Canadian Journal of Fisheries and Aquatic Sciences

OPEN ACCESS | Research Article

Bayesian species recognition and abundance estimation: unravelling the mysteries of salmonid migration in the Teno River

Antti Räty[a,b], **Henni Pulkkinen**[a], **Jaakko Erkinaro**[a], **Panu Orell**[a], **Morten Falkegård**[c], **and Samu Mäntyniemi**[d]

[a]Natural Resources Institute Finland, Paavo Havaksen tie 3, PO Box 413, 90014 Oulu, Finland; [b]Research Unit of Mathematical Sciences, University of Oulu, 90014 Oulu, Finland; [c]Norwegian Institute for Nature Research, Fram Centre, PO Box 6606, Langnes, NO-9296 Tromsø, Norway; [d]Natural Resources Institute Finland, Latokartanonkaari 9, 00790 Helsinki, Finland

Corresponding author: **Antti Räty** (email: antti.raty@luke.fi)

Abstract

In Teno River, annual sonar monitoring is used to estimate the abundance of three salmonid species: Atlantic salmon, pink salmon, and sea trout. However, the size distribution of these species is partially overlapping making species recognition impossible from plain sonar data. A Bayesian model was developed to tackle this problem and to estimate abundance and migration timing for these three species. The model integrates multiple sources of data including catch, video count, daily average school sizes, and expert knowledge. Given the limited catch and video statistics for 2021, the use of school size data and expert knowledge on migration intensity enhanced the estimation when other data sources were unavailable. The model estimated a median of 11.8 thousand Atlantic salmon, 6.6 thousand sea trout, and 52.0 thousand pink salmon migrating into the river during 2021. These findings offer a more accurate representation of species distribution, support future conservation and management efforts, and provide a modelling-based solution for distinguishing similarly sized species from sonar counting data.

Key words: Atlantic salmon, pink salmon, sonar monitoring, Bayesian modelling, Teno River

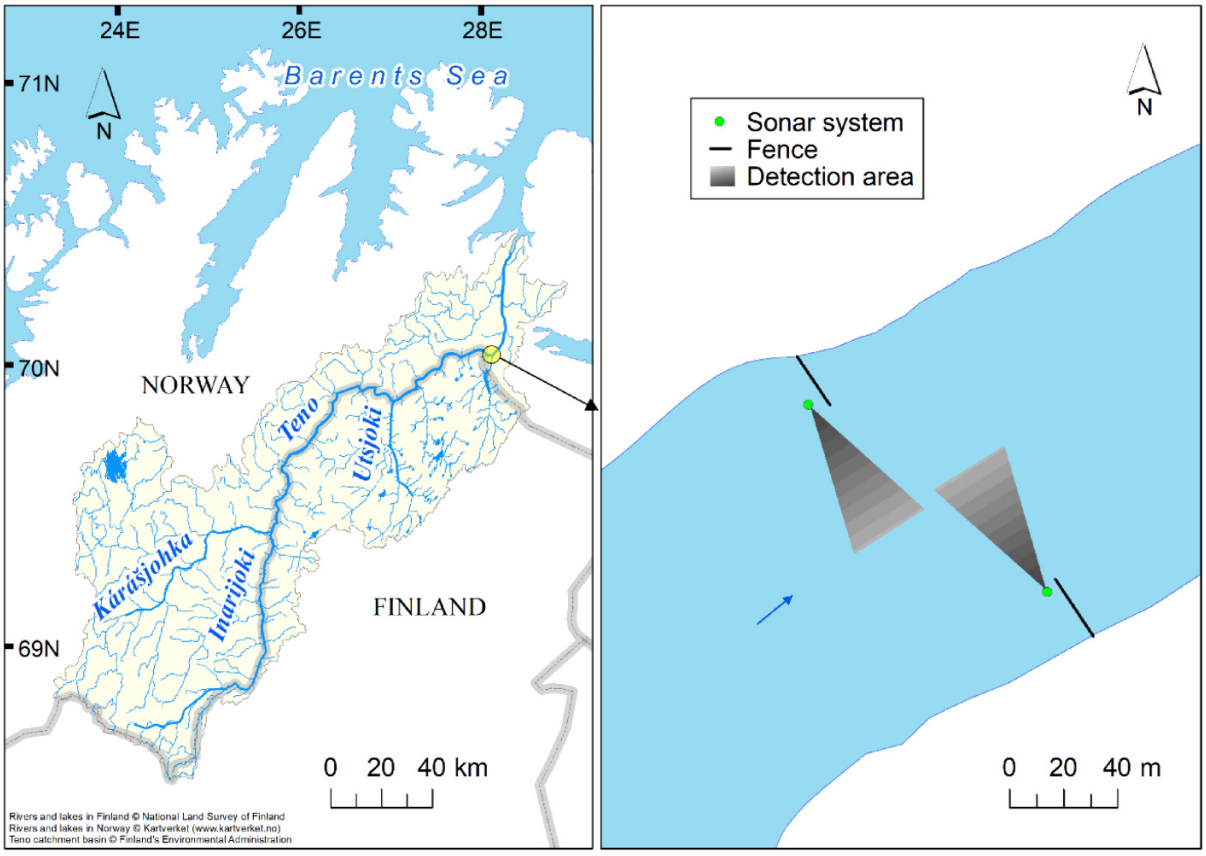
Introduction

Management of migratory fish populations is typically based on research and monitoring programmes that provide scientific information about the characteristics and status of the stocks (Beddington et al. 2007), and form the basis for scientific advice for management (Forseth et al. 2013). Monitoring programmes should ideally be based on long-term data collection on dynamics of the spawning stock size and subsequent recruitment.

Management strategies based on biological reference points are typically used for migratory fish such as Atlantic salmon (*Salmo salar*) (Prévost and Chaput 2001). The precautionary approach in a science- and target-based management is the principle of the internationally agreed management system for Atlantic salmon in an intergovernmental convention, established by the North Atlantic Salmon Conservation Organization (NASCO 1998, 2009). Biological reference points used in assessment of salmon stock status can be defined as conservation limits (CL) or spawning targets (NASCO 2009; Forseth et al. 2013). Traditionally indices for stock status have often been derived from catch data, but in recent decades, other and non-intrusive methods have been increasingly at use. In small- and medium-sized rivers, applicable monitoring methods include several alternatives, e.g., optical methods such as snorkelling (Orell and Erkinaro 2007) and video surveillance (Orell et al. 2007*b*) but population census in larger rivers pose particular challenges. One of the monitoring methods proven practical and successful in large rivers is sonar counting (Boswell et al. 2008; Helminen et al. 2020) that produces a time series of numbers of fish passing the sonar site, data on fish length, swimming direction, and distance from the transducer.

Sonar counting is an active are area of development and research, with significant advancements such as automatic fish detection (Helminen and Linnansaari 2021) improving its efficiency. However, one of the main challenges remains accurately recognising different fish species (Horne 2008; Wei et al. 2022). In some cases, species recognition can be done based on significant length difference between fish species (Burwen et al. 2010). However, migrations may involve multiple species with strongly overlapping size ranges. In such cases, other information sources must be used for estimation of species distribution, such as good-quality catch statistics (Fox and Starr 1996; Alglave et al. 2022) or video surveillance (Lamberg and Imsland 2022). Since the goal is to make probability statements about unobserved quantities (true species distribution) based on observed quantities (sonar data), the problem can be identified as a natural case of Bayesian infer-

Fig. 1. The Teno system in northernmost Europe (left), with the Polmak sonar counting site marked by a yellow circle. The right panel illustrates the sonar setup used at the site.



ence (O'Hagan 1994). In contrast, the classical frequency approach is bound to assess how potential new estimates would vary when some known values for the parameters have first been assumed (Pierce 1973; Romakkaniemi 2015), which is not of direct interest when assessing fish populations based on observed data (Mäntyniemi et al. 2005; Kuparinen et al. 2012; Romakkaniemi 2015). Bayesian modelling is able to combine multiple sources of information and expert knowledge in the same modelling framework, simultaneously providing proper estimates of uncertainty and a better understanding of the phenomenon as a whole (Buckland et al. 2007; Kuparinen et al. 2012).

In the large Teno River in the northernmost Europe (Erkinaro et al. 2019), three anadromous salmonid species of relatively similar size range, Atlantic salmon, pink salmon (*Oncorhynchus gorbuscha*), and sea trout (*Salmo trutta*), ascend during the short subarctic summer period between June and August. Status assessment of the multiple Atlantic salmon populations (Anon 2024) requires reliable estimation of run sizes, and in recent odd years, the increasing numbers of the alien pink salmon (Erkinaro et al. 2022) have severely complicated this task. So far, species composition from sonar data has been roughly estimated using catch information from areas near the sonar site, relating the catches of different species with sonar data on a daily basis, and no quantitative modelling has been used before. Since the salmon fishing ban in 2021, catch data have become scarce (some small-scale experimental catches only) or non-existing.

In this study, we developed a Bayesian model to estimate the distribution of fish species ascending the Teno River. Our model uses sonar data, catch statistics, video data, school size data, and expert knowledge of migratory patterns, resulting in species-specific daily and total abundance estimates of the fish passing sonar counting site. The estimates have proper quantification of uncertainty and provide a better understanding of the total abundances of different species.

Materials and methods

Study area

The subarctic Teno River (Tana in Norwegian, Deatnu in Sámi) forms the border between northern Norway and Finland being one of the largest salmon rivers in both countries (Fig. 1). The river supports wide diversity of Atlantic salmon populations with c. 30 genetically distinct stocks (Vähä et al. 2017) and extremely high diversity of life histories (Erkinaro et al. 2019). The main stem, large headwater branches, and smaller tributaries comprise > 1100 km of accessible stretches for ascending salmon. The Teno River is one of the last remaining large salmon rivers with a virtually pristine and abundant distribution area for the Atlantic salmon,

largely free from major man-made obstructions (Erkinaro et al. 2017).

Until recent years, Teno River has been known for its massive wild Atlantic salmon production with annual river catches varying between 80 and 250 metric tons or 20 000 and 60 000 fish (Erkinaro et al. 2019). Unfortunately, a steep decline in salmon stock status has been documented in recent years, especially since 2019. The cause for this decline is believed to be a sum of multiple factors, including climate change, which has the most pronounced effect on higher latitudes, changing the environmental conditions for Atlantic salmon during its life stages in freshwater and at sea (Alioravainen et al. 2023). Strong management measures have been implemented to facilitate salmon stock recovery by decreasing fishing pressure, and eventually, a complete fishing ban has been issued for 2021–2024 (Anon 2024).

The invasive pink salmon has recently started to reproduce in the Teno system (Erkinaro et al. 2022, 2024). The pink salmon was introduced to the Barents Sea–White Sea region through a Russian introduction programme between 1956 and 2001 (Sandlund et al. 2019). The number of pink salmon entered and captured in the Teno in odd years has been mostly low since their introduction to Northwestern Russia, but in 2017, numbers started to strongly increase: approximately 4600, 50 000, and 170 000 pink salmon were estimated to migrate into the Teno system in 2019, 2021, and 2023, respectively (Erkinaro and Orell 2022; Domaas et al. 2024).

