Kvantificering og geostatistisk simulering af usikkerhederne på den nationale hydrostratigrafiske model

Rasmus Bødker Madsen, Ingelise Møller, Frederik Falk & Anne-Sophie Høyer



DE NATIONALE GEOLOGISKE UNDERSØGELSER FOR DANMARK OG GRØNLAND KLIMA-, ENERGI- OG FORSYNINGSMINISTERIET

Kvantificering og geostatistisk simulering af usikkerhederne på den nationale hydrostratigrafiske model

Rasmus Bødker Madsen, Ingelise Møller, Frederik Falk & Anne-Sophie Høyer



DE NATIONALE GEOLOGISKE UNDERSØGELSER FOR DANMARK OG GRØNLAND KLIMA-, ENERGI- OG FORSYNINGSMINISTERIET

Indholdsfortegnelse

1.	Indledning og baggrund	5
1.1	Rammebetingelser for den hydrostratigrafiske model	5
2.	Nationalt koncept for håndtering af usikkerhed	12
2.1	Anvendelse af de hydrostratigrafiske modeller	12
2.2	Usikkerhedsestimering af boringer	13
2.3	Usikkerhedsestimering af geofysiske data	15
2.4	Usikkerhedsestimering relateret til modelleringsområder	20
2.5	Usikkerhedsestimering relateret til geologisk kompleksitet	23
2.6	Samlet usikkerhedsestimering	24
2.7	Eksempler på kvantificering af usikkerhed	30
3.	Geostatistiske simuleringer af usikkerhederne	44
3.1	Kalibreringsmodellen	46
3.2	Masker	49
3.3	Udtagning af punkter	53
3.4	Estimering af en ikke-stationær zoneret statistisk model	61
3.5	Geostatistisk modellering og generering af hydrostratigrafiske modeller	70
3.6	Algoritme	73
3.7	Resultater fra modelleringen	75
4.	Diskussion af modelresultater	80
4.1	Automatisering	80
4.2	Effektivitet	80
4.3	Konsistens og usikkerheder	81
4.4	Samlet vurdering	81
5.	Perspektiver for metoden (GDM-N)	82
5.1	Usikkerhedstemaer	82
5.2	Opsætning af simuleringsalgoritme	83
6.	Referenceliste	84

1. Indledning og baggrund

Denne rapport relaterer sig til arbejde udført i relation til kvælstofretentionskortet version 2025 (Højbjerg et al., 2025). Som et led i projektet blev der arbejdet med den nationale hydrostratigrafiske model (FOHM), hvor arbejdet især fokuserede på to centrale opgaver: 1) Lokale opdateringer af FOHM-modellen (Mortensen et al. 2025) og 2) Kvantificering og simulering af usikkerhederne på modellen (denne rapport).

Formålet med at kvantificere og simulere usikkerheden i FOHM-modellen var at inddrage usikkerheden af hydrostratigrafien i de samlede usikkerhedsberegninger af kvælstofretentionen. Målet er således at opnå et ensemble af hydrostratigrafiske modeller, der repræsenterer FOHM-modellen inklusiv de ny-modellerede lag (Mortensen et al., 2025), og samtidigt repræsenterer de kvantificerede usikkerheder lokalt i modellen.

Til at beskrive usikkerheden af hydrostratigrafien er der i projektet videreudviklet på en metode introduceret af Madsen et al. (2022). Den grundlæggende tilgang indebærer, at usikkerheden estimeres for hvert tolkningspunkt. På baggrund heraf genereres alternative lagflader, som varierer indenfor de estimerede usikkerheder. Resultatet er et sæt plausible alternative hydrostratigrafiske modeller, der alle ligger indenfor den observerede usikkerhed.

FOHM er baseret på nationale kortlægningsmodeller, som er opstillet over mange år af forskellige aktører. Dette har medført, at de originale tolkningspunkter og tilhørende usikkerhedsestimater ikke har været tilgængelige. Derfor er der i dette projekt udviklet nye metoder til at udtrække punkter fra eksisterende flader og efterfølgende tilskrive disse punkter usikkerheder. For at muliggøre simulering af lagfladerne er der desuden udviklet en metode til inddeling af modellen i ensartede geologiske domæner.

I denne rapport introduceres rammebetingelserne for FOHM-modellen i Kapitel 1, hvorefter arbejdet med at kvantificere usikkerhederne og simulere de hydrostratigrafiske modellag præsenteres i henholdsvis Kapitel 2 og 3. Afslutningsvis diskuteres resultaterne i Kapitel 4 og den udviklede simuleringsmetode perspektiveres i Kapitel 5.

1.1 Rammebetingelser for den hydrostratigrafiske model

Som led i den nationale grundvandskortlægning er der løbende, i takt med kortlægningen af nye områder, blevet opbygget geologiske og/eller hydrostratigrafiske modeller på baggrund af eksisterende og ny-indsamlede geofysiske data samt boringsdata. Det har ført til en række hydrostratigrafiske modeller af forskellig alder.

I 2017 blev de hydrostratigrafiske modeller på Fyn, Tåsinge, Langeland og Ærø samlet til én model (herefter kaldet Fyn) og året efter i 2018, blev alle grundvandskortlægningsmodeller i Jylland og på Als samlet i FOHM-projektet. Efterfølgende er FOHM-modellen i Jylland løbende blevet opdateret med nye modeller. Grundvandskortlægningsmodeller på Sjælland, Lolland, Falster, Møn og mindre øer i Smålandsfarvandet og Storebælt er løbende blevet inkorporeret i DK-modellens hydrostratigrafi for Sjællandsmodellen. De nyeste opdateringer

som er medtaget i nærværende arbejde, er udført inden starten af april 2023 for Jylland og Sjælland. Derudover er der lavet lokale opdateringer i relation til kvælstofretentionskortet version 2025, der er nærmere beskrevet i Mortensen et al., (2025). De lokale opdateringer inkluderer blandt andet opdatering af lavbundsområder samt to mindre områder, hvor der er udført ny hydrostratigrafisk modellering efter tTEM-kortlægninger (Østsalling og LOOP-2 i Himmerland) samt hydrostratigrafiske modeller for Nordsamsø, Tunø, Endelave, Anholt og Læsø.

1.1.1 Den hydrostratigrafiske model for Fyn

Eftersom det meste af Fyn er udpeget som OSD-områder (Områder af Særlige Drikkevandsinteresser), er der udført national grundvandskortlægning over det meste af Fyn. De ældste modeller er fra 2004 (Kerteminde) og 2005 (Assens syd), hvorefter hovedparten af de resterende områder af Fyn er modelleret mellem 2011 og 2015 (Figur 1-1). I 2017 blev de hydrostratigrafiske modeller fra Fyn samlet i én model. Orbicon (2018) beskriver i detaljer, hvilke modeller der er anvendt, hvordan lag fra de forskellige lag er korreleret, og hvordan der er prioriteret mellem de enkelte modeller. I arbejdet med at sammenstille alle modellerne blev enkelte af de eksisterende hydrostratigrafiske modeller eller enkelte lag i en model opdateret og redigeret. Ligeledes blev områder mellem eksisterende modeller og i modellers randområder justeret for at få modellag til at blive sammenhængende. Den hydrostratigrafiske model består af fire kvartære sandlag med lerlag over og under, samt et prækvartært lerlag og kalk, hvilket i alt giver 8 lag udover toppen ved terrænoverfladen (topo) og bunden af modellen (bund) placeret 50 m under Kalk-lagfladen, se Tabel 1-1.

	<u> </u>	, , , , , , , , , , , , , , , , , , ,	r	
Overordnet	DK-model lagflade	DK-model lag	Litologi	Hydrogeo-
gruppe	betegnelse*	betegnelse		logi
	Торо			
Kvartær	ks1t	KL1	Ler og silt	Akvitard
Kvartær	ks1b	KS1	Sand og grus	Akvifer
Kvartær	ks2t	KL2	Ler og silt	Akvitard
Kvartær	ks2b	KS2	Sand og grus	Akvifer
Kvartær	ks3t	KL3	Ler og silt	Akvitard
Kvartær	ks3b	KS3	Sand og grus	Akvifer
Kvartær	preq	KL4	Ler og silt	Akvitard
Prækvartær	kalk	PL7	Ler	Akvitard
Kalk	bund	Kalk	Kalk	Akvifer

Tabel 1-1: Modellag i den hydrostratigrafiske model for Fyn.

*Lagflade knyttet til bund af lag i samme række



Figur 1-1: Oprindelige hydrostratigrafiske modeller, som indgår i den samlede Fyn model. Modelleringsåret fremgår i legenden.

1.1.2 Den hydrostratigrafiske model for Sjælland

Den hydrostratigrafiske model for Sjælland og øerne tager udgangspunkt i en hydrostratigrafisk model fra DK-model 2014 og efterfølgende opdateringer med grundvandskortlægningsmodeller og andre modeller frem til marts 2023 (Stisen et al. 2020). Der er sket en opdatering med lokale og regionale modeller fra ca. 50 områder (Figur 1-2). Den hydrostratigrafiske model består af fem kvartære sandlag med lerlag over og under, samt et prækvartært lerlag og kalklag inddelt i 3 enheder, hvilket i alt giver 11 lagflader udover toppen ved terrænoverfladen (topo) og bunden af modellen (bund) placeret 50 m under bunden af det Palæogene ler (pl1b-lagfladen), se Tabel 1-2.

Overordnet	DK-model lagflade	DK-model lag	Litologi	Hydrogeo-
gruppe	betegnelse*	betegnelse		logi
	Торо			
Kvartær	ks1t	KL1	Ler og silt	Akvitard
Kvartær	ks1b	KS1	Sand og grus	Akvifer
Kvartær	ks2t	KL2	Ler og silt	Akvitard
Kvartær	ks2b	KS2	Sand og grus	Akvifer
Kvartær	ks3t	KL3	Ler og silt	Akvitard
Kvartær	ks3b	KS3	Sand og grus	Akvifer
Kvartær	ks4t	KL4	Ler og silt	Akvitard

Tabel 1-2: Modellag i den hydrostratigrafiske model for Sjælland.

Kvartær	ks4b	KS4	Sand og grus	Akvifer
Kvartær	preq	KL5	Ler og silt	Akvitard
Prækvartær	pl1b	PL1	Ler	Akvitard
Kalk	gk1b	GK	Kalk	Akvifer
Kalk	dk1b	BK	Kalk	Akvifer
Kalk	Bund**	SK	Kalk	Akvifer

*Lagflade knyttet til bund af lag i samme række. **Bund er sat til 50 m under pl1b.



Figur 1-2: Oprindelige hydrostratigrafiske modeller, som indgår i den samlede hydrostratigrafiske model for Sjælland og øerne omkring Sjælland. Farveskalaen er tilfældig og markerer udstrækningen af de enkelte modeller.

1.1.3 Den hydrostratigrafiske model for Jylland

Omkring 90 hydrostratigrafiske modeller fra Jylland er samlet i FOHM-modellen, hvoraf de ældste er fra 2004 og nyeste fra 2022 (Figur 1-3). Ved samling af den hydrostratigrafiske model for Jylland ud fra modeller fra Grundvandskortlægningen i 2018, er der opstillet en samlet hydrostratigrafisk model for Jylland: FOHM-modellen (Miljøstyrelsen 2020). Den består af syv kvartære sandlag med lerlag over og under, 14 Prækvartære sandlag med lerlag over og under samt to kalklag. Samlet svarer det til 45 lagfalder, se Tabel 1-3. I områder med højtliggende stensalt er toppen af dette modelleret. De Prækvartære lag følger stort set den miocæne stratigrafi, opstillet i Kristensen et al. (2015).

Overordnet FOHM-model lagflade betegnelse gruppe		Litologi	Hydrogeologi	
	Торо			
Kvartær	0100_Postglacial_toerv_Bund Ler, silt og tørv Akvitard		Akvitard	
Kvartær	0200_Kvartaer_sand_Bund Sand og grus Akvifer		Akvifer	
Kvartær	0300_Kvartaer_ler_Bund	Ler og silt	Akvitard	
Kvartær	0400_Kvartaer_sand_Bund	Sand og grus	Akvifer	
Kvartær	1100_Kvartaer_ler_Bund	Ler og silt	Akvitard	
Kvartær	1200_Kvartaer_sand_Bund	Sand og grus	Akvifer	
Kvartær	1300_Kvartaer_ler_Bund	Ler og silt	Akvitard	
Kvartær	1400_Kvartaer_sand_Bund	Sand og grus	Akvifer	
Kvartær	1500_Kvartaer_ler_Bund	Ler og silt	Akvitard	
Kvartær	2100_Kvartaer_sand_Bund	Sand og grus	Akvifer	
Kvartær	2200_Kvartaer_ler_Bund	Ler og silt	Akvitard	
Kvartær	2300_Kvartaer_sand_Bund	Sand og grus	Akvifer	
Kvartær	2400_Kvartaer_ler_Bund	Ler og silt	Akvitard	
	5100_Maadegrup-	Ler og silt	Akvitard	
Miocæn	pen_Gram_og_Hodde_Bund			
Miocæn	5200_Oevre_Odderup_ODS3_Bund	Sand og grus	Akvifer	
Miocæn	5300_Oevre_Arnum_ARL3_Bund	Ler og silt	Akvitard	
Miocæn	5400_Nedre_Odderup_ODS2_Bund	Sand og grus	Akvifer	
Miocæn	5500_Nedre_Arnum_ARL2_Bund	Ler og silt	Akvitard	
Miocæn	5600_Bastrup_BADS6_Bund	Sand og grus	s Akvifer	
Miocæn	5700_Klintinghoved_KRL6_Bund	Ler og silt	Akvitard	
Miocæn	5800_Bastrup_BADS5_Bund	I Sand og grus Akvifer		
Miocæn	5900_Klintinghoved_KRL5_Bund Ler og silt Akvitard		Akvitard	
Miocæn	6000_Bastrup_BADS4_Bund	Sand og grus	Akvifer	
Miocæn	6100_Klintinghoved_KRL4_Bund Ler og silt Akvitard		Akvitard	
Miocæn	6200_Bastrup_BADS3_Bund	Sand og grus	Akvifer	
Miocæn	6300_Klintinghoved_KRL3_Bund Ler og silt Akvitard		Akvitard	
Miocæn	6400_Bastrup_BADS2_Bund Sand og grus Akvifer		Akvifer	
Miocæn	6500_Klintinghoved_KRL2_Bund	Ler og silt	r og silt Akvitard	
Miocæn	6600_Bastrup_BADS1_Bund	Sand og grus	Akvifer	
	6700_Klintingho-	Ler og silt	Akvitard	
Miocæn ved_KRL1_Vejle_Fjord_Bund				
Miocæn	6800_Billund_BDS6_BDS9_Bund Sand og grus Akvifer		Akvifer	
Miocæn	6900_Vejle_Fjord_VFL6_Bund Ler og silt Akvitard		Akvitard	
Miocæn	7000_Billund_BDS4_BDS5_Bund Sand og grus Akvifer		Akvifer	
Miocæn	7100_Vejle_Fjord_VFL4_Bund Ler og silt Akvitard		Akvitard	

Tabel 1-3: Modellag i den hydrostratigrafiske FOHM-model for Jylland.

Miocæn	7200_Billund_BDS3_Bund Sand og grus Akvifer		Akvifer
Miocæn	7300_Vejle_Fjord_VFL3_Bund	Ler og silt	Akvitard
Miocæn	7400_Billund_BDS2_Bund	Sand og grus	Akvifer
Miocæn	7500_Vejle_Fjord_VFL2_Bund	Ler og silt	Akvitard
Miocæn	7600_Billund_BDS1_Bund	Sand og grus	Akvifer
Miocæn	7700_Vejle_Fjord_VFL1_Bund Ler og silt Akvitard		Akvitard
Miocæn	7800_Billund_BDS0_Bund	Sand og grus	Akvifer
Prækvartær	8000_Palaeogen_ler_Bund	Ler	Akvitard
Kalk	8500_Danien_Kalk_Bund	Kalk	Akvifer
Kalk	9000_Skrivekridt_Bund Kalk		Akvifer
Stensalt	9500_Stensalt_Top	Salt	Akvitard



Figur 1-3: Oprindelige modeller fra Jylland samlet i FOHM modellen. Farveskalaen er tilfældig og markerer udstrækningen af de enkelte modeller. Den grå baggrund indeholder kyst- og andre støttepunkter samt punkter indlæst i 2021, hvor i blandt der kan være nye modeller fra 2021, som ikke har fået en label om oprindelse.

2. Nationalt koncept for håndtering af usikkerhed

De hydrostratigrafiske modeller for Jylland, Fyn og Sjælland består som beskrevet i ovenstående afsnit af samlinger af lokale til regionale geologiske og/eller hydrostratigrafiske modeller. En del af disse modeller fik evalueret usikkerheder under modelleringen ud fra flere forskellige tilgange. Eftersom de originale modeller har forskellige aldre, har været lavet i forskelligt software, og er samlet ved brug af forskellige procedurer og arbejdsgange har det ikke været muligt på national skala at tilgå og identificere de oprindelige usikkerhedsevalueringer, hvis de var udført.

