Knut Midttun

Brown trout recruitment from natural streams with residual regulated flow

Lake population dynamics in the Upper Kova River system, Telemark

Appendix 2: Supplement to Methods

Contents

1 Formal requirements and ethics underlying methods

The Animal Welfare Act is the overriding legislation for the treatment of animals in Norway, and fish are among the animals covered by this law (Dyrevelferdsloven, 2009, p. § 2). To undertake animal research, the responsible institution and persons need permission from the appropriate authority (Dyrevelferdsloven, 2009, p. § 13). Mattilsynet is this authority; its role and more detailed rules and regulations for animal research are detailed in a separate bylaw (Forskrift om bruk av dyr i forsøk, 2015). Key principles are "the 3 Rs" (Russel & Burch, 1959): replace (avoid using animals when possible), reduce (use power analysis to ensure the use of as few animals as possible), and refine (maximum animal welfare and improvement of methods). Permissions presuppose appropriate training and accreditation. Another condition is that the research quality should be as high as possible, and that the results are published in appropriate channels. Gillnet test fisheries require a permit from the regional federal environmental authority (*Statsforvalteren*), but no separate handling permit from Mattilsynet.

Given the scope of the present study, the use of animals could not be avoided. Gillnet fishing is a destructive sampling method, killing the fish, and the welfare of the trapped fish is probably low up to the point of death. Less harmful alternative methods such as hydroacoustic sampling should thus be considered in principle but was judged not to be feasible in the present case. One important reason was that several of the data collection methods required dead samples. A substantial change of methods would also limit the temporal-comparative aspect of the study. Hopefully, new technologies will pave the way for better options in the future, but for the moment, test fishing with gillnets seemed the only option for this project. An ameliorating factor in this respect was perhaps the relatively long period since the last test fishery in 2009. As for the number of fish collected, this is not directly controllable but influenced by the fishing effort. In this project, the number chosen was in line with the relevant recommendation in NS-14757 (and its underlying power analysis) and approved by the project supervisor.

2 Sampling, data collection and registration

2.1 Gillnets: type, locations, setting/lifting times with weather conditions

As it turned out, five of the eight gillnets used in Nedre Urdetjønn were of a slightly different type, having a height of 1.4 meters instead of the standard 1.5. This was accounted for in the analyses.

Replication of the 2009 gillnet placement was not done in one case: gillnet serial number 8 in Bjønntjønn 2009 was set in the separate lake section upstream from the "stream-like reach"; in 2022, we opted to drop this section and instead set our gillnet number 4 (Fig. 3.3) in a part of the main Bjønntjønn lake poorly covered by gillnets in 2009.

As the way the gillnets had been packed (and thus how they were set) varied, we noted which part (panel 1 or 12) was facing inward and outward, respectively (Table A1 3.1). The setting and lifting times of gillnets recommended in the NS-EN 14757 were adapted to summer day length at our latitudes. In Øvre- and Nedre Urdetjønn, the gillnets were set and water measurements other than with the multiparameter logger were taken in the evening (18 – 21 o'clock) of June 20. Weather conditions were sunny, with air temperatures around 15° C, and a light breeze (< 5 m/s) from west/north-west. These gillnets were lifted the next day (11 – 13 o'clock). Weather conditions then were cloudy with light rain, the air temperature around 10° C, and the winds about the same as the day before. In Bjønntjønn, the gillnets were set in the evening of June 21 (20 – 21 o'clock). Weather conditions were about the same as in the morning, except that the rain had stopped. These gillnets were lifted and all water measurements for Bjønntjønn were taken the next day (the latter around 13 – 14 o'clock). The weather was then partly cloudy, with temperatures around 15° C, and wind as before. Later that afternoon, measurements were made with the multiparameter logger in Øvre- and Nedre Urdetjønn, at which time it was cloudy and with some heavier gusts of wind.

2.2 Data registration in the field

Removing the caught fish from the gillnets, it was sometimes difficult to decide which panel particularly entangled individuals were caught in, but the conclusion was based on the mesh size found in the fish mouth (if any). The adipose fin of the unexpectedly small individual of apparent hatchery stock caught in Bjønntjønn was carefully assessed by both fieldwork participants in unison. More generally, the in-situ fish registrations for Øvre Urdetjønn and Bjønntjønn were made about

50/50 by the participants. All the on-site fish registrations for Nedre Urdetjønn were made by the field assistant, as I returned to Bø to place the catch from Øvre Urdetjønn in a deep freezer.

With the handheld devices for water measurements, temperatures were slow to instill, and the pH measurements varied so much and quickly as to seem unreliable. In Bjønntjønn, some of the measurements with these devices were not obtained or even disregarded, as we assumed we would later get more reliable readings of the same variables from the multiparameter logger.

Field data on fish and water were annotated first by hand, and later handled and stored in spreadsheets and csv files through MSO 365 Excel Version 2206. .

2.3 Gillnets and depth profiles

When transferring and plotting the GPS coordinates after fieldwork, it became apparent that a few gillnet coordinates especially for Øvre Urdetjønn were missing or faulty and thus had to be estimated (Table A1 3.1).