Sonar counting of ascending fish in the Teno main stem has been carried out since 2018 when the Natural Resources Institute Finland (LUKE) installed two echo sounders 55 km from the Teno estuary (Anon 2024) (Fig. 1). Three similarly sized salmonids are ascending the river with overlapping migratory patterns: one-sea-winter (1SW) Atlantic salmon, sea trout, and the pink salmon (Erkinaro et al. 2022).

Data

In this study, we utilized sonar data from the Teno River in 2021. The data collection was conducted with two ARIS explorer 1200 (Sound Metrics, Bellewue, WA, USA) multibeam sonar units (see Fig. 1) equipped with ARIS telephoto lenses (Sound Metrics, Bellewue, WA, USA) and 14 °C spreader lenses. Sonar data were recorded and stored to external hard drives by using ARIScope software (version 2.8.0) and collected sonar data were analysed by using ARISFish software (version 2.8.0). Four underwater custom-made (Lamberg Bio-Marin Service) video cameras with 3.5 mm wide angle lenses were installed to the sonar counting site for species identification purposes. Camera data were both recorded and analysed with Timespace X300 digital video recorder (Timespace Technology Ltd, Huntingdon, UK). All sonar and camera data analysis were performed by experienced personnel.

Final sonar data were generated by analysing raw sonar data from two echo sounders, which requires lots of manual labour because all fish observations are detected and measured manually. Analysing daily raw data from both echo sounders completely would require an immense number of working hours. To ease the workload, 12 h of raw sonar data were analysed from both echo sounders each day.

The analysed raw data comprised 30 + 30 min alternating between two sonars each hour. Based on the analysed hours, the complete daily number of fish passing the sonar counting site was estimated by extrapolating fish counts and measurements over the unanalysed hours, assuming the same volume of migration. This resulted in a daily time series of the number of fish passing the sonar counting site, grouped by swimming direction and length classes. All downstream migrants were considered kelts before 26 June, which were not subtracted from upstream counts. After this date, the subtracted proportion of downstream migrants was increased stepwise over 9 days from 10% to 100%. These proportions were based on earlier catch information (Niemelä et al. 2000).

Since the vast majority of ascending pink salmon and sea trout belong to the body length range of 45–65 cm, the size group of interest in Atlantic salmon is <65 cm, which are mostly 1SW fish (Alioravainen et al. 2023). Ascending fish that exceed 65 cm in body length were considered multi-sea-winter Atlantic salmon, and fish shorter than 45 cm consist mostly of other local species, like grayling (*Thymallus thymallus*) and whitefish (*Coregonus lavaretus*). The resulting sonar data comprised 70 350 observations from ascending fish between 45 cm and 65 cm in length, observed between 27 May and 1 September 2021 (Fig. 2C).

Statistics concerning fish school sizes and the number of schools were calculated from the analysed hours of raw sonar data. Two or more fish were assumed to belong to the same school if the horizontal distance between fish was less than 15 m and the time difference between two consecutive detections of individual fish was less than or equal to 1 s. The daily number of schools were calculated by dividing the daily fish count by the daily average school size (Fig. 2D).

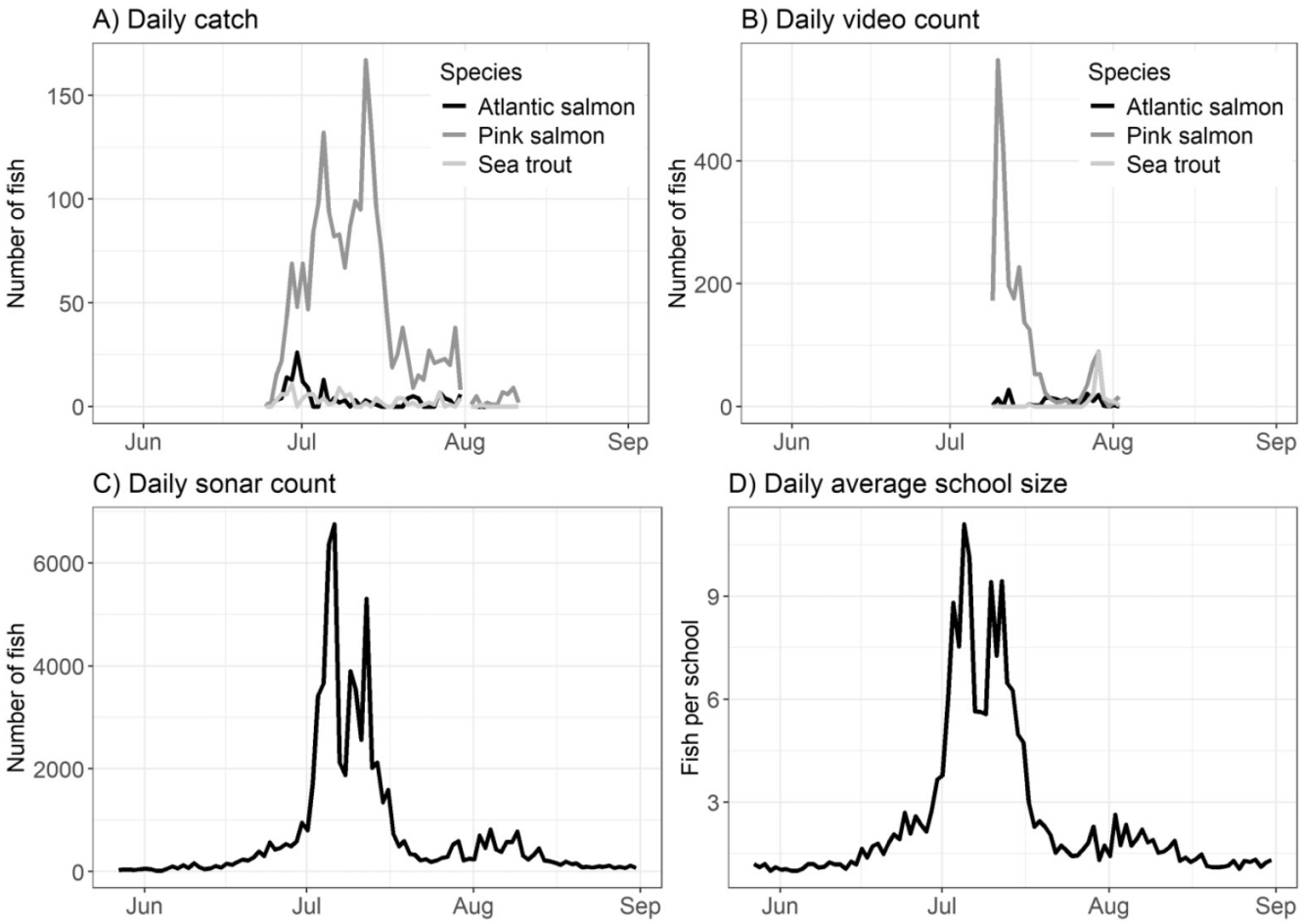
In 2021, a salmon fishing ban was introduced for the Teno River, and therefore catch data from regular fisheries were not available. However, small-scale special fishing was allowed in the lower part of the river for collecting information on pink salmon migration and their catches in relation to Atlantic salmon and sea trout. Data were selected to include Atlantic salmon, pink salmon, and sea trout catches of individuals between 45 and 65 cm in length from six fishing sites between the Teno estuary and the Polmak sonar counting site. The resulting catch data included 2345 observations from 48 days between 24 June and 11 August 2021 (Fig. 2A).

Video observations of ascending fish were gathered for 25 days between 9 July and 2 August 2021, using four video cameras installed at the southern side of the river at the Polmak sonar counting site, covering c. 10 m of the transect and 20% of the sonar window length. A total of 2779 Atlantic salmon, pink salmon, and sea trout between 45 and 65 cm in length were observed. These video observations were backed by simultaneous sonar observations of the same schools and individuals, allowing reliable species identification. The time frame for the video surveillance was selected so that large numbers of pink salmon could still be observed (Fig. 2B).

Probability model of the problem

The model assumes three species in the sonar data: Atlantic salmon, pink salmon, and sea trout. The relative share be-

Fig. 2. Daily observations of salmonids between 45 and 65 cm in length in 2021. (A) Number of Atlantic salmon, pink salmon, and sea trout captured between the Teno estuary and the Polmak sonar counting site. (B) Number of Atlantic salmon, pink salmon, and sea trout counted on video at the Polmak sonar counting site. (C) Number of ascending salmonid fish counted on sonar at the Polmak sonar counting site. (D) Average school sizes of ascending salmonid fish at the Polmak sonar counting site.



tween these species is believed to be reflected by catch statistics and video data. The model follows a four-part structure described as follows:

(1) Model for migration process of all three species
(2) Observation model for catch statistics
(3) Observation model for video counts
(4) Observation model for the number of schools

JAGS (Plummer 2003) code for the model can be found in the supplementary material, and a visual representation of the model as a directed acyclic graph is shown in Fig. 3. Table 1 contains a list of model's parameters, their explanations, and prior distributions.

Model for migration process of all three species

On day i, the total number of fish passing the sonar counting site, each fish being observed, comprises the sum of Atlantic salmon, pink salmon, and sea trout passing through:

$$(1) \quad n_i^{\text{tot}} = n_i^{\text{at}} + n_i^{\text{pi}} + n_i^{\text{tr}}$$

where n_i^{at}, n_i^{pi}, and n_i^{tr} are the daily migration numbers (at for Atlantic salmon, pi for pink salmon, and tr for trout).

The daily numbers n_i^{at}, n_i^{pi}, and n_i^{tr} at day i calculated from the total amounts:

