Contents

Summary

Gangsei, L.E. (2013) A Bayesian method for estimating moose (*Alces alces*) population size based on hunter observations and killed at age data.
Master Thesis 2013. Norwegian University of Life Sciences.

Lots of wild species, fish and mammals, are heavy harvested through fishing and hunting. Reliable population size estimates are valuable management tools for these species. In cases where killed at age data are available, models outlined under the framework known as "cohort analysis" or "virtual population analysis (VPA)" are used extensively. In fish stock management several models using Bayesian techniques have been developed through the last two decades.

In this study a model using a Bayesian approach for estimating moose population size is examined. The model combines principles from *discrete time series analysis*, where basic cohort analysis based on killed at age data constitutes the bulk, and *analysis in continuous time* for each hunting season based on data from hunter observations. The analysis in continuous time aims to find age- and year-specific expressions for the hunting mortality rate. In the discrete time series analysis, the hunting yield is viewed as a binomially distributed variable, with pre-harvest population size as "number of trials" and mortality rate derived from the analysis in continuous time as "probability parameter". All basic principles are known from previous surveys, but the way they are assembled is, to the authors knowledge, innovative.

The model performed very well when tested against simulated populations with known parameter values. For real data tests are conducted through cross-validation based on spatial subsets and by comparing results from temporal subsets. Generally the model performed well in these test. However, an issue is revealed by comparing results from different temporal subsets, since the hunters ability to spot moose seems to develop over time (years) and/or depend on moose density. This issue should not terminate the practical implementation of the model. If a satisfying solution to the issue is achieved, it might have a possible positive impact on other methods for estimating abundance of wild species based on *effort*, a very prevalent class of models.

The real data used for testing the model, and to demonstrate some practical interpretations, are from the municipality of Ringerike in southern Norway. Killed at age data are available from 1988 till 2012 in combination with hunter observations. The estimates

show a moose population size rapidly increasing in the period from 1988 till its peak in 1993 at a posterior mean population size of approximately 3900 individuals. Thereafter, in line with large hunting yields, reduced reproductivity rate and increased rate of natural mortality, the population size declined rapidly till an estimated pre-harvest population size of approximately 1700 individuals in year 2000. Thereafter the total population size has been estimated as quite stable, but with a declining trend over the last couple of years.

Usually the natural (non harvest) mortality rate is assumed fixed and known when cohort analysis methods are used for estimating abundance of wild species. The model presented in this study is capable of producing reliable, and to some extent practical beneficial, posterior distributions for the natural mortality rate based on an informative prior distribution and an adequate amount of data. These posterior distributions for natural mortality rates indicate surprisingly high rates for the years around 1993.

Sammendrag

(*Norwegian summary*)

Populasjonene av en rekke ville arter, såvell fisk som pattedyr, blir hardt beskattet gjennom jakt og fiske. For forvaltningen av slike arter vil sikre bestandestimater være et verdifullt verktøy. I tilfeller hvor aldersregistreringer av uttaket er tilgjengelig benyttes modeller kjent som "kohort-analyse" regelmessig. I fiske-forvaltningen har ulike Bayesianske modeller blitt utviklet gjennom de siste 20 år.

I denne studien presenteres og evalueres en Bayesiansk modell for å estimere bestandsstørrelse hos elg. Modellen kombinerer prinsipper fra analyse av *diskrete tidsserier*, hvor klassisk kohort-analyse utgjør hoveddelen, og prinsipper for analyse av *kontinuerlige tidsserier* basert på jegerdata fra de enkelte jaktsesongene. Analysen i kontinuerlig tid tar sikte på å finne alder- og års-spesifikke uttrykk for jaktdødelighetsraten til elg. I analysen av den diskrete tidsserien er jaktuttaket sett på som en binomialt fordelt variabel, med populasjonsstørrelsen før jakt som "antall-parameter" og jaktdødeligheten utledet i analysen gjennom kontinuerlig tid som "sjanse-parameter". Alle basis-prinsippene er kjent fra tidligere studier, men måten prinsippene er kombinert på er, så langt forfatteren kjenner til, en nyvinning.

Modellen fungerte svært tilfredsstillende når den ble testet mot simulerte populasjoner med kjente parametere. Med grunnlag i ekte data ble modellen testet gjennom kryssvalidering (cross-validation) basert på data i romlige (spatial) undergrupper. Modellen fungerte gjennomgående svært tilfredsstillende også i disse testene. Imidlertid ble et avvikende forhold avdekket gjennom bruk av data fra ulike tidsperioder. Jegernes estimerte evne til å oppdage elg er avhengig av hvilken tidsperiode datagrunnlaget hentes fra. Problemet bør ikke forhindre at modellen taes i bruk, men avviket bør utredes ytterligere. Dersom man finner en tilfredsstillende løsning kan det få følger for en rekke metoder hvor populasjonsstørrelser estimeres med grunnlag i jakt-/fiske-innsats.

Dataene benyttet i denne studien kommer fra Ringerike kommune i Buskerud. Data fra aldersregistreringer og jegerobservasjoner er tilgjengelige fra 1988. Estimatene viser en elgbestand i rask vekst fra 1988 til 1993 da elgtettheten nådde toppen med en posteriori gjennomsnittlig totalpopulasjon på ca. 3900 elg før jakt. Etter dette, samtidig med at jaktuttakene var høye, den estimerte naturlige dødelighetsraten var høy, og reproduksjonsraten (kalveraten) minket, falt elgtettheten jevnt frem til år 2000, da estimert

bestandsstørrelse før jakt var på ca. 1700 elg. Etter dette har den beregnede elgtettheten
vært relativt stabil, dog med en markert nedgang de siste par årene.

Når kohorts-analyse benyttes for ville arter er det er vanlig å anta at den naturlige
dødelighetsraten er konstant og kjent. Modellen som presenters i denne studien gir tro-
verdige, og i stor grad praktisk nyttige, posteriori fordelinger for den naturlige dødelighetsraten.
Riktig nok må informative "prior"-fordelinger benyttes, og datamaterialets størrelse må
være betydelig. Beregningene av den naturlige dødelighetsraten viser overraskende store
verdier for årene rundt 1993.

Preface

Killed at age data for moose have been an essential part of my profession for the last decade as an employee at the management agency Faun Naturforvaltning AS. Variations of the model presented in this study were routinely applied to several moose populations in southern Norway, including the Ringerike moose population. As time went by my interest for the theoretical statistical aspects of the model, and statistics in general, became dominant compared to biological challenges. Eventually this led me to apply for a Master degree in Applied Statistics at Norwegian University of Life Sciences (UMB).

In the present thesis a draft for a scientific paper focusing on the statistical aspects of the method constitute the bulk. As an introduction to the paper there are added some thoughts of how improved statistical methods might contribute to practical moose management, and possibly also to the management of other species. Further, the introduction contains some paragraphs about Bayesian statistic with emphasis on Marcov Chain Monte Carlo sampling, since these methods are essential in the present study.

My supervisors Trygve Almøy and in particular Solve Sæbø have been most helpful whenever I have had questions during my period as student. Even more importantly, during my professions as a biologist, I was inspired by them to study statistics, in particular as a consequence of Solves's totally unselfish contributions to the development of the model presented in this study.

The data from Ringerike exist due to meticulously registrations performed by hunters over a period of 25 years, for which their reward is restricted to my heartfelt thanks. I am also grateful to my former employer, Faun Naturforvaltning AS, for providing me the opportunity to work with the theme and making all data easily available.

Åsmund Pålerud has made a huge contribution to my knowledge and engagement by teaching me the practical age determination technique and basic principles for cohort analysis.

Finally I will give Hans Bergan, former secretary for the Ringerike Wildlife Board, a big thanks. He has been the driving force for collecting killed at age data in Ringerike. Further he has taught me all I know about practical moose management, which in the name of justice might be glaring inadequate, at least if judged by some hunters.

Fyresdal, May 2013
Lars Erik Gangsei

Chapter 1

Introduction

Are moose population size estimates demanded?

Results from a preliminary version of the model presented in this study was an essential foundation for advises regarding moose management given by the author in his former carrier as consultant in the management agency Faun Naturforvaltning AS. The feedback from the customers was good and the population size estimates derived from the model were regarded valuable. The need for methods providing absolute population size estimates for moose, and other deers-species (*Cervidae*), is also emphasized by others, see for instance Pedersen (2009, pp.27–28).

The model described in the present study should, in combination with the large dataset from Ringerike, form a solid basis for further biological research on moose, and thereby contribute with valuable knowledge for larger scale areas. In the following sections possible applications for the model are outlined.

Moose population - Costs and assets

During the second half of the 20th century the moose density in south-eastern Norway increased sharply as a result of changed harvesting strategies and higher food-production (Direktoratet for Naturforvaltning, 1995). Currently moose is the most important hunting game in Norway, at least if measured as meat yield.

Moose populations are managed primarily by the municipalities and the landowners. Various municipalities and landowners may have different targets for their management. A consensual ideal target is a moose management which maximizes the benefits and minimizes the cost for society. However, in practice this ideal comes with a great variety of

interpretations.

A common management goal, in order to preserve moose hunting as a valuable asset, is to stabilize moose population sizes at levels where they contribute with large and sustainable hunting yields. A major challenge in this context has been a generally decreasing production capacity in the moose populations over the last two decades, probably caused by density dependent factors (Solberg et al., 2006).

Moose populations contribute with substantial costs for society. The most important is traffic accidents involving moose (Solberg et al., 2009). There is a tight connection between moose density and the expected number of traffic accidents involving moose (Rolandsen et al., 2011). In some areas damage from moose browsing on young pine (*Pinus sylvesteris*) forest causes large economical losses for landowners (Direktoratet for Naturforvaltning, 1995). Over-harvesting of some tree species might have important influence on ecosystems and their biodiversity (Aanderaa et al., 1996). Parasites like deer ked (*Lipoptena cervi*) and sheep tick (*Ixodes ricinus*), both nuisances and potentially dangerous for humans, have their main hosts in the deer family. In practice moose is a crucial host for deer ked in Scandinavia (Välimäki et al., 2010), whereas sheep tick has a wider specter of hosts (Jaenson et al., 2012).

Due to these considerable costs, the main management policy in south-eastern Norway over the last two decades has been to reduce moose density. This reduction must be carried out through hunting since the population sizes of natural moose predators, wolf (*Canis lupus*) and brown bear (*Ursus arctos*), are insufficient to regulate the moose density over large areas (Wabakken et al., 2011, Tobiassen et al., 2012).

If, and from a political point of view there certainly exist an "if", the moose management should aim at managing the moose population in a manner where the gain for society is maximized, the need for further knowledge is substantial.

From general biology its known that to some extent the production capacity in the moose population will be density dependent. However, quantified knowledge about the interactions between moose density and production capacity is scarce. Some questions in need of quantified answers are:

- When moose density increases, will production capacity decrease steadily or do thresholds exist?

- Are time-lags present for the interaction between moose density and production capacity? If so can they be quantified?

- Some areas might be more productive areas regarding moose than others. Which

factors are decisive for the "moose production capacity" in an area?

- To what extent do the composition of the hunting yield influence subsequent years moose population, and its production capacity?

Questions similar to the ones listed above might be asked for the interactions between moose density and traffic-accidents involving moose, moose density and forest damage, moose density and the prevalence of deer ked and sheep tick, etc.

The theme for this thesis is statistic and not biology. Therefore it will be made no attempt to make the list of questions exhaustive. Neither will the questions listed above be answered. However, answering these kind of questions prerequisite reliable population size estimates, like the ones obtained from the model presented in this study.

Further work - development of the model

Validation on other moose populations

Through Faun Naturforvaltning AS killed at age data are available for a number of municipalities in southern Norway. However, the data-series from Ringerike are the better ones in the sense of having large average annual hunting yields, covering the longest period of time and containing killed at age data close to completeness. Consequently, using data from other municipalities in southern Norway for model validation is anticipated to be of limited utility.

Norwegian Institute for Nature Research (NINA) disposes killed at age data for several areas in Norway. Of special interest is the data-series from the municipalities Grane, Vefsn and Hattfjelldal in the county of Nordland in Northern Norway, where killed at age data are available from 1967, see for instance Solberg & Sæther (1999). Testing the model on these data would be anticipated to contribute substantially to the model's validity.

Model adjustment for areas lacking killed at age data

Hunter observations are present for all parts of Norway. Most of these data are archived in the the National deer register, at the website *www.hjorteviltregisteret.no* (National deer register, 2013), where the access is free. Hunter observations contain registrations of number of hunting days, and the number of observed moose in different age-, sex- and reproductive categories. Hunter observations give valuable information about sex ratios, reproduction rates and relative changes in moose density.

Killed at age data is a prerequisite for the model presented in this study. The limited prevalence of killed at age data for moose, and most other species, restricts the implementation of the model. Even if killed at age data lack for most areas, all municipalities have registrations of the total hunting yield, where the individuals are classified as calves, yearlings or older. Of course all individuals registered as "older" had a certain age when they were shot, this age is just unknown. By handling these unknown ages as "missing data" an adjusted version of the model presented in this study can be applied to areas lacking killed at age data. It might be necessary to use informative priors for some of the parameters, and especially the new variable, i.e. the age distribution in the hunting yield. It is also likely that parameter estimates from areas lacking killed at age data will have larger variability than estimates from areas where killed at age data are present.

Migration, natural mortality and production capacity

Surveys on moose marked with GPS- and/or radio collars in Norway generally show a clear pattern with well defined and stable home-ranges for adult moose. In many areas the majority of individuals has separate "summer-" and "winter-areas". In the present study population sizes during the hunting seasons are analyzed. Then seasonal moose movements may be ignored since the moose only at rare occasions will be in the "winter-areas" during the hunting season. True migration, i.e. permanent shift of home-range, is a matter of greater concern. When the offspring is chased from their mothers, at the age of approximately one year, they tend to establish their new home-ranges a considerable distance away from their mothers home-range. Moose in the county Nord-Trøndelag showed a pattern where 50% ($n = 20$) migrated, i.e. established their home-range apart from their mothers (Rolandsen et al., 2010, pp. $73 - 74$). Even though the typical distance between home-ranges for offspring and mothers is less than 70 km, the probability of offspring establishing home-range fully or partly in a neighbouring municipality is considerable.

Even if the migration is random in space, a situation with different moose-densities among neighbouring municipalities, would lead, or at least be expected to lead, to net immigration in the areas with the lower moose density and vice versa. In Gangsei (1999) an attempt to estimate the net migration size was carried out. The basic idea was to estimate the net migration as the difference between cohort sizes for calves estimated by basic cohort analysis, and estimates based on population size for adult females (by cohort analysis) and reproduction rates as observed in hunter observations. These calculations show a net emigration from Ringerike in the period from 1988 till around 1992 followed by a period with high immigration during most of the 90-s. This pattern seems intuitively

right since Ringerike had relatively low hunting quotas during the 80-s and high quotas during the 90-s compared with its neighbouring municipalities.

However, the migration estimates from Gangsei (1999) have several weaknesses. The estimates are based on a fixed and non-stochastic natural mortality rate at 5% between hunting seasons for all age-classes and years. As demonstrated in the present analysis, this assumption is dubious at best. Secondly Gangsei (1999) seems to underestimate the population size, which probably causes too high estimates for immigration during the 90-s.

Even though the migration estimates from Gangsei (1999) might be criticised, a natural next step for the present study is to model migration as outlined there. A likely outcome would be lower estimates for the natural mortality rate in years around 1993, combined with substantial immigration in the following years.

Statistics Norway's registrations of moose perishing from natural causes (Statistics Norway, 2013) are divided into 7 different categories. For most of these categories, for instance individuals killed in traffic incidents, the registrations are assumed to be close to exhaustive. However, for a couple of categories, "poaching" and "other causes", it is likely that just an unknown fraction of the real number of deaths is registered.

There are several opportunities for further investigation of the relationship between results from Statistics Norway and the estimated number of natural dead moose. These options seems to be suitable in an biological context, i.e. if the goal is to investigate variation in natural mortality among moose.

Killed at age data contain huge amounts of information unexploited in the present study. Under a Bayesian regime it should be straight forward to apply carcass-weights to all individuals in the population, based on carcass-weights from the hunting yield as registered in the killed at age data. Then estimates for meat production per -cohort, -age-class, -year etc. would be achievable. Such estimates would be interesting by virtue of themselves. Of even greater interest is the interaction between production estimates and -density, -age distribution, -sex ratio, etc. in the moose population.

Transferability to other species

The origin of cohort analysis is fishing management, and model improvement for cohort analysis are attributed primarily to research on fish stocks. Cohort analysis forms the management foundation for some fish stocks of large economical interest. Even though data for *effort* might be present for commercial fisheries, the fisheries lack data corresponding to the hunter observations present for moose. Consequently some of the results from the analysis in this study, especially the results regarding the variations in natural mortality rate, and

the results showing discrepancy when assuming constant instantaneous observation rate, might shed some new light on the models used for fish and possibly also other species. An equal reasoning might be applied to the huge class of population abundance models, not necessarily relying on cohort analysis, using *effort* as an explanatory variable.

Bayesian statistics

Background

This section lists an overview over some basic principles and methods in Bayesian statistics and estimation, with emphasis on Marcov Chain Monte Carlo (MCMC) sampling. MCMC is not a Bayesian method per se, but is used extensively in Bayesian estimation. A lot of the principles presented are from canonical studies. However, the whole section might be regarded as a brief overview of chapters 1 to 5 in Gilks et al. (1996), and central parts of Dellaportas & Roberts (2003).

Bayesian inference

The term "Bayesian" refers to Thomas Bayes (1702-1761), who proved a special case of what is now known as Bayes' theorem:

$$P(A|B) = \frac{P(B|A) \times P(A)}{P(B)}$$

In classical frequentist statistics "A" and "B" are seen as events. Bayes' theorem shows the connection between the conditional and unconditional probabilities for the two events "A" and "B".

Under a Bayesian inference the elements in Bayes' theorem are interpreted as probability density functions (pdf-s). A common notation is to denote "parameters" by $\boldsymbol{\theta}$, "data" by $\boldsymbol{y}$, and use π as function indicator, giving:

$$\pi(\boldsymbol{\theta}|\boldsymbol{y}) = \frac{\pi(\boldsymbol{y}|\boldsymbol{\theta}) \times \pi(\boldsymbol{\theta})}{\pi(\boldsymbol{y})}$$

The probability density function $\pi(\boldsymbol{\theta}|\boldsymbol{y})$ is known as the *posterior distribution*. Notice that the posterior distribution has the interpretation *probability of parameter given data*, which is the exact target for most statistical analyses. A related interpretation is also common, though slightly erroneous, for frequentist confidence intervals, a point emphasized by a lot of Bayesian statisticians.

The posterior distribution is a function of $\pi(\boldsymbol{y}|\boldsymbol{\theta})$, known as the *likelihood*, $\pi(\boldsymbol{\theta})$, known as the *prior*, and $\pi(\boldsymbol{y})$, known as the *marginal likelihood*.

The marginal likelihood is a q-dimensional integral, where q is the dimension of $\boldsymbol{\theta}$. The marginal likelihood might be viewed as a constant making sure that the posterior

distribution integrates to 1. Luckily, since calculation of the marginal likelihood often is impossible or extremely demanding, analysis of the marginal density is usually not really required, as analysis is rather driven by the posterior distribution, $\pi(\boldsymbol{\theta}|\boldsymbol{y})$. However, calculation of $\pi(\boldsymbol{y})$ is nevertheless of importance in some situations. One example is the need of numerical values for $\pi(\boldsymbol{y})$ when dealing with issues of different candidate models.

The likelihood, $\pi(\boldsymbol{y}|\boldsymbol{\theta})$, is the basis for traditional likelihood based frequentist statistics, where it is analysed for instance to provide *maximum likelihood estimates*. The likelihood also plays a major role in Bayesian statistics. Besides being a central part in Bayes' theorem the likelihood is central when calculating different model selection criteria, which will be discussed briefly later in this section.

The last element in Bayes' theorem is the prior, $\pi(\boldsymbol{\theta})$, a probability density function reflecting the knowledge about parameters, $\boldsymbol{\theta}$, before the data, $\boldsymbol{y}$, are taken into account. The prior might be viewed both as the Achilles heel and Silver bullet in Bayesian statistics. The prior has, even if we would want to avoid it, influence on the posterior distribution. Since the prior, at least ideally, should be independent of data, there will always be a part of subjectivity reflected in the prior. This is a major point in traditional frequentist criticism of Bayesian statistics, see for instance Gelman (2008). On the other hand the prior offers a well defined option for integrating knowledge not reflected in the data into the model. Ideally the prior distribution should be a proper probability distribution, i.e. positive over the whole parameter-space and integrate to 1. However, so-called flat priors are commonly applied to parameters with undefined endpoints, making the priors improper. Mostly such priors enables satisfactory results, under the prerequisite of a proper posterior distribution.

Several different types of priors, not always clearly distinguished, are utilized. The following list shows some main groups, but is by no means exhaustive:

- Flat priors: That is priors $\propto 1$. Even if the "gut-feeling" is that such priors should have no influence on the posterior distribution, this is not always the case. However, their influence on the posterior distribution are mostly minor.

- Vague priors: Proper priors carrying little information.

- Jeffrys prior: Has the advantage of being invariant under reparametrization of the likelihood.

- Informative priors:

    - Empirical Bayes: Using data to find prior. "Cheating, but sometimes effective".

  – Using moments, i.e. expectation, variance or higher moments, from knowledge independent of data, to find prior distributions.

  – Using percentiles from knowledge independent of data, to find prior distributions.

- Conjugate priors: When using a conjugate prior the posterior distribution has the same distribution as the prior, but with adjusted parameter-values based on data. Models where the likelihood comes from the exponential family represent the majority of distributions where conjugate priors are achievable. For all exponential family likelihoods there exists a conjugate prior (Diaconis & Ylvisaker, 1979), which is often also in the exponential family.

Decision theory - The Master Recipe for finding the Bayes solution

If we want to make decisions, a simple example might be to estimate ("decide") a parameter value, the Bayesian framework fits extremely well since there exist a well defined "Bayes' solution" (Carlin & Louis, 2011, pp.429–430, Hjort, 2012). This solution depends on known:

- Likelyhood function, $\pi(\boldsymbol{y}|\boldsymbol{\theta})$,

- Prior function, $\pi(\boldsymbol{\theta})$,

- Data, $\boldsymbol{y}$,

- Loss function, $L(\theta, a)$, where a is the "action" or "decision"

Then the Bayes' solution, $\hat{a}_B$, is given by $\hat{a}_B(\boldsymbol{y}) = arg_{min}(E(L(\theta, a)|\boldsymbol{y}))$ (Hjort, 2012). That is the estimator *minimizing expected posterior loss*. It might be shown that the Bayes' solution is given by:

- $\hat{a}_B(\boldsymbol{y}) = $ "Posterior mean" under quadratic loss funcion, i.e. $L(\theta, a) = (\theta - a)^2$

- $\hat{a}_B(\boldsymbol{y}) = $ "Posterior median" under absolute loss funcion, i.e. $L(\theta, a) = |\theta - a|$

- $\hat{a}_B(\boldsymbol{y}) = $ "Posterior mode" under the loss function $L(\theta, a) = 0$ if $|\theta - a| < C_1$ and $L(\theta, a) = C_2$ if $|\theta - a| \geq C_1$, as $C_1 \to 0$ and $C_2 > 0$.

Different "ad-hoc" loss functions might be applied to the "Master Recipe". For instance; in practice moose managers might prefer an underestimation of true population size rather than an overestimation. A loss function reflecting this fact might be added to the model

presented in this study. Such a loss function would affect the Bayes estimates. However, posterior means, medians etc. would be unaffected, they would just not be the the Bayes estimates under the "ad-hoc" loss function.

Marcov Chain Monte Carlo (MCMC) - sampling

Even though Bayesian statistics has many theoretical benifits, like the "Master Recipe" shown in the preceding paragraphs, benefits that have been well known for a long time, its practical implementation was restricted to situations where analytical solutions were available, mainly very simple situations with conjugate priors or marginal likelihoods available in closed form.

In the 1980s, there was a dramatic growth in research and applications of Bayesian methods, mostly attributed to the discovery of Markov chain Monte Carlo (MCMC) methods, which removed many of the computational problems. This discovery, combined with a rapid increase in computer power, changed the situation for Bayesian statisticians dramatically. Suddenly the Bayesian framework offered a practical solution to complex models which could be hard to solve using traditional frequentist methods.

The principle for MCMC-methods is, as stated by its name, a combination of Marcov Chains and Monte Carlo sampling. The posterior distribution is not analyzed analytically, but the MCMC algorithm allows for sampling of values for $\boldsymbol{\theta}$ from the true posterior distribution, $\pi(\boldsymbol{\theta}|\boldsymbol{y})$. By sampling an adequate number of values the posterior distribution can be analyzed by Monte Carlo methods. In addition the MCMC approach has a straightforward solution for handling missing data, denoted $\boldsymbol{y}_{mis}$ (Dellaportas & Roberts, 2003).

A Marcov Chain is memoryless. That is the next state of a variable θ (under Bayesian inference the parameters, $\boldsymbol{\theta}$, might be viewed as random variables), denoted θ_{n+1}, of the chain depends only on the current state, denoted θ_n, and not on the sequence of events that preceded it, or formally it might be written; $\pi(\theta_{n+1} = \theta_0|\theta_n, \theta_{n-1}, \ldots, \theta_1) = \pi(\theta_{n+1} = \theta_0|\theta_n)$.

The main idea in MCMC sampling is to update the elements of $\boldsymbol{\theta}$ and $\boldsymbol{y}_{mis}$ in blocks or one at a time by sampling values from the conditional posterior distribution assuming all other elements of $\boldsymbol{\theta}$ and $\boldsymbol{y}_{mis}$ to be known. Even though the full conditional distribution might be very complex for the full set of parameters it usually simplifies substantially for each element updated through the Marcov Chain. Parts of the full conditional distribution independent of the parameter- (or missing data) element in question should be excluded from the sampling algorithm when a new value is drawn.

Sampling directly from the full conditional distribution is known as *Gibbs sampling* (Gelfand et al., 1990; Gelfand & Smith, 1990; Geman & Geman, 1984). Surprisingly often it is possible to find prior distributions for a given parameter that is a conjugate prior to the full conditional distribution. Then, as long as the prior in question is a known distribution, sampling a value from the full conditional is straightforward. If a conjugate prior is not available, Gibbs-sampling might still be possible, for instance by *rejection sampling* or *adaptive rejection sampling* (Gilks & Wild, 1992).

Sometimes it is not possible to draw values directly from the full conditional distribution. By applying an algorithm known as Metropholis-Hastings algorithm (MH-algorithm) (Hastings, 1970; Metropolis et al., 1953) to the Marcow Chain, the sampled values will still come from the true posterior distribution. The algorithm is easily applied to Marcov Chains, and makes use of a proposal function, commonly denoted $q(\cdot)$. The proposal function samples a candidate value, denoted θ^p for the parameter-element (or elements) in question, based on the current state, denoted θ_n, of the element. The candidate value is accepted with probability as shown below. If θ^p is accepted then $\theta_{n+1} = \theta^p$, and if θ^p is not accepted then $\theta_{n+1} = \theta^n$.

$$P\left(\theta^p \text{ is accepted}\right) = min\left(1, \frac{\pi(\boldsymbol{y}|\theta^p) \cdot \pi(\theta^p) \cdot q(\theta_n|\theta^p)}{\pi(\boldsymbol{y}|\theta_n) \cdot \pi(\theta_n) \cdot q(\theta^p|\theta_n)}\right)$$

If the proposal-function is symmetric, i.e. $q(\theta^p|\theta_n) = q(\theta_n|\theta^p)$, then the $q(\cdot)$-terms in the equation shown above are skipped. Then the algorithm is a pure "Metropolis" algorithm. There are a number of different principles for making effective proposal functions. If the acceptance rate, i.e. the probability that a candidate value is accepted, is too high, mixing is slow due to "small steps" and high autocorrelation in the Marcov Chain. On the other hand if the acceptance rate is very low, then mixing is also unsatisfying since the same parameter-values tend to be sampled for several iterations. Under some ideal conditions acceptance-rates between 23% ($dim(\boldsymbol{\theta}) \to \infty$) and 44% ($dim(\boldsymbol{\theta}) = 1$) are preferable (Gelman et al., 1996).

