Udvikling af en statistisk analysemetode til brug for validering af Miljøstyrelsens AI-model til genkendelse af habitatnaturtyper

Fagligt notat fra DCE – Nationalt Center for Miljø og Energi

Dato: 24. August 2023 | **46**

AARHUS
UNIVERSITET
DCE – NATIONALT CENTER FOR MILJØ OG ENERGI

Datablad

	Fagligt notat fra DCE – Nationalt Center for Miljø og Energi
Kategori:	Rådgivningsnotat
Titel:	Udvikling af en statistisk analysemetode til brug for validering af Miljøstyrelsens AI-model til genkendelse af habitatnaturtyper
Forfattere:	Thorsten J. S. Balsby, Jesper Bladt, Bettina Nygaard
Institution:	Institut for Ecoscience
Faglig kommentering:	Jesper E. Moeslund
Kvalitetssikring, DCE:	Jesper Fredshavn
Ekstern kommentering:	Miljøstyrelsen. Kommentarerne findes her: http://dce2.au.dk/pub/komm/N2023_46_komm.pdf
Rekvirent:	Miljøstyrelsen
Bedes citeret:	Balsby, TJS., Bladt, J., Nygaard, B. 2023. Udvikling af en statistisk analysemetode til brug for validering af Miljøstyrelsens AI-model til genkendelse af habitatnaturtyper. Aarhus Universitet, DCE – Nationalt Center for Miljø og Energi, 20 s. - – Fagligt notat nr. 2023|46

Gengivelse tilladt med tydelig kildeangivelse

Foto forside:	Colorbox
Sideantal:	20

Indhold

1 Forord **4**

2 Feltbaseret kortlægning **5**

3 AI-modellen **6**

4 Den statistiske valideringsmodel **7**

4.1 Databehandling 7

4.2 Statistisk analyse 10

5 Eksempel på en dataanalyse **12**

5.1 Samlet analyse (test 1) 14

5.2 Identifikation af habitatnaturtyper (test 2) 15

5.3 Hvor meget skal der samples? 16

5.4 Arealafhængighed 16

6 Konklusion **17**

7 Litteratur **19**

8 Appendix **20**

1 Forord

Dette notat beskriver DCE's – Nationalt Center for Miljø og Energi, Aarhus Universitet – arbejde med udvikling af en statistisk metode, der kan validere Miljøstyrelsens AI-model til genkendelse af danske habitatnaturtyper. AI-modellen er en del af den digitale naturovervågning (DNO). Den statistiske validering skal sammenholde AI-modellens forudsigelser med en feltbaseret kortlægning efter traditionelle kortlægningsmetoder (se krav nedenfor). Analyserne har taget udgangspunkt i en tidligere udgave af AI-modellen, og DCE har derfor lagt vægt på metodeudvikling fremfor de konkrete resultater af valideringen af AI-modellen.

MST har bedt DCE udvikle en statistisk metode, der om muligt kan opfylde følgende krav:

- Den statistiske valideringsmetode skal som udgangspunkt angive en sikkerhed pr. naturtype. Hvis det giver mening, må der også gerne angives en sikkerhed pr. naturtype pr. geografisk meningsfyldte område, f.eks. efter den inddeling, der bliver brugt i NOVANA sammenhæng. Desuden skal valideringsmetoden også angive en sikkerhed for de situationer, hvor der ikke er forudsagt habitatnatur i de udlagte valideringsplots og det skal fastslås, hvor mange feltindsamlinger der er nødvendigt pr. naturtype for at få et statistisk holdbart resultat.
- I beskrivelsen af valideringsmetoden er det vigtigt for Miljøstyrelsen, at analyserne bliver beskrevet på en måde, så det kan forstås, hvad man kan tolke ud af resultaterne med tanke på usikkerheder og dataindsamlingsmetode.
- MST beder også om at inddrage de geografiske unøjagtigheder i hhv. AI-klassifikationerne og feltobservationerne.

MST's ønske om at inddrage de geografiske unøjagtigheder er ikke muligt med de eksisterende data, da det kræver et estimat på nøjagtigheden for både feltobservationerne og for AI-modelforudsigelserne i form af estimater for vegetationsændringerne inden for en given kort afstand fra observationerne.

Det skal understreges, at notatet primært handler om udviklingen af den statistiske valideringsmetode, og at de konkrete resultater i notatet er eksempler på metoden ud fra den foreliggende, tidligere version af AI-modellen. Der er siden bygget videre på AI-modellen, og disse ændringer er der selvsagt ikke taget højde for i beskrivelserne.

2 Feltbaseret kortlægning

Den feltbaserede kortlægning efter den traditionelle kortlægningsmetode (teknisk anvisning, TA N03) er udført på en anden skala end den opløsning, som AI-forudsigelserne kommer i (10*10 m). Det betyder, at man ikke kan bruge den hidtidige kortlægning som direkte reference for, hvor godt AI-modellen forudsiger, hvis man er interesseret i hvor godt AI-modellen fungerer på større rumlig skala, fx i et landskab eller inden for en hel NOVANA station. Der har derfor været behov for at indsamle nye feltdata til at validere AI-modellens forudsigelser, således at feltindsamlingerne er foretaget på en sammenlignelig størrelsesmæssig skala, i dette tilfælde en cirkel med 15 m radius. Hvis der ønskes viden om modellens forudsigelser på større skala end dette, fx for hele NOVANA-stationer eller andre større polygoner, er det relevant også at udføre den beskrevne statistiske validering med udgangspunkt i data for sådanne hele områder. Med udgangspunkt i 15 m cirklerne er det sandsynligt at valideringen vil vise, at modellen er dårligere til at forudsige de korrekte naturtyper for et givet, afgrænset område end den er på større rumlig skala.

