where a lake specific k is unknown). To explore the efficacy of this approach to estimate abundance in lakes (equation 8) without pre-existing baseline data, we additionally fitted two

models: i) a model with lake specific k but no sampling period specific I_0 to assess the importance of accounting for local, lake-specific eDNA degradation rates (referred to as equation 8.1 in the results) and ii) a model without a lake specific eDNA disappearance constant (k) nor sampling period specific I_0 to assess how a generalizable model would perform as well as the importance of a sampling period specific eDNA generation coefficient (referred to as equation 8.2 in results). The performance of the different models was compared using Bayesian R^2 values. All analyses were performed in *R* (49) using the *brms* package (version 2.18.0) (57).

Results

Fish abundance varied significantly between lakes with estimates ranging from 72 (46 – 117) to 3173 (2644 – 3965) individuals based on mark-recapture. Temporal variation in population size was particularly pronounced in the size-selectively harvested lakes, where annual harvest ranged from 46.8% to 92.7% of the population. Despite these high harvest levels, compensatory effect allowed populations to partially recover between harvesting treatments (36). Natural mortality also contributed to population changes. For instance, in Dog Lake population abundance declined by 32% between the summers of 2018 and 2019, reflecting substantial non-harvest losses.

Comparing the mass balance eDNA approach to other existing models

The eDNA mass balance model explained more variability in eDNA concentration and was more accurate than a linear regression with biomass or allometrically scaled biomass (Table 1, Fig. 2). While all three models had little bias, biomass alone explained only 24% of the variation in predicted eDNA values (Bayes R^2). This increased to 44% after integrating the effect of allometry

on eDNA production (e.g., allometrically scaled mass). Integrating bioenergetics and mass-balance modelling approaches, however, explained 66% of the variation in observed eDNA concentrations.

Evaluating the efficacy of eDNA to estimate population abundance

Total abundance estimates derived from eDNA concentration data (equation 8) effectively differentiated populations across spatial and temporal scales (Fig. 3). The study populations exhibited a substantial gradient in total abundance across lakes, with the smallest population exhibiting a 32-fold difference in abundance relative to the largest population. At this scale, the integrated eDNA model was very effective at differentiating abundance across populations (Fig 3).

Variations in abundance across populations were significantly greater than within population changes, even in populations that exhibited substantial declines due to size selective harvesting manipulations. Within a population, abundance levels were more consistent across seasons and years (Fig. 3 and Fig. 4). Yet, even within populations, the accuracy and precision of the integrated eDNA model was sufficient to detect changes across seasons and years in several of the harvested lakes (e.g., McNair and Temple), as well as natural fluctuations in abundance across control populations (e.g., Margaret and Dog lakes).

The majority (31/35 or 89%) of the total Brook Trout abundance estimates derived from the eDNA mass balance model (equation 8) fell within the credible intervals for total abundance estimated by mark-recapture (Fig. 3 and 4). Yet, five estimates showed no overlap between the 80% credible intervals for total abundance from both eDNA and mark-recapture, leaving 26 (74%) estimates that exhibited a high correspondence. The integrated eDNA model similarly proved effective at differentiating population biomass across space and time. The biomass estimates from the integrated eDNA model were highly correlated with estimates derived from mark-recapture

(Bayes R^2 = 0.80) (Fig. 5) and, as with abundance, enabled the clear differentiation of biomass both across and within populations (Fig. SM1).

Examining potential sources of bias

When testing for different bias of the eDNA total abundance estimates, there was no significant systematic bias detected related to the total Brook Trout abundance nor lake size. There was a non-significant tendency to generate larger estimates in lakes experiencing experimental harvesting during the fall seasons (e.g., Ross and Mud at fall 2019) and smaller abundance estimates in control lakes and some harvest lakes outside of harvest periods (e.g. Helen Lake at fall 2018 and 2019, Dog Lake at fall 2018, and Mud Lake at summer 2018) compared to mark-recapture (Fig 4, Table 2, and Appendix 5). There was also a non-significant trend for the integrated eDNA model to produce estimates of abundance lower than mark-recapture estimates for samples taken during lakes that were experiencing thermal mixing, especially at high elevation. However, this was largely driven by Helen Lake, for which total abundance estimates from the integrated eDNA model were lower than mark-recapture while mixing occurred during fall 2018 and 2019.

Can our nine lakes calibrate a model to predict abundance in novel systems?

Incorporating a lake specific eDNA disappearance constant (*k*; the sum of the sedimentation and degradation constant) significantly improved model performance, as demonstrated by the large standard deviation on a general *k* estimated across lakes and the reduction in model performance when removing lake specific *k* (equation 8.2), with the Bayes R^2 dropping from 0.70 to 0.51 (Table 1 and Appendix 6). The effect of accounting for a sampling period specific eDNA generation coefficient (I_0; equation 8.1) also reduced model performance, but to a substantially lesser extent than including a lake-specific *k*. Removing the sampling period specific eDNA generation

coefficient (I_0) decreased the Bayes R^2 value from 0.72 to 0.70. This mainly impacted predictions based on the summer 2018 samples, making the temporal trends in McNair and Olive less congruent to the mark-recapture (Table 1 and Fig SM2). Brook Trout abundances predicted from the eDNA mass balance model that lacked a sampling period specific I_0 (equation 8.1) also fell within the 95% credible intervals for total abundance estimated by mark-recapture for 31/35 estimates (Appendix 5).

Discussion

Despite the increasing demand for abundance data to understand and manage ecosystems (34), the application of eDNA to quantify the abundance or biomass of aquatic organisms remains a controversial topic (8, 9, 58). This stems from the complex interplay of biotic and abiotic factors that influence eDNA production, degradation, and transportation in natural aquatic environments (12), which are considerably more variable across space and time in natural ecosystems relative to controlled laboratory settings (9).

Herein we demonstrate, that eDNA can provide estimates of organism abundance and biomass comparable in precision and accuracy to mark-recapture. However, achieving this accuracy required accounting for biotic and abiotic factors that affect its distribution in natural ecosystems, namely the ecology of its production and the dynamics of the physical eDNA particles. Applying our integrated bioenergetics and mass balance approach to model quantitative eDNA signals greatly improved the relationship between eDNA concentration and total abundance estimates in the study ecosystems, compared to a basic linear correlation between either eDNA and biomass or eDNA and allometrically scaled biomass. Our integrated approach allowed for the

accurate differentiation of population abundance and biomass across study populations (i.e., space) based on lake eDNA concentrations. Quantifying the concentration of eDNA within the study lakes also facilitated the prediction or monitoring of temporal changes in fish abundance/biomass over time within several populations.

While previous research demonstrated the influence of several factors on eDNA concentration and dynamics, including transportation (7, 11, 22), degradation (11, 13, 15, 16), temperature (28, 59), and fish metabolic rate (12, 23, 26), this study is the first to integrate mass balance processes with bioenergetics to relate eDNA concentration to absolute population abundance in natural ecosystems. We empirically demonstrate that eDNA could represent a time- and cost-effective tool for generating total abundance estimates with sufficient precision to detect changes in abundance by integrating relatively simple-to-collect environmental and ecological data.

Evaluating and comparing the efficacy of eDNA to estimate population abundance

eDNA abundance estimates corresponded to mark-recapture abundance estimates in 31 out of 35 cases when considering the 95% credible interval, and in 26 out of 35 cases when considering the 80% credible intervals. Most discrepancies (6 out of 9) were likely due to issues associated with our unique study design causing the violation of assumptions in our mass-balance modelling (i.e. experimental harvesting), and/or the violation of assumptions associated with the mark-recapture estimates for populations in the fall sampling periods (e.g., no natural mortality or emigration).