In Bayesian MCMC-sampling it is fully acceptable to mix Gibbs-sampling and MH-sampling. However, in my experience, effort should be allocated to making Gibbs-sampling possible. For the model presented in this study Gibbs-sampling is possible for all parameter elements and missing data, a result of meticulous work for finding conjugating prior functions.

The Marcow Chain has to be initiated, that is some initial values for $\boldsymbol{\theta}$ and $\boldsymbol{y}_{mis}$ must be set. These initial values do not need to be close to the asymptotic value for the Markov Chain. It can be shown that the asymptotic value for the Marcow chain is the true posterior

distribution. When the Marcow Chain has reached the asymptotic level its said to have converged. Values sampled prior to the time of convergence has to be eliminated from the Monte Carlo analysis. The steps between initialization and convergence are commonly known as *burn-in period*.

The results are evaluated through Monte Carlo methods, which might be most common in, but not exclusive to, Bayesian statistic. First the burn-in period has to be excluded from further analysis. There exist some formal methods for deciding how much of the trace that should be regarded as burn-in. However, a common, and easy, way to decide the length of burn-in is actually to "analyse" a trace plot of the Marcow Chain in question by eye, and "see" where the chain converges.

The output from a Marcov Chain is potentially highly autocorrelated. Sometimes a set of approximately independent estimates from the Marcov Chain is required. Then a method called "thinning" might be applied. Thinning is actually to pick values from the Marcov Chain with a given interval between them. The interval size should be set so the sample autocorrelation for the selected values is approximately zero.

Parameter estimates are mainly presented as *point estimates*, *Credibility intervals* or *percentiles*. Credibility intervals are based on MCMC output and are analogues to the frequentist Confidence interval. The limits for a $100(1 - \alpha)\%$ credibility interval is given by the two values (or average of two values) from the output excluding $100(\alpha/2)\%$ and $100(1 - \alpha/2)\%$ of the output values when sorted in increasing order.

Model selection - Deviance Information Criteria (DIC)

There are numerous methods for model selection. When comparing k different models, $M_1, \ldots, M_k$, under Bayesian inference, $\pi(M_j)$, $j = 1, \ldots, k$, might be seen as the prior (unaffected by data) probability of model M_j being the best model. It might be shown (Hjort, 2012), that the posterior probability for model M_j, $\pi(M_j|\boldsymbol{y})$ is given by:

$$\pi(M_j|\boldsymbol{y}) = \frac{\pi_j(\boldsymbol{y}) \cdot \pi(M_j)}{\sum_{i=1}^{k} \pi_i(\boldsymbol{y}) \cdot \pi(M_i)}, \quad j = 1, \ldots, k$$

where:

$$\pi(\boldsymbol{y}) = \int_\theta \pi(\boldsymbol{y}|\boldsymbol{\theta})\pi(\boldsymbol{\theta})\delta\boldsymbol{\theta}, \quad \text{i.e. the marginal likelihood for the given model.}$$

The setup shown above is excellent and has a very nice interpretation of $\pi(M_j|\boldsymbol{y})$ as *the probability of model M_j being the best model conditional on data and prior belief*. However,

for complex models a serious matter occurs since this setup relies on the ability to calculate the marginal likelihoods.

Mostly, as is the case of the model presented in this study, computation of the marginal likelihood is not possible. Then so-called *penalized likelihood criteria* might be used. Examples are the *Akaike Information Criteria* (AIC) (Akaike, 1974), *Bayesian Information Criteria* (BIC) (Schwarz, 1978) and *Deviance Information Criteria* (DIC) (Spiegelhalter et al., 2002). The main principle for all these criteria is that they reward good model fit, but penalizes the model if to many parameters are included.

For Bayesian models based on MCMC-output DIC is commonly applied since it is easily computable. The DIC is used for the model selection in this study and is calculated as:

$$DIC = p_D + \overline{D}$$

where:

$$\overline{D} = \overline{D\left(\hat{\boldsymbol{\theta}}\right)}, \quad D\left(\overline{\theta}\right) = D\left(\overline{\hat{\boldsymbol{\theta}}}\right), \quad p_D = \overline{D} - D\left(\overline{\theta}\right)$$

$$D(\boldsymbol{\theta}) = -2\log(p(\boldsymbol{y}|\boldsymbol{\theta})): \quad \text{"The deviance"},$$

$$\hat{\boldsymbol{\theta}}: \quad \text{"The output"}$$

All elements above might be calculated from MCMC-output. A good model fit causes small values for $\overline{D}$, and $\overline{D}$ will decrease as more parameters are added to the model, at least if the models are nested. On the other hand p_D increases as number of parameters increase. Finally the model with lowest DIC-value should be selected.

References

AANDERAA, R., ROLSTAD, J. & SØGNEN, S. (1996). *Biological diversity in forests - knowledge for a sustainable management.* Oslo, Norway: The Norwegian Forest Owners' Federation and Landbruksforlaget.

AKAIKE, H. (1974). A new look at the statistical model identification. *Automatic Control, IEEE Transactions on* 19 716–723.

CARLIN, B. P. & LOUIS, T. A. (2011). *Bayesian methods for data analysis*, vol. 78. Chapman and Hall/CRC.

DELLAPORTAS, P. & ROBERTS, G. O. (2003). An introduction to MCMC. In J. Moller, ed., *Spatial Statistics and Computational Methods.*, vol. 173 of *Lecture Notes in Statistics.* Berlin: Springer, 1–41.

DIACONIS, P. & YLVISAKER, D. (1979). Conjugate priors for exponential families. *The Annals of statistics* 7 269–281.

DIREKTORATET FOR NATURFORVALTNING (1995). Forvaltning av hjortevilt mot år 2000. Handlingsplan. DN-rapport 1995-1, Direktoratet for naturforvaltning, Trondheim, Norway.

GANGSEI, L. E. (1999). *Er alle elgar like lette å observere? Evaluering av sett-elg metodikk ut frå samanlikning med modellpopulasjonar bygd på kohortanalyse.* Master's thesis, Norwegian University of Life Sciences (UMB), Ås, Norway.

GELFAND, A., HILLS, S., RACINE-POON, A. & SMITH, A. (1990). Illustration of bayesian inference in normal data models using Gibbs sampling. *Journal of the American Statistical Association* 85 972–985.

GELFAND, A. & SMITH, A. (1990). Sampling-based approaches to calculating marginal densities. *Journal of the American statistical association* 85 398–409.

GELMAN, A. (2008). Objections to bayesian statistics. *Bayesian Analysis* 3 445–449.

GELMAN, A., ROBERTS, G. & GILKS, W. (1996). Efficient metropolis jumping rules. *Bayesian statistics* 5 599–608.

GEMAN, S. & GEMAN, D. (1984). Stochastic relaxation, Gibbs distributions, and the Bayesian restoration of images. *Pattern Analysis and Machine Intelligence, IEEE Transactions on* 721–741.

Gilks, W. R., Richardson, S. & Spiegelhalter, D. J. (1996). *Markov Chain Monte Carlo in practice*. Chapman & Hall/CRC.

Gilks, W. R. & Wild, P. (1992). Adaptive rejection sampling for Gibbs sampling. *Applied Statistics* 337–348.

Hastings, W. K. (1970). Monte carlo sampling methods using Markov Chains and their applications. *Biometrika* 57 97–109.

Hjort, N. L. (2012). STK 4020: Bayesian Statistics - Autumn 2012 - Course Notes and Exercises. URL `www.uio.no/studier/emner/matnat/math/STK4020/h12/exercises_stk4020_2012c.pdf`.

Jaenson, T. G., Hjertqvist, M., Bergström, T. & Lundkvist, Å. (2012). Why is tick-borne encephalitis increasing? A review of the key factors causing the increasing incidence of human TBE in Sweden a. *Parasites & Vectors* 5 1–13.

Metropolis, N., Rosenbluth, A. W., Rosenbluth, M. N., Teller, A. H. & Teller, E. (1953). Equation of state calculations by fast computing machines. *The journal of chemical physics* 21 1087.

National deer register (2013). Sett elg–felt elg. URL `http://www.hjorteviltregisteret.no/`.

Pedersen, P. H. (2009). Modern wild cervid management with a new use of measures up to 2015. Proposals for a modern strategy for sustainable management of moose, red der, roe deer and wild reindeer. DN-rapport 2009-4, Direktoratet for Naturforvaltning, Trondheim, Norway.

Rolandsen, C., Solberg, E., Bjørneraas, K., Heim, M., Van Moorter, B., Herfindal, I., Garel, M., Pedersen, P., Sæther, B., Lykkja, O. & Os, Ø. (2010). Moose in Nord-Trøndelag, Bindal and Rissa 2005–2010. final report. NINA Rapport 588, Norwegian Institute for Nature Research (NINA), Trondheim, Norway.

Rolandsen, C. M., Solberg, E. J., Herfindal, I., Van Moorter, B. & Sæther, B.-E. (2011). Large-scale spatiotemporal variation in road mortality of moose: Is it all about population density? *Ecosphere* 2.

Schwarz, G. (1978). Estimating the dimension of a model. *The annals of statistics* 6 461–464.

SOLBERG, E., ROLANDSEN, C., HEIM, M., GRØTAN, V., GAREL, M., SÆTHER, B., NILSEN, E., AUSTRHEIM, G. & HERFINDAL, I. (2006). Elgen i norge sett med jegerøyne. En analyse av jaktmaterialet fra overvåkningsprogrammet for elg og det samlede sett elgmaterialet for perioden 1968-2004. NINA Rapport 125, Norwegian Institute for Nature Research (NINA), Trondheim, Norway.

SOLBERG, E. J., ROLANDSEN, C., HERFINDAL, I. & HEIM, M. (2009). Hjortevilt og trafikk i norge. NINA Rapport 463, Norwegian Institute for Nature Research (NINA), Trondheim, Norway.

SOLBERG, E. J. & SÆTHER, B.-E. (1999). Hunter observations of moose *Alces alces* as a management tool. *Wildlife Biology* 5 43–53.

SPIEGELHALTER, D. J., BEST, N. G., CARLIN, B. P. & VAN DER LINDE, A. (2002). Bayesian measures of model complexity and fit. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)* 64 583–639.

STATISTICS NORWAY (2013). table 03501. URL `http://statbank.ssb.no/statistikkbanken/`.

TOBIASSEN, C., BRØSETH, H., BAKKE, B. B., AARNES, S. G., HAGEN, S. B. & EIKEN, H. G. (2012). Populasjonsovervåking av brunbjørn 2009-2012: DNA-analyse av prøver samlet i Norge i 2011. Bioforsk report 57, Bioforsk, Svanhovd, Norway.

VÄLIMÄKI, P., MADSLIEN, K., MALMSTEN, J., HÄRKÖNEN, L., HÄRKÖNEN, S., KAITALA, A., KORTET, R., LAAKSONEN, S., MEHL, R., REDFORD, L. ET AL. (2010). Fennoscandian distribution of an important parasite of cervids, the deer ked (*Lipoptena cervi*), revisited. *Parasitology research* 107 117–125.

WABAKKEN, P., ARONSON, Å., STRØMSETH, T. H., SAND, H., MAARTMANN, E., SVENSSON, L., ÅKESSON, M., FLAGSTAD, Ø., LIBERG, O. & KOJOLA, I. (2011). Ulv i skandinavia: Statusrapport for vinteren 2010-2011. Oppdragsrapport 2011-1, Høgskolen i Hedmark, Evenstad, Norway.

Chapter 2

Paper

A Bayesian method for estimating moose (*Alces alces*) population size based on hunter observations and killed at age data.

Lars Erik Gangsei

Abstract: Lots of wild species, fish and mammals, are heavy harvested through fishing and hunting. Reliable population size estimates are valuable management tools for these species. In cases where killed at age data are available models outlined under the framework known as *cohort analysis* or *virtual population analysis* (VPA) are used extensively.

In this study a model using a Bayesian approach for estimating moose population size is examined. The model combines principles from *discrete time series analysis*, where basic cohort analysis based on killed at age data constitutes the bulk, and *analysis in continuous time* for each hunting season based on data from hunter observations.

The model performed very well when tested against simulated populations with known parameter values. The real data are gathered from the municipality of Ringerike in southern Norway. For these data tests are conducted through spatial cross-validation and temporal test-set validation. Generally the model performed well in these tests. However, a considerable issue is revealed by comparing results from different temporal subsets, since the hunters ability to spot moose seems to develop over time (years) and/or depend on moose density. This issue should not terminate the practical implementation of the model.

Keywords: Cohort analysis, moose, *Alces alces*, Bayesian, Markov Chain Monte Carlo (MCMC)

Introduction

The model described in this study consists of principles known from previous surveys, but the way they are assembled in a Bayesian framework is, to the authors knowledge, innovative. The main principle is cohort analysis which requires killed at age data. Auxiliary information from hunter observations is incorporated in the model.

Population estimates from cohort analyses are widely used on fish populations (Hilborn & Walters, 1992; Quinn & Deriso, 1999). Over the last two decades the use of Bayesian methods has increased both within the framework of cohort analysis (Quinn & Deriso, 1999) and for other methods in population ecology (King et al., 2009). Cohort analysis has also been used for terrestrial mammals, including moose and red deer (*Cervus elaphus*) in Norway, see for instance Ueno et al. (2009), Solberg & Sæther (1999) and Mysterud et al. (2007). Other articles of essential importance for outlining basic parts of the model presented in this study are Ricker (1940), Dupont (1983) and Deriso et al. (1985).

The regulation of moose density in Norway is predominantly carried out through hunting since the population sizes of natural moose predators, wolf (*Canis lupus*) and brown bear (*Ursus arctos*), are insufficient to regulate the moose density over large areas (Wabakken et al., 2011, Tobiassen et al., 2012). The main managers and policymakers are municipalities and landowners whose main instrument for controlling moose density development is the hunting quotas. The quotas should be set in a manner where the balance between costs and benefits from the moose population is assessed. In this process reliable estimates for moose density and production capacity would be valuable tools.

Preferably such estimates should be derived from hunter observations which are available for all parts of Norway at the website *www.hjorteviltregisteret.no* (National deer register, 2013) where the access is free. Hunter observations contain registrations of number of hunting days, and the number of observed and shot moose in different age-, sex- and reproductive categories. Hunter observations are currently widely utilized in practical moose management. The prevalence of killed at age data is, opposed to hunter observations, limited.

The Bayesian framework offers several benefits compared to alternative methods. One of them is the ability to utilize "all" commonly available data. Further the method elucidates estimate uncertainties, a possibility which is absent, or at least cumbersome, for the bulk of alternative methods. Finally the Bayesian framework, and its utility to handle missing data, forms a suitable basis in order to avail a slightly adjusted version of the model in areas lacking killed at age data.

Material and Methods

Data

The model is tested on data from Ringerike municipality in Norway. All statistics are performed using R 2.15.1 (R Core Team, 2012), including the packages "LearnBayes" (Albert, 2011) and "xtable" (Dahl, 2012). All data and code can be received from the author upon request.

Ringerike wildlife board is the origin for all data. Hunter observations are available in the period from 1984 till 2012 and killed at age data are available in the period from 1988 till 2012. The time span used for testing the model is limited to the period covered by the killed at age data, i.e. 1988 till 2012. The hunter observations for the period 1984 till 1995 are available as paper copies, and for the latter years via the National deer register. The killed at age data are received via the management agency Faun Naturforvaltning AS.

The ordinary hunting season in Norway starts on September 25th and ends on October 31th. For the years 1997 and 2008 till 2012 the hunting seasons were extended with an extra winter hunting season in December. Pooled data for both ordinary and winter hunting seasons are used for these years. Data from the National deer register contains pooled values for the years 2008 till 2012. For 1997 data for the winter hunting season from paper copies are added to the data from the National deer register.

Hunter observations are registered for each hunting field. Only hunting fields that have been administrated from the Ringerike wildlife board for the whole period in question is included in the study. Hunting fields are organized in so-called valds (Norwegian term). Each vald is an administrative unit containing at least one hunting field. When validating the model by spatial subsets the data are split according to the 22 valds present by 2011.

The age determination of moose shot during hunting seasons has been conducted by Åsmund Pålerud in the period from 1988 till 2000 and by Faun Naturforvaltning AS in the period from 2001 till 2012. All individuals registered as old animals (2 years or older) by hunters, and individuals classified as yearlings by hunters, but whose incisors by visually inspection are classified as older, have their age determined by counting the number of annuli in the cementum of their incisors (i.e. Rolandsen et al., 2008).

The killed at age data are not totally complete. In addition some of the individuals in the killed at age data are not counted on the hunting quota for the actual vald due to bad condition, inflammations, wounds etc., and thereby not registered as shot in the hunter observations. This causes some minor discrepancy between killed at age data and number of killed from hunter observations. In total 13740 moose are registered shot from

the hunter observations. The killed at age data contains data for 3994 calves and 9746 adult moose, i.e. 95.7% of the total hunting yield. For the adult moose 9559 (98.1%) are of known age and 187 of unknown age. These individuals are distributed according to the sex- and age distribution of adult moose with known age for the concerned valds and years.

The areas for the different Valds are obtained from an online map (Ringerike municipality, 2012). Statistics Norway registers the number of reported moose perishing from other causes than ordinary hunting for all municipalities in Norway (Statistics Norway, 2013). These registrations are not used directly in the model, but contributes in the validation process as independent benchmarks.

Model

Discrete Time Hazard Model

The model for population estimation presented in this study is based on the assumption of no net migration. In addition no assumptions on, or estimates of, reproductivity rate are made.

Let I denote the number of age-classes and let J denote the number of years in the analysis. Further let $\boldsymbol{N}_{I \times J}$ denote the $I \times J$ matrix of pre-harvest population sizes for all age-classes and across all years. Similarly, let $\boldsymbol{K}_{I \times J}$ denote the matrix of moose killed during hunting seasons, and $\boldsymbol{D}_{(I-1) \times (J-1)}$, denote the matrix of moose perishing between hunting-seasons from causes other than hunting. Both $\boldsymbol{N}$ and $\boldsymbol{D}$ are unobservable variables, but $\boldsymbol{K}$ is an observable variable containing the killed at age data.

All individuals with coinciding birth year belong to the same cohort. There are $L = I + J - 1$ cohorts present in the analysis. Let $l = 1$ denote the youngest cohort, i.e. the cohort with birth year J and let $l = L$ denote the oldest cohort, i.e. the cohort of age I in year 1. Each cohort is present in the analysis for a given number of years, M_l, where $M_l = min(l, L + 1 - l, J, I)$, for $1 \le l \le L$.

Some parts of the model are easier explained using a complementary notation to the "age-year" $(_{ij})$-notation. The complementary notation, "cohort-years present" $(_{lm})$ is marked by $*$. The relationship between the two notations is illustrated by equation (1) where elements for cohort $l = 3$, i.e. the cohort with birth-year $J - 2$, is <u>underlined</u>. The numeric relationship between the two notations is given by equation (2). In both equations, (1) and (2), $\boldsymbol{X}$ denotes any variable having an "age-year" structure.

$$\boldsymbol{X} = \begin{bmatrix} X_{11} & \cdots & \underline{X_{1(J-2)}} & X_{1(J-1)} & X_{1J} \\ \vdots & & & \underline{X_{2(J-1)}} & \vdots \\ \vdots & & & & \underline{X_{3J}} \\ \vdots & & & & \vdots \\ X_{I1} & \cdots & \cdots & \cdots & X_{IJ} \end{bmatrix},$$

$$\boldsymbol{X}^* = \begin{bmatrix} X_{J1}^* & \cdots & \underline{X_{31}^*} & X_{21}^* & X_{11}^* \\ \vdots & & & \underline{X_{32}^*} & \vdots \\ \vdots & & & & \underline{X_{33}^*} \\ \vdots & & & & \vdots \\ X_{L1}^* & \cdots & \cdots & \cdots & X_{I(min(I,J))}^* \end{bmatrix}$$

$$(1)$$

$$X_{lm}^* = X_{(max(0,l-J)+m)(max(0,J-l)+m)}, \quad l \le L, m \le M_l, \tag{2}$$

Under the assumption of no net migration for any cohort the deterministic relationship between pre-harvest population sizes ($\boldsymbol{N}$), number shot during the hunt ($\boldsymbol{K}$), and the number of natural dead individuals ($\boldsymbol{D}$) is given by equation (3).

$$N_{ij} = N_{(i+1)(j+1)} + K_{ij} + D_{ij}, \quad 1 \le i < I, \quad 1 \le j < J,$$

or equivalently: $\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad$ (3)

$$N_{lm}^* = N_{l(m+1)}^* + K_{lm}^* + D_{lm}^*, \quad 1 < l < L, \quad 1 \le m \le M_l - 1,$$

Equation (3) is the basis for all cohort analysis, and simply shows that the difference in cohort size from one year to another equals the number of deaths in the intermediate period. All natural mortality is assumed to take place between hunting seasons.

For this framework there are three different and disjunct possible outcomes for a random moose; (i) it might be shot, (ii) it might die of natural causes, or (iii) it might still be alive after the hunting season in year J. This framework fits into a *discrete time hazard model* (see Congdon, 2010, pp.435–441) for each cohort.

For all moose, conditional on being alive at the start of the time period in question, the probabilities of being shot, $\boldsymbol{p}_{I \times J}$, or suffering a natural death, $\boldsymbol{\nu}_{(I-1) \times (J-1)}$, are assumed equal and independent for all individuals from the same cohort. These assumptions lead to binomial distributions for the variables $\boldsymbol{K}$ and $\boldsymbol{D}$ as shown in equations (4) and (5).

$$K_{lm}^* \sim Bin\left(p_{lm}^*, \ N_{lm}^*\right), \quad 1 \le l \le L, \quad 1 \le m \le M_l, \tag{4}$$

$$D_{lm}^* \sim Bin\left(\nu_{lm}^*, \ N_{lm}^* - K_{lm}^*\right), \quad 1 < l < L, \quad 1 \le m < M_l, \tag{5}$$

Let p_{lm}^{c*} and ν_{lm}^{c*} denote the probabilities for an individual from cohort l of being shot or die a natural death in the m-th year the cohort is present in the analysis. Then p_{lm}^{c*} and ν_{lm}^{c*} are given by equations (6) and (7), equations derived by minor expansions of the equation given in Congdon (2010, p.436).

$$p_{lm}^{c*} = p_{lm}^* \prod_{n=1}^{m-1} (1 - p_{ln}^*)(1 - \nu_{ln}^*), \quad l \le L, m \le M_l \tag{6}$$

$$\nu_{lm}^{c*} = \nu_{lm}^*(1 - p_{lm}^*) \prod_{n=1}^{m-1} (1 - p_{ln}^*)(1 - \nu_{ln}^*), \quad 1 < l < L, \quad m < M_l \tag{7}$$

The cumulative probabilities for a moose from cohort l of being shot and of suffering a natural death, denoted p_l^{CC} and ν_l^{CC} respectively, are given in equation (8). Cohorts reaching an age larger than I before year J, i.e. cohorts where $l > I$, reach the age I in the year $L - l + 1$. Notice that for these cohorts elements from $[p_{I(L-l+2)}, \ldots, p_{IJ}]$ and $[\nu_{(I-1)(L-l+1)}, \ldots, \nu_{(I-1)(J-1)}]$, i.e. elements from the oldest age-class in succeeding cohorts, are included in the cumulative probabilities. These inclusions are justified by the fact that if an individual should reach an age older than I it would still be exposed to both hunting and natural mortality.

$$p_l^{CC} = \sum_{m=1}^{M_l} p_{lm}^{c*} \ + \ I(l) \cdot \sum_{j=L-l+2}^{J} p_{Ij}^c, \quad 1 \le l \le L,$$

$$\nu_l^{CC} = \sum_{m} \nu_{lm}^{c*} \ + \ I(l) \cdot \sum_{j=L-l+1}^{J-1} \nu_{(I-1)j}^c, \quad 1 < l < L, \tag{8}$$

where:

$$I(l) = \begin{cases} 1 \text{ if } l > I \\ 0 \text{ else} \end{cases}$$

The initial population size for cohort l, N_{l1}^*, consist of (i) shot moose, $K_l^{CC} = \sum_m K_{lm}^*$ (known number), (ii) natural dead moose, $D_l^{CC} = \sum_m D_{lm}^*$ (unknown number), and (iii)

"survivors", i.e. moose still alive post hunt in year J, denoted N_l^s (unknown number). Then obviously $N_{l1}^* = K_l^{CC} + D_l^{CC} + N_l^s$ holds for all cohorts. Further these assumptions lead to a negative binomial distribution for the sum of natural dead moose and survivors if $K_l^{CC} > 0$ and a geometric distribution if $K_l^{CC} = 0$ as given in equation (9).

$$(D_l^{CC} + N_l^s) \sim \begin{cases} Neg.Bin \left(p_l^{CC}, \ K_l^{CC} \right) & , \text{ if } K_l^{CC} > 0 \\ Geom \left(p_l^{CC} \right) & , \text{ if } K_l^{CC} = 0 \end{cases}, \quad 1 \le l \le L \tag{9}$$

Let $\boldsymbol{d}_l^+$ denote a vector consisting of the numbers of natural dead moose and survivors from cohort l, i.e. $\boldsymbol{d}_l^+ = [D_{l1}^*, \ldots, D_{l(M_l-1)}^*, N_l^s]^t$. Further let $\boldsymbol{\nu}_l^{C+}$ denote the probabilities associated with $\boldsymbol{d}_l^+$, i.e. $\boldsymbol{\nu}_l^{C+} = [\nu_{l1}^{c*}, \ldots, \nu_{l(M_l-1)}^{c*}, 1 - p_l^{CC} - \nu_l^{CC}]^t$. Under these assumptions $\boldsymbol{d}_l^+$ is a multinomal distributed variable as shown in equation (10).

$$\boldsymbol{d}_l^+ \sim Multinom \left(\boldsymbol{\nu}_l^{C+}, \ N_{l1}^* - K_l^{CC} \right), \quad 1 \le l \le L \tag{10}$$

From (3) it is obvious that if $\boldsymbol{N}$ is known, so is $\boldsymbol{D}$ and vice versa. Consequently if $\boldsymbol{d}_l^+$ is known so is N_{lm} for all m.

Analysis in Continuous Time

During any random hunting season, j, the cumulative probability of a random individual of age i of being shot by time t, $p_{ij}(t)$, is modeled in continuous time, t ($0 \le t \le h_j$), where h_j is the (known) hunting effort in year j, measured as number of hunting days per km^2 hunting area.

The instantaneous probability for a random hunter of observing a moose still alive at time t is assumed independent of age and year, denoted α, and referred to as *instantaneous observation rate*. A practical interpretation of α is *the expected number of moose observed by a random hunter on a random day if the moose density is* 1 *moose per* km^2.

Further, the probability of a moose of age i in year j being shot, conditional on being observed, is given by $\gamma_j \lambda_i$. Here $\boldsymbol{\gamma}_J$, referred to as *year-specific hunting mortality*, is a vector of year dependent factors, and $\boldsymbol{\lambda}_I$, referred to as *age-specific hunting mortality*, is a vector of age dependent factors.

