

797

Statistiske egenskaper til Naturindeks for Norge.

Usikkerhet i datagrunnlaget og sensitivitet

Bård Pedersen
Olav Skarpaas

NINA Rapport



NINAs publikasjoner

NINA Rapport

Dette er en elektronisk serie fra 2005 som erstatter de tidligere seriene NINA Fagrapport, NINA Oppdragsmelding og NINA Project Report. Normalt er dette NINAs rapportering til oppdragsgiver etter gjennomført forsknings-, overvåkings- eller utredningsarbeid. I tillegg vil serien favne mye av instituttets øvrige rapportering, for eksempel fra seminarer og konferanser, resultater av eget forsknings- og utredningsarbeid og litteraturstudier. NINA Rapport kan også utgis på annet språk når det er hensiktsmessig.

NINA Temahefte

Som navnet angir behandler temaheftene spesielle emner. Heftene utarbeides etter behov og serien favner svært vidt; fra systematiske bestemmelsesnøkler til informasjon om viktige problemstillinger i samfunnet. NINA Temahefte gis vanligvis en populærvitenskapelig form med mer vekt på illustrasjoner enn NINA Rapport.

NINA Fakta

Faktaarkene har som mål å gjøre NINAs forskningsresultater raskt og enkelt tilgjengelig for et større publikum. De sendes til presse, ideelle organisasjoner, naturforvaltningen på ulike nivå, politikere og andre spesielt interesserte. Faktaarkene gir en kort framstilling av noen av våre viktigste forskningstema.

Annen publisering

I tillegg til rapporteringen i NINAs egne serier publiserer instituttets ansatte en stor del av sine vitenskapelige resultater i internasjonale journaler, populærfaglige bøker og tidsskrifter.

Statistiske egenskaper til Naturindeks for Norge.

Usikkerhet i datagrunnlaget og sensitivitet

Bård Pedersen
Olav Skarpaas

Pedersen, B. og Skarpaas, O. 2012. Statistiske egenskaper til Naturindeks for Norge. Usikkerhet i datagrunnlaget og sensitivitet. – NINA Rapport 797. 55 s.

Trondheim, februar 2012

ISSN: 1504-3312

ISBN: 978-82-426-2392-8

RETTIGHETSHAVER

© Norsk institutt for naturforskning

Publikasjonen kan siteres fritt med kildeangivelse

TILGJENGELIGHET

Åpen

PUBLISERINGSTYPE

Digitalt dokument (pdf)

KVALITETSSIKRET AV

Steinar Engen

ANSVARLIG SIGNATUR

Ass. Forskningssjef Signe Nybø (sign.)

OPPDRAKSGIVER(E)

Direktoratet for naturforvaltning

KONTAKTPERSON(ER) HOS OPPDRAGSGIVER

Knut Simensen

FORSIDEBILDE

Jaktfalk.

Foto: © Jan Ove Gjershaug

NØKKEWORD

Norge, Naturindeks, biologisk mangfold, målefeil, sensitivitet

KEY WORDS

Norway, Nature Index, biodiversity, measurement error, sensitivity

KONTAKTOPPLYSNINGER

NINA hovedkontor

Postboks 5685 Sluppen
7485 Trondheim
Telefon: 73 80 14 00
Telefaks: 73 80 14 01

NINA Oslo

Gaustadalléen 21
0349 Oslo
Telefon: 73 80 14 00
Telefaks: 22 60 04 24

NINA Tromsø

Framsenteret
9296 Tromsø
Telefon: 77 75 04 00
Telefaks: 77 75 04 01

NINA Lillehammer

Fakkeltgården
2624 Lillehammer
Telefon: 73 80 14 00
Telefaks: 61 22 22 15

www.nina.no

Sammendrag

Pedersen, B. og Skarpaas, O. 2012. Statistiske egenskaper til Naturindeks for Norge. Usikkerhet i datagrunnlaget og sensitivitet. – NINA Rapport 797. 55 s.

Naturindeks for Norge er et sammensatt mål for biologisk mangfold som beregnes for et geografisk område og et gitt tidspunkt ut fra et antall biodiversitetsindikatorer. Hver indikator representerer et aspekt av biologisk mangfold. Rent matematisk defineres naturindeksen som et veid gjennomsnitt av skalerte indikatorer. Skalering av indikatorene skjer ved hjelp av tre ulike skaleringsfunksjoner, LOW, MAX og OPT. Funksjonene inneholder en parameter kalt referanseverdien. Alle observasjoner eller målinger av en indikator angis med usikkerhet. Ved beregning av naturindeksen behandles denne usikkerheten ved å tilordne en sannsynlighetsfordeling (her kalt samplingfordeling) til hver enkelt observasjon av en indikator. Ut fra dette beregnes en tilsvarende fordeling for naturindeksen. Medianen i denne fordelingen benyttes som et estimat for naturindeksens verdi, og et 95 % konfidensintervall angir usikkerheten i dette estimatet.

Denne rapporten undersøker og forklarer hvordan usikkerhet i datagrunnlaget påvirker estimatene av naturindeksens verdi. Her kalles en slik effekt for lokasjonsforskyvning. Videre undersøkes hvordan eventuell lokasjonsforskyvning påvirker indeksens sensitivitet overfor endringer i biodiversitetsindikatorene. Det undersøkes også i hvilken grad naturindeksen er robust i forhold til feil i referanseverdiene. I rapporten studeres enkle, konstruerte eksempler med identisk fordelte indikatorobservasjoner, hvor en er i stand til å beregne sannsynlighetsfordelingen til naturindeksen numerisk. Det foretas også en matematisk analyse av lokasjonsforskyvning.

Analysene viser at størrelsen på lokasjonsforskyvning avhenger av hvilken skaleringsfunksjon som benyttes, indikatorenes forventningsverdier før skalering, spredningen i deres fordelinger og antall indikatorer som indeksen beregnes fra. Forskyvningen kan være både positiv og negativ. Lokasjonsforskyvning øker svakt i absoluttverdi med antall indikatorer, men nærmer seg raskt en grense. Absoluttverdien til forskyvningen øker med økende usikkerhet og er større jo nærmere de oppgitte verdiene for indikatorene er referanseverdiene. Skaleringsfunksjonene LOW og MAX gir generelt mindre forskyvning enn OPT. Vi viser at økende usikkerhet i dataene både kan øke og redusere følsomheten til estimatet av naturindeksen overfor endringer i biodiversitetsindikatorene. Dette avhenger av om endringene i indikatorene er assosiert med økninger eller reduksjoner i lokasjonsforskyvning. Naturindeksen ser ut til å være robust overfor feil i referanseverdiene.

Naturindeksen vil vanligvis bli beregnet fra observasjoner med ulike forventningsverdier og usikkerhet, og fra indikatorer som skaleres med ulike skaleringsfunksjoner. Både positive og negative effekter på estimatet av naturindeksen vil kunne forekomme ved samme beregning, og disse vil dermed delvis kansellere hverandre. Det er så langt ikke gjort analyser for å beregne effekten av lokasjonsforskyvning på indeksverdier beregnet fra ekspertenes input data. Med utgangspunkt i analysene av de konstruerte eksemplene vil vi imidlertid anta at lokasjonsforskyvning medfører en underestimering av naturindeksen når estimatene nærmer seg maksimalverdien 1, og at indeksens følsomhet overfor endringer i indikatorene generelt sett er redusert i dette verdiområdet når det forekommer usikkerhet i datagrunnlaget.

Vi anbefaler at det utvikles prosedyrer for estimering av naturindeksen som korrigerer for lokasjonsforskyvning og tar hensyn til mulige målefeil i referanseverdiene. Slike estimat av indeksen vil bedre reflektere endringer i indikatorene enn den etablerte beregningsmåten. Vi argumenterer at det vil være enklere å utvikle estimatorer med utgangspunkt i forventningsverdien til indeksens samplingfordeling framfor medianen. Videre anbefaler vi å fase ut OPT modellen og skisserer hvordan dette kan gjøres samtidig som de aktuelle indikatorene beholdes som del av naturindeksens datagrunnlag.

Bård Pedersen (bard.pedersen@nina.no), NINA, Postboks 5685 Sluppen, 7485 Trondheim
Olav Skarpaas (olav.skarpaas@nina.no), NINA, Gaustadalléen 21, 0349 Oslo

Abstract

Pedersen, B. and Skarpaas, O. 2012. Statistical properties of the Nature index for Norway. Measurement error and sensitivity. – NINA Report 797. 55 pp.

The Nature Index for Norway is an aggregate of biodiversity indicators where each indicator represents an aspect of biodiversity. The index is calculated for a delimited geographical area and a given year. Mathematically, the Nature Index is defined as a weighted average of scaled biodiversity indicators. Indicators are scaled by one of three scaling functions, LOW, MAX, and OPT, which each contain one parameter called the reference value. Observations of indicators are provided by experts and are associated with measurement error. Thus, when estimating the Nature Index each observation is assumed to be randomly drawn from a statistical distribution. The corresponding sampling distribution for an observation of the Nature Index is simulated using Monte Carlo methods. The median in this distribution is used to estimate the index, and the 0,025- and 0,975-quantiles are used as limits in a 95%-confidence interval that indicates the estimate's uncertainty.

The present report describes and explains how measurement error in input data affects the estimates of the Nature Index. Here, such effects are referred to as location displacement. The report describes how location displacement affects the sensitivity of the Nature Index with respect to changes in indicators through time. The report also studies whether the Nature Index is robust with respect to errors in the reference values. Conclusions are based on both formal, mathematical analyses of location displacement and on numerical analyses of simple, constructed examples where all indicators' sampling distributions are identical.

The analyses show that the size of location displacement depends on scaling model, location and dispersion of the indicator observations' sampling distributions, and the number of indicators. Displacement may be both positive and negative. It increases in absolute value with increasing dispersion of the observations' sampling distributions, and when the observations' expected values approach the reference values. Location displacement shows small increases in absolute value with increasing number of indicators before approaching a limit. LOW and MAX scaling cause less displacement than OPT. It is shown that measurement error may both increase and decrease the sensitivity of the Nature Index with respect to changes in the biodiversity indicators, depending on whether the changes cause increased or reduced location displacement. The Nature Index seems to be robust with respect to errors in the reference values.

The Nature Index will usually be estimated from observations that vary in expected value and dispersion and which are scaled with different models. These may cause both positive and negative displacements that partly cancel each other. So far, we have not estimated location displacement in index values calculated from experts' input data. However, based on the present analyses, it seems reasonable to assume that location displacement leads to negatively biased estimates of the Nature Index when most indicator observations are near their reference values. In such cases, uncertainty in data input reduces the index sensitivity with respect to changes in the indicators.

It is recommended that future estimates of the Nature Index correct for location displacement when they are calculated from indicator data with measurement error. Such estimates will better reflect changes in the observed values of indicators than current estimates. It is argued that it is easier to obtain corrected estimates when they are based on the index' expected value rather than its median. A further recommendation is not to use the OPT scaling model. It is described how this may be accomplished while keeping the affected indicators as part of the Nature Index database.

Bård Pedersen (bard.pedersen@nina.no), NINA, Postboks 5685 Sluppen, NO-7485 Trondheim, Norway
Olav Skarpaas (olav.skarpaas@nina.no), NINA, Gaustadalléen 21, NO-0349 Oslo, Norway

Innhold

Sammendrag	3
Abstract	4
Innhold	5
Forord	7
1 Innledning	8
2 Naturindeks for Norge – definisjon og estimering	10
2.1 Matematisk rammeverk.....	10
2.2 Økologisk rammeverk	12
2.3 Usikkerhet i datagrunnlaget	12
2.4 Estimering av naturindeksen.....	13
3 Hvordan usikkerhet i datagrunnlaget påvirker estimatet av naturindeksen	15
3.1 Metoder.....	15
3.1.1 Teoretisk analyse	15
3.1.2 Analyse av spesialtilfeller	15
3.1.3 Numerisk beregning av sannsynlighetsfordelinger	15
3.2 Resultat.....	16
3.2.1 Skalering med LOW modellen	16
3.2.2 Skalering med MAX modellen	17
3.2.3 Skalering med OPT modellen.....	18
3.2.4 Hvordan oppstår lokasjonsforskyvning?	18
4 Naturindeksens sensitivitet med hensyn til endring i indikatorene	21
4.1 Metoder.....	21
4.2 Resultat.....	22
5 Naturindeksens sensitivitet med hensyn til endringer og usikkerhet i referanseverdiene	24
5.1 Metoder.....	24
5.2 Resultat.....	24
6 Diskusjon og anbefalinger	28
7 Referanser	31
8 Vedlegg	32
8.1 Sannsynlighetsfordelinger – en kort oversikt.....	32
8.1.1 Stokastiske variabler	32
8.1.2 Sannsynlighetsfordeling	32
8.1.3 Diskrete, kontinuerlige og blandete sannsynlighetsfordelinger	32
8.1.4 Kvantiler og kvartiler	34
8.1.5 Sentralitet og spredning.....	34
8.1.6 Fordelingens form – skjevhet	35
8.2 Sannsynlighetsfordelingen til målinger av skalerte indikatorer.....	35
8.2.1 Skalering med LOW modellen	35
8.2.2 Skalering med MAX modellen	38
8.2.3 Skalering med OPT modellen.....	40
8.3 Sannsynlighetsfordelingen til målinger av naturindeksen.....	43

8.3.1	Indikatorer skalert med LOW modellen	43
8.3.2	Indikatorer skalert med MAX modellen.....	45
8.3.3	Indikatorer skalert med OPT modellen	48
8.4	Lokasjonsforskyvning og sensitivitet	50
8.4.1	Definisjon av lokasjons-, skalerings- og summeringsforskyvning	50
8.4.2	Skaleringsforskyvning.....	51
8.4.3	Summeringsforskyvning	52
8.4.4	Konklusjoner – lokasjonsforskyvning	52
8.4.5	Sensitivitet med hensyn til endring i indikatorenes forventningsverdier	53
8.4.6	Sensitivitet med hensyn til feil i referanseverdiene	54

Forord

Denne rapporten inngår i NINAs arbeid med å utvikle det metodiske rammeverket for Naturindeks for Norge. Dette arbeidet er presentert i en serie av NINA rapporter (NINA Rapport 347, 542, 794), og andre dokumenter (se Direktoratet for naturforvaltnings nettsider, www.dirnat.no/naturindeks).

Forfatterne er takknemlige for konstruktive kommentarer og diskusjoner med Iulie Aslaksen, Grégoire Certain, Steinar Engen, Silvio Funtowicz, Per Arild Garnåsjordet, Signe Nybø og Andrea Saltelli. Arbeidet er finansiert av Direktoratet for naturforvaltning.

Februar 2012

Bård Pedersen
Olav Skarpaas

1 Innledning

Naturindeks for Norge er et sammensatt mål for biologisk mangfold som skal gjenspeile tilstanden i terrestre-, limniske- og marine natursystemer i Norge inklusive kulturlandskapet. Indeksen beregnes for et geografisk område og en eller flere økosystemtyper ut fra et varierende antall indikatorer. Indikatorene representerer hver for seg et aspekt av biologisk mangfold. Naturindeksen sammenfatter på denne måten en omfattende mengde av relevant informasjon fra økologisk forskning og overvåkning og formidler denne til miljøforvaltningen og allmennheten på en forenklet måte (Nybø *et al.* 2008, Certain og Skarpaas 2010, Certain *et al.* 2011). Ved å beregne indeksens verdi ved flere ulike tidspunkt, er naturindeksen ment å gi en oversikt over positive og negative trender i utviklingen i biologisk mangfold i Norge og mer spesifikt avklare hvorvidt tapet av biologisk mangfold har stanset. Naturindeks for Norge ble første gang publisert i 2010 (Nybø (red.) 2010a,b). Indeksen var da beregnet for 9 økosystemtyper basert på drøyt 300 indikatorer.

Naturindeksen defineres rent matematisk som et veid gjennomsnitt av skalerte biodiversitetsindikatorer (Nybø *et al.* 2008, Certain og Skarpaas 2010, Certain *et al.* 2011). Indikatorene måles imidlertid på mange ulike skalaer som i utgangspunktet ikke er sammenliknbare. Indikatorene reskaleres derfor til en felles skala for biologisk mangfold med 0 og 1 som ekstremverdier. Dette skjer ved hjelp av tre ulike skaleringsfunksjoner. Valg av skaleringsfunksjon bestemmes av hvorvidt indikatoren samvarierer positivt eller negativt med det aspektet av biologisk mangfold som indikatoren representerer, eller om det er en unimodal sammenheng.

Skaleringsfunksjonene inneholder kun en parameter kalt referanseverdien. Referanseverdien tilsvarer maksimumsverdien 1 etter reskalering. Referanseverdiene definerer hvilke verdier for de ulike indikatorene som representerer samme tilstand med hensyn til biologisk mangfold. Samtidig angir referanseverdien maksimums- og minimumsverdier for det intervallet av indikatorens måleskala hvor endringer påvirker naturindeksen. Referanseverdiene har derfor en felles økologisk tolkning. Referanseverdier diskuteres videre i Nybø *et al.* (2008), Certain og Skarpaas (2010) og Certain *et al.* (2011).

Denne konstruksjonen gir et fleksibelt konsept. Ved å variere vektor og utvalget av indikatorer kan naturindeksen beregnes for ulike geografiske områder, ulike "økosystem", ulike funksjonelle grupper eller trofiske nivåer. Konseptet kan på samme vis benyttes til å utvikle tematiske indekser som fokuserer på bestemte utfordringer og påvirkningsfaktorer på det biologiske mangfoldet og måler effekten av disse (Certain *et al.* 2011).

Datagrunnlaget, dvs. indikatorenes verdier og referanseverdier for norske kommuner, er framskaffet av ca. 125 forskere som alle er nasjonale eksperter på de ulike indikatorene (Nybø (red.) 2010b). Data er enten overvåkingsdata, fra modeller, eller generert som såkalte ekspertvurderinger (Nybø (red.) 2010a). Alle observasjoner angis med usikkerhet. Denne usikkerheten behandles i naturindekssammenheng med å tilordne en sannsynlighetsfordeling til hver enkelt observasjon av en indikator (Certain og Skarpaas 2010, Certain *et al.* 2011). Observasjonene og alle størrelser beregnet fra dem er derfor stokastiske variabler. Dette gjelder observasjoner/målinger av skalerte indikatorer og selve naturindeksen. Ettersom en ikke kjenner de aktuelle sannsynlighetsfordelingene, må naturindeksens verdi estimeres snarere enn beregnes nøyaktig.

Ved beregning av naturindeksen er det så langt ikke tatt hensyn til usikkerhet i referanseverdiene.

Naturindeksen er et eksempel på en sammensatt indeks som for å være vellykket må være i stand til å fange opp og påvirkes av endringer i biologisk mangfold over tid. Dette avhenger bl.a. av at slike reflekteres i noen av de utvalgte indikatorene, og at indeksen beregnes og estimeres på en slik måte at endringene påvirker indeksen også når det er usikkerhet i data-

grunnlaget. Samtidig er det viktig for indeksens troverdighet at det tas hensyn til usikkerhetene i datagrunnlaget slik at disse reflekteres i en tilsvarende usikkerhet i angivelsen av naturindeksen (Certain og Skarpaas 2010).

I denne rapporten undersøker vi hvordan usikkerhet i datagrunnlaget påvirker estimatene av naturindeksens verdi, og hvordan usikkerhet påvirker indeksens evne til å måle endringer i biologisk mangfold over tid slik den i dag estimeres fra dataene. Vi studerer enkle, konstruerte eksempler med kjente egenskaper der vi er i stand til å beregne sannsynlighetsfordelingen til observasjoner av naturindeksen ved hjelp av numeriske metoder. Vi viser at usikkerhet i datagrunnlaget har en effekt på estimatet av naturindeksen under visse betingelser, beskriver disse betingelsene og forklarer hvordan effektene oppstår. De tre skaleringsfunksjonene viser seg å oppføre seg forskjellig med hensyn på dette. Et viktig moment er derfor å vise og forklare forskjellene mellom de tre skaleringsfunksjonene. Vi viser videre at disse forholdene påvirker naturindeksens evne til å fange opp tilstandsendringer i biologisk mangfold når mange av indikatorene er nær referanseverdiene. Vi undersøker også i hvilken grad naturindeksen, slik den i dag estimeres, er robust i forhold til feil i referanseverdiene. Til slutt skisserer vi og gir anbefalinger om hvordan metodene for å estimere naturindeksen kan revideres med tanke på å korrigere for effektene av usikkerhet. Vi starter med å gi en nærmere beskrivelse av hvordan indeksen rent matematisk er definert og hvordan den i dag estimeres.

Med rapporten følger fire vedlegg. Det første gir en kort introduksjon til stokastiske variabler og sannsynlighetsfordelinger. De neste to behandler fordelingene til observasjoner av skalerte biodiversitetsindikatorer og naturindeksen. Det siste vedlegget inneholder en teoretisk analyse av effektene av usikkerhet på estimatet av naturindeksen. I tillegg gjengis en del tekniske detaljer som ligger til grunn for analysen av indeksens sensitivitet med hensyn til endringer i indikator- og referanseverdier.

2 Naturindeks for Norge – definisjon og estimering

2.1 Matematisk rammeverk

Naturindeksen kan beregnes på ulike nivå av aggregering, for en spesifikk økosystemtype i en kommune, for alle økosystemene i kommunen, for en økosystemtype i hele landet, eller en verdi for all natur i hele landet, osv. (Certain og Skarpaas 2010, Certain *et al.* 2011). Vi skal i det følgende ignorere dette aspektet og i stedet se på en generell situasjon hvor indeksen (NI) beregnes på et eller annet nivå av aggregering for et ikke nærmere angitt geografisk område og tidspunkt. Indeksen er definert som en veid sum av et antall (n) skalerte biodiversitetsindikatorer (S_1, \dots, S_n)

$$(1) \quad NI = \sum_{i=1}^n S_i w_i,$$

der vektene (w_i) oppfyller

$$(2) \quad \sum_{i=1}^n w_i = 1.$$

De skalerte biodiversitetsindikatorer er funksjoner av ikke skalerte indikatorer (U_i)

$$(3) \quad S_i = S_i(U_i; U_i^{ref}),$$

der U_i som oftest er målt på kontinuerlige måleskalaer. Diskrete indikatorer forekommer. Skaleringsfunksjonenes (f_i) verdiområde er intervallet $[0,1]$. Sammen med betingelsen likning 2 medfører dette at naturindeksens verdiområde også er $[0,1]$. Det er tre typer skaleringsfunksjoner, LOW, MAX, eller OPT (Figur 1):

LOW:

$$(4a) \quad S_i(U_i) = \begin{cases} \frac{U_i}{U_i^{ref}}, & 0 \leq U_i \leq U_i^{ref} \\ 1, & U_i > U_i^{ref} \end{cases},$$

MAX:

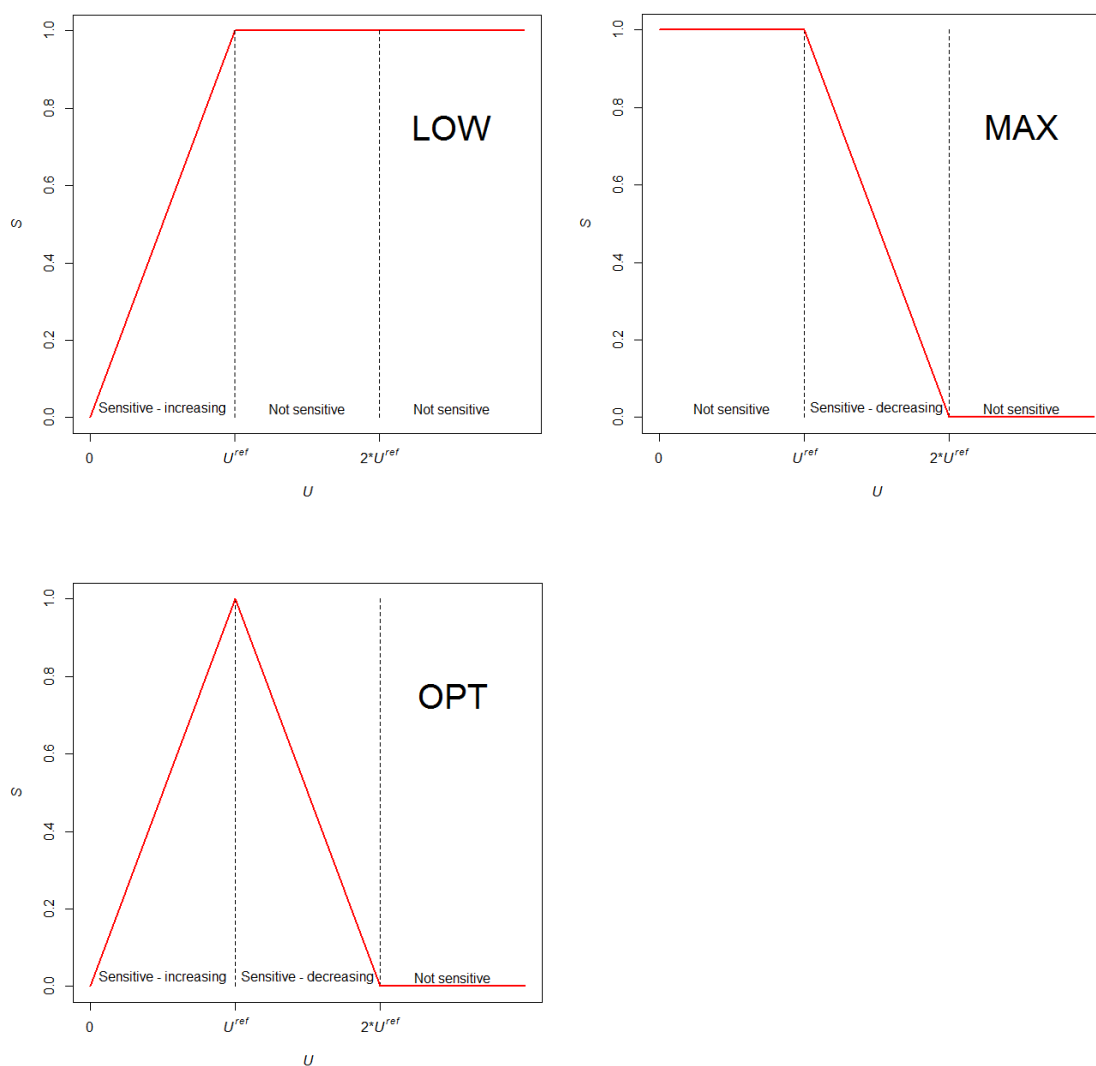
$$(4b) \quad S_i(U_i) = \begin{cases} 1, & 0 \leq U_i \leq U_i^{ref} \\ 2 - \frac{U_i}{U_i^{ref}}, & U_i^{ref} < U_i \leq 2U_i^{ref} \\ 0, & U_i > 2U_i^{ref} \end{cases}.$$

OPT:

$$(4c) \quad S_i(U_i) = \begin{cases} \frac{U_i}{U_i^{ref}}, & 0 \leq U_i \leq U_i^{ref} \\ 2 - \frac{U_i}{U_i^{ref}}, & U_i^{ref} < U_i \leq 2U_i^{ref} \\ 0, & U_i > 2U_i^{ref} \end{cases},$$

LOW modellen benevnes også MIN i noen publikasjoner (se Certain og Skarpaas 2010). Referanseverdien, U_i^{ref} , er eneste parameter i skaleringsfunksjonene. Rent matematisk har parameteren flere "roller" som varierer mellom de tre typene av skaleringsfunksjoner. Referanseverdien definerer en skaleringskonstant, men representerer i tillegg grenseverdier mellom intervaller av definisjonsområdet til indikatoren U_i , hvor skaleringsfunksjonene oppfører seg forskjellig.

