

2570

NINA Report

## Experimental harvest of willow ptarmigan, black grouse and capercaillie:

### Summary of the planning stage

Erlend B. Nilsen, Lasse Frost Eriksen, Christoffer Høyvik Hilde and Matthew Grainger



# NINA Publications

## **NINA Report (NINA Rapport)**

This is NINA's ordinary form of reporting completed research, monitoring or review work to clients. In addition, the series will include much of the institute's other reporting, for example from seminars and conferences, results of internal research and review work and literature studies, etc. NINA

## **NINA Special Report (NINA Temahefte)**

Special reports are produced as required and the series ranges widely: from systematic identification keys to information on important problem areas in society. Usually given a popular scientific form with weight on illustrations.

## **NINA Factsheet (NINA Fakta)**

Factsheets have as their goal to make NINA's research results quickly and easily accessible to the general public. Fact sheets give a short presentation of some of our most important research themes.

## **Other publishing.**

In addition to reporting in NINA's own series, the institute's employees publish a large proportion of their research results in international scientific journals and in popular academic books and journals.

# Experimental harvest of willow ptarmigan, black grouse and capercaillie:

Summary of the planning stage

Erlend B. Nilsen  
Lasse Frost Eriksen  
Christoffer Høyvik Hilde  
Matthew Grainger

Nilsen, E. B., Eriksen, L.F., Hilde, C. H. & Grainger, M. 2025. Experimental harvest of willow ptarmigan, black grouse and capercaillie: Summary of the planning stage. NINA Report 2570. Norwegian Institute for Nature Research.

Trondheim, February 2025

ISSN: 1504-3312

ISBN: 978-82-426-5387-1

COPYRIGHT

© Norwegian Institute for Nature Research

The publication may be freely cited where the source is acknowledged

AVAILABILITY

Open

PUBLICATION TYPE

Digital document (pdf)

QUALITY CONTROLLED BY

Knut Morten Vangen

SIGNATURE OF RESPONSIBLE PERSON

Research director Signe Nybø (sign.)

CLIENT(S) REFERENCE(S)

M-2929|2025

CLIENTS/SUBSCRIBER CONTACT PERSON(S)

Erik Lund

COVER PICTURE

Ptarmigan hunter with hunting dog © Erlend B. Nilsen

KEY WORDS

- Norway
- Willow ptarmigan
- Capercaillie
- Black grouse
- Harvest
- Experimental

NØKKELOORD

- Norge
- Lirype
- Storfugl
- Orrfugl
- Høsting
- Eksperimentell

CONTACT DETAILS

**NINA head office**

P.O.Box 5685 Torgarden  
NO-7485 Trondheim  
Norway  
P: +47 73 80 14 00

**NINA Oslo**

Sognsveien 68  
0855 Oslo  
Norway  
P: +47 73 80 14 00

**NINA Tromsø**

P.O.Box 6606 Langnes  
NO-9296 Tromsø  
Norway  
P: +47 77 75 04 00

**NINA Lillehammer**

Vormstuguvegen 40  
NO-2624 Lillehammer  
Norway  
P: +47 73 80 14 00

**NINA Bergen:**

Thormøhlens gate 55  
NO-5006 Bergen.  
Norway  
P: +47 73 80 14 00

[www.nina.no](http://www.nina.no)

## Abstract

Nilsen, E. B., Eriksen, L.F., Hilde, C. H. & Grainger, M. 2025. Experimental harvest of willow ptarmigan, black grouse and capercaillie: Summary of the planning stage. NINA Report 2570. Norwegian Institute for Nature Research.

Sustainable harvest management of game resources is an important goal in Norwegian nature management. A central concept for predicting the effect of hunting on wildlife populations is whether hunting mortality is additive or compensatory under different harvest levels. Grouse species such as willow ptarmigan (*Lagopus l. lagopus*) and the forest grouse species black grouse (*Lyrurus tetrix*) and capercaillie (*Tetrao urogallus*) are popular game species in Norway. The most robust knowledge about sustainable harvest for willow ptarmigan comes from experimental studies carried out in Norway in the late 1990s. However, it is unknown whether past results are directly transferable to a broader set of new environmental conditions induced by climate change. Also, no similar studies exist for black grouse or capercaillie. Here, we aim to address these uncertainties by conducting a conceptual replication of these experiments for the three above-mentioned species in hunting areas managed by Statskog across Norway (from Agder in the south to Troms in the north). This report summarises the planning stage of the project and includes two main sections: i) A systematic mapping of the literature related to effects of harvest on galliformes globally, and ii) a thorough evaluation of the experimental design including both the sampling of study subjects and assignment of treatment levels.

The systematic mapping of the literature revealed that several key aspects of the effects of harvest on galliformes warrant further attention and research. First, the literature is dominated by research with relatively "weak" study designs, and few were based on Randomized Controlled Trials (RCT). In fact, only one study based on RCT was found in addition to the above-mentioned Norwegian research. Second, most studies had a relatively small or moderate spatial extent. This might limit our ability to draw firm conclusions that could be transferred to other ecological conditions (and species). Third, the amount of research on black grouse and capercaillie in relation to harvest management is very limited.

To select an adequate number and location of included areas (study units), we first used mathematical simulations to assess the number of study units that would be required to detect effects of the prescribed harvest treatment levels (i.e., no harvest, 15% and 30% harvest for willow ptarmigan, and no harvest, 10% and 20% harvest for the forest grouse species black grouse and capercaillie). We suggest recruiting 34 study units for willow ptarmigan and 30 study units for forest grouse, each subject to a randomised three-year harvest experiment. Black grouse and capercaillie will be treated as two distinct experiments. To recruit specific study units, we applied a set of criteria and algorithms based on an approach resembling "*proportional quota sampling*". In addition to certain key requirements concerning practical suitability, we selected study units to obtain balance in terms of geographical coverage and environmental gradients (*bioclimatic sections and zones*). Each study unit will be subject to a randomly allocated harvest treatment, following an RCT approach and a "*crossover-design*", where each unit is subject to a randomly selected *sequence* of harvest treatments. Because we are planning for a three-year experiment with three harvest levels, each unit will be exposed to all treatments.

This experimental study of the effects of harvest on willow ptarmigan, black grouse and capercaillie, will be a unique and substantial contribution to the field. The systematic mapping of the literature clearly revealed that the knowledge created by this study will fill important knowledge gaps. The proposed study units for the experiment will be a good basis for a replication study with high external validity. The planned study will be unique in that it both has a (near) national spatial extent, and that it covers three related but ecologically distinct species.

Erlend B. Nilsen – [erlend.nilsen@nina.no](mailto:erlend.nilsen@nina.no), Lasse Frost Eriksen – [lasse.eriksen@nina.no](mailto:lasse.eriksen@nina.no), Christoffer Høyvik Hilde – [christoffer.hilde@nina.no](mailto:christoffer.hilde@nina.no), Matthew Grainger – [matthew.grainger@nina.no](mailto:matthew.grainger@nina.no), Norwegian Institute for Nature Research, NINA

## Sammendrag

Nilsen, E. B., Eriksen, L.F., Hilde, C. H. & Grainger, M. 2025. Eksperimentell jakt på lirype, orrfugl og storfugl: Oppsummering av planleggingsprosjektet. NINA Rapport 2570. Norsk institutt for naturforskning.

Høsting av viltressurser på en bærekraftig måte er et viktig mål i norsk naturforvaltning. Et sentralt konsept for å forutsi effekten av jakt på viltbestander er om jaktmortaliteten er additiv eller kompensatorisk under forskjellige høstingsnivåer. Hønsefugler som lirype (*Lagopus l. lagopus*), storfugl (*Lyrurus tetrix*) og orrfugl (*Tetrao urogallus*) er populære jaktbare arter i Norge. Den mest robuste kunnskapen om bærekraftig høsting av lirype kommer fra eksperimentelle studier utført i Norge på slutten av 1990-tallet. Det er imidlertid ukjent om resultatene herfra er direkte overførbare til et bredere sett med nye miljøforhold som følge av klimaendringer. Det finnes ingen lignende studier for storfugl eller orrfugl. I denne rapporten adresserer vi disse usikkerhetene ved å gjennomføre en konseptuell replikasjon av høstingseksperimentet, for alle de tre ovennevnte artene, i jaktområder som forvaltes av Statskog over hele Norge (fra Agder i sør til Troms i nord). Denne rapporten oppsummerer planleggingsfasen av prosjektet og inkluderer to hoveddeler: i) En systematisk kartlegging av litteraturen knyttet til effekter av høsting på hønsefugler globalt, og ii) en grundig evaluering av det eksperimentelle designet, inkludert både utvalg av studieobjekter og tildeling av behandlingsnivåer.

Den systematiske kartleggingen av litteraturen avslørte at flere sentrale aspekter ved effektene av høsting av hønsefugler krever økt oppmerksomhet og forskning. For det første er litteraturen dominert av forskning med relativt "svake" studiedesign, og få var basert på randomiserte kontrollerte forsøk (RCT). Kun én studie i våre funn var basert på RCT, i tillegg til den ovennevnte norske forskningen. For det andre hadde de fleste studiene et relativt lite eller moderat romlig omfang. Dette kan begrense muligheten til å trekke sikre konklusjoner som kan overføres til andre økologiske forhold (og arter). For det tredje er det per i dag svært lite forskning på jakteffekter på storfugl og orrfugl.

For å velge et tilstrekkelig antall og plassering av inkluderte områder (studieenheter), brukte vi først matematiske simuleringer for å vurdere antall studieenheter som ville være nødvendig for å oppdage effekter av de foreskrevne høstingsnivåene (dvs. ingen høsting, 15 % og 30 % høsting for lirype, og ingen høsting, 10 % og 20 % høsting for skogshønsartene storfugl og orrfugl). Vi foreslår å rekruttere 34 studieenheter for lirype og 30 studieenheter for skogshøns, hver underlagt et randomisert treårig høstingseksperiment. Storfugl og orrfugl vil bli behandlet som to separate eksperimenter. For å rekruttere spesifikke studieenheter brukte vi et sett med kriterier og algoritmer basert på en tilnærming som ligner "proporsjonal kvoteutvelging". I tillegg til visse nøkkelkrav angående praktisk egnethet, valgte vi studieenheter hvor vi oppnår balanse i geografisk dekning og miljøgradienter (bioklimatiske seksjoner og soner). Hver studieenhet vil bli underlagt et tilfeldig tildelt høstingsnivå, etter en RCT-tilnærming og et "crossover-design", der hver enhet er underlagt en tilfeldig valgt sekvens av høstingsnivåer. Ettersom vi planlegger et treårig eksperiment med tre høstingsnivåer, vil hver enhet få ulike høstingsnivå i ulike år.

Denne eksperimentelle studien av høstingseffekter på lirype, storfugl og orrfugl vil gi et unikt og betydelig bidrag til kunnskapsfeltet. Den systematiske kartleggingen av litteraturen avslørte tydelig at kunnskapen som skapes her vil fylle viktige kunnskapshull. De foreslåtte studieenheterne for eksperimentet vil være et godt grunnlag for en replikasjonsstudie med høy ekstern validitet. Den planlagte studien vil også være unik ved at den både har et (nær) nasjonalt romlig omfang, og at den dekker tre beslektede, men økologisk distinkte arter.

Erlend B. Nilsen – [erlend.nilsen@nina.no](mailto:erlend.nilsen@nina.no), Lasse Frost Eriksen – [lasse.eriksen@nina.no](mailto:lasse.eriksen@nina.no), Christoffer Høyvik Hilde – [christoffer.hilde@nina.no](mailto:christoffer.hilde@nina.no), Matthew Grainger – [matthew.grainger@nina.no](mailto:matthew.grainger@nina.no), Norsk institutt for naturforskning, NINA

# Contents

<b>Abstract .....</b>	<b>3</b>
<b>Sammendrag .....</b>	<b>4</b>
<b>Contents .....</b>	<b>5</b>
<b>Foreword .....</b>	<b>6</b>
<b>1 Introduction.....</b>	<b>7</b>
<b>2 Systematic mapping of the literature.....</b>	<b>9</b>
2.1 Methods .....	9
2.1.1 Search for articles .....	9
2.1.2 Benchmark papers.....	10
2.1.3 Article screening .....	10
2.1.4 Eligibility criteria .....	11
2.1.5 Title and abstract screening stage.....	11
2.1.6 Full text eligibility.....	11
2.1.7 Data coding .....	11
2.2 Results from the literature review .....	12
2.3 Summary of the literature mapping .....	23
<b>3 Experimental design .....</b>	<b>24</b>
3.1 Experimental harvest treatment .....	26
3.1.1 Random allocation of harvest levels.....	27
3.1.2 Outcome variables .....	27
3.1.3 Estimation of hunting quota .....	27
3.2 Simulation to assess sample size .....	28
3.2.1 Population parameters.....	28
3.2.2 Harvest treatment and covariate .....	28
3.2.3 Results .....	29
3.3 Inclusion of study units .....	31
3.3.1 Criteria for inclusion of study units.....	32
3.3.2 Key characteristics of study units .....	32
3.3.3 Calculation of area sizes.....	36
<b>4 Conclusions .....</b>	<b>37</b>
<b>5 References .....</b>	<b>39</b>

## Foreword

We here report from the planning stage of a large-scale project where the aim is to generate new knowledge for sustainable management, through experimental harvest of willow ptarmigan, black grouse and capercaillie. Such a project has been discussed for many years between several of the project participants, and between researchers at NINA and managers from Statskog and Fjellstyrene. In wildlife research it is rare that we are granted funding at the planning stage of a project, and we are grateful to the Norwegian Environment Agency for funding this pre-project.

As project leader, I would like to thank my co-authors for outstanding contributions to this report. Moreover, I would like to thank Chloé R. Nater, Hans Chr. Pedersen, James Martin and Fjellstyrene i Lierne for fruitful discussions and input to this report. Finally, it has been very productive to collaborate with Statskog (Jo Inge Breisjøberget, Jonas Hagen, Kristian Eiken Olsen, Lars Movik, Stian Sundsvik and Eva Kristin Grøndal) in this pre-project, and we are grateful that we can collaborate on carrying out the planned experiments on their land.

March 12<sup>th</sup> 2025, Erlend B. Nilsen



# 1 Introduction

Sustainable harvest management of our game resources is a key goal in Norwegian nature management. In order to achieve sustainable management, it is generally recommended to use an adaptive approach, where one actively searches for new knowledge that is necessary to fill existing knowledge gaps (Johnson et al. 1997; Månsson et al. 2023). In the last couple of decades a series of environmental challenges have arisen, which are expected to largely affect hunting management. These challenges are generally linked to climate change and other man-made changes to the natural environment, such as habitat loss and degradation, but also changes in other components of the socio-ecological systems wildlife species are part of and which are expected to influence how sustainable management is defined (Eriksen et al. 2018; Bowler et al. 2020; Henden et al. 2020; Månsson et al. 2023). With increased focus in society on the loss of biological diversity and degradation of ecosystems on a global and national scale, there are increased expectations from the society that recreational hunting is carried out in a sustainable way with minimal risk of overharvesting. This can eventually be assumed to lead to stricter requirements for documentation of sustainability. More recent conceptual studies of adaptive management have discussed how these new situations not only force new knowledge needs, but that it can also lead to changed objectives and measurement parameters for adaptive management (Månsson et al. 2023). In the light of the drastic changes happening across the world's ecosystems, largely driven by anthropogenic activity, it is vital to simultaneously investigate and quantify the effects of harvest and environmental variability on wildlife populations.

A central concept for predicting the effects of hunting on wildlife populations is the relationship between additive hunting mortality (which comes in addition to other mortality) vs compensatory hunting mortality (which is compensated by other mortality being reduced in proportion to the harvest mortality - see Péron (2013) and Sandercock et al. (2011)). Our knowledge about this phenomena in our small game populations comes largely from experimental studies carried out at the end of the 1990s under the auspices of the "Rypejaktprosjektet" (with Hans Chr. Pedersen, NINA, as project leader). Here it was concluded that hunting mortality was partially compensatory at low harvest rates, while it was additive or even super-additive at high harvest rates (Pedersen et al. 2004; Sandercock et al. 2011). This knowledge base was produced by carrying out large-scale experimental studies, where the harvest treatments (0%, 15% and 30%, respectively) were distributed randomly between the study areas. Effects were investigated both directly in terms of survival (using radio-tagged birds - see Sandercock et al. (2011)) and indirectly via population densities (see Pedersen et al. (2004)). This project focused exclusively on willow ptarmigan (*Lagopus l. lagopus*), and no similar knowledge exists for e.g. black grouse (*Lyrurus tetrix*) or capercaillie (*Tetrao urogallus*).

