RESEARCH ARTICLE

Methods in Ecology and Evolution | BRITISH ECOLOGICAL SOCIETY

Integrating genetic analysis of mixed populations with a spatially explicit population dynamics model

Rebecca Whitlock[1] [ID] | Samu Mäntyniemi[2] | Stefan Palm[1] | Marja-Liisa Koljonen[3] |
Johan Dannewitz[1] | Johan Östergren[1]

[1]Institute of Freshwater Research, Swedish University of Agricultural Sciences, Drottningholm, Sweden

[2]Department of Environmental Sciences, Bayesian Environmental Modelling Group, University of Helsinki, Helsinki, Finland

[3]Natural Resources Institute Finland (Luke), Helsinki, Finland

Correspondence
Rebecca Whitlock
Email: rebecca.whitlock@slu.se

Abstract

1. Inferring the dynamics of populations in time and space is a central challenge in ecology. Intra-specific structure (for example genetically distinct sub-populations or metapopulations) may require methods that can jointly infer the dynamics of multiple populations. This is of particular importance for harvested species, for which management must balance utilization of productive populations with protection of weak ones.

2. Here we present a novel method for simultaneous learning about the spatio-temporal dynamics of multiple populations that combines genetic data with prior information about abundance and movement, akin to an integrated population modelling approach. We apply the Bayesian genetic mixed stock analysis to 17 wild and 10 hatchery-reared Baltic salmon (*S. salar*) stocks, quantifying uncertainty in stock composition in time and space, and in population dynamics parameters such as migration timing and speed.

3. The genetic data were informative about stock-specific movement patterns, updating priors for migration path, timing and speed. Use of a population dynamics model allowed robust interpolation of expected catch composition at areas and times with no genetic observations. Our results indicate that the commonly used "equal prior probabilities" assumption may not be appropriate for all mixed stock analyses: incorporation of prior information about stock abundance and movement resulted in more plausible and precise estimates of mixture compositions in time and space.

4. The model we present here forms the basis for optimizing the spatial and temporal allocation of harvest to support the management of mixed populations of migratory species.

KEYWORDS
Baltic salmon, Bayesian approach, genetic mixed stock analysis, integrated population models, spatial models

1 | INTRODUCTION

Understanding population-specific space use patterns has broad relevance to a number of questions in ecology, including metapopulation dynamics (Chandler et al., 2015; Gilpin, 1996; Hanski, 1998; Hanski & Thomas, 1994), the environmental underpinnings of spatial distributions that shape responses to future environmental change (Hazen et al., 2013) and the mechanisms for migration (Dittman & Quinn, 1996; Somveille, Rodrigues, & Manica, 2015; Whitlock et al., 2015). The dynamics of populations in time and space are also of central

importance to the management of exploited migratory populations, which may be subject to the same harvesting effort, despite potential differences in abundance and productivity. Simultaneous harvest of multiple populations, exemplified by mixed stock fisheries, presents both benefits and challenges for management and conservation. Yields may be more stable across years because of the buffering effect of interpopulation variability in life history and phenology (Schindler et al., 2010; Utter & Ryman, 1993). On the other hand, targeting multiple populations can lead to high harvest rates on depleted or endangered populations (Branch, Lobo, & Purcell, 2013; Clayton, Keeling, & Milner-Gulland, 1997; Crozier et al., 2004). Catches should thus be apportioned among populations in order to quantify harvest rates and allow implementation of measures to fully utilize productive populations, while protecting weak ones.

Mixed stock fisheries serve as a widespread example of spatially structured population interactions. Typically, the contributions of different populations (stocks) to catches are unknown, but must be quantified to estimate stock-specific harvest rates and specific appropriate management measures. Stock composition in mixed catches has been estimated using a variety of approaches (see Bradbury et al., 2015; Cadrin, Friedland, & Waldman, 2005), including tagging (e.g. Brodziak, 1993; Hoenig, Latour, & Olney, 2008), age structure (e.g. Chasco, Hilborn, & Punt, 2007) and genetic methods (Grant, Milner, Krasnowski, & Utter, 1980; Koljonen, 2006; Milner, Teel, Utter, & Burley, 1981; Utter & Ryman, 1993). Genetic methods use markers such as allozymes, mitochondrial DNA, microsatellites or single nucleotide polymorphism (SNPs) that show variation among stocks. Genetic mixed stock analysis (MSA) estimates the proportions of contributing baseline stocks to a mixed sample and may also assign individuals of unknown origin within the mixture to a stock or stock group. It essentially compares an individual's observed multilocus genotype with a "baseline" that comprises multilocus genotype distributions from individuals of known stock of origin.

Genetic MSA has traditionally been performed as a separate analysis, external to models of population dynamics, whereby estimates of stock composition from MSA are used as inputs to population models (either as observations or as known fixed quantities) (e.g. Bradbury et al., 2016; ICES, 2017; Michielsens, Mäntyniemi, & Koljonen, 2004). Integrated population models provide a statistical framework for simultaneously analysing multiple datasets (Abadi, Gimenez, Arlettaz, & Schaub, 2010; Besbeas, Freeman, Morgan, & Catchpole, 2002; Buckland, Newman, Thomas, & Koesters, 2004; Schaub, Gimenez, Sierro, & Arlettaz, 2007), linking changes in demographic rates and abundance via a model of population dynamics while accounting for both process and observation error (Besbeas & Morgan, 2014; Chandler & Clark, 2014; Schaub & Abadi, 2011). Information from multiple datasets is combined using the product of likelihoods for independent data sources (Maunder & Punt, 2013; Schaub & Abadi, 2011). Integrated population models are now widespread in the field of fisheries stock assessment, where they provide a rigorous framework for dealing with the multiple data types and uncertainty related to sampling and process variability that are frequently encountered in this context (Maunder & Punt, 2013).

Frequentist and Bayesian implementations have been developed for both genetic MSA and individual assignment (e.g. Kalinowski, Manlove, & Taper, 2007; Manel, Gaggiotti, & Waples, 2005; Millar, 1987; Pella & Masuda, 2001) and integrated population models (Besbeas, Lebreton, & Morgan, 2003; Besbeas et al., 2002; Brooks, King, & Morgan, 2004; Taylor, McAllister, Lawson, Carruthers, & Block, 2011). In the context of genetic MSA, Bayesian approaches assign individuals to populations probabilistically, accounting for uncertainty about mixture composition and population-specific allele frequencies (Bolker, Okuyama, Bjorndal, & Bolten, 2007; Pella & Masuda, 2001; Pritchard, Stephens, & Donnelly, 2000). Bayesian statistical methods are more robust to estimation problems caused by missing data and rare alleles (Corander, Marttinen, & Mäntyniemi, 2006) and offer improved performance in both mixture analysis and individual assignment (Bolker et al., 2007; Bradbury et al., 2015). Bayesian approaches also provide solutions for cases where the number of genetically diverged sources contributing to the observed mixture data is unknown (Corander, Waldmann, & Sillanpää, 2003, 2004; Corander et al., 2006; Dawson & Belkhir, 2001).

A key advantage of the Bayesian approach in general is the possibility to incorporate prior information, for example about demographic rates, or individual origins, which may enhance the statistical power of individual assignment. However, a majority of MSA studies published to date make an assumption of equal prior proportions of different populations in the mixture (e.g. Bradbury et al., 2015; Corander et al., 2003; Pella & Masuda, 2001; but see Bolker et al., 2007; and Dann, Habicht, Baker, and Seeb, 2013). Here, we use prior knowledge about fish movement and stock-specific pre-migration abundances from earlier studies, allowing information flow in both directions: stock composition estimates from MSA are used to learn about population dynamics parameters and vice versa.

Atlantic salmon (*Salmo salar* L.) is a keystone migratory species in the Baltic Sea and an important regional resource, supplying provisioning services (i.e. catch), cultural services (recreational fishing) and ecosystem services (nutrient cycling, river sediment reduction) (Karlsson & Karlström, 1994; Kulmala et al., 2013). Wild Baltic salmon populations persist in less than 30 of ~100 former salmon rivers (ICES, 2015). Damming, habitat destruction, pollution and intensive fishing have been identified as the main causes of population decline (ICES, 2015; Kuikka, Vanhatalo, Pulkkinen, Mäntyniemi, & Corander, 2014; Kulmala et al., 2013). In addition to the remaining wild salmon rivers, hatchery-produced salmon are released annually in several exploited rivers with extinct or severely depleted wild populations, to compensate for the loss of natural production.

Sixteen wild Baltic salmon river stocks are currently assessed by the Baltic Salmon and Trout Assessment Working Group within ICES (ICES, 2015); these have been shown to support genetically differentiated populations (e.g. Koljonen, Jansson, Paaver, Vasin, & Koskiniemi, 1999; Säisä et al., 2005). Baltic salmon have a complex life history, spending 1–5 years in rivers before undergoing physiological adaptations for a saltwater environment (smoltification) and migrating to sea. Individuals usually spend 1 to 4 years feeding at sea before returning to their natal river to spawn (Karlsson & Karlström, 1994) and may repeat the feeding migration to spawn multiple times. While river fisheries generally target specific stocks as a result of homing behaviour,

coastal fisheries targeting reproductively mature salmon on their spawning migration typically exploit a mixture of stocks (Crozier et al., 2004). The composition of samples taken at different coastal locations and times is expected to reflect the differential abundances, migration timings and migration routes of multiple source populations (e.g. Jutila, Jokikokko, Kallio-Nyberg, Saloniemi, & Pasanen, 2003; Kallio-Nyberg, Romakkaniemi, Jokikokko, Saloniemi, & Jutila, 2015).

Quantitative assessment of wild Baltic salmon stocks indicates large variation in status; several stocks have recovered during the last few decades and are currently underexploited according to the present management target, while others are still depleted with a low probability of recovery under status quo conditions (ICES, 2017). Management measures are prescribed and implemented at an aggregate stock level, but the large difference in the status of stocks suggests that controlling harvest rates on a stock-specific basis would allow the most efficient and sustainable use of this resource. To maintain coastal and marine mixed stock fisheries, new assessment tools are needed to integrate genetic information from catch samples with other types of data to inform stock-specific management measures.

We present a hierarchical Bayesian model that utilizes knowledge about migration timing, speed and direction from earlier tagging studies and information about pre-migration stock abundances from ICES' assessment to provide a prior for the expected stock composition of mixed Baltic salmon catches in space and time. Our analysis resembles the integrated population model framework, combining raw genetic data (microsatellite allele frequencies and genotypes) with auxiliary data distinguishing reared and wild populations and linking both to the spatio-temporal dynamics of multiple populations. We evaluate the effect of the stock composition prior on estimated mixture proportions and individual assignments in a genetic MSA. We also quantify the improvements in accuracy that can be achieved using auxiliary data on stock of origin group (reared vs. wild salmon stocks), when mixed catches comprise partly of reared fish.

2 | MATERIALS AND METHODS

2.1 | Study system

We apply Bayesian genetic MSA to mixed catches from the coastal fisheries in the Gulf of Bothnia and Baltic proper, which target salmon on their spawning migrations (Figure 1). The central and southern parts of the Baltic's main basin form the primary feeding areas for Swedish and Finnish salmon stocks (Carlin, 1969; Karlsson & Karlström, 1994; Siira, Erkinaro, Jounela, & Suuronen, 2009). In late spring, Gulf of Bothnia salmon stocks leave the feeding grounds and begin their spawning migrations, moving northwards (Siira et al., 2009).
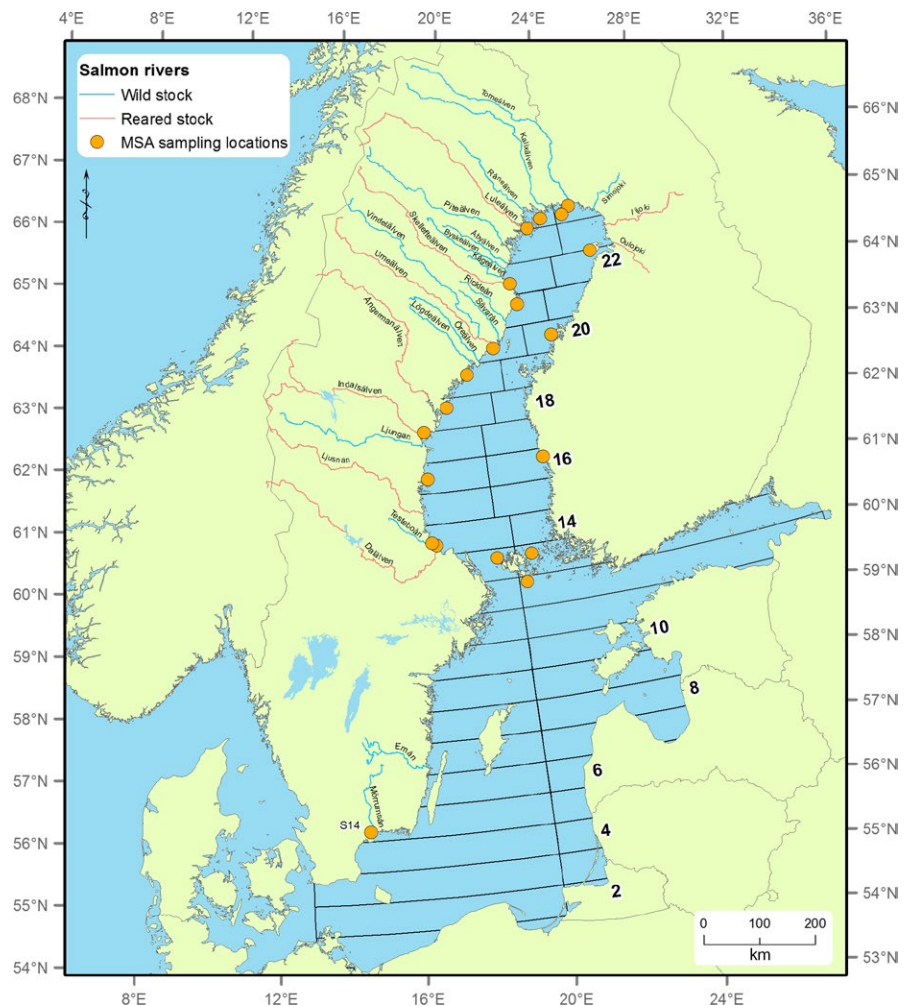