Der er derfor udviklet et koncept til vurdering af usikkerheder i modellerne, som er baseret på de boringsdata (Afsnit 2.2) og geofysiske data (Afsnit 2.3), som har været til rådighed ved modelleringen samt en senere vurdering af de originale modellers kvalitet. Da det er vanskeligt og for stort et arbejde tidsmæssigt at identificere præcist hvilke data, der er anvendt i modelleringen, vurderes dette ved at anvende information om boringsdata rapporteret til Jupiter databasen og geofysik rapporteret til GERDA databasen. Det tjekkes således om data var tilgængelige på modelleringstidspunktet, og at de havde tilstrækkelig indtrængning til at informere det aktuelle hydrostratigrafiske lag i modellen. Alle data tilskrives en usikkerhed afhængig af datatype, som beskrevet i afsnit 2.2 og 2.3.

De enkelte datapunkters usikkerhed udbredes til alle celler i modellen og hvorledes dette gøres afhænger af typen af modelleringsområde og den geologiske kompleksitet for området som beskrevet i hhv. Afsnit 2.4 og 2.5. I Afsnit 2.6 beskrives, hvorledes alle usikkerhedsestimater lægges sammen til det samlede usikkerhedsestimat og dette illustreres med eksempler på den samlede kvantificerede usikkerhed i Afsnit 2.7.

2.1 Anvendelse af de hydrostratigrafiske modeller

For hver af de tre modelområder er alle tolkningspunkter blevet samlet i fælles tolkningspunktsdatabaser i Geoscene3D format, hvor der er en tabel for hvert modellag. Fra tolkningspunktsdatabasernes tabeller anvendes information om position (X, Y), dybde (Z), SnapID (vedrørende snap til boringer), forskellige felter som kan identificere de oprindelige modeller og evt. deres alder. For Fynsmodellen er de originale modellers alder taget fra modelrapportens beskrivelser af hver enkelt model (Orbicon 2018).

I forbindelse med sammenstilling af modeller hhv. på Fyn og i Jylland er der foretaget en kvalitetsvurdering og prioritering af de oprindelige modeller. Da der for Fyns modellen i 2017 blev lavet en detaljeret re-modellering af områder, hvor de oprindelige modeller ikke kunne anvendes, er det vurderet at hele det fynske modelområde har samme kvalitet (Orbicon 2018). For det jyske modelområde er de oprindelige modeller opdelt i tre kvaliteter i forbindelse med sammenstillingen (Miljøstyrelsen 2020). Denne kategorisering er udtrukket og anvendt i dette arbejde. Da der ikke er gemt en sammenstilling af de oprindelige modellers kvalitet for det sjællandske område, er alle oprindelige modeller tilskrevet samme modelkvalitet.

Ud fra tolkningspunkterne og deres oplysninger om modellernes alder laves et modelleringsaldergrid med de samme specifikationer som griddene i de respektive modeller. Hvis der er overlap mellem tolkningspunkter fra flere oprindelige modeller indenfor en gridcelle vælges den yngste alder. Dette modelleringsaldersgrid anvendes i bearbejdningen af boringsdata og geofysiske data.

2.2 Usikkerhedsestimering af boringer

Boringsdatasættet er baseret på et udtræk fra Jupiter fra november 2022. For dette boringsudtræk er der udført en kvalitetsvurdering med hensyn til litologi ved brug af MSTs procedure for dette (Damsgaard & Makwarth 2020). Usikkerhedsestimering af boringer tager udgangspunkt i en kvalitetsvurdering af boringer og en kvantitativ vurdering af usikkerhed på bestemmelse af laggrænser i relation til udførsel af boringer og tilhørende prøvetagning. Til analysen af boringsusikkerheder indgår kun kvalitetsvurderingen mht. litologi og der ses bort fra kvalitetsvurderingen af boringspositionen, da usikkerhedsvurderinger vil blive anvendt på grids med 100x100 m cellestørrelser. Boringer med en kvalitetsvurderingsscore på mindst 3 på MST's score, der går fra 0 (dårligst) til 7 (bedst) indgår i usikkerhedsvurderingen. MST's anvendte 5 scores klassificeres om til 4 kategorier ('Nret scores'), hvor 1 er bedst og 4 er dårligst. Tabel 2-1 beskriver sammenhængen mellem MST' kvalitetsvurderingsscores og Nret scores, samt kriterierne for de forskellige kvalitetsklasser.

MST borings-	Nret borings-	Beskrivelse af kriterier til forespørgsel
kvalitetskode kvalitetskode		
3	4	Der findes en litologisk beskrivelse, ikke DAPCO boring
4	3	+ Litologisk beskrivelse fra GEUS jordprøvelab
5	3	+ boringsdybde registreret
6	2	+ mindst 1 prøve per 2 m (men anvendt omvendt, så
		højest 1 prøve per 2 m)
7	1	+ boremetode er snegleboring (men i anvendte skript
		indgår også, tør- eller slagboring, omvendt skylleboring,
		lufthæveboring og rotaryboring)

Tabel 2-1: Beskrivelse af boringskvalitetsvurdering fra MST med kvalitetskoder

Efter arbejdet er afsluttet er der desværre fundet fejl i MST's SQL-forespørgsel, som udfører kvalitetsvurderingen. Det betyder, at alle boringer, som får MST boringskvalitetskode 6 eller 7 ikke er de boringer, som er prøvetaget med den højeste frekvens og med *mindst* 1 prøve pr. 2 m, men omvendt er det dem, som *højest* er prøvetaget med 1 prøve pr. 2 m. Desuden beskriver den tekstlige dokumentation, at snegleboringer med mindst 1 prøve pr. 2 m bliver tildelt den bedste MST boringskvalitetskode, men når der ses på hvilke boremetodekoder, som er medtaget, inkluderer dette også tør- og slagboringer, omvendt skylleboring, lufthæveboring og rotaryboring (direkte skylleboring). Disse fejl i SQL-forespørgslen betyder desværre, at de boringer, der burde have den bedste kvalitet, kun kommer i den tredje bedste kvalitetsklasse.

Ud over boringskvalitetsvurderinger flettes information om modellens alder samt dybden af hvert af den hydrostratigrafiske models lagflader på boringspositionen med hver boring i

Jupiter. Ligeledes tilføjes information om, hvorvidt et tolkningspunkt er snappet til boringen. Dette aflæses enten; 1) direkte i tolkningspunktsdatabasen eller 2) ud fra en antagelse om, at hvis der ligger et tolkningspunkt i samme position (indenfor 5 m) og dybde (dybde af tolkningspunkt er mindre end boringens dybde), som der er boringsinformation, så har det været snappet til boringen. Boringer snappet til tolkningspunkter er tilskrevet mere sikkerhed end boringer, der ikke er snappt til tolkningspunkter. Dette er gjort, da det antages, at modelløren har vægtet disse boringer højere og direkte brugt en laggrænse i boringen i bestemmelsen af et lags dybde. Information om snappede boringer anvendes også i forberedelsen af den geostatiske simulering, se Afsnit 3.3.

Ud fra oplysninger om boringsposition, boringskvalitet og lagdybde beregnes boringsusikkerhed ud fra følgende formel for variansen i den dybde, hvor et given modellag skærer boringen:

 $var_{0, boringer}(z) = K^2 + (b^*z)^2$

hvor z er dybden til en given modelflade i boringspositionen, b er en procentvis stigende standardafvigelse og K er en minimumsstandardafvigelse. Til usikkerhedsestimering er anvendt b = $0.5*[0.02\ 0.03\ 0.07\ 0.1]$ og K = $[0.15\ 1\ 2\ 2.5]$ m for snappede boringer, b = $0.5*[0.025\ 0.035\ 0.075\ 0.105]$ og K = $[0.4\ 1.5\ 2.5\ 3]$ m for andre boringer, og elementerne i b og K relaterer sig til GEUS boringskvalitet 1, 2, 3 og 4 i respektive rækkefølger. Boringer vil kun bidrage til det samlede usikkerhedsestimat, hvis de er dybere end den givne modellagflade i boringspunkterne.

Som illustreret i Figur 2-1, angiver b hvor meget standardafvigelsen øges med dybden og K er en minimumstandardafvigelse. Hvis man tager boringskvalitet kategori 2 for en boring snappet til et tolkningspunkt som eksempel, så vil usikkerheden ved terrænoverfladen starte med at have en standardafvigelse på 1 m, hvilket er afstemt efter at prøvetagningsafstanden for denne kategori maksimalt må være 2 m. For boringskvalitet kategori 1, der henviser til snegleboringer, antages at laggrænser kan ses direkte på sneglen, så her sættes K til 0.15 cm for en boring snappet til et tolkningspunkt. For boringskategori 3 og 4, som hhv. er grovere prøvetaget og indeholder mindre information om litologien, antages laggrænser at være mere usikkert bestemt og K sættes til henholdsvis 2 m og 2,5 m. Der er valgt en procentvis stigning med dybden, da skifte eller forlængelse af borestænger kan medføre unøjagtigheder i bestemmelsen af dybden. Ved boremetoder, hvor prøvemateriale transporteres med strømmen af boremudder, giver dette også en større usikkerhed jo større dybde, prøvematerialet transporteres op fra. For boringskategori 1 antages en 1 % stigning med dybden, hvilket svarer til, at man i 100 m dybde i mere end 68 % af tilfældene vil have bestemt prøvedybden til at være 100 m ± 1 m under antagelse om, at fejlen på dybdebestemmelsen kan beskrives ved en normalfordeling (Ditlefsen et al. 2008). Som for K antages, at den procentvise stigning med dybden, b, er større for de boringskategorier med mindre sikker litologisk information.



Figur 2-1: Boringsusikkerhed angivet som standardafvigelse (til venstre) og varians (til højre) for de fire boringskvalitetskategorier (Tabel 2-1). Fuldt optrukket linje er for boringer, som er snappet og stiplet linje for andre boringer.

2.3 Usikkerhedsestimering af geofysiske data

På Fyn er det meste af arealet dækket med geofysiske målinger. SkyTEM dækker store områder og ældre jordbaseret TEM dækker nogle mindre arealer. I langt de fleste områder er SkyTEM og TEM supplereret med PACES- eller PACEP-målinger. Stedvis er der målt MEP. Refleksionsseismik og borehulslogs er ikke medtaget i usikkerhedsanalysen. Disse to datatyper, blev af tidsmæssige årsager ikke prioriteret i første omgang, da der er tæt dækning af SkyTEM og boringer, hvor der hhv. er refleksionsseismiske linjer og borehulslogs. Der indgår ikke tTEM, da alle tTEM kortlægninger er nyere end 2018 og derfor ikke kan være anvendt ved modellering af de hydrostratigrafiske lag på Fyn.

For Sjælland er datadækningen med geofysiske data ikke helt så høj, som for Fyn og MEP udgør en større andel af datadækningen. Der er udover SKYTEM, TEM, MEP, PACES og PACEP målinger medtaget refleksionsseismik og borehulslogs, da disse to datatyper supplerer den geofysiske datadækning. Der indgår ikke tTEM, da kortlægninger med tTEM er for nye til at have indgået i den hydrostratigrafiske modelleringen på Sjælland.

I Jylland er der en varierende datadækning med geofysiske metoder, hvor der er større områder, hvor ældre jordbaseret TEM eller MEP udgør datagrundlaget. For Jylland er alle datatyper medtaget. Der er desuden en række nyere mindre modelleringsområder, hvor der indgår tTEM.

Det første skridt i præpareringen af geofysiske data til usikkerhedsvurderingen er at få et overblik over hvilke tolkningsmodeller, der skal anvendes. Det gøres via forespørgsler i GERDA databasen. Hvis TEM/SkyTEM data er tolket med mangelagsmodeller vælges disse.

Det er tilfældet for alle SkyTEM målinger på Fyn, mens der på Sjælland og i Jylland er ældre SkyTEM kortlægninger, som kun er tolket med fålagsmodeller. Alle jordbaserede TEM målinger er tolket med fålagsmodeller på Fyn, mens der er enkelte kortlægninger med jordbaseret TEM, der er tolket med mangelagsmodeller på Sjælland og i Jylland. PACES, PACEP og MEP-data er typisk tolket med flere modeller med forskellige antal lag. Hvis der er udført en tolkning med en mangelagsmodel, er denne valgt. For Sjælland og Jylland anvendes refleksionsseismik kun til at informere om de prækvartære lag, da det kan være meget svært at tolke interne strukturer i den kvartære lagserie.

På samme vis, som for boringer, flettes et geofysisk datapunkt med en modelleringsalder og lagdybder fra den hydrostratigrafiske model fra den gridcelle, hvori det geofysiske datapunkt er beliggende.

2.3.1 SkyTEM, TEM og tTEM

Mangelagsmodeller diskretiseres typisk sådan, at lagtykkelsen stiger med dybden. Når der udføres geologisk tolkning med støtte fra mangelagsmodellerne kan et tolkningspunkt ikke sættes med større nøjagtighed end lagtykkelsen i resistivitetsmodellen i den pågældende dybde. Derfor har vi valgt at lade resistivitetsmodellens lagtykkelse indgå i udtrykket for Sky-TEM og tTEM datas usikkerhed. Variansen for SkyTEM og tTEM, tolket med mangelagsmodeller er således givet ved:

 $var_{0, SkyTEM}(z) = (0.5 \cdot 1.2 \cdot Resistivitets_model_tykkelse(z))^2 og$

 $var_{0, tTEM}(z) = (0.5 \cdot 1.2 \cdot \text{Resistivitets}_model_tykkelse(z))^2$,

hvor Resistivitets_model_tykkelse er resistivitetsmodellens lagtykkelse i den dybde, z, hvor variansen beregnes.

TEM og SkyTEM fålagsmodeller har stor variation i lagtykkelsen, så det er ikke meningsfuldt at anvende samme princip som for mangelagsmodeller. Derimod anvendes samme princip, som for boringer, så variansen har en konstant værdi tæt på terræn og en procentvis stigning med dybden. Den procentvise stigning er valgt, så den følger standardafvigelsen fra den parameterisering af SkyTEM mangelagsmodellerne, som har de tykkeste resistivitetslag. Variansen for jordbaseret TEM tolket med fålagsmodeller er givet ved

 $var_{0,TEM}(z) = (max(1.2, 0.5 \cdot 0.15 \cdot z) + 2)^2$,

hvor z er dybden til det punkt på en modellagflade, hvor variansen beregnes.

I de øverste meter af jorden lige under terrænoverfladen integreres jordlagenes resistivitet til et bidrag i første gatetid. Der er derfor kun meget lille information fra den helt terrænære del fra data indsamlet med TEM systemer. Det vil afhænge af opsætningen af de forskellige TEM-systemer og resistivitetsforholdene i jorden, hvor dybt denne zone går. Vi har for Sky-TEM valgt at definere denne dybde som de øverste 5 m under terræn, for tTEM 2 m og jordbaseret TEM 10 m, hvor variansen sættes til en meget stor værdi; matematisk formuleret som

 $var_{0, SkyTEM}(z)$ for (z<5) = 100000, $var_{0, tTEM}(z)$ for (z<2) = 100000 og $var_{0, TEM}(z)$ for (z<10) = 10000,

hvor z er dybden til det punkt på en modellagflade, hvor variansen beregnes.

SkyTEM og tTEM tolket med mangelagsmodeller og TEM tolket med fålagsmodeller kan kun bidrage til det samlede usikkerhedsestimat for en given lagflade, hvis datapunktet har en lower DOI (Christiansen & Auken 2012), som ligger dybere end det givne modellag. Da der er en del SkyTEM data uden beregnet DOI i GERDA-databasen, er der lavet et estimat af DOI, hvor der tages hensyn til, at der vil være en begrænset indtrængning i meget konduktive jordlag.

Figur 2-2 viser, hvordan standardafvigelse og varians varierer med dybden for 3 SkyTEM systemer tolket med mangelagsmodeller, tTEM mangelagsmodeller og jordbaseret TEM tolket med fålagsmodeller. De tre kurver for standardafvigelser og varians for SkyTEM-systemer baserer sig på mangelagsmodeller fra Fyn og viser de kurver, som udspænder rummet for mulige dybdevariationer for standardafvigelses- og variansestimater.



Figur 2-2: Standardafvigelse (venstre) og varians (højre) for SkyTEM (mangelagsmodeller), TEM (fålagsmodeller) og tTEM (mangelagsmodeller).