Whether results from the former large scale experiment is transferable in time (i.e. to the current environmental conditions) and to other species (i.e. black grouse and capercaillie) is an empirical question that cannot be answered directly by existing data. For example, one can expect that climate change will have significant effects on the trophic interactions that largely affect grouse population dynamics and densities (Bowler et al. 2020; Henden et al. 2020). Changes in these interactions will have the potential to directly affect grouse populations' potential to compensate for hunting mortality (see discussions in e.g. Péron 2013). Similarly, large changes in the size of the grouse population could have similar effects, since the degree of compensation can be closely linked to density dependence (Sinclair and Pech 1996). One way to address these uncertainties would be to conduct a conceptual replication of the experiments carried out in the 1990s. Such conceptual replication studies (*sensu* Nakagawa and Parker 2015) are rare in ecology (Fraser et al. 2020), but have recently been

strongly recommended to examine the validity and generalisability of previously conducted studies (Nakagawa and Parker 2015; Fraser et al. 2020).

Several aspects should be carefully investigated before carrying out such a conceptual replication. Initially, an up-to-date mapping of the research-based literature on the sustainable management of galliformes from a climate change perspective (see also Grainger et al. 2020) will be suitable to identify gaps in the literature and to ensure that the replication rests on the best available knowledge. Further, a successful implementation of a replication study requires thorough planning of how best to design such an experiment in a way that sample sizes and climatic gradients are sufficiently estimated in order to close the most pressing knowledge gaps. This report is a central part of the planning stage of such a project, where the intent is to apply experimental harvest treatments to populations of willow ptarmigan, black grouse and capercaillie across Norway. The project is planned to take place across a three-year experiment within a select number of hunting areas managed by Statskog across Norway (from Agder in South to Troms in the North). The report includes two main sections:

- A systematic mapping of the literature related to effects of harvest on galliformes
- A thorough evaluation of the experimental design including both the sampling of study subjects and assignment of treatment levels

We will first present the results from the systematic mapping of the literature, and then the evaluation of experimental design. In the latter chapter, we will present a detailed description of the suggested design for the new harvest experiment.

## 2 Systematic mapping of the literature

We aimed to identify studies that have quantitatively assessed the effects of legal harvest (i.e., managed hunting) on galliformes, to ensure that the planned experimental studies on the sustainability of galliform hunting are based on the best available evidence. This mapping aims to identify, categorise, and synthesise existing research on how hunting mortality interacts with ecological and environmental factors. By systematically compiling relevant studies, we provide a foundation for assessing knowledge gaps and determining the generalisability of past research.

The protocol (<https://osf.io/ny2it/>) i.e., the workflow of identifying, categorising and synthesising the existing literature, was prepared prior to conducting the literature searches. Here, we include the relevant information and report any deviation from the protocol.

### 2.1 Methods

#### 2.1.1 Search for articles

We used a PIO (Population, Intervention, Outcome) question structure to help develop the search strategy.

- Population (P): Bird species within the order Galliformes
- Intervention (I): Legal (managed) hunting or harvest
- Outcome (O): The ecological impacts of managed hunting on population dynamics, demography, genetics or behaviour

We searched for academic literature (publications) in two search platforms; Web of Science and lens.org. Specifically, we searched the Web of Science Core Collection (WOS.SCI: 1987 to 2025, WOS.AHCI: 1987 to 2025, WOS.ESCI: 2020 to 2025 and WOS.SSCI: 1987 to 2025) using the NINA library subscription and the following search string:

((TS=(galliform\* OR grouse OR ptarmigan OR Phasianidae OR pheasant OR "lagopus lagopus" OR "lagopus muta" OR "wild turkey" OR "bobwhite" OR partridge OR "tetrao" OR "lyrurus" OR "capercaillie")

AND (TS=("hunt" OR "harvest"))

AND SU=("Ecology" OR "Wildlife" OR "Environment"))

For lens.org we used the following search string:

(abstract:("galliform" OR "grouse" OR "ptarmigan" OR "Phasianidae" OR "pheasant" OR "lagopus lagopus" OR "lagopus muta" OR "wild turkey" OR "bobwhite" OR "partridge" OR "tetrao" OR "lyrurus" OR "capercaillie") OR title:("galliform" OR "grouse" OR "ptarmigan" OR "Phasianidae" OR "pheasant" OR "lagopus lagopus" OR "lagopus muta" OR "wild turkey" OR "bobwhite" OR "partridge" OR "tetrao" OR "lyrurus" OR "capercaillie")) AND abstract:("hunt" OR "harvest") OR title:("hunt" OR "harvest") AND field\_of\_study:("Ecology" OR "Wildlife" OR "Environment").

In addition, a backward and forward citation chasing (using CitationChaser) was performed based on the three first benchmark papers listed below (Pedersen et al. 2004; Seding and Rotella 2004; Sandercock et al. 2011).

### 2.1.2 Benchmark papers

The following benchmark papers were used to assess the effectiveness of the search string. We expected to find all these papers in a review of this question. We recorded if each benchmark paper was found in the searches (Wo and lens.org).

- Sandercock, B. K., E. B. Nilsen, H. Brøseth, and H. C. Pedersen. 2011. Is hunting mortality additive or compensatory to natural mortality? Effects of experimental harvest on the survival and cause-specific mortality of willow ptarmigan. *Journal of Animal Ecology* 80:244-258.
- Pedersen, H. C., H. Steen, L. Kastdalen, H. Brøseth, R. A. Ims, W. Svendsen, and N. G. Yoccoz. 2004. Weak compensation of harvest despite strong density-dependent growth in willow ptarmigan. *Proceedings of the Royal Society of London Series B-Biological Sciences* 271:381-385.
- Sedinger, J.S. & Rotella, J.J. 2005. Effect of harvest on sage-grouse *Centrocercus urophasianus* populations: what can we learn from the current data? *Wildlife Biology*, 11, 371–375.
- Pedersen, Å. Ø., E. M. Soininen, S. Unander, M. H. Willebrand, and E. Fuglei. 2014. Experimental harvest reveals the importance of territoriality in limiting the breeding population of Svalbard rock ptarmigan. *European Journal of Wildlife Research* 60:201-212.
- Brøseth, H., J. Tufto, H. C. Pedersen, H. Steen, and L. Kastdalen. 2005. Dispersal patterns in a harvested willow ptarmigan population. *Journal of Applied Ecology* 42:453-459.
- Brøseth, H., E. B. Nilsen, and H. C. Pedersen. 2012. Temporal quota corrections based on timing of harvest in a small game species. *European Journal of Wildlife Research* 58:797-802.
- Brøseth, H., and H. C. Pedersen. 2000. Hunting effort and game vulnerability studies on a small scale: a new technique combining radio-telemetry, GPS and GIS. *Journal of Applied Ecology* 37:182-190.
- Henden, J.-A., R. A. Ims, N. G. Yoccoz, E. J. Asbjørnsen, A. Stien, J. P. Mellard, T. Tveraa, F. Marolla, and J. U. Jepsen. 2020. End-user involvement to improve predictions and management of populations with complex dynamics and multiple drivers. *Ecological Applications* 30(6):e02120.
- Small, R.J., Holzward, J.C. & Rusch, D.H. 1991. Predation and hunting mortality of ruffed grouse in central Wisconsin. *Journal of Wildlife Management*, 55, 512–520.

In our searches, all benchmark papers were found. Henden et al. (2020) and Small et al. (1991) were not indexed in lens.org but were found in Web of Science.

### 2.1.3 Article screening

We used the online screening tool Rayyan (Ouzzani et al. 2016) for screening the results of the article search. A pre-determined inclusion/exclusion criteria was used when screening eligible studies. Excluded papers will be shown in a flowchart using the ROSES flow chart

(Haddaway et al. 2018). Included studies was de-duplicated based on DOI and title matches using the deduplication tool in Rayyan.

After de-duplication, title and abstract screening was performed using the eligibility criteria. If uncertainty remained, the full text version was visited to confirm/reject. A set of 54 papers was reviewed by three members of the review team to ensure consistency.

#### 2.1.4 Eligibility criteria

We reviewed scientific articles from all parts of the world covering all galliform species. Studies of both native and introduced populations were eligible. Studies of all spatial scales and across all relevant ecosystems were considered eligible. Only studies documenting the effects of legal harvest on galliform populations were eligible. Studies of poaching were not eligible. The effects must have been assessed quantitatively for the study to be eligible. Eligible outcomes included effects of harvest on population abundance, density, sex- and age structure, recruitment, survival, behaviour, genetics, physiology and distribution. Studies focusing on heavy metal (e.g. lead) poisoning or content in wild birds from shooting were not eligible. Both observational and experimental study designs were eligible. Mathematical simulation studies were eligible.

#### 2.1.5 Title and abstract screening stage

We used a decision tree to make decisions about inclusion and exclusion at the *Title* and *Abstract* stage. The tree is presented in the protocol and in the *Appendix* of this report. The decision tree helped the reviewers to make decisions about eligibility of each paper in a structured way. Dual screening was carried out on 20% of papers ( $n = 183$ ). Agreement between reviewers was assessed using the kappa statistic which was  $> 0.7$ , which is deemed acceptable.

#### 2.1.6 Full text eligibility

Papers included at the *Title* and *Abstract* stage was taken forward to the full text screening stage. An RIS file of included articles was downloaded from Rayyan and added to a Zotero database. PDF versions of each paper were searched using Zotero's "Find Available PDF" tool, ResearchGate, Google Scholar and Google. If a PDF was not available online (or was not open to us), the paper was excluded at this stage of the review. Full text papers were excluded if they did not meet the following criteria:

- Written in English, Norwegian, Swedish or Danish
- Focused on species from the order Galliformes
- Documented the effect of legal or managed hunting on galliformes
- Assessed the effect of hunting quantitatively

#### 2.1.7 Data coding

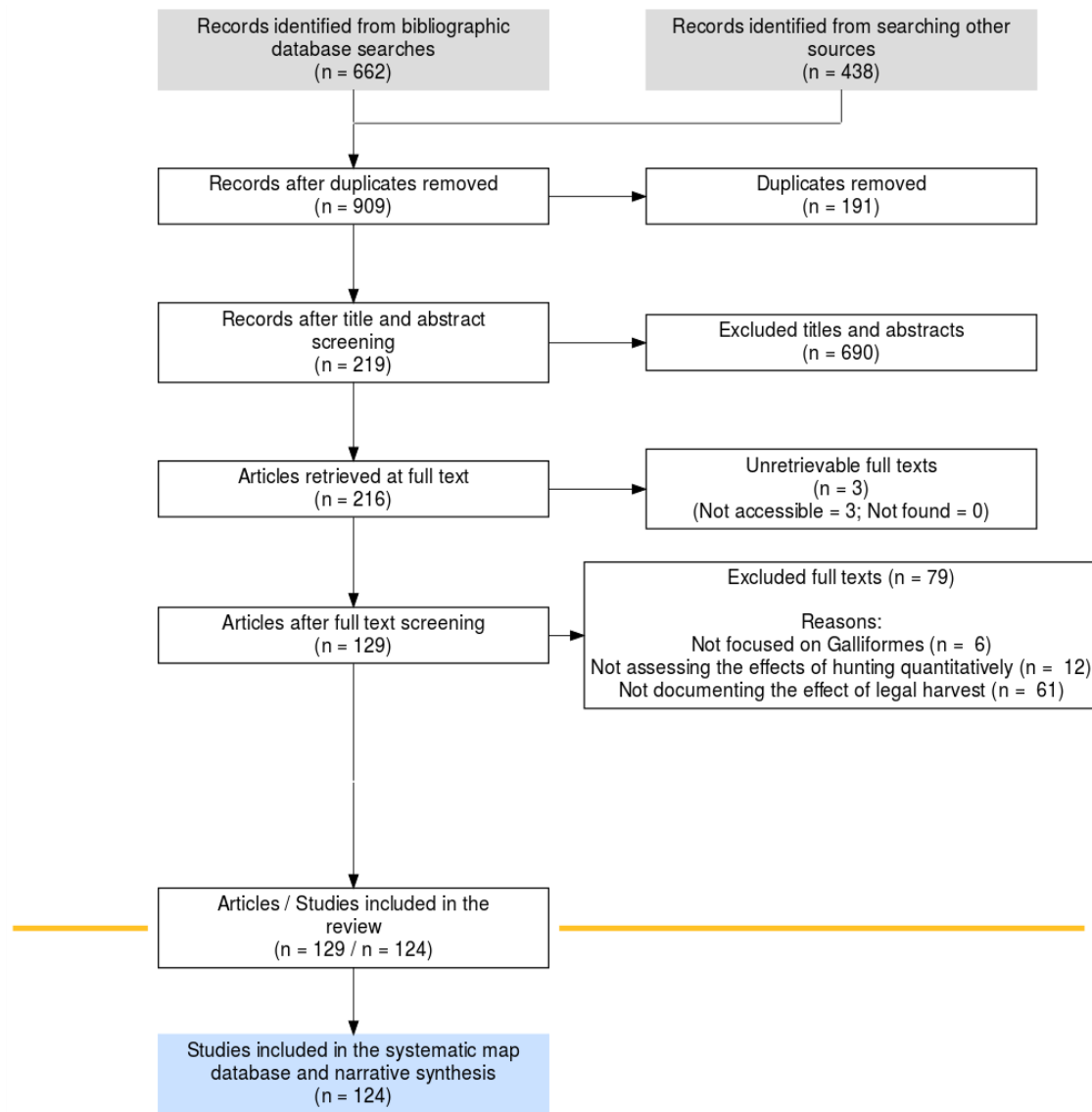
From all papers that met the eligibility criteria, we extracted key variables (a.k.a. "data coding"). In addition to metadata about the publication (publication year, journal name etc.) we extracted from the full-text information about study species, study area location and extent, research approach and experimental design, outcome variables and to which extent the

paper reported harvest rates and to which extent the study assessed whether harvest was additive or compensatory. For outcome variables, we used the Essential Biodiversity Variable (EBV) framework (Pereira et al. 2013; Navarro et al. 2017), and coded both for EBV class (high order classification) and EBV name (lower order classification). Some of the EBV names were modified to suit our study, and these marked with an asterisk in the relevant tables and figures. A detailed description of data coding is presented in the protocol and in the *Appendix 1*.

