

Anders Gravbrøt Finstad, Ivar Herfindal, Sam Wenaas Perrin, Bob O`Hara, Joseph Chipperfield, Joachim Paul Töpfer

## Modellering av heildekkande utbreiingskart for arter

Bakgrunn, konsept og arbeidsflyt



Rapport

# Modellering av heildekkande utbreiingskart for arter

Bakgrunn, konsept og arbeidsflyt

Historikk

**RAPPORTNUMMER**

1/2024

**RAPPORTSERIE**

NTNU Gjørevollsentret

**DATO**

22.01.2024



Kunnskap for ei betre verd

**VERSJON**

1.1

**DATO**

22.01.2024

**FORFATTARAR**

Anders Gravbrøt Finstad  
Ivar Herfindal  
Sam Wenaas Perrin  
Bob O`Hara  
Joseph Chipperfield  
Joachim Paul Tøpper

**ANTAL SIDER**

56

**OPPDRAKSGIVAR**

Miljødirektoratet, kontaktperson ved  
miljødirektoratet har vore Tomas Holmern

**OPPDRAKSGIVAR REF.**

M-2665 | 2023

**UTARBEIDDE AV**

Anders Gravbrøt Finstad

**KONTROLLERT AV**

Aline M. Lee

**GODKJENT AV**

Bernt-Erik Sæther

**RAPPORTNUMMER**

1/2024

**RAPPORTSERIE**

NTNU Gjærevollsenteret

**ISBN**

ISBN-nummer

**GRADERING**

Open

# Samandrag

Finstad AG, Herfindal I, Perrin SW, O'Hara B, Chipperfield J, Tøpper JP. (2023). Modellering av heildekkande utbreiingskart for arter: Bakgrunn, konsept og arbeidsflyt. NTNU Gjærevollsenteret, rapport 1/2023, Norges Teknisk Naturvitenskaplege Universitet

Ei heilskapleg kartlegging av naturverdiar er avgjerande for ei kunnskapsbasert forvaltning av natur. Avvegingar mellom naturverdiar og andre samfunnsinteresser kan ikkje gjerast verken på internasjonalt, nasjonalt, regionalt eller kommunalt nivå utan at utbreiinga av naturverdiar er kjent. Dette krev at me kjenner utbreiinga av artar og artsmangfald på relevante romlege skalaer.

Modellering av artar si utbreiing er eit viktig tillegg til feltobservasjonar av artsmangfald. Artsutbreiingskart er prediksjonar frå modellerte samanhengar mellom førekomst av ulike artar, og dei miljøfaktorane som styrer artane si geografiske utbreiing. Slike artsutbreiingskart er avhengige av feltobservasjonar. Dersom observasjonar som ligg til grunn for desse samhengane er mangelfulle vil resultatet bli usikkert. Likeins, viss kartlegginga som ligg til grunn for observasjonane er skeivfordelt og korrelert med miljøfaktorane som styrer utbreiinga, må dette takast omsyn til når artsutbreiinga blir modellert. For bruk i praktisk forvaltning må sluttproduktet ha ei tilstrekkeleg høg geografisk oppløysning og vera fleksibelt slik at ulike indikatorar på naturverdi kan tilpassast relevante forvaltingsmål.

Skal resultatane ha tillit må dei også kunna gjenskapast av andre, uavhengige fagmiljø som ikkje har bindingar til eller restriksjonar frå opphavsmiljøet. Modellar og data må difor vera dokumentert og tilgjengelege til både innsyn og etterbruk. I praksis er Artskart / GBIF den einaste nasjonalt dekkande datakjelda med bra nok taksonomisk oppløysing (artsnivå) som er open og tilgjengeleg til dette formålet. Det er ikkje mogleg å bruke aggregerte datakjelder (samledatabaser) som til dømes Artskart / GBIF utan å ta omsyn til skeivskap i kartleggingsinnsatsen. Det er òg særskilt viktig at Artskart / GBIF vert behandla som ei samling av ulike datakjelder, alle med sine metodiske særtrekk som det kan vera naudsynt å ta omsyn til i modelleringa.

Ei viktig metodologisk utvikling som forbetrar artsutbreiingsmodellar er moglegheita til å kombinere data som inneheld informasjon om innsamlingsintensitet, og opportunistisk innsamla data som berre har informasjon om førekomst av artar. Når data med ulik innsamlingsprotokoll vert analysert innanfor ein felles statistisk modell får ein som regel betydeleg betre prediksjonar, då dette gjev moglegheit til å korrigere for variasjon og skeivskap i innsamlingsintensiteten.

Prosjektet "Modellering av heildekkande utbredelseskart for artar" har som formål å teste ut og vidareutvikle heildekkande artsutbreiingskart for fastlands-Noreg. Desse karta skal predikere sannsyn for artsførekomster og gjennom deriverte produkt peika på områder med høg artsrikdom og endringar i artsrikdom basert på predikerte arealendringar.

Sidan oppstarten i september 2023 har prosjektet hatt fokus på å etablere teknisk arbeidsflyt, datagrunnlag og konsept for interaktiv visualisering av sluttprodukta. Alle resultat frå prosjektet er ope tilgjengelege med teknisk dokumentasjon og køyrbar

kjeldekode. Denne rapporten skildrar teknisk arbeidsflyt og datagrunnlag for prosjektet slik den føreligg per desember 2023.

Konseptet heildekkande utbreiingskart av enkeltartar og avleia sluttprodukt vil gje ein fleksibilitet i indikatorar på naturverdi som er naudsynt når ein må gjera prioriteringar mellom verneinteresser og utbygging. Det er i dag eit stort fokus frå både politisk nivå og norsk naturforvaltning på å etablera ulike verktøy for å talfesta tilstand, utbreiing, funksjon og tenester frå ulike økosystem. Heildekkande utbreiingskart for artar vil, saman med kunnskap om funksjon og eigenskapar til dei ulike artene, gje eit naudsynt grunnlag for å lage funksjonelle, heildekkande og ikkje minst skalerbare kart over økosystem og deira naturverdiar. Andre moglege vidareutviklingar kan være å inkludera kvantitativ informasjon for å lage universale indikatorar for sårbarheit og påverknad. Arbeidsflyten kan også gjerast heilt eller delvis automatisk oppdaterbar ved tilfang av ny informasjon, og slik utgjere ein digital tvilling av norsk natur.

Anders Gravbrøt Finstad, NTNU Gjærevollsenteret, Institutt for Naturhistorie, NTNU Vitskapsmuseet, 7491 Trondheim, [anders.finstad@ntnu.no](mailto:anders.finstad@ntnu.no) ; Ivar Herfindal NTNU Gjærevollsenteret, Fakultet for naturvitskap, Realfagbygget, NTNU, 7491 Trondheim, [ivar.herfindal@ntnu.no](mailto:ivar.herfindal@ntnu.no) ; Bob O'Hara, NTNU, Gjærevollsenteret, Institutt for matematisk fag, Fakultet for informasjonsteknologi og elektroteknikk, 7491 Trondheim; Sam Wenaas Perrin, NTNU Gjærevollsenteret, Fakultet for naturvitskap, Realfagbygget, NTNU, 7491 Trondheim, [sam.perrin@ntnu.no](mailto:sam.perrin@ntnu.no); Joachim Paul Tøpper, NINA, Thormøhlens gate 55, 5006 Bergen, [joachim.topper@nina.no](mailto:joachim.topper@nina.no); Joseph Chipperfield, NINA, Thormøhlens gate 55, 5006 Bergen, [joseph.chipperfield@nina.no](mailto:joseph.chipperfield@nina.no)

## Summary

Finstad AG, Herfindal I, Perrin SW, O'Hara B, Chipperfield J, Töpper JP. (2023). Modelling of comprehensive distribution maps for species: Background, concept and work flow. NTNU Gjørevoll centre report 1/2023, Norwegian University of Science and Technology.

A comprehensive survey of nature values is crucial for knowledge-based nature management. Trade-offs between nature management and other social interests cannot be made on an international, national or municipal scale without a comprehensive understanding of natural ecosystems. This requires knowledge of the spread of species diversity at relevant spatial scales.

Modeling the distribution of species is an important supplement to field observations of species diversity. Species distribution maps are predictions from modeled relationships between the occurrence of different species and the environmental factors that drive species distributions. Such maps are dependent on field observations. If the observations that form the basis for these relationships are insufficient, the results will be unreliable. Likewise, if the mapping that is the basis for the observations is skewed and correlated with the environmental factors that control the distribution, this must be accounted for when modelling these relationships. For use in practical management, the end product must have a sufficiently high geographical resolution and be flexible enough to ensure that different indicators can be adapted to different management goals.

A challenge with many approaches to modeling the distribution of species is that the models must be based on a range of different data sources, each with their own strengths and weaknesses. If results are to be reliable, they must also be possible to recreate independently on a range of geographical scales. It is therefore imperative that both modelling and data processing frameworks are documented and available for both inspection and subsequent use.

In practice, the Norwegian service Artskart (Species Map) and the Global Biodiversity Information Facility (henceforth referred to collectively as Artskart / GBIF) are the only nationally comprehensive species data sources with sufficient taxonomic resolution (species level) available for this purpose. It is not possible to use aggregated data sources such as Artskart / GBIF without taking account of bias in the mapping effort. In practice, there is no way around it when creating a comprehensive distribution map. It is therefore particularly important that Artskart / GBIF is not itself treated as a data source, but as a collection of many different data sources, each with their own characteristics.

An important methodological step in improving species distribution models is to combine data that contain information on collection intensity with opportunistically collected data that only contain information on the occurrence of species. When data with different collection protocols are analyzed within a common statistical model, significantly better predictions can be achieved, as this gives the opportunity to correct for variation in collection intensity.

The project "Modelling of comprehensive distribution maps for species" aims to test and further develop comprehensive species distribution maps for mainland Norway. These maps will predict the likelihood of species occurrences and through derived products

highlight areas with high species richness, and high predicted change in species richness based on changes in land use.

Since its start in September 2023, the project has focused on establishing a technical workflow, data basis and proof of concept for interactive visualization of the end products. All results from the project are openly available through publication of technical documentation and executable code. This report describes the technical workflow and data basis for the project as it exists as of December 2023.

The concept of a comprehensive distribution map of individual species and derived end products will enable the necessary flexibility in producing biodiversity indicators and setting biological conservation priorities. There is currently a strong focus from both the national government and Norwegian nature management institutes on the production of various tools to quantify the state, extent, function and services of various ecosystems. Comprehensive distribution maps for species will, together with knowledge of the function and characteristics of the various species, provide a necessary basis for creating functional, comprehensive and not at least scalable maps of ecosystems and biodiversity value.

Anders Gravbrøt Finstad, NTNU Gjørevollsentret, Institutt for Naturhistorie, Vitenskapsmuseet, 7491 Trondheim, [anders.finstad@ntnu.no](mailto:anders.finstad@ntnu.no) ; Ivar Herfindal, NTNU Gjørevollsentret, Fakultet for naturvitenskap, Realfagbygget, NTNU, 7491 Trondheim, [ivar.herfindal@ntnu.no](mailto:ivar.herfindal@ntnu.no) ; Bob O'Hara, NTNU Gjørevollsentret, Institutt for matematisk fag, Fakultet for informasjonsteknologi og elektroteknikk, 7491 Trondheim; Sam Wenaas Perrin, NTNU Gjørevollsentret, Fakultet for naturvitenskap, Realfagbygget, NTNU, 7491 Trondheim, [sam.perrin@ntnu.no](mailto:sam.perrin@ntnu.no); Joachim Paul Tøpper, NINA, Thormøhlens gate 55, 5006 Bergen, [joachim.topper@nina.no](mailto:joachim.topper@nina.no); Joseph Chipperfield, NINA, Thormøhlens gate 55, 5006 Bergen, [joseph.chipperfield@nina.no](mailto:joseph.chipperfield@nina.no)

# Innholdsliste

<b>Samandrag</b>	<b>4</b>
<b>Summary</b>	<b>6</b>
<b>Forord</b>	<b>9</b>
<b>1 Innleiing</b>	<b>10</b>
1.1 Bakgrunn	10
1.2 Utbreiingskart som forvaltingsverktøy	11
1.3 Utfordringar med heildekkande utbreiingskart	12
1.4 Hovudmål og delmål for prosjektet	15
<b>2 Data</b>	<b>16</b>
2.1 Strategiar for bruk av data	16
2.2 Artsdata	17
2.2.1 Avgrensing av artsdata	17
2.2.2 Hovuddatasett	22
2.3 Miljødata	24
2.3.1 Moglege forklaringsvariabler	25
2.3.2 Data for endringsanalysar	33
2.3.3 Nokre tankar om val av miljøvariablar	34
2.3.5 Forslag til val av forklaringsvariablar	36
<b>3 Arbeidsflyt</b>	<b>38</b>
3.1 Dokumentasjon av arbeidsflyt	38
3.2 Skildring av arbeidsflyt	39
<b>4. Avleia produkt</b>	<b>43</b>
<b>5 Formidlingspakke</b>	<b>44</b>
5.1 Føremålet med formidlingspakka	44
5.2 Produktskildring	44
<b>6 Diskusjon</b>	<b>47</b>
6.1. Oppsummering og vegen vidare for prosjektet	47
6.2. Viderutvikling	47
6.3. Datatilgang	50
<b>Litteratur</b>	<b>51</b>



## Forord

Denne rapporten skildrar fyrste del av arbeidet prosjektet “Modellering av heildekkande utbreiingskart for arter” slik det står per desember 2023. Prosjektet starta september 2023 og fokus i denne rapporten er på prosjektbakgrunn, konseptuell tilnærming og arbeidsflyt. Prosjektet er eit oppdrag frå Miljødirektoratet. Arbeidet i prosjektet bygger på resultat frå ulike prosjekt ved Senter for Biodiversitetsdynamikk ved Noregs Teknisk Naturvitskaplege Universitet, Senter for berekraftig arealbruk (CeSAM) ved UiB og Norsk Institutt for Naturforskning.

NTNU Gjærevollsenteret har hatt prosjektleing og Anders G. Finstad har fungert som prosjektleiar og Ivar Herfindal som nestleiar. Bob O’Hara har hatt det overordna ansvaret for utbreiingsmodelleringa sammen med Joseph Chipperfield. Sam Perrin har vore ansvarleg for programmering av arbeidsflyten og visualiseringsapplikasjonen. Joseph Chipperfield og Joachim Tøpper har bidratt i utviklinga av prosjektkonsept og i kjeldekoden dette prosjektet bygger på.

Ein stor takk til Lisbeth Pedersen ved NTNU Gjærevollsenteret som har vore ansvarleg for koordinering og økonomi, og til Aline Magdalena Lee for kvalitetssikring.

Tomas Holmern har vore prosjektet sin kontaktperson ved Miljødirektoratet. Vi vil nytte anledninga til å takka personalet ved Miljødirektoratet for gode diskusjoner og konstruktive tilbakemeldingar.

Trondheim 10/12/2023



Anders G. Finstad

# 1 Innleiing

## 1.1 Bakgrunn

Ei heilheitleg kartlegging av naturverdier er avgjerande for ei kunnskapsbasert naturforvaltning. Avvegingar mellom naturverdier og andre samfunnsinteresser kan ikkje gjerast verken på internasjonalt, nasjonalt, regionalt eller kommunalt nivå utan at utbreiinga av naturverdier er kjent. Slike avvegingar er også heilt sentrale i samfunnsomstillingar som følge av det grønne skiftet, blant anna når det gjeld energiomstillingar og næringsutvikling (t.d. NOU 2023: 3, 2023). Det ligg også føringar i internasjonale avtalar som Noreg har ratifisert, slik som Naturavtalen (Kunming-Montreal-rammeverket for naturmangfald) der Noreg forpliktar seg til å verna og restaurera store naturområde. Utan eit heilskapleg bilete over kvar naturverdiar finst, vil det ikkje vera mogleg å fastslå den samla effekten av inngrep, vern eller restaurering. Resultatet vert ei stykkevis forvaltning av natur.

Uansett korleis ein definerer naturverdi er det biologisk mangfald som legg grunnlaget for denne (Blicharska et al. 2019). Det er altså naudsynt å forstå korleis biologisk mangfald er fordelt geografisk om ein ønskjer eit heilskapleg bilete av den romlege utbreiinga av naturverdier. Det er i dag store kunnskapshol når det gjeld utbreiinga av norsk flora og fauna (Elven and Sørli 2021). Dette kjem delvis av at det er ei umogleg oppgåve å observere alt artsmangfaldet på alle stader i landet. Dei siste åra har Noreg fått eit stort tilfang av artsdata. Dette skuldast blant anna folkeforskningsprogram som Artsobservasjonar, og ny teknologi som eDNA. Sjølv med ei massiv oppskalering av folkeforskning eller ny teknologi, er det lite truleg at ein i nær eller fjern framtid vil kunna kartlegga biologisk mangfald på alle relevante areal.

Modellering av artar si utbreiing er eit nyttig supplement til bakkeobservasjonar av artsmangfald (Sæther 2021). Artsutbreiingskart er prediksjonar frå modellerte samanhengar mellom førekomst av ulike artar og dei miljøfaktorane som styrer artsutbreiinga. Slike artsutbreiingskart er avhengige av bakkeobservasjonar. Viss observasjonar som ligg til grunn for desse samhengane er mangelfulle vil resultatet bli usikkert. Likeins, viss kartlegginga som ligg til grunn for observasjonane er skievfordelt og korrelert med miljøfaktorane som styrer utbreiinga, må dette takast omsyn til når artsubreiinga vert modellert. Ein stor del av kartlegginga av artsmangfald i Noreg er frå folkeforskningsinitiativ som Artsobservasjonar, materiale frå naturhistoriske museum eller frå lokale kartleggingar i samband med planlagde eller gjennomførte utbyggingar. Felles for dei er at dei ikkje er samla inn med ein strukturert kartleggingsinnsats. Strukturerte innsamlingar er i stor grad avgrensa til nokre få program, som arealrepresentativ naturovervaking, ANO (Tingstad et al. 2019). Desse har generelt for liten kartleggingsinnsats og dekkjer for få taksonomiske grupper til at dei eignar seg som grunnlag for og modellering av utbreiing åleine. Ei hovudutfordring med praktisk bruk av artsutbreiingsmodellar som forvaltingsverktøy har difor vore mangel på metodar for å ta omsyn til observasjonsprosessen i artsdata, med følgjande usikkerheit i presisjonen til sluttprodukta.

## 1.2 Utbreiingskart som forvaltingsverktøy

Modellerte utbreiingskart for enkeltarter kan vera ein viktig reiskap for forvaltning av natur (Baker et al. 2021). Til dømes må analysar som vurderer konsekvensar av ulike inngrep ta inn over seg sannsynet for at ein truga art er til stades i området som er vurdert nedbygd. Dagens praksis med å bruka kart som viser kvar aktuelle artar er observert som ledd i slik vurdering, gjev ikkje denne informasjonen. Observasjonar av kvar artar er til stades, også kalt førekomstdata<sup>1</sup>, kan ikkje seie noko om kvar artane ikkje er til stades. Kart som viser førekomstdata, som til dømes artskart<sup>2</sup>, gjev difor ikkje informasjon om sannsynet for at ein art er til stades eller ikkje. Ei hovudutfordring er at innsatsen bak kartlegginga av dei ulike artsfunna er skeivt fordelt. Dette gjeld både profesjonelt innsamla data og data frå folkeforskningsprosjekt (Mandeville, Nilsen, and Finstad 2022). Funn av artar må difor tolkast før dei kan brukast i praktisk forvaltning. Ein transparent og etterprøvbar metode for tolking av artsfunn i ein geografisk samanheng er bruk av artsubreiingsmodellering (Baker et al. 2021). Artsutbreiingsmodellering er eit stort felt med mange ulike tilnærmingar. Dette gjer at utvalet av metodar som er utvikla er stort, og svært mange val må gjerast før desse modellane kan køyrast. Val av metode må alltid sjåast i samanheng med korleis resultatata skal brukast (Araújo et al. 2019). Dette gjeld ikkje minst korleis resultat frå artsutbreiingsmodellar vert nytta til å laga avleia produkt og statistikkar om artsmangfaldet.

I tillegg til ubreiingskart for kvar einskild art, kan avleia produkt frå desse vera svært nyttige reiskap for forvaltninga. Det mest openberre er kart over kor mange artar som er forventa å finna innanfor eit gitt område, også kalt artsrikdom<sup>3</sup>. Slike kart gjer det mogleg å identifisera stader med høg og låg artsrikdom. Det finst fleire rammeverk for å estimera artsrikdom. Nokre modellar legg saman resultat frå fleire einartsmodellar, medan andre estimerer funn av ulike artar i same modell. Desse modellane kan ha ulike krav til data og også ulik evne til å predikera førekomst av artar (Norberg et al. 2019). Kart over artsrikdom er grunnlag for definering av spesielt artsrike område, også kalla “hotspots” (Myers 1988; Myers et al. 2000). Den opphavlege definisjonen av slike område frå Myers fokuserer på større geografiske område med eit stort mangfald av truga endemiske artar. Seinare har omgrepet vorte brukt meir generisk om område som inneheld sjeldent stort biologisk mangfald generelt og i mindre skala. Konseptet har vorte brukt som grunnlag for prioritering av område med særskilt høgt artsmangfald for bevaring, til dømes i samband med miljøkartlegging av skog (MiS-kartlegging) (Gjerde, Sætersdal, and Blom 2007). Det er likevel ingen einsarta svar på kva som gjer at eit område er definert som sær artsrikt (“hotspot”) og fleire ulike indikatorar er brukt (Reid 1998).