2.3.2 Geoelektriske metoder

MEP-data er opdelt i ældre data målt i Wenneropstilling og nyere data målt overvejende med gradientkonfigurationer eller Schlumbergerkonfigurationer, hvor der er en større vertikal datadækning. Derfor tildeles MEP data i gradient- og Schlumbergerkonfigurationer en lavere varians end MEP data i Wennerkonfiguration. Variansen for de to typer af MEP-målinger er henholdsvis

 $var_{0, MEP-gradient}(z) = (0.5 \cdot max(1.25, 0.15 \cdot z))^2 \text{ og}$

 $var_{0, MEP-Wenner}(z) = (0.5 \cdot max(2.4, 0.25 \cdot z))^{2}$

hvor z er dybden til det punkt på en modellagflade, hvor variansen beregnes. Enten anvendes lower DOI eller en konstant værdi på 50 m som den maksimale dybde, hvortil MEP-data kan bidrage til det samlede usikkerhedsestimat.

Ældre PACEP-målinger er typisk tolket med faste laggrænser i en 3-lagsmodel, hvilket gør disse data til de mest usikre af de geoelektiske metoder. PACES data kan både være tolket med 3-lagsmodeller med varierede laggrænser eller mangelagsmodeller med stigende lagtykkelser for stigende dybde. Den procentvise stigning i standardafvigelse sættes til det samme som for MEP i Wenneropstillinger. Variansen for PACES og PACEP-målinger er henholdsvis

 $var_{0,PACES}(z) = (0.5 \cdot max(2, 0.25 \cdot z))^2 \text{ og}$

 $var_{0,PACEP}(z) = (0.5 \cdot max(5, 0.5 \cdot z))^2$,

hvor z er dybden til det punkt på en modellagflade, hvor variansen beregnes. PACES- og PACEP-data kan bidrage til det samlede usikkerhedsestimat til henholdsvis 18 m og 15 m dybde.

Figur 2-3 viser, hvorledes standardafvigelse og varians ændrer sig som funktion af dybden i målepunktet for de geoelektriske metoder. Vi har valgt at den procentvise stigning for nyere MEP i gradientopstilling har samme hældning som TEM tolket med fålagsmodeller, mens den procentvise stigning for PACES og ældre MEP i Wenneropstilling sættes til det samme og med en lidt højere procentvis stigning, da der er færre datapunkter til at beskrive dybdevariationen. Alle standardvariansfunktionerne starter med en konstant værdi som afspejler, at det øverste lag i resistivitetsmodellen har en vis tykkelse.



Figur 2-3: Standardafvigelses- (venstre) og variansestimat (højre) for de geoelektriske metoder; MEP, PACES og PACEP.

2.3.3 Refleksionsseismik

Refleksionsseismik bliver kun brugt til at informere de prækvartære lag i de hydrostratigrafiske modeller, da seismiske data primært er anvendt i modelleringen af disse lag. Da opløsningen af refleksioner af laggrænser ikke ændrer sig væsentlig med dybden indenfor de dybdeintervaller som modellagene ligger i, anvendes en konstant standard afvigelse på 7,5 m. Variansen formuleres derfor således

 $var_0, SEISMIK(z) = (0.5.15))^2,$

hvor z er dybden til det punkt på en modellagflade, hvor variansen beregnes.

2.3.4 Borehulslogs – gamma- og resistivitetslog

Da naturlig gammalog og resistivitetslog er de to mest udbredte borehulsloggingmetoder og typisk disse som er anvendt i geologisk modellering, er det kun disse to logtyper, som er medtaget her. Da opløsningen af laggrænser ikke ændrer sig med dybden, er det mest meningsfuldt at formulere usikkerheden som konstant med dybden. Det er også sådan usikkerheden er formuleret for gammalog med en konstant standard afvigelse på 1,5 m, mens usikkerheden for resistivitetslog er formuleret som boringer svarende til noget som ligger mellem GEUS boringskategori 3 og 4. Matematisk er det formuleret for gammalog som

 $var_{0, GAMMA}(z) = (0.5 \cdot 3))^{2}$

og for resistivitetslog som

 $var_{0, RESISTIVITET}(z) = 0.25 + (0.5 \cdot 0.09 \cdot z))^{2}$

hvor z er dybden til det punkt på en modellagflade, hvor variansen beregnes.

2.4 Usikkerhedsestimering relateret til modelleringsområder

Modelleringsområder bruges som en basis baggrundsusikkerhed. En gængs modelleringsprocedure ved hydrostratigrafisk modellering af et område er, at man modellerer et centralt område omkranset af en bufferzone. Ud fra tolkningspunkter fra hvert af de originale modeller foretages en beregning af hvilke områder, der hører til den centrale del af en model, hvilke dele som hører til randzonen af modellen og hvilke dele, som ligger uden for et originalt modelleringsområde. Figur 2-4 viser fordelingen af de forskellige typer af modelleringsområder på Fyn og Sjælland.

Ved samlingen af alle grundvandskortlægningsmodellerne i Jylland i FOHM-projektet blev der lavet en kvalitetsvurdering af hver enkelt model fordelt i klasserne god, mellem og dårlig sammen med en prioritering af modellerne (Figur 2-5). Denne kvalitetsvurdering bruges i en opdeling af de centrale områder i tre klasser. Figur 2-5 viser MST' kvalitetsvurdering af modellerne i Jylland inden den anvendes i beregningerne af usikkerheden for modelleringsom-råderne.



Figur 2-4: Modelleringsområder på Fyn og Sjælland ud fra originalmodeller.



Figur 2-5: Oprindelige modelområder i Jylland vist med kvalitetsvurdering udført i forbindelse med samlingen af modellerne i 2018 (Miljøstyrelsen, 2020). Nye modelområder som er tilføjet efter 2018 er til dette arbejde tildelt en god kvalitet.

Vi har foretaget opdeling i forskellige typer af modelleringsområder, da vi antager, at de centrale dele af en model må være mindre usikker end et randområde eller områder, som ligger uden for de originale modelleringsområder. En variansfunktion er sat op for de tre typer af modelleringsområder med forskellige minimumsvarians ved terrænoverfladen

$$Var_{m}(z) = (0.5 \cdot ((z < 50) \cdot (0.3 \cdot z + T_{std}) + (z \ge 50) \cdot (0.15 \cdot z + 50 \cdot (0.3 - 0.15) + T_{std})))^{2},$$

hvor z er dybden til det punkt på en modellagflade hvor variansen beregnes og T_{std} er en minimumsstandardafvigelse ved terrænoverfladen på 10 m, 12.5 m og 20 m henholdsvis inden for, i randen og uden for modelleringsområdet for Fyn og Sjælland, mens T_{std} for Jylland er hhv. 10, 11,25 og 12,5 m for modelkvalitet god, middel og dårlig inden for modelleringsområdet og T_{std} er hhv. 12,5 og 20 m i randen og uden for modelleringsområdet.

Figur 2-6 viser, hvorledes standardafgivelse og varians aftager med dybden for de tre typer af modelleringsområde. Den stejlere procentvise stigning i standardafvigelse i de øverste 50 m er valgt for at indhylle variansfunktionen for de mest usikre geoelektriske datatyper, mens den reducerede procentvise stigning i dybder større end 50 m svarer til den procentvise stigning i standardafvigelse for TEM-fålagsmodellerne og de mest usikre SkyTEM-mangelagsmodeller. Usikkerhedsestimatet for Jyllands middel modelkvalitet fremgår ikke af figuren, men ligger mellem de to laveste kurver.

I Figur 2-7 er standardafvigelse og variansestimater for boringer, geofysik og modelleringsområder samlet i et plot, hvor variansstørrelsesforholdene mellem metoderne kan ses. Boringers varians ligger lavest og modelleringsområdernes varians ligger som et loft over variansestimaterne for alle metoderne.



Figur 2-6: Standardafvigelses- (venstre) og variansestimat (højre) for modelleringsområder.



Figur 2-7: Standardafvigelses (venstre) og variansestimat (højre) for alle typer: boringer, geofysik og modelleringsområder. Legende for kurverne findes i Figur 2-1, Figur 2-2, Figur 2-3 og Figur 2-6. Derudover er refleksionsseismik markeret med sort stiplet (--) linje, gammalog med rød stiplet (--) linje og resistivitetslog med grøn stiplet (--) linje.

2.5 Usikkerhedsestimering relateret til geologisk kompleksitet

Eftersom brugen af lagmodeller generelt er en bedre approksimation i områder med simpel geologi i forhold til områder med kompleks geologi, har vi inddraget et tema relateret til den geologiske kompleksitet. Der var ikke tid til at udarbejde en national oversigt over den geologiske kompleksitet i dette projekt, og der blev derfor benyttet et kompleksitetskort udarbejdet af Sandersen (2021) i regi af Mapfield projektet. Dette kort beskriver den geologiske kompleksitet af de øverste 30 m af undergrunden vurderet i en målestok på ca. 1:300.000 (Figur 2-8). Kortet fokuserer kun på kompleksiteten af den øverste del af lagserien, og er derfor kun benyttet for de kvartære aflejringer.

Kompleksitetskortet er udviklet på baggrund af information fra eksisterende geologisk information: landskabstype (Aarhus Universitet 2005) og sekundært på topografi, jordartskort (Jacobsen & Tougaard 2020), geomorfologisk kort (Smed 1978), borehulsdata, og diverse geologiske publikationer om overfladenær geologi i Danmark. Proceduren er beskrevet i Sandersen (2021). Kompleksitetsklasse 1 har mindst geologisk kompleksitet (fx yngre marin flade over Yoldia marin flade) og kompleksitetsklasse 4 har højst kompleksitet (fx randmoræne i yngre glacialt landskab over ældre landskaber). Vi anvender de fire geologiske kompleksitetsklasser i beskrivelsen af, hvorledes information fra de enkelte datapunkter udbredes lateralt. I områder, hvor den geologiske kompleksitet er lav, kan informationen fra et datapunkt udbredes over et større område, end hvor den geologiske kompleksitet for de kvartære lag i den hydrostratigrafiske model. Alle prækvartære lag antages at have lav geologisk kompleksitet.



Figur 2-8: Geologisk kompleksitet, fordelt på fire klasser, hvor klasse 4 er højest kompleksitet. Kortet er efter af Sandersen (2021).

2.6 Samlet usikkerhedsestimering

I første del af vores procedure til usikkerhedsestimering har vi vurderet usikkerheden i det enkelte datapunkt og lavet en variansfunktion som beskriver, hvorledes usikkerheden varierer med dybden i datapunktet for hver datatype (afsnit 2.2 og 2.3). I den næste del af proceduren til usikkerhedsestimering beskriver vi, hvordan usikkerhedsestimatet fra de enkelte datapunkter udbredes lateralt og kombineres, så vi kan beregne en samlet usikkerhed i et tolkningspunkt knyttet til en modelflade, se Figur 2-9.



Figur 2-9: Illustration af usikkerhed fra tre forskellige kilder til information repræsenteret i et punkt på fladen.

Figur 2-9 viser, hvordan usikkerheden er defineret omkring fladens værdi. I dette eksempel er der tre kilder til information (lilla, blå og grøn kurve). I nogle tilfælde vil én type information være meget god til at indikere en laggrænses placering og kurven bliver smal. I andre tilfælde er der flere stykker information, der er gode og alle bidrager til at mindske usikkerheden. I enkelte tilfælde er alle stykker information meget usikre.

2.6.1 Vished

Vi baserer vores beregninger af den samlede usikkerhed på begrebet "vished", $1/\sigma^2$, som vi definerer som den inverse af variansen (=standardafvigelse²), σ^2 . Dernæst tager vi udgangspunkt i, at to sandsynlighedsfordelinger kan kombineres ved multiplikation under en antagelse om uafhængighed mellem informationskilderne I_1 og I_2

 $f(x|I_1, I_2) = k \cdot f(x|I_1) \cdot f(x|I_2).$

Hvis vi dernæst antager, at vi kan beskrive alle usikkerheder som normalfordelinger med samme middelværdi, μ , er givet

$$\begin{split} N(\mu, \sigma_{combined}^2) &= k \cdot N(\mu, \sigma_1^2) \cdot N(\mu, \sigma_2^2) \\ N(\mu, \sigma_{combined}^2) &= k \cdot \frac{1}{\sqrt{2\pi\sigma_1^2}} e^{-\frac{1}{2} \left(\frac{(x-\mu)^2}{\sigma_1^2}\right)} \cdot \frac{1}{\sqrt{2\pi\sigma_2^2}} e^{-\frac{1}{2} \left(\frac{(x-\mu)^2}{\sigma_2^2}\right)} \\ N(\mu, \sigma_{combined}^2) &= k \cdot \frac{1}{2\pi\sigma_1\sigma_2} e^{-\frac{1}{2} \left(\frac{(x-\mu)^2}{\sigma_1^2} + \frac{(x-\mu)^2}{\sigma_2^2}\right)} \\ N(\mu, \sigma_{combined}^2) &= \frac{k}{2\pi\sigma_1\sigma_2} e^{-\frac{1}{2} (x-\mu)^2 \left(\frac{1}{\sigma_1^2} + \frac{1}{\sigma_2^2}\right)} \end{split}$$

$$N(\mu, \sigma_{combined}^2) = \frac{k}{2\pi\sigma_1\sigma_2} e^{-\frac{1}{2}\left(\frac{(x-\mu)^2}{\sigma_{combined}^2}\right)}$$

Givet disse antagelser kan man se, at visheder kan summeres

$$\frac{1}{\sigma_{combined}^2} = \frac{1}{\sigma_1^2} + \frac{1}{\sigma_2^2}$$

Eftersom de anvendte informationskilder i tolkningen ikke er betingede af hinanden, anser vi antagelsen om uafhængighed for værende opfyldt. Således forventer vi for eksempel ikke at datausikkerheden på en boring ændres af, at der på samme lokalitet også findes TEM sonderinger. Hvis disse kilder var korrelerede, ville vores borehulsbeskrivelse blive bedre og bedre hver gang vi indsamlede ny geofysik på samme lokalitet.

Den anden antagelse om, at usikkerheden kan beskrives Gaussisk med centrum i laget, kan delvist retfærdiggøres ved at tolkningsusikkerheden må forventes at tilhøre en laggrænse og være henført til laggrænsen af den geologiske modellør. Den Gaussiske antagelse er ydermere en nemmere abstraktion at operere med end mange andre fordelinger, da den er meget intuitiv for modelløren, da det må forventes, at muligheden for fejlplacering af laget må aftage med afstanden til laggrænsen, og at det er en meget brugt fordeling. Valget kan dog ikke retfærdiggøres til fulde, da det er muligt at forestille sig andre fordelinger, hvor usikkerheden falder med afstanden til laggrænsen. Hvis en alternativ fordeling skal indføres, vil dette dog ligeledes kræve enten information om formen på denne fordeling eller et aktivt subjektivt valg.

For at udbrede visheden $1/\sigma^2$ lateralt har vi brug for at definere funktioner, der beskriver hvorledes vished vil aftage (henfalde) med afstanden fra datapunktet. Hvor hurtigt visheden aftager vil afhænge af den geologiske kompleksitet således at høj geologisk kompleksitet vil føre til en hurtig reduktion i visheden. Hvor hurtigt visheden aftager beskrives med udtrykket range. Vi definerer en generel funktion til beskrivelse af vishedens henfald

 $V(x) = 1/var_0 \cdot exp(-3 \cdot x^2/r^2),$

hvor x er afstanden lateralt fra et datapunkt med variansen var $_0$ og r er range bestemt af den geologiske kompleksitet ved datapunktet.

For geologisk kompleksitet givet fra lav til høj (klasse 1 til klasse 4) er range henholdsvis 500 m, 400 m, 250 m og 100 m. Disse størrelsesordner er valgt ud fra erfaringer i, hvor langt man kan ekstrapolere data i forskellige geologiske miljøer, når man laver geologisk modellering. For prækvartære lag er range altid 500 m, da den geologiske kompleksitet anses for at være lav i disse lag.

Der laves en række temakort ($[T_1, T_2, ..., T_n]$) for usikkerheden udbredt til alle celler i den diskretiserede modelflade, således at der er et temakort for hver kilde til information og for hvert modellag. Den samlede usikkerhed for et modellag er så en kombination af alle usikkerheder, beskrevet i afsnittene 2.2, 2.3 og 2.4.

2.6.2 Vished-henfaldsfunktion for boringer

For boringer bruges vished-henfaldsfunktionen direkte, så

$$V_{\text{boringer}}(x,z) = 1/var_{0, \text{ boringer}}(z) \cdot exp(-3 \cdot x^2/r^2),$$

hvor var_{0, boringer} er defineret i afsnit 2.2. Figur 2-10 illustrerer, hvordan visheds-henfaldsfunktionen både afhænger af afstanden, x, fra datapunktet til tolkningspunkt og dybden, z, til den lagflade i den hydrostratigrafiske model, hvortil tolkningspunktet er knyttet. Figur 2-1 viser variansfunktionen for boringskvalitetsklasse 2 for snappede boringer.



Figur 2-10: Vished-henfaldsfunktion for boringer med tolkningspunkt i dybderne i 0 m, 15 m, 50 m og 100 m for fire klasser af geologisk kompleksitet (klasse 4 =blå kurve, klasse 3 = grøn kurve, klasse 2 = rød kurve og klasse 1 = magenta kurve).