A post-field look at the recorded depth data supported the on-site indications that several of our gillnets in 2022 probably crossed the standard 3-meter depth stratums, thus deviating from assumptions in NS-EN 14757 and annotations in the raw data file from the 2009 test fishery (where no crossing of depth strata was indicated). Likewise, it was substantiated that many of our gillnets were likely placed partly or entirely at depths shallower than the height of the gillnet (1.5 m), unavoidably so in the shallow Bjønntjønn. This might have been the case also in 2009, but this was not possible to ascertain from the available information.

Based on measured water depths at gillnet ends and at locations for water sampling, I estimated the fishing depth of each gillnet and of the respective panels in each of them. Here, I assumed that the lake bottom had an approximately linear decline from a measured shallower to a deeper point (examples in Figs. A2 2.1-2.3), and that each gillnet was set in a tight, straight line (not curved) from shore outwards, fishing vertically in the water column.

For the rough lake depth profiles, I connected where possible measured depth points with straight lines, thus constructing approximate isodepths visualized on maps presented in the main text (Figs. 3.1-3.3). Where relevant depth measurements were lacking, I made even more uncertain projections with straight lines, marked with different dotting/color on the map, and reflected in the map legend.

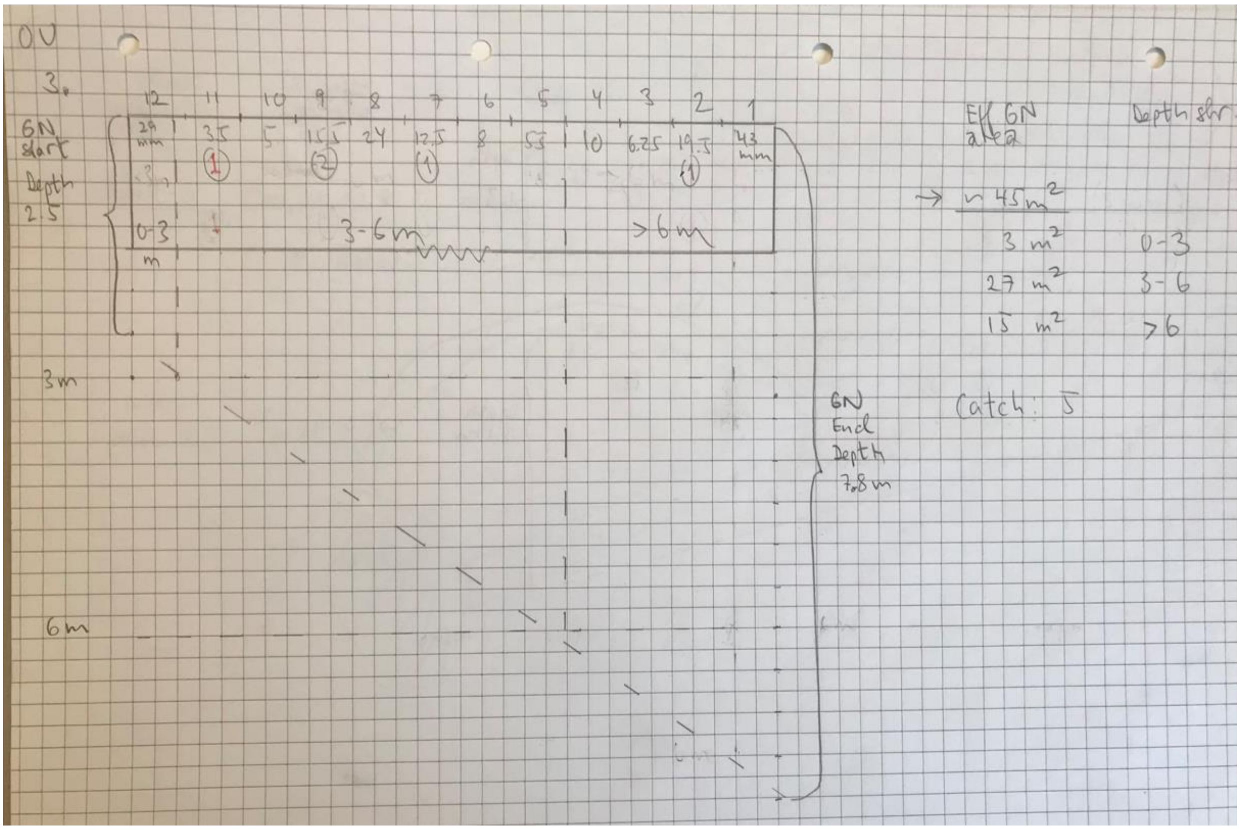


Figure A2 2.1. Øvre Urdetjønn: sketch of estimated depths and catch per panel in gillnet #3 in the 2022 test fishery. All the gillnets were sketched in the same way. The information was used to calculate catches per depth stratum, reflecting also the assumption of gradual, evenly decreasing depth between gillnet ends used to construct the rough depth profiles per lake in the present study.