For bruk i praktisk forvaltning må “hot-spot”-konseptet ha høg nok geografisk oppløysing og vera fleksibelt slik at indikatorar kan tilpassast ulike forvaltingsmål (Marchese 2015). Tidlege heildekkande kart over artsmangfald og område med særskilt høg artsrikdom i Noreg (Olsen et al. 2018, 2020) har brukt ei romleg oppløysing på 1\*1km<sup>2</sup>. Dette har vist seg ikkje å vera høgt nok for mange forvaltingsrelevante formål, særleg kommunale planprosesser og anna forvaltning på lokalt nivå.

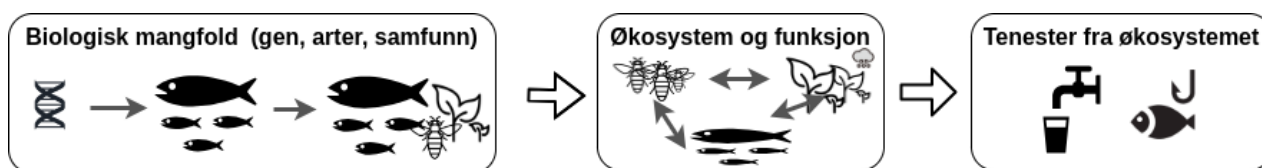
---

<sup>1</sup> Vanleg engelsk brukt terminologi er “occurrence only data” eller “presence only data”

<sup>2</sup> <https://artskart.artsdatabanken.no/>

<sup>3</sup> artsrikdom (antallet artar) saman med antallet (individ eller anna abundansmål) av kvar art, definerer artsmangfald.

Heildekkande utbreiingskart av enkeltartar og avleia produkt derfrå vil gje den naudsynte fleksibiliteten i indikatorar for naturmangfald slik at ein kan etablere kunnskap om naturverdi og gjera prioriteringar om korleis ein forvalta naturen. Både politisk og i norsk naturforvaltning er det i dag eit stort fokus på å få etablert ulike verktøy for å tallfeste tilstand, utbreiing, funksjon og tenester i norske økosystem. Eit slikt naturrekneskap krev andre typer data enn dei som i dag vert samla inn gjennom overvakingsprogram (Anon. 2023). Heildekkande utbreiingskart for artar utgjer, saman med kunnskap om funksjon og eigenskapar til dei ulike artane, eit naudsynt grunnlag for å laga funksjonelle, heildekkande og ikkje minst skalerbare kart over økosystema. Funksjonen til eit økosystem er bestemt utifrå ein kombinasjon av kva artar som finst der, eigenskapane deira, og abiotiske miljøfaktorar (Figur 1).



**Figur 1.** Samanhengen mellom menneska sin velferd og økosystemet, definert som alle dei grunnleggande goder og tenester som kan haustast frå naturen, er eit resultat av kva økosystem ein har og korleis det fungerer. Økosystemet og funksjonen er påverka av det biologiske mangfaldet - kva for artar som er til stades og mengda av dei.

Skal ein setja opp eit naturrekneskap basert på areal må ein ha kunnskap om kvar ein finn naturverdiar i form av naturmangfald. Dette krev ei kvantifisering av artsriktom og førekomst av artar i ulike økosystem og naturtypar. For dei fleste artar i Noreg, spesielt insekt og andre invertebratar, er det i dag lite eller ingen kunnskap om førekomst og artsriktom i ulike økosystem og naturtypar. Denne kunnskapen kan ikkje konstruerast frå fjernmålingsdata basert på til dømes satellittar, fly eller dronar.

Den romlege fordelinga av artsriktom er, og kan fåast frå utbreiingskart. Eit område sin naturverdi må setjast opp mot kor stor del av artsmangfaldet området held samanlikna med det som finst i eit større område. Såleis kan eit artsfattig område ha stor naturverdi, viss dei artane som finst der ikkje er til stades i andre område. Naturverdien av eit område er også påverka av samansetjinga og fordelinga av ulike naturtypar. Til dømes kan isolerte område med høg artsriktom, ha lågare bevaringsverdi enn samanhengande område, då isolerte område over tid vil miste artsmangfald gjennom utdøying. Samanhengende område kan ha svært høg bevaringsbiologisk verdi då slike område minkar sannsynet for at bestandar av artar dør ut.

Indikatorar for truga natur er ein føresetnad for arealplanar som tek omsyn til naturverdiar og naturmangfald. I Noreg er kommunale arealplanar det viktigaste offentlege styringsverktøyet for arealforvaltning. Heildekkande utbreiingskart og avleia produkt, som kart over område og identifisering av naturtypar med høg artsriktom og høg førekomst av truga artar, kan bli eit svært viktig verktøy på alle forvaltingsnivå. Dette er særskilt viktig for å kvantifisera samla belastning av ulike typar inngrep på naturmangfald, på tvers av artsgrupper og økosystem.

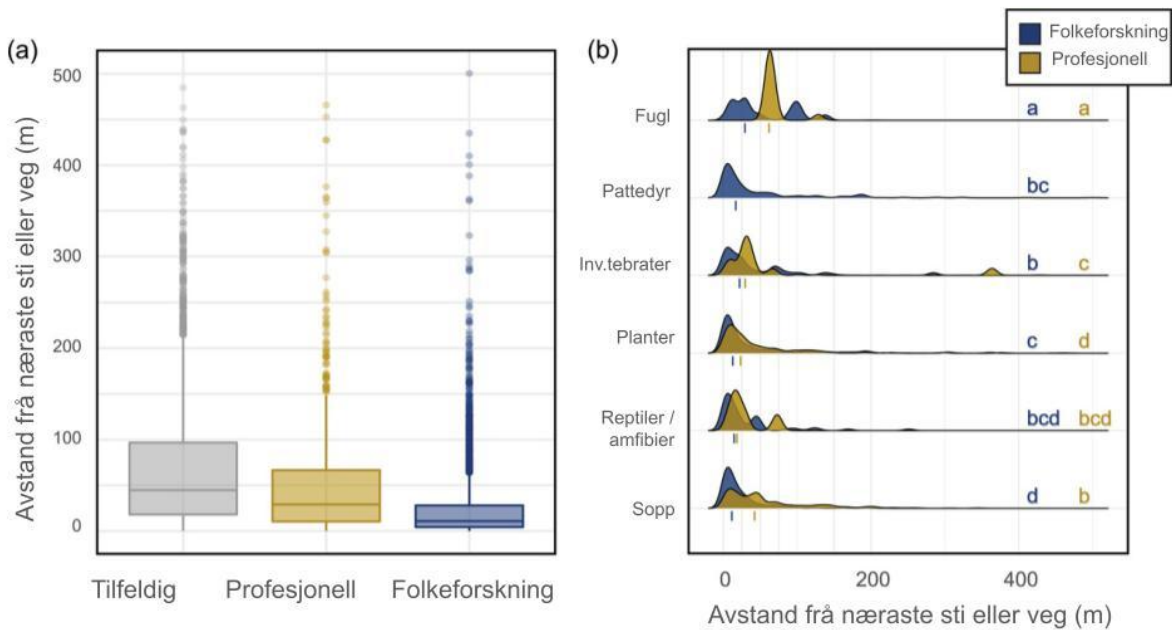
### 1.3 Utfordringar med heildekkande utbreiingskart

Ei av dei største utfordringane som gjeld for mange måtar å modellera utbreiinga av artar og artsrikdom, er at dei baserer seg på fleire datakjelder som har ulike styrkar og svakheiter. Ein kan grovt klassifisera desse utfordringane som

- I. Store delar av datagrunnlaget er opportunistisk innsamla frå folkeforsking eller data frå naturhistoriske samlingar som manglar informasjon om innsamlingsprotokoll og kartleggingsinnsats.
- II. Data kjem frå svært mange ulike datasett samla av ulike folk frå ulike institusjonar, med mange ulike metodar.
- III. Stadig større datamengder med ikkje-dokumentert innsamlings skeivskap i tid, rom og taksonomisk utval vert tilgjengelege, noko som aukar sannsynet for feil konklusjon om denne skeivskapen ikkje vert handsama på ein god måte.

Data som skildrar romleg lokalisert naturmangfald kjem tradisjonelt frå naturhistoriske samlingar, folkeforskningsprogram, eller økologiske kartleggingsprogram. Desse har ulike typer svakheitar. Overvåkingsprogram har ofte ei strukturert og godt dokumentert innsamling etter standardiserte metodar, men data frå desse programma er ofte avgrensa i romleg utstrekking og i taksonomisk breidde eller oppløysing. Mange grupper er til dømes ikkje bestemt til art, men berre skildra på overordna taksonomisk nivå. Data fra naturhistoriske samlingar har ofte ei svært god taksonomisk oppløysing på artsnivå og data er også kvalitetssikra og dokumentert gjennom lagring av fysisk materiale. Data fra naturhistoriske samlingar manglar ofte informasjon om kartleggingsinnsats. Likeins har data frå folkeforsking som regel ei stor geografisk utbreiing, men svært skeiv kartleggingsinnsats som ikkje er dokumentert. I tillegg er folkeforskningsdata ofte avgrensa til et utvalg taksona med varierende grad av sikkerheit i artsbestemminga. Dei siste åra har det også vorte eit stort tilfang av andre innsamlingsmetodar, drive fram av teknologiske nyvinningar (e.g. Tuia et al. 2022, Maclean and Triguero 2023, Ruppert, Kline and Rahman 2019). Dette inkluderer både artsbestemming gjennom molekylære metodar som metabarcoding, og artsidentifisering frå lyd eller bilete basert på maskinlæring.

Alle datakjelder om artsførekomstar har store geografiske skeivskapar i innsamlinga (Mandeville, Nilsen, and Finstad 2022). For dei fleste artsgrupper er det berre opportunistisk innsamla førekomstinformasjon ("førekomstdata") som finst over stor romleg skala (heile Noreg) og med høg taksonomisk oppløysing og breidde. Slike datasett har òg stor samvariasjon mellom innsamlingsintensiteten og miljøvariablar på lokal skala (Mandeville, Nilsen, and Finstad 2022). Data er oftest samla inn nær infrastruktur som veg og sti (Figur 2), og plassering av sti og veg i landskapet er heller ikkje vilkårleg (Mandeville, Nilsen, and Finstad 2022). Ein kan difor ikkje nytta opportunistisk innsamla førekomstdata i artsmodellering åleine, eller bruka metodar som ikkje gjev uavhengig informasjon på kartleggingsinnsats.



**Figur 2:** (a) Avstand frå observasjonslokalitet til næraste stig/veg for data frå folkeforskning og profesjonelle innsamlingar, og eit vilkårleg utval av lokalitetar i studieområdet. (b) Avstand frå observasjonslokalitetar til næraste stig/veg for ulike taksonomisk grupper. Området under linja viser andel, loddrette strekar viser medianverdiar. Bokstavar fortel om det er signifikant forskjell mellom gruppene (Dunns post hoc-test,  $\alpha = 0,05$ ). Figuren viser tilgjengelege artsdata frå Bymarka i Trondheim kommune og er omarbeidd frå Mandeville, Nilsen og Finstad (2022) (CC BY 4.0).

Store datamengder er ikkje det same som mykje informasjon. Med større utval kjem også ein større risiko for å trekke feil konklusjon som følgje av skeivskap i data, også kjent som “the big data paradox” (Meng 2018). Presisjonen til eit estimat basert på eit stort utval vil ofte være særskilt god. Paradokset er difor at så lenge skeivskapen til estimatet ikkje endrar seg med datamengda, vil risikoen for å dra feil konklusjon bli større med eit større utval.

Det er difor ikkje mogleg å bruke datakjelder som i hovudsak er basert på førekomstdata, som til dømes Artskart / GBIF utan å ta omsyn til skeivskap i kartleggingsinnsatsen. I praksis er Artskart / GBIF er den einaste nasjonalt dekkende datakjelda med tilstrekkelig taksonomisk oppløsing (artsnivå) som er tilgjengeleg. Ei fallgrop er å behandle desse databasane som ei einskild datakjelde når dei i realiteten består av mange ulike datasett, alle med sine særtrekk og innsamlingsmetoder og ulik kartleggingsinnsats.

Ei viktig metodeutvikling som gjer artsutbreiingsmodellar meir presise er moglegheita til å kombinere opportunistisk innsamla data som berre inneheld informasjon om førekomst av artar med data som inneheld informasjon om innsamlinga (“innsatsdata”)<sup>4</sup> (Isaac et al. 2019). I sin enklaste form er innsatsdata data som inneheld informasjon om kor ein har observert eller ikkje observert data (“funn / ikkje-funn”), men også det som i økologien kallast “kvantitative data” som inneheld informasjon om antal eller anna mål på mengde av organismar samla i ein prøve eller oppdaga i ein observasjonsperiode. Når data med

<sup>4</sup> I GBIF samanheng blir ofte innsatsdata på engelsk nemnt som “sampling-event data”. Merk at denne terminologien kan være forvirrende da “sampling-event data” både referer til at innsamlingsinformasjon er til stades i datasettet og til ein spesifikk datastandard.

ulik innsamlingsprotokoll vert analysert i ein felles statistisk modell, får ein som regel mykje betre prediksjonar sidan dette gjev moglegheit for å korrigera for variasjon i innsamlingsintensiteten.

#### 1.4 Hovudmål og delmål for prosjektet

Prosjektet “Modellering av heildekkande utbreiingskart for artar” har som formål å teste ut og videreutvikle heildekkande artsutbreiingskart for fastlands-Noreg. Desse karta skal predikera sannsynet for artsførekomst og gjennom avleia produkt peika på områder med høg artsriksdom, og eventuelle endringar i artsriksdom basert på arealtypeendringar. Prosjektet har som målsetjing å produsera kart med ei romleg oppløysing som gjer kunnskapsgrunnlaget relevant for kommunalt forvaltingsnivå.

Prosjektet har følgande delmål:

- Utvikla og bruka relevante variablar for å forklåra førekomst av einskildartar og artsriksdom, og estimerer variablane si betyding for dette
- Utvikla heildekkande kart som viser sannsynet for førekomst av terrestre artar, truga artar og truga ansvarsartar. Dette inneber også kart som kan peika på områder med stor artsriksdom
- Kartlegga områder med låg kartleggingsinnsats, men med stor forventa artsriksdom
- Identifisera samanhengande område med høg artsriksdom
- Bruka fjernmålingsdata for å identifisera fråfall av habitat eller naturtypar som har konsekvensar for artsriksdomen, på ein skala som er relevant for kommunal forvaltning
- Identifisera overlapp mellom område med høg artsriksdom og høg forvaltingsinteresse (t.d. verneområde)
- Bruka validering for å estimerer nøyaktigheita til utbreiingskarta
- Koma med anbefalingar på korleis artsutbreiingsmodellering kan utviklast vidare som eit målretta og kostnadseffektivt verktøy for forvaltning av norsk natur.

Prosjektet har frå oppstarten i september 2023 fokusert på å etablere teknisk arbeidsflyt, datagrunnlag og konsept for interaktiv visualisering av sluttprodukt. Alle resultat frå prosjektet er ope tilgjengelege med teknisk dokumentasjon og køyrbar kode som vert kontinuerleg oppdatert og arkivert på nettstaden GitHub<sup>5</sup>. Teknisk dokumentasjon og kode vert skrive på engelsk for å opna for at internasjonale miljø skal kunne ettergå og nytta resultatane.

---

<sup>5</sup> <https://github.com/gjearevoll/BioDivMapping>



## 2 Data

### 2.1 Strategiar for bruk av data

Ein hovudpremiss i dette prosjektet er at alle resultat skal kunna gjenskapast og ettergåast av uavhengige fagmiljø utan bindingar til, eller restriksjonar frå, opphavsmiljøet. Bruk av modellar og datasynteser har auka kraftig dei seinare åra, i takt med aukande tilgang til data og digitale hjelpemiddel. Inntoget av kunstig intelligens vil truleg akselerera denne utviklinga. Dette gjeld også i miljøforvaltinga og er i tråd med ein ønskt strategi<sup>6</sup>. Skal analyser vera truverdige og verta oppfatta som pålitelege, er det heilt naudsynt at modellar er godt dokumentert og at data er tilgjengelege (Feng et al. 2019; Alston and Rick 2021; Araújo et al. 2019). Forvaltingsverktøy som er finansiert av offentlege midlar bør også vera mogleg å vidareutvikla av andre enn opphavsmiljøet. For å sikra dette må inndata til modellar ikkje berre vera mogleg å finna att, men også tilgjengelege for gjenbruk utan andre vilkår enn sitering. Slik sikrar ein at uavhengige miljø har moglegheit til å analysere og testa resultat sjølv dersom dei er fagleg ueinige i konklusjonen. Slike prinsipp er no praktisert i dei aller fleste vitenskaplege tidsskrift, men påfallande lite forankra i offentleg finansiert anvendt forskning og forvalting i Noreg.

Dette prosjektet vil sikra resultat som kan etterprøvast og vidareutviklast gjennom i så stor grad som mogleg å bruka opne kjelder til både artsdata og miljødata. Som ein start vil det vera eit krav at data følgjer FAIR datastandardar; “Findable, Available, Interoperable and Reproducible” (Wilkinson et al. 2016). Det betyr at data skal vere både mogleg å finna, vera tilgjengelege, vera interoperabile og mogleg å reprodusere. Det vil også vera eit krav at dei skal vera opne, altså mogleg å gjenbruka utan vilkår frå opphavsperson eller -institusjon.

For å sikra etterprøvbarheit vil prosjektet bruka artsdata nedlasta gjennom GBIF (<https://www.gbif.org/>). GBIF kan skildrast som “det globale Artskart” og for utbreiingsmodellering også i praksis eit spegelbilde av data som ligg i Artskart sine tenester. GBIF har godt dokumenterte og opne API-ar kor data er tilgjengeleg i ein standardisert maskinlesbar form, og, ikkje minst ei velutvikla datasiteringsteneste. Kvar einaste nedlastning gjev ein DOI som kan gjerast permanent peikande mot originalt søk og data. Dette gjer GBIF til ein svært god plattform for reproduserbare og transparente datasyntesar som også er nasjonalt og internasjonalt kompatible.

I tillegg til data publisert på GBIF vil det også verta brukt data frå arealrepresentativ naturovervaking, ANO, (Tingstad et al. 2019). Data frå dette programmet er per i dag ikkje gjort tilgjengeleg på GBIF / Artskart, men kan lastast ned manuelt frå Miljødirektoratet sine nettsider, og må sidan verta omarbeidd til kompatibel standard.

Prosjektet inneheld ikkje ressursar til å ta i bruk nye datakjelder. Dette gjeld både relevante data som ikkje er samla inn, data som er samla inn men ikkje tilrettelagt for bruk, eller data som er samla inn men ikkje gjort offentleg tilgjengeleg. Dette kan også gjelde deler av datasett der relevant informasjon ikkje er offentleg tilgjengeleg. Likeins har heller ikkje prosjektet ressursar til storskala feilretting og formatering av datasett som ikkje har naudsynt kvalitet.

---

<sup>6</sup><https://www.miljodirektoratet.no/publikasjoner/2019/desember-2019/digitalt-malbilde-2019-2025/>



Det kan bli naudsynt og bruke datakjelder som ikkje er åpent tilgjengeleg der det ikkje er mogleg og produsere tolkbare resultat utan. Dette gjeld særskilt data som inneheld informasjon om innsamlingsintensitet (“innsamlingsdata”). Særtilfeller vil bli diskutert i kvart enkelt tilfelle.

## 2.2 Artsdata

### 2.2.1 Avgrensing av artsdata

Geografisk og taksonomisk er avgrensinga for prosjektet godt definert frå oppdragsgjevar (Miljødirektoratet). Det står likevel att ein del val med tanke på kva data innanfor desse rammene som skal inngå i artsmodelleringa. Her drøftar me dei viktigaste kriteria basert på **innsamlingsprotokoll, tilgjengelegheit, tid, romleg presisjon, taksonomisk presisjon og utval**. For ei fullstendig liste av utvalsprosedyrer viser me til dokumentasjonen som er ope tilgjengeleg på GitHub<sup>7</sup>.

#### Innsamlingsprotokoll

Artsdata som inkluderast i prosjektet vil bli delt i to hovudgrupper i) data som inneheld informasjon som kan omsetjast til kunnskap om kartleggingsdata (“innsatsdata”) og ii) data som ikkje inneheld informasjon om kartleggingsinnsats (“førekomstdata”). I praksis betyr (i) data som inneheld skildring av innsamlingsprotokoll og som har nok informasjon til at “funn” / “ikkje-funn” av arter eller kvantitet av ulike arter på dei ulike innsamlingshendingane kan stadfestast. I GBIF samanheng blir slike data ofte referert til som “sampling-event data”<sup>8</sup> da kjernen i datasettet krinsar rundt innsamlingshendinga. I praksis blir mykje data som originalt er samla inn med informasjon om funn / ikkje-funn eller kvantitativ informasjon ikkje nytta på denne måten i vidare i arbeidsflyten. Hovudårsakane til dette er enten manglande metadata eller manglande datastruktur kor innsamlingshendinga ikkje lar seg rekonstruere. Det er også svært mykje data med informasjon om funn / ikkje-funn og kvantitativ informasjon som ikkje er åpent tilgjengeleg, anten som rådata eller i ein strukturert form via GBIF / Artskart. Det er også døme på slike data som blir gjort tilgjengeleg i GBIF / Artskart, men verdifull informasjon blir halden tilbake før publisering. I praksis vil mykje av data som eigentleg er kvantitative data eller funn / ikkje-funn data brukt som førekomst-data siden at kvaliteten på data skildring ikkje er god nok eller at datasett ikkje inneheld den relevante informasjonen som trengs. Som vi vil drøfte vidare er difor tilgang på slike “hendelsesdata” (“sampling-event data”) ofte det som er med på og begrense kvalitet på dei heildekkande artskarta.