FIGURE 1 Map of the study area, wild and reared salmon rivers in the analysis and model spatial boxes. Primary (latitudinal) box numbers are shown to the right of the Baltic Sea

2.2 | Data

2.2.1 | Genetic baseline

The current baseline dataset for Atlantic salmon stocks in the Baltic Sea includes information on 17 DNA microsatellite loci for 39 Baltic salmon stocks, totalling 4,453 individuals (ICES, 2015; Koljonen, 2006). In this study, however, we only use the baseline samples for Finnish and Swedish Baltic salmon stocks, excluding stocks in the Gulf of Finland, which follow a partially different migratory route (Karlsson & Karlström, 1994), and have not been observed in the Gulf of Bothnia (Koljonen, 2006). The baseline used in this study thus comprises 3,444 individuals from 27 salmon stocks, of which 17 are wild and 10 reared (Table 1), spanning from River Torne in the north to River Mörrumsån in the south (Figure 1).

TABLE 1 Gulf of Bothnia and Baltic proper baseline (Swedish and Finnish river stocks) used in the analysis. (R) after the stock names indicates a reared stock

River	Sample size (number of individuals)	Fin-clipping prior proportion
Abyälven	102	0.01
Ångermanälven (R)	79	0.63
Byskeälven	105	0.01
Dalälven (R)	98	0.95
Emån	182	0.01
Iijoki (R)	105	0.02
Indalsälven (R)	144	0.95
Kågeälven	44	0.01
Kalixälven	336	0.01
Ljungan	137	0.01
Ljusnan (R)	135	0.95
Lögdeälven	102	0.01
Luleälven (R)	131	0.95
Mörrumsån	321	0.01
Öreälven	54	0.01
Oulujoki (R)	167	0.01
Piteälven	53	0.01
Råneälven	183	0.01
Rickleån	52	0.01
Sävarån	74	0.01
Simojoki	174	0.01
Skellefteälven (R)	58	0.95
Testeboån	79	0.01
Torneälven	333	0.01
Torneälven Hatchery (R)	109	0.03
Umeälven (R)	87	0.95
Vindelälven	149	0.01

2.2.2 | Mixture data

We analyse a total of 2,058 adult individuals sampled from coastal trap nets between 5 May and 11 August in 2014 (Table 2; Figure 1) (Östergren et al., 2015). In addition to scale samples for DNA microsatellite analysis (17 loci), we obtained individual data on catch date, location and adipose fin status (present/absent). Alleles found in a mixture sample but not in the baseline are excluded from analyses as in Bolker et al. (2007). Samples from traps located in the same model box (see Figure 1) were combined for the purposes of statistical analysis.

Hatcheries in Sweden routinely remove the adipose fin from hatchery-reared salmon smolts released into exploited rivers with some exceptions (e.g. for experimental purposes), thus providing a further means to distinguish between wild and reared stocks (Table 2). This is expected to be particularly useful for rivers that support both a wild and a hatchery stock, and for some geographically close wild and reared stocks, which tend to be genetically similar. We investigate the utility of adipose fin presence/absence as a means to improve discrimination between genetically similar fish from reared and wild stocks in the same or neighbouring river systems.

2.3 | Genetic analyses

Total DNA was extracted from dry scales and tissue samples (fin clips), followed by PCR and genotyping of 17 polymorphic microsatellite markers. Baseline (river stock) samples were genetically processed in Finland (Jarmo Koskiniemi, University of Helsinki) and Sweden (SLU Aqua); Swedish and part of the Finnish coastal catch samples were analysed in Sweden. To assure fully comparable genotypes scored at the two laboratories, a marker calibration (i.e. replicated allele length scoring of same individuals) was performed. Details on laboratory procedures, microsatellites screened and marker calibrations are provided in the Supporting Information.

2.4 | Genetic MSA for Baltic salmon

We present results from three Bayesian genetic MSAs that differ in the amount of prior information and types of data used. We use a slightly modified version of the genetic MSA approach introduced by Pella and Masuda (2001) (the MSA model is described below under Section 2.6). In model 1 (uninformative prior, genetic data only), we follow a standard assumption in MSA and assume equal prior probabilities for the proportions of different baseline stocks for each mixture sample. In model 2 (informative prior, genetic data only), we integrate the genetic data with the population dynamics model prior for spatio-temporal stock composition. Models 1 and 2 do not use auxiliary information about stock type from adipose fin-clipping. In model 3 (informative prior, genetic and fin-clipping data), we use the population dynamics prior and incorporate observation models for both multilocus genotypes and the presence of an adipose fin on sampled fish, indicating whether they are of wild or hatchery origin.

TABLE 2 Mixture samples collected in 2014, with sample size and the proportion of individuals with fin-clipping information

Sampling location and model box	Sample size (number of individuals)	Proportion with fin-clipping information	Proportion with adipose fin intact
Sweden S1 24, 1	58	1.00	0.93
Sweden S2 24, 2	135	0.97	0.93
Sweden S3 24, 2	49	1.00	1.00
Sweden S4 23, 1	50	1.00	0.00
Finland F1 23, 2	183	0.01	0.50
Sweden S5 22, 1	178	0.93	0.02
Sweden S6 21, 1	17	1.00	0.29
Sweden S7 20, 1	33	1.00	0.61
Finland F2 20, 2	135	0.99	0.94
Sweden S8 19, 1	108	0.99	0.81
Sweden S9 18, 1	141	0.94	0.83
Sweden S10 17, 1	37	1.00	0.19
Sweden S11 16, 1	91	1.00	0.16
Finland F3 16, 2	96	0.99	0.93
Sweden S12 14, 1	63	0.98	0.23
Sweden S13 14, 1	86	0.97	0.12
Finland F4 13, 2	8	1.00	0.88
Finland F5 13, 1	185	0.99	0.98
Finland F6 12, 2	127	1.00	0.98
Sweden S14 5, 1	278	0.00	NA

2.5 | Prior for mixture stock composition

The spatially and temporally structured population dynamics model provides a prior for mixture stock composition, defined by the predicted relative abundances of migrating salmon from different stocks in a given time step and area. We divide the Baltic Sea latitudinally into 24 areas, most of which span 0.5° latitude, and longitudinally into east and west domains within each latitudinal box, creating a total of 48 model areas (Figure 1). We model the period corresponding to the spawning migration for Baltic salmon, from mid-April to the end of August, with a fortnightly time step. The timing of the start of migration from the Baltic Sea feeding grounds and the proportion of fish that depart in each time step are allowed to differ among stocks.

For flexibility, the model is set up to enable movement in two directions, the first of which corresponds to the main direction of migration (in our case south to north), and the second of which allows movement perpendicular to the first direction (in our case east–west and west–east). In the following text, we refer to the main direction of migration as *primary movement* and other movements as *secondary movement*. To implement unidirectional migration (e.g. a spawning run in a river), secondary movement rates could simply be set equal to 0.

In the following equations, subscript y denotes year, t denotes time step, a denotes age and s denotes stock, $s = 1, ... ,S$ (below we present results from a single year of genetic data, but the model's implementation allows for multiple years). We define two stock groups (g) in the migration model: wild stocks, $g = 1$; and reared stocks, $g = 2$. Subscripts i and j are used to index latitudinal boxes (1 to 24); i denotes originating box and j denotes destination box, while k is used to index longitudinal boxes (1 to 2). For each stock, we additionally denote the latitudinal box corresponding to the natal river as SP1 and the longitudinal box as SP2 (Table 3).

Following the logic of Bayesian inference, we define the full probability model for all the variables that are unknown prior to observing the dataset by specifying either a marginal or conditional distribution for each variable of the model. Marginal prior distributions are denoted as $x \sim Distribution$ (parameters) and distributions that are thought to depend on other uncertain variables are denoted as $x|parameters \sim Distribution$ (parameters) or as a deterministic equation if no uncertainty is assumed.

2.5.1 | Initial conditions $t = 1$

The number of salmon from stock s in western box i on 15 April ($t = 1$) is given by:

$$N_{y,i,1,s,1} = B_{y,s}\gamma_i\xi_s$$

The number of salmon from stock s in eastern box i on 15 April is given by:

$$N_{y,i,2,s,1} = B_{y,s}\gamma_i(1-\xi_s)$$

where γ_i is the proportion of individuals in primary box i on 15 April and ξ_s is the proportion of individuals in secondary box 1 (i.e. in the

TABLE 3 Indices and symbols used in the population dynamics model

Symbol	Description
Indices	
a	Age
AU	ICES assessment unit (1–4)
g	Stock group: 1 = wild, 2 = reared
i, j	Latitudinal box, 1–24
k	Longitudinal box, 1–2
SP1	Latitudinal spawning area box, 1–24
SP1	Longitudinal spawning area box, 1–2
Model variables	
$N_{y,i,k,s,t}$	Number of salmon from stock s in primary area i and secondary area k in time step t of year y
$N_{init_{y,i,k,s}}$	Number of salmon from stock s in primary area i and secondary area k on 15 April in year y
$N_{mov_{y,i,j,k,s,t}}$	Number of migrating salmon from stock s moving from primary area i to primary area j in secondary area k in time step t of year y
$N_{new_{y,i,k,s,t}}$	Number of migrating salmon in primary area i and secondary area k after movement in time step t of year y
$N_{start_{y,i,k,s,t}}$	Number of salmon in primary area i and secondary area k that begin their migration in time step t of year y
Model parameters	
h	Rate at which salmon move from the spawning box into the river in one time step
Z_g	Total annual instantaneous mortality rate (the sum of fishing and natural mortality) for group g
$\delta_{s,t}$	Proportion of salmon from stock s that begins migrating at time t
γ_i	Proportion of individuals in primary area i on 15 April
ξ_s	Proportion of individuals in secondary area 1 on 15 April for stock s
$\pi_{i,j,s}$	Probability of moving from box i to box j in the primary direction (north) for stock s
$\pi_{N,j}$	Probability of moving $j-3$ boxes in the primary direction (north) from boxes north of the spawning area
$\pi_{SP,j}$	Probability of moving $j-2$ boxes in the primary direction (north) from the spawning area
$\rho_{k,Bl,G}$	Probability of moving from secondary area k to the other secondary area, for secondary area block Bl and movement mode G
$\rho_{SP,k,SP2}$	Probability of moving from secondary area k to the other secondary area, for a stock with secondary spawning area SP2

western domain of box i). The initial proportions of salmon from each stock in each primary box are assumed to follow a Dirichlet distribution:

$$\gamma_{1:I} \sim \text{Dirichlet}(\alpha_{\gamma,1:I}),$$

We based the initial spatial distribution of mature salmon on 15 April on the distribution of tag recaptures reported in Carlin (1969) (see Supporting Information for details and Table S3 for $\alpha_{\gamma,i}$ values).

The number of fish that begin their migration at time $t = 1$ is then given by:

$$N_{start_{y,i,k,s,1}} = N_{y,i,k,s,1} e^{-Z_{g(s)}/T} \delta_{s,1}$$

where $Z_{g(s)}$ is the annual rate of total instantaneous mortality (the sum of fishing and natural mortality) for the group to which stock s belongs. We use a discretized normal distribution to describe the proportion of salmon from stock s that begins migrating (i.e. moves from the initial spatial distribution) at time t, $\delta_{s,t}$, starting from the initial spatial distribution. Mean migration start date is allowed to occur earlier than the first date modelled (15 April), in which case the proportion of animals migrating in the first model time step is given as the sum of the proportions starting their migrations over the previous and current time steps (see Supporting Information for details).

We model movement as a deterministic process: primary movement is assumed to follow one of three sets of rules, according to the location of a fish in relation to its natal river:

1. South of their natal river, migrating salmon move in a relatively fast, directed fashion (π).
2. Within the natal river box, residency is most probable a priori, with movements of one box allowed to the south or north (π_{SP}).
3. North of the natal river box, the (prior) net movement direction is southwards, although movement farther northward can occur. Movements of up to two boxes to the south or north are allowed (π_N).