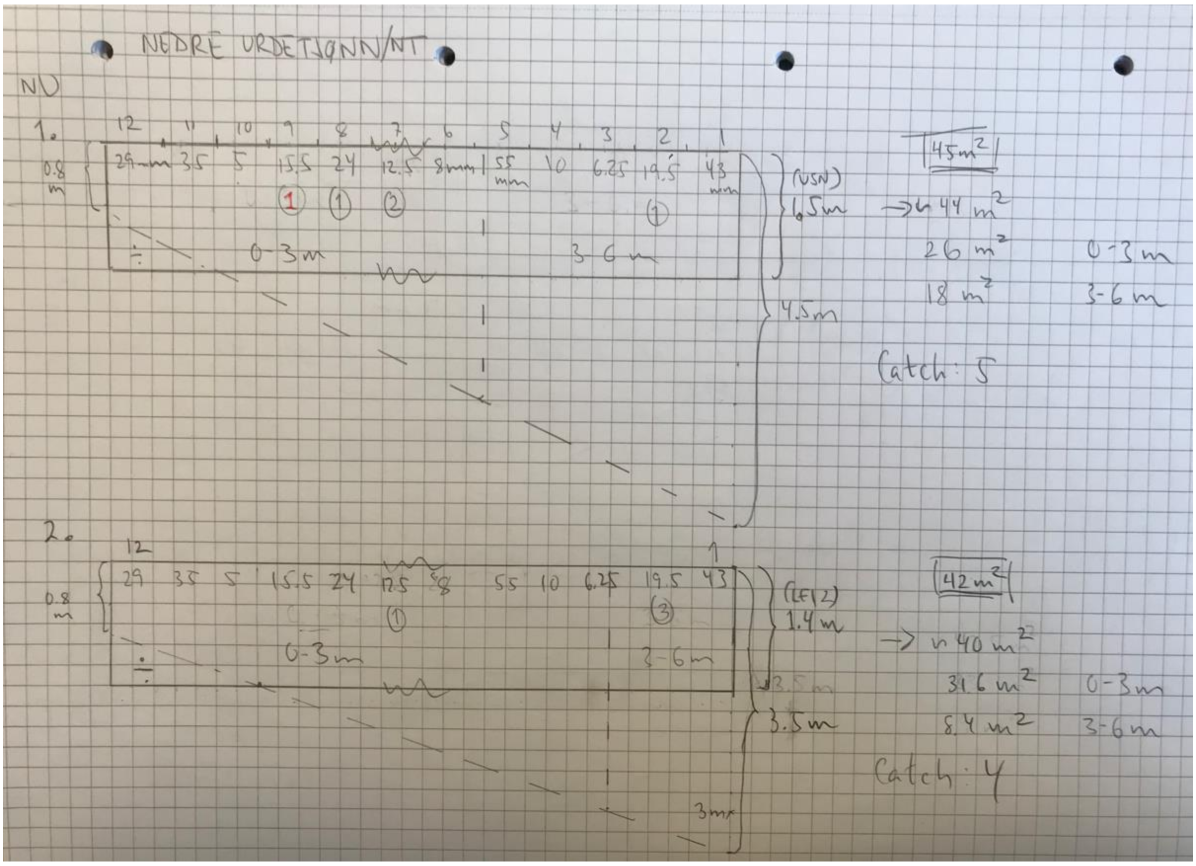


Figure A2 2.2. Nedre Urdetjønn: sketch of estimated depths and catch per panel in gillnets #1 and #2 in the 2022 test fishery. All the gillnets were sketched in the same way. The information was used to calculate catches per depth stratum, reflecting also the assumption of gradual, evenly decreasing depth between gillnet ends used to construct the rough depth profiles per lake in the present study.