#### Tilgjengelegheit

Det er eit krav at alle data som inngår i dette prosjektet er opne og tilgjengelege. Data som ikkje er det, uavhengig av kvalitet, vil ikkje verta brukt i artsmodelleringa med mindre det er gode grunner til at data ikkje er offentleg tilgjengelege. “Gode grunnar” er i samanheng med artsdata i hovudsak situasjonar med sensitive data, til dømes der nøyaktig posisjon til truga artar ikkje er offentleg men kor denne informasjonen er under

<sup>7</sup> <https://github.com/gjearevoll/BioDivMapping/tree/main/data>

<sup>8</sup> <https://www.gbif.org/data-quality-requirements-sampling-events>

upartisk kontroll og tilgjengeleg for formålet. I praksis vil dette gjelde sensitive artsdata tilgjengeleg gjennom Miljødirektoratet sin løysing for innsyn i sensitive artsdata<sup>9</sup>

Tilfang av data er derfor avgrensa både av mangel på innsamlingar og av mangel på deling. Denne avgrensninga i tilfang av data gjeld spesielt "innsatsdata". Pattedyr er som gruppe eit godt døme på denne avgrensninga i tilfang av data. Store systematiske overvåkingsprogram av store pattedyr som store rovdyr, fjellrev og hjortevilt, er ofte ikkje gjort tilgjengeleg i ein strukturert form. Dette vil etter vår vurdering utgjera eit lite problem i dette prosjektet, då kunnskapen om utbreiing av desse artane er svært god og det dreier seg om få artar som vert forvalta på artsnivå. Små pattedyr som hare, flaggermus og smågnagarar er derimot meir eller mindre ignorert av offentleg forvaltning, og regelmessig overvaking er så godt som fråverande. Dette er grupper som inneheld truga artar og der utbreiingsmodellering er begrensa av eit lite datagrunnlag.

Det vil også, som drøfta under, vera spesielle forhold når det gjeld avgrensa tilgjengelegheit i tid og rom.

#### **Tid: År**

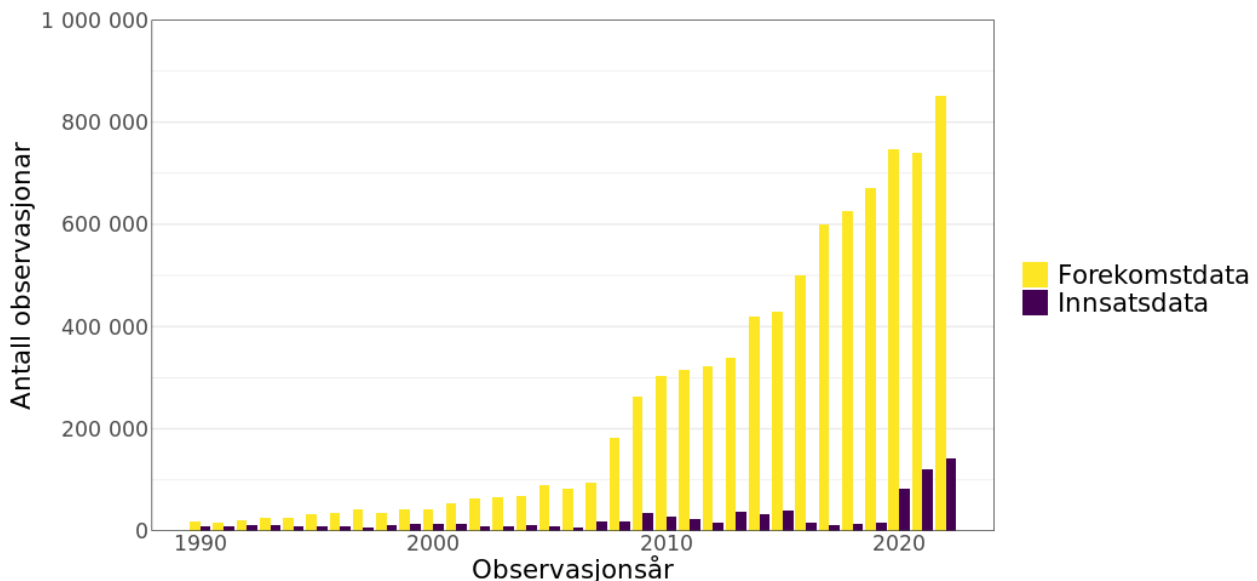
Føremålet med artsmodelleringa er å gje ein oversikt over kvar i Noreg ein finn område med høg artsrikdom, og eventuelle deriverte produkt basert på slike kart. Dette inneber å modellera kvar artar finst i dag, og ikkje kvar dei historisk har hatt si utbreiing. Artar si utbreiing endrar seg av ei rekke årsaker, der påverknad frå menneske, slik som arealendring eller overhausting, er dei viktigaste. Å bruka historiske observasjonar vil difor kunna gje feil bilete av artsrikdom, fordi naturtypen i eit gitt område kan ha endra seg, fordi andre miljøtilhøve slik som klima eller forureining har endra seg, eller fordi den er utrydda lokalt gjennom jakt eller fangst. Det vart til dømes gjort svært omfattande undersøkingar i samband med verneplanar for vassdrag, også i område som vart bygde ned. Data om førekomst av artar frå desse undersøkingane er ikkje relevante i dag, når til dømes tidlegare bjørkeskog er neddemt (Figur 3). Dette er sjølvsagt eit karikert døme, og historiske data frå områder som har endra karakter dramatisk, for eksempel i forhold til bygging av infrastruktur eller byutvikling kan relativt enkelt filtreres bort. Prinsippet gjeld også i område med mindre grad av miljøendring, slik som hogst, hyttebygging, eller endring av skjøtsel av seminaturleg mark. Her er det vanskelegare og definere klare kriterier for når data skal filtrerast bort.

---

<sup>9</sup> <https://sensitive-artsdata.miljodirektoratet.no/Contentpages/Om.aspx>



**Figur 3:** Funn av karplantar i Nesjøen, Trøndelag. Dei fleste av observasjonane er i samband med forundersøkingar i samband med utbygginga, som stod ferdig i 1971. Det gjev lita meining å bruka observasjonar frå område som i dag ligg under vatn til å modellera og predikera førekomst av terrestriske karplanter som funksjon av dagens arealdekkkart.



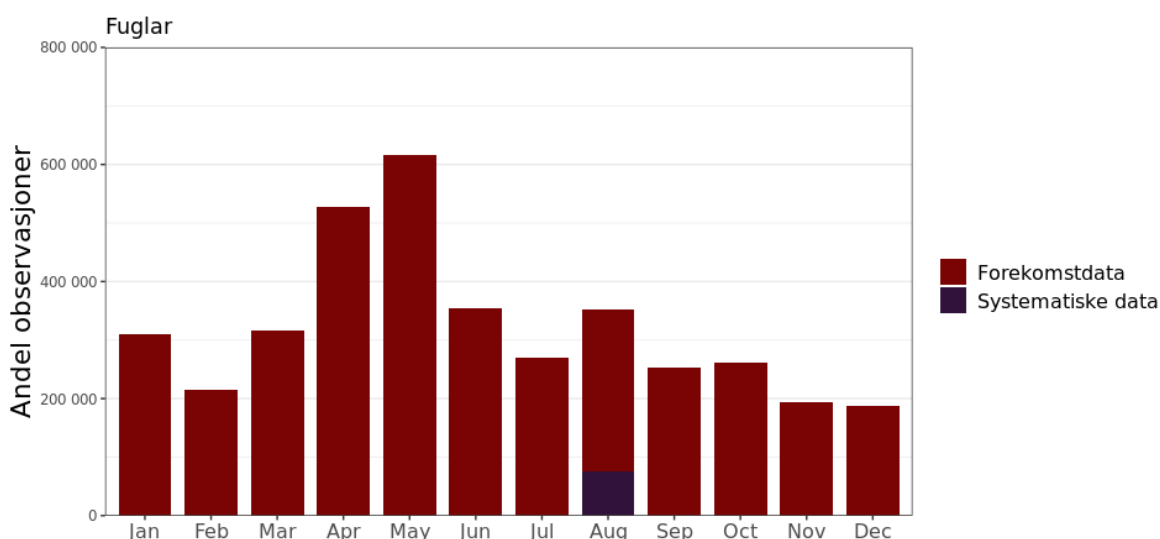
**Figur 4:** Talet på artsobservasjonar i Noreg per år frå opportunistisk innsamla data (gult) og data med informasjon om funn / ikkje-funn eller kvantitet av artar (svart). Den store auken i opportunistisk innsamla data skuldast i all hovudsak auka tilfang av folkeforskningsdata gjennom Artsobservasjonar.

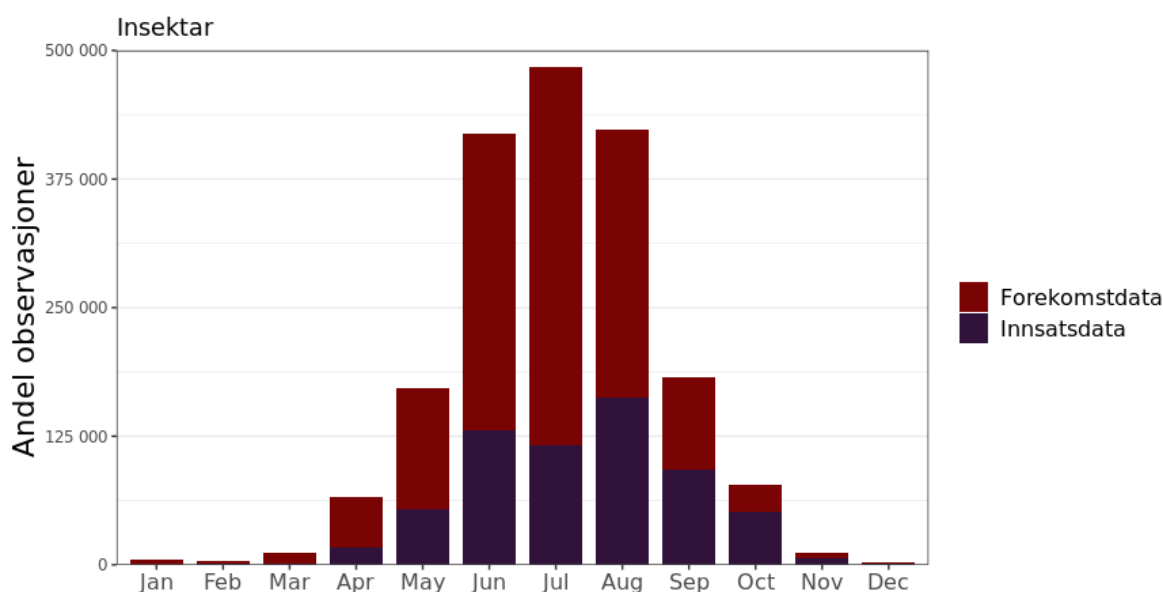
Eit anna argument for å avgrensa kva år som vert inkludert i modelleringa er at historiske data om viktige miljøfaktorar, slik som naturtypar eller produktivitet, sjeldan er tilgjengelege. Å tilpassa modellar der tidspunkt for artskartlegging ikkje samsvarar med tidspunkt for å skildra miljøet, vil kunna gje svært usikre artsutbreiingskart med store feil. Det er difor gjort eit val å ta med data frå 1991, og å utelata eldre data, noko som også samsvarar med avgrensning på klimaperioder.

## Tid: Årsted

For mange artar er årsted uproblematisk for modellering av utbreiinga deira, då dei held seg i same området gjennom heile året, men mange artar, spesielt fuglar og insekt, trekker mellom område for reproduksjon og overvintring. Om ein modellerer heile året i ei vending, vil ein få utbreiingskart som ikkje skil mellom ulike livshistoriestadium der habitatkrava ofte kan vera vidt forskjellige. Til dømes: Fossekall (*Cinclus cinclus*) hekkar ofte langt til fjells, men også i lågareliggande område. Om vinteren er fossekallen heilt avhengig av ope vatn og trekker til låglandet og nær kysten for å finna område med god mattilgang. Samanhengen mellom førekomst og høgde over havet vil difor vera svært ulik mellom overvintringsområda og hekkeområda til fossekallen. Det er difor eit ynskje om å utarbeida sesongbaserte modellar der det er grunnlag for dette i datamaterialet, og der det gjev biologisk mening for arten. I praksis er dette vanskeleg. Den einaste gruppa som realitetsvurderast her er fugl. Dette er kanskje den einaste relevante artsgruppa som er så godt kartlagt at det kan vera datagrunnlag for ei slik inndeling. Avgrensing av sesong må basera seg på ei avveging mellom biologisk mening og tilgjengelegheit på data.

Tilgjengelegheit på data kan bli ei utfordring sjølv for dei mest kartlagde fugleartane. Årstidsvariasjon i habitatkrav hjå arten vil påverka sannsynet for at den vert observert, til dømes fordi vinterområda oftare ligg i område med høgare tettleik av folk enn kor fugl oftare vert observert resten av året. I tillegg er ikkje fordelinga av observatørar den same gjennom året. Dette tyder på at observasjonsprosessen må modellerast separat for kvar sesong. Dette er spesielt utfordrande då det er lite systematisk kartlegging vinterstid i Noreg (Figur 5).





**Figur 5:** Talet på observasjoner frå datasett som inneheld informasjon om fråvær eller kvantitet (mørk blå eller forekomstdata som berre inneheld informasjon om tilstedeværelse (raud) for fuglar (øvre pannel) og insekt (nedre pannel) sortert på måned. For fuglar kjem data som inneheld informasjon om fråvær eller kvantitet kun frå to ulike datasett tilknytta hønsefuglportalen<sup>10</sup> (Nilsen m.fl. 2022 a,b,c).

For å få eit artsutbreiingskart der observasjonar er knytt til rett miljø må den geografiske presisjonen til artsobservasjonane samsvara med den romlege oppløysinga til miljødata. Miljødata som er tenkt brukt kjem med ulik romleg oppløysing, men vil ikkje vera høgare enn 100\*100 m<sup>2</sup>. Det er difor sett eit kriterium om minimum 100 m koordinatpresisjon for artsdata som skal brukast i modelleringane. Data frå sårbare eller truga artar der nøyaktig posisjon ikkje er offentleg tilgjengeleg, men kuratert av offentlige instansar og tilgjengeleg etter søknad<sup>11</sup>, vil verta inkludert. Vi vil også forsøke å få inkludert data frå Norsk hekkefuglovervåking (Kålås m.fl. 2023). Desse er i dag publisert med koordinatpresisjon på 5km for å skjerme overvåkningslokaliteter, men nøyaktig posisjon ikkje tilgjengeleg frå uavhengig institusjon. Data herfrå vil derfor ikkje i utgangspunktet tilfredsstillende inklusjonskriteria.

### Taksonomisk presisjon og utval

For å modellera utbreiinga til artar krev det at observasjonane er gjeve på artsnivå. For mange artar er dette uproblematisk, men somme artar krev svært høg taksonomisk kompetanse for å skilja frå kvarandre. I slike tilfelle kan observasjonar vera gjeve til slekt- eller familienivå. Det gjev lita mening å modellera på høgare taksonomisk nivå, då desse artene kan ha vidt ulike økologiske nisjer sjølv om dei er like av utsjånad. Data som ikkje har informasjon på artsnivå vert difor forkasta.

<sup>10</sup> <https://honsefugl.nina.no/Innsyn/nb>

<sup>11</sup> <https://sensitive-artsdata-innsyn.miljodirektoratet.no/>

Prosjektet skal gje modellar og resultat der artar lista som kritisk trua (CR), sterkt trua (EN), eller sårbar (VU) i den norske raudlista for artar (Artsdatabanken 2021). For planter brukar raudlista stundom underartar og varietetar for å skilja ut truga artar. For raudlista 2021 er 48 av 343 karplanter lista med varietet eller underart i dei tre kategoriane CR, EN og VU (Artsdatabanken 2021). Til dømes er hundepersille (*Aethusa cynapium*) kategorisert som Livskraftig (LC), medan varieteten skogpersille (*Aethusa cynapium* var. *cynapium*) har kategori Sårbar (VU). For modellering av utbreiing til artar på raudlista kan det difor brukast finare taksonomisk nivå enn art. Det som talar mot dette er at det er utfordrande å blanda ulike taksonomiske nivå inn i eit mål på artsrikdom som “hot-spot”-konseptet. Det er vidare mogleg at det ikkje er tilstrekkeleg datagrunnlag for å modellera utbreiing med finare taksonomisk oppløysing enn art, då mange datakjelder, som til dømes folkeforskningsdata frå Artsobservasjonar, ikkje kan reknast som pålitelege når det gjeld identifisering av underartar og varietetar som er vanskelege å skilje frå hovudarten. Prosjektet vil derfor ikkje bruke underart / varietet inn i modellane for truga artar.

Artar som er lista på framandartslista (Artsdatabanken 2023) vil verta ekskludert i modelleringa av artsrikdom då desse ikkje reknast som ein del av norsk natur. Dette gjeld også regionalt framande artar utanfor naturleg utbreiingsområde.

### 2.2.2 Hovuddatasett

Artsdata vert lasta ned frå GBIF via GBIF API og asynkron nedlasting. Data vert grovfiltret før nedlasting via eit API søk som i hovudsak avgrensar data taksonomisk, ved hjelp av ei liste over høgare takson (Tabell 1), geografisk (landområde i Noreg, utan Jan-Mayen og Svalbard), og usikkerheita til den oppgjevne koordinaten ( $\leq 100$  m). Det vert òg filtret på ei rekke kvalitetsparametrar som til dømes ugyldig koordinatinformasjon. Kvar nedlasting gjev ein DOI som identifiserer kjelldata (GBIF.org 2023a).

**Tabell 1:** Talet på observasjonar (# obs) og datasett som er tilgjengelege i GBIF per 30.11.2023 (GBIF.org 2023a) for dei ulike artsgruppene som er vurdert i prosjektet og etter gjeldende søkekriterier.

Norsk namn	GBIF taxon ID	NBIC taxon ID	Taksonomisk nivå	Vitskapleg namn	Talet på observasjonar	Talet på datasett
Amfibiar og reptilar	11592253, 131	NA, 243	Klasse	Squamata, Amphibia	31217	12, 13
Edderkoppar	367	42	Klasse	Arachnida	61366	10
Insekt	1470	89	Klasse	Insecta	1969825	60
Karplantar	7707728	NA	Rekke	Tracheophyta	3191118	87
Mosar	35	1158	Rekke	Bryophyta	171184	31
Pattedyr	359	293	Klasse	Mammalia	104220	37
Sopp og lav	5	975	Rike	Fungi	906444	60
Fuglar	212	252	Klasse	Aves	3859742	36

Etter nedlasting vert data stykka opp etter kva som er relevant for ulike modellkøyningar (sjå kapittel 3.2). I tillegg vert det gjort ein del avgrensingar og kvalitetskontrollar etter nedlasting. Observasjonar som ikkje er identifisert til art vert ekskludert. Lav er ikkje ei einiskild takonomisk gruppe. For lav vert det difor brukt ei sjekkliste for artsnamn for å avgrensa denne gruppa (Santesson 1993). Etter postprosesseringa vert ferdige datasett registrert i GBIF som eit derivert datasett (GBIF.org 2023a) og sitering frå det vil følgja data og produkt vidare i arbeidsflyten.

Ein del av datasetta inneheld informasjon frå strukturerte innsamlingar der det kan vera mogleg å dra ut data om artar si frávære for lokalitetar innanfor ei gitt innsamling. Dette er informasjon som kan være koda på ulike måtar og kan vera vanskeleg å tyda. Der det er mistanke om at datasett kan innehalde slik informasjon vert dataeigar kontakta og spurt om tolkinga av data er riktig.

## Distribusjon av observasjoner

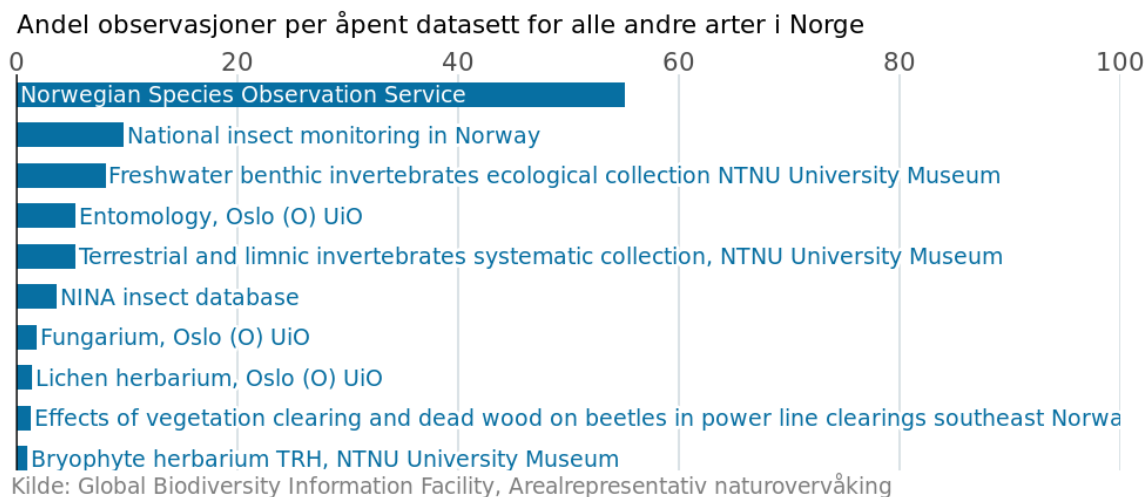
Andel observasjoner per åpent datasett for alle karplanter i Norge



Andel observasjoner per åpent datasett for alle fugler i Norge







**Figur 6:** Fordeling av observasjoner mellom ulike datasett for a) karplantar, b) fuglar og c) andre takson i Noreg. Dei 10 største datasetta målt i antall førekomster, er vist. Kjelde: GBIF.org (30 November 2023) GBIF Occurrence Download <https://doi.org/10.15468/dl.ebuzfh>

## 2.3 Miljødata

Å finna miljøvariablar som forklårar førekomst og utbreiing av ein art er utfordrande. Ideelt sett skal desse variablane fanga opp abiotiske og biotiske tilhøve som på ein god måte kan skildra den økologiske nisja til arten. Førekomst av artar vil også vera eit resultat av andre faktorar, som innvandringshistorikk eller påverknad frå menneske. Overhausting, forureining eller mangel på ressursar kan påverka utbreiinga av artar på ein måte som i liten grad kan fangast opp av miljøvariablar. Det er difor eit omfattande arbeid å finna gode variablar som forklårar kvifor ein finn ein art enkelte stader, og ikkje andre stader. Når oppgåva er å modellera utbreiinga av mange artar, må ein ha ei pragmatisk tilnærming til kva variablar som kan vera relevante for artsgruppa, og ikkje forventa å kunna fanga opp den økologiske nisja til kvar einskild art.