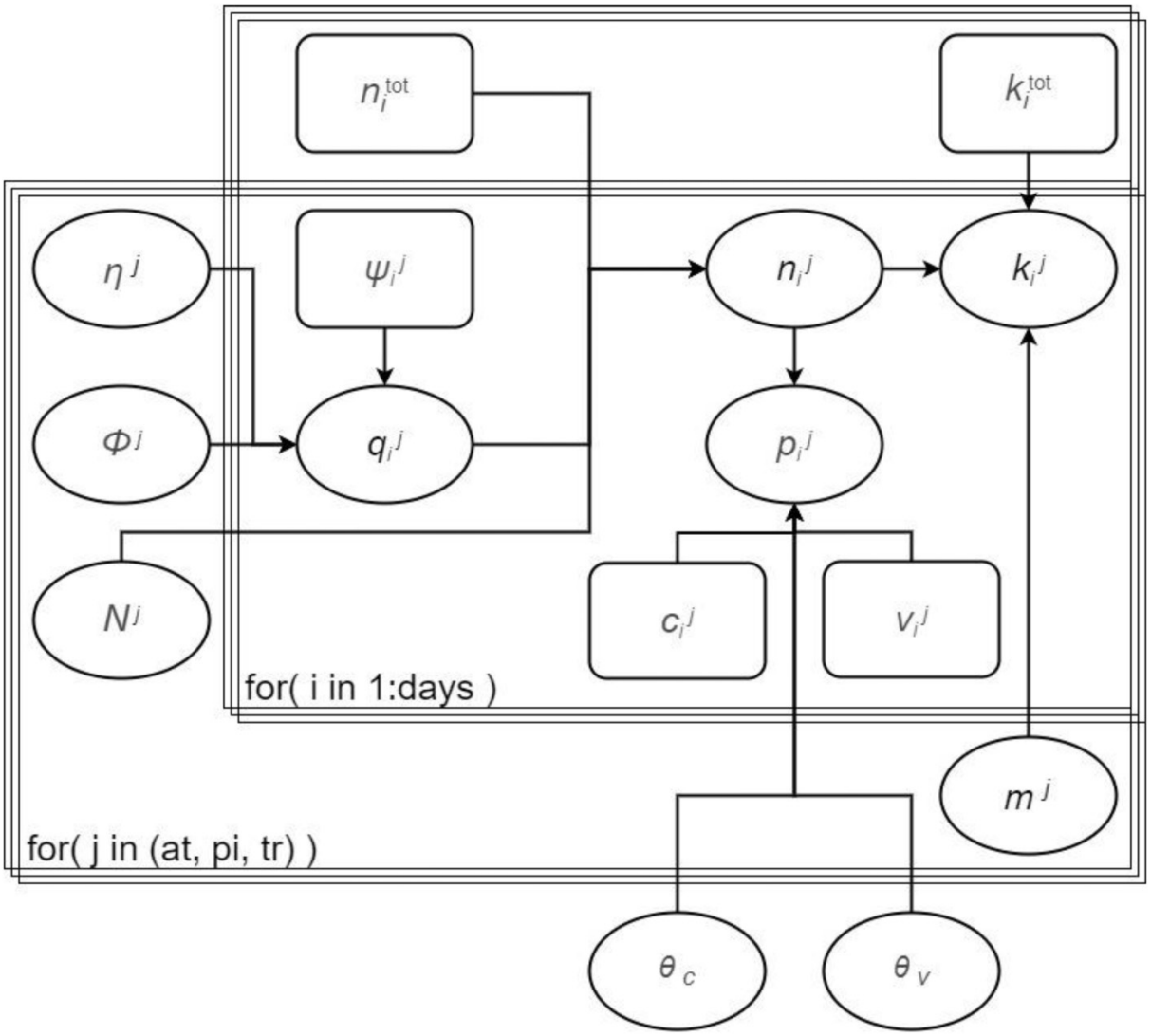
$$(2) \quad n_i{}^{\text{at}} = q_i^{\text{at}} \cdot N^{\text{at}}$$

$$(3) \quad n_i{}^{\text{pi}} = q_i^{\text{pi}} \cdot N^{\text{pi}}$$

$$(4) \quad n_i{}^{\text{tr}} = q_i^{\text{tr}} \cdot N^{\text{tr}}$$

Here, N^{at}, N^{pi}, and N^{tr} represent the total numbers of Atlantic salmon, pink salmon, and sea trout, respectively. The values q_i^{at}, q_i^{pi}, and q_i^{tr} correspond to the proportions migrating on day i, derived from the respective vectors q^{at}, q^{pi}, and q^{tr} representing the daily migration proportions.

Fig. 3. A simplified visual representation of the model structure represented as a directed acyclic graph. Variable names and descriptions are detailed in Table 1.



The total amounts of each species are given an uninformative uniform prior on the log-scale allowing for a wide, minimally informative prior (Pulkkinen et al. 2019):

$$\log\left(N^{at}\right), \log\left(N^{pi}\right), \log\left(N^{tr}\right) \sim \text{Unif}(0,\ 20) \quad (5)$$

Daily migration proportions are assumed to be affected by the total population size and a scaling factor, reflecting variability in migration intensity. At its foundation, the probability vectors can be described using Dirichlet distribution:

$$q^{at} \sim \text{Dir}\left(\psi^{at} \cdot \eta^{at} \cdot N^{at}\right) \quad (6)$$

$$q^{pi} \sim \text{Dir}\left(\psi^{pi} \cdot \eta^{pi} \cdot N^{pi}\right) \quad (7)$$

$$q^{tr} \sim \text{Dir}\left(\psi^{tr} \cdot \eta^{tr} \cdot N^{tr}\right) \quad (8)$$

Dirichlet weights are determined by the expected value vectors ψ^{at}, ψ^{pi}, ψ^{tr}, dispersion scale parameters η^{at}, η^{pi}, η^{tr}, and total amounts N^{at}, N^{pi}, N^{tr}.

For better computational performance, the Dirichlet distributions are approximated with lognormal distributions (see Appendix A). Within the Dirichlet approximation, the expected value vectors are derived from an autoregressive AR(1) structure with autoregression coefficients ϕ and expert elicited expected values μ^q (details in Appendix A). The expert elicited expected values are formed from the expert elicited migration intensity graphs (Fig. 4). Autoregression is used to capture the migration trends assuming that deviation from the expected number of daily migration is not random and subsequent days are likely to be similar.

While the Dirichlet distribution provides a foundational structure for the probability vectors, the incorporation of AR(1) structure using the expert elicited values adds a layer of temporal dependency, reflecting the non-random nature of migration trends over time.

The dispersion scale parameters η^{at}, η^{pi}, and η^{tr} are given uninformative priors. These parameters control the degree of overdispersion in the Dirichlet distribution by scaling the total amounts N. Specifically, when η is 1, the Dirichlet dis-

Table 1. List of symbols, descriptions, and prior distributions.

Symbol	Description	Prior distribution
$\log(N^{at})$	Total number of Atlantic (45–65 cm) salmon on a logarithmic scale	Unif (0, 20)
$\log(N^{pi})$	Total number of pink salmon on a logarithmic scale	Unif (0, 20)
$\log(N^{tr})$	Total number of sea trout on a logarithmic scale	Unif (0, 20)
N^{tr}	Total number of trout on real scale	–
n_i^{at}	Daily migration of Atlantic salmon	–
n_i^{pi}	Daily migration of pink salmon	–
n_i^{tr}	Daily migration of sea trout	–
η^{at}	Dispersion scale parameter for the total number of Atlantic salmon (45–65 cm)	Unif (0, 1)
η^{pi}	Dispersion scale parameter for the total number of pink salmon	Unif (0, 1)
η^{tr}	Dispersion scale parameter for the total number of sea trout	Unif (0, 1)
q^{at}	Daily portions to migrate for Atlantic salmon	–
ψ^{at}	Experts' view on the expected migration intensity of Atlantic salmon	–
q^{pi}	Daily portions to migrate for pink salmon	–
ψ^{pi}	Experts' view on the expected migration intensity of pink salmon	–
q^{tr}	Daily portions to migrate for trout	–
ψ^{tr}	Experts' view on the expected migration intensity of sea trout	–
ϕ^{at}	Autoregressive coefficient of daily migration intensity for Atlantic salmon	Unif (0, 0.95)
ϕ^{pi}	Autoregressive coefficient of daily migration intensity for pink salmon	Unif (0, 0.95)
ϕ^{tr}	Autoregressive coefficient of daily migration intensity for sea trout	Unif (0, 0.95)
$p_i = (p_i^{at}, p_i^{pi}, p_i^{tr})$	True daily species distribution	
$\widehat{p}_{1,i} = \left(\widehat{p}_{1,i}^{at}, \widehat{p}_{1,i}^{pi}, \widehat{p}_{1,i}^{tr}\right)$	Daily species distribution within catch statistics	
$\widehat{p}_{2,i} = \left(\widehat{p}_{2,i}^{at}, \widehat{p}_{2,i}^{pi}, \widehat{p}_{2,i}^{tr}\right)$	Daily species distribution within video data	
θ_c	Dispersion scale parameter for species distribution within catch data	Unif (0, 1)
θ_v	Dispersion scale parameter for species distribution within video data	Unif (0, 1)
k_i^{at}	Daily number of schools for Atlantic salmon (45–65 cm)	–
k_i^{pi}	Daily number of schools for pink salmon	–
k_i^{tr}	Daily number of schools for sea trout	–
m^{at}	Average school size for Atlantic salmon (45–65 cm)	LogN (0, 0.2)
m^{pi}	Average school size for pink salmon	LogN (0.75, 0.17)
m^{tr}	Average school size for sea trout	LogN (0, 0.2)
Data		
n_i^{tot}	Daily observations	–
c_i^{at}	Daily number of caught Atlantic salmon (45–65 cm)	–
c_i^{pi}	Daily number of caught pink salmon	–
c_i^{tr}	Daily number of caught sea trout	–
c_i^{tot}	Daily total catch	–
v_i^{at}	Daily number of Atlantic salmon (45–65 cm) observed on video	–
v_i^{pi}	Daily number of pink salmon observed on video	–
v_i^{tr}	Daily number of sea trout observed on video	–
v_i^{tot}	Total number of fish observed on video daily	–
k_i^{tot}	Daily number of schools observed	–

tribution follows the expected values ψ without additional overdispersion and the uncertainty around expected values is as expected given the N. If η is less than one, overdispersion is present indicating greater than expected uncertainty around the expected values for the given N. By assuming a uniform prior between 0 and 1:

$$\eta^{at}, \eta^{pi}, \eta^{tr} \sim \text{Unif}(0, 1) \quad (9)$$

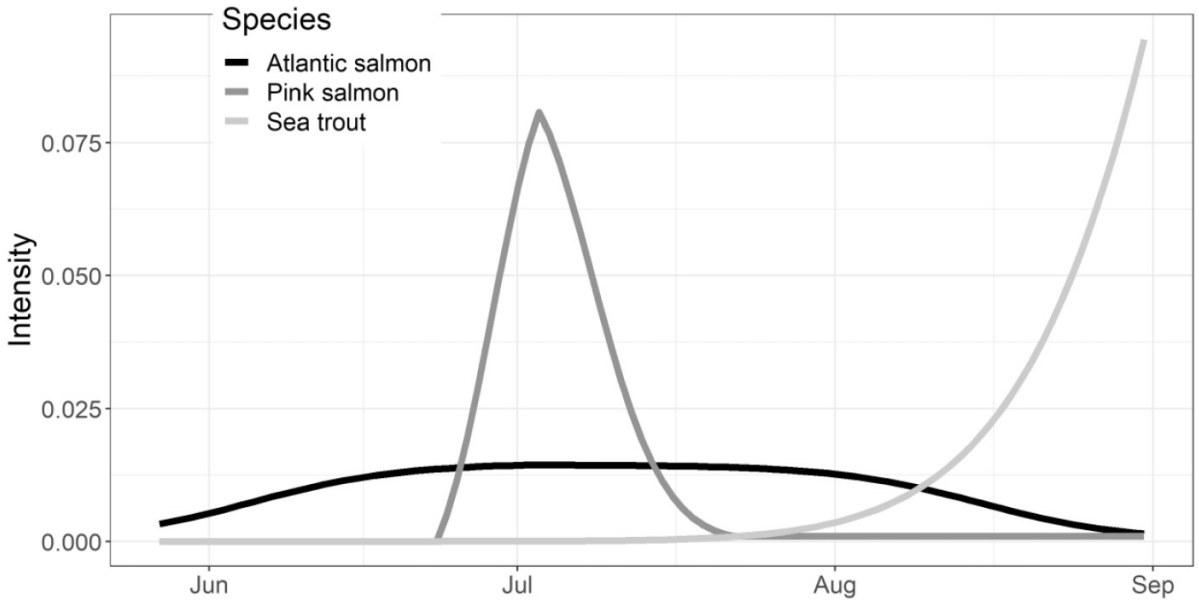
we assume no prior knowledge about the degree of overdispersion.