There was a non-significant trend wherein estimates of abundance derived from eDNA tended to be higher than for mark-recapture in harvested lakes in the fall. The time necessary for eDNA concentrations to stabilize depends on eDNA degradation rate, which can be impacted by

amplicon size, temperature, and potentially UV radiation, pH, and microbial activity (14). The values estimated by the eDNA mass balance model for the *k* degradation parameter (the eDNA "disappearance rate", or the sum of the sedimentation and degradation rates) corresponded to expectations based on previous controlled experiments. Notably, a meta-analysis reported degradation constant values between 0–0.2 h-1 or 0 – 2.5 day-1 in freshwater ecosystems (Lamb et al., 2022), and we obtained an average *k* value estimate of 0.49 (0.41 – 0.6) day-1. Based on the eDNA disappearance constant *k* estimated by the model, the outflow rate, and the summer eDNA concentration in each lake and each year, we estimated that three to twenty days would be necessary for eDNA turnover in this study system (Appendix 8). For logistical reasons (i.e., the harvesting treatment took a substantial amount of time and extended late in the season), there was generally less than seven days between the end of harvesting and eDNA sampling within the harvested lakes which is less than the turnover rate in most cases. A tendency for the integrated eDNA model to overestimate fish abundance following harvesting could also be explained by many eDNA copies released during harvesting. Carcasses can be a significant source of eDNA (60) and often remained in nets for 24 - 48h, thus increasing the initial amount of eDNA and time before stabilisation back to a steady state concentration.

There were also four cases where eDNA estimates were lower than mark-recapture estimates during fall sampling periods when looking at 80% credible intervals (two in Helen Lake and two in Dog Lake). An instinctive response may be to attribute the discrepancies to error in the integrated eDNA model. However, the discrepancies between the eDNA-derived and mark-recapture abundance estimates could be attributable to violations of the assumptions of the mark-recapture methodology employed. In Helen Lake, the discrepancies could be attributed to mixing, but no other lakes experiencing mixing exhibited an underestimation of Brook Trout abundance

from the integrated eDNA model. Alternatively, Helen Lake also has an outflow stream where the Brook Trout are known to spawn. The lower estimates of total fall abundance from the integrated eDNA model in Helen Lake could therefore be due to fish leaving the lake to spawn, especially since spawning is known to occur earlier in high-altitude alpine lakes such as Helen Lake (61). Similarly, mark-recapture estimates demonstrated that Dog Lake exhibited a substantial natural decline in abundance between 2018 and 2019. While mortality was assumed to be negligeable between summer and fall, we do not actually know when this mortality event occurred because mark-recapture studies occurred in the spring and summer of each year. Results from the eDNA model could imply that the natural mortality event may have occurred between the Summar and fall sampling periods both in 2018 and 2019, as the estimate of abundance derived from the integrated eDNA model in fall 2018 corresponds well to the mark-recapture estimates of abundance obtained during summer 2019. In these cases, eDNA may more accurately reflect the abundance of fish in the study lakes than mark recapture estimates, for which we assumed there was no natural mortality nor emigration for the fall estimates.

Generalisation of the integrated bioenergetic and mass balance model

While 25% (9 out of 35) of the estimates from the integrated bioenergetic and mass balance model exhibit some discrepancy between eDNA-derived and mark-recapture estimates of abundance when looking at the 80% CI, only 3 out of 35 (9%) of the discrepancies lack potential explanations, which is reduced to 1 of 35 (3%) when looking at the 95% CI. Additionally, although discrepancies between eDNA and mark-recapture derived estimates in abundance existed, these discrepancies still reflect a relatively high level of precision *across* study populations, for which eDNA permitted broad differentiation across high- and low-abundance populations. Although we advise caution, these results imply that eDNA can potentially be used to reliably estimate

abundance and biomass in nature across populations, especially given that mark-recapture estimates themselves can exhibit biases or fail when assumptions are violated. Our study design does not allow us to identify which estimate is closer to the 'true' population size, and these issues would only be compounded in larger natural ecosystems. Overall, we detected no systematic bias in eDNA-derived estimates related to lake size, mixing, elevation, and population size. The integrated bioenergetics and mass balance approach that we applied to the analysis of eDNA concentrations produced mostly accurate population abundance estimates in small to medium-sized oligotrophic lakes. With further validation, such models have the potential to be extended to other aquatic ecosystems.

Accounting for lake mixing was an important component of the mass-balance modelling because the volume of the waterbody is a critical parameter in mass-balance modelling approaches. Previous research has demonstrated that eDNA disperses throughout a lake and is roughly homogenous during thermal mixing (54). During lake stratification, we found that the volume of the epilimnion is the appropriate volume input data, likely because Brook Trout remains close to the surface during summer (41). In contrast, when a lake is undergoing thermal mixing, the volume of the entire lake is the appropriate volume input data in the mass-balance model as the signal is diluted in the entire lake. We detected no systematic bias related to lake thermal mixing when data for Helen Lake was removed from the model; in all other lakes mixing explained no discrepancies between eDNA and mark-recapture abundance estimates. The absence of overestimation of Brook Trout abundance during mixing implies that the resuspension of sedimentary DNA during lake mixing likely had a negligible impact on abundance estimates, despite sediment representing a potentially potent source of eDNA (62). However, the amount of eDNA stored within sediments and how it exchanges with the water column also depends on an organism's spatial distribution, as

well as biotic and abiotic conditions in water and at the water-sediment interface (63). Sedimentary DNA re-suspension may be more important in other types of lakes or for other species, particularly for systems inhabited by benthivorous fish species that disturb sediment during feeding (e.g. Carp; Parkos et al., 2003).

Future directions: Can our nine lakes calibrate a model to predict abundance in novel systems?

Our perspective on the future application of eDNA to estimate population abundance and biomass is not to replace conventional methods, but rather to complement and/or improve upon estimates generated by them. Notably, the integrated bioenergetic and mass-balance model developed herein requires population size structure data which can only be obtained using mixed mesh index gillnetting. These methods, often used in state and provincial monitoring programs to estimate relative abundance based on CPUE (e.g. Zimmerman & Palo, 2011), require fewer resources than mark-recapture techniques. To broadly apply eDNA to estimate abundance using the methods and models developed herein also requires the establishment of a 'calibration' system, in which abundance estimated using conventional methods on a subset of ecosystems are used to calibrate an eDNA/abundance model. Once calibrated, this model can be applied to estimate abundance in other systems or time points lacking conventional estimates, by sampling eDNA and the required physicochemical parameters.

Our findings imply that the capacity to identify local predictors of *k* would be particularly critical for such applications, as baseline data to estimate *k* rates within an individual system substantially improved our model performance. To increase the reliability of abundance estimates in freshwater systems derived from eDNA, further research is needed on eDNA dynamics in lentic systems, with a particular focus on drivers of sedimentation and degradation rates. A quantitative understanding of these interactions would allow us to better predict sedimentation and degradation

rates specific to local and temporal conditions, which would in turn likely improve the integrated eDNA bioenergetic and mass-balance model performance as such dynamics could be integrated into the modelling process. Additionally, such knowledge would facilitate the estimation of the disappearance *k* parameter for lakes and time points outside of any putative "calibration system" lakes simply by measuring the relevant environmental variables.