These movement rules reflect observations of relatively direct and fast movements once salmon begin migrating to their natal river to spawn. As the fish approach their home river, their speed decreases and they seek the river mouth (Westerberg, 1982), remaining there for some time and undergoing adaptation to fresh water before entering the river (Karlsson & Karlström, 1994). Limited movement beyond the natal river is also supported by results from earlier tagging studies (e.g. Carlin, 1969; Siira et al., 2009). We assume that the speed of movement in the primary direction (northwards) is the same for all stocks. The expected number of fish moving from primary area i to primary area j at time $t = 1$ is given by:

$$N_{mov_{y,i,j,k,s,1}} = I_{i \notin SP1,N} \, N_{start_{y,i,k,s,1}} \pi_{i,j-i+1,s} + I_{i \in SP1} N_{start_{y,i,k,s,1}} \pi_{SP,j-i+2}$$
$$+ I_{i \in N} \, N_{start_{y,i,k,s,1}} \pi_{N,j-i+3}$$

where $\pi_{i,l,s}$ is the probability of moving $l-1$ boxes in the primary direction (north) from box i, for stock s: $\pi_{i,1,s}$ thus corresponds to residency, while $\pi_{i,16,s}$ corresponds to a maximum movement of 15 boxes north. $\pi_{SP,1:3}$ is a vector of movement probabilities originating from the spawning area (denoted by the subscript SP), where $\pi_{SP,2}$ corresponds to residency in the spawning box. $\pi_{N,1:5}$ is a vector of movement probabilities originating from boxes north of the spawning area (denoted

by the subscript N), where $\pi_{N,3}$ denotes residency. I terms are indicator variables; for example, $I_{i \in SP1}$ takes the value 1 if i is equal to the (primary) spawning area, and 0 otherwise. Movement for an example stock (Luleälven) is illustrated in Figure 2. See Supporting Information for details of movement parameter priors and Table S1 for stock-specific spawning areas.

To simplify the model structure, we allowed secondary (east–west and west–east) movements only within certain boxes that have been identified from tagging studies as crossing points for salmon migrating along the coasts of Sweden and Finland, such as the Åland Sea and the Northern Quark (Siira et al., 2009) (Figure 2). Instead of allowing secondary movements to be wholly stock-specific, we assume three different modes $G(s)$ for secondary movement, whereby each mode is a distinct pattern of movement with its own estimated parameters, and stock membership to a movement mode is stochastic. Mode 1 is set to no secondary (longitudinal) movement. The expected number of salmon in primary area i in secondary area 1 after movement is given by:

$$N_{\text{new}_{y,i,1,s,1}} = \sum_{j=i-b_i}^{i} N_{\text{mov}_{y,j,i,1,s,1}} \left(1 - \rho_{1,\text{Bl}(i,s),G(s)}\right) + \sum_{j=i-b_i}^{i} N_{\text{mov}_{y,j,i,2,s,1}} \rho_{2,\text{Bl}(i,s),G(s)}$$

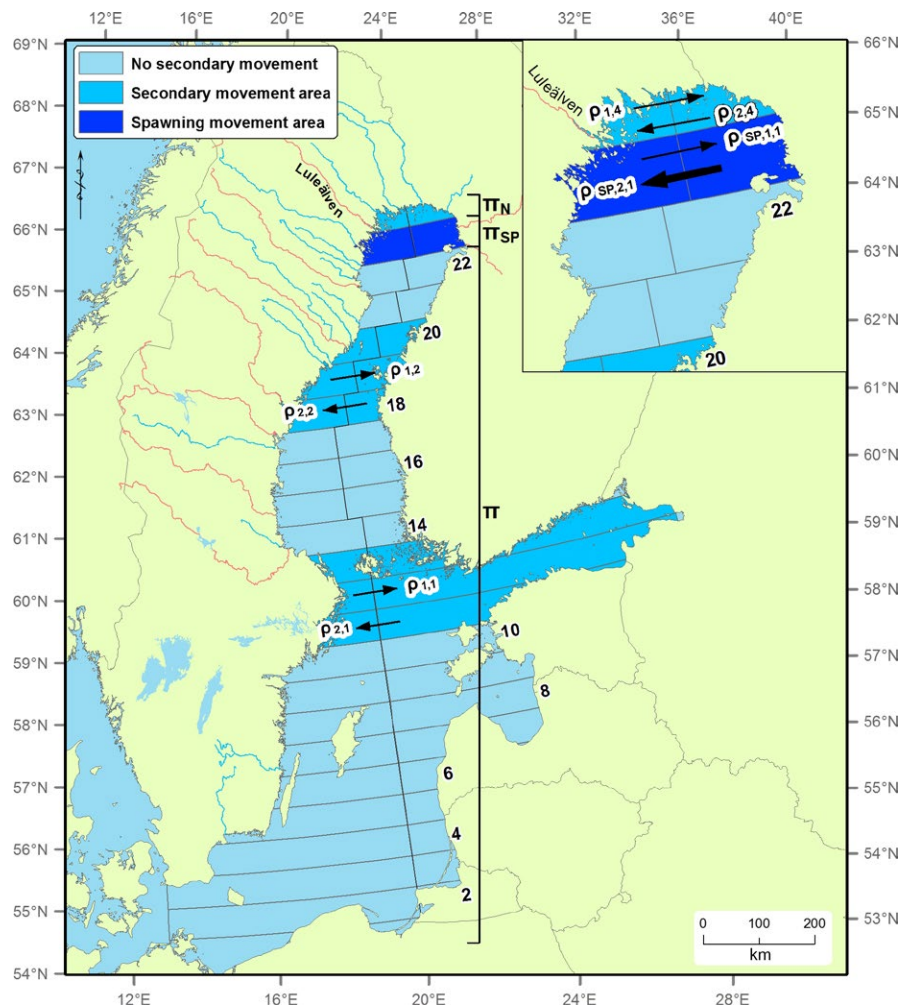The expected number of salmon in primary area i in secondary area 2 after movement is given by:

$$N_{\text{new}_{y,i,2,s,1}} = \sum_{j=i-b_i}^{i} N_{\text{mov}_{y,j,i,2,s,1}} \left(1 - \rho_{2,\text{Bl}(i,s),G(s)}\right) + \sum_{j=i-b_i}^{i} N_{\text{mov}_{y,j,i,1,s,1}} \rho_{1,\text{Bl}(i,s),G(s)}$$

where $\rho_{1,\text{Bl}(i,s),G(s)}$ is the probability of moving from secondary area 1 to secondary area 2, for area block Bl and secondary movement mode G. $1 - \rho_{1,\text{Bl}(i,s),G(s)}$ is the probability of remaining in secondary area 1. b_i is the lower bound for the box from which movement to area i can occur, defined as $b_i = \max(1, i - 16 + 1)$.

In addition to fixed secondary movement areas for all stocks (Figure 2), salmon are assumed to cross to the side of their natal river in the secondary movement area closest to the spawning box with high prior probability, ρ_{SP} (Figure 2). For these areas, the expected number of salmon from stocks with a spawning box in secondary area 1 (Swedish coast) in primary area i and secondary area 1 after movement is given by:

$$N_{\text{new}_{y,i,1,s,1}} = \sum_{j=i-b_i}^{i} N_{\text{mov}_{y,j,i,1,s,1}} \left(1 - \rho_{SP,1,1}\right) + \sum_{j=i-b_i}^{i} N_{\text{mov}_{y,j,i,2,s,1}} \rho_{SP,2,1}$$

FIGURE 2 Illustration of movement areas and parameters for an example stock (Lulealven). Longitudinal movement can take place in latitudinal boxes 11–13, 18–20, 23 and 24 (ρ parameters). Latitudinal movement parameters (π) applying to different boxes are shown alongside the square bracket. In this example, box 23 is the spawning box, within which spawning movement rules apply (inset)

The expected number of salmon from stocks with a spawning box in secondary area 1 (Swedish coast) in primary area i and secondary area 2 after movement is given by:

$$N_{\text{new}_{y,i,2,s,1}} = \sum_{j=i-b_i}^{i} N_{\text{mov}_{y,j,i,2,s,1}} (1-\rho_{\text{SP},2,1}) + \sum_{j=i-b_i}^{i} N_{\text{mov}_{y,j,i,1,s,1}} \rho_{\text{SP},1,1},$$

where $\rho_{\text{SP},k,\text{SP2}}$ is the probability of moving out of area k for stocks with secondary spawning area SP2. Spawning movements are defined analogously for stocks with spawning boxes along the Finnish coast (secondary area 2).

2.5.2 | Population dynamics $t \geq 2$

Equations for the second and later time steps are the same as equations presented for $t = 1$, unless defined below. The total number of fish (non-migrating and migrating) in primary area i and secondary area k at time t is given by:

$$N_{y,i,k,s,t} = N_{y,i,k,s,1} e^{-\sum_1^{t-1} Z_{g(s)}} \left(1 - \sum_1^{t-1} \delta_{s,t}\right) + N_{\text{new}_{y,i,k,s,t-1}}$$

The number of migrating salmon surviving after total mortality in primary area i and secondary area k in time step $t \geq 2$ in year y is given by:

$$N_{\text{surv}_{y,i,k,s,t}} = (N_{\text{start}_{y,i,k,s,t}} e^{-\sum_2^{t-1} Z_{g(s)}} + N_{\text{new}_{y,i,k,s,t-1}}) e^{-Z_{g(s)}}$$

The expected number of fish moving from primary area i to primary area j in time step t is given by:

$$N_{\text{mov}_{y,i,j,k,s,1}} = I_{i \notin S1,N} N_{\text{surv}_{y,i,k,s,t}} \pi_{i,j-i+1,s} + I_{i \in S1} N_{\text{surv}_{y,i,k,s,t}} \pi_{S,i,j-i+3} (1 - hI_{k \in S2})$$
$$+ I_{i \in N} N_{\text{surv}_{y,i,k,s,t}} \pi_{N,i,j-i+2}$$

where h is the rate at which salmon move from the spawning box (see Table 1) into the river in one time step. I terms are indicator variables; for example, $I_{i \in SP1}$ takes the value 1 if i is equal to the (primary) spawning area, and 0 otherwise.

2.5.3 | Mixed stock sample

The unknown quantities in the mixed stock sample are the proportions $q_{y,i,k,s,t}$ of each stock s in the sample taken in primary area i and secondary area k in time step t of year y. The migration model provides a prior for the expected proportion of stock s in the sample:

$$q_{y,i,k,s,t} = \frac{N_{y,i,k,s,t}}{\sum_{u=1}^{S} N_{y,i,k,u,t}}$$

In the run where we omit the population dynamics model and make the standard assumption of equal prior probabilities for the proportion of each stock in the mixture, we use a Dirichlet prior probability distribution to ensure that the proportions $q_{y,i,k,s,t}$ sum to unity:

$$(q_{y,i,k,1,t}, \ldots, q_{y,i,k,S,t}) \sim \text{Dirichlet}((\beta_1, \ldots, \beta_s)),$$

where the parameters β_s are given the value 1/S.

2.6 | Observation models

2.6.1 | Microsatellite alleles

We assume that Hardy–Weinberg genotype relative frequencies (RFs) hold in all baseline populations, specifying the model in terms of the allele RFs from which genotype RFs are derived. This assumption requires the following conditions to be met: (1) mating within each stock happens at random, (2) the size of each stock is infinite, (3) there is no linkage between loci and (4) there is no temporal variation in baseline allele frequencies. Assumption (2) is clearly not met in reality; however, combined with (1), it means that allele frequencies in a given stock are not subject to random variation and that the two alleles at a given locus in an individual are inherited independently from each other. In practice, we assume that these same effects are achieved if it can be assumed that all subpopulations (baseline stocks) are large enough not to undergo noticeable genetic change over the time frame during which mixture samples are collected.

The number of genetic loci used in the analysis is denoted by L, and a specific locus is indexed by $l = 1, \ldots, L$. The number of alternative alleles at locus l is denoted by K_l, and a particular allele for that locus is identified by a and can take integer values from 1 to K_l. The RF of allele a at locus l in stock s is denoted by $p_{l,s,a}$. We use a Dirichlet distribution to describe our prior knowledge about allele RFs, that is the proportions of each allele at a locus (as allele proportions at a locus must sum to one):

$$(p_{l,s,1}, \ldots, p_{l,s,K_l}) \sim \text{Dirichlet}((\alpha_{p,l,s,1}, \ldots, \alpha_{p,l,s,K_l}))$$

If there is no prior knowledge about the allele proportions, a standard strategy is to set $\alpha_{p,l,s,a} = 1/K_l$, $a = 1, \ldots K_l$. Given knowledge about allele proportions, the baseline sample $(x_{l,s,1}, \ldots, x_{l,s,K_l})$ for locus l in stock s can be modelled as a sample from a multinomial distribution:

$$(x_{l,s,1}, \ldots, x_{l,s,K_l}) \mid (p_{l,s,1}, \ldots, p_{l,s,K_l}) \sim \text{Multinomial}((2W_s, p_{l,s,1}, \ldots, p_{l,s,K_l}))$$

where $x_{l,s,a}$ is the count of allele a at locus l for stock s and W_s is the number of individuals from stock s in the baseline sample. We set the number of alleles K_l for each locus equal to the observed number of different alleles for that locus in the baseline data. Because of the theory of conjugate distributions, the conditional distribution of the allele proportions, given the baseline sample, also follows a Dirichlet distribution, with parameters equal to the sums of the prior parameters (α) and the corresponding observations (x); thus:

$$(p_{l,s,1}, \ldots, p_{l,s,K_l}) \mid (\alpha_{p,l,s,1}, \ldots, \alpha_{p,l,s,K_l}), (x_{l,s,1}, \ldots, x_{l,s,K_l})$$
$$\sim \text{Dirichlet}((\alpha_{p,l,s,1} + x_{l,s,1}, \ldots, \alpha_{p,l,s,K_l} + x_{l,s,K_l}))$$

This Dirichlet posterior distribution includes all the information about the allele proportions at each locus in each stock, before observing samples from an unknown mixture of the stocks. We use this posterior distribution as the prior for the rest of the model.