LOW modellen benyttes når det er en positiv sammenheng mellom indikatoren og biologisk mangfold. Dette gjelder for de fleste indikatorene. Referanseverdien i denne modellen deler indikatoraksen i to intervaller. I intervallet mellom 0 og referanseverdien er S_{LOW} , og dermed også NI , sensitiv til endringer i indikatoren. S_{LOW} er i dette intervallet en monotont stigende funksjon av U_i med konstant stigningstall lik $\frac{1}{U_i^{ref}}$. For indikatorverdier større enn U_i^{ref} er naturindeksen ikke sensitiv i forhold endringer i indikatoren. Hele dette intervallet av indikatorens definisjonsområde tilsvarer den maksimale verdien etter skalering, dvs. den beste tilstand for det biologiske mangfoldet som indikatoren representerer, som kan bli målt i naturindekssammenheng.



Figur 1. Skaleringsmodellene LOW, MAX og OPT.

MAX modellen benyttes når det er en negativ sammenheng mellom indikatoren og biologisk mangfold. Referanseverdien i denne modellen deler indikatorens definisjonsområde i tre intervaller. I intervallet mellom 0 og referanseverdien er den skalerte indikatoren lik 1 og naturindeksen er ikke sensitiv med hensyn til endringer i indikatoren U_i . Hele intervallet tilsvarer den maksimale verdien indikatoren kan få etter skalering. I intervallet mellom U_i^{ref} og $2U_i^{ref}$ er S_{MAX} en monotont avtagende funksjon av U_i med konstant stigningstall $\frac{-1}{U_i^{ref}}$. For indikatorverdier større enn $2U_i^{ref}$ er S_i lik 0, og naturindeksen er ikke sensitiv i forhold til endringer i indikatoren. Intervallet tilsvarer den minimale verdien for tilstanden til biologisk mangfold representert av indikatoren.

OPT modellen benyttes når det er en unimodal sammenheng mellom indikatoren og biologisk mangfold. Som for MAX modellen deler referanseverdien indikatorens definisjonsområde i tre intervaller. I intervallet mellom 0 og referanseverdien er naturindeksen sensitiv til endringer i indikatoren U_i . S_{OPT} er i dette en monotont stigende funksjon av U_i med konstant stigningstall lik $\frac{1}{U_i^{ref}}$. I intervallet mellom U_i^{ref} og $2U_i^{ref}$ er S_{OPT} en monotont avtagende funksjon av U_i med stigningstall $\frac{-1}{U_i^{ref}}$. For indikatorverdier større enn $2U_i^{ref}$ er S_{OPT} lik 0.

Ekspertene har valgt skaleringsfunksjon for de ulike indikatorene ut fra deres kunnskap om indikatorene. Av indikatorene som lå til grunn for den første publikasjonen av Naturindeks for Norge ble 239 skalert med LOW modellen, 10 med MAX og 59 med OPT modellen (Nybø (red.) 2010a).

2.2 Økologisk rammeverk

Med det økologiske rammeverket for NI menes her de økologiske definisjonene av begrepene biomangfold, biodiversitetsindikator og referanseverdi som benyttes i NI-sammenheng. Disse definisjonene gir meningsinnhold til naturindeksen og beskriver hva indeksen måler. Rammeverket omfatter også vurderinger knyttet til utvalget av indikatorer som inngår i beregningen av naturindeksen, allokering av type skaleringsfunksjon til den enkelte indikator, hvordan referansetilstanden til den enkelte indikator skal beregnes og fastsetting av vektene (w_i , likningene 1 og 2). Det økologiske rammeverket er presentert og diskutert i Nybø *et al.* (2008), Certain og Skarpaas (2010) og Certain *et al.* (2011), mens de ulike biodiversitetsindikatorene omtales i Nybø (red., 2010b).

2.3 Usikkerhet i datagrunnlaget

Datagrunnlaget for beregning av naturindeksen er beskrevet i Nybø (red., 2010b). Målinger av de ulike indikatorene er forbundet med usikkerhet (Certain og Skarpaas 2010, Certain *et al.* 2011). Som forklart i Vedlegg 8.1.1 modelleres usikkerheten ved hjelp av en sannsynlighetsfordeling for hver observasjon. Slike fordelinger kalles gjerne samplingfordelinger. En sannsynlighetsfordeling beskrives ved hjelp av en fordelingsfunksjon, $F_{U_i}(u)$, som for hver mulig verdi (u) av indikatoren angir sannsynligheten for at målingens sanne verdi er mindre eller lik den aktuelle verdien (Vedlegg 8.1.2). Det er altså hver enkelt observasjon, og ikke indikatoren som sådan som tilordnes en sannsynlighetsfordeling. Likevel vil vi i det følgende for enkelthets skyld ofte benytte feilaktige formuleringer som «indikatorens fordeling» når vi egentlig refererer til observasjonens fordeling.

På grunn av potensielle målefeil betraktes en observasjon eller måling av en biodiversitetsindikator som en realisering av et stokastisk forsøk eller en stokastisk variabel (Vedlegg 8.1.1). Det medfører at alle størrelser avledet fra disse observasjonene, så som observasjoner av skalerte indikatorer og naturindeksen, også er realiseringer av stokastiske variabler, og har sannsynlighetsfordelinger som beskrives av fordelingsfunksjoner ($F_{S_i}(s)$, $F_{NI}(x)$, $0 \leq s, x \leq 1$). Naturindeksen kan derfor ikke beregnes fra datagrunnlaget som en entydig verdi. I stedet beregnes parametere eller størrelser som beskriver viktige egenskaper ved indeksens samplingfordeling, dvs. fordelings plassering på tallinja og dens spredning. De størrelsene som benyttes er fordelings median, m_{NI} , som mål på dens plassering på tallinja, og et 95%-konfidensintervall for en tilfeldig observasjon fra fordelingen som mål på fordelings spredning.

Vedlegg 8.1 gir en kort introduksjon til sannsynlighetsfordelinger, fordelingsfunksjoner, fordelingenes egenskaper og noen parametere og statistiske størrelser som benyttes til å beskrive egenskapene. I vedlegg 8.2 analyseres fordelingene til observasjoner av skalerte indikatorer. Vedlegg 8.3 inneholder en analyse av sannsynlighetsfordelingen til observasjoner av naturindeksen i de spesielle tilfellene hvor indeksen beregnes fra n kontinuerlige biodiversitetsindikatorer som alle har samme fordeling, skaleres med samme modell og gis lik vekt.

Ettersom vi ikke vet hvilken samplingfordeling observasjoner av de ulike indikatorene har, blir parameterne som beskriver naturindeksens fordeling estimert snarere enn beregnet nøyaktig. Estimeringsmetoden er beskrevet i avsnitt 2.4.

Referanseverdiene er også forbundet med målefeil. Ved beregning av naturindeksen er det imidlertid så langt ikke tatt hensyn til usikkerhet i referanseverdiene.

2.4 Estimering av naturindeksen

Naturindeksen presenteres ved å oppgi estimat av medianen til naturindeksens samplingfordeling (\hat{m}_{NI} , Certain og Skarpaas 2010, Certain *et al.* 2011). Dette kan ses på som et estimat av naturindeksens verdi. Usikkerheten i estimatet angis ved et 95%-konfidensintervall for en tilfeldig observasjon fra fordelingen, $[\hat{q}_{NI,0.025}, \hat{q}_{NI,0.975}]$ (Vedlegg 8.1.4).

Basis for estimering av naturindeksen er ekspertenes estimat av forventningsverdien ($\hat{\mu}_{U_i}$), samt nedre ($\hat{q}_{U_i,0.25}$) og øvre ($\hat{q}_{U_i,0.75}$) kvartil til alle observasjoner/målinger av biodiversitetsindikatorene som inngår i beregningen av NI . Observasjonene er enten basert på data, modeller eller ekspertvurderinger (Nybø (red.) 2010a, Certain og Skarpaas 2010, Certain *et al.* 2011). For hver enkelt observasjon blir $F_{\hat{U}_i}(u)$ estimert ved å tilpasse en teoretisk modellfordeling til de oppgitte verdiene for $\hat{\mu}_{U_i}$, $\hat{q}_{U_i,0.25}$ og $\hat{q}_{U_i,0.75}$. Tilpasningen skjer ved å finne den modellfordelingen som minimaliserer kvadratsummen av avstandene mellom de oppgitte parameterverdiene og de tilsvarende i modellfordelingene (Certain og Skarpaas 2010). De aktuelle modellfordelingene i denne sammenheng er normal-, trunkert normal-, lognormal-, Weibull-, Gumbell- og gamma fordeling for kontinuerlige indikatorer, og Poisson-, negative binomial-, og "zero-inflated" Poisson fordeling for diskrete biodiversitetsindikatorer.

Parametrisk bootstrapping (Manly 2007) benyttes til å generere 999 tilfeldige observasjoner fra fordelingen gitt ved $F_{NI}(x)$. En observasjon genereres ved å trekke tilfeldig en observasjon fra hver indikators estimerte samplingfordeling, og deretter beregne naturindeksens verdi som gitt ved likningene 1, 3 og 4. Parametrene $\hat{q}_{NI,0.025}$, \hat{m}_{NI} , $\hat{q}_{NI,0.975}$, beregnes fra fordelingen av de 999 observasjonene.

Framgangsmåten kan oppfattes å benytte elementer fra bayesiansk inferens ved estimering av naturindeksen fra usikre data. Input data behandles som om ekspertene ut fra sine kunnskaper oppgir parametere i såkalte *posteriori* fordelinger til observasjoner av indikatorene, og disse parameterne benyttes til å estimere *posteriori* fordelingene til observasjoner av både indikatorer og naturindeksen. Fokus er på *posteriori* fordelingen til observasjonen av naturindeksen der medianen og 95%-konfidensintervallet oppgis for å beskrive dens egenskaper.

Alternativt kan estimeringsprosedyren forstås i et tradisjonelt, frekventistisk perspektiv basert på en modell der ekspertenes input er realiseringer av stokastiske variabler, der $\hat{q}_{NI,0.025}$ og $\hat{q}_{NI,0.975}$ oppfattes som grensene i et 50%-konfidensintervall, der fordelingene til observasjonene estimeres fra en mengde modellfordelinger, og der \hat{m}_{NI} oppfattes som et punktestimat for naturindeksens sanne verdi. I denne tradisjonen vil det være naturlig å undersøke om punktestimatet er medianrett eller ikke.

3 Hvordan usikkerhet i datagrunnlaget påvirker estimatet av naturindeksen

3.1 Metoder

Vårt formål var å undersøke om usikkerhet eller målefeil i datagrunnlaget hadde en effekt på estimatet av naturindeksen. Mer presist, vi ønsket å undersøke om usikkerhet i indikatorenes verdier påvirket plasseringen av naturindeksens samplingfordeling på tallinja (dvs dens sentral-tendens, vedlegg 8.1.5). Vi målte denne effekten, som vi her kaller lokasjonsforskyvning eller bare forskyvning, b_{NI} , som differansen mellom m_{NI} og verdien indeksen ville ha fått hvis den ble beregnet direkte ved hjelp av likningene (1), (3), og (4) fra forventningsverdiene (μ_{U_i}) til indikatorene som inngikk, dvs verdien en fikk for naturindeksen hvis μ_{U_i} representerte indikatorenes «sanne» verdier og det ikke var målefeil i dataene. Sett i et frekventistisk perspektiv er lokasjonsforskyvning en mulig årsak til at en estimator ikke er median- eller forventningsrett. Den presise definisjonen av lokasjonsforskyvning som vi benyttet, samt mulige alternative definisjoner, er gjengitt og diskutert i vedlegg 8.4.1.

3.1.1 Teoretisk analyse

Lokasjonsforskyvning kan oppstå ved skalering av biodiversitetsindikatorene (her kalt skaleringsforskyvning) og ved summering av de skalerte indikatorene (summeringsforskyvning). Den totale forskyvning ved beregning av m_{NI} er summen av skaleringsforskyvning og summeringsforskyvning (likning 31a). Vi foretok en teoretisk analyse av lokasjonsforskyvning i naturindeksen ved hjelp av analytiske, matematiske metoder. Gjennom analysen ønsket vi å forstå bedre om og hvordan forskyvning oppstod, å sammenlikne skaleringsmodellene med hensyn til skaleringsforskyvning, å finne ut om naturindeksens forventningsverdi, μ_{NI} , også viste tilsvarende forskyvning og hvordan slik forventningsverdiforskyvning eventuelt oppstod, og om forskyvning kunne unngås ved å beregne indeksen fra indikatorenes medianer m_{U_i} i stedet for deres forventningsverdier.

3.1.2 Analyse av spesialtilfeller

Vi ønsket videre å studere hvordan lokasjonsforskyvning varierer med forventningsverdiene til indikatorene, usikkerheten i datagrunnlaget, valg av skaleringsmodell og antall biodiversitetsindikatorer som inngår. Vi gjennomførte dette ved å undersøke konstruerte spesialtilfeller der naturindeksen ble beregnet fra uavhengige observasjoner av indikatorene som alle hadde samme samplingfordeling, lik referanseverdi, ble skalert med samme skaleringsfunksjon, og gitt de samme vektene. Vi antok også at observasjonene fulgte en lognormal fordeling.

Vi kunne dermed for hver skaleringsmodell, gitt begrensningene over, beregne forskyvning som funksjon av indikatorenes felles forventningsverdi, av spredningen i deres fordeling målt som kvartilbredden, og av antall indikatorer som inngikk.

3.1.3 Numerisk beregning av sannsynlighetsfordelinger

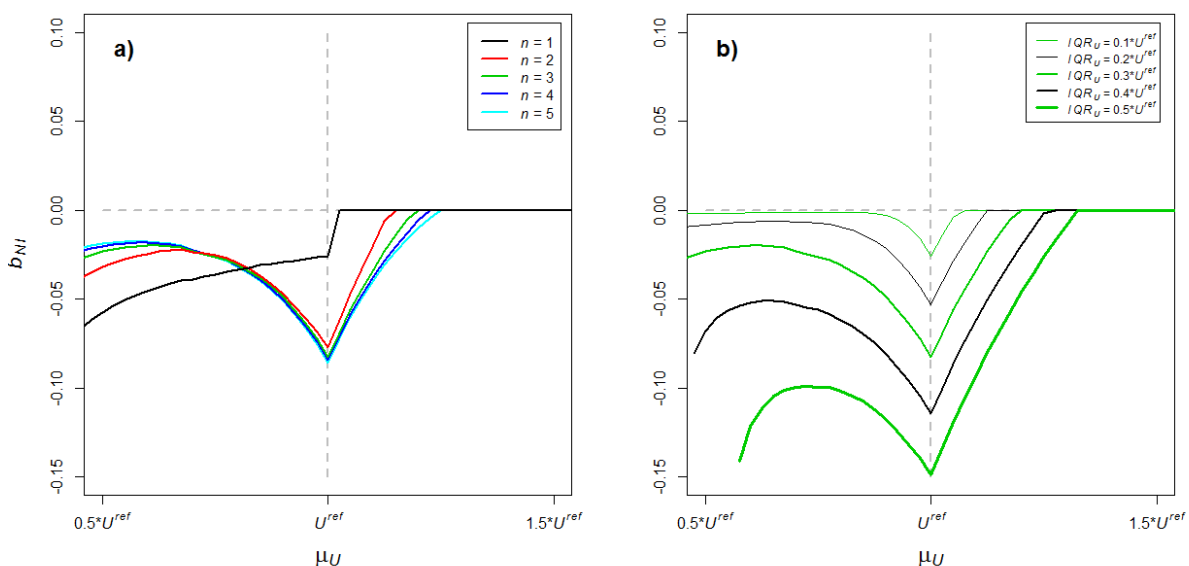
Beregning av lokasjonsforskyvning, slik som definert her, forutsetter at en kjenner m_{NI} i spesialtilfellene beskrevet over. Vi regnet ut medianen slik som definert i vedlegg 8.1.4 direkte fra $F_{NI}(x)$. Vi utviklet et program i R (R Development Core Team 2008) som beregnet fordelingen numerisk basert på likningene 19 - 27 avhengig av hvilken skaleringsmodell som ble benyttet. Integralene som inngår i likningene ble beregnet med midtpunkts integrasjon (Andreassen et

al. 1981) basert på en oppdeling av verdiområdet til U som tilsvarte integrasjonsintervaller med bredde lik $0,0001 * U^{ref}$ eller mindre. Konvulsjonene ble også beregnet numerisk med R-funksjonen “convolve”.

3.2 Resultat

3.2.1 Skalering med LOW modellen

De numeriske beregningene viste at b_{NI} er mindre eller lik null når naturindeksen beregnes fra indikatorer som skales med LOW modellen (Figur 2, se også ulikhet 36). Skaleringsforskyvning, som er lik lokasjonsforskyvningen når naturindeksen beregnes fra kun en indikator, avtar i absoluttverdi med økende indikatorverdi og er null når både medianen og forventningsverdien til indikatorens samplingfordeling er større enn referanseverdien (vedlegg 8.4.2).



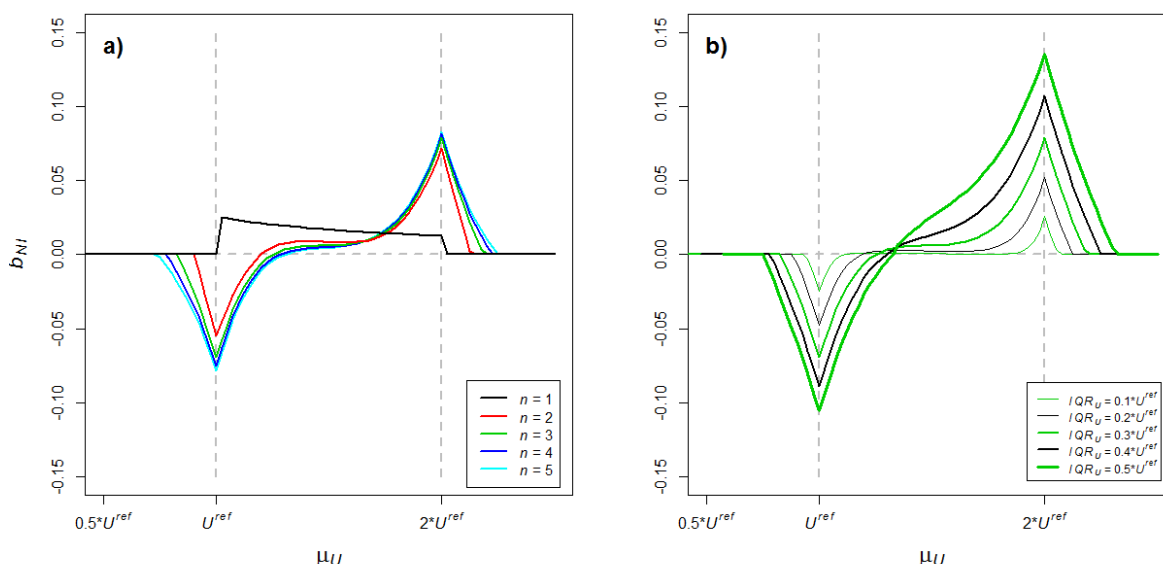
Figur 2. Lokasjonsforskyvning i naturindeksen, $b_{NI} = m_{NI} - \sum_{i=1}^n S_i(\mu_{U_i}) w_i$, som funksjon av indikatorenes forventningsverdier, μ_U , kvartilbredde, IQR_U , samt antall indikatorer, n , når indeksen beregnes som gjennomsnittet av n skalerte biodiversitetsindikatorer der indikatorene før skalering har samme lognormal fordeling og referanseverdi, og alle skales med LOW-modellen. I a) er naturindeksen beregnet fra $n = 1 \dots 5$ indikatorer der deres fordeling har kvartilbredde lik 0,3 ganger referanseverdien U^{ref} . Heltrukken svart linje representerer dermed kun skaleringsforskyvning. I b) er indeksen beregnet fra 3 indikatorer der IQR_U varierer fra 0,1 til 0,5 ganger U^{ref} . Lokasjonsforskyvning er beregnet numerisk slik som beskrevet i kapittel 3.1.4.

Lokasjonsforskyvningen varierer noe forskjellig fra dette mønsteret når indeksen beregnes fra to eller flere indikatorer. Forskyvningen er da mindre i absoluttverdi enn skaleringsforskyvning for lave verdier av indikatorene, men absoluttverdien øker igjen når indikatorenes forventningsverdier nærmer seg referanseverdien. Lokasjonsforskyvning forekommer også for verdier større enn referanseverdien. Lokasjonsforskyvning avtar i absoluttverdi med antall indikatorer for lave indikatorverdier, men øker med antall indikatorer for verdier som nærmer seg og er større enn referanseverdien. Effekten av å legge til en ny indikator avtar imidlertid etter hvert som antallet øker.

De numeriske beregningene viste at b_{NI} under LOW modellen varierer mellom 2,5% og 10% av den estimerte verdien til naturindeksen når indeksen beregnes fra to eller flere indikatorer med forventningsverdier mellom en halv og en gang referanseverdien og usikkerheten i dataene tilsvarer en kvartilbredde lik 3 tideler av referanseverdien (figur 2a). Beregningene viste også at b_{NI} varierer mye med usikkerheten i datagrunnlaget. Forskyvningen øker i absoluttverdi med økende usikkerhet i datagrunnlaget (figur 2b).

3.2.2 Skalering med MAX modellen

For MAX modellen viste de numeriske beregningene at skaleringsforskyvning er positiv når forventningsverdien til indikatoren er mellom referanseverdien og drøyt 2 ganger referanseverdien, og lik null ellers (Figur 3a, se også ulikhet 37 og vedlegg 8.4.2). Skaleringsforskyvning er stort sett en avtagende funksjon av indikatorenes forventningsverdier i intervallet hvor skaleringsforskyvningen er positiv. Under denne modellen er skaleringsforskyvning gjerne mindre enn under LOW modellen ved samme usikkerhet i datamaterialet.

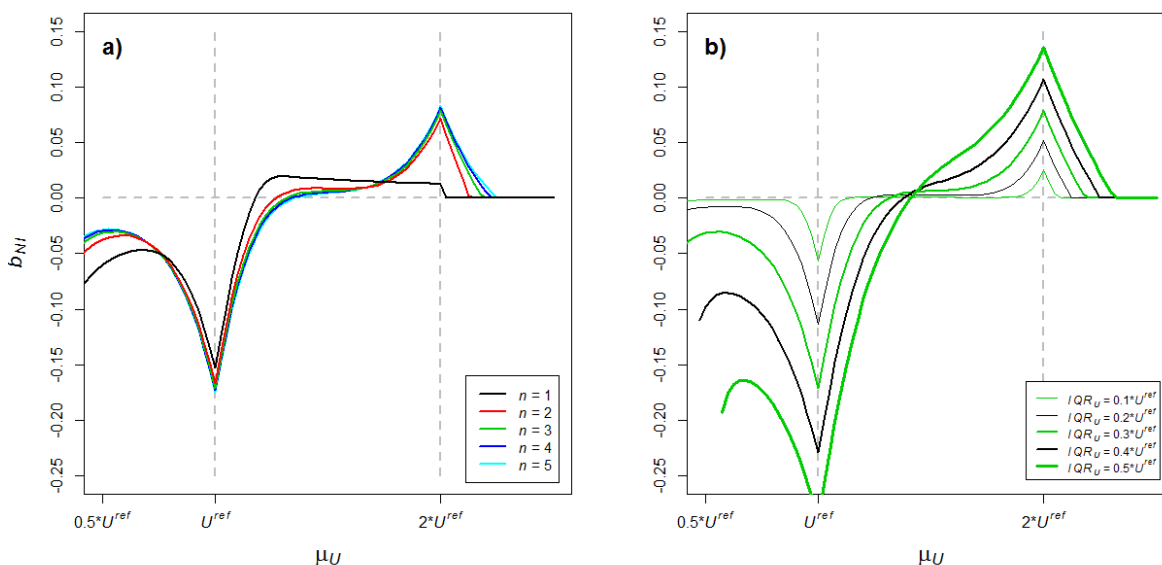


Figur 3. Lokasjonsforskyvning i naturindeksen, $b_{NI} = m_{NI} - \sum_{i=1}^n S_i(\mu_{U_i}) w_i$, som funksjon av indikatorenes forventningsverdier, μ_U , kvartilbredde, IQR_U , samt antall indikatorer, n , når indeksen beregnes som gjennomsnittet av n skalerte biodiversitetsindikatorer der indikatorene før skalering har samme lognormal-fordeling og referanseverdi, og alle skaleres med MAX-modellen. a) og b) som i Figur 2.