These assumptions lead to a constant hazard function, denoted $g_{ij}(t)$, given in equation (11), for a moose of age i in year j.

$$g_{ij}(t) = \alpha \gamma_j \lambda_i, \quad 0 \le t \le h_j, \quad 1 \le i \le I, \quad 1 \le j \le J \tag{11}$$

Since the hazard function $g_{ij}(t)$ is independent of t an expression for $p_{ij} = p_{ij}(t = h_j)$ is easily derived from (11) (see Congdon, 2010, pp.414–418) and is given by equation (12). Further, due to assumed independence between individuals of being killed during the hunting season, the hunting yield, $\boldsymbol{K}$, is binomial distributed as shown in equation (4).

$$p_{ij} = 1 - \exp\left(-\alpha\gamma_j\lambda_i \cdot h_j\right), \quad 1 \leq i \leq I, \quad 1 \leq j \leq J \tag{12}$$

The model described in equations (11), (12) and (4) might also be written in an alternative way, involving augmented data (Tanner & Wong, 1987). This alternative, given in equations (14) and (15), might look cumbersome. However, it enables conjugate prior distributions for all parameters under the Bayesian inference. Then Gibbs-sampling (Gelfand et al., 1990; Gelfand & Smith, 1990; Geman & Geman, 1984) might be used for all parameters in the estimation algorithm. Dellaportas & Roberts (2003) describes attainment of simple conjugate priors as one of two main justifications for using data augmentation, the other one being handling of missing data.

Let the unobservable variable $\boldsymbol{S}^1_{I \times J}$ denote the number of observed moose for all age-classes across all years. The yearly sums of $\boldsymbol{S}^1$, i.e. the total number of observed moose across all age-classes, $\boldsymbol{s}_J$, is observed every year. The change in moose density during the hunting season is modelled to be proportional to the actual moose density at the time. Consequently the average number of moose present during the hunting season across all years and age-classes, $\bar{\boldsymbol{N}}_{I \times J}$, is given by equation (13).

$$\bar{N}_{ij} = \begin{cases} -K_{ij}\left(\log\left(1 - \frac{K_{ij}}{N_{ij}}\right)\right)^{-1} & , \text{ if } 0 < K_{ij} < N_{ij} \\ N_{ij} & , \text{ if } K_{ij} = 0 \text{ or } K_{ij} = N_{ij} \end{cases}, \quad 1 \leq i \leq I, \ 1 \leq j \leq J \tag{13}$$

As stated earlier the instantaneous observation rate, α, is modelled to be constant. Under these assumptions the numbers of observed moose, given in $\boldsymbol{S}^1$, are results of non-homogeneous poisson processes as given in equation (14).

$$\begin{aligned} S^1_{ij} &\sim Pois\left(\alpha h_j \bar{N}_{ij}\right), \quad 1 \leq i \leq I, \quad 1 \leq j \leq J, \\ &\Rightarrow s_j \sim Pois\left(\alpha h_j \sum_{i=1}^{I} \bar{N}_{ij}\right), \quad 1 \leq j \leq J \end{aligned} \tag{14}$$

A fraction, $\boldsymbol{S}^2_{I \times J}$, of the observed moose, $\boldsymbol{S}^1$, is considered shot by the hunters. The probability of a random observed moose of being considered shot is assumed to be equal

for moose from all age-classes during each hunting season. These probabilities are given by the year-specific hunting mortality, $\boldsymbol{\gamma}$. Consequently $\boldsymbol{S^2}$ is a binomial distributed variable as given in equation (15).

Likewise a fraction, $\boldsymbol{K}$, of $\boldsymbol{S^2}$ is actually shot. The age-specific hunting mortalities, $\boldsymbol{\lambda}$, are assumed to be constant, giving a binomial distribution for $\boldsymbol{K}$ also shown in equation (15). The last equation shown in (15) is a simple result of the first two equations, and has no practical application, but offers a somewhat more intuitive interpretation of the parameters $\boldsymbol{\lambda}$ and $\boldsymbol{\gamma}$, since their product might be viewed as the probability of a moose being shot conditional on being observed.

$$
\begin{aligned}
S_{ij}^2 &\sim Bin\left(\gamma_j,\ S_{ij}^1\right), & i \leq I, & \quad j \leq J, \\
K_{ij} &\sim Bin\left(\lambda_i,\ S_{ij}^2\right), & i \leq I, & \quad j \leq J, \\
\Rightarrow K_{ij} &\sim Bin\left(\lambda_i\gamma_j,\ S_{ij}^1\right), & i \leq I, & \quad j \leq J
\end{aligned}
\tag{15}
$$

Posterior distribution

A Markov Chain Monte Carlo (MCMC) approach is applied in order to estimate posterior distributions for the unknown parameters, $\boldsymbol{\theta} = \{\boldsymbol{\lambda}, \boldsymbol{\gamma}, \alpha, \boldsymbol{\nu}\}$, and to predict the missing data, $\boldsymbol{Y}_{mis} = \{\boldsymbol{N}, \boldsymbol{D}, \boldsymbol{S}^1, \boldsymbol{S}^2\}$. A general introduction to MCMC can be found in Dellaportas & Roberts (2003). The MCMC algorithm enables simulations from the joint posterior distribution for the unknown parameters, $\boldsymbol{\theta}$, and missing data, $\boldsymbol{Y}_{mis}$, conditional on the observed data, $\boldsymbol{Y}_{obs} = \{\boldsymbol{K}, \boldsymbol{s}\}$ and the prior distributions, $\pi(\boldsymbol{\theta}, \boldsymbol{Y}_{mis}|\boldsymbol{\Phi})$, where $\boldsymbol{\Phi} = \{\phi_2^\lambda, \phi_2^\gamma, \phi_2^\alpha, \phi_2^\nu\}$ denotes the prior hyperparameters. The hunting effort, $\boldsymbol{h}$, is viewed as a fixed and observed covariate. The general form of the joint distribution of $\boldsymbol{\theta}$ and $\boldsymbol{Y}_{mis}$ conditional on $\boldsymbol{Y}_{obs}$ and $\boldsymbol{\Phi}$, is given in equation (16).

$$
\begin{aligned}
\pi\left(\boldsymbol{\theta}, \boldsymbol{Y}_{mis}|\boldsymbol{Y}_{obs}, \boldsymbol{\Phi}, \boldsymbol{h}\right) &= \frac{\pi\left(\boldsymbol{Y}_{obs}|\boldsymbol{\theta}, \boldsymbol{Y}_{mis}, \boldsymbol{h}\right) \cdot \pi\left(\boldsymbol{\theta}, \boldsymbol{Y}_{mis}|\boldsymbol{\Phi}\right)}{\pi\left(\boldsymbol{Y}_{obs}, \boldsymbol{\Phi}\right)} \\
&\propto \pi\left(\boldsymbol{Y}_{obs}|\boldsymbol{\theta}, \boldsymbol{Y}_{mis}, \boldsymbol{h}\right) \cdot \pi\left(\boldsymbol{\theta}, \boldsymbol{Y}_{mis}|\boldsymbol{\Phi}\right)
\end{aligned}
\tag{16}
$$

In equation (16) $\pi(\boldsymbol{Y}_{obs}|\boldsymbol{\theta}, \boldsymbol{Y}_{mis}, \boldsymbol{h})$ is known as the likelihood function. The likelihood function might be rewritten as equation (17). The main idea in MCMC sampling is to update the elements of $\boldsymbol{\theta}$ and $\boldsymbol{Y}_{mis}$ in blocks or one at a time by sampling values from the conditional posterior distribution assuming all other elements of $\boldsymbol{\theta}$ and $\boldsymbol{Y}_{mis}$ to be known, see Gilks et al. (1996, pp.75–79). Expressions for these full conditional posterior

distributions will be shown in the following sections for all elements in $\boldsymbol{\theta}$ and $\boldsymbol{Y}_{mis}$.

$$\pi\left(\boldsymbol{Y}_{obs}|\boldsymbol{\theta},\boldsymbol{Y}_{mis},\boldsymbol{h}\right) = f_4\left(\boldsymbol{K}|\boldsymbol{\lambda},\boldsymbol{\gamma},\alpha,\boldsymbol{N},\boldsymbol{h}\right) \cdot f_{14}\left(\boldsymbol{s}|\boldsymbol{K},\alpha,\boldsymbol{N},\boldsymbol{h}\right) \tag{17}$$

The prior distributions, $\pi\left(\boldsymbol{\theta},\boldsymbol{Y}_{mis}|\boldsymbol{\Phi}\right)$, are given in equation (18). The hyperparameters, $\boldsymbol{\Phi}$, are quite uninformative for α ($\boldsymbol{\phi}^{\alpha} = [10^{-4}, 10^{-4}]^t$). For $\boldsymbol{\gamma}$ and $\boldsymbol{\lambda}$ hyperparameters with some vague prior information are chosen ($\boldsymbol{\phi}^{\lambda} = \boldsymbol{\phi}^{\gamma} = [5,5]^t$), whereas the hyperparameters for $\boldsymbol{\nu}$ are informative with prior median at 0.05 and prior 99%-percentile at 0.15, i.e. $\boldsymbol{\phi}^{\nu} = [2.96, 50.41]^t$.

$$
\begin{aligned}
\lambda_i &\sim Beta\left(\phi_1^{\lambda},\ \phi_2^{\lambda}\right), & 1 \leq i \leq I, \\
\gamma_j &\sim Beta\left(\phi_1^{\gamma},\ \phi_2^{\gamma}\right), & 1 \leq j \leq J, \\
\nu_j &\sim Beta\left(\phi_1^{\nu},\ \phi_2^{\nu}\right), & 1 \leq j < J, \\
\alpha &\sim Gamma\left(\phi_1^{\alpha},\ \phi_2^{\alpha}\right) \\
S_{ij}^1 &\propto 1, & 1 \leq i \leq I \quad 1 \leq j \leq J, \\
S_{ij}^2 &\propto 1, & 1 \leq i \leq I \quad 1 \leq j \leq J,
\end{aligned}
\tag{18}
$$

Data augmentation

The missing data, $\boldsymbol{Y}_{mis}$, is updated through data augmentation where the conditional posterior distributions for the elements $\boldsymbol{S}^1$ and $\boldsymbol{S}^2$ are categorical distributed as shown in equation (19). For simplicity uniform prior distributions are used for $\boldsymbol{S}^1$ and $\boldsymbol{S}^2$, as shown in equation (18). Since $\boldsymbol{S}^1$ and $\boldsymbol{S}^2$ both are categorical distributed variables with finite sample spaces as shown in equation (19) the prior distributions as given in equation (18) are proper probability distributions.

The number of observed moose, $\boldsymbol{S}^1$, is part of equations (14) and (15). The lower and upper bounds for $\boldsymbol{S}^1$ are provided by $\boldsymbol{S}^2$ and $\boldsymbol{s}$ respectively. In equation (19) $\tilde{\boldsymbol{s}}_{ij}^1$ is the possible outcomes for S_{ij}^1 conditional on S_{ij}^2 and s_j with the associated probabilities $\tilde{\boldsymbol{q}}_{ij}^1$.

Likewise the number of moose considerd shot, $\boldsymbol{S}^2$, is part of equation (15) "two times". Obviously every element of $\boldsymbol{S}^2$ has an upper limit given by the elements of $\boldsymbol{S}^1$ and a lower limit given by the elements of $\boldsymbol{K}$. In equation (19) $\tilde{\boldsymbol{s}}_{ij}^2$ is the possible outcomes for S_{ij}^2 conditional on K_{ij} and S_{ij}^1 with the associated probabilities $\tilde{\boldsymbol{q}}_{ij}^2$. Details for equation (19) are given in Appendix 2.

$$S_{ij}^1 \sim Cat\left(\tilde{\boldsymbol{s}}_{ij}^1, \ \tilde{\boldsymbol{q}}_{ij}^1\right), \quad 1 \le i \le I, \quad 1 \le j \le J,$$
$$S_{ij}^2 \sim Cat\left(\tilde{\boldsymbol{s}}_{ij}^2, \ \tilde{\boldsymbol{q}}_{ij}^2\right), \quad 1 \le i \le I, \quad 1 \le j \le J, \tag{19}$$

Parameter estimation

Generally the main goal for the MCMC algorithm is to estimate the posterior distribution for $\boldsymbol{\theta}$. In this study reliable estimates for the population size, $\boldsymbol{N}$, is of even larger importance.

Gibbs-sampling is used for estimating all parameters, $\boldsymbol{\theta}$. The Gibbs-sampling procedure is simplified by implementation of conjugate prior distributions for all parameters in question. Most of them are vectors of probability parameters for binomial distributed variables, see equations (15) and (5). The conjugate prior distributions for these parameters are beta distributions as given in equation (18). The full conditional posterior distributions for the parameters $\boldsymbol{\lambda}$, $\boldsymbol{\gamma}$ and $\boldsymbol{\nu}$ are new beta distributions given in equation (20).

$$\lambda_i \sim Beta\left(\phi_1^\lambda + \sum_{j=1}^J K_{ij}, \ \phi_2^\lambda + \sum_{j=1}^J \left(S_{ij}^2 - K_{ij}\right)\right), \quad 1 \le i < I_2,$$

$$\lambda_{I_2} \sim Beta\left(\phi_1^\lambda + \sum_{j=1}^J \sum_{i=I_2}^I K_{ij}, \ \phi_2^\lambda + \sum_{j=1}^J \sum_{i=I_2}^I \left(S_{ij}^2 - K_{ij}\right)\right),$$

$$\gamma_j \sim Beta\left(\phi_1^\gamma + \sum_{i=1}^I S_{ij}^2, \ \phi_2^\gamma + \sum_{i=1}^I \left(S_{ij}^1 - S_{ij}^2\right)\right), \quad 1 \le j \le J,$$

$$\nu_j \sim Beta\left(\phi_1^\nu + \sum_{i=1}^{I-1} D_{ij}, \ \phi_2^\nu + \sum_{i=1}^{I-1} \left(N_{ij} - K_{ij} - D_{ij}\right)\right), \quad 1 \le j < J \tag{20}$$

It is also possible to simulate values for α using Gibbs-sampling. The full conditional posterior distribution for α is a gamma distribution given by equation (21). Further details are given in Appendix 2.

$$\alpha \sim Gamma\left(\phi_1^\alpha + \sum_{j=1}^J s_j, \ \phi_2^\alpha + \sum_{i=1}^I \sum_{j=1}^J \left(\bar{N}_{ij} h_j\right)\right) \tag{21}$$

Several models with different parameter restrictions are applied to the data. For all models the length of $\boldsymbol{\lambda}$ is restricted to $I_2 = 3$. That is all individuals of age corresponding to I_2 or older are assumed to have equal age-specific hunting mortality. Further all age-

classes are assumed to have equal probability off suffering a natural death, that is $\nu_{ij} = \nu_j$ for all $i \leq I$. Additional simplifications are implemented by running the $2^3 = 8$ model combinations with assumptions about equal values for all elements in $\boldsymbol{\lambda}$ (i.e. $I_2 = 1$), $\boldsymbol{\gamma}$ and $\boldsymbol{\nu}$.

The conditional posterior distributions for $\boldsymbol{\gamma}$, $\boldsymbol{\lambda}$ and $\boldsymbol{\nu}$ are given by equation (20). In the situations where all elements are assumed equal for $\boldsymbol{\gamma}$ and $\boldsymbol{\nu}$ the parameters given in equation (20) are calculated by summarizing across both year and age as demonstrated for λ_{I_2} in equation (20).

Estimation algorithm

The estimation algorithm is an iterative algorithm consisting of data augmentation (step 4) and parameter estimation (step 5). Notice that equations (9) and (10) enables sampling for all $\boldsymbol{N}_{l1}^{(w)}$ and $\boldsymbol{D}^{(w)}$ directly without any need of prior distributions for these missing data.

1. Initiate $\boldsymbol{\theta}^0$

2. Set initial value for $\boldsymbol{N}^{(0)}$ and $\boldsymbol{S^2}^{(0)}$ consistent with (3) and (19),

3. Set the iteration number $w = 1$,

4. Data-augmentation,

   (a) for $\boldsymbol{S}^{1(w)}$ by equation (19),

   (b) for $\boldsymbol{S}^{2(w)}$ by equation (19),

   (c) for $N_{l1}^{(w)*}$ for $1 \leq l \leq L$ by equation (9),

   (d) for $\boldsymbol{D}^{(w)}$ by equation (10),

   (e) deterministic calculation of $\boldsymbol{N}^{(w)}$ by equation (3),

5. Parameter estimation by Gibbs-sampling,

   (a) for $\boldsymbol{\lambda}^{(w)}$ by equation (20),

   (b) for $\boldsymbol{\gamma}^{(w)}$ by equation (20),

   (c) for $\boldsymbol{\nu}^{(w)}$ by equation (20),

   (d) for $\alpha^{(w)}$ by equation (21),

6. Increment w by one,

7. Repeat steps 4 till 6 for $W = n_{burn-in} + n_{sim}$ iterations.

8. The parameter estimates are obtained as the posterior means computed as the Monte Carlo averages from the n_{sim} values sampled after burn-in.

Simultaneous estimation of female- and male populations

Moose observed by hunters are separated in three different classes; calves, females and males. These classes are indexed ^{C}, ^{F} and ^{M} respectively.

Calves is the youngest age group, i.e. for calves $i = 1$. The total number of calves from both sexes observed each year, $\boldsymbol{s}^C$, is an observable variable. Since the vast majority of calves are observed together with their mother, calves of both sexes and adult females are assumed to have equal instantaneous observation rate, α^F. Consequently the observed, but not shot, number of calves of a given sex, might be viewed as a binomial distributed variable as shown in equation (22), where the parametrization is set up for female calves.

$$\left(S_{1j}^{1F} - K_{1j}^F\right) \sim Bin\left(\bar{N}_{1j}^F/\bar{N}_{1j}^M, \ S_j^C - K_{1j}^F - K_{1j}^M\right), \quad 1 \le j \le J, \tag{22}$$

The probability of a calf being shot conditional on being observed, $\boldsymbol{\gamma}^C$, might vary between years. Since the number of observed calves is observed directly, an age-specific hunting mortality parameter for calves is redundant. Therefore the first elements of the $\boldsymbol{\lambda}$- vectors, λ_1^F and λ_1^M, belongs to the yearlings whose age-class is $i = 2$. Then, as a consequence of equation (15), the number of calves considered shot equals the number of shot calves, i.e. $S_{1j}^2 = K_{1j}$ for all j (years) and both sexes.

Since calves and females are assumed to have equal instantaneous observation rate, α^F, the conditional posterior distributions for α^F and α^M are given by equation (23).

$$\begin{aligned} \alpha^F &\sim Gamma\left(\phi_1^\alpha + \sum_{j=1}^J \left(s_j^F + s_j^C\right), \ \phi_2^\alpha + \sum_{i=1}^I \sum_{j=1}^J \left(\bar{N}_{ij}^F h_j\right) + \sum_{j=1}^J \left(\bar{N}_{1j}^M h_j\right)\right) \\ \alpha^M &\sim Gamma\left(\phi_1^\alpha + \sum_{j=1}^J s_j^M, \ \phi_2^\alpha + \sum_{i=2}^I \sum_{j=1}^J \left(\bar{N}_{ij}^M h_j\right)\right) \end{aligned} \tag{23}$$

The age of the oldest individual in the killed at age data differs between sexes, thus they are given separate values for I. Values for $\boldsymbol{\lambda^F}$ and $\boldsymbol{\lambda^M}$, and, $\boldsymbol{\gamma^F}$ and $\boldsymbol{\gamma^M}$, are estimated on the basis of age-classes older than calves and separately for each sex, and for calves $\boldsymbol{\gamma^C}$ is estimated on the basis of pooled values from both sexes. Likewise, the natural mortality

rate, ν, is assumed to be equal for the two sexes, and therefore estimated on the basis of pooled values across age- and sex-classes.

Model validation

Four different methods are applied for model validation; (i) calculating Deviance Information Criterion (DIC) (Spiegelhalter et al., 2002), (ii) testing the model's ability to estimate known parameters from simulated model populations, (iii) splitting the data into different spatial subsets and evaluating the results by cross-validation, and (iv) splitting the data into different temporal subsets and compare parameter estimates. Methods (ii), (iii) and (iv) are described in the following sections.

Model selection, i.e. choosing the number of elements included in the parameters $\boldsymbol{\lambda}$, $\boldsymbol{\gamma}$ and $\boldsymbol{\nu}$, is primarily based on DIC-values for the real data set and the simulated populations.

Test variables

For the simulated populations and the cross-validation, different test variables are calculated. For the simulated population parameters estimates are compared with known values. In principle the same approach is used for the cross-validation in spatial subsets where predicted values for observed moose, $\hat{\boldsymbol{s}}$, are compared with known observations, $\boldsymbol{s}$.

In general let n_{sim} denote the number of simulations after the initial burn-in period of the MCMC-algorithm. Further let $\hat{\boldsymbol{\theta}}^x_{n_{sim}}$ denote the simulated results for parameter x where x might be any element from $\boldsymbol{\theta}$, $\boldsymbol{N}$ or $\boldsymbol{s}$.

Three different test variables are calculated. The test variables are (i) average estimation error, ϵ^x_μ, (ii) posterior variance, ϵ^x_σ, and (iii) Monte Carlo p-value, ϵ^x_p. Equation (24) shows the formulas for the test parameters in question.

The Monte Carlo p-value, ϵ^x_p, is the p-value for the hypothesis $\theta^x = \theta^x_0$, where θ^x_0 is the true/observed value, tested on basis of $\hat{\boldsymbol{\theta}}^x$. If the estimate, $\hat{\theta}^x$, is unbiased, ϵ^x_p is approximately uniform distributed between 0 and 1 (Besag & Clifford, 1991). The approximation is a result of influence from the prior distributions. Further, if the approximation is ignored, Besag & Clifford (1991) states that for a sample of m different ϵ^x_p-s the test-variable $-2 \times \sum_m \log(\epsilon^x_{p(m)})$ will be χ^2_{2m}-distributed under the null hypothesis that $\epsilon^x_p \sim \text{Uniform}(0, 1)$.

$$\epsilon_\mu^x = \overline{\hat{\theta}^x - \theta_0^x},$$

$$\epsilon_\sigma^x = var(\hat{\theta}^x - \theta_0^x) = var(\hat{\theta}^x),$$

$$\epsilon_p^x = 2 \times min \left(0.5, \ \frac{\sum_n I\left(\hat{\theta}_n^x \geq \theta_0^x\right) + 1}{n_{sim} + 1}, \ \frac{\sum_n I\left(\hat{\theta}_n^x \leq \theta_0^x\right) + 1}{n_{sim} + 1} \right) \quad (24)$$

where:

$$I(\text{argument}) = \begin{cases} 1 \text{ if argument is true} \\ 0 \text{ if argument is false} \end{cases}$$

Simulated population

In the following description sex indexes are omitted. The model is tested by applying the estimation algorithm to two simulated populations with known parameter values. The parameters, $\boldsymbol{\theta}$, the initial adult population, $[N_{21}, \ldots, N_{I1}]^t$, and the number of calves, $[N_{11}, \ldots, N_{1J}]$ for the full period, used in the simulation process, are set equal to the posterior means from real data based on the "full model" and the model with reduced $\boldsymbol{\nu}$, respectively. The hunting effort, $\boldsymbol{h}$, are taken directly from the real data.

Next the probabilities of being shot, $\boldsymbol{p}$, are calculated by equation (12). Then the initial hunting yield, $[K_{11}, \ldots, K_{I1}]^t$, and the initial number perishing from natural causes, $[D_{11}, \ldots, D_{(I-1)1}]^t$, is simulated in sequence by (4) and (5). These numbers form a basis for calculating pre-harvest population size for adults in year 2 by equation (3). This procedure is repeated for the following years resulting in simulations for the observable variable $\boldsymbol{K}$ and the unobservable variables $\boldsymbol{N}$ and $\boldsymbol{D}$. Finally the number of observed moose, $\boldsymbol{s}$ is simulated based on (14). The final result is realisations for $\boldsymbol{Y}_{obs}$, $\boldsymbol{\theta}$, and the unobserved variables of interest, $\boldsymbol{N}$ and $\boldsymbol{D}$.

Spatial and temporal subsets

The implementation of spatial- and temporal subsets is restricted to the model with parameter combinations possessing the lowest DIC-value, which turns out to be the "full model".

In a "leave one out" strategy simulations are conducted on subsets leaving one vald out

at a time. By 2011 22 different valds were present. Let the index $^{-k}$ denote the subset consisting of all valds except vald k, $k = 1, \ldots, 22$. Correspondingly the index $^{+k}$ denote the subset consisting of data solely from vald k.

Model validation done on basis of vald subsets relies on the assumption that $\boldsymbol{\theta}^{+k} = \boldsymbol{\theta}^{-k}$. For all subsets leaving one vald out at a time all parameters are simulated and named $\hat{\boldsymbol{\theta}}^{-k}_{n_{sim} \times dim(\theta)}$. Thereafter the number of observed moose in all subsets is estimated and stored in $\hat{\boldsymbol{S}}^{+k}_{n_{sim} \times J}$ by applying the following algorithm, where the sex-/age indexes are omitted for simplicity.

1. Set $n=1$,

2. $\boldsymbol{p}^{+k}$ is calculated using equation (12) and setting $\boldsymbol{\theta}^{+k} = \hat{\boldsymbol{\theta}}^{-k}_n$,

3. a value for $\boldsymbol{N}^{+k}$ is drawn using equations (9), (10) and (3),

4. a value for $\boldsymbol{S}^{1(+y)}$ is drawn using equation (14), and a value for $\boldsymbol{s}^{+k}$ is calculated,

5. the result is stored by setting $\hat{S}^{+k}_n = \boldsymbol{s}^{+k}$,

6. increase n by 1,

7. step 2 till 6 are repeated until $n = n_{sim}$.

The result of this algorithm, $\hat{\boldsymbol{S}}^{+k}$, is a function of $\boldsymbol{Y}^{(+y)-\boldsymbol{s}}_{obs}$ and $\hat{\boldsymbol{\theta}}^{-k}$, i.e. $\hat{\boldsymbol{S}}^{+k}$ and $\boldsymbol{s}^{+k}$ are independent. The test variables presented in equation (24) are calculated by substituting $\hat{\boldsymbol{\theta}}$ with $\hat{\boldsymbol{S}}$ and $\boldsymbol{\theta}$ with $\boldsymbol{s}$.

For this situation with observed values for $\boldsymbol{s}$ an additional Monte Carlo inspired method are applied to test the strength of the model. There are 66 different $\boldsymbol{s}^{+k}$-vectors (3 sex-/age-classes from 22 valds), each of length $J = 25$. For every $\boldsymbol{s}^{+k}$ the sequence of elements are randomly reshuffled and then new values for ϵ^s_p are calculated based these reshuffled vectors and $\hat{\boldsymbol{S}}$ for the corresponding subsets with columns in their original sequence. This reshuffling procedure is conducted for 10^4 simulations. For every simulation an average value for ϵ^s_p, denoted $\overline{\epsilon^s_p}$, is calculated for each sex-/age-class. Consequently every $\overline{\epsilon^s_p}$ is based on 550 separate ϵ^s_p values (25 years for 22 valds). The values for $\overline{\epsilon^s_p}$ are stored in vectors of length 10^4. For all sex-/age-classes the Monte Carlo distributions for $\overline{\epsilon^s_p}$ based on reshuffled elements are compared with observed $\overline{\epsilon^s_p}$. A distinct difference between observed $\overline{\epsilon^s_p}$ and its Monte Carlo distributions based on reshuffled elements indicates a good model fit.

The data from Ringerike spans over the period 1988 till 2012, totally 25 years of data. Of course $\boldsymbol{\theta}$ might be estimated using any consecutive time period. The full data set is divided into 14 different subsets, each of them with 12 subsequent years of data. The validation based on time subsets is conducted by evaluating potential trend in estimates for the time-independent parameters α and $\boldsymbol{\lambda}$.