Dropping subscripts other than s for clarity of presentation, suppose that the stock proportions q_s were known. If the mixed population

was assumed to be infinitely large, or sampling was performed with replacement, then each individual in the sample has the probability q_s to be a member of stock s. Denoting the origin of individual $i = 1,..., I$ as $O_i \in \{1,...,S\}$, if individual 1 belongs to stock 2, then $O_1 = 2$, and $P(O_1 = 2|q_1,...,q_s) = q_2$. In other words, the origin of each individual in the sample follows a categorical distribution, conditional on stock proportions:

$$O_i \mid (q_1, \ldots, q_S) \sim \text{Categorical}\left((q_1, \ldots, q_s)\right)$$

If the origin of individual i is known ($O_i = s$), then it is possible to assess the probability to find a particular allele a from the locus l of that individual. This probability is simply the baseline allele proportion $p_{l,s,a}$ in stock s. Each individual has two alleles at locus l, which have equal probabilities: the allele of the first chromatid of the chromosome for individual i at locus l is denoted by $A_{i,l,1}$ and the allele of the second chromatid at the same locus by $A_{i,l,2}$. If the individual belongs to stock $s = 2$ and the allele proportions of that stock are known, then $P\left(A_{i,1,1} = 2 \mid (p_{1,2,1}, \ldots, p_{1,2,K_l})\right) = p_{1,2,2}$. Thus, the distribution of alleles can also be expressed using two conditionally independent categorical distributions:

$$A_{i,l,1} \mid O_i, \left(p_{l,s,1}, \ldots, p_{l,s,K_l}\right) \sim \text{Categorical}\left((p_{l,O_i,1}, \ldots, p_{l,O_i,K_l})\right)$$

$$A_{i,l,2} \mid O_i, \left(p_{l,s,1}, \ldots, p_{l,s,K_l}\right) \sim \text{Categorical}\left((p_{l,O_i,1}, \ldots, p_{l,O_i,K_l})\right)$$

The Hardy–Weinberg assumption allows the alleles for the two chromatids to be treated as conditionally independent, meaning that the joint probability of the two alleles at a single locus is the product of their baseline population allele RFs. Ignoring which chromatid each allele came from, genotype probabilities are obtained by summing up the appropriate combinations from this joint distribution.

2.6.2 | Fin-clipping

We used a categorical observation model for fin-clipping observations (fin, wild; no fin, reared), assuming that for the majority of rivers, on average 1% of salmon from wild rivers will be misreported as fin-clipped, and that 5% of salmon from reared rivers in Sweden will have an intact adipose fin. This 5% arises from a combination of smolts that are missed during fin-clipping, accidental releases of fry or parr, and successful reproduction by reared parents. For reared Finnish stocks and reared Swedish stocks with specific information on proportions of fin-clipped smolts by year (Torneälven Hatchery, Iijoki, Oulujoki, Ångermanälven), we used the scalar product of proportions-at-age at sea and the annual proportions of fin-clipped smolts between 4 and 1 years earlier (fin-clipping information for Finnish stocks was provided by Tapani Pakarinen, Luke). See Table S1 for stock-specific prior fin-clipping proportions.

The observation model for fin-clipping follows the same principles as the observation model for allelic data. Let the state of a salmon's adipose fin be denoted by F, where $F = 1$ represents an intact adipose fin and $F = 2$ represents removal of the adipose fin. Then:

$$F_i \mid O_i \sim \text{Categorical}\left((u_{O_i,1}, u_{O_i,2})\right),$$

where $u_{s,1}$ and $u_{s,2}$ denote the proportions of fish from stock s with intact and clipped adipose fins, respectively.

2.7 | Simulation study: effect of fin-clipping data

To obtain an indication of the potential gains in assignment accuracy when using fin-clipping observations for Baltic Sea salmon stocks, we estimated stock of origin using the microsatellite observation model described above with 10 sets of 200 individuals (of known stock of origin). Mixture genotypes were sampled without replacement from the baseline and subsequently removed from the baseline for estimation. For simplicity, we omitted the population dynamics prior, using a model with equal prior probabilities for stock proportions, and performing estimation with and without the fin-clipping observation model for each sample. We assumed that fin-clipping information was available for 89% of individuals, the average among mixture samples in the Gulf of Bothnia case study (Table 2). Underlying prior proportions of fin-clipped individuals from each stock were set equal to values used in analyses with real data (Table 1). Assignment accuracy for each individual was defined as the proportion of posterior samples in which the individual was assigned to the correct stock.

Models were implemented in JAGS version 4.00 (Just Another Gibbs Sampler; Plummer, 2015) using the rjags interface (Plummer, 2016) to R version 3.2.3 (R Core Team, 2015). A burn-in period of 320,000 iterations was used, after which 480,000 more iterations were kept and thinned at an interval of 250 to yield a final sample of 1,920 iterations. Four chains were run in parallel for all models. Convergence to the posterior distribution was assessed using visual inspection of trace plots and using the Gelman–Rubin diagnostic (Gelman & Rubin, 1992).

3 | RESULTS

5.54% of Gelman–Rubin diagnostics were greater than 1.01, while only 0.75% were greater than 1.05 (continuous model parameters and variables) indicating convergence of MCMC chains to the posterior distribution. Sample trace and Gelman–Rubin plots (Figures S1 and S2) and posterior predictive checks are provided in the Supporting Information.

Spatio-temporally structured MSAs for Baltic salmon revealed strong variation in estimated stock composition between areas and over time within a single model area (Figures 3 and 4). The temporal variation within a given area meant that the mixture could be dominated by different stocks at different times during the 3.5-month period we studied (Figures 3 and 4).

Moving along a gradient from a less to a more informative situation, differences in inferences about stock composition in time and space became apparent. Under the uninformative prior scenario (i.e. genetic data only), stock composition parameters are only updated in boxes and time steps with genetic marker observations (e.g. compare Figure 3b, with genetic observations, to 3a
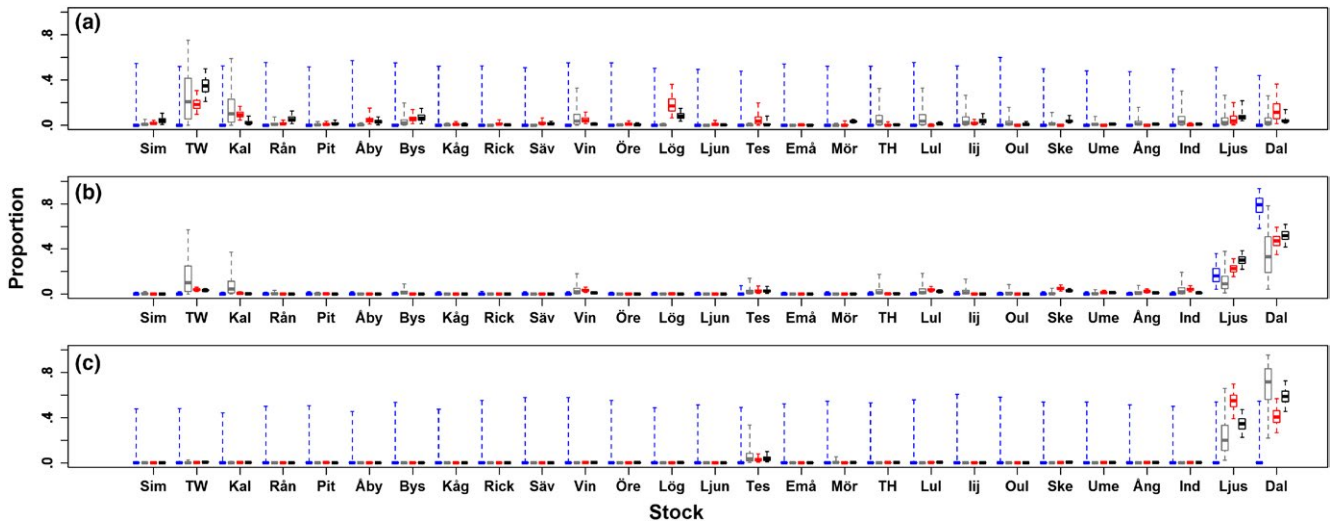
FIGURE 3 Predicted stock composition in primary area 14 (west), spawning box for the Testeboån ("Tes," wild) and Dalälven ("Dal," reared) stocks. Blue boxes, equal stock prior probabilities plus genetic marker data; grey boxes, population dynamics prior; red boxes, posterior distribution with only genetic marker data; black boxes, posterior distribution with both genetic marker and fin-clipping data. (a) 13–26 May 2014, N = 0; (b) 8–21 July 2014, N = 22; and (c) 5–18 August 2014, N = 0. Stocks to the right of Mörrumsån ("Mör") are reared. Stocks are listed in order of river latitude (from left/north to right/south) within wild and reared groups
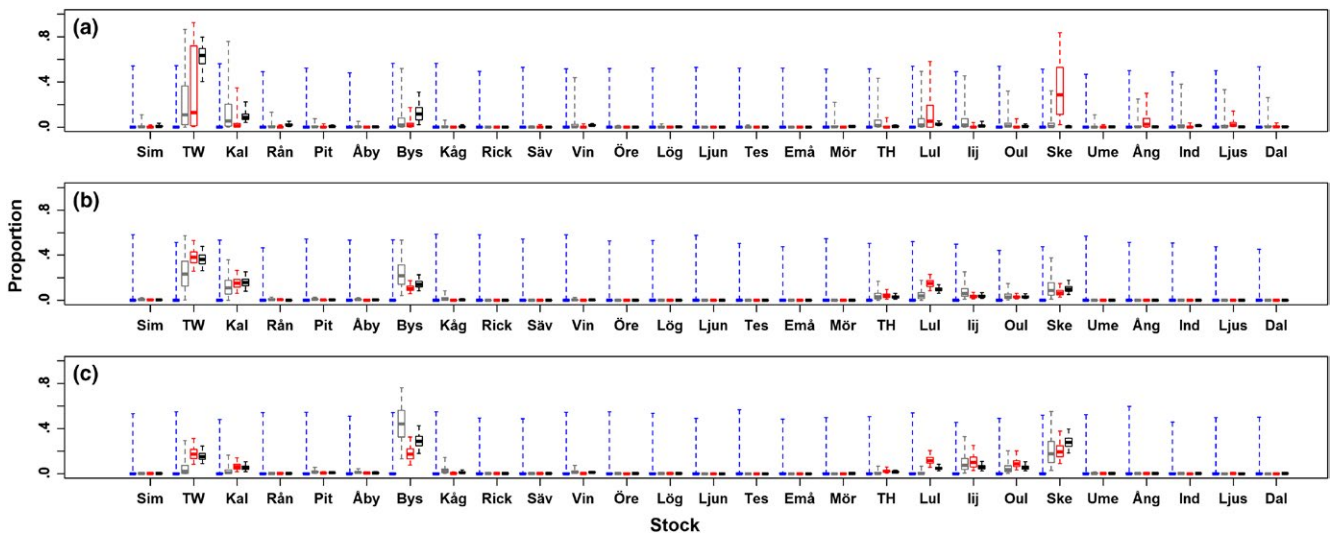
FIGURE 4 Predicted stock composition in primary area 22 (east). Blue boxes, equal stock prior probabilities plus genetic marker data; grey boxes, population dynamics prior; red boxes, posterior distribution with only genetic marker data; black boxes, posterior distribution with both genetic marker and fin-clipping data. (a) 13–26 May 2014, N = 0; (b) 8–21 July 2014, N = 0; and (c) 5–18 August 2014, N = 0. Stocks to the right of Mörrumsån ("Mör") are reared. Stocks are listed in order of river latitude (from left/north to right/south) within wild and reared groups

and 3c, without). Formalizing available information and knowledge about migration into a prior for population dynamics allowed prediction of stock compositions in all areas and time steps, but with greater uncertainty compared to the uninformative prior-/genetic data-only case (Figure 3). Adding genetic marker data (but no fin-clipping data) generally decreased the uncertainty associated with stock composition estimates (Figures 3 and 4). Finally, integrating the population dynamics prior with genetic and fin-clipping data led to appreciable gains in the precisions of stock composition estimates compared with either a prior-only or data-only scenario

(Figures 3 and 4) and small gains in precision compared with a prior-plus-genetic data-only scenario (Figures 3 and 4). It appeared that in some cases where there is a lack of agreement between the prior and genetic data, fin-clipping information may be important in resolving stock composition and increasing the precision of estimates (e.g. Figure 4a).