## 2.2 Results from the literature review

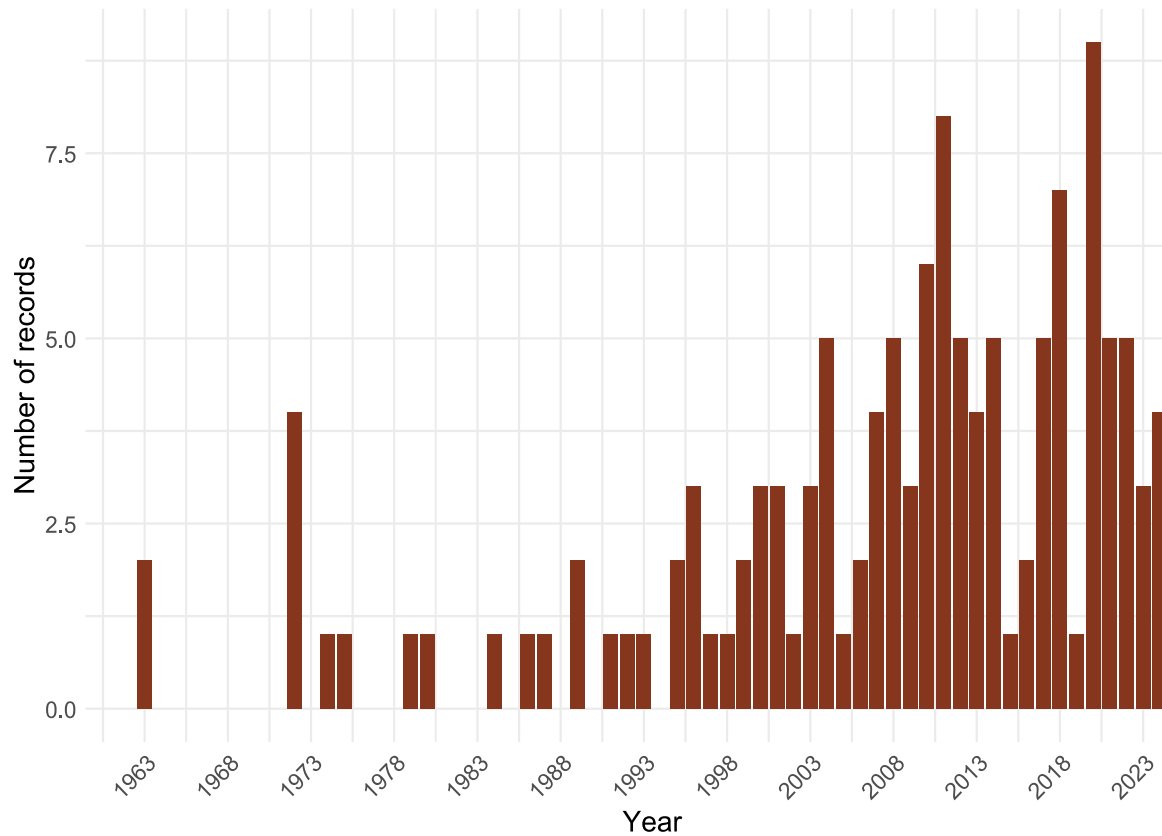
Our searches returned 1100 articles in total, with 549 from Web of Science, 113 from lens.org, 9 benchmark articles and 168 from backward-, and 192 from forward citation chasing from the three benchmark articles mentioned above (i.e., Sandercock et al. (2011); Pedersen et al. (2004); Sedinger and Rotella (2004)). Of these, 191 were duplicates which were removed from the database. We screened the remaining 909 articles for eligibility at the *Abstract* and *Title* stage. Of the 909 included articles, a total of 219 were moved further to full-text screening, of which 176 papers that had an “include” decision and 34 a “maybe” decision. There were 9 conflicts (where reviewers disagreed on eligibility). All these papers were taken through to the fulltext stage and assessed for eligibility. Exclusion decisions (690 records) were due to the research not focusing on the effects of legal or managed hunting (368 records), not being focused on galliformes (280 records), not quantitative (23 records), abstract not found online (16 records), or abstract not available in English, Norwegian, Swedish or Danish (3 records).

At the fulltext stage 90 articles were excluded (3 with no available pdf, 6 not focused on galliformes, 12 not assessing the effects of hunting quantitatively and the rest, 61, not documenting the effects of legal harvest). This left 129 papers that went forward to the data extraction stage (Figure 1). A full list of included papers is available from a repository at Open Science Framework (<https://osf.io/ny2jt/>).



**Figure 1:** ROSES Flowchart showing the number of papers included or excluded in the review

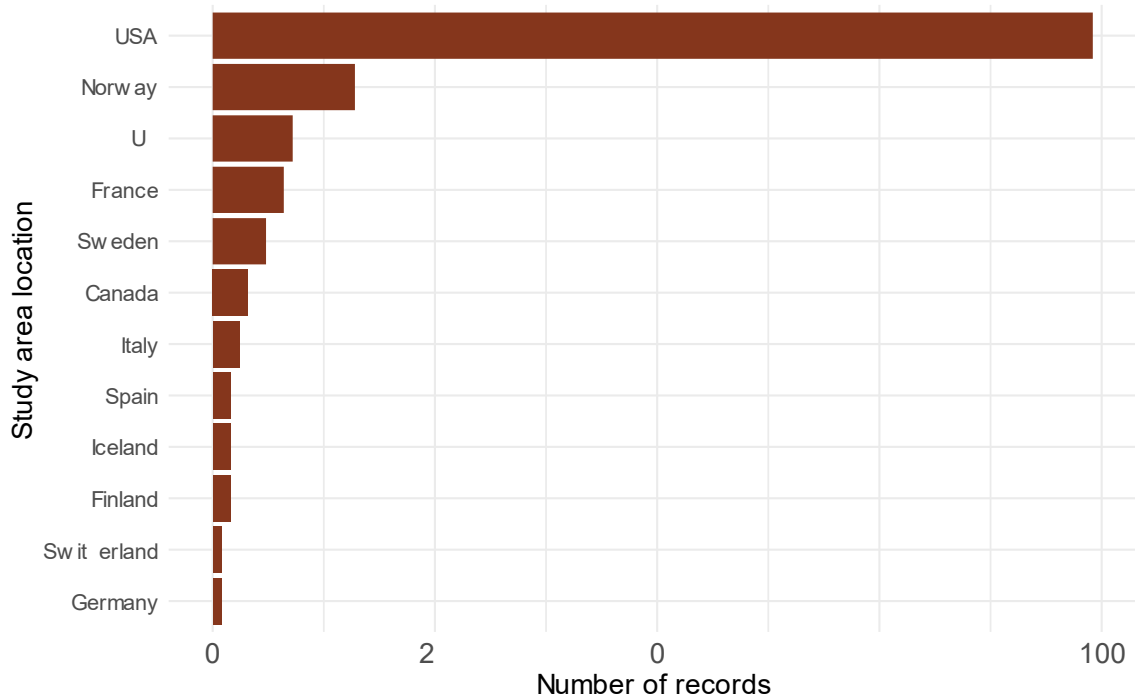
Legal hunting of galliformes has been subject to extensive research. In our searches we identified studies published as far back as 1963, but we are aware that research on this topic was carried out even before that; these studies were however not covered by our searches. The number of studies per year peaked in 2020 (8 studies) and the long-term average was 2.93 studies per year (Figure 2).



**Figure 2:** *The number of studies per year*

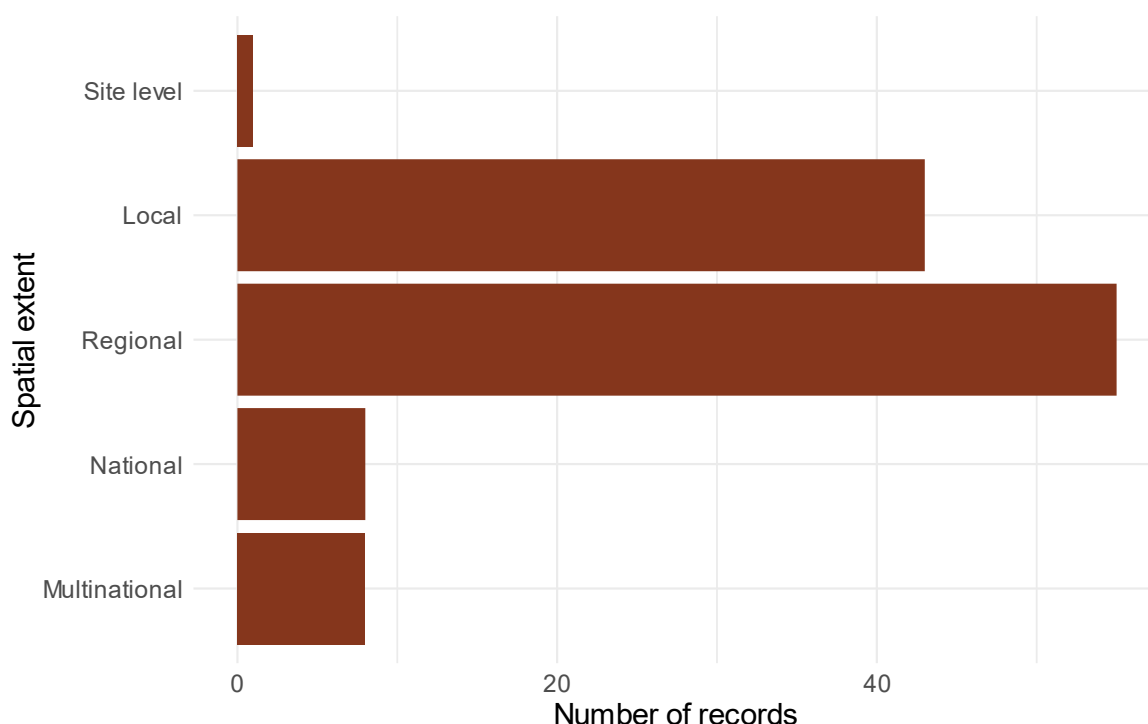
The most common geographic location for studies was in the USA (99 studies), followed by Norway (16) and the UK (9) (Figure 3). In the USA studies were carried out in 33 States (Alaska, California, Colorado, Florida, Georgia, Idaho, Illinois, Indiana, Kansas, Kentucky, Louisiana, Maine, Michigan, Mississippi, Missouri, Montana, Nevada, New York, North Carolina, Ohio, Oklahoma, Oregon, Pennsylvania, Poland, Rhode Island, South Carolina, South Dakota, Tennessee, Texas, Utah, Virginia, West Virginia, Wisconsin).





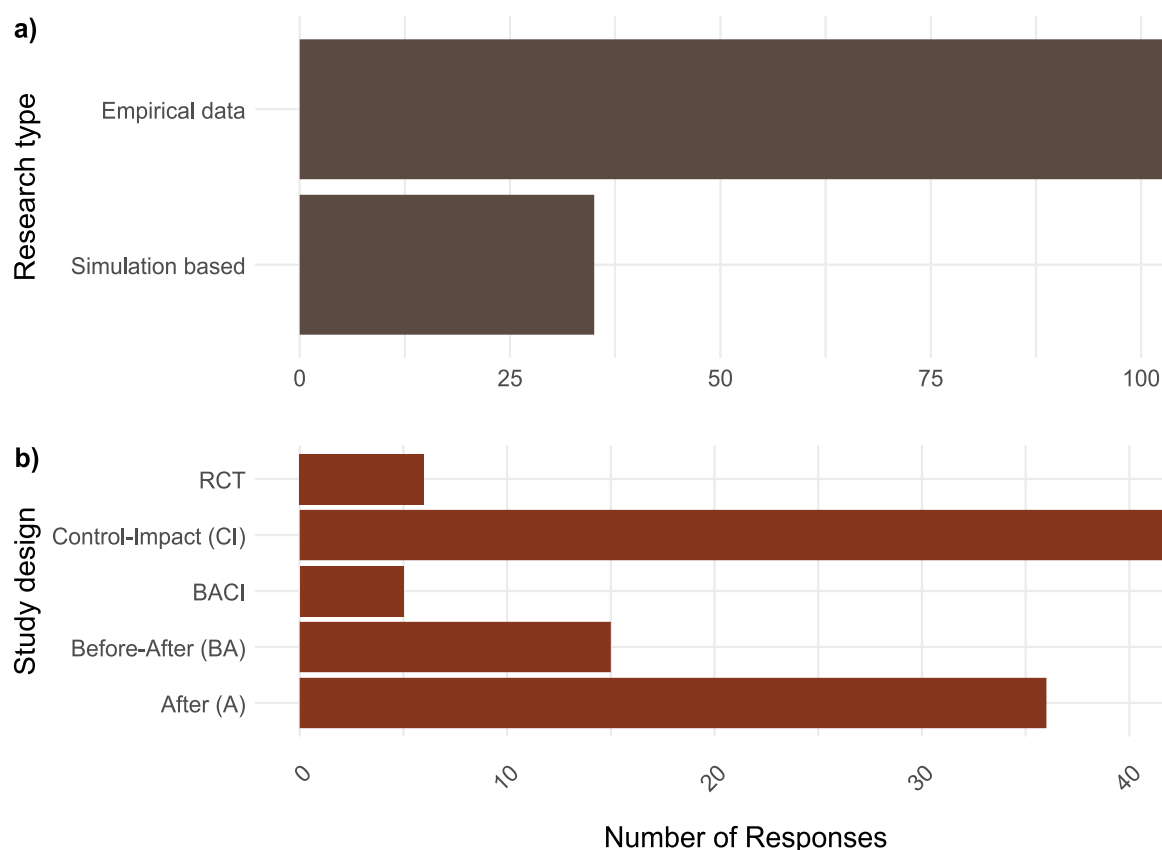
**Figure 3:** Geographic location (country) of included studies

The extent of the study areas were predominately regional (e.g. a county, a larger mountain area etc) or local (e.g. a municipality or part of municipality), with very few studies at the site level (Figure 4). Few studies were of national or multinational scale. Note that this contrasts previous calls for more large-scale research (Estes et al. 2018) that might be needed to uncover the ecological dynamics and to provide input to natural resource management policy at larger spatial scales.



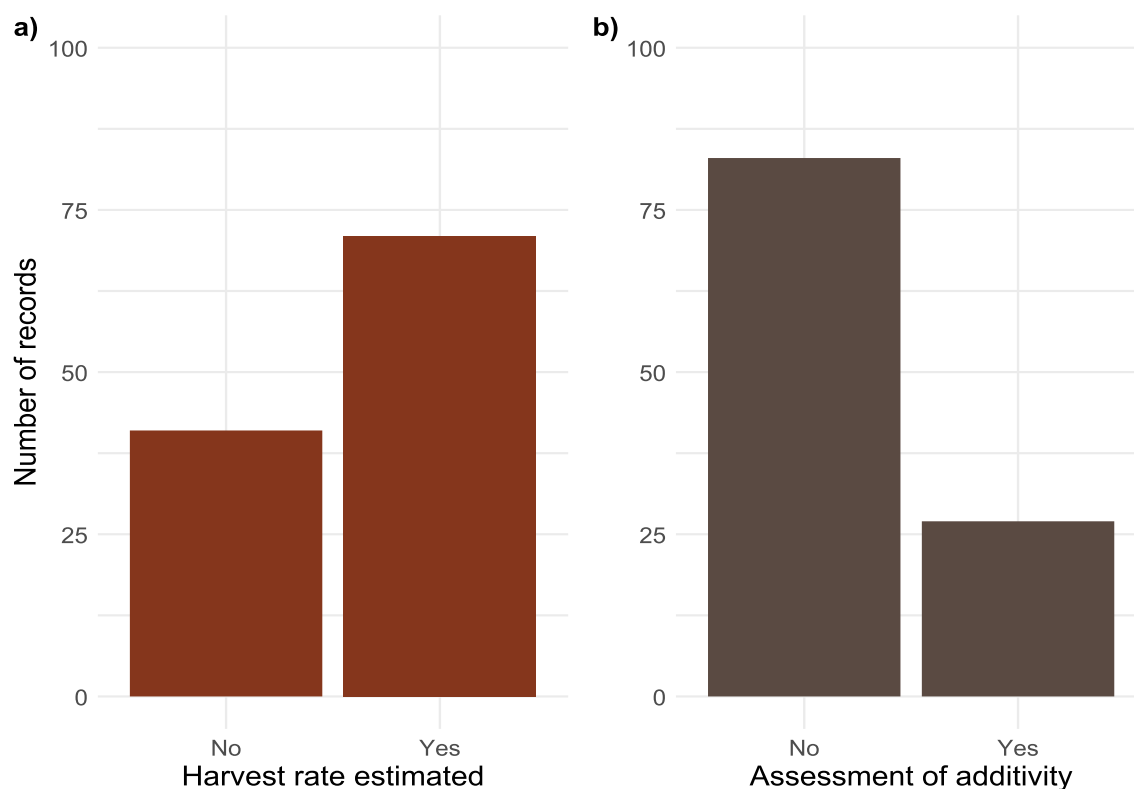
**Figure 4:** The study area extent

Most studies covered in our mapping were observational (i.e. based on analysis of empirical data) ( $n = 106$ ), but there were also a substantial number of studies that relied on simulations or scenario analyses ( $n = 35$ ; Figure 5 a). Note that the sum of observational studies based on empirical data and studies based on simulations is higher than the total of 129 papers that were found eligible. This is because some studies used both research approaches and thus were included in both categories. In agreement with previous assessments of study designs in ecological research most of the studies that presented empirical data were based on simple (and less robust) study designs (Figure 5 b). There were only 6 papers using a “Randomised Controlled Trial (RCT)” approach, the study design that leads to the highest level of evidence for causal understanding (Christie et al. 2019). These were limited to Norway (5) and the USA (1). The Norwegian articles were all from the same research project/experiment conducted in the late 1990s that was mentioned above. Note that the study from the USA was based on “simulated harvest”, but did include randomisation.