Results

For all model runs the number of iterations in the estimation algorithm are 10^4 of which 2500 are used as burn-in period ($n_{sim} = 7500$). All model runs with real and simulated data mixed well, as the trace plot in Figure 1 indicates.
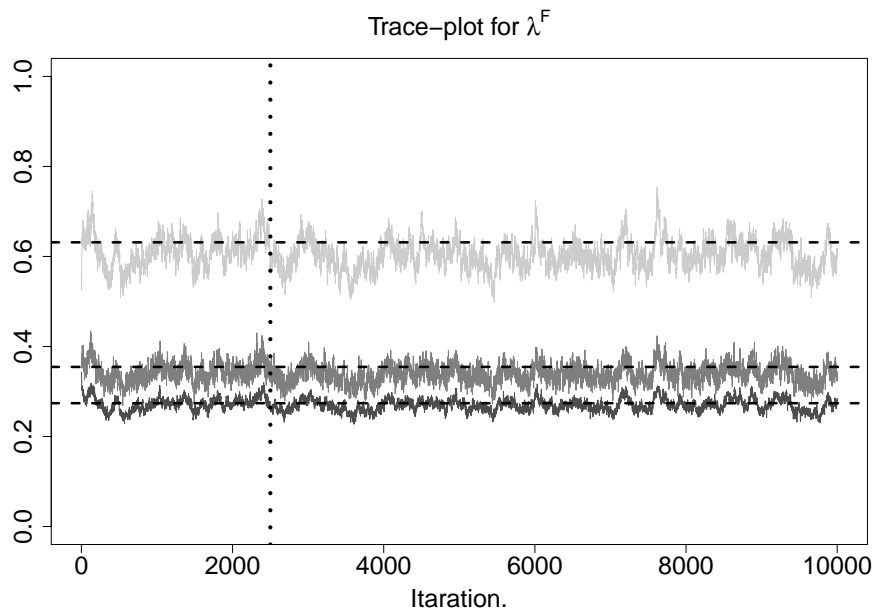
Figure 1: Trace plot for the 3 elements in the parameter $\boldsymbol{\lambda}^F$ estimated by "full model" and based on the simulated data with "full $\boldsymbol{\nu}$". Real values shown with horizontal dashed lines, and limit for burn-in period shown with vertical dotted line.

Model selection

As shown in Table 1 the parameter-combinations with full $\boldsymbol{\gamma}$ generally stand out as preferable compared to the combinations with reduced $\boldsymbol{\gamma}$. For the real data and the simulated data where the elements of $\boldsymbol{\nu}$ vary over years, the "full model" has lowest DIC-value. For the simulated data where the elements of $\boldsymbol{\nu}$ were kept equal, the parameter-combination with reduced $\boldsymbol{\nu}$ and full $\boldsymbol{\gamma}$ and $\boldsymbol{\lambda}$ is preferable. This model will be referred to as the "reduced model".

The model selection has a considerable impact on the population size estimates as indicated in Figure 2. The difference is caused by high estimates for the natural mortality rate in the years around 1993 when using the "full model", see Figure 7. In turn the high natural mortality rates causes large estimates for the number of natural dead moose which

Table 1: DIC-values for models with different combinations of number of elements for the parameters $\boldsymbol{\lambda}$, $\boldsymbol{\gamma}$ and $\boldsymbol{\nu}$. The three first columns defines the parameter combination. The "full model" has a parameter combinations where all parameters are full, whereas the "reduced model" has reduced $\boldsymbol{\nu}$-parameter, but full $\boldsymbol{\lambda}$ and $\boldsymbol{\gamma}$-parameters. The next 6 columns show DIC-values and the ranking of models based on DIC-values for "real data" and the two simulated data-sets.

Redused (=1)			Real data		Simulated from "Full model"		Simulated from "Redused ν"	
γ	λ	ν	DIC	Rank	DIC	Rank	DIC	Rank
"Full model"			5049	1	4630	1	4360	2
"Reduced model"			6163	3	5307	3	4336	1
-	1	-	5542	2	5195	2	4791	5
-	1	1	6680	7	5871	6	4780	4
1	-	-	6188	4	5689	5	5071	6
1	-	1	6427	5	5663	4	4672	3
1	1	-	6665	6	6180	8	5425	8
1	1	1	6827	8	6104	7	5077	7

in turn increases the estimates for total population size for the years prior to 1993/94. For the period successive to 1993/94 differences in estimated natural mortality rate between the "full model" and the "reduced model" are minor, and thereby differences for population size estimates are small for this period.

In general all models seem to estimate moose population size, $\boldsymbol{N}$, and the parameters, $\boldsymbol{\theta}$, reasonably well as seen from Table 2. It is worth noticing that the "reduced model" estimates the very determining parameter α poorly for the simulated data with varying elements for $\boldsymbol{\nu}$, where the true value for α_M is less than all simulated values, i.e. $\epsilon_p^\alpha = 0$. However, in general the "full model" and "reduced model" has higher ("better") values, close to the expected number 0.5, for $\overline{\epsilon_p}$, and smaller average error, $\overline{\epsilon_\mu}$ for the parameters α and $\boldsymbol{\lambda}$, compared to the other relevant parameter-combinations, see Table 2. The posterior variance, $\overline{\epsilon_\sigma}$, varies little between models for the parameter α, but for the parameter $\boldsymbol{\lambda}$ the "full model" and the "reduced model" stands out once again with lower posterior parameter variances.

All relevant models seem to estimate $\boldsymbol{\gamma}$ well, with $\overline{\epsilon_p}$ values close to 0.5, and quite similar values for average error, $\overline{\epsilon_\mu}$, and average posterior parameter variance, $\overline{\epsilon_\sigma}$. The test-variables for $\boldsymbol{\gamma}^C$ vary very little between the different relevant models. As a consequence of rounding no differences are expressed in Table 2, though present. The result is as expected since $\boldsymbol{\gamma}^C$ might be viewed as a vector of probability-parameters for an observed binomial

variable, $[K_1^C, \ldots, K_J^C]$, where the other parameter in question, $\boldsymbol{s}^C$, also is observed.

When testing different models ability to simulate the natural mortality rate the two relevant models with full $\boldsymbol{\gamma}$ stands out as better since they have higher values for $\overline{\epsilon_p}$, and lower values for average error, $\overline{\epsilon_\mu}$, and average posterior parameter variance, $\overline{\epsilon_\sigma}$.

Posterior means for the yearly total numbers of natural dead moose from the "full model" and the "reduced model" are both highly correlated with adjusted non-harvest mortality numbers registered by Statistics Norway (2013). The Pearson correlations are 0.758 and 0.772 respectively, and the corresponding p-values for the null hypothesis of no correlation between estimated number- and registered number of dead moose are $2.8 \cdot 10^{-05}$ and $1.6 \cdot 10^{-05}$.

The estimates of the population size, $\boldsymbol{N}$, from the "full model" stands out as better than estimates from other parameter-combinations, especially since it has substantially lower values for average error, $\overline{\epsilon_\mu}$. The "reduced model" has higher values for average error, $\overline{\epsilon_\mu}$, as a consequence of its poor model fit for the simulated population with full $\boldsymbol{\nu}$. The average values for ϵ_p^N are very high, i.e. larger than 0.5 for all, except one, parameter-combinations, and the biases seems to be very small. However, these results might be slightly misleading since the elements of $\boldsymbol{N}$, unlike the parameters in $\boldsymbol{\theta}$, are discrete variables, many of them with true value 0. The high average values for ϵ_p^N are partly a result of many 0-s estimated correctly in a high proportion of the simulations.

On the basis of the results presented in this section the "full model" is selected as the best model. The "full model" is already substantially reduced by restricting the length of $\boldsymbol{\lambda}$ till $I_2 = 3$ and by pooling $\boldsymbol{\nu}$ over all age-classes and both sexes. The results presented in the following sections are, if not noted differently, from the "full model" and with real data.

Table 2: Test variables for the different parameters. The Table shows average results for the 2 simulated populations, each simulation ran once. The number of simulations as basis for averages, $n^\theta = 2$ (simulated populations) $\times dim(\theta)$, and average true value, $\bar{\theta}$ are given for all parameters. The right columns for both females and males shows the fraction of $\epsilon_p = 0$, that is the fraction of simulations where the true value is outside the extreme bounds reached by the simulations. For each parameter only relevant model are included, i.e. models where the parameter in question is not reduced.

	Redused (=1)			Females				Males			
	γ	λ	ν	$\overline{\lvert\epsilon_\mu\rvert}$	$\overline{\epsilon_\sigma}$	$\overline{\epsilon_p}$	ϵ_p	$\overline{\lvert\epsilon_\mu\rvert}$	$\overline{\epsilon_\sigma}$	$\overline{\epsilon_p}$	ϵ_p
$\boldsymbol{\alpha}$											
	-	-	-	4.6e-03	2.9e-05	0.47	0	4.0e-03	1.6e-05	0.40	0
	-	-	1	8.7e-03	3.1e-05	0.39	0	7.7e-03	1.6e-05	0.45	50
$n^\alpha = 2$	-	1	-	1.5e-02	2.5e-05	0.036	50	7.8e-03	1.4e-05	0.14	0
	-	1	1	1.7e-02	2.6e-05	0.089	50	1.1e-02	1.4e-05	0.29	50
$\overline{\alpha^F} = 0.21$	1	-	-	1.8e-02	3.6e-05	0.067	50	6.5e-03	2.0e-05	0.21	0
$\overline{\alpha^M} = 0.20$	1	-	1	1.9e-03	3.4e-05	0.78	0	8.6e-03	1.7e-05	0.14	0
	1	1	-	1.3e-02	3.1e-05	0.20	50	9.2e-03	1.6e-05	0.026	0
	1	1	1	1.2e-02	2.5e-05	0.089	0	1.2e-02	1.3e-05	0.073	50
$\boldsymbol{\lambda}$											
	-	-	-	1.4e-02	6.9e-04	0.57	0	9.7e-03	1.4e-03	0.78	0
$n^\lambda = 6$	-	-	1	1.4e-02	8.4e-04	0.58	0	1.1e-02	1.3e-03	0.77	0
$\overline{\lambda^F} = 0.42$	1	-	-	5.1e-02	3.7e-03	0.45	0	1.0e-01	4.3e-03	0.14	0
$\overline{\lambda^M} = 0.65$	1	-	1	5.7e-02	3.0e-03	0.32	0	9.3e-02	3.1e-03	0.12	0
$\boldsymbol{\gamma}$											
	-	-	-	3.1e-02	2.0e-03	0.53	0	3.1e-02	2.0e-03	0.53	0
$n^\gamma = 50$	-	-	1	3.2e-02	2.2e-03	0.54	0	3.2e-02	2.2e-03	0.54	0
$\overline{\gamma^F} = 0.51$	-	1	-	3.1e-02	1.9e-03	0.54	0	3.1e-02	1.9e-03	0.54	0
$\overline{\gamma^M} = 0.56$	-	1	1	3.1e-02	2.3e-03	0.57	0	3.1e-02	2.3e-03	0.57	0
$\boldsymbol{N}$											
	-	-	-	2.5	21	0.65	0	1.9	16	0.76	0
$n^{N^F} = 1500$	-	-	1	3.8	16	0.61	2	3.1	10	0.72	2
$n^{N^M} = 950$	-	1	-	3.8	17	0.47	0	2.1	15	0.73	0
	-	1	1	4.7	13	0.47	4	3.2	9.6	0.70	3
$\overline{N^F} = 41$	1	-	-	6.0	27	0.55	3	4.1	17	0.66	2
$\overline{N^M} = 52$	1	-	1	4.4	17	0.57	2	3.4	9.5	0.68	2
	1	1	-	6.1	23	0.50	4	4.3	15	0.64	3
	1	1	1	5.0	13	0.48	4	3.5	8.5	0.66	3
$\boldsymbol{\gamma^C, \nu}$				γ^C, (Calves)				ν			
$n^{\gamma^C} = 50$	-/-	-/-	-/-	2.0e-02	3.5e-04	0.39	0	8.5e-03	6.0e-04	0.75	0
$n^\nu = 48$	-/-	-/1	1/-	2.0e-02	3.5e-04	0.39	0	1.0e-02	6.2e-04	0.72	0
$\overline{\gamma^c} = 0.26$	-/1	1/-	-/-	2.0e-02	3.5e-04	0.39	0	3.4e-02	8.0e-04	0.43	0
$\overline{\nu} = 0.056$	-/1	1/1	1/-	2.0e-02	3.5e-04	0.39	0	3.6e-02	8.0e-04	0.42	0

Total population size.

Number of natural dead moose.
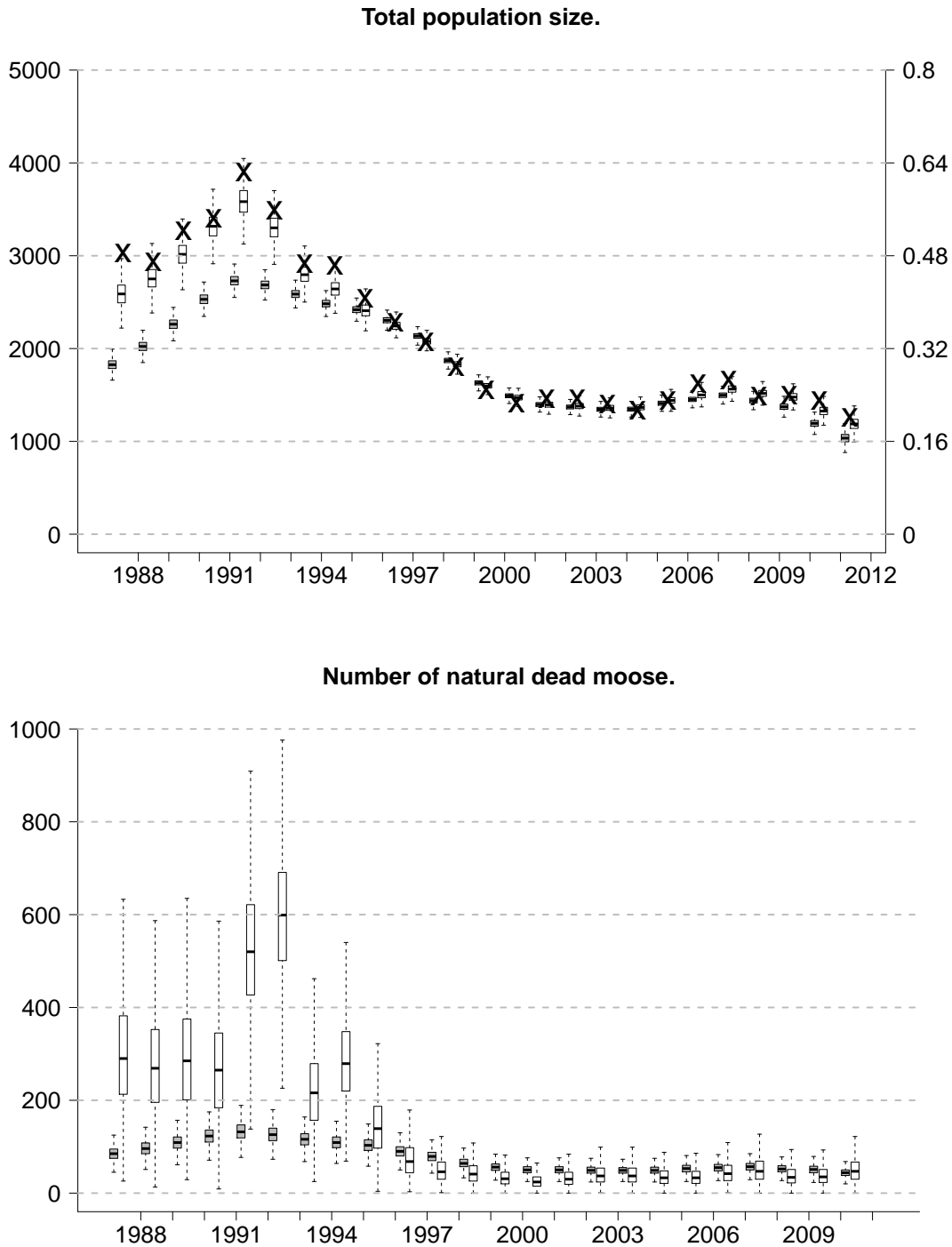
Figure 2: Boxplots for the Ringerike moose population over the period 1988 till 2012 based on $n_{sim} = 7500$ Monte Carlo simulations from the "full model" (white) and "reduced model" (grey). Upper panel: Total average population size. i.e yearly averages of $\bar{N}$. "Moose seen per hunter-day" from hunter-observations are shown with black x-s and scale at the right axis. Lower panel: Total number of natural dead moose, i.e. the number perishing from other causes than hunting between hunting seasons.

Sex-ratio.

Calf-ratio.
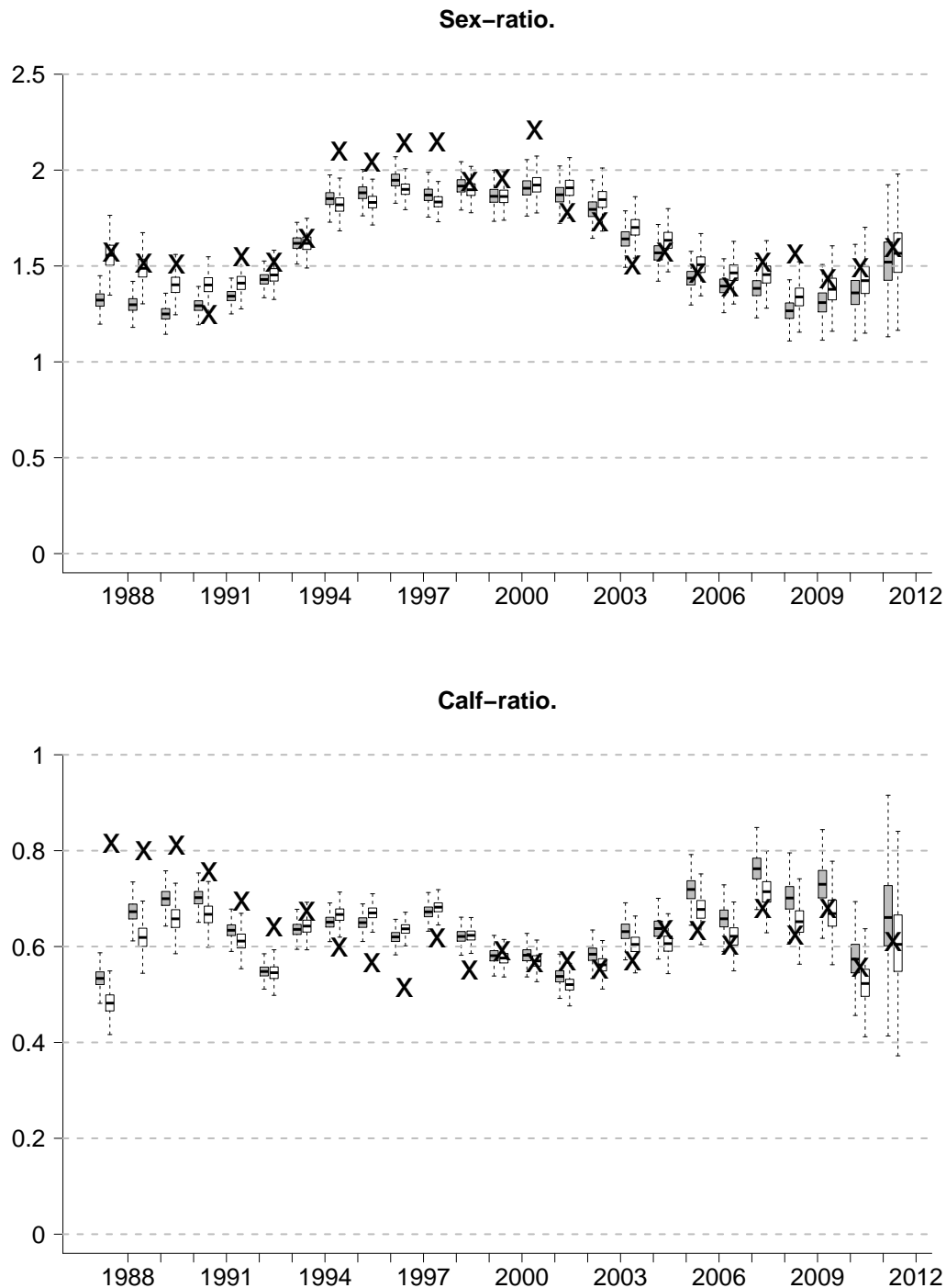
Figure 3: Boxplots for the Ringerike moose population over the period 1988 till 2012 based on $n_{sim} = 7500$ Monte Carlo simulations from the "full model" (white) and "reduced model" (grey). Results from hunter-observations are shown with black x-s. Upper panel: Average sex-ratio, i.e. the number of adult females per adult male, during the hunting seasons, and "female observed per male" from hunter-observations. Lower panel: Average calf-ratio, i.e. the number of calves per adult female, during the hunting seasons, and "calf observed per female" from hunter-observations.
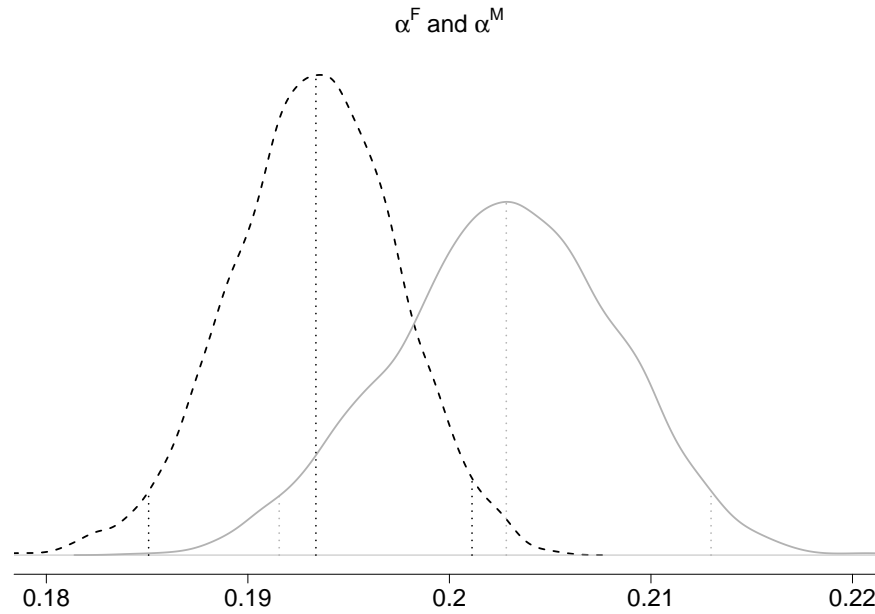
Figure 4: Monte Carlo posterior probability distributions based on $n_{sim} = 7500$ simulations from real data and the "full model" for the unknown parameters α^F (grey-, full line) and α^M (black-, dashed line). Posterior 2.5%-, 50%- and 97.5% percentiles are shown with dotted lines.

Parameter estimates, population development

The basic development in the moose population are shown in Figures 2 and 3, and are given in some more detail in Appendix 2. Notice that Figures 2 and 3 are based on average population sizes during hunting seasons, i.e. yearly averages for $\bar{N}$.

The result, based on the "full model", shows a rapid increase in moose density from 1988 till 1992, when the pre-harvest density peaked at a posterior mean of approximately 3900 individuals in total. This increase is followed by a steady decrease in moose density during the period from approximately 1992 till 2002. In the period 2002 till 2010 the moose density stayed quite stable with posterior mean pre-harvest population sizes at approximately 1700 individuals. Increased hunting yields over the last three years have resulted in reduced estimates for pre-harvest population sizes, to a posterior mean at approximately 1400 individuals for 2012. The development in moose density over the whole period corresponds well with the hunting yield, which is high in the periods with decreasing moose density and vice versa.

Population development estimates derived directly from hunter-observations, i.e. *moose seen per hunter-day, observed female per male* and *observed calf per female*, fits very well

$$\lambda^F \text{ and } \lambda^M$$
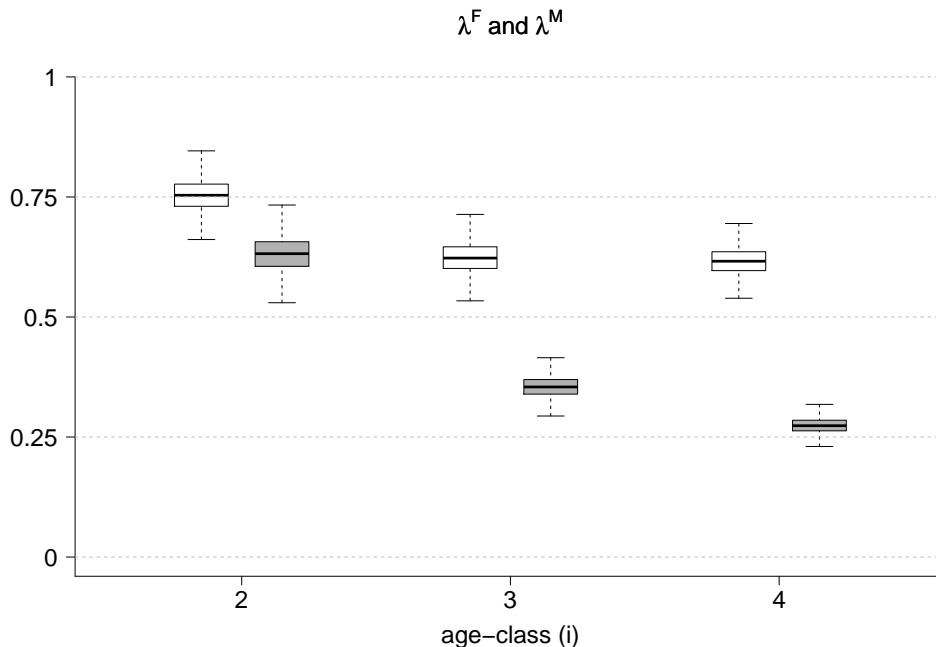
Figure 5: Boxplot for the parameters $\boldsymbol{\lambda}^F$ (grey) and $\boldsymbol{\lambda}^M$ (white) for the age-classes $i = 2, 3, 4$ in the Ringerike moose population based on $n_{sim} = 7500$ Monte Carlo simulations from real data and the "full model".

with the corresponding absolute estimates from the model, i.e. *population size*, *sex-ratio* and *calf-ratio* respectively. The fit between hunter-observations and model estimates is visualized in Figures 2 and 3. These figures also indicates that the fit between "full model" and hunter-observations are better than the fit between "reduced model" and hunter-observations. Further there seem to be some discrepancy regarding the calf-rate for both models, as the model estimates indicates higher calf-rates during the period 1988 till 1993 and lower calf-rates during the period 1995 till 1998 than the corresponding rates derived directly from hunter-observations.

The strong fit between hunter-observations and model-estimates should not be emphasized as important for model validation, since the model estimates are heavily influenced by hunter observations.