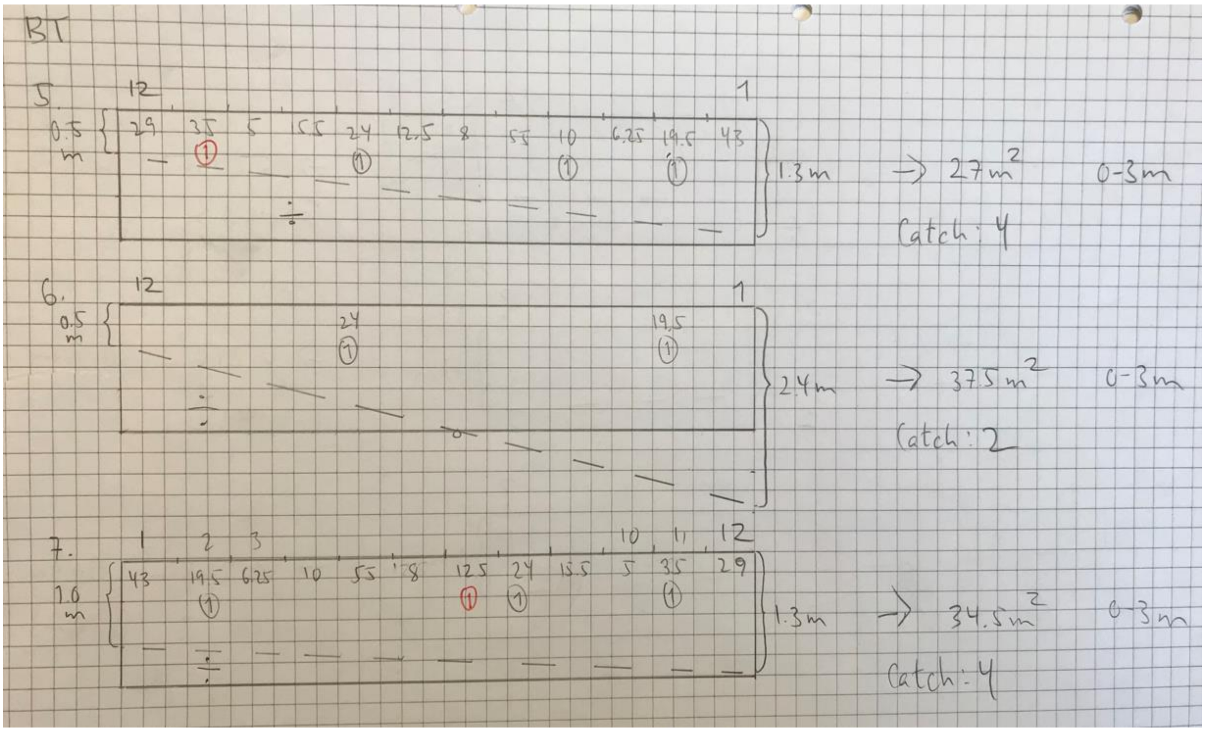
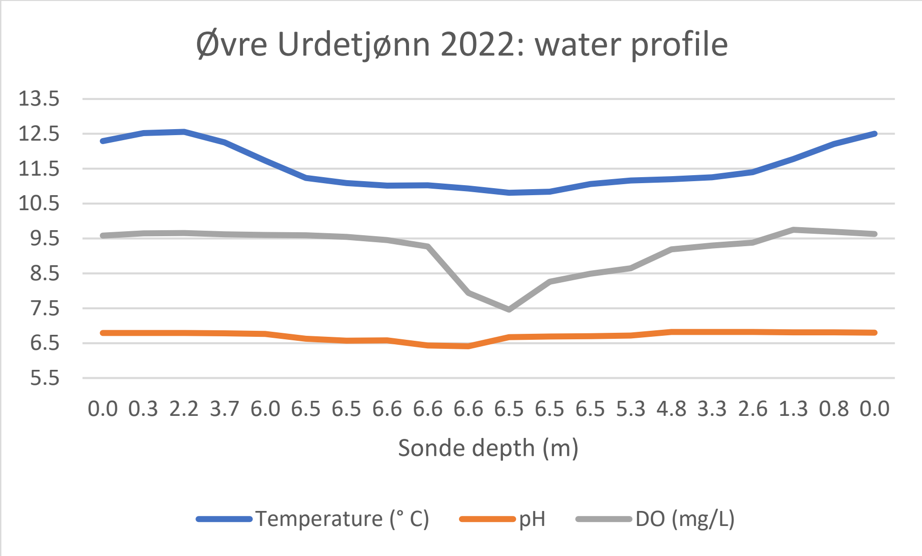
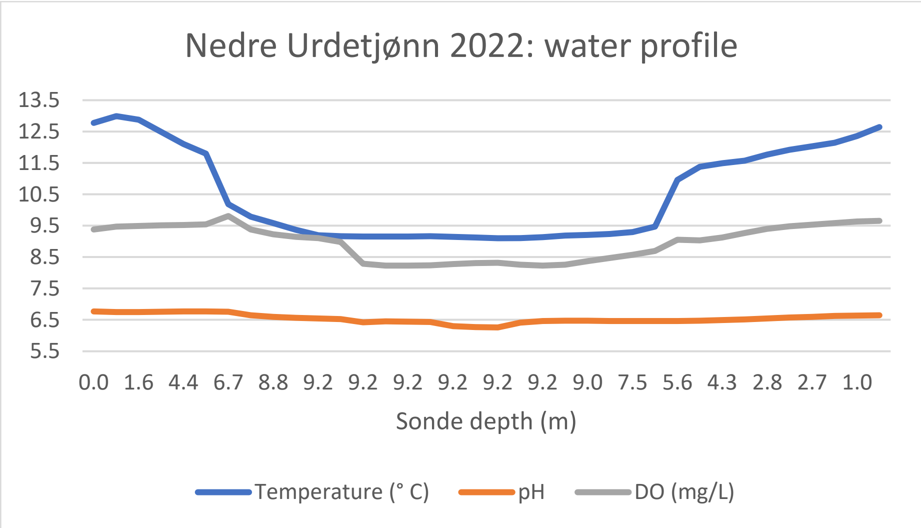


Figure A2 2.3. Bjønntjønn: sketch of estimated depths and catch per panel in gillnets #5, #6 and #7 in the 2022 test fishery. All the gillnets were sketched in the same way. The information was used to calculate catches per depth stratum, reflecting also the assumption of gradual, evenly decreasing depth between gillnet ends used to construct the rough depth profiles per lake in the present study.