The improvement in inference gained by the inclusion of fin-clipping information is perhaps best illustrated at the level of individual assignments (Figure 5). Assignment of individuals to a particular stock was affected by both information about population

dynamics and information about stock type (wild vs. reared) from adipose fin observations (Figure 5). For example, using an uninformative prior, the most probable stock of origin for the third individual in Figure 5 is Byskeälven (wild), while embedding the genetic observation model within a prior for population dynamics results in Skellefteälven (reared) being the most probable stock of origin. In this case (third row of Figure 5), there appears to be some conflict between the genetic marker and fin-clipping information: the fin-clipping observation (intact adipose fin) suggests that the individual comes from a wild stock, whereas the most probable stock with genetic marker data only (Skellefteälven) is reared. The relative weight

of the fin-clipping data relative to the genetic marker data and population dynamics prior appeared to vary among individuals (cf. first and third rows, Figure 5).

Unless stated otherwise, the remaining results in this section are from model 3 (informative prior, genetic and fin-clipping data).

Estimated dates for the onset of migration from feeding grounds in the southern and central main basin of the Baltic Sea were updated to varying degrees by the microsatellite and fin-clipping data (Figure 6). On average, the timing of migration was generally later for reared stocks compared to wild stocks (Figure 6). Among wild stocks, there also appeared to be a north–south cline in migration timing (with
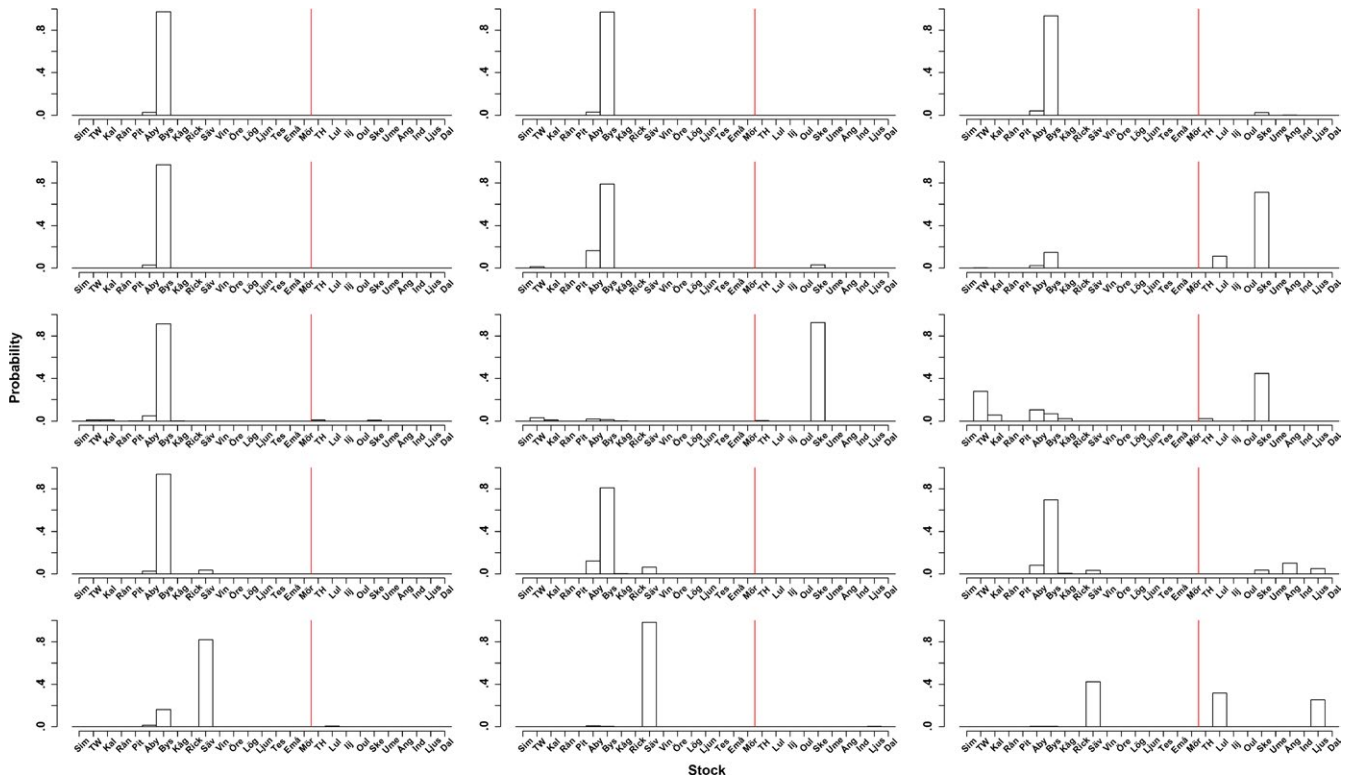

FIGURE 5 Example posterior distributions for individual assignments from the sample taken in box 21 (west) in 2014. Each row represents one individual. Left column, equal prior probabilities (no fin-clipping information); central column, population dynamics prior, no fin-clipping information; rightmost column, population dynamics prior with fin-clipping information. Actual fin-clipping observations: row 1, no adipose fin (hatchery); row 2, no adipose fin (hatchery); row 3, adipose fin (wild); row 4, no adipose fin (hatchery); row 5, no adipose fin (hatchery). Stocks to the left of the vertical red line are wild; those to the right are reared. Stocks are listed in order of river latitude (from left/north to right/south) within wild and reared groups
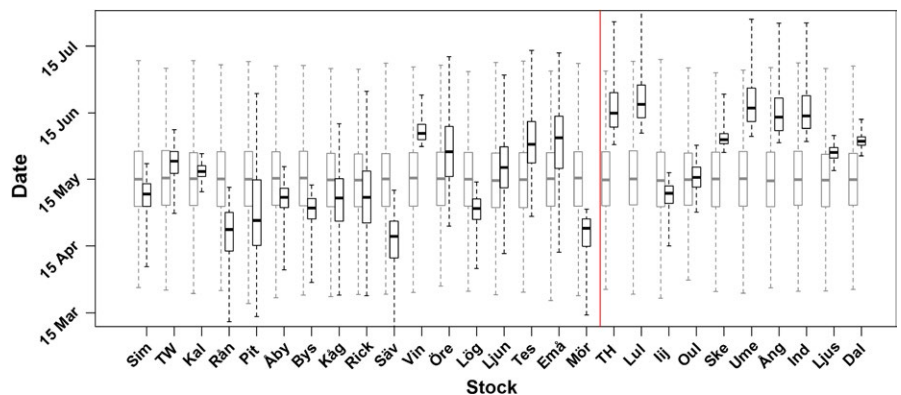

FIGURE 6 Timing of the onset of migration from feeding grounds in the main basin. Stocks to the left of the vertical red line are wild; those to the right are reared. Stocks are listed in order of river latitude (from left/north to right/south) within wild and reared groups

several exceptions, including the Mörrumsån stock from the southern Main Basin; Figure 1), so that stocks from rivers farther north tended to start migrating earlier (Figure 6). Differences in the onset of migration appeared to be countered by migration distance to some extent, so that arrival times at more southerly spawning areas were often earlier than those for more northern rivers (Figure 7).

Updating of the prior for latitudinal movement in each 2-week time step was limited, with the most updating occurring for the probability

of remaining in the same primary box (π_1), to a higher value than under the prior (Figure 8a). The genetics and fin-clipping data were informative about patterns of longitudinal movement, with two movement modes (Figure 8b and c cf. Figure 8d and e) emerging in addition to a movement mode consisting of 0 longitudinal movement (mode 1). Under movement mode 2 (Figure 8b and c), there was a pattern of net west-to-east movement around the Quark (longitudinal movement area 2; Figure 2), followed by net east-to-west movement in
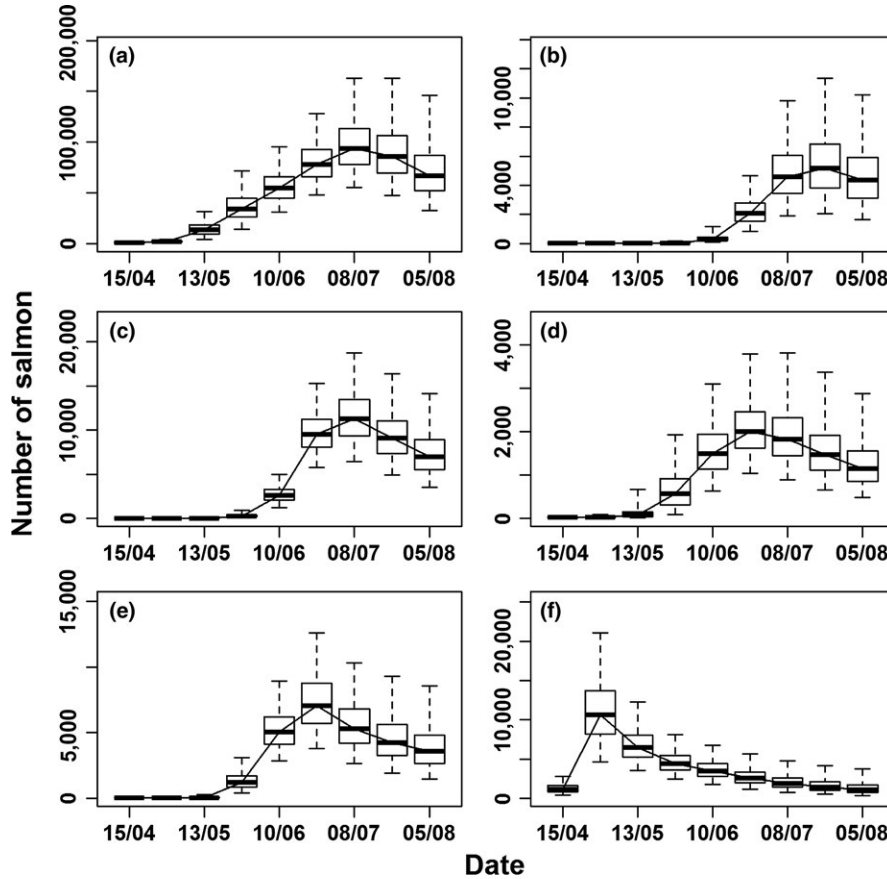

FIGURE 7 Estimated abundances of salmon from different stocks in their respective spawning boxes between 15 April and 5 August 2014. Net arrivals are indicated by the difference in numbers from one time step to the next. (a) Torneälven wild, (b) Torneälven hatchery, (c) Vindelälven, (d) Oulujoki, (e) Dalälven and (f) Mörrumsån
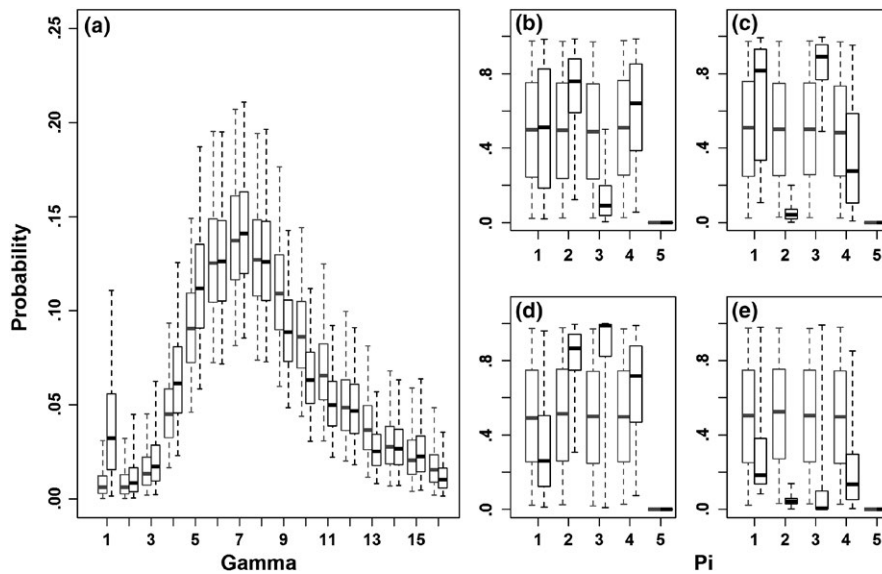

FIGURE 8 Prior (grey boxes) and posterior (black boxes) probability distributions for (a) latitudinal movement (number of boxes north in each time step) and (b–e) longitudinal movement: (b) west–east movement rate, movement mode 2; (c) east–west movement rate, movement mode 2; (d) west–east movement rate, movement mode 3; and (e) east–west movement rate, movement mode 3

the northern Bothnian Bay (longitudinal movement area 3; Figure 2). Movement mode 3 was associated with net west-to-east movement at both the Quark and the northern Bothnian Bay (Figure 8d and e). The estimated pattern of movement during the coastal migration is illustrated for the Luleälven reared stock (Figure 9). This stock is estimated to have migrated north primarily along the Finnish coast, before crossing back to the Swedish coast in the northern Bothnian Bay to reach the natal river (Figure 9).