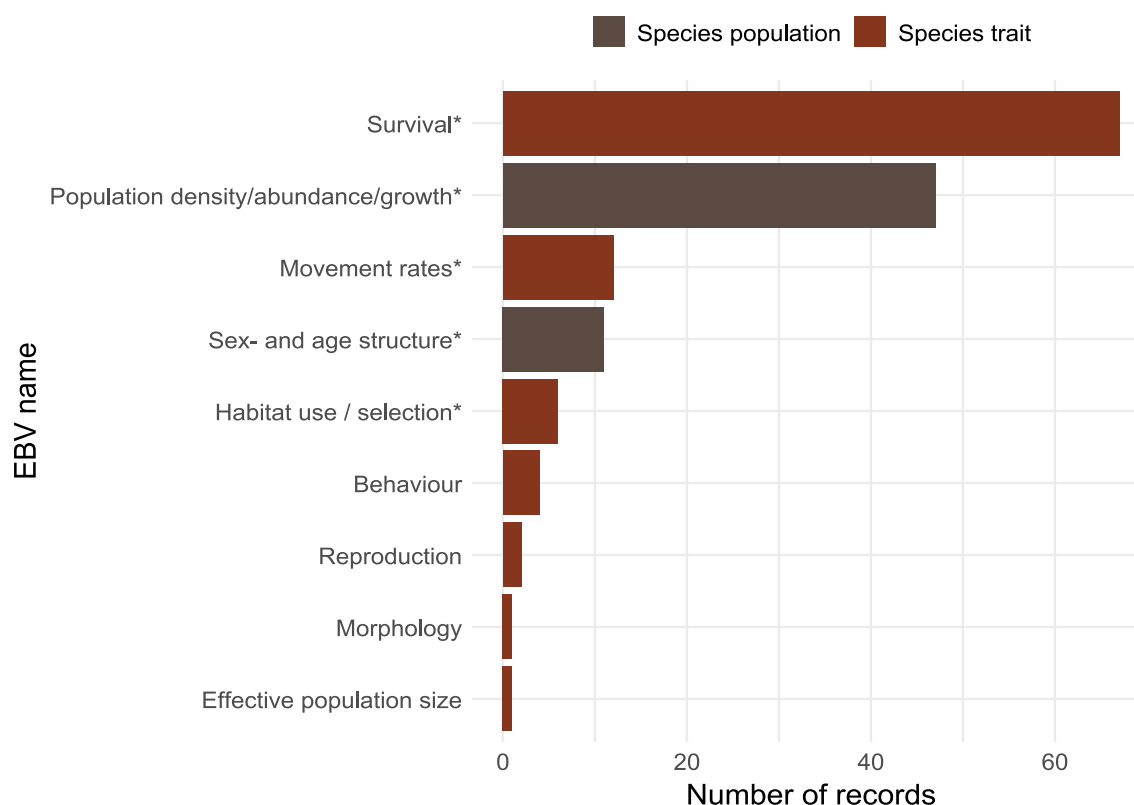
**Figure 5:** Assessment of which type of research has been reported in the covered literature (a), and the study design for studies that relied on analysis of empirical data (i.e. observational studies; RCT = Random(-ised) Controlled Trial).

We also extracted information about harvest rates (or percentage of the population removed by harvest) in 71 papers. The remaining papers did not explicitly report such information (Figure 6 a). Moreover, among the papers that reported harvest rate or harvest mortality, only 27 made any assessment to which extent harvest was additive or not (Figure 6 b). Note however that this is somewhat ambiguous, and it might depend on how the concept of additivity is defined: although we did not quantify this by direct coding we note that much fewer studies were in fact able to properly document additivity vs compensation via a robust quantitative approach.



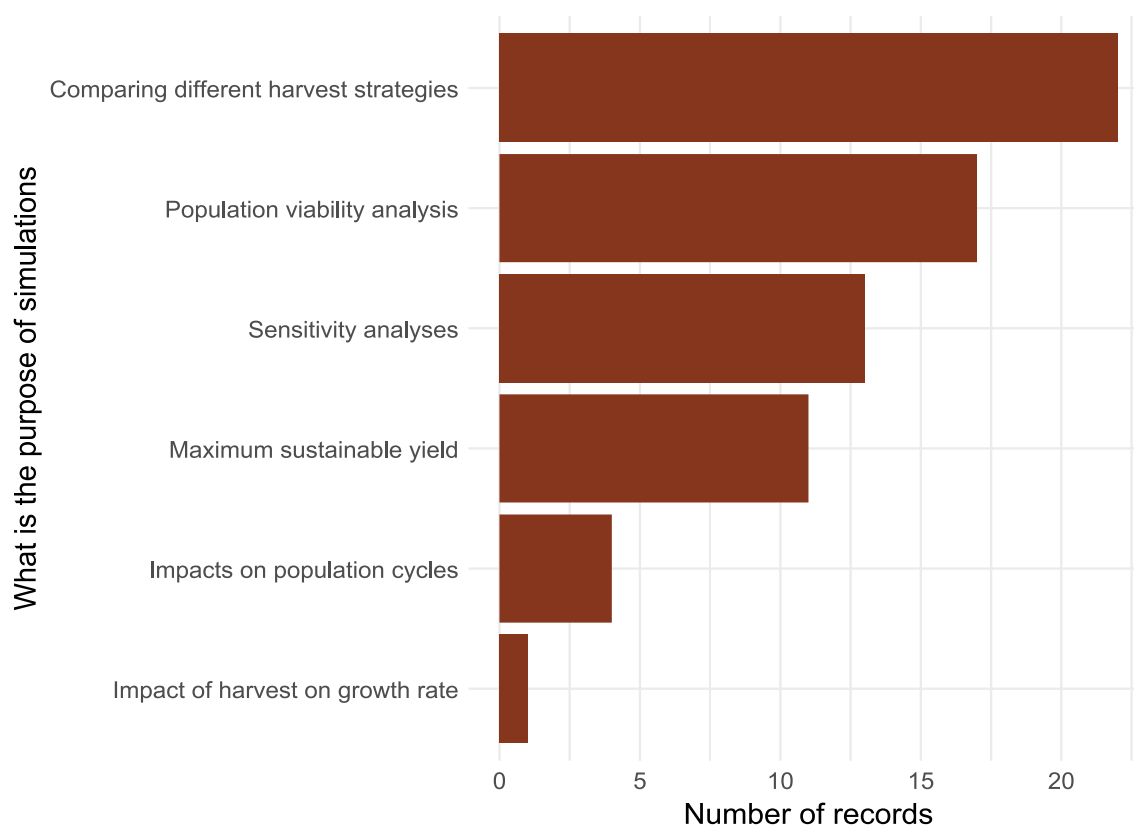
**Figure 6:** Number of studies explicitly reporting harvest rate (a) or did an assessment of to which extent the harvest mortality was additive (b).

From each of the eligible studies, we assessed which outcome variables were examined (Figure 7), at two different levels (EBV class and EBV names, respectively). In general, there were about equal number of studies focusing on the EBV classes “Species traits” (6 records) and “Species populations” (60 records). We did not record this variable for simulation based studies. Separating the EBV classes into EBV names, we found that the outcome variable most commonly studied was “Survival” (6 records), followed by “Population density/abundance/growth\*” (4 records). In addition, several studies focused on effects of “Movement rates” or “Habitat use/selection” (12 and 6 records, respectively), and 11 papers reported effects of harvest on “Sex- and age-structure”.



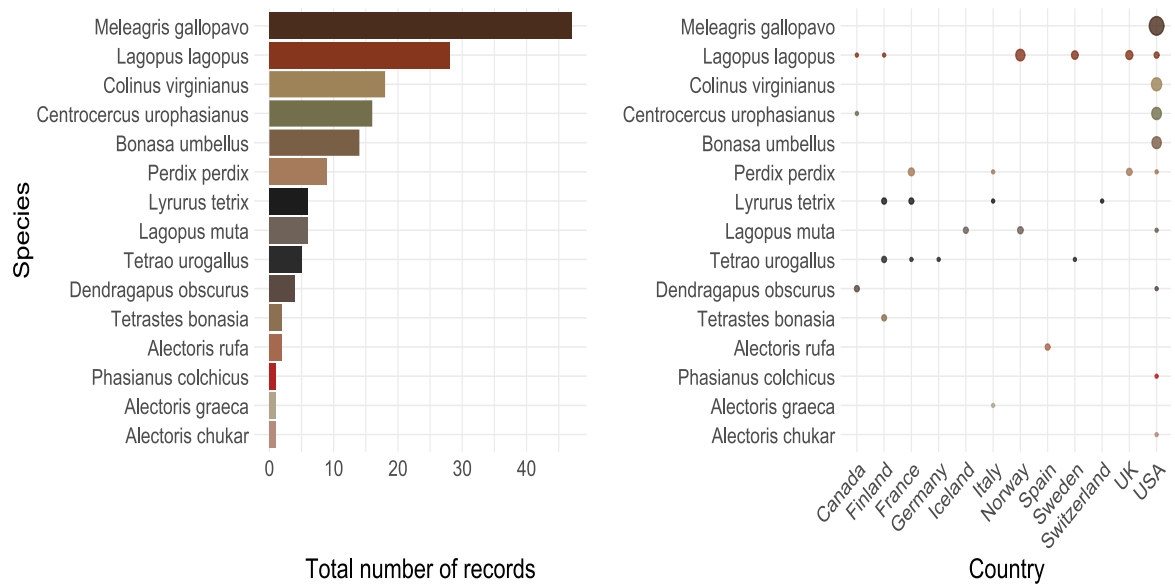
**Figure 7:** The number of studies for each EBV (Essential Biodiversity Variable) name and EBV class. EBV name is given on the vertical axis, whereas EBV class is indicated by the color of the bar.

Among studies that were based on (mathematical or statistical) simulations (Figure 8), the most common approach was to compare different harvest strategies (22 records), followed by conducting population viability analyses (17 records) and various forms of sensitivity analyses (13 records). In total, 11 studies were aiming to estimate/calculate maximum sustainable yield. Note that many studies used several of these approaches, so that the combined numbers will be higher than the total of 35 papers that included simulations or scenario analysis.



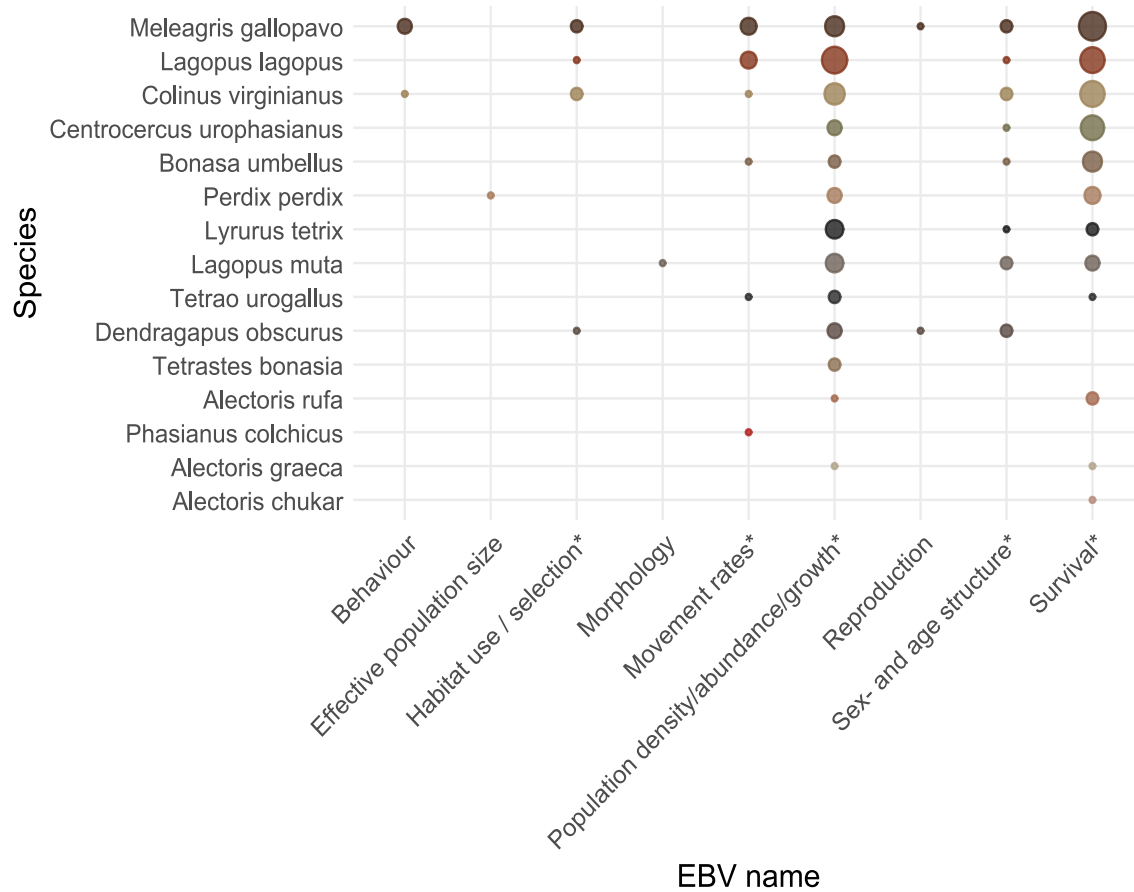
**Figure 8:** The purpose / outcome variable of interest of included simulation studies

The number of studies differed notably among species (Figure 9): The most frequently studied species were wild turkeys (*Meleagris gallopavo*; 33 records), followed by willow ptarmigan/grouse (*Lagopus lagopus*; 28 records), bobwhite (*Colinus virginianus*; 18 records) and sage grouse (*Centrocercus urophasianus*; 18 records). The remaining 11 species that were represented among the eligible papers each had less than 10 papers. Both capercaillie (*Tetrao urogallus*; 4 records) and black grouse (*Lyrurus tetrix*; 6 records) are among these poorly studied species.



**Figure 9:** The number of studies for each species (left panel) and country-by-species (right panel)

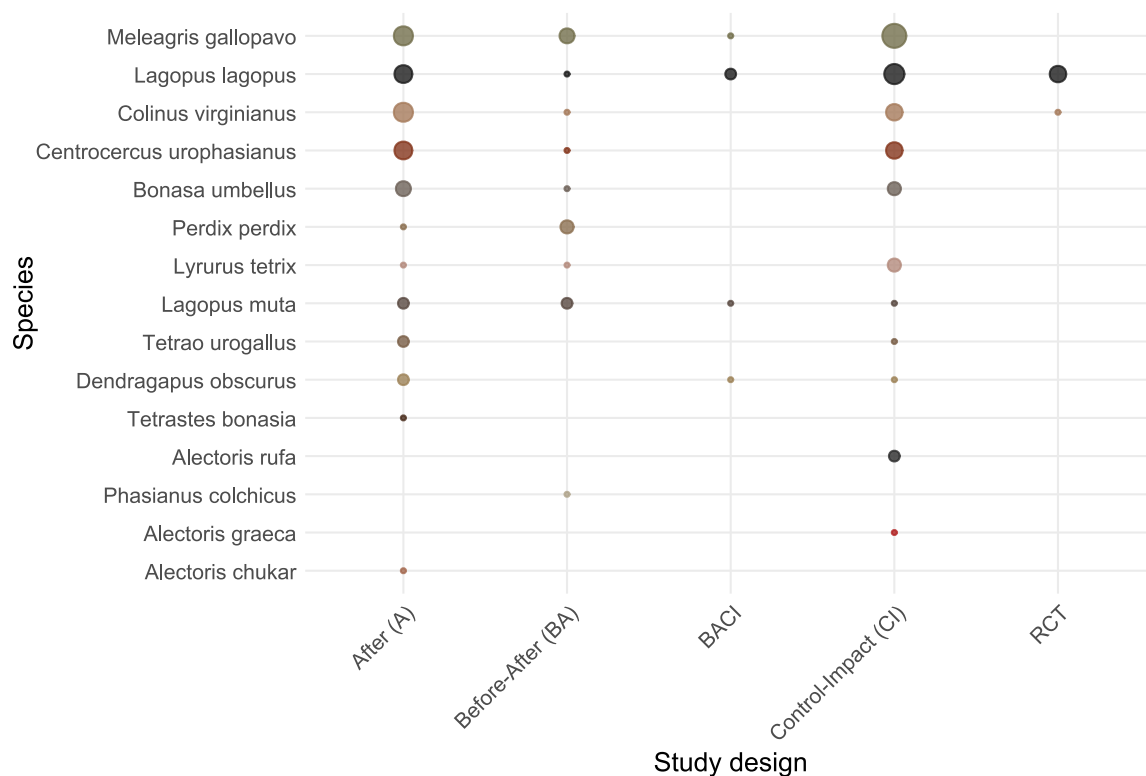
Different species were also subject to research focusing on different outcome variables (Figure 10). Wild turkey was the most well-studied species in the sense of number of different outcome variables (7 outcome variables), with bobwhite (6 outcome variables) and willow ptarmigan/grouse (5 outcome variables) following closely. Research on capercaillie and black grouse included 3 outcome variables. Two of the outcome variables had only been focused upon in two studies on grey partridge (*Perdix perdix*; effective population size) and rock ptarmigan (*Lagopus muta*; morphology).