The general pattern for cohort analysis estimates is increasing variation for the population size estimates for the latter years of the analysis. As seen from Figure 2 this pattern is valid for the analysis of the Ringerike moose population back till approximately 1998. However, the variation for population sizes prior to 1998, especially prior to 1994, are higher than for the last years. The obvious reason for this inconsistency is the high, and
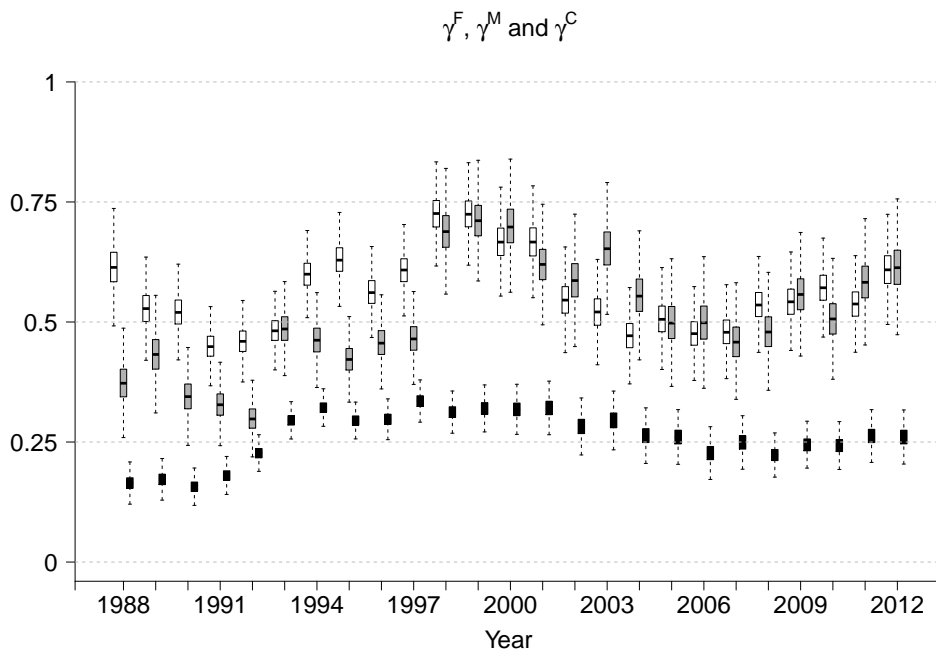
Figure 6: Boxplot for the parameters $\boldsymbol{\gamma}^F$ (grey), $\boldsymbol{\gamma}^M$ (white) and $\boldsymbol{\gamma}^C$ (black) for the years $1988, \ldots, 2012$ in the Ringerike moose population based on $n_{sim} = 7500$ Monte Carlo simulations from real data and the "full model".

quite variable, estimates for number of natural dead moose, $\boldsymbol{D}$, in the years prior to 1998.

Results for posterior estimates of unknown parameters, $\boldsymbol{\theta}$, are presented in Figure 4, 5, 6 and 7. The results are given in some more detail in Tables in Appendix (2).

The posterior distributions for the instantaneous observation rate, α, are very similar and not substantially different for the two sexes. The distribution for males is a little more "peaked" than the distribution for females, see Figure 4. Median posterior values for α are 0.203 and 0.193 for females and males respectively, i.e. close to 0.2 for both sexes. Consequently as a rule of thumb the average moose density (individuals per km^2) during the hunting season may be estimated by "$5\times$ *moose seen per hunter day*", where *moose seen per hunter day* is the total number of observed moose during the hunting season divided by total number of hunting days, a very common measurement for moose density.

When evaluating results for age-specific hunting mortality, $\boldsymbol{\lambda}$, notice the somewhat non-intuitive connection between age-class and real age. Calves has age 0 years, but their age-class is set to $i = 1$, yearlings has age 1 year, but their age-class is set to $i = 2$ etc. Further note that the elements of $\boldsymbol{\lambda}^F$ and $\boldsymbol{\lambda}^M$ are not directly comparable, since the probability of being shot conditional on being observed also depends on year-specific

$$\nu$$
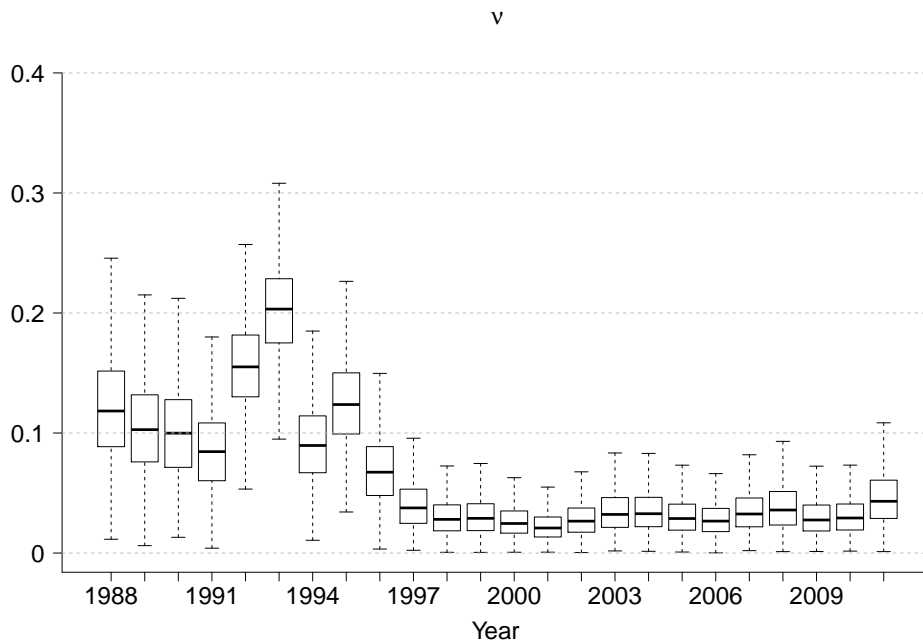
Figure 7: Boxplot for the parameter $\boldsymbol{\nu}$ for the years $1988, \ldots, 2011$ in the Ringerike moose population based on $n_{sim} = 7500$ Monte Carlo simulations from real data and the "full model"

hunting mortality, $\boldsymbol{\gamma}$, which differs for the two sexes. The results for $\boldsymbol{\lambda}$ shows that yearlings (age-class $i = 2$) has high age-specific hunting mortality for both sexes.

For females the age-specific hunting mortalities rapidly decreases with age, see Figure 5. For males the results show higher age-specific hunting mortality for yearlings than older age-classes. However, the difference between 2 year old males and older males seems to be minor.

The year-specific hunting mortality, $\boldsymbol{\gamma}$, is estimated for females, males and calves. The year-specific hunting mortality for females was very low in the period 1988 till 1992, see Figure 6. Thereafter it increased substantially from 1992 till 1993, stayed quite stable over a few years, and then increased rapidly from 1997 till 1998. After 1998 the year to year differences have been less highlighted, but there is a clear pattern of steadily decreasing values from 1999 till 2007 and thereafter increasing values till 2012.

For males a pattern comparable to females emerges. The year-specific hunting mortality increased rapidly for the years 1993 till 1994 and 1997 till 1998. The trend from 1998 till 2012 shows the same general pattern as the trend for females.

In the years prior to 1992 calves had a low year-specific hunting mortality. From 1991 till 1993 age-specific hunting mortality for calves increased rapidly followed by a period of

quite stable values in the period till 2001. From 2001 till 2008 the year-specific hunting mortality for calves decreases steadily followed by a minor increase in the period from 2008 till 2012.

In general the hunting pressure is higher for males than for females since the instantaneous observation rates, α, are close to equal combined with higher average values for both year-specific hunting mortality, $\boldsymbol{\gamma}$, and in particular age-specific hunting mortality, $\boldsymbol{\lambda}$, for males.

It is of essential importance to remember that the probability for a moose of being shot, given by $\boldsymbol{p}$, is a function of both unknown parameters, $\boldsymbol{\theta}$, and the observed hunting effort, $\boldsymbol{h}$, see equation (12). For instance, the year-specific hunting mortality for males, $\boldsymbol{\gamma}^M$, are high in the first years of the analysis. But the hunting mortality for males these years are not especially high compared with later years due to low hunting effort. For females and calves the total hunting mortalities these years are very low as they benefitted from both low hunting effort and low year-specific hunting mortality.

The natural mortality rate, $\boldsymbol{\nu}$, is pooled over all age-classes and both sexes. From the posterior distributions it is clear that the natural mortality rate is substantially higher than the prior median at 5% for the years 1992, 1993 and 1995, see Figure 7. The natural mortality rate is especially high for 1993, with 20.3% as its posterior median value. The high natural mortality rates corresponds with the peak for total moose population size, resulting in very high estimates for total number of moose perishing from natural causes during this period.

Model validation by subsets

Implementation of spatial- and temporal subsets is restricted to the "full model".

Figure (8) indicates a strong relationship between estimated values, $\hat{\boldsymbol{s}}$, and real values, $\boldsymbol{s}$, for the number of observed moose in the 22 spatial subsets (valds). The Pearson correlation between posterior mean for $\hat{\boldsymbol{s}}$ and $\boldsymbol{s}$ is 0.896. Even though the relationship is strong, it is not perfect, since the distributions for ϵ_p^s-s shows an evident pattern where they are skewed towards 0 compared to the standard uniform distribution (Figure 9). For all sex-/ age-classes the distribution of ϵ_p^s differed significantly from the standard uniform distribution, when tested by a χ^2-test (Besag & Clifford, 1991) ($p-value \approx 0$ for for females, calves and in total, and $p-value = 1.14 \times 10^{-12}$ for males). The skewed pattern is most evident for females and least for males, indicating a better model fit for males than females. Average values for ϵ_p^s were 0.336 for females (n=550), 0.447 for males (n=550), 0.416 for calves (n=550) and 0.400 for all classes in total (n=1650).
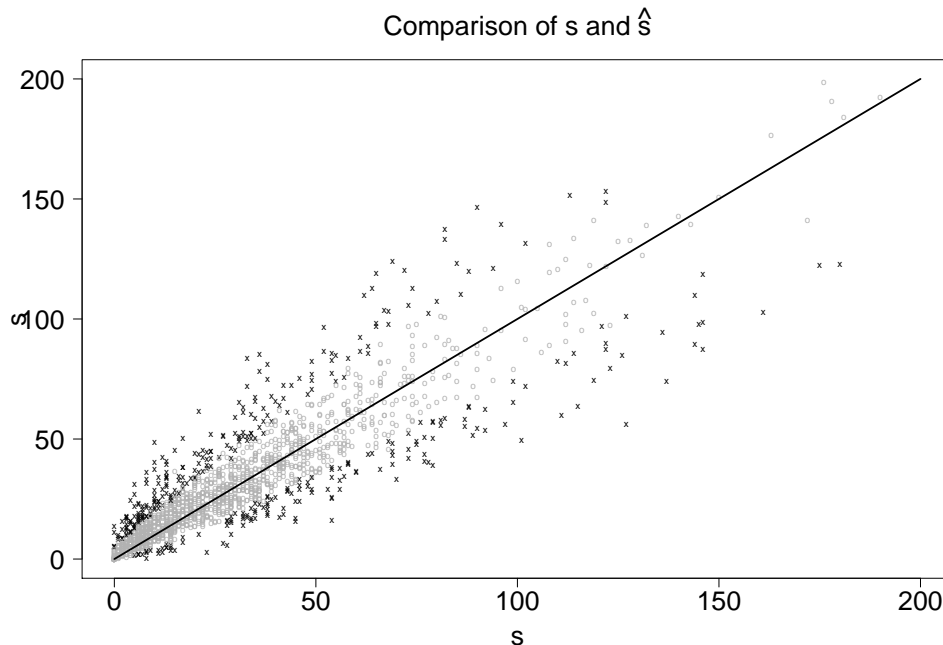
Figure 8: Posterior mean for number of observed moose against real observations ($n = 1650$). Grey circles (1269 in total, or 76.9%) indicates that the associated ϵ_p^s is less than 0.05 and black x-s (381 in total, or 23.1%) indicates the opposite.

As seen from Figure 9 there are substantial differences between distributions for $\overline{\epsilon_p^s}$ based on reshuffled elements for $\boldsymbol{s}$ and the observed values for $\overline{\epsilon_p^s}$, i.e. the values calculated with the elements from all $\boldsymbol{s}$-s in their original sequence. In fact for all sex-/age-classes no value for $\overline{\epsilon_p^s}$ based on reshuffled elements are as high as the observed value, resulting in Monte Carlo p-values at $\approx 10^{-4}$ for the hypothesis that *the model has no predictive effect for $\boldsymbol{s}$*. This result strongly supports the validity of the model.

A linear model with ϵ_p as response-variable and *sex-/age- class*, *vald*, *year*, and the interactions *sex-/age-class - year* and *sex-/age-class - vald* as predictor variables shows a significant effect of both *sex-/age- class* ($p-value = 7.22 \times 10^{-8}$), *vald* ($p-value = 2.31 \times 10^{-11}$) and *year* ($p-value = 4.33 \times 10^{-4}$), but not for the interaction terms (p-values 0.265 and 0.268 respectively). This result strongly indicates that the model-fit vary with both vald and year. When splitting ϵ_p^s according to the valds the average ϵ_p^s ($n = 75$ for each vald) vary between 0.279 and 0.627, and when splitting according to years the average ϵ_p^s ($n = 66$ for each year) vary between 0.315 and 0.524. There is a pattern with smaller average values for ϵ_p^s during the period $1993 - 2001$, indicating a poorer model-fit for this period.

Density distributions for ϵ_p^s

Pdf–s vs. observed values for $\overline{\epsilon_p^s}$
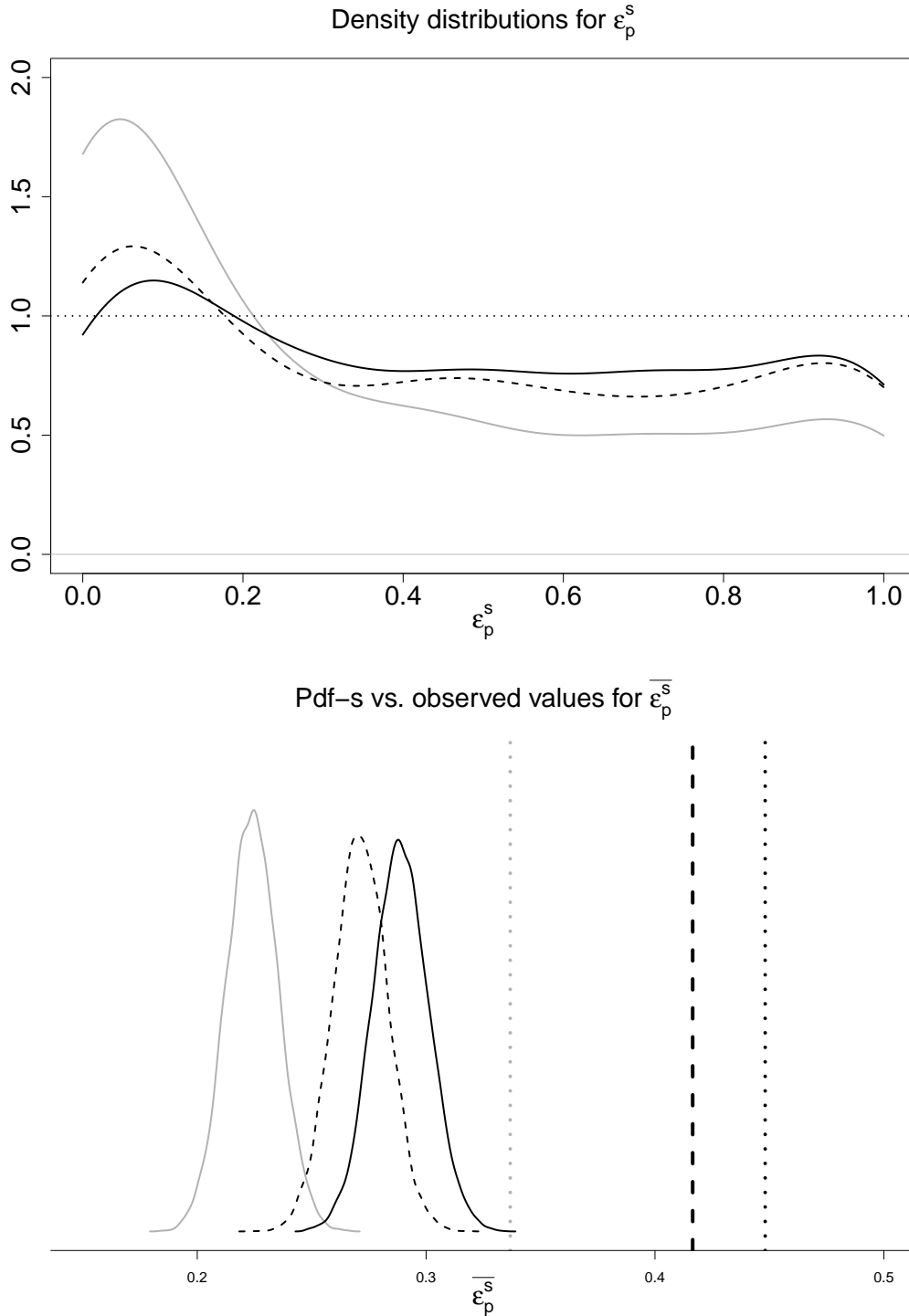
Figure 9: Upper panel: Densities for ϵ_p^s-s for females (grey, $n = 550$), males (black, $n = 550$) and calves (dashed black, $n = 550$). Lower panel: Densities for $\overline{\epsilon_p^s}$-s based on 10^4 reshuffled elements for females (grey), males (black) and calves (dashed black). Average value for ϵ_p^s-s based on $\boldsymbol{s}$ in original sequence are shown with dotted lines for females and males, and dashed line for calves.
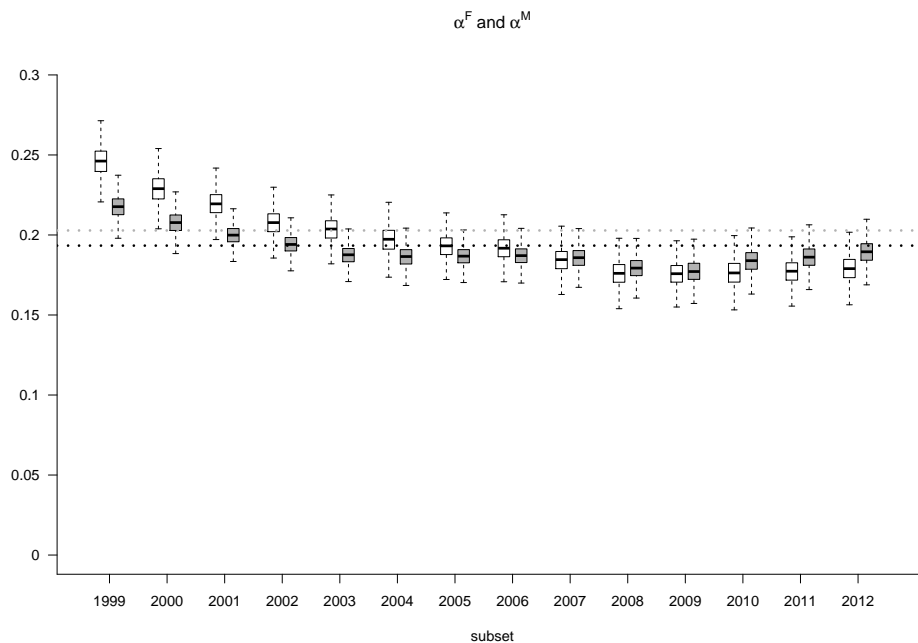
Figure 10: Boxplot showing posterior Monte Carlo distributions for α^F (grey) and α^M (white) from the different time subsets. Posterior median values from the full dataset are shown with dotted lines.

Comparisons between the vast majority of the temporal subsets have limited value due to dependency between them. Hence formal validation tests performed on these subsets are omitted. However, Figure 10 clearly demonstrates a pattern of decreasing estimates for instantaneous observation rate, α, with increasing end year for subsets, until the subset ending in year 2009. The slope in this trend is steeper for females than for males, and steeper during the first years than the latter years. The subsets ending in the years 1999 and 2012 respectively, are totaly independent. For both sexes values for α is significantly different for these two subsets, with Monte Carlo $p-values \approx 0$ and 0.007 for females and males respectively.

The estimates for age-specific hunting mortality, $\boldsymbol{\lambda}$, do not depend heavily on time subset. When comparing estimates for $\boldsymbol{\lambda}$ from the subsets ending in 1999 and 2012 there exist significant differences, i.e. Monte Carlo $p-value < 0.05$, for λ_3^F and λ_2^M.

Discussion

First, and most important, the model used and tested in this study performed very satisfactorily. Especially convincing is the cross-validation test from the spatial subsets. Even though the distribution of the Monte Carlo $p - values$ (ϵ_p-s) are skewed compared to the ideal standard uniform distribution, the results strongly supports the model's validity.

There are two main challenges for practical implementation of the model, one of them being the dependency between instantaneous observation rate, α, and temporal subset. Further estimates for the natural mortality, $\boldsymbol{D}$, vary substantially between models. Both of these challenges might have huge impact on population size estimates, and are discusses in detail later in this section.

Model selection is an important issue of this study. In this respect DIC-values for different models have been emphasised. Using DIC for model selection seems to favor complex models, and an alternative model selection criteria, *Bayesian predictive information criterion* (BPIC), has been proposed by Ando (2007) to solve this problem. Since the model selected in this study is the most complex model tested, implementation of BPIC instead of DIC as selection criteria might have some minor influence on model selection.

The model seems to give a better model fit for males than females, at least if judged by the results from cross-validation by different spatial subsets. Since estimates for the two sexes are pooled for calves, and thereby depends on each other, the better model fit for males might have positive impact also on female estimates.

Model assumptions and errors

To obtain simple and easily computable formulaes in the estimation algorithm some simplifying model assumptions are made, which in general are considered to be of minor importance for the parameter estimates. However, the general effect of such simplifications is reduced variation in the posterior parameter distributions.

The reduction in the population size during the hunting season is a stepwise (discrete) process. Consequently the assumption about constant relative reduction in moose density during the hunting season, an assumption in line with traditional principles for harvesting mortality (Ricker, 1940), leading to equation (13), is a simplification.

Further, the assumption of no natural mortality during the hunting season is an obvious simplification. In practise the opposite situation is more likely to emerge, i.e. increased natural mortality rate during the hunting season due to individuals shot and fatally wounded, but vanishing from the hunters. Under the model outlined in this study these individ-

uals would be regarded as perished from natural causes since they are not registered in the hunting yield. However, hunting is by far the dominating mortality factor during the hunting season, making also this simplification of minor importance.

In this study it is assumed no aging errors for moose killed during the hunt. Rolandsen et al. (2008) show that age determination by counting dental cement layers in incisors is inaccurate to some extent. However, the inaccuracies are minor, and if errors are not systematically wrong, cohort analysis methods are quite robust against aging errors when cohort sizes are quite stable, which is the case for moose populations.

Interpretation of model parameters, θ

The results from the real data indicate a dependency between the instantaneous observation rate, α, and the temporal subsets. This dependency is a violation of the model assumptions and might cause severely biased population estimates especially for the latter years in the analysis. The dependency might be due to development of α over time (years) and/or due to dependency between α and moose density. Such dependencies can be given credible practical explanations.

In the model it is assumed that any random hunter has a constant instantaneous chance of observing a random moose in the whole study area. In practice every hunter/hunting team disposes a limited hunting area, and the moose density might vary considerable between different areas. Indications of differences between areas, not necessarily restricted to the α-parameter, is indicated by the fact that there are a significant *vald-effect* in the cross-validation for spatial subsets.

The irregularities for α is in line with observations from Solberg et al. (2010), where an imperfect fit between moose density and *moose seen per hunter day* is reported. The population size estimates in Solberg et al. (2010) are highly reliable since they are based on marked moose in a population where the large majority of moose are marked. However, the study object was a "young" moose population founded in 1985 (Solberg et al., 2011) at the island Vega, on the cost outside the county of Nordland. The nature on Vega and in Ringerike is extremely different. Consequently the transfer value of the results from Vega till Ringerike might be questioned.

The different estimates for α in different temporal subsets imply that any population size estimate derived from *effort* will be biased unless the variation in α is implemented in the model. For a lot of wild spices, both fish, birds and mammals, the most common and important population size estimates depends, to varying degrees, of *effort*. The limitations and potential biases by using such estimates are well known (Quinn & Deriso,

1999, pp.1–49). However, a better understanding of, and improved methods for modelling, the relationship between effort and yield/observations would be a valuable contribution to improve a broad specter of models used for estimating abundance of wild species.

The high estimates for age-specific hunting mortality, $\boldsymbol{\lambda}$, for age-class $i = 2$ (yearlings) are expected. In most valds hunting quotas for calves and yearlings are pooled. Hunters mostly prefer to shot yearlings before calves due to higher meat yield. Solberg et al. (2010) also found that yearlings seemed to expose themselves easier for hunters.

The rapid decrease in age-specific hunting mortality for females from age-classes $i = 2$ till $i = 4$ is also as expected. The present data show that females carcass weight and reproductive rate, both factors potentially recognizable by hunters, increases until the age of approximately 3 years ($i = 4$), a pattern consistent with other studies, for instance Solberg et al. (2006). For older age-classes these factors are quite stable. Unreported results, with $I_2 > 3$, show a pattern where the age-specific hunting mortality for females are stable for age-classes 3 years and older.

In contradiction to females there are several factors, for instance weight and antler size, both potentially recognizable by hunters, developing until at least 6 years of age ($i = 7$) for males. Unreported results, with $I_2 > 3$, show a pattern where the age-specific hunting mortality for males decreases until 3 years of age, but thereafter increases until $6 - 7$ years of age. Consequently the decision to set $I_2 = 3$ might be questioned for males. Ideally different combinations for I_2 between the two sexes should be tested and evaluated using DIC-values and test-variables. However, the model is quite complex as it is, and the hard restriction on the $\boldsymbol{\lambda}$-parameter is justifiable.

For the hunting seasons 1984 till 2009 yearly reports about hunter observations, hunting yield and general moose management are available from the Ringerike wildlife board. In retrospective these reports gives valuable knowledge about the contemplations performed at the time. By consulting the reports for the two periods 1992 till 1994 and 1997 till 1998 (Bergan, 1994, 1995, 1998) it turns out that the increase in year-specific hunting mortality during these periods coincides very well with the guidelines given to hunters by the Ringerike wildlife board.

The observed moose, $\boldsymbol{s}$, is an element of the observed data, $\boldsymbol{Y}_{obs}$, and has an year-structure. Consequently these data, combined with the yearly hunting yields, contain considerable information about the year specific hunting mortality, $\boldsymbol{\gamma}$. Oppositely these data lack age-structure for adult moose, and thereby also lack any direct information about the age-specific hunting mortality, $\boldsymbol{\lambda}$. Therefore it makes sense to estimate $\boldsymbol{\gamma}$ for all years, but to restrict the number of elements in $\boldsymbol{\lambda}$ till I_2. Likely, the pattern where models with "full $\boldsymbol{\gamma}$" in general scores lower DIC-values than models with "reduced $\boldsymbol{\gamma}$", is a result of

the year-structure in $\boldsymbol{s}$.

As shown in (15) K_{ij} is a binomial distributed variable with $\lambda_i \gamma_j$ as probability parameter. Consequently the estimates for $\boldsymbol{\lambda}$ and $\boldsymbol{\gamma}$ might be severely biased and at the same time the matrix $(\boldsymbol{\lambda} \times \boldsymbol{\gamma^t})_{I_2 \times J}$ might be only minor biased. The same argumentation is valid for $\boldsymbol{p}$ as seen from equation (12). When flat prior distributions for $\boldsymbol{\lambda}$ and $\boldsymbol{\gamma}$ were applied, i.e. $\phi_1^\lambda = \phi_2^\lambda = \phi_1^\gamma = \phi_2^\gamma = 1$, this situation in fact occurred for the simulated populations (unreported results), resulting in underestimation of $\boldsymbol{\lambda}$ and correspondingly overestimation of $\boldsymbol{\gamma}$. By adjusting the prior distributions to be vaguely informative, symmetric, and with prior expectation equal to 0.5, i.e. $\phi_1^\lambda = \phi_2^\lambda = \phi_1^\gamma = \phi_2^\gamma = 5$, the posterior distributions for the parameters in question performed close to perfection for the simulated populations. A possible, not verified, reason for the unsatisfying results using flat prior distributions is the somewhat unwise choice of initial value for $\boldsymbol{S}^2$, which is set equal to $\boldsymbol{K}$.

The natural mortality rate, $\boldsymbol{\nu}$, is a key parameter for estimating "on target" population sizes. Most studies using "cohort-analysis" or related methods for mammals, but also for fish (see Quinn & Deriso, 1999, pp.355–357), operates with fixed, assumed known, values for natural mortality rates, $\boldsymbol{\nu}$. The ability to estimate natural mortality rate, demonstrated in this study, should be considered a step forward even if the estimates lean heavily on the informative prior distribution.