The prior for the annual instantaneous rate of natural mortality in reared salmon was updated slightly by the genetic and fin-clipping data, although there was no update of the prior for the natural mortality rate in wild Baltic salmon (Figure 10). The posterior distribution for the rate of natural mortality in reared fish had a median of 0.17 and coefficient of variation (CV) of 0.28 (prior median 0.20 and CV 0.31). The prior for the proportion of salmon in the spawning box that enters the natal river in each time step was updated to yield a posterior distribution with median 0.49 and CV of 0.26 (Figure 10) (prior median 0.50 and CV 0.58). The genetic data indicated a lower average migration speed in the primary direction (northwards) than the prior based on earlier tagging studies (28 km/day compared with 35 km/day).
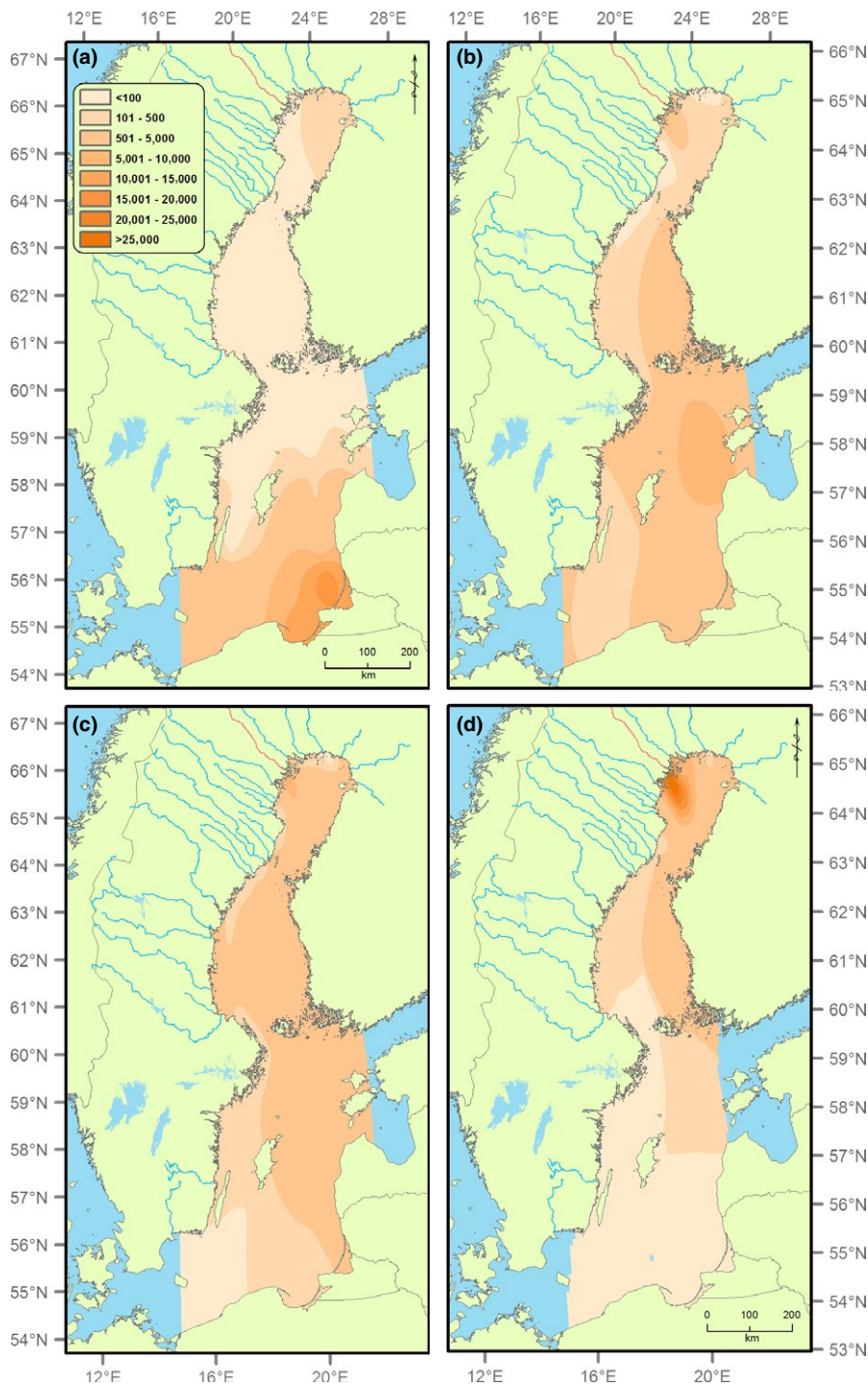

FIGURE 9 Posterior spatio-temporal distribution for the Luleälven stock (smoothed box-specific abundance estimates). (a) 27 May–9 June, (b) 10–23 June, (c) 24 June–7 July and (d) 22 July–4 August

The estimated stock composition in samples taken from the coastal fishery along the Swedish and Finnish coasts of the Baltic Sea showed strong spatial variation (Figure 11). Overall, the estimated stock composition results showed a pattern of predominance of stocks with the spatially closest spawning boxes (Figure 11). Stock composition at sampling sites along the Finnish coast of the Gulf of Bothnia was dominated by the wild Tornionjoki stock ("*TW*"), the most abundant Baltic salmon stock (ICES, 2015), while along the Swedish coast stock composition was more heterogeneous with greater spatial variation in stock composition (Figure 11).

In our study, the population dynamics prior had a noticeable effect on individual assignments, as well as on estimated mixture proportions (Figure 5, Figure 11 vs. Figure S5). Adding fin-clipping data had a less marked effect (Figure 5, Figure 11 vs. Figure S6), but still led to differences in estimated mixture proportions, particularly for genetically similar stocks. In some cases, adding fin-clipping data either did not change an individual's assignment or led to a more or less even distribution of the probability between wild and reared stocks.

3.1 | Effect of fin-clipping data

Analyses with simulated data indicated that an auxiliary marking such as fin-clipping can impact estimates of mixture population proportions (Figure 12) and improve the accuracy of individual assignments. Mean (over individuals) assignment accuracies without fin-clipping observations ranged between 59% and 69%, while mean assignment accuracies with fin-clipping observations ranged between 65% and 74% for the 10 simulated datasets evaluated. The overall average gain in assignment accuracy using fin-clipping data was 5% (assignment accuracy with fin-clipping data minus assignment accuracy without) for Gulf of Bothnia salmon stocks. This increased accuracy was most pronounced for genetically similar stocks, particularly in the case of reared and wild pairs originating from the same river or neighbouring rivers [e.g. the wild Vindelälven (18% gain in accuracy with fin-clipping data) and reared Umeälven (30% gain) stocks; the wild Piteälven (8% gain) and reared Luleälven

(18% gain) stocks and the wild ("*TW*," −1%) and hatchery ("*TH*," 17%) Tornionjoki River stocks].

4 | DISCUSSION

We have presented an integrated model for the joint inference of the spatial dynamics of multiple populations. This approach is expected to have utility for a number of applications in ecology (e.g. metapopulation dynamics and management of species with genetically distinct subpopulations), particularly where a hierarchical structure can be used to describe differences among members of the same group (in our case genetically distinct stocks of Baltic salmon).

Performing an integrated analysis where the raw genetic data are embedded into the population dynamics model (rather than using a two-step approach) avoids loss of information that can occur when raw data are summarized and allows uncertainty to be appropriately accounted for. Empirical studies have shown that integrated population models can yield more precise estimates of demographic parameters than separate models (Besbeas et al., 2003). In some cases, combining multiple datasets and diverse knowledge in an integrated approach is the only feasible solution for parameter estimation (Besbeas et al., 2002; Schaub et al., 2007)

The advantages of Bayesian statistical methods for mixed stock analysis are widely recognized, although the potential to incorporate prior information is often overlooked. We developed a mechanistic model of population dynamics as a prior for spatio-temporal stock compositions. Our study demonstrates the benefits of utilizing available prior knowledge in the context of genetic MSA, both in reducing uncertainty in stock composition estimates in areas and at times when observations have been made, and in making mechanistically based inferences about stock composition in areas and at times when data are lacking. This represents a step forward in the use of prior information in MSA problems with spatial and/or temporal structure from earlier studies that assumed equal prior mixture proportions or applied a sequential Bayesian approach to a time series of mixture samples (e.g. Bradbury et al., 2015; Dann et al., 2013). Our approach
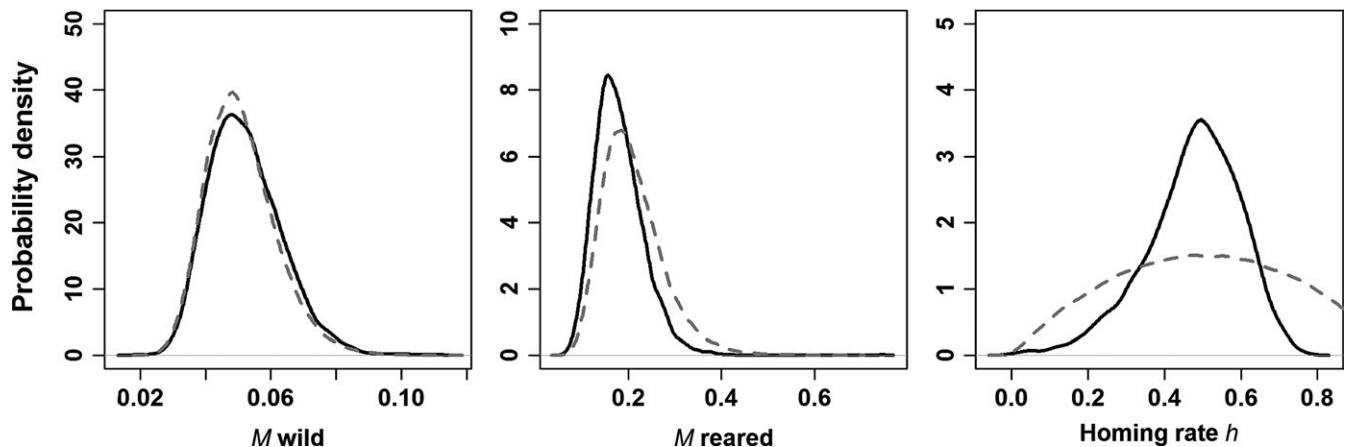

FIGURE 10 Prior (dashed grey lines) and posterior (solid black line) probability density functions for the annual instantaneous rate of natural mortality (M) in wild and reared salmon, and the proportion of salmon in the spawning box that enters the natal river in each time step (fortnight)

also shows that spatio-temporal difference in stock composition can be used to learn about underlying demographic processes, particularly population-specific movements.

Populations that undergo simultaneous harvest present a unique set of challenges for management and conservation. Together, developments in genetic marker technology and MSA methods are facilitating the design and implementation of population-specific management measures (Bradbury et al., 2016; Dann et al., 2013). Our analysis demonstrates complex stock-specific patterns of migration along the Swedish and Finnish coasts that create strong spatio-temporal variation in stock composition in Baltic salmon catches. The pattern of temporal variation in stock composition within a given spatial area (Figures 2 and 3) can be explained by variation among stocks in the onset of migration (Figure 6), together with the characteristics of migration (salmon slow down becoming more spatially aggregated closer to the natal river and

are eventually lost from coastal areas as they enter the river). The results from our study provide valuable information for management of Baltic salmon stocks that vary markedly in their status, and for setting potential catch quotas under an MSY-based management approach. These issues are priorities for future work.