2.4 Water measurements with multiparameter logger

The main operational problem with the multiparameter logger reflected in the data was that the sensors had evidently not had a stable, directly vertical descent through the water column, probably due to drifting (especially in L. Bjønntjønn). The result was that much of the data for each lake were from similar depths (near the surface and lakebed), while large parts of the water column were without measurements (Fig. A2 2.4). A general puzzle for all three lakes was the results on conductivity. The output file contained conductivity data in µS/cm, both as absolute values and corrected for temperature. I selected the latter for consideration; however, the values were 5 – 10 times the magnitude of µS/cm values both in our readings with handheld devices and in the reported results from the 2009 test fishery. Despite my best efforts, I was unable solve the puzzle, but problems with calibration, sensors and/or non-harmonized units/decimal errors were key hypotheses. In the end, I decided to exclude these conductivity data from the analyses but listed the recorded ranges in the main text, for reference.





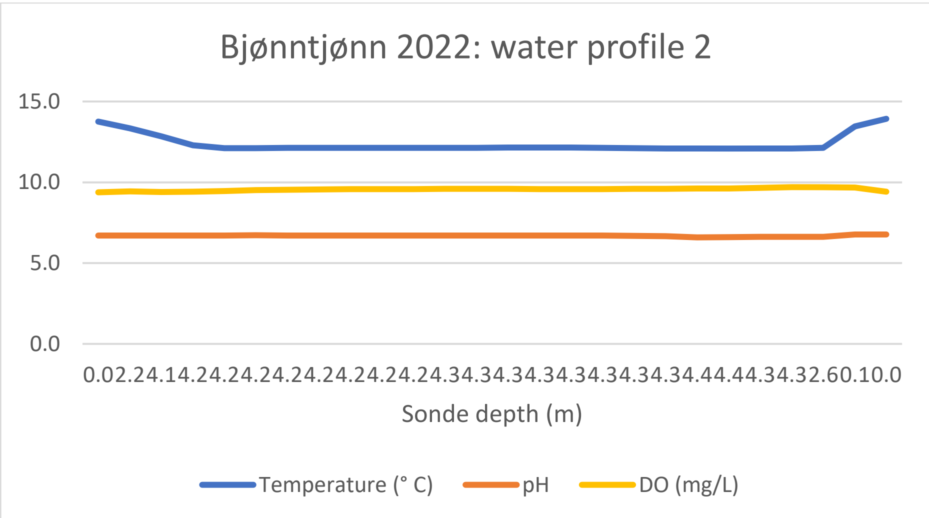


Figure A2 2.4. Three Upper Kova lakes: three water variables plotted against depth in the 2022 test fishery. Input: complete recorded data sequence (excluding several measurements at 0.0 meters at each end), measured with multiparameter logger at location of assumed maximum lake depth, on June 22, 2022. DO: dissolved oxygen.

2.5 Laboratory work (general)

I had no previous experience in the data collection from individual fish to be done in the laboratory. No timely systematic training was available, but I got general tips from a co-student and was shown a procedure for burning and breaking otoliths by a student assistant.

Since low catch in Bjønntjønn would limit the possibilities of analyses anyway (especially with sub-divisions), I started with fish from this lake to gain experience and improve my procedures and evaluation schemes before moving on to the fish from the other two lakes. Throughout the lab work, I documented my estimated level of uncertainty for the data assessments made, summarized below (Table A2 2.1). Among the first fish I collected data from, some structures had started to decompose (skin and internal organs), probably due to an unfortunately long thawing period combined with small body size. As this appeared to make determination of especially sex/maturation stage (possibly a bit more challenging on small fish early in the season, as well?) and stomach filling more difficult, I decided to record the preservation status of each individual examined, as a possible source of error and/or contributor to lower data quality in some cases.

Lack of experience also influenced my use of predefined classification schemes. Starting out, I had to keep the possibility open that a more "extreme" specimen might turn up, so I was hesitant about using especially the top category. Re-classification after the whole exercise was not always feasible; in the end, I did not use some of the top categories neither for maturation stage (value V/5) nor for stomach filling (value 5).

Table A2 2.1. Three Upper Kova lakes: self-evaluated quality of data collected in the lab in the 2022 test fishery. Values are means. Level of uncertainty was rated on a scale of 0 – 2, where 0 = low, 2 = high, except for preservation status, where the scale was reversed, based on intactness of exterior and interior structures (0 = both intact, 2 = both aspects less than ideal). ØU: Øvre Urdetjønn, NU: Nedre Urdetjønn, BT: Bjønntjønn. Corrected data: used in comparison with test fisheries in 2009 and 1997. Mean = mean of means.