For hvert af de hydrostratigrafiske lag, konstrueres et temakort i et grid med samme udbredelse og cellestørrelse, hvor der i hver node beregnes en samlet borings-vished. Dette gøres ved at summere visheden for de tre nærmeste snappede boringer og de tre nærmeste af de øvrige boringer inden for en radius af 600 m fra nodepunkter, da visheden for en boring med længste range på 500 er tæt på nul for en afstand på mere end 600 m. Hvis der ikke findes boringer indenfor en 600 meters radius af nodepunktet, sættes visheden til 0 for denne node i griddet. Figur 2-11 viser varianskort (det inverse af visheds-griddet) for Fyn for hvert af de hydrostratigrafiske lag. I de mest terrænnære lag giver boringer eller klynger af boringer anledning til de laveste varianser, hvilket også er forventet ud fra variansfunktionen for boringerne, som beskriver stigende varians med dybden. De forskellige varianser i Figur 2-11 afspejler, at boringerne har forskellig boringskvalitet.



Figur 2-11: Varianskort for boringer for alle lag i den hydrostratigrafiske model for Fyn.

2.6.3 Vished-henfaldsfunktion for geofysiske data

Vished-henfaldsfunktionen for de geofysiske data er en modificeret version af den generelle Vished-henfaldsfunktion, hvor Vishedsfunktionen kun gælder inden for en kernelfunktion, som approksimerer en geofysisk målings følsomhed. Kernelfunktionen er defineret som $Kernel(z) = max(2 \cdot z, 75)$

for TEM-metoderne og

Kernel(z) = 75

for de geoelektriske metoder og

 $V_{geofysik}(x) = (x \le Kernel(z)) \cdot 1/var_{0,geofysik} \cdot exp(-3 \cdot (x - Kernel(z))^2/r^2),$

hvor var_{0,geofysik} er en af variansfunktionerne defineret i afsnit 2.3. Vished-henfaldsfunktionen afhænger både af afstanden, x, fra datapunktet til tolkningspunkt og dybden, z, til den lagflade i den hydrostatigrafiske model, hvortil tolkningspunktet er knyttet. Information fra en geofysisk måling bidrager kun til den samlede usikkerhed inden for Kernel-funktionen. I forbindelse med justering af funktionen under udviklingen af procedurerne til kvantificering af den samlede usikkerhed i tolkningspunkterne er der desværre sneget sig en fortegnsfejl ind, så det er ved randen af kernelfunktionen og ikke i centrum af kernelfunktionen, at visheden er størst. Ved høj datatæthed har fejlen ikke en stor betydning for den samlede datatæthed.

Figur 2-12 viser som eksempel på varianskort, varianskort for SkyTEM relateret til de hydrostratigrafiske lag på Fyn. Når man sammenligner forskellige områder på de enkelte varianskort, kan man se variationer, som er relateret til, at den hydrostratigrafiske flade varierer i dybden. Samme dybderelaterede variationer kan man se, når man sammenligner varians i det samme område for de forskellige hydrostratigrafiske lag. På det nordlige Fyn kan man for de to øverste lag (ks1t og ks1b) se en sporadisk lav varians, som bliver mere sammenhængende fra det tredje lag (ks2t). Den meget høje varians i de to øverste lag skyldes, at lagene ligger så tæt på terræn, at SkyTEM i variansfunktion er defineret til ikke at bidrage med information om lagfølgen, se afsnit 2.3.1. I de underliggende lag stiger variansen, eftersom lagenes dybde stiger.



Figur 2-12: Varianskort for SkyTEM for alle lag i den hydrostratigrafiske model for Fyn.

2.7 Eksempler på kvantificering af usikkerhed

Figur 2-13 viser det samlede varianskort for lagfladen 1200_Kvartær_sand_bund for Jylland og den tilsvarende lagflade, ks3_bund for Fyn og Sjælland, som i alle landsdele repræsenterer dybden til bunden af det tredje kvartære sandlag. Variansen er den reciprokke til den summerede vished beskrevet i afsnit 2.6.1. Som hjælp til at illustrere at de fleste af de

individuelle variansfunktioner har en dybdeafhængighed er dybden til de anvendte lagflader vist i Figur 2-14. Gennemgående for lagfladerne kan det ses, at det er boringer, som bidrager mest til en lav varians på mindre end 1-5 m² (standardafvigelse på 1- 2,5 m). For de øvre, terrænnære lagflader bidrager PACES til at reducere variansen til et niveau omkring 5-10 m² (standardafvigelse på 2,5-3 m). Hvor lagfladerne befinder sig terrænnært i dybder op til 20-30 m (Figur 2-14) bidrager SkyTEM og de øvrige geofysiske metoder til et variansniveau på omkring 5-15 m² (standardafvigelse på 2,5-4 m), mens variansniveauet for de dybest liggende områder med dybder over 70 m (Vestfyn) stiger til omkring 50-100 m² (standardafvigelse på 7-10 m). For de mest terrænnære områder på lagfladerne ses det tydeligt, at modelleringsområdevariansen bidrager til, at den samlede varians maksimalt bliver 30-50 m² (standardafvigelse på c. 5-7 m). For de mellemste og dybest liggende områder af lagfladerne er det overvejende boringer og SkyTEM, der bidrager til at reducere variansen til et niveau omkring 30-50 m² (standardafvigelse på c. 5-7 m).

Figur 2-15 og Figur 2-16 viser hhv. den samlede varians og dybden til Prækvartæroverfladen (lagfladerne '2400_kvartær_bund_ler' i Jylland, og 'preq' på Fyn og Sjælland). Figurerne viser de samme tendenser for den samlede varians som i Figur 2-13 og med større variationer i den samlede varians, da der er større variationer i dybden til lagfladen (Figur 2-16).

Figur 2-17 til Figur 2-24 viser profilsnit med dels a) datagrundlaget, b) de beregnede standardafvigelser fra hver af de bidragende datatyper og den samlede standardafvigelse for udvalgte hydrostratigrafiske lag i modellen og c) hvordan den samlede standardafvigelse varierer omkring lagfladen i et dybdeprofil. Fra fx Figur 2-17b og Figur 2-18b fremgår det, som også illustreret i Figur 2-13, at det er boringer som bidrager til de laveste standardafvigelser, og jo tættere de lateralt er placeret på beregningsnoden, jo lavere standardafvigelse (c. 0.3-1 m). Hvor der er indflydelse fra SkyTEM data ligger standardafvigelserne i et niveau mellem det laveste niveau for boringer og et niveau, der kommer fra modelleringområdestandardafvigelsen, som lægger loftet for størrelsen af standardafvigelsen. Hvor de forskellige niveauer ligger afhænger af dybden til lagfladen (Figur 2-17b og Figur 2-18b; Figur 2-19 og Figur 2-20 samt Figur 2-23 og Figur 2-24) og ca. med halvdelen af det niveau, der kommer fra modelleringområde-standardafvigelsen. På Figur 2-23 og Figur 2-24 kan man se forskellen på tTEM (omkring profilkoordinat 500.000 m) med tyndere modellag og dermed lavere varians (se afsnit 2.3.1) end SkyTEM (omkring profilkoordinater 510.000–513.000 m).

Samlet set viser den totale standardafvigelse, at lagfladerne får en lille usikkerhed (ca. ± 1 m) tæt ved boringer og har et mellemniveau tæt ved geofysiske data og det højeste niveau, hvor der ikke er bidrag fra nogen data. Størrelsen på standardafvigelsen i områder uden tætliggende data, der kan bidrage til reduktion af standardafvigelsen, afhænger af dybden til lagfladen (Figur 2-17c til Figur 2-24c).



Figur 2-13: Samlet varianskort (reciprokke af vished summeret for alle typer af data og modelleringsområder) for lagfladen '1200_kvartær_sand_bund' i Jylland og 'ks3-bund' på Fyn og Sjælland, som i alle landsdele repræsenterer bunden til det tredje kvartære sand-lag. De fire sorte streger markerer V-Ø-profiler vist i Figur 2-17 til Figur 2-24.



Figur 2-14: Dybden under terræn i meter for lagfladen '1200_kvartær_sand_bund' i Jylland og 'ks3-bund' på Fyn og Sjælland, som i alle landsdele repræsenterer bunden til det tredje kvartære sandlag. De fire sorte streger markerer V-Ø-profiler vist i Figur 2-17 til Figur 2-24.



Figur 2-15: Samlet varianskort (reciprokke af vished summeret for alle typer af data og modelleringsområde) for lagfladen '2400_kvartær_ler_bund' i Jylland og 'preq' på Fyn og Sjælland, som i alle landsdele repræsenterer Prækvartæroverfladen. De fire sorte streger markerer V-Ø-profiler vist i Figur 2-17 til Figur 2-24.



Figur 2-16: Dybde under terræn i meter for lagfladen '2400_kvartær_ler_bund' i Jylland og 'preq' på Fyn og Sjælland, som i alle landsdele repræsenterer Prækvartæroverfladen. De fire sorte streger markerer V-Ø-profiler vist i Figur 2-17 til Figur 2-24. Bemærk at farveskalaen spænder over et større dybdeinterval end i Figur 2-14.



Figur 2-17: VØ-profil for hydrostratigrafisk lag ks1-bund på Vestfyn. a) kort over data som er til rådighed i estimering af usikkerheden langs det viste profil. b) Standardafvigelse for datatyper, som bidrager til den samlede standardafvigelse og den samlede standardafvigelse. c) V-Ø profilsnit i den hydrostratigrafiske model, hvor lagfladen ks1bund er fremhævet og plottet sammen med ± 1 standardafvigelse (svarer til den sorte stiplede kurve i b). Placeringen af profilet er markeret på Figur 2-13 til Figur 2-16.


Figur 2-18: VØ-profil for hydrostratigrafisk lag ks3-bund på Vestfyn. a) kort over data som er til rådighed i estimering af usikkerheden langs det viste profil. b) Standardafvigelse for datatyper, som bidrager til den samlede standardafvigelse og den samlede standardafvigelse. c) V-Ø profilsnit i den hydrostratigrafiske model, hvor lagfladen ks3-bund er fremhævet og plottet sammen med ± 1 standardafvigelse (svarer til den sorte stiplede kurve i b). Placeringen af profilet er markeret på Figur 2-13 til Figur 2-16.



Figur 2-19: VØ-profil for hydrostratigrafisk lag ks3-bund på Vestsjælland. a) kort over data som er til rådighed i estimering af usikkerheden langs det viste profil. b) Standardafvigelse for datatyper, som bidrager til den samlede standardafvigelse og den samlede standardafvigelse. c) V-Ø profilsnit i den hydrostratigrafiske model, hvor lagfladen ks3-bund er fremhævet og plottet sammen med ± 1 standardafvigelse (svarer til den sorte stiplede kurve i b). Placeringen af profilet er markeret på Figur 2-13 til Figur 2-16.



Figur 2-20: VØ-profil for hydrostratigrafisk lag 'preq', repræsenterende Prækvartæroverfladen på Vestsjælland. a) kort over data som er til rådighed i estimering af usikkerheden langs det viste profil. b) Standardafvigelse for datatyper, som bidrager til den samlede standardafvigelse og den samlede standardafvigelse. c) V-Ø profilsnit i den hydrostratigrafiske model, hvor lagfladen 'preq' er fremhævet og plottet sammen med ± 1 standardafvigelse (svarer til den sorte stiplede kurve i b). Placeringen af profilet er markeret på Figur 2-13 til Figur 2-16.



Figur 2-21: VØ-profil for hydrostratigrafisk lag '1200_kvartær_sand_bund', repræsenterende bunden af det tredje sandlag i det sydlige Jylland. a) kort over data som er til rådighed i estimering af usikkerheden langs det viste profil. b) Standardafvigelse for datatyper, som bidrager til den samlede standardafvigelse og den samlede standardafvigelse. c) V-Ø profilsnit i den hydrostratigrafiske model, hvor lagfladen '1200_kvartær_sand_bund' er fremhævet og plottet sammen med ± 1 standardafvigelse (svarer til den sorte stiplede kurve i b). Placeringen af profilet er markeret på Figur 2-13 til Figur 2-16.



Figur 2-22: VØ-profil for hydrostratigrafisk lag '2400_kvartær_ler_bund', repræsenterende Prækvartæroverfladen i det sydlige Jylland. a) kort over data som er til rådighed i estimering af usikkerheden langs det viste profil. b) Standardafvigelse for datatyper, som bidrager til den samlede standardafvigelse og den samlede standardafvigelse. c) V-Ø profilsnit i den hydrostratigrafiske model, hvor lagfladen '2400_Kvartær_ler_bund' er fremhævet og plottet sammen med ± 1 standardafvigelse (svarer til den sorte stiplede kurve i b). Placeringen af profilet er markeret på Figur 2-13 til Figur 2-16.



Figur 2-23: VØ-profil for hydrostratigrafisk lag '1200_kvartær_sand_bund', repræsenterende bunden af det tredje sandlag i det nordlige Jylland. a) kort over data som er til rådighed i estimering af usikkerheden langs det viste profil. b) Standardafvigelse for datatyper, som bidrager til den samlede standardafvigelse og den samlede standardafvigelse. c) V-Ø profilsnit i den hydrostratigrafiske model, hvor lagfladen '1200_kvartær_sand_bund' er fremhævet og plottet sammen med ± 1 standardafvigelse (svarer til den sorte stiplede kurve i b). Placeringen af profilet er markeret på Figur 2-13 til Figur 2-16.



Figur 2-24: VØ-profil for hydrostratigrafisk lag '2400_kvartær_ler_bund', repræsenterende Prækvartæroverfladen i det nordlige Jylland. a) kort over data som er til rådighed i estimering af usikkerheden langs det viste profil. b) Standardafvigelse for datatyper, som bidrager til den samlede standardafvigelse og den samlede standardafvigelse. c) V-Ø profilsnit i den hydrostratigrafiske model, hvor lagfladen '2400_Kvartær_ler_bund' er fremhævet og plottet sammen med ± 1 standardafvigelse (svarer til den sorte stiplede kurve i b). Placeringen af profilet er markeret på Figur 2-13 til Figur 2-16.

3. Geostatistiske simuleringer af usikkerhederne

Den overordnede idé med den geostatistiske modellering i dette projekt er at tage udgangspunkt i en allerede defineret hydrostratigrafisk 3D lagflademodel og lade variationen inden for tolkningsusikkerheden på de enkelte flader komme til udtryk gennem en række realisationer. Modelgrundlaget er DK-modellen (FOHM i Jylland) med lokale opdateringer, som er beskrevet i Rapporten 'Lokale opdateringer af hydrostratigrafien i FOHM-modellen til Kvælstofretentionskortet version 2025' (Mortensen et al. 2025).

Metoden til at simulere lagfladerne blev første gang beskrevet i Madsen et al. (2022) under navnet GDM. Her blev GDM anvendt på en lokal hydrostratigrafisk model omkring Egebjerg (lige nord for Horsens), i Danmark. I det følgende vil det blive beskrevet, hvordan GDM er blevet opskaleret til at virke på større områder (national skala). Udvidelsen benævnes i det følgende GDM-N, som står for "Geology Driven Modelling - National". GDM-N er udviklet specifikt til dette projekt, da der til dato ikke findes geostatistiske metoder, der er anvendelige på national skala, og derfor måtte en ny tilgang designes til formålet.

Det overordnede workflow for den geostatiske simulering er kort skitseret i det følgende og vist overordnet i Figur 3-1. Her ses, at der tages udgangspunkt i de eksisterende flader plus nytolkninger af lavbundsjorde og tTEM-områder (Mortensen et al., 2025), samt en opdateret topografi fra april 2023 (gule bokse). Disse flader danner grundlag for den kalibreringsmodel (orange boks), der skal benyttes dels som udgangspunkt for den hydrologiske kalibrering, men også til den efterfølgende geostatistiske modellering. Kalibreringsmodellen kombineres med masker for land og udbredelse (lilla bokse) til at minimere områder, hvor modellen skal beregnes. Lagenes udbredelsesmasker beregnes gennem de eksisterende flader og tolkningspunkter.

GDM her to krav der skal opfyldes (de blå bokse i Figur 3-1):

- 1. Usikkerhederne for modelfladernes punkter er beskrevet kvantitativt
- 2. Fladernes forløb er karakteriseret gennem en sandsynlighedsfordeling

I projektet er der udviklet et nationalt koncept og metode til kvantificering af tolkningsusikkerheder på lagflader i lagmodeller (se kapitel 2). Ved at anvende varianskort for hvert lag er det muligt at opskalere GDM til GDM-N ved at tilskrive hvert tolkningspunkt den estimerede usikkerhed, som er angivet for den pågældende placering i de respektive varianskort.