Setting informative priors for the natural mortality rate, $\boldsymbol{\nu}$, is justified by at least two causes. Primarily a sufficient "narrow" prior distribution for $\boldsymbol{\nu}$ is required for the model to converge. If this prior distribution is set uninformative the model tend to produce very large, and not converging, estimates for $\boldsymbol{N}$, $\boldsymbol{D}$ and $\boldsymbol{\nu}$ along with small estimates for $\boldsymbol{p}$ (unreported results), that is, a very large moose population where the main mortality cause is natural death, not hunting. Additionally the prior knowledge for natural mortality rate is quite high due to a number of studies with radio-marked moose, see for instance Rolandsen et al. (2010), Roer & Gangsei (2008) or the website *www.dyreposisjoner.no* (Norwegian Institute for Nature Research, 2013).

The estimated mortality rates around year 1993 are very high with posterior mean at 20% for the year 1993/94. The winter 1993/94 was a winter with deep snow and stable cold temperatures after a series of previous succeeding winters with little snow and warm weather in the late 80-s and early 90-s. From another study (Roer & Gangsei, 2008) in the municipality of Vegårshei in southern Norway an estimated natural mortality rate among radio-marked moose at 14% ($n = 47$) is observed during the years 2006 till 2008, a period with harsh snow conditions in the area. Even if data are scarce, this mortality rate is in line with estimated rates for Ringerike around 1993.

The number of parameters to be estimated is large, and further restrictions by using the

"reduced model" might be wise in a number of occasions. The evaluation with simulated populations showed, beyond doubt, that the "full model" is superior when the natural mortality rate vary substantially between years. On the other hand the simulations showed that "reduced model" is better when natural mortality rate is constant or close to constant.

The effects of varying data foundation are untested. Generally more complex models might be applied if a huge amount of data is available. Consequently if the data basis is large, like in Ringerike, and the prior biological knowledge indicates a varying natural mortality rate, the "full model" should be preferred. Oppositely if data basis is scarce, and the prior biological knowledge indicates near constant natural mortality rate over years the "reduced model" is the better choice. Of course the intersection point for which model to choose lies somewhere in between these two opposite situations. A practical solution when in doubt about what model to choose is to use both models and evaluate them by DIC-value.

Further restrictions on number of unknown parameters could be conducted by reducing the number of age-specific hunting mortalities to be estimated. I.e. further reduction in the length of $\boldsymbol{\lambda}$ by setting $I_2 < 3$. Obviously the gain is limited. The results indicate that the year-specific hunting mortality, $\boldsymbol{\gamma}$, should be estimated for every year, i.e. $\boldsymbol{\gamma}$ should have length J. Of course the instantaneous observation rate, α, should always be estimated. However, α might be pooled for the two sexes. The parameter estimates for areas like Ringerike provides a valuable basis for more informative prior distributions in other areas.

Population-size estimates. Interaction between factors

There are three factors that potentially might have huge impact at the population size estimates, (i) inconsistency in the assumption about constant instantaneous observation rate, α, (ii) severely biased estimates for natural mortality rate, $\boldsymbol{\nu}$, and (iii) the presence of migration which is modelled to be absent. Of course these three factors might interact.

Figure 2 demonstrates that the huge estimate for the natural mortality rate, $\boldsymbol{\nu}$, around 1993 is likely to be a result of the assumption about constant α, since constant α is expected to give a very tight fit between moose density and "moose seen per hunter day". This fit is much better when the natural mortality rate is allowed to vary between years. As a consequence the yearly estimates for natural mortality rate from the "full model" will be reliable if the assumption about constant α is correct, and vice versa.

Surveys from the county Nord-Trøndelag (Rolandsen et al., 2010, pp. 73–74) indicates that the assumption of no net migration for any age-class any year might be dubious. In Gangsei (1999) an attempt to estimate the net migration size was carried out. Even if

these migration estimates have several weaknesses, a pattern related to these estimates is apparent in Figure 3. The discrepancies between calf-rates estimated by the model, and the calf-rates from hunter observations, might be a result of migration, see Gangsei (1999) for further explanation.

In the cross-validation for spatial subsets there are a significant *year-effect*. It seems like the model's prediction accuracy for observed moose, tested in the cross-validation, is poorer for the period 1993 till 2001. This period coincides mostly with the period crucial for the discrepancies between population size estimates from the "full model" and the "reduced model", which are most evident in the period prior to 1995. Further, during this period the hunting yields were high and so was the estimated hunting mortalities. It seems likely that circumstances with large hunting quotas and hunting yields, possibly combined by somewhat different management strategies among the valds, paves the way for a poorer model fit/violation of the model assumptions.

Impacts for practical moose management

The method used in this study has not been tested against alternative methods for estimating moose population size, for instance methods based on hunter observations alone (Solberg & Sæther, 1999) or changes in sex ratio (Solberg et al., 2005). Neither has the method been tested on other moose populations than the Ringerike population. Such tests and comparisons are obvious task for further work.

For the particular case in this study, moose in Norway, several studies have shown that *moose seen per hunter day*, i.e. the relative moose density estimate derived from the instantaneous observation rate, α, generally responds well to changes in moose density, but the validity differs between areas (Gangsei, 1999; Solberg et al., 2006, 2010; Solberg & Sæther, 1999). Solberg & Sæther (1999) and Solberg et al. (2010) questions to some extent the validity of estimates for absolute population sizes from *moose seen per hunter day*. The present study challenge this view since *moose seen per hunter day* fits extremely well with the estimated population size, especially the estimates from the "full model". Using results from Ringerike derived from a cohort analysis method independent of hunter observations, for instance the method described by Ueno et al. (2009), to compare absolute population size estimates and *moose seen per hunter day* would most likely result in a better fit than reported in Solberg & Sæther (1999) and Solberg et al. (2010).

Estimating moose population size based on change in sex ratio is described by Solberg et al. (2005). Even though the method has several benefits, it is less intuitive than using estimates derived from *moose seen per hunter day*. Further, the method leans heavily on

an assumption of known sex ratio among calves each year. As pointed out by Solberg et al. (2005) the sex ratio among calves might be affected by paternal age and the sex ratio in the adult population (Sæther et al., 2004). Further a hypothesis known as the *Trivers-Willard hypothesis* (Trivers & Willard, 1973) states that the general fitness of females in the population might influence the sex ratio among calves.

For areas where killed at age data are present methods for estimating population size based on cohort analysis combined with hunter observations, like the one presented in this study, should be superior to methods based purely on hunter observations. Since the method incorporates data from hunter observations in the estimation procedure it should also perform better than cohort analysis methods based solely on killed at age data, like for instance Ueno et al. (2009).

Conclusion

The model described in this study produces reliable estimates for moose population size. Its utility is to some extent limited by the shortage of killed at age data in most parts of Norway. Anyhow the Bayesian framework forms a suitable basis for applying a slightly adjusted version of the model to areas lacking killed at age data.

Further work should be concentrated on developing improved methods for modelling variation in the instantaneous observation rate, α, over time. Such methods could be a valuable contribution to improve a broad specter of population size models.

References

ALBERT, J. (2011). *LearnBayes: Functions for Learning Bayesian Inference.* R package version 2.12, URL `http://CRAN.R-project.org/package=LearnBayes`.

ANDO, T. (2007). Bayesian predictive information criterion for the evaluation of hierarchical Bayesian and empirical Bayes models. *Biometrika* 94 443–458.

BERGAN, H. (1994). Fellings- og tellingsresultater 1993. Note 1, Ringerike wildlife board, Hønefoss, Norway.

BERGAN, H. (1995). Fellings- og tellingsresultater 1994. Note 1, Ringerike wildlife board, Hønefoss, Norway.

BERGAN, H. (1998). Fellings- og tellingsresultater 1997. Note 1, Ringerike wildlife board, Hønefoss, Norway.

BESAG, J. & CLIFFORD, P. (1991). Sequential Monte Carlo p-values. *Biometrika* 78 301–304.

CONGDON, P. (2010). *Applied Bayesian hierarchical methods.* CRC Press.

DAHL, D. B. (2012). *xtable: Export tables to LaTeX or HTML.* R package version 1.7-0, URL `http://CRAN.R-project.org/package=xtable`.

DELLAPORTAS, P. & ROBERTS, G. O. (2003). An introduction to MCMC. In J. Moller, ed., *Spatial Statistics and Computational Methods.*, vol. 173 of *Lecture Notes in Statistics.* Berlin: Springer, 1–41.

DERISO, R. B., QUINN, T. J., II & NEAL, P. (1985). Catch-age analysis with auxiliary information. *Canadian Journal of Fisheries and Aquatic Sciences* 42 815–824.

DUPONT, W. (1983). A stochastic catch-effort method for estimating animal abundance. *Biometrics* 1021–1033.

GANGSEI, L. E. (1999). *Er alle elgar like lette å observere? Evaluering av sett-elg metodikk ut frå samanlikning med modellpopulasjonar bygd på kohortanalyse.* Master's thesis, Norwegian University of Life Sciences (UMB), Ås, Norway.

GELFAND, A., HILLS, S., RACINE-POON, A. & SMITH, A. (1990). Illustration of Bayesian inference in normal data models using Gibbs sampling. *Journal of the American Statistical Association* 85 972–985.

Gelfand, A. & Smith, A. (1990). Sampling-based approaches to calculating marginal densities. *Journal of the American statistical association* 85 398–409.

Geman, S. & Geman, D. (1984). Stochastic relaxation, Gibbs distributions, and the Bayesian restoration of images. *Pattern Analysis and Machine Intelligence, IEEE Transactions on* 721–741.

Gilks, W. R., Richardson, S. & Spiegelhalter, D. J. (1996). *Markov Chain Monte Carlo in practice.* Chapman & Hall/CRC.

Hilborn, R. & Walters, C. J. (1992). *Quantitative Fisheries Stock Assessment: Choice, Dynamics and Uncertainty.* New York: Chapman and Hall.

King, R., Morgan, B. J., Gimenez, O. & Brooks, S. P. (2009). *Bayesian analysis for population ecology.* Chapman & Hall/CRC.

Mysterud, A., Meisingset, E., Veiberg, V., Langvatn, R., Solberg, E., Loe, L. & Stenseth, N. (2007). Monitoring population size of red deer *Cervus elaphus*: an evaluation of two types of census data from Norway. *Wildlife Biology* 13 285–298.

National deer register (2013). Sett elg–felt elg. URL `http://www.hjorteviltregisteret.no/`.

Norwegian Institute for Nature Research (2013). Dyreposisjoner. URL `www.dyreposisjoner.no`.

Quinn, T. J., II & Deriso, R. B. (1999). *Quantitative fish dynamics.* New York: Oxford University Press, USA.

R Core Team (2012). *R: A Language and Environment for Statistical Computing.* R Foundation for Statistical Computing, Vienna, Austria. ISBN 3-900051-07-0, URL `http://www.R-project.org/`.

Ricker, W. (1940). Relation of "catch per unit effort" to abundance and rate of exploitation. *Journal of the Fisheries Board of Canada* 5 43–70.

Ringerike municipality (2012). Hunting areas. URL `http://kart.ringerike.kommune.no/Content/Main.asp?layout=ringerike&time=1341958204&vwr=asv`.

Roer, O. & Gangsei, L. E. (2008). Sluttrapport elgens områdebruk Vegårshei & omegn. Faun-rapport 020-2008, Faun Naturforvaltning AS, Fyresdal, Norway.

ROLANDSEN, C., SOLBERG, E., BJØRNERAAS, K., HEIM, M., VAN MOORTER, B., HERFINDAL, I., GAREL, M., PEDERSEN, P., SÆTHER, B., LYKKJA, O. & OS, Ø. (2010). Moose in Nord-Trøndelag, Bindal and Rissa 2005–2010. *Final report. NINA Report* 588.

ROLANDSEN, C. M., SOLBERG, E. J., HEIM, M., HOLMSTRØM, F., SOLEM, M. I. & SÆTHER, B.-E. (2008). Accuracy and repeatability of moose (*Alces alces*) age as estimated from dental cement layers. *European journal of wildlife research* 54 6–14.

SÆTHER, B., SOLBERG, E., HEIM, M., STACY, J., JAKOBSEN, K. & OLSTAD, R. (2004). Offspring sex ratio in moose *Alces alces* in relation to paternal age: an experiment. *Wildlife Biology* 10 51–57.

SOLBERG, E., GRØTAN, V., ROLANDSEN, C., BRØSETH, H. & BRAINERD, S. (2005). Change-in-sex ratio as an estimator of population size for Norwegian moose *Alces alces*. *Wildlife Biology* 11 163–172.

SOLBERG, E., ROLANDSEN, C., HEIM, M., GRØTAN, V., GAREL, M., SÆTHER, B., NILSEN, E., AUSTRHEIM, G., HERFINDAL, I. ET AL. (2006). Elgen i norge sett med jegerøyne. En analyse av jaktmaterialet fra overvåkningsprogrammet for elg og det samlede sett elg-materialet for perioden 1968-2004. *NINA Rapport 125: 197 pp.* 125.

SOLBERG, E. J., HEIM, M., ROLANDSEN, C., SÆTHER, B.-S., ARNEMO, J. M. ET AL. (2011). Immobilisering og merking av elg på Vega, 1992-2010. NINA Rapport 658, Norwegian Institute for Nature Research (NINA), Trondheim, Norway.

SOLBERG, E. J., ROLANDSEN, C. M., HEIM, M., LINNELL, J. D., HERFINDAL, I. & SÆTHER, B.-E. (2010). Age and sex-specific variation in detectability of moose (*Alces alces*) during the hunting season: implications for population monitoring. *European Journal of Wildlife Research* 56 871–881.

SOLBERG, E. J. & SÆTHER, B.-E. (1999). Hunter observations of moose *Alces alces* as a management tool. *Wildlife Biology* 5 43–53.

SPIEGELHALTER, D. J., BEST, N. G., CARLIN, B. P. & VAN DER LINDE, A. (2002). Bayesian measures of model complexity and fit. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)* 64 583–639.

STATISTICS NORWAY (2013). table 03501. URL `http://statbank.ssb.no/statistikkbanken/`.

Tanner, M. & Wong, W. (1987). The calculation of posterior distributions by data augmentation. *Journal of the American statistical Association* 82 528–540.

Tobiassen, C., Brøseth, H., Bakke, B. B., Aarnes, S. G., Hagen, S. B. & Eiken, H. G. (2012). Populasjonsovervåking av brunbjørn 2009-2012: DNA-analyse av prøver samlet i Norge i 2011. Bioforsk report 57, Bioforsk, Svanhovd, Norway.

Trivers, R. L. & Willard, D. (1973). Natural selection of parental ability to vary the sex ratio of offspring. *Science* 179 90–92.

Ueno, M., Matsuishi, T., Solberg, E. & Saitoh, T. (2009). Application of cohort analysis to large terrestrial mammal harvest data. *Mammal study* 34 65–76.

Wabakken, P., Aronson, Å., Strømseth, T. H., Sand, H., Maartmann, E., Svensson, L., Åkesson, M., Flagstad, Ø., Liberg, O. & Kojola, I. (2011). Ulv i skandinavia: Statusrapport for vinteren 2010-2011. Oppdragsrapport 2011-1, Høgskolen i Hedmark, Evenstad, Norway.

Appendix

Data ($\boldsymbol{Y}_{obs}$) for Ringerike (total)

The total hunting area in Ringerike is 1205 km^2.

Table 3: Number of observed moose for females, males and calves, and number of hunting days in the period 1988 till 2011.

Year	Number of observed			Number of hunterdays
	Females ($\boldsymbol{S^F}$)	Males ($\boldsymbol{S^M}$)	Calves ($\boldsymbol{S^C}$)	
1988	599	381	488	3031
1989	641	423	513	3360
1990	766	508	621	3623
1991	888	712	672	4173
1992	1225	792	850	4590
1993	1464	963	938	6021
1994	1440	875	970	7054
1995	1653	787	991	7416
1996	1477	724	837	7485
1997	1659	775	854	9003
1998	1258	586	777	7902
1999	1201	619	662	8608
2000	936	479	553	7929
2001	894	405	507	8013
2002	697	392	397	6383
2003	712	411	394	6494
2004	734	487	418	7294
2005	666	423	423	7094
2006	645	441	409	6497
2007	702	505	423	6287
2008	830	547	564	7293
2009	836	535	522	7945
2010	773	540	525	7690
2011	851	571	473	8198
2012	700	439	427	7794

Table 4: Number of shot female moose per age-class over the period 1988 till 2012.

Age	88	89	90	91	92	93	94	95	96	97	98	99	00	01	02	03	04	05	06	07	08	09	10	11	12
0	37	39	36	53	80	125	154	134	119	137	110	109	95	87	50	51	53	52	40	54	53	61	61	66	55
1	17	34	40	58	78	103	78	107	91	92	94	111	86	68	52	52	40	33	42	47	41	48	51	62	43
2	8	17	6	15	10	39	38	28	31	44	61	41	36	23	15	21	27	16	17	18	26	24	26	23	22
3	8	11	9	6	10	24	19	20	22	22	23	37	21	21	14	22	19	12	10	11	12	16	13	14	16
4	3	5	8	6	6	22	27	18	11	20	15	24	19	19	15	21	8	11	3	5	6	2	9	14	16
5	9	3	5	2	5	11	13	13	17	19	25	20	16	8	7	11	10	10	5	5	6	12	7	10	7
6	8	4	3	1	4	14	6	8	13	14	22	19	11	8	7	6	6	3	4	5	7	13	12	6	9
7	3	3	9	5	6	10	11	8	6	12	14	12	7	7	7	4	4	5	3	5	5	5	6	14	8
8	0	4	4	2	1	6	4	4	4	11	9	9	7	5	5	3	4	4	7	3	5	6	4	7	4
9	5	2	0	3	2	4	3	5	11	8	6	6	5	5	2	3	3	3	4	1	8	6	1	3	6
10	0	2	1	1	3	5	7	7	6	4	6	3	5	3	4	0	3	5	3	7	3	2	5	7	2
11	3	0	0	1	3	6	5	3	2	4	4	3	4	10	2	1	1	1	1	2	2	6	3	7	3
12	1	6	1	2	2	2	2	7	6	6	5	5	6	8	4	4	6	1	7	3	5	4	0	1	7
13	1	1	2	1	0	3	4	4	3	5	5	4	2	3	4	5	5	2	2	2	6	7	0	3	3
14	2	3	0	0	0	1	4	6	6	3	8	7	4	4	0	3	0	3	2	0	4	4	0	1	2
15	3	0	2	0	1	2	5	4	4	5	3	6	3	4	4	0	0	3	0	2	3	3	0	0	0
16	2	0	1	0	1	2	2	6	5	5	2	4	2	1	1	2	1	1	0	1	1	4	3	1	2
17	0	0	1	1	1	1	2	1	1	2	2	2	1	3	1	2	1	2	1	0	1	1	0	2	1
18	1	0	0	0	0	0	2	0	1	1	0	1	2	1	1	3	0	0	2	2	2	2	0	0	0
19	0	0	1	0	1	0	0	1	1	0	2	0	1	1	1	0	0	2	0	0	0	0	0	1	0
20	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	2	0	0	0	0	0
21	0	0	0	0	0	1	1	0	0	0	0	0	0	1	0	1	0	0	0	1	0	2	0	1	1
22	0	0	0	0	0	2	1	0	1	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0
23	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
24	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
25	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
26	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
27	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
28	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
29	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0

Table 5: Number of shot male moose per age-class over the period 1988 till 2012.

Age													Year												
	88	89	90	91	92	93	94	95	96	97	98	99	00	01	02	03	04	05	06	07	08	09	10	11	12
0	40	46	58	65	110	150	156	156	128	148	131	101	79	74	60	63	55	56	50	49	70	64	64	56	54
1	45	51	58	69	98	116	101	132	126	131	128	135	107	62	61	48	64	60	72	75	73	76	72	81	54
2	31	29	23	52	42	62	71	57	38	78	75	76	32	35	21	24	29	32	28	30	37	42	64	43	37
3	28	18	20	22	34	46	42	40	17	34	26	30	27	35	25	19	20	15	30	26	25	21	28	28	28
4	15	21	18	24	23	31	37	34	19	19	21	25	16	18	23	23	18	13	8	20	20	25	21	18	24
5	14	12	23	15	15	14	36	25	19	9	15	12	12	8	8	8	13	11	10	11	16	10	18	24	15
6	9	7	19	13	9	12	22	17	19	15	5	12	10	5	5	8	8	11	6	5	8	10	11	7	6
7	4	9	5	7	10	7	14	7	7	9	10	3	4	3	4	2	2	4	4	5	6	9	7	5	6
8	1	0	2	4	5	9	6	6	6	3	6	3	3	3	0	2	1	2	0	1	3	5	4	3	2
9	6	0	1	3	2	4	6	4	3	7	2	2	2	1	0	2	1	4	0	1	1	1	1	0	1
10	0	0	2	0	2	3	2	2	2	2	2	1	1	4	1	0	0	0	2	0	3	4	1	1	3
11	1	0	3	2	1	1	2	1	2	2	2	1	1	3	0	0	1	0	1	0	1	3	0	2	0
12	1	0	1	0	0	1	2	2	3	3	3	1	1	0	0	0	0	0	1	1	1	0	0	0	0
13	0	0	0	0	0	0	0	2	5	1	1	0	0	0	1	0	0	0	0	0	0	0	0	0	0
14	0	0	0	0	0	0	1	0	1	0	0	0	1	0	1	0	0	1	1	1	1	0	0	0	0
15	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
16	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
17	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
18	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0

Tables with results for the moose population in Ringerike

Table 6: 2.5%, 50% and 97.5% posterior percentiles for the instantaneous observing rates, α, for females and males. Estimates based on $n_{sim} = 7500$ simulations from real data and the "full model".

Females (α^F)			Males (α^M)		
2.5%	50%	97.5%	2.5%	50%	97.5%
0.192	0.203	0.213	0.185	0.193	0.201

Table 7: 2.5%, 50% and 97.5% posterior percentiles for the elements of age-specific hunting mortality ($\boldsymbol{\lambda}$) for females and males over the age-classes $i = 2$ till $i = 4$. Estimates based on $n_{sim} = 7500$ simulations from real data and the "full model".

Age-class(i)	Females ($\boldsymbol{\lambda^F}$)			Males ($\boldsymbol{\lambda^M}$)		
	2.5%	50%	97.5%	2.5%	50%	97.5%
2	0.56	0.63	0.70	0.69	0.75	0.82
3	0.31	0.35	0.40	0.56	0.62	0.69
4	0.24	0.27	0.31	0.56	0.62	0.67

Table 8: 2.5%, 50% and 97.5% posterior percentiles for the natural mortality rate (ν) over the period 1988 till 2011. Estimates based on $n_{sim} = 7500$ simulations from real data and the "full model".

Year	ν 2.5%	50%	97.5%	Year	ν 2.5%	50%	97.5%
1988	0.040	0.118	0.211	2000	0.006	0.025	0.059
1989	0.034	0.103	0.193	2001	0.005	0.021	0.053
1990	0.030	0.100	0.180	2002	0.006	0.027	0.061
1991	0.025	0.084	0.152	2003	0.008	0.032	0.080
1992	0.076	0.155	0.225	2004	0.009	0.033	0.075
1993	0.121	0.203	0.272	2005	0.007	0.029	0.070
1994	0.034	0.090	0.161	2006	0.006	0.027	0.063
1995	0.059	0.124	0.201	2007	0.008	0.033	0.081
1996	0.019	0.067	0.135	2008	0.009	0.036	0.086
1997	0.009	0.038	0.084	2009	0.006	0.028	0.067
1998	0.007	0.028	0.074	2010	0.007	0.029	0.067
1999	0.007	0.029	0.068	2011	0.011	0.043	0.105

Table 9: 2.5%, 50% and 97.5% posterior percentiles for the elements of year-specific hunting mortality (γ) for females, males and calves over the period 1988 till 2012. Estimates based on $n_{sim} = 7500$ simulations from real data and the "full model".

Year	Females (γ^F)			Males (γ^M)			Calves (γ^C)		
	2.5%	50%	97.5%	2.5%	50%	97.5%	2.5%	50%	97.5%
1988	0.29	0.37	0.47	0.53	0.61	0.70	0.13	0.16	0.20
1989	0.35	0.43	0.53	0.45	0.53	0.61	0.14	0.17	0.21
1990	0.28	0.34	0.43	0.45	0.52	0.60	0.13	0.16	0.19
1991	0.27	0.33	0.40	0.39	0.45	0.51	0.15	0.18	0.21
1992	0.25	0.30	0.36	0.40	0.46	0.52	0.20	0.23	0.26
1993	0.42	0.49	0.56	0.42	0.48	0.55	0.27	0.30	0.32
1994	0.39	0.46	0.54	0.54	0.60	0.67	0.29	0.32	0.35
1995	0.36	0.42	0.49	0.56	0.63	0.70	0.27	0.29	0.32
1996	0.39	0.46	0.53	0.50	0.56	0.63	0.27	0.30	0.33
1997	0.40	0.46	0.54	0.54	0.61	0.68	0.30	0.34	0.37
1998	0.60	0.69	0.79	0.65	0.73	0.81	0.28	0.31	0.35
1999	0.62	0.71	0.81	0.65	0.72	0.81	0.29	0.32	0.36
2000	0.60	0.70	0.80	0.59	0.67	0.75	0.28	0.32	0.36
2001	0.53	0.62	0.72	0.58	0.67	0.75	0.28	0.32	0.36
2002	0.49	0.59	0.69	0.47	0.55	0.63	0.24	0.28	0.33
2003	0.55	0.65	0.76	0.44	0.52	0.60	0.25	0.30	0.34
2004	0.47	0.55	0.66	0.40	0.47	0.55	0.22	0.26	0.31
2005	0.41	0.50	0.60	0.43	0.51	0.59	0.22	0.26	0.30
2006	0.40	0.50	0.61	0.41	0.48	0.55	0.19	0.23	0.27
2007	0.38	0.46	0.55	0.41	0.48	0.55	0.21	0.25	0.29
2008	0.40	0.48	0.57	0.47	0.54	0.61	0.19	0.22	0.26
2009	0.47	0.56	0.65	0.47	0.54	0.62	0.21	0.24	0.28
2010	0.42	0.51	0.61	0.50	0.57	0.65	0.21	0.24	0.28
2011	0.49	0.58	0.68	0.47	0.54	0.61	0.22	0.26	0.30
2012	0.52	0.61	0.73	0.53	0.61	0.70	0.22	0.26	0.30

Table 10: 2.5%, 50% and 97.5%-percentiles for pre-harvest population sizes for females, males and calves over the period 1988 till 2012.

Year	Females			Males			Calves		
	2.5%	50%	97.5%	2.5%	50%	97.5%	2.5%	50%	97.5%
1988	1106	1255	1429	788	861	950	560	626	707
1989	1117	1249	1409	814	883	970	714	787	878
1990	1184	1316	1481	921	997	1083	803	883	978
1991	1300	1447	1606	1025	1102	1189	906	988	1079
1992	1455	1612	1788	1129	1219	1319	957	1042	1137
1993	1477	1611	1766	1099	1177	1269	886	952	1027
1994	1253	1359	1474	900	949	1006	908	959	1022
1995	1227	1322	1452	792	833	890	895	948	1019
1996	1136	1213	1311	707	738	779	820	859	911
1997	1130	1183	1254	701	719	746	790	813	846
1998	1053	1097	1156	652	666	689	743	764	796
1999	978	1019	1072	605	618	637	623	642	668
2000	843	883	934	512	524	541	511	528	550
2001	768	809	862	450	463	480	469	487	511
2002	714	758	815	419	434	454	394	413	437
2003	697	741	798	414	431	454	408	428	455
2004	651	694	752	429	445	470	411	431	459
2005	632	676	734	435	453	478	408	430	458
2006	627	674	732	462	482	508	438	463	495
2007	659	711	776	507	530	561	433	458	492
2008	670	724	790	526	551	583	499	529	566
2009	664	720	784	547	575	609	448	478	512
2010	630	690	762	526	558	596	445	479	518
2011	622	690	768	488	528	575	340	377	421
2012	534	608	693	386	433	486	305	377	468

Table 11: 2.5%, 50% and 97.5%-percentiles for number of natural dead individuals for females, males and calves over the period 1988 till 2011.