Når naturindeksen beregnes fra to eller flere indikatorer skalert med MAX modellen er b_{NI} tilnærmet lik null bortsett fra når indikatorenes forventningsverdier er rundt referanseverdien eller 2 ganger referanseverdien, eller når det er stor usikkerhet i dataene. Lokasjonsforskyvningen er negativ når forventningsverdiene er rundt referanseverdien og positivt når forventningsverdiene er rundt 2 ganger referanseverdien (Figur 3). I disse intervallene øker b_{NI} med antall indikatorer. Men som for LOW modellen avtar effekten av å legge til en ny indikatorer etter hvert som antallet øker. Lokasjonsforskyvning forekommer for forventningsverdier mindre enn referanseverdien og for verdier større enn 2 ganger referanseverdien.

3.2.3 Skalering med OPT modellen

Våre numeriske beregninger av skaleringsforskyvning under OPT modellen viste at forskyvningen er negativ når forventningsverdien til indikatoren er mindre og lik referanseverdien, etter hvert blir positiv når forventningsverdien blir høyere enn dette. Skaleringsforskyvning blir til slutt null når forventningsverdien er drøyt 2 ganger referanseverdien (Figur 4). Under OPT modellen er skaleringsforskyvning større i absoluttverdi enn under de andre modellene når indikatorens forventningsverdi er rundt referanseverdien. En negativ skaleringsforskyvning ved referanseverdien innebærer at indikatorens maksimale verdi etter skalering er mindre enn 1 når det er usikkerhet i dataene. Dette er forskjellig fra de andre to skaleringsmodellene. Når usikkerheten tilsvarer en kvartilbredde lik 3 tideler av referanseverdien er denne maksimalverdien ca. 0,85.



Figur 4. Lokasjonsforskyvning i naturindeksen, $b_{NI} = m_{NI} - \sum_{i=1}^n S_i(\mu_{U_i}) w_i$, som funksjon av indikatorenes forventningsverdier, μ_U , kvartilbredde, IQR_U , samt antall indikatorer, n , når indeksen beregnes som gjennomsnittet av n skalerte biodiversitetsindikatorer der indikatorene før skalering har samme lognormal-fordeling og referanseverdi, og alle skaleres med OPT-modellen. a) og b) som i Figur 2.

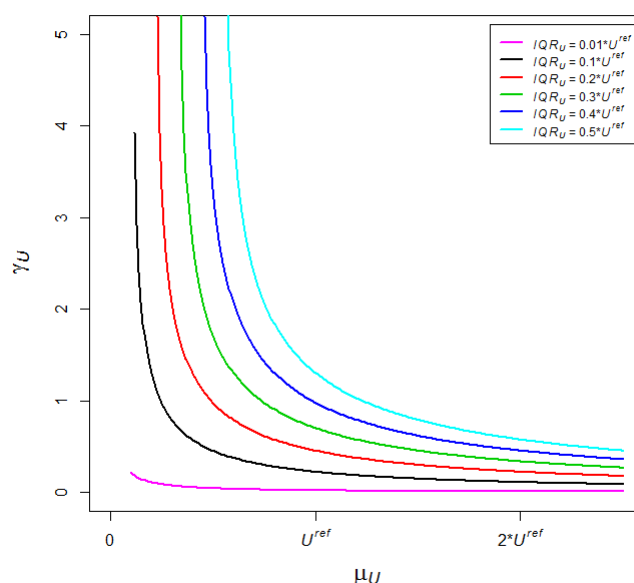
Når naturindeksen beregnes fra to eller flere indikatorer skalert med OPT modellen varierer lokasjonsforskyvning med indikatorenes forventningsverdier på samme måte som under LOW modellen når forventningsverdiene er mindre enn referanseverdien, og som under MAX modellen når de er høyere enn referanseverdien (Figur 4). Imidlertid er forskyvningen større i absoluttverdi for forventningsverdier rundt referanseverdien sammenliknet med de andre modellene.

3.2.4 Hvordan oppstår lokasjonsforskyvning?

Den teoretiske analysen av lokasjonsforskyvning er gjengitt i vedlegg 8.4.1 – 8.4.4. Analysen viste at forskyvning som opptrer i naturindeksen har tre årsaker.

Forskyvning oppstår for det første fordi fordelingene til observasjoner av indikatorene gjerne har positiv skjevhet, γ_U (vedlegg 8.1.6). Skjevheten skyldes at observasjonene er ikke-negative og dermed måles på skala med absolutt nullpunkt. Dette innebærer at medianen i fordelingen,

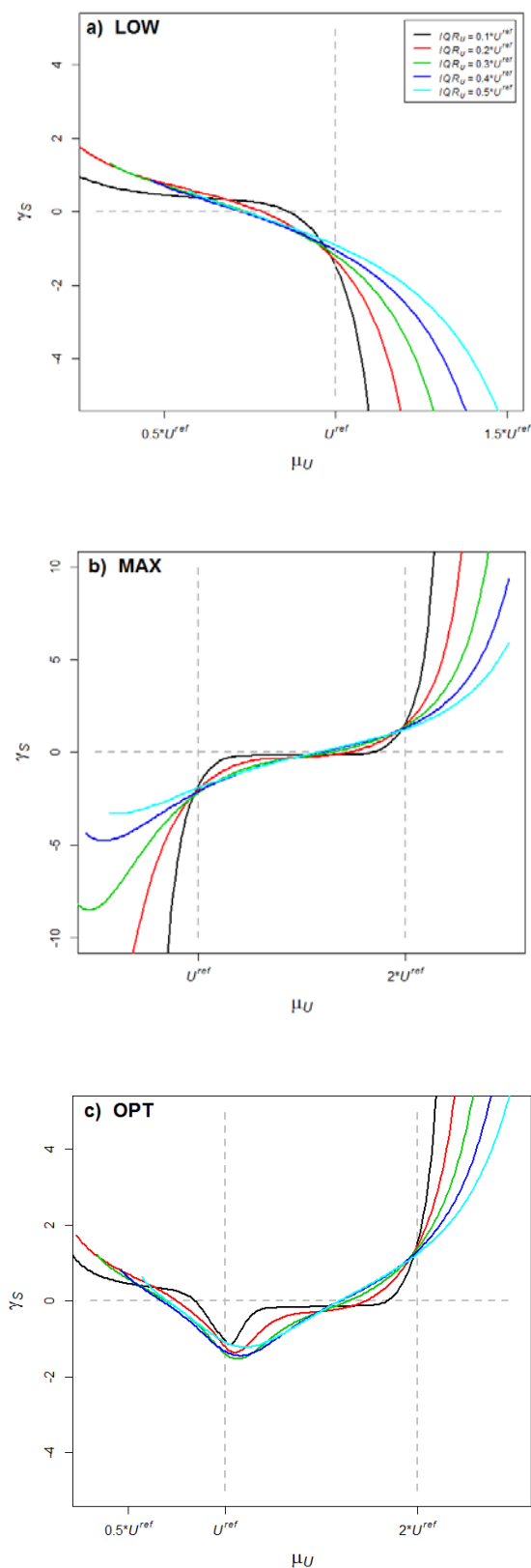
m_U , gjerne er mindre enn forventningsverdien, μ_U . Skjevheten avtar med økende verdier av μ_U , men øker med fordelings spredning (Figur 5). Naturindeksen estimeres som medianen i dens fordeling, mens indikatorene estimeres som forventningsverdien i deres fordeling. Dette skiftet fra forventningsverdi til median gir en negativ forskyvning under LOW modellen når $m_U < U^{ref}$ (ulikhet 36), og en positiv forskyvning under MAX modellen når $\mu_U > U^{ref}$ og samtidig $m_U < 2U^{ref}$ (ulikhet 37). I figurene 2a og 3a ses dette bidraget til forskyvning under LOW og MAX modellene som den heltrukne, svarte linjen. Under OPT modellen er dette bidraget negativt når $\mu_U, m_U < U^{ref}$ og positivt når $U^{ref} < \mu_U, m_U < 2U^{ref}$.



Figur 5. Skjevhet i lognormal fordelingen, γ_U , som funksjon av fordelings forventningsverdi, μ_U , og kvartilbredde, IQR_U .

Under OPT modellen er det i tillegg et annet bidrag til skaleringsforskyvning (likning 38). Dette skyldes at skaleringsmodellen ikke er en monoton funksjon av U . Når m_U er nær referanseverdien vil den skalerte verdien, $S_{OPT}(m_U)$, bli større enn medianen til den skalerte indikatoren, m_S , noe som gir en negativ forskyvning. Størrelsen på bidraget ses ved å sammenlikne den heltrukne, svarte linjen i figur 4a med tilsvarende linje i figur 2a for $\mu_U < U^{ref}$, og i figur 3a for $U^{ref} < \mu_U < 2U^{ref}$.

Et tredje bidrag oppstår ved summasjon av de skalerte indikatorene. Bidraget skyldes at fordelingen til observasjoner av naturindeksen, som er gjennomsnittet av de skalerte indikatorene, endrer form etter hvert som antall indikatorer øker, selv om alle indikatorer har samme fordeling (se Figur 18 – 20 i vedlegg 8.3). Skalering med de tre modellene medfører at de skalerte indikatorene får fordelinger med enten positiv eller negativ skjevhet, avhengig av skaleringsmodell samt plassering av og spredning til indikatorenes fordelinger. Skalering med LOW modellen gir fordelinger med positiv skjevhet for lave verdier av μ_U , og fordelinger med negativ skjevhet når μ_U nærmer seg eller er større enn referanseverdien (Figur 6a). Skalering med MAX modellen gir motsatt fordelinger med negativ skjevhet for lave verdier av μ_U , og positiv skjevhet ved høye verdier (Figur 6b). Skalering med OPT modellen gir fordelinger med positiv skjevhet ved både høye og lave verdier av μ_U , og negativ skjevhet når μ_U er nær referanseverdien (Figur 6c). Størrelsen på summeringsforskyvning for ulike verdier av μ_U og antall indikatorer som inngår i summen, ser en som differansen mellom de fargete og de svarte linjene i figurene 2a, 3a og 4a.

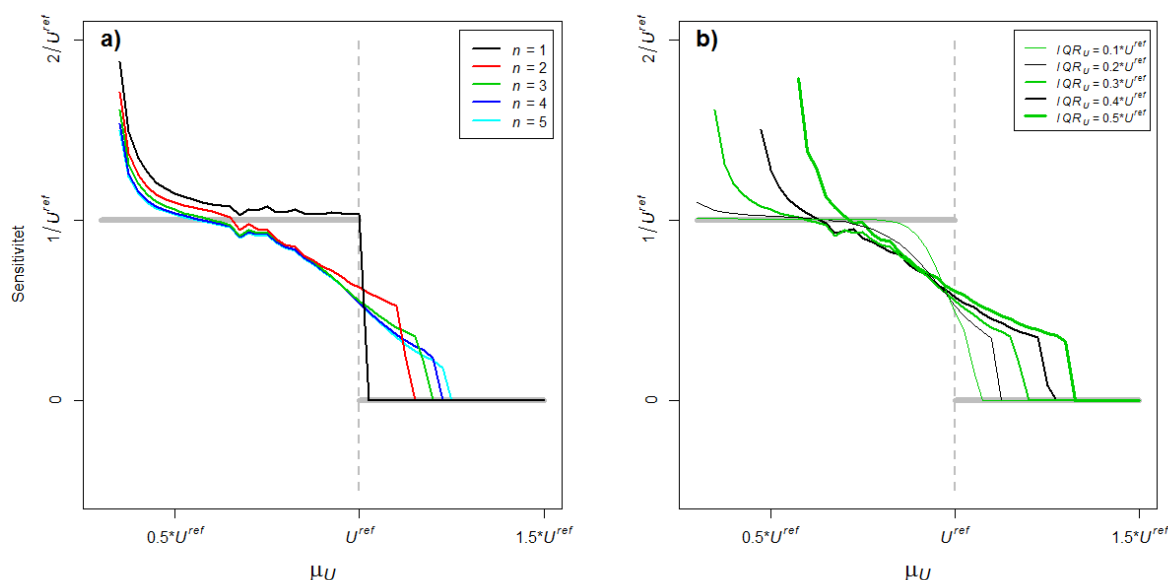


Figur 6. Skjevhet i fordelingene til skalerte biodiversitetsindikatorer, γ_S , som funksjon av log-normalfordelte indikatorers forventningsverdier (μ_U) og kvartilbredde (IQR_U) før skalering med hhv. **a)** LOW-, **b)** MAX- og **c)** OPT modellene.

4 Naturindeksens sensitivitet med hensyn til endring i indikatorene

4.1 Metoder

Sensitivitet har flere, nært beslektede betydninger. Med sensitivitet i forhold til endringer i indikatorenes verdier mener vi her hvor følsom naturindeksen er i forhold til endringer i en, flere, eller alle indikatorenes forventningsverdier. Mer formelt kan sensitivitet i denne betydningen defineres som den retningsderivate av medianen i naturindeksens samplingfordeling, $m_{NI}'(\mu_U; \mathbf{y})$, med hensyn til indikatoren(e)s forventningsverdier (μ_U) når disse endres i retningen gitt av vektoren \mathbf{y} , hvor \mathbf{y} bestemmes av den relative størrelsen på og fortegnet til endringene i de ulike indikatorene.



Figur 7. Naturindeksens sensitivitet med hensyn til endring i indikatorenes forventningsverdi, her definert som $m_{NI}'(\mu_U; \mathbf{y})$, som funksjon av indikatorenes forventningsverdier, μ_U , kvartilbredde, IQR_U , samt antall indikatorer, n . Dette for spesialtilfeller hvor indeksen beregnes som gjennomsnittet av n skalerte biodiversitetsindikatorer, indikatorene har samme referanseverdi og samme lognormal fordeling før skalering, alle skales med LOW-modellen og alle indikatorene endres like mye ($\mathbf{y} = \mathbf{1}$). I **a**) er naturindeksen beregnet fra $n = 1 \dots 5$ indikatorer der deres fordeling har kvartilbredde lik 0,3 ganger referanseverdien U^{ref} . I **b**) er indeksen beregnet fra 3 indikatorer der IQR_U varierer fra 0,1 til 0,5 ganger U^{ref} . Sensitivitet er beregnet numerisk slik som beskrevet i kapittel 4.1. Horisontale, grå, heltrukne, tykke linjer representerer naturindeksens sensitivitet med hensyn til endringer i indikatorverdiene når det ikke er usikkerhet i dataene.

Vi ønsket å beskrive hvordan sensitiviteten varierer med forventningsverdiene til indikatorene, usikkerheten i datagrunnlaget, valg av skaleringsmodell og antall biodiversitets-indikatorer som inngår i beregningen av naturindeksen. Som ved studien av lokasjonsforskyvning studerte vi derfor også her spesialtilfeller hvor alle indikatorene hadde samme fordeling og skaleringsmodell, ble veid likt og endret like mye ($\mathbf{y} = \mathbf{1}$). I slike tilfeller vil naturindeksen ha samme sensitivitet i forhold til endringer i alle indikatorene, og sensitiviteten vil være n ganger større enn

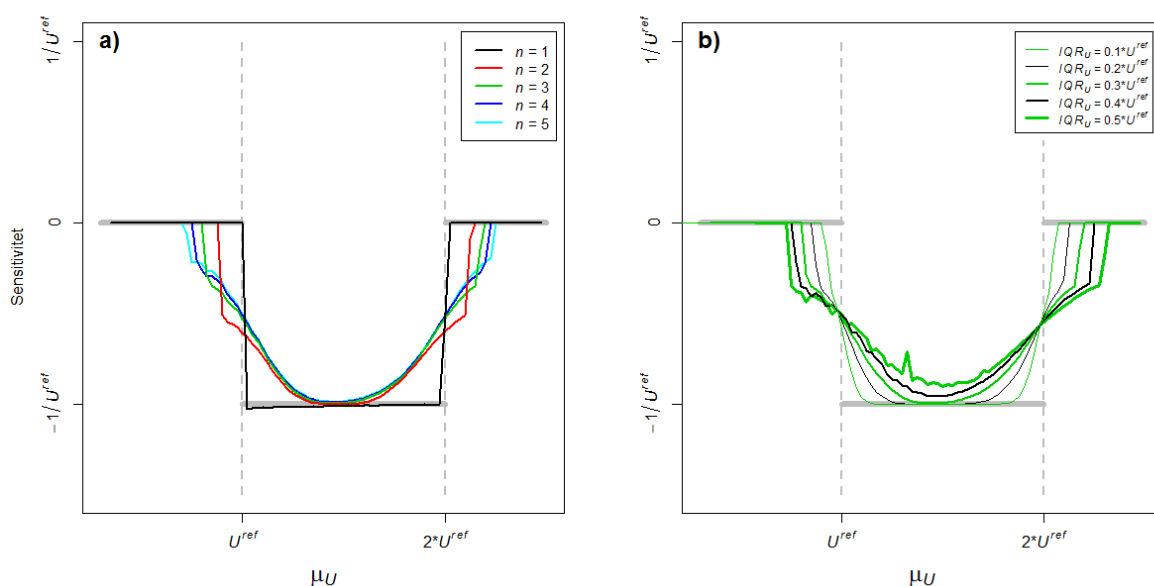
sensitiviteten med hensyn til endring i bare en av indikatorene, der n er antall indikatorer (likning 43, vedlegg 8.4).

Videre, hvis det ikke er usikkerhet i datagrunnlaget, vil sensitivitet i våre spesialtilfeller være lik stigningstallet til skaleringsfunksjonene i den aktuelle indikatorverdien (likning 44), uavhengig hvor mange indikatorer som inngår i beregningen av naturindeksen. Hvordan stigningstallet varierer i de ulike skaleringsmodellene er beskrevet i kapittel 2.1. Vi kunne dermed sammenlikne sensitiviteten med stigningstallet for å se på hvilken effekt usikkerhet i datagrunnlaget har på naturindeksens følsomhet overfor endringer i indikatorene.

Vi beregnet naturindeksens sensitivitet numerisk som $\frac{\Delta m_{NI}}{\Delta \mu}$ (se likning 45), der $\Delta \mu = 0.025 U^{ref}$ var endring i indikatorenes forventningsverdi, en endring som var lik for alle indikatorer, $\Delta m_{NI} = m_{NI}(\mu_U + \Delta \mu \mathbf{1}) - m_{NI}(\mu_U)$, og m_{NI} ble regnet ut ved hjelp av metodene beskrevet i kapittel 3.1.4.

4.2 Resultat

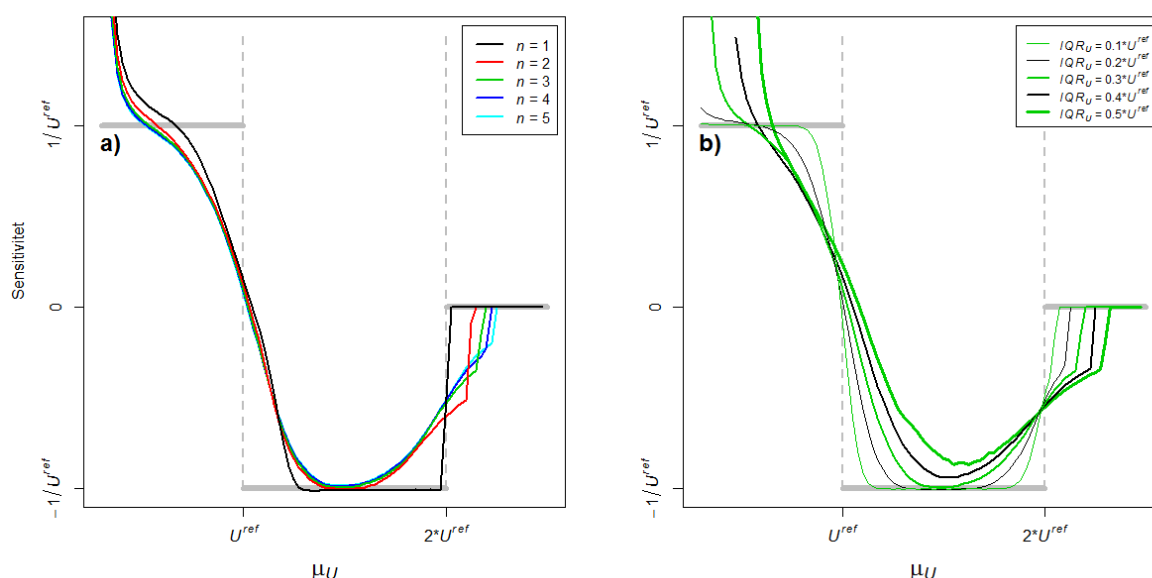
De numeriske beregningene av naturindeksens sensitivitet viste at usikkerhet i datagrunnlaget påvirker hvor mye indeksen endres som følge av endringer i indikatorenes forventningsverdier. Effekten av usikkerhet varierte med skaleringsmodell, indikatorenes forventningsverdier før endring, spredningen i deres fordelinger og antall indikatorer som inngår i indeksen (Figurene 7 - 9). Som gitt av definisjonen av sensitivitet var effekten av usikkerhet i dataene relatert til variasjon i lokasjonsforskyvningen som ble beskrevet i kapittel 3. Betingelser som gir økende positiv (eller avtagende negativ) forskyvning medfører en positiv effekt av usikkerhet på sensitivitet, dvs. betingelsene medfører at sensitiviteten er større når det er usikkerhet i datagrunnlaget enn uten usikkerhet. I slike tilfeller er endringene i indeksens median større enn det endringene i indikatorene skulle tilsi. Betingelser som gir økende negativ (eller avtagende positiv) forskyvning medfører en negativ effekt av usikkerhet på sensitivitet. Dvs. betingelsene medfører at sensitiviteten er mindre med usikkerhet i datagrunnlaget enn uten slik usikkerhet (sammenlikn figurene 2 - 4 og 7 - 9). Endringene i indeksens median er mindre enn det endringene i indikatorene skulle tilsi.



Figur 8. Naturindeksens sensitivitet med hensyn til endring i indikatorenes forventningsverdi, $m_{NI}'(\mu_U; \mathbf{y})$, når indikatorene skales med MAX-modellen. Ellers som figur 7.

Figurene 7 - 9 viser for de ulike skaleringsmodellene hvilke intervaller av indikatorverdier hvor naturindeksen er sensitiv for endringer i indikatorene når det ikke er usikkerhet i dataene (se også Figur 1). Med usikkerhet i dataene blir sensitiviteten i disse intervallene stadig mindre i absoluttverdi når indikatorenes forventningsverdier nærmer seg referanseverdiene. Dette gjelder for LOW og MAX for to eller flere indikatorer og for OPT modellen uansett antall indikatorer naturindeksen bygger på (Figur 7 - 9). Det betyr at når det er usikkerhet i dataene vil positive endringer i tilstanden til biologisk mangfold i stadig mindre grad reflekteres i positive endringer i naturindeksen jo nærmere tilstanden i utgangspunktet er referansetilstanden. Nær referanseverdien er sensitiviteten under LOW og MAX ca. halvparten så stor med usikkerhet i dataene som uten usikkerhet, mens under OPT modellen er sensitiviteten tilnærmet null nær referanseverdien.

Det motsatte gjelder i de intervaller der naturindeksen ikke er sensitiv når det ikke er målefeil i dataene. Med målefeil øker indeksens sensitivitet i absoluttverdi når indikatorenes forventningsverdier nærmer seg referansetilstanden. Ved svært stor usikkerhet i dataene, opptrer mer komplekse sammenhenger mellom sensitivitet og indikatorenes forventningsverdier under MAX- og OPT modellene (se vedlegg 8.4.5).



Figur 9. Naturindeksens sensitivitet med hensyn til endring i indikatorenes forventningsverdi, $m_{NI}'(\mu_U; \mathbf{y})$, når indikatorene skales med OPT-modellen. Ellers som figur 7.

Effekten av økende usikkerhet i dataene på naturindeksens sensitivitet avhenger av skaleringsmodell og indikatorenes forventningsverdier. Under LOW- og OPT modellene øker naturindeksens sensitivitet med økende usikkerhet i dataene for lave indikatorverdier. Sensitiviteten øker også for indikatorverdier større enn referanseverdien under LOW modellen og for verdier mindre enn referanseverdien under MAX modellen. For MAX og OPT avtar sensitivitetens absoluttverdi med økende usikkerhet for indikatorverdier mellom U^{ref} og $2U^{ref}$, mens den øker med usikkerheten for forventningsverdier større enn $2U^{ref}$ (Figur 7b, 8b, 9b).