Overall, our results are consistent with the findings of earlier studies on the migration of salmon in the Baltic Sea. Later arrival of reared fish compared with wild ones has been noted by several authors, including Siira et al. (2009) and Karlsson and Karlström (1994), who reported an approximately 10-day later arrival date for reared salmon stocks in the northern Gulf of Bothnia, compared with wild ones. In our study, the posterior median of the hyper-prior mean migration start date for reared stocks was ~15 days later than that for wild stocks. Siira et al. (2009) estimated peak arrival for salmon homing to the Oulujoki River in the Bothnian Bay between early June and early July in 2001 and 2002.
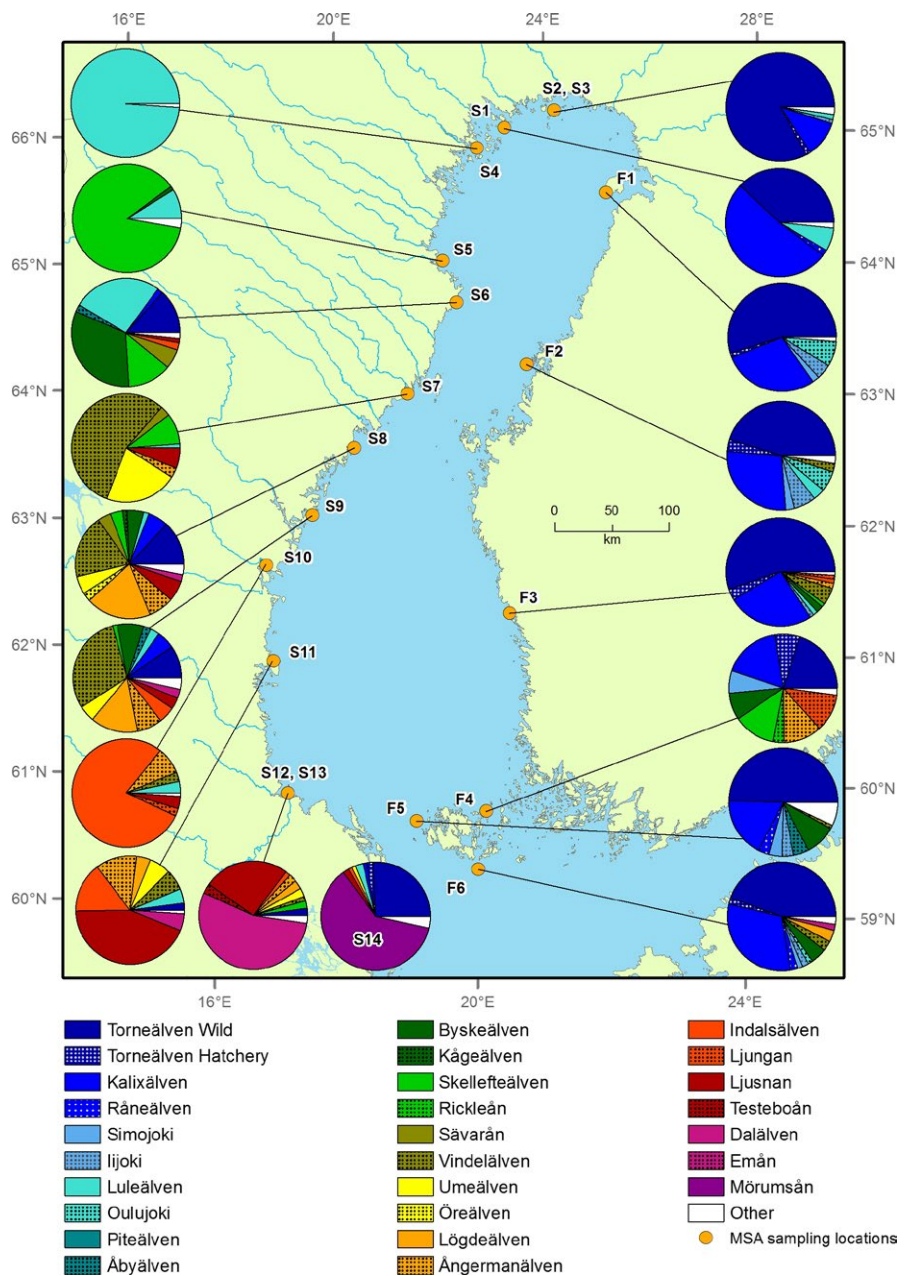

FIGURE 11 Estimated posterior median stock compositions in samples taken from coastal trap nets in 2014, population dynamics prior plus genetic and fin-clipping data
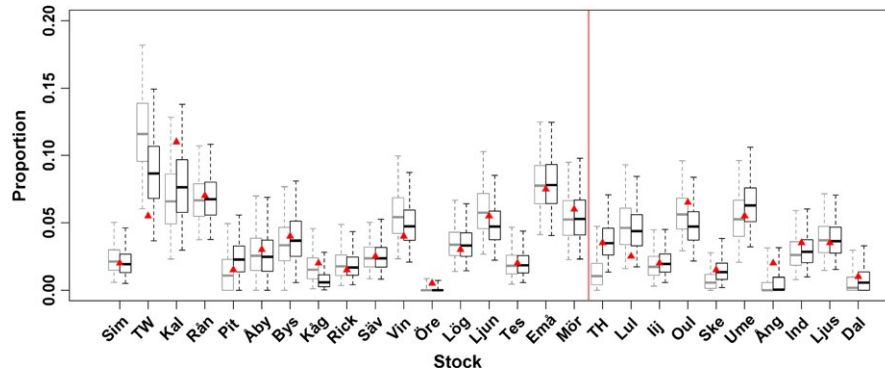
FIGURE 12 Estimated stock proportions for one simulated dataset (of 10), consisting of 200 individuals sampled without replacement from the baseline. Grey boxes, no fin-clipping data; black boxes, fin-clipping data. True proportions in the sample are indicated by red triangles. Stocks to the left of the vertical red line are wild; those to the right are reared. Stocks are listed in order of river latitude (from left/north to right/south) within wild and reared groups

Our results suggest a slightly earlier arrival time to the river mouth in 2014, between late May and mid-June (Figure 6d). However, some annual fluctuation in arrival timing for adult Baltic salmon at river mouths is expected, as arrival time has been found to correlate with temperature (colder winters/springs tend to result in later arrival and vice versa; Karlsson, Karlström, & Hasselborg, 1995). Estimated arrival times at the river mouth (Figure 7) can also be compared with counts of ascending spawners in some rivers. For example, in the Torne River, peak counts of ascending spawners at Kattilakoski were made between 24 June and 7 July in 2014, with the second largest count in the previous fortnight. Peak arrivals at the river mouth from this study were estimated to have occurred in the fortnights 10–23 June, followed by 27 May–9 June (Figure 7a). This fits well with a travel time of 1 to 2 weeks between the river mouth and Kattilakoski, based on in-season development of daily catches at different locations within the river (Atso Romakkaniemi, Luke, pers. comm.). In Vindelälven, fish ladder counts show peak arrivals to the river in 2014 during the fortnights 8–21 July and 22 July–4 August (Kjell Leonardsson, SLU, unpubl. data), with an estimated travel time of 40–44 days (~3 fortnights) from the river mouth to the ladder (Lundqvist, Rivinoja, Leonardsson, & McKinnell, 2008; McKinnell, Lundqvist, & Johansson, 1994). This is consistent with our estimated peak arrivals at the river mouth during the fortnight 10–24 June in 2014 (Figure 7c).

The microsatellite data were informative about population dynamics parameters governing migration timing and patterns of movement in the Baltic Sea, but less so about mortality parameters. This likely resulted from a combination of informative priors and a weak signal about absolute rates of mortality in the genetic data—the data may however be informative about relative rates of mortality for different populations or population types (e.g. wild vs. reared salmon stocks), according to the rates at which numbers of fish from different groups decay over time. We used posterior distributions for annual mortality rates from ICES' assessment (ICES, 2015) as priors for total mortality. This is expected to result in conservative estimates of total mortality, owing to the fact that catch data are not yet included. The prior for the initial (pre-migration) spatial distribution is based on tag recapture data from the 1960s (Carlin, 1969), which were down-weighted owing to the long period of elapsed

time and possibility of an altered initial spatial distribution in recent years. While this prior is not expected to have a large effect on the overall results (because of the modest prior sample size (Table S3), coarse spatial resolution of the model and flexible description of movement), substantial changes in the spatial distribution in early May could be expected to lead to slight biases in estimates of model parameters such as migration start times.

Addition of an observation model for catch data from the coastal trap net fishery, allowing estimation of stock-specific harvest rates and escapement for the coastal fishery, is planned as the next step in this work. Extending the analysis to multiple years within a hierarchical framework would facilitate predictions about stock-specific movements and catch composition in future years. We did not account for process uncertainty in our model—in reality, both rates of survival and movement can be expected to show random variability. Explicitly modelling process and observation error can reduce bias in parameter estimates, confidence intervals and hypothesis tests (Maunder, Deriso, & Hanson, 2015; de Valpine & Hastings, 2002). Extending the model to a state–space framework is a further avenue for model development.

In summary, we have developed a model for jointly inferring the movement dynamics of multiple populations, demonstrating the method using data from fisheries on Gulf of Bothnia salmon stocks. This tool can potentially be used to evaluate spatio-temporal management actions for mixed stock fisheries. In the case of Baltic salmon it will enable allocation of fishing effort to target reared and healthy wild stocks while avoiding weak ones. Overall, genetic marker data appear to have strong potential to inform population-specific management of migratory species, with enhanced utility when integrated with knowledge about population status and movement dynamics.

provided data on ascending spawner counts in the Torne and Vindelälven Rivers in 2014. Tore Prestegaard and Emma Lind performed laboratory work and pre-processing of genetic data. Anders Asp created all map figures in the manuscript. R.W., S.P., J.D. and J.Ö. received financial support from the Swedish Agency for Marine and Water Management. J.Ö. received funding from the Swedish Research Council Formas (Dnr. 2013-1288).

AUTHORS' CONTRIBUTIONS

J.D., J.Ö., S.M., S.P. and R.W. conceived and developed ideas; J.Ö. organized collection and analysis of scales; M.-L.K. maintained the Baltic Atlantic salmon microsatellite DNA baseline; S.M. developed the prototype Bayesian MSA model for microsatellite DNA; R.W. developed the population dynamics model; J.D., J.Ö. and S.P. contributed to model parameterization; R.W. performed MSAs for Baltic salmon; and R.W. led writing of the manuscript with text contributions from M.-L.K., S.M. and S.P. All authors contributed critically to the drafts and gave final approval for publication.

DATA ACCESSIBILITY

Data have been deposited in the Dryad Digital Repository https://doi.org/10.5061/dryad.4pg37 (Östergren et al., 2017).

ORCID

Rebecca Whitlock 🄳 http://orcid.org/0000-0003-0058-634X

REFERENCES

Abadi, F., Gimenez, O., Arlettaz, R., & Schaub, M. (2010). An assessment of integrated population models: Bias, accuracy, and violation of the assumption of independence. *Ecology*, *91*, 7–14. https://doi.org/10.1890/08-2235.1

Besbeas, P., Freeman, S. N., Morgan, B. J., & Catchpole, E. A. (2002). Integrating mark–recapture–recovery and census data to estimate animal abundance and demographic parameters. *Biometrics*, *58*, 540–547. https://doi.org/10.1111/j.0006-341X.2002.00540.x

Besbeas, P., Lebreton, J. D., & Morgan, B. J. (2003). The efficient integration of abundance and demographic data. *Journal of the Royal Statistical Society: Series C (Applied Statistics)*, *52*, 95–102. https://doi.org/10.1111/1467-9876.00391

Besbeas, P., & Morgan, B. J. (2014). Goodness-of-fit of integrated population models using calibrated simulation. *Methods in Ecology and Evolution*, *5*, 1373–1382.

Bolker, B. M., Okuyama, T., Bjorndal, K. A., & Bolten, A. B. (2007). Incorporating multiple mixed stocks in mixed stock analysis: 'many-to-many' analyses. *Molecular Ecology*, *16*, 685–695. https://doi.org/10.1111/j.1365-294X.2006.03161.x

Bradbury, I. R., Hamilton, L. C., Hamilton, Lorraine. C., Sheehan, T. F., Chaput, G., Robertson, M. J., … Bernatchez, L. (2016). Genetic mixed-stock analysis disentangles spatial and temporal variation in composition of the West Greenland Atlantic Salmon fishery. *ICES Journal of Marine Science*, *73*, 2311–2321. https://doi.org/10.1093/icesjms/fsw072

Bradbury, I. R., Hamilton, L. C., Rafferty, S., Meerburg, D., Poole, R., Dempson, J. B., … Bernatchez, L. (2015). Genetic evidence of local exploitation of Atlantic salmon in a coastal subsistence fishery in the Northwest Atlantic. *Canadian Journal of Fisheries and Aquatic Sciences*, *72*, 83–95. https://doi.org/10.1139/cjfas-2014-0058

Branch, T. A., Lobo, A. S., & Purcell, S. W. (2013). Opportunistic exploitation: An overlooked pathway to extinction. *Trends in Ecology and Evolution*, *28*, 409–413. https://doi.org/10.1016/j.tree.2013.03.003

Brodziak, J. (1993). An extension of stock composition analysis to include marking data. *Canadian Journal of Fisheries and Aquatic Sciences*, *50*, 251–257. https://doi.org/10.1139/f93-029

Brooks, S. P., King, R., & Morgan, B. J. T. (2004). A Bayesian approach to combining animal abundance and demographic data. *Animal Biodiversity and Conservation*, *27*, 515–529.

Buckland, S. T., Newman, K. B., Thomas, L., & Koesters, N. B. (2004). State-space models for the dynamics of wild animal populations. *Ecological Modelling*, *171*, 157–175. https://doi.org/10.1016/j.ecolmodel.2003.08.002

Cadrin, S. X., Friedland, K. D., & Waldman, J. R. (2005). *Stock identification methods: Applications in fishery science*. Amsterdam, the Netherlands: Elsevier Academic Press.

Carlin, B. (1969). Salmon tagging experiments. *Swedish Salmon Research Institute Report*, *3*, 8–13.

Chandler, R. B., & Clark, J. D. (2014). Spatially explicit integrated population models. *Methods in Ecology and Evolution*, *5*, 1351–1360.

Chandler, R. B., Muths, E., Sigafus, B. H., Schwalbe, C. R., Jarchow, C. J., & Hossack, B. R. (2015). Spatial occupancy models for predicting metapopulation dynamics and viability following reintroduction. *Journal of Applied Ecology*, *52*, 1325–1333. https://doi.org/10.1111/1365-2664.12481

Chasco, B., Hilborn, R., & Punt, A. E. (2007). Run reconstruction of mixed-stock salmon fisheries using age-composition data. *Canadian Journal of Fisheries and Aquatic Sciences*, *64*, 1479–1490. https://doi.org/10.1139/f07-105

Clayton, L., Keeling, M., & Milner-Gulland, E.-J. (1997). Bringing home the bacon: A spatial model of wild pig hunting in Sulawesi, Indonesia. *Ecological Applications*, *7*, 642–652. https://doi.org/10.1890/1051-0761(1997)007[0642:BHTBAS]2.0.CO;2

Corander, J., Marttinen, P., & Mäntyniemi, S. (2006). A Bayesian method for identification of stock mixtures from molecular marker data. *Fishery Bulletin*, *104*, 550–558.