For at udvide GDM til GDM-N er der også behov for at opfylde krav nr. 2 om, at fladernes forløb er beskrevet ved hjælp af en sandsynlighedsfordeling. Da det ikke kan forventes, at fladerne opfører sig ens over hele landet, er GDM-N udviklet til at kunne håndtere en lokalt varierende statistik ved hjælp af en estimeret Machine Learning (ML)-model. ML-modellen kan tage en punktserie fra en flade (f.eks. tolkningspunkter) og oversætte til en passende statistisk model indenfor et vindue af fladen. ML-modellen bruges på alle flader til at lave en punktvis statistisk model for alle gridpunkter i alle lag. Dernæst benyttes en clustering-rutine til at efter-processere den punktvise statistik og sammensætte i større områder med tilnær-melsesvis samme statistisk. Resultatet er en beskrivelse af den ikke-stationære, altså lokalt-

varierende, statistik (blå kasse til venstre, krav nr. 2). Denne procedure er beskrevet i yderligere detalje i (Falk og Madsen, 2023).



Flowchart - Geostastisk setup til simulering af 3D hydrostratigrafi

Figur 3-1: Skematisk overview af modelleringsarbejdet

Når usikkerhedstemaerne er kvantificeret, og der er lavet en zonering på fladerne (blå bokse), kan der udføres en klassisk "kriging" inden for de enkelte områder, som har sammenlignelig statistik baseret på den model, der er kommet ud af ML-modellen (lyserøde bokse til venstre). I grænseområdet mellem to eller flere forskellige områder med forskellige statistiske egenskaber, smeltes fladen sammen ved afstandsvægtning baseret på euklidiske afstande. På baggrund af middelværdien og variansen fra kriging-estimatet kan der laves realisationer ved Cholesky dekomposition (Higham, 2009, lyserød boks til højre). Disse realisationer af hver flade kan sættes sammen til en 3D model og derefter lagfladekorrigeres (sort boks til højre) til et danne én 3D realisation af den samlede model. Et sæt af 3D realisationer udgør det endelige stokastiske ensemble over variationen i den hydrostratigrafiske model på basis af de usikkerhedstemaer, der er genereret i Kapitel 2. Fra ensemblen kan der også beregnes en middelmodel og varians. I de følgende afsnit beskrives først den overordnede databearbejdning og opsætning for kørsel af GDM-N. Dette inkluderer:

Afsnit 3.1: Opstilling af kalibreringsmodellen.

- Afsnit 3.2: Hvordan maskerne er produceret.
- Afsnit 3.3: Punktudtagningsstrategien anvendt i projektet.

Herefter følger en detaljeret gennemgang af den geostatistiske simulering i GDM-N, som består af to hoveddele:

Afsnit 3.4: Estimering af en ikke-stationær zoneret statistisk model baseret på kalibreringsmodellen.

Afsnit 3.5: Generering af hydrostratigrafiske lag vha. geostatistisk simulering samt sammensmeltning og lagfladekorrektion.

Til sidst beskrives de implementerede algoritmer og resultaterne fra simuleringerne:

Afsnit 3.6: En gennemgang af den implementerede algoritme. Afsnit 3.7: Et udvalg af de opnåede resultater.

3.1 Kalibreringsmodellen

Kalibreringsmodellen er en hydrostratigrafisk model, der fungerer som strukturelt input i den hydrologiske model. Da den hydrologiske model skal kalibreres inden modelleringen af de enkelte 3D realisationer af hydrostratigrafien påbegyndes, er det derfor vigtigt at kalibreringsmodellen også fungerer som basis for den geostatistiske hydrostratigrafiske modellering.

Til opstilling af kalibreringsmodellen blev den mest opdaterede version af DK- og FOHMmodellen anvendt. Da den hydrostratigrafiske modellering skulle udføres forud for den efterfølgende grundvandsmodellering, blev denne proces gennemført i foråret 2023. Derfor er kalibreringsmodellen baseret på april 2023-versionen og den tilhørende topografi. Lagene fra disse modeller blev derefter opdateret med de lokale forbedringer af FOHM-modellen, udført i regi af kvælstofretentionsprojektet og beskrevet i Mortensen et al., (2025). Indarbejdelsen af disse er beskrevet i afsnit 3.1.1 og 3.1.2.

3.1.1 Sen og postglaciale lag på Fyn og i Jylland

En af de lokale opdateringer i kvælstofretentionsprojektet omhandlede modellering af senog postglaciale lag (Mortensen et al., 2025) tolket i programmet Geoscene3D. Nord for Limfjorden (Vendsyssel) er dette arbejde udført af GEO for Miljøstyrelsen, mens områderne syd for Limfjorden er tolket af GEUS som beskrevet i Mortensen et al., (2025). Syd for Limfjorden er modelleringen begrænset af, at den rumlige udbredelse er større end den datamængde, som programmet (Geoscene3D) kan holde i hukommelsen ad gangen. Derfor er alle lag tolket i mindre områder (se Figur 3-3), defineret som såkaldte "tiles". I det udviklede workflow, kræves der derfor en række skridt for at generere de samlede nytolkede flader:

- 1. Alle tiles syd for Limfjorden indlæses og placeres rigtigt i det overordnede grid for at danne flader. I tiles hvor der ikke er modelleret indsættes terrænet.
 - a. Evt. små uoverensstemmelser i kote mellem tiles glattes ud.
 - b. Da lagene er modelleret efter FOHM-griddet og dermed 50 m forskudt af DKmodellen, laves der en re-interpolation af fladerne, samt en korrektion for tabt topografisk information.
- 2. Vendsyssel-fladerne indlæses.
- 3. Evt. interpolationsartefakter fra Geoscene3D-modelleringen fjernes fra begge sæt flader.
- 4. Vendsyssel smeltes sammen med fladerne syd for Limfjorden ved en maske (se Figur 3-2) med en glidende overgang hen over Limfjorden.



Figur 3-2: Maske til sammensmeltning af Vendsyssels model med nytolkninger syd

Den ovenstående procedure er regionsspecifik, så der for Fyn springes over punkt 2 og 4 i den ovenstående algoritme, hvor al modellering er foretaget af GEUS og er ensrettet. På Sjælland er der ikke modelleret Sen- og postglaciale lag.

Figur 3-3 viser hvordan Danmark er inddelt i 25 km x 25 km tiles samt andelen af landet, der er modelleret med sen- og postglaciale lag (Mortensen et al., 2025).



Figur 3-3: Tilegrid for Danmark per 31/04-2023 plus udbredelse og mægtigheder af tolkninger af sen- og postglaciale lag. a-e) Mægtigheder af sen- og postglaciale lag tolket af GEO og GEUS. f) Samlet mægtighed af ny-modellerede lag. Bemærk den ændrede farveskala i figur f).

3.1.2 Hydrostratigrafi i tTEM kortlagte områder

Den nytolkede hydrostratigrafi i de tTEM-kortlagte områder fra Mapfield (Mortensen et al., 2025) er indlejret i modellen ved, at de enkelte lag er opdateret indenfor randzonen omkring områderne. Specifikt klippes den nye model ind i de flader, der er dannet ved først at indføre alle post- og senglaciale lag som beskrevet i foregående afsnit. I tilfælde af, at der er små uoverensstemmelser mellem de enkelte grids, bruges en vægtet midling, der er baseret på euklidiske afstande. Disse uoverensstemmelser er ikke markante, men kan opstå som følge af, at det kan være svært at genskabe fladerne helt præcist lige uden for tTEM-områderne, idet der typisk kræves en anden interpolationsparameterisering i tTEM-områderne end i det omkringliggende pga. den finere opløsning på tTEM data.

3.1.3 Korrektion for terrænvariationer

FOHM- og DK-modellen benytter ikke samme griddefinition på grund af forskellige filformater (.grd og .xyz), hvor griddet enten defineres gennem cellernes centrum eller i nederste venstre hjørne. Dette kan føre til forskydninger mellem griddene afhængigt af, hvordan filerne indlæses og gemmes i det anvendte program. Derudover findes der forskellige versioner af modellerne, da højdemodellen over Danmark løbende forfines og ændres over tid.

Afhængigt af modelløren kan forskellige terrænmodeller være anvendt som baggrund for modelleringen. Dette kan i nogle tilfælde resultere i lag, der er tolket højere end terrænkoten i topografien fra 2023, som bruges som basis for kalibreringsmodellen. For at løse dette problem og undgå komplikationer i den efterfølgende hydrologiske modellering, er topografien fra 2023 anvendt til at korrigere ved at påtrykke de resterende lag i kalibreringsmodellen.

3.2 Masker

For at sikre at beregningerne kun udføres i områder, hvor lag forventes at eksistere eller kunne tænkes at optræde, selvom de ikke er modelleret i FOHM og DK-modellen, er der lavet masker for alle relevante lag. Dette er primært gjort for at håndhæve den geologiske information, men har samtidig også en positiv indvirkning på den samlede beregningstid.

3.2.1 Masker for landmasser

Fælles for alle modelområder er, at der er lavet masker, der definerer forholdet mellem landmasse og hav. Disse masker sikrer, at der kun modelleres det absolut nødvendige fra kystzonen. Denne afstand er sat til 2 km euklidisk afstand fra kystzonen i Danmark. Denne afstand er sat tilpas langt ude i vandet til at den efterfølgende hydrologiske modellering ikke påvirkes af resultatet, og at de eksisterende randbetingelser kan benyttes. Efter denne afstand sammensmeltes modellen med den eksisterende DK-model på åbent vand. Figur 3-4a viser et eksempel på et maskelag for Sjælland, hvor også fjordene er inkluderet i landmassen. Figur 3-4c viser den kombinerede maske, hvor den lokale maske for lagudbredelse for KS4 på Sjælland (Figur 3-4b) er taget i betragtning.



Figur 3-4: Maskelag for kvartært sand 4 - Eksempel fra Sjælland. a) onshore-maske, b) lagets udbredelse, c) kombineret maske.

3.2.2 Lagspecifikke masker

For de lagspecifikke masker på land følger her en gennemgang af de forskellige modelleringsområder:

3.2.2.1 Jylland, Endelave, Tunø og Samsø

Brugen af masker er i særdeleshed vigtig i Jylland, hvor mange kvartære lag dækker store dele af Jylland, mens de Miocæne lag kun dækker sparsomme områder af det samlede areal. Et eksempel på dette ses i Figur 3-5 og Figur 3-6. Maskerne for de oprindelige FOHM-lag er lavet ved følgende fremgangsmåde:

3.2.2.1.1 Kvartære laggrænser

Kvartære lag er sat med en udbredelse på 5 km fra et positivt tolkningspunkt. Et positivt tolkningspunkt forstås i denne henseende som et punkt, der er sat, hvor laget har en mægtighed i modellen. Hvis der inden for disse fem kilometer findes negative tolkningspunkter, der omvendt angiver en 'ikke-tykkelse' af et lag, benyttes disse til at afgrænse området. Maskerne er konstrueret af Hydrologisk afdeling på GEUS. Valget på 5 km kommer af et ønske om at give lag mulighed for at have en udbredelse ud over den mægtighed, der er angivet i FOHM. Den er her sat som 2 gange den valgte interpolationsradius på 2,5 km i den oprindelige FOHM-model.

3.2.2.1.2 Prækvartæret

Bunden af det kvartære lagpakke (2400: Kvartær Ler) definerer prækvartæroverfladen. Det er vigtigt at hele denne flade simuleres. Derfor er der ingen lokal maske på dette lag.



Figur 3-5: Maske for "1400 Kvartær Sand Bund" i FOHM, der viser et lag med forventning om stor udbredelse i den samlede model.



Figur 3-6: Maske for "5900 Klintinghoved KRL5 Bund" i FOHM, der viser et lag med forventning om lille lokal udbredelse i den samlede model.

3.2.2.1.3 Miocæne laggrænser

Miocæne lag er lavet på samme måde som i kvartære lag, dog med en udbredelse på 10 km fra et positivt tolkningspunkt. Maskerne er ligeledes konstrueret af Hydrologisk afdeling på GEUS. Valget på 10 km kommer af et ønske om at give lag mulighed for at have en udbredelse ud over den mægtighed, der er angivet i FOHM. Den er her sat som 4 gange den valgte interpolationsradius på 2,5 km i den oprindelige FOHM-model.

3.2.2.2 Fyn

Kvartære lag i DK-modellen har en udbredelse over det meste af Fyn. Derfor er det ikke hensigtsmæssigt at fastlægge specifikke områder, hvor lagene kan simuleres, som det er gjort i Jylland. I stedet kan alle kvartære lag simuleres i hele området på Fyn – altså med en fuld maske for alle kvartære lag. Det samme gælder for det Paleogene ler og kalken. Principielt kan det Paleogene ler muligvis have en mere begrænset udbredelse i den sydøstligste del af området, nær toppen af Langeland. På grund af det begrænsede datagrundlag er dette dog usikkert, og derfor er det vurderet at det Paleogene ler også bør kunne simuleres overalt.

Den eneste flade med en lokal begrænsning er KS3, som kun findes i den vestlige og nordlige del af Fyn i DK-modellen. Denne begrænsning fastholdes i hydrostratigrafien for at sikre, at den efterfølgende hydrologiske modellering ikke afviger væsentligt fra det eksisterende modelgrundlag i DK-modellen. Masken til KS3 er designet af Hydrologisk afdeling på GEUS.

3.2.2.3 Sjælland

På Sjælland er de masker, der normalt anvendes i hydrologisk modellering, også taget i brug i den geostatistiske modellering til dannelse af hydrostratigrafien. Dette sikrer, at lagene får udbredelser, der er ensartede med DK-modellen og dermed skaber en vis konsistens med det eksisterende modelgrundlag.

3.2.2.4 Anholt og Læsø

Der er ikke anvendt lagspecifikke masker for Anholt og Læsø. Dette skyldes dels, at der i den eksisterende model ikke er en lokal udbredelse, der kræver særlig håndtering, og dels, at introduktion af sådanne masker ikke ville give en væsentlig beregningsmæssig gevinst, da øerne har en begrænset størrelse.

3.3 Udtagning af punkter

FOHM-modellen og resten af DK-modellen indeholder sammensatte modeller af mindre regionale og lokale modeller, samt modeller hvor der er brugt forskellig tolknings- og interpolationsstrategi. Dette er eksemplificeret i Figur 3-7, der viser et lille udsnit af "tolkningspunkter", der ligger i et af de kvartære lag omkring Frederikshavn. Dette udsnit er udtaget helt tilfældigt for at illustrere problemet, som er gældende for alle lag og i hele modelområdet. Fordelingen af tolkningspunkter antyder, at der er brugt forskellige tolkningsstrategier i de oprindelige modeller, der ligger til grund for FOHM-modellen: Både for områderne markeret med A og B tyder det på, at tolkningspunkterne er sat tæt langs tolkningsprofiler. Det er samtidig tydeligt, at tolkningsprofilerne er fordelt med den største gennemsnitlige afstand i område A, hvilket forårsager en meget ujævn fordeling af tolkningspunkterne i dette område. I området markeret med C ligger tolkningspunkterne ret spredt, men relativt jævnt fordelt. Disse punkter har sandsynligvis været tolket i relevante boringer. I område D er der tolket meget tæt, mens der også findes områder som E, hvor de originale tolkningspunkter er mistet, og hele modellen derfor er indlæst i et regulært grid. Slutteligt ses stedvist spredte punkter mellem originale tolkningsområder (F), der primært har været tolket som støttepunkter. Eftersom punkterne derfor ikke repræsenterer det samme, er det derfor nødvendigt at antage en punktudtagningsstrategi, der sikrer en mere homogen punktfordeling.



Figur 3-7: Udsnit af punkter for '1400 kvartært sand' i FOHM-modellen. Kortudsnittet dækker et område af Nordjylland tæt ved Frederikshavn.

3.3.1 Udtagningsstrategien

I stedet for at bruge tolkningspunkter som basis for den geostatistiske beregning, benyttes punkter fra fladerne, idet disse repræsenterer:

- 1. **Ensartede egenskaber** (at de alle er punkter fra den samme model). Som kontrast repræsenterer tolkningspunkter både gridværdier, reelle tolkningspunkter, lokale støttepunkter, interpolationspunkter, regionale støttepunkter, m.fl.
- 2. Den maksimale geologiske tolkningsinformation. Tolkningspunkterne i sig selv repræsenterer ikke den samlede tolkning. Det gør modelfladen derimod. Altså har tolkningspunkterne kun en delvis mængde information, al anden information har tolkeren/tolkerne lagt i interpolationen.

For at udvælge punkter i modelfladerne kan man udtage dem uniformt i rummet med en vis sandsynlighed. Alternativt kan man anvende en ikke-uniform rumlig fordeling. En sådan vægtet tilgang tillader, at antallet af udvalgte punkter kan reduceres (og dermed mindskes påvirkningen fra den oprindelige interpolation), samtidig med at informative punkter bibeholdes til beskrivelse af signifikante strukturer. Den informerede punktudvælgelse, der anvendes i dette projekt, er baseret på INPOX-metoden (Madsen et al. 2024) og beskrives i de følgende paragraffer.

3.3.2 INPOX

Vægtning af punkter skal styres af et princip. Vi anvender et princip om at prioritere områder med særlige karakteristika. Antagelsen i INPOX-metoden er, at informationsmængden er størst i områder, hvor fladen ændrer karakter og højde over kortere afstande (f.eks. ved en begravet dal). Begrundelsen er, at disse strukturer sjældent bliver modelleret uden et tilsvarende datagrundlag, hvilket antyder, at de repræsenterer reel information i fladen.