Dette prosjektet skal ikkje berre forklåra førekomsten av artar, men også predikera kvar det er sannsynleg at artar finst, også der det ikkje er gjort kartlegging. Det inneber å føreseie artsførekomst for verdier av miljøvariablar som ofte er utanfor dei verdiane som vart brukt i modelleringa.

Ein del data, spesielt som skildrar arealbruk, er produsert for ein eller eit par økosystem. Døme på dette er skogressurskart (SR16) som har ei rekke variablar som skildrar skogen på ein god måte (ref), men som manglar informasjon for områder som ikkje er definert som skog. For ein del av variablane gjev det inga mening for anna enn skog, slik som treslag og høgde, men andre variablar kan også relevant for andre økosystem, slik som "leaf area index". Det er også ei utfordring å bruka data med "manglande informasjon" for store delar av det geografiske området ein ønskjer å modellera på. Likevel kan slike økosystemspesifikke kartprodukt gje verdifull informasjon for arealklassar som elles er ganske grovt skildra, som jo er tilfelle for til dømes skog. Ein måte å bruka slike data kan då vera å auka den tematiske presisjonen til arealklassane i andre datakjelder.



Noreg er eit langstrakt land med ei total riksgrense på land på over 2550 km. Det betyr at store delar av norsk natur kan vera påverka av miljøtilhøve utanfor Noreg. Dette er spesielt relevant for ein del landskapsøkologiske variablar, slik som avstand til elv/bekk/innsjø og habitatvariasjon. For ein del miljøvariablar kan det då vera naudsynt å sjå utover dei nasjonale produkta slik at ein fangar opp relevant miljøinformasjon også langs riksgrensa.

Når ein skal finna relevante miljøvariablar for reproduserbare artsutbreiingskart er det eit krav at data er opent tilgjengeleg for ulisensiert bruk. Heldigvis vert meir og meir miljødata publisert ope, spesielt produkt frå fjernmåling. Ein del nasjonale kartprodukt er likevel lisensierte, noko som avgrensar bruk og vidarepublisering av desse. Når prosjektet har som føremål å laga heildekkande kart over artsførekomst og artsriksdom for Noreg, der sluttprodukta er tenkt brukt i praktisk forvaltning og prioriteringar om norsk natur, har me kome til at presisjonen til sluttproduktet er viktigare enn at alle data som inngår har ubegrensa bruk. Dei fleste data som er nemnd i kapitlet nedanfor er ope, men nokre har begrensingar i bearbeiding og bruk. Alle miljøvariablar har gode metadata slik at dokumentasjonen av arbeidsflyt og dei val og kriterium som inngår der, vil vera god.

I kapittel 2.3.1 drøftar me dei variablane som det er vanleg å bruka i artsutbreiingsmodellering, og nemner i tillegg nokre andre miljøvariablar. Undervegs i prosessen med modelleringa kan det dukka opp nye produkt eller lisensieringa på eksisterande produkt kan verta endra slik at det stettar krava om openheit. Desse nye produkta kan forbetra modellane og presisjonen til prediksjonane. I fall dette skjer vil det verta dokumentert i den opne kjeldekoda. Lista nedanfor må difor sjåast på som eit minimum av det som vil verta prøvd ut.

I statistisk modellering ønskjer ein eit kompromiss mellom ein god modell som fangar variasjonen i responsvariabelen bra, samstundes som den ikkje skal vera overparametrisert. Mange av variablane som er lista nedanfor samvarierer geografisk, eller dei bidrar i liten grad til å forklåra førekomst av artar for enkelte artsgrupper. Lista må difor sjåast på som eit utkast, der det vidare arbeidet vil avgjera kva som er naudsynt å ta med for å fanga opp den variasjonen i artsførekomst som finst i norsk natur for å gje gode prediksjonar om artsriksdom.

### *2.3.1 Moglege forklaringsvariablar*

Tidlegare forsøk på modellering av artsubreining og artsmangfald har brukt ei rekke biotiske og abiotiske variablar som er tenkt å vera knytt til den økologiske nisja til arten. I tillegg er ein del variablar tatt med for å ta omsyn til usystematisk innsamling av artsdata, slik som større innsamlingsintensitet i folketette områder. Dette er også utgangspunktet for dei vurderingane som er gjort for vårt prosjekt. For dei fleste av tema som er diskutert nedanfor, vil fleire variablar vera naudsynt for å forklåra miljøvariasjonen på ein god måte. Lista av moglege forklaringsvariablar kan difor verta lang, og det vil vera naudsynt å gjera eit utval basert på økologisk relevans, samvariasjon mellom variablar, og kvaliteten på dei data som skildrar miljøvariabelen.

## **Klima**

Klimatiske tilhøve har vist seg å vera svært viktig for å forklåra og predikera førekomst av artar. Nedbør og temperatur verkar å fanga det meste av den geografiske variasjonen i klima som er viktig for artar si utbreiing, men også snøtilhøve kan vera viktig.

Artar vert ulikt påverka av klima, både når det gjeld kva som er gode klimatiske tilhøve, og kva tid på året som klima har størst påverknad på arten (Sæther et al. 2022). Kva tid av året som best forklårar førekomst av ei heil artsgruppe vert difor eit kompromiss mellom kva som er statistisk mogleg å gjennomføra, og kva som er biologisk relevant. Dette vil også avhenga av kva tid på året ein ønskjer at artsutbreiingsmodellen skal vera gyldig. For dei fleste artsgrupper som det skal modellerast utbreiing for, er tilhøva om vinteren mindre viktig, då artane då har kvilestadium. Av den grunn er det biologisk mest aktuelt å bruka klima om vår og sumar i utbreiingsmodelleringa, og då begrensa seg til temperatur og nedbør. Snøtilhøve kan vera viktig for enkelte artar, og for dei fleste artane vil det vera lengde på perioden med snø på bakken som er viktig. Unnataket er ein del fuglar og pattedyr som vil få problem med mattilgang i djup snø. For mobile artar er det også relevant å modellera utbreiinga for andre tider på året enn sumaren, i den grad artsdata tilgjengeleg, og då vil vinterklima og snøtilhøve vera dei føretrekte klimavariablane.

Kva mål som er viktig kan også variera mellom artar. Ein tenkjer ofte at det er dei rådande tilhøva som er viktige for utbreiinga til ein art, til dømes snittemperaturen om sumaren dei siste 30 åra (normal-perioden). Men ekstrem-tilhøva kan også vera viktige for å forklåra artar sin førekomst (Stewart m.fl. 2021). Ein svært kald eller tørr periode i eit elles normalt år kan vera nok til at ein art forsvinn frå eit område, og viss slike ekstreme hendingar skjer ofte nok, vil ikkje re-etablering skje raskt nok til å oppretthalda levedyktige bestandar over tid. Ekstrem-vêr vil truleg verta meir vanleg i åra som kjem som følgje av klimaendringar.

Ekstrem-vêr vert vanlegvis skildra og kvantifisert som eit temporært fenomen, altså noko som seier noko om avviket frå det normale for eit gitt geografisk område, men det finst også definisjonar som tar utgangspunkt i tolegrensene til ein spesifikk art (van de Pol m.fl. 2017). Det er ei stor utfordring å gå frå ein temporær definisjon av ekstrem-vêr til noko som skildrar variasjon i forekomst av ekstreme hendingar romleg. Det er til dømes store ulikskapar mellom økosystem i kor stor variasjon i temperatur dei opplever; enkelte terrestriske system har gjerne ein forskjel på opptil 60° C mellom sumar og vinter, medan botnlevande organismar i havet opplever eit par graders variasjon gjennom året (Herfindal m.fl. 2022). Denne ulikskapen i opplevd temperaturvariasjon mellom økosystem eller mellom områder av det same økosystemet er i liten grad fanga opp av dei definisjonane som er brukt for ekstremver i dag (van de Pol m.fl. 2017).

Det er i dag lite kunnskap om kva rolle ekstremvêr har for utbreiinga av artar, og det gjer det vanskeleg å finna økologisk relevante mål på korleis ekstremvê varierer geografisk på ein måte som påverkar heile artssamfunn (Feldmeier m.fl. 2018). Det vil verta gjort ei vidare vurdering av i kva grad det er mogleg å få gode romlege kvantifiseringar av ekstrem-vêr som ikkje krev detaljkunnskap om tolegrensene til artane.

Noreg har relativt gode data om klimatiske tilhøve ope tilgjengeleg. Dei mest relevante produkta er observasjonsdata på temperatur og nedbør som er nedskalert til 1\*1 km<sup>2</sup> romleg oppløysing og aggregert på månad og siste normal-periode (1991-2020), og der endringar i forekomst av observasjonslokalitetar er tatt omsyn til (Kuya m.fl., 2022).

Nedskalerte produkt finst også basert på døgnobservasjonar. Frå desse kan tidspunkt for vårstart estimerast basert på den kumulative summen av døgngrader når døgnmiddeltemperaturen er minst 5 grader. Liknande mål kan brukast for kortid vekstsesongen sluttar, og dimed også lengda på vekstsesongen. Vekstsesongen for vegetasjon kan definerast på fleire måtar, og ulike datakjelder skildrar ikkje alltid dei same økologiske prosessane (Körner, Möhl and Hiltbrunner 2023). Det må difor gjerast ei nærare vurdering av i kva grad temperatur fangar den same romlege variasjonen i vekstsesong som t.d. fenologi basert på NDVI frå MODIS-satellittfoto. Det finst også ein del globale produkt om vêr og klima, slik som Worldclim (Fick m.fl. 2017) og Chelsa (Karger m.fl. 2017). Ei utfordring med ein del globale nedskalerte modellar er at presisjonen lokalt kan vera monaleg dårlegare enn for lokale (nasjonale) produkt.

## Topografi

Høgde over havet, helningsvinkel og helningsretning er dei mest vanlege topografiske variablane i artsmodellering. Høgde over havet fangar opp fleire faktorar som kan vera viktige for førekomst av artar, spesielt klimatiske tilhøve. Samanhengen mellom høgde over havet og klima varierer sterkt frå sør til nord, og kyst til innland. Til dømes er den klimatiske skoggrensa over 1300 m.o.h. i sentrale Sør-Noreg, og på havnivå (0 m.o.h.) på Varangerhalvøya i Finnmark (Moen 1999, Bryn og Potthoff 2018). Av den grunn er det vanleg å bruka forholdet mellom høgde over havet og den lokale klimatiske skoggrensa for å få eit mål som har nokolunde same meining over heile Noreg.

Helningsretning og -vinkel avgjer solinnstrålinga. Høg solinnstråling vil påverka temperatur, og det kan vera store forskjellar over små avstandar, spesielt i variert terreng. Slike lokale tilhøve er i liten grad fanga opp i nedskaleringa av meteorologiske observasjonar, og vil difor vera viktige bidrag til å forklara finskala variasjon i artsførekomst. Solinnstråling kan også reknast med ein digital terrengmodell. I stor grad vil dette gje same resultat som å bruka helningsretning og -vinkel, men det er også råd å ta omsyn til fjellskugge og atmosfæriske tilhøve som påverkar innstrålinga. Fjellskugge kan vera relevant for ein del område i Noreg. Atmosfæriske tilhøve er i hovudsak knytt til høge nivå av luftforureining, noko som i liten grad er relevant for Noreg.

Alle topografivariablane nemnt ovanfor vert rekna frå ein digital terrengmodell. Dette er tilgjengeleg for heile Noreg i ulik romleg oppløysing, der  $10*10m^2$  og  $100*100m^2$  er dei mest aktuelle for artsmodelleringa.

## Geologi, jordsmonn

For dei fleste planter er eigenskapar til jordsmonnet avgjerande for om dei kan leva ein stad eller ikkje. Dei viktigaste eigenskapane som i dag er kartfesta er geologi/berggrunn og hydrologi. Frå berggrunnskart får ein informasjon om kalkinnhald. Kalkinnhald er ein svært viktig indikator for vegetasjon, og brukt til dømes i NiN-systemet (Körner, Möhl and Hiltbrunner 2023). Kart over berggrunn og lausmassar er opent tilgjengelege gjennom Geonorge frå NGU som vektordata i ulik målestokk. NGU har også eit kartprodukt som viser kalkinnhald i berggrunn i 5 klassar, også dette ope tilgjengeleg for nedlasting som vektordata. Det er her verdt å merka seg at data om kalkinnhald gjev dette for

berggrunnen, og at dette ikkje naudsynleg er det same som tilgjengeleg kalkinnhald for vegetasjon. Kombinasjonen av lausmassedata og kalkinnhald i berggrunnen kan gje eit betre bilete på dei jordkjemiske tilhøva for vegetasjonen. Desse data finst i vektorformat, og kan rasteriserast til ønska romleg oppløysing. Presisjone på vektordata, dvs målefeilen på grensa mellom ulike berggrunntypar, er ikkje oppgjeve for berggrunnsdata i målestokk 1:250 000, som er den høgaste oppløysinga som dekker heile Noreg. For 1:50 000 berggrunnsdata er presisjonen gjeve til 50 meter. Ei rasteroppløysing på 100\*100m<sup>2</sup> verkar difor høveleg for berggrunnsdata i målestokk 1:250 000.

Sjølv om hydrologiske tilhøve er tett knytt opp mot nedbør og topografi, finst det lokale tilhøve som ikkje vert fanga opp med desse parametrane. Det finst så vidt me veit ingen gode kartprodukt som viser hydrologiske tilhøve for heile Noreg. NGU har eit produkt over lausmassar, men dette kartlaget vil truleg ikkje gje relevant informasjon utover det som finst i andre miljøvariablar. Ein del satellittprodukt har indeksar på jordfukt, til dømes MODIS-programmet. Desse produkta har relativt dårleg romleg oppløysing (1\*1km<sup>2</sup>), men det er utvikla algoritmar som kan brukast til å produsera indeksar på jordfuktigheit frå satellittfoto med betre romleg oppløysing. Det må må gjerast ei nærare evaluering av i kva grad dei aukar presisjonen på prediksjonane av artsutbreiing.

### **Primærproduktivitet**

Deriverte produkt om primærproduktivitet finst frå dei fleste satellittfoto, basert på at forholdet mellom raud og infraraud refleksjon er eit svært godt mål på fotosynteseaktiviteten til dei fleste planter. NDVI, EVI og FPAR er døme på slike indeksar. Ein del satellittar produserer slike data med kort tidsintervall, slik at ein får mål på primærproduksjon gjennom året. Basert på dette kan ein rekna netto primærproduksjon (NPP), men også viktige variablar som tidspunkt for vårstart og lengde på vekstsesong. Tidsseriar av t.d. NDVI gjennom sesongen er også viktig om ein ønskjer å bruka maksimal fotosynteseaktivitet, "peak NDVI" som variabel, fordi ulike naturtypar kan nå maksimalverdien til ulik tidspunkt på sumaren. Satellittar med høg romleg oppløysing, slik som Landsat og Sentinel, vil ikkje ha slike tidsseriar av vegetasjonsindeksar gjennom sesongen. Av den grunn er MODIS-produkta mest brukt til å skildra primærproduktivitet. Desse produkta har ei romleg oppløysing på 250\*250m<sup>2</sup>, temporær oppløysing på 1 dag som i praksis er aggregert til 15 dagar for å redusera periodar med skydekke, og finst frå 2001 og er framleis aktive. Det kan vera stor variasjon i NPP og fenologi mellom år, og det er difor viktig med ein lang tidsserie for å få eit godt mål på dei rådande tilhøva når det gjeld primærproduktivitet i eit område.

I område med sterk påverknad av menneskeleg aktivitet kan satellittbaserte mål på primærproduktivitet vera svært ulik nærliggande område med meir naturleg vegetasjon. Område med stor grad av nedbygging (byar og tettbygde strom) vert difor ofte maskert vekk. Men også i jordbruksområde der pløying og hausting gjennom vekstsesongen har stor påverknad for primærproduktiviteten, vil satellittbaserte mål på NPP eller fenologi vera svært annleis enn i naturlege habitat. Kulturlandskapet har ein stor andel av dei truga artane i Noreg. Den menneskelege aktiviteten som er viktig for at landskapet er såpass artsrikt kan samstundes vera problematisk for å få gode mål på primærproduktivitet i desse områda.

AR5 (Tenke 2016) har informasjon om bonitet, som seier noko om potensiell produktivitet for eit areal. Sidan AR5 har som hovudformål å seia noko om arealressursar er bonitet kun gjeve for produktiv skogsmark og AR5 er heller ikkje heildekkande for Noreg. Det same gjeld SatSkog, som også har informasjon om produktivitet. Desse to produkta kan difor ikkje brukast i heildekkande artsutbreiingsmodellering. Det beste produktet for primærproduktivitet og andre relaterte vegetasjonsmål, kjem difor frå MODIS-data.

## Arealbruk/habitattypar

Gode data om habitattypar er av dei viktigaste variablane som forklårar førekomst av artar. Årsaka til dette er at habitattypar og arealbruk i stor grad er gjeve av dei fysiske tilhøva på ein stad slik som klima og jordsmonn, saman med påverknad av menneske gjennom noverande eller historisk bruk av området. Tradisjonelt har arealbruk eller habitattypar i naturen ofte vore klassifisert basert på vegetasjon og eigenskapen deira, det vil seia basert på vegetasjonskartleggingar. Eit slikt grunnlag for modellering og predikering av førekomst av planteartar kan verka sirkulært, då det er førekomsten av ein del planteartar som indikerar vegetasjonstypen, som igjen vert brukt til å predikera førekomsten av dei same artane. For dei fleste planteartar, og alle andre artsgrupper, er dette ikkje tilfelle, og det er difor ønskjeleg med gode og presise kart over arealtypar i artsmodelleringa.

Tidlegare kartlegging av arealklassar baserte seg på tolking av flyfoto og kartlegging i felt. Av den grunn vart område av marginal økonomisk interesse ofte nedprioritert, og det tok gjerne lang tid før karta vart oppdaterte. Delar av AR5 baserer seg på slike kartleggingar. Dette gjer at AR5 ikkje er heildekkande, og at delar av AR5 er basert på gamle flyfoto der det er usikkert om kartet stemmer med dagens terreng. I dag er det aller meste av kartlegging av arealbruk basert på digitale fjernmålingsdata, anten flyfoto, satellittfoto eller ein kombinasjon. SatSkog er eit døme der båe datatypar er nytta (Ørka et al. 2019). Satskog har svært høg romleg oppløysing og har også fleire økologiske miljøvariablar. Satskog dekkjer stort sett produktive skogareal i Noreg, og er heller ikkje heildekkande for desse områda. Av den grunn kan Satskog ikkje brukast til heildekkande utbreiingsmodellering av artar i Noreg.

Romleg oppløysing, tematisk oppløysing og temporær oppløysing er tre viktige eigenskapar til arealklassekart, og desse er ofte avhengige av kvarandre. Årsaka til det er at det krev meir informasjon om ein ønskjer eit kart med mange ulike klassar, og for fjernmålingsdata betyr det å ha mange ulike spektrale band. Dette går ut over storleiken på området på bakken som kan gje informasjon til eit piksel, altså den romlege oppløysinga. Den temporære oppløysinga seier noko om kor ofte kartet kan oppdaterast, noko som er avhengig av kor ofte det finst nye fjernmålingsdata. Også her vil satellittar med høg romleg oppløysing ha lenger tid mellom nye data enn t.d. geostasjonære satellittar som gjev nye data gjerne fleire gonger om dagen, men med monaleg dårlegare romleg oppløysing.