		Assessed per individual (0-2)				General/retrospective (0-2)		
Lake	**Data**	**Age**	**Sex**	**Maturity**	**Preservation**	**Flesh color**	**Stomach filling**	**Mean**
ØU	Corrected	1.3	0.3	0.1	0.2	0.5	0.5	0.5
ØU	Complete	1.3	0.3	0.0	0.2	0.5	0.5	0.5
NU	Corrected	1.4	0.1	0.1	0.2	0.5	0.5	0.5
NU	Complete	1.4	0.2	0.1	0.1	0.5	0.5	0.5
BT	Corrected	1.6	1.5	1.2	1.0	1.0	1.0	1.2
BT	Complete	1.6	1.5	1.3	1.0	1.0	1.0	1.2
Mean		1.4	0.6	0.5	0.5	0.7	0.7	0.7

2.6 Individual fish length and weight re-measured

Individual fish lengths and weights measured in the laboratory were on average ca. 2% and 5% lower than the field measurements (Table A2 2.2). The reason why the latter were used in all analyses was that I found it likely that the deviances was more due to freezing and thawing effects than measurement error or bias in the field due to equipment or use thereof.

In the re-measurement process, it became apparent that the length of one individual caught in Nedre Urdetjønn must have been incorrectly registered in the field notes (38 cm versus my control in the lab: 315 mm). In all analyses, I used the laboratory-measured length for this individual.

Table A2 2.2. Three Upper Kova lakes: mean differences between field and (lower) laboratory measurements of length and weight per individual in the 2022 test fishery. Corrected data: used in comparison with test fisheries in 2009 and 1997.

		Assessed per individual	
Lake	**Data**	**Length difference (%)**	**Weight difference (%)**
Øvre Urdetjønn	Corrected	1.7	3.3
Øvre Urdetjønn	Complete	1.6	3.1
Nedre Urdetjønn	Corrected	1.4	5.1
Nedre Urdetjønn	Complete	1.3	5.5
Bjønntjønn	Corrected	3.5	6.1
Bjønntjønn	Complete	3.3	5.3
Mean of means		2.1	4.7

2.7 Otoliths and age analysis

Otoliths were extracted with a success rate of over 99 % for Øvre- and Nedre Urdetjønn and somewhat less for Bjønntjønn. After some experience on the materials from Bjønntjønn, I established a routine for the other two lakes whereby both unmanipulated otoliths from each fish were inspected, then I burnt and cut one of them for each fish for further scrutiny. On average, I kept 3-4 otolith photos for each fish, later selecting the most promising one (in most cases a burnt one) for formal age analysis and documentation.

Considering possible improvement of my data quality on fish age, I thought about examining other available material (still unburnt otoliths, or scales, which had been collected as a possible backup) and/or or reanalysis of already assessed materials. Ideally, this could be set up as blinded (also for body lengths) intra- and/or interrater exercises with calculation of Cohen's kappa (Burghardt et al., 2012). Unfortunately, I had to drop this for capacity reasons.

As an alternative to renewed or more age analyses, I considered using only the age data that I was least uncertain about. For Øvre Urdetjønn, this would have had a limited effect on the overall distribution pattern (Fig. A2 2.5). For Nedre Urdetjønn, the resulting age distribution seemed intuitively not to be representative for the younger fish (Fig. A2 2.6), so I dropped this option primarily for this reason.

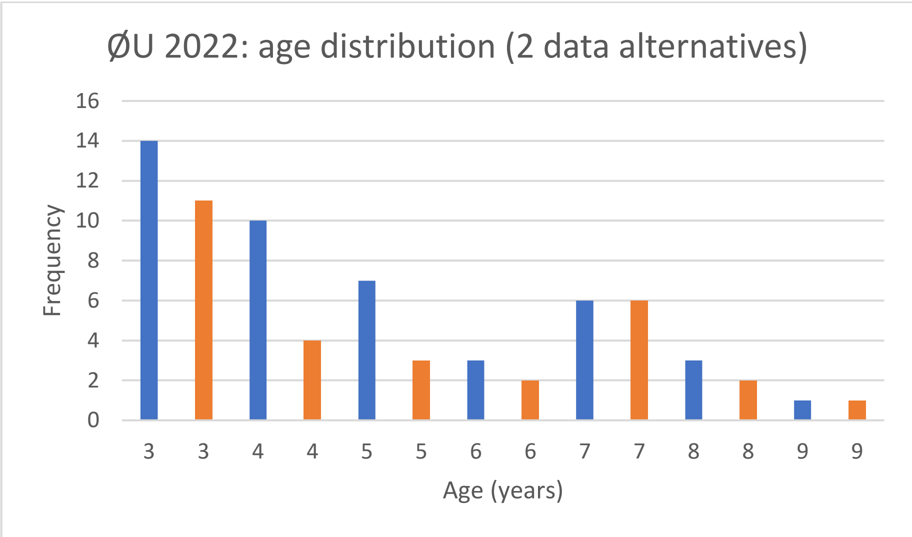


Figure A2 2.5. Øvre Urdetjønn (ØU): frequency distribution of fish age based on the 2022 test fishery, with 2 data alternatives: blue bars = all data (n=44, corrected for comparisons with 2009 and 1997) and orange bars (n=29) = excluding from these 44 the individuals with the most uncertain age estimates.