Figur 3-8: Udtagning af punkter enten med uniform fordeling (venstre) eller vægtet fordeling (højre). Den stiplede linje er en interpolation.

Som det er illustreret øverst i Figur 3-8 kan seks punkter tages ud enten med en uniform udtagning eller en vægtet udtagning. I det sidste tilfælde bliver en lineær interpolation efterfølgende væsentlig bedre til at repræsentere originalfladen. Punkterne har således arvet mere informationsmængde om fladens forløb end det er tilfældet for den uniforme fordeling af punkter.

Det samme gør sig også gældende når antallet af punkter bliver mindre, som det ses nederst i Figur 3-8. Den uniforme fordelings evne til at beskrive informationen i fladen skalerer mere lineært med antallet af punkter, der tages ud, hvorimod den vægtede fordeling kan opretholde en rimelig repræsentation af fladen selv for få antal punkter, da de enkelte punkter er mere informative. Altså, for at udtage så meget information fra fladerne som muligt uden at overtage interpolationsinformation er det derfor vigtigt at prioritere hvor stor en mængde punkter, der skal udtages samt hvilken metode, der skal informere om lokale ændringer i punktudtagningen.

3.3.2.1 Valg af operator

Baseret på koten ved stedkoordinaten z(x, y) er der opstillet tre forskellige punktudtagningsoperatorer til testning af den mest informative metode. Disse tre er: den første afledte til stedkoordinaten $\|\nabla z(x, y)\|$, den anden afledte til stedkoordinaten $\|\nabla (\|\nabla z(x, y)\|)\|$ og Laplace-operatoren $\Delta = \nabla^2$ på stedkoordinaten. Udtryk for de forskellige operatorer er vist i Tabel 3-1.

Figur 3-9 illustrerer, hvordan Laplace-operatoren (nederste række) kan udvælge punkter, der indeholder mere information end ved en uniform udtagning (øverste række), den første afledte i forhold til stedkoordinaten (næstøverste række), eller den anden afledte i forhold til stedkoordinaten (næstnederste række). I dette eksempel er der taget lige mange punkter ud fra den eksisterende flade og derefter lineært interpoleret som illustreret i Figur 3-8. Den relative fejl (RMSE) er markant lavere for Laplace-operatoren, mens den største gennem-snitlige fejl opstår ved en uniform punktudtagning (RMSE = 2,19 m). Fejlen kan dog reduce-res med næsten en halv meter i gennemsnit ved at anvende punkter udtaget med Laplace-operatoren. Alle RMSE for Figur 3-9 er vist i Tabel 3-1. Især fremstår signifikante geologiske strukturer bedre opløst.

På baggrund af dette og yderligere forsøg udført på andre lokaliteter og i andre lag, er Laplace-operatoren valgt til at guide punktudtagningen, da den konsekvent leverer mest information, så længe punktantallet ikke bliver for lavt. I så fald er en uniform punktudtagning at foretrække. I projektet er der taget udgangspunkt i punktætheden i den eksisterende FOHMmodel til at fastsætte antallet af punkter, der skal udtages. Da denne tæthed ligger langt over tærsklen for, hvornår den informerede punktudtagning er relevant, kræver dette ikke yderligere håndtering.

Operator	Udtryk	Række i Figur 3-9	RMSE i Figur 3-9
1. ordens afledte	$\ \nabla z(x,y)\ = \sqrt{\left\ \frac{\partial z(x,y)}{\partial x} + \frac{\partial z(x,y)}{\partial y}\right\ }$	Næst- øverst	2,01m
2. ordens afledte	$\ \nabla(\ \nabla z(x,y)\)\ = \sqrt{\left\ \frac{\partial\ \nabla z(x,y)\ }{\partial x} + \frac{\partial\ \nabla z(x,y)\ }{\partial y}\right\ }$	Næstne- derst	1,99m
Laplace	$\nabla^2 z(x,y) = \nabla \cdot \nabla z(x,y) = \frac{\partial^2 z(x,y)}{\partial x^2} + \frac{\partial^2 z(x,y)}{\partial y^2}$	Nederst	1,72m

Tabel 3-1: Operatorer til vægtet udtagning af punkter



Figur 3-9: Eksempel fra Jylland, hvor Laplace operatoren viser sig at indeholde mest information ift. en udtagning baseret på enten uniform, første afledte til stedkoordinaten og anden afledte af stedkoordinaten.

3.3.2.2 Oversættelsesfunktion mellem sandsynligheder og Laplace

Den rumlige fordeling af Laplace, som er vist i Figur 3-10 med rød ovenpå koten til KS3, er ikke i sig selv nok til punktudtagning. Der skal vælges hvor stor informationsmængde Δ indeholder og om den er værd at bibeholde – en oversættelsesfunktion. Vi antager en basisudtagning af punkter, der skaleres op som funktion af stigende Laplace. På denne måde fås en større sandsynlighed for at trække værdier ved større Laplace.

Der kan dog også opstå områder med meget høje Laplace, som relaterer sig til interpolationsartefakter fra den oprindelige flade. Dette er visualiseret i Figur 3-11. Ved interpolationsartefakter bliver Laplace ekstremt høj og meget lokalt forankret (se evt. $\nabla^2 > 5$ i Figur 3-10). Denne effekt ses primært langs kystzonen og i nogle runde cirkler omkring punkter, der svarer til den radius, der er brugt til interpolationen). I venstre side af Figur 3-11 (gradienten i stedkoordinaten) ses denne interpolationsradius som "spaghetti-lignende" strukturer, der tydeligt fremstår som utilsigtede interpolationsartefakter.

For at undertrykke disse interpolationsartefakter er der, som vist i Figur 3-12, behov for at vælge et forløb af en kurve, der giver en øget vægtning af høje ∇^2 , hvilket regnes for at være geologisk signifikante og som samtidigt kan undertrykke uønskede interpolationsartefakter. Da $\nabla^2 < 3$ i fladen typisk indeholder reelle strukturer, $\nabla^2 < 5$ er en blanding af artefakter og strukturer og $\nabla^2 > 5$ nærmest udelukkende består af artefakter, er den viste vægtning brugt (blå kurve i Figur 3-12).



Figur 3-10: Illustration af den rumlige fordeling af Laplace for stedkoordinaten for KS3 i Jylland for forskellige bins af ∇^2 .



Figur 3-11: Eksempel på interpolationsartefakter i en flade fra FOHM. Nederste højre hjørne er tydeligt påvirket af artefakter.

Funktionens analytiske definition ses på Figur 3-13. Der sættes en basisudtagning P₀, der sikrer en vis udtagning af punkter uanset det strukturelle input fra laplace-værdierne. Denne basissandsynlighed bliver derefter manipuleret gennem en funktion, der er afhængig af Laplace-værdien. Funktionen består af summen af to separate dele: funktion 1 (cyan område) og funktion 2 (lilla område). Funktion 1 øger basisudtagningen op til et maksimum på P₀ + C

og gælder for ∇^2 < B. Ved ∇^2 >= B tager funktion 2 over og reducerer punktudtagningen asymptotisk fra P₀+C til P₀-D. B markerer dermed den Laplace-værdi, hvor udtagningen prioriteres mest.



Figur 3-12: Funktioner til udtagning af punkter, den røde linje angiver, at alle punkter har sandsynlighed 1/8 for at blive udtaget og den blå linje angiver, hvordan en udtagning kan vægtes med laplacian værdierne i FOHM fladen. Ved analyse af Laplace er det bestemt at naturlige strukturer typisk har værdier under 3 imens artefakter fra interpolation sagtens kan nå op over 20.



Figur 3-13: Til venstre ses funktionen til den vægtede udtagning af punkter skitseret og til højre ses det analytiske udtryk for funktionen.

3.3.2.3 Punktudtagning fra sandsynlighedsgrids

Gennem ovenstående procedure kan der genereres et sandsynlighedsgrid for punktudtagning for samtlige flader (se Figur 3-14a for en simpel illustration). Vha. dette grid kan der udtrækkes en punktserie, hvor der lægges vægt på geologiske features og nedprioriteres interpolationsartefakter. Udover sandsynligheder fra fladens forløb introduceres også (faste) punkter til det grid, der skal udtages. Dette kan opnås ved at sætte en sandsynlighed på 1 i griddet.

Når de endelige sandsynlighedsgrids er konstrueret, udtages der punkter for alle lagflader. Dette gøres ved at trække et tilfældigt tal fra en ukorreleret uniform fordeling mellem 0 og 1 for hver gridcelle. Sandsynligheden fra sandsynlighedsgriddet fungerer som en tærskelværdi for, om et punkt skal udtages eller ej. Figur 3-14b illustrerer, hvordan der for alle gridceller er trukket et tilfældigt tal mellem 0 og 1. Hvis et tal er mindre end den tilsvarende værdi i sandsynlighedsgriddet i Figur 3-14a, udtages punktet (gule celler).



Figur 3-14: Illustration i princippet i konvertering fra sandsynlighedsgrid til udtrukne punkter i et grid. Til venstre (a) er sandsynlighedsgriddet dannet. Den gennemsnitlige punktudtagning med dette grid vil være 0.35. Til højre er der trukket tilfældige tal fra en uniform fordeling mellem 0 og 1. Alle gule celler har en tilfældig værdi under gridcellesandsynligheden og udtages (gul). Forholdet mellem gule celler (6) og det samlede antal celler (16) er 0.38.

3.3.3 Opsætning af sandsynlighedsgrids

Den samme oversættelsesfunktion blev valgt for samtlige lag i modellen, da det viste sig for omstændigt at vælge specifikke udtagnings-sandsynligheder for alle lag. Der er derfor anvendt en konservativ oversættelsesfunktion.

Udgangspunktet var en punktudtrækning med en basisværdi P₀=0,2, mens sandsynligheden blev hævet til cirka 0,45 i intervallet mellem 5 og 12. Erfaringen viser, at Laplace over 12 i FOHM-fladerne ofte er forbundet med simuleringsartefakter. Derfor blev sandsynligheden reduceret til 0,15 for disse højere Laplace for at bevare en konservativ tilgang, der ikke fuldstændigt udelukker at høje Laplace kan være geologisk relevante. Ulempen ved en konservativ punktudtrækning er, at den ikke er lige så effektiv til at fjerne interpolationsartefakter for de enkelte lag, som en mere lagspecifik opsætning ville være.

Udover punktudtrækningen, der baserede sig på fladernes forløb, blev sandsynligheden sat til 1, når modelpunkter faldt sammen med en laggrænse i et boreprofil, da disse punkter blev vurderet som de mest pålidelige i modellen. De originale tolkningspunkter, der var 'snappet' til en boring fra den oprindelige model, blev medtaget som faste punkter. Dette sikrede, at punkter, hvor geologen havde været mest sikker på placeringen, blev prioriteret, og at laggrænser fra gode boringer blev anvendt direkte. Endelig, for at sikre udtrækning af punkter i flade områder, blev sandsynligheden sat til 1 for hver 1 km i begge retninger.

På Sjælland blev en række geologiske features oprindeligt trykket ned i DK-modellen, hvilket resulterede i meget skarpe gradienter (og høje Laplace), som havde geologisk relevans, men som naturligt blev nedprioriteret af algoritmen, da den var designet til at anse disse højere Laplace-værdier som interpolationsartefakter. Da disse rent faktisk relaterede sig til geologisk information, var ønsket at bevare den i modellen. Elementerne var tilgængelige via shapefiler i DK-modellen og blev konverteret til enten faste punkter (med en sandsynlighed på 1) eller meget høje sandsynligheder for at sikre, at informationen blev bibeholdt.

3.4 Estimering af en ikke-stationær zoneret statistisk model

Fra de udtrukne punkter kan der beregnes en korrelationslængde og en varians baseret på Gaussisk statistik, altså en kovariansbaseret tilgang. Normalt bestemmes kovariansmodellen ved at fitte et semivariogram til punkterne, enten manuelt eller gennem en automatisk procedure, hvor fittet optimeres via en misfit-funktion. På landsplan, hvor fladens opførsel varierer betydeligt, og der derfor kræves mange statistiske kovariansmodeller, er begge tilgange imidlertid for tidskrævende. Manuel estimering er upraktisk i så stort et omfang, og automatisk optimering bliver særligt tidskrævende, når mange punkter bruges til at opbygge semivariogrammerne. Hvis antallet af punkter reduceres for at øge effektiviteten, kan det medføre ustabilitet i visse regioner og føre til bias i korrelationslængderne. I værste tilfælde kan dette kræve en manuel efterbehandling, hvilket yderligere øger kompleksiteten.

3.4.1 Machine Learning-model til prædiktion af kovariansmodeller

Som vist i Falk og Madsen (2023) kan der i stedet estimeres en kovariansmodel vha. superviseret Machine Learning (ML). Specifikt benyttes et "convolutional neural network" eller et foldnings-neuralt netværk på dansk. Superviseret læring baserer sig på, at der eksisterer en række kendte eksempler på både input og output, hvor der kan estimeres en model, der beskriver dette forhold. I dette projekt er netværket trænet til at gå fra et sæt punkter til en kovariansmodel, som skitseret med den sorte pil i Figur 3-15. Det er muligt at generere en række realisationer fra en kovarians model vha. kendte metoder såsom Cholesky-dekomposition og Fast Fourier Transform Moving Average (blå pil). Dette giver i princippet mulighed for at træne på et uendeligt stort datasæt af input (punkter) og et tilhørende svar (kovariansmodel).



Figur 3-15: Skematisk opsætning af arbejdsgangen til estimering af ML-modellen

3.4.1.1 Implementering og træning

For at sikre, at der kan findes mange forskellige kovariansmodeller, genereres der punkter fra en suite af forskellige kovariansmodeller. I Figur 3-16 er der vist 16 eksempler på, hvordan der er udtrukket punkter fra vidt forskellige kovariansmodeller. Som det ses, varierer realisationerne i korrelationslængde (længere eller kortere mellem skift i farver) og varians (farvespektret). Der tages punkter vilkårligt ud efter det valgte punktudtagningsprincip skitseret i foregående afsnit. Dette sikrer den bedst mulige overensstemmelse mellem det, som netværket er trænet på, og det der kommer fra den hydrostratigrafiske flade.



Figur 3-16: 16 eksempler på Gaussiske fordelinger med støj. Disse er vist for et 4 km x 4 km grid.

I alt er netværket trænet på 150.000 kombinationer af punktudtagning og unikke kovariansmodeller. Valget på 150.000 kommer efter en række forsøg, der viste, at der efter denne grænse opnås en asymptotisk opførsel i evnen til at træne et bedre netværk, selvom der tilføjes flere kombinationer. Inputpunkterne til det neurale netværk har en middelværdi på 0 for at sikre, at der ikke er højdeafhængighed i det trænede netværk. Dette skal forstås som at netværket skal kunne bruges alle steder i FOHM-modellen og ikke kun i områder, der har én bestemt kote. Desuden er der lagt støj på punkterne med samme korrelationslængde som i den valgte kovariansmodel. Niveauet af støjen trækkes tilfældigt mellem 0 og 1 gange amplituden af det originale punktsæt. På denne måde bliver netværket mere upræcist, men kan samtidig også bedre honorere korrelerede fejl i fladen. Det neurale netværks arkitektur under træningen er 69 lag med 76 forbindelser. Der er krydsvalidering under træningen af netværket, hvor 10% af data er udtaget til valideringen.

3.4.1.2 Machine Learning-modellens nøjagtighed og præcision

Figur 3-17 viser ML-modellens præcision gennem et krydsplot af de faktiske værdier for korrelationslængde (øverst) og sill (nederst) sammenlignet med de prædikterede værdier. For korrelationslængderne ses det tydeligt, at de prædikterede værdier (røde prikker) ligger jævnt fordelt omkring ideallinjen, hvilket indikerer, at der ikke er noget markant bias i modellen. Det kan derfor konkluderes, at ML-modellen præcist forudsiger korrelationslængderne. Til sammenligning er der også plottet et netværk, der er trænet på hele realisationen og ikke kun på punkter (blå prikker), som vist i Figur 3-16. At benytte hele fladen repræsenterer et optimalt scenarie, og som forventet bliver netværket en smule mere stabilt med den fulde information. Det er dog bemærkelsesværdigt, at præcisionen ikke falder markant, når der i stedet trænes på punkter frem for den fulde flade. Det kan således konkluderes, at ML-modellen er præcis, selv når den anvender punkter, og at størstedelen af informationen om fladens forløb bevares i de udtagne punkter, hvilket stemmer overens med forventningerne til den informative punktudtagningsstrategi.