I figur 7 er det vist døme på nokre av dei mest brukte heildekkande arealkarta over Noreg. Det nyaste av desse er ESA WorldCover, som er eit verdsdekkande arealklassekart med 10\*10m<sup>2</sup> oppløysing. Corine Land Cover (CLC) +Backbone er ei vidareutvikling av Corine-data, med 10\*10m<sup>2</sup> oppløysing. ELC10 er eit anna nytt produkt med 10\*10m<sup>2</sup> oppløysing (Venter og Sydenham 2021). Også Miljødirektoratet sitt kart

over Hovedøkosystem i Norge har høg romleg oppløysing. Utfordringa med desse produktet er ei dårlegare tematisk oppløysing samanlikna med produkt som har dårlegare romleg oppløysing. ESA WorldCover skil til dømes ikkje mellom ulike skogtypar, og veldig mange naturlege habitattypar som ikkje er skog, vert klassifisert som graseng (grassland). Dette gjeld til dømes det meste av myr og våtmark, men også beitemark og jordbruksområde for grasproduksjon. Også CLC +Backbone har 11 ulike arealklassar, medan Corine 100\*100m<sup>2</sup> oppløysing har 44 ulike arealklassar. ELC10 har 8 arealklassar. Det er difor lite ekstra informasjon i desse produkta samanlikna med t.d. markslagskartet frå N50-kartserien til Kartverket. Produkt med høg romleg oppløysing, men relativt dårleg tematisk oppløysing, kan likevel gje verdifull informasjon til dømes om fragmentering og heterogeneitet av naturlege arealklassar. Dette er informasjon som ofte forsvinn ved lågare romleg oppløysing. På same vis kan LiDAR-data auka informasjonen om ein del arealklassar, spesielt skog, ved å gje informasjon om høgde på vegetasjonen (t.d. Nybø m.fl. 2023).

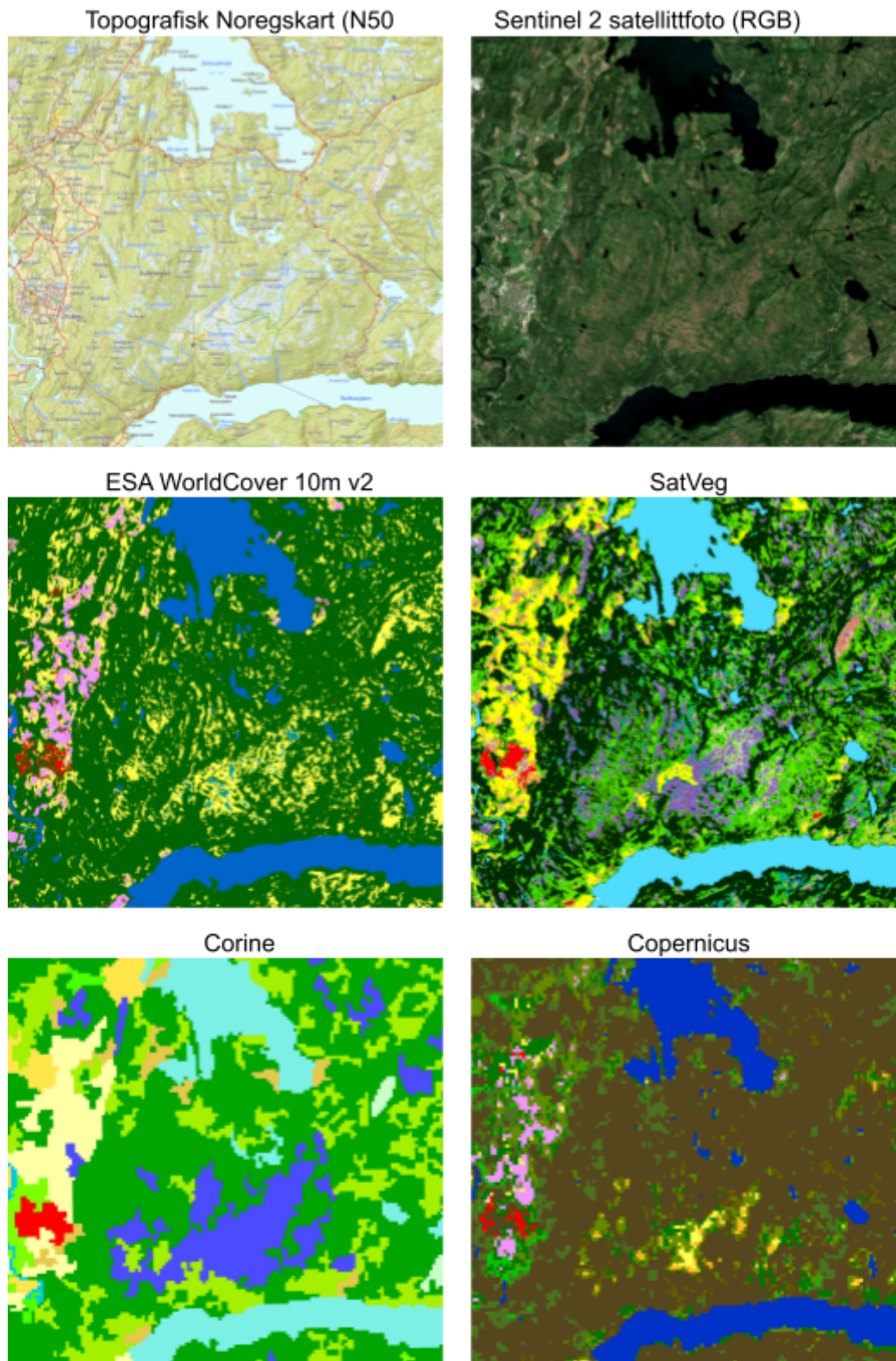
Corine og Copernicus er produkt med betre tematisk oppløysing (44 og 23 klassar), men lågare romleg oppløysing (100\*100m<sup>2</sup> for båe to). Båe desse produkta er utvikla for Europa. Utfordringa når klassifisering vert gjort over store område, er at ein prøver å skilja svært mange ulike arealklassar. Mange av desse kan vera ulike med tanke på vegetasjon og fysiske tilhøve, men likevel framstå like i den informasjonen som ligg i satellittdata. Ein risikerar då eit produkt som jamt over er riktig for dei vanlege arealklassane, men som lokalt kan ha feilklassifisering. Noreg utgjør ein liten del av Europa, og viktige habitattypar i Noreg kan verta nedprioritert for å auka presisjonen til andre og meir vanlege habitattypar i resten av Europa. Dette er tilfelle for tidlege versjonar av Copernicus, der presisjonen var god for vanlege habitattypar som jordbruksland, eviggrøn lauvskog, område utan vegetasjon, og vatn, medan typar som blandingsskog hadde låg presisjon. Den samla presisjonen for Corine sitt arealklassekart i 2015 er rundt 75% .

SatVeg er eit arealklassekart utvikla spesielt for Noreg og der arealtypane var knytt til tradisjonelle vegetasjonsklassar. SatVeg baserer seg på Landsat-satellittane, med ei romleg oppløysing på 30\*30 m<sup>2</sup>. SatVeg har ei høgare tematisk oppløysing spesielt for habitattypar som er viktige i Noreg (myr, skog og område over tregrensa). SatVeg vart ferdigstilt i 2006, og er basert på dels gamle satellittdata (det eldste er frå 13. juli 1993, det nyaste frå 29. juli 2006). Det meste av SatVeg er dermed over 20 år gammalt. SatVeg har også ein del utfordringar med feilklassifisering eller ukjent arealklasse i bratt terreng.

Det finst fleire data over arealbruk og habitattypar enn dei som er lista ovanfor, men dei utfordringane som er diskutert vil også gjelda desse andre kjeldene. Desse utfordringane er tematisk oppløysing, romleg oppløysing, feilklassifisering, og kor oppdaterte dataa er, eller den temporære overlappen med dei artsdataa som skal koplast til arealbruken. Vår vurdering er at tematisk oppløysing, til dømes at det er kartlagt fleire ulike skogstypar, er viktigare enn å ha ei romleg oppløysing på 10\*10 m<sup>2</sup>. Arealklassekart frå Corine og Copernicus finst båe på 100\*100 m<sup>2</sup>, som også er den romlege oppløysinga som er valgt for artsdata. Desse to produkta finst også tilbake i tid til år 2000, noko som gjer det mogleg å gjera endringsanalysar. Vi har gjort førebelse undersøkingar som tyder på at Corine har den mest presise arealklasseskildringa for norsk natur, og vil vera det føretrekte arealklasseproduktet for modellering av artsutbreiingskart.

Som drøfta ovanfor er det vanleg at arealklasseprodukt inneheld ein del feilklassifisering. Sjølv om denne usikkerheita kan kvantifiserast, er den sjeldan stadfesta, slik at den ikkje kan takast omsyn til når produktet skal brukast i andre samanhengar. Arealklassekart kan altså sjåast på som prediksjonar frå ein modell basert på rådata frå fjernmåling, der rådata er mengda av refleksjon av ulike bølgelengder, skalert til verdiar mellom 0 og 255. Når føremålet med fjernmålingsdata er å modellera og predikera artsutbreiing, kan det vera førmålstenleg å bruka rådata frå satellittfoto, utan å gå via ein klassifiseringsprosedyre som både forenkler naturen til eit begrensa tal arealklassar, og som inneheld usikkerheit. Å bruka dei mest relevante banda frå satellittfoto som kontinuerlege forklaringsvariablar har vist seg å predikera artsførekost like bra som klassifiserte produkt, spesielt om ein tillet ikkje-lineære samanhengar mellom refleksjon og artsførekost (Sæther m.fl. 2021). Fordelen med å bruka arealklasse er at parametraner til modellen for artsutbreiing er enklare å forstå, og at resultatane kan brukast i praktisk forvaltning. Det gjev meir mening å forvalta natur basert på arealklassar som lauvskog eller myr, enn å basera resultatane basert på refleksjon av nær-infraraud, raud, grøn og blå bølgelengder.





**Figur 7:** Nokre kjelder for arealklassar. N50 topografisk Noregskart er vektorbasert og har høg romleg presisjon, men eit begrensa tal arealklassar. Sentinel satellittdata har romleg oppløysing på  $10 \times 10 \text{ m}^2$ , og er brukt blant anna til eit heildekkande verdskart med same romlege oppløysing og 11 arealklassar. Satveg er eit vegetasjonskart tilpassa norske naturtypar, med relativt høg romleg oppløysing på  $30 \times 30 \text{ m}^2$ , og god tematisk oppløysing. Corine og Copernicus dekkjer Europa og verda, og har  $100 \times 100 \text{ m}^2$  oppløysing og ganske god tematisk oppløysing.



## Infrastruktur og andre variablar knytt til menneske

Mange artar er tett knytt til menneske og drar nytte av tilhøve som me skapar. Av den grunn er informasjon om folketettleik, bygningar og vegar og anna infrastruktur, vanleg å bruka i artsmodellering. Sjølv om desse variablane dels har økologisk relevans for ei rekke artar, er det ofte påverknaden på korleis data vert innsamla som vert framheva som den viktigaste årsaka til å ta med denne type variablar. Av den grunn er avstand til museumssamlingar også brukt til å forklåra artsforekomstar. Årsaka ligg i at folkeforsking og anna innsamling generelt, skjer der folk er, og då er folketettleik og avstand til veg viktigast. Utfordringa i tidlegare artsutbreiingsmodellar er at dei ikkje skil mellom den økologiske effekten av desse variablane, og effekten dei har på observasjonsprosessen.

Vårt forslag for artsmodellering skil ut observasjonsprosessen i ein eigen modell, noko som gjer at det ikkje skal vera naudsynt å ta omsyn til variablar som kan forklåra innsamlingsprosessen. Ein sit då ideelt att med ein rein økologisk effekt av desse variablane på artsutbreiinga, viss modellen for innsamlingsprosessen er god. Ei utfordring med variablar knytt til menneske er at desse variablane ikkje er vilkårleg plassert i landskapet. Desse variablane vil difor samvariera sterkt med ei rekke andre variablar, slik som produktivitet, høgd over havet, og avstand til vatn. Det vil likevel vera interessant å bruka variablar på infrastruktur og andre mål på menneskeleg aktivitet, for å få eit inntrykk av kor godt innsamlingsprosessen er modellert. Alle desse variablane er anten ope tilgjengeleg eller kan produserast frå opne kartdata.

### Andre miljøvariablar

Viktige faktorar som ikkje er fanga opp av variablane skildra ovanfor er spesielt knytt opp mot landskapet rundt ein gitt lokalitet. Til dømes er tilgang til andre naturtypar eller miljøtilhøve, enn det som finst på ein gitt lokalitet, viktig for mange artar. Av den grunn er til dømes avstand til ope vatn (elv, bekk, innsjø) ein mykje brukt variabel. Det same er kortaste avstand til tregrensa eller kortaste avstand til anna naturtype/arealklasse enn den lokaliteten ligg i. Felles for desse type variablar er at dei fortel noko om landskapet rundt lokaliteten. Dette kan vera svært viktig då mange artar er avhengige av ulike miljø for ulike livsstadier. Desse variablane baserar seg i stor grad på kartdata som er ope tilgjengeleg, slik som N50-kartdata frå Kartverket. Det er likevel viktig å sikra at heile landskapet vert vurdert også for områder nær riksgrensa. Difor kan internasjonale data vera aktuelle, sjølv om den romlege presisjonen ikkje er like bra. Eit slikt dataset er EU-Hydro River Network Database frå European Union's Copernicus Land Monitoring Service information.

### 2.3.2 Data for endringsanalysar

Arealbruksendringar har den siste tida fått stor merksemd, og eit viktig spørsmål både for forskning og forvaltning er kva effekt slike arealbruksendringar har på naturmangfaldet og førekomsten av artar. Å svara på dette krev gode data på arealbruksendringar, altså tidsseriar med arealbrukskart. Som andre tidsseriar er det viktig med ein konsistent metodikk over tid, slik at ein er sikra at dei endringane som vert kvantifisert i produktet, er faktiske endringar og ikkje som følgje av endra metodologi t.d. i klassifisering av satellittdata.

Det er stor variasjon mellom artar i kva grad dei vert påverka av arealbruksendring, og kor raskt ei endring av leveområda deira resulterer i endra utbreiing. Likeins er det stor forskjell mellom ulike typar endring. Nokre endringar, ofte reversible endringar, har mindre økologisk effekt enn dei irreversible endringane som ofte har umiddelbar effekt på førekomsten av artar. Produkt som viser arealbruksendringar bør ha ei tematisk oppløysing som viser att desse ulike typane endring.

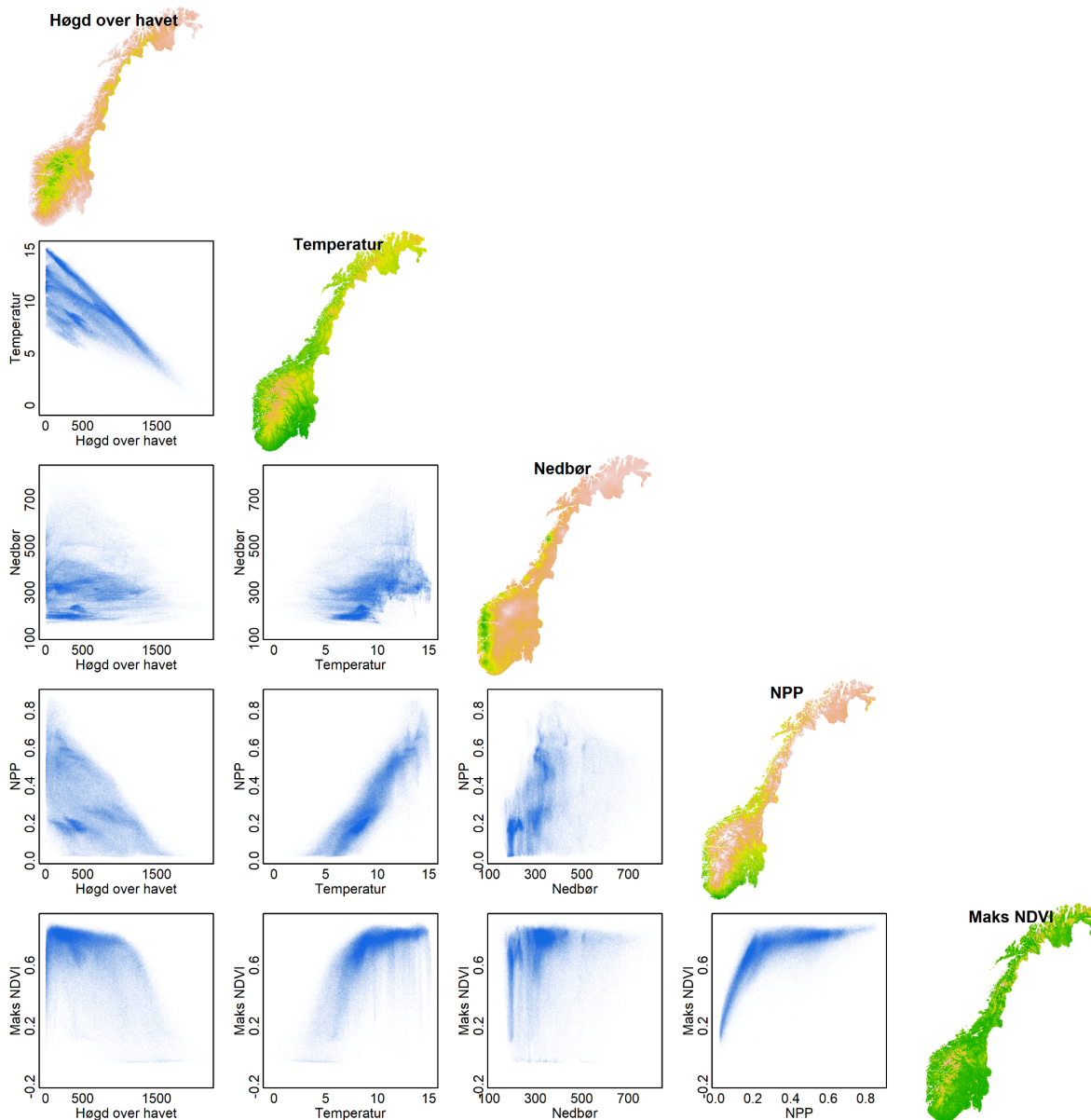
Fjernmålingsdata har vist seg å vera svært godt eigna for å kartleggja og kvantifisera endringar i arealbruk. Til dømes produserer Corine land cover oppdaterte data om arealbruk kvart sjette år frå 2000, og har også eit produkt frå 1990. Dette representerar såleis ein av dei lengste standardiserte tidsseriane over arealbruksendringar som er tilgjengeleg. Ei utfordring er at mykje av dei endringane som skjer i landskapet kan vera små i utstrekning, og ikkje verta fanga opp av 100\*100m<sup>2</sup> ruter. Med ny sensorteknologi vert produkt med betre romleg oppløysing tilgjengeleg, slik som Dynamic Wold (Brown m.fl. 2022) og ELC10 (Venter og Sydenham 2021), som vert hevda å gje ei nær kontinuerleg oppdaterting av arealbruken og endringar. Dette er spesielt relevant for ein del små-skala endringar slik som skogsbilveggar, og teknologien gjer det også mogleg å fanga opp reversible endringar slik som hogst. Utfordringa med desse produkta er korte tidsseriar, slik at det ikkje kan brukast til å seia noko om endringar før 2020. Dei viser likevel kva potensiale som ligg i ny teknologi for å starta overvåka arealbruksendringar i tida som kjem.

### 2.3.3 Nokre tankar om val av miljøvariablar

#### Samvariasjon

I artsmodellering, som all anna statistisk modellering, er målet å forklåra så mykje av variasjonen i ein variabel, her førekomst av ein art, på ein så enkel måte som mogleg. I praksis betyr det at ein unngår å ta med variablar som i liten grad aukar presisjonen til modellen. Ei av dei viktigaste årsakene til dette er at jo fleire variablar, jo meir sannsynleg er det at dei samvarierer og dimed "sloss" om å forklåra den same variasjonen i førekomst av ein art. Dette vert spesielt viktig når ein skal predikera frå denne modellen. Samvariasjon mellom forklåringsvariablar vil kunna gje store utslag i estimert effekt av ein av variablane, spesielt for områder der dei to variablane ikkje samvarierer i like stor grad. Predikeringar kan då verta uventa og lite pålitelege i desse områda.

Basert på dette er det difor viktig å gjera ei vurdering på førehand om kva som er relevante variablar for den artsgruppa ein ønskjer modellera og predikera utbreiinga til, og i kva grad desse variablane skildrar den same romlege variasjonen i miljøfaktorar i landskapet. Eit døme på miljøvariablar som i stor grad er samvarierer geografisk er klimavariablar, slik som temperatur og nedbør, og netto primærproduksjon (NPP). Likeins finn ein ofte variablar som skildrar vegetasjonsutvikling gjennom sesongen (fenologi), slik som dato for vårstart eller lengde på vekstsesong. Dette er i stor grad påverka av temperatur og snøtilhøve, som også er forklåra av vinternedbør. Det er difor viktig å gjera ei vurdering av samvariasjon mellom miljøvariablane for å unngå overparametrisering og at lokale områder med avvik mellom denne samvariasjonen får stor påverknad på dei predikerte artsutbreiingane.



**Figur 8:** Nokre forklaringsvariablar og deira geografiske samvariasjon. For detaljar om datasetta, sjå Tabell 2. Temperatur og nedbør er for månadane mai til august. Stor samvariasjon kan ha uventa effekt på prediksjonar, ved at områder der samvariasjonen er låg, får uforholdsmessig stor påverknad på den statistiske modellen som ligg til grunn for predikeringa.

### Økologisk relevans

Det vil i praksis ikkje vera råd å gjera artsspesifikke vurderingar for alle artar som skal modellerast, både fordi det er eit svært omfattande arbeid, og fordi grunnleggjande økologi er ukjent for mange artar. Ein lyt difor freista finna variablar som har økologisk relevans for artsgruppa som heilheit, og såleis sjå bort frå artsspesifikke samanhengar. Det betyr ikkje at modelleringa ikkje tillet ulik effekt av ein variable på ulike artar, men at det er dei same forklaringsvariablane som inngår i modelleringa av kvar einskild art.