Year	Females			Males			Calves		
	2.5%	50%	97.5%	2.5%	50%	97.5%	2.5%	50%	97.5%
1988	44	139	274	26	84	165	20	68	136
1989	36	119	243	22	76	157	21	73	150
1990	33	122	238	23	82	159	22	80	156
1991	30	113	215	20	77	150	19	73	144
1992	103	228	359	71	155	241	60	136	216
1993	149	272	392	100	182	265	77	143	211
1994	33	99	193	19	57	113	19	58	114
1995	59	133	236	29	66	118	34	81	146
1996	16	66	144	8	33	71	9	40	87
1997	7	35	83	3	15	38	3	18	46
1998	4	22	64	1	10	29	2	14	41
1999	3	20	53	1	8	23	2	11	31
2000	2	15	40	1	7	19	1	8	24
2001	2	12	34	0	5	15	0	6	20
2002	2	15	39	1	7	20	1	8	21
2003	3	18	49	1	9	26	1	10	28
2004	3	17	44	2	10	26	2	10	27
2005	2	15	41	1	9	25	1	9	25
2006	2	14	37	1	9	24	1	9	26
2007	3	18	50	2	12	34	2	11	32
2008	4	20	53	2	13	35	2	14	39
2009	2	15	39	1	10	28	1	9	27
2010	2	15	39	1	9	25	1	10	26
2011	4	21	58	2	13	37	2	11	31

Table 12: 2.5%, 50% and 97.5%-percentiles for pre-harvest mean-age of adult females and males, and pre-harvest sex ratio for the period 1988 till 2012.

| | Mean-age | | | | | | Sex ratio | | |
| | Females | | | Males | | | | | |
Year	2.5%	50%	97.5%	2.5%	50%	97.5%	2.5%	50%	97.5%
1988	4.76	5.00	5.28	2.86	2.96	3.08	1.33	1.46	1.60
1989	4.82	5.05	5.30	2.93	3.02	3.13	1.30	1.41	1.54
1990	4.78	4.99	5.22	2.93	3.02	3.11	1.22	1.32	1.43
1991	4.65	4.83	5.03	2.88	2.96	3.05	1.22	1.31	1.41
1992	4.60	4.77	4.94	2.89	2.97	3.05	1.24	1.32	1.42
1993	4.67	4.81	4.96	3.03	3.10	3.18	1.29	1.37	1.45
1994	4.88	5.00	5.14	3.29	3.35	3.42	1.36	1.43	1.51
1995	4.84	4.96	5.09	3.15	3.21	3.28	1.52	1.59	1.67
1996	4.87	4.97	5.08	2.98	3.02	3.08	1.58	1.64	1.72
1997	4.78	4.87	4.98	2.72	2.76	2.80	1.59	1.64	1.71
1998	4.86	4.96	5.07	2.58	2.61	2.66	1.59	1.65	1.71
1999	4.76	4.87	4.98	2.53	2.56	2.61	1.59	1.65	1.72
2000	4.88	5.00	5.14	2.54	2.58	2.63	1.62	1.68	1.76
2001	4.94	5.09	5.25	2.63	2.68	2.73	1.67	1.75	1.84
2002	4.93	5.10	5.27	2.56	2.61	2.67	1.66	1.75	1.85
2003	5.02	5.19	5.38	2.65	2.71	2.78	1.63	1.72	1.82
2004	5.15	5.34	5.55	2.61	2.66	2.73	1.47	1.56	1.66
2005	5.18	5.39	5.61	2.64	2.70	2.77	1.40	1.49	1.60
2006	5.18	5.40	5.64	2.66	2.72	2.80	1.31	1.40	1.50
2007	5.05	5.28	5.52	2.71	2.78	2.86	1.25	1.34	1.45
2008	5.15	5.39	5.64	2.80	2.87	2.96	1.22	1.31	1.42
2009	5.00	5.26	5.54	2.62	2.70	2.80	1.15	1.25	1.36
2010	4.74	5.04	5.35	2.60	2.70	2.82	1.12	1.24	1.37
2011	4.84	5.17	5.53	2.65	2.79	2.94	1.16	1.30	1.46
2012	4.91	5.33	5.76	2.86	3.06	3.28	1.20	1.40	1.64

Calculating formulas

Categorical distributions for S^1 and S^2 in equation (19)

The full conditional posterior distribution for any element of S^1, might be rewritten as follows:

$$\pi\left(S_{ij}^1|\boldsymbol{\theta}, \boldsymbol{Y}_{obs}, \boldsymbol{Y}_{mis}^{-S^1}\right) \propto f_{14}\left(S_{ij}^1|\boldsymbol{h}, \alpha, \bar{\boldsymbol{N}}\right) \cdot f_{15}\left(\boldsymbol{S}^2|\boldsymbol{\gamma}, S_{ij}^1, \boldsymbol{s}\right) \cdot \pi\left(S_{ij}^1|\boldsymbol{\Phi}\right)$$

The function $f_{15}\left(\boldsymbol{S}^2|\boldsymbol{\gamma}, S_{ij}^1, \boldsymbol{s}\right)$ is a bit of a challenge. For a given year j the total number of observed moose is given by s_j. Let S_{-ij}^1 denote the number of observed moose in year j from all age-classes except i, i.e. $s_j = S_{ij}^1 + S_{-ij}^1$. Then it is obvious that the element S_{ij}^1 in $f_{15}\left(\boldsymbol{S}^2|\boldsymbol{\gamma}, S_{ij}^1, \boldsymbol{s}\right)$ influences all elements of $\boldsymbol{S}_{:j}^2$, and not only the element S_{ij}^2. Further let S_{-ij}^2 denote the sum of $\boldsymbol{S}^2$ in year j for all age-classes except i. Then, since the element γ_j is common for all age-classes in year j, S_{ij}^2 might be seen as a hypergeometric variable as shown in (25).

$$\left(S_{ij}^2|\boldsymbol{s}, \boldsymbol{S}^2, S_{ij}^1\right) \sim Hypergeometric\left(s_j, \ S_{ij}^1, \sum_{i=1}^{I_2} S_{ij}^2\right) \tag{25}$$

The data augmentation for $\boldsymbol{S}^1$ is conducted through a stepwise process for each year. In the first step $\boldsymbol{S}^1$ is assumed unknown for all age-classes and a value for age-class 1 is drawn by equation (19) and assumed fixed. In the second step the elements in (25) is adjusted by subtracting age-class 1 and a value for age-class 2 is drawn by equation (19). This algorithm continues until age $I_2 - 1$. $S_{I_2 j}^1$ is simply set to the difference between s_j and the sum of $\boldsymbol{S}_{:j}^1$ for the younger age-classes.

For all i and j the formulas for $\tilde{\boldsymbol{s}}_{ij}^1$ and $\tilde{\boldsymbol{q}}_{ij}^1$ in (19) are given by the following formulaes:

$$\tilde{\boldsymbol{s}}_{ij}^1 = \left[S_{ij}^2, \ldots, \left(S_{ij}^2 + M_1 - 1\right)\right]^t, \quad \tilde{\boldsymbol{q}}_{ij}^1 = \left[\tilde{q}_{ij(1)}^1, \ldots, \tilde{q}_{ij(M_1)}^1\right]^t, \text{ where}$$

$$M_1 = \left(s_j - \sum_{l=1}^{i-1} S_{lj}^1\right) - \sum_{l=1}^{I} S_{lj}^2 + 1, \text{ and}$$

$$\tilde{q}_{ij(m)}^1 \propto f_{25}\left(S_{ij}^2|\boldsymbol{s}, \boldsymbol{S}^2, S_{ij}^1 = \tilde{s}_{ij(m)}^1\right) \cdot f_{14}\left(S_{ij}^1 = \tilde{s}_{ij(m)}^1|\bar{\boldsymbol{N}}, \boldsymbol{h}, \alpha\right), \quad 1 \le m \le M_1,$$

$$\tilde{\boldsymbol{s}}_{ij}^2 = \left[K_{ij}, \ldots, (K_{ij} + M_2 - 1)\right]^t, \quad \tilde{\boldsymbol{q}}_{ij}^2 = \left[\tilde{q}_{ij(1)}^2, \ldots, \tilde{q}_{ij(M_2)}^1\right]^t, \text{ where}$$

$$M_2 = S_{ij}^1 - S_{ij}^2 + 1, \text{ and}$$

$$\tilde{q}^2_{ij(m)} \propto f_{15}\left(K_{ij}|S^2_{ij} = \tilde{s}^2_{ij(m)}, \lambda_i\right) \cdot f_{15}\left(S^2_{ij} = \tilde{s}^2_{ij(m)}|S^1_{ij}, \gamma_j\right), \quad 1 \le m \le M_2$$

Conjugate prior distribution for α

Since the different elements of $\boldsymbol{S^1}$ are assumed to be independent the posterior distribution of $\boldsymbol{S^1}$ conditional on $\boldsymbol{\theta}^{-\alpha}, \boldsymbol{Y}_{obs}, \boldsymbol{Y}_{mis}$ might be written as shown below. Then the full conditional distribution for α is gamma distribution.

$$\pi\left(\boldsymbol{S^1}|\boldsymbol{\theta}^{-\alpha}, \boldsymbol{Y}_{obs}, \boldsymbol{Y}_{mis}\right) \propto_\alpha \left(\prod_{i=1}^{I}\prod_{j=1}^{J} f_{14}\left(S^1_{ij}|\alpha, \boldsymbol{N}, \boldsymbol{K}, \boldsymbol{H}\right)\right) \cdot \pi\left(\alpha|\boldsymbol{\Phi}\right)$$

$$= C \cdot e^{-\left(\phi^\alpha_2 + \sum_{i=1}^{I}\sum_{j=1}^{J}\left(\bar{N}_{ij}h_j\right)\right)\alpha} \cdot \alpha^{\left(\phi^\alpha_1 + \sum_{i=1}^{I}\sum_{j=1}^{J} S^1_{ij}\right)-1}$$

$$\text{where } C = \left(\prod_{i=1}^{I}\prod_{j=1}^{J}\frac{\left(\bar{N}_{ij}h_j\right)^{S_{ij}}}{S_{ij}!}\right) \cdot \frac{\phi_2^{\alpha\phi_1^\alpha}}{\Gamma(\phi_1^\alpha)}$$

Then:

$$\pi\left(\alpha|\boldsymbol{\theta}^{-\alpha}, \boldsymbol{Y}_{obs}, \boldsymbol{Y}_{mis}\right) \sim Gamma\left(\phi_1^\alpha + \sum_{j=1}^{J} s_j, \quad \phi_2^\alpha + \sum_{i=1}^{I}\sum_{j=1}^{J}\left(\bar{N}_{ij}h_j\right)\right)$$

Chapter 3

R-code

Script for running models

```
###########################################################################
#####  Modell for moose                                                 ##
#####  Lars Erik Gangsei, 17. april 2013.                               ##

rm(list=ls());setwd("C:/Users/Eier/Dropbox/LEGMasterThesis/R_Scripts");
setwd("C:/Users/LarsErik/Dropbox/LEGMasterThesis/R_Scripts")

##### Prior-hyperparameters                                             ##
Phi<-
list(DelA=c(10^(-4),10^(-4)),TauL=c(5,5),TauG=c(5,5),TauNu=c(2.96,50.41))
Iclass<-3;W<-10^4

##### Loading Data and initiating values                                ##
##### Setting working directory, loading packages and own-made functions ##

for (nm in list.files(paste(getwd(),"Funksjoner",sep="/")))
{source(paste(getwd(),"Funksjoner",nm,sep="/"))}

load(paste(getwd(),"Data/Data",sep="/"));
load(paste(getwd(),"Data/ValdAreal",sep="/"));
Vald<-as.character(ValdAreal$Valdnavn)
load(paste(getwd(),"Data/ThetaInit",sep="/"));
ThetaInit$alfa<-0.1;ThetaInit$alfaM<-0.1
ThetaInit$lambda<-ThetaInit$lambda[1:Iclass]
ThetaInit$lambdaM<-ThetaInit$lambdaM[1:Iclass]
load(paste(getwd(),"Data/Nstart",sep="/"))
```

```
##############################################################################
##### Model with real data                                                 ##

Yobs<-Data$Totalt;Yobs$H<-1000*Yobs$H/sum(ValdAreal$Areal)

##### Making a list, ResultsAll, for storing results                        ##
ResultsAll<-vector("list",2)

for(GG in c(1,0))
{
ResultsAll[[GG+1]]<-vector("list",2)
for(LL in c(1,0))
{
ResultsAll[[GG+1]][[LL+1]]<-vector("list",2)
for(NN in c(1,0))
{
##### Kohort_func "runs the model", and returns estimates                   ##
ResultsAll[[GG+1]][[LL+1]][[NN+1]]<-Kohort_func(Yobs=Yobs,
Iclass=Iclass,W=W,Phi=Phi,ThetaInit=ThetaInit,Nstart=Nstart,
Gsim=GG,Lsim=LL,Nsim=NN)

save(ResultsAll,file=paste(getwd(),"Resultat","ResultsAll.R",sep="/"))
}
}
}

##############################################################################
##### Modell with simulated data                                           ##

load(file=paste(getwd(),"SimData/Simdata",sep="/"))

##### Making a list, ResultsSim, for storing results                        ##
ResultsSim<-vector("list",2)

Yobs<-list(K=SimData$K_sim,KM=SimData$KM_sim,S=SimData$s_sim,
SM=SimData$sM_sim,SC=SimData$sC_sim,H=SimData$H_sim)

for(GG in c(1,0))
{
ResultsSim[[1]][[GG+1]]<-vector("list",2)
for(LL in c(1,0))
{
ResultsSim[[1]][[GG+1]][[LL+1]]<-vector("list",2)
```

```
for(NN in c(1,0))
{
##### Kohort_func "runs the model", and returns estimates            ##
ResultsSim[[1]][[GG+1]][[LL+1]][[NN+1]]<-Kohort_func(Yobs=Yobs,
Iclass=Iclass,W=W,Phi=Phi,ThetaInit=ThetaInit,Nstart=Nstart,
Gsim=GG,Lsim=LL,Nsim=NN,SIM=SimData)

print(c(GG,LL,NN))

save(ResultsSim,file=paste(getwd(),"Resultat","ResultsSim.R",sep="/"))
}
}
}


##### Testing the single elements of theta. Results unreported in the   ##
##### Master-thesis                                                     ##
ResultsSim[[2]]<-vector("list",5)
for (ii in 1:5)
{
ResultsSim[[2]][[ii]]<-Kohort_func(Yobs=Yobs,Iclass=Iclass,
W=W,Phi=Phi,ThetaInit=ThetaInit,Nstart=Nstart,
Gsim=ifelse(ii==1,1,2),Lsim=ifelse(ii==2,1,2),
Nsim=ifelse(ii==3,1,2),Asim=ifelse(ii==4,1,2),
NNsim=ifelse(ii==5,1,2),SIM=SimData)
save(ResultsSim,file=paste(getwd(),"Resultat","ResultsSim.R",sep="/"))
}

################################################################################
##### Modell with simulated data B                                      ##

load(file=paste(getwd(),"SimData/Simdata_B",sep="/"))

##### Making a list, ResultsSim_B, for storing results                  ##
ResultsSim_B<-vector("list",2)

Yobs<-list(K=SimData_B$K_sim,KM=SimData_B$KM_sim,S=SimData_B$s_sim,
SM=SimData_B$sM_sim,SC=SimData_B$sC_sim,H=SimData_B$H_sim)

for(GG in c(1,0))
{
ResultsSim_B[[1]][[GG+1]]<-vector("list",2)
for(LL in c(1,0))
{
```

```
ResultsSim_B[[1]][[GG+1]][[LL+1]]<-vector("list",2)
for(NN in c(1,0))
{
##### Kohort_func "runs the model", and returns estimates               ##
ResultsSim_B[[1]][[GG+1]][[LL+1]][[NN+1]]<-Kohort_func(Yobs=Yobs,
Iclass=Iclass,W=W,Phi=Phi,ThetaInit=ThetaInit,Nstart=Nstart,
Gsim=GG,Lsim=LL,Nsim=NN,SIM=SimData_B)

print(c(GG,LL,NN))

save(ResultsSim_B,file=paste(getwd(),"Resultat","ResultsSim_B.R",sep="/"))
}
}
}


##### Testing the single elements of theta. Results unreported in the    ##
##### Master-thesis                                                      ##
ResultsSim_B[[2]]<-vector("list",5)
for (ii in 1:5)
{
ResultsSim_B[[2]][[ii]]<-Kohort_func(Yobs=Yobs,Iclass=Iclass,
W=W,Phi=Phi,ThetaInit=ThetaInit,Nstart=Nstart,
Gsim=ifelse(ii==1,1,2),Lsim=ifelse(ii==2,1,2),
Nsim=ifelse(ii==3,1,2),Asim=ifelse(ii==4,1,2),
NNsim=ifelse(ii==5,1,2),SIM=SimData_B)
save(ResultsSim_B,file=paste(getwd(),"Resultat","ResultsSim_B.R",sep="/"))
}



################################################################################
##### Modell with different time-subsets                                 ##

##### Making a list, ResultsTime, for storing results. Only full model   ##
##### i.e. Gsim=Lsim=Nsim=1, is applied.                                 ##
ResultsTime<-vector("list",14)
names(ResultsTime)<-1999:2012

for(k in 1:14)
{
Yobs<-Data$Totalt;
for(l in 1:2){Yobs[[l]]<-Yobs[[l]][,k:(k+11)]}
for(l in 3:6){Yobs[[l]]<-Yobs[[l]][k:(k+11)]}
Yobs$H<-1000*Yobs$H/sum(ValdAreal$Areal)
```

```
ThetaInit_Time<-ThetaInit

for(l in c(2,6,8)){ThetaInit_Time[[l]]<-ThetaInit_Time[[l]][1:12]}
ThetaInit_Time[[4]]<-ThetaInit_Time[[4]][1:11]

##### Kohort_func "runs the model", and returns estimates               ##
ResultsTime[[k]]<-Kohort_func(Yobs=Yobs,Iclass=Iclass,W=W,Phi=Phi,
ThetaInit=ThetaInit_Time,Nstart=Nstart[,14:25],Gsim=1,Lsim=1,Nsim=1)

save(ResultsTime,file=paste(getwd(),"Resultat","ResultsTime.R",sep="/"))
}

#########################################################################
##### Modell with different geographical-subsets                        ##

##### Making a list, ResultsVald, for storing results. Only full model   ##
##### i.e. Gsim=Lsim=Nsim=1, is applied.                                 ##

ResultsVald<-vector("list",22);names(ResultsVald)<-ValdAreal$Valdnavn

for(k in 1:22)
{
Yobs<-
mapply("-",Data$Totalt,Data[[which(names(ResultsVald[k])==names(Data))]]);

Yobs$H<-1000*Yobs$H/(sum(ValdAreal$Areal)
-ValdAreal$Areal[which(names(ResultsVald[k])==ValdAreal$Valdnavn)])

##### Kohort_func "runs the model", and returns estimates               ##
ResultsVald[[k]]<-Kohort_func(Yobs=Yobs,Iclass=Iclass,W=W,Phi=Phi,
ThetaInit=ThetaInit,Nstart=Nstart,Gsim=1,Lsim=1,Nsim=1)

save(ResultsVald,file=paste(getwd(),"Resultat","ResultsVald.R",sep="/"))
}#End for
```

Model-script

```
############################################################################
######   Main model for Cohort-analysis of moose                       ##
######   Lars Erik Gangsei, 26.04.2013                                 ##

Kohort_func<-function(Yobs,Iclass,W,Phi,ThetaInit,Nstart,Gsim=1,Lsim=1,
Nsim=1,Asim=1,NNsim=1,SIM=NULL)

{#start function
attach(Yobs);attach(Phi);Imax<-dim(K)[1];ImaxM<-dim(KM)[1];Jmax<-dim(K)[2]

###### Matrixes for storing of simulatet values and data-augmented values##

lambda.<-matrix(NA,W,ifelse(Lsim==1,Iclass,1));
gamma.<-matrix(NA,W,ifelse(Gsim==1,Jmax,1));alfa.<-rep(NA,W);
nu.<-matrix(NA,W,ifelse(Nsim==1,(Jmax-1),1));
lambdaM.<-matrix(NA,W,ifelse(Lsim==1,Iclass,1));
gammaM.<-matrix(NA,W,ifelse(Gsim==1,Jmax,1));alfaM.<-rep(NA,W);
gammaC.<-matrix(NA,W,ifelse(Gsim==1,Jmax,1));N.<-matrix(NA,Imax*W,Jmax);
NM.<-matrix(NA,ImaxM*W,Jmax);S1.<-matrix(NA,(Iclass+1)*W,Jmax);
S1M.<-matrix(NA,(Iclass+1)*W,Jmax);S2.<-matrix(NA,(Iclass+1)*W,Jmax)
S2M.<-matrix(NA,(Iclass+1)*W,Jmax);DIC.<-matrix(NA,W,2)

############################################################################
##### Cumulative values for shoot moose, used in calculations         ##
cumK<-K;cumKM<-KM
for(i in (Imax-1):1){cumK[i,1:(Jmax-1)]<-cumK[i+1,2:Jmax]+K[i,1:(Jmax-1)]};
for(i in (ImaxM-1):1){cumKM[i,1:(Jmax-1)]<-cumKM[i+1,2:Jmax]+KM[i,1:(Jmax-1)]}

##### Cohort-indexes (l) put up in a matrix. Used in calculations, for  ##
##### instance in data-augmentation of N.                              ##
Index<-NULL;IndexM<-NULL;
for(j in 1:Jmax)
{Index<-cbind(Index,c((Jmax+1-j):(Jmax+Imax-1),rep(99,Jmax-j)))}
for(j in 1:Jmax)
{IndexM<-cbind(IndexM,c((Jmax+1-j):(Jmax+ImaxM-1),rep(99,Jmax-j)))}

############################################################################
##### MCMC algorithm                                                   ##

############################################################################
##### Step 1: Initiation of parameters                                 ##
```

```
attach(ThetaInit)
alfa.[1]<-alfa;alfaM.[1]<-alfaM

###############################################################################
##### Step 2: Start value for N, Nbar, D and S2                            ##
N<-Nstart[1:Imax,]+cumK;Nbar<--K/(log(1-K/N));Nbar[K==N]<-N[K==N];
Nbar[K==0]<-N[K==0];

NM<-Nstart[1:ImaxM,]+cumKM;NbarM<--KM/(log(1-KM/NM));
NbarM[KM==NM]<-NM[KM==NM];NbarM[KM==0]<-NM[KM==0];

S2<-rbind(K[1:Iclass,],colSums(K[(Iclass+1):Imax,]));
S2M<-rbind(KM[1:Iclass,],colSums(KM[(Iclass+1):ImaxM,]))

S1<-matrix(NA,Iclass+1,Jmax);S1M<-matrix(NA,Iclass+1,Jmax)

###############################################################################
##### Step 3: Initiate the model                                           ##
w=1
while(w<=W)
{#start while

###############################################################################
##### Step 4:Data-augmentasjon                                             ##

##### a) Data-augmentasjon N, "data.augment.NC" is a custom function       ##
#####    returning values for N and D.                                     ##
#####    NNsim = 1 in all situations except the situation when single      ##
#####    elements from theta are tested, results unrep. in the Master-th.##


N_D<-data.augment.NC(Index=Index,cumK=cumK,K=K,H=H,D=D,alfa=alfa,
alfaC=alfa,gamma=gamma,gammaC=gammaC,lambda=lambda,nu=nu,
Iclass=Iclass,Imax=Imax,Jmax=Jmax)

N_D_M<-data.augment.NC(Index=IndexM,cumK=cumKM,K=KM,H=H,D=DM,alfa=alfaM,
alfaC=alfa,gamma=gammaM,gammaC=gammaC,lambda=lambdaM,nu=nu,
Iclass=Iclass,Imax=ImaxM,Jmax=Jmax)

if(NNsim==1){N<-N_D$N;D<-N_D$D;NM<-N_D_M$N;DM<-N_D_M$D}

else{N<-SIM$N_sim;D<-SIM$D_sim;NM<-SIM$NM_sim;DM<-SIM$DM_sim}
```

```
##### Update Nbar for every step.                                        ##
Nbar<--K/(log(1-K/N));Nbar[K==N]<-N[K==N];Nbar[K==0]<-N[K==0]
NbarM<--KM/(log(1-KM/NM));NbarM[KM==NM]<-NM[KM==NM];NbarM[KM==0]<-NM[KM==0]


##### b) Data-augmentasjon S1 (categorical based on gamma and lambda).   ##
#####     "data.augment.s1" is a custom function returning value for S1   ##
#####     for adult moose.                                               ##

S1[2:(Iclass+1),]<-data.augment.S1(Nbar=Nbar[2:Imax,],
s=S,S2=S2[2:(Iclass+1),],gamma=gamma,lambda=lambda,Iclass=Iclass,
Imax=Imax-1,Jmax=Jmax)

S1M[2:(Iclass+1),]<-data.augment.S1(Nbar=NbarM[2:ImaxM,],
s=SM,S2=S2M[2:(Iclass+1),],gamma=gammaM,lambda=lambdaM,
Iclass=Iclass,Imax=ImaxM-1,Jmax=Jmax)

##### Calves.                                                            ##
S1[1,]<-rbinom(Jmax,SC-K[1,]-KM[1,],Nbar[1,]/(Nbar[1,]+NbarM[1,]))+K[1,];
S1M[1,]<-SC-S1[1,]

##### c) Data-augmentasjon S2 (categorical based on gamma and lambda).   ##
#####     "data.augment.s2B" is a custom function returning value for S2  ##
#####     for adult moose.                                               ##

#c) Data-augmentasjon S2, (categorical based on gamma and lambda)
S2[2:(Iclass+1),]<-data.augment.S2B(K=K[2:Imax,],S1=S1[2:(Iclass+1),],
lambda=lambda,gamma=gamma,Iclass=Iclass,Imax=Imax-1,Jmax=Jmax);

S2M[2:(Iclass+1),]<-data.augment.S2B(K=KM[2:ImaxM,],S1=S1M[2:(Iclass+1),],
lambda=lambdaM,gamma=gammaM,Iclass=Iclass,Imax=ImaxM-1,Jmax=Jmax);

##### Calves.                                                            ##
S2[1,]<-K[1,];S2M[1,]<-KM[1,]

#########################################################################
##### Step 5: Gibbs sampling                                             ##

##### a) lambda                                                          ##

##### Full length of lambda:                                             ##
if(Lsim==1)
{
Hyp1L<-c(TauL[1]+rowSums(K[2:Iclass,]),
```

```
TauL[1]+sum(K[(Iclass+1):Imax,]));
Hyp2L<-c(TauL[2]+rowSums(S2[2:Iclass,])-rowSums(K[2:Iclass,]),
TauL[2]+sum(S2[(Iclass+1),])-sum(K[(Iclass+1):Imax,]))

Hyp1LM<-c(TauL[1]+rowSums(KM[2:Iclass,]),
TauL[1]+sum(KM[(Iclass+1):ImaxM,]));
Hyp2LM<-c(TauL[2]+rowSums(S2M[2:Iclass,])-rowSums(KM[2:Iclass,]),
TauL[2]+sum(S2M[(Iclass+1),])-sum(KM[(Iclass+1):ImaxM,]))

lambda<-rbeta(Iclass,Hyp1L,Hyp2L);lambdaM<-rbeta(Iclass,Hyp1LM,Hyp2LM)
}

##### Setting lambda equal to true value. Unrep. results, simulated pop. ##
else if(Lsim==2)
{
lambda<-SIM$lambda_sim;lambdaM<-SIM$lambdaM_sim;
}

##### Reduced length of lambda:                                       ##
else
{
Hyp1L<-TauL[1]+sum(K[2:Imax,]);
Hyp2L<-TauL[2]+sum(S2[2:(Iclass+1),])-sum(K[2:Imax,])

Hyp1LM<-TauL[1]+sum(KM[2:ImaxM,]);
Hyp2LM<-TauL[2]+sum(S2M[2:(Iclass+1),])-sum(KM[2:ImaxM,])

lambda_val<-rbeta(1,Hyp1L,Hyp2L);lambda<-rep(lambda_val,Iclass);
lambdaM_val<-rbeta(1,Hyp1LM,Hyp2LM);lambdaM<-rep(lambdaM_val,Iclass)
}

##### b) gamma                                                        ##

##### Full length of gamma:                                           ##
if(Gsim==1)
{
Hyp1G<-TauG[1]+colSums(S2[2:(Iclass+1),]);
Hyp2G<-TauG[2]+S-colSums(S2[2:(Iclass+1),])

Hyp1GM<-TauG[1]+colSums(S2M[2:(Iclass+1),]);
Hyp2GM<-TauG[2]+SM-colSums(S2M[2:(Iclass+1),])

Hyp1GC<-TauG[1]+(S2[1,]+S2M[1,]);Hyp2GC<-TauG[2]+SC-(S2[1,]+S2M[1,])
```

```
gamma<-rbeta(Jmax,Hyp1G,Hyp2G);
gammaM<-rbeta(Jmax,Hyp1GM,Hyp2GM);
gammaC<-rbeta(Jmax,Hyp1GC,Hyp2GC)
}

##### Setting gamma  equal to true value. Unrep. results, simulated pop. ##
else if(Gsim==2)
{
gamma<-SIM$gamma_sim;gammaM<-SIM$gammaM_sim;gammaC<-SIM$gammaC_sim
}

##### Reduced length of gamma:                                        ##
else
{
Hyp1G<-TauG[1]+sum(S2[2:(Iclass+1),]);
Hyp2G<-TauG[2]+sum(S)-sum(S2[2:(Iclass+1),])

Hyp1GM<-TauG[1]+sum(S2M[2:(Iclass+1),]);
Hyp2GM<-TauG[2]+sum(SM)-sum(S2M[2:(Iclass+1),])

Hyp1GC<-TauG[1]+sum(S2[1,]+S2M[1,]);
Hyp2GC<-TauG[2]+sum(SC)-sum(S2[1,]+S2M[1,])

gamma_val<-rbeta(1,Hyp1G,Hyp2G);gamma<-rep(gamma_val,Jmax);
gammaM_val<-rbeta(1,Hyp1GM,Hyp2GM);gammaM<-rep(gammaM_val,Jmax);
gammaC_val<-rbeta(1,Hyp1GC,Hyp2GC);gammaC<-rep(gammaC_val,Jmax)
}

##### c)nu                                                            ##

##### Full length of nu:                                              ##
if(Nsim==1)
{
Hyp1Nu<-TauNu[1]+colSums(D)+colSums(DM);
Hyp2Nu<-(TauNu[2]+colSums((N-K)[1:(Imax-1),1:(Jmax-1)])
+colSums((NM-KM)[1:(ImaxM-1),1:(Jmax-1)])-colSums(D)-colSums(DM))

nu<-rbeta((Jmax-1),Hyp1Nu,Hyp2Nu)
}

##### Setting nu  equal to true value. Unrep. results, simulated pop.    ##
else if(Nsim==2)
```

```
{
nu<-SIM$nu_sim;
}

##### Reduced length of nu:                                                    ##
else
{
Hyp1Nu<-Jmax*TauNu[1]+sum(D)+sum(DM);
Hyp2Nu<-(Jmax*TauNu[2]+sum((N-K)[1:(Imax-1),1:(Jmax-1)])
+sum((NM-KM)[1:(ImaxM-1),1:(Jmax-1)])-sum(D)-sum(DM))

nu_val<-rbeta(1,Hyp1Nu,Hyp2Nu);nu<-rep(nu_val,Jmax-1)
}

##### d) alfa                                                                   ##

if(Asim==1)
{

Hyp1A<-DelA[1]+sum(S1)+sum(S1M[1,]);
Hyp2A<-(DelA[2]+sum(Nbar*matrix(rep(H,Imax),Imax,Jmax,byrow=TRUE))
+sum(NbarM[1,]*H))

Hyp1AM<-DelA[1]+sum(S1M)-sum(S1M[1,]);
Hyp2AM<-(DelA[2]+sum(NbarM*matrix(rep(H,ImaxM),ImaxM,Jmax,byrow=TRUE))
-sum(NbarM[1,]*H))

alfa<-rgamma(1,Hyp1A,Hyp2A);alfaM<-rgamma(1,Hyp1AM,Hyp2AM)
}

##### Setting alpha equal to true value. Unrep. results, simulated pop.  ##
else
{
alfa<-SIM$alfa_sim;alfaM<-SIM$alfaM_sim;
}

#################################################################################
##### Storing data                                                             ##

if(Lsim==1){lambda.[w,]<-lambda;lambdaM.[w,]<-lambdaM}
else{lambda.[w,]<-lambda[1];lambdaM.[w,]<-lambdaM[1]};

if(Gsim==1){gamma.[w,]<-gamma;gammaM.[w,]<-gammaM;gammaC.[w,]<-gammaC}
```

```
else{gamma.[w,]<-gamma[1];gammaM.[w,]<-gammaM[1];gammaC.[w,]<-gammaC[1]}

if(Nsim==1){nu.[w,]<-nu}else{nu.[w,]<-nu[1]}

alfa.[w]<-alfa;alfaM.[w]<-alfaM;
N.[((w-1)*Imax+1):(w*Imax),]<-N;
NM.[((w-1)*ImaxM+1):(w*ImaxM),]<-NM;
S1.[((w-1)*(Iclass+1)+1):(w*(Iclass+1)),]<-S1;
S1M.[((w-1)*(Iclass+1)+1):(w*(Iclass+1)),]<-S1M
S2.[((w-1)*(Iclass+1)+1):(w*(Iclass+1)),]<-S2;
S2M.[((w-1)*(Iclass+1)+1):(w*(Iclass+1)),]<-S2M

##########################################################################
##### Step 6: Increase iteration (w) by 1.                            ##

w=w+1;print(w)

if(w%%100==0){plot(alfa.,type="l");if(!is.null(SIM)){abline(h=SIM$alfa_sim)}}

}#End while

##########################################################################
##### Storing result                                                  ##
ResSub<-list(lambda.=lambda.,lambdaM.=lambdaM.,gamma.=gamma.,
gammaM.=gammaM.,gammaC.=gammaC.,nu.=nu.,alfa.=alfa.,alfaM.=alfaM.,
N.=N.,NM.=NM.,S1.=S1.,S1M.=S1M.,S2.=S2.,S2M.=S2M.,DIC.=DIC.)

return(ResSub)
}#End function
```