Corander, J., Waldmann, P., & Sillanpää, M. J. (2003). Bayesian analysis of genetic differentiation between populations. *Genetics*, *163*, 367–374.

Corander, J., Waldmann, P., & Sillanpää, M. J. (2004). Bayesian analysis of genetic differentiation between populations. *Genetics*, *163*, 367–374.

Crozier, W. W., Schön, P.-J., Chaput, G., Potter, E. C. E., Maoiléidigh, N. Ó., & MacLean, J. C. (2004). Managing Atlantic salmon (*Salmo salar* L.) in the mixed stock environment: Challenges and considerations. *ICES Journal of Marine Science*, *61*, 1344–1358. https://doi.org/10.1016/j.icesjms.2004.08.013

Dann, T. H., Habicht, C., Baker, T. T., & Seeb, J. E. (2013). Exploiting genetic diversity to balance conservation and harvest of migratory salmon. *Canadian Journal of Fisheries and Aquatic Sciences*, *70*, 785–793. https://doi.org/10.1139/cjfas-2012-0449

Dawson, K. J., & Belkhir, K. (2001). A Bayesian approach to the identification of panmictic populations and the assignment of individuals. *Genetics Research*, *78*, 59–77.

de Valpine, P., & Hastings, A. (2002). Fitting population models incorporating process noise and observation error. *Ecological Monographs*, *72*, 57–76. https://doi.org/10.1890/0012-9615(2002)072[0057:FPMIPN]2.0.CO;2

Dittman, A., & Quinn, T. (1996). Homing in Pacific salmon: Mechanisms and ecological basis. *Journal of Experimental Biology*, *199*, 83–91.

Gelman, A., & Rubin, D. B. (1992). Inference from iterative simulation using multiple sequences. *Statistical Science*, *7*, 457–511. https://doi.org/10.1214/ss/1177011136

Gilpin, M. (1996). Metapopulations and wildlife conservation: approaches to modeling spatial structure. In D.R. McCullough (Ed.), *Metapopulations and wildlife conservation*. (pp. 11–27). Washington, DC. Island Press.

Grant, W., Milner, G., Krasnowski, P., & Utter, F. (1980). Use of biochemical genetic variants for identification of sockeye salmon (*Oncorhynchus nerka*) stocks in Cook Inlet, Alaska. *Canadian Journal of Fisheries and Aquatic Sciences*, 37, 1236–1247. https://doi.org/10.1139/f80-159

Hanski, I. (1998). Metapopulation dynamics. *Nature*, 396, 41–49. https://doi.org/10.1038/23876

Hanski, I., & Thomas, C.D. (1994). Metapopulation dynamics and conservation: A spatially explicit model applied to butterflies. *Biological Conservation*, 68, 167–180. https://doi.org/10.1016/0006-3207(94)90348-4

Hazen, E. L., Jorgensen, S., Rykaczewski, R. R., Bograd, S. J., Foley, D. G., Jonsen, I. D., … Block, B. A. (2013). Predicted habitat shifts of Pacific top predators in a changing climate. *Nature Climate Change*, 3, 234–238.

Hoenig, J. M., Latour, R. J., & Olney, J. E. (2008). Estimating stock composition of anadromous fishes from mark–recovery data: Possible application to American shad. *North American Journal of Fisheries Management*, 28, 507–515. https://doi.org/10.1577/M07-072.1

ICES. (2015). Report of the Baltic Salmon and Trout Assessment Working Group (WGBAST), 23-31 March 2015, Rostock, Germany. ICES CM 2015\ACOM:08. 362 pp.

ICES. (2017). Report of the Baltic Salmon and Trout Assessment Working Group (WGBAST), 27 March–4 April 2017, Gdańsk, Poland. ICES CM 2017/ACOM:10. 298 pp.

Jutila, E., Jokikokko, E., Kallio-Nyberg, I., Saloniemi, I., & Pasanen, P. (2003). Differences in sea migration between wild and reared Atlantic salmon (*Salmo salar* L.) in the Baltic Sea. *Fisheries Research*, 60, 333–343. https://doi.org/10.1016/S0165-7836(02)00169-8

Kalinowski, S. T., Manlove, K. R., & Taper, M. (2007). ONCOR: a computer program for genetic stock identification. Bozeman, MT: Montana State University.

Kallio-Nyberg, I., Romakkaniemi, A., Jokikokko, E., Saloniemi, I., & Jutila, E. (2015). Differences between wild and reared Salmo salar stocks of two northern Baltic Sea rivers. *Fisheries Research*, 165, 85–95. https://doi.org/10.1016/j.fishres.2014.12.022

Karlsson, L., & Karlström, Ö. (1994). The Baltic salmon (*Salmo salar* L.): Its history, present situation and future. *Dana*, 10, 61–85.

Karlsson, L., Karlström, Ö., & Hasselborg, T. (1995). *Laxens lekvandringstid i Bottniska vikens kustområden och dess samband med havsvattentemperaturen*. Älvkarleby, Sweden: Laxforskningsinstitutet Meddelande 1/1995.

Koljonen, M.-L. (2006). Annual changes in the proportions of wild and hatchery Atlantic salmon (*Salmo salar*) caught in the Baltic Sea. *ICES Journal of Marine Science*, 63, 1274–1285. https://doi.org/10.1016/j.icesjms.2006.04.010

Koljonen, M.-L., Jansson, H., Paaver, T., Vasin, O., & Koskiniemi, J. (1999). Phylogeographic lineages and differentiation pattern of Atlantic salmon (*Salmo salar*) in the Baltic Sea with management implications. *Canadian Journal of Fisheries and Aquatic Sciences*, 56, 1766–1780. https://doi.org/10.1139/f99-104

Kuikka, S., Vanhatalo, J., Pulkkinen, H., Mäntyniemi, S., & Corander, J. (2014). Experiences in Bayesian Inference in Baltic salmon management. *Statistical Science*, 29, 42–49. https://doi.org/10.1214/13-STS431

Kulmala, S., Haapasaari, P., Karjalainen, T. P., Kuikka, S., Pakarinen, T., Parkkila, K., … Vuorinen, P. J. (2013). TEEB Nordic case: Ecosystem services provided by the Baltic salmon. Retrieved from http://www.teebweb.org/

Lundqvist, H., Rivinoja, P., Leonardsson, K., & McKinnell, S. (2008). Upstream passage problems for wild Atlantic salmon (*Salmo salar* L.) in a regulated river and its effect on the population. *Hydrobiologia*, 602, 111–127. https://doi.org/10.1007/s10750-008-9282-7

Manel, S., Gaggiotti, O. E., & Waples, R. S. (2005). Assignment methods: Matching biological questions with appropriate techniques. *Trends*

in Ecology and Evolution, 20, 136–142. https://doi.org/10.1016/j.tree.2004.12.004

Maunder, M. N., Deriso, R. B., & Hanson, C. H. (2015). Use of state-space population dynamics models in hypothesis testing: Advantages over simple log-linear regressions for modeling survival, illustrated with application to longfin smelt (*Spirinchus thaleichthys*). *Fisheries Research*, 164, 102–111. https://doi.org/10.1016/j.fishres.2014.10.017

Maunder, M. N., & Punt, A. E. (2013). A review of integrated analysis in fisheries stock assessment. *Fisheries Research*, 142, 61–74. https://doi.org/10.1016/j.fishres.2012.07.025

McKinnell, S., Lundqvist, H., & Johansson, H. (1994). Biological characteristics of the upstream migration of naturally and hatchery-reared Baltic salmon, *Salmo salar* L. *Aquaculture and Fisheries Management*, 25, 45–63.

Michielsens, C. G. J., Mäntyniemi, S., & Koljonen, M.-L. (2004). The use of genetic stock identification results for the assessment of wild Baltic salmon stocks. ICES CM 2004/EE:03.

Millar, R. B. (1987). Maximum likelihood estimation of mixed stock fishery composition. *Canadian Journal of Fisheries and Aquatic Sciences*, 44, 583–590. https://doi.org/10.1139/f87-071

Milner, G. B., Teel, D. J., Utter, F. M., & Burley, C. L. (1981). National Marine Fisheries Service, Columbia River stock identification study: validation of genetic method. Final Report to Bonneville Power Administration (Contract No. 1980BP18488), BPA Report DOE/BP-18488-1.

Whitlock, R., Mäntyniemi, S., Palm, S., Koljonen, M.-L., Dannewitz, J., & Östergren, J. (2017). Data from: Integrating genetic analysis of mixed populations with a spatially explicit population dynamics model. *Dryad Digital Repository*, https://doi.org/10.5061/dryad.4pg37

Östergren, J., Lind, E., Palm, S., Tärnlund, S., Prestegaard, T., & Dannewitz, J. (2015). Stamsammansättning av lax i det svenska kustfisket 2013 & 2014 – Genetisk provtagning och analys. Sveriges lantbruksuniversitet (SLU), institutionen för akvatiska resurser. 30 sidor.

Pella, J., & Masuda, M. (2001). Bayesian methods for analysis of stock mixtures from genetic characters. *Fishery Bulletin*, 99, 151–167.

Plummer, M. (2015). *JAGS version 4.0.0 user manual*. Retrieved from https://sourceforge.net/projects/mcmc-jags/files/Manuals/4.x/jags_user_manual.pdf.

Plummer, M. (2016). *rjags: Bayesian Graphical Models using MCMC*. R package version 4-6. Retrieved from https://CRAN.R-project.org/package=rjags

Pritchard, J. K., Stephens, M., & Donnelly, P. (2000). Inference of population structure using multilocus genotype data. *Genetics*, 155, 945–959.

R Core Team. (2015). *R: A language and environment for statistical computing*. Vienna, Austria: R Foundation for Statistical Computing. Retrieved from https://www.R-project.org/

Säisä, M., Koljonen, M.-L., Gross, R., Nilsson, J., Tähtinen, J., Koskiniemi, J., & Vasemägi, A. (2005). Population genetic structure and postglacial colonization of Atlantic salmon (*Salmo salar*) in the Baltic Sea area based on microsatellite DNA variation. *Canadian Journal of Fisheries and Aquatic Sciences*, 62, 1887–1904. https://doi.org/10.1139/f05-094

Schaub, M., & Abadi, F. (2011). Integrated population models: A novel analysis framework for deeper insights into population dynamics. *Journal of Ornithology*, 152, 227–237. https://doi.org/10.1007/s10336-010-0632-7

Schaub, M., Gimenez, O., Sierro, A., & Arlettaz, R. (2007). Use of integrated modeling to enhance estimates of population dynamics obtained from limited data. *Conservation Biology*, 21, 945–955. https://doi.org/10.1111/j.1523-1739.2007.00743.x

Schindler, D. E., Hilborn, R., Chasco, B., Boatright, C. P., Quinn, T. P., Rogers, L. A., & Webster, M. S. (2010). Population diversity and the portfolio effect in an exploited species. *Nature*, 465, 609–612. https://doi.org/10.1038/nature09060

Siira, A., Erkinaro, J., Jounela, P., & Suuronen, P. (2009). Run timing and migration routes of returning Atlantic salmon in the Northern Baltic Sea: Implications for fisheries management.

Fisheries Management and Ecology, *16*, 177–190. https://doi.org/10.1111/j.1365-2400.2009.00654.x

Somveille, M., Rodrigues, A. S., & Manica, A. (2015). Why do birds migrate? A macroecological perspective. *Global Ecology and Biogeography*, *24*, 664–674. https://doi.org/10.1111/geb.12298

Taylor, N. G., McAllister, M. K., Lawson, G. L., Carruthers, T., & Block, B. A. (2011). Atlantic bluefin tuna: A novel multistock spatial model for assessing population biomass. *PLoS ONE*, *6*, e27693. https://doi.org/10.1371/journal.pone.0027693

Utter, F. M., & Ryman, N. (1993). Genetic markers and mixed stock fisheries. *Fisheries*, *18*, 11–21. https://doi.org/10.1577/1548-8446(1993)018<0011:GMAMSF>2.0.CO;2

Westerberg, I. (1982). Ultrasonic tracking of Atlantic salmon (*Salmo salar* L.) 1. Movements in coastal regions. *Report: Institute of Fresh-water Research. Drottningholm*, *60*, 82–101.

Whitlock, R. E., Hazen, E. L., Walli, A., Farwell, C., Bograd, S. J., Foley, D. G., ... Block, B. A. (2015). Direct quantification of energy intake in an apex marine predator suggests physiology is a key driver of migrations. *Science Advances*, *1*, e1400270. https://doi.org/10.1126/sciadv.1400270

SUPPORTING INFORMATION

Additional Supporting Information may be found online in the supporting information tab for this article.

How to cite this article: Whitlock R, Mäntyniemi S, Palm S, Koljonen M-L, Dannewitz J, Östergren J. Integrating genetic analysis of mixed populations with a spatially explicit population dynamics model. *Methods Ecol Evol*. 2018;9:1017–1035. https://doi.org/10.1111/2041-210X.12946