A growing number of individual studies on eDNA degradation have found that eDNA degradation is the result of a complex combination of many variables such as pH, temperature, UV radiation and microbial activity, and eDNA properties such as amplicon size and whether the targeted gene is nuclear or mitochondrial (13, 15–17, 66). Although temperature has emerged as a consistently impactful parameter affecting eDNA degradation rates, the importance of other parameters such as UV, microbial activity, eDNA source, and amplicon size remains to be confirmed (Lamb et al. 2022, Jo and Minamoto 2021, Saito and Doi 2021). Collectively, we are still far from producing effective eDNA degradation models, as few studies to date have examined the interaction between these different environmental variables and eDNA properties (especially in natural ecosystems) and even fewer have examined eDNA sedimentation. However, in the absence of controlled empirical studies, further large-scale observational studies in natural ecosystems could facilitate the identification of drivers of local eDNA decay rates. Experiments studying degradation dynamics in natural ecosystems are thus crucial to facilitate and improve the adoption of eDNA to estimate abundance in natural ecosystems.

Theoretical and empirical evidence suggests that eDNA production rate is related to fish metabolic rate (23, 38, 67, 68). However, to our knowledge, only one study to date has experimentally verified this relationship under controlled empirical conditions and found a positive correlation between fish activity rate, energy use, and eDNA shedding rate (68). The mass balance

model assumes eDNA production scales with body mass and is influenced by temperature-dependent metabolic rates in poikilothermic Brook Trout (Hervé et al., 2023; Thalinger et al., 2021; Yates, Cristescu, et al., 2021; Yates et al., 2023). We used a temperature-dependent metabolic function (Hartman & Cox, 2008) to adjust for temperature effects on eDNA, though this may not fully reflect eDNA production dynamics. While many studies have documented an effect of temperature on eDNA production (e.g., Caza-Allard et al., 2022; T. Jo et al., 2019), whether this effect directly reflects metabolic responses to temperature remains to be tested. Investigating this could improve model assumptions and performance.

Our study additionally included mostly only headwater lakes, so eDNA inputs from upstream were considered equal to zero. However, this is rarely the case within a watershed. To add eDNA transportation into the lake within the integrated model, data on both the inflow rate and the average eDNA concentration in the inflow are required. In theory, these data can be implemented in equation 4 (31) such that:

$$[eDNA]_{Lake} \sim \frac{I_0 \sum_{i=1}^{N}(M_i^b) \times f(T) + Q_{in}[inflow]}{Q_{out} + kV} \tag{9}$$

Where Q_{in} is the inflow rate, Q_{out} is the outflow rate, and $[inflow]$ is the mean eDNA concentration of the inflow.

Conclusion

Our study proposes an integrated bioenergetics and mass-balance approach to relate eDNA concentration to abundance estimation. Our integrated eDNA model performed well at estimating abundance, facilitating total abundance and biomass estimates comparable to conventional mark-recapture methodologies, often considered a ‘gold-standard’ for population estimates. We show that eDNA can detect both spatial and temporal differences in fish abundance, provided the

equilibrium assumptions of the model are respected. To respect these assumptions, we recommend avoiding sampling ten to twenty days after known fluctuations in abundance (e.g. harvest), depending on the eDNA degradation rate and outflow flux. Overall, results of this study represent an important advancement in the use of eDNA to quantify total abundance, with strong implications for the application of eDNA for abundance estimates in a context of ecosystem management. The quantitative analysis of eDNA, applied with an integrated bioenergetics and mass-balance approach, could provide a cost-effective, precise, and accurate tool for managers and researchers, in an era where the acquisition of population abundance data is increasingly important for biodiversity management (e.g. Canals et al., 2021; Fonseca et al., 2023; Kelly et al., 2024).

Acknowledgements

We thank Parks Canada for logistic support, Andrew Macdonald for statistical advice, Maria Gheta for the help with the Taylor expansions, Yves Prairie for math validation, Jérôme Lemelin for the map, and Mélanie Desrochers for GIS advice. We also express our gratitude to Louis Astorg, Dylan Glaser, Brent Brooks, Shannon Clarke, Jean-Michel Matte, Justin Budyk, Mélia Lagacé, Mathilde Salamon, Sydney Urschel, and Ahmad Yaghi for field work. This research was funded by a Fonds de Recherche du Québec-Nature et Technologie (FRQNT) team grant to AMD, MEC, and DJF, a grant from the Natural Sciences and Engineering Research Council of Canada (NSERC Strategic Project grant, AMD, DJF), and a NSERC Discovery grant to AMD.

Authors contribution

JB designed the models and performed the analysis. JB and MCY contributed to the writing of the manuscript. MCY and JB collected the eDNA samples. MCY extracted the eDNA samples and

conducted laboratory analyses. All authors contributed to the study design and subsequent edits to the manuscript.

References

1. J. S. Albert, *et al.*, Scientists' warning to humanity on the freshwater biodiversity crisis. *Ambio* **50**, 85–94 (2021).

2. L. T. Ho, P. L. M. Goethals, Opportunities and Challenges for the Sustainability of Lakes and Reservoirs in Relation to the Sustainable Development Goals (SDGs). *Water (Basel)* **11**, 1462 (2019).

3. J. R. Post, *et al.*, Canada's Recreational Fisheries: The Invisible Collapse? *Fisheries (Bethesda)* **27**, 6–17 (2002).

4. A. Howarth, V. M. Nguyen, S. J. Cooke, Managing Canadian freshwater fisheries: persistent challenges and emerging opportunities. *Canadian Journal of Fisheries and Aquatic Sciences* **80**, 1436–1455 (2023).

5. G. F. Ficetola, C. Miaud, F. Pompanon, P. Taberlet, Species detection using environmental DNA from water samples. *Biol Lett* **4**, 423–425 (2008).

6. C. L. Jerde, Can we manage fisheries with the inherent uncertainty from eDNA? *J Fish Biol* **98**, 341353 (2019).

7. A. J. Sepulveda, N. M. Nelson, C. L. Jerde, G. Luikart, Are Environmental DNA Methods Ready for Aquatic Invasive Species Management? *Trends Ecol Evol* **35**, 668–678 (2020).

8. M. L. Rourke, *et al.*, Environmental DNA (eDNA) as a tool for assessing fish biomass: A review of approaches and future considerations for resource surveys. *Environmental DNA* **4**, 9–33 (2022).

9. M. C. Yates, D. J. Fraser, A. M. Derry, Meta-analysis supports further refinement of eDNA for monitoring aquatic species-specific abundance in nature. *Environmental DNA* **1**, 5–13 (2019).

10. K. A. Stewart, Understanding the effects of biotic and abiotic factors on sources of aquatic environmental DNA. *Biodivers Conserv* **28**, 983–1001 (2019).

11. L. Carraro, H. Hartikainen, J. Jokela, E. Bertuzzo, A. Rinaldo, Estimating species distribution and abundance in river networks using environmental DNA. *Proceedings of the National Academy of Sciences* **115**, 11724–11729 (2018).

12. M. C. Yates, M. E. Cristescu, A. M. Derry, Integrating physiology and environmental dynamics to operationalize environmental DNA (eDNA) as a means to monitor freshwater macro-organism abundance. *Mol Ecol* 1–20 (2021). https://doi.org/10.1111/mec.16202.

13. K. M. Strickler, A. K. Fremier, C. S. Goldberg, Quantifying effects of UV-B, temperature, and pH on eDNA degradation in aquatic microcosms. *Biol Conserv* **183**, 85–92 (2015).

14. P. D. Lamb, V. G. Fonseca, D. L. Maxwell, C. C. Nnanatu, Systematic review and meta-analysis: Water type and temperature affect environmental DNA decay. *Mol Ecol Resour* **22**, 2494–2505 (2022).

15. T. Jo, T. Minamoto, Complex interactions between environmental DNA (eDNA) state and water chemistries on eDNA persistence suggested by meta-analyses. *Mol Ecol Resour* **21**, 1490–1503 (2021).

16. T. Jo, H. Murakami, S. Yamamoto, R. Masuda, T. Minamoto, Effect of water temperature and fish biomass on environmental DNA shedding, degradation, and size distribution. *Ecol Evol* **9**, 1135–1146 (2019).