**Figure 10:** The distribution of EBV names by species

Only 6 studies were based the robust RCT approach, of which 5 were about willow ptarmigan/grouse and one about wild turkeys (Figure 11). Four species (wild turkey, willow and rock ptarmigan and the dusky grouse (*Dendragapus obscurus*)) had been studied using a before-after-control-impact (BACI) approach.





**Figure 11:** The design of observational studies by species (RCT = Randomised Controlled Trial, BACI = Before-After-Control-Impact).

## 2.3 Summary of the literature mapping

Overall, there has been a substantial amount of research on the effects of harvest on galliform species. However, our mapping of the literature revealed that the literature is dominated by non-experimental research and often conducted at a local or regional scale. We identified only two research projects that were based on an RCT approach, of which one of the projects (represented with five papers) was the former Norwegian project on willow ptarmigan carried out in the late 1990s. Moreover, while the willow ptarmigan (also known as willow grouse) was well represented in the literature (28 records), there was little research on capercaillie (4 records) and black grouse (6 records), and no experimental assessments of the effects of harvest were identified.

### 3 Experimental design

The planned harvest experiment will rely on the existing line transect surveys gathered through the grouse inventory portal Hønsefuglportalen (<https://honsefugl.nina.no/Innsyn>). In cooperation with Statskog, the state enterprise managing state-owned forest and mountain properties in Norway, the planned experiment will have access to a wide range of hunting areas in Norway covering a latitudinal gradient between 58 and 70 degrees north (from Agder in South to Troms in North).

In this chapter, we will provide the basis for the experimental study design for the planned hunting experiment. We will start by defining some key terms (**BOX: Glossary of terms**), and then put this into the context of the planned experiment. First, we will define the harvest treatments in terms of harvest levels. Then, we will present the results from a simulation exercise that was conducted to assess the required sample sizes. Finally, we will present a detailed description of how study units (i.e., hunting areas) were included based on a pre-determined set of criteria. A complete list of study units will be presented together with key characteristics as an *Appendix 2*.

The setup of the experiment must be carefully designed so that the potential effects of both climate and harvest on ptarmigan and forest grouse populations will be detectable in the study. This means that the experimental design must include 1) areas with enough climatic variation and 2) large enough samples, i.e., number of study units, to statistically estimate the effects of the treatments (harvest rates) and covariates (climatic variables) on survival and population growth rate. Here, we have done a power analysis using simulations, with reported values of survival and population density from the literature, to estimate the number of areas needed to get the statistical power to test the effects of experimental treatments and climate.

**BOX: Glossary of terms****General terms**

**Outcome variable:** This is the variable of interest, i.e., the one to draw inference about. Defining the outcome is part of the planning of an experiment. In our harvest experiment, the outcome variables of interest include the effects of experimental harvest on i) survival and ii) population growth.

**Study population:** The population we will draw inference about. In strict terms, this is the willow ptarmigan population within the geographical boundaries of our study. Assuming that the willow ptarmigan population in our study area responds in a similar way to harvest mortality as other populations, we can make predictions about how the response will be in other areas. By including *covariates* that affect the outcome (response to harvest) these predictions can be made more accurate and less biased.

**Statistical population:** A statistical population is the set of subjects that we sample from and make inference from. In our case, the statistical population is the harvest areas that are included within the study area. One subject in the statistical population from which we draw some measurement is typically named a statistical unit or a sampling unit.

**Sampling frame:** The sampling frame is the “collection” of all study subjects that are available for inclusion. A structured way of thinking about this is to think of a regular grid covering the whole study area. However, in our case the grid is not regular but consists of harvest areas of different sizes. In addition, we will have to make other adjustments to the sampling frame: i) only areas with “sufficient” line transect survey data are available for inclusion, and ii) only areas larger than a cut-off size. These adjustments and their consequences should be considered before we make the final decision, as they can induce bias (see below).

**Blocked study design:** In blocked designs, the study units are grouped (blocked) by some characteristics. In a *randomized block design*, the treatment level is randomly assigned within the blocks. A block design typically means that the sample size must be larger in order to obtain the same statistical power as in non-blocked designs.

**Internal and external validity:** The term *internal validity* refers to the extent the inference drawn from the sample, e.g. in terms of cause-effect relationships, are true for the sample itself. The term external validity relates to the extent one can generalise from the current sample to the larger population that we usually want to make statements about. While the former is largely determined by the robustness of the methods and how treatment levels are assigned etc., the latter is largely determined by the sampling design and to which extent the included units are representative for the population.

**Recruitment/inclusion of study subjects:** The rules or algorithms chosen when recruiting / including study subjects. Since we cannot usually measure all subjects in the population, this involves some form of probabilistic or non-probabilistic sampling. These should be constructed to minimize the potential for *biased* estimates of inference. For instance, if all included subjects share a common characteristic that is not shared with other members of the statistical population, then the results might be biased. This is termed selection bias. Selection bias occurs whenever the subjects that are included in the study systematically deviate from the study population. Note that selection bias might also apply when assigning treatment level to a study unit in an experiment (see below).

**Probabilistic sampling to recruit study subjects:** This involves recruiting units into the study by some sort of random (probabilistic) sampling. In the basic form (i.e. *random sampling*), all units have the same probability of being recruited. In *stratified random sampling*, the units are sorted along one or more covariates to ensure that all values along the covariate are represented with a certain

proportion in the sample. Under proportional stratification, all strata are represented with a proportion that is similar to their proportion in the statistical population. Under balanced stratification, the number of recruited units are balanced (e.g. similar) across strata. Under restricted random sampling, the initial sampling might be random but it is combined with rules that make the next decisions non-probabilistic. Examples include “do not include units that are located next to each other” etc. Another sampling strategy that combines probabilistic and non-probabilistic sampling is systematic sampling when the selection of the first unit is decided by a random draw.

**Non-probabilistic selection to recruit study subjects:** These approaches include methods that use non-probabilistic criteria to select study subjects. Such criteria may include geography, certain characteristics (i.e, placement along a covariate axis but with no random selection, as opposed to stratified random sampling) and expert knowledge. A very specific type of non-probabilistic selection is convenience sampling where the units that for some reason are “easiest” to include are selected. One method that further resembles stratified random sampling is what is often termed proportional quota sampling in the medical literature. Here, the number of units that are included are proportional to the size of the strata, but the selection of units to include are not done at random. In cases where the total size of the population from which to draw sampling units from is unknown, this method will be known as non-proportional quota sampling. Finally, purposive sampling is a rather broad class of approaches where each study unit is recruited into the study based on one or more specific characteristics. In strict terms, a non-probabilistic selection of study subjects does not rely on a sampling frame.

**Assignment of treatment level to each study unit:** In manipulative experiments, there also needs to be rules for assigning treatment level to the study subjects. This can be done by random assignment or non-random (quasi random) assignment. Under random assignment of experimental treatment level, the treatment level (or control vs treatment) is randomly assigned to each unit. This is efficient in terms of assuring internal validity, as there will be no confounding effects from covariates. In blocked designs, treatment levels are randomly assigned within blocks. In a crossover study or trial, each study subject obtains a series of treatments allocated at random. In most cases, the study is designed so that all included subjects receive all treatments in a random order. This will result in a balanced study design, where the number of observations for each treatment are the same. This is in contrast to parallel studies or non-crossover studies where each unit is exposed to only one treatment level. For heterogeneous samples, one advantage with such a balanced design (depending on the analyses methods) is that each study subject can serve as their own control.

#### Key references

Christie et al. (2019) and Morrison et al. (2008) were used to develop the above text.

### 3.1 Experimental harvest treatment

In the planned experiment, we will base the harvest treatment for willow ptarmigan on the same experimental levels as the former large-scale experiment carried out in the 1990s in Norway (Sandercock et al. 2011): 0%, 15% and 30% harvest, respectively. This will make our results directly comparable to those results. For black grouse and capercaillie, we will apply 0%, 10% and 20% harvest. The rationale for this is that these species have a slower life history and higher natural survival rates, and thus we expect that harvest to a larger extent is additive for these species. For these species, we will also consider to carry out a parallel experiment where we assess the effects of protecting females (i.e. 0% harvest of females, but harvest allowed for males). This will be considered in year 2 and 3 of the planned experiment, and will not interfere with the detailed plans presented here.

### 3.1.1 Random allocation of harvest levels

Harvest treatment for each study unit will be randomly allocated based on a full crossover study design. As the planned experiment will have three experimental harvest levels (see above) and last for three seasons, each study unit will be subject to each of the three harvest treatment. The sequence of the treatments will be randomly decided for each study unit. Since there are three harvest treatments there will be six possible sequences of treatments for any given area (e.g., 30% first year, 0% second year and 15% third year, or 30% first year 15% second year and 0% third year, etc.). To ensure balance between northern and southern parts of the study area in terms of the relative frequency of treatment sequences, the study units were first sorted by latitude and separated into groups of six study units. This was done separately for willow ptarmigan and forest grouse units. Then, a set of the six unique sequences were generated, and random draws of treatment sequences were made without replacement within each group of study units. Thus, if there were e.g. 30 study units included in the experiment, each treatment sequence would be given for  $30/6 = 5$  study units spread across the study area. The random draws of treatment sequences are presented in *Appendix 3*.

### 3.1.2 Outcome variables

Based on the distance sampling line transect survey data collected from the study units, we will focus on the effects of harvest on two outcome variables, namely i) (annual) survival probability under different harvest treatments and ii) population growth under different harvest treatments. In addition to including information about harvest treatment, we will include information about climatic and biotic factors (in particular rodent dynamics) known to affect grouse population dynamics or that is expected to affect the effect of harvest on the outcome variables of interest.

We will base our inference on the model presented by Nilsen and Nater (2024) and Nater et al. (2024) to assess effects on survival. This model allows us to directly assess effects of harvest on survival rates, as it includes information about the population structure to estimate survival and recruitment rates separately. In this model, population growth rate is not estimated directly but is a derived parameter; thus we will use a modelling framework similar to that presented in (Bowler et al. 2020) to investigate effects directly on population growth.

### 3.1.3 Estimation of hunting quota

Each year, prior to the hunting season, a hunting quota that corresponds to the randomly allocated harvest level will be calculated based on the annual line transect surveys. The population density in a given hunting area is estimated using distance sampling techniques (Nilsen and Nater 2024, Bowler et al. 2020). The transect surveys are coordinated through Hønsfuglportalen (<https://honsfugl.nina.no/Innsyn>) and the results can be used to estimate the total population density as well as the number of adults and juveniles (that year's production of young). In many hunting areas the yearly production is a frequently used parameter in deciding the yearly quotas, but in the harvest experiment the quotas will solely be based on the total population size. For example, if the harvest treatment for a given year is 30% and the density is estimated as 15 birds/km<sup>2</sup> in an area of 100 km<sup>2</sup>, that year's hunting quota will be  $density \cdot area \cdot harvestlevel = 15 \cdot 100 \cdot 0.3 = 450$  birds.

## 3.2 Simulation to assess sample size

### 3.2.1 Population parameters

The willow ptarmigan populations that will be used in the harvest experiment are all surveyed in late summer, when the chicks are large enough to take flight. To simulate the effects of harvest and climate on willow ptarmigan survival and population growth rate, we used a post breeding projection matrix ( $A_t$ ) with two age classes: juveniles (0-1 year) and adults (1+ years):

$$A_t = \begin{bmatrix} f_{juv,t} & f_{ad,t} \\ s_{juv,t} & s_{ad,t} \end{bmatrix}$$

where,  $s_{juv,t}$  and  $s_{ad,t}$  are age-specific survival rates and  $f_{juv,t}$  and  $f_{ad,t}$  are recruitment (number of juveniles produced per female) multiplied with age-specific survival. A projection matrix contains both survival and reproduction for a population and can be used to calculate the population size for the next time-step,  $t + 1$ , directly from the age-specific population size at the last time step,  $t$ .

$$N_{t+1} = A_t N_t$$

where  $A$  is the projection matrix and  $N_t$  is a vector containing age-specific population sizes at time  $t$ . In these simulations we assumed equal recruitment for the two age classes. Annual adult survival estimates ( $S_{ad} = 0.54$ ,  $S_{juv} = 0.3$ ) were collected from Sandercock et al. (2011), while recruitment was modelled as a normally distributed variable with an average of two juveniles per female and SD = 0.3. The yearly population growth rate,  $r_t$ , was calculated as:

$$r_t = \log(N_{juv,t+1} + N_{ad,t+1}) - \log(N_{juv,t} + N_{ad,t})$$

where  $N_{juv}$  and  $N_{ad}$  is the number of juveniles and adults at time  $t$  and  $t+1$ . Both survival rates and population sizes were simulated with environmental stochasticity, i.e., year-to-year variation, and observation error to better reflect the real parameters that will be collected during the harvest experiment.