### Romleg skala

Dei fleste artar er avhengige av ulike miljø i ulike livsstadier. Unnataket er fastsitande organismar, slik som planter, men også mange av desse er indirekte avhengig av andre miljø gjennom artsinteraksjonar slik som pollinering. Det er også ein tett samanheng mellom storleiken på eit område med ein viss naturtype, og levedyktigheita til bestandar som er avhengig av denne naturtypen. Dette er omsyn som sjeldan vert tatt med i artsmodellering. Utfordringa er å velja den riktige romlege skala for å vurdere t.d. kva ulike habitat som finst tilgjengeleg, og kor store områder desse dekkjer. Det er stor variasjon mellom artar i kor mykje dei flyttar på seg, og for mange artar er kunnskapen om områdebruk mangelfull. Det er difor utanfor prosjektet sitt formål å gjera ei vurdering av slike landskapsøkologiske faktorar for artsutbreiing. Det ville likevel vera interessant å gjera vurdering av effekten av romleg skala, til dømes ved å bruka den romlege oppløysinga til ANO-rutenettet (500\*500m<sup>2</sup>) for å kvantifisera ein del variablar knytt til romleg miljøvariasjon.

### 2.3.5 Forslag til val av forklaringsvariablar

Tabell 2 listar opp dei variablane som til no er vurdert aktuelle for modellering av utbreiing av artar og artsmangfald som skal gjerast i dette prosjektet. Som nemnd ovanfor kan den endelege lista verta mindre, men nye variablar kan også verta lagt til. Merk at me her tar for oss variablane som inngår i modelleringa, og ikkje etteranalysane som til dømes skal sjå på potensielle effektar av endra habitat og arealbruk (sjå kap. 2.3.2 *Data for endringsanalysar*)

**Tabell 2:** Tema og miljøvariablar som er vurdert å inngå i artsutbreiingmodelleringa i prosjektet.

Tema	Variabel	Egenskapar	Kjelde
Klima	Sumar-temperatur, snitt	snitt over perioden mai - august, 1*1km <sup>2</sup> , 1991-2020	<a href="https://thredds.met.no/thredds/catalog/KSS/Gridded_climate_normals_1991-2020/temperature/catalog.html">https://thredds.met.no/thredds/catalog/KSS/Gridded_climate_normals_1991-2020/temperature/catalog.html</a>
	Sumar-nedbør, sum	sum over perioden mai - august, 1*1km <sup>2</sup> , 1991-2020	<a href="https://thredds.met.no/thredds/catalog/KSS/Gridded_climate_normals_1991-2020/precipitation/catalog.html">https://thredds.met.no/thredds/catalog/KSS/Gridded_climate_normals_1991-2020/precipitation/catalog.html</a>
	Lengde på snødekke	dagar med snødekke > 5cm eller 25 cm, 1*1km <sup>2</sup> , 1991-2020	<a href="https://kartkatalog.geonorge.no/metadata/klima-dager-med-snoe-paa-bakken/d560cb5d-6464-4df5-8e10-6c2df77f95a7">https://kartkatalog.geonorge.no/metadata/klima-dager-med-snoe-paa-bakken/d560cb5d-6464-4df5-8e10-6c2df77f95a7</a>
	Ekstrem-temperatur	dagar med avvik > 1SD frå månadsnormal, 1*1km <sup>2</sup> , 1991-2020	
	Ekstremnedbør	Dagar med avvik 1SD frå månadsnormal, 1*1km <sup>2</sup> , 1991-2020	
Topografi	Høgde over havet	10*10m <sup>2</sup> eller 100*100m <sup>2</sup>	<a href="https://register.geonorge.no/produktark/dtm-10">https://register.geonorge.no/produktark/dtm-10</a>
	Helningsretning	10*10m <sup>2</sup> eller 100*100m <sup>2</sup>	Basert på Høgde over havet
	Helningsvinkel	10*10m <sup>2</sup> eller 100*100m <sup>2</sup>	Basert på Høgde over havet
	Solinnstråling	10*10m <sup>2</sup> eller 100*100m <sup>2</sup>	Basert på Høgde over havet

Tema	Variabel	Egenskapar	Kjelde
Geologi	Berggrunn	Vektordata, variabel nøyaktighet	<a href="https://kartkatalog.geonorge.no/metadata/berggrunn-n250/7c39be66-77b6-4b74-b58d-53b6bee90067">https://kartkatalog.geonorge.no/metadata/berggrunn-n250/7c39be66-77b6-4b74-b58d-53b6bee90067</a>
	Kalkinnhald	Vektordata, variabel nøyaktighet	<a href="https://kartkatalog.geonorge.no/metadata/berggrunn-n250/7c39be66-77b6-4b74-b58d-53b6bee90067">https://kartkatalog.geonorge.no/metadata/berggrunn-n250/7c39be66-77b6-4b74-b58d-53b6bee90067</a>
	Jordsmonn	Vektordata, variabel nøyaktighet	<a href="https://kartkatalog.geonorge.no/metadata/loesmasser/3de4ddf6-d6b8-4398-8222-f5c47791a757">https://kartkatalog.geonorge.no/metadata/loesmasser/3de4ddf6-d6b8-4398-8222-f5c47791a757</a>
	Jordfuktighet	1*1km <sup>2</sup> , avhengig av satellitt	
Primærproduktivitet	Netto primærproduksjon (NPP)	250*250m <sup>2</sup> , 2001-2022	<a href="https://modis-land.gsfc.nasa.gov/npp.html">https://modis-land.gsfc.nasa.gov/npp.html</a>
	Maksimum NDVI	250*250m <sup>2</sup> , 2001-2022	<a href="https://lpdaac.usgs.gov/products/mod13q1v061/">https://lpdaac.usgs.gov/products/mod13q1v061/</a>
	Vårstart, hauststart, lengde på vekstsesong	250*250m <sup>2</sup> , 2001-2022	Basert på <a href="https://lpdaac.usgs.gov/products/mod13q1v061/">https://lpdaac.usgs.gov/products/mod13q1v061/</a> eller <a href="https://kartkatalog.geonorge.no/metadata/klima-vekstsesong/5566c9b-814b-4474-829c-eee13045ad66">https://kartkatalog.geonorge.no/metadata/klima-vekstsesong/5566c9b-814b-4474-829c-eee13045ad66</a>
Arealklasse	Corine Arealdekkekart	23 klassar, 100*100m <sup>2</sup>	<a href="https://land.copernicus.eu/en/products/corine-land-cover">https://land.copernicus.eu/en/products/corine-land-cover</a>
	SR16	17 variablare som skildrar skogen, brukast i kombinasjon med Corine arealdekkekart	<a href="https://kart8.nibio.no/nedlasting/dashboard">https://kart8.nibio.no/nedlasting/dashboard</a>
	LiDAR høgdemodell	Høgde på vegetasjon (overflatehøgde - terrenghøgde)	<a href="https://hoydedata.no/LaserInnsyn2/">https://hoydedata.no/LaserInnsyn2/</a>
Infrastruktur og andre menneskeskapte faktorar	Avstand til veg	Basert på vektorkart, 2023	<a href="https://register.geonorge.no/produktspesifikasjoner/n50-kartdata">https://register.geonorge.no/produktspesifikasjoner/n50-kartdata</a>
	Folketettleik	250*250m <sup>2</sup> , 2016-2019	<a href="https://kartkatalog.geonorge.no/metadata/befolkning-paa-rutenett-250-m/0c0ad0ce-55e8-4d73-9c12-0eb0e2454acb">https://kartkatalog.geonorge.no/metadata/befolkning-paa-rutenett-250-m/0c0ad0ce-55e8-4d73-9c12-0eb0e2454acb</a>
	Tettheit av bygningar	Basert på vektorkart, 2023	
	Avstand til museums-samlingar	Basert på vektorkart, 2023	
Andre variablar	Habitatvariasjon	Talet på habitattypar og deira fordeling innan 500*500m <sup>2</sup>	Basert på Corine arealdekkekart, CLC+Backbone eller andre høgoppløysede arealbrukskart
	Avstand til vatn	Basert på vektorkart over innsjø, bekk og elv	<a href="https://register.geonorge.no/produktspesifikasjoner/n50-kartdata">https://register.geonorge.no/produktspesifikasjoner/n50-kartdata</a>

Tema	Variabel	Eigenskapar	Kjelde
	Avstand til tregrensa	Forholdet mellom lokal høgd over havet og den klimatiske tregrensa, basert på digital terrengmodell og vektorkart.	Basert på Høgd over havet og data frå Bryn og Potthoff 2018.

## 3 Arbeidsflyt

### 3.1 Dokumentasjon av arbeidsflyt

All programmeringskode for prosjektet er opent tilgjengeleg på delingsplattforma GitHub<sup>12</sup>. Ei oversikt over struktur og innhald på prosjektsida ligg på README-dokumentet. All kode er skriva i programmeringsspråket R (R Core Team, 2023). Språket er svært mykje brukt innan vitenskapleg arbeid og statistikk. Ein utfyllande teknisk dokumentasjon av arbeidsflyt, som også inkluderer køyrbar kode og lenker til data, er tilgjengeleg på prosjektarkivet på GitHub. Nedanfor gjev me ei oppsummering av arbeidsflyten.

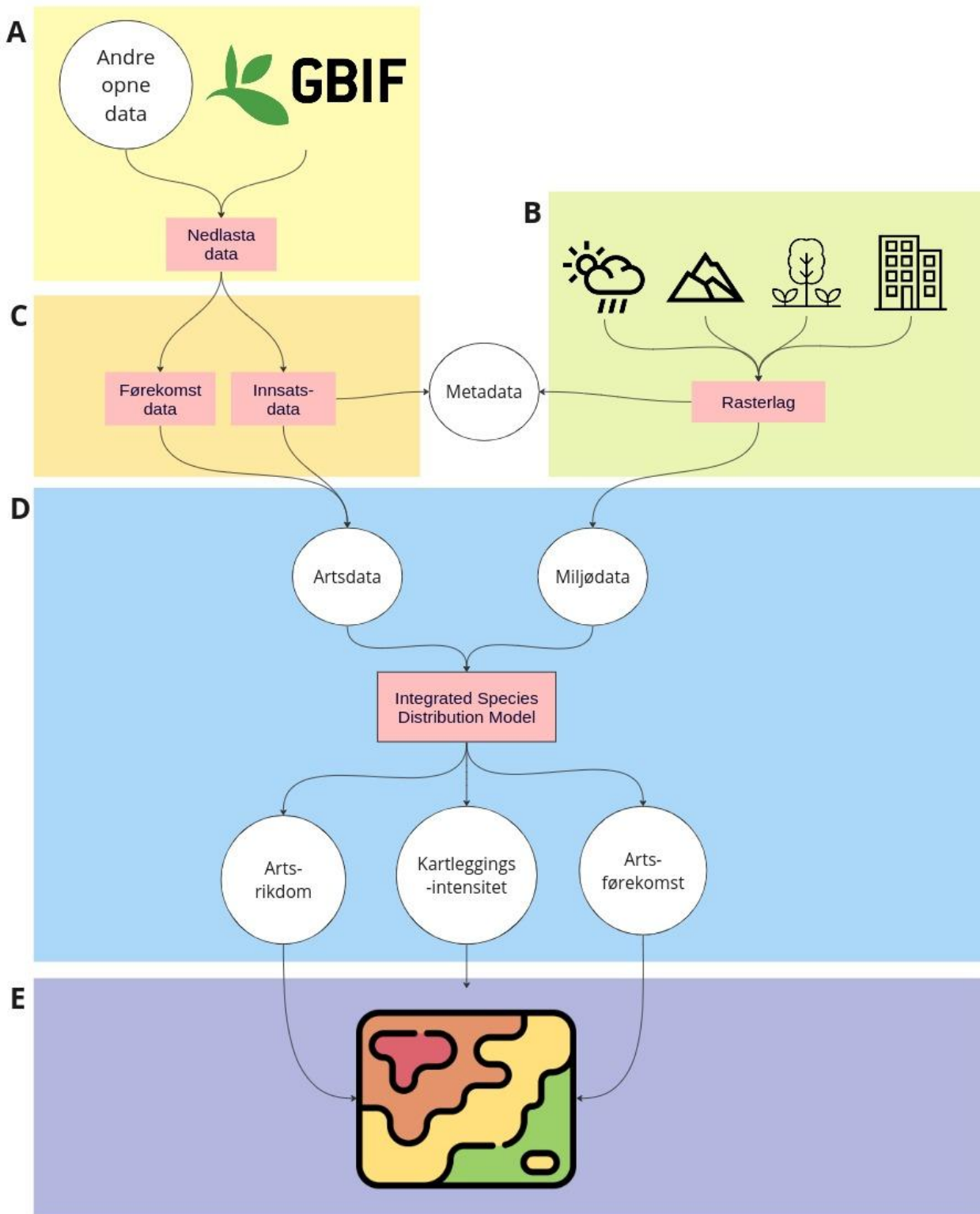
### 3.2 Skildring av arbeidsflyt

Tre variablar trengs for å starta arbeidsflyten og som (i) definerer kva for artar som skal være i fokus, (ii) kva for miljøvariablar som skal forklara utbreiinga til desse artane og (iii) ei liste over kva for datasett som skal inkluderast, og om nokon av desse datasetta må behandlast med eigne formateringskript før dei vert inkludert:

- Liste over takson
- Liste over miljøvariablar med kjelde
- Liste over datasett med merknadar om spesiell prosessering er naudsynt samt lenke til prosesseringskript.

Takson og miljøvariablar er skildra i seksjon 2.1.1 og 2.2.1. Liste over datasett vert manuelt kuratert på gjeldande tidspunkt for å ha mest mogleg oppdaterte data. Det er enno ikkje mogleg å gjera kvalitetskontroll og kontroll av innhald i dei ulike datasetta automatisk. Dette må difor gjerast manuelt for kvart datasett. Manuell gjennomgang vert stoppa når 99% av alle observasjonar av ei artsgruppe er gjennomgått. Dei resterande 1% er ofte fordelt på svært mange datasett og den manuelle arbeidsmengda vert difor uforholdsmessig stor i høve til vinsten på å inkludera data i modelleringa.

<sup>12</sup> <https://github.com/gjearevoll/BioDivMapping>



Figur 9: Visualisering av Hotspots-prosjekt sin arbeidsflyt

Arbeidsflyten kan kjørest ved bruk av eit "masterscript"<sup>13</sup>. Den tekniske skildringa under følger rekkefølga til dette. Kvar av undertitlane er lenka til ei korresponderande fil som køyrer relevant kode i R.

<sup>13</sup> <https://github.com/gjearevoll/BioDivMapping/blob/main/masterScript.R>



## A. Import av artsdata

Data er lasta ned fra GBIF ved bruk av GBIF API og asynkron nedlastning via *rgbif*-pakka (Chamberlain et al. 2022). Denne pakka er en applikasjon for programmeringsspråket *R* som lager en “wrapper” rundt GBIF API’et.

Data er filtrert før nedlastning basert på takson, som skildra i lista over takson som skal nedlastast (*focalTaxa.csv*). Det vert også filtrert på geometri (landområde i Norge, minus Svalbard og Jan Mayen) og ulike kvalitetsparametre som koordinatusikkerheit. Original filtrering vil vera tilgjengeleg via lenke til nedlastningssiteringa<sup>14</sup>.

Etter nedlastning er datasettet frå GBIF delt i ulike tabellar basert på kva datasett dei høyrer til. Basert på referanse vert så metadata frå ein skilde datasett lasta ned. Metadata inkluderer også ei lenke til opphavlege datasett slik det er publisert frå opphavskjelda (“Darwin Core Archive”; DwC-A). I nokre tilfeller er det naudsynt å bruke data frå kvart DwC-A der GBIF sentralt filtrerar bort relevant informasjon. Dette gjeld i hovudsak informasjon om kvantitative mål og kartleggingsinnsats. Andre opne datasett blir også lagt til i arbeidsflyten no, etter å ha blitt standardisert og gjort interoperabile med GBIF data. No gjeld dette berre Arealrepresentativ Overvåkning (ANO).

På dette steget i arbeidsflyten vil også informasjon om raudlistestatus bli lagt til datasettet basert på ei nedlastning frå Artsdatabanken sin API. Artsnamn frå raudlista er resolvert og sjekka mot det den taksonomiske ryggrada til GBIF for å sikre inkludera av synonymiserte artar. Aksepterte nam frå GBIF si taksonomiske ryggrad er då utgangspunkt for resten av den tekniske arbeidsflyten. Det er verdt å merke at GBIF si taksonomiske ryggrad er basert også på Artsnavnebasen, men dei tekniske løysningane rundt namneresolvering er betre fasilitert ved å jobba mot GBIF API’et.

## B. Import av miljødata

Per i dag er eit utval av miljødata lasta ned lokalt. Data er standardisert og konvertert til rasterdata ved nedlastning, og klippa til relevant utstrekning. Miljødata vert vidare samla i eit objekt der alle rasterlaga har same romlege eigenskapar (oppløysing og utstrekning). Oppløysing kan definerast særskilt for kvar enkelt køyring av modellen og også nedskaleras seinare om det er hensiktsmessig til dømes for visualisering.

## C. Prosessering av artsdata

Delar av artsdata har meir informasjon enn berre førekomst av artert, til dømes informasjon om kva artar som ikkje leita etter, men ikkje funne (“absence-data”). Desse datasetta er manuelt identifisert frå metadata. I utgangspunktet har GBIF ein eige datatype “Sampling-event data” for slike data. Det er derimot ikkje ei konsekvent bruk av dette, og data med informasjon om funn / ikkje-funn blir også skildra i andre datatyper

---

<sup>14</sup> Til dømes <https://doi.org/10.15468/dl.ebuzfhz>



brukt for opportunistiske innsamla data (“occurrence data”). Dersom det ikkje er eintydig frå metadata korleis data er koda, vert opphavsperson kontakta.

Den vanlegaste situasjonen er at på datasettnivå er dokumentasjon av innsamlinga avgrensa til ei gitt taksonomisk gruppe. Innanfor kvar innsamlingshending vert då fråvær av artar som ikkje er lista som funne lagt til som data om fråvær. Informasjon om fråvær av artar i innsamlingshending og kvantitativ informasjon er ofte koda på ulikt frå datasett til datasett. Kvart datasett vert derfor ofte omforma med skreddarsydde skript. Der det er tvil om struktur og innhald vert opphavsperson kontakta, og om ikkje dette er klårgjerande vert datasettet brukt som førekomstdata.

På dette steget i dataprosesseringa er det fleire funksjonar som legg til eller gjer om data til bruk i modellering eller visualisering.

- Trusselkategori frå raudlistevurdering
- Ei vurdering av om det er nok data til å inkludera arten i modelleringa
- Generering av funn / ikkje-funn og innsatsinformasjon i datasett der slik informasjon har blitt fjerna før publisering men er mogleg og rekonstruere frå informasjon i metadata eller frå opphavsperson
- Produsera metadata for kvart datasett slik at omforming og opphav kan reproduseres
- Nedlastning av anna dokumentasjon, til dømes foto, for bruk i visualisering.

#### **D. Ubreeingsmodellering**

Modelleringa er basert på ei føresetnad om at det for kvar art er eit sant tal individ til stades på ein kvar lokalitet til ei kvar tid. Alle datasett har sine egne eigenskapar som må modellerast på sin måte. Me utviklar difor ein eigen observasjonsmodell for kvart datasett og ein prosessmodell for den faktiske utbreiinga av kvar art.

På grunn av måten me skil mellom prosess- og observasjonsdel av modellen, for kvart datasett, kan me enkelt kombinera desse og ha fleire observasjonsmodellar, ein for kvart datasett, for kvar art. Dette gjev ein integrert populasjonsmodell (Isaac et al. 2020), der datasetta kan tilføra informasjon til kvarandre. Til dømes vil strukturerte data med kvantitative observasjonar eller data for artar som ikkje er observerte, men søkt etter, ikkje ha skeivheiter for om ein art er observert eller ikkje. Desse data kan då samanliknast med førekomstdata og skeivheita i førekomstdata kan estimeres (Simmonds et al. 2020).

Prosessmodellen er ein modell for kvar arten er og korleis dei ulike miljøvariablane påverkar dette. Modellen er samanhengande i rom, og forenkler det romlege ved å anta at dette er bygd opp av trekantar: storleiken på desse trekantane bestemmer oppløysinga til den endelege modellen vår. Artsfordelinga er overflata på dette rommet: jo høgare den er, jo større sannsyn for at arten er der. Høgda er modellert med miljøvariablar og gjev ein ekstra romleg effekt.

I praksis er heile modellen tilpassa med ei Bayesiansk tilnærming ved bruk av INLA-programvare. Dette gjer at modellkøyringane går relativt raskt. Denne metodikken

er også fleksibelt slik at me kan tilpassa modellane til ulike datakjelder og måten dei er samla inn på. Me har utvikla ein R-pakke, PointedSDMs (Mostert and O'Hara 2023) for å forenkla modellutviklinga. Teknisk modellgrunnlag samt forutsetninger er skildra i Mostert og O'Hara (2023)

Utbreiingsmodelleringsmodulen gjev tre hovudprodukt:

- Romlege estimater for førekomst av ein-skilde artar
- Romlege estimater for artsrikdom
- Romlege estimater for kartleggingsintensitet

#### **E. Post-prosessering**

Resultat frå artsutbreiingsmodelleringa kan bearbeidast internt i R ved hjelp av ulike romlige program, eller lastas ned og jobbast vidare med i ulike GIS-program.

## 4. Avleia produkt

Det primære resultatet frå prediksjonene er førekomstar for artar, artsrikdom , og kartleggingsinnsats. Desse produkta kan åleine eller saman med ulike miljø- og kartprodukt, gje fleire relevante produkt. Dei viktigaste er lista opp i Tabell 3.