The prior distribution for the autoregression coefficients of daily migration intensities ϕ^{at}, ϕ^{pi}, ϕ^{tr} assumes a positive correlation between consecutive days:

$$\phi^{at}, \phi^{pi}, \phi^{tr} \sim \text{Unif}(0, 0.95) \quad (10)$$

With daily migration numbers n_i^{at}, n_i^{pi}, n_i^{tr}, and their sum, the true daily species distribution is calculated and subsequently used as the expected values for catch and video observations:

Fig. 4. Expert-defined expected values for migration intensities of Atlantic salmon, pink salmon, and sea trout.



$$p_i^{at} = \frac{n_i^{at}}{n_i^{at} + n_i^{pi} + n_i^{tr}} \quad (11)$$

$$p_i^{pi} = \frac{n_i^{pi}}{n_i^{at} + n_i^{pi} + n_i^{tr}} \quad (12)$$

$$p_i^{tr} = \frac{n_i^{tr}}{n_i^{at} + n_i^{pi} + n_i^{tr}} \quad (13)$$

The individual values p_i^{at},p_i^{pi}, and p_i^{tr} are combined into vector $p_i = \left(p_i^{at}, p_i^{pi}, p_i^{tr}\right)$.

Observation model for the catch statistics

The daily number of caught Atlantic salmon, pink salmon, and sea trout represents a sample from the actual species distribution and is modelled using a Dirichlet-multinomial distribution:

$$c_i^{at}, c_i^{pi}, c_i^{tr} \sim \text{Multinom}\left(\widehat{p}_{1,i}, c_i^{tot}\right) \quad (14)$$

where c_i^{at}, c_i^{pi}, and c_i^{tr} are the number of Atlantic salmon, pink salmon, and sea trout caught at day i and c_i^{tot} is the total number of fish caught on that day. The expected species distribution within the c_i^{tot}, denoted as $\widehat{p}_{1,i} = \left(\widehat{p}_{1,i}^{at}, \widehat{p}_{1,i}^{pi}, \widehat{p}_{1,i}^{tr}\right)$ is assumed to follow Dirichlet distribution:

$$\widehat{p}_{1,i}^{at}, \widehat{p}_{1,i}^{pi}, \widehat{p}_{1,i}^{tr} \sim \text{Dir}\left(p_i \cdot \theta_c \cdot c_i^{tot}\right) \quad (15)$$

where the actual species distribution p_i is used as the expected value.

The dispersion scale parameter θ_c, given a minimally informative prior distribution:

$$\theta_c \sim \text{Unif}(0, 1) \quad (16)$$

functions similarly to η parameters discussed in earlier part, controlling the degree of over dispersion. While the multinomial distribution accounts for sample variation, θ_c introduces additional uncertainty through overdispersion, allowing the model to better capture the variation in catch.

For computational efficiency, the Dirichlet distribution is approximated using lognormal distributions (Appendix B), and the multinomial distribution is represented using independent Poisson distributions by conditioning on their sum (Gelman et al. 2013).

Observation model for the video counts

Video observations are also a sample of the actual daily species distribution and are modelled in the same way as caught salmon, using the Dirichlet-multinomial distribution:

$$v_i^{at}, v_i^{pi}, v_i^{tr} \sim \text{Multinom}\left(\widehat{p}_{2,i}, v_i^{tot}\right) \quad (17)$$

where v_i^{at}, v_i^{pi}, and v_i^{tr} are the number of Atlantic salmon, pink salmon, and sea trout observed in video at day i and v_i^{tot} is the total number of fish observed in video on that day. The species distribution within the video$_i$, denoted as $\widehat{p}_{2,i} = \left(\widehat{p}_{2,i}^{at}, \widehat{p}_{2,i}^{pi}, \widehat{p}_{2,i}^{tr}\right)$ is assumed to follow Dirichlet distribution:

$$\widehat{p}_{2,i}^{at}, \widehat{p}_{2,i}^{pi}, \widehat{p}_{2,i}^{tr} \sim \text{Dir}\left(p_i \cdot \theta_v \cdot v_i^{tot}\right) \quad (18)$$

where the actual species distribution p_i is used as the expected value. The dispersion scale parameter θ_v, which is given an uninformative prior distribution, similarly accounts for the additional variation through overdispersion akin to θ_c discussed previously:

$$\theta_v \sim \text{Unif}(0, 1) \quad (19)$$

Like with catch data, the Dirichlet distribution is approximated using lognormal distributions (Appendix B), and the multinomial distribution is represented using independent Poisson distributions by conditioning on their sum (Gelman et al. 2013) for computational efficiency.

Observation model for the number of schools

The daily number of schools for each species k_i^{at}, k_i^{pi}, and k_i^{tr} are assumed to follow gamma distributions:

$$k_i^{at} \sim \text{Gamma}\left(n_i^{at},\ m_{at}\right) \quad (20)$$

$$k_i^{pi} \sim \text{Gamma}\left(n_i^{pi},\ m_{pi}\right) \quad (21)$$

$$k_i^{tr} \sim \text{Gamma}\left(n_i^{tr},\ m_{tr}\right) \quad (22)$$

where m^{at}, m^{pi}, and m^{tr} represent average school sizes for each species. The average school sizes m^{at}, m^{pi}, and m^{tr} are constricted to support values in the interval $[1, \infty)$. This constraint is implemented through a transformation using latent variables Lm^{at}, Lm^{pi}, and Lm^{tr}, each following a log-normal distribution:

$$Lm^{at} \sim \text{LogN}\left(0, 0.2\right),\ m^{at} = \exp\ \left(Lm^{at}\right) \quad (23)$$

$$Lm^{pi} \sim \text{LogN}\left(0.75, 0.17\right),\ m^{pi} = \exp\left(Lm^{pi}\right) \quad (24)$$

$$Lm^{tr} \sim \text{LogN}\left(0, 0.2\right),\ m^{tr} = \exp\left(Lm^{tr}\right) \quad (25)$$

The total daily number of schools, k_i, is the sum of gamma-distributed variables k_i^{at}, k_i^{pi}, and k_i^{tr} and is assumed to follow a gamma distribution:

$$k_i^{tot} \sim \text{Gamma}\left(\frac{E\left(k_i^{tot}\right)^2}{\text{Var}\left(k_i^{tot}\right)}, \frac{E\left(k_i^{tot}\right)}{\text{Var}\left(k_i^{tot}\right)}\right) \quad (26)$$

where the expected value $E\left(k_i^{tot}\right)$ and the variance $\text{Var}\left(k_i^{tot}\right)$ are given by

$$E\left(k_i^{tot}\right) = \frac{n_i^{at}}{m_{at}} + \frac{n_i^{pi}}{m_{pi}} + \frac{n_i^{tr}}{m_{tr}} \quad (27)$$

$$\text{Var}\left(k_i^{tot}\right) = \frac{n_i^{at}}{m_{at}^2} + \frac{n_i^{pi}}{m_{pi}^2} + \frac{n_i^{tr}}{m_{tr}^2} \quad (28)$$

The expected value of $E\left(k_i^{tot}\right)$ represents the expected total number of schools observed, aggregating species-specific counts k_i^{at}, k_i^{pi}, and k_i^{tr}. Similarly, the variance $\text{Var}\left(k_i^{tot}\right)$ is calculated as sum of variances from the same components.

Prior distributions for average daily school size for Atlantic salmon and sea trout (m^{at} and m^{tr}) are chosen based on earlier observations from the Teno River monitoring programs (unpub. data). According to the data, Atlantic salmon and sea trout tend to migrate alone or in small schools and thus priors for m^{at} and m^{tr} mainly support average school sizes less than five individuals (Table 1). In contrast, pink salmon prefer larger school sizes (Kaev and Rudnev 2007, unpub. Teno data). Prior given for average school size of pink salmon (m^{pi}) mainly supports values over five individuals. Prior distributions for m^{at}, m^{pi}, and m^{tr} are parameterized to support values greater than one because an average school size of less than one is impossible.

Expert elicitation

Informative priors about the timing and intensity of spawning migrations for three species were elicited from two competent experts. Experts were consulted separately and asked to base their views on their existing knowledge on timing and intensity of Atlantic salmon, sea trout, and pink salmon's spawning migration in earlier years. Three blank graphs, one for each species, with a time axis between May and September, were supplied to each expert. Experts drew a single solid continuous line to present the most likely intensity curve of the migration.

Upon comparing the results, it was found that the experts' views on the migration intensities were highly similar. Due to this high agreement, their individual intensity curves were combined into a single representative curve for each species. These combined curves were later shown to the experts, who reviewed and agreed upon them being accurate representations of their knowledge. This approach ensured that the informative priors accurately reflected the consensus of expert opinion. The expected value vectors used in eqs. 6–8 were based on these combined curves (Fig. 4) (Appendix A).

Sensitivity analysis on influence of school size observations

Additionally, the model was run without the school size data to study its influence on the estimates. Results of this model run can be found from the supplementary material.

Details of the model run

The model was fitted using Markov chain Monte Carlo sampling with JAGS 4.3.0 (just another Gibbs sampler) (Plummer 2003) and runjags R-package (Denwood 2016) to control the JAGS software. The model was run for 12 million iterations with two chains, taking around 1 week. A burn-in period of one million samples was removed from the beginning. The sample was post-processed by thinning it by 750, resulting in a final sample of eight thousand iterations per chain. Convergence for parameters of interest was diagnosed using Gelman–Rubin statistics (Gelman and Rubin 1992) by making sure the potential scale reduction factor was under 1.05 and the Monte Carlo standard error was under 5% of standard deviation for these parameters. More detailed convergence diagnostics and trace plots can be found from the supplementary material Figs. S1a and S1b and Table S1.

Results

Based on our model's posterior distributions, the median for the total number of small, between 45 and 65 cm long, Atlantic salmon migrating into Teno River in 2021 was 11.8 k individuals, with 90% highest density probability interval (PI) being between 10.5 and 13.3 k. For the pink salmon the median is 52.0 k with 90% PI from 51.4 to 52.5 k, and for sea trout the median 6.6 k with 90% PI between 5.3 and 7.8 k. Table 2 shows prior and posterior results for all key parameters.

All key parameters updated when comparing prior distributions to their corresponding posteriors (Figs. 5*a* and 5*b*). Posterior distributions for dispersion scale parameters on migration intensities η^{at}, η^{pi}, and η^{tr} showed significant updates from their uninformative priors to values near zero. This indicates a substantial difference from the expert-defined ex-

Table 2. Prior and posterior medians and 5% and 95% quantiles of key parameters from the 2021 model run.