Nederst i Figur 3-17 ses det, at sill-værdien er velbestemt mellem 0 og 2000 m². Ved højere sill er der en bias i prædiktionen mod lavere værdier end den faktisk værdi. For at undersøge dette nærmere er der trænet yderligere netværk med både et større og mindre sill-spektrum (ikke vist her). Resultaterne fra disse netværk viser en tilsvarende afvigelse fra ideallinjen efter cirka 2/3 af træningsintervallet og frem mod den maksimale sill-værdi. Dette indikerer, at fejlen ikke er knyttet til en specifik absolut sill-værdi, men snarere er en iboende begrænsning ved foldningsnetværk. Fejlen opstår sandsynligvis som følge af vanskeligheder med at håndtere kontrasterne mellem de laveste og højeste sill-værdier. Til yderligere sammenligning er sample-variansen for punkterne også plottet (grønne prikker). Dette plot viser fuld nøjagtighed, men samtidig et upræcist resultat, hvilket fremgår af den brede og jævne spredning omkring ideallinjen.

Der er også testet, om et netværk trænet udelukkende på korrelationslængder kunne erstatte det koblede netværk, hvor både sill og korrelationslængder estimeres. Resultaterne viste dog, at dette netværk er mindre præcist i sine prædiktioner sammenlignet med det koblede netværk. Det er ikke blevet testet, om information om støj kunne eliminere bias i sillprædiktionen. Dette skyldes, at støjen er skaleret relativt til den udtrukne sill-værdi, hvilket snarere forventes at øge spredningen omkring ideallinjen end at introducere bias.



Figur 3-17: Præcision og nøjagtighed af det neurale netværk. Øverst er den rigtige korrelationslængde (x-akse) plottet mod netværkets prædikterede korrelationslængde (y-akse). Ideallinjen er markeret med sort, de røde prikker repræsenterer det neurale netværk trænet på de udtrukne punkter, og de blå prikker viser et neuralt netværk trænet på hele fladen som sammenligning. Nederst er den rigtige sill (x-akse) plottet mod netværkets prædikterede sill (y-akse). Farvekodningen er den samme som i det øverste plot, med tilføjelse af grønne punkter, der angiver variansen af punkterne som reference.

Vi konkluderer samlet set, at bias i sill-værdierne er uacceptabel og vanskelig at korrigere for. Det endelige netværk er derfor trænet til at prædiktere både korrelationslængder og sill, men kun korrelationslængderne anvendes i den efterfølgende modellering og clustering af ensartede områder. Sample-variansen anvendes som erstatning for sillen, da det er vigtigere at opnå en nøjagtig sill-værdi. Denne løsning har dog den ulempe, at variansen er mere upræcis end sill-prædiktionen.

Når netværket var trænet, blev det brugt på fladen til at estimere kovariansmodellens parametre (range/"korrelationslængde" og sill/"varians"). Et gridvindue på 31 x 31 blev designet og kørt over fladen som et "sliding window". Middelværdien for punkterne inden for vinduet blev beregnet og fratrukket punkterne. På denne måde blev det sikret, at ML-modellen kunne anvendes overalt, da den var trænet til en middelværdi på nul. Outputtet fra at køre vinduet over hele fladen blev tre grids: et med korrelationslængde, et med sill og et med alle middelværdier. Da gridvinduet på grund af sin firkantede form og størrelse på 3100 x 3100 meter ikke kunne dække kystområderne, blev værdier ekstrapoleret ud fra kysten ved hjælp af "nearest neighbour"-metoden.

3.4.2 Clustering

Efter at hver flade i de hydrostratigrafiske lagmodeller er konverteret til et grid med range, sill og middelværdier er der behov for at zonere (clustre) områder af fladen, der har relativt ens statistisk. Figur 3-18 viser et eksempel på, hvordan der kan findes zoner for, hvordan fladen opfører sig lokalt. Da den prædikterede sill ikke var pålidelig (se foregående afsnit) er der clusteret på korrelationslængden (Figur 3-18a). Til denne clustering er der brugt ML-teknikken "Self-organizing map". Modsat ML-modellen der var brugt til at estimere kovariansmodellen som baserede sig på træningsdata, hvor svaret var kendt, altså superviseret læring, er et self-organizing map en læringsteknik, hvor svaret ikke er givet under træningen – altså ikke-superviseret læring. Da vi ikke på forhånd kan sige noget konkret om sammenhængen mellem den rumlige udbredelse af de enkelte clustre, er det oplagt at benytte en ikke-superviseret metode. Valget på self-organizing maps skyldes, at mange andre gængse clustering-algoritmer, såsom k-means, ikke kan håndtere den rumlige information, der er essentiel, idet korrelationslængderne er fordelt geografisk. Der er brugt implementeringen af selv-organizing maps fra Matlab til denne øvelse.



Figur 3-18: Eksempel på hvordan der kan estimeres zoner, hvor en flade har de samme statistiske egenskaber (figur taget fra Falk og Madsen (2023)). a) Et oversigtskort over den korrelationslængde der er fundet med ML-modellen, b) zoner hvor korrelationslængde opfører sig ensartet. Den sorte streg på a) og b) angiver et tværsnit der er afbilledet på c). Farverne viser de enkelte zoner som fladen er inddelt i og ovenpå er vist fladens forløb og den estimerede korrelationslængde.



Figur 3-19: Til venstre: Eksempel på clustering-outputtet fra self-organizing maps på en flade fra Sjællandsmodellen. Til højre: Clustering-resultatet efter den udførte post-processering, hvor der er tilskrevet mode-clusters og potentielle store clustre er splittet i mindre områder.

Figur 3-19 viser et eksempel på outputtet fra clustering-algoritmen. Som det ses på den venstre figur, har self-organizing maps algoritmen en tendens til producere nogle små og usammenhængende clustre. I visse tilfælde kan der endda være en enkelt gridcelle, der er tildelt et separat cluster indeks inde i et større område, der ellers er defineret til at tilhøre et andet cluster. Dette er upraktisk for den videre modellering, idet det betyder, at disse mini-clustre skal behandles separat og syes sammen med resten af modellen. Derfor er der indført en post-processeringsrutine til at lægge nogle af disse mini-clustre ind under nogle af de større clustre. Fremgangsmåden var, at der blev konstrueret et cirkulært sliding-window med en bestemt søgeradius (defineret i gridceller). Inden for dette vindue blev der beregnet det mest forekommende cluster indeks – et såkaldt "mode-cluster indeks". Til højre i Figur 3-19 ses, hvordan clustering-resultatet ser ud efter postprocessering – altså et kort der viser modecluster indeks. Det er tydeligt, at postprocesseringen resulterer i blødere og bedre afgrænsede zoner end det, der kommer direkte ud af self-organizing maps. Mængden af mini-clustre er samtidig reduceret markant. Valget af søgeradius og dermed vinduets størrelse har afgørende betydning for mængden af både blødheden i resultatet og størrelsen på mini-clustre, der bibeholdes i resultatet. Som vist i Figur 3-19 valgte vi en søgeradius på 4 celler for at komme de allermindste clustre til livs uden at gribe for meget ind i informationen, der kommer direkte ud af self-organizing maps. Derudover blev der indført en automatisk opdeling af clustre over en bestemt tærskelværdi i antallet af gridceller. Dette er gjort af numeriske årsager, da tidsforbruget af den efterfølgende modellering skalerer med mindst n², hvor n er antallet af gridceller i et givent cluster. Denne tærskelværdi er sat til 8000 gridceller. Hvis et cluster er større end 8000 gridceller deles det i to områder med det samme antal gridceller, baseret på den euklidiske afstand. Hvis disse nye områder stadig ikke overholder tærskelværdien, fortsættes proceduren rekursivt indtil ingen af de nye clustre overskrider tærskelværdien. Disse opdaterede clustre tilskrives unikke cluster-indekser, så der efterfølgende kan tilskrives den samme statistik til opsplittede clustre.



Figur 3-20: Eksempel på den endelige clustering af 1400 Kvartært Sand Bund i Jylland.



Figur 3-21: Sammenhæng mellem lagareal, antal unikke clustre og gennemsnitlig clusterstørrelse for tre regioner i Danmark (Jylland, Fyn og Sjælland). a) Antal unikke clustre som funktion af det samlede lagareal viser en lineær sammenhæng, hvor hældningen ikke varierer nævneværdigt mellem regionerne. b) Gennemsnitlig clusterstørrelse som funktion af lagdybde viser en konstant størrelse i Jylland, mens der ses en stigning med dybden for Fyn og Sjælland. De ny-modellerede lag på Fyn og i Jylland er ikke taget med.

Figur 3-21 illustrerer clusterstørrelser og -antal for alle lag, der ikke er ny-modellerede på tværs af de tre regioner Jylland, Fyn og Sjælland. Generelt set afhænger clusterdannelsen af lagenes variabilitet, hvor større variabilitet kræver flere clustre. I Figur 3-21a ses en tydelig lineær sammenhæng mellem antallet af unikke clustre og lagarealet, hvor de tre regioner følger næsten identiske tendenser. Dette indikerer, at dannelsen af nye clustre med en unik rumlig variation er proportional med lagets areal, uafhængigt af region. Med andre ord bety-der større lagudbredelse, at der er flere typer rumlig variabilitet, der skal beskrives.

Da datatætheden og opløsningsevnen generelt falder med dybden, bliver lagene ofte mere glatte og ensartede, hvilket resulterer i større clustre. Denne tendens ses tydeligt i Figur 3-21b på Fyn og Sjælland, hvor resultaterne flugter med forventningen om større clustre i dybere lag. For Jylland stiger clusterstørrelsen også som funktion af lagets kote, men med en lavere rate sammenlignet med Fyn og Sjælland. Forskellen kan skyldes, at nogle af lagene i Jylland ligger væsentligt dybere, og vi kunne forvente, at Fyn og Sjælland ville minde mere om Jylland, hvis de havde lag med tilsvarende dybder. Jyllands resultater påvirkes sandsynligvis også af den absolutte tærskelværdi på 5000 celler pr. cluster, som især har betydning for store lag med stor udbredelse. Da Jylland har et større modelareal, har tærskelværdien haft en større effekt i Jylland end på Fyn og Sjælland. Havde tærskelværdien været dynamisk og skaleret efter regionernes areal, ville resultatet formentlig have været mere ensartet. Alt i alt viser Figur 3-21, at clustering-delen er udført tilfredsstillende og generelt stemmer overens med vores forventninger.

3.5 Geostatistisk modellering og generering af hydrostratigrafiske modeller

3.5.1 Kriging- og simuleringsstrategi

Figur 3-22 viser et eksempel på, hvordan standardafvigelser og range-værdier fordeler sig for 9 forskellige clustre. Som det ses i figuren, kan der være stor variation i hvilke korrelationer og varianser, der gælder inden for de enkelte områder af lagfladen. Dette viser, hvorfor det er vigtigt at underinddele lagfladen i mindre områder. Hvert cluster tilskrives en repræsentativ sill (i dette tilfælde repræsenteret ved punktvariansen) og range, der fortæller hvilken statistisk model, der er gældende inden for dette område.



Figur 3-22: Fordelinger af standardafvigelser og range-værdier for 9 forskellige clustre. Middelværdien er vist med en tynd streg

I dette projekt er hvert område tilskrevet middelværdien for korrelationslængden og 75 percentilen for standardafvigelsen. Dette er gjort ud fra et konservativt princip for at sikre, at variationen på fladen ikke underestimeres. Hvis variationen overestimeres vil tolkningsusikkerheden alligevel være den dominerende kilde til usikkerhed og den kombinerede usikkerhed lav. Hvis tolkningsusikkerheden derimod er høj, er det et problem, hvis variationen i området er for lille, da den kombinerede usikkerhed igen vil blive lav.

For et givet område i lagfladen er proceduren følgende, når der er tilskrevet en sill og range: Tolkningspunkterne trækkes ud af fladen og fratrækkes ML-gennemsnittet (d_{obs}) og tilskrives en statistisk model på baggrund af clustrene (C_M). Udenom hvert individuelt element lægges en buffer på 200 m. Alle elementer for fladerne gennemgås én efter én, og der laves en lokal kriging, hvor der findes en posterior kovarians:

og bias/mean-vekor ($\widehat{\mathbf{m}}$):

$$\widehat{\mathbf{m}} = \mathbf{C}_{\mathrm{M}}\mathbf{G}^{\mathrm{T}}(\mathbf{G}\mathbf{C}_{\mathrm{M}}\mathbf{G}^{\mathrm{T}} + \mathbf{C}_{\mathrm{d}})^{-1}\mathbf{d}_{\mathrm{obs}}$$

 $\hat{\mathbf{C}}_{\mathrm{M}} = (\mathbf{G}^{\mathrm{T}} {\mathbf{C}_{\mathrm{d}}}^{-1} \mathbf{G} + \mathbf{C}_{\mathrm{M}})^{-1}$

G er en forward operator, der går mellem punktrummet og modelrummet. Da data i dette tilfælde er lagenes dybde ligesom modelrummet, er **G** en binær matrice med ettaller, hvor tolkningspunkter og modelgriddet er sammenfaldende. **C**_d er lavet vha. usikkerhedstemaet for de punktdata, der er udtaget. Halvdelen af støjen er antaget at være korreleret med samme korrelationslængde som i **C**_M, mens resten af usikkerheden er antaget af være ukorreleret. Vha. \hat{C}_M og \hat{m} kan der udtrækkes lige mange realisationer inden for området uden nævneværdigt tidsforburg. Tidsforbruget er koblet til selve kriging-delen og beregningen af \hat{C}_M , da det kræver invertering af både **C**_M og **C**_d. I dette projekt er der lavet 100 realisationer inden for hvert område. Middelmodellen lægges til alle realisationer for at flytte dem til den korrekte dybde.

3.5.2 Sammensmeltning

Hvert område, der er simuleret med 100 realisationer, har behov for at blive sammensat med realisationerne, for de andre områder. Til dette formål benyttes bufferzonerne rundt om de individuelle elementer. Inden for disse bufferzoner vægtes bidragene fra de forskellige elementer som funktion af afstanden til elementet. Sammensmeltning er lineær som funktion af den euklidiske afstand til simuleringsområderne. Til sidst normeres alle bidrag til én for at skabe en sammenhængende flade. Kriging, generering af realisationer og sammensmeltning for én flade på Fyn tager ca. 7-8 minutter på en kraftig workstation. På Sjælland tager processen ca. 15 minutter, mens den i Jylland (inkl. Endelave, Tunø og Samsø) typisk tager 45 minutter. I nogle tilfælde kan det dog tage op til halvanden time i Jylland, afhængigt af de individuelle lags udbredelse. Anholt og Læsø sammensmeltes ikke, da de betragtes som hele områder. For disse områder tager kriging og generering af realisationer kun få sekunder.

3.5.3 Lagfladekorrektion

Når der er dannet det ønskede antal sammensatte realisationer af lagfladerne, er der behov for to forskellige lagkorrektioner. Den første korrektion skyldes, at der i nogle boringer er truffet et aktivt valg fra den tolkende geolog om, at et eller flere lag ikke eksisterer. Et eksempel kunne f.eks. være et lerlag, der ses i én boring og ikke i en anden et stykke derfra. Antagelsen er således i tolkningsøjemed, at laget forsvinder mellem de to boringer. Denne information er normalt indlejret i tolkningen ved at omkringliggende sandlag er "snappet" i samme dybde ved boringen. Hvis dette er tilfældet, lader vi dette punkt markere et nultykkelsespunkt af lerlaget. I de realisationer, hvor der er opstået en tykkelse for en realisation i et nultykkelsespunkt, føres det nederste sandlag op til bunden af det overliggende sandlag for at gennemtvinge nultykkelser i fladerne. Væk fra dette punkt laves en vægtet overgang mod de oprindelige realisationer, så nultykkelsen ikke honoreres som et spike midt i en flade.

Den anden nødvendige korrektion tager højde for, at lag kan være overlappende. Dette problem opstår, fordi fladerne er genereret uafhængigt af hinanden. Der opstilles nogle regler for hvilke lag, der "eroderer" eller "onlapper" på hinanden som vist på Figur 3-23. Her har vi implementeret et workflow, der starter i terræn og samler lagene med stigende dybde/alder konsistent med lagenes rækkefølge i FOHM-modellen. Ved det første lag under terræn (lagflade B) vil områder med overlap blive naturligt skåret af terrænet (lagflade A) efter princippet øverst i Figur 3-23. Alle lag bliver derefter behandlet efter samme princip indtil bunden af modellen nås. Det vil altid være udbredelsen af det overliggende lag, der dikterer udbredelsen af det underliggende lag. Det er dog simpelt manuelt at gå ind at udvælge specifikke lag, hvor man ønsker en "onlap" sammenhæng som vist nederst i Figur 3-23, hvis dette passer bedre med den geologiske forståelse. Dette er gjort for den prækvartære flade (Top PreQ adjusted0).


Figur 3-23: Konceptuel oversigt over lagfladeinteraktion. Øverst "eroderer" grænse A og nederst "onlapper" grænse A.

