

Hvor kommer data fra?

af Emma Emilie Andersen, Carsten Rahbek, Jacob Heilmann-Clausen og Karsten Elmoose Vad.



En dansk bøgeskov i efterårsfarver. Foto: Natasja Corfixen.

Alle undervisningsmaterialer, som udvikles i 'Biodiversitet i gymnasiet', er baseret på den bedste tilgængelige forskningsviden. Det gælder også forløbet her om biodiversitet i danske skove.

Alle de data, som eleverne arbejder med, er altså ægte forskningsdata, som stammer fra forskningsprojektet 'Bøgeskovsprojektet' (med den officielle titel 'Urørt skov eller naturnær skovdrift – hvad skal der til for at bevare biodiversiteten i de danske skove?').

Når man arbejder med ægte forskningsdata fra biologiens verden, så er det vigtigt at forstå, hvor data kommer fra, hvordan det er indsamlet og hvilke begrænsninger, der følger med. Det vil vi gennemgå for Bøgeskovsprojektet i denne artikel.

Bøgeskovsprojektet

Bøgeskovsprojektet blev udført ved Center for Makroøkologi, Evolution og Klima (CMEC) på Københavns Universitet i perioden 2014 til 2018 under ledelse af lektor Jacob Heilmann-Clausen og professor Carsten Rahbek i samarbejde med Naturstyrelsen og med støtte fra 15. Juni Fonden.

Formålet med projektet var at udvikle en stærk data- og vidensbaseret platform for forvaltning af biodiversitet i danske skove. Første del af denne opgave bestod i at besvare spørgsmålet: Hvordan fordeler biodiversiteten sig langs gradienten fra traditionel produktionskov til gammel urørt skov?

Det spørgsmål søgte forskerne at besvare gennem en grundig empirisk kortlægning af artsdiversitet og struktur i bøgetræsdominerede løvskove fordelt rundt i landet.

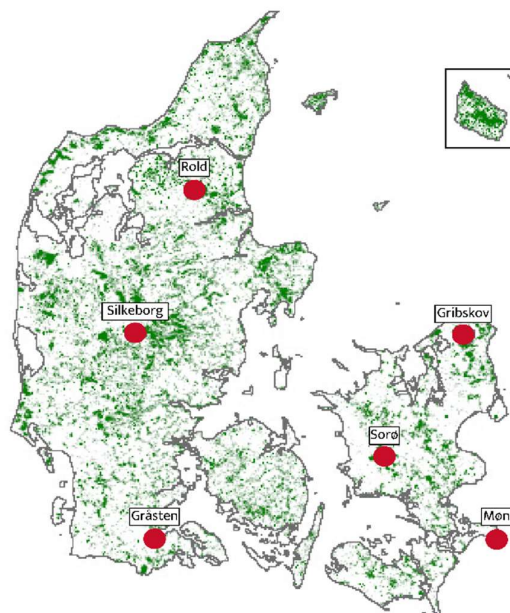
Forsøgsdesignet

Forskerne udvalgte seks bevoksninger i Danmark, med Gribskov som hovedområde med 5 replikater og de resterende 5 områder som satellitter med et enkelt replikat per område (figur 1).

Her kortlagde forskerne forekomsten og hyppigheden af arter samt habitatstruktur i to områder med produktionsdrift og to områder med urørt skov. I alt har der altså været undersøgt 40 forskellige skovområder. Alle områderne var ca. 3 hektar store og mindst 80 år gamle.

For at plottene bedst kan sammenlignes med hinanden, så er de blevet udvalgt, så de ligner hinanden så meget som muligt.

I hvert område kortlagde forskerne forekomsten og antal arter af vedboende svampe, epifytiske laver, mosser, karplanter, biller og fugle samt en række parametre, som beskriver skovens habitatstruktur, fx områdernes volumen af dødt ved pr. hektar, antal af veterantræer og dækning af hhv. vådområder og lysåbent areal.



Figur 1: De seks områder i Danmark, hvor Bøgeskovsprojektets undersøgelser fandt sted. I alt undersøgte forskerne 40 forskellige skovområder. Kort fra rapporten 'Bevarelse af biodiversiteten i de danske skove' udgivet af CMEC i 2016.