**Tabell 3:** Oversikt over nokre moglege deriverte produkt basert på modellering av artsutbreiing. Produkta er anten basert på overlegging eller sone-analysar frå eitt eller fleire kartprodukt.

Produkt	Beskriving
Områder med særskild høg artsrikdom.	Heildekkande kart som identifiserer områder med særskild høg artsrikdom, særskild mange truga artar og truga ansvarsartar. Verdier er gjeve på kontinuerleg skala. Framsyning og kategoriar, til dømes kvar grensa går for kva skal definerast som eit særskilt artsrikt område ("hotspot"), kan enkelt tilpassast til behov.
Kartleggingsinnsats	Heildekkande kart som identifiserer områder med låg kartleggingsinnsats og eventuelt høg artsrikdom. Verdier er gjeve på kontinuerleg skala, som kan klassifiserast og forenklast etter behov.
Samanhengande områder	Heildekkande kart som syner samanhengande intakte områder med særskild høg artsrikdom.
Overlapp mellom områder med særskild høg artsrikdom og områder med relevans for forvaltning	Identifisering og kvantifisering av områder med høg artstriktheit, utbreiing av enkeltartar eller kartleggingsinnsats, innanfor områder av forvaltingsmessig interesse. Slike områder kan vera hovedøkosystem, kartlagte naturtypar, inngrepsfrie naturområder, nøkkelbiotopar frå MiS eller verneområder.
Endring over tid	Basert på arealendringar dokumenter i tidsseriar over arealdekkkart, kan ein utforska moglegheita for å estimere endring over tid i artsrikdom i områder med særskild betydning for artsrikdom. Produktet er i nok begrensa av overlapp i tid mellom kart og biologiske data (sjå seksjon 2.3 og 2.4).
Samla naturverdi	Er alle områder med høg artsrikdom like mykje verdt? Ein type avleia produkt kan peika på kor stor andel av naturmangfald som vil gå tapt i eit aktuelt forvaltingsområde (t.d. ein kommune) ved inngrep i eit mindre område. I nokre område vil artsrikdom ha ein meir jamn fordeling enn i andre og artsutvekslinga kan vera større. Slike samanhengar kan undersøkast med ein analyse av betadiversitet ("romleg artsomsetning").

## 5 Formidlingspakke

### 5.1 Føremålet med formidlingspakka

For å skildra og visa fram moglegheitene som ligg i produkta frå prosjektet lyt me ta i bruk dedikerte formidlingsverktøy. Ei slik formidling må både visa produkta, og på lettfatteleg vis forklåra kva produkta kan, og ikkje kan, brukast til. Dette inneber at sluttbrukaren sit att med ei forståing av kva dei ulike produkta fortel om artsrikdom, og ikkje minst kva usikkerheit som ligg i dei. Med desse måla i tankane har, er Shiny-app valgt som verktøy for formidling. Dette valet er ikkje minst tufta på den tette integreringa mellom Shiny-appar og programmeringsspråket i R. Formidlingsverktøyet slik det føreligg i dag har to hovudmål:

1. Gje sluttbrukarar ei kortfatta og intuitiv oversikt av prosjektet sine moglegheiter og målsettingar
2. Gje sluttbrukarar ei forståing om styrker og usikkerheitar i data og sluttprodukta, og kva konsekvensar dette har til dømes om produkta skal brukast i praktisk forvaltning

Shiny-appen gjev brukaren moglegheit til å utforska data og produkt på ein rask, intuitiv og fleksibel måte. Ulik informasjon om artsmangfald vert vist i ulike faner. Innan kvar fane kan ein navigera seg gjennom data og produkt på ein måte som viser att arbeidsflyten som ligg bak sluttproduktet. Det er også ei kort skildring produktet og korleis ein skal tolka det, og av dei data som produktet er basert på. Dette gjev ein rask og lettfatteleg framsyning av moglegheitane som ligg i prosjektet.

Målet med ei slik transparent og lettfatteleg presentasjon og forklåring av data, arbeidsflyt og sluttprodukt gjer at sluttbrukarane skal få ei klår oppfatning om kva delar av prosjektet som er relevante for dei, og korleis dei kan brukast. Når bakgrunnsdata vert vist fram saman med sluttprodukta, karleggingsinnsats og usikkerheita, gjev det også ein opnare prosess rundt dei avgjerdsler som vert tatt om norsk natur, fordi kunnskapsgrunnlaget er meir ope for alle interessentar. Viktigheita av tillit til myndigheiter og den kunnskapen dei baserar sine avgjerdsler på, har vorte svært tydeleg dei siste åra i ljøs av til dømes klimakrisa og COVID-19 pandemien.

Formidlingsverktøyet er laga primært for sluttbrukarar som er involverte i avgjerdsler om norsk naturmangfald. Utforminga av verktøyet er tenkt å vera intuitiv og gje ei lettfatteleg forståing av innhaldet, slik at brukarar ikkje treng kompetanse i statistisk modellering eller økologisk teori for å forstå og utnytta verktøyet.

Shiny-app vart valgt som teknologisk ramme for formidlingspakka i hovudsak fordi den støttar ei svært tett integrering med prosjektet sin arbeidsflyt i R. Koden i Shiny-appen er enkel å omsetja til meir komplekse programmeringspakker som gjev større moglegheitar for skreddarsydde løysingar når det gjeld design.

### 5.2 Produktskilddring

Formidlingspakka slik den framstår i dag er opent tilgjengeleg her:

<https://swp-data-projects.shinyapps.io/hotspotApp/> . Noverande versjon er produsert for

fire artsgrupper i Noreg: Tovengjer, amfibiar og reptilar, augestikkarar, og pattedyr. Informasjon om arbeidsflyt og modellkøyeringa finst på fana "Metadata". Alle data og produkt som per i dag er i figuren er testdata som ikkje skal brukast som kunnskapsgrunnlag i nokon samanheng.

## Innleiing

Startsida gjev ei kort oversikt over hovudmålet til prosjektet, ei kort skildring av strukturen på formidlingsverktøyet, og kvar ein kan finna meir informasjon om prosjektet og data som vert nytta. Startsida finst både på norsk og engelsk.

## Artsdata

Denne sida viser modellerte intensitetar av førekomst og innsamling (karleggingsinnsats) i Noreg for kvar truga art av den valde artsgruppa. Sida gjev også litt informasjon om artane sin rødlitestatus (Artsdatabanken 2021), og eit bilete av arten som er nedlasta frå iNaturalist, viss dette finst. Det er også mogleg å sjå registrerte observasjonar av arten i Noreg. Sidan uttrykket "intensitet" kan vera nytt for mange, gjev me også ei kort skildring av omgrepet slik at faren for feiltolkning vert mindre.

Velg art

**Taxa:**

mammals

**Art:**

Mustela putorius Linnaeus, 1758

Vis forekomster


Velg intensitetstype

**Intensitetstypee:**

Forekomstintensitet

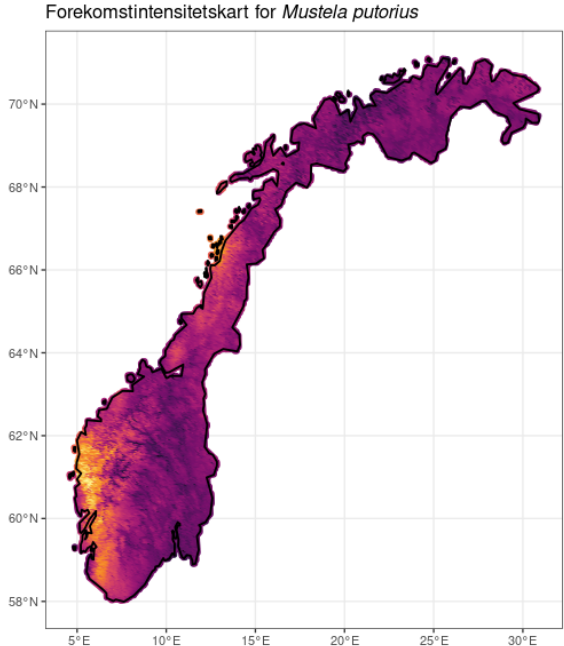
Artsinformasjon

**Vitenskapelig navn:** Mustela putorius  
**Tall observasjoner:** 53  
**Status på rødlista:** Sårbar  
**Bildekreditt:** [naturalist25051](#)



Artsintensitetskart

Forekomstintensitetskart for *Mustela putorius*



Skalert intensitet

1.00  
0.75  
0.50  
0.25  
0.00

Forekomstintensitetskart for *Mustela putorius*. Datakilde er Global Biodiversity Information Facility (DOI: 10.15468/dl.z8ugm7). Intensitet er modellert med bruk av Hotspots data arbeidsflyt: <https://github.com/gjearevoll/BioDivMapping>.

**Figur 10:** Skjermdump av Hot-spots-appen som viser intensitetskart for førekomst av *Mustela putorius* (Ilder).

### **Artsrikdom**

I denne fana vert det vist fram observert artsrikdom for kvart takson, utrekna frå førekomstdata frå GBIF og Artsdatabanken. Her vil det etter kvart også vera mogleg å sjå modellert artsrikdom for dei ulike takson.

### **Metadata**

Det ligg inne ei omfattande skildring av metadata, slik at sluttbrukaren kan sjå og validera data om det er naudsynt. Alle data som er brukt gjennom arbeidsflyten fram til sluttprodukta vert lista og skildra, slik at resultatata er i tråd med FAIR-prinsippa (Wilkinson et al. 2016).

## 6 Diskusjon

### 6.1. Oppsummering og vegen vidare for prosjektet

Sjølv om arbeidsflyten som er skildra i denne rapporten i praksis er operasjonell, står det att eit omfattande arbeid for å produsera sluttprodukta til prosjektet. Den neste fasen av prosjektet vil bruka arbeidsflyten som er skildra i rapporten til å produsera utbreiingskart for artar, kart over artsriksdom, kart over kartleggingsusikkerheit, og ulike avleia produkt og statistikkar som er skildra i kapittel 4. Stega som står att er:

- Kvalitetssikring og val av artsdata
- Kvalitetssikring og val av miljøvariablar
- Køyering av utbreiingsmodellar
- Generering av avleia produkt og framstilling av kart

Desse stega må køyrast iterativt der ei modellkøyering vert brukt for å evaluera eignaheita til datagrunnlaget og eventuelt gjera endringar i dette. I den samanheng vil det også verta gjort ei nærare vurdering av kva miljøvariablar som skal brukast for ulike artsgrupper.

I tillegg må formidlingskonseptet ferdigstillast og operasjonaliserast. Det skal også gjerast ei feltvalidering av prediksjonane for karplanter.

### 6.2. Vidareutvikling

Fleire moglege vidareutviklingar av arbeidsflyt og -omfang av modellering ligg ikkje inne i dei rammene som prosjektet har i dag. Dette kan vera nye tilnærmingar i form av algoritmar, prosedyrar og data, men også auka geografisk utstrekning og fleire typar økosystem (til dømes ferskvatn og marint).

Mange av dei store samfunnsutfordringane nasjonalt og internasjonalt krev betre kunnskap om natur og naturmangfald. Metodar og verktøy for modellering av artsbreiding og artsriksdom er difor i rask utvikling. Det er til dømes publisert nærare sju ulike R-pakker for slike typar analysar kvart år dei siste ti åra (Sillero et al. 2023). Det vert også publisert fleire hundre vitenskaplege artiklar kvart år som anten presenterer bruk eller metodeutvikling innan feltet. Ei tilsvarande rask utvikling finn me også innan andre relevante fagfelt, til dømes når det gjeld informasjonshandtering gjennom bruk av kunstig intelligens (Johaadien and Torma 2023). Vidareutvikling av metoder og enkeltelementer i arbeidsflyten vil difor vera ein kontinuerlig prosess.

Prinsippa om fri kildekode og uavhengighet til enkeltleverandørar er drøfta tidligare i denne rapporten. Prinsippet om at programvare som vert utvikla for offentlege midlar bør være fri kildekode, vil på sikt også gje store fordelar for forvaltinga si moglegheit til å velje leverandør og styre vidareutvikling av fagfelt, produkt og forvaltningsverktøy. Dette er grunnleggjande problemstillingar som er relevante på tvers av nasjonale grenser,. Ein bør difor i større grad vurdere internasjonalt samarbeid, etter ei vurdering om moglege konflikter mellom optimalisering av produkt på nasjonal skala og internasjonal utvikling. Dette gjeld i hovudsak kva for miljøvariablar som blir nytta. Naturmangfaldsdata som vert henta frå GBIF er allereie del av eit globalt informasjonsnettverk, og når GBIF blir nytta som datastrøm for artsdata (både "innsatsdata" og "förekomstdata") er internasjonal



kompatibilitet godt ivaretatt. Merk at dette gjeld primærprodukt slik som artsutbreiingar og artsrikdom. Deriverte produkt som til dømes overlapp mellom truga natur og ulike typer arealbruk kan relativt enkelt verta tilpassa lokale tilhøve.

Nye produkt eller bruk av produkta er også ei form for vidareutvikling. Dette kan være nye deriverte produkt basert på dei same grunnkarta over artsrikdom, det kan vera nye utvikla modular i arbeidsflyten, eller det kan være andre måter å bruke rammeverket på. I tabell 4 gjev me ein oversikt over nokre døme på mogleg vidareutvikling.

**Tabell 4:** *Moglege vidareutviklingar av forvaltningsrelevante verktøy basert på leveransane i dette prosjektet*

Navn	Skildring
Ferskvatn	<p>Rammeverket som er skildra i denne rapporten kan med små grep overførast til ferskvatnsystem. Det vil krevja i) ein del andre miljøvariablar og ii) noko annleis avgrensing / maskering av inngangsdata om artar. Dette er mogleg å løysa relativt raskt. Ei større utfordring er tilgangen på maskinlesbare innsatsdata (artsdata med funn / ikkje-funn eller kvantitative data). Mykje, men langt frå alt, av data som er samla i ferskvatn vert gjort tilgjengeleg via vannmiljøsystemet. Det er ikkje heilt rett fram korleis data på innsats blir handtert i denne dataflyten. I tillegg er det ofte doble datastrømmer der noko, men langt frå alt, av data som går inn i Vannmiljø vert publisert ut på Artskart / GBIF først og så publisert vidare frå Vannmiljø ut på Artkart / GBIF på ny. Manglande krav til å bestemma prøver til artsnivå for mykje av overvåkningsmetodikken som vert nytta i akvatiske miljø vil også gje moglege problem for enkelte grupper. Det siste kan til ein viss grad verta løyst med vidareutvikling av modellgrunnlag der ein nyttar modellar som utnyttar informasjon i data som skildrar funn på høgare taksonomisk nivå enn art.</p>
Hav og kyst	<p>Mykje av dei same problema som er skildra for ferskvatn gjeld også for hav og kyst. Ei foreløpig vurdering er at det truleg er dårlegare tilgang til maskinlesbare data på artsmangfald frå marine miljø, anten fordi data ikkje er gjort tilgjengelege i ein maskinlesbar form og / eller fordi data ikkje er gjort tilgjengeleg for gjenbruk for uavhengige fagmiljø. Til dømes er data frå einskilde store program som sjøfugelovervakninga ikkje tilgjengeleg for allmenn bruk (<a href="https://www2.nina.no/seapop/seapophtml/">https://www2.nina.no/seapop/seapophtml/</a>). Produkt, slik som oversikt over kartleggingsinnsats, vil gje ein god oversikt over data- og kunnskapsmanglar i det marine økosystemet. Ei anna utfordring med marine miljø er at nokre habitat er tredimensjonalt i struktur og svært dynamiske, og difor kan krevja andre tilnærmingar. Alternativt kan ein fokusere på einskilde habitat, til dømes bentiske miljø.</p>

Digital tvilling av Norsk Natur	<p>Ein Digital tvilling av norsk natur kan sjåast på som einn kontinuerleg oppdatert oversikt over det biologiske mangfaldet i Noreg. Arbeidsflyt som er satt opp i dag vil relativt enkelt kunne vidareutviklast til dette. Dei største utfordring i dag er:</p> <ul style="list-style-type: none"> <li>• Forbetra algoritmar for raskare oppdatering av utbreiingsmodellar</li> <li>• Stabilitet til kjeldekode</li> <li>• Prosedyrar for oppdatering og tilgjengeleggjing av data.</li> </ul> <p>Det fyrste bør være mogleg å løyse da dette vert gjort innanfor andre sektorar i dag. Stabilitet i kjeldekode er eit større problem når utviklinga skjer på eit begrensa budsjett. Ei løysing er å utvikla digital tvilling som eit ope kjeldekodeprojekt, og då helst med eit internasjonalt perspektiv / deltaking. Prosedyrar for oppdatering og tilgjengeleggjing gjeld i hovudsak artsdata. Det er i mindre grad eit teknisk problem, men krev at data vert levert i maskinlesbar form ved rapportering av prosjektresultat. Det vil også krevja at det vert etablert rutinar, enten manuelle eller semi-manuelle ved bruk av KI, som kan kontrollera, kvalitetssikra og standardisera nytt datatilfang.</p>
Naturbudsjett	<p>Metodane og produkta som er skildra i denne rapporten legg grunnlaget for eit funksjonelt naturbudsjett. I dag eksisterer det ikkje ei heildekkande kopling mellom økosystem- og naturtypekartlegging og artsmangfald. Etter vår meining vil ikkje eit naturbudsjett som baserer seg på kartlegging av økosystem slik det er skildra i til dømes anon. (2023) fanga opp konsekvensar av arealbruk på artsmangfald, inkludert trua natur, på ein god måte. Bruk av reiskap som er skildra i denne rapporten kan gje slik innsikt. På sikt kan dette skje saman med kontinuerleg oppdaterte produkt i form av ein digital tvilling for natur.</p>
Kvantitative predikasjonar	<p>Det noverande modellrammeverket handterer ikkje kvantitative predikasjonar på ein god måte. Det finst rammeverk som kan gjera dette (t.d. Pagel m.fl. 2014) og dette kan implementerast i den eksisterande arbeidsflyten. Det vil krevja at ein har tilgjengeleg kvantitative data for dei artane ein ynskjer modellert, i tillegg til førekomstdata og innsatsdata med funn / ikkje-funn. Vår vurdering er at dette vil krevje ein del utviklingstid av eksisterande modellrammeverk, men, dersom dette vert implementert opnar det for store moglegheiter i form av relevante reiskap for forvaltninga til dømes for å rekna sensitiviteten til artssamfunn for endringar i miljøet (Sæther m.fl. 2023).</p>
Romlege-temporære modellar	<p>Dagens implementering av tid inneber i hovudsak ei oppstyking av predikasjonane, kor ein endrar miljøvariablane og inngangsdata etter tidsperiodar. Det er ei kategorisering og ikkje det ein eigentleg forstår med romlege-temporære modellar. Implementering av dette er etter vår vurdering høgst mogleg teknisk og vil ha ei medium tidshorisont på utvikling. Det er derimot ikkje sikkert det per i dag eksisterer artsdata som støtter dette. Det er truleg eit mindre problem med miljødata, då desse vert samla inn meir kontinuerleg i tid og rom, til dømes med fjernmåling.</p>

Tilfanget av artsdata aukar stadig ved at nye data vert lagt til, eller at tidlegare innsamla data vert gjort tilgjengeleg. Når det gjeld nye innsamlingar, vil informasjon om kartleggingsinnsats frå dette prosjektet gjera det mogleg å peika på områder som bør prioriterast. Ein interaktiv prosess der sluttprodukt vert oppdatert etter kvart som nye data kjem til i Artskart / GBIF, og det kan gjerast synleg korleis presisjonen til modellen endrar seg, vil vera ei svært nyttig vidareutvikling av formidlingspakka, men då må køyretida til modellane vert monaleg redusert. Dette ligg truleg litt fram i tid.

Arbeidsflyten som er presentert i denne rapporten er tilrettelagt for at ei slik kontinuerleg køyring med oppdaterte artsdata eller miljødata skal vera mogleg i framtida. Dette vil gjera at data, modellane og sluttprodukta samla vil utgjera ein digital tvilling av norsk natur. Ein slik oppdaterbar og dynamisk digital tvilling av natur har ei rekkje bruksmåtar åleine eller tolka saman med andre typar data. Produkta frå ein operasjonell digital tvilling vil til dømes vera heilt grunnleggjande for eit fagleg godt og transparent naturrekneskap, og vil vera heilt uvurderleg når til dømes effektar av endringar av arealbruk skal utreia. Eit naturrekneskap krev heildekkande kart som gjer det mogleg å talfesta status og endringar i naturverdi over tid. Det ligg i premissane for interoperabilitet med internasjonale system at ein slik rekneskap skal basere seg på førekomst av ulike økosystemtypar. Dette krev gode indikatorar som gjer det mogleg å vurdere tilstanden til førekomsten av økosystem, og samanlikna slike tilstandsvurderingar mellom områder med same økosystem, eller mellom ulike økosystem.

Det gjev lita økologisk meining å laga indikatorar som seier noko om naturverdi av eit einskild avgrensa areal utan å ta omsyn til artsmangfald og tilstand både i det avgrensa økosystemet og i omliggande økosystem. Eit økosystem sin funksjon vert i stor grad bestemt av eigenskapar til dei ulike artane i systemet, samspelet mellom desse artane, og i kva grad og korleis artane er avhengig av andre økosystem (Cardinale et al. 2012; Kinzig, Pacala, and Tilman 2001; Millennium Ecosystem Assessment 2023). Kva artar, og kor mange av dei som er til stades vil truleg variera innan økosystem avhengig av geografi. Dette er eit resultat både av miljøfaktorar, som til dømes klima, men også innvandringshistorikk. Det også vera heilt avgjerande for funksjon og sårbarheit til eit gitt avgrensa økosystem kva for omliggande økosystem som finst og kva artsmangfald desse har.