### 3.2.2 Harvest treatment and covariate

To evaluate and analyse the effects of harvest, we assigned the simulated hunting areas with a harvest treatment, i.e., a harvest rate of either 0%, 15% or 30% each year over the study period of three years. Because we aim to test the effects of all treatments in all areas, the treatments were randomly given to each hunting area without replacement to ensure that each area cannot have the same treatment twice. We assumed equal harvest rate for both juveniles and adults. Harvest mortality was added to the simulation by multiplying the age-specific survival rates by 1-harvest rate. In addition to harvest treatment, we simulated the effect of a climate covariate directly on survival. The “climatic variable” chosen was latitude and it was simulated with a slope of 0.003. This does not mean that we expect to see any direct effect of latitude in the real experiment, but rather it was included as a proof-of-concept to test the effects of a climatic covariate. The age-specific survival with harvest and the covariate effect were thus calculated as:

$$s'_{juv/ad,t} = s_{juv/ad,t} \times (1 - h_t) \times (\beta \times \text{latitude}),$$

where  $s'_{juv/ad,t}$  is the juvenile or adult survival after harvest and the effect of latitude. This gives a modified projection matrix ( $A'_t$ ) including the effects of harvest:

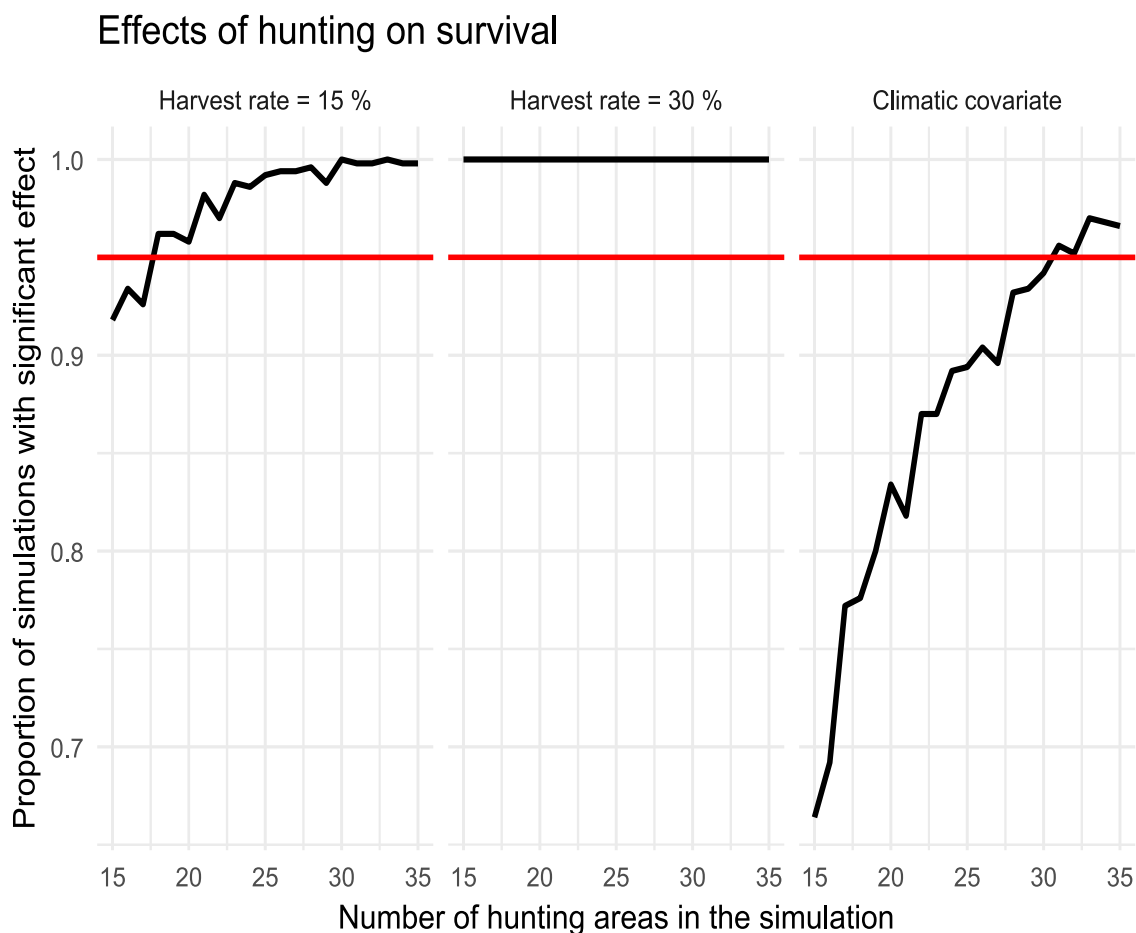
$$A'_t = \begin{bmatrix} f_{juv,t} & f_{ad,t} \\ s'_{juv,t} & s'_{ad,t} \end{bmatrix}$$

To investigate how many areas are needed in the harvest experiment to be able to estimate the effects of both harvest treatments and climate, we simulated ptarmigan populations over a range of hunting areas, from 15 to 35. For each number of hunting areas, we did 500 simulations over three years.

### 3.2.3 Results

#### Survival

The effects of harvest treatment and the climatic variable on survival was modelled using a generalized linear model with a quasibinomial error function. In the simulated harvest experiment, harvest treatment had a clear effect on survival, that was evident for all sample sizes (number of study units) for the 30% harvest, while the effect of 15% harvest was found in >95% of the simulations from 17 study units and upwards (Figure 12). The effect of the climatic variable was found in >84% of the simulations from 20 study units and upwards.



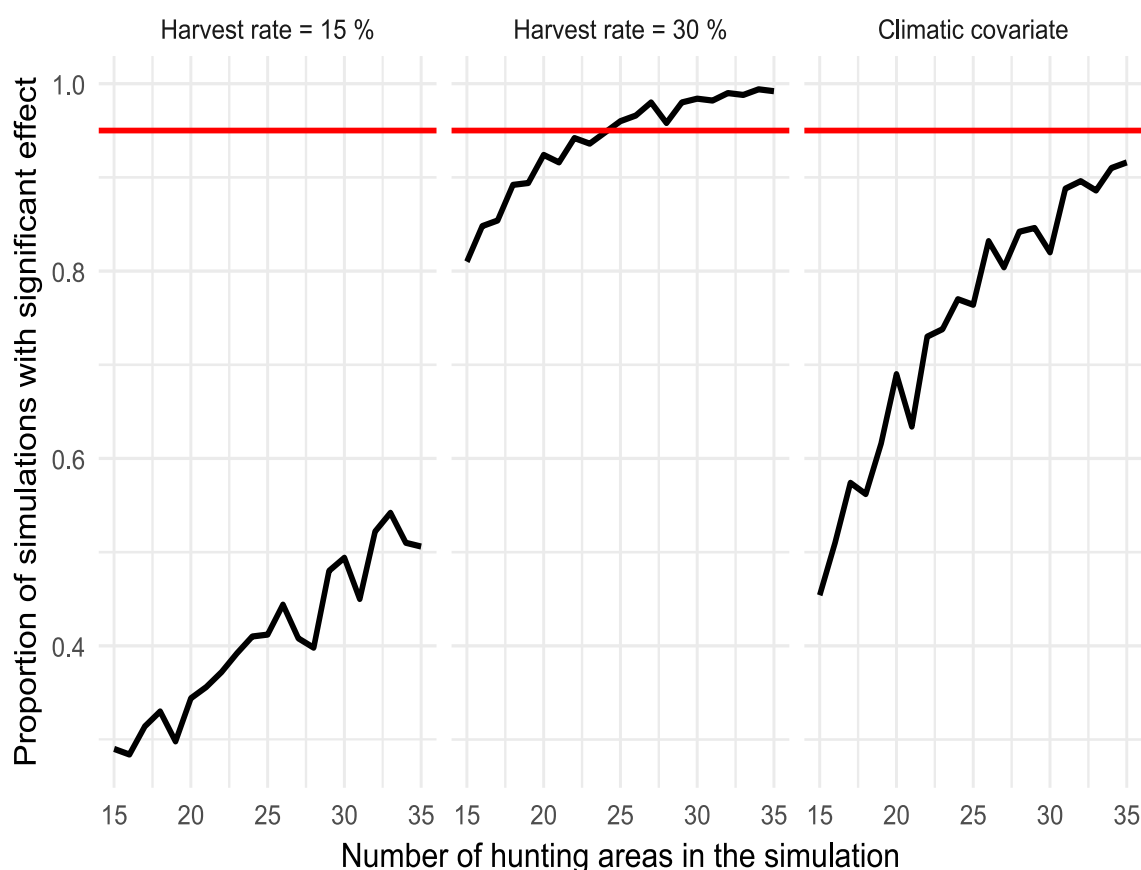
**Figure 12:** The proportion of simulations ( $n = 500$ ), over a range of hunting areas, that has a significant effect of harvest treatment (15 and 30 %) and a climatic variable on survival. The horizontal red line indicates 95% proportions of simulations with a significant effect.

### Population growth rate

The effects of harvest treatment and the climatic variable on the population growth rate was modelled using a linear regression. The 30% harvest treatment had an effect on the population growth rate in >95% of the simulations when the number of study units exceeded 21 (Figure 13). The effect of 15% harvest was harder to detect and it had an effect in about 50% of the simulations when the number of study units exceeded 28. The effect of the climatic variable on the population growth rate was found in >80% of the simulations from 27 study units and upwards.



## Effects of hunting on the population growth rate



**Figure 13:** The proportion of simulations ( $n = 500$ ), over a range of hunting areas, that has a significant effect of harvest treatment (15 and 30 %) and a climatic variable on the population growth rate. The horizontal red line indicates 95% proportions of simulations with a significant effect.

### 3.3 Inclusion of study units

Guided by the simulations above, we estimated a number of required study units necessary to detect the effects of harvest treatment and climatic variables on survival and population growth rates. As black grouse and capercaillie have slower life histories, with higher natural survival rates and lower reproductive potential than willow ptarmigan, we expect that it will be easier to detect effects of harvest for the two forest grouse species (see above). However, the planned harvest treatments for the forest grouse are lower (0, 10 and 20%) than for willow ptarmigan (0, 15 and 30%), and the lower harvest level should be expected to require a higher number of study units to detect effects. Based on the results from the simulations, we assumed that a minimum of 25 study units for each species would be satisfactory. In order to generalize from the sample to the larger population and to ensure high external validity, study units should be selected to avoid selection bias (see *Box: Glossary of terms*). For forest grouse we did not include protecting only female birds (i.e. 0% harvest on females) as a harvest treatment in the experiment, but we will investigate options and feasibility for such an expansion of the experiment after gaining experience from the first year of the study.

### 3.3.1 Criteria for inclusion of study units

We used specific rules when making decisions about inclusion of study units from the larger sampling frame. This was done to avoid selection bias, but the rules also included management considerations that disqualified units. Our approach resembles the “*Proportional quota sampling*” outlined in Box 1. The set of criteria we applied to the entire sampling frame of Statskog’s harvest management units ( $n = 390$ ) were the following:

1. Based on current knowledge of the ranges of the species, we set a minimum area size of 25 km<sup>2</sup> to be eligible for inclusion. This was done to reduce the risk of a bias in population growth estimates caused by dispersal between areas, especially by juvenile females (Brøseth et al. 2005; Hörnell-Willebrand et al. 2014). Pedersen et al. (2004) were able to find effects of harvest when using study units of 20 to 54 km<sup>2</sup> area size, which should imply that our minimum area size criteria is adequate. However, many hunting areas are also notably larger than this, with well over 100 km<sup>2</sup> of size. We did not set a maximum area size criteria, but we also did not include areas that were too large, as there would be challenges related to obtaining a representative distribution of line transects and harvest effort throughout the unit. In general, we opted to select a variety of area sizes that would reflect the variation in managed hunting areas in Norway.
2. We disqualified units where access to the unit is limited (e.g. because they are located far from roads or larger lakes where hunters could be transported with boats), as this would clearly interfere with the ability to obtain a sufficient amount of distance sampling line transects, or sufficient harvest effort to reach the desired treatment.
3. We disqualified units where it was not possible due to management concerns to perform a full randomization between the treatment levels for the full three-year period.
4. We disqualified units where the experimental set-up could be assumed to impose conflicts with indigenous sami reindeer herding.

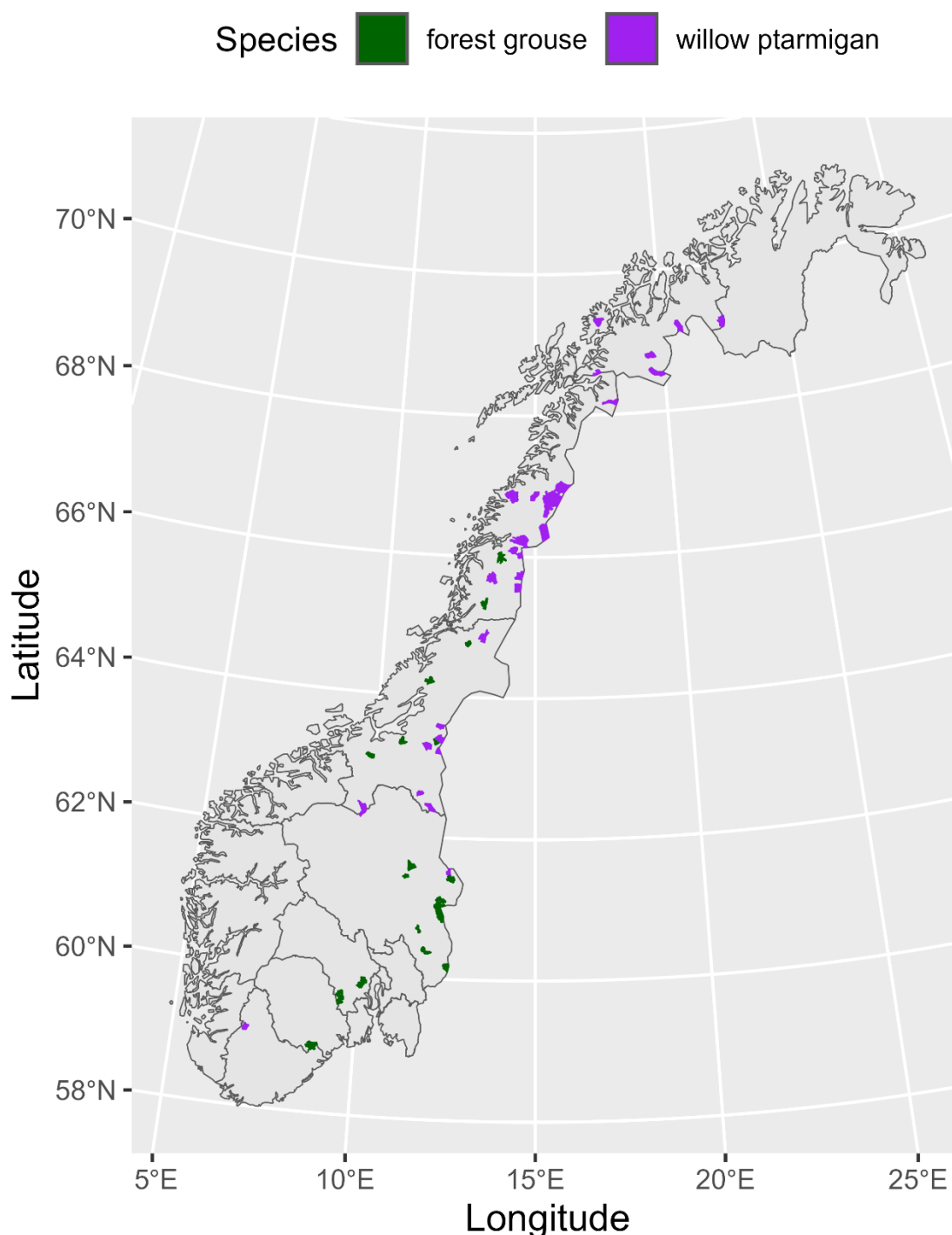
In cooperation with Statskog, we investigated options for qualifying study units that initially did not meet criteria 1 or 2, in particular when such units would give better coverage of geographical or climatic variation (cf. next section). This implies that if single management units did not meet the area size criteria, we created new units by combining two or more adjacent units when possible, as long as the combined unit constituted a coherent area for the species. Further, we investigated possibilities for re-allocating personnel resources for transect lines during the study period, to qualify a higher number of study units. In general, we opted for 40 km or more of line transect surveys in each included unit, but practical issues might result in slightly lower effort for some units.

Following the above set of criteria, 30 forest grouse units and 34 willow ptarmigan units were included as study units.

### 3.3.2 Key characteristics of study units

An important step in avoiding selection bias and ensuring high external validity is to ensure that statistical units do not systematically deviate from the study population. As such, we opted to select study units where the variation in central characteristics will reflect the general population of units in the sampling frame. The characteristics we considered included population density, area size and environmental variables in terms of climate and habitat.

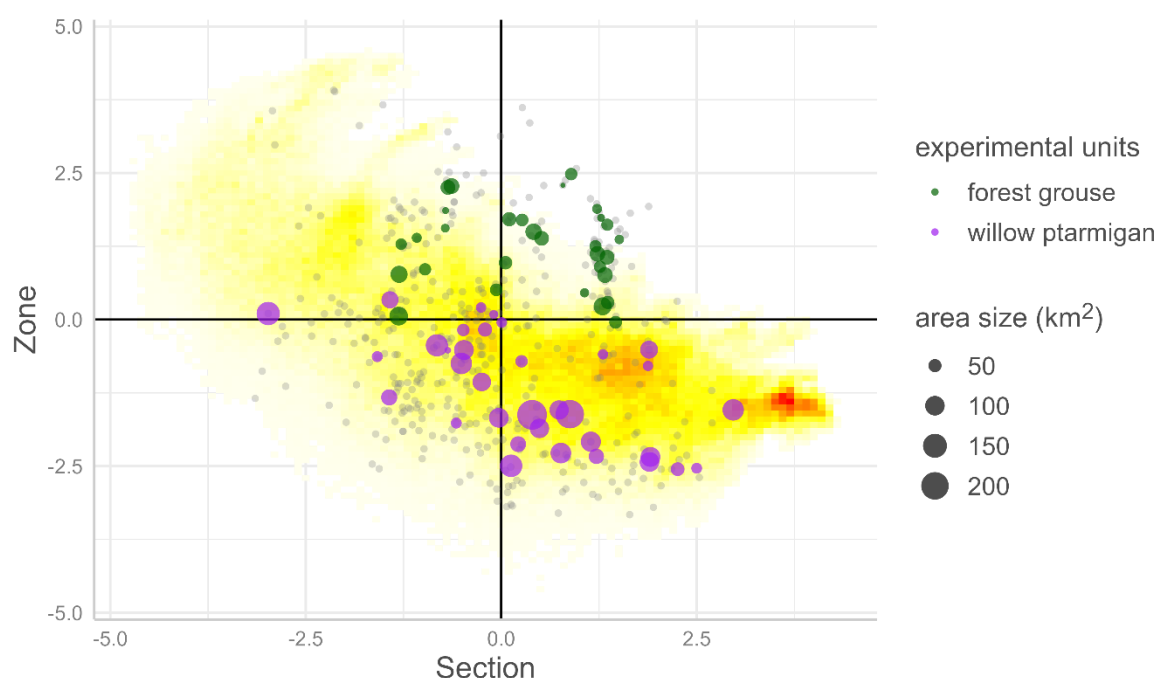
The included study units (Figure 14) were geographically distributed across a latitude gradient. As there tends to be spatial variation in willow ptarmigan population dynamics between the northern and southern parts of the country (Bowler et al. 2020), this will most likely mean that our ptarmigan study units will cover a wide range of population densities each year. Although the empirical basis is not as strong for the forest grouse species, any such geographical variation will be accounted for by the inclusion of study units along much of the latitudinal range of these species in Norway.