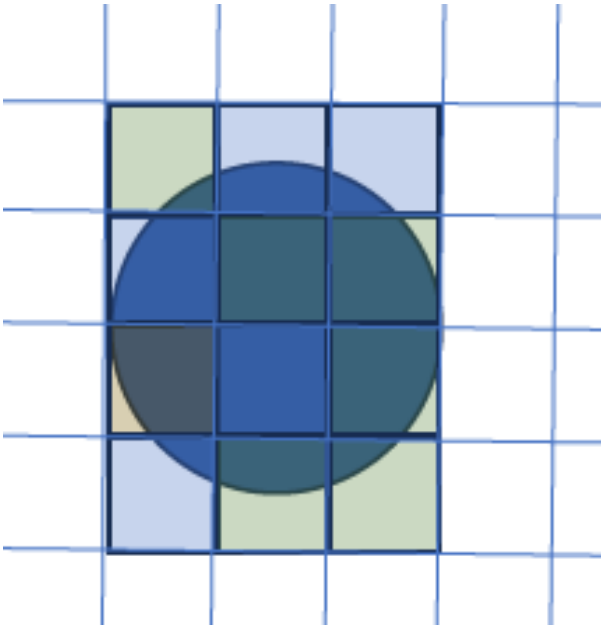
I 2021 blev der i samarbejde med Fagdatacenter for Biodiversitet og Terrestrisk Natur beskrevet en metode til indsamling af feltvalideringsdata (reference) og der blev lavet en teknisk anvisning til indsamling af feltvalideringer. I 2022 blev der indsamlet feltdata om habitatnaturtyperne efter denne tekniske anvisning.

Feltmetoden giver mulighed for, ved en heterogen vegetationssammensætning, at opdele den konkrete 15 m cirkel i delarealer, typisk op til tre delarealer, og kun undtagelsesvist op til fire delarealer, med angivelse af den procentvise udbredelse af habitatnaturtyperne. For at sikre, at tvivlstilfælde ved naturtypebestemmelserne ikke afhænger af inventørens valg, er der desuden givet mulighed for at kunne angive ligestillede, alternative naturtyper for hvert givet delareal.

3 AI-modellen

AI-modellen kombinerer adskillige GIS lag vedr. jordbund, hydrologi, geografisk placering og luftfotos til at klassificere habitatnaturtyperne. AI-modellen er under fortsat udvikling, og det er uden for dette notats ramme at give en yderligere beskrivelse af modellen.

AI-modellens estimering af habitatnaturtype baseres på de pixels (10x10m), som rammes af den kortlagte 15 m cirkel. Udbredelsen af de enkelte habitatnaturtyper i en 15 m cirkel bestemmes på basis af arealet af den del af pixels, der er inden for cirklen (se Figur 1).



Figur 1. Gridnet med 10 x 10 m pixels fra AI-modellens klassifikation. Indlagt er en 15 m cirkel fra feltobservationerne. I dette tilfælde berører i alt 12 pixels 15 m cirklen, og AI-modellen har klassificeret 6 grønne pixels, 5 blå pixels og 1 gul pixel. Med andre ord indgår 3 forskellige AI-klassifikationer fra AI-modellen. Fra feltobservationerne er angivet 1 eller flere observerede habitatnaturtyper med hver deres procentvise udbredelse i 15 m cirklen. Der har samtidig været mulighed for at angive alternative, ligestillede habitatnaturtyper i feltobservationerne for hver af de opdelte partier. I forhold til AI-modellens udbredelse af de tre habitatnaturtyper (grøn, blå og gul) beregnes de ud fra de mørkere dele af de 12 pixels, svarende til overlappet med 15 m cirklen.

4 Den statistiske valideringsmodel

Et perfekt match mellem AI-modellens forudsigelser af habitatnaturtype og dens udbredelse i 15 m cirklen og feltobservationernes angivelse af habitatnaturtype og dens udbredelse, dvs. hvor resultatet af de to modeller viser præcis samme resultat, ville betyde, at et plot, hvor de to datasæts procentvise udbredelse af habitatnaturtyperne i 15 m cirklen plottes mod hinanden, ville resultere i en ret linje med hældningen 1.

Da feltobservationerne er virkeligheden og dermed det, man ønsker information om og AI-modellen forsøger at reproducere virkeligheden, behandles feltobservationerne som sandheden ("ground-truth") dvs. i modelsammenhæng den afhængige variabel (responsvariablen) og AI-modellens forudsigelser, som prediktor-variabel.

Afvigelser fra perfekt overensstemmelse mellem de to datasæt kan skyldes flere forhold, der dels handler om AI-modellens evne til korrekt habitatnaturtypebestemmelse, og dels handler om modellens aggregering af resultaterne til et samlet resultat for 15 m cirklen. AI-modellen kan fx bestemme en eller flere af habitatnaturtyperne i 15 m cirklen til andre habitatnaturtyper end feltobservationerne. AI-modellen kan også estimere udbredelsen af de forskellige habitatnaturtyper i 15 m cirklen på en anden måde end feltobservationerne. Begge forhold vil give en afvigelse fra den perfekte hældning på 1.

Foruden de ovenstående årsager betyder den geografiske usikkerhed for en del af AI-modellens parametre, at modelresultaterne også har en betydelig geografisk usikkerhed. Hertil kommer at feltobservationerne indsamles med GPSer, som har en usikkerhed på op til 10 m. AI-modellen kan derfor i princippet klassificere et område, der ikke, eller kun delvist overlapper med den 15 m cirkel, som feltobservatøren har klassificeret. Det er dog vurderingen at den geografiske unøjagtighed ikke betyder det helt store, da der arbejdes på 10+m skala. Således kan selv unøjagtigheder i størrelsesordenen 10 m formentlig ikke rykke meget ved resultaterne, såfremt den typiske rumlige autokorrelation af habitatnaturtyperne i landskabet medregnes. Men det er klart, at det i nogle tilfælde, særligt i mosaiklandskaber, kan betyde noget.

4.1 Databehandling

AI-modellen valideres på to måder:

1: I første test testes det mere detaljerede fit mellem AI-modellens forudsigelse af de enkelte habitatnaturtypers procentuelle udbredelse i 15 m cirklen, sammenholdt med feltobservatørernes (se ovenfor).

2: I andet test testes om AI-modellen rammer rigtigt for bare én af de pixels (10x10m), der overlapper med cirklen.