Metoder til dataindsamling i felten



Lektor Jacob Heilmann-Clausen måler diameteren på et stykke dødt ved som led i udregningen af skovområdets volumen af dødt ved pr. hektar.
Foto: Karsten Elmoose Vad

I Bøgeskovsprojektets feltarbejde er der anvendt forskellige dataindsamlingsmetoder for hver organismegruppe og strukturparameter, som datasættene indeholder.

En del af indsamlingen af data om skovenes struktur er udført med remote sensing-metoder, fx Lidar, hvor man fra fly kan danne tredimensionelle billeder af skovens overflade og træer. Andre struktur-parametre er opmålt ved feltarbejde. Det gælder fx mængden af dødt ved, som forskerne har opmålt vha. transekter på 50 meter (10 styk pr. skovområde). Stående dødt træ er opmålt i 5 meters afstand fra samme transekter.

For alle de undersøgte organismegrupper har forskerne identificeret og registreret de tilstedeværende arter samt – i nogle tilfælde – opgjort antallet af individer.

Eksempelvis har forskerne for vedboende svampe tre gange på en sæson optalt alle de arter, som dannede frugtlegemer på det døde ved, som blev opmålt vha. transektmetoden. For de øvrige organismegrupper foregik dataindsamlingen naturligvis anderledes, fx blev fugle identificeret ud fra både observationer og lyd.

Projektets forskere har indsamlet data sammen med eksterne organismeeksperter, for at sikre den højst mulige kvalitet af data for alle organismegrupper.

Fra forskningsdata til undervisningsmateriale

De datasæt, som forskerne skabte under Bøgeskovsprojektet, er enormt store. Alene datasættet for svampene indeholder mere end 27.400 datapunkter bestående af bl.a. artsnavne, antal registreringer af hver art på hvert stykke dødt ved i alle

skovområder samt arternes rødlistestatus. Dertil skal lægges rådata for alle andre organismegrupper samt for de struktur-parametre, som også blev undersøgt.

Størrelsen på – og kompleksiteten af – datasættene har betydet, at vi måtte foretage nogle justeringer og finpudsnings af data, for at de har kunnet anvendes i en undervisningssammenhæng.

For eksempel har vi valgt kun at skelne mellem to skovdriftstyper (produktionsskov og urørt skov; figur 2), mens man i

dataindsamlingen har skelnet mellem fire typer (konventionel produktionsskov, naturnær produktionsskov, nyligt udlagt urørt skov og længe udlagt urørt skov). Desuden har vi begrænset undervisningsmaterialet til de fire organismegrupper planter, svampe, biller og fugle.

Finpudsningen af datasættene er foregået i tæt sparring med de forskere, som har indsamlet data. Vi har fokuseret på at gøre data pædagogisk tilgængelige uden at gå på kompromis med datasættenes autenticitet eller indhold.

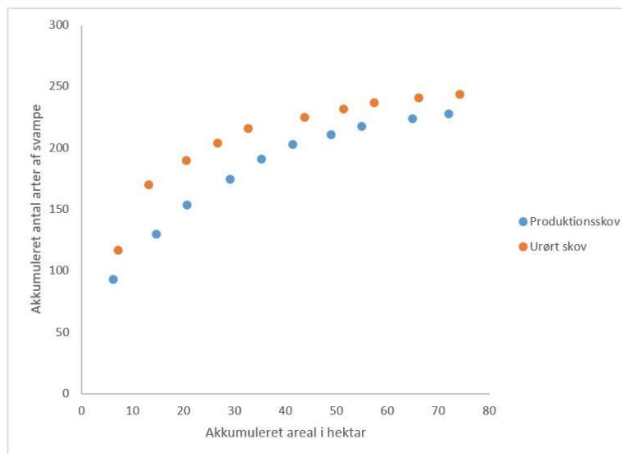
Ægte data giver gode faglige diskussioner

Undervisningsmaterialet er som bekendt baseret på ægte forskningsdata. I vores mange afprøvninger af øvelserne med både lærere og elever er det tydeligt, at de ægte data kan give nogle faglige diskussioner og erkendelser, som man ikke nødvendigvis får, når data er gjort eksemplariske til brug i undervisningen.