**Figure 14:** Study units in the harvest experiment in Norway, separated on willow ptarmigan and the two forest grouse species. In total there are 30 study units for forest grouse and 34 study units for willow ptarmigan.

Bakkestuen et al. (2008) assessed major environmental gradients in Norway using principal components analyses (PCA). They found that the two most important gradients were a regional gradient from humid to continental and coast to inland (PCA1) and a regional gradient from north to south and from high to low altitudes (PCA2). These two PCA axes corresponded with the previous classification of Norwegian biogeographical regions into a

humidity gradient, *bioclimatic sections*, and a temperature gradient, *bioclimatic zones* (Moen et al. 1999). We used these two gradients to assess whether our available study units represented the environmental bioclimatic variation in the larger population. To achieve this, we used the raster map from Bakkestuen et al. (2008) with 1x1 km resolution to calculate mean values of both the PCA axes for each study unit, as well as each unit in the overall sampling frame of Statskog units. In addition, we used the habitat suitability map from Kvasnes et al. (2018) to quantify the bioclimatic variation within suitable willow ptarmigan habitat in all of Norway. The end result can be seen in Figure 15.



**Figure 15:** Willow ptarmigan and forest grouse experimental study units, placed along gradients of bioclimatic sections and zones based on the mean value for each unit. Higher section values = drier bioclimatic region, and higher zone values = lower altitude and/or latitude. Grey dots show the total Statskog harvest management units to select from (not scaled by size). The heat map in the background shows amount of suitable habitat across Norway for willow ptarmigan only, with darker colours signifying higher amount of habitat within a given bioclimatic region.

Overall, Figure 15 gives confidence that we may generalise from our results beyond our study units. By visualising our willow ptarmigan study units together with the assumed bioclimatic variation in the larger population, we can verify that our sample of willow ptarmigan units cover much of the overall variation. There may be a slight underrepresentation among highly humid lowland areas (i.e. the upper left corner of the plot), which is usually considered less suitable among willow ptarmigan habitats (Kvasnes et al. 2018), and thus without Statskog areas that qualified for inclusion. It is also worth noting that the study units are presented with the mean of each bioclimatic gradient for each area, signifying that there is in many cases a proportion of values above and below the mean which is also represented in low amounts (i.e. lighter colour) in the heat map in Figure 15.

For forest grouse, the plot enables us to verify that the study units are satisfactory scattered across the vegetation sections and the boreal forest vegetation zones, within the range of Statskog forest grouse harvest units. There is no existing habitat suitability map for forest grouse in Norway, such as for willow ptarmigan. Nevertheless, the Statskog management units are located over vast parts of Norway, thus it seems fair to assume that they cover much of the bioclimatic variation of forest grouse hunting areas in Norway.

### **3.3.3 Calculation of area sizes**

The willow ptarmigan study units naturally include various amounts of habitat unsuitable for willow ptarmigan (e.g. lakes, lowland forest in the unit edges, or high mountain peaks). Thus, for willow ptarmigan we calculated area size of suitable habitat within the study units, following the classification by Kvasnes et al. (2018). For forest grouse, the study units are usually smaller and the total area of the selected management units are generally representative of the forest grouse habitat within each unit. Thus, for the forest grouse species we use the total area of the units as basis for the abundance calculations. A complete list of included study units is found in *Appendix 2*.

## 4 Conclusions

The systematic mapping of the literature revealed that, although there has been relatively extensive research on the effects of harvest on galliform species, there are several important key aspects that require further research and emphasis. First, the literature is dominated by research with study designs with a relatively weak ability to detect causal relationships (Christie et al. 2019), including *After*, *Before-After* and *Control-Impact*. Although these are often the only options that are available due to ethical, logistic and practical reasons, it is well known that manipulative experiments and in particular *Randomized Controlled Trials* (RCT) are better suited for drawing firm conclusions about causal relationships. This holds true even though recent advances in statistical methodology also allow to a larger extent to make causal inference from non-manipulative experiments (Law et al. 2017). In our systematic literature mapping, we only identified two research projects that relied on RCT: One study from Norway represented with five publications, and one study from USA on bobwhite. Note that the study on bobwhite applied “simulated” harvest, whereby the researchers removed birds from the population rather than “real” harvesting from the population. As documenting the relative additivity of harvest mortality under different environmental and demographic conditions is fraught with statistical difficulties, manipulative randomized experiments would be particularly important in order to progress our understanding of this topic.

Second, most studies had a relatively small or moderate spatial extent. As ecological dynamics are often context dependent (Bowler et al. 2020) and might vary both in time and space, and because some ecological properties play out over larger areas, this might limit our ability to draw firm conclusions that could be transferred to other ecological conditions (and species). Our planned study will be unique not only because it has a (near) national spatial extent, but also because it covers three related but ecologically distinct species. While there has been a fair amount of research on willow ptarmigan related to effects of harvest (including the before-mentioned RCTs from the 1990s), the amount of research on black grouse and capercaillie related to effects of harvest is very limited. This highlights that the current management of these species is not based on a solid knowledge base about harvest sustainability, as a contrast to willow ptarmigan management. Overall, the project described here is likely to fill several key knowledge gaps in the literature, and will be highly relevant for harvest management well beyond the study context. In addition, the design of the study allow us to also make substantial contributions about the population dynamics of grouse species in general.

We used mathematical simulations to assess the sample sizes (i.e., the number of study units) that would be needed in order to detect effects of the prescribed experimental treatments (harvest levels of 0%, 15% and 30% for willow ptarmigan, and 0%, 10% and 20% for the forest grouse species). The exact power to detect effects will depend on both the consistency of the treatment effects in time and space (e.g. due to ecological correlates), the precision and accuracy of the line transect survey data, and the specific statistical models that will be applied. In addition, we only performed simulations for willow ptarmigan, as we did not have comparable data to construct simulation models for forest grouse without relying on additional assumptions. However, as both forest grouse species have a slower life history than willow ptarmigan, we assume that sample sizes that are sufficient for willow ptarmigan are also sufficient for forest grouse. We found, as expected, that sample sizes needed to detect an effect of 30% harvest (compared to 0% harvest) was smaller than those needed to detect an effect of 15% harvest. Moreover, the sample sizes needed to detect a direct effect on population growth rate was higher than those for detecting effects on survival. In the former harvest experiment in Norway, the sample size was approximately 13 areas followed for 3-4 years (most areas followed for 4 years), and it was concluded that such large sample sizes would be needed to detect the relevant effects due to large heterogeneity among areas (Pedersen et al. 2004). Thus, in our planned experiments we opted to recruit

34 study units for willow ptarmigan and 30 study units for forest grouse, each subject to a randomised three year harvest experiment.

To select study units we applied a set of criteria and algorithms. We did not randomly select study units from the sampling grid, but used an approach resembling the *proportional quota sampling* described in Box 1. In addition to certain key requirements in terms of area size and ability to achieve the required amount of distance sampling line transect data, we actively selected study units to give a balanced geographical coverage both in terms of a north-south axis, but also in terms of environmental covariates and gradients. As the main environmental gradients against which we stratified our sampling of study units, we used the *bioclimatic sections* and *bioclimatic zones* in Norway (see Bakkestuen et al. 2008; Moen et al. 1999). Overall, we think these variables should give us an adequate distribution of study units in terms of key environmental gradients.

Each study unit will be subject to a randomly allocated harvest treatment. In the planned experiment, we use a crossover-design where each unit is subject to a randomly selected sequence of harvest treatments. Because we are planning for a three-year experiment and there are three harvest levels, each unit will be exposed to each treatment. If there is substantial heterogeneity, crossover-designs could be powerful as each unit serves as “it’s own control”. In addition, by applying a sequence of treatments we avoid any adverse effects on population level that could arise if for instance one unit was subject to the highest harvest treatment across all years.

In general, we believe that the experimental study of the effects of harvest on willow ptarmigan, black grouse and capercaillie as outlined here will be a unique and substantial contribution to the field. The systematic mapping of the literature clearly revealed that the knowledge that would be created by this study will fill important knowledge gaps. The final list of study units for inclusion also suggest that the study will be of adequate power and cover important environmental gradients in a robust way, so that the results will be highly transferable across environmental and geographical space.



## 5 References

- Bakkestuen, V., Erikstad, L., and Halvorsen, R. 2008. Step-less models for regional environmental variation in Norway. *Journal of biogeography* **35**(10): 1906–1922. Wiley Online Library.
- Bowler, D.E., Kvasnes, M.A.J., Pedersen, H.C., Sandercock, B.K., and Nilsen, E.B. 2020. Impacts of predator-mediated interactions along a climatic gradient on the population dynamics of an alpine bird. *Proc Biol Sci* **287**: 20202654. doi:10.1098/rspb.2020.2653.
- Brøseth, H., Tufto, J., Pedersen, H.C., Steen, H. and Kastedalen, L. 2005. Dispersal patterns in a harvested willow ptarmigan population. *Journal of Applied Ecology* **42**: 453–459.
- Christie, A.P., Amano, T., Martin, P.A., Shackelford, G.E., Simmons, B.I., and Sutherland, W.J. 2019. Simple study designs in ecology produce inaccurate estimates of biodiversity responses. *Journal of Applied Ecology* **56**(12): 2742–2754. Wiley Online Library.
- Eriksen, L.F., Moa, P.F., and Nilsen, E.B. 2018. Quantifying risk of overharvest when implementation is uncertain. *Journal of Applied Ecology* **55**(2): 482–493. doi: <https://doi.org/10.1111/1365-2664.12992>.
- Estes, L., Elsen, P.R., Treuer, T., Ahmed, L., Caylor, K., Chang, J., Choi, J.J., and Ellis, E.C. 2018. The spatial and temporal domains of modern ecology. *Nature Ecology & Evolution* **2**(5): 819–826. doi:10.1038/s41559-018-0524-4.
- Fraser, H., Barnett, A., Parker, T.H., and Fidler, F. 2020. The role of replication studies in ecology. *Ecology and Evolution* **10**(12): 5197–5207. doi: <https://doi.org/10.1002/ece3.6330>.
- Grainger, M.J., Bolam, F.C., Stewart, G.B., and Nilsen, E.B. 2020. Evidence synthesis for tackling research waste. *Nature Ecology and Evolution* **4**: 495–497. doi: <https://doi.org/10.1038/s41559-020-1141-6>.
- Haddaway, N.R., Macura, B., Whaley, P., and Pullin, A.S. 2018. ROSES RepOrting standards for systematic evidence syntheses: Pro forma, flow-diagram and descriptive summary of the plan and conduct of environmental systematic reviews and systematic maps. *Environmental Evidence* **7**. doi:10.1186/s13750-018-0121-7.
- Henden, J.-A., Ims, R.A., Yoccoz, N.G., Asbjørnsen, E.J., Stien, A., Mellard, J.P., Tveraa, T., Marolla, F., and Jepsen, J.U. 2020. End-user involvement to improve predictions and management of populations with complex dynamics and multiple drivers. *Ecological Applications* **30**(6): e02120. doi: <https://doi.org/10.1002/eap.2120>.
- Hörnelt-Willebrand, M., Willebrand, T. and Smith, A.A. 2014. Seasonal Movements and Dispersal Patterns: Implications for Recruitment and Management of Willow Ptarmigan (*Lagopus lagopus*). *Journal of Wildlife Management* **78**: 194–201. doi: 10.1002/jwmg.650.
- Johnson, F.A., Moore, C.T., Kendall, W.L., Dubovsky, J.A., Caithamer, D.F., Kelley, J.R., and Williams, B.K. 1997. Uncertainty and the management of mallard harvests. *Journal of Wildlife Management* **61**: 202–216. doi:10.2307/3802429.
- Kvasnes, M.A.J., Pedersen, H.C., and Nilsen, E.B. 2018. Quantifying suitable late summer brood habitats for willow ptarmigan in Norway. *BMC ecology* **18**: 1–13. Springer.

- Law, E. A., P. J. Ferraro, P. Arcese, B. A. Bryan, K. Davis, A. Gordon, M. H. Holden, G. Iacona, R. M. Martinez, C. A. McAlpine, J. R. Rhodes, J. S. Sze, and K. A. Wilson. 2017. Projecting the performance of conservation interventions. *Biological Conservation* **215**:142–151. doi: 10.1016/j.biocon.2017.08.029
- Månsson, J., Eriksson, L., Hodgson, I., Elmberg, J., Bunnefeld, N., Hessel, R., Johansson, M., Liljebäck, N., Nilsson, L., Olsson, C., Pärt, T., Sandström, C., Tombre, I., and Redpath, S.M. 2023. Understanding and overcoming obstacles in adaptive management. *Trends in Ecology & Evolution* **38**(1): 55–71. doi: <https://doi.org/10.1016/j.tree.2022.08.009>.
- Moen, A., Lillethun, A., and Odland, A. 1999. *Vegetation: National atlas of norway*. Norwegian Mapping Authority, Hønefoss, Norway.
- Morrison, M.L., Block, W.M., Strickland, M.D., Collier, B.A., and Peterson, M.J. 2008. *Wild-life study design*. Springer Science & Business Media.
- Nakagawa, S., and Parker, T.H. 2015. Replicating research in ecology and evolution: Feasibility, incentives, and the cost-benefit conundrum. *BMC Biology* **88**(18): 164–173. doi:10.1186/s12915-015-0196-3.
- Nater, C. R., F. Frassinelli, J. A. Martin, and E. B. Nilsen. 2024. Large-scale spatio-temporal variation in vital rates and population dynamics of an alpine bird. *Peer Community Journal* **4**. doi: 10.24072/pcjournal.494
- Navarro, L.M., Fernandez, N., Guerra, C., Guralnick, R., Kissling, W.D., Londoño, M.C., Muller-Karger, F., Turak, E., Balvanera, P., Costello, M.J., and others. 2017. Monitoring biodiversity change through effective global coordination. *Current opinion in environmental sustainability* **29**: 158–169. Elsevier.
- Nilsen, and Nater, C.R. 2024. An integrated open population distance sampling approach for modelling age-structured populations. *EcoEvoRxiv*. doi: <https://doi.org/10.32942/X2Q899>.
- Ouzzani, M., Hammady, H., Fedorowicz, Z., and Elmagarmid, A. 2016. Rayyan—a web and mobile app for systematic reviews. *Systematic reviews* **5**: 1–10. Springer.
- Pedersen, H.C., Steen, H., Kastedalen, L., Brøseth, H., Ims, R.A., Svendsen, W., and Yoccoz, N.G. 2004. Weak compensation of harvest despite strong density-dependent growth in willow ptarmigan. *Proceedings of the Royal Society of London. Series B: Biological Sciences* **271**(1537): 381–385. doi:10.1098/rspb.2003.2599.
- Pereira, H.M., Ferrier, S., Walters, M., Geller, G.N., Jongman, R.H., Scholes, R.J., Bruford, M.W., Brummitt, N., Butchart, S.H., Cardoso, A., and others. 2013. Essential biodiversity variables. *Science* **339**(6117): 277–278. American Association for the Advancement of Science.
- Péron, G. 2013. Compensation and additivity of anthropogenic mortality: Life-history effects and review of methods. *Journal of Animal Ecology* **82**(2): 408–417. doi: <https://doi.org/10.1111/1365-2656.12014>.
- Sandercock, B., Nilsen, E.B., Brøseth, H., and Pedersen, H.C. 2011. Is hunting mortality additive or compensatory to natural mortality? Effects of experimental harvest on the survival and cause-specific mortality of willow ptarmigan. *Journal of Animal Ecology* **80**(1): 244–258. doi: 10.1111/j.1365-2656.2010.01769.x.