17. T. Saito, H. Doi, A Model and Simulation of the Influence of Temperature and Amplicon Length on Environmental DNA Degradation Rates: A Meta-Analysis Approach. *Front Ecol Evol* **9**, 1–8 (2021).

18. P. F. P. Brandão-Dias, *et al.*, Particle size influences decay rates of environmental DNA in aquatic systems. *Mol Ecol Resour* **23**, 756–770 (2023).

19. J. E. Littlefair, M. D. Rennie, M. E. Cristescu, Environmental nucleic acids: a field-based comparison for monitoring freshwater habitats using eDNA and eRNA. *bioRxiv* 2021.12.01.469845 (2021).

20. S. F. Jane, *et al.*, Distance, flow and PCR inhibition: EDNA dynamics in two headwater streams. *Mol Ecol Resour* **15**, 216–227 (2015).

21. D. Pont, Predicting downstream transport distance of fish eDNA in lotic environments. *Mol Ecol Resour* **24**, 1–14 (2024).

22. A. J. Shogren, *et al.*, Controls on eDNA movement in streams: Transport, Retention, and Resuspension. *Sci Rep* **7**, 5065 (2017).

23. M. C. Yates, T. M. Wilcox, M. Y. Stoeckle, D. D. Heath, Interspecific allometric scaling in eDNA production among northwestern Atlantic bony fishes reflects physiological allometric scaling. *Environmental DNA* **5**, 1105–1115 (2023).

24. M. J. Vanni, P. B. McIntyre, Predicting nutrient excretion of aquatic animals with metabolic ecology and ecological stoichiometry: A global synthesis. *Ecology* **97**, 3460–3471 (2016).

25. J. E. Allegier, S. J. Wenger, A. D. Rosemond, D. E. Schindler, C. A. Layman, Metabolic theory and taxonomic identity predict nutrient recycling in a diverse food web. *PNAS* **112 (20)** (2015).

26. M. C. Yates, *et al.*, Allometric scaling of eDNA production in stream-dwelling brook trout (Salvelinus fontinalis) inferred from population size structure. *Environmental DNA* **3**, 553–560 (2021).

27. K. W. Thornton, A. S. Lessem, A Temperature Algorithm for Modifying Biological Rates. *Trans Am Fish Soc* **107**, 284–287 (1978).

28. I. Caza-Allard, M. Laporte, G. Côté, J. April, L. Bernatchez, Effect of biotic and abiotic factors on the production and degradation of fish environmental DNA: An experimental evaluation. *Environmental DNA* **4**, 453–468 (2022).

29. D. Deslauriers, S. R. Chimpps, J. E. Breck, J. A. Rice, C. P. Madenjian, Fiash Bioenergetic 4.0: An R-based Modeling Application. *Fisheries (Bethesda)* **42**, 586–596 (2017).

30. P. C. Hanson, T. B. Johnson, D. E. Schindler, J. F. Kitchell, Fish Bioenergetics 3.0. *World Wide Web Internet And Web Information Systems* (1997).

31. S. C. Chapra, K. H. Reckhow, *Ingeneering Approaches for Lake Management. Volume 2: Mechanistic Modeling* (Butterworth, 1983).

32. L. M. Sassoubre, K. M. Yamahara, L. D. Gardner, B. A. Block, A. B. Boehm, Quantification of Environmental DNA (eDNA) Shedding and Decay Rates for Three Marine Fish. *Environ Sci Technol* **50**, 10456–10464 (2016).

33. K. Kagzi, R. M. Hechler, G. F. Fussmann, M. E. Cristescu, Environmental RNA degrades more rapidly than environmental DNA across a broad range of pH conditions. *Mol Ecol Resour* **22**, 2640–2650 (2022).

34. S. Huang, K. Yoshitake, S. Watabe, S. Asakawa, Environmental DNA study on aquatic ecosystem monitoring and management: Recent advances and prospects. *J Environ Manage* **323**, 116310 (2022).

35. J. S. Messner, M. M. Maclennan, R. D. Vinebrooke, Higher temperatures enhance the effects of invasive sportfish on mountain zooplankton communities. *Freshw Biol* **58**, 354–364 (2013).

36. J. M. O. Matte, D. M. Glaser, J. R. Post, D. J. Fraser, Contrasting demographic responses to size-selective harvesting among neighbouring wild fish populations. *Journal of Applied Ecology* (2023). https://doi.org/10.1111/1365-2664.14420.

37. D. Trépanier-Leroux, *et al.*, Density-dependent effects of exotic brook trout on aquatic communities in mountain lakes revealed by environmental DNA and morphological taxonomy. *Hydrobiologia* **851**, 1489–1512 (2023).

38. M. C. Yates, *et al.*, The relationship between eDNA particle concentration and organism abundance in nature is strengthened by allometric scaling. *Mol Ecol* **30**, 3068–3082 (2021).

39. T. M. Wilcox, *et al.*, Robust detection of rare species using environmental DNA: the importance of primer specificity. *PLoS One* **8**, e59520 (2013).

40. C. J. Krebs, "Estimating Abundance in Animal and Plant Populations" in *Ecology: The Experimental Analysis of Distribution and Abundance*, 6th ed., (Pearson Benjamin Cummings, 2009).

41. R. Tiberti, L. Nelli, S. Brighenti, R. Iacobuzio, M. Rolla, Spatial distribution of introduced brook trout Salvelinus fontinalis (Salmonidae) within alpine lakes: Evidences from a fish eradication campaign. *European Zoological Journal* **84**, 73–88 (2017).

42. S. A. Bonar, W. A. Hubert, D. W. Willis, *Standard methods for sampling North American freshwater fishes* (American Fisheries Society, 2009).

43. W. A. Hubert, K. L. Pope, J. M. Dettmers, "Passive capture techniques" in *A. V. Zale, D. L. Parrish, & T. . Sutton (Eds.), Fisheries Techniques*, 3rd ed., (American Fisheries Society, 2012).

44. L. Johnson, Homeostatic Characteristics of Single Species Fish Stocks in Arctic Lakes. *Canadian Journal of Fisheries and Aquatic Sciences* **40**, 987–1024 (1983).

45. J. R. Post, E. A. Parkinson, N. T. Johnston, Density-dependent processes in structured fish populations: Interaction strengths in whole-lake experiments. *Ecol Monogr* **69**, 155–175 (1999).

46. H. G. M. Ward, P. J. Askey, J. R. Post, D. A. Varkey, M. K. McAllister, Basin characteristics and temperature improve abundance estimates from standard index netting of rainbow trout (Oncorhynchus mykiss) in small lakes. *Fish Res* **131–133**, 52–59 (2012).

47. G. E. Morgan, Manual of instructions - fall walleye index netting (FWIN). (2002).

48. Z. E. Schnabel, The Estimation of the Total Fish Population of a Lake. *The American Mathematical Monthly* **45**, 348–352 (1938).

49. R. CoreTeam, R: A language and environment for statistical computing. [Preprint] (2020). Available at: https://www.r-project.org.

50. D. H. Ogle, *Introductory Fisheries Analyses with R* (Chapman & Hall, 2015).

51. G. A. F. Seber, *The Estimation of Animal Abundance and Related Parameters*, B. Publishing, Ed., 2 nd (2002).

52. K. J. Hartman, M. K. Cox, Refinement and Testing of a Brook Trout Bioenergetics Model. *Trans Am Fish Soc* **137**, 357–363 (2008).

53. Environment and Natural Ressources Canada, Real-Time Hydrometric Data. (2024). Available at: https://wateroffice.ec.gc.ca/map/index_e.html.