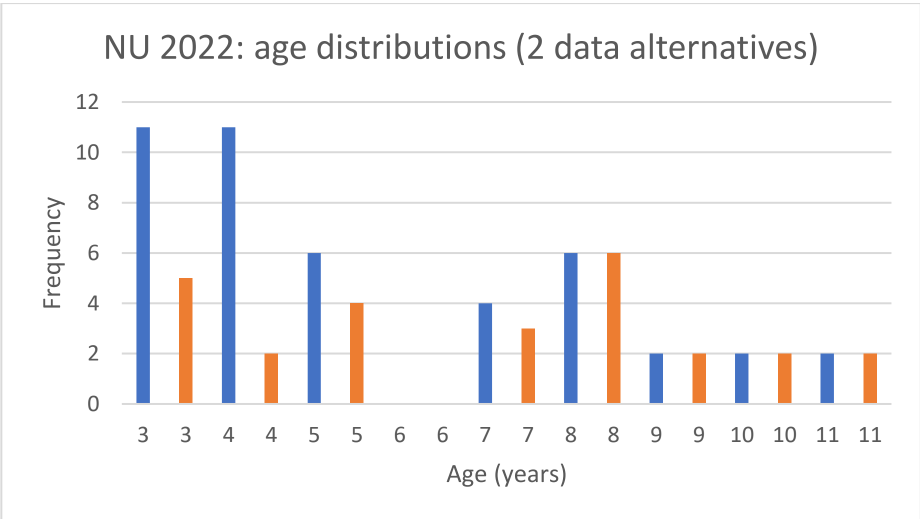


Figure A2 2.6. Nedre Urdetjønn (NU): frequency distribution of fish age based on the 2022 test fishery, with 2 alternatives: blue bars = all data (n=44, corrected for comparisons with 2009 and 1997) and orange bars (n=26) = excluding from these 44 the individuals with most uncertain age estimates.

2.8 Sex and sexual maturity

Based on challenges in determining sex and maturation stage for the first batch of small, mostly poorly preserved individuals from Bjønntjønn, I tried taking photos of head/snout and anal fin of each fish to see if they could be used as controls for my later sex assessments. As it turned out, the match between my photo assessments and fish I had classified for sex based on clearly visible gonads was not good enough to be helpful in the most challenging cases, so I discarded this photo method altogether.

2.9 Flesh color

From Bjønntjønn, I initially placed one individual with what I would normally call orange or mid-level color in the category light red, and the rest of the catch in the lake in the category white. As it turned out, no other fish in the other two lakes came close to this strength of red flesh pigmentation, but some of the larger individuals in these lakes had a slight pink hue. Based on this, I retrospectively reclassified the flesh color of all the individuals for which I had retained relevant information. The mentioned individual from Bjønntjønn was thus classified as having red flesh color, and all individuals with a slight pink hue as having light red flesh color.

3 Data analysis and displays

3.1 General concepts and considerations

Possible normal distribution was assessed through q-q plots, histograms, similarity of central tendency values, and skewness (± 2), as well as by formal testing (Shapiro-Wilk). Homoscedasticity was assumed when the ratio for the largest/smallest standard deviation was less than 2.5. When square root transformation was attempted, I used a formula from Whitlock and Schluter's textbook on statistics in biology (2020):

$$Y' = \sqrt{Y + 0.5} \qquad \text{(A2 3.1)}$$

To assess whether certain data points were formal outliers, I used the definition of "extreme" values in Whitlock and Schluter's textbook (2020): those lying more than 1.5 times the interquartile range from the interquartile bounds or box edges in a boxplot.

3.2 Estimates based on catch per lake, gillnet and mesh size

The formal stratum-based weighting procedure outlined in NS-EN 14757 was not used, as our randomization and depth-related data seemed inappropriate for this mode of analysis. The alternative I chose instead is outlined in the main text (Results, 4.2).

3.3 Comparisons with test fisheries using single-mesh Jensen series

I found the guidelines of Ugedal and colleagues (2005) for calculation of CPUE (catch per 100m² relevant gillnet area per night) as the basis of main density classes quite clear in that catch in mesh sizes < 15.5 mm (typically catching fish < 150 mm) are not relevant in comparisons with standard Jensen series and should thus be excluded. However, it was not clear to me how fish deviating from modal size should be handled. In the end I excluded all fish ≤ 150 mm *and* all fish caught in panels with mesh size < 15.5 mm (even if actual size was > 150 mm). The number of excluded fish > 150 mm in panels with mesh size < 15.5 mm from the catches in 2022 were 6 in Øvre Urdetjønn, 7 I Nedre Urdetjønn, and 2 in Bjønntjønn. The resultant total relevant gillnet areas were 210 m^2 (Øvre Urdetjønn and Bjønntjønn) and 201 m^2 (Nedre Urdetjønn) in the 2022 test fishery. The rules for comparison of multi-mesh catches with catches *per Jensen series* (in numbers and weight) (Table 5.1, middle/right columns) were less clear to me; after expert consultation and to avoid inappropriate extrapolation, I used the full, uncorrected gillnet area of the eight Nordic multi-mesh gillnets used per lake to calculate these latter numbers for 2022.