Ægte data er sjældent perfekte, og det giver troværdighed, men det gør også, at eleverne får brug for deres biologiske viden, når de skal forklare data, der ikke nødvendigvis passer 1:1 med teorien.

Når data ikke passer med teorien...

I arbejdet med større biologiske datasæt er det nemlig ikke altid, at ens resultater ser ud præcis som teorien foreskriver. Når man løber ind i en datamæssig uoverensstemmelse med teori, så skal man være bevidst om, hvad det skyldes.



Figur 2: Et eksempel på en besvarelse fra øvelsen om habitat-heterogenitetssammenhængen. Figuren viser det akkumulerede antal arter af svampe for hhv. produktionsskov (blå) og urørt skov (orange).

For Bøgeskovsprojektet findes der to overordnede fejlkilder, som har påvirket resultaterne: Uperfekt forsøgsdesign og biologiske tilfældigheder.

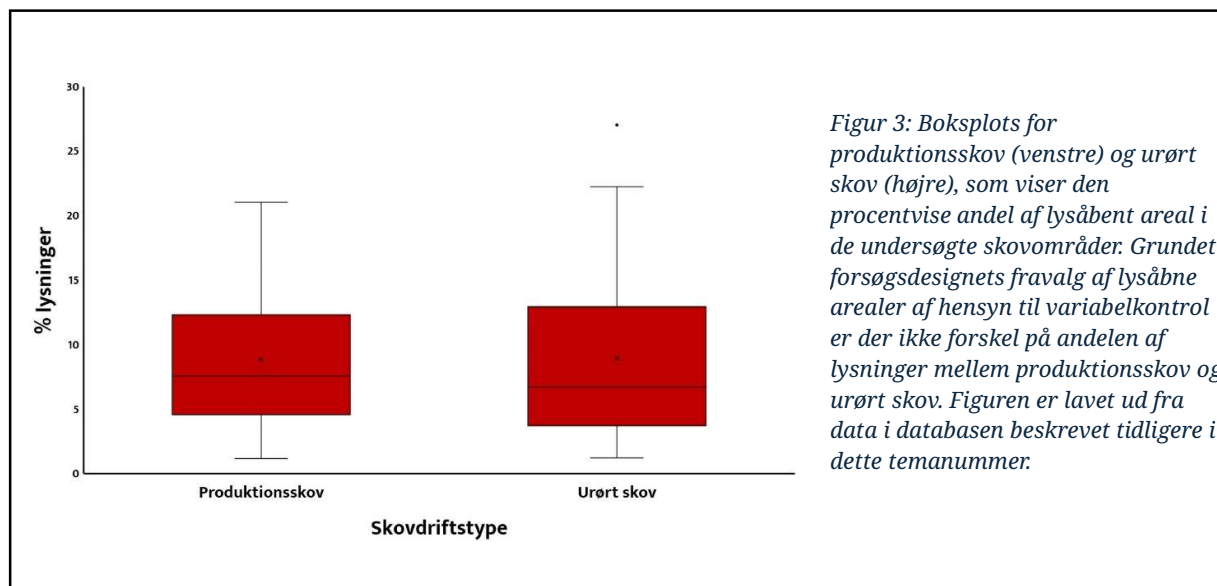
Forsøgsdesignet i projektet kan ikke laves perfekt, fordi graden af skovområdernes urørthed varierer meget – nogle af de dyrkede skove har fx haft et overraskende højt habitatudbud sammenlignet med nogle af de urørte skovområder (og omvendt). Den biologiske virkelighed er altså mere rodet, end man kunne ønske sig med et perfekt forsøgsdesign. Desuden er skovenes arter individuelle agenter, som bevæger sig frit, og det skaber en vis mængde tilfældig støj i ens observationer.

Herunder følger et par eksempler med forklaringer på, hvorfor nogle data i mindre grad passer med teori. Der findes med sikkerhed flere mulige forklaringer.