	Prior			Posterior		
Variable	Median	5%	95%	Median	5%	95%
N^{at}	20 000	1	66 526 511	11 796	10 486	13 283
N^{pi}	20 000	1	66 526 511	51 957	51 358	52 522
N^{tr}	20 000	1	66 526 511	6590	5254	7805
η^{at}	0.50	0.05	0.95	0.011	9.8×10^{-6}	0.021
η^{pi}	0.50	0.05	0.95	1.4×10^{-4}	8.9×10^{-9}	3.7×10^{-4}
η^{tr}	0.50	0.05	0.95	1.6×10^{-4}	1.3×10^{-9}	5.6×10^{-4}
θ_c	0.50	0.05	0.95	2.3×10^{-3}	2.6×10^{-7}	7.6×10^{-3}
θ_v	0.50	0.05	0.95	1.0×10^{-3}	5.6×10^{-8}	3.8×10^{-3}
m^{at}	2.71	1.94	3.75	1.40	1.36	1.44
m^{pi}	8.33	4.18	14.32	12.81	11.95	13.82
m^{tr}	2.71	1.94	3.75	1.21	1.17	1.25
ϕ^{at}	0.47	0.02	0.88	0.62	0.37	0.84
ϕ^{pi}	0.47	0.02	0.88	0.69	0.62	0.76
ϕ^{tr}	0.47	0.02	0.88	0.92	0.90	0.95

pected values for migration intensities, as visualized in Figs. 6–8. These results were expected because the expert's migration intensity curves were designed to be general and thus cannot cover year specific details in, for example, environmental conditions that affect the migration intensity.

The estimated daily abundance of sea trout (Fig. 8) suggests a bimodal migration, with the first peak occurring near the start of July and the second peak at the start of August. The first peak in July is unexpected and potentially caused by the catch statistics, which mostly record over-wintering sea trout during that time. Considering this, if the estimated sea trout up to 24 June are ignored, a more realistic and revised estimate for sea trout would have a median of 5.2 k individuals and 90% PI of 4.1–6.2 k.

The parameters for daily average school size m^{at}, m^{pi}, and m^{tr} have updated, the one for pink salmon (m^{pi}) being the largest with a median of 12.8 and 90% PI of 12.0–13.8 fish per school. Parameters of average school size for Atlantic salmon and sea trout are significantly smaller, the median for m^{at} is 1.40 with 90% PI between 1.36 and 1.44 fish per school, and the median for m^{tr} is 1.21 with 90% PI between 1.17 and 1.25 fish per school. These differences are in line with known differences of the schooling behaviour of these species.

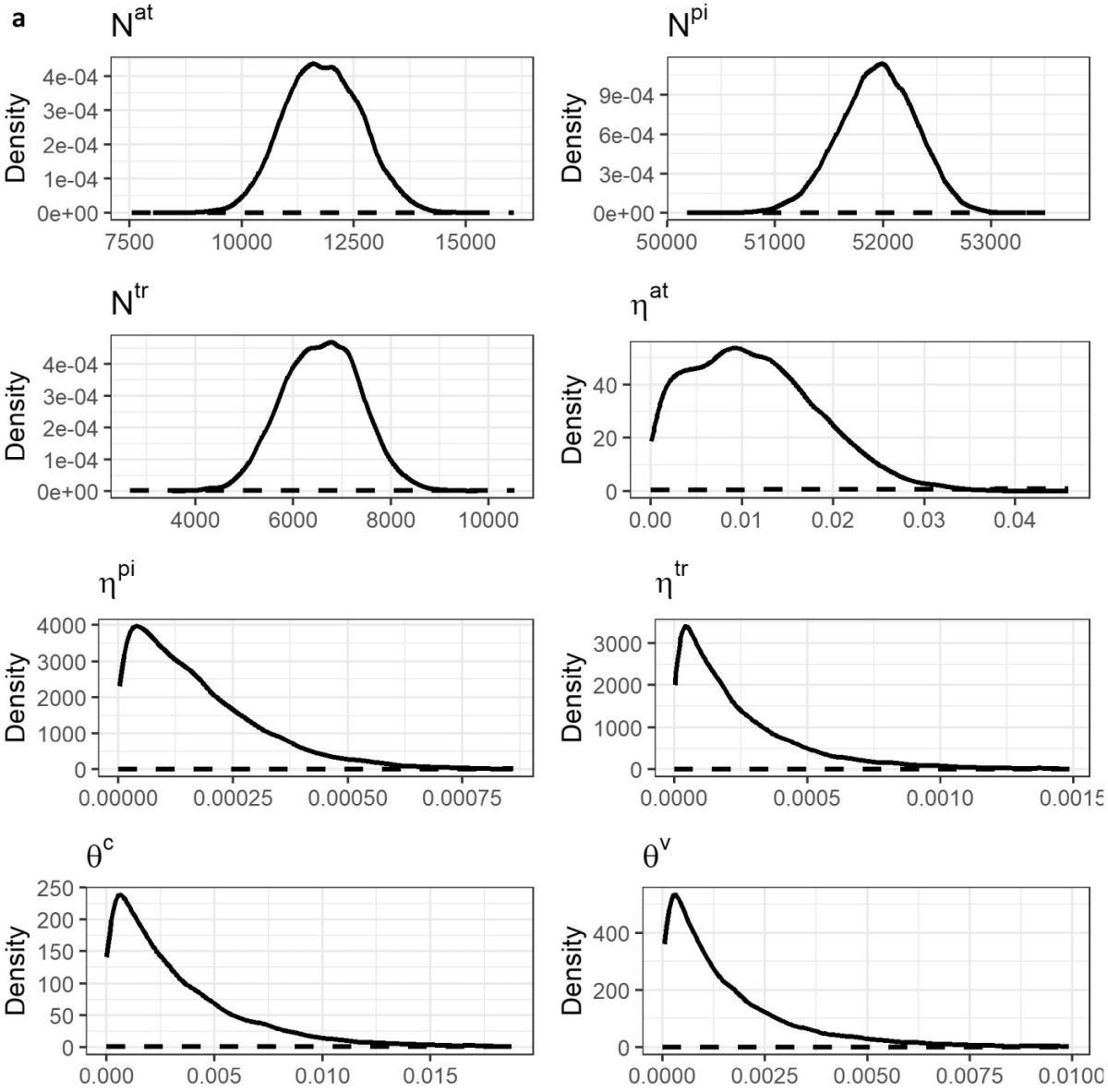
Autoregression coefficient of daily migration intensity for sea trout (ϕ^{tr}) is the largest among the three species, with a median of 0.92 and 90% PI from 0.90 to 0.95. This indicates that sea trout migration shows stronger autocorrelation compared to other species. Parameter ϕ^{pi} is smaller, with a median of 0.69 and 90% PI of 0.62−0.76 and ϕ^{at} is the smallest with a median of 0.62 and 90% PI of 0.37−0.84.

Discussion

Sonar-based monitoring methods are gaining popularity in fish stock monitoring and management (Wei et al. 2022). In rivers where several similarly sized species coexist, as in the Teno River, the challenges of species recognition become apparent. In this study, we developed a Bayesian model combining three data sources (individual count and school size data from the sonar, catch statistics, and video data) and expert knowledge on migration intensity to estimate number of individuals passing the sonar site for Atlantic salmon, sea trout, and pink salmon of size range 45–65 cm. Our approach offers clear advantage over the previously used (2018–2020) method that relied solely on catch statistics to estimate species distribution in sonar counts and did not measure uncertainty of abundance. Although the species distribution can be estimated from high-quality catch statistics (Fox and Starr 1996; Alglave et al. 2022), they can be biased and not representative. At Teno River, the quality of catch statistics has dramatically decreased after the introduced fishing ban in 2021, and only a limited amount of fishing licenses were given out for selected people to conduct small-scale experimental fishing. This prevented the use the previous estimation method relying on catch statistics in 2021. By combining multiple sources of information in a Bayesian model, the estimated species distribution is not strictly driven by single uncertain source of data and will provide a better representation of the true species distribution.

A direct comparison between our model and the previous method to estimate small salmon abundance in 2021 is challenging. To compare results, the previous method was applied using early season sonar counts (1–29 June; before pink salmon run starts) corrected with daily species distributions from catch data 2018–2020. The corrected early season sonar counts were then extrapolated based on the annual proportions (average and years separately) of small salmon migrating past the sonar monitoring site during the same early season period in 2018–2020. This method yielded a small salmon estimate of 16.4 k salmon when using average proportion from 2018 to 2020 and estimates between 13.8 k and 21.2 k salmon when using single year proportions. The model-based estimate is more reliable because it integrates annually updated information, multiple data sources, and expert-based

Fig. 5. Posterior distribution to prior distribution comparisons for the key parameters. Solid line represents the posterior distribution, and dashed line represents the prior distribution.

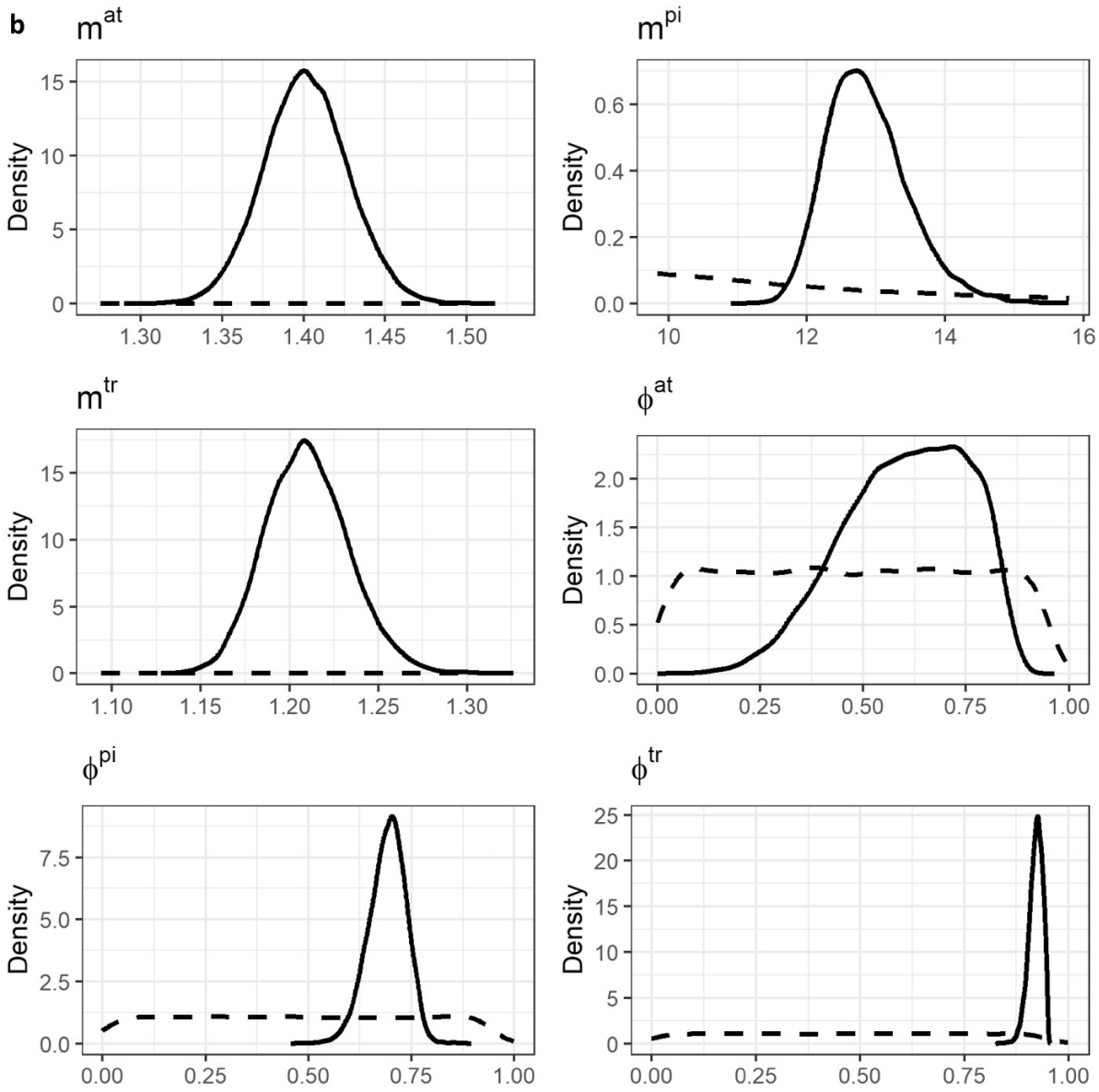


expected values, providing a more comprehensive representation of the actual migration.