I den samlede analyse (test 1), som angår habitatnaturtypernes udbredelse i 15 m cirklen, matches AI-modellens estimater procentuelt med feltobservationerne. Dvs. hvis AI-modellen siger, der er 15 % og 85 % af hhv. habitatnaturtype 1 og 2, så betragtes det kun som accepteret, hvis feltobservationerne har angivet de samme to habitatnaturtyper OG med samme procentuelle udbredelse i 15 m cirklen. For AI-modellen benyttes alle de pixels, der rammer

15 m cirklen, og for hver klassificeret habitatnaturtype blandt de pixels, der er indeholdt i 15 m cirklen, beregnes arealet. Kun de dele af pixels, der er indeholdt i 15 m cirklen medtages i arealberegningen (se Figur 1). Hvis der for en habitatnaturtype fx indgår tre pixels, der hver indgår med halvdelen af deres areal, så udgør det samlede areal 3 x 0,5 x 100 m2 = 150 m2. Hele cirklen udgør 707 m2, så i dette tilfælde udgør habitatnaturtypens udbredelse i AI-modellen altså 21 % af cirklen.

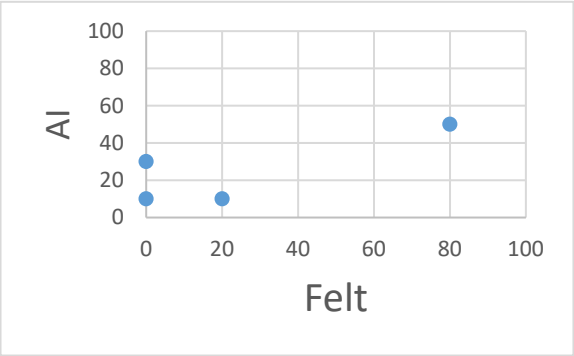
For hver 15 m cirkel har feltobservatørerne angivet udbredelsen af habitatnaturtyperne i felten. Men datasættet med feltobservationerne indeholder foruden feltobservatørernes angivelse af hvilke habitatnaturtyper, der indgår i 15 m cirklen, også op til to alternative angivelser af habitatnaturtyperne for hver forekomst i 15 m cirklen. Den første angivelse af habitatnaturtype betegnes som den primære feltobservation, det andet og tredje alternativ som hhv. den sekundære og tertiære feltobservation. Feltobservationerne af habitatnaturtyperne er ligestillede. Udbredelsen gælder altså for både den primære og evt. sekundære eller tertiære feltobservation. For hver 15 m cirkel checkes om enten primær, sekundær eller tertiær feltobservation for en udbredelse i 15 m cirklen matcher AI-modellens forudsigelse. Hvis den gør, opfattes det i denne valideringssammenhæng som en accepteret modelforudsigelse. På den baggrund genereres et datasæt, der for hver 15 m cirkel angiver om modellens forudsigelse er accepteret eller ej. Arealudbredelsen af AI-modellens pixels, der indgår i 15 m cirklen, beregnes for hver klassificeret habitatnaturtype og det tjekkes at summen af habitatnaturtypernes udbredelse giver 100 %, for at sikre at dette er gjort korrekt.

Klassifikationen af habitatnaturtyper (test 2) testes ved at sammenholde den eller de habitatnaturtyper, der er konstateret i feltobservationerne med AI-modellens forudsigelser. Fra feltobservationerne benyttes således alle de observerede habitatnaturtyper, uanset deres udbredelse i 15 m cirklen. Fra AI-forudsigelsessættet benyttes alle de klassifikationer af habitatnaturtyper, der foreligger fra alle de pixels, der måtte overlappe med 15 m cirklen. Typisk vil en cirkel med 15 m radius, og dermed 30 m diameter, overlappe med minimum 9 pixels a 10x10 m. Disse ni eller flere pixels kan altså have op til ni eller flere klassificerede habitatnaturtyper, der matches med feltobservationernes habitatnaturtyper fra den pågældende 15 m cirkel (se Figur 1). Hvis bare én af de habitatnaturtyper AI-modellen estimerer som værende tilstede rent faktisk også er fundet i felten inden for 15 m cirklen betragtes AI-modellens estimat i denne test som accepteret.

Box 1. *Eksempel på alignment mellem de to datasæt.*

I feltdata for en given 15 m cirkel er konstateret to delarealer på hhv. 20 % og 80 %, og for delarealet med 20 % har inventøren angivet naturtype A, eller alternativt naturtype B, og for delarealet med 80 % er angivet naturtype C.
AI-modellens forudsigelser omfatter 12 pixels, med naturtyperne A, C, D og E. De har procentvise udbredelser på hhv 10, 50, 30 og 10.
Når de to datasæt alignes, sammenstilles for hver naturtype de udbredelser naturtypen har i hhv det ene og det andet sæt.
Naturtype A (20, 10)
Naturtype B (20, 0)
Naturtype C (80, 50)
Naturtype D (0, 30)
Naturtype E (0, 10)

I de tilfælde, hvor feltdatasættet har alternative muligheder (her naturtype A og B), undersøges først om den primære naturtype (A) forekommer i AI-datasættet for 15 m cirklen. Hvis den gør, benyttes dette datapunkt, og de alternative muligheder udelades. Hvis den primære naturtype ikke forekommer i AI-datasættet, undersøges de alternative naturtyper successivt. Hvis ingen af feltdatasættets alternative naturtyper forekommer i AI-datasættet, benyttes den primære naturtype med udbredelsen 0 i AI-datasættet: A (x, 0).



Når feltobservationens angivelse af naturtypen med dens udbredelse (x-aksen) plottes mod AI-forudsigelserne (y-aksen) giver det i dette eksempel anledning til fire punkter på plottet, ét for hver af naturtyperne A, C, D og E.

De naturtyper, der kun optræder i feltbestemmelserne, ligger altså alle på x-aksen og de naturtyper, der kun optræder i AI-forudsigelserne, ligger alle på y-aksen. De naturtyper, der både optræder i feltobservationerne og i AI-forudsigelserne befinder sig i rummet udspændt af de to akser, jf. Figur 3.

Hvis der havde været et perfekt match, ville der kun have været to punkter på plottet, der enten kunne have været
Naturtype A (20, 20)
Naturtype C (80, 80)
Eller
Naturtype B (20, 20)
Naturtype C (80, 80)

Alle andre kombinationer, inklusive eksemplet ovenfor, ville give anledning til afvigelser fra den rette linje med hældningen 1.

4.2 Statistisk analyse

Det ovenfor beskrevne datasæt med primære, sekundære og tertiære feltobservationer, og AI-modellens forudsigelser for de berørte pixels, og deres procentuelle andel af 15 m cirklen, er udgangspunktet for de følgende analyser.

Samlet analyse (test 1)

For at teste overensstemmelsen mellem feltobservationerne og AI-modellens forudsigelse af habitatnaturtyper og opdelingen af hver 15 m cirkel har vi udført nedenfor beskrevne analyse. Udgangspunktet for analysen var, at ved et perfekt match mellem feltobservationerne og AI-modellens forudsigelser (dvs. AI-modellen rammer samme procent af en given habitatnaturtype som angivet i felten) ville hældningen – hvis disse plottedes mod hinanden – være 1. Dette testedes med en random coefficient mixed model, hvor feltobservation var respons variabel og AI-forudsigelse var prediktor variabel. Den enkelte 15 m cirkel indgik i modellen som subject for at tage højde for den afhængighed, der var mellem observationerne inden for hver 15 m cirkel Vi brugte her en unbiased covarians struktur. Denne model vægtede alle observationer ligeligt, uanset habitatnaturtypens udbredelse. Det betød at alle observationer/ forudsigelser talte ligeligt, uanset habitatnaturtypens udbredelse.

For at tage hensyn til, at AI-modellen muligvis havde sværere ved at detektere habitatnaturtyper med lille udbredelse, kørte vi også modellen på et reduceret datasæt, hvor vi frasorterede alle de datapunkter, hvor naturtypen ikke var at finde i enten feltdatasættet eller i AI-datasættet. Med andre ord, hvor udbredelsen af naturtypen var lig nul. En udbredelse lig nul i AI-datasættet betyder, at habitatnaturtypen er bestemt i feltobservationerne, men er ikke klassificeret i AI-modellen. Jævnfør beskrivelsen i Box 1, betyder det at ingen af de alternative naturtyper, der er angivet for det pågældende delareal, i feltobservationerne er at finde i AI-klassifikationerne. Omvendt betyder en udbredelse lig nul i feltobservationerne, at habitatnaturtypen er klassificeret i ét eller flere pixels i AI-modellen, men er ikke fundet i felten. Foruden udbredelse lig nul blev udbredelser mindre eller lig hhv. 5 % og 20 % udeladt i senere step i analyserne for hhv. feltobservationerne og for AI-forudsigelserne. Hvis disse udeladelser fra datasættet betyder at hældningen for relationen mellem feltobservationerne og AI-modellens forudsigelser bliver tættere på 1 end det fulde datasæt, så indikerer det, at AI-modellen enten ikke kan detektere habitatnaturtyper med lille udbredelse, eller at AI-modellen klassificerer en række naturtyper, der ikke er fundet i feltobservationerne.

Test af habitatnaturtypebestemmelsen og deres udbredelse kan gennemføres naturtypevist, og dermed får man afklaret systemets styrker og svagheder i forhold til de enkelte habitatnaturtyper. Denne mulighed er ikke præsenteret i dette notat af hensyn til overskueligheden og omfanget af notatet, men bør selvfølgelig indgå i den videre udvikling og validering af systemet.

Der er i denne analyse ikke beregnet en adjusted R^2, der kunne kvantificere, hvor stor en del af variationen i feltobservationerne, der kunne forklares med AI-observationerne (som beskrevet her: https://www.ars.usda.gov/ARSUserFiles/80000000/SpatialWorkshop/19kramersupplrsq.pdf). En evt. opfølgende analyse vil kunne levere denne information.

Klassifikation af habitatnaturtyper (test 2)

For at undersøge om AI-modellen havde sværere ved at identificere specifikke habitatnaturtyper testedes det om korrekte og ikke-korrekte vegetationsforudsigelser afhang af habitatnaturtypen. Dette testedes med en generalized linear model. Ideelt skulle denne model have inkluderet landsdel (Bornholm, Nordjylland, Østjylland+ Fyn, Vestjylland, Sjælland) som tilfældig effekt, men modellen kunne i så fald ikke konvergere og derfor undlod vi det. Det betyder, at modellen ikke tager hensyn til eventuelle regionale forskelle i forudsigelsessikkerhed. Årsagen til at modellen ikke kunne konvergere med region som tilfældig effekt var formentlig, at nogle habitatnaturtyper ikke var repræsenteret i alle regioner. Klassifikationen af habitatnaturtyper testedes også med en chi^2 test, hvor antallet af korrekte og ikke-korrekte forudsigelser per habitatnaturtype testedes for at kunne identificere de habitatnaturtyper, som AI-modellen enten identificerede for få eller for mange af i forhold til feltobservationerne. Antagelser vedrørende uafhængighed af observationer fra samme 15 m cirkel kan ikke opfyldes, men analysen giver dog stadig et bud på hvilke habitatnaturtyper AI modellen har svært ved at klassificere.

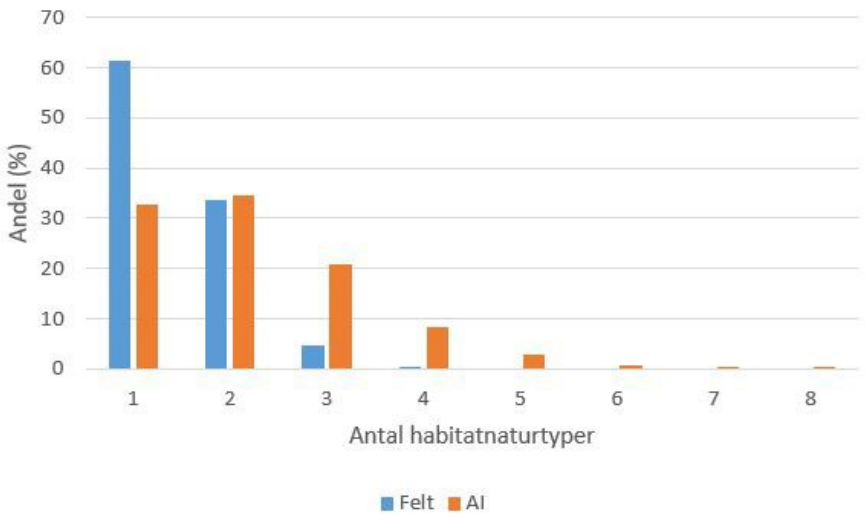
Vi testede endvidere, om sandsynligheden for at en habitatnaturtype klassificeredes korrekt afhang af udbredelsen af habitatnaturtypen. Til dette brugte vi en generalized linear model, hvor sandsynligheden for at en variabel klassificeres korrekt følger en binomial fordeling og habitatnaturtypens udbredelse (kontinuert) var den uafhængige variabel. Desuden indgik region som en tilfældig effekt i modellen.

Analyserne er lavet i SAS 9.4 (SASInstitute, Cary, NC) under procedurerne proc mixed, proc freq, proc glm og proc glimmix.

5 Eksempel på en dataanalyse

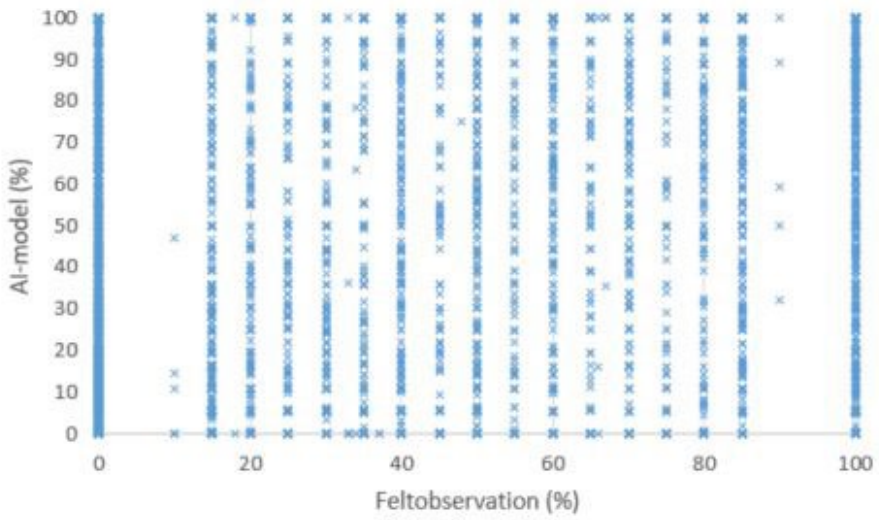
For at illustrere analysen har vi her brugt et datasæt fra en tidligere version af AI-modellen [Version 4, downloaded 25/1-23]. Bemærk, at det udtrukne datasæt ikke indeholder habitatnaturtype=2 (ikke en habitatnaturtype). Denne er derfor også udeladt fra feltobservationerne for at kunne opnå sammenlignelige datasæt. I fremtidige valideringer bør habitatnaturtype 2 også indgå, da dette tester modellens evne til at skelne habitatnatur fra ikke-habitatnatur. Dette har dog ikke nogen praktisk betydning for udviklingen af valideringsmetoden, der er bare en habitatnaturtype mindre.

Det kombinerede datasæt består af 9808 observationer fra 3828 15 m cirkler fordelt på 48 habitatnaturtyper. Der er 41 habitatnaturtyper i feltobservationerne og 48 habitatnaturtyper i AI-forudsigelsen. I feltobservationerne fordeler antallet af habitatnaturtyper per 15 m cirkel sig lidt forskelligt fra AI-modellen, idet andelen af 15 m cirkler med kun én habitatnaturtype er lavere for AI-forudsigelserne end for feltobservationerne (Figur 2, chi^2 =996, df=5, $p<0.001$).

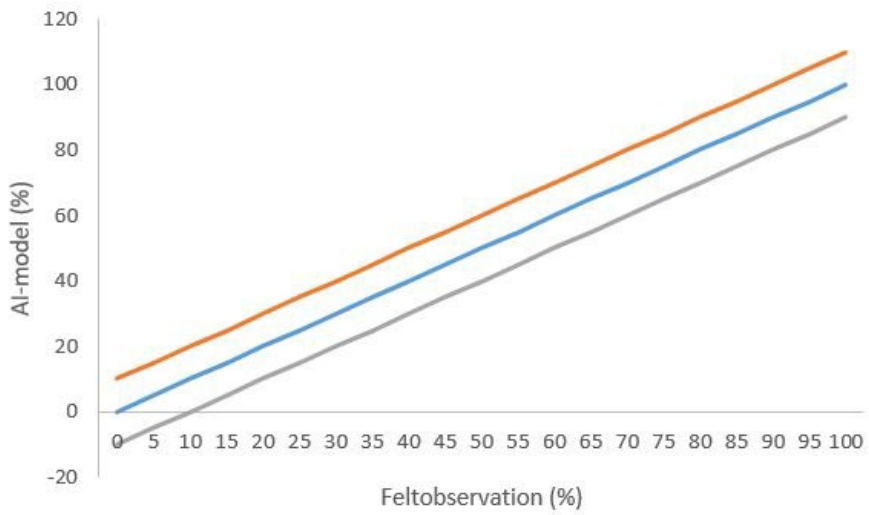


Figur 2. Andel af 15 m cirkler i forhold til antallet af habitatnaturtyper, der indgår i cirklen for hhv feltobservationerne (blå) og AI-forudsigelserne (orange). Der er i feltobservationerne mulighed for at opdele cirklen i op til tre delarealer og undtagelsesvist op til fire, med hver deres alternative habitatnaturtyper, og for AI-forudsigelserne indgår typisk 9-16 pixels, med tilsvarende antal mulige forskellige habitatnaturtyper.

Figur 3. Scatterplot for AI-modellens forudsigelser af habitatnaturtype OG dens udbredelse versus feltobservationernes angivelser. Habitatnaturtyper i en 15 m cirkel, der kun er fundet i feltobservationerne, men ikke i AI-modellen, ligger alle på x-aksen (x, 0). Habitatnaturtyper, der ikke er fundet i felten, men klassificeret af AI-modellen, ligger alle på y-aksen (0, y). Alle de habitatnaturtyper, der både er fundet i felten og i AI-modellen er fordelt mellem de to akser (x,y). Hvis feltobservationen fx har angivet at habitatnaturtype 4030 med en udbredelse på 40 % af 15 m cirklen, og AI-modellen har estimeret at 4030 udgør 25 % af 15 m cirklen, så vil dette match give et punkt på (40, 25) i koordinatsystemet. Hvis habitatnaturtypen udgør hele 15 m cirklen i feltobservationen ligger punktet yderst til højre på figuren (100, y), og hvis den også udgør hele AI-modellens forudsigelse, så ligger punktet allerøverst i højre side af figuren (100,100). Bemærk de mange punkter ved feltobservation = 0 %, altså på Y-aksen, der svarer til, at AI-modellen har angivet én eller flere pixels med habitatnaturtyper, der ikke er fundet i feltobservationerne for den givne 15 m cirkel.



Figur 4. Ideel hældning for relationen (blå linje) mellem feltobservationer og AI-modellen samt fejllinjer (orange og grå), der er forskudt 10 enheder (%) fra den ideelle relation. Disse fejllinjer illustrerer det rum som en afvigelse på 10 procentpoint tillader. Hvis AI-modellens forudsigelser generelt ligger nogenlunde inden for disse forskudte linjer så indikerer det en god relation. Er man villig til at acceptere en større afvigelse end 10 enheder, så vil disse fejllinjer befinde sig længere fra ideallinjen (blå) og formentlig inkludere flere observationer.



Hvis der er en god overensstemmelse mellem feltobservationer og AI-modellen forventes hældningen i Figur 3 at være tæt på 1, dvs. at hovedparten af punkterne skal helst ligge nogenlunde inden for de forskudte linjer, der angiver et usikkerhedsinterval (Figur 4). Den udgave af AI-modellen, der bruges i denne rapport, er meget langt fra idealet (Figur 3). Hvis habitatnaturtyper med stor udbredelse er nemmere for AI-modellen at detektere, så vil der være mindre variation for feltobservationer med udbredelser større end 50 % end for udbredelser mindre end 50 %. Grafer som Figur 3 og 4 er således meget anvendelige til hurtigt at danne et overblik over, hvor godt AI-modellen overordnet performer.

5.1 Samlet analyse (test 1)

I dette eksempel er der betydelige uoverensstemmelser mellem feltobservationerne og AI-modellens forudsigelse af enkelte habitatnaturtyper og deres udbredelse (Figur 3). Alle de observationer, der ikke ligger på diagonalen med hældning lig 1, er eksempler på habitatnaturtyper, hvor AI-modellen enten har klassificeret habitatnaturtypen forkert, eller udbredelsen afviger fra feltobservationernes angivelser.

Andelen af habitatnaturtyper, som findes i feltobservationerne, men som AI-modellen ikke finder, og vice versa (nul-udbredelser), for hhv. AI-forudsigelser og feltobservationer, illustrerer AI-modellens evne til at detektere habitatnaturtyper. Der er 1605 (16,4 %) nul-udbredelser for AI-modellen (Figur 3), dvs. at i 16,4 % af feltobservationerne findes der ikke den tilsvarende habitatnaturtype i AI-modellen. Tilsvarende er der 4474 (45,6 %) nul-udbredelser for feltobservationerne, hvilket svarer til, at i 45,6 % af AI-forudsigelserne, estimerer AI-modellen en habitatnaturtype, der ifølge feltobservationerne ikke findes i den tilsvarende 15 m cirkel.

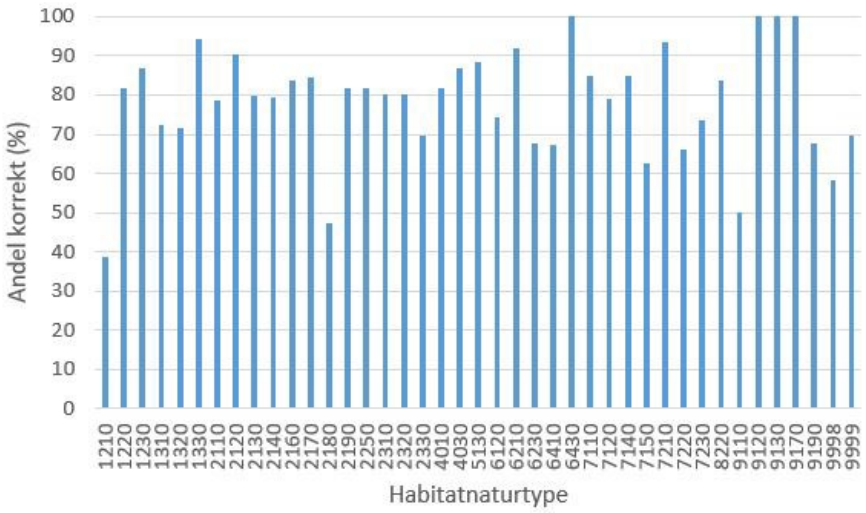
Overensstemmelsen mellem feltobservationerne og AI-modellens forudsigelser er 0,29 (Tabel 1, Figur 3). Det skyldes tilsyneladende tildels, at AI-modellen ikke formår at detektere habitatnaturtyper med lille udbredelse, da udeladelse af habitatnaturtyper med lille udbredelse i feltobservationerne forbedrer relationen mellem AI-forudsigelserne og feltobservationerne (hhv. 0,46 og 0,47, Tabel 1). Desuden viser det sig, at når habitatnaturtyper med lille udbredelse i AI-forudsigelserne udelades, så forbedres relationen mellem feltobservationer og AI-observationer til 0,45 (Tabel 1). Med andre ord, så klassificerer AI-modellen en lang række habitatnaturtyper, der ikke findes i virkeligheden (feltobservationerne).

Tabel 1. Random coefficient mixed model for relationen mellem feltobservationer og AI-modellens forudsigelser. Et estimat på 1.00 svarer til en fuldstændig overensstemmelse mellem feltobservationer og AI-forudsigelser. Resultatet af testen er vist for hhv. det fulde datasæt og for datasæt, hvor habitatnaturtyper, der ikke detekteres (0 %) eller med lave dækningsgrader (<=5 % eller <=20 %) i enten feltobservationerne eller i AI-modellens forudsigelser udelades. Df_{errror}: Antallet af frihedsgrader for residualerne, SE: standard error

	$df_{residual}$	F	P	Estimat	SE
Fuldt datasæt	3000	877.6	<.0001	0.29	0.01
Feltobservationer >0 % (0, y)	1314	779.2.3	<0.001	0.46	0.02
Feltobservationer >20 % (0.2, y)	672	438.5	<.0001	0.47	0.02
AI-forudsigelser, hvor AI>0% (x, 0)	2184	2486.7	<0.001	0.45	0.01
AI-forudsigelser, hvor AI > 5% (0, >0.05)	1948	2021.1	<.0001	0.41	0.01
AI-forudsigelser, hvor AI > 20% (0, >0.20)	925	550.9	<.0001	0.19	0.01

5.2 Identifikation af habitatnaturtyper (test 2)

Når man tester hvor god AI-modellen er til at ramme rigtigt for bare én af de pixels, der er i en 15 m cirkel, bliver det klart, at der er forskel på hvor succesfuldt AI-modellen kan identificere habitatnaturtyperne. I Figur 5 er for hver habitatnaturtype angivet hvor mange procent af 15 m cirklerne, hvor der er fundet en overensstemmelse mellem en af de habitatnaturtyper, der er angivet i feltobservationerne, både for arealandelene og de evt. alternative habitatnaturtyper, og blot én af de pixels, der berører 15 m cirklen.



Figur 5. Andel af AI-forudsigelser, hvor habitatnaturtypen klassificeres på samme måde som i datasættet med feltobservationer. En korrekt klassifikation er hvis blot én af AI-modellens pixels, der berører 15 m cirklen, er i overensstemmelse med en af de habitatnaturtyper, der er angivet i feltobservationen. Tallene der ligger til grund for denne graf findes i Appendiks.

Andelen af habitatnaturtyper, der klassificeres korrekt, varierer signifikant mellem habitatnaturtyperne (Generalized linear model $F_{40, 6287}$=7.04, p<0.001). I denne version af AI-modellen er habitatnaturtyperne 1210, 2180 og 9110 alle under 50 % (Figur 5). Da denne analyse kun repræsenterer én version af AI-modellen, foretages der ikke yderligere analyse på post hoc pairwise comparisons for at identificere de habitatnaturtyper, der klassificeres signifikant dårligere eller bedre end de øvrige. Denne analyse kunne dog foretages med en Least square means comparison eller en Tukey post hoc test (Zar 2004).

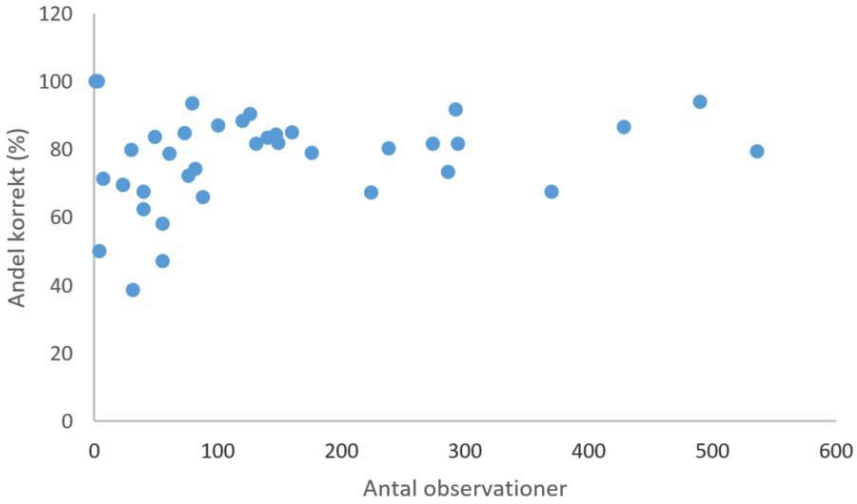
Fordelingen af korrekte forudsigelser kan også testes med en chi^2-test. Denne tester om habitatnaturtyperne fordeles tilfældigt i forhold til række og søjlesummer, med andre ord, om de fordeles tilfældigt uafhængigt af antallet af 15 m cirkler med naturtypen i feltdatasættet og antallet af 15 m cirkler med naturtypen i AI-datasættet. Andelen af habitatnaturtyper, der klassificeres korrekt, er ikke tilfældigt fordelt mellem habitatnaturtyperne (Chi^2=324, df= 41, p<0.001). For nogle habitatnaturtyper er der flere forudsigelser af habitatnaturtypen i AI-modellen end der er observeret i feltobservationerne. Det gælder især habitatnaturtyperne surt overdrev (6230) og tidvis våd eng (6410) (Figur 6). Hvorimod der i AI-modellen er færre 15 m cirkler med habitatnaturtyperne strandeng (1330), bræmmer (4030) og kalkoverdrev (6210) end der forventes ud fra fordelingen af feltobservationerne (Figur 6). Dette er således de habitatnaturtyper, som AI-modellen har sværere ved at detektere end andre.

Figur 6. Forskel mellem antallet af observerede og forventede antal korrekte observationer of hver habitatnaturtype fra chi^2 testen. Positive værdier indikerer at AI-modellen predikterer flere korrekte observationer af den pågældende habitatnaturtype end det forventes ud fra fordelingen af korrekte versus ikke korrekte observationer. Negative værdier indikerer at AI-modellen har færre korrekte identifikationer af den pågældende habitatnaturtype end det forventes ud fra fordelingen af korrekte versus ikke korrekte observationer.

5.3 Hvor meget skal der samples?

Andelen af AI-modellens korrekte forudsigelser per habitatnaturtype er tilsyneladende delvist relateret til antallet af feltobservationer, således at habitatnaturtyper med et lavt antal feltobservationer har en lavere andel af korrekte observationer (Figur 7). For at kunne validere AI-modellens mulighed for at klassificere en habitatnaturtype bør der derfor minimum være omtrent 100 observationer af hver habitatnaturtype.

Figur 7. Andelen af korrekte AI-forudsigelser per habitatnaturtype i relation til andelen af feltobservationer.



5.4 Arealafhængighed

En sammenhæng mellem andelen af korrekt klassificerede habitatnaturtyper i AI-modellen og antallet af feltobservationer indikerer at sandsynligheden for at en habitatnaturtype klassificeres korrekt af AI-modellen forøges signifikant med arealudbredelsen af habitatnaturtypens udbredelse i feltobservationerne (Generalized linear mixed model $F_{1,\ 6322}$=158.8, p<0.001, estimat=0.014).

6 Konklusion

Den samlede analyse baseret på en random coefficient mixed model viser, at AI-modellen tilsyneladende klassificerer flere habitatnaturtyper end der forekommer i feltobservationerne. Overensstemmelsen mellem feltobservationerne og AI-modellens forudsigelser forbedres derfor, når man ser bort fra de habitatnaturtyper, der enten ikke forekommer, eller har meget lav udbredelse, i feltobservationerne. Dette ses blandt andet ved, at overensstemmelsen mellem AI-modellens forudsigelser og feltobservationerne forbedres fra 0,29 til 0,47, når habitatnaturtyper, der ikke forekommer i feltobservationerne, udelades.

Udviklingen af egnede analysemetoder omfatter muligheden for at køre disse analyser habitatnaturtypevist. Dermed får man afklaret systemets styrker og svagheder i forhold til de enkelte habitatnaturtyper. Denne mulighed er ikke præsenteret i dette notat af hensyn til overskueligheden og omfanget af notatet, men bør selvfølgelig indgå i den videre udvikling og validering af systemet.

Resultaterne viser også, at der bør indsamles minimum 100 observationer af hver habitatnaturtype i felten, hvis valideringen skal være så troværdig som mulig og uafhængig af valideringsdatamængden. Dette resultat kan være en konsekvens af, at nogle habitatnaturtyper er sjældne og måske sværere for AI-modellen at detektere.

Den geografiske usikkerhed på både AI-forudsigelserne og feltobservationerne introducerer muligheden for, at det ikke er præcis de samme områder, der observeres i feltobservationerne og i AI-forudsigelserne. Denne usikkerhed kan disse analyser ikke kvantificere og dermed heller ikke tage højde for.

I evalueringen af AI-modellen bør det tages i betragtning, hvilken rumlig skala AI-forudsigelserne ønskes anvendt på. For det analyserede datasæt er det tydeligt, at AI-modellen er bedre til at identificere habitatnaturtyper med større udbredelse end de habitatnaturtyper, der har en relativt lille udbredelse. Så hvis AI-modellen skal håndtere habitatnaturtyper med stor udbredelse, så er den mere velegnet, end hvis den også skal kunne opfange habitatnaturtyper med lille udbredelse. Det er således sandsynligt at evaluering af AI-modellen på et feltdatasæt, som er indsamlet på større rumlig skala (fx 100 x 100 m) ville vise, at modellen er bedre end ovenstående eksempel viser. Omvendt skal modellen også fungere i mosaiksamfund, som fx klitter med klitlavninger, som typisk har ganske lille udbredelse og det taler for at evaluere modellen på den rumlige skala, som det er gjort her.

En anden mulighed er at foretage en aggregering af de mange informationer i hver 15 m cirkel, både i feltdatasættet og i AI-datasættet, fx ved at teste forekomsten af den dominerende habitatnaturtype i den pågældende 15 m cirkel, ved forinden at foretage en udvælgelse heraf, både i feltdatasættet og i AI-klassifikationerne. Det vil selvfølgelig ignorere den faktisk observerede variation i feltobservationerne, og ignorere de mange detaljer AI-klassifikationerne tilbyder i form af pixelstørrelser på 10x10 m, men det vil givetvis give en mere realistisk, og sandsynligvis bedre vurdering af AI-systemets evne til at klassificere på en grovere skala.

Til evaluering af andre udgaver af AI-modellen vil estimatet af hældningen for den samlede analyse og evt. et estimat af R^2 være illustrative indikatorer. Hvis analyserne på den samlede model indikerer et mindre godt fit, så kan analyserne af afhængigheden af arealudbredelse, og sandsynligheden for korrekt forudsigelse af den enkelte habitatnaturtype bruges til at detektere hvor AI-modellen fejler.

7 Litteratur

Zar. J.H 2005. Biostatistical analysis 4th edition. Prentice Hall.

8 Appendix

Tabel A1. Oversigt over antallet af overensstemmende og ikke-overensstemmende klassifikationer mellem AI forudsigelser og feltobservationer for hver habitatnaturtype.

Habitatnaturtype	*Ikke overens-stemmende klassificering*	*Overensstem-mende klassifi-cering*	*Total*
1210	19	12	31
1220	27	122	149
1230	13	87	100
1310	21	55	76
1320	2	5	7
1330	29	461	490
2110	13	48	61
2120	12	114	126
2130	166	657	823
2140	110	426	536
2160	23	117	140
2170	23	124	147
2180	29	26	55
2190	54	240	294
2250	24	107	131
2310	6	24	30
2320	47	191	238
2330	7	16	23
4010	50	224	274
4030	57	371	428
5130	14	106	120
6120	21	61	82
6210	24	268	292
6230	120	250	370
6410	73	151	224
6430	0	2	2
7110	11	62	73
7120	37	139	176
7140	24	136	160
7150	15	25	40
7210	5	74	79
7220	30	58	88
7230	76	210	286
8220	8	41	49
9110	2	2	4
9120	0	1	1
9130	0	2	2
9170	0	3	3
9190	13	27	40
9998	23	32	55
9999	7	16	23