LYSNINGER

Som nævnt er områderne udvalgt så strukturen er forholdsvis ens, for at kunne sammenligne faktorerne mellem produktionsskov og urørt skov. Det betyder bl.a., at forskerne for den urørte skov har fravalgt områder med større lysninger.

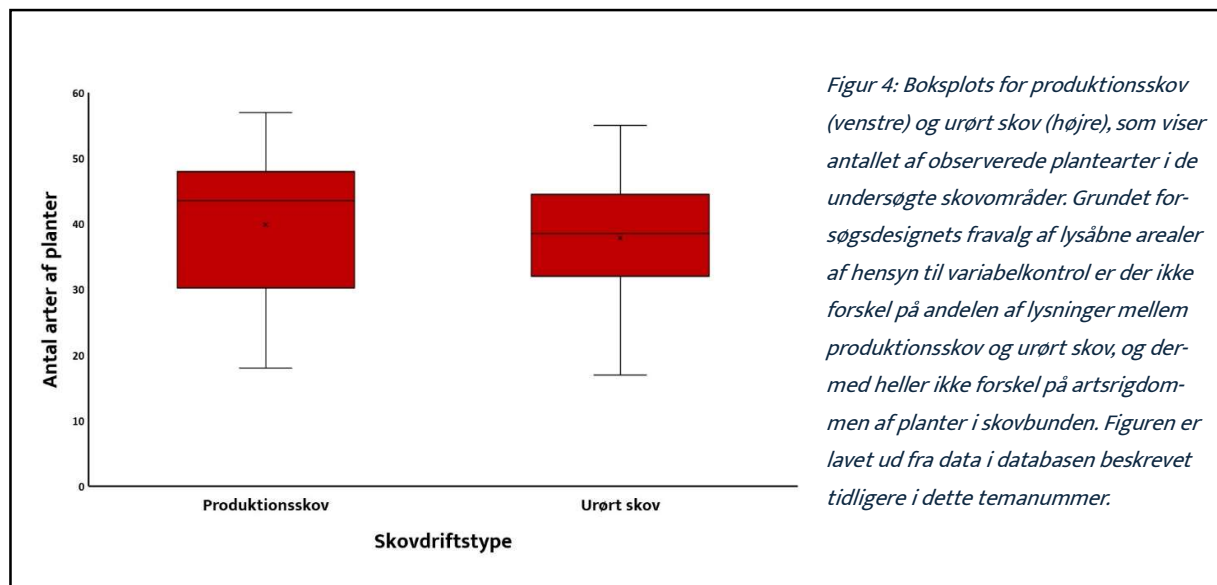
Ift. variabelkontrol ville det fx være uhensigtsmæssigt at sammenligne 3 hektar tæt produktionsskov med 3 hektar helt lysåbent areal i en urørt skov – så er skovdriftsformen pludselig ikke den eneste variabel i sammenligningen. Det betyder, at de urørte skovområder ud fra datasættene kan fremstå mindre varierede, end hvad man ellers kunne forvente i en urørt skov (figur 3).



PLANTER

På tværs af de fire organismegrupper, som eleverne kan arbejde med i databasen (fugle, svampe, biller og planter) er der naturligvis forskel på, hvordan deres artsrigdom påvirkes af skovdriftstypen – helt som vi vil forvente ud fra teori. Men planterne viser ikke lige så tydeligt denne forskel, som de andre organismegrupper (figur 4).

Det handler i bund og grund om lys – for lys er jo en kritisk ressource for planter vækst. Og i en ung urørt skov, som får lov til at vokse frit, vil lyset i store områder have svært ved at trænge ned til de mindre planter på skovbunden. Det faktum – samt selve forsøgsdesignet, som grundet variabelkontrol har fravalgt lysåbne arealer af hensyn til variabelkontrol er der ikke forskel på andelen af lysninger mellem produktionsskov og urørt skov, og dermed heller ikke forskel på artsrigdommen af planter i skovbunden. Det er biologisk set forventeligt, og planterne er af netop samme årsag heller ikke de mest relevante indikatorer for de urørte skoves biodiversitet.