54. J. E. Littlefair, L. E. Hrenchuk, P. J. Blanchfield, M. D. Rennie, M. E. Cristescu, Thermal stratification and fish thermal preference explain vertical eDNA distributions in lakes. *Mol Ecol* **30**, 3083–3096 (2021).

55. J. Beaulieu, M. Gheta, A. A. M. Macdonald, Using Taylor Expansions to fit power law functions. (2024). Available at: https://julien-beaulieu.github.io/taylor-expansions-stats.github.io/vignette_taylor_series.html.

56. R. McElreath, *Statistical rethinking: A Bayesian course with examples in R and Stan*, Chapman an (2018).

57. P. C. Bürkner, Advanced Bayesian multilevel modeling with the R package brms. *R Journal* **10**, 395–411 (2018).

58. M. B. Parsley, E. J. Crespi, T. A. G. Rittenhouse, J. L. Brunner, C. S. Goldberg, Environmental DNA concentrations vary greatly across productive and degradative conditions, with implications for the precision of population estimates. *Sci Rep* **14**, 17392 (2024).

59. T. S. Jo, Correlation between the number of eDNA particles and species abundance is strengthened by warm temperature: simulation and meta-analysis. *Hydrobiologia* **850**, 39–50 (2023).

60. K. J. Dunker, *et al.*, Potential of environmental DNA to evaluate northern pike (Esox lucius) eradication efforts: An experimental test and case study. *PLoS One* **11**, 1–21 (2016).

61. J. Beaulieu, *et al.*, Environmental DNA for assessing impact and recovery of aquatic communities in an invaded mountain lake. *Lake Reserv Manag* **40**, 111–131 (2024).

62. C. R. Turner, K. L. Uy, R. C. Everhart, Fish environmental DNA is more concentrated in aquatic sediments than surface water. *Biol Conserv* **183**, 93–102 (2015).

63. E. Capo, *et al.*, Lake Sedimentary DNA Research on Past Terrestrial and Aquatic Biodiversity: Overview and Recommendations. *Quaternary* **4**, 6 (2021).

64. J. J. Parkos, V. J. Santucci, D. H. Wahl, Effects of adult common carp (Cyprinus carpio) on multiple trophic levels in shallow mesocosms. *Canadian Journal of Fisheries and Aquatic Sciences* **60**, 182–192 (2003).

65. J. K. M. Zimmerman, R. T. Palo, Reliability of catch per unit effort (CPUE) for evaluation of reintroduction programs – A comparison of the mark-recapture method with standardized trapping. *Knowl Manag Aquat Ecosyst* 07 (2011). https://doi.org/10.1051/kmae/2011016.

66. T. Saito, H. Doi, Degradation modeling of water environmental DNA : Experiments on multiple DNA sources in pond and seawater. 850–860 (2021). https://doi.org/10.1002/edn3.192.

67. A. Hervé, *et al.*, Do temperature and population size structure influence the quantity of fish eDNA in water? *Front Environ Sci* **11**, 1–6 (2023).

68. B. Thalinger, *et al.*, The Effect of Activity, Energy Use, and Species Identity on Environmental DNA Shedding of Freshwater Fish. *Front Ecol Evol* **9**, 1–13 (2021).

69. R. P. Kelly, *et al.*, Toward a national eDNA strategy for the United States. *Environmental DNA* **6**, 1–10 (2024).

70. V. G. Fonseca, *et al.*, The Application of eDNA for Monitoring Aquatic Non-Indigenous Species: Practical and Policy Considerations. *Diversity (Basel)* **15**, 1–17 (2023).

71. O. Canals, I. Mendibil, M. Santos, X. Irigoien, N. Rodríguez-Ezpeleta, Vertical stratification of environmental DNA in the open ocean captures ecological patterns and behavior of deep-sea fishes. *Limnol Oceanogr Lett* **6**, 339–347 (2021).

Table 1. Parameter estimates and Bayes R^2 for the linear regression, allometrically scaled biomass regression and the mass balance model. The formula column refers to the formula of the model specified as equations in the method section. The numbers in parentheses represent the 95 % credible interval on parameter estimates. The Bayes R^2 reported are the mean of the Bayes R^2 of each model iteration. *b* is the allometric coefficient, *k* is the sum of the degradation and sedimentation constant, I_0 is the eDNA generation coefficient, β is the slope of the linear model, sigma is the variance. Sd stands for standard deviation.

Model	**Bayes R^2**	***b***	***k***	**Sd(ln(*k*))**	**Ln(I_0) or β**	**Sd(ln(I_0))**	**Sigma**	**Formula**
Linear regression	0.24	-	-	-	0.13 (0.05-0.21)	-	2387 (1889-3037)	equation 6
Allometrically scaled biomass	0.44	0.58 (0.25-0.89)	-	-	0.03 (-1.4-1.43)	-	2162.93 (1707-2775)	equation 7
Mass balance eDNA	0.66	0.61 (0.31-0.92)	0.5 (0.41-0.6)	0.17 (0.01-0.59)	9.71 (8.18-11.11)	-	θ = 5.01 (4.35-5.77) α = 0.14 (0.09-0.18)	equation 5
Mass balance with N as dependent variable, with lake specific *k* and sampling period specific I_0	0.72	0.59 (0.18-0.93)	0.49 (0.40-0.6)	0.40 (0.14-0.82)	10.06 (8.33-11.98)	0.33 (0.03-0.83)	θ = 2.49 (1.26-3.75) α = 0.16 (0.10-0.21)	equation 8
Mass balance with N as dependent variable without sampling period specific I_0	0.70	0.65 (0.21-0.96)	0.49 (0.4-0.59)	0.5 (0.21-0.97)	9.7 (8.13-11.76)	-	θ = 2.84 (1.69-4.02) α = 0.15 (0.09-0.2)	equation 8.1
Mass balance with N as dependent variable without lake specific *k* nor sampling period specific I_0	0.51	0.47 (0.06-0.96)	0.49 (0.4-0.59)	-	10.60 (8.23-12.47)	-	θ = 2.75 (1.55-3.98) α = 0.17 (0.11-0.21)	equation 8.2

Table 2. Parameter estimates for bias testing. The bayes R^2 reported is the mean of the Bayes R^2 of each draw. The dependent variable is the z-score of each total Brook Trout population abundance estimate, calculated as the estimate from eDNA minus the estimate from mark-recapture divided by the standard deviation of the draws of the estimate from eDNA. After running the model with all the lakes, the same model was fitted without Helen Lake to see if some of the results were driven only by this one lake. CI stands for credible interval.

	With Helen Lake			**Without Helen Lake**		
Variable	**Parameter estimate**	**CI (95 %)**	**Bayes R^2**	**Parameter estimate**	**CI (95 %)**	**Bayes R^2**
Total abundance	-0.00045	-0.0087 – 0.0018	0.40	-0.00054	-0.0008 – 0.0018	0.31
Lake area	0.02	-0.18 – 0.21		-0.01	-0.21 – 0.19	
Mixing (yes)	0.20	-1.72 – 2.10		-0.14	-2.12 – 1.88	
Elevation	-0.00024	-0.0021 – 0.0026		-0.00008	-0.0031 – 0.0029	
Season (fall)	0.15	-1.67 – 1.98		-0.28	-2.17 – 1.62	
Treatment (Harvest)	-0.06	-1.38 – 1.30		-0.35	-1.76 – 1.12	
Elevation*Mixing	-0.001	-0.004 – 0.002		-0.00009	-0.0026 – 0.0024	
Elevation*Season	-0.0005	-0.0025 – 0.0021		0.0001	-0.002 – 0.003	
Season*Treatment	1.17	-0.39 – 2.70		0.80	-0.76 – 2.35	