The model's sensitivity to school size data was assessed by running the model without it. Comparison of the results (Table 2 and Table S1) reveals significant differences. Autoregression coefficients for daily migration intensities (ϕ^{at}, ϕ^{pi}, and ϕ^{tr}) and dispersion scale parameters for migrations intensities (η^{at}, η^{pi}, and η^{tr}) and for catch and video data (θ_c and θ_v) are consistently higher in the absence of school size data. Parameters θ_c and θ_v show the largest increase, indicating greater reliance on catch and video data for estimating daily species distributions. Abundance estimates for Atlantic salmon and sea trout are lower, while abundance estimate for pink salmon becomes higher, when the school size data are not included. Figures 6–8 and S4–S6 show substantial differences in daily abundance estimates, with Atlantic salmon migration appearing to decline earlier, and sea trout migration peaking earlier. The 90% PI for sea trout abundance in the model without the school size data nearly overlaps with our revised estimate. However, the Atlantic salmon abundance estimate (median: 7.2 k, 90% PI: 6.3 k–8.2 k) appears unrealistically low, considering that additional monitoring data from tributaries Utsjoki, Inarijoki, and Kárášjohka (Fig. 1) indicate that 4.3 k small Atlantic salmon migrated into these tributaries alone in 2021. These findings highlight the critical role of school size data in improving the accuracy of model's abundance estimates, particularly for Atlantic salmon.

In general, video monitoring can provide reliable data on migration of different fish species and their life stages (Orell et al. 2007*a*). However, if the cameras do not cover the whole river width, certain species may be under- or overestimated

Fig. 5. (*concluded*).



if they prefer a specific part of the river during migration (Thorstad et al. 2008). The four cameras used in Teno River in 2021 covered different depths and parts of the river channel, but still a small proportion of the river width. In addition, the video cameras were operational for less than a half of the migration period. Preferably, video data should cover the entire migration periods and a larger proportion of the channel width.

Our model's total abundance estimates have relatively small standard deviations, resulting in tight highest-density PIs. The main cause for this is the assumption that all fish are observed, leading to situation where the total number of fish does not contribute to the uncertainty. Instead, the uncertainty is solely due to the uncertainty between the proportions of different fish species. While the small uncertainty is generally beneficial, acknowledging this assumption is crucial.

The dispersion scale parameters, related to the daily number of each species, estimated into values near zero. This suggests an incompatibility between the expert-defined expected migration intensity and the estimated migration intensity. Such a deviance is not unexpected, as accurately capturing migration intensity is challenging for most years due to the variable nature of river conditions. Factors such as discharge and temperature, which significantly affect the migration behaviour of Teno River salmon, can change drastically from year to year (Erkinaro et al. 1999; Karppinen et al. 2004).

To further enhance the model's predictive capabilities, future developments should focus on integrating a migration model that includes additional data sources from environmental and in-river variables, such as those used for modelling smolt migration in Utsjoki (Pulkkinen et al. 2019). These variables could include water temperature, discharge, and tidal rhythms, helping to capture the autocorrelation be-

Fig. 6. Estimated daily abundance of Atlantic salmon. Boxplots represent the 5%, 25%, 50%, 75%, and 95% quantiles of estimated daily abundancies. The dashed line represents the case if the estimated total abundance of Atlantic salmon followed the expert-based expected values.

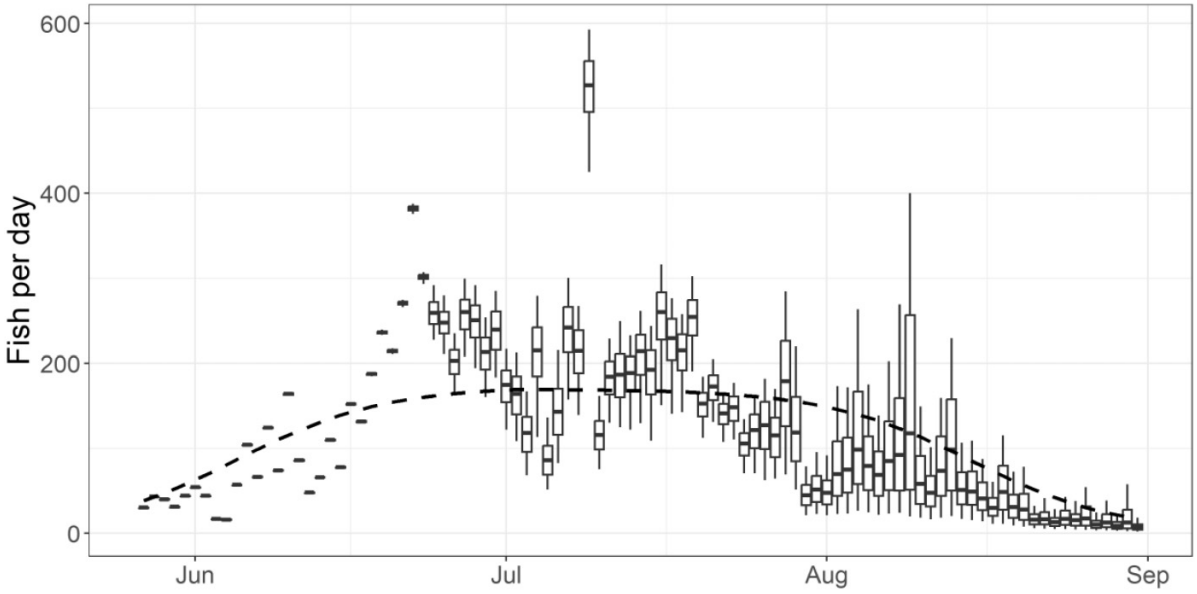
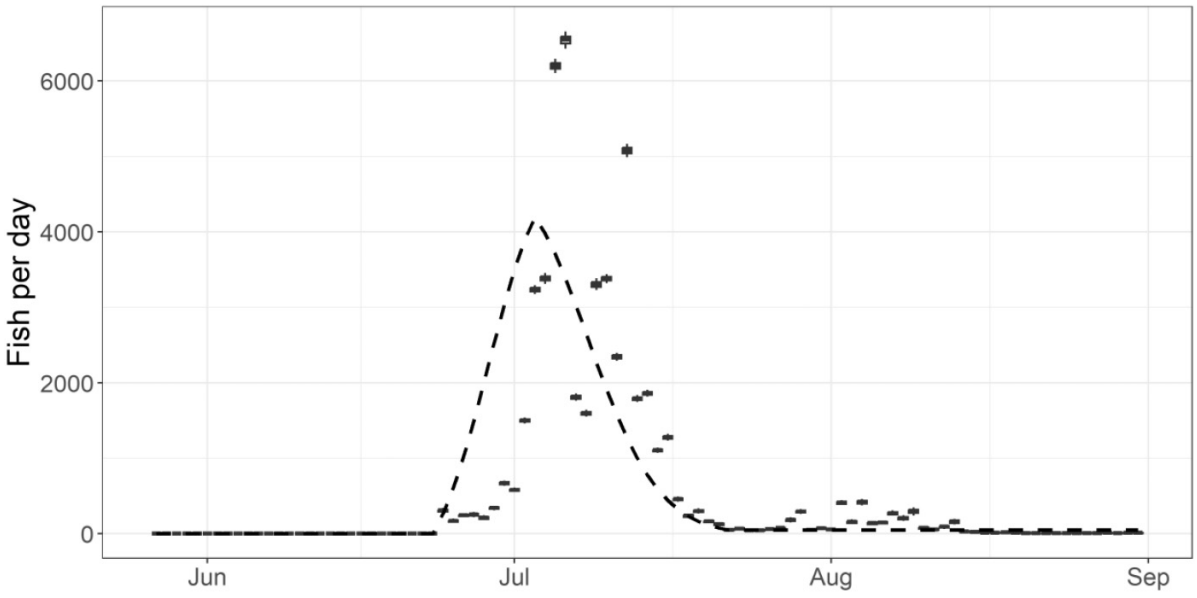


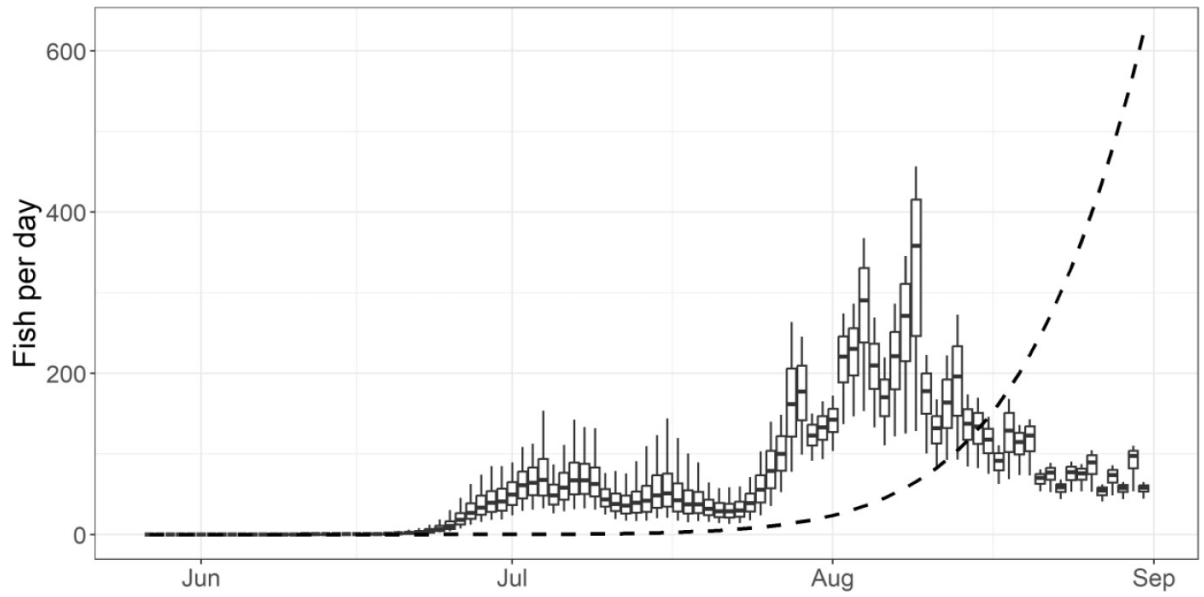
Fig. 7. Estimated daily abundance of pink salmon. Boxplots represent the 5%, 25%, 50%, 75%, and 95% quantiles of estimated daily abundancies. The dashed line represents the case if the estimated total abundance of pink salmon followed the expert-based expected values.



tween days and provide a more accurate representation of migration patterns. Moreover, fish length data from sonar measurements could provide an additional source of information, as it has been shown to aid with species recognition (Gurney et al. 2014). These refinements would allow the model to better explain uncertainties and improve its overall performance.