5 Naturindeksens sensitivitet med hensyn til endringer og usikkerhet i referanseverdiene

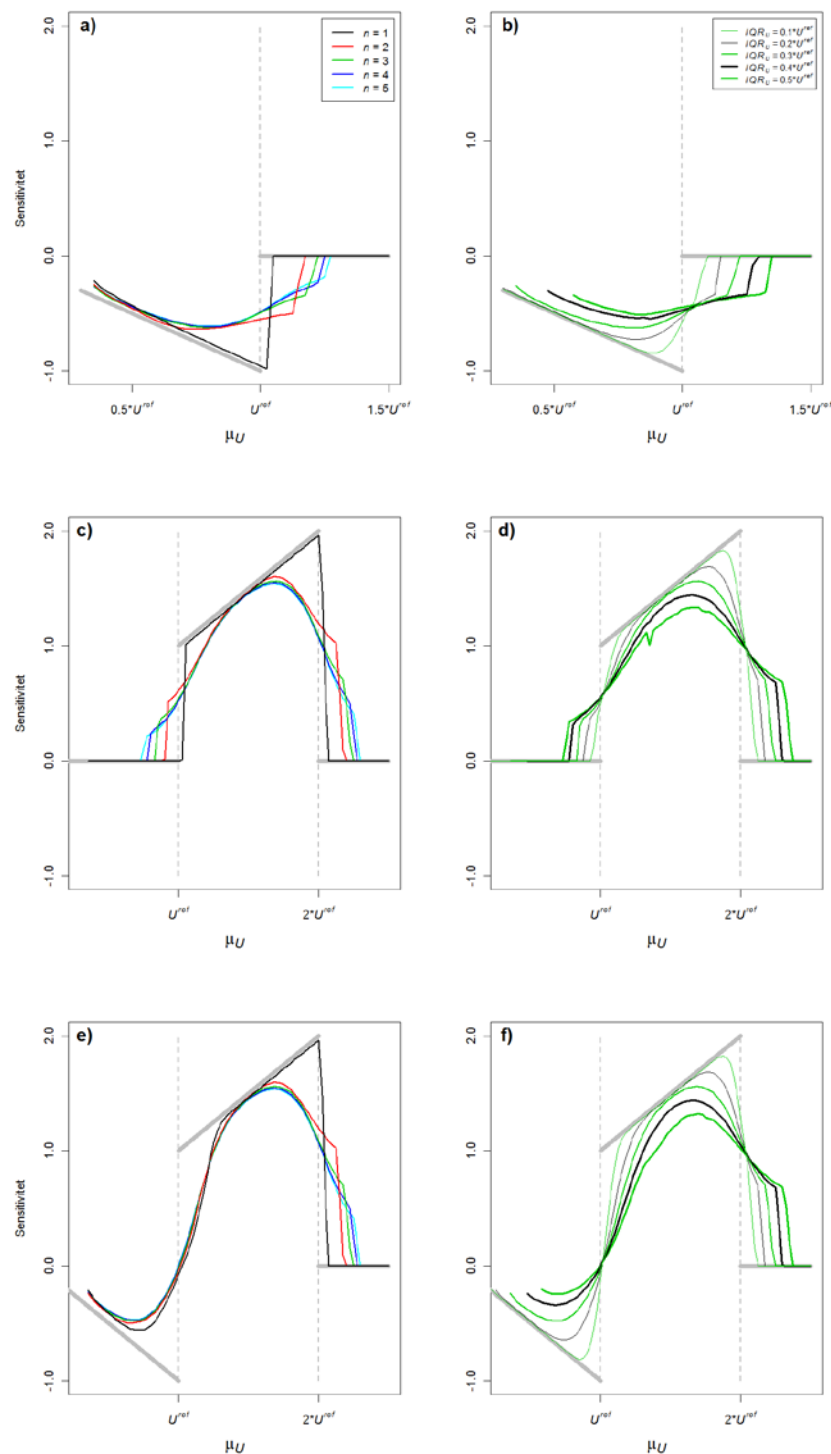
5.1 Metoder

Slik som i kapittel 4 ønsket vi å studere hvordan naturindeksens sensitivitet i forhold til endringer i referanseverdiene varierer med forventningsverdiene til indikatorene, usikkerheten i datagrunnlaget, valg av skaleringsmodell og antall biodiversitets-indikatorer som inngår i beregningen av naturindeksen. Vi studerte derfor de samme spesialtilfellene som beskrevet i kapittel 3.1.3. Som mål på sensitivitet benyttet vi den retningsderiverte av naturindeksens median med hensyn til (den naturlige) logaritmen av indikatorenes referanseverdier, $m_{NI}'(\mathbf{LU}^{ref}; \mathbf{y}), \mathbf{LU}^{ref} = [\ln(U_1^{ref}), \dots, \ln(U_n^{ref})]$. Vi benyttet logaritmeforandringen for å forenkle sammenlikningen med naturindeksens sensitivitet når det ikke er usikkerhet i datagrunnlaget (Figur 10). Vi analyserte tilfellene der referanseverdiene ble endret like mye ($\mathbf{y} = \mathbf{1}$), og med tilsvarende motivasjon som i kapittel 4.1 beregnet vi sensitivitet numerisk som $\frac{\Delta m_{NI}}{\Delta \ln U^{ref}}$, der $\Delta \ln U^{ref} = 0,0025 U^{ref}$ var endringen i de ln-transformerte referanseverdiene, $\Delta m_{NI} = m_{NI}(e^{\Delta \ln U^{ref}} \mathbf{U}^{ref}) - m_{NI}(\mathbf{U}^{ref})$, og m_{NI} ble regnet ut ved hjelp av metodene beskrevet i kapittel 3.1.4.

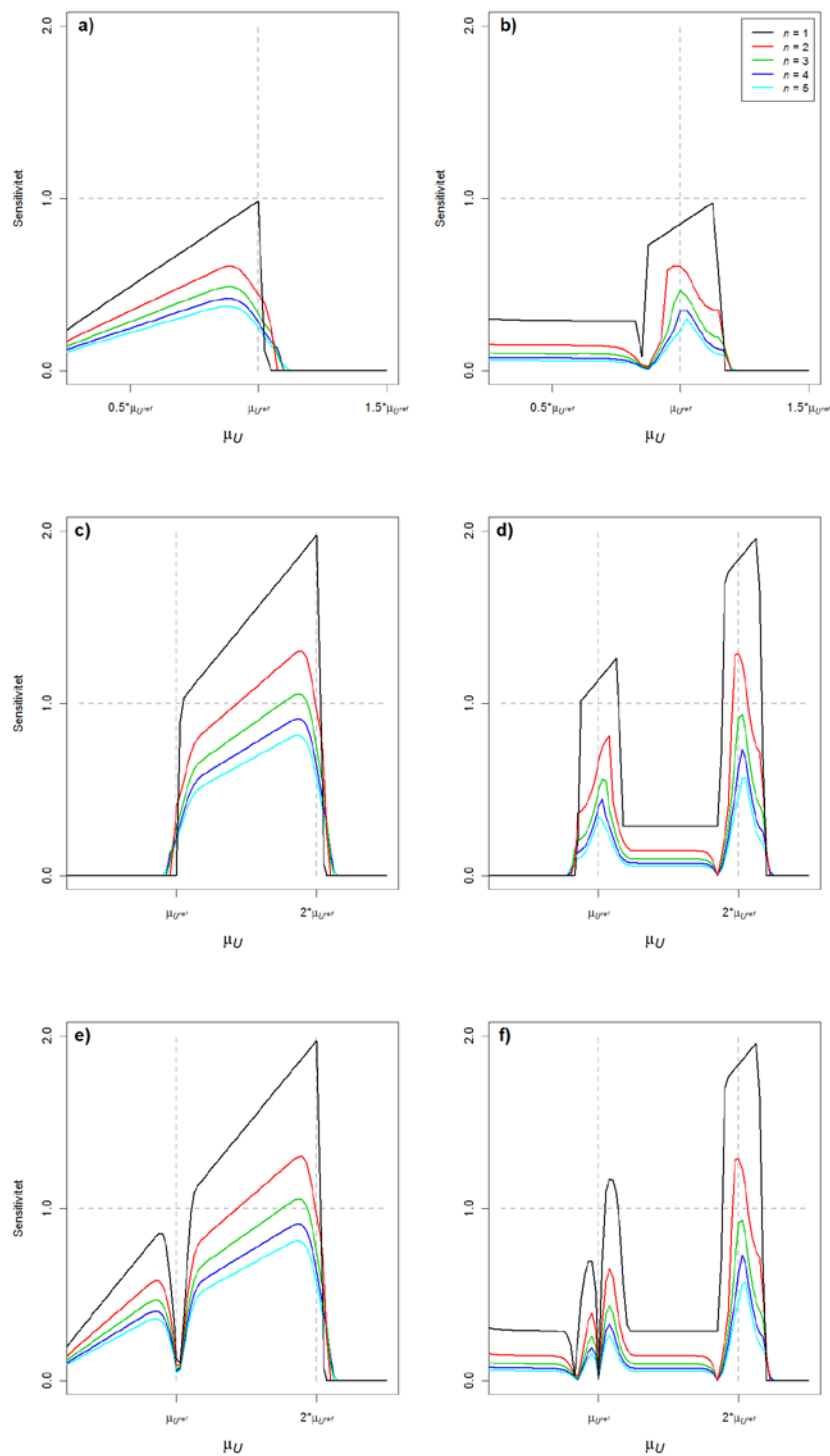
Vi ønsket også å studere hvor robust naturindeksen, slik den i dag beregnes og estimeres, var i forhold til usikkerhet/målefeil i referanseverdiene. Dette for de samme spesialtilfellene som beskrevet tidligere. Som for biodiversitetsindikatorene modellerte vi slik usikkerhet ved å betrakte referanseverdiene som stokastiske variabler med sannsynlighetsfordelinger. Vi antok at usikkerheten var den samme i alle referanseverdiene. Vi ønsket med andre ord å sammenlikne den resulterende spredningen til parameterne i naturindeksens fordeling, dvs. medianen, m_{NI} , og 95%-konfidensintervallet $CI_{NI,0,95} = q_{NI,0,975} - q_{NI,0,025}$, med spredningen i fordelingene til referanseverdiene. Som mål for sensitivitet i denne sammenheng benyttet vi derfor hhv. $\frac{\sigma_{m_{NI}}}{cv_{U^{ref}}}$ og $\frac{\sigma_{CI_{NI}}}{cv_{U^{ref}}}$, der $cv_{U^{ref}}$ er referanseverdienes variasjonskoeffisient og $\sigma_{m_{NI}}, \sigma_{CI_{NI}}$ er standardavviket i hhv. naturindeksens median og konfidensintervall når fordelingene til de n biodiversitetsindikatorene er gitt. I vedlegg 8.4.6 viser vi at for små verdier for $cv_{U^{ref}}$ gjelder tilnærmelsen $\frac{\sigma_{m_{NI}}}{cv_{U^{ref}}} \approx \frac{1}{\sqrt{n}} |m_{NI}'(\mathbf{LU}_1^{ref}; \mathbf{1})|$. Tilsvarende kan også vises for konfidensintervallet. Vi beregnet derfor medianens og konfidensintervallets sensitivitet numerisk som hhv $\frac{1}{\sqrt{n}} \left| \frac{\Delta m_{NI}}{\Delta \ln U^{ref}} \right|$ og $\frac{1}{\sqrt{n}} \left| \frac{\Delta CI_{0,95}}{\Delta \ln U^{ref}} \right|$.

5.2 Resultat

Naturindeksens sensitivitet med hensyn til endringer og feil i referanseverdiene øker i absoluttverdi med økende indikatorverdier når det ikke er usikkerhet i datagrunnlaget. Dette gjelder de intervaller hvor indeksen er sensitiv i forhold til endringer i indikatorene (Figur 10). Dvs. at indeksen generelt er mer sensitiv for endringer i referanseverdier under MAX modellen enn LOW modellen. Positive endringer i referanseverdiene har en negativ effekt på indeksen i intervaller der naturindeksen er en voksende funksjon av indikatorene, og motsatt en positiv effekt der naturindeksen er en avtagende funksjon av indikatorene.



Figur 10. Naturindeksens sensitivitet med hensyn til endringer i referanseverdiene, U^{ref} , her definert som $m_{NI}'(\mathbf{L}U^{ref}; \mathbf{y})$, som funksjon av indikatorenes forventningsverdier, μ_U , kvartilbredde, IQR_U , samt antall indikatorer, n , når indeksen beregnes fra indikatorer skalert med LOW- (a, b), MAX- (c, d) og OPT modellene (e, f). Dette for de samme spesialtilfeller som i Figurene 7 – 10, dvs. $\mathbf{y} = \mathbf{1}$. I a), c), e) er $IQR_U = 0,3 * U^{ref}$, mens i b), d), f) er indeksen beregnet fra 3 indikatorer. Sensitivitet er beregnet numerisk slik som beskrevet i kapittel 5.1. Horisontale, grå, heltrukne, tykke linjer representerer naturindeksens sensitivitet med hensyn til endringer i referanseverdiene når det ikke er usikkerhet i dataene.



Figur 11. Sensitivitet hos naturindeksens median, $\frac{\sigma_{m_{NI}}}{cv_{U^{ref}}}$, **(a, c, e)** og 95% konfidensintervall $\frac{\sigma_{CI_{NI}}}{cv_{U^{ref}}}$, **(b, d, f)** med hensyn til feil i referanseverdiene, som funksjon av indikatorernes forventningsverdier, μ_U , og antall indikatorer, n , når indeksen beregnes fra indikatorer skalert med LOW- **(a, b)**, MAX- **(c, d)** og OPT modellene **(e, f)**. Dette for spesialtilfeller hvor indeksen beregnes som gjennomsnittet av n skalerte biodiversitetsindikatorer med samme lognormal fordeling med kvartilbredde lik $0,1 * U^{ref}$ før skalering, alle indikatorene skaleres med samme modell, og referanseverdiene har samme feilfordeling.

Økende usikkerhet i datagrunnlaget reduserer sensitivitetsens absoluttverdi i de samme intervallene. Denne reduksjonen er størst når indikatorene er nær referanseverdiene (og nær to ganger referanseverdiene under MAX- og OPT modellene). Økende usikkerhet øker indeksens sensitivitet i intervallene der indeksen i utgangspunktet ikke er følsom for endringer i indikatorverdiene (Figur 10). Liknende effekter fås ved å øke antall indikatorer i beregningen av indeksen.

Naturindeksen er generelt robust overfor feil i referanseverdiene. Standardavviket i indeksens median er generelt sett mindre enn referanseverdiens variasjonskoeffisient (Figur 11). Dette gjelder spesielt LOW modellen som benyttes til å skalere de fleste indikatorene. Dette gjelder i enda større grad 95% konfidensintervallet (Figur 11) med unntak når indikatorverdiene og referanseverdiene er om lag like (og når indikatorverdiene er nær to ganger referanseverdiene under MAX- og OPT modellene).

6 Diskusjon og anbefalinger

Målinger av de ulike indikatorene er forbundet med usikkerhet (Certain og Skarpaas 2010, Certain *et al.* 2011). Som forklart i Vedlegg 8.1.1 er standard metode for å handtere slik usikkerhet å betrakte hver observasjon av biodiversitetsindikatorene som en stokastisk variabel med en sannsynlighetsfordeling. Dette betyr videre at også observasjoner av skalerte indikatorer og naturindeksen er forbundet med feil og betraktes som stokastiske variabler. Skaleringsmodellene er ikke-lineære funksjoner av biodiversitetsindikatorene (likning 4, Figur 1). Vi viser i vedlegg 8.2 at skalerte observasjoner av indikatorene kan ha fordelinger som er svært forskjellige fra observasjonenes fordeling før skalering. Fordelingene er blandet (vedlegg 8.1.3) og har gjerne enten positiv eller negativ skjevhet (vedlegg 8.1.6). Naturindeksen er et veid gjennomsnitt av skalerte biodiversitetsindikatorer. I vedlegg 8.3 viser vi at observasjoner av naturindeksen også har en blandet fordeling, og demonstrerer gjennom eksempler at fordelingene gjerne blir mer symmetriske etterhvert som antall indikatorer øker.

Naturindeksen er i Nybø (red., 2010a) estimert som medianen i indeksens samplingfordeling (Certain og Skarpaas 2010, Certain *et al.* 2011). Usikkerhet i datagrunnlaget må nødvendigvis medføre at alle slike estimat også er forbundet med usikkerhet. Usikkerheten i naturindeksen er i Nybø (red., 2010a) beskrevet vha. et 95%-konfidensintervall for en tilfeldig observasjon fra samplingfordelingen. Våre analyser viser imidlertid at usikkerheten i datagrunnlaget i tillegg påvirker medianen i fordelingen. Det vil si at estimatet av naturindeksen vil kunne endres både når observasjonene (de oppgitte forventningsverdiene) av de ulike indikatorene endres, men også når usikkerheten (størrelsen på de oppgitte konfidensintervallene) i dem endres.

Effekten av usikkerhet på medianen har vi her kalt lokasjonsforskyvning. Effekten er utilsiktet og oppstår ved beregning/estimering av naturindeksen. Vi har vist at lokasjonsforskyvning skyldes tre forhold: det er positiv skjevhet i indikatorenes feilfordelinger før skalering, OPT modellen er ikke en monoton funksjon av indikatorene, og det er en summeringsforskyvning ved beregning av naturindeksen fra observasjoner av skalerte indikatorer med skjeve fordelinger.

Vår analyse av hypotetiske spesialtilfeller har vist at størrelsen på lokasjonsforskyvningen avhenger av hvilken skaleringsmodell som benyttes, indikatorenes forventningsverdier før skalering, spredningen i deres fordelinger og antall indikatorer som utgjør grunnlaget for indeksen. Lokasjonsforskyvning øker svakt med antall indikatorer, men nærmer seg raskt en grense. Forskyvningen øker naturlig nok med økende usikkerhet og er større jo nærmere indikatorene er referanseverdiene (Figurene 2 – 4). De hypotetiske tilfellene vi har analysert skiller seg imidlertid fra beregningene av naturindeksen basert på ekspertenes input data. Naturindeksen vil vanligvis bli beregnet fra observasjoner med ulike forventningsverdier og usikkerhet, og bli beregnet fra indikatorer som skaleres med ulike modeller. I disse tilfellene vil både positive og negative forskyvninger kunne forekomme ved skalering av indikatorverdiene, og forskyvningene vil dermed delvis kunne utjevne hverandre ved summering av de skalerte indikatorene til en indeks. Det er så langt ikke gjort analyser for å beregne den potensielle effekten av lokasjonsforskyvning på indeksverdier beregnet fra ekspertenes input data. Men det virker rimelig å anta at lokasjonsforskyvning medfører en underestimering av naturindeksen når estimatene nærmer seg maksimalverdien 1. Slike verdier for naturindeksen forutsetter at de fleste indikatorene har verdier nær referanseverdiene. For slike indeksverdier vil lokasjonsforskyvningen være negativ og av en viss størrelse. Vi anbefaler derfor at det utvikles prosedyrer for estimering av naturindeksen som korrigerer for lokasjonsforskyvning.

Når data oppgis uten målefeil, avhenger naturindeksens følsomhet overfor endringer i indikatorverdier av om verdiene er større eller mindre enn indikatorenes referanseverdi, skaleringskonstanten (= den inverse av referanseverdien) og vekten (w_i) gitt indikatoren (jfr. likningene 1 og 4, figur 1). Vi demonstrerer i kapittel 4 at usikkerhet i datagrunnlaget gjennom lokasjonsforskyvning påvirker medianen i fordelingen til naturindeksen sin følsomhet for endringer i biodiversitetsindikatorenes forventningsverdier. Vi har vist at usikkerhet i dataene både kan øke og

redusere følsomheten til medianen sammenliknet med situasjonen uten målefeil i datagrunnlaget. Som vist i kapittel 4 avhenger dette av skaleringsmodell, indikatorenes forventningsverdier før endring, spredningen i deres fordelinger og antall indikatorer som inngår i indeksen.

Vi foreslår at det tas utgangspunkt i forventningsverdien i naturindeksens samplingfordeling i stedet for medianen når en skal utvikle estimatorer som korrigerer for lokasjonsforskyvning. Riktignok viser vi i vedlegg 8.4.2 og 8.4.3 at lokasjonsforskyvning forekommer både i medianen og forventningsverdien. Skalering medfører forskyvning av forventningsverdien til de skalerte indikatormålingene, mens summering av de skalerte målingene medfører en summeringsforskyvning av medianen til naturindeksen. Imidlertid oppgis indikatormålingenes sentraltendens som forventningsverdier. Dessuten kan eksplisitte uttrykk for størrelsen av skaleringsforskyvning i forventningsverdien utledes for de ulike skaleringsmodellene fra likningene 8, 13 og 17 i vedlegg 8.2. Det synes derfor langt enklere å utvikle forventningsrette estimatorer som korrigerer for lokasjonsforskyvningen i deres forventningsverdi enn å utvikle medianrette estimatorer for naturindeksen.

OPT modellen skiller seg fra de andre to skaleringsmodellene på to måter. OPT er en unimodal og ikke en monoton funksjon av biodiversitetsindikatorene. Dessuten er det i OPT modellen kun referanseverdien som tilsvarer maksimalverdien 1 for skalerte indikatorer. For de andre to modellene tilsvarer alle verdier større og lik (LOW) eller mindre og lik (MAX) referanseverdien denne maksimalverdien. Disse forholdene medfører at lokasjonsforskyvning er større under OPT modellen enn under LOW og MAX når indikatorenes verdier er nær referanseverdiene. Derfor reduseres også indikatorenes maksimalverdier etter skalering under OPT modellen. Lokasjonsforskyvning medfører videre at naturindeksen ikke er sensitiv overfor endringer i OPT-skalerte indikatorer når deres forventningsverdi er nær referanseverdien. Rutiner som korrigerer for lokasjonsforskyvning vil delvis kunne rette på dette, men ikke på det forhold at kun referanseverdien til indikatoren tilsvarer verdien 1 etter skalering.

Av indikatorene som lå til grunn for den første publikasjonen av Naturindeks for Norge, ble 59 skalert med OPT modellen (Nybo (red.) 2010a). Tetthetene av elg (*Alces alces*), hjort (*Cervus elaphus*) og rådyr (*Capreolus capreolus*) er tre slike indikatorer som måler tilstanden til skog-økosystemenes biomangfold. Referanseverdiene for de tre indikatorene er de tettheter av de ulike hjortedyrene som norsk natur, etter ekspertenes vurdering, kan tolerere uten at det biologiske mangfoldet forringes (Eide *et al.* 2010). For indikatorverdier større enn referanseverdien representerer indikatorene tilstanden i andre deler av skogens biomangfold enn hjorteviltet, da først og fremst populasjoner av beiteplanter. I denne forstand fungerer tetthetene av hjortevilt som substitutter for tilstandene i plantepopulasjonene. Høyere tettheter enn referansetilstanden har dermed en annen tolkning enn lavere tettheter. Et alternativ til å skalere med OPT modellen kunne derfor være å la hver av de tre representere to indikatorer, en som representerer tilstanden blant primærkonsumentene, og som skaleres med LOW modellen, og en som representerer tilstanden blant primærprodusentene som påvirkes av beiting, og som skaleres med MAX modellen. De to indikatorenes referanseverdier kunne dermed fastsettes uavhengig av hverandre. Indikatorer som benytter MAX modellen er i hovedsak såkalte indirekte indikatorer eller surrogater. Dette er indikatorer som gjerne samvarierer negativt med utviklingen til andre arter som inngår i det aktuelle økosystemet, og som dermed gjenspeiler denne utviklingen. Ideelt sett burde indirekte indikatorer erstattes av de artene som de representerer ved beregning av naturindeksen, men i mange tilfeller mangler tilstrekkelig data til å gjøre dette. Et alternativ i eksemplet over kunne imidlertid være å la tilstanden i populasjonene hos flere av beiteplantene, slik som rogn (*Sorbus aucuparia*), osp (*Populus tremula*) og selje (*Salix caprea*), inngå som egne indikatorer i stedet for å benytte tetthetene av hjortevilt som indirekte indikatorer.

Referanseverdiene er som indikatorverdiene forbundet med målefeil, noe som øker usikkerheten i estimatene av naturindeksen. Våre analyser indikerer imidlertid at indeksen er robust overfor feil i referanseverdiene, spesielt når indeksen beregnes fra mange indikatorer som skaleres med LOW modellen og disse ikke har verdier nær referanseverdiene. Likevel anbefaler vi

at framtidige estimat av usikkerheten i naturindeksens verdi tar hensyn til eventuelle målefeil i referanseverdiene. Dette vil øke indeksens troverdighet. Det bør derfor utvikles rutiner for dette som samtidig ikke introduserer ny lokasjonsforskyvning ved estimering av indeksen.

Ut fra dette gir vi følgende anbefalinger for en revisjon av prosedyrene for å estimere verdien til naturindeksen:

- 1) Det bør utvikles og inkorporeres rutiner for korreksjon av lokasjonsforskyvning ved estimering av naturindeksen.
- 2) En bør satse på å utvikle forventningsrette estimatorer som korrigerer for lokasjonsforskyvningen i estimatorenes forventningsverdi.
- 3) Usikre målinger/observasjoner av indikatorer bør som nå oppgis ved estimat av målingenes forventningsverdier.
- 4) Det bør foretas en gjennomgang av alle indikatorer som skaleres med OPT modellen med tanke på å erstatte hver av disse med to eller flere indikatorer slik som beskrevet over for hjortedyrsindikatorene i skog.
- 5) Det bør utvikles rutiner for estimering av naturindeksens verdi og tilhørende konfidensintervall som tar hensyn til målefeil i referanseverdiene.

7 Referanser

Andreassen, D., Håvie, T., Krogstad, H.E., Nørsett, S.P., og Aasen, J.O. 1981. Numeriske metoder. – Tapir, Trondheim.

Apostol, T.M. 1969. Calculus volume II, Multivariable calculus and linear algebra with applications to differential equations and probability, 2. ed. – John Wiley & Sons, New York.

Certain, G. og Skarpaas, O. 2010. Nature Index: General framework, statistical method and data collection for Norway – NINA Rapport 542. 47 s.

Certain, G., Skarpaas, O., Bjerke, J.-W., Framstad, E., Lindholm, M., Nielsen, J.-E., Norderhaug, A., Oug, E., Pedersen, H.-C., Schartau, A.-K., Storaunet, K.O., Van der Meeren, G.I., Aslaksen, I., Engen, S., Garnåsjordet P.-A., Kvaløy P., Lillegård M, Yoccoz N.G., og Nybø, S. 2011. The Nature Index: A General Framework for Synthesizing Knowledge on the State of Biodiversity. – PLoS ONE 6: e18930.

Eide, N.E., Solberg, E., Brøseth, H., Strand, O., og Framstad, E. 2010. Pattedyr. – I: Nybø, S. (red.). Datagrunnlaget for "Naturindeks i Norge 2010", DN-utredning 4-2010, s. 23 – 30.

Høyland, A. 1979. Sannsynlighetsregning og statistisk metodelære. Del I sannsynlighetsregning. – Tapir, Trondheim.

Løvås, G.G. 2004. Statistikk for universiteter og høyskoler, 2. utg. – Universitetsforlaget, Oslo.

Manly, B.F.J. 2007. Randomization, bootstrap and Monte Carlo methods in biology. - Chapman & Hall, London.