List of figures

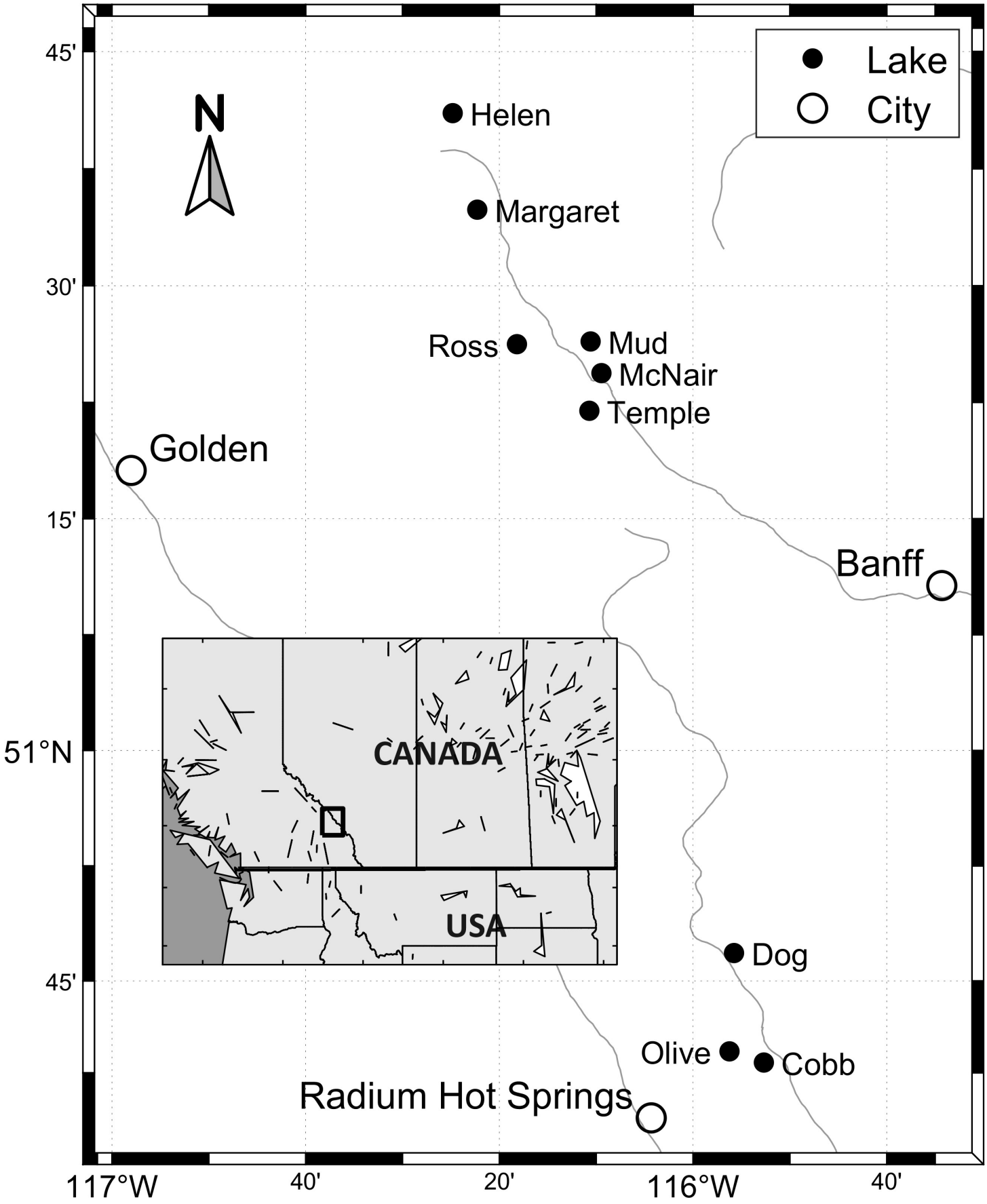
Figure 1. Map of the study area showing the locations of the lakes included in the research.

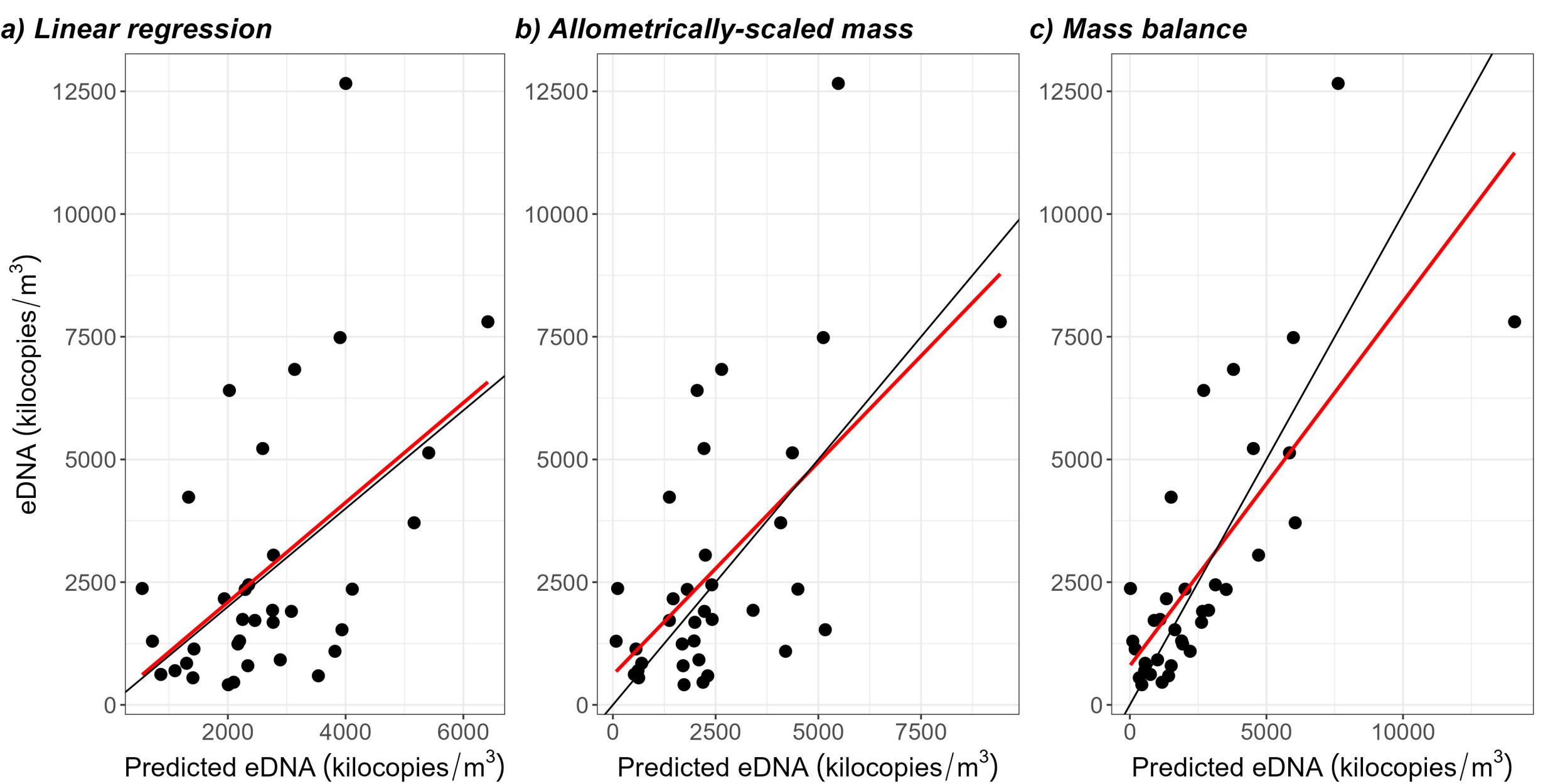
Figure 2. Comparison of model performance relating eDNA concentration to total abundance. The black line represents the 1:1 slope, while the red line shows the linear regression between the predicted and measured eDNA concentrations. Parameter estimates and R^2 values for each model are provided in Table 2.