Expert knowledge on migration intensities of different species varied a lot. For Atlantic salmon, long-term monitoring in different parts of the Teno River (catch, video, snorkelling, sonar data) provides a strong basis for understanding the general pattern for migration (Borgstrøm et al. 2010; Orell et al. 2011). For pink salmon, very little information is available before 2017 and understanding their migration patterns is still rather weak. Migration of the Teno sea trout is better understood and clearly different from the two other species. Orell et al. (2017) have documented a 2-year spawning migration of sea trout and extensive overwintering-feeding migrations between sea and river during the spring and summer.

Fig. 8. Estimated daily abundance of sea trout. Boxplots represent the 5%, 25%, 50%, 75%, and 95% quantiles of estimated daily abundancies. The dashed line represents the case if the estimated total abundance of sea trout followed the expert-based expected values.



The decision to analyse 12 h of raw sonar data per echo sounder each day, while necessary to manage workload, results in half of the daily raw sonar data being unanalysed. To mitigate this, the analysed hours are alternated between the two echo sounders, and the results are extrapolated to cover the unanalysed hours. This extrapolation assumes consistent migration patterns, which may not necessarily be a totally valid assumption due to varying diel variations in environmental conditions such as illumination, temperature, and discharge (Erkinaro et al. 1999; Karppinen et al. 2004; Davidsen et al. 2005). Generally, this approach works well, but it relies on the assumption that no regular 2 h based patterns exist in salmonid migration. A regular 2 h based migration pattern would cause the extrapolation to become biased and to ensure unbiased extrapolation, the hours selected for analysis should be randomized, giving each hour 50% probability to be chosen for analysis. This would mitigate any potential sources for bias.

Additionally, a straightforward solution is to analyse all 24 h of raw data from both echo sounders. Advancements in automatic fish detection and measurement techniques, such as the use of deep learning and other neural network approaches, could make this feasible by reducing manual labour (Kandimalla et al. 2022; Tang et al. 2023). Alternatively, incorporating uncertainty about daily sonar observations into the model could also mitigate this issue. This incorporation would likely increase the uncertainty attached to the abundance estimates, but the estimates would more likely be a better representation of the uncertainty about true abundances.

The proportion of sea trout was somewhat higher than expected, especially during mid-summer at late June and July. Possible cause for this may be linked to the complicated migration patterns of sea trout (Orell et al. 2017) and the use of catch data that also include individuals that are not on their spawning migration and should not be accounted while estimating the total abundance. By excluding the abundance estimates before 25 June, a more realistic estimate for sea trout abundance was gained. This adjustment also caused the abundance estimates for the Atlantic salmon and the pink salmon to raise slightly. This revised estimate likely represents a more realistic spawning population size estimate for the sea trout. To avoid this issue a quick and straightforward solution is to calculate the revised estimate by excluding early counts. A more sophisticated approach would be to divide sea trout into two groups: over-wintering and spawning. Each group would have its own expert-defined expected values for migration intensity and be modelled as separate species. This two-group solution would allow the model to distinguish between over-wintering and spawning populations of sea trout, rather than relying on a selected date to make this division.

Reliable modelling of the multi-species salmonid run at the Teno River is crucial for robust assessment of the status of the numerous Atlantic salmon populations in the catchment. As the status of Atlantic salmon populations of the Teno River has rapidly declined and is currently at the all-time low (Anon 2024) and the ascending pink salmon numbers have concurrently been increasing dramatically (Domaas et al. 2024; Erkinaro et al. 2024), better methods for estimating the species-specific run sizes are badly needed. Moreover, as the use of sonars in estimating fish population sizes has increased globally (Wei et al. 2022), and become popular in assessing anadromous salmonid runs (Helminen et al. 2020), there are obvious needs for reliable modelling methods for species-identification that can be applied for various situations. Our model for estimating the number of individuals of different species from sonar counts is flexible and can be adapted to function in different rivers with varying species

composition. The number of species can be changed, and different migration intensity priors can be specified for each species. While less accurate methods can be used for species recognition, in most cases, a more precise modelling-based approach with proper uncertainty estimates becomes essential to fully utilise the collected data. As data over multiple years accumulate, the concept of forming a hierarchical structure for migration intensities becomes more apparent. This approach not only improves the understanding of migration dynamics over time but also refines the estimation of migration intensity parameters through the utilization of data from multiple years, ultimately improving the precision of estimates.

Acknowledgements

We thank numerous field workers enabling the sonar monitoring program and collecting sonar data on the Polmak sonar counting site. We acknowledge the pivotal know-how of Jan-Peter Pohjola in organizing the analysis of raw sonar data and helping to understand the analysis process. Additionally, we acknowledge Olli van der Meer for their assistance in creating visual illustrations of study area.

Article information

History dates
Received: 4 October 2024
Accepted: 27 January 2025
Accepted manuscript online: 20 February 2025
Version of record online: 26 March 2025

Data availability
The data underlying this article are available in the Dryad Digital Repository at https://doi.org/10.5061/dryad.cvdncjtds.

Author information

Author ORCIDs
Antti Räty https://orcid.org/0009-0004-2277-0353
Henni Pulkkinen https://orcid.org/0000-0002-0926-0285
Jaakko Erkinaro https://orcid.org/0000-0002-7843-0364
Panu Orell https://orcid.org/0000-0003-4294-5048
Morten Falkegård https://orcid.org/0000-0003-2853-2749
Samu Mäntyniemi https://orcid.org/0000-0002-3367-6280

Author contributions
Conceptualization: AR, HP, JE, PO, SM
Data curation: PO, MF
Formal analysis: AR, HP
Funding acquisition: JE
Investigation: JE, PO, MF
Methodology: AR, HP, SM
Project administration: JE
Resources: JE, PO, MF, SM
Software: AR, HP, SM
Supervision: HP
Visualization: AR
Writing – original draft: AR
Writing – review & editing: AR, HP, JE, PO, SM

Competing interests
The authors declare there are no competing interests.

Funding information
Funding for this project was provided by Luke's thematic funding (project PINKS), Research Council of Finland (project 346981, DEATNU), and Biodiversa+ (project 2021-987, RecoSal).

Supplementary material

Supplementary data are available with the article at https://doi.org/10.1139/cjfas-2024-0309.

References

Alglave, B., Rivot, E., Etienne, M.-P., Woillez, M., Thorson, J.T., and Vermard, Y. 2022. Combining scientific survey and commercial catch data to map fish distribution. ICES J. Mar. Sci. **79**: 1133–1149. doi:10.1093/icesjms/fsac032.

Alioravainen, N., Orell, P., and Erkinaro, J. 2023. Long-term trends in freshwater and marine growth patterns in three Sub-Arctic Atlantic salmon populations. Fishes, **8**: 441. doi:10.3390/fishes8090441.

Anon. 2024. Status of the Tana/Teno River salmon populations in 2023. Tana Monitoring and Research Group.

Beddington, J.R., Agnew, D.J., and Clark, C.W. 2007. Current problems in the management of marine fisheries. Science, **316**: 1713–1716. doi:10.1126/science.1137362. PMID: 17588923.

Borgstrøm, R., Opdahl, J., Svenning, M.-A., Länsman, M., Orell, P., Niemelä, E., et al., 2010. Temporal changes in ascendance and in-season exploitation of Atlantic salmon, *Salmo salar*, inferred by a video camera array. Fish. Manage. Ecol. **17**: 454–463. doi:10.1111/j.1365-2400.2010.00744.x.

Boswell, K.M., Wilson, M.P., and Cowan, J.H. 2008. A semiautomated approach to estimating fish size, abundance, and behavior from dual-frequency identification sonar (DIDSON) data. N. Am. J. Fish. Manage. **28**: 799–807. doi:10.1577/M07-116.1.

Buckland, S.T., Newman, K.B., Fernández, C., Thomas, L., and Harwood, J. 2007. Embedding population dynamics models in inference. Stat. Sci. **22**: 44–58. doi:10.1214/088342306000000673.

Burwen, D.L., Fleischman, S.J., and Miller, J.D. 2010. Accuracy and precision of salmon length estimates taken from DIDSON sonar images. Trans. Am. Fish. Soc. **139**: 1306–1314. doi:10.1577/T09-173.1.

Davidsen, J., Svenning, M.-A., Orell, P., Yoccoz, N., Dempson, J.B., Niemelä, E., et al., 2005. Spatial and temporal migration of wild Atlantic salmon smolts determined from a video camera array in the sub-Arctic River Tana. Fish. Res. **74**: 210–222. doi:10.1016/j.fishres.2005.02.005.

Denwood, M.J. 2016. runjags: an R package providing interface utilities, model templates, parallel computing methods and additional distributions for MCMC models in JAGS. J. Stat. Softw. **71**: 1–25. doi:10.18637/jss.v071.i09.

Domaas, S., Orell, P., Kytökorpi, M., Myklebost, M.R., Erkinaro, J., and Gjelland, K.Ø. 2024. Evaluation of fish trap and guiding fence efficiency in the River Tana in 2023, 57. Norwegian Institute for Nature Research (NINA).

Erkinaro, J., and Orell, P. 2022. Pink salmon (*Oncorhynchus gorbuscha*) in the northernmost Atlantic area—with special emphasis on the River

Teno/Tana, Finland/Norway. *In* Pink salmon and Red Skin Disease: emerging threats for Atlantic salmon, blue book series. *Edited by* K. Whelan and T.A. Mo. AST. pp. 8–10.

Erkinaro, J., Czorlich, Y., Orell, P., Kuusela, J., Falkegård, M., Länsman, M., et al., 2019. Life history variation across four decades in a diverse population complex of Atlantic salmon in a large subarctic river. Can. J. Fish. Aquat. Sci. **76**: 42–55. doi:10.1139/cjfas-2017-0343.

Erkinaro, J., Erkinaro, H., and Niemelä, E. 2017. Road culvert restoration expands the habitat connectivity and production area of juvenile Atlantic salmon in a large subarctic river system. Fish. Manage. Ecol. **24**: 73–81. doi:10.1111/fme.12203.

Erkinaro, J., ØKland, F., Moen, K., Niemelä, E., and Rahiala, M. 1999. Return migration of Atlantic salmon in the River Tana: the role of environmental factors. J. Fish Biol. **55**: 506–516. doi:10.1111/j.1095-8649.1999.tb00695.x.

Erkinaro, J., Orell, P., Kytökorpi, M., Pohjola, J.-P., and Power, M. 2024. Active feeding of downstream migrating juvenile pink salmon (*Oncorhynchus gorbuscha*) revealed in a large Barents Sea river using diet and stable isotope analysis. J. Fish Biol. **104**: 797–806. doi:10.1111/jfb.15625. PMID: 37986023.