Nybø, S. (red.) 2010a. Naturindeks for Norge 2010. – DN-utredning 3-2010. 164 s.

Nybø, S. (red.) 2010b. Datagrunnlaget for "Naturindeks i Norge 2010". – DN-utredning 4-2010. 145 s.

Nybø, S., Skarpaas, O., Framstad, E. og Kålås, J.A. 2008. Naturindeks for Norge, forslag til rammeverk – NINA Rapport 347. 68 s.

R Development Core Team. 2008. R: A language and environment for statistical computing. – R Foundation for Statistical Computing, Vienna, Austria.

Zar, J.H. 1999. Biostatistical analysis, fourth edition. – Prentice Hall, Upper Saddle River, New Jersey.

Zar, J.H. 2010. Biostatistical analysis, fifth edition. – Pearson Education International, London.

8 Vedlegg

8.1 Sannsynlighetsfordelinger – en kort oversikt

8.1.1 Stokastiske variabler

Begrepet **stokastisk variabel** er sentralt i sannsynlighetsteori og sannsynlighetsregning. En stokastisk variabel er en funksjon som tilordner verdier til alle mulige utfall eller resultater av et stokastisk (tilfeldig) forsøk (Høyland 1979). For eksempel kan en stokastisk variabel beskrive utfallet av et terningkast, og de mulige utfallene er da elementene i mengden $\{1,2,3,4,5,6\}$. De mulige verdiene som en stokastisk variabel kan tilordne enkeltutfall, kalles variabelens **verdiområde**.

Observasjoner, ekspertvurderinger eller målinger av biodiversitetsindikatorer (U) modelleres som stokastiske variable ved beregning av naturindeksen. Dette fordi den enkelte observasjon av en indikator er forbundet med usikkerhet og er dermed muligens feil. Målefeil er forskjellen mellom den observerte verdien og den sanne verdien som vi ikke kjenner. En tenker seg en måling eller observasjon av indikatoren som et stokastisk forsøk som ved gjentatte målinger kan gi forskjellige verdier på grunn av variasjon i målefeil. For de fleste indikatorene er verdiområdet et intervall på den ikke-negative delen av tallinja fra og med 0 og til en maksimalverdi som kan være uendelig.

Størrelser som er funksjoner av (avledes fra) stokastiske variabler, er selv stokastiske variabler. Observasjoner av indikatorer etter skalering (S) og målinger av naturindeksen (NI) er begge funksjoner av stokastiske biodiversitetsindikatorer (jfr. likningene 1 og 4). Begge har verdiområde fra og med 0 til og med 1.

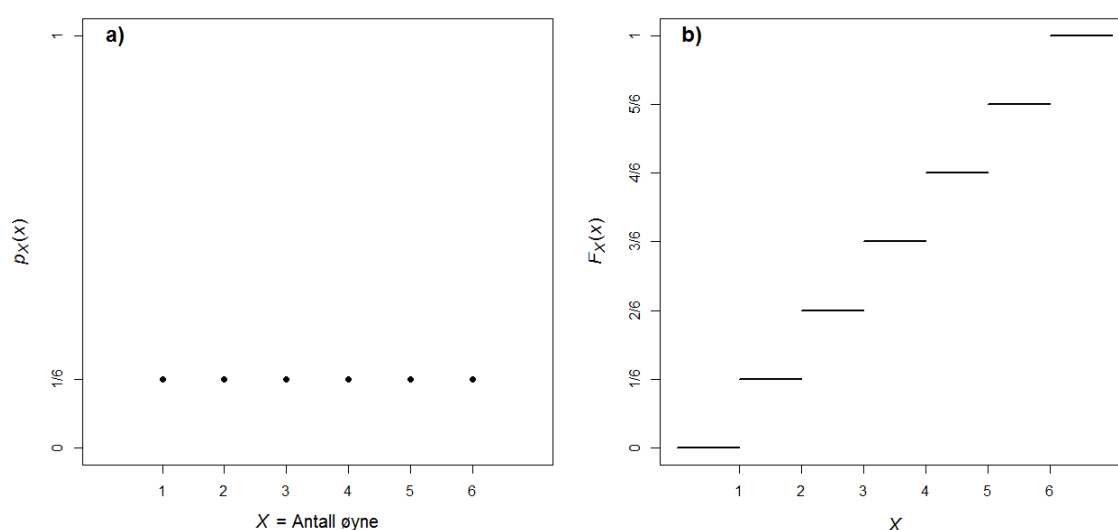
8.1.2 Sannsynlighetsfordeling

Sannsynlighetsfordelingen til en stokastisk variabel beskriver hvordan variabelen fordeler seg i verdiområdet. Hvis X er en stokastisk variabel, vil sannsynlighetsfordelingen tildele en sannsynlighet $P[a \leq X \leq b]$ til intervallet $[a, b]$ som er sannsynligheten for at X tar en verdi i dette intervallet. Fordelingen til X kan entydig beskrives av variabelens **fordelingsfunksjon**, $F_X(x)$, som er definert som $F_X(x) = P[X \leq x]$, der x er et vilkårlig reelt tall, og P er sannsynlighet. Fordelingsfunksjonen er en ikke-avtagende funksjon som oppfyller $F_X(x) \rightarrow 1$ når $x \rightarrow \infty$, og $F_X(x) \rightarrow 0$ når $x \rightarrow -\infty$ (Høyland 1979).

8.1.3 Diskrete, kontinuerlige og blandete sannsynlighetsfordelinger

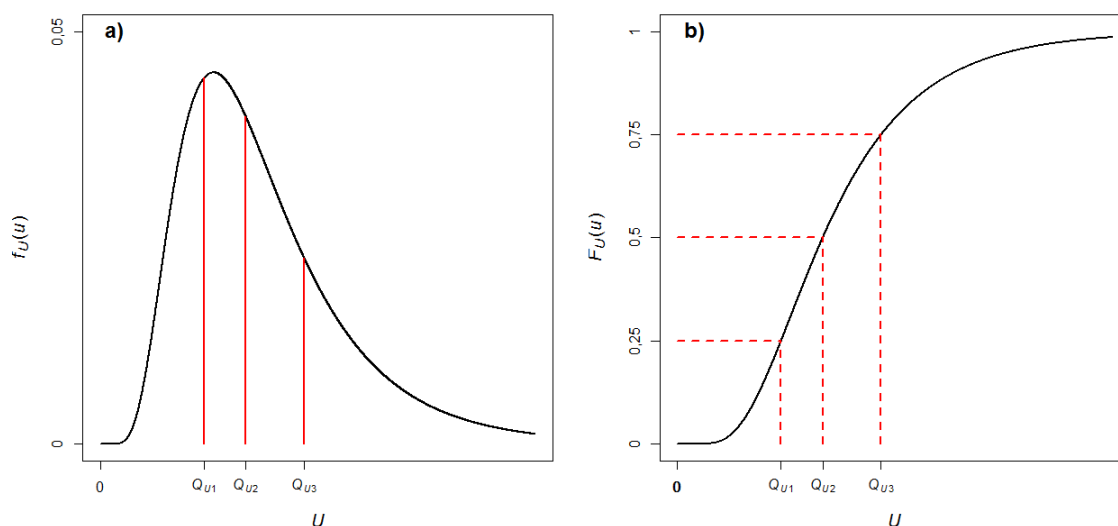
En stokastisk variabel sies å være diskret, og å ha en **diskret fordeling**, dersom verdiområdet høyst består av et nummererbart antall verdier (Høyland 1979). Sannsynlighetsfordelingen beskrives ved for hver mulige verdi, x_i , å angi sannsynligheten for at den stokastiske variabelen er lik verdien $p_X(x_i) = P[X = x_i]$. Funksjonen $p_X(x)$ kalles **punktsannsynligheten** til X . Punktsannsynligheten oppfyller $0 \leq p_X(x_i) \leq 1$ og $\sum_i p_X(x_i) = 1$. Fordelingsfunksjonen til en diskret variabel er en stykkevis kontinuerlig trappefunksjon der sprangene mellom trinnene er gitt av punktsannsynligheten (Figur 12).

En stokastisk variabel X sies å være kontinuerlig, og å ha en **kontinuerlig fordeling**, hvis dens fordelingsfunksjon er kontinuerlig, og i tillegg har en stykkevis kontinuerlig derivert for alle verdier i verdiområdet med unntak av høyst et endelig antall verdier (Høyland 1979). Fordelingsfunksjonens deriverte, $f_X(x)$, kalles **sannsynlighetstettheten** til X . Sannsynlighetstettheten har følgende egenskaper: $F_X(x) = \int_{-\infty}^x f_X(u) du$, $\int_{-\infty}^{\infty} f_X(u) du = 1$ og $f_X(x) \geq 0$ (Figur 13).

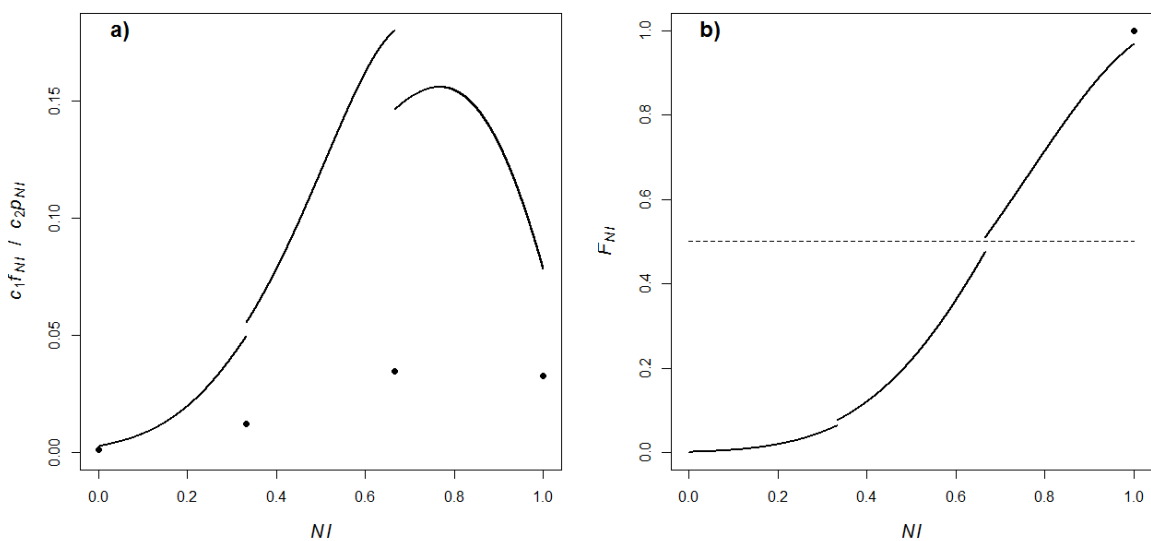


Figur 12. Eksempel på stokastisk variabel med diskret fordeling. **a)** Punktsannsynlighet, p_X , og **b)** fordelingsfunksjon, F_X , til et terningkast, X .

Hvis sannsynlighetsfordelingen til X kan skrives på formen $F_X(x) = c_1 F_1(x) + c_2 F_2(x)$ der $F_1(x)$ beskriver en kontinuerlig fordeling, mens $F_2(x)$ beskriver en diskret, og $c_1 + c_2 = 1$, har X en **blandet fordeling** (Apostol 1969). Den totale sannsynlighetsmassen (= 1) til en blandet variabel består med andre ord av to deler hvor den ene delens fordeling beskrives av en sannsynlighetstetthet og den andre delen av en punktsannsynlighet. Skalerte målinger av biodiversitetsindikatorer og målinger av naturindeksen er eksempler på stokastiske variable med blandede fordelinger (Vedlegg 8.2 og 8.3). Figur 14 viser fordelingen til naturindeksen i et hypotetisk tilfelle der indeksen beregnes fra tre lognormal-fordelte indikatorer som alle er målt med stor usikkerhet og er skalert med MAX modellen. I figur 14a er sannsynlighetsmassen fordelt på en sannsynlighetstetthet og en punktsannsynlighet, mens figur 14b framstiller fordelingsfunksjonen. Fordelingsfunksjonen til blandede fordelinger er stykkevis kontinuerlig.



Figur 13. Eksempel på stokastisk variabel med kontinuerlig fordeling. **a)** Sannsynlighetstetthet, f_U , og **b)** fordelingsfunksjon, F_U , til en kontinuerlig, lognormalfordelt måling av en indikator, U . De røde linjene markerer kvartilene i fordelinga. Medianen, m_X , er lik midtre kvartil, $Q_{U,2}$.



Figur 14. Eksempel på stokastisk variabel med blandet fordeling. **a)** Sannsynlighetstetthet, f_{NI} , punktsannsynlighet, p_{NI} , og **b)** fordelingsfunksjon, F_{NI} , til en måling av naturindeksen, NI , i et hypotetisk tilfelle der indeksen beregnes fra tre indikatorer som alle er målt med stor usikkerhet og er skalert ved MAX modellen. c_1 og c_2 angir andelen av sannsynlighetsmassen som er beskrevet med hhv. f_{NI} og p_{NI} , $c_1 + c_2 = 1$. Stiplet linje representerer $F_{NI} = 0,5$ som «krysser» fordelingsfunksjonen i medianen.

8.1.4 Kvantiler og kvartiler

α -kvantilen til en kontinuerlig stokastisk variabel X er den verdien $q_{X,\alpha}$ i verdiområdet til X som oppfyller at $F_X(q_{X,\alpha}) = \alpha$, der $0 < \alpha < 1$ (Zar 2010). α -kvantilen deler den totale sannsynlighetsmassen i to deler der den ene delen er lik α og angir sannsynligheten for at X er mindre enn $q_{X,\alpha}$. Den andre delen, $1 - \alpha$, er sannsynligheten for at X er større enn verdien $q_{X,\alpha}$.

0,5-kvantilen til X kalles **medianen** ($m_X = q_{X,0.5}$, figur 13). Medianen deler sannsynlighetsmassen i en kontinuerlig fordeling i to like deler. I en diskret fordeling er medianen den verdien m_X som oppfyller både at $P[X \leq m_X] \geq 0,5$ og $P[X \geq m_X] \geq 0,5$. 0,25-, 0,5-, og 0,75-kvantilene kalles **kvartiler** og betegnes $Q_{X,1}$, $Q_{X,2}$, $Q_{X,3}$ (Zar 2010). Sammen deler kvartilene sannsynlighetsmassen i en kontinuerlig fordeling i fire like store deler lik 0,25, hhv. $P[X < Q_{X,1}]$, $P[Q_{X,1} < X < Q_{X,2}]$, $P[Q_{X,2} < X < Q_{X,3}]$ og $P[X > Q_{X,3}]$ (Figur 13).

Ved beregning av naturindeksen inngår kvantilene og kvartilene i fordelingene til målinger av indikatorene og naturindeksen på forskjellig vis. I datagrunnlaget som benyttes til å beregne naturindeksen, er $Q_{U,1}$ og $Q_{U,3}$ estimert for de fleste observasjoner av biodiversitetsindikatorer. Videre angis naturindeksen som et estimat av m_{NI} . Usikkerheten i estimatet av naturindeksen beskrives ved hjelp av 0,025- og 0,975-kvantilene til NI , som sammen angir grensene for et 95%-konfidensintervall til indeksen (Certain og Skarpaas 2010, Certain *et al.* 2011).

8.1.5 Sentralitet og spredning

En sannsynlighetsfordeling kan karakteriseres ved dens plassering på tallinja (sentraltendens), dens spredning og form. Her og i kapittel 8.1.6 omtales ulike størrelser og parametere som beskriver disse egenskapene.

Forventningsverdi og median beskriver begge en fordelings sentraltendens. Forventningen, $E(X)$ eller μ_X , til en stokastisk variabel X , er den verdi som gjennomsnittet av utfallene vil nær-

me seg når en gjentar eksperimentet som ligger til grunn for X , uendelig antall ganger. I data-grunnlaget for naturindeksen inngår estimat av μ_{U_i} for alle indikatorene.

En fordelings varians, standard avvik og **kvartilbredde** er eksempler på statistikker som beskriver fordelings spredning eller utstrekning på tallinja. Kvartilbredden (IQR_X) er differansen mellom øvre og nedre kvartil, $IQR_X = Q_{X,3} - Q_{X,1}$ (Zar 2010) og utgjør et 50%-konfidensinterval for en tilfeldig observasjon av X . IQR_U benyttes i naturindekssammenheng til å beskrive usikkerheten i enkeltobservasjoner av biodiversitetsindikatorer. Derfor har en også i denne rapporten stort sett benyttet kvartilbredden som spredningsmål i stedet for de mer kjente størrelsene varians og standard avvik. **Variansen** til X , σ_X^2 , er forventningen til den kvadrerte avstanden mellom X og forventningen til X , $\sigma_X^2 = E((X - \mu_X)^2)$, **standardavviket**, $\sigma_X = \sqrt{\sigma_X^2}$, måles på samme skala som X , mens **variasjonskoeffisienten**, $cv_X = \frac{\sigma_X}{|\mu_X|}$, er et dimensjonsløst mål for spredning (Løvås 2004).

8.1.6 Fordelingens form – skjevhet

En fordelings form refererer til andre egenskaper enn plassering og utstrekning på tallinja. Et eksempel på slike formegenskaper er **skjevhet**.

Skjevhet angir grad av asymmetri og kan være både positiv og negativ (Zar 1999). Hos en symmetrisk fordeling er sannsynlighetstettheten/punktsannsynligheten aksesympetrisk om medianen (fordelingens "halvdeler" er speilbilder av hverandre), og fordelingsfunksjonen er punktsymmetrisk om punktet $(m_X, 0.5)$. Forventningsverdien til en symmetrisk fordeling er lik medianen. Skjevetskoeffisienten (γ) er en av flere statistikker som er foreslått som mål for skjevhet. Den defineres som $\gamma_X = E\left(\left(\frac{X - \mu_X}{\sigma_X}\right)^3\right)$ (Zar 1999). Negativ skjevhet indikerer rent kvalitativt at fordelings venstre hale er lengre og/eller "tykkere" enn dens høyre, og medianen er ofte, men ikke alltid, større enn forventningsverdien. Positiv skjevhet indikerer at fordelings høyre hale er lengre og/eller "tykkere" enn dens venstre, og medianen er ofte, men ikke alltid, mindre enn forventningsverdien. En symmetrisk fordeling har skjevhet null, men en fordeling med skjevhet null er ikke nødvendigvis symmetrisk (Zar 2010).

8.2 Sannsynlighetsfordelingen til målinger av skalerte indikatorer

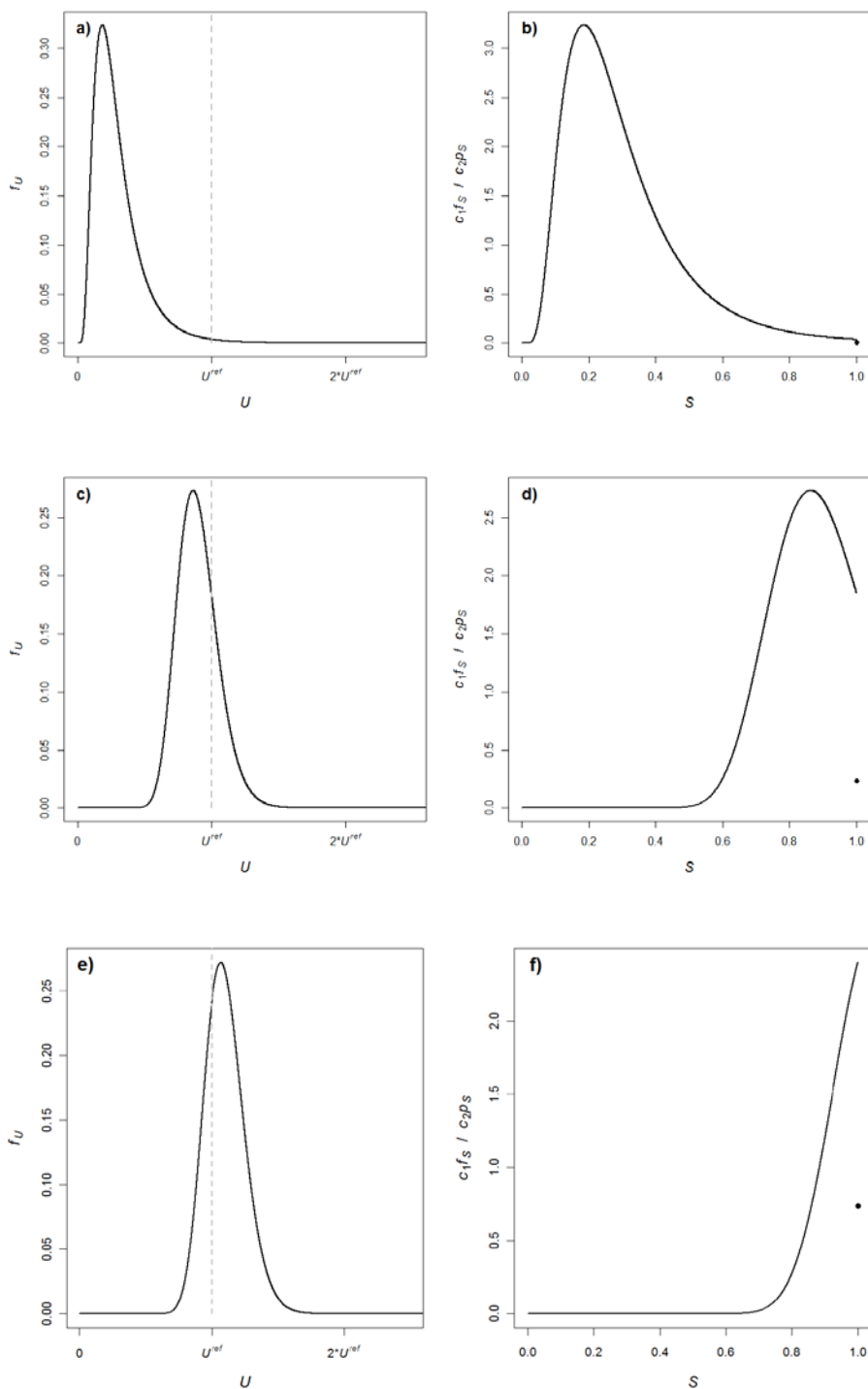
Her studeres feilfordelingen til skalerte målinger av biodiversitetsindikatorer. Vi viser at feilfordelingen er en blandet fordeling under alle skaleringsmodellene. Vi utleder punktsannsynligheten, sannsynlighetstettheten, fordelingsfunksjonen, forventningsverdien og medianen til den skalerte målingen som funksjoner av fordelingen til observasjonen før skalering. Vi starter med å definere funksjonene og parameterne som beskriver fordelingen til målingen av indikatoren.

La $f_U(u)$ være sannsynlighetstettheten til en vilkårlig måling av en kontinuerlig biodiversitetsindikator, U , definert på intervallet $[0, u_{max}]$, $u_{max} \in (0, \infty)$, $F_U(u)$ den tilhørende fordelingsfunksjonen og U^{ref} indikatorens referanseverdi, $0 < U^{ref} < u_{max}$. La videre m_U og μ_U være observasjonens median og forventningsverdi.

8.2.1 Skalering med LOW modellen

La S_{LOW} være den skalerte målingen, $S_{LOW} = S_{LOW}(U; U^{ref})$, gitt ved Likning 4a. S_{LOW} har verdiområde $s \in [0, 1]$ og er en strengt voksende funksjon av U når $U \in [0, U^{ref})$ og lik 1 når $U \in [U^{ref}, u_{max}]$. S_{LOW} har dermed en blandet fordeling gitt ved en punktsannsynlighet med

verdiområde $\{1\}$ og en sannsynlighetstetthet med verdiområde $[0,1]$. La $f_{S_{LOW}}(s)$, $p_{S_{LOW}}(s)$, $F_{S_{LOW}}(s)$ betegne sannsynlighetstetthet, punktsannsynlighet og fordelingsfunksjon for S_{LOW} :



Figur 15. Tre eksempler på sannsynlighetstettheter (f_U) til målinger av biodiversitetsindikatorer (U) a), c), e), og målingenes (S) tettheter (f_S) og punktsannsynligheter (p_S) etter skalering med LOW modellen b), d), f). Målingenes forventningsverdi før skalering er hhv. i a) b) $0,3 \cdot U^{ref}$, i c) d) $0,9 \cdot U^{ref}$, og i e) f) $1,1 \cdot U^{ref}$, der U^{ref} er referanseverdien. Kvartilbredden er i alle tre tilfeller $0,2 \cdot U^{ref}$. c_1 og c_2 angir andelen av sannsynlighetsmassen som er beskrevet med hhv. f_S og p_S , $c_1 + c_2 = 1$.

$$(5a) \quad c_1 f_{S_{LOW}}(s) = U^{ref} * f_U(U^{ref} s), \quad 0 \leq s < 1,$$

$$(5b) \quad c_2 p_{S_{LOW}}(s) = \int_{U^{ref}}^{u_{max}} f_U(u) du = 1 - F_U(U^{ref}), \quad s = 1,$$

$$(6) \quad F_{S_{LOW}}(s) = \begin{cases} c_1 \int_0^s f_{S_{LOW}}(x) dx + c_2 \sum_{y \leq s} p_{S_{LOW}}(y) \\ \int_0^{sU^{ref}} f_U(u) du = F_U(U^{ref} s), & 0 \leq s < 1 \\ c_1 + c_2 = 1, & s = 1 \end{cases}$$

hvor $c_1 = F_U(U^{ref})$ og $c_2 = 1 - F_U(U^{ref})$. Figur 15 viser fordelingene før og etter skalering med LOW modellen for tre ulike målinger av U . Av figuren ser en at fordelings skjevhet etter skalering, γ_S , avhenger bl.a. av plasseringen av fordelingen til U i forhold til referanseverdien. Se også Figur 6a som viser γ_S som funksjon av μ_U og IQR_U .