3.4 Correction for the timing of the 2022 test fishery

I assumed that the early timing of the 2022 test fishery meant that the lengths and weights of most of the caught fish were lower than they would have been at the end of the growth season. On this basis, I hypothesized that some or even most of the lower fish lengths, weights and k-factors in the 2022 test fishery compared to the 2009 and 1997 surveys (Table 5.3) might be attributed to the timing of the fieldworks. The 2022 test fishery was probably undertaken near the beginning of the growth season and the 2009 and 1997 occasions towards or even after the end. In my attempt to project what weight and length each caught individual in 2022 might have had at the end of the growth season (if not caught by us or dead by some other reason), a basic premise was that most or all annual growth occurs during the growth season. I assumed that interannual variation may be somewhat evened out over a ten-year period (i.e., the approximate potential lifespan of brown trout in the Kova lakes), and that the 2022 data on fish length, weight, and age could be used to make projections if one assumed that each individual had a growth rate not too different from average values. My procedure was to first calculate mean length and weight for each age class, then a factor between two consecutive age classes by dividing the mean of the latter on the mean of the former, and, finally, to multiply each individual's measured length and weight with the corresponding factor. To estimate how much of the difference in weight and k-factor might be "explained by" the timing of fieldwork according to this model, I divided the difference between the "observed" and projected values for 2022 on the "observed" difference between the latter and the previous test fisheries.

The method did not lead to the results I initially anticipated, especially for k-factor but partly also for weight. Apart from the possibility that the reasoning behind the model was faulty, I assume that my somewhat uncertain age estimates could have influenced the results. Lower predicted k-factor than expected could technically be explained by underestimation of weight or overestimation of length or a combination of the two. Trying to adjust the model, I wanted to stick to empirical data and avoid extrapolation. Thus, I found it better to reduce the length input (by assuming arbitrarily that 25% of the expected length gain per season had already been achieved by mid-June) than increasing the weight input by some factor related to an unknown influence external to the model.

On the other hand, I find it likely that one or more such "external" influences could be part of the explanation for the gap between projections for 2022 and the previous Kova test fisheries. Since all the input to my model was based on mid-June weights, it did for instance not account for possible influences on weight and k-factor by gonadal development in mature fish towards the end of the season. I assume that especially mature females on average would experience a net weight gain due to "inflated" gonads before spawning. Interestingly, in mid-June and based on the 2022 data corrected for comparison with the previous test fisheries, mean k-factor was significantly *lower* for mature than for immature females in both Øvre Urdetjønn and Nedre Urdetjønn, tested with two-sample Wilcoxon (Øvre Urdetjønn: W = 162, $P_{0.05}$ = 0.007, n = 12/17; Nedre Urdetjønn: W = 75, $P_{0.05}$ = 0.02, n = 13/7). If this difference had to do with maturity status, I imagine it could be related to the energetic priority given to gonadal development vis-à-vis general body growth, and that the gonads in mid-June had not gained enough weight yet to "compensate" for this (perhaps due to a lower water content, for instance). On the other hand, the mature females also had higher average length than immatures. As k-factor decreased significantly with length in both of these lakes based on the 2022 data, length in itself could also be a factor in the different k-factors of mature and immature females. In either case, to meet my assumption of net weight gain due to gonads on average at the end of growth season, the rate of gain for mature females would have to be higher through the season than for the immatures.

When I decided to make some use of my prediction model despite its probable weaknesses, it was because the basic input was data from the lakes (reflecting processes working within a time range of about a decade) and because the results seemed to be illustrative of an important point and to be at least within plausible ranges.

3.5 Nomenclature for fish age

In the report from the test fishery in 2009, it was stated in the section on age distribution that fish age was "rounded up" based on the reasoning that the growth season was over when they were caught (3+ was for instance handled as 4 years old) (2010). I assumed this also applied to the growth analyses in the same study. The results in the report from the 1997 test fishery made most sense to me if the same approach had been applied there as well, so I assumed that this was the case. As this type of "rounding up" is unconventional (J. Heggenes, personal communication, 20 February 2023), I converted all age references from the 2009 and 1997 test fisheries to the nomenclature I used for the 2022 results.

References in Appendix 2

Burghardt, G. M., Bartmess-LeVasseur, J. N., Browning, S. A., Morrison, K. E., Stec, C. L., Zachau, C. E., & Freeberg, T. M. (2012). Perspectives - Minimizing Observer Bias in Behavioral Studies: A Review and Recommendations: Minimizing Observer Bias in Behavioral Research. *Ethology*, *118*(6), 511-517. https://doi.org/10.1111/j.1439-0310.2012.02040.x

Dyrevelferdsloven. (2009). *Lov om dyrevelferd*. (LOV-2009-06-19-97). Lovdata. Retrieved from https://lovdata.no/lov/2009-06-19-97

Forskrift om bruk av dyr i forsøk. (2015). *Forskrift om bruk av dyr i forsøk*. (FOR-2015-06-18-761). Lovdata. Retrieved from https://lovdata.no/forskrift/2015-06-18-761

Russel, W. M. S., & Burch, R. L. (1959). *The Principles of Humane Experimental Technique*. Johns Hopkins, Bloomberg School of Public Health. https://caat.jhsph.edu/principles/the-principles-of-humane-experimental-technique

List of figures and tables in Appendix 2

Figures

Tables