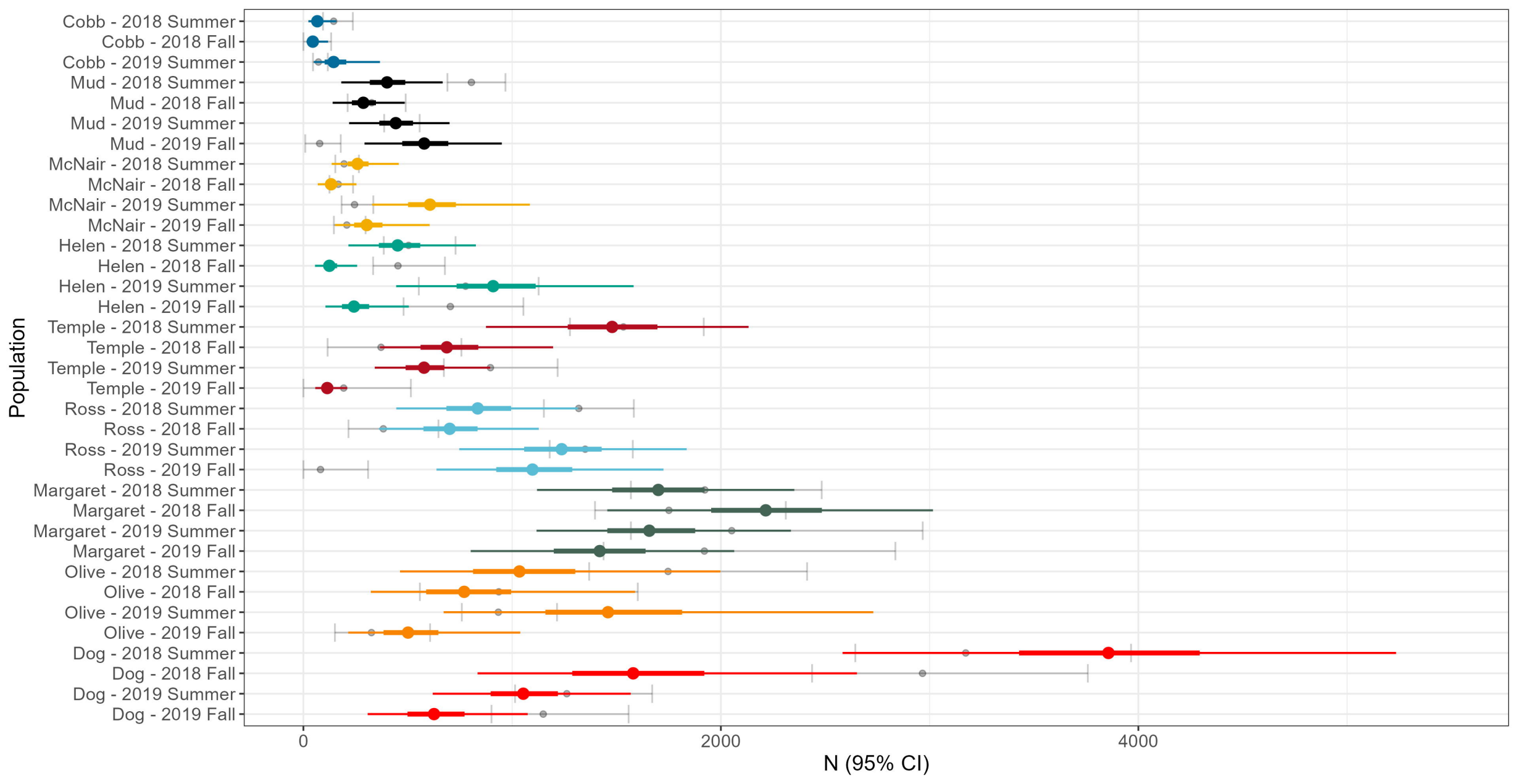
Figure 3. Total abundance estimates from eDNA for all the populations and sampling periods. The thick colored line represents the 50 % credible interval of eDNA estimates, while the thin colored lines represent the 95 % credible intervals. Grey lines and dots represent the 95% credible interval and mean from the mark-recapture abundance estimates. eDNA estimates are derived from the full mass balance model (Equation 8).

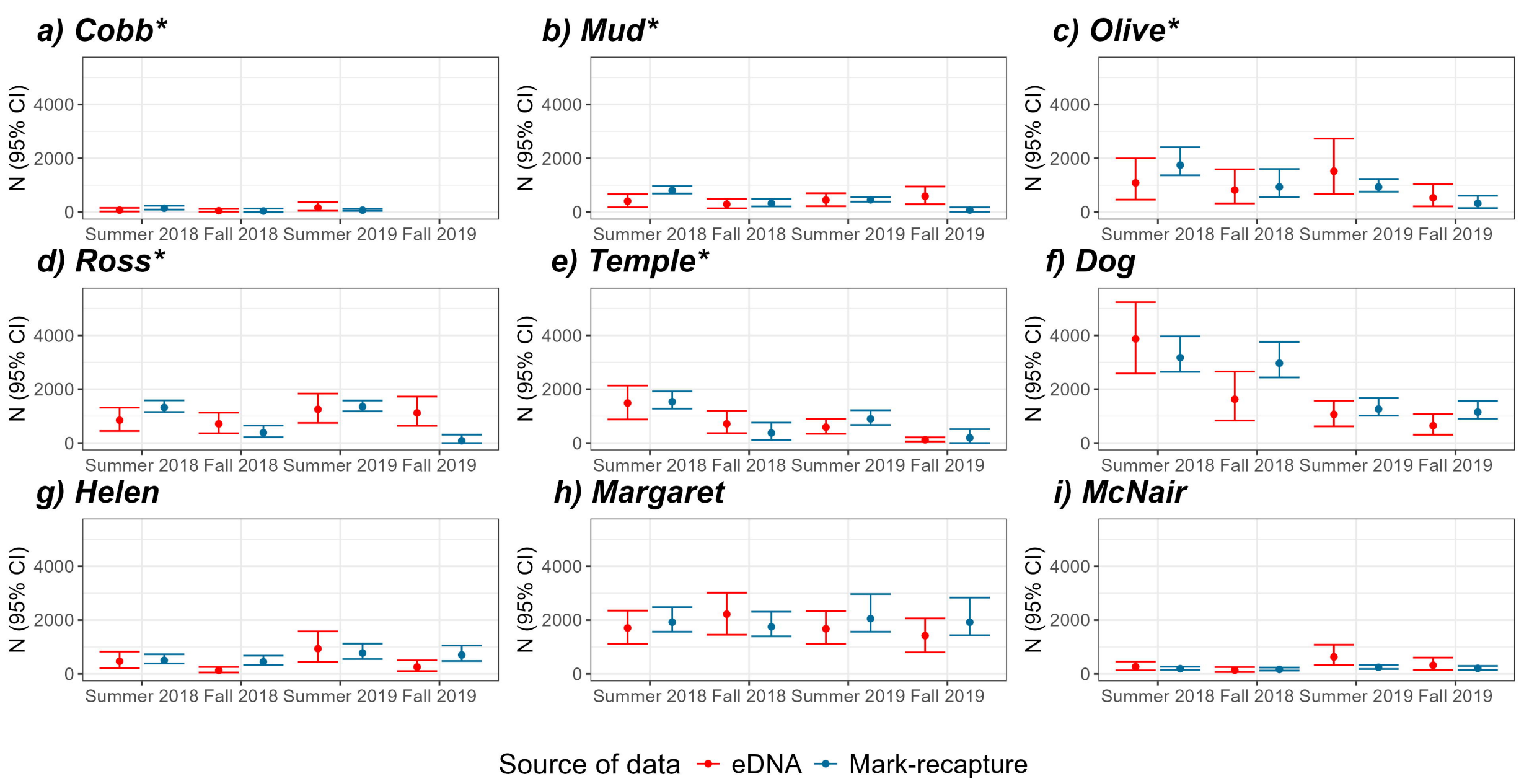
Figure 4. Total abundance estimates from the eDNA mass balance model (red) and mark-recapture (blue). eDNA estimates are produced from the fitted mass balance model with sampling period specific eDNA generation coefficient (I_0) (Equation 8). Lakes that were harvested are marked with an asterisk (*).