Function: Data-augmentation *N*

```
##############################################################################
##### Funksjon for returning Data-augmented values for N and D         ##
#####                                                                  ##
##### The data are split into different cohorts using functions "mapply" ##
##### and the matrix "Index", where cohort "l" (cohort number) are given ##
##### in matrix form.                                                  ##
#####                                                                  ##
##### Lars Erik Gangsei 27.12.2012                                     ##

data.augment.NC<-function(Index,cumK,K,H,D,alfa,alfaC,gamma,gammaC,lambda,
nu,Iclass,Imax,Jmax)


{#Start function

##### Number of Cohorts, Lmax                                          ##

Lmax<-Imax+Jmax-1

##### pp: probability of being shot, pp-inv: probability of survive    ##

pp<-rbind(1-exp(-alfaC*gammaC*H),
1-exp(-alfa*(c(lambda,rep(lambda[Iclass],Lmax-Iclass-1))%*%t(gamma*H))))
pp_inv_list<-tapply(1-pp,Index,c)[1:Lmax]
pp_list<-tapply(pp,Index,c)[1:Lmax]

##### nu: natural mortality rate                                       ##
nu_mat<-rbind(rep(0,Jmax),matrix(c(0,nu),Lmax-1,Jmax,byrow=TRUE))
nu_inv_list<-c(tapply(1-nu_mat,Index,c)[1:Lmax])

nu_mat_B<-rbind(matrix(c(nu,1),Lmax-1,Jmax,byrow=TRUE),rep(1,Jmax))
nu_list<-c(tapply(nu_mat_B,Index,c)[1:Lmax])

###### Probability of surviving from hunting end to next years hunt. end ##
nu_pp_vec<-mapply("*",nu_inv_list,pp_inv_list)

###### Probability of being alive post-hunt                            ##
pp_cum_post_surv_list<-sapply(nu_pp_vec,cumprod)

###### Probability of being alive pre-hunt                             ##
pp_cum_pre_surv_list<-mapply("/",pp_cum_post_surv_list,pp_inv_list)
```

```
###### Unconditional probability of being shot a given year          ##
pp_t_list<-mapply("*",pp_cum_pre_surv_list,pp_list)

###### Cumulative probability of being shot                         ##
cum_pp_t_list<-lapply(pp_t_list,sum)

###### Number of shot in every cohort     ##
cumK_list<-tapply(K,Index[1:Imax,],sum)[1:Lmax]

###### Number not shot in every cohort                              ##
LEG_rnbin<-function(ss,pp)
{if(ss>0){res<-rnbinom(1,size=ss,prob=pp)}
else{res<-rgeom(1,pp)};return(res)}

N0_cum_list<-mapply(LEG_rnbin,ss=cumK_list,pp=cum_pp_t_list)

###### Unconditional natural mortality rate                         ##
nu_t_list<-mapply("*",pp_cum_post_surv_list,nu_list)

###### Distribution in natural dead and survivors                   ##
DD_N0_list<-mapply(rmultinom,n=1,size=N0_cum_list,prob=nu_t_list)

###### Finding D og N                                               ##
DD<-matrix(NA,Lmax,Jmax)
DD[1,Jmax]<-DD_N0_list[[1]]

for(l in 2:Lmax)
{diag(DD[max(1,l+1-Jmax):l,max(1,Jmax+1-l):Jmax])<-DD_N0_list[[l]]}

for (j in 1:(Jmax-1)){DD[Imax,j]<-sum(diag(DD[Imax:(Lmax+1-j),j:Jmax]))}

DD<-DD[1:Imax,]

cumDD<-DD
for(i in (Imax-1):1)
{cumDD[i,1:(Jmax-1)]<-cumDD[i+1,2:Jmax]+DD[i,1:(Jmax-1)]}

NN<-cumK+cumDD

res<-list(N=NN,D=DD[1:(Imax-1),1:(Jmax-1)])

return(res)
}#End function
```

Function: Data-augmentation S^1

```
##########################################################################
##### Function for returning S1 from data-augmentation             ##
#####                                                              ##
##### The function is based on categorical distribution            ##
#####                                                              ##
##### Lars Erik Gangsei 16.02.2013                                 ##

data.augment.S1<-function(Nbar,S2,s,gamma,lambda,Iclass,Imax,Jmax)

{#Start function
S1<-matrix(NA,Iclass,Jmax)

for (j in 1:Jmax)
{#start for
for (i in 1:(Iclass-1))
{#start i
X<-0:ifelse(i>1,s[j]-sum(S1[1:(i-1),j])-
sum(S2[i:Iclass,j]),s[j]-sum(S2[,j]))

if(max(X)==0){S1[i:(Iclass-1),j]<-S2[i:(Iclass-1),j];break()}
else
{#start else
propS2<-dhyper(x=S2[i,j],m=X+S2[i,j],n=s[j]-
ifelse(i>1,sum(S1[1:(i-1),j]),0)-X-S2[i,j],k=sum(S2[i:Iclass,j]))
propNbar<-dbinom(X,max(X),Nbar[i,j]/sum(Nbar[i:Imax,j]))
PropT<-propS2*propNbar;PropT[PropT=="NaN"]=0;
if(sum(PropT)==0){PropT<-rep(1/length(PropT),length(PropT))}
S1[i,j]<-sample(X,size=1,prob=PropT)+S2[i,j]
}#end else
}#end for (i)
}#end for (j)

S1[Iclass,]<-s-colSums(S1[1:(Iclass-1),])
return(S1)
}#End function
```

Function: Data-augmentation S^2

```
###########################################################################
##### Function for returning S2 from data-augmentation              ##
#####                                                               ##
##### Function is based on categorical distribution. The function   ##
#####  "mapply" is applied to "estimate all elements at once".      ##
#####                                                               ##
##### Lars Erik Gangsei 16.02.2013                                  ##

data.augment.S2B<-function(K,S1,lambda,gamma,Iclass,Imax,Jmax)

{#Start function

KK<-rbind(K[1:(Iclass-1),],colSums(K[Iclass:Imax,]))

X<-mapply(seq,KK,S1)
PropG<-mapply(dbinom,X,as.list(S1),as.list(matrix(rep(gamma,Iclass),
Iclass,Jmax,byrow=TRUE)))
PropL<-mapply(dbinom,as.list(KK),X,as.list(matrix(rep(lambda,Jmax),
Iclass,Jmax,byrow=FALSE)))
Prop<-mapply("*",PropG,PropL)
S2<-sapply(X,sum);
S2[mapply(length,X)>1]<-mapply(sample,X[mapply(length,X)>1],
prob=Prop[mapply(length,X)>1],size=1)

S2<-matrix(S2,Iclass,Jmax,byrow=FALSE)

return(S2)

}#End function
```

Function: DIC.1

```
#############################################################################
##### Function returning values bar(DEV), p_DEV, DEV_bar(th) and DIC    ##
##### for a list containing Monte Carlo realizations for parameter      ##
##### estimates and single realizations for DEV for every iteration. The ##
##### names for parameter estimates must be proper. The algorithm is    ##
##### restricted to use on the moose population model only. The         ##
##### Burnin-period should be eliminated from Theta.                     ##
#####                                                                    ##
##### Lars Erik gangsei 23.04.2013                                       ##

DIC_function_theta<-function(GG=1,LL=1,NN=1,Theta,Data,Iclass,hh)
{
mm<-length(Theta$alfa.);Imax<-dim(Data$K)[1];
ImaxM<-dim(Data$KM)[1];Jmax<-dim(Data$K)[2]



##### Calculating the deviance for each element in the output using the  ##
##### custom function "DIC_func_1_B" which returns the deviance for the  ##
##### moose model as function of data (Y_obs), missing data             ##
##### (Y_mis, i.e. N) and parameters (theta).                           ##

Bar_Dev_mat<-rep(NA,mm)

for(xx in 1:mm)
{
Bar_Dev_mat[xx]<-DIC_func_1_B(KK=Data$K,
KKM=Data$KM,NN=Theta$N.[((xx-1)*Imax+1):(xx*Imax),],
NNM=Theta$NM.[((xx-1)*ImaxM+1):(xx*ImaxM),],
ss=Data$S,ssM=Data$SM,ssC=Data$SC,
if(GG==1){gg=Theta$gamma.[xx,]}
else{gg=rep(Theta$gamma.[xx],Jmax)},

if(GG==1){ggM=Theta$gammaM.[xx,]}
else{ggM=rep(Theta$gammaM.[xx],Jmax)},

if(GG==1){ggC=Theta$gammaC.[xx,]}
else{ggC=rep(Theta$gammaC.[xx],Jmax)},

if(LL==1){ll=Theta$lambda.[xx,]}
else{ll=rep(Theta$lambda.[xx],Iclass)},
```

```
if(LL==1){llM=Theta$lambdaM.[xx,]}
else{llM=rep(Theta$lambdaM.[xx],Iclass)},

if(NN==1){nn=Theta$nu.[xx,]}
else{nn=rep(Theta$nu.[xx],Jmax-1)},

aa=Theta$alfa.[xx],aaM=Theta$alfaM.[xx],hh=hh)
}

##### Calculating "bar_D", a central part of DIC                  ##
Bar_Dev<-mean(Bar_Dev_mat)


##### Calculating the deviance for average parameter/missing data output.##
##### First calculating average values for parameters and miss. data (N).##
if(LL==1)
{Bar_lambda<-colMeans(Theta$lambda.)}
else{Bar_lambda<-rep(mean(Theta$lambda.),Iclass)}

if(LL==1)
{Bar_lambdaM<-colMeans(Theta$lambdaM.)}
else{Bar_lambdaM<-rep(mean(Theta$lambdaM.),Iclass)}

if(GG==1)
{Bar_gamma<-colMeans(Theta$gamma.)}
else{Bar_gamma<-rep(mean(Theta$gamma.),Jmax)}

if(GG==1)
{Bar_gammaM<-colMeans(Theta$gammaM.)}
else{Bar_gammaM<-rep(mean(Theta$gammaM.),Jmax)}

if(GG==1)
{Bar_gammaC<-colMeans(Theta$gammaC.)}
else{Bar_gammaC<-rep(mean(Theta$gammaC.),Jmax)}

if(NN==1){Bar_nu<-colMeans(Theta$nu.)}
else{Bar_nu<-rep(mean(Theta$nu.),Jmax-1)}

Bar_alfa<-mean(Theta$alfa.)
Bar_alfaM<-mean(Theta$alfaM.)

Bar_N<-NULL;Bar_NM<-NULL;
for(i in 1:Imax)
```

```
{Bar_N<-rbind(Bar_N,colMeans(Theta$N.[seq(i,mm*Imax,by=Imax),]))}
for(i in 1:ImaxM)
{Bar_NM<-rbind(Bar_NM,colMeans(Theta$NM.[seq(i,mm*ImaxM,by=ImaxM),]))}

##### Bar_dev(bar_theta), a central part of DIC                         ##
Dev_Theta_Bar<-
DIC_func_1_B(KK=Data$K,KKM=Data$KM,NN=round(Bar_N),NNM=round(Bar_NM),
ss=Data$S,ssM=Data$SM,ssC=Data$SC,gg=Bar_gamma,ggM=Bar_gammaM,
ggC=Bar_gammaC,ll=Bar_lambda,llM=Bar_lambdaM,nn=Bar_nu,
aa=Bar_alfa,aaM=Bar_alfaM,hh=hh)

##### P_d and returning results                                         ##
p_Dev<-Bar_Dev-Dev_Theta_Bar; DIC<-Bar_Dev+p_Dev

return(c(GG,LL,NN,Bar_Dev,Dev_Theta_Bar,p_Dev,DIC))
}
```

Function: DIC.2

```
################################################################################
##### Functions returning the deviance based on data, missing data and   ##
##### parameters.                                                         ##
#####                                                                     ##
##### Lars Erik Gangsei 20.04.2013                                        ##

DIC_func_1_B<-function(KK,KKM,NN,NNM,ss,ssM,ssC,#SS1,SS1M,SS2,SS2M,
gg,ggM,ggC,ll,llM,nn,aa,aaM,hh)
{
NNBB<--KK/(log(1-KK/NN));NNBB[KK==NN]<-NN[KK==NN];
NNBB[KK==0]<-NN[KK==0];NNMMBB<--KKM/(log(1-KKM/NNM));
NNMMBB[KKM==NNM]<-NNM[KKM==NNM];NNMMBB[KKM==0]<-NNM[KKM==0]

II<-dim(KK)[1];JJ<-dim(KK)[2];IICC<-length(lambda)+1;IIM<-dim(KKM)[1]

Like_s<-
(sum(dpois(ss,colSums(NNBB[2:II,])*hh*aa,log=TRUE))
+sum(dpois(ssM,colSums(NNMMBB[2:IIM,])*hh*aaM,log=TRUE))
+sum(dpois(ssC,(NNBB[1,]+NNMMBB[1,])*hh*aa,log=TRUE)))

ppF<-1-exp(-aa*(c(ll,rep(ll[IICC-1],II-IICC))%*%t(gg*hh)))
ppM<-1-exp(-aaM*(c(llM,rep(llM[IICC-1],IIM-IICC))%*%t(ggM*hh)))
ppC<-1-exp(-aa*ggC*hh)

Like_p<-
(sum(dbinom(KK[2:II,],NN[2:II,],ppF,log=TRUE))
+sum(dbinom(KKM[2:IIM,],NNM[2:IIM,],ppM,log=TRUE))
+sum(dbinom(KK[1,]+KKM[1,],NN[1,]+NNM[1,],ppC,log=TRUE)))

return(-2*(Like_p+Like_s))
}
```

Script for simulating data

```
################################################################################
##### Script returning a simulated moose populations                      ##
##### The script has to be run 2 times, one time for each simulation.     ##
##### Names for results has to be changed manually then.                  ##
#####                                                                     ##
##### Lars Erik Gangsei 20.04.2013                                        ##


##### Cleaning working space, settting working directory.                 ##
rm(list=ls());
setwd("C:/Users/Eier/Dropbox/LEGMasterThesis/R_Scripts")
setwd("C:/Users/LarsErik/Dropbox/LEGMasterThesis/R_Scripts")

##### Loading results from the real data.                                 ##
load(paste(getwd(),"Resultat/ResultsAll.R",sep="/"))
load(paste(getwd(),"Data/ValdAreal",sep="/"));
load(paste(getwd(),"Data/Data",sep="/"));

##### Setting parameters                                                  ##
Par<-list(Jmax=25,Imax=30,ImaxM=19,Iclass=3,Burnin=2500,W=10^4);attach(Par)

##### Calculating mean numbers for "startng populatins", i.e calves and   ##
##### first year. Then calculating mean parametervalues.                  ##

Omega_N<-
round(tapply(ResultsAll[[2]][[2]][[1]]$N.[(Burnin*Imax+1):(W*Imax),1],
rep(1:Imax,W-Burnin),mean))[2:Imax]

Omega_NM<-
round(tapply(ResultsAll[[2]][[2]][[1]]$NM.[(Burnin*ImaxM+1):(W*ImaxM),1],
rep(1:ImaxM,W-Burnin),mean))[2:ImaxM]
Omega_N2<-round(
colMeans(ResultsAll[[2]][[2]][[1]]$
N.[seq(Burnin*Imax+1,W*Imax,by=Imax),]))
Omega_NM2<-round(
colMeans(ResultsAll[[2]][[2]][[1]]$
NM.[seq(Burnin*ImaxM+1,W*ImaxM,by=ImaxM),]))


hh_sim<-Yobs<-1000*Data$Totalt$H/sum(ValdAreal$Areal)
```

```
ll_sim<-colMeans(ResultsAll[[2]][[2]][[1]]$lambda.[(Burnin+1):W,])
llM_sim<-colMeans(ResultsAll[[2]][[2]][[1]]$lambdaM.[(Burnin+1):W,])

gg_sim<-colMeans(ResultsAll[[2]][[2]][[1]]$gamma.[(Burnin+1):W,])
ggM_sim<-colMeans(ResultsAll[[2]][[2]][[1]]$gammaM.[(Burnin+1):W,])
ggC_sim<-colMeans(ResultsAll[[2]][[2]][[1]]$gammaC.[(Burnin+1):W,])

aa_sim<-mean(ResultsAll[[2]][[2]][[1]]$alfa.[(Burnin+1):W])
aaM_sim<-mean(ResultsAll[[2]][[2]][[1]]$alfaM.[(Burnin+1):W])

##### If full length of nu:                                              ##
#nn_sim<-colMeans(ResultsAll[[2]][[2]][[1]]$nu.[(Burnin+1):W,])

##### If reduced length of nu:                                           ##
nn_sim<-rep(mean(ResultsAll[[2]][[2]][[1]]$nu.[(Burnin+1):W]),Jmax-1)

##### Calculating p based on parameters and h:                           ##
pp<-rbind(1-exp(-aa_sim*ggC_sim*hh_sim),
1-exp(-aa_sim*(c(ll_sim,
rep(ll_sim[Iclass],Imax-Iclass-1))%*%t(gg_sim*hh_sim))))
ppM<-rbind(1-exp(-aa_sim*ggC_sim*hh_sim),
1-exp(-aaM_sim*(c(llM_sim,
rep(llM_sim[Iclass],ImaxM-Iclass-1))%*%t(ggM_sim*hh_sim))))

##### Setting up matrixes for simulaing data:                            ##
N_sim<-matrix(NA,Imax,Jmax);N_sim[2:Imax,1]<-Omega_N
N_sim[1,]<-Omega_N2
NM_sim<-matrix(NA,ImaxM,Jmax);NM_sim[2:ImaxM,1]<-Omega_NM;
NM_sim[1,]<-Omega_NM2

K_sim<-matrix(NA,Imax,Jmax);D_sim<-matrix(NA,Imax-1,Jmax-1)
KM_sim<-matrix(NA,ImaxM,Jmax);DM_sim<-matrix(NA,ImaxM-1,Jmax-1)

##### Simulating process, N and K:                                       ##
for(j in 1:Jmax)
{#start for
K_sim[,j]<-rbinom(Imax,N_sim[,j],pp[,j])
KM_sim[,j]<-rbinom(ImaxM,NM_sim[,j],ppM[,j])

if(j==Jmax){break()}
D_sim[,j]<-rbinom((Imax-1),(N_sim[-Imax,j]-K_sim[-Imax,j]),nn_sim[j])
DM_sim[,j]<-rbinom((ImaxM-1),
(NM_sim[-ImaxM,j]-KM_sim[-ImaxM,j]),nn_sim[j])
```

```
N_sim[2:Imax,(j+1)]<-N_sim[1:(Imax-1),j]-K_sim[1:(Imax-1),j]-D_sim[,j]
NM_sim[2:ImaxM,(j+1)]<-
NM_sim[1:(ImaxM-1),j]-KM_sim[1:(ImaxM-1),j]-DM_sim[,j]
}#end for

##### Simulating process, s and S:                                          ##
N_bar<--K_sim/(log(1-K_sim/N_sim));
N_bar[K_sim==N_sim]<-N_sim[K_sim==N_sim];
N_bar[K_sim==0]<-N_sim[K_sim==0];
NM_bar<--KM_sim/(log(1-KM_sim/NM_sim));
NM_bar[KM_sim==NM_sim]<-NM_sim[KM_sim==NM_sim];
NM_bar[KM_sim==0]<-NM_sim[KM_sim==0];

S1_sim<-matrix(
rpois(Imax*Jmax,(aa_sim*matrix(hh_sim,Imax,Jmax,byrow=TRUE)*N_bar)),
Imax,Jmax,byrow=FALSE);ss_sim<-colSums(S1_sim[2:Imax,])

S1M_sim<-rbind(rpois(Jmax,aa_sim*hh_sim*NM_bar[1,]),
matrix(rpois((ImaxM-1)*Jmax,(aaM_sim*matrix(hh_sim,(ImaxM-1),
Jmax,byrow=TRUE)*NM_bar[2:ImaxM,])),
(ImaxM-1),Jmax,byrow=FALSE));ssM_sim<-colSums(S1M_sim[2:ImaxM,])

ssC_sim<-S1_sim[1,]+S1M_sim[1,]

##### Results returned:                                                     ##
SimData_B<-list(lambda_sim=ll_sim,lambdaM_sim=llM_sim,gamma_sim=gg_sim,
gammaM_sim=ggM_sim,gammaC_sim=ggC_sim,alfa_sim=aa_sim,alfaM_sim=aaM_sim,
H_sim=hh_sim,nu_sim=nn_sim,N_sim=N_sim,NM_sim=NM_sim,s_sim=ss_sim,
sM_sim=ssM_sim,sC_sim=ssC_sim,K_sim=K_sim,KM_sim=KM_sim,D_sim=D_sim,
DM_sim=DM_sim)

save(SimData_B,file=paste(getwd(),"SimData/SimData_B",sep="/"))
```