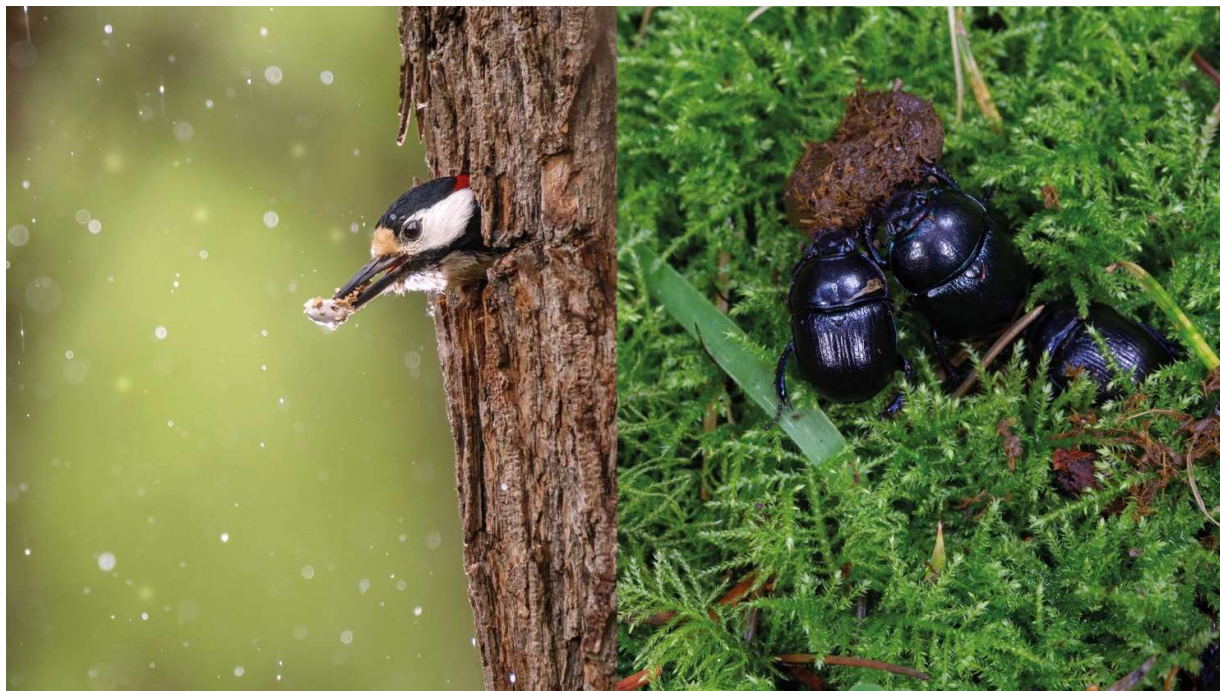
SKALA

Noget andet, som gør sig gældende, når vi betragter forskellige organismegrupper, er, at der er forskel på organismernes størrelse. Det betyder, at organismerne opfatter skala forskelligt.

Undersøgelserne i Bøgeskovsprojektet foregik i skovområder på ca. 3 hektar. For en skarnbasse er 3 hektar enormt stort. For en fugl, fx en stor flagspætte, er 3 hektar meget småt.

En konsekvens af denne skalaeffekt er, at fordi arealet opfattes som større af skarnbassen, så vil vi også forvente en mere tydelig arts-arealkurve for den, mens kurven vil være mindre kraftig for den store flagspætte, som opfatter 3 hektar som et meget småt område.

Uanset hvordan man arbejder med data, så er det altid vigtigt at forholde sig til, hvorfra og hvordan disse data er skabt. Vi har i denne artikel forsøgt at bidrage med netop denne information, med et fokus på, at arbejdet med ægte data ikke altid går lige ud ad landevejen – men at det til gengæld kan være fundamentet for gode faglige diskussioner om naturvidenskabelig metode og forsøgsdesign.



En stor flagspætte, Dendrocopus major og skovskarnbasser, Anoplotrupes stercorosus.

Fotos: Mads Hagen (venstre) og Thomas Kehlet (højre)