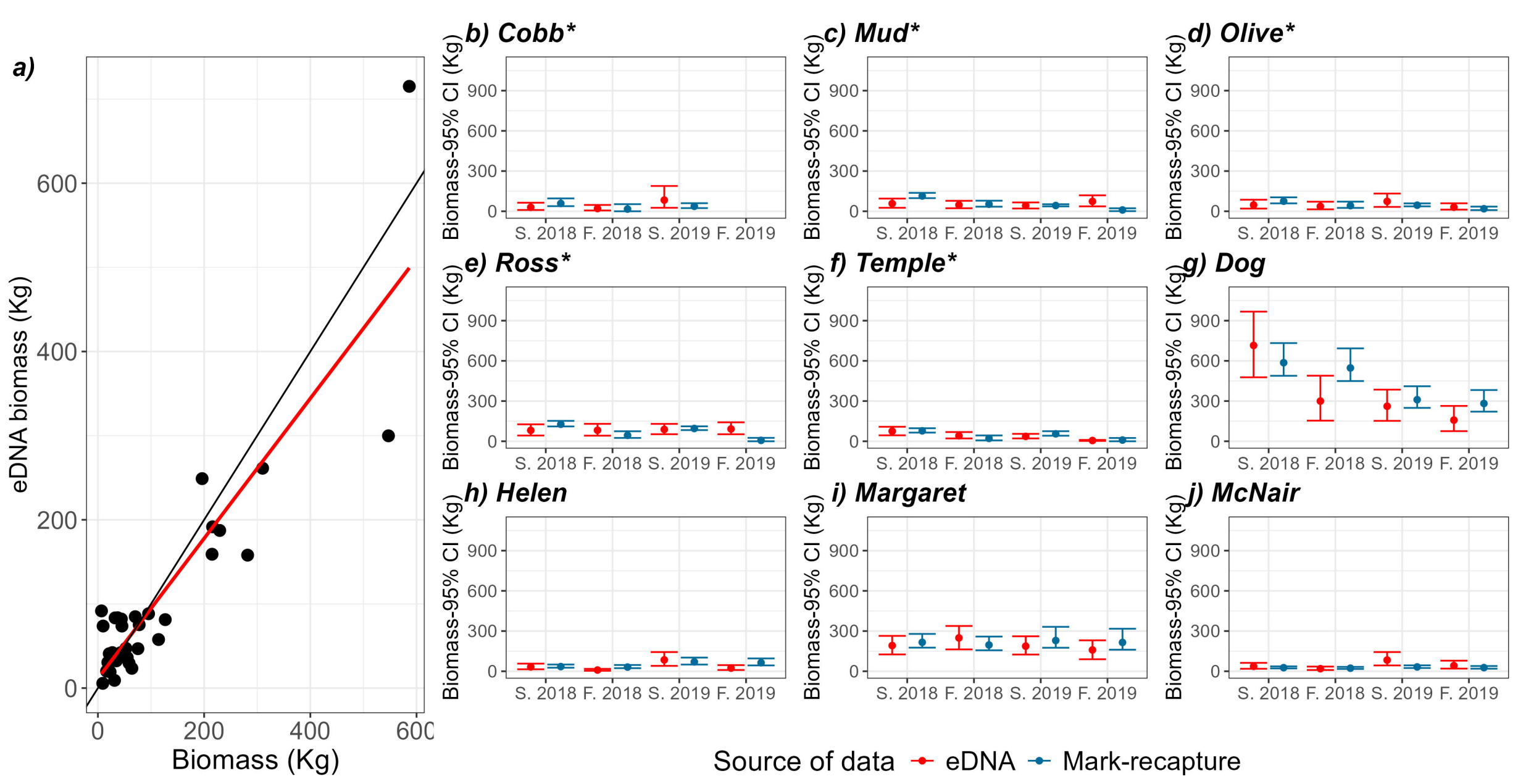
Figure 5. Biomass estimated from the eDNA mass balance model and mark-recapture. In a) the x axis is the biomass estimated from the mark-recapture abundance estimates and the y axis is the biomass estimated from the eDNA mass balance model. The red line is the linear regression between the two, and the black line is the 1:1 slope. In b) to j) the biomass estimated from the eDNA mass balance model (red) and the biomass estimated from the mark-recapture (blue). eDNA estimates are produced from the fitted mass balance model with sampling period specific eDNA generation coefficient (I0) (equation 8.2). Lakes that were harvested are marked with an asterisk (*).


Lake
City
N
Helen
Margaret
Ross
Mud
McNair
Temple
Golden
Banff
CANADA
USA
Dog
Olive
Cobb
Radium Hot Springs
45'
30'
15'
51°N
45'
117°W
40'
20'
116°W
40'


a) Linear regression
b) Allometrically-scaled mass
c) Mass balance
eDNA (kilocopies/m3)
Predicted eDNA (kilocopies/m3)
0
2500
5000
7500
10000
12500
2000
4000
6000


Population
Cobb - 2018 Summer
Cobb - 2018 Fall
Cobb - 2019 Summer
Mud - 2018 Summer
Mud - 2018 Fall
Mud - 2019 Summer
Mud - 2019 Fall
McNair - 2018 Summer
McNair - 2018 Fall
McNair - 2019 Summer
McNair - 2019 Fall
Helen - 2018 Summer
Helen - 2018 Fall
Helen - 2019 Summer
Helen - 2019 Fall
Temple - 2018 Summer
Temple - 2018 Fall
Temple - 2019 Summer
Temple - 2019 Fall
Ross - 2018 Summer
Ross - 2018 Fall
Ross - 2019 Summer
Ross - 2019 Fall
Margaret - 2018 Summer
Margaret - 2018 Fall
Margaret - 2019 Summer
Margaret - 2019 Fall
Olive - 2018 Summer
Olive - 2018 Fall
Olive - 2019 Summer
Olive - 2019 Fall
Dog - 2018 Summer
Dog - 2018 Fall
Dog - 2019 Summer
Dog - 2019 Fall
0
2000
4000
N (95% CI)


a) Cobb*
b) Mud*
c) Olive*
d) Ross*
e) Temple*
f) Dog
g) Helen
h) Margaret
i) McNair
N (95% CI)
0
2000
4000
Summer 2018
Fall 2018
Summer 2019
Fall 2019
Source of data
eDNA
Mark-recapture


a)
eDNA biomass (Kg)
Biomass (Kg)
0
200
400
600
b) Cobb*
c) Mud*
d) Olive*
e) Ross*
f) Temple*
g) Dog
h) Helen
i) Margaret
j) McNair
Biomass-95% CI (Kg)
0
300
600
900
S. 2018
F. 2018
S. 2019
F. 2019
Source of data
eDNA
Mark-recapture