Medianen til den skalerte målingen $m_{S_{LOW}}$ er lik den skalerte medianen til U etter skalering med LOW modellen, $S_{LOW}(m_U)$, fordi

$$F_{S_{LOW}}(S_{LOW}(m_U)) = F_{S_{LOW}}\left(\frac{m_U}{U^{ref}}\right) = F_U\left(U^{ref} \frac{m_U}{U^{ref}}\right) = F_U(m_U) = 0,5$$

når $m_U < U^{ref}$, og

$$\begin{aligned} P[S_{LOW} \geq S_{LOW}(m_U)] &= P[S_{LOW} \geq 1] = 1 - F_U(U^{ref}) \geq 0,5, \\ P[S_{LOW} \leq S_{LOW}(m_U)] &= P[S_{LOW} \leq 1] = 1 > 0,5 \end{aligned}$$

når $m_U \geq U^{ref}$. Dermed får vi fra definisjonen av median at

$$(7) \quad m_{S_{LOW}} = S_{LOW}(m_U).$$

Tilsvarende relasjon gjelder imidlertid ikke for forventningsverdien til S_{LOW} :

$$(8) \quad \mu_{S_{LOW}} = \begin{cases} \int_0^{u_{max}} S_{LOW}(u) f_U(u) du = \frac{1}{U^{ref}} \int_0^{U^{ref}} u f_U(u) du + \int_{U^{ref}}^{u_{max}} f_U(u) du \\ = \frac{\mu_U}{U^{ref}} - \int_{U^{ref}}^{u_{max}} \left(\frac{u}{U^{ref}} - 1\right) f_U(u) du \\ S_{LOW}(\mu_U) - \int_{U^{ref}}^{u_{max}} \left(\frac{u}{U^{ref}} - 1\right) f_U(u) du, & \mu_U \leq U^{ref} \\ S_{LOW}(\mu_U) - \int_0^{U^{ref}} \left(1 - \frac{u}{U^{ref}}\right) f_U(u) du, & \mu_U > U^{ref} \end{cases}$$

ettersom $S_{LOW}(\mu_U) = 1$ når $\mu_U > U^{ref}$. Dette betyr at

$$(9) \quad \mu_{S_{LOW}} < S_{LOW}(\mu_U).$$

8.2.2 Skalering med MAX modellen

La S_{MAX} være den skalerte målingen $S_{MAX} = S_{MAX}(U; U^{ref})$ gitt ved Likning 4b. S_{MAX} har verdiområde $s \in [0, 1]$, er lik 1 når $u \in [0, U^{ref}]$, er en strengt avtagende funksjon av U når $U \in (U^{ref}, 2 * U^{ref})$, og lik 0 når $U \in [2 * U^{ref}, u_{max}]$. S_{MAX} har dermed en blandet fordeling gitt ved en punktsannsynlighet med verdiområde $\{0, 1\}$ og en sannsynlighetstetthet med verdiområde $(0, 1)$. La $f_{S_{MAX}}(s)$, $p_{S_{MAX}}(s)$, $F_{S_{MAX}}(s)$ være sannsynlighetstetthet, punktsannsynlighet og fordelingsfunksjon for S_{MAX} :

$$(10a) \quad c_1 f_{S_{MAX}}(s) = U^{ref} * f_U(U^{ref}(2-s)), \quad 0 < s < 1,$$

$$(10b) \quad c_2 p_{S_{MAX}}(s) = \begin{cases} \int_{2U^{ref}}^{u_{max}} f_U(u) du = 1 - F_U(2U^{ref}), & s = 0 \\ \int_0^{U^{ref}} f_U(u) du = F_U(U^{ref}), & s = 1 \end{cases},$$

$$(11) \quad F_{S_{MAX}}(s) = c_1 \int_0^s f_{S_{MAX}}(x) dx + c_2 \sum_{y \leq s} p_{S_{MAX}}(y)$$

$$F_{S_{MAX}}(s) = \begin{cases} \int_{U^{ref}(2-s)}^{2U^{ref}} f_U(u) du + 1 - F_U(2U^{ref}) = 1 - F_U(U^{ref}(2-s)), & 0 \leq s < 1 \\ c_1 + c_2 = 1, & s = 1 \end{cases}$$

hvor $c_1 = F_U(2U^{ref}) - F_U(U^{ref})$ og $c_2 = 1 - (F_U(2U^{ref}) - F_U(U^{ref}))$. Figur 16 viser fordelinger før og etter skalering med MAX modellen for tre ulike målinger av en biodiversitetsindikator. Se også Figur 6b som viser γ_S som funksjon av μ_U og IQR_U .

Som for LOW modellen kan en vise at medianen til den skalerte målingen $m_{S_{MAX}}$ er lik den skalerte medianen til U etter skalering med MAX modellen, $S_{MAX}(m_U)$. Når $m_U < U^{ref}$ gjelder

$$P[S_{MAX} \geq S_{MAX}(m_U)] = P[S_{MAX} \geq 1] = F_U(U^{ref}) \geq 0,5,$$

$$P[S_{MAX} \leq S_{MAX}(m_U)] = P[S_{MAX} \leq 1] = 1 > 0,5,$$

når $U^{ref} < m_U < 2 * U^{ref}$ får vi fra likning 11 at

$$F_{S_{MAX}}(S_{MAX}(m_U)) = F_{S_{MAX}}\left(2 - \frac{m_U}{U^{ref}}\right) = 1 - F_U(m_U) = 0,5,$$

og når $m_U \geq 2 * U^{ref}$

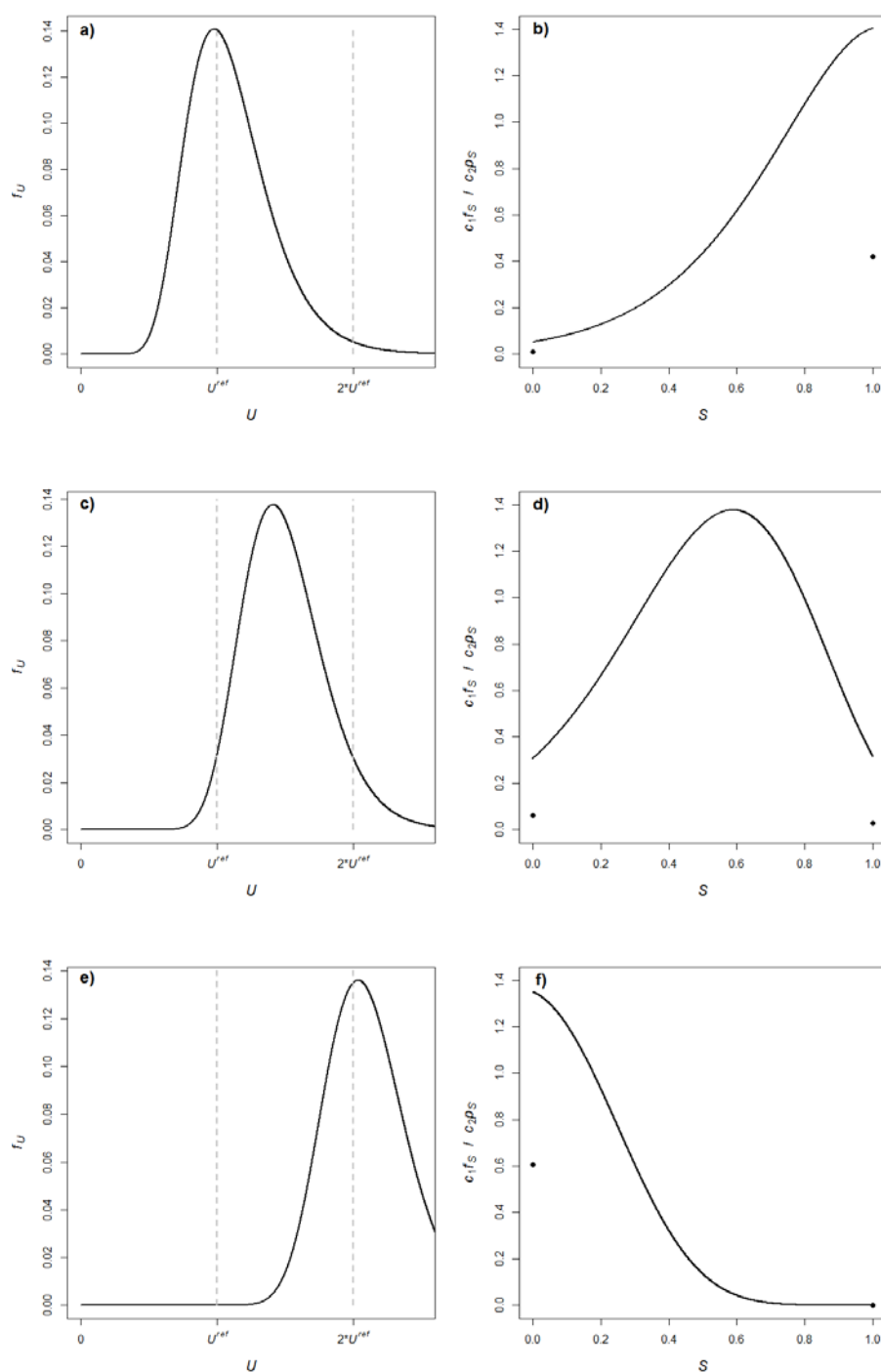
$$P[S_{MAX} \geq S_{MAX}(m_U)] = P[S_{MAX} \geq 0] = 1 > 0,5,$$

$$P[S_{MAX} \leq S_{MAX}(m_U)] = P[S_{MAX} \leq 0] = 1 - F_U(2U^{ref}) \geq 0,5.$$

Dermed kan vi fra definisjonen av median konkludere at

$$(12) \quad m_{S_{MAX}} = S_{MAX}(m_U).$$

Slik som vist for LOW modellen over, gjelder ikke tilsvarende relasjon for forventningsverdien til S_{MAX} .



Figur 16. Tre eksempler på sannsynlighetstettheter (f_U) til målinger av biodiversitetsindikatorer (U) **a)**, **c)**, **e)**, og målingenes (S) tettheter (f_S) og punktsannsynligheter (p_S) etter skalering med MAX modellen **b)**, **d)**, **f)**. Målingenes forventningsverdi før skalering er hhv. i **a)** **b)** $1,1 \cdot U^{ref}$, i **c)** **d)** $1,5 \cdot U^{ref}$, og i **e)** **f)** $2,1 \cdot U^{ref}$, der U^{ref} er referanseverdien. Kvartilbredden er i alle tre tilfeller $0,4 \cdot U^{ref}$. Ellers som figur 15.

$$\begin{aligned}
\mu_{S_{MAX}} &= \int_0^{u_{max}} S_{MAX}(u) f_U(u) du = \int_0^{U^{ref}} f_U(u) du + \frac{1}{U^{ref}} \int_{U^{ref}}^{2U^{ref}} (2U^{ref} - u) f_U(u) du \\
&= \int_0^{2U^{ref}} f_U(u) du + \int_{U^{ref}}^{2U^{ref}} f_U(u) du - \frac{1}{U^{ref}} \int_{U^{ref}}^{2U^{ref}} u f_U(u) du \\
&= 2 - \frac{\mu_U}{U^{ref}} - \int_0^{U^{ref}} \left(1 - \frac{u}{U^{ref}}\right) f_U(u) du - \int_{2U^{ref}}^{u_{max}} \left(2 - \frac{u}{U^{ref}}\right) f_U(u) du
\end{aligned}
\tag{13}$$

$$\mu_{S_{MAX}} = \begin{cases} S_{MAX}(\mu_U) - \int_{U^{ref}}^{2U^{ref}} \left(\frac{u}{U^{ref}} - 1\right) f_U(u) du - \int_{2U^{ref}}^{u_{max}} f_U(u) du, & \mu_U \leq U^{ref} \\ S_{MAX}(\mu_U) - \int_0^{U^{ref}} \left(1 - \frac{u}{U^{ref}}\right) f_U(u) du + \int_{2U^{ref}}^{u_{max}} \left(\frac{u}{U^{ref}} - 2\right) f_U(u) du, & U^{ref} < \mu_U < 2U^{ref} \\ S_{MAX}(\mu_U) + \int_0^{U^{ref}} f_U(u) du + \int_{U^{ref}}^{2U^{ref}} \left(2 - \frac{u}{U^{ref}}\right) f_U(u) du, & \mu_U \geq 2 * U^{ref} \end{cases}$$

ettersom $S_{MAX}(\mu_U) = 1$ når $\mu_U \leq U^{ref}$ og $S_{MAX}(\mu_U) = 0$ når $\mu_U \geq 2 * U^{ref}$. Dette betyr at $\mu_{S_{MAX}} < S_{MAX}(\mu_U)$ når $\mu_U \leq U^{ref}$, og $\mu_{S_{MAX}} > S_{MAX}(\mu_U)$ når $\mu_U \geq 2 * U^{ref}$. Når $U^{ref} < \mu_U < 2 * U^{ref}$ kan $\mu_{S_{MAX}}$ være større enn, lik, eller mindre enn $S_{MAX}(\mu_U)$, avhengig av om integralet $\int_0^{U^{ref}} \left(1 - \frac{u}{U^{ref}}\right) f_U(u) du \geq 0$ er mindre enn, lik, eller større enn integralet $\int_{2U^{ref}}^{u_{max}} \left(\frac{u}{U^{ref}} - 2\right) f_U(u) du \geq 0$.

8.2.3 Skalering med OPT modellen

La S_{OPT} være den skalerte målingen $S_{OPT} = S_{OPT}(U; U^{ref})$ gitt ved Likning 4c. S_{OPT} har verdiområde $s \in [0, 1]$, er en strengt voksende funksjon av U når $U \in [0, U^{ref}]$, en strengt avtagende funksjon av U når $U \in [U^{ref}, 2 * U^{ref}]$, og lik 0 når $U \in [2 * U^{ref}, u_{max}]$. S_{OPT} har dermed en blandet fordeling gitt ved en punktsannsynlighet med verdiområde $\{0\}$ og en sannsynlighetstetthet med verdiområde $[0, 1]$. La $f_{S_{OPT}}(s)$, $p_{S_{OPT}}(s)$, $F_{S_{OPT}}(s)$ være sannsynlighetstetthet, punktsannsynlighet og fordelingsfunksjon for S_{OPT} :

$$(14a) \quad c_1 f_{S_{OPT}}(s) = U^{ref} \left(f_U(U^{ref} s) + f_U(U^{ref} (2 - s)) \right), \quad 0 \leq s \leq 1,$$

$$(14b) \quad c_2 p_{S_{OPT}}(s) = \int_{2U^{ref}}^{u_{max}} f_U(u) du = 1 - F_U(2U^{ref}), \quad s = 0,$$

$$\begin{aligned}
F_{S_{OPT}}(s) &= c_1 \int_0^s f_{S_{OPT}}(x) dx + c_2 \sum_{y \leq s} p_{S_{OPT}}(y) = \int_0^{sU^{ref}} f_U(u) du + \int_{U^{ref}(2-s)}^{2U^{ref}} f_U(u) du + 1 - F_U(2U^{ref}) \\
(15) \quad F_{S_{OPT}}(s) &= 1 + F_U(U^{ref} s) - F_U(U^{ref} (2 - s)) = 1 - \int_{sU^{ref}}^{U^{ref}(2-s)} f_U(u) du, \quad 0 \leq s \leq 1
\end{aligned}$$

hvor $c_1 = F_U(2U^{ref})$ og $c_2 = 1 - F_U(2U^{ref})$. Figur 17 viser fordelinger før og etter skalering med OPT modellen for tre ulike målinger av en indikator. Se også Figur 6c som viser γ_S som funksjon av μ_U og IQR_U .

For OPT modellen er medianen til den skalerte målingen, $m_{S_{OPT}}$, mindre eller lik den skalerte medianen til U etter skalering med OPT modellen, $S_{OPT}(m_U)$. Dette er forskjellig fra LOW og MAX modellene. Når $m_U \leq U^{ref}$ gjelder

$$F_{S_{OPT}}(S_{OPT}(m_U)) = F_{S_{OPT}}\left(\frac{m_U}{U^{ref}}\right) = 1 + F_U(m_U) - F_U(2U^{ref} - m_U) > 0,5,$$

og når $U^{ref} < m_U < 2 * U^{ref}$

$$F_{S_{OPT}}(S_{OPT}(m_U)) = F_{S_{OPT}}\left(2 - \frac{m_U}{U^{ref}}\right) = 1 - F_U(m_U) + F_U(2U^{ref} - m_U) > 0,5.$$

og når $m_U \geq 2 * U^{ref}$

$$\begin{aligned} P[S_{OPT} \geq S_{OPT}(m_U)] &= P[S_{OPT} \geq 0] = 1 > 0,5, \\ P[S_{OPT} \leq S_{OPT}(m_U)] &= P[S_{OPT} \leq 0] = 1 - F_U(2U^{ref}) \geq 0,5. \end{aligned}$$

Dermed får vi at

$$(16a) \quad m_{S_{OPT}} < S_{OPT}(m_U), \quad m_U < 2 * U^{ref},$$

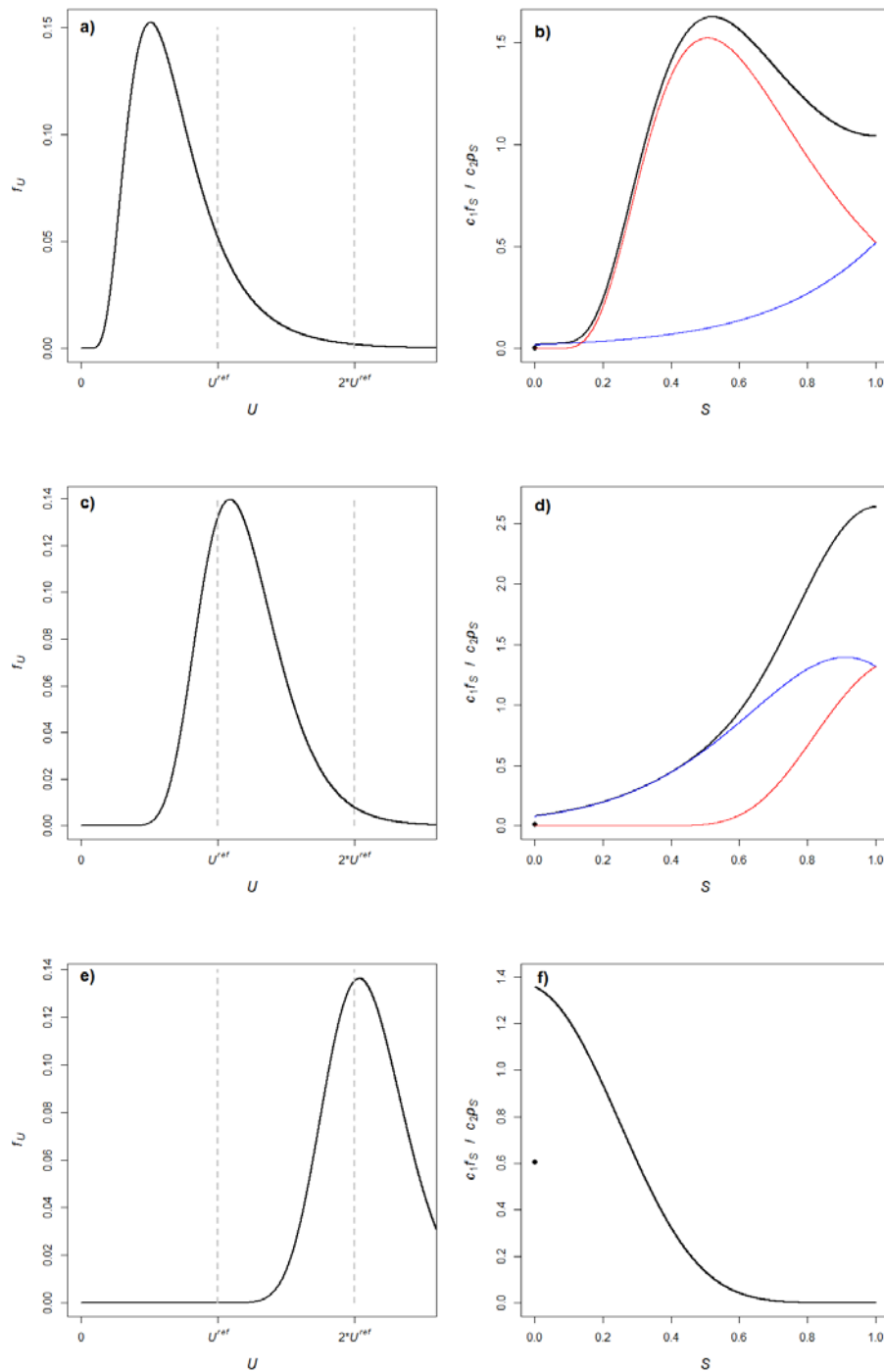
$$(16b) \quad m_{S_{OPT}} = S_{OPT}(m_U) = 0, \quad m_U \geq 2 * U^{ref}.$$

Sammenhengen mellom forventningsverdiene før og etter skalering med OPT modellen er gitt ved:

$$\begin{aligned} \mu_{S_{OPT}} &= \int_0^{u_{max}} S_{OPT}(u) f_U(u) du = \frac{1}{U^{ref}} \int_0^{U^{ref}} u f_U(u) du + \frac{1}{U^{ref}} \int_{U^{ref}}^{2U^{ref}} (2U^{ref} - u) f_U(u) du \\ &= \frac{\mu_U}{U^{ref}} + 2 \int_{U^{ref}}^{2U^{ref}} \left(1 - \frac{u}{U^{ref}}\right) f_U(u) du - \frac{1}{U^{ref}} \int_{2U^{ref}}^{u_{max}} u f_U(u) du \end{aligned} \quad (17)$$

$$\mu_{S_{OPT}} = \begin{cases} S_{OPT}(\mu_U) - 2 \int_{U^{ref}}^{2U^{ref}} \left(\frac{u}{U^{ref}} - 1\right) f_U(u) du - \frac{1}{U^{ref}} \int_{2U^{ref}}^{u_{max}} u f_U(u) du, & \mu_U \leq U^{ref} \\ S_{OPT}(\mu_U) - 2 \int_0^{U^{ref}} \left(1 - \frac{u}{U^{ref}}\right) f_U(u) du + \int_{2U^{ref}}^{u_{max}} \left(\frac{u}{U^{ref}} - 2\right) f_U(u) du, & U^{ref} < \mu_U < 2U^{ref} \\ S_{OPT}(\mu_U) + \frac{1}{U^{ref}} \int_0^{U^{ref}} u f_U(u) du + \int_{U^{ref}}^{2U^{ref}} \left(2 - \frac{u}{U^{ref}}\right) f_U(u) du, & \mu_U \geq 2 * U^{ref} \end{cases}$$

ettersom $S_{OPT}(\mu_U) = 0$ når $\mu_U \geq 2 * U^{ref}$. Dette betyr at $\mu_{S_{OPT}} < S_{OPT}(\mu_U)$ når $\mu_U \leq U^{ref}$, og $\mu_{S_{OPT}} > S_{OPT}(\mu_U)$ når $\mu_U \geq 2 * U^{ref}$. Når $U^{ref} < \mu_U < 2 * U^{ref}$ kan $\mu_{S_{OPT}}$ være større enn, lik, eller mindre enn $S_{OPT}(\mu_U)$, avhengig av om integralet $\int_0^{U^{ref}} \left(1 - \frac{u}{U^{ref}}\right) f_U(u) du \geq 0$ er mindre enn, lik, eller større enn integralet $\int_{2U^{ref}}^{u_{max}} \left(\frac{u}{2U^{ref}} - 1\right) f_U(u) du \geq 0$.



Figur 17. Tre eksempler på sannsynlighetstettheter (f_U) til målinger av biodiversitetsindikatorer (U) **a)**, **c)**, **e)**, og målingenes (S) tettheter (f_S) og punktsannsynligheter (p_S) etter skalering med OPT modellen **b)**, **d)**, **f)**. Målingenes forventningsverdi før skalering er hhv. i **a)** **b)** $0,7 * U^{ref}$, i **c)** **d)** $1,2 * U^{ref}$, og i **e)** **f)** $2,1 * U^{ref}$, der U^{ref} er referanseverdien. Kvartilbredden er i alle tre tilfeller $0,4 * U^{ref}$. Rød og blå linjer i **c)** og **d)** er hhv. $U^{ref} * f_U(U^{ref} s)$ og $U^{ref} * f_U(U^{ref}(2 - s))$, «bidraget» til sannsynlighetstettheten til S fra hhv. verdier mindre enn U^{ref} og verdier mellom U^{ref} og $2U^{ref}$. Ellers som figur 15.

8.3 Sannsynlighetsfordelingen til målinger av naturindeksen

Her utledes feilfordelingen til målinger av naturindeksen i spesielle tilfeller hvor indeksen beregnes fra n kontinuerlige biodiversitetsindikatorer som alle har observasjoner med lik fordeling gitt ved tetthet $f_U(u)$ og fordelingsfunksjon $F_U(u)$, like referanseverdier U^{ref} , skales med samme modell og gis lik vekt ved summering, $w_i = w = \frac{1}{n}$. Vi antar videre at målingenes feil er uavhengige.