3.6 Algoritme

Den samlede algoritme er kodet i Matlab og er i grove træk skitseret således:

- 1. **Initial opsætning:** Forskellige inputparametre sættes, der styrer eksekveringen, såsom den region, der skal simuleres, forskellige flag for præprocessering, antallet af realisationer samt andre indstillinger. Standardværdier sættes for manglende parametre.
- Regionsinitialisering: Baseret på den valgte region, initialiserer koden relevante variabler og forbereder systemet til behandling. Den indlæser også boredata for den valgte region.
- 3. Præprocessering:
 - a) Læsning af sen- og postglaciale lag: Scriptet læser og behandler sen- og postglaciale lag, trykker dem ned i de eksisterende flader og gemmer dem i filer, der fungerer som kalibreringsmodel.
 - b) Præprocessering af Lag: Præprocessering involverer at forberede lag til Kriging. Her køres ML-modellen til at prædiktere den statistiske model for hver lagflade. Derefter laves opdeling af lagfladen gennem clustering af områder med lignende statistiske egenskaber. Scriptet håndterer hvert lag for sig.

- c) Forberedelse af boringer: Koden forbereder boredata og skaber to datasæt: ét med "snappede" boringer (de, der matcher tolkningspunkter) og ét med ikkesnappede boringer. Information om boringsdybde, kvalitet og andre egenskaber gemmes, og boringerne organiseres til senere brug i modellen.
- d) Initialisering af usikkerhedstemaer: Matricen wellthemes initialiseres til lagring af brøndusikkerheder på tværs af forskellige lag (*Nlay*). En correction_maps-matrix initialiseres også for at håndtere "tie"-korrektion, hvor simuleringer muligvis tvinges til at overholde nultykkelse på visse lokaliteter.
- e) Lag-specifik forbehandling: For hvert lag behandler koden boredataene. Hvis der er manglende eller ugyldige datapunkter (boringer uden for grid), fjernes disse. Derefter laves et usikkerhedstemakort på basis af boringerne. I kvartære lag bruges kompleksitetskortet til at styre udbredelsen af usikkerheden. I de dybere lag fikseres kompleksiteten.
- f) Generelle geofysiske temaer: Scriptet henter alle de meget generelle geofysiske data (SkyTEM, PACEP, MEP, Seismik) og genererer geofysiske usikkerhedstemaer baseret på usikkerhedskonceptet.
- g) Region-specifikke temaer: Forskellige regioner (f.eks. 'Jylland', 'Fyn', 'Anholt-Læsø') kan have specifikke geofysiske datakilder, som ikke findes over hele landet. Der konstrueres geofysiske usikkerhedstemaer for disse geofysiske metoder efter usikkerhedskonceptet.
- h) **Lokale sen- og postglaciale lag:** For Jylland og Fyn er der foretaget vurderinger af usikkerheden på de enkelte introducerede sen- og postglaciale lag under modelleringen. Disse usikkerheder omsættes til et usikkerhedstema for disse lag.
- i) **Clusterstatistik:** Histogrammer over range og varians for hvert cluster beregnes og én værdi for range og varians tilskrives hvert cluster.
- 4. **Kriging og simulering:** Kriging-processen starter med indlæsning af nødvendige variable og data fra præprocesseringen. For hvert lag udføres kriging og simulering gennem Cholesky dekomposition.

5. Lagbehandling:

- a) Realisationer integreres med baggrundsmodellen: Realisationerne fra kriging-processen integreres med kalibreringsmodellen uden for maskeområderne vha. en glidende overgang. Dette sikrer konsistens i den videre hydrologiske modellering.
- b) Anvendelse af ties ved absolutte nultykkelsespunkter: Hvis der af modelløren er sat punkter, hvor tykkelsen af geologiske lag er nul, samles laggrænser i realisationer, der har simuleret eksistens af dette lag ved disse punkter. Disse justeringer sikrer, at modellen opfører sig konsistent med den geologiske fortolkning.

- c) Lagfladekorrektion: Alle områder med overlap i lagflader i realisationer justeres gennem det valgte laghierarki.
- Samling af 3D modeller og eksport: Alle lag, der hører til en bestemt realisation, samles sammen til én 3D model og eksporteres i .XYZ format. Dette gøres for samtlige realisationer.

Koden sørger generelt for at gemme relevante resultater efter hvert trin i algoritmen. Dette gør det nemt at finde information om betydningen af de enkelte trin og giver samtidig fleksibilitet til at starte koden fra ethvert punkt. Efter hver større opgave (f.eks. præprocessering, kriging eller eksport) sendes automatiserede e-mails med statusopdateringer om processen. E-mails indeholder detaljer om, hvor resultaterne er gemt (både lokalt og på serveren), og eventuelle fejl eller mangler rapporteres også.

3.7 Resultater fra modelleringen

Der er udvalgt profiler til at repræsentere de store modelleringsområder: Jylland, Fyn og Sjælland. I Jylland er der valgt to profiler for at illustrere de geologiske forskelle mellem regionerne. Det ene profil løber nord-syd i Nordjylland og viser de ny-modellerede lag samt tykke kvartære aflejringer, der præger området. Det andet profil løber øst-vest i Sydvestjylland, hvor større mægtigheder af Miocæne lag dominerer. På Fyn repræsenterer profilet primært kvartære aflejringer, mens profilet fra Sjælland fremhæver flere mægtige kalkaflejringer, som her findes tættere på overfladen.



Figur 3-24: Placering af profiler i modelleringsområderne Jylland N-S (blå), Jylland V-Ø (rød), Fyn (gul) og Sjælland (lilla).

Figur 3-24 viser placeringen af de fire profiler, som er valgt som repræsentative tværsnit til visualisering af resultaterne fra de hydrostratigrafiske modeller (Figur 3-25 til Figur 3-28). Realisationerne for udvalgte lag er placeret ovenpå den underliggende hydrostratigrafi fra DK-modellen. De hvide linjer repræsenterer 25 realisationer, der viser forskellige mulige fortolkninger af laggrænserne. Profilet i Sydvestjylland, på Fyn og på Sjælland er delvist sammenfaldende med profilerne vist i afsnit 2.7. For alle fire profiler gælder det, at der anvendes en kraftig vertikal overhøjning i størrelsesordenen 100 til 150. Det er derfor vigtigt, at læseren af denne rapport har for øje, at forskellene mellem de enkelte realisationer, som vises i de følgende gennemgange af profilerne, fremstår stærkt overdrevne på grund af denne overhøjning.

Figur 3-25 viser et profil fra Fyn, der illustrerer usikkerhederne i modelleringen af følgende hydrostratigrafiske laggrænser: Toppen af KS1, Bunden af KS2 og prækvartærfladen. For at billedet ikke skulle blive for mudret, er der kun visualiseret 25 realisationer, selvom der i det endelige ensemble blev genereret 100 realisationer til at udspænde usikkerheden. Figuren viser, at usikkerheden generelt stiger med dybden som forventet. Dette fremgår også af Tabel 6, der viser procentdelen af gridceller med datainformation i forhold til de givne lags gennemsnitsdybder på Fyn.



Figur 3-25: Profil fra Fyn med 25 realisationer (hvide streger), hvor usikkerheden på tre udvalgte laggrænser (Toppen af KS1, Bunden af KS2 og prækvartærfladen) illustreres. Pilen markerer en position, hvor der er relativt lav usikkerhed på grund af en nærtliggende boring med information om laggrænserne. Overhøjningen er 93,4 og placeringen af profilet ses i Figur 3-24. UTMX-koordinater og højde i meter over havniveau bruges som akser.

Tabel 2: Tabel over gennemsnitlige dybder og procentdelen af gridceller der har datainformation fra Fyn.

Lag nummer	1	2	3	4	5	6	7	8	9	10	11	12
Gennemsnitlige	-2.9	-3.2	-6.9	-9.1	-14.6	-17.2	-27.0	-27.7	-39.3	-52.6	-61.3	-92.4
dybde af lag												
Procent af da-	79.6	77.0	62.4	62.8	57.1	54.8	51.6	50.3	27.5	39.0	36.9	16.6
tainformerede												
celler												

Variationerne indenfor hvert lag følger generelt de oprindelige laggrænser. Realisationernes korrelationslængder ser ud til at stemme overens med dem i den originale hydrostratigrafi, hvilket er vigtigt for at bevare det geologiske udtryk fra den oprindelige model i de enkelte realisationer. Samtidig er det tydeligt, at usikkerheden på laggrænserne varierer afhængigt af lokaliteten. For eksempel er usikkerheden lav omkring UTM X = 586000 (sort pil), hvor en nærliggende boring giver høj sikkerhed på laggrænserne. Omvendt ses længere mod vest, hvor datadækningen er lavere, en markant større variation i realisationerne. Dette er særligt tydeligt ved grænsen til de prækvartære lag, hvor variationen kan være op til 30 meter.

Figur 3-26 viser et profil fra Sjælland med 25 realisationer, der illustrerer usikkerhederne i modelleringen af følgende hydrostratigrafiske laggrænser: Toppen af KS1, Bunden af KS3 og bunden af Grønsandskalken. Disse realisationer er placeret ovenpå den underliggende hydrostratigrafi fra DK-modellen. De hvide linjer repræsenterer realisationerne, der viser forskellige mulige fortolkninger af laggrænserne. Som på Fyn ses det, at usikkerheden generelt er dybdeafhængig, hvor de to kvartære laggrænser har markant lavere usikkerhed end grænsen mellem de to kalkaflejringer. Dog er usikkerheden mindre for kalkfladen i den østlige del af modellen, hvor kalken findes tættere på overfladen.



Figur 3-26: Profil fra Sjælland med 25 realisationer (hvide streger), hvor usikkerheden på tre udvalgte laggrænser (Toppen af KS1, Bunden af KS3 og bunden af Grønsandskalken) illustreres. Overhøjningen er 107,6 og placeringen af profilet ses i Figur 3-24. UTMX-koordinater og højde i meter over havniveau bruges som akser.

Generelt følger usikkerheden den originale hydrostratigrafi, men mindre afvigelser kan observeres enkelte steder. Disse afvigelser skyldes to faktorer: For det første opstår de i områder med høj datausikkerhed, som indgår i den geostatistiske modellering. For det andet er der i disse områder få tolkningspunkter til at binde realisationerne i modellen. Det er derfor forventeligt, at mindre afvigelser opstår i områder, der i forvejen er dårligt beskrevet. Den underliggende hydrostratigrafi skal således ikke opfattes som en sandhedsmodel, hvor variationen nødvendigvis altid skal ligge tæt på den. Det vigtigste er at observere, at realisationerne generelt er i overensstemmelse med den oprindelige model, hvilket sikrer, at den overordnede hydrostratigrafi forbliver konsistent.

Figur 3-27 viser det nordsydgående profil fra Jylland med 25 realisationer, der illustrerer usikkerhederne i modelleringen af følgende hydrostratigrafiske laggrænser: Bunden af det postglaciale ler (ny-modelleret), bunden af kvartært sand (1400 i FOHM-stratigrafien) og bunden af kvartært ler (2400 i FOHM-stratigrafien). Generelt set følger mange af de samme pointer som profilerne i Fyn og Sjælland – usikkerheden er dybdeafhængig, hvor de øvre laggrænser generelt har lavere usikkerhed end de dybere. Dette skyldes dels den højere tæthed af data i de øvre lag og den generelle dybdeafhængighed, der må forventes af geoinformationskilderne, hvilket reducerer variationen mellem realisationerne. Dette kan kvantificeres gennem den procentvise del af et lag, der har datadækning, som eksemplificeret i Tabel 6.

I profilforløbet vist i Figur 3.27 er det bemærkelsesværdigt, at ny-modelleringen af det postglaciale ler medfører relativt stor usikkerhed, selv nær overfladen. Dette skyldes, at lagene primært er tolket ud fra boringsdata i områder uden supplerende information, hvilket ifølge det nationale koncept giver anledning til høj usikkerhed – i tråd med realisationerne. Usikkerheden aftager dog markant, hvor laget findes tættest ved overfladen, da der er flest boringer her. På denne måde viser realisationerne, at GDM-N har været i stand til at håndtere det hurtige skift mellem høj sikkerhed tæt på observationer og stor usikkerhed i det omkringliggende område, hvor der reelt ikke eksisterer meget information om lagenes placering.



Figur 3-27: Profil fra Jylland i N-S retning med 25 realisationer (hvide streger), hvor usikkerheden på tre udvalgte lag-grænser (Bunden af det postglaciale ler (ny-modelleret), bunden af kvartært sand (1400 i FOHM-stratigrafien) og bunden af kvartært ler (2400 i FOHMstratigrafien)) illustreres. Overhøjningen er 147,9 og placeringen af profilet ses i Figur 3-24. UTMY-koordinater og højde i meter over havniveau bruges som akser.

Figur 3-28 viser det vestøstgående profil fra Jylland med 25 realisationer, der illustrerer usikkerhederne i modelleringen af følgende hydrostratigrafiske laggrænser: bunden af kvartært ler (2400 i FOHM-stratigrafien), bunden af Øvre Odderup Formation (5200 i FOHM-stratigrafien) og bunden af Klintinghoved Formation (5700 i FOHM-stratigrafien).

Da dette profil domineres af miocæne aflejringer (lyseblå farver), kan usikkerheden for de miocæne lag tydeligt anskueliggøres. TEM-metoderne har generelt svært ved at skelne mellem ler og sand på de dybder, grundet indtrængningsdybden og følsomheden, hvor miocænet findes, hvilket fører til stor usikkerhed i disse lag. Dette illustreres ved den betydelige variation mellem realisationerne for de miocæne aflejringer, hvor forskelle på op til 50 meter observeres. Samtidigt er det tydeligt at identificere områder med dybe boringer, da disse er associeret med lokaliteter med smalle usikkerhedsbånd på få meter, selv ved større dybder (for eksempel ved UTMX = 483000 og UTMX = 507000, se sorte pile). Modellen genskaber dog ikke den originale models placering af den begravede dal mod øst. Dette kan potentielt føre til en anderledes hydrologisk model sammenlignet med anvendelse af den originale hydrostratigrafi.



Figur 3-28: Profil fra Jylland i V-Ø retning med 25 realisationer (hvide streger), hvor usikkerheden på tre udvalgte lag-grænser (bunden af kvartært ler (2400 i FOHM-stratigrafien), bunden af Øvre Odderup Formation (5200 i FOHM-stratigrafien) og bunden af Klintinghoved Formation (5700 i FOHM-stratigrafien)) illustreres. Pilene markerer positioner med relativt lave usikkerheder på grund af dybe boringer med information om laggrænserne. Overhøjningen er 172,5 og placeringen af profilet ses i Figur 3-24. UTMX-koordinater og højde i meter over havniveau bruges som akser.

4. Diskussion af modelresultater

Det beskrevne arbejde i denne rapport omfatter opgaver udført som en del af kvælstofretentionsprojektet. Udviklingen af et nationalt koncept til kvantificering af usikkerheder og en metode til geostatistisk simulering af disse på landsplan repræsenterer et vigtigt gennembrud, da der ikke eksisterer lignende metoder til modellering på national skala. Det er derfor udfordrende at sammenligne resultaterne med andre løsninger, og modelleringen bør primært vurderes i forhold til de opnåede resultater. I denne rapport vurderer vi arbejdsgangens resultater i denne kontekst og foreslår forbedringer til både kvantificeringen af tolkningsusikkerhedskortene og videreudviklingen af GDM-N. Generelt vurderer vi, at enhver metode udviklet til probabilistisk modellering af hydrostratigrafi på landsplan skal opfylde tre fundamentale krav:

- 1. **Automatisering:** For at en metode på national skala skal være praktisk anvendelig, skal de forskellige trin i kvantificeringen og simuleringen være automatiserede, mens modelløren efterfølgende foretager kvalitetskontrol.
- 2. **Effektivitet:** Modelleringen skal kunne gennemføres inden for en rimelig tidsramme, der sikrer, at ventetiden på modelleringsresultater ikke bliver for omfattende. Dette giver mulighed for at eksperimentere med forskellige parametervalg i simuleringsal-goritmen.
- 3. **Konsistens og usikkerheder:** Den usikkerhed, der udspændes af de genererede realisationer, bør være i overensstemmelse med de geologiske og hydrologiske forventninger, og modellen som helhed skal være konsistent med udgangspunktet, herunder den oprindelige hydrostratigrafi.

4.1 Automatisering

Ved at lave en punktudtagning baseret på fladerne i stedet for at bruge tolkningspunkter direkte er det muligt at ensarte punktfordelingen ud over landet og dermed hjælpe til at simuleringsmetoden kunne automatiseres. Det overordnede koncept for usikkerhed, beskrevet i Kapitel 2, der baserer sig primært på det tilstedeværende data og modellernes kvalitet sikrer ligeledes at inputtet til algoritmen er ensartet. Med udviklingen af GDM-N har vi med succes automatiseret centrale procedurer i arbejdsgangen med generering af en hydrostratigrafisk ensemble og det første krav er opfyldt.

4.2 Effektivitet

Vi vurderer også, at vi opfylder det andet krav, selvom denne vurdering afhænger af tilgængelig computerkraft, antