



SØER UDEN VELDEFINEREDE AFLØB – TOPOGRAFISK OPLAND OG INDSATSER

Videnskabelig rapport fra DCE – Nationalt Center for Miljø og Energi

nr. 698

2026



AARHUS
UNIVERSITET

DCE – NATIONALT CENTER FOR MILJØ OG ENERGI

SØER UDEN VELDEFINEREDE AFLØB – TOPOGRAFISK OPLAND OG INDSATSER

Videnskabelig rapport fra DCE – Nationalt Center for Miljø og Energi

nr. 698

2026

Tuba Bucak
Robert Ladwig
Eti E. Levi
Liselotte Sander Johansson
Ane Kjeldgaard
Martin Søndergaard

Aarhus Universitet, Institut for Ecoscience



AARHUS
UNIVERSITET

DCE – NATIONALT CENTER FOR MILJØ OG ENERGI

Datablad

Serietitel og nummer:	Videnskabelig rapport fra DCE - Nationalt Center for Miljø og Energi nr. 698
Kategori:	Rådgivningsrapporter
Titel:	Søer uden veldefinerede afløb – topografisk opland og indsatser
Forfattere:	Tuba Bucak, Robert Ladwig, Eti E. Levi, Liselotte Sander Johansson, Ane Kjeldgaard, Martin Søndergaard
Institution:	Aarhus Universitet, Institut for Ecoscience
Udgiver:	Aarhus Universitet, DCE – Nationalt Center for Miljø og Energi ©
URL:	https://dce.au.dk
Udgivelsesår:	Februar 2026
Redaktion afsluttet:	12. februar 2026
Faglig kommentering:	Mikkel Boel
Kvalitetssikring, DCE:	Signe Jung-Madsen
Sproglig kvalitetssikring:	Anne Mette Poulsen
Ekstern kommentering:	Kommentarerne findes her:
Finansiel støtte:	Styrelsen for Grøn Arealomlægning og Vandmiljø (SGAV)
Bedes citeret:	Bucak, T., Ladwig, R., Levi, E.E., Johansson, L.S., Kjeldgaard, A., Søndergaard, M. 2026. Søer uden veldefinerede afløb – topografisk opland og indsatser. Aarhus Universitet, DCE – Nationalt Center for Miljø og Energi, 47 s. - Videnskabelig rapport nr. 698
	Gengivelse tilladt med tydelig kildeangivelse
Sammenfatning:	For mange af de danske søer, der indgår i Vandområdeplan 2021-2027 (VP3), er der ingen veldefinerede afløb. Dette umuliggør belastningsberegninger og planlægning af målrettede restaureringstiltag i søerne, selv om de ikke opfylder målet om mindst god økologisk tilstand. Ved hjælp af multivariate statistiske metoder (RDA) og maskinlæringsmetoder (Random Forest og Conditional Inference Trees) blev der i dette projekt undersøgt, hvordan landskabskarakteristika påvirker søers vandkvalitet, økologiske tilstand og typer af næringsstofbegrænsning. Alle metoder udviste lav forklaringsværdi, hvilket indikerer, at søernes tilstand kun delvist kan forklares af de undersøgte oplandskarakteristika og sandsynligvis også påvirkes af f.eks. intern næringsstofbelastning, historiske påvirkninger og hydrologiske processer. Ikke desto mindre var søer i lav økologisk tilstand konsekvent forbundet med jorde med højt organisk indhold, intensiv dræning og relativt våde oplande, mens større andele af ikke-dyrkede arealer og større sødybde var forbundet med bedre økologisk tilstand. Resultaterne peger på, at reduktion af kunstig dræning og beskyttelse af naturlige arealer kan bidrage til forbedringer i søers vandkvalitet og økologisk tilstand.
Emneord:	Opland; søafløb; økologisk tilstand; næringsstofbegrænsning; landskabskarakteristika; dræning; Random Forest; Conditional Inference Trees; RDA
Foto forside:	Sem Sø ved Skrødstrup, Region Nordjylland (Kilde: GeoDanmark Ortofoto (forår 2025))
ISBN:	978-87-7648-029-5
ISSN (elektronisk):	2244-9981
Sideantal:	47
Supplerende oplysninger:	Rapporten er også udgivet på engelsk – se den her

Indhold

Forord	5
Sammenfatning	6
Summary	8
1 Baggrund og formål	10
1.1 Baggrund	10
1.2 Formål	10
2 Data	11
2.1 Data fra søerne	11
2.2 Landskabskarakteristika	11
3 Metoder	14
3.1 Klassificering af næringsstofbegrænsning	14
3.2 Udtræk af landskabskarakteristika	14
3.3 Redundansanalyse (RDA)	14
3.4 Random forest-analyser	16
3.5 Conditional inference trees – beslutningstræer	16
4 Resultater	18
4.1 Overblik over søerne	18
4.2 Oversigt over landskabskarakteristika	19
4.3 Effekter af landskabskarakteristika på søernes tilstand	22
4.4 N- eller P-begrænsning	33
4.5 Palæøkologiske undersøgelser	35
5 Konklusioner og anbefalinger	38
Anbefalinger	40
6 Referencer	41
7 Bilag	43

Forord

Denne rapport undersøger, hvordan landskabskarakteristika kan påvirke danske søer uden veldefinerede topografiske oplande. Analyserne er gennemført både for søernes estimerede oplande og for forskellige bufferzoner omkring søerne for at belyse sammenhængen mellem landskabskarakteristika, søernes vandkvalitet og økologiske tilstand. Rapporten er udarbejdet af Aarhus Universitet på foranledning af Styrelsen for Grøn Arealomlægning og Vandmiljø (SGAV), som har haft mulighed for at kommentere et udkast til rapporten.

Som en del af projektleverancen har SGAV, udover denne rapport, modtaget baggrundsdata og beregninger, som rapporten bygger på for hver enkelt sø.

Rapporten er delvist en opdatering af en tidligere publikation af Søndergaard et al. (2023).

Sammenfatning

Mange søer i Danmark, som er omfattet af vandområdeplan 2021–2027 (VP3), i alt 418 ud af 985 søer, har ikke et klart defineret opland på grund af manglende eller dårligt definerede tilløb og afløb. Som følge heraf har det ikke været muligt at gennemføre belastningsopgørelser og dermed heller ikke at estimere den nødvendige indsats for at forbedre disse søers økologiske tilstand, selv om størstedelen af dem ikke opfylder målet om god økologisk tilstand. Et tidligere studie udført af Søndergaard et al. (2023) viste, at der kun var en svag sammenhæng mellem de nuværende oplandskarakteristika og den økologiske tilstand i disse søer.

På baggrund af ovenstående er analyserne i dette studie blevet opdateret i forhold til Søndergaard et al. (2023) ved anvendelse af en kombination af multivariate statistiske metoder (reduktionsanalyse, RDA) og maskinlæringsmetoder (Random Forest, RF og Conditional Inference Trees, CIT) med henblik på at undersøge vigtigheden af landskabskarakteristika for vandkvalitet, økologisk tilstand og klassificering af næringsstofbegrænsning. Tilsammen giver disse metoder komplementære indsigter i de potentielle effekter af landskabskarakteristika på søernes næringsstofniveauer og økologiske tilstand. Derudover inkluderede analyserne flere størrelser af bufferzoner omkring søerne (10, 50, 100 og 250 m) for at kunne sammenligne med – eller understøtte – resultaterne baseret på søernes estimerede oplande.

Alle tre analytiske tilgange havde en forholdsvis lav forklaringsevne, hvilket stemmer overens med det tidligere studie baseret på lineær regression (Søndergaard et al., 2023). Den lave forklaringsevne tyder på, at andre faktorer også spiller ind, såsom intern næringsstofbelastning, historiske næringstilførsler, vandsøjlels opblandingsdynamik og grundvandsbidrag, som ikke opfanges ved overfladebaserede oplandsafgrænsninger og -karakteristik.

Alle tilgange identificerede de samme landskabskarakteristika som centrale determinanter for søernes økologiske tilstand (baseret på næringsstofkoncentrationer eller økologisk status). Dårligere økologisk tilstand (højere næringsstofkoncentrationer og ringe/dårlig økologisk status) var forbundet med højere andele af organiske jordtyper, større drænedes arealer og vådere oplande. Omvendt var større roddybde og højere andele af ikke-dyrkede arealer forbundet med bedre økologisk tilstand.

RF indikerede endvidere, at dybere og større søer med lav sammenhæng mellem vandløbsafstrømning og nedbør er karakteriseret ved en forholdsvis god økologisk tilstand. CIT understøttede disse resultater og identificerede klare tærskler, idet dybere søer (især >5,7 m) eller lavvandede søer med >60 % ikke-dyrkede arealer havde lavere næringsstofkoncentrationer og en relativt god økologisk tilstand. CIT viste desuden, at fosforbegrænsede søer hyppigst forekom i sandede, svagt drænedes oplande, og kvælstofbegrænsede søer i stærkt drænedes systemer.

Samlet set viser resultaterne, at selv om landskabskarakteristika kun forklarer en del af variationen i søernes tilstand, spiller processer på oplandsniveau alligevel en væsentlig rolle for næringsstofniveauer og økologisk tilstand i tilsyneladende hydrologisk isolerede søer. Projektets resultater viser, at søer uden veldefinerede afløb kan have gavn af, at kunstig dræning bliver

reduceret samt af at øge eller beskytte naturlige (ikke-dyrkede) arealer. Da højere næringsstofniveauer er forbundet med intensiv dræning og organiske jordtyper, kan søer beliggende i sådanne oplande forventes at reagere bedre på målrettede indsatser. Andre faktorer som fx intern belastning, samspil med grundvand og søspecifikke processer vil dog sandsynligvis også være af betydning. Palæolimnologiske analyser kan bidrage til at identificere historiske drivkræfter bag økologiske ændringer (og dermed forklare noget af den resterende uforklarede variation), hvilket kan understøtte beslutningsprocesser og mere målrettede indsatser.

Summary

Many lakes in Denmark included in the 3rd River Basin Management Plan (VP3) 2021-2027, 418 out of 985 lakes, lack a clearly defined catchment due to missing or poorly defined inflows and outflows. As a result, load assessments could not be carried out, and thus it has not been possible to estimate the effort needed to improve their ecological conditions, even though the majority of these lakes fail to meet the environmental target of good ecological status. A previous study by Søndergaard et al. (2023) found that the current catchment characteristics and ecological status of these lakes only showed weak correlations.

Considering the above information, the analysis in the current study was updated compared to Søndergaard et al. (2023) using a combination of multivariate statistical (redundancy analysis, RDA) and machine learning approaches (Random Forest, RF, and Conditional Inference Trees, CIT) to investigate the influence of landscape characteristics on water quality, ecological status, and nutrient limitation classification. Together, these methods provide complementary insights into the potential effects of landscape features on lake nutrients and ecological status. Additionally, the analyses were performed using several buffer zones (10, 50, 100, and 250 m) to compare with, or support, the results obtained from the estimated catchment area.

All three analytical approaches had relatively low explanatory power, which was consistent with the previous linear regression-based study (Søndergaard et al., 2023). This low explanatory power likely points to the influence of additional factors, such as internal nutrient loading, historical nutrient inputs, mixing dynamics, and groundwater inputs, which were not captured by surface-based catchment delineations.

All approaches identified the same set of landscape characteristics as key determinants of lake ecological conditions (based on nutrient concentrations or ecological status). Poorer ecological conditions (higher nutrient concentrations and Poor/Bad ecological status) were associated with higher proportions of organic soils, larger drained areas, and wetter catchments. In contrast, greater root depth and larger proportions of non-cultivated land were linked with improved ecological conditions.

RF further indicated that deeper and larger lakes with low streamflow precipitation elasticity are characterized by relatively good ecological conditions. CIT supported these findings and suggested clear thresholds, showing that deeper lakes (especially >5.7 m), or shallower lakes with >60% non-cultivated land, had lower nutrient concentrations and relatively good ecological status. CIT also showed that P-limited lakes were most common in sandy, weakly drained catchments, whereas N-limited lakes occurred in heavily drained systems.

In summary, although landscape characteristics explain only part of the variation in lake conditions, catchment-scale processes still play a major role in shaping nutrient levels and ecological status in hydrologically isolated lakes. Based on our results, lakes without well-defined outlets could benefit from reducing artificial drainage and increasing or protecting natural (non-cultivated) areas. Because higher nutrient levels were associated with intensive drainage and organic soils, lakes situated in such catchments may respond better to targeted interventions. However, other factors, such as internal

loading, groundwater interactions, and lake-specific processes are also likely to be important. Palaeolimnological analyses may further help identify historical drivers of ecological change (thus explaining some of the remaining unexplained variation), thereby contributing to decision-making and targeted interventions.

1 Baggrund og formål

1.1 Baggrund

Omtrent halvdelen af alle søer, der indgår i genbesøget af Vandområdeplan 2021–2027 (VP3), er søer uden veldefinerede oplande inden for størrelsesintervallet 1–629 ha, og størstedelen af disse (ca. 80 %) opfylder ikke miljømålet om god økologisk tilstand. I VP3 har det ikke været muligt at vurdere den nødvendige indsats for at opnå forbedrede økologiske forhold i disse søer, da en forudsætning for udarbejdelse af en belastningsopgørelse er, at et topografisk opland kan afgrænses.

I forbindelse med en tidligere rapport fra DCE om søer uden veldefinerede oplande (Søndergaard et al., 2023) blev det topografiske opland estimeret for 418 søer. Rapporten understregede imidlertid, at disse oplandsestimater var behæftede med betydelig usikkerhed, og analyser af sammenhænge mellem de aktuelle oplandskarakteristika og søernes økologiske tilstand viste kun svage korrelationer. Anbefalingerne i rapporten omfattede blandt andet en vurdering af vigtigheden af fosfor og kvælstof i søerne samt, i forbindelse med forundersøgelser af søer udpeget til restaurering, gennemførelse af palæolimnologiske analyser for at beskrive søernes udvikling og tidligere tilstand.

1.2 Formål

Formålet med dette projekt er at anvende nye metoder til på ny at undersøge de empiriske sammenhænge, som Søndergaard et al. (2023) identificerede mellem søernes økologiske tilstand og arealanvendelsen i både estimerede oplande og bufferzoner af varierende bredde omkring de enkelte søer. Disse analyser vil – sammen med resultaterne fra et igangværende projekt om kvælstofs betydning i søer – danne grundlag for at vurdere, hvilke indsatser det kan være relevante at iværksætte for at forbedre den økologiske tilstand af disse søer.

Projektets resultater og konklusioner kan bidrage til SGAV's arbejde med at identificere behovet for indsats for søer uden veldefinerede oplande og afløb samt til at fastlægge de tiltag, der er nødvendige for at opfylde miljømålet om god økologisk tilstand i vandområdeplanerne 2027–2033.

2 Data

2.1 Data fra søerne

En liste over de pågældende søer blev stillet til rådighed af SGAV. Data for vandkemi blev indhentet fra det nationale overvågningsprogram, NOVANA, lagret i databasen VanDa (<https://vanda.miljoeportal.dk/>) og tilgået via Overfladevanddatabasen (ODA – ODA.dk). For søer uden veldefinerede oplande blev koncentrationer af totalfosfor (TP), totalkvælstof (TN) og klorofyl *a* (Chla) beregnet som årlige sommermiddelværdier (maj-september) til brug i de efterfølgende analyser. De anvendte sødata dækkede perioden 2010–2023. I de videre analyser blev der for hver sø anvendt den gennemsnitlige sommermiddelværdi for hver søparameter beregnet over årene. Søernes økologiske tilstandsklasser var baseret på Chla-værdier leveret af SGAV. Denne klassificering afspejler ikke nødvendigvis den endelige økologiske tilstand, da den ikke inddrager de øvrige biologiske kvalitetselementer, som kræves i henhold til vandrammedirektivet.

2.2 Landskabskarakteristika

Arealanvendelse/markdata

I denne rapport er anvendt datasættet *Marker og Markblokke* (IMK-Markkort) fra Landbrugsstyrelsen, som indeholder de officielle afgrænsninger af landbrugsmarkblokke i Danmark. Data foreligger fra 2010 og fremad og er blevet opdateret årligt siden 2018. Datasættet omfatter mere end 350 klasser af markblokke. I dette projekt er disse dog aggregeret til fire hovedklasser i overensstemmelse med den fremgangsmåde, der er beskrevet i Søndergaard et al. (2023).

Ikke-dyrkede arealer (*ej_mark*): Arealer, der ikke dyrkes, hverken i omdrift eller som varigt græs. Skov- og naturområder er også inkluderet i denne klasse.

Dyrkede arealer i omdrift (*i_omdrift*): Arealer, der regelmæssigt dyrkes ved pløjning eller harvning (1–2-årige sædskifter), herunder arealer anvendt til grøntsagsproduktion og gartneri.

Varigt græs (*varig_græs*): Arealer, der aldrig omlægges, dvs. hverken pløjes eller harves. Græsset skal afgræsses eller slås årligt for at forhindre tilgroning med buske.

Miljøordninger (*miljøordning_ov*): Arealer, der omfatter en række forskellige ordninger, herunder pleje af græs- og naturområder, typisk med græsning og uden gødsning; miljøvenlige landbrugspraksisser, som oftest med reduceret gødsning; etablering og vedligeholdelse af vådområder samt vedligeholdelse af modificerede drænsystemer, som også kan omfatte vådlægning.

Jorddata

De danske jordartskort med en opløsning på 10 × 10 m, anvendt i nærværende studie (version 1.1; Møller et al., 2024), er udarbejdet af Aarhus Universitet. I analyserne anvendtes datasættet 'JB2024_opdelt', som indeholder et revideret

kort over jordtyper i overjorden (0–30 cm). I denne version er JB4klassen yderligere opdelt på baggrund af teksturen i underjorden i overensstemmelse med den klassifikation, der er beskrevet i Greve (2024). De oprindelige jordklasser i databasen blev aggregeret til følgende hovedklasser:

Ler: Lerjord samt svær lerjord eller siltjord

Sand: Grovsand, finsand samt lerblandet sandjord

Organisk (humus): Humusrig og organisk jord

Potentielt drænet areal

Til beregning af andelen af oplandet, der forventes at være drænet (*drænet%*), anvendtes et datasæt udarbejdet af Møller et al. (2018). Datasættet har en rumlig opløsning på 30 m, og dræningsprocenterne er beregnet som et arealvægtet gennemsnit af oplandets landareal, eksklusive søarealet.

Hældning

Oplandets hældning og bufferområder blev beregnet på baggrund af DHM-data (Danmarks HøjdeModel) med en rumlig opløsning på 10 m. Hver sø blev tildelt den gennemsnitlige hældning i oplandene og bufferzonerne.

CAMELS-DK datasæt

Datasættet CAMELS-DK blev anvendt til at integrere oplandskarakteristika (Liu et al., 2025). CAMELS-datasættet (Catchment Attributes and Meteorological Time Series for Large Samples) er en ramme til fremme af hydrologiske analyser og oplandsanalyser (Addor et al., 2017; Coxon et al., 2017; Höge et al., 2023) og er udviklet for en række regioner verden over. CAMELS-DK er udviklet af forskere fra GEUS og Aarhus Universitet på baggrund af data fra DMI, GEUS, VanDa og BASEMAP samt åbne datasæt såsom CORINE Land Cover. Datasættet omfatter 304 målte oplande og 3026 umålte oplande og dækker i alt 3300 danske ID-15-oplande.

Dynamiske oplandskarakteristika omfatter klimavariabler fra Danmarks Meteorologiske Instituts (DMI) daglige griddatasæt, observationer af vandføring fra overfladevandsdatabasen VanDa samt simulerede data fra DK-modellen, herunder vandføring og grundvandsspejl. Datasættet indeholder endvidere oplysninger om jordbund, geologi, arealanvendelse og topografi for hvert ID15opland. I dette projekt indgik udelukkende klima, hydrologi og jordbundskarakteristika i analyserne. Datasættet indeholder endvidere oplysninger om jordbund, geologi, arealanvendelse og topografi for hvert ID-15-opland. I dette projekt indgik udelukkende klima-, hydrologi- og jordbundskarakteristika.

For CAMELSDKvariable blev der ikke beregnet separate værdier for hver bufferafstand (10–250 m), da opløsningen af CAMELSDKdatasættet er ca. 15 km². I stedet blev CAMELSDKværdierne udtrukket én gang ved anvendelse af en buffer på 1000 m omkring hver sø, hvilket var tilstrækkeligt til at overlape den eller de relevante CAMELSDKoplandsenheder. De resulterende arealvægtede CAMELSDKværdier blev herefter ensartet anvendt for alle bufferafstande (10, 50, 100 og 250 m) samt for variabelsættet på oplandsskala, da

de ikke varierer med bufferafstanden. En oversigt over variable fra CAMELS-DKdatasættet samt deres beskrivelser er vist i tabel 2.1.

Tabel 2.1. Oversigt over oplands- og vandføringsvariable fra CAMELS-DKdatasættet, herunder nedbør, jordbunds og hydrologiske indekser anvendt i studiet.

Variabel	Beskrivelse
high_prec_freq (KN_fre)	Hyppighed af dage med kraftig nedbør (≥ 5 gange den gennemsnitlige daglige nedbør) (dage $\cdot$ år $^{-1}$).
high_prec_dur (KN_var)	Gennemsnitlig varighed af perioder med kraftig nedbør (antal på hinanden følgende dage med ≥ 5 gange den gennemsnitlige daglige nedbør).
p_seasonality (n_seas)	Sæsonvariation og timing af nedbør, hvor positive værdier indikerer maksimum om sommeren, negative værdier indikerer maksimum om vinteren, og værdier tæt på nul afspejler en jævn fordeling af nedbør gennem året.
root_depth (r_dyb)	Dybde i jordprofilen, hvorfra planterødder kan optage vand (m). Dybtgående nedsivning af nedbør og veldrænedede forhold fremmer dybere rodsystemer, øget perkolation og reduceret overfladeafstrømning, mens dårligt drænedede jorde begrænser rødderne til lavere dybder, hvilket øger jordmætning og afstrømningen (Fan et al., 2017).
tawc (V_kap)	Samlet vandindhold i jorden (mm).
high_Q_duration (HQ_var)	Gennemsnitlig varighed af perioder med højt vandflow (antal på hinanden følgende dage med vandflow > 9 gange medianen af det daglige flow) (dage).
low_Q_duration (LQ_var)	Gennemsnitlig varighed af perioder med lavt vandflow (antal på hinanden følgende dage med vandflow $< 0,2$ gange det gennemsnitlige daglige vandflow) (dage).
Q_var	Varians i vandføring (mm $^2\cdot$ d $^{-2}$).
BFI	Baseflowindeks (forholdet mellem gennemsnitlig daglig baseflow og gennemsnitlig daglig vandføring).
QP_elasticity (QP_el)	Årlig sammenhæng mellem vandføring og nedbør (følsomhed af vandføringen over for ændringer i nedbør, beregnet med gennemsnitlig daglig vandføring som reference).
Storage_fraction (lag_fraq)	Forholdet mellem aktiv vandbeholdning og samlet vandbeholdning.

3 Metoder

3.1 Klassificering af næringsstoffbegrænsning

Til klassificering af næringsstoffbegrænsning blev der anvendt en metode baseret på støkiometrisk ubalance mellem fytoplankton og næringsstoffer i overensstemmelse med Moon, Scott og Johnson (2021) samt en nylig implementering for det kontinentale USA (Rock & Collins, 2024). En gennemgang af metoden findes i Ladwig et al. (2026, under udgivelse). Metoden bygger på kvantilregressioner og er baseret på den antagelse, at fosforbegrænsede søer udviser et højere Chla-niveau pr. enhed fosfor, mens kvælstofbegrænsede søer giver et højere Chla-niveau pr. enhed kvælstof.

3.2 Udtræk af landskabskarakteristika

Målet med denne rapport er at undersøge sammenhængen mellem landskabskarakteristika og søernes økologiske tilstand. Med dette formål blev der udtrukket landskabskarakteristika for fire bufferafstande – 10 m, 50 m, 100 m og 250 m omkring søerne – ud over de tidligere beregnede oplandsafgrænsninger. Som led i projektet skulle DCE vurdere, om de oplandsdefinitioner, der blev anvendt i Søndergaard et al. (2023), kunne forbedres. Da der ikke forelå nye data fra digitale højdemodeller, blev den samme oplandsafgrænsning som i Søndergaard et al. (2023) anvendt i denne rapport, mens yderligere landskabsinformation blev indhentet ved brug af bufferzoner.

For arealanvendelses- og jordbundskort blev klasseandele beregnet for det samlede opland eller bufferareal efter fratrækning af søarealet. For potentielt dræned arealer blev andelen af det samlede potentielle dræningsareal beregnet, ligeledes eksklusive søarealet inden for oplandet. For hældning blev der beregnet gennemsnitlige hældningsværdier for hver bufferafstand samt for oplandet.

3.3 Redundansanalyse (RDA)

Redundansanalyse (RDA) blev gennemført for at vurdere søernes generelle responsmønstre ved at undersøge påvirkningen og den relative vigtighed af fem typer forklarende variable: arealanvendelse, jordbund, dræning, oplandskarakteristika (CAMELS-DK) samt søareal i forhold til de interne søvariable (tabel 3.1). Separate analyser blev udført for hver bufferzone samt for det estimerede opland.

Tabel 3.1. Oversigt over forklarende variable og responsvariable anvendt i RDA analysen. Secchi/Maks_dyb: Forholdet mellem sigtdybde og maksimumsdybde. Forklaring på øvrige forkortelser fremgår af Kapitel 2.

Forklarende variable							Responsvariable
Arealanvendelse	Jord	Dræning	Camels_DK	Arealhældning	Lagringsfraktion	Sørelateret	Søbaserede parametre
ej_mark	humus	drænet	m_temp	hældning	lag_frak	Søareal	sommer_TN
varig_græs	ler		n_sæs				sommer_TP
i_omdrift	sand		KN_freke				Klorofyl-a
miljøordning_ov			KN_var				Secchi/Maks_dyb
			HQ_var				
			LQ_var				
			Q_var				
			R_dyb				
			V_kap				
			BFI				
			QP_el				

RDA-analysen blev gennemført ved anvendelse af *vegan* pakken (v. 2.72; Oksanen et al., 2025) i R Statistical Software (v. 4.5.1; R Core Team, 2025). Med henblik på at identificere statistisk signifikante forklarende variable og samtidig reducere modellernes kompleksitet blev der anvendt baglæns variabelselektion baseret på funktionen *ordiR2step*. For hver RDA-analyse er justerede R^2 -værdier angivet i resultatafsnittet. For at minimere effekten af multikollinearitet mellem de forklarende variable blev variable med en variansinflationsfaktor (VIF) > 10 udeladt forud for analyserne. Efter gennemførelsen af baglæns selektion, blev den statistiske signifikans af de forklarende variable bestemt ved hjælp af permutationsbaseret ANOVA (*anova.cca*, *by* = "term").

Responsvariablene blev $\log_{10}(x + 1)$ -transformeret efter frasortering af observationer med $TP > 3 \text{ mg L}^{-1}$ og $TN > 5 \text{ mg L}^{-1}$. Disse grænseværdier svarer omtrent til 95.-99. percentil i datasættet ($TN > 5 \text{ mg L}^{-1} \approx 95.$ percentil; $TP > 3 \text{ mg L}^{-1} \approx 99.$ percentil) og repræsenterer således ekstreme værdier. Oplandskarakteristika fra CAMELS-DK samt dræningsvariable blev centreret og skaleret forud for analyserne. Sammensatte forklarende variable (arealanvendelses- og jordbundsvariable) blev transformeret ved hjælp af centreret log-ratio-transformation (CLR) ved anvendelse af R-pakken *compositions* (van den Boogaart et al., 2025). CLR-transformationen indebærer, at hver komponent divideres med det geometriske gennemsnit af alle komponenter i en observation, hvorefter den naturlige logaritme beregnes. Denne tilgang reducerer risikoen for misvisende sammenhænge, som ofte opstår ved analyse af sammensatte data ved brug af euklidisk afstandsbaserede metoder såsom redundansanalyse.

Forud for CLR-transformationen blev nulværdier i de sammensatte datasæt håndteret ved først at udelade variable med mange nulværdier og derefter erstatte resterende nulværdier ved hjælp af funktionen *cmultRepl* fra R-pakken *zCompositions*. Denne funktion anvender en bayesiansk-multiplikativ erstatningsmetode (Palarea-Albaladejo & Martín-Fernández, 2025). Efter CLR-transformationen blev de sammensatte forklarende variable ligeledes centreret og skaleret.

3.4 Random forest-analyser

Random forest-modeller er et robust valg til analyse af store, multivariate datasæt i klassifikationsopgaver, idet metoden konstruerer et stort antal uafhængige beslutningstræer (Breiman, 2001). Ved random forest-analyser blev der anvendt oplandskarakteristika og generelle sødata (se kapitel 2 for forklarende variable).

Datasættet blev opdelt i trænings- og testdata med et forhold på 4:1 (80–20), stratificeret efter økologisk tilstand. Forud for modelleringen blev kovariater med nulvarians frasortet, og data blev normaliseret. Modellen anvendte de kategoriske økologiske tilstandsklasser som målvariable. For at sikre en optimal modelopsætning blev hyperparametre tilpasset ved hjælp af k-fold krydsvalidering, ligeledes stratificeret efter økologisk tilstand. De optimerede hyperparametre blev herefter anvendt i den endelige klassifikationsmodel, som blev estimeret ved brug af R-pakkerne *parSNIP* og *ranger* (Kuhn, 2025; Wright & Ziegler, 2017). Modellens præstation blev evalueret på testdatasættet ved beregning af ROCkurven (Receiver Operating Characteristic) og AUCværdien (areal under ROCkurven). ROCanalysen illustrerer forholdet mellem sande positive og falske positive klassifikationer og indikerer dermed, hvor godt modellen er i stand til at klassificere de kategoriske mål. AUC er også en vigtig parameter, der måler en models evne til at skelne mellem klasser. En værdi på 0,5 angiver ingen bedre diskrimination end tilfældig forventning, og værdier tæt på 1 angiver stigende diskriminationsevne.

For at vurdere modellens følsomhed over for inputdata og dermed identificere, hvilke forklarende variable der har størst betydning for klassificeringen af en søs økologiske tilstand, blev der gennemført en analyse af variabelbetydning ved hjælp af R-pakken *iml* (Molnar, Bischl & Casalicchio, 2018). Der blev udført en global permutationsanalyse, som kvantificerer, hvor følsomt modellens tab (opgjort ved krydsentropi) er over for ændringer i de enkelte input variable. For vigtigheden af lokale variable blev der beregnet akkumulerede lokale effekter (ALE) for de variable, der lå over 90-percentilen i den globale permutationsanalyse. ALE-analyserne viser den gennemsnitlige effekt af hver variabel på modellens estimerede sandsynligheder. Random forest-analyserne blev gennemført separat for hver bufferzonekategori.

3.5 Conditional inference trees – beslutningstræer

Der blev konstrueret beslutningstræer (conditional inference trees, CIT) ved anvendelse af R-pakken *partykit* (Hothorn, Seibold & Zeileis, 2015) med henblik på at identificere vigtige landskabsvariable samt deres potentielle tærskelværdier. Beslutningstræerne blev anvendt på det fulde multivariate datasæt med TN, TP, økologisk tilstand og næringsstofbegrænsningstype som responsvariable for hver bufferzone samt for oplandsafgrænsningen. Denne tilgang muliggør identifikation af de mest betydende oplands- og vandkvalitetsrelaterede tærskelværdier, som bedst adskiller responsvariablene.

Beslutningstræer er særligt velegnede til at identificere hierarkiske strukturer og ikke-lineære sammenhænge mellem forklarende variable og giver fortolkelige beslutningsregler, som viser, hvordan kombinationer af næringsstofkoncentrationer og oplandskarakteristika påvirker søernes tilstand. CITanalyserne blev primært anvendt med forklarende formål, selv om der også blev beregnet præstationsmål for at vurdere modellernes ydeevne. Nøjagtigheden (Accuracy) og Kappa-værdien (som måler overensstemmelsen mellem forudsigelser og observationer ud over, hvad der kan forventes ved tilfældigt gæt;

<0: ingen overensstemmelse; 0,01–0,20: ingen til ringe; 0,21–0,40: rimelig; 0,41–0,60: moderat; 0,61–0,80: betydelig; 0,81–1,00: perfekt overensstemmelse (McHugh (2012)) blev beregnet for kategoriske responser (økologisk tilstand, type af næringsstofbegrænsning). For kontinuerte responsvariable (TN og TP) blev R^2 beregnet som mål for beslutningstræernes forklaringsevne. Forud for CITanalyserne blev ekstreme værdier frasorteret, idet observationer med TP > 3 mg L⁻¹ og TN > 5 mg L⁻¹ blev fjernet for at fokusere på de mest relevante næringsstofniveauer.

Analyserne blev gentaget for fire forskellige bufferzoner samt for hele oplandet. I rapportens kapitel 4 præsenteres dog udelukkende oplandsresultaterne, da den forklarede varians generelt var størst for landskabskarakteristika på oplandsskala. Resultaterne fra bufferzoneanalyserne findes i bilaget.

4 Resultater

4.1 Overblik over søerne

De væsentligste karakteristika for søerne, der indgår i analyserne i denne rapport, er præsenteret i tabel 4.1. Koncentrationen af totalfosfor (TP) varierede fra 0,003 til 16,7 mg L⁻¹ med en gennemsnitsværdi på 0,3 mg L⁻¹, mens koncentrationen af totalkvælstof (TN) varierede fra 0,3 til 11,4 mg L⁻¹ med en gennemsnitsværdi på 1,8 mg L⁻¹.

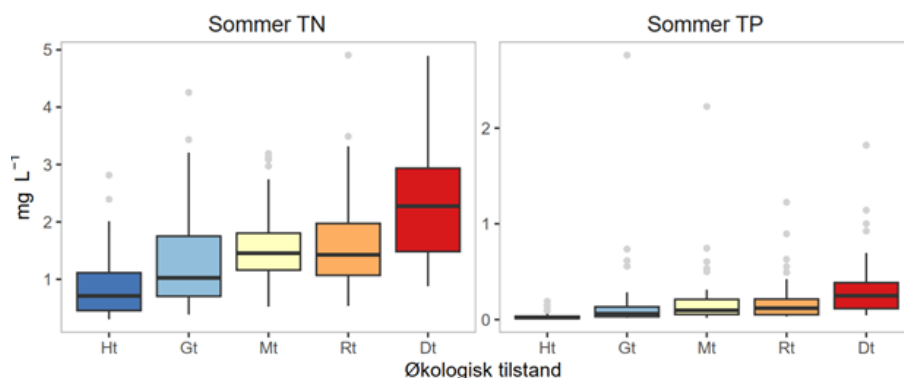
På baggrund af klassificeringen af økologisk tilstand blev 74 af de 418 søer klassificeret som Ht (høj tilstand), 50 som Gt (god tilstand), 53 som Mt (moderat tilstand), 43 som Rt (ringe tilstand) og 65 som Dt (dårlig tilstand). For 60 søer var der ikke nogen fastlagt økologisk tilstand, og disse blev derfor ikke inkluderet i analyserne. Efter frasortering af data fra før 2010 samt udelukkelse af søer med usædvanligt høje TP- og TN-koncentrationer (TN > 5 mg L⁻¹ og TP > 3 mg L⁻¹) omfattede det endelige datasæt 331 søer.

Tabel 4.1. Det fulde spektrum af interne søkarakteristika i det ufiltrerede datasæt. Ekstreme værdier er udeladt fra analyserne, men vises her for at afspejle den oprindelige variation i datasættet.

Variable	Gennemsnit	Median	Min.	Q25	Q75	Maks.
Sommer TN (mg L ⁻¹)	1,8	1,3	0,3	0,8	2,1	11,4
Sommer TP (mg L ⁻¹)	0,3	0,1	0,003	0,03	0,2	16,7
Maksimumsdybde (m)	3,9	1,9	0,1	0,9	5,7	45,0
Sigtdybde (m)	1,5	0,8	0,1	0,4	2,0	7,3
Søareal (ha)	11,8	5,9	1,0	2,5	9,4	576,7

Både TN- og TP-koncentrationerne steg konsistent fra søer i høj økologisk tilstand (Ht) mod søer i dårlig tilstand (Dt) (figur 4.1). Søer i Dt-klassen udviste markant forhøjede næringsstofniveauer, mens søer i Ht- og Gt-klasserne generelt havde lave TN- og TP-koncentrationer med snævre variationsintervaller.

Figur 4.1. Boksplots af sommerkoncentrationer af totalkvælstof (Sommer TN) og totalfosfor (Sommer TP) fordelt på økologiske tilstandsklasser (Ht, Gt, Mt, Rt, Dt). Hver boks viser 1. kvartil, median og 2. kvartil. Whiskers viser det sidste punkt inden for 1,5*interkvartilområdet og cirkler viser outliers udenfor dette område.



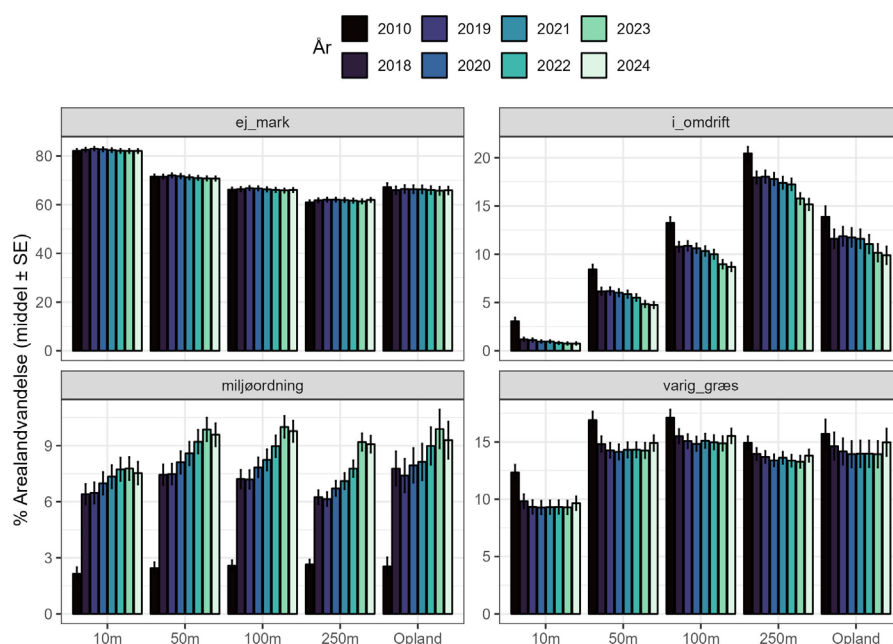
4.2 Oversigt over landskabskarakteristika

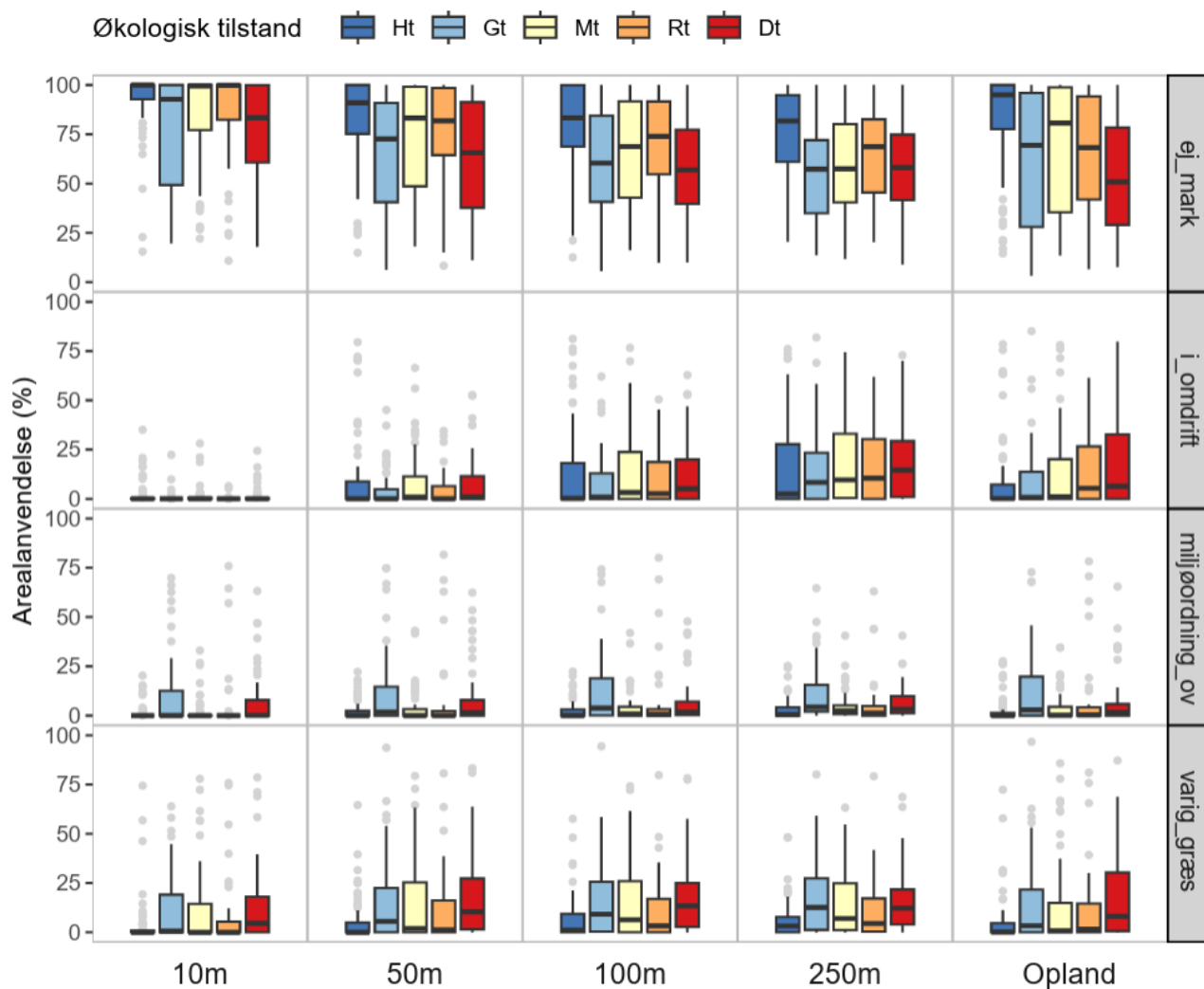
Marker

Ifølge *Markblokke*-datasættet har andelen af ikke-dyrkede arealer (*ej-mark*) været relativt stabil gennem de seneste 15 år (figur 4.2). Derimod er der sket markante ændringer i kategorierne landbrugsarealer (*i_omdrift*), miljøordninger og græsarealer (*varig_græs*). Landbrugs- og græsarealer har udvist en faldende tendens over tid, mens andelen af arealer omfattet af miljøordninger er steget jævnt. Figuren viser endvidere den gennemsnitlige andel af de væsentligste arealanvendelseskategorier inden for fire bufferafstande (10 m, 50 m, 100 m og 250 m) samt for det samlede opland for årene 2010 og 2018–2024. I gennemsnit udgjorde den ikke-dyrkede klasse – som bl.a. repræsenterer naturområder – den største andel blandt kategorierne, men andelen aftog med stigende bufferafstand. Landbrugsarealer udviste det modsatte mønster, idet deres andel steg med øget bufferstørrelse.

Figur 4.3 viser fordelingen af arealanvendelsesklasser i procent på tværs af de økologiske tilstandsklasser for søer i forskellige bufferzoner. Overordnet set havde søer med høj økologisk tilstand en større andel af ikke-dyrkede arealer på tværs af alle bufferafstande, mens søer i dårlig tilstand havde en lavere andel af ikke-dyrkede arealer. For landbrugsarealer og græsarealer var mønstret omvendt ved 250-m-bufferen og på oplandsskala, idet søer med ringe/dårlig økologisk tilstand generelt havde højere andele af landbrug og græs end søer med høj økologisk tilstand. Mønstret for arealer omfattet af miljøordninger var mindre entydigt, men søer i god tilstand havde en svagt højere andel af disse arealer.

Figur 4.2. Gennemsnitlige andele af arealanvendelsestyper inden for forskellige bufferafstande (10 m, 50 m, 100 m og 250 m) samt for estimerede oplande beregnet over otte år (2010–2024). Søjlerne viser middelværdi $\pm$ standardfejl (SE) på tværs af søer for hver arealanvendelsesklasse. Forklaring af variable fremgår af kapitel 2.

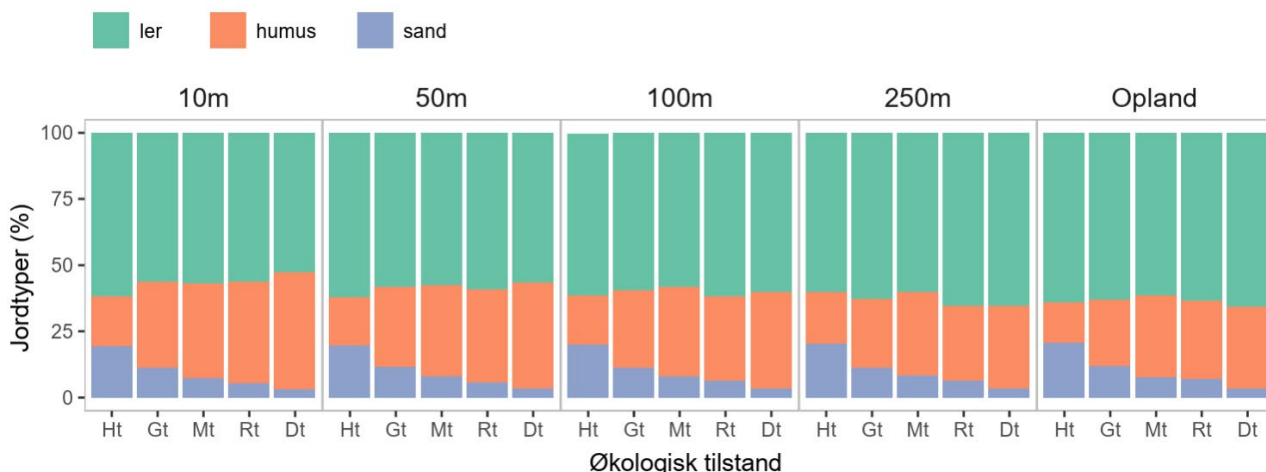




Figur 4.3. Andele af arealanvendelsestyper inden for søernes bufferzoner og oplande fordelt på økologiske tilstandsklasser. Hver boks viser 1. kvartil, median og 2. kvartil. Whiskers viser det sidste punkt inden for 1,5*interkvartilområdet og cirkler viser outliers udenfor dette område. Forklaring af variable fremgår af kapitel 2.

Jordtypedata

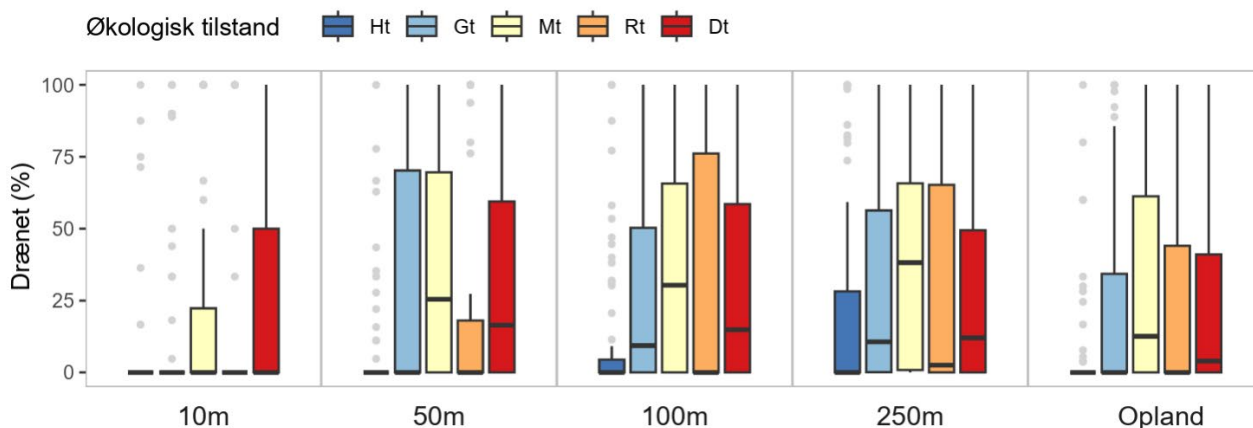
På tværs af alle bufferafstande og oplande var lerjord den dominerende jordtype og udgjorde typisk mere end halvdelen af det omkringliggende areal (figur 4.4). Organiske jorde (*humus*) udgjorde en betydelig, men mindre andel, mens sandjord generelt havde den laveste andel på tværs af alle økologiske tilstandsklasser. Andelen af organisk jord steg med faldende økologisk tilstand, med de højeste værdier i søer med ringe eller dårlig økologisk tilstand (Rt og Dt). Omvendt havde søer med høj eller god økologisk tilstand (Ht og Gt) generelt lavere andele af organisk jord. Fordelingen af jordtyper varierede ikke mellem de forskellige bufferzoner. Søer i bedre økologisk tilstand havde en højere andel af sandjord sammenlignet med søer i dårligere økologisk tilstand.



Figur 4.4. Andele af jordtyper inden for søernes bufferzoner og oplande fordelt på økologiske tilstandsklasser.

Potentielle dræningsområder

Dræningsintensiteten (*drænet%*) på tværs af forskellige bufferzoner for økologisk tilstand er vist i figur 4.5. Søer med høj (Ht) og god (Gt) økologisk tilstand havde generelt en lavere andel af drænet areal i deres omkringliggende oplande. I modsætning hertil havde søer med ringere økologisk tilstand tendens til at ligge i oplande med en højere procentandel af drænet areal.

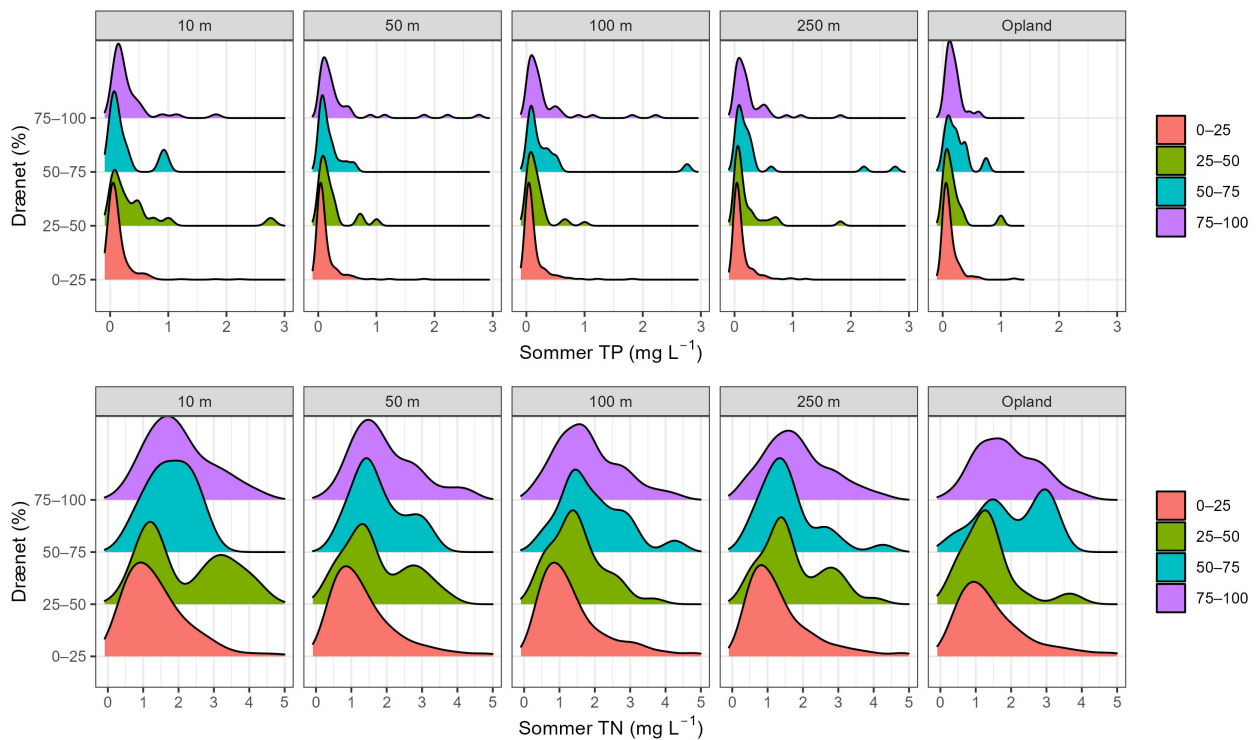


Figur 4.5. Andelen af potentielt drænedes arealer (drænet %) i søernes bufferzoner og oplande fordelt på økologiske tilstandsklasser. Hver boks viser 1. kvartil, median og 2. kvartil. Whiskers viser det sidste punkt inden for $1,5 \times$ interkvartilområdet og cirkler viser outliers uden for dette område.

Sammenhængen mellem dræningsintensitet (*drænet%*) og næringsstofkoncentrationer i søer er vist i figur 4.6. Med stigende dræningsintensitet i oplandet forskydes maksimum i tæthedsfordelingerne for totalkvælstof (TN) og totalfosfor (TP) sig mod højere koncentrationer, hvilket indikerer højere næringsstofniveauer i søer med mere omfattende dræning i bufferzonen og oplandet. Denne forskydning medfører, at fordelingerne er højreskæve – har en lang højre hale, hvilket illustrerer, at søer i stærkt drænedes oplande oftere udviser forhøjede TN- og TP-koncentrationer.

For TN udviste søer i oplande med 0–25 % dræning en unimodal fordeling med et maksimum omkring ca. 1 mg L^{-1} , mens søer i oplande med højere dræningsintensitet havde forskudte fordelinger i retning af højere koncentrationer og i visse tilfælde en bimodal fordeling. Samlet set indikerer resultaterne, at søer beliggende i oplande med en høj grad af kunstig dræning (75–100 %) har højere næringsstofkoncentrationer og udviser større variation

sammenlignet med søer i mindre drænedede oplande. Dette mønster var særligt udtalt for TN, hvor oplande med høj dræningsintensitet var kendetegnet ved fordelingerne med en lang højre hale, hvilket indikerer forekomst af høje TN-koncentrationer.



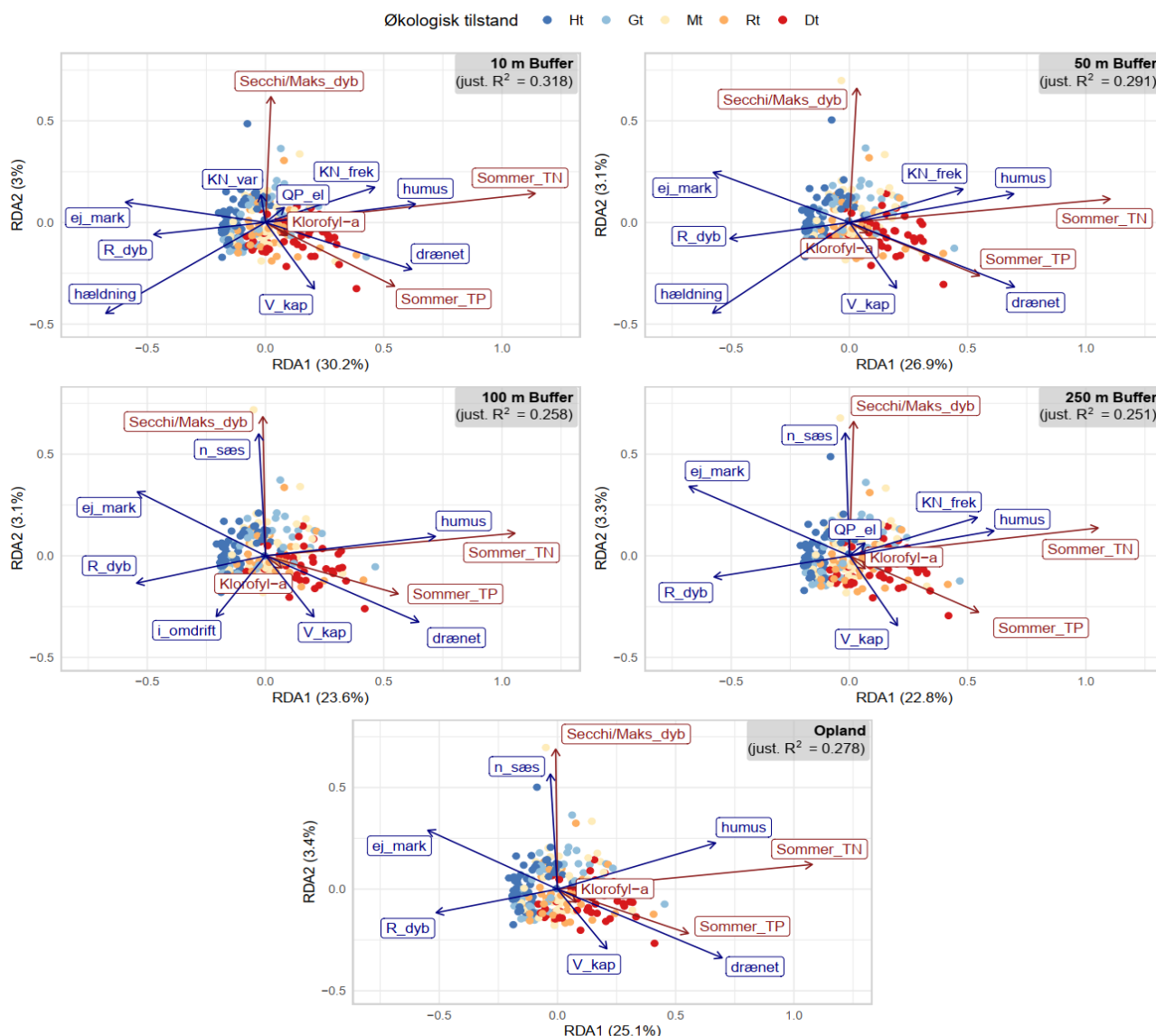
Figur 4.6. Fordeling af koncentrationer af totalfosfor (TP) og totalkvælstof (TN) inden for forskellige dræningsklasser i oplandet omkring de enkelte søer. Dræningsklasserne repræsenterer andelen af opland eller bufferzone, der er kunstigt drænet (0–25 %, 25–50 %, 50–75 % og 75–100 %). Koncentrationer over 3 mg L⁻¹ for TP og 5 mg L⁻¹ for TN er ikke medtaget i figurene.

4.3 Effekter af landskabskarakteristika på søernes tilstand

RDA-resultater

RDA blev anvendt til at vurdere de overordnede sammenhænge mellem variablene arealanvendelse, jordbundsforhold, dræning, oplandskarakteristika og areal med søernes interne responsvariable, herunder koncentrationerne af TN, TP, Chla og forholdet mellem sigtdybde og maksimumsdybde (secchi/maksimumsdybde), som et mål for søernes turbiditet. Observationer med manglende værdier blev udeladt fra analyserne, da RDA ikke kan håndtere manglende data. Som følge heraf varierede prøvestørrelsen (antal søer) i RDA-analyserne mellem 222 og 231, afhængigt af bufferafstand eller opland.

Alle globale RDA-modeller samt deres to første akser var signifikante. De justerede R²-værdier viste, at de valgte forklarende variable forklarede 25–32 % af den (begrænsede) variation i søernes vandkvalitet (figur 4.7). RDA1 var den dominerende akse og forklarede 22–30 % af variationen, mens RDA2 bidrog med 3,0–3,4 %. RDA-modellerne blev efterfølgende reduceret ved baglæns selektion, og alle forklarende variable, der ikke fremgik af permutationstest-tabellen (tabel 4.2), blev frasorteret enten som ikke-signifikante eller på grund af multikollinearitet.



Figur 4.7. Resultater af RDA-analyser for forskellige bufferafstande (10 m, 50 m, 100 m og 250 m) og oplande. Blå pile angiver statistisk signifikante forklarende variable, som er udvalgt automatisk ved baglæns selektion, mens røde pile repræsenterer respons-variable i søerne. Punkter repræsenterer søer, og farverne angiver økologisk status fra dårlig (rød) til høj (mørkeblå). Forklaring af variable fremgår af kapitel 2.

Eftersom arealanvendelses- og jordbundsdata blev CLR-transformeret, skal deres effekter fortolkes i forhold til alternative klasser inden for hver sammensatte gruppe frem for som individuelle effekter, også selv om nogle klasser ikke var individuelt signifikante i RDA-analyserne. For eksempel var en højere andel af ikke-dyrkede arealer (*ej_mark*) i forhold til andre arealanvendelseskategorier (fx *i_omdrift*, *varig_græs*, *miljøordning_ov*) forbundet med bedre søtilstand (økologisk status: høj og god). Omvendt indikerede en højere andel af organisk jord (*humus*) i forhold til sand og ler en større risiko for eutrofiering (økologisk status: ringe og dårlig).

Ved bufferafstandene 10 m og 50 m samt for hele oplandet repræsenterede RDA1 en eutrofieringsgradient. Responsvariablene i søerne, TN (*Sommer_TN*), TP (*Sommer_TP*) og Chla (*Klorofyl-a*), var stærkt og positivt relateret til denne akse, hvilket indikerer, at RDA1 afspejler stigende næringsstofkoncentrationer. Disse responsvariable var positivt relateret til forklarende variable, herunder andelen af potentielt drænet areal (*drænet*), andelen af organisk jord (*humus*) og høj total plantetilgængelig vandkapacitet (*V_kap*) samt i mindre grad hyppigheden af perioder med intens nedbør (*KN_freke*).

I modsætning hertil viste roddybde (*R_Dyb*) en stærk negativ sammenhæng med denne gradient, og ikke-dyrkede arealer (*ej_mark*) udviste en tilsvarende, men svagere, negativ sammenhæng. Den gennemsnitlige hældning inden for bufferzonerne viste ligeledes en markant modsat relation til RDA1, især med *KN_frek*. Søerne, der farvekodet efter økologisk status, viste, at forklarende variable med en negativ sammenhæng til RDA1 var forbundet med søer med bedre økologisk status.

Samlet set indikerer resultaterne i figur 4.7, at vådere oplande med en højere andel af organisk jord, mindre hyppig nedbør og større dræned arealer er forbundet med ringere økologisk kvalitet i søerne.

Tabel 4.2. Permutationstest for forklarende RDA-variable; separat signifikanstest for hver enkelt forklarende variabel (*p*). Ikke-justeret varians angiver den ikke-justerede variation i sø-responsdata, der forklares af hver forklarende variabel. Residual varians er ikke vist. Se kapitel 2 for forklaring af variablene.

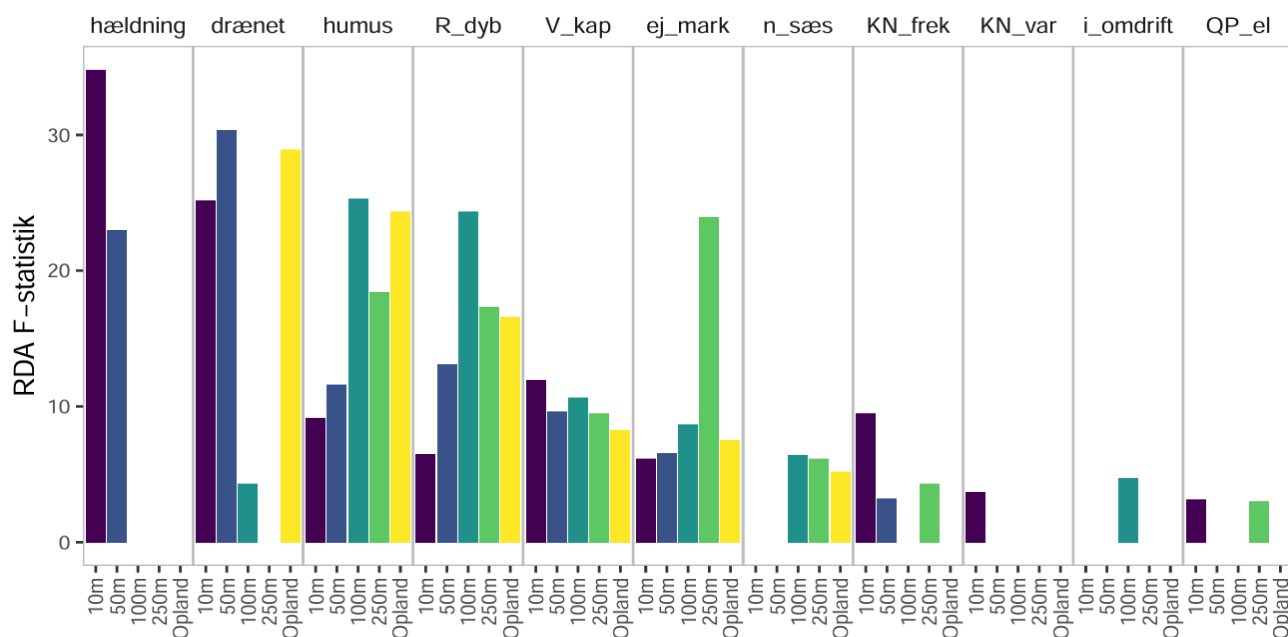
Buffer/Opland	Forklarende variable	Ikke-justeret varians	% Total varians	F	<i>p</i>
10m	<i>hældning</i>	0,0043	10,9	34,8	0,001
10m	<i>drænet</i>	0,0031	7,9	25,2	0,001
10m	<i>V_kap</i>	0,0015	3,8	11,9	0,002
10m	<i>KN_frek</i>	0,0012	3,0	9,5	0,001
10m	<i>humus</i>	0,0011	2,9	9,2	0,001
10m	<i>R_dyb</i>	0,0008	2,0	6,5	0,002
10m	<i>ej_mark</i>	0,0008	1,9	6,2	0,004
10m	<i>KN_var</i>	0,0005	1,2	3,7	0,037
10m	<i>QP_el</i>	0,0004	1,0	3,2	0,049
50m	<i>drænet</i>	0,0038	9,8	30,4	0,001
50m	<i>hældning</i>	0,0029	7,4	23,0	0,001
50m	<i>R_dyb</i>	0,0016	4,2	13,1	0,001
50m	<i>humus</i>	0,0015	3,7	11,6	0,001
50m	<i>V_kap</i>	0,0012	3,1	9,6	0,002
50m	<i>ej_mark</i>	0,0008	2,1	6,6	0,003
50m	<i>KN_frek</i>	0,0004	1,0	3,3	0,036
100m	<i>humus</i>	0,0033	8,4	25,3	0,001
100m	<i>R_dyb</i>	0,0032	8,1	24,4	0,001
100m	<i>V_kap</i>	0,0014	3,5	10,7	0,001
100m	<i>ej_mark</i>	0,0011	2,9	8,7	0,001
100m	<i>n_sæs</i>	0,0008	2,1	6,5	0,005
100m	<i>i_omdrift</i>	0,0006	1,6	4,7	0,014
100m	<i>drænet</i>	0,0006	1,4	4,3	0,012
250m	<i>ej_mark</i>	0,0031	7,9	23,9	0,001
250m	<i>humus</i>	0,0024	6,1	18,4	0,001
250m	<i>R_dyb</i>	0,0022	5,8	17,4	0,001
250m	<i>V_kap</i>	0,0012	3,1	9,5	0,001
250m	<i>n_sæs</i>	0,0008	2,0	6,1	0,006
250m	<i>KN_frek</i>	0,0006	1,4	4,3	0,023
250m	<i>QP_el</i>	0,0004	1,0	3,0	0,065
Opland	<i>drænet</i>	0,0037	9,5	28,9	0,001
Opland	<i>humus</i>	0,0031	8,0	24,3	0,001
Opland	<i>R_dyb</i>	0,0021	5,4	16,6	0,001
Opland	<i>V_kap</i>	0,0011	2,7	8,3	0,003
Opland	<i>ej_mark</i>	0,0010	2,5	7,5	0,002
Opland	<i>n_sæs</i>	0,0007	1,7	5,2	0,008

Ved de større bufferafstande (100 m og 250 m) (bilagsfigur 7.1) var eutrofieringsgradienten fortsat til stede, men den relative vigtighed af de forklarende variable ændrede sig. Ved 100 m- og 250 m-bufferne bidrog andelen af potentielt drænet areal (*drænet*) mindre end ved de mindre bufferzoner, mens organisk jord (*humus*) og roddebyde (*R_dyb*) fortsat var vigtige forklarende variable.

Det bemærkes, at *ej_mark* fik stigende forklaringssevne med øget bufferafstand, og at både F-værdi og forklaret varians var højest ved 250 m (figur 4.8, tabel 4.2). På trods af dette faldt af RDA-modellernes samlede forklaringssevne svagt med stigende bufferafstand. På 250 m-skalaen var *KN_fre* (hyppighed af dage med kraftig nedbør) ligeledes en af de signifikante forklarende variable, som var forbundet med ringere økologisk status og forhøjede næringsstofkoncentrationer, om end variabelens forklaringssevne var lav.

Den anden RDA-akse var domineret af forholdet mellem sigtdebyde og maksimumsdybde (*Secchi/Maks_dyb*) og var sammenfaldende med varigheden af perioder med hyppig nedbør (*KN_var*) samt nedbørens sæsonmæssige fordeling (*n_sæs*; sommermaksimum). Denne akse forklarede kun $\leq 3,4$ % af variationen på tværs af alle modellerne, hvilket indikerer, at dette mønster næppe afspejler en stærk økologisk effekt. Endvidere var *i_omdrift* (intensivt dyrket landbrug) ganske vist signifikant ved 100 m-bufferen (tabel 4.2; $F = 4,7$, $p = 0,014$), men bidraget var begrænset sammenlignet med øvrige forklarende variable.

Samlet set indikerer RDA-analyser baseret på forskellige arealer af bufferzonerne og oplandsafgrænsninger, at ringere økologisk kvalitet i søer, karakteriseret ved højere næringsstofkoncentrationer, er forbundet med vådere oplande, hvilket afspejles i en højere samlet vandtilgængelighed, større andele af organisk jord og mere omfattende dræning. Omvendt peger en høj roddebyde i bufferzoner eller oplande i modsat retning og indikerer bedre økologiske forhold. Desuden - med en forøgelse af bufferzonens areal bliver den relative påvirkning fra ikke-dyrkede områder (*ej_mark*) højere, sammenlignet med andre kategorier af arealanvendelse, forstået på den måde, at en relativt højere dækning af *ej_mark*-arealet er forbundet med forbedret økologisk tilstand.

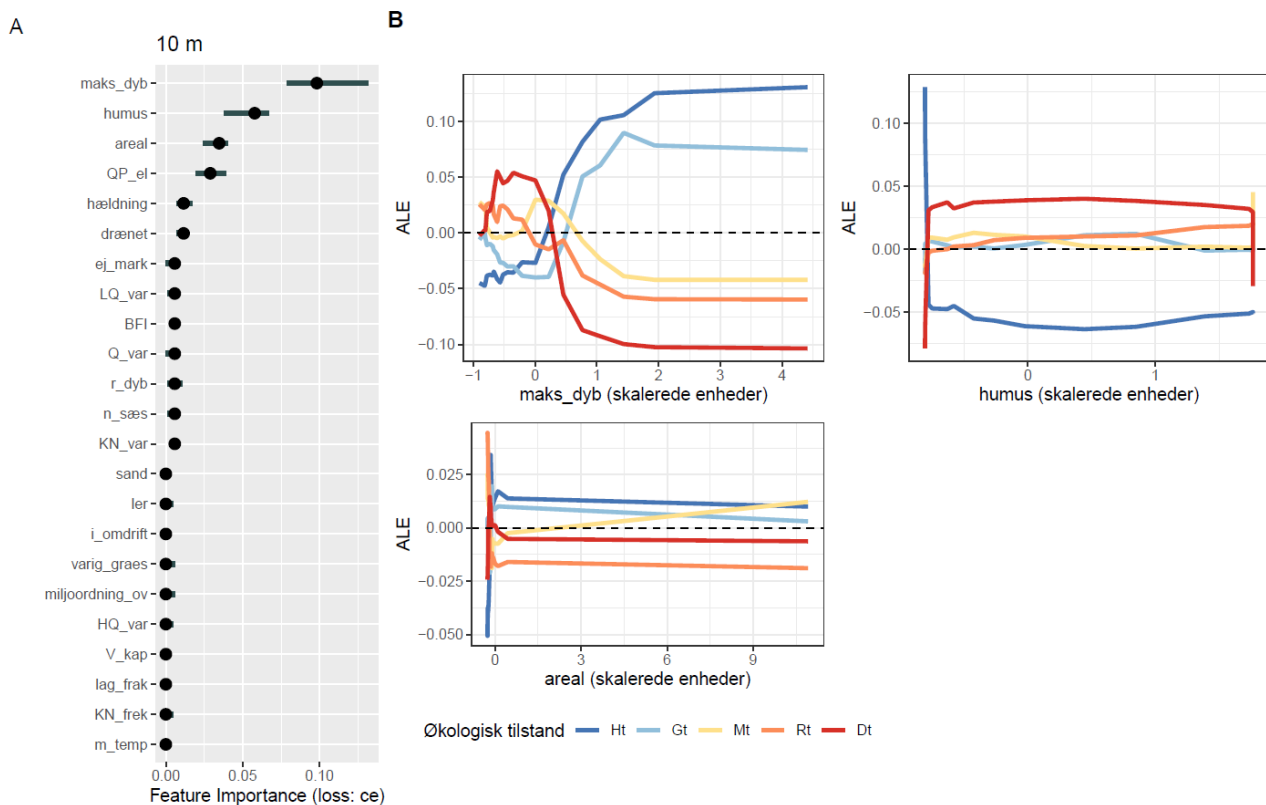


Figur 4.8. F værdier fra permutationstesten for forklarende variable i RDA, som sammenligner bidraget fra hver variabel inden for de enkelte bufferzoner og oplande. Forklaring af variable fremgår af kapitel 2.

Random forest

For en bufferzone på 10 m (figur 4.9) opnåede random forest-modellen en AUC på 0,6 (individuelle AUC-værdier for Ht = 0,88, Gt = 0,77, Mt = 0,56, Rt = 0,47 og Dt = 0,77). Modellen udviste en begrænset forklaringsevne for moderat og ringe økologisk status, hvilket kun i begrænset omfang blev opvejet af en højere præcision for søer med god og høj økologisk status. Dette resulterede samlet set i den relativt lave AUC-værdi på 0,682.

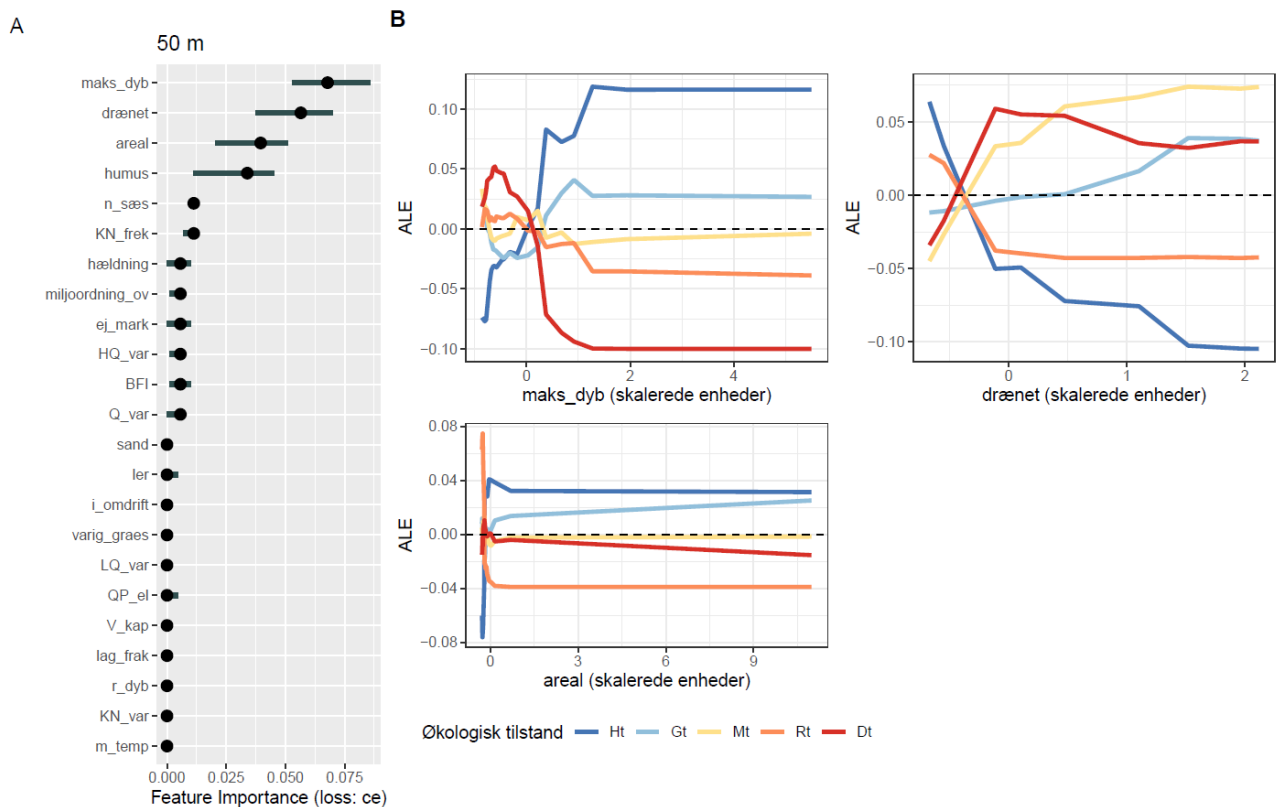
Den globale vigtighed af forklarende variable (figur 4.9A) viste, at maksimumsdybde (*maks_dyb*), andelen af organisk jord i oplandet (*humus*) og areal var blandt de vigtigste forklarende variable for økologisk status. Sandsynligheden for god eller høj økologisk status steg med stigende dybde. Omvendt reducerede større dybdesandsynligheden for ringere økologisk status, som udviste maksimumsværdier ved lave dybdeværdier. Sandsynligheden for høj økologisk status var højest ved meget lave værdier af organisk jord, mens ringe økologisk status var forbundet med et højere indhold af organisk materiale. Tilsvarende var et øget areal, som for dybde, forbundet med høj, god og moderat økologisk status (figur 4.9B).



Figur 4.9. Vigtigheden af variable for bufferzonen 10 m. A: Global vigtighed af variable (kvantificeret ved hjælp af "cross-entropy loss" ved klassificering). B: Akkumulerede lokale effekter (ALE), som viser effekten af specifikke kovariater på forudsigelserne i RF-modellerne. Positive ALE-værdier (y-aksen) indikerer en "over gennemsnit" forudsigelse af den pågældende tilstand over kovariatens normaliserede område (x-aksen). Omvendt – negative ALE-værdier viser et "under gennemsnit" bidrag til den pågældende kovariat. Her viser plottet modellens gennemsnitlige forklaringsevne for hver variabel opdelt efter økologisk status. Forklaring af variable fremgår af kapitel 2.

For en bufferzone på 50 m (figur 4.10) opnåede random forest-modellen en AUC på 0,6 (individuelle AUC-værdier for Ht = 0,8, Gt = 0,6, Mt = 0,49, Rt = 0,42 og Dt = 0,76). Modellen udviste en begrænset forklaringsevne for moderat økologisk status, men en højere nøjagtighed for søer med henholdsvis høj/god og lav/ringe økologisk status. Den samlede AUC var dog fortsat lav og kun en smule højere end tilfældig klassifikation (0,5).

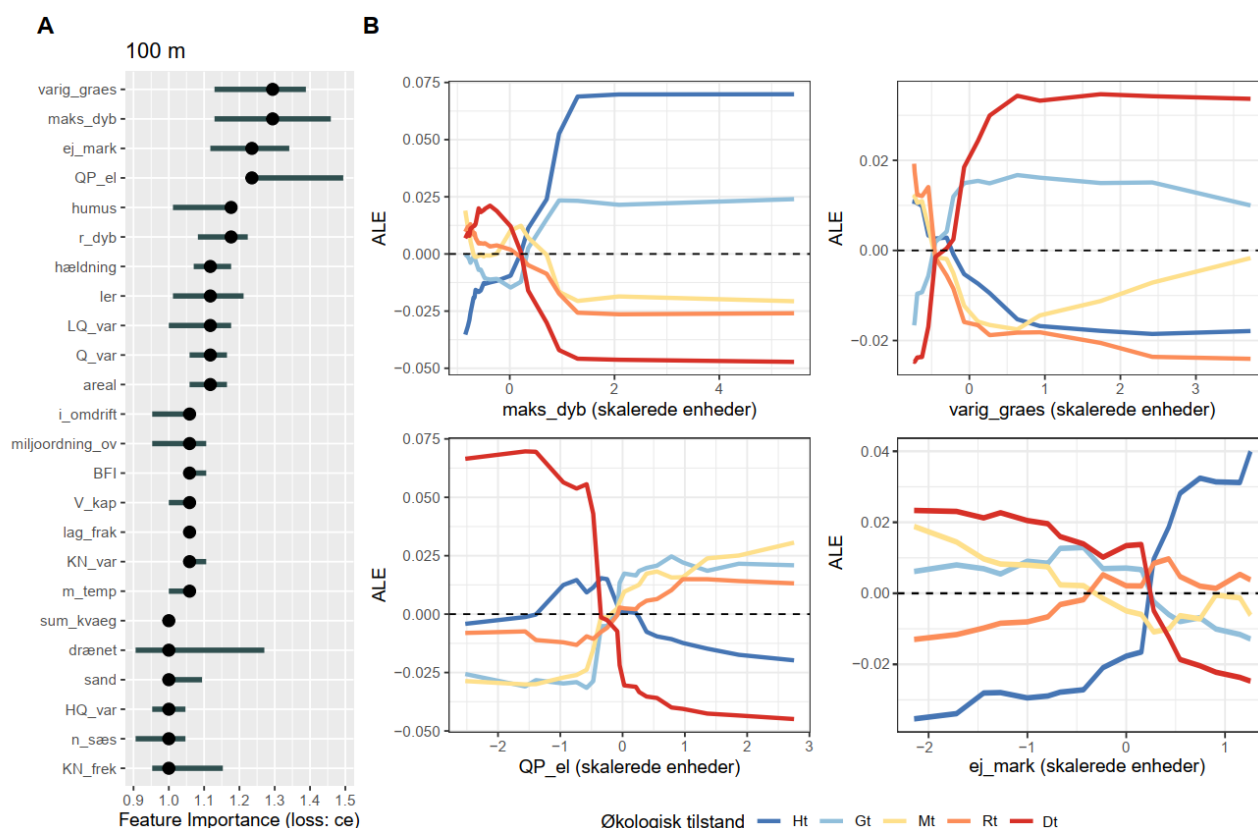
Den globale vigtighed af variable (figur 4.10) viste, at maksimumsdybde (*maks_dyb*), dræningsprocent (*drænet*) og areal var blandt de vigtigste variable til klassifikation af økologisk status. Som ved 10 m-bufferzonen var forbedret økologisk status positivt relateret til større dybde og større areal. Dræningsprocenten var positivt relateret til ringe økologisk status og negativt relateret til høj økologisk status. Samtidig steg sandsynligheden for, at en sø blev klassificeret som god eller moderat, med stigende dræningsprocent, mens sandsynligheden for høj økologisk status faldt. Disse resultater er indbyrdes modstridende og sandsynligvis påvirket af modellens lave forklaringsevne for god, moderat og ringe økologisk status.



Figur 4.10. Vigtigheden af variable for bufferzonen 50 m. A: Global vigtighed af variable. B: Akkumulerede lokale effekter, som viser modellens gennemsnitlige forklaringsevne for hver variabel opdelt efter økologisk status. Forklaring af variable fremgår af kapitel 2, og forklaring af figurer fremgår af figur 4.9.

For en bufferzone på 100 m (figur 4.11) opnåede random forest-modellen en AUC på 0,6 (individuelle AUC-værdier for Ht = 0,85, Gt = 0,64, Mt = 0,68, Rt = 0,73, Dt = 0,83). Modellen udviste en dårlig evne til at forudsige dårlig økologisk tilstand, hvilket dog blev opvejet af en høj nøjagtighed for søer med høj økologisk tilstand. AUC-værdien forblev igen relativt lav og kun en smule end ved tilfældig klassifikation (0,5). Den globale vigtighed af variable (figur 4.11) identificerede maksimumsdybde, græsarealer (*variggræs*), sammenhæng mellem vandføring og nedbør (QP_el) og ikke-dyrkede arealer (*ej-mark*) som vigtige indikatorer for økologisk tilstand. Her var søer med høj økologisk tilstand positivt forbundet med større maksimumsdybder og højere andele af ikke-dyrkede arealer, mens de var negativt forbundet med græsarealer og sammenhængen mellem vandføring og nedbør. Sandsynligheden for, at en sø blev klassificeret som værende i dårlig økologisk tilstand, steg med mere græsareal i bufferzonen, men var lav ved højere indekser for vandføring og nedbør – på samme måde som for søer i høj økologisk tilstand.

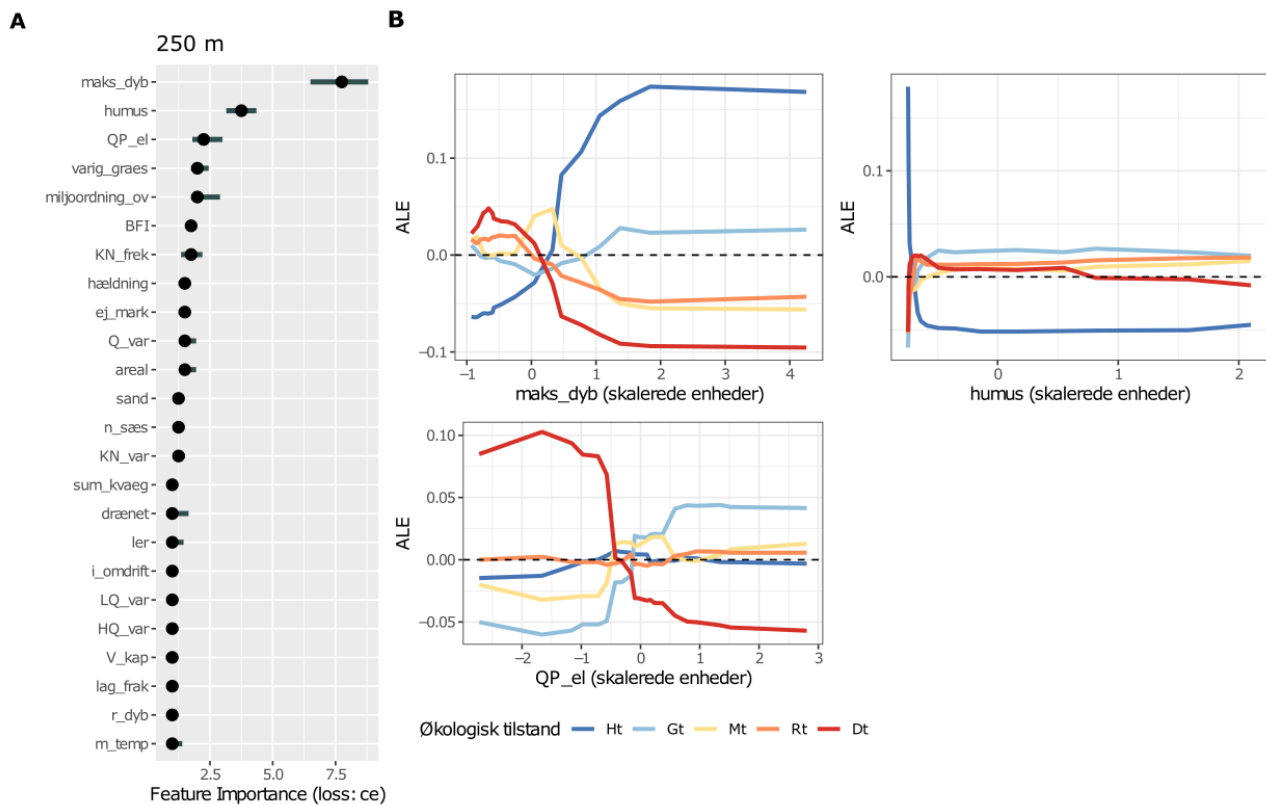
Sandsynligheden for at være i god, moderat eller dårlig tilstand steg med højere indekser for vandføring og nedbør. Sær i god, moderat og dårlig økologisk tilstand viste modstridende dynamikker i de akkumulerede lokale effekter, formentlig på grund af den lave AUC-værdi i den bagvedliggende modelklassifikation.



Figur 4.11. Vigtigheden af variable for bufferzonen 100 m. A: Global vigtighed af variable. B: Akkumulerede lokale effekter, som viser modellens gennemsnitlige forklaringsevne for hver variabel opdelt efter økologisk status. Forklaring af variable fremgår af kapitel 2, og forklaring af figurer fremgår af figur 4.9.

For en bufferzone på 250 m (figur 4.12) opnåede random forest-modellen en AUC på 0,5 (individuelle AUC-værdier for Ht = 0,69, Gt = 0,54, Mt = 0,56, Rt = 0,51 og Dt = 0,86). Modellen havde moderat/dårlig klassifikationsevne for alle økologiske tilstande bortset fra dårlig økologisk status. Den samlede AUC var således meget lav og tæt på tilfældig klassifikation (0,5).

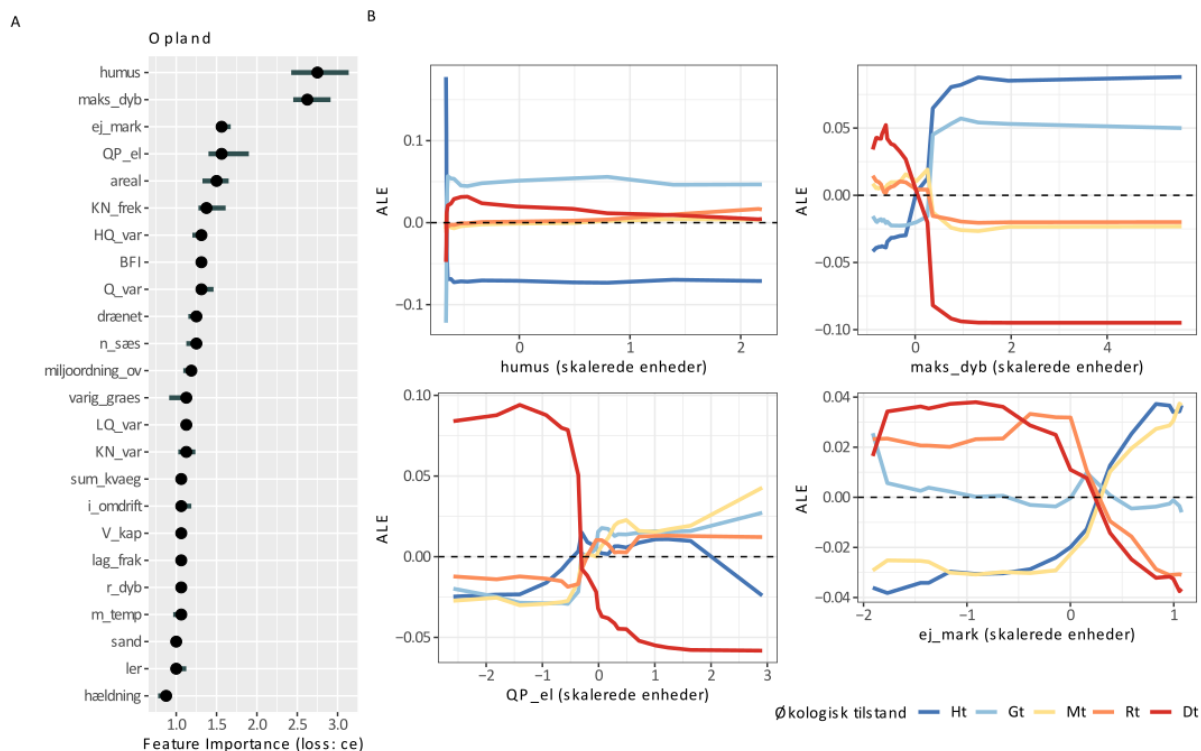
Den globale vigtighed af variable (figur 4.12) fremhævede maksimumsdybde, organisk jord og sammenhængen mellem nedbør og vandføring som vigtige forklarende variable, med de samme sammenhænge som beskrevet ovenfor.



Figur 4.12. Vigtigheden af variable for bufferzonen 250 m. A: Global vigtighed af variable. B: Akkumulerede lokale effekter, som viser modellens gennemsnitlige forklaringsevne for hver variabel opdelt efter økologisk status. Forklaring af variable fremgår af kapitel 2, og forklaring af figurer fremgår af figur 4.9.

For et bufferområde svarende til hele oplandet (figur 4.13) opnåede random forest-modellen en AUC på 0,6 (individuelle AUC-værdier for Ht = 0,87, Gt = 0,67, Mt = 0,53, Rt = 0,44 og Dt = 0,82). Modellen havde en begrænset evne til at klassificere moderat og ringe økologisk status korrekt. Samlet set var AUC-værdien lav og tæt på tilfældig klassifikation (0,5).

Den globale vigtighed af variable (figur 4.13) viste, at organisk jord, maksimumsdybde, sammenhængen mellem vandføring og nedbør samt ikke-dyrkede arealer var blandt de vigtigste forklarende variable. Mønstrene i de akkumulerede lokale effekter svarede til dem, der er beskrevet ovenfor.



Figur 4.13. Vigtigheden af variable for oplandet. A: Global vigtighed af variable. B: Akkumulerede lokale effekter, som viser modellens gennemsnitlige forklaringssevne for hver variabel opdelt efter økologisk status. Forklaring af variable fremgår af kapitel 2, og forklaring af figurer fremgår af figur 4.9.

Conditional inference trees – beslutningstræer

CIT-analyser (conditional inference trees, beslutningstræer) blev anvendt for hver bufferzone og hele oplandet til at belyse de faktorer og tærskelværdier, der bestemmer sommer TN, sommer TP, økologisk status og næringsstofbe- grænsning i søer uden veldefinerede oplande. Resultaterne for oplandet præ- senteres nedenfor, mens beslutningstræer for bufferzonerne er vist i bilagsfi- gurerne 7.3–7.5.

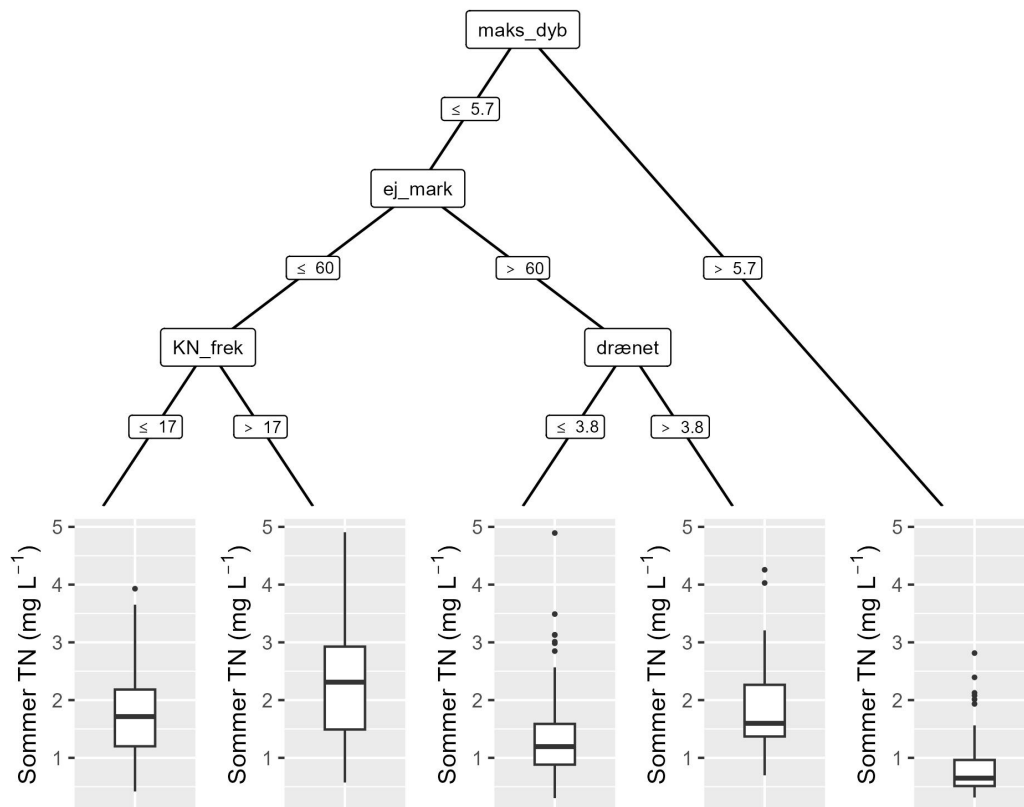
Sommer-TN

Beslutningstræer blev konstrueret med henblik på at identificere de vigtigste parametre og tærskelværdier, der regulerer TN-koncentrationerne om som- meren. Blandt alle landskabs- og søbaserede forklarende variable fremstod maksimumsdybde, andelen af ikke-dyrkede arealer (*ej_mark*), andelen af po- tentielt drænet areal (*drænet*) samthyppegheden af kraftige nedbørshændelser (*KN_frek*) som de vigtigste variable (figur 4.14).

Maksimumsdybde (*maks_dyb*) var den stærkeste forklarende variabel for som- merTN. I lavvandede søer (< 5,7 m) forekom yderligere opdelinger baseret på andelen af ikkedyrkede arealer (*ej_mark*) og hyppigheden af kraftig nedbør. Søer med færre kraftige nedbørshændelser (< 17) havde lavere TNkoncentra- tioner end søer, der var udsat for hyppigere ekstreme nedbørshændelser. Blandt lavvandede søer med en højere andel af ikke-dyrkede arealer (*ej_mark* > 54%) blev TNniveauerne yderligere påvirket af oplandets dræningsforhold. Specifikt havde søer med en drænet arealandel $\leq 3,8$ relativt lave TNværdier, mens søer over denne tærskel udviste markant højere TNkoncentrationer.

I modsætning hertil udviste dybere søer (*maks_dyb* > 5,7 m) konsekvent lave TN-niveauer, og ingen yderligere forklarende variable forbedrede modellens opdeling signifikant. Dette indikerer, at dybden alene er tilstrækkelig til at

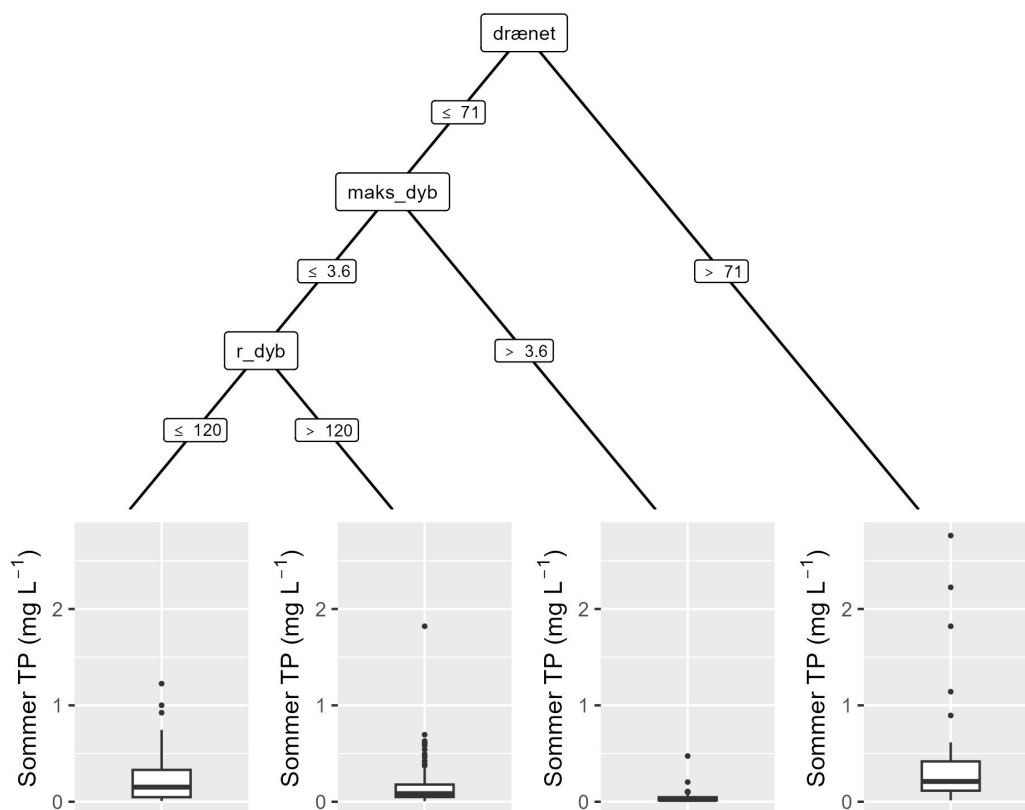
forklare den observerede variation i sommer-TN i dybere søer. Den forklarende varians var 0,31, hvilket afspejler en rimelig evne til at forklare variationen i TN-koncentrationer.



Figur 4.14. Beslutningstræ, der forklarer koncentrationer af sommer totalkvælstof (sommer TN) i søer. Forklaring af variable fremgår af kapitel 2.

Sommer-TP

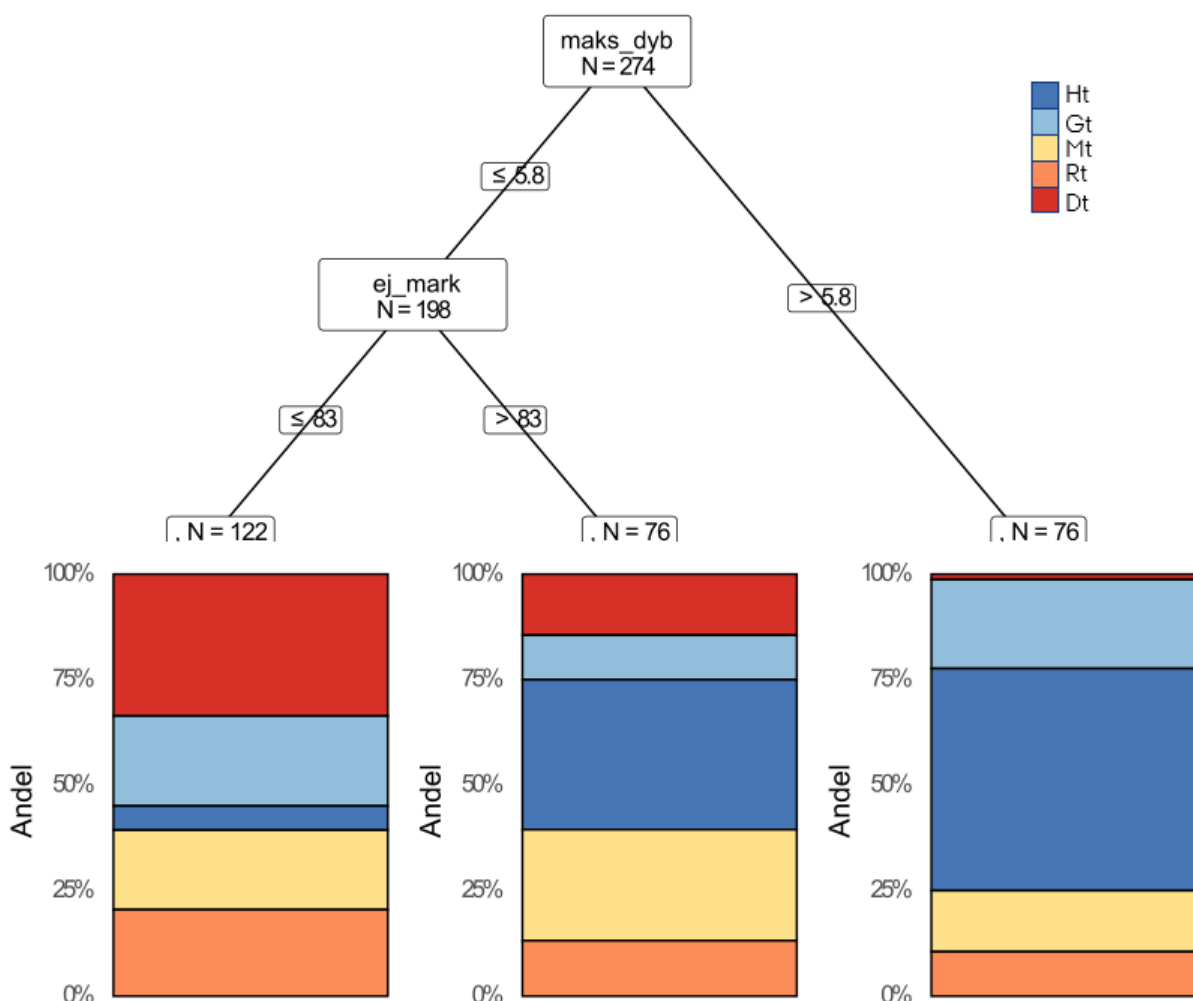
Beslutningstræer med sommer-TP som responsvariabel identificerede andelen af potentielt drænet areal som den primære faktor, der differentierede TP-koncentrationerne (figur 4.15). Søer med en højere dræningsprocent ($>71\%$) udviste konsekvent højere TP-koncentrationer. Blandt søer med en lavere andel af potentielt drænet areal ($\leq 71\%$) ledte maksimumsdybde til en yderligere opdeling. Søer, der var dybere end 3,6 m, havde forholdsvis lave TP-koncentrationer, mens lavvandede søer havde højere sommer-TP-koncentrationer. For de lavvandede søer var *R_dyb* (roddybe, som afspejler jordens infiltrationsevne i oplandet) også signifikant. Søer med oplande med lavere roddybe (≤ 120 cm) havde tendens til højere sommer-TP. Den forklarende varians var 0,16, hvilket indikerer, at modellen har en begrænset forklarende evne.



Figur 4.15. Beslutningstræ, der forklarer koncentrationer af sommer totalkvælstof (sommer TP) i søer. Forklaring af variable fremgår af kapitel 2.

Økologisk tilstand

Beslutningstræer med økologisk tilstand som responsvariabel identificerede maksimumsdybde som den primære faktor, der påvirkede den økologiske tilstand (figur 4.16). Dybere søer (>5,8 m) havde generelt en bedre økologisk tilstand. Blandt de lavvandede søer bidrog andelen af ikkedyrket areal yderligere til at differentiere den økologiske tilstand. Søer med en højere andel af ikkedyrket areal (>83 %) udviste tendens til at have en højere økologisk tilstand, mens søer med lavere andele af ikkedyrket areal typisk havde en lavere økologisk tilstand. Modellens samlede nøjagtighed var 0,35, med en kap-paværdi på 0,16, hvilket indikerer, at modellens præstation var dårlig.



Figur 4.16. Beslutningstræ, der forklarer den økologiske tilstand i søer. Forklaring af variable fremgår af kapitel 2.

Sammenfattende viste CIT-analyserne tydelige landskabseffekter til trods for modellens lave nøjagtighed. Maksimumsdybde, andelen af drænet areal og andelen af ikke-dyrkede arealer forklarede størstedelen af variationen i søernes næringsstofniveauer og økologiske tilstand. En højere andel af ikke-dyrkede arealer og en lavere dræningsintensitet var forbundet med lavere TN- og TP-koncentrationer. Sammenhængen mellem landskabskarakteristika og økologisk tilstand var svagere, idet kun dybde og andelen af ikke-dyrkede arealer adskilte søgrupperne. Overordnet understreger resultaterne vigtigheden af dybde og arealanvendelse (andel af ikke-dyrkede arealer og drænet areal) for næringsstofdynamikken.

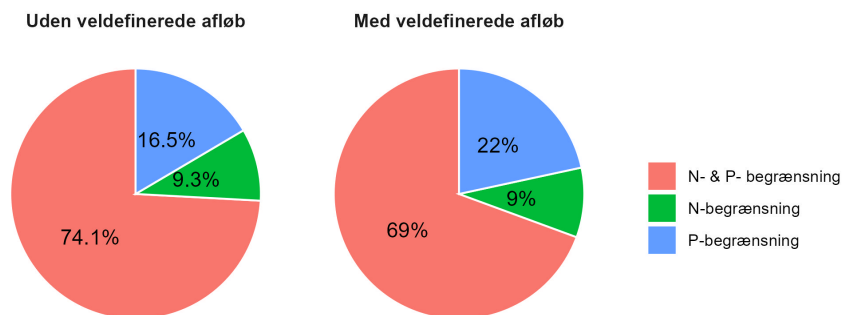
4.4 N- eller P-begrænsning

Den støkiometriske metode

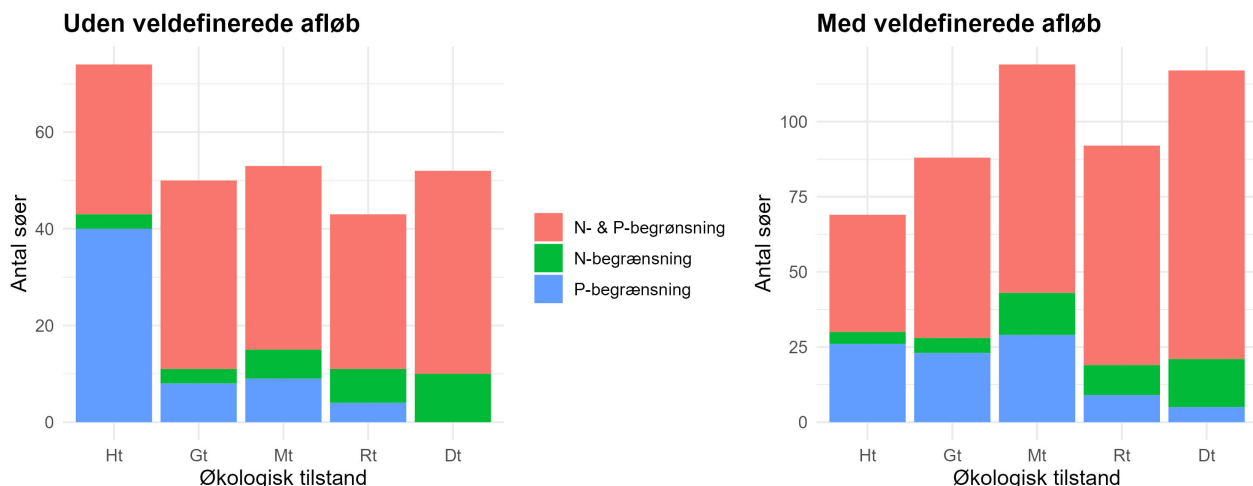
Ifølge klassifikationen af næringsstofbegrænsning baseret på Ladwig et al. (2026, under udarbejdelse) blev størstedelen af søerne uden veldefineret afløb (74 %) vurderet til at være begrænset af både N og P, mens 17 % var P-begrænsede og 9 % var N-begrænsede. Til sammenligning havde søer med veldefineret afløb en højere andel af P-begrænsede søer og en lavere andel af N&P-begrænsede søer end søer uden veldefineret afløb (figur 4.17). I denne gruppe var 69 % af søerne N&P-begrænsede, 9 % N-begrænsede og 22 % P-begrænsede. I henhold til Pearson's Chi-squared test var forskellen mellem søer med og uden veldefineret afløb signifikant ($\chi^2 = 9,7632$, $df = 2$, $p = 0,0076$).

Selvom forskellen mellem søer med og uden veldefineret afløb ikke er stor, bekræfter resultaterne tidligere resultater beskrevet i Søndergaard et al. (2023), som tyder på, at søer uden veldefineret afløb oftere er N-begrænsede.

Figur 4.17. Fordeling af typer af næringsstofbegrænsning for søer uden (venstre) og med veldefineret afløb (højre).



For søer uden veldefineret afløb var N-begrænsede søer overvejende forbundet med ringe og dårlig økologisk tilstand (figur 4.18, Dt og Rt). Omvendt var P-begrænsning mere udbredt blandt søer med moderat, god og høj økologisk tilstand (Mt, Gt og Ht). N&P-begrænsede søer udviste ikke noget klart mønster på tværs af de økologiske tilstandsklasser, men var forholdsvis jævnt fordelt i alle kategorier. Søer med veldefineret afløb (figur 4.18, højre panel) udviste et tilsvarende overordnet mønster, idet N&P-begrænsning var klart dominerende på tværs af alle økologiske tilstande. Der var dog markante forskelle mellem de to søgrupper. Søer med veldefineret afløb havde færre P-begrænsede søer i høj økologisk tilstand (Ht) og flere N&P-begrænsede søer fra moderat til dårlig tilstand (Mt-Dt). Dette indikerer, at P-begrænsning relativt set er mere udbredt i søer med høj økologisk tilstand uden klart definerede oplande, mens N&P-begrænsning dominerer i søer med veldefineret afløb uanset økologisk tilstand.

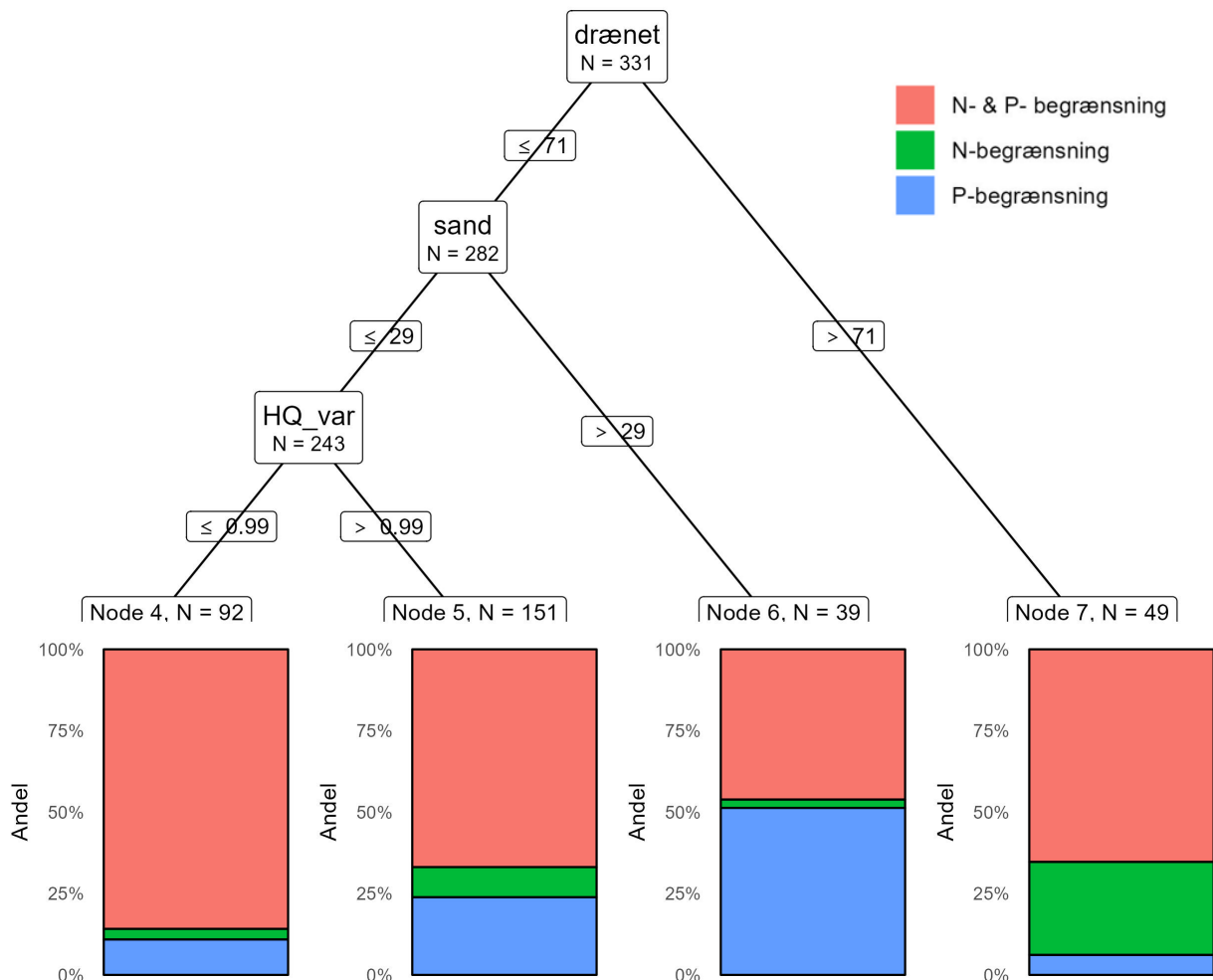


Figur 4.18. Fordeling af typer af næringsstofbegrænsning på tværs af økologiske tilstandsklasser for søer uden (venstre) og med (højre) veldefineret afløb.

Conditional inference tree – beslutningstræer

Beslutningstræer identificerede andelen af potentielt drænet areal (*drænet*) som den vigtigste variabel til adskillelse af klasserne for næringsstofbegrænsning (figur 4.19). Søer beliggende i oplande med mere end 71 % drænet areal havde en højere andel af N-begrænsede og N&P-begrænsede søer. For søer

med ≤ 71 % drænet areal bidrog andelen af sandede jorde i oplandet (*sand*) yderligere til at strukturere responsen. Blandt disse søer var søer med ≤ 29 % sandet jord overvejende N&P-begrænsede, mens søer med > 29 % sandet jord ud over N&P-begrænsning også omfattede P-begrænsede søer. For søer med < 29 % sandet jord viste en yderligere opdeling baseret på varigheden af perioder med højt vandflow (*HQ_var*), at andelen af P-begrænsede søer steg ved *HQ_var* $> 0,99$. Samlet set indikerer modellen, at P-begrænsede søer primært forekommer i oplande med en højere andel af sandede jorde og lavere dræningsintensitet, mens N-begrænsede søer er forbundet med oplande karakteriseret ved en højere andel af drænet areal. På trods af disse mønstre var den overordnede modelpræstation lav, med en nøjagtighed på 0,65 og en kappa-værdi på -0,01, hvilket indikerer ringe overensstemmelse ud over den, der kan forventes ved en tilfældighed.



Figur 4.19. Beslutningstræ, der forklarer næringsstofbegrænsningsklasser i søer. Se kapitel 2 for forklaring af variablene.

4.5 Palæoekologiske undersøgelser

I den tidligere rapport foreslog Søndergaard et al. (2023), at palæoekologiske undersøgelser kan være et nyttigt redskab til at identificere reference- eller nær-reference tilstande i søer. Dette er særligt vigtigt i forbindelse med vandrammedirektivet, da det bidrager til at fastlægge den økologiske tilstand, der bør sigtes mod i forvaltningen af søer. Ved at analysere makrofossile rester af planter og dyr i forskellige sedimentdybder i kombination med sedimentdatering kan sådanne undersøgelser give indsigt i tidligere økologiske forhold, fx for 100 eller 200 år siden.

Den historiske tilstand og referenceforholdene for søer uden veldefinerede oplande (og uden tilløb og afløb) er i høj grad ukendt. Det er således muligt, at deres nuværende "påvirkede" tilstand også kan repræsentere deres naturlige tilstand; der findes imidlertid kun begrænset dokumentation herfor. Da overvågningsdata ofte er for kortvarige til at indfange langsigtet økologisk variation, kan palæolimnologiske tilgange bidrage til at rekonstruere tidligere forhold og skelne mellem naturlig variation og menneskeskabte ændringer. Sådanne undersøgelser kan derfor give indsigt ikke blot i den historiske tilstand eller referenceforholdet for en sø eller en søtype, men også i det nuværende påvirkningsniveau.

En palæoøkologisk undersøgelse kunne omfatte (se også bemærkninger nedenfor):

1. Indsamling af lange (>100 cm) sedimentkerner fra en sø/søer) uden veldefineret opland, ideelt set både fra det dybeste område og fra littoralzonen.
2. Gennemførelse af røntgenfluorescensanalyse (XRF) på kernerne kombineret med analyser af tørvægt, organisk kulstof (glødetab) og våd densitet.
3. Datering af sedimentkernerne ved hjælp af Pb-210 og Cs-137.
4. Analyse af sedimentære pigmentkoncentrationer i kernerne. Dette kan udføres som en sammenligning mellem bund- og overfladekoncentrationer; dog er analyser af regelmæssigt fordelte prøver at foretrække.
5. Analyse af makrofossile planterester i sedimenterne.
6. Analyse af rester af cladocéer i sedimenterne.

Analyser af tørvægt og organisk materiale giver væsentlige baggrundsdata for tolkningen af sedimentkerner.

XRF er en højt opløselig elementanalyse af sedimentkerner, hvilket kan påvise historiske forureningsniveauer.

Datering af sedimentkernerne gør det muligt at bestemme de forskellige sedimenttybders alder og dermed fastlægge referenceperioden samt at vurdere hastigheden af og tidspunktet for væsentlige ændringer.

Analyse af sedimentære pigmenter kan indikere historiske niveauer af klorofyl *a* og pigmenter fra blågrønalger (cyanobakterier), som begge er vigtige for at definere søernes tidligere økologiske tilstand.

Makrofossile planterester giver information om de historiske plantesamfund i søen og deres ændringer over tid, hvilket bidrager til fortolkningen af økologiske skift. Dette giver et unikt historisk perspektiv på søens nuværende tilstand i relation til tidligere ændringer.

Cladocé-samfund formes både af *bottom-up*-processer, såsom næringsstofbelastning, og *top-down*-processer, såsom prædation fra fisk. Ændringer i cladocé-samfundene kan derfor anvendes til en integreret vurdering af en søs økologiske tilstand, idet de afspejler balancen mellem bentisk og pelagisk produktion samt intensiteten af fiskeprædation. Kombinationen af plante-makrofossiler og analyser af cladocéer er særligt velegnet til at spore tidligere økologiske ændringer.

En detaljeret palæolimnologisk undersøgelse vil kunne give det bedste overblik over historiske ændringer. Det anbefales at udtage mindst to sedimentkerner pr. sø – én fra littoralzonen og én fra den pelagiske zone. Begge kerner

bør analyseres for tørvægt, glødetab og sammensætningen af grundstoffer (XRF) og efterfølgende dateres. Den pelagiske kerne bør analyseres for cladocé-rester og sedimentære pigmenter, mens den littorale kerne bør analyseres for planterester, begge fra cirka 10 dybder i kernen.

Af hensyn til omkostninger og gennemførlighed kan der også udføres en indledende kvalitativ vurdering baseret på et udvalg af prøver fra hver kerne (fx nogle få fra de nederste dybder kombineret med et par stykker fra overfladen). Hvis potentielt vigtige ændringer påvises, kan en mere detaljeret analyse som beskrevet ovenfor gennemføres. Denne indledende tilgang kan give en grov indikation af økologiske ændringer i søen over de seneste par hundrede år og bidrage til at vurdere, om en mere detaljeret analyse er anbefalelsesværdig.

5 Konklusioner og anbefalinger

Dette studie er en opdateret analyse af, hvordan landskabskarakteristika kan påvirke vandkvalitet, økologisk tilstand og klassificering af næringsstofbe-grænsning i danske søer uden veldefinerede topografiske oplande. En multi-variat statistisk tilgang (reduktionsanalyse, RDA) samt to maskinlæringsme-toder (Random Forest og Conditional Inference Trees/beslutningstræer) blev anvendt til at afdække potentielle effekter af landskabskarakteristika på søer-nes næringsstoffer og økologiske tilstand. Samlet set bidrager disse metoder med komplementære indsigter i de potentielle effekter af landskabskakteri-stika på søernes næringsstofniveauer og økologiske tilstand.

RDA viste, at 25–32 % af variationen i søparametre kunne forklares af land-skabsvariable. Andelen af organisk jord, roddebyde, andelen af ikke-dyrkede arealer samt andelen af potentielt drænede arealer fremstod som vigtige for-klarende variable for forholdene i søerne. Organisk jord (*humus*) og drænede arealer (*drænet*) var konsekvent forbundet med højere næringsstofkoncentra-tioner og lavere økologisk tilstand, mens en højere andel af ikke-dyrkede are-aler (*ej_mark*) og større roddebyde (*R_dyb*) var knyttet til lavere næringsstofni-veauer og bedre økologisk tilstand. Derimod var overvejende landbrugsarea-ler (*i_omdrift*) og arealer med varigt græs (*varig_græs*) forbundet med søer med ringe eller dårlig økologisk tilstand. Dette stemmer overens med generelle fund i litteraturen, som viser, at landbrug og intensivt forvaltede græsarealer øger tabet af kvælstof og fosfor gennem gødskning, jorderosion og dræning, hvilket efterfølgende fører til øgede næringsstofkoncentrationer i søerne og forringer den økologiske tilstand (Nielsen et al., 2012; Jeppesen et al., 2007). Studier fra danske og andre europæiske søer fremhæver ligeledes, at nærings-stofekspert fra landbrugsarealer er en væsentlig kilde til næringsstoffer i dan-ske overfladevande, og at oplande domineret af intensivt landbrug har en ten-dens til at understøtte eutrofe forhold (Jeppesen et al., 2007).

Oplandskarakteristika, såsom den samlede vandtilgængelighed, var ligeledes signifikante ved alle bufferafstande, og nedbørsfrekvensen var signifikant i bufferzoner på 10 m, 50 m og 250 m. Overordnet set var vådere oplande med en højere andel af organiske jorde og en større grad af dræning forbundet med ringere vandkvalitet. Tilsvarende viste et studie af 52 tempererede søer – her-under danske søer – at oplandets jordbundsforhold, særligt indholdet af or-ganisk kulstof i jorden, var stærke forklarende variable for søernes DOC (op-løst organisk kulstof), TON (totalt organisk kvælstof) og TP. Højere andele af organisk jord var forbundet med forhøjede næringsstofniveauer, hvilket indi-kerer, at oplandets jordtyper har stor betydning for søernes næringsstoffil-stand (Sepp et al., 2022).

Bufferzonens størrelse påvirkede de forklarende variables signifikansni-veauer; dog forblev de samme forklarende variable overvejende signifikante, og de overordnede mønstre var ens på tværs af skalaer. Efterhånden som buf-ferstørrelsen øgedes, blev effekten af potentielt drænede arealer svagere, mens vigtigheden af ikke-dyrkede arealer blev stærkere. Vigtigheden af dræ-ning er ligeledes fremhævet i tidligere rapporter om vandkvalitet (Sønder-gaard, 2023; Nielsen et al., 2012), idet Nielsen et al. (2012) viste, at kunstige markdræn kan stå for en stor andel af det kvælstof, der eksporteres fra land-brugsarealer. Dette skyldes, at en betydelig del af de danske landbrugsjorde

er undergrundsdrænede, og sådanne dræn går uden om de naturlige transportveje gennem jord og grundvand (Olesen, 2009; Møller et al., 2018).

Random forest-modellen havde vanskeligt ved præcist at klassificere søer med god, moderat og ringe økologisk tilstand, men opnåede generelt en tilstrækkelig adskillelse mellem høj og dårlig økologisk tilstand. Dette kan enten afspejle manglende information om vigtige kovariater, der styrer de økologiske forhold i søerne, eller omfattende støj i data, som vanskeliggør en præcis klassifikation. Ikke desto mindre identificerede modellerne konsekvent en række vigtige variable for klassifikationen: maksimumsdybde, andel af organisk jord, areal, sammenhængen mellem vandføring og nedbør, dræningsprocent, andel af ikke-dyrkede arealer samt græsarealer. Søer i høj økologisk tilstand var kendetegnet ved større maksimumsdybde, lav andel af organisk jord, større areal, lav sammenhæng mellem vandføring og nedbør, lav andel af drænede arealer, højere andel af ikke-dyrkede arealer og mindre græsareal. Søer i dårlig økologisk tilstand viste det modsatte mønster, dog med undtagelse af sammenhængen mellem vandføring og nedbør.

CIT-analysen (*Conditional Inference Trees* – beslutningstræer) fremhævede tydelige tærskelrelationer, selvom modellens præcision og forklarede varians var lav. Maksimumsdybde var den stærkeste forklarende variabel for både TN og økologisk tilstand, idet dybere søer (maksimumsdybde >5,7–5,8 m) konsekvent udviste lavere næringsstofniveauer og bedre økologisk tilstand. Blandt lavvandede søer bidrog dræningsintensitet, andelen af ikke-dyrkede arealer samt ekstreme nedbørshændelser yderligere til at strukturere næringsstofkoncentrationer og økologisk tilstand. En stigende andel af ikke-dyrkede arealer, reduceret dræning og færre kraftige nedbørshændelser var forbundet med lavere TN-koncentrationer. For TP var andelen af drænede arealer og dybde ligeledes vigtige faktorer, idet dybere søer generelt havde lavere TP-koncentrationer og en lavere andel af drænede arealer. Sammenhængen mellem landskabskarakteristika og økologisk tilstand var svagere, idet kun dybde og andelen af ikke-dyrkede arealer bidrog til at opdele søerne i grupper. Søer med en maksimumsdybde over 5,8 m havde generelt en bedre økologisk tilstand, mens andelen af ikke-dyrkede arealer var afgørende for lavvandede søer. Lavvandede søer med mindre end 83 % ikke-dyrkede arealer havde dårligere økologisk tilstand. Søtypologi er vigtig for forståelsen af de underliggende mekanismer, idet lavvandede søer typisk er polymiktiske med hyppig opblanding, hvor interaktioner mellem sediment og vand øger tilgængeligheden af næringsstoffer. Dybere søer er derimod stratificerede i længere perioder, hvilket ændrer den interne næringsstoffdynamik. Da størstedelen af søerne i datasættet var lavvandede, undersøgte vi, om dybde blot adskilte lavvandede og dybe søtyper, eller om variabelen også var informativ inden for lavvandede systemer. Beslutningstræerne viste, at dybde var bestemmende for den første opdeling, men at yderligere landskabs- og oplandsvariable strukturerede variationen blandt de lavvandede søer. Dette indikerer, at dybde bestemmer den primære respons, mens oplandsforhold modulerer næringsstofniveauerne i lavvandede systemer.

Typen af næringsstoffbegrænsning (N, P eller N&P) blev associeret med dræningsintensitet, andelen af sandjorde samt hydrologisk aktivitet. P-begrænsede søer forekom hyppigst i sandede oplande med lavere grad af kunstig dræning, mens N-begrænsede søer primært fandtes i kraftigt drænede systemer, hvilket igen er i overensstemmelse med den forventede hydrologiske kontrol af næringsstofftransporten.

Martinsen & Sand-Jensen (2022) fandt tilsvarende klare sammenhænge mellem landskabet og søer på tværs af danske søer i det nationale overvågningsprogram for overfladevand. I deres analyse var de vigtigste forklarende variable dog primært terrænbaserede såsom krumning og ujævnhed i terrænet, mens virkningen af arealanvendelsesvariable var forholdsvis begrænset. Denne forskel kan forklares ved de bredere topografiske gradienter i deres nationale datasæt. I nærværende studie analyseredes kun et deludsnit af danske søer beliggende i svagt kuperede landskaber, hvor variationen i terrænvariable er minimal. Isådanne områder bliver terrænbaserede forklarende variable mindre informative, og arealanvendelses- og jordbundskaraktistika forklarer en større del af den observerede variation. I overensstemmelse med tidligere resultater baseret på lineær regression (Søndergaard et al., 2023) havde anvendelsen af ikke-lineære maskinlæringsmetoder også en lav forklaringssevne. Samlet set bekræfter resultaterne, at landskabsvariable kun forklarer en del af variationen i søerne tilstand, og at der er flere mekanismer på oplandsskala, som har stor betydning for næringsstofniveauerne i søer uden veldefineret opland. Disse omfatter andelen af drænede arealer, jordbundens sammensætning, ikke-dyrkede arealer (ej-mark), ekstreme nedbørshændelser samt søernes morfologi. Den lave forklaringssevne understreger samtidig den fortsatte vigtighed af andre faktorer, såsom intern belastning, vandsøjleens opblandingsdynamik og bidrag fra grundvand, som ikke i tilstrækkelig grad fanges af overfladebaserede oplandsafgrænsninger. Desuden kan søernes nuværende tilstand være påvirket af den historiske næringsstofftilførsel, som ikke afspejles i de nuværende oplandskaraktistika. Det bør desuden fremhæves, at den økologiske tilstand anvendt i denne rapport udelukkende er baseret på klorofyl *a* koncentrationer, mens den formelle vurdering af økologisk tilstand omfatter alle fem biologiske kvalitetselementer. På grund af de komplekse interaktioner i søkosystemer kan sammenhænge mellem økologisk tilstand og oplandskaraktistika sløres af interne processer og feedbackmekanismer, hvilket gør statistiske sammenhænge vanskelige at påvise.

Anbefalinger

Selvom modellernes forklaringssevne er lav, identificerer resultaterne konsekvent flere landskabskaraktistika, som er vigtige for at forklare vandkvalitet og økologisk tilstand i søer uden veldefinerede topografiske oplande.

For det første kan en reduktion af dræningsintensiteten – særligt i oplande med en høj andel af organiske jorde – bidrage til at begrænse transporten af næringsstoffer til søerne, idet drænede organiske jorde gentagne gange er blevet associeret med forhøjede næringsstofkoncentrationer. En reduktion af kunstige dræningsnetværk kan derfor give målbare forbedringer.

For det andet kan en forøgelse eller beskyttelse af eksisterende af ikke-dyrkede arealer (*ej-mark*) ligeledes fungere som buffer mod næringsstofftilførsel, da disse arealer har lavere næringsstofniveauer og bedre økologisk tilstand.

Endelig bør forvaltningsstrategier eksplicit tage hensyn til jordbundens sammensætning. Jorde med et højt indhold af organisk materiale kombineret med intensiv dræning havde de højeste næringsstofkoncentrationer, hvilket indikerer, at sørestaureringstiltag i disse områder kan være særligt effektive til at reducere næringsstoffbelastningen og derfor bør prioriteres. På grund af den relativt lave forklarede variation bør disse resultater betragtes som forebyggende, supplerende tiltag, der skal kombineres med vurderinger af intern belastning, grundvandsforbindelser, søspecifikke processer samt palæolimnologisk evidens.

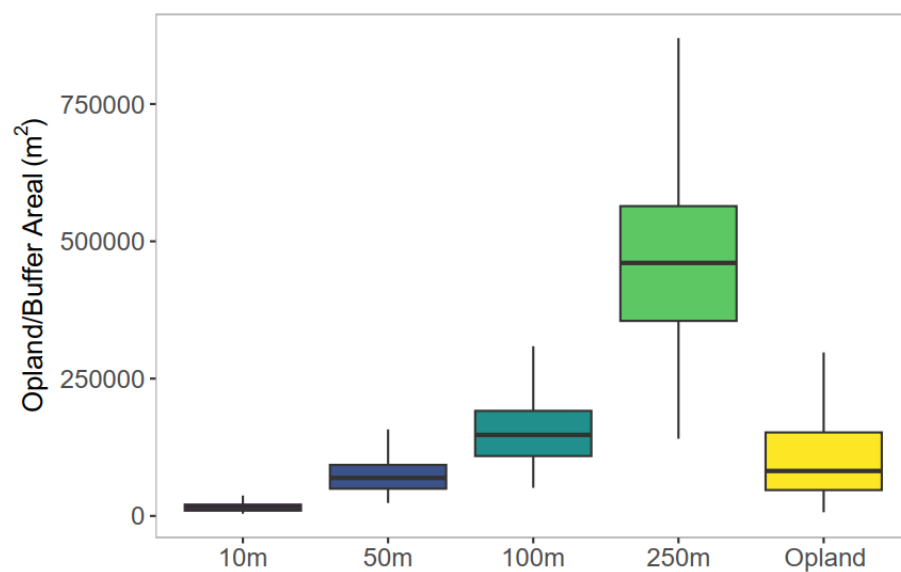
6 Referencer

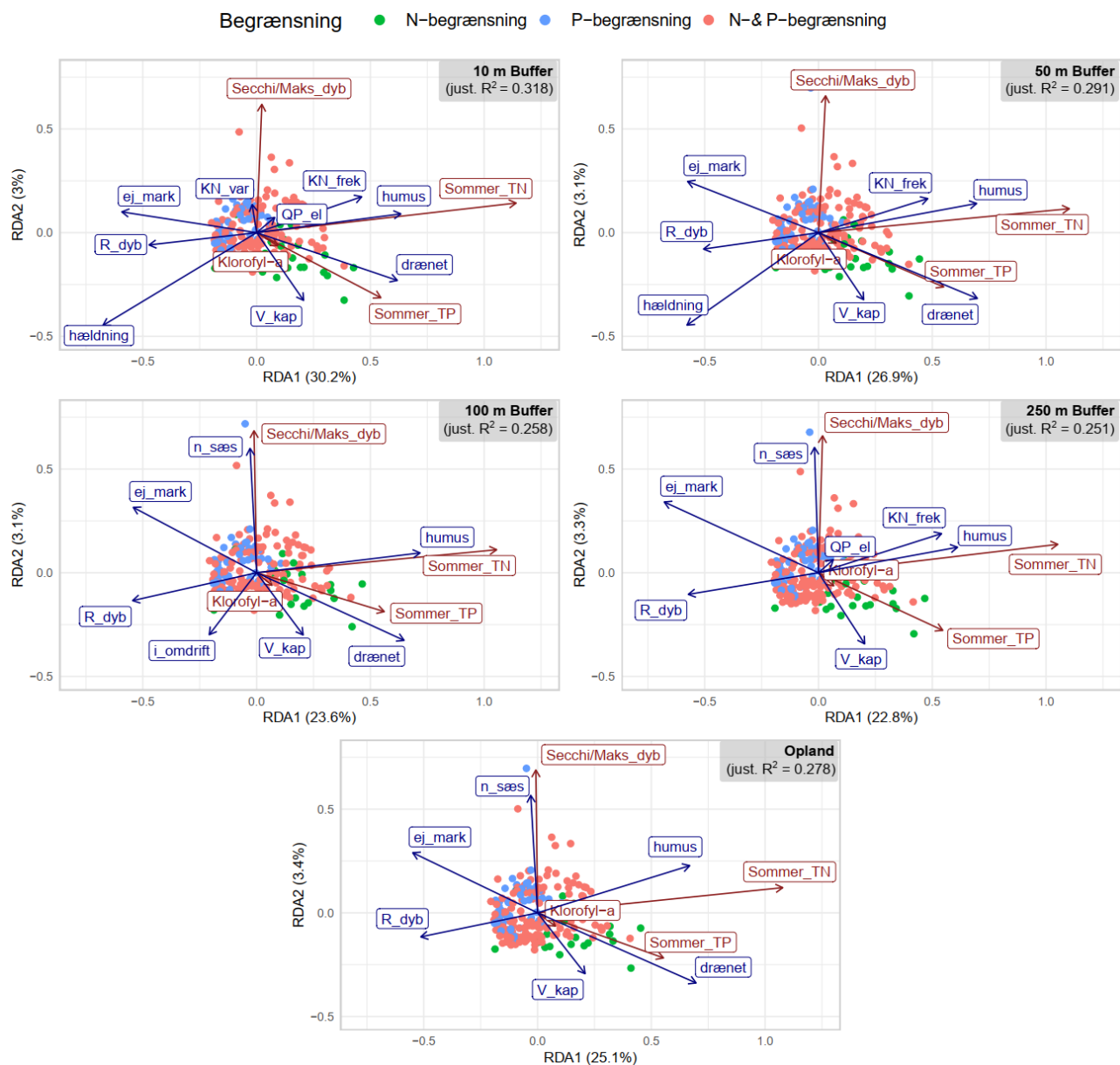
- Addor, N., Newman, A. J., Mizukami, N., & Clark, M. P. (2017). The CAMELS data set: Catchment attributes and meteorology for large-sample studies. *Hydrology and Earth System Sciences*, 21(10), 5293–5313.
- Breiman, L. (2001). Random forests. *Machine learning*, 45(1), 5–32.
- Coxon, G., et al. (2020). CAMELS-GB: Hydrometeorological time series and landscape attributes for 671 catchments in Great Britain. *Earth System Science Data*, 12(4), 2459–2483.
- Fan, Y., Miguez-Macho, G., Jobbágy, E. G., Jackson, R. B., & Otero-Casal, C. (2017). Hydrologic regulation of plant rooting depth. *Proceedings of the National Academy of Sciences*, 114(40), 10572–10577.
- Greve, M. B. (2024). *Revidering af jordbundstypekort 2024*. Rådgivningsnotat fra DCA – Nationalt Center for Fødevarer og Jordbrug, Aarhus Universitet, 4 sider. Leveret 17.05.2024.
- Hothorn, T., Seibold, H., & Zeileis, A. (2015). *partykit: A toolkit for recursive partitioning*.
- Höge, M., et al. (2023). CAMELS-CH: Hydrometeorological time series and landscape attributes for 331 catchments in Switzerland. *Earth System Science Data*, 15, 5755–5784.
- Jeppesen, E., Søndergaard, M., Lauridsen, T. L., Kronvang, B., Beklioglu, M., Lammens, E., ... & Tátrai, I. (2007). Danish and other European experiences in managing shallow lakes. *Lake and Reservoir Management*, 23(4), 439–451.
- Kuhn, M., & Vaughan, D. (2025). *parsnip: A common API to modeling and analysis functions* (R package version 1.4.0). <https://github.com/tidymodels/parsnip>
- Ladwig, R., Bucak, T., Johansson, L.S., Berthelsen, A., Søndergaard, M. (in prep). Nitrogen as a limiting nutrient in Danish lakes. Aarhus Universitet, DCE – Nationalt Center for Miljø og Energi.
- Liu, J., Koch, J., Stisen, S., Troldborg, L., Højberg, A. L., Thodsen, H., ... & Schneider, R. J. (2025). CAMELS-DK: hydrometeorological time series and landscape attributes for 3330 Danish catchments with streamflow observations from 304 gauged stations. *Earth System Science Data*, 17(4), 1551–1572.
- Martinsen, K. T., & Sand-Jensen, K. (2022). Predicting water quality from geospatial lake, catchment, and buffer zone characteristics in temperate lowland lakes. *Science of the Total Environment*, 851, 158090.
- McHugh, M. L. (2012). Interrater reliability: the kappa statistic. *Biochemia medica*, 22(3), 276–282.
- Molnar C, Bischl B, Casalicchio G (2018). “iml: An R package for Interpretable Machine Learning.” *JOSS*, 3(26), 786. <https://doi.org/10.21105/joss.00786>, <https://joss.theoj.org/papers/10.21105/joss.00786>.

- Moon, D. L., Scott, J. T., & Johnson, T. R. (2021). Stoichiometric imbalances complicate prediction of phytoplankton biomass in U.S. lakes: Implications for nutrient criteria. *Limnology and Oceanography*, 66(8), 2967–2978.
- Møller, A. B., Greve, M. H., & Beucher, A. M. (2024). *Opdateret jordbundstypekort*. Rådgivningsnotat fra DCA – Nationalt Center for Fødevarer og Jordbrug, Aarhus Universitet, 33 sider. Leveret 19.03.2024.
- Møller, A. B., Beucher, A., Iversen, B. V., & Greve, M. H. (2018a). Predicting artificially drained areas by means of a selective model ensemble. *Geoderma*, 320, 30–42.
- Møller, A. B., Børgesen, C. D., Bach, E. O., Iversen, B. V., & Moeslund, B. (2018b). *DCA rapport nr. 135*. DCA – Nationalt Center for Fødevarer og Jordbrug.
- Nielsen, A., Trolle, D., Søndergaard, M., Lauridsen, T. L., Bjerring, R., Olesen, J. E., & Jeppesen, E. (2012). Watershed land use effects on lake water quality in Denmark. *Ecological Applications*, 22(4), 1187–1200.
- Oksanen, J., Simpson, G., Blanchet, F., Kindt, R., Legendre, P., et al. (2025). *vegan: Community ecology package*. <https://doi.org/10.32614/CRAN.package.vegan>
- Palarea-Albaladejo, J., & Martín-Fernández, J. A. (2015). zCompositions: R package for multivariate imputation of left-censored data under a compositional approach. *Chemometrics and Intelligent Laboratory Systems*, 143, 85–96. <https://doi.org/10.1016/j.chemolab.2015.02.019>
- R Core Team (2025). R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna, Austria.
- Rock, L. A., & Collins, S. M. (2024). A broad-scale look at nutrient limitation and a shift toward co-limitation in United States lakes. *Environmental Science & Technology*, 58(26), 11482–11491.
- Sepp, M., Kõiv, T., Nõges, P., Nõges, T., Newell, S. E., & McCarthy, M. J. (2022). Catchment soil characteristics predict organic carbon, nitrogen, and phosphorus levels in temperate lakes. *Freshwater Science*, 41(1), 1–17.
- Søndergaard, M., Johansson, L. S., & Kjeldgaard, A. (2023). *Sammenhænge mellem oplandskarakteristika og tilstanden i søer uden veldefinerede afløb*. Aarhus Universitet, DCE – Nationalt Center for Miljø og Energi, 32 s. Videnskabelig rapport nr. 548.
- van den Boogaart K.G., Tolosana-Delgado R. and Bren M. (2025). *compositions: Compositional Data Analysis*. doi:10.32614/CRAN.package.compositions. R package version 2.0-9.
- Wright MN, Ziegler A (2017). “ranger: A Fast Implementation of Random Forests for High Dimensional Data in C++ and R.” *Journal of Statistical Software*, 77(1), 1–17. <https://doi.org/10.18637/jss.v077.i01>.

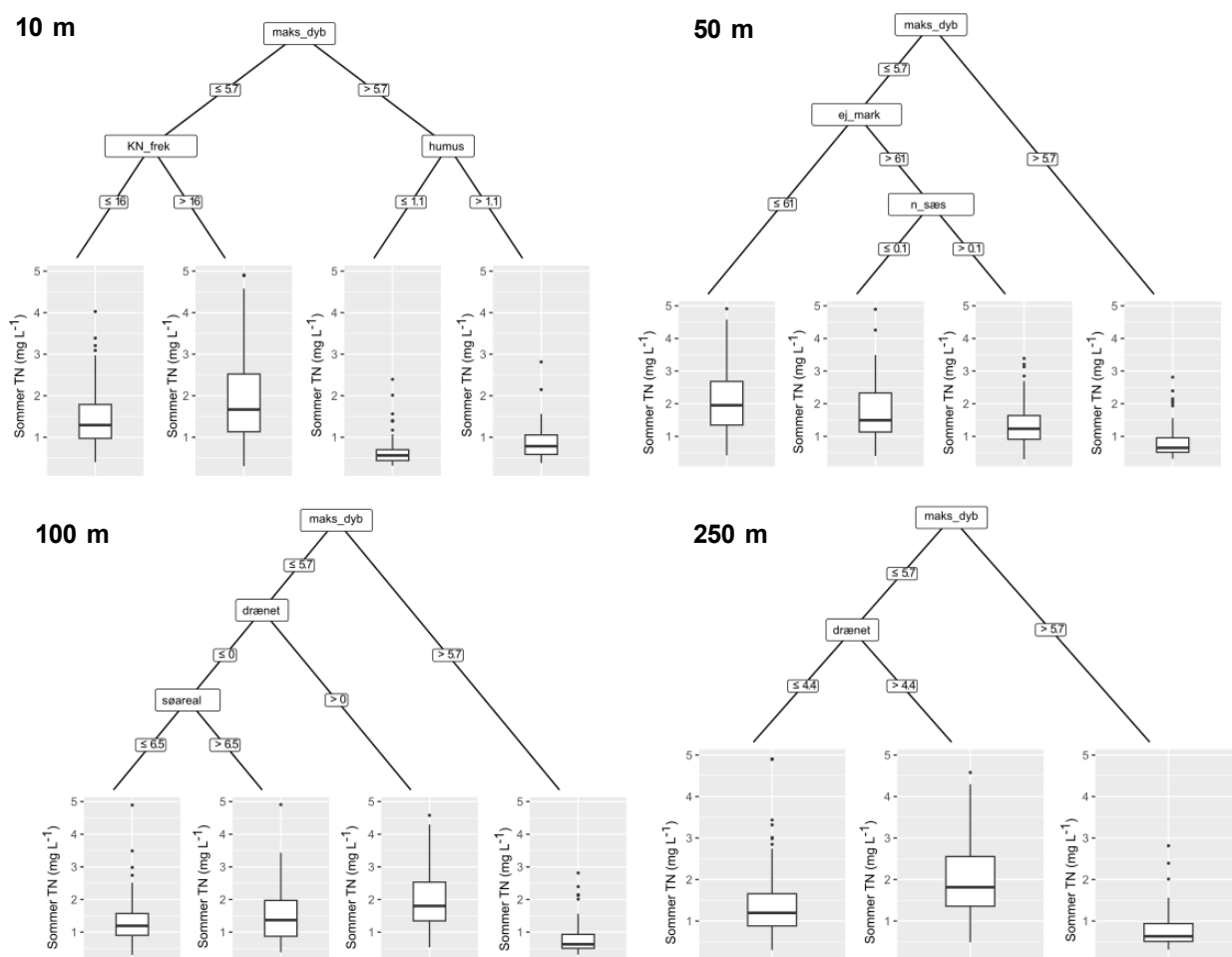
7 Bilag

Figur 7.1. Boksplot, der sammenligner arealerne af bufferzoner og oplande. Ekstreme værdier er fjernet af hensyn til visualiseringen. Boksplottet viser første (Q1) og tredje (Q3) kvartil, og den sorte linje repræsenterer medianværdien.

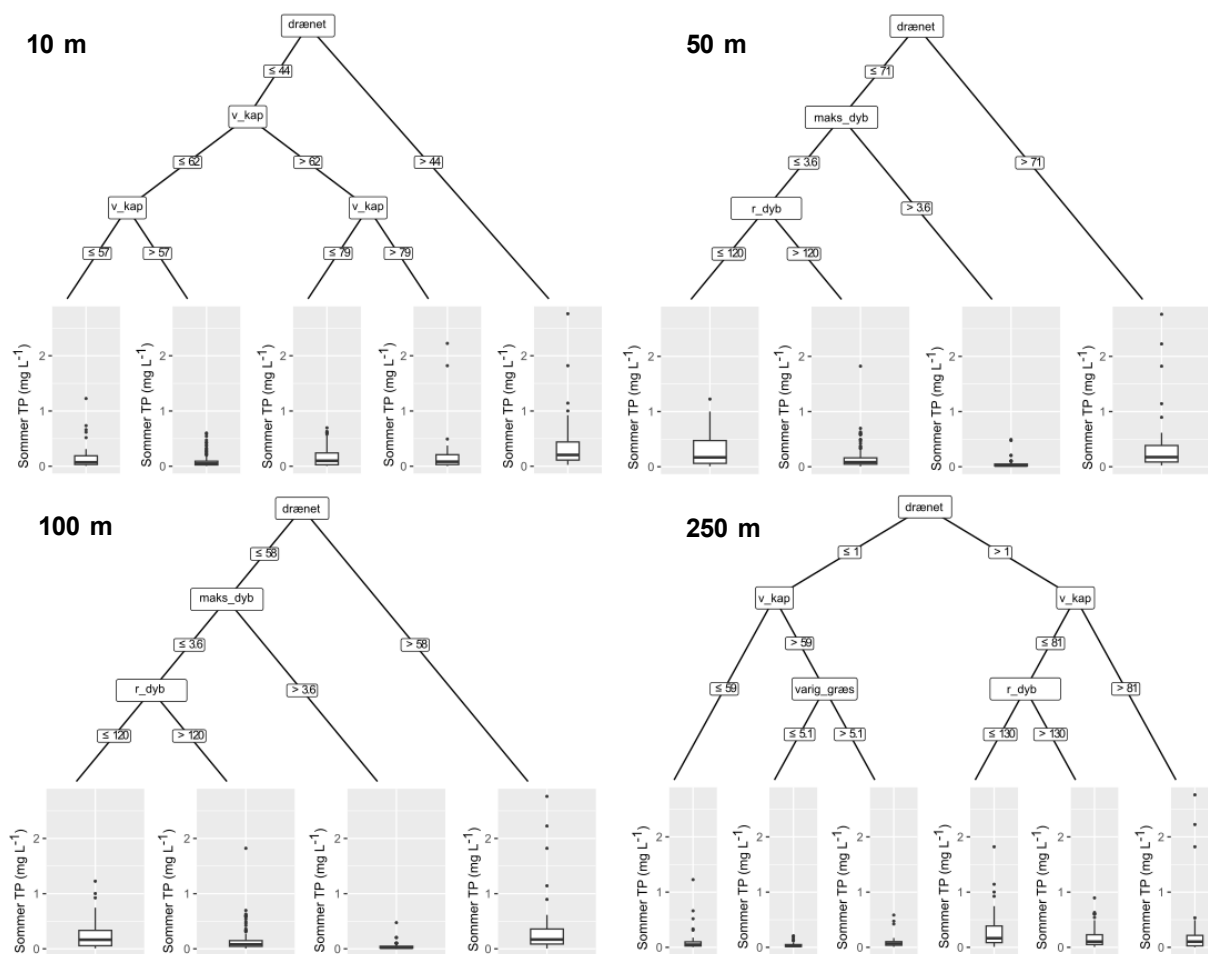




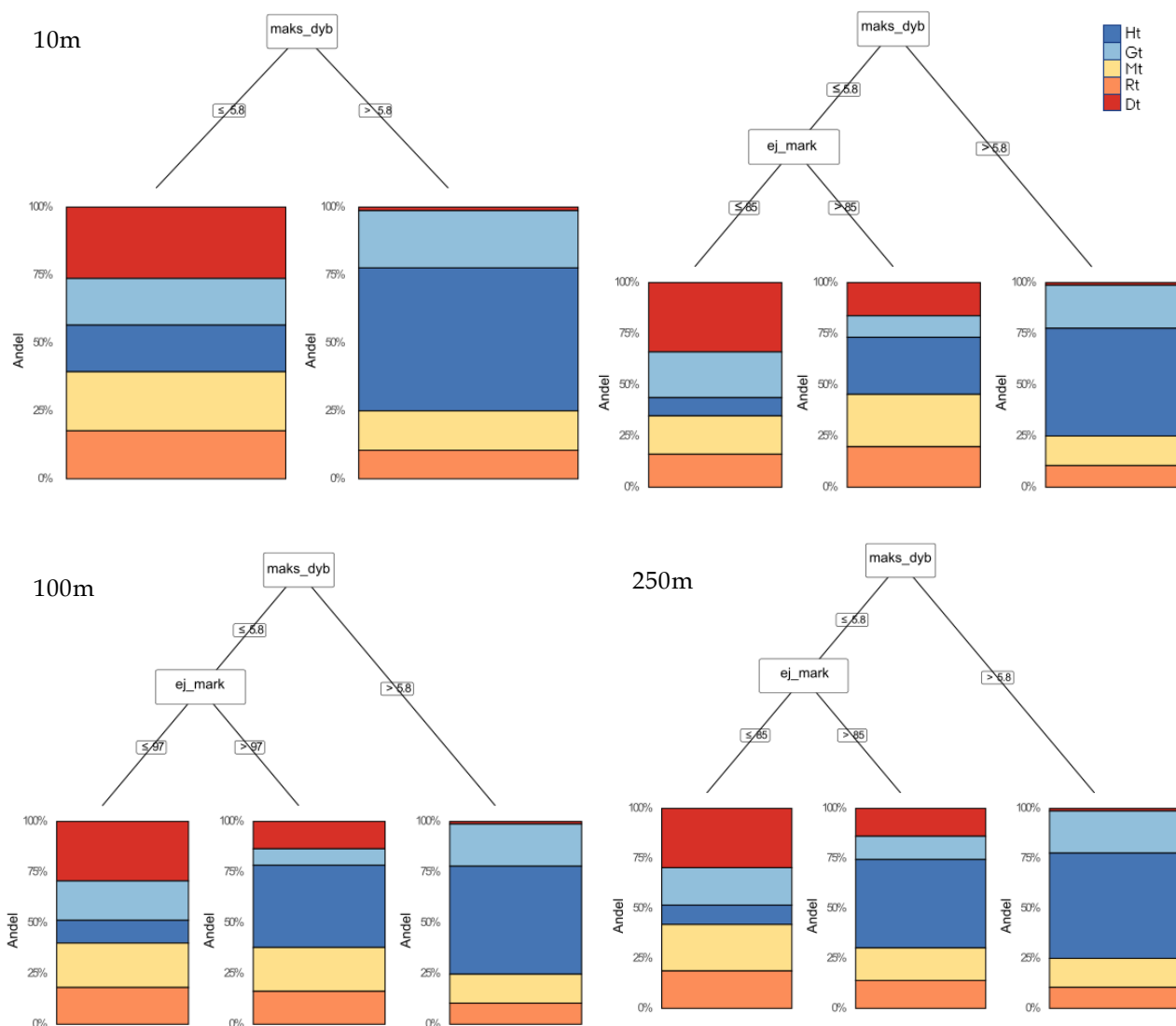
Figur 7.2. Resultater fra RDA-analysen for forskellige bufferafstande (10 m, 50 m, 100 m og 250 m) samt oplande. Blå pile angiver statistisk signifikante forklarende variable, som automatisk er udvalgt ved hjælp af baglæns selektion, mens de røde pile er responsvariable. Hvert punkt repræsenterer en sø, og farverne angiver klasser af næringsstofbegrænsning: N-begrænsning (grøn), P-begrænsning (blå) og N&P-begrænsning (rød).



Figur 7.3. Beslutningstræer, der forklarer sommerkoncentrationer af totalkvælstof (sommer TN) i søer for forskellige bufferafstande.



Figur 7.4. Beslutningstræ, der forklarer sommerkoncentrationer af totalfosfor (sommer TP) i søer for forskellige bufferafstande.



Figur 7.5. Beslutningstræ, der forklarer den økologiske tilstand i søer for forskellige bufferafstande.



SØER UDEN VELDEFINEREDE AFLØB – TOPOGRAFISK OPLAND OG INDSATSER

For mange af de danske søer, der indgår i Vandområdeplan 2021-2027 (VP3), er der ingen veldefinerede afløb. Dette umuliggør belastningsberegninger og planlægning af målrettede restaureringstiltag i søerne, selv om de ikke opfylder målet om mindst god økologisk tilstand. Ved hjælp af multivariate statistiske metoder (RDA) og maskinlæringsmetoder (Random Forest og Conditional Inference Trees) blev der i dette projekt undersøgt, hvordan landskabskarakteristika påvirker søers vandkvalitet, økologiske tilstand og typer af næringsstofbegrænsning. Alle metoder udviste lav forklaringsværdi, hvilket indikerer, at søernes tilstand kun delvist kan forklares af de undersøgte oplandskarakteristika og sandsynligvis også påvirkes af f.eks. intern næringsstofbelastning, historiske påvirkninger og hydrologiske processer. Ikke desto mindre var søer i lav økologisk tilstand konsekvent forbundet med jorde med højt organisk indhold, intensiv dræning og relativt våde oplande, mens større andele af ikke-dyrkede arealer og større sødybde var forbundet med bedre økologisk tilstand. Resultaterne peger på, at reduktion af kunstig dræning og beskyttelse af naturlige arealer kan bidrage til forbedringer i søers vandkvalitet og økologisk tilstand.