Vi viser at fordelingen til en måling av naturindeksen er en blandet fordeling under alle skaleringsmodellene. Vi utleder videre uttrykk for punktsannsynligheten, sannsynlighetstettheten og fordelingsfunksjonen. Forventningsverdien til NI i disse tilfellene er

$$(18) \quad \mu_{NI} = \mu_{S_i},$$

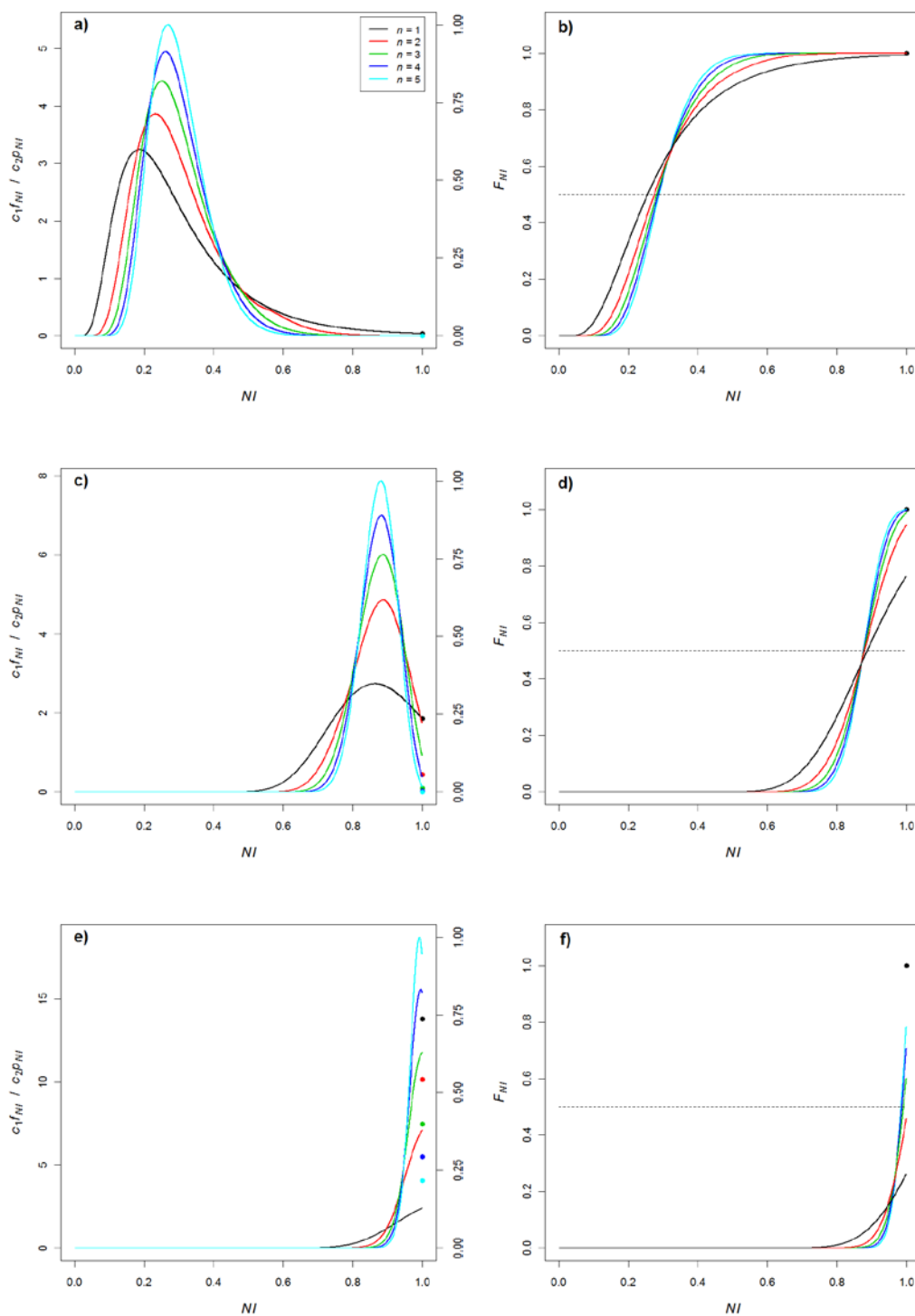
der μ_{S_i} er gitt ved likningene 8, 13 eller 17, avhengig av hvilken skaleringsmodell som benyttes. For medianen gjelder $\lim_{n \rightarrow \infty} m_{NI} = \mu_{S_i}$.

8.3.1 Indikatorer skalert med LOW modellen

Betrakter først tilfellet der NI_{LOW} beregnes som gjennomsnittet av n identisk fordelte, skalerte indikatormålinger S_1, \dots, S_n , skalert med LOW-modellen slik at deres sannsynlighetstetthet og punktsannsynlighet er gitt ved $c_{1s}f_{S_1}(s) = \dots = c_{1s}f_{S_n}(s) = c_{1s}f_{S_{LOW}}(s) = U^{ref} * f_U(U^{ref}s)$, $0 \leq s < 1$, og $c_{2s}p_{S_1}(1) = \dots = c_{2s}p_{S_n}(1) = c_{2s}p_{S_{LOW}}(1) = 1 - F_U(U^{ref}) = \varphi$, hvor $c_{1s} = F_U(U^{ref})$, og $c_{2s} = 1 - F_U(U^{ref})$ (likning 5). Fordelingen til NI_{LOW} kan da beskrives ved hjelp av en punktsannsynlighet definert for $NI_{LOW} = 1$ som representerer hendelsen der alle de skalerte målingene har verdien 1, og en stykkevis kontinuerlig sannsynlighetstetthet som representerer alle hendelser der minst en $S_i < 1$.

La E være antall skalerte målinger som oppfyller at $S_i = 1$, $0 \leq E \leq n$. Dermed er $E \sim \text{binomisk}(n, \varphi)$. Ved å anvende loven om total sannsynlighet (Høyland 1979) kan en dedusere at fordelingsfunksjonen til målingen av naturindeksen, $F_{NI_{LOW}}(x)$, $0 \leq x < 1$, kan skrives som følgende sum:

$$\begin{aligned}
 F_{NI_{LOW}}(x) &= P[0 \leq NI_{LOW} < x] = P\left[0 \leq \frac{1}{n} \sum_{i=1}^n S_i < x\right] \\
 &= \sum_{j=0}^n \left(P\left[0 \leq \frac{1}{n} \sum_{i=1}^n S_i < x \mid E = j\right] P[E = j] \right) \\
 &= \sum_{j=0}^n P\left[0 \leq \sum_{i=1}^n S_i < nx \mid E = j\right] \binom{n}{j} \varphi^j (1 - \varphi)^{n-j} \\
 &= \sum_{j=0}^n P\left[0 \leq \sum_{i=1}^{n-j} S'_i < nx - j \mid S'_i < 1, \forall i\right] \binom{n}{j} \varphi^j (1 - \varphi)^{n-j} \\
 (19) \quad F_{NI_{LOW}}(x) &= \sum_{j=0}^{n-1} \binom{n}{j} \varphi^j (1 - \varphi)^{n-j} I\left(x > \frac{j}{n}\right) \int_0^{nx-j} (c_{1s}f_{S_{LOW}}(s))^{[(n-j)*]} ds,
 \end{aligned}$$



Figur 18. **a), c), e)**: Sannsynlighetstettheter (f_{NI} , venstre skala) og punktsannsynligheter (p_s , høyre skala) til målinger av naturindeksen (NI) når indeksen beregnes som gjennomsnitt av $n = 1 \dots 5$ biodiversitetsindikatorer der alle indikatorene er skalert med LOW modellen, og indikatormålingene har samme fordeling. c_1 og c_2 angir andelen av sannsynlighetsmassen som er beskrevet med hhv. f_{NI} og p_{NI} , $c_1 + c_2 = 1$. **b), d), f)**: Fordelingsfunksjon, F_{NI} , for de samme fordelingene. Stiplet linje representerer $F_{NI} = 0,5$ som «krysser» fordelingsfunksjonen i medianen. Indikatormålingenes forventningsverdi, μ_U , er i **a) b)** $0,3 * U^{ref}$, i **c) d)** $0,9 * U^{ref}$, og i **e) f)** $1,1 * U^{ref}$, der U^{ref} er referanseverdien. Kvartilbredden er i alle tre tilfeller $0,2 * U^{ref}$.

hvor $\{S_i^j\} = \{S_i: S_i < 1\}$, $\left(x > \frac{j}{n}\right) = \begin{cases} 0, & \leq \frac{j}{n} \\ 1, & x > \frac{j}{n} \end{cases}$, og $f^{[k*]}$ er $k - 1$ ganger konvulsjonen av f med seg selv, $f^{[1*]} = f$. Merk at antall ledd i summen som er ulik 0, er τ når $\frac{\tau-1}{n} \leq x < \frac{\tau}{n}$, $\tau = 1, \dots, n$. Sannsynlighetstettheten for $0 < x < 1$ er gitt ved

$$(20) \quad c_1 f_{NI_{LOW}}(x) = \sum_{j=0}^{n-1} n \binom{n}{j} \varphi^j (1 - \varphi)^{n-j} I\left(x > \frac{j}{n}\right) \left(c_{1s} f_{S_{LOW}}(nx - j)\right)^{[(n-j)*]},$$

og punktsannsynligheten ved

$$(21) \quad c_2 p_{NI_{LOW}}(1) = \varphi^n,$$

hvor $c_1 = 1 - c_2$ og $c_2 = \left(1 - F_U(U^{ref})\right)^n$ (Figur 18). Merk også at $c_1 f_{NI_n}(x)$ er diskontinuerlig i $x = \frac{\tau}{n}$, $\tau = 1, \dots, n - 1$, når $\left(c_{1s} f_{S_{LOW}}(0)\right)^{[(n-\tau)*]} > 0$.

8.3.2 Indikatorer skalert med MAX modellen

La NI_{MAX} beregnes som gjennomsnittet av n identisk fordelte, skalerte indikatormålinger S_1, \dots, S_n , skalert med MAX-modellen. Deres sannsynlighetstetthet og punktsannsynlighet er da gitt ved $c_{1s} f_{S_1}(s) = \dots = c_{1s} f_{S_n}(s) = c_{1s} f_{S_{MAX}}(s) = U^{ref} * f_U\left(U^{ref}(2 - s)\right)$, $0 < s < 1$, $c_{2s} p_{S_1}(0) = \dots = p_{S_n}(0) = c_{2s} p_{S_{MAX}}(0) = 1 - F_U(2U^{ref}) = \vartheta$, og $c_{2s} p_{S_1}(1) = \dots = c_{2s} p_{S_n}(1) = c_{2s} p_{S_{MAX}}(1) = F_U(U^{ref}) = \vartheta$, hvor $c_{1s} = F_U(2U^{ref}) - F_U(U^{ref})$, og $c_{2s} = 1 - \left(F_U(2U^{ref}) - F_U(U^{ref})\right)$ (likning 10). Fordelingen til NI_{MAX} kan da beskrives ved hjelp av en punktsannsynlighet, $p_{NI_{MAX}}$, definert for $NI \in \left\{\frac{i}{n}: i = 0, 1, \dots, n\right\}$, som representerer hendelsene der alle de skalerte målingene tar verdiene 0 eller 1, og en stykkevis kontinuert sannsynlighetstetthet som representerer alle hendelser der minst en S_i oppfyller $0 < S_i < 1$.

La D være antall skalerte målinger som oppfyller at $S_i = 0$, E være antall skalerte målinger som oppfyller at $S_i = 1$, og G vektoren $[D \ E]$. Da er $G \sim \text{trinomisk}(n, \theta, \vartheta)$, og for $i = 0, 1, \dots, n$

$$(22) \quad c_2 p_{NI_{MAX}}\left(\frac{i}{n}\right) = P\left(NI_{MAX} = \frac{i}{n}\right) = P(G = [n - i \ i]) = \binom{n}{n - i, i, 0} \theta^{n-i} \vartheta^i,$$

der $c_2 = \sum_{i=0}^n \binom{n}{n - i, i, 0} \theta^{n-i} \vartheta^i$. Fordelingsfunksjonen, $F_{NI_{MAX}}(x)$, $0 \leq x \leq 1$, kan skrives:

$$\begin{aligned} F_{NI_{MAX}}(x) &= P[0 \leq NI_{MAX} < x] = P\left[0 \leq \frac{1}{n} \sum_{i=1}^n S_i < x\right] \\ &= \sum_{j=0}^n \sum_{k=0}^{n-j} P\left[0 \leq \frac{1}{n} \sum_{i=1}^n S_i < x \mid D = j, E = k\right] P[G = [j \ k]] \\ &= \sum_{j=0}^n \sum_{k=0}^{n-j} P\left[0 \leq \sum_{i=1}^n S_i < nx \mid D = j, E = k\right] P[G = [j \ k]] \\ &= \sum_{j=0}^{n-1} \sum_{k=0}^{n-1-j} P\left[0 \leq \sum_{i=1}^{n-j-k} S_i'' < nx - k \mid S_i'' < 1, \forall i\right] P[G = [j \ k]] \end{aligned}$$

$$\begin{aligned}
& + \sum_{k=0}^n I(0 < nx - k) * P[G = [n - k \quad k]] \\
= & \sum_{j=0}^{n-1} \sum_{k=0}^{n-1-j} I\left(x \geq \frac{n-j}{n}\right) * P[G = [j \quad k]] \\
& + \sum_{j=0}^{n-1} \sum_{k=0}^{n-1-j} P\left[0 \leq \sum_{i=1}^{n-j-k} S_i'' < nx - k \mid S_i'' < 1, \forall i\right] * \\
& I\left(x < \frac{n-j}{n}\right) P[G = [j \quad k]] \\
& + \sum_{k=0}^n I(0 < nx - k) * P[G = [n - k \quad k]]
\end{aligned}$$

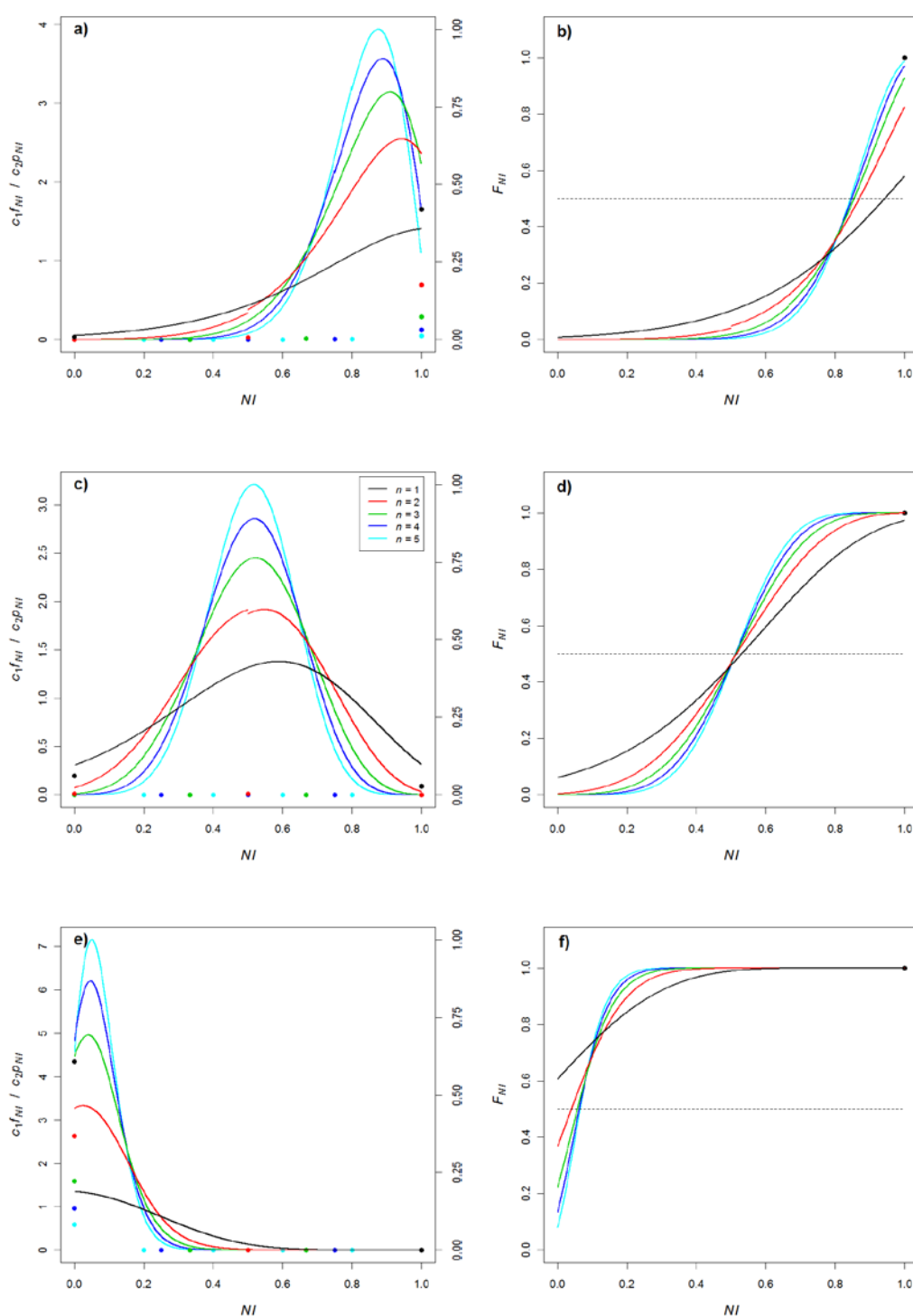
fordi $P(0 \leq \sum_{i=1}^{n-j-k} S_i'' < nx - k \mid S_i'' < 1, \forall i) = 1$ når $x \geq \frac{n-j}{n}$. Dermed får vi følgende uttrykk for fordelingsfunksjonen

$$\begin{aligned}
(23) \quad F_{NI_{MAX}}(x) = & \sum_{j=0}^{n-1} \sum_{k=0}^{n-1-j} \binom{n}{j, k, n-j-k} \theta^j \vartheta^k (1 - \theta - \vartheta)^{n-j-k} * I\left(x \geq \frac{n-j}{n}\right) \\
& + \sum_{j=0}^{n-1} \sum_{k=0}^{n-1-j} \binom{n}{j, k, n-j-k} \theta^j \vartheta^k (1 - \theta - \vartheta)^{n-j-k} * \\
& I\left(x < \frac{n-j}{n}\right) I\left(x > \frac{k}{n}\right) \int_0^{nx-k} (c_{1s} f_{S_{MAX}}(s))^{[(n-j-k)*]} ds \\
& + \sum_{k=0}^n \binom{n}{n-k, k, 0} \theta^{n-k} \vartheta^k I\left(x > \frac{k}{n}\right),
\end{aligned}$$

der $\{S_i''\} = \{S_i : 0 < S_i < 1\}$, $I\left(x \geq \frac{n-j}{n}\right) = \begin{cases} 0, & x < \frac{n-j}{n} \\ 1, & x \geq \frac{n-j}{n} \end{cases}$, $I\left(x > \frac{k}{n}\right) = \begin{cases} 0, & x \leq \frac{k}{n} \\ 1, & x > \frac{k}{n} \end{cases}$. Sannsynlighets-
tettheten, $f_{NI_{MAX}}$, for $0 < x < 1$ er da gitt ved

$$\begin{aligned}
(24) \quad c_{1s} f_{NI_{MAX}}(x) = & \sum_{j=0}^{n-1} \sum_{k=0}^{n-1-j} \binom{n}{j, k, n-j-k} \theta^j \vartheta^k (1 - \theta - \vartheta)^{n-j-k} * \\
& I\left(\frac{k}{n} < x < \frac{n-j}{n}\right) (c_{1s} f_{S_{MAX}}(nx - k))^{[(n-j-k)*]},
\end{aligned}$$

hvor $I\left(\frac{k}{n} < x < \frac{n-j}{n}\right) = \begin{cases} 0, & x \leq \frac{k}{n} \\ 1, & \frac{k}{n} < x < \frac{n-j}{n} \\ 0, & x \geq \frac{n-j}{n} \end{cases}$, og $c_1 = 1 - c_2$ (Figur 19).



Figur 19. **a), c), e)**: Sannsynlighetstettheter (f_{NI} , venstre skala) og punktsannsynligheter (p_S , høyre skala) til målinger av naturindeksen (NI) når indeksen beregnes som gjennomsnittet av $n = 1 \dots 5$ biodiversitetsindikatorer der alle indikatorene er skalert med MAX modellen, og indikatormålingene har samme fordeling. c_1 og c_2 angir andelen av sannsynlighetsmassen som er beskrevet med hhv. f_{NI} og p_{NI} , $c_1 + c_2 = 1$. **b), d), f)**: Fordelingsfunksjon, F_{NI} , for de samme fordelingene. Stiplet linje representerer $F_{NI} = 0,5$ som «krysser» fordelingsfunksjonen i medianen. Indikatormålingenes forventningsverdi, μ_U , er i **a) b)** $1,1 * U^{ref}$, i **c) d)** $1,5 * U^{ref}$, og i **e) f)** $2,1 * U^{ref}$, der U^{ref} er referanseverdien. Kvartilbredden er i alle tre tilfeller $0,4 * U^{ref}$.

8.3.3 Indikatorer skalert med OPT modellen

La NI_{OPT} beregnes som gjennomsnittet av n identisk fordelte, skalerte indikatormålinger S_1, \dots, S_n , skalert med OPT-modellen slik at deres sannsynlighetstetthet og punktsannsynlighet er gitt ved $c_{1s}f_{S_1}(s) = \dots = c_{1s}f_{S_n}(s) = c_{1s}f_{S_{OPT}}(s) = U^{ref} \left(f_U(U^{ref}s) + f_U(U^{ref}(2-s)) \right)$, $0 \leq s \leq 1$, og $c_{2s}p_{S_1}(0) = \dots = c_{2s}p_{S_n}(0) = c_{2s}p_{S_{OPT}}(0) = 1 - F_U(2U^{ref}) = \omega$, hvor $c_{1s} = F_U(2U^{ref})$, og $c_{2s} = 1 - F_U(2U^{ref})$ (likning 14). Fordelingen til NI_{OPT} kan da beskrives ved hjelp av en punktsannsynlighet definert for $NI_{OPT} = 0$ som representerer hendelsen at alle de skalerte målingene har verdien 0, og en stykkevis kontinuerlig sannsynlighetstetthet som representerer alle hendelser der minst en $S_i > 1$.

La D være antall skalerte målinger som oppfyller at $S_i = 0$, $0 \leq D \leq n$. Dermed er $D \sim \text{binomisk}(n, \omega)$,

$$(25) \quad c_2 p_{NI_{OPT}}(0) = \omega^n,$$

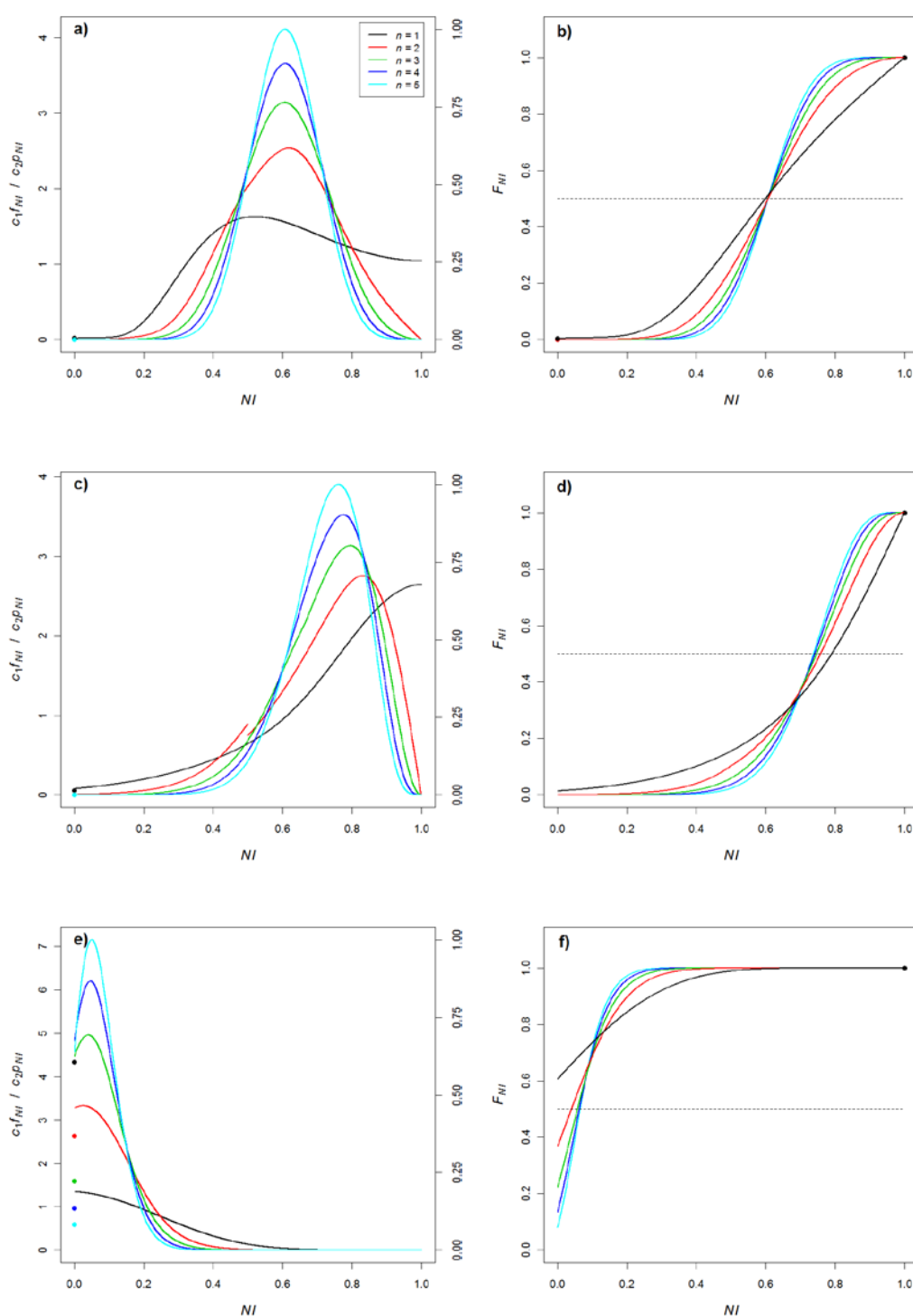
$c_2 = \left(1 - F_U(2U^{ref})\right)^n$, og fordelingsfunksjonen, $F_{NI_{OPT}}(x)$, $0 \leq x \leq 1$, kan skrives som:

$$\begin{aligned}
 F_{NI_{OPT}}(x) &= P[0 \leq NI_{OPT} < x] = P\left[0 \leq \frac{1}{n} \sum_{i=1}^n S_i < x\right] \\
 &= \sum_{j=0}^n P\left[0 \leq \frac{1}{n} \sum_{i=1}^n S_i < x \mid D = j\right] P[D = j] \\
 &= \sum_{j=0}^n P\left[0 \leq \sum_{i=1}^n S_i < nx \mid D = j\right] \binom{n}{j} \omega^j (1 - \omega)^{n-j} \\
 &= \omega^n + \sum_{j=0}^{n-1} P\left[0 \leq \sum_{i=1}^{n-j} S_i''' < nx \mid S_i''' > 0, \forall i\right] \binom{n}{j} \omega^j (1 - \omega)^{n-j} \\
 &= \omega^n + \sum_{j=0}^{n-1} I\left(x \geq \frac{n-j}{n}\right) \binom{n}{j} \omega^j (1 - \omega)^{n-j} \\
 &\quad + \sum_{j=0}^{n-1} I\left(x < \frac{n-j}{n}\right) \binom{n}{j} \omega^j (1 - \omega)^{n-j} * \\
 &\quad P\left[0 \leq \sum_{i=1}^{n-j} S_i''' < nx \mid S_i''' < 1, \forall i\right] \\
 (26) \quad F_{NI_{OPT}}(x) &= \omega^n + \sum_{j=0}^{n-1} I\left(x \geq \frac{n-j}{n}\right) \binom{n}{j} \omega^j (1 - \omega)^{n-j} \\
 &\quad + \sum_{j=0}^{n-1} \binom{n}{j} \omega^j (1 - \omega)^{n-j} I\left(x < \frac{n-j}{n}\right) \int_0^{nx} \left(c_{1s}f_{S_{OPT}}(s)\right)^{[(n-j)*]} ds,
 \end{aligned}$$

der $\{S_i'''\} = \{S_i : S_i > 0\}$. Sannsynlighetstettheten, $f_{NI_{OPT}}$, for $0 < x < 1$ er

$$(27) \quad c_1 p_{NI_n}(x) = \sum_{j=0}^{n-1} n \binom{n}{j} \omega^j (1 - \omega)^{n-j} I\left(x < \frac{n-j}{n}\right) \left(c_{1s}f_{S_{OPT}}(nx)\right)^{[(n-j)*]},$$

der $c_1 = 1 - c_2$ (Figur 20).