Sedinger, J.S., and Rotella, J.J. 2004. Effect of harvest on sage-grouse *centrocercus urophasianus* populations: What can we learn from the current data? *Wildlife Biology* **11**: 371–375. doi:10.2981/0909-6396(2005)11[371:EOHOSC]2.0.CO;2.

Sinclair, A.R.E., and Pech, R.P. 1996. Density dependence, stochasticity, compensation and predator regulation. *Oikos* **75**(2): 164–173. [Nordic Society Oikos, Wiley]. doi:10.2307/3546240.

Small, R.J., Holzward, J.C., and Rusch, D.H. 1991. Predation and hunting mortality of ruffed grouse in central Wisconsin. *Journal of Wildlife Management* **55**: 512–520.

# Supplementary Information

## Appendix 1

Additional information about the systematic review:

### Decision tree for title and abstract screening:

5. Is there an abstract for the article?
  - a) Yes - Go to 3.
  - b) No - Go to 2
6. Can you find the abstract online? (Use google.com or Google Scholar to search for the title of the article)
  - a) Yes - Enter the url for the website where you found the abstract in the 'Add note' box and Go to 3
  - b) No – Add/select 'Abstract not found online' in the 'Exclude with Reason' box.
7. Is the abstract in English/Norwegian/Danish/Swedish?
  - a) Yes - Go to 4.
  - b) No – Add/select 'Abstract not in English/Norwegian/Danish/Swedish' in the 'Exclude with Reason' box.
8. Does the abstract represent a research study focused on Galliformes?
  - a) Yes - Go to 5
  - b) No – Add/select 'Not Galliformes' in the 'Exclude with Reason' box.
9. Does the study document the effects of legal and/or managed harvest on Galliformes populations
  - a) Yes - Go to 6
  - b) No – Add/select 'Not Hunting' in the 'Exclude with Reason' box.
10. Are effects assessed quantitatively?
  - a) Yes – Press the include button
  - b) No - Add/select 'Not quantitative' in the 'Exclude with Reason' box.

## Data coding

From all papers that met the eligibility criteria, we extracted key variables (aka “data coding”). In addition to metadata about the publication (publication year, journal name etc) we extracted the following information:

- Species
- Study area location
- Country/countries
- County/counties
- Municipality/municipalities
- Study extent / size of study area
  - Multinational
  - National
  - Regional (e.g. a county, a larger mountain area etc)
  - Local (e.g. a municipality or part of municipality)
  - Site level (< 1km<sup>2</sup>)
- Research type
- Observational / analyses of empirical data
- Simulation based / scenario
- Study design – observational studies
- Randomized Controlled Trials (RCTs)
- Before-After-Control-Impact (BACI)
- Control-Impact (CA)
- Before-After (BA)
- After (A)
- Harvest level
  - Boolean (Yes / No)
- Assessment of additivity of harvest mortality
  - Boolean (Yes / No)
- Outcome variables: EBV class
  - Genetic composition
  - Species population
  - Species trait
- Outcome variable: EBV name [not exhaustive; \* indicate name that is not included in original framework]
  - Inbreeding
  - Effective population size
  - Species distribution
  - Species abundance
  - Population density \*
  - Morphology

- Physiology
- Movement rates\*
- Habitat use / selection\*
- Reproduction
- Survival\*
- Sex- and age structure\*

## Appendix 2

**Table S2:** Overview of included study units. WP area refers to the area of suitable willow ptarmigan habitat (not relevant for forest grouse areas). PCA1 and PCA2 refers to the axes defining bioclimatic sections and bioclimatic zones.

Experimental unit	Species	Total area	WP area	PCA1 (mean)	PCA2 (mean)
Bangdalen1_E_FG	forest grouse	35.8	NA	-1.08	1.40
Bangdalen2_E_FG	forest grouse	41.6	NA	-1.28	1.29
Finnemarka1_E_FG	forest grouse	47.7	NA	0.27	1.70
Finnemarka2_E_FG	forest grouse	56.4	NA	0.10	1.71
Gjerstad1_E_FG	forest grouse	66.5	NA	-0.64	2.28
Gjerstad2_E_FG	forest grouse	60.2	NA	-0.68	2.25
Gran19_E_FG	forest grouse	77.1	NA	-1.31	0.77
Gravberget1_E_FG	forest grouse	65.9	NA	1.33	0.76
Gravberget2_E_FG	forest grouse	44.9	NA	1.26	0.90
Gravberget3_E_FG	forest grouse	59.8	NA	1.35	1.06
Gravberget4_E_FG	forest grouse	65.4	NA	1.23	1.12
Gravberget5_E_FG	forest grouse	44.6	NA	1.21	1.25
Gravberget6_E_FG	forest grouse	44.8	NA	1.36	1.62
Hemn3_E_FG	forest grouse	91.0	NA	-1.31	0.06
Ljørdalen_Drevja_E_FG	forest grouse	84.5	NA	1.30	0.23
Løten_E_FG	forest grouse	34.1	NA	1.51	1.37
Malvik1_E_FG	forest grouse	28.2	NA	-0.71	1.86
Malvik2_E_FG	forest grouse	32.0	NA	-0.72	1.56
Meitsjøen1_E_FG	forest grouse	30.0	NA	1.28	1.74
Meitsjøen2_E_FG	forest grouse	35.6	NA	1.23	1.89
Meråker_Dalavola_E_FG	forest grouse	47.4	NA	-0.06	0.51
Namsskogan1_E_FG	forest grouse	45.4	NA	-0.97	0.86
Orkland_E_FG	forest grouse	51.9	NA	0.05	0.97
Rendalen1_E_FG	forest grouse	33.3	NA	1.07	0.46
Rendalen2_E_FG	forest grouse	50.6	NA	1.36	0.29
Rendalen3_E_FG	forest grouse	48.0	NA	1.46	-0.05
Sølvverket1_E_FG	forest grouse	59.8	NA	0.52	1.39
Sølvverket2_E_FG	forest grouse	71.7	NA	0.41	1.50
Varaldskogen1_E_FG	forest grouse	46.4	NA	0.89	2.48
Varaldskogen2_Særkilampi_E_FG	forest grouse	27.0	NA	0.79	2.29
Altevatnet_E_WP	willow ptarmigan	146.0	100.5	1.90	-2.43
Bangfjellet_Dividalen_E_WP	willow ptarmigan	101.5	63.7	1.22	-2.33
Drivstua_N_E_WP	willow ptarmigan	76.6	53.4	2.26	-2.55
Drivstua_Sør_E_WP	willow ptarmigan	48.2	38.7	2.50	-2.54
Gran20_E_WP	willow ptarmigan	155.1	128.4	-0.82	-0.44

Hatt2_E_WP	willow ptarmigan	104.0	96.4	0.74	-1.54
Hatt4_E_WP	willow ptarmigan	106.4	98.0	0.49	-1.86
Helligskogen_E_WP	willow ptarmigan	118.7	101.9	1.91	-2.34
Hemn11_E_WP	willow ptarmigan	45.5	39.4	-0.57	-1.76
HemnKj_E_WP	willow ptarmigan	86.7	69.7	-1.43	-1.33
Ljørdalen_Drevfjellet_E_WP	willow ptarmigan	37.5	37.1	1.30	-0.59
Meråker1_E_WP	willow ptarmigan	55.8	46.5	-0.48	-0.18
Meråker2_E_WP	willow ptarmigan	39.6	32.8	-0.10	0.09
Meråker3_E_WP	willow ptarmigan	47.2	46.1	0.26	-0.71
Meråker4_E_WP	willow ptarmigan	60.1	55.6	-0.21	-0.17
Meråker5_E_WP	willow ptarmigan	45.1	37.4	-0.26	0.21
Meråker6_E_WP	willow ptarmigan	46.8	38.0	0.01	-0.05
Namsskogan2_E_WP	willow ptarmigan	115.2	103.1	-0.48	-0.52
Njardarheim_E_WP	willow ptarmigan	57.8	39.4	-1.59	-0.63
Nordalen_E_WP	willow ptarmigan	92.7	66.8	0.22	-2.12
Ran10_E_WP	willow ptarmigan	142.6	110.7	0.76	-2.28
Ran2_E_WP	willow ptarmigan	135.9	119.8	-0.51	-0.75
Ran3_E_WP	willow ptarmigan	117.0	101.1	-0.03	-1.67
Ran9_E_WP	willow ptarmigan	115.1	105.5	1.15	-2.08
Reisavann_E_WP	willow ptarmigan	122.4	119.3	2.97	-1.54
Røros_Flensmarka_E_WP	willow ptarmigan	91.7	84.0	1.89	-0.51
Røros_Vestre_E_WP	willow ptarmigan	38.8	36.6	1.88	-0.79
Salt1_E_WP	willow ptarmigan	230.4	141.5	-2.98	0.10
Salt2_E_WP	willow ptarmigan	95.1	90.1	-0.25	-1.06
Salt3_E_WP	willow ptarmigan	277.4	212.9	0.88	-1.61
Salt4_E_WP	willow ptarmigan	297.0	131.1	0.13	-2.49
Salt5_E_WP	willow ptarmigan	260.5	233.7	0.39	-1.62
Senja_Svandalen_E_WP	willow ptarmigan	102.0	77.9	-1.42	0.34
Skånland_Øst_E_WP	willow ptarmigan	47.0	27.8	-0.68	-0.52



## Appendix 3

**Table S3:** Overview of harvest treatment across the three-year experiment for each included study unit. The experiment is based on a complete crossover design, where each study unit is subject to a sequence of three different harvest treatments. Because there are three different harvest levels for each species, there is a total of six unique sequences for willow ptarmigan and forest grouse, respectively. Prior to making the random draws, the study units were first sorted by latitude, and grouped into groups of six units. This was done separately for willow ptarmigan and forest grouse units. Prior to the draws, a set of the six unique sequences were generated, and random draws of treatment sequences were made without replacement within each group. This was done to ensure balance between northern and southern parts of the study area in terms of the relative frequency of treatment sequences.

Experimental unit	Species	Sequence of harvest treatments	Group
Gjerstad2_E_FG	forest grouse	20-10-0	1
Gjerstad1_E_FG	forest grouse	0-20-10	1
Sølvverket2_E_FG	forest grouse	10-0-20	1
Sølvverket1_E_FG	forest grouse	0-10-20	1
Finnemarka2_E_FG	forest grouse	10-20-0	1
Finnemarka1_E_FG	forest grouse	20-0-10	1
Varaldskogen1_E_FG	forest grouse	10-0-20	2
Varaldskogen2_Særkilampi_E_FG	forest grouse	0-20-10	2
Meitsjøen2_E_FG	forest grouse	10-20-0	2
Meitsjøen1_E_FG	forest grouse	20-10-0	2
Løten_E_FG	forest grouse	20-0-10	2
Gravberget6_E_FG	forest grouse	0-10-20	2
Gravberget5_E_FG	forest grouse	0-20-10	3
Gravberget4_E_FG	forest grouse	10-0-20	3
Gravberget3_E_FG	forest grouse	20-10-0	3
Gravberget2_E_FG	forest grouse	20-0-10	3
Gravberget1_E_FG	forest grouse	10-20-0	3
Ljørdalen_Drevja_E_FG	forest grouse	0-10-20	3
Rendalen1_E_FG	forest grouse	20-10-0	4
Rendalen3_E_FG	forest grouse	10-0-20	4
Rendalen2_E_FG	forest grouse	0-10-20	4
Orkland_E_FG	forest grouse	0-20-10	4
Malvik2_E_FG	forest grouse	20-0-10	4
Meråker_Dalavola_E_FG	forest grouse	10-20-0	4
Malvik1_E_FG	forest grouse	20-10-0	5
Bangdalen1_E_FG	forest grouse	10-20-0	5
Bangdalen2_E_FG	forest grouse	0-20-10	5

Namsskogan1_E_FG	forest grouse	10-0-20	5
Gran19_E_FG	forest grouse	20-0-10	5
Hemn3_E_FG	forest grouse	0-10-20	5
Njardarheim_E_WP	willow ptarmigan	0-30-15	1
Ljørdalen_Drevfjellet_E_WP	willow ptarmigan	30-0-15	1
Drivstua_Sør_E_WP	willow ptarmigan	15-30-0	1
Røros_Flensmarka_E_WP	willow ptarmigan	30-15-0	1
Drivstua_N_E_WP	willow ptarmigan	15-0-30	1
Røros_Vestre_E_WP	willow ptarmigan	0-15-30	1
Meråker3_E_WP	willow ptarmigan	15-30-0	2
Meråker2_E_WP	willow ptarmigan	15-0-30	2
Meråker1_E_WP	willow ptarmigan	0-15-30	2
Meråker6_E_WP	willow ptarmigan	0-30-15	2
Meråker5_E_WP	willow ptarmigan	30-15-0	2
Meråker4_E_WP	willow ptarmigan	30-0-15	2
Namsskogan2_E_WP	willow ptarmigan	15-0-30	3
Hatt4_E_WP	willow ptarmigan	0-30-15	3
Gran20_E_WP	willow ptarmigan	0-15-30	3
Hatt2_E_WP	willow ptarmigan	30-0-15	3
Hemn11_E_WP	willow ptarmigan	30-15-0	3
HemnKj_E_WP	willow ptarmigan	15-30-0	3
Ran3_E_WP	willow ptarmigan	30-15-0	4
Ran2_E_WP	willow ptarmigan	15-0-30	4
Ran9_E_WP	willow ptarmigan	0-15-30	4
Ran10_E_WP	willow ptarmigan	0-30-15	4
Salt3_E_WP	willow ptarmigan	30-0-15	4
Salt4_E_WP	willow ptarmigan	15-30-0	4
Salt2_E_WP	willow ptarmigan	0-30-15	5
Salt1_E_WP	willow ptarmigan	30-0-15	5
Salt5_E_WP	willow ptarmigan	15-0-30	5
Nordalen_E_WP	willow ptarmigan	15-30-0	5
Skånland_Øst_E_WP	willow ptarmigan	0-15-30	5
Altevatnet_E_WP	willow ptarmigan	30-15-0	5
Bangfjellet_Dividalen_E_WP	willow ptarmigan	0-15-30	6
Helligskogen_E_WP	willow ptarmigan	15-0-30	6
Senja_Svandalen_E_WP	willow ptarmigan	30-15-0	6
Reisavann_E_WP	willow ptarmigan	15-30-0	6



*The Norwegian Institute for Nature Research, NINA, is as an independent foundation focusing on environmental research, emphasizing the interaction between human society, natural resources and biodiversity.*

*NINA was established in 1988. The headquarters are located in Trondheim, with branches in Tromsø, Lillehammer, Bergen and Oslo. In addition, NINA owns and runs the aquatic research station for wild fish at lms in Rogaland and the arctic fox breeding center at Oppdal.*

*NINA's activities include research, environmental impact assessments, environmental monitoring, counselling and evaluation. NINA's scientists come from a wide range of disciplinary backgrounds that include biologists, geographers, geneticists, social scientists, sociologists and more. We have a broad-based expertise on the genetic, population, species, ecosystem and landscape level, in terrestrial, freshwater and coastal marine ecosystems.*

ISSN: 1504-3312  
ISBN: 978-82-426-5387-1

## Norwegian Institute for Nature Research

NINA head office

Postal address: P.O. Box 5685 Torgarden,  
NO-7485 Trondheim, NORWAY

Visiting address: Høgskoleringen 9, 7034 Trondheim

Phone: +47 73 80 14 00

E-mail: [firmapost@nina.no](mailto:firmapost@nina.no)

Organization Number: 9500 37 687

<http://www.nina.no>



Cooperation and expertise for a sustainable future