Indikatorar må baserast på underliggjande økologisk funksjon til dei ulike artane (Sæther 2021). Moglege tilnærmingar kan vera å nytta eigenskapar til mengde-art fordelinga til artene innan eit område (Sæther et al. 2022). Dette kan brukast blant anna til å kvantifisera kor robust artssamfunnet i eit område er for endringar i miljøet, slik som tilgangen på viktige habitattypar (Sæther m.fl. 2023). Modellrammeverket som er presentert i denne rapporten kan utvidast til å produsera også slik kunnskap, gitt at det finst artsdata med tilstrekkeleg informasjon om kor mykje det finst av ulike artar innan eit område.

### 6.3. Datatilgang

Produkta som vert skildra i denne rapporten vil ofte verta avgrensa av tilgang til relevante datakjelder. Dette gjeld også i høg grad vidareutvikling som til dømes ei utviding av heildekkande utbreiingskart til også å dekkje ferskvatn, hav og kyst, som skildra i tabell 4. I første rekke er dette ei utfordring for artsdata, og da særskilt innsatsdata av ulike slag. Eit av dei primære produkta frå modelleringa som er skildra i denne rapporten vil være

oversikt over kartleggingsinnsats. Dette kan da nyttast til å planlegge aktivitet for ny overvaking eller kartlegging. Det kan i den samanhengen være nyttig å peike på årsakene til datamangel.

Vi kan dele mangel på data inn i to hovudgrupper:

- Manglande innsamling og observasjonar; slik som at informasjon aldri er innhenta for relevante artar, områder eller tidsperiodar
- Manglande tilgjengeleggjing

Manglande kartlegging og overvaking er i hovudsak eit spørsmål om ressursar og prioriteringar. Vi trur produkta frå dette prosjektet, særskilt da ei oversikt over kartleggingsinnsats, vil verta viktige bidrag til å få ei best mogleg utnytting av dei avgrensa ressursane miljøforvaltninga har til kartlegging og overvaking. Det kan likevel vera utfordrande å skilje mellom manglande data og manglande tilgjengeleggjing. Dette vil krevja eit eige og omfattande arbeid med en “data-revisjon” av naturmangfaldsdata i Noreg, noko som ikkje er mogleg innafor dette prosjektet.

Manglande tilgjengelegheit på data òg eit spørsmål om bestillerkompetanse hjå myndigheiter som pålegg eller bestiller ei undersøking eller kartlegging. Svært ofte vert data som skildrar naturmangfald og som er samla inn i samband med kartlegging, overvaking eller konsekvensutredningar, ikkje offentleg tilgjengeleg etter at undersøkinga er ferdig. For artsmangfaldsdata, både “innsatsdata” og “førekomstdata”, eksisterer det standardar og system for å få desse tilgjengelege etter anerkjente prinsipp gjennom Artsdatabanken og GBIF. Med prinsipp tenker vi her i første rekke på FAIR prinsippa (Wilkinson m.fl. 2016). Dette står for “Findable, Accessible, Interoperable og Re-usable”, noko som på norsk ofte blir omsett til “mogleg å finne, tilgjengeleg, samhandlande og gjenbrukbar”.

No har tilgjengeleggjing av data rett nok nokre kostnader. For nye data vil det i starten vera litt meirarbeid, spesielt i ein implementeringsfase før personell vert kjent med system og standardar. Det vil også vera indirekte kostnader for både enkeltpersonar og institusjonar då eksklusiv tilgang til data kan omsetjast både i nye prosjekt og i vitskapleg status (medforfattarskap på artklar som nyttar data). Tilgjengeleggjing av gamle data kan ha større kostnader. Dersom data eksisterer i det heile tatt, kan mangel på dokumentasjon, manglande standardisering eller digitalisering gjere det vanskeleg å få desse opp til FAIR standard. Då det i dag ikkje eksisterar vesentlege positive initiativ for å gjera data tilgjengeleg, vil mangel på krav frå oppdragsgjevar difor i stor grad føre til at data om naturmangfald ikkje blir gjort offentleg tilgjengeleg.

## Litteratur

Alston, Jesse M., and Jessica A. Rick. (2021). “A Beginner’s Guide to Conducting Reproducible Research.” *Bulletin of the Ecological Society of America* 102 (2). <https://doi.org/10.1002/bes2.1801>.

Anon. (2023). Etablering av naturregnskap i Norge Eksisterende data og utviklingsbehov i møte med internasjonale standarder og krav. Miljødirektoratet rapport. M-2599|2023  
Araújo, Miguel B., Robert P. Anderson, A. Márcia Barbosa, Colin M. Beale, Carsten F. Dormann, Regan Early, Raquel A. Garcia, et al. (2019). “Standards for Distribution

- Models in Biodiversity Assessments.” *Science Advances* 5 (1): eaat4858.  
<https://doi.org/10.1126/sciadv.aat4858>
- Artsdatabanken (2021). “Norsk rødliste for arter 2021.”  
<http://www.artsdatabanken.no/lister/rodlisteforarter/2021>
- Artsdatabanken. (2023). “Fremmede arter i Norge - med økologisk risiko 2023.”  
<http://www.artsdatabanken.no/lister/fremmedartslista/2023>
- Astrup, R., Rahlf, J., Bjørkelo, K., Debella-Gilo, M., Gjertsen, A.-K., Breidenbach, J.,  
 2019. Forest information at multiple scales: development, evaluation and application  
 of the Norwegian forest resources map SR16. *Scandinavian Journal of Forest  
 Research* 34, 484–496. <https://doi.org/10.1139/cjfr-2020-0170>
- Baker, David J., Ilya M. D. Maclean, Martin Goodall, and Kevin J. Gaston. (2021).  
 “Species Distribution Modelling Is Needed to Support Ecological Impact  
 Assessments.” *The Journal of Applied Ecology* 58 (1): 21–26.  
<https://doi.org/10.1111/1365-2664.13782>
- Blicharska, Malgorzata, Richard J. Smithers, Grzegorz Mikusiński, Patrik Rönnbäck,  
 Paula A. Harrison, Måns Nilsson, and William J. Sutherland. (2019). “Biodiversity’s  
 Contributions to Sustainable Development.” *Nature Sustainability* 2 (12): 1083–93.  
<https://doi.org/10.1038/s41893-019-0417-9>
- Brown, C.F., Brumby, S.P., Guzder-Williams, B. et al. Dynamic World, Near real-time  
 global 10 m land use land cover mapping. *Sci Data* 9, 251 (2022).  
<https://doi.org/10.1038/s41597-022-01307-4>
- Bryn, A., Potthoff, K. Elevational treeline and forest line dynamics in Norwegian mountain  
 areas – a review. *Landscape Ecol* 33, 1225–1245 (2018).  
<https://doi.org/10.1007/s10980-018-0670-8>
- Cardinale, Bradley J., J. Emmett Duffy, Andrew Gonzalez, David U. Hooper, Charles  
 Perrings, Patrick Venail, Anita Narwani, et al. (2012). “Biodiversity Loss and Its  
 Impact on Humanity.” *Nature* 486 (7401): 59–67. <https://doi.org/10.1038/nature11148>
- Chamberlain, Scott, Damiano Oldoni, Laurens Geffert, Peter Desmet, Vijay Barve,  
 Karthik Ram, Matt Blissett, et al. (2022). *Ropensci/rgbif: Rgbif v3.7.0*. Zenodo.  
<https://doi.org/10.5281/ZENODO.6023735>.
- Elven, Hallvard and Geir Sjøli (2021) “Kunnskapsstatus for arts mangfoldert i Norge 2020”.  
 Naturhistorisk Museum, Universitet i Oslo.
- Feng, Xiao, Daniel S. Park, Cassondra Walker, A. Townsend Peterson, Cory Merow, and  
 Monica Papeş. (2019). “A Checklist for Maximizing Reproducibility of Ecological  
 Niche Models.” *Nature Ecology & Evolution* 3 (10): 1382–95.  
<https://doi.org/10.1038/s41559-019-0972-5>
- Feldmeier, S., Schefczyk, L., Hochkirch, A., Lötters, S., Pfeifer, M. A., Heinemann, G., &  
 Veith, M. (2018). Climate versus weather extremes: Temporal predictor resolution  
 matters for future rather than current regional species distribution models. *Diversity &  
 Distributions*, 24(8), 1047–1060. <https://doi.org/10.1111/ddi.12746>
- Fick, S.E. and R.J. Hijmans, 2017. WorldClim 2: new 1km spatial resolution climate  
 surfaces for global land areas. *International Journal of Climatology* 37 (12):  
 4302-4315.
- GBIF.org. (2023a). “Citation Guidelines.” <https://www.gbif.org/citation-guidelines>.
- GBIF.org. (2023b). GBIF Occurrence Download <https://doi.org/10.15468/dl.ebuzfh>.  
 Accessed November 30, 2023-
- Gjerde, Ivar, Magne Sætersdal, and Hans H. Blom. (2007). “Complementary Hotspot  
 Inventory – A Method for Identification of Important Areas for Biodiversity at the  
 Forest Stand Level.” *Biological Conservation* 137 (4): 549–57.  
<https://doi.org/10.1016/j.biocon.2007.03.007>

- Isaac, Nick J. B., Marta A. Jarzyna, Petr Keil, Lea I. Dambly, Philipp H. Boersch-Supan, Ella Browning, Stephen N. Freeman, et al. (2020). "Data Integration for Large-Scale Models of Species Distributions." *Trends in Ecology & Evolution* 35 (1): 56–67. <https://doi.org/10.1016/j.tree.2019.08.006>
- Herfindal, I., Aanes, S., Benestad, R. E., Finstad, A. G., Salthaug, A., Stenseth, N. C., & Sæther, B.-E. (2022). Spatiotemporal variation in climatic conditions across ecosystems. *Climate Research*, 86, 9-19. <https://doi.org/10.3354/cr01641>
- Johaadien, Rukaya, and Michal Torma. (2023). "Publish First": A Rapid, GPT-4 Based Digitisation System for Small Institutes with Minimal Resources." *Biodiversity Information Science and Standards* 7 (November): e112428. <https://doi.org/10.3897/biss.7.112428>
- Karger, D.N., Conrad, O., Böhner, J., Kawohl, T., Kreft, H., Soria-Auza, R.W., Zimmermann, N.E., Linder, P., Kessler, M. (2017): Climatologies at high resolution for the Earth land surface areas. *Scientific Data*. 4 170122. <https://doi.org/10.1038/sdata.2017.122>
- Kinzig, Ann P., Stephen Pacala, and David Tilman. (2001). *The Functional Consequences of Biodiversity: Empirical Progress and Theoretical Extensions*. Princeton University Press.
- Kuya, E. K., Gjelten, H. M., and Tveito, O. E.: Homogenization of Norwegian monthly precipitation series for the period 1961–2018, *Adv. Sci. Res.*, 19, 73–80, <https://doi.org/10.5194/asr-19-73-2022>
- Körner, Christian, Patrick Möhl, & Erika Hiltbrunner. (2023). "Four ways to define the growing season". *Ecology Letters*, 26(8), 1277–1292. <https://doi.org/10.1111/ele.14260>
- Kålås J A, Øien I J, Stokke B, Vang R (2023). Norwegian breeding bird monitoring scheme. Version 1.8. Norwegian Institute for Nature Research. Sampling event dataset <https://doi.org/10.15468/6jmw2e> accessed via GBIF.org on 2024-01-04.
- Macleán, Kyle, Isaac Triguero. (2023). Identifying bird species by their calls in Soundscapes. *Appl Intell* 53, 21485–21499. <https://doi.org/10.1007/s10489-023-04486-8>
- Mandeville, Caitlin P., Erlend B. Nilsen, and Anders G. Finstad. (2022). "Spatial Distribution of Biodiversity Citizen Science in a Natural Area Depends on Area Accessibility and Differs from Other Recreational Area Use." *Ecological Solutions and Evidence* 3 (4). <https://doi.org/10.1002/2688-8319.12185>.
- Marchese, Christian. (2015). "Biodiversity Hotspots: A Shortcut for a More Complicated Concept." *Global Ecology and Conservation* 3 (January): 297–309. <https://doi.org/10.1016/j.gecco.2014.12.008>
- Meng, Xiao-Li. (2018). "Statistical Paradises and Paradoxes in Big Data (I): Law of Large Populations, Big Data Paradox, and the 2016 US Presidential Election." *The Annals of Applied Statistics* 12 (2): 685–726. <https://doi.org/10.1214/18-AOAS1161SF>
- Millennium Ecosystem Assessment. (2023). "Ecosystems and Human Well-Being: Current State and Trends." Island Press. Accessed December 6, 2023.
- Moen, Asbjørn. (1999). Norges nasjonalatlas: Vegetasjon. Statens kartverk, Hønefoss.
- Mostert, Philip S., and Robert B. O'Hara. (2023). "PointedSDMs: An R Package to Help Facilitate the Construction of Integrated Species Distribution Models." *Methods in Ecology and Evolution / British Ecological Society* 14 (5): 1200–1207. <https://doi.org/10.1111/2041-210X.14091>
- Myers, Norman. (1988). "Threatened Biotas: 'Hot Spots' in Tropical Forests." *The Environmentalist* 8 (3): 187–208. <https://doi.org/10.1007/BF02240252>
- Myers, Norman, Russell A. Mittermeier, Cristina G. Mittermeier, Gustavo A. B. da



- Fonseca, and Jennifer Kent. 2000. "Biodiversity Hotspots for Conservation Priorities." *Nature* 403 (6772): 853–58. <https://doi.org/10.1038/35002501>
- Nilsen Erlend B, Vang Roald, Breisjøberget Jo Inge (2022a). Tetraonid line transect surveys from Norway: Data from Statskog. Version 1.8. Norwegian Institute for Nature Research. Sampling event dataset <https://doi.org/10.15468/q2ehlq> accessed via GBIF.org on 2024-01-04.
- Nilsen Erlend B, Vang Roald, Kjøsberg, Marius, Kvasnes, Mikkel Andreas Jønsøn (2022b). Tetraonid line transect surveys from Norway: Data from Fjellstyrene. Version 1.7. Norwegian Institute for Nature Research. Sampling event dataset <https://doi.org/10.15468/975ski> accessed via GBIF.org on 2024-01-04.
- Nilsen Erlend B, Vang Roald, Asbjørnsen, Einar (2022c). Tetraonid line transect surveys from Norway: Data from Finnmarkseiendommen (FeFo). Version 1.12. Norwegian Institute for Nature Research. Sampling event dataset <https://doi.org/10.15468/s7c8qd> accessed via GBIF.org on 2024-01-04.
- Norberg, Anna, Nerea Abrego, F. Guillaume Blanchet, Frederick R. Adler, Barbara J. Anderson, Jani Anttila, Miguel B. Araújo, et al. 2019. "A Comprehensive Evaluation of Predictive Performance of 33 Species Distribution Models at Species and Community Levels." *Ecological Monographs* 89 (3): e01370. <https://doi.org/10.1002/ecm.1370>
- NOU 2023: 3 (2023). Mer av alt – raskere — Energikommisjonens rapport. Olje- og energidepartementet. <https://www.regjeringen.no/no/dokumenter/nou-2023-3/id2961311/>
- Nybø, S., Kolstad, A., Sandvik, H., Bakkestuen, V., Evju, M., Framstad, E., Grainger, M., Jepsen, J., Nater, C., Speed, J.D.M., Töpfer, J., Venter, Z., Åström, J., Åström, S., Bargmann, T., Silvennoinen, H.M. & Sydenham, M.A.K. 2023. Indikatorer for økologisk tilstand i våtmark, semi-naturlig mark og naturlig åpne områder under skoggrensa. NINA Rapport 2336. Norsk institutt for naturforskning. <https://hdl.handle.net/11250/3096724>
- Olsen, Siri Lie, Richard D. Hedger, Ditte Hendrichsen, Benno Dillinger, Alexander Venter & Marianne Evju. (2020). "Geografisk utbredelse av truede insekter og edderkoppdyr, sopp, lav og moser: modellering av hotspots" *Norsk institutt for naturforskning*.
- Olsen, Siri Lie, Richard D. Hedger, Megan Nowell, Ditte Hendrichsen & Marianne Evju. (2018). "Geografisk utbredelse av truede karplanter i Norge: modellering av hotspots" *Norsk institutt for naturforskning*.
- Pagel, J., Anderson, B.J., O'Hara, R.B., Cramer, W., Fox, R., Jeltsch, F., Roy, D.B., Thomas, C.D. and Schurr, F.M. (2014), Quantifying range-wide variation in population trends from local abundance surveys and widespread opportunistic occurrence records. *Methods Ecol Evol*, 5: 751-760. <https://doi.org/10.1111/2041-210X.12221>
- R Core Team (2023). R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna, Austria. <<https://www.R-project.org/>>.
- Reid, Walter V. (1998). "Biodiversity Hotspots." *Trends in Ecology & Evolution* 13 (7): 275–80. [https://doi.org/10.1016/S0169-5347\(98\)01363-9](https://doi.org/10.1016/S0169-5347(98)01363-9)
- Ruppert, Krista M., Richard J. Kline, Saydur Rahman. (2019). "Past, present, and future perspectives of environmental DNA (eDNA) metabarcoding: A systematic review in methods, monitoring, and applications of global eDNA". *Global Ecology and Conservation*, 17, E00547, <https://doi.org/10.1016/j.gecco.2019.e00547>.
- Santesson, Rolf. 1993. *The Lichens and Lichenicolous Fungi of Sweden and Norway*.
- Sillero, Neftalí, João Carlos Campos, Salvador Arenas-Castro, and A. Márcia Barbosa. (2023) "A Curated List of R Packages for Ecological Niche Modelling." *Ecological Modelling* 476 (February): 110242. <https://doi.org/10.1016/j.ecolmodel.2022.110242>
- Simmonds, Emily G., Susan G. Jarvis, Peter A. Henrys, Nick J. B. Isaac, & Robert B.



- O'Hara. 2020. "Is More Data Always Better? A Simulation Study of Benefits and Limitations of Integrated Distribution Models." *Ecography* 43 (10): 1413–22.  
<https://doi.org/10.1111/ecog.05146>
- Stewart, S.B., Elith, J., Fedrigo, M., Kasel, S., Roxburgh, S.H., Bennett, L.T., Chick, M., Fairman, T., Leonard, S., Kohout, M., Cripps, J.K., Durkin, L. and Nitschke, C.R. (2021), Climate extreme variables generated using monthly time-series data improve predicted distributions of plant species. *Ecography*, 44: 626-639.  
<https://doi.org/10.1111/ecog.05253>
- Sæter, Bernt-Erik. (2021) "Prediksjonsmodellering av arter og samfunn i forvaltning av biologisk mangfold: utfordringer og muligheter". M-1864|2020. CBD rapport 1/2023, Norge Teknisk Naturvitenskaplige Universitet
- Sæter, Bernt-Erik, Ivar Herfindal, Erik Blystad Solbu, Anna Nordberg, Steinar Engen. (2022) Prediksjonsmodellering av samfunn i et endre miljø: praktiske eksempler. M-2305|2022. CBD rapport 1/2022, Norges Teknisk Naturvitenskapelige Universitet.
- Sæther, Bernt-Erik, Steinar Engen, & Erik Blystad Solbu. (2023). Assessing the sensitivity and resistance of communities in a changing environment. *Journal of Animal Ecology*, 00, 1–13. <https://doi.org/10.1111/1365-2656.14003>
- Tenge, IM (2016) Arealressurskart AR5, NIBIO, 7 s.
- Tingstad, Lise, Marianne Evju, Hanne Sickel, and Joachim Töpper. (2019). "Utvikling av nasjonal arealrepresentativ naturovervåking (ANO). Forslag til gjennomføring, protokoller og kostnadsvurderinger med utgangspunkt i erfaringer fra uttesting i Trøndelag". NINA rapport. Norsk Institutt for Naturforskning (NINA).
- Tuia, Devis, Benjamin Kellenberger, Sara Beery, Blair R. Costelloe, Silvia Zuffi, Benjamin Risse, Alexander Mathis, et al. (2022) "Perspectives in machine learning for wildlife conservation." *Nature Communications* 13, 792.  
<https://doi.org/10.1038/s41467-022-27980-y>
- van de Pol M, Jenouvrier S, Cornelissen JHC, Visser ME. Behavioural, ecological and evolutionary responses to extreme climatic events: challenges and directions. *Philos Trans R Soc Lond B Biol Sci.* 2017 Jun 19;372(1723):20160134. doi: 10.1098/rstb.2016.0134. PMID: 28483865; PMCID: PMC5434086.
- Venter, Z.S.; Sydenham, M.A.K. Continental-Scale Land Cover Mapping at 10 m Resolution Over Europe (ELC10). *Remote Sens.* 2021, 13, 2301.  
<https://doi.org/10.3390/rs13122301>
- Wilkinson, Mark D., Michel Dumontier, I. Jsbrand Jan Aalbersberg, Gabrielle Appleton, Myles Axton, Arie Baak, Niklas Blomberg, et al. (2016). "The FAIR Guiding Principles for Scientific Data Management and Stewardship." *Scientific Data* 3 (March): 160018.  
<https://doi.org/10.1038/sdata.2016.18>
- Ørka, Hans Ole, Marie-Claude Jutras-Perreault, Janis Gailis, Arvid Svensson, Endre Hanse, Terje Gobakken (2019) "Fjernmålingsbasert Kartlegging Og Overvåkning Av økosystemet Skog–på Veg Mot et Nasjonalt Heldekkende Skogøkologisk Grunnkart." *Project Report. Miljødirketoratet.*