Figur 20. **a), c), e)**: Sannsynlighetstettheter (f_{NI} , venstre skala) og punktsannsynligheter (p_s , høyre skala) til målinger av naturindeksen (NI) når indeksen beregnes som gjennomsnittet av $n = 1 \dots 5$ biodiversitetsindikatorer der alle indikatorene er skalert med OPT modellen, og alle indikatormålingene har samme fordeling. c_1 og c_2 angir andelen av sannsynlighetsmassen som er beskrevet med hhv. f_{NI} og p_{NI} , $c_1 + c_2 = 1$. **b), d), f)**: Fordelingsfunksjon, F_{NI} , for de samme fordelingene. Stiplet linje representerer $F_{NI} = 0,5$ som «krysser» fordelingsfunksjonen i medianen. Indikatormålingenes forventningsverdi, μ_U , er i **a) b)** $0,7 * U^{ref}$, i **c) d)** $1,2 * U^{ref}$, og i **e) f)** $2,1 * U^{ref}$, der U^{ref} er referanseverdien. Kvartilbredden er i alle tre tilfeller $0,4 * U^{ref}$.

8.4 Lokasjonsforskyvning og sensitivitet

8.4.1 Definisjon av lokasjons-, skalerings- og summeringsforskyvning

Følgende tre likninger kunne tenkes som mulige, alternative definisjoner av lokasjonsforskyvning i naturindeksen ($b_{NI,i}$):

$$(28a) \quad b_{NI,1} = m_{NI} - \sum_{i=1}^n S_i(\mu_{U_i}) w_i$$

$$(28b) \quad b_{NI,2} = m_{NI} - \sum_{i=1}^n S_i(m_{U_i}) w_i$$

$$(28c) \quad b_{NI,3} = \mu_{NI} - \sum_{i=1}^n S_i(\mu_{U_i}) w_i.$$

Ettersom μ_{U_i} inngår i datagrunnlaget, ikke m_{U_i} , samt at \hat{m}_{NI} benyttes når naturindeksen presenteres, benyttes definisjonen 28a som mål for lokasjonsforskyvning i naturindeksen i denne rapporten. $b_{,2}$ og $b_{NI,3}$ måler lokasjonsforskyvning i hhv. median og forventningsverdi. Vi vil referere til disse som medianforskyvning og forventningsverdiforskyvning. Selv om disse ikke direkte måler lokasjonsforskyvning i naturindeksen slik den i dag beregnes, er de verdt en nærmere analyse fordi de illustrerer hvorvidt forekomsten av lokasjonsforskyvning i naturindeksen kun er et spørsmål om valg av mål for sentraltendens.

Lokasjonsforskyvning kan oppstå ved skalering av biodiversitetsindikatorene (likningene 3 og 4) og/eller ved summering av de skalerte indikatorene (likning 1). Vi kaller her disse to typene for skaleringsforskyvning og summeringsforskyvning. Som mål for skaleringsforskyvning i indikatormåling i ($b_{S_i,1}$), målingens median ($b_{S_i,2}$) og forventningsverdi ($b_{S_i,3}$) defineres

$$(29a) \quad b_{S_i,1} = m_{S_i} - S_i(\mu_{U_i}),$$

$$(29b) \quad b_{S_i,2} = m_{S_i} - S_i(m_{U_i}),$$

$$(29c) \quad b_{S_i,3} = \mu_{S_i} - S_i(\mu_{U_i}),$$

og som mål for summeringsforskyvning i hhv median ($b_{\Sigma 2}$) og forventningsverdi ($b_{\Sigma 3}$) defineres

$$(30a) \quad b_{\Sigma 2} = m_{NI} - \sum_{i=1}^n m_{S_i} w_i$$

$$(30b) \quad b_{\Sigma 3} = \mu_{NI} - \sum_{i=1}^n \mu_{S_i} w_i.$$

Lokasjonsforskyvning kan dermed uttrykkes som summen av summeringsforskyvning og skaleringsforskyvning summert over alle indikatorene:

$$(31a) \quad b_{NI,1} = b_{\Sigma 2} + \sum_{i=1}^n b_{S_i,1} w_i$$

$$(31b) \quad b_{NI,2} = b_{\Sigma 2} + \sum_{i=1}^n b_{S_i,2} w_i$$

$$(31c) \quad b_{NI,3} = b_{\Sigma 3} + \sum_{i=1}^n b_{S_i,3} w_i.$$

Nedenfor utledes en serie resultater vedrørende lokasjonsforskyvning i naturindeksen.

8.4.2 Skaleringsforskyvning

Likningene 7 og 12 medfører at

$$(32) \quad b_{S_{LOW},2} = b_{S_{MAX},2} = 0$$

for alle verdier av m_U , og fra likning/ulikhet 16 får vi

$$(33a) \quad b_{S_{OPT},2} < 0, m_U < 2 * U^{ref}$$

$$(33b) \quad b_{S_{OPT},2} = 0, m_U \geq 2 * U^{ref}.$$

Videre, likning 8 medfører

$$(34) \quad b_{S_{LOW},3} < 0$$

for alle verdier av μ_U , og fra vedlegg 8.2.2. følger at

$$(35a) \quad b_{S_{MAX},3} < 0, \mu_U \leq U^{ref},$$

$$(35b) \quad b_{S_{MAX},3} > 0, \mu_U \geq 2 * U^{ref},$$

mens fortegnet til $b_{S_{MAX},3}$ når $U^{ref} < \mu_U < 2U^{ref}$ avhenger av størrelsen til integralene som inngår i likning 13. Tilsvarende konklusjoner fås for OPT modellen fra vedlegg 8.2.3.

For indikatorer skalert med LOW og MAX modellene er $b_{S_i,1} = S_i(m_{U_i}) - S_i(\mu_{U_i})$ fordi det ikke er skaleringsforskyvning i medianen under disse modellene. Fordelingene til indikatormålingene har imidlertid gjerne positiv skjevhet fordi de måles på skala med absolutt nullpunkt. Dvs. m_{U_i} er ofte mindre enn μ_{U_i} . Skjevheten avtar med fordelings sentraltendens, men øker med dens spredning (Figur 5).

Dvs.

$$(36) \quad b_{S_{LOW},1} \leq 0$$

og

$$(37) \quad b_{S_{MAX},1} \geq 0$$

fordi S_{LOW} er voksende og S_{MAX} er avtagende funksjoner av U . $b_{S_{LOW},1} = 0$ når $m_{U_i} > U^{ref}$ fordi da er $S_i(m_{U_i}) = S_i(\mu_{U_i}) = 1$. Tilsvarende gjelder at $b_{S_{MAX},1} = 0$ når $\mu_{U_i} < U^{ref}$ og når $m_{U_i} > 2U^{ref}$.

For indikatorer skalert med OPT modellen er

$$(38) \quad b_{S_{OPT},1} = b_{S_{OPT},2} + S_{OPT}(m_{U_i}) - S_{OPT}(\mu_{U_i}).$$

Dvs. indikatorens skaleringsforskyvning er summen av skaleringsforskyvning i medianen og forskyvning som skyldes skjevhet i indikatorens fordeling.

8.4.3 Summeringsforskyvning

Det er ikke summeringsforskyvning i forventningsverdiene. Dette fordi $NI = \sum_{i=1}^n S_i w_i$ som medfører at $\mu_{NI} = \sum_{i=1}^n \mu_{S_i} w_i$. Som en direkte konsekvens av dette følger at

$$(39) \quad b_{NI,3} = \sum_{i=1}^n w_i b_{S_i,3},$$

der $b_{S_i,3}$ fås fra likningene 8, 13 og 17 avhengig av hvilken skaleringsmodell som blir anvendt på de ulike indikatorene og indikatormålingenes forventningsverdier og referanseverdier. Forventningsverdiforskyvning skyldes derfor kun skaleringsforskyvning.

Lokasjonsforskyvning studeres i denne rapporten bl.a. ved å undersøke spesialtilfeller der naturindeksen beregnes fra uavhengige målinger av biodiversitetsindikatorer som alle har samme fordeling, lik referanseverdi, skaleres med samme skaleringsfunksjon, og gis de samme vektene. Spesialtilfellene oppfyller dermed at $w_i = w = \frac{1}{n}$, og at $S_i(\mu_{U_i}; U_i^{ref}) = k$, for alle i . Videre er $\mu_{NI} = \sum_{i=1}^n \mu_{S_i} w_i = \mu_{S_i}$. Dette betyr at under disse betingelsene er

$$(40) \quad b_{NI,3} = \frac{1}{n} \sum_{i=1}^n (\mu_{S_i} - S_i(\mu_{U_i})) = b_{S,3}.$$

Fra Lindeberg - Levys sentralgrenseteorem (Høyland 1979) følger for disse spesielle tilfellene at

$$(41) \quad \lim_{n \rightarrow \infty} m_{NI} = \mu_{NI} \Rightarrow \lim_{n \rightarrow \infty} b_{NI,1} = b_{S,3}$$

fordi NI da er snittet av identisk fordelte stokastiske variable. Dette betyr at lokasjonsforskyvning i naturindeksen i våre spesialtilfeller nærmer seg en grense når antall indikatorer indeksen beregnes fra øker. Denne grensen er lik skaleringsforskyvningen i forventningsverdien.

For spesialtilfellet gjelder under LOW og MAX modellene at

$$(42) \quad \begin{aligned} \lim_{n \rightarrow \infty} b_{NI,2} &= \lim_{n \rightarrow \infty} \left(m_{NI} - \sum_{i=1}^n S_i(m_{U_i}; U_i^{ref}) w_i \right) = \lim_{n \rightarrow \infty} \left(m_{NI} - \sum_{i=1}^n m_{S_i} w_i \right) \\ &= \lim_{n \rightarrow \infty} (m_{NI} - m_{S_i}) = \mu_S - m_S, \end{aligned}$$

ettersom $S_i(m_{U_i}; U_i^{ref}) = m_{S_i}$ (likningene 7 og 12) under disse skaleringene. Dvs. at medianforskyvning under LOW- og MAX modellene har en grense gitt av skjevheten i fordelingene til de skalerte indikatormålingene. Ettersom disse fordelingene som oftest er skjeve (figur 6) er μ_S og m_S forskjellige. Likning 42 må derfor innebære at det er en summeringsforskyvning i naturindeksens median ettersom det ikke er skaleringsforskyvning under disse to modellene.

8.4.4 Konklusjoner – lokasjonsforskyvning

Vi har dermed vist at lokasjonsforskyvning i naturindeksen forekommer og har følgende tre årsaker: 1) Fordelingene til målinger av indikatorene har positiv skjevhet fordi de er på en skala med absolutt nullpunkt. Skjevheten medfører skaleringsforskyvning under alle skaleringsmodellene. 2) En negativ skaleringsforskyvning i medianen under OPT modellen. 3) De tre skaleringsmodellene gir fordelinger med positiv eller negativ skjevhet etter skalering. Denne skjev-

heten gir en positiv eller negativ summeringsforskyvning som skyldes at fordelingen til gjennomsnittet av de skalerte indikatormålingene endrer form etter hvert som antall målinger som inngår i snittet øker. Summeringsforskyvning i de tre modellene er illustrert gjennom eksempler i figurene 18 – 20 som viser hvordan fordelingen til målinger av naturindeksen endrer form mot en symmetrisk fordeling etter hvert som antall indikatormålinger øker, og hvordan medianen i fordelingen endres parallelt.

Vi har vist for vårt spesialtilfelle at lokasjonsforskyvning i naturindeksen nærmer seg en grense når antall indikatorer indeksen beregnes fra øker. Denne grensen er lik skaleringsforskyvningen i forventningsverdien. Dette fordi naturindeksens fordeling blir stadig mer symmetrisk jo flere identisk fordelte indikatormålinger som inngår.

Videre har vi også vist at forekomst eller fravær av lokasjonsforskyvning ikke kun er et spørsmål om hvilket mål for sentraltendens en velger som estimat for naturindeksens og indikatorernes verdier. Skalering medfører forskyvning av forventningsverdien til de skalerte indikatormålingene, mens summering av de skalerte målingene medfører en summeringsforskyvning av medianen til naturindeksen. Lokasjonsforskyvning forekommer derfor både når en estimerer naturindeksens verdi som fordelingsmedian og når en estimerer verdien som fordelingsforventningsverdi. Dessuten vil lokasjonsforskyvning forekomme enten indikatormålingenes sentraltendens oppgis som fordelingsmedian eller forventningsverdi.

8.4.5 Sensitivitet med hensyn til endring i indikatorernes forventningsverdier

Sensitivitet med hensyn til endring i indikatormålingenes forventningsverdier defineres som den retningsderiverte av naturindeksens median, $m_{NI}'(\mu_U; \mathbf{y})$, med hensyn til målingenes forventningsverdier og deres endringsretning \mathbf{y} .

Betrakter her tilfellet der naturindeksen beregnes fra målinger av n indikatorer (U_1, \dots, U_n) som alle har samme fordeling ($F_{U_1}(u) = \dots = F_{U_n}(u)$, $\mu_{U_1} = \dots = \mu_{U_n} = \mu_U$), skaleringsmodell, veies likt ($w_1 = \dots = w_n = \frac{1}{n}$) og endres like mye. Da er $\frac{\partial m_{NI}(\mu_U)}{\partial \mu_{U_1}} = \dots = \frac{\partial m_{NI}(\mu_U)}{\partial \mu_{U_n}}$, og for tilfeldig valgt U_j blant målingene gjelder:

$$(43) \quad m_{NI}'(\mu_U; \mathbf{y}) = m_{NI}'(\mu_U; \mathbf{1}) = \sum_{i=1}^n \frac{\partial m_{NI}(\mu_U)}{\partial \mu_{U_i}} = n \frac{\partial m_{NI}(\mu_U)}{\partial \mu_{U_j}} = n * m_{NI}'(\mu_U; \mathbf{e}_j),$$

der \mathbf{e}_j er enhetsvektoren langs dimensjon j . Videre, hvis det ikke er usikkerhet i målingene av indikatorerne, er naturindeksen en deterministisk funksjon av indikatorerne gitt ved likningene 1 til 4, og for U_j gjelder

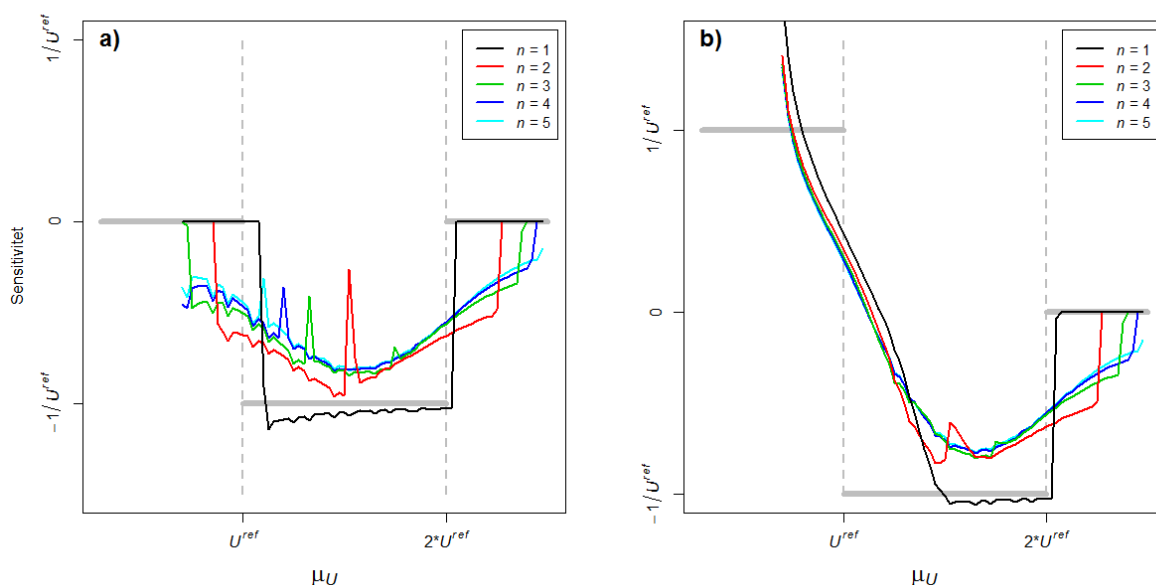
$$(44) \quad NI'(\mathbf{u}; \mathbf{1}) = \sum_{i=1}^n \frac{\partial NI(\mathbf{u})}{\partial U_i} = \sum_{i=1}^n \frac{\partial (\sum_{i=1}^n S_i w_i)(\mathbf{u})}{\partial U_i} = \sum_{i=1}^n w_i \frac{\partial S_i(\mathbf{u})}{\partial U_i} = \frac{dS_j(\mathbf{u})}{dU_j}$$

i alle punkter \mathbf{u} der de ulike deriverte er definert. La $\Delta m_{NI}(\mu_U) = m_{NI}(\mu_U + k\mathbf{1}) - m_{NI}(\mu_U)$ være endringen i naturindeksens median for en liten felles endring k i målingenes forventningsverdier. Ved å anvende en første ordens Taylor tilnærming til $\Delta m_{NI}(\mu_U)$ får vi

$$(45) \quad \Delta m_{NI}(\mu_U) \approx k\mathbf{1} * \nabla m_{NI}(\mu_U) = \sum_{i=1}^n k \frac{\partial m_{NI}(\mu_U)}{\partial \mu_{U_i}} = k * m_{NI}'(\mu_U; \mathbf{1}),$$

dvs. $m_{NI}'(\mu_U; \mathbf{1})$ kan tilnærmet beregnes som $\frac{\Delta m_{NI}(\mu_U)}{k}$ under disse betingelsene.

Under MAX modellen og ved svært stor usikkerhet i dataene opptrer brå endringer i naturindeksens følsomhet for endringer i dataene for spesifikke verdier av indikatorernes forventningsverdier (Figur 21a). Indeksen blir brått mindre sensitiv (i absoluttverdi) for μ_U ca lik $\frac{1}{n}$, en effekt som avtar med n . En liknende, men svakere effekt ses også under OPT modellen (Figur 21b). For MAX modellen har dette sammenheng med at en av diskontinuitetene som opptrer i fordelingsfunksjonen til en måling av naturindeksen sammenfaller med medianen for disse verdiene av μ_U (Figur 14b). For OPT modellen sammenfaller medianen med en stor endring i fordelingsfunksjonen.



Figur 21. Naturindeksens sensitivitet med hensyn til endring i indikatorernes forventningsverdi, $m_{NI}'(\mu_U; \mathbf{y})$, som funksjon av indikatorernes forventningsverdier, μ_U , og antall indikatorer, n . Dette for spesialtilfeller hvor indeksen beregnes som gjennomsnittet av n skalerte biodiversitetsindikatorer, indikatorerne har samme referanseverdi og samme lognormal-fordeling med kvartilbredde lik 0,6 ganger referanseverdien før skalering, alle indikatorerne skaleres med MAX-modellen (a) eller OPT modellen (b), og alle indikatorerne endres like mye ($\mathbf{y} = \mathbf{1}$). Sensitivitet er beregnet numerisk slik som beskrevet i kapittel 4.1. Horisontale, grå, heltrukne, tykke linjer representerer naturindeksens sensitivitet med hensyn til endringer i indikatorverdiene når det ikke er usikkerhet i dataene.

8.4.6 Sensitivitet med hensyn til feil i referanseverdiene

Betrakter igjen tilfellet der naturindeksen beregnes fra målinger av n indikatorer (U_1, \dots, U_n) som alle har samme fordeling ($F_{U_1}(u) = \dots = F_{U_n}(u)$, $\mu_{U_1} = \dots = \mu_{U_n} = \mu_U$) og skaleringsmodell, og veies likt ($w_1 = \dots = w_n = \frac{1}{n}$). La referanseverdiene være målt med feil og gitt ved $U_i^{ref} = U_0^{ref} Y_i$, $i = 1, \dots, n$, der U_0^{ref} er lik for alle indikatorerne og $Y_i \sim \text{lognormal}\left(-\frac{1}{2} \ln(1 + q^2), (1 + q^2)\right)$ slik at $E(Y_i) = 1$ og $\text{var}(Y_i) = q^2$ og $0 < q \leq 0,1$. $E(U_i^{ref}) = U_0^{ref}$ og $SD(U_i^{ref}) = U_0^{ref} q$.

La videre $\Delta m = m_{NI_1} - m_{NI_0}$ der $NI_1 = \sum_{i=1}^n S_i(U_{1,i}^{ref} | U_i)$, $\{U_{1,i}^{ref}\}$ er n observasjoner av U^{ref} og $NI_0 = \sum_{i=1}^n S_i(U_0^{ref} | U_i)$. La \mathbf{LU}_1^{ref} betegne vektoren $[\ln(U_{1,1}^{ref}), \dots, \ln(U_{1,n}^{ref})]$ og g_n være den ret-

ningsderiverte til medianen med hensyn til LU_1^{ref} i punktet LU_0^{ref} , $g_n = m_{NI}'(LU_1^{ref}; \mathbf{1})$. Ved å benytte 1. ordens Taylor-tilnærming til Δm fås

$$\begin{aligned}\Delta m &\approx \sum_{i=1}^n \frac{\partial m_{NI}(LU_0^{ref})}{\partial \ln(U_i^{ref})} (\ln(U_{1,i}^{ref}) - \ln(U_0^{ref})) \\ &= \sum_{i=1}^n \frac{1}{n} m_{NI}'(LU_1^{ref}; \mathbf{1}) (\ln(U_{1,i}^{ref}) - \ln(U_0^{ref})) = \sum_{i=1}^n \frac{1}{n} g_n \ln \frac{U_{1,i}^{ref}}{U_0^{ref}} = \frac{1}{n} g_n \sum_{i=1}^n \ln Y_i,\end{aligned}$$

fordi i det spesialtilfellet vi studerer her er $m_{NI}'(LU_1^{ref}; \mathbf{1}) = \sum_{i=1}^n \frac{\partial m_{NI}(LU_0^{ref})}{\partial \ln(U_i^{ref})} = n \frac{\partial m_{NI}(LU_0^{ref})}{\partial \ln(U_j^{ref})}$ for vilkårlig j . For $0 < x \leq 0,1$ benytter vi tilnærmelsen $\ln(1+x) \approx x$ og får følgende tilnærmede uttrykk for variansen til en måling av naturindeksen under betingelsene over og for gitt $F_{U_i}(u)$

$$\sigma_{m_{NI}}^2 \approx \frac{1}{n^2} g_n^2 \sum_{i=1}^n \text{var}(\ln Y_i) = \frac{n}{n^2} g_n^2 \ln(1+q^2) \approx \frac{1}{n} g_n^2 q^2,$$

og standardavviket er under de betingelsene som er beskrevet over, tilnærmet lik

$$(46) \quad \sigma_{m_{NI}} \approx \frac{1}{\sqrt{n}} |g_n| cv_{U^{ref}}.$$



Norsk institutt for naturforskning (NINA) er et nasjonalt og internasjonalt kompetansesenter innen naturforskning. Vår kompetanse utøves gjennom forskning, utredningsarbeid, overvåking og konsekvensutredninger.

NINAs primære aktivitet er å drive anvendt forskning. Stikkord for forskningen er kvalitet og relevans, samarbeid med andre institusjoner, tverrfaglighet og økosystemtilnærming. Offentlig forvaltning, næringsliv og industri samt Norges forskningsråd og EU er blant NINAs oppdragsgivere og finansieringskilder.

Virksomheten er hovedsakelig rettet mot forskning på natur og samfunn, og NINA leverer et bredt spekter av tjenester gjennom forskningsprosjekter, miljøovervåking, utredninger og rådgiving.

ISSN:1504-3312
ISBN: 978-82-426-2392-8

Norsk institutt for naturforskning

NINA Hovedkontor

Postadresse: Postboks 5685 Sluppen, NO-7485 Trondheim

Besøks/leveringsadresse: Tungasletta 2, NO-7047 Trondheim

Telefon: 73 80 14 00, Telefaks: 73 80 14 01

E-post: firmapost@nina.no

Organisasjonsnummer 9500 37 687

<http://www.nina.no>

Samarbeid og kunnskap for framtidens miljøløsninger