Erkinaro, J., Orell, P., Pohjola, J.-P., Kytökorpi, M., Pulkkinen, H., and Kuusela, J. 2022. Development of invasive pink salmon (*Oncorhynchus gorbuscha* Walbaum) eggs in a large Barents Sea river. J. Fish Biol. **101**: 1063–1066. doi:10.1111/jfb.15157. PMID: 35790001.

Forseth, T., Fiske, P., Barlaup, B., Gjøsæter, H., Hindar, K., and Diserud, O. 2013. Reference point based management of Norwegian Atlantic salmon populations. Environ. Conserv. **40**: 356–366. doi:10.1017/S0376892913000416.

Fox, D.S., and Starr, R.M. 1996. Comparison of commercial fishery and research catch data. Can. J. Fish. Aquat. Sci. **53**: 2681–2694. doi:10.1139/f96-230.

Gelman, A., and Rubin, D.B. 1992. Inference from iterative simulation using multiple sequences. Stat. Sci. **7**: 457–472. doi:10.1214/ss/1177011136.

Gelman, A., Carlin, J.B., Stern, H.S., Dunson, D.B., Vehtari, A., and Rubin, D.B. 2013. Bayesian data analysis. *In* Chapman & Hall/CRC texts in statistical science. 3rd ed. Taylor & Francis.

Gurney, W.S.C., Brennan, L.O., Bacon, P.J., Whelan, K.F., O'Grady, M., Dillane, E., and McGinnity, P. 2014. Objectively assigning species and ages to salmonid length data from dual-frequency identification sonar. Trans. Am. Fish. Soc. **143**: 573–585. doi:10.1080/00028487.2013.862185.

Helminen, J., and Linnansaari, T. 2021. Object and behavior differentiation for improved automated counts of migrating river fish using imaging sonar data. Fish. Res. **237**: 105883. doi:10.1016/j.fishres.2021.105883.

Helminen, J., Dauphin, G.J.R., and Linnansaari, T. 2020. Length measurement accuracy of adaptive resolution imaging sonar and a predictive model to assess adult Atlantic salmon (*Salmo salar*) into two size categories with long-range data in a river. J. Fish Biol. **97**: 1009–1026. doi:10.1111/jfb.14456. PMID: 32652539.

Horne, J. 2008. Acoustic approaches to remote species identification: a review. Fish. Oceanogr. **9**: 356–371. doi:10.1046/j.1365-2419.2000.00143.x.

Kaev, A.M., and Rudnev, V.A. 2007. Population dynamics of pink salmon *Oncorhynchus gorbuscha* (Salmonidae) from the southeastern coast of Sakhalin island. J. Ichthyol. **47**: 228–240. doi:10.1134/S0032945207030058.

Kandimalla, V., Richard, M., Smith, F., Quirion, J., Torgo, L., and Whidden, C. 2022. Automated detection, classification and counting of fish in fish passages with deep learning. Front. Mar. Sci. **8**. doi:10.3389/fmars.2021.823173.

Karppinen, P., Erkinaro, J., Niemelä, E., Moen, K., and Økland, F. 2004. Return migration of one-sea-winter Atlantic salmon in the River Tana. J. Fish Biol. **64**: 1179–1192. doi:10.1111/j.0022-1112.2004.00380.x.

Kuparinen, A., Mäntyniemi, S., Hutchings, J.A., and Kuikka, S. 2012. Increasing biological realism of fisheries stock assessment: towards hierarchical Bayesian methods. Environ. Rev. **20**: 135–151. doi:10.1139/a2012-006.

Lamberg, A., and Imsland, A.K.D. 2022. Using merged pre-fishery abundance as a parameter evaluating the status of atlantic salmon and anadromous brown trout populations: a Norwegian case study. Fishes, **7**: 264. doi:10.3390/fishes7050264.

Mäntyniemi, S., Romakkaniemi, A., and Arjas, E. 2005. Bayesian removal estimation of a population size under unequal catchability. Can. J. Fish. Aquat. Sci. **62**: 291–300. doi:10.1139/f04-195.

NASCO. 1998. Agreement on the adoption of a precautionary approach. NASCO Council Document CNL 98, 3.

NASCO. 2009. NASCO Guidelines for the management of Salmon Fisheries. NASCO Council Document CNL 09, 12.

Niemelä, E., Makinen, T.S., Moen, K., Hassinen, E., Erkinaro, J., Lansman, M., and Julkunen, M. 2000. Age, sex ratio and timing of the catch of kelts and ascending Atlantic salmon in the subarctic River Teno. J. Fish Biol. **56**: 974–985. doi:10.1111/j.1095-8649.2000.tb00886.x.

O'Hagan, A. 1994. Bayesian inference, Kendall's advanced theory of statistics. E. Arnold, London.

Orell, P., and Erkinaro, J. 2007. Snorkelling as a method for assessing spawning stock of Atlantic salmon, *Salmo salar*. Fish. Manage. Ecol. **14**: 199–208. doi:10.1111/j.1365-2400.2007.00541.x.

Orell, P., Erkinaro, J., and Karppinen, P. 2011. Accuracy of snorkelling counts in assessing spawning stock of Atlantic salmon, *Salmo salar*, verified by radio-tagging and underwater video monitoring. Fish. Manage. Ecol. **18**: 392–399. doi:10.1111/j.1365-2400.2011.00794.x.

Orell, P., Erkinaro, J., Kanniainen, T., and Kuusela, J. 2017. Migration behavior of sea trout (*Salmo trutta*, L.) in a large sub-arctic river system: evidence of a two-year spawning migration. Troubador Publishing Limited.

Orell, P., Erkinaro, J., Svenning, M., Davidsen, J., and Niemelä, E. 2007a. Synchrony in the downstream migration of smolts and upstream migration of adult Atlantic salmon in the subarctic River Utsjoki. J. Fish Biol. **71**: 1735–1750. doi:10.1111/j.1095-8649.2007.01641.x.

Orell, P., Länsman, M., Kylmäaho, M., Niemelä, E., Erkinaro, J., Brørs, S., et al. 2007b. Teno- ja näätämöjoen lohikantojen seurantatutkimukset vuosina 2001–2005. Riista- ja kalatalouden tutkimuslaitos.

Pierce, D.A. 1973. On some difficulties in a frequency theory of inference. Ann. Stat. **1**: 241–250. doi:10.1214/aos/1176342362.

Plummer, M. 2003. JAGS: a program for analysis of bayesian graphical models using gibbs sampling. Working Papers.

Prévost, E., and Chaput, G. 2001. Stock, recruitment and reference points : assessment and management of Atlantic salmon.

Pulkkinen, H., Orell, P., Erkinaro, J., and Mäntyniemi, S. 2019. Bayesian arrival model for Atlantic salmon smolt counts powered by environmental covariates and expert knowledge. Can. J. Fish. Aquat. Sci. **77**: 462–474. doi:10.1139/cjfas-2018-0352.

Romakkaniemi, A. (*Editor*). 2015. Best practices for the provision of prior information for Bayesian stock assessment. doi:10.17895/ICES.PUB.5496.

Sandlund, O.T., Berntsen, H.H., Fiske, P., Kuusela, J., Muladal, R., Niemelä, E., et al., 2019. Pink salmon in Norway: the reluctant invader. Biol. Invasions, **21**: 1033–1054. doi:10.1007/s10530-018-1904-z.

Tang, Z., Li, J., Wang, Z., Huang, J., Li, Y., and Wang, C. 2023. Research on underwater target measurement technology based on sonar image and artificial landmark. Multimed. Tools Appl. **82**: 29713–29732. doi:10.1007/s11042-023-14822-2.

Thorstad, E.B., Økland, F., Aarestrup, K., and Heggberget, T.G. 2008. Factors affecting the within-river spawning migration of Atlantic salmon, with emphasis on human impacts. Rev. Fish Biol. Fish. **18**: 345–371. doi:10.1007/s11160-007-9076-4.

Vähä, J.-P., Erkinaro, J., Falkegård, M., Orell, P., and Niemelä, E. 2017. Genetic stock identification of Atlantic salmon and its evaluation in a large population complex. Can. J. Fish. Aquat. Sci. **74**: 327–338. doi:10.1139/cjfas-2015-0606.

Wei, Y., Duan, Y., and An, D. 2022. Monitoring fish using imaging sonar: capacity, challenges and future perspective. Fish Fish. **23**: 1347–1370. doi:10.1111/faf.12693.

Appendix A. Lognormal approximation for Dirichlet distribution with AR(1) autoregression

Let $q_{1:n}$ to be a set of Dirichlet-distributed parameters:

$$q_{1:n} \sim \text{Dirich}\left(\psi_1 \cdot \eta \cdot N,\ \ldots,\ \psi_n \cdot \eta \cdot N\right)$$

where ψ_i is the expected probability for index i, η is the overdispersion parameter between 0 and 1, and N is the total amount. The expected probability ψ_i is calculated using AR(1) autoregression from expert-defined expected value μ_i and autoregression coefficient ϕ. With a set of lognormally distributed z_i, Dirichlet distribution can be approximated for index $i = 1$ by assuming

(A1) $$z_1 = \exp\left(\log(\mu_1) - 0.5 \cdot \log\left(\frac{1}{\mu_1 \cdot \eta \cdot N} + 1\right) + v_1\right)$$

(A2) $$v_1 \sim N\left(0, \sigma_1^2\right)$$

(A3) $$\sigma_1^2 = \log\left(\frac{1}{\mu_1 \cdot \eta \cdot N} + 1\right)$$

and for indexes$i = [2 : n]$:

(A4) $$z_i = \exp\left(\log(\mu_i) - 0.5 \cdot \log\left(\frac{1}{\mu_i \cdot \eta \cdot N} + 1\right) + v_i\right)$$

(A5) $$v_i \sim \mathrm{N}\left(\phi \cdot v_{i-1}, \sigma_i^2\right)$$

(A6) $$\sigma_i^2 = \left(1 - \phi^2\right) \cdot \log\left(\frac{1}{\mu_i \cdot \eta \cdot N} + 1\right)$$

Values for q_i are calculated, after $z_{1:n}$ have been resolved:

(A7) $$q_i = \frac{z_i}{\sum_{i=1}^{n} z_i}$$

Appendix B. Lognormal approximation for Dirichlet distribution

Let $q_{1:n}$ to be a set of Dirichlet-distributed parameters:

$$q_{1:n} \sim \mathrm{Dirich}\left(\mu_1 \cdot \eta,\ \mu_2 \cdot \eta,\ \ldots,\ \mu_n \cdot \eta\right)$$

where μ_i is the expected probability for index i and η is the dispersion parameter. With a set of lognormally distributed z_i with location parameter M and precision T, which are parametrized with the expected value μ and dispersion parameter η. Dirichlet distribution can be approximated by assuming

(B1) $$q_i = z_i / \sum_{i=1}^{n} z_i$$

(B2) $$z_i \sim \mathrm{logN}\left(M_i,\ T_i\right)$$

(B3) $$M_i = \log(\mu_i) - 0.5/T_i$$

(B4) $$T_i = \log\left(\frac{1}{\mu_i \cdot \eta} + 1\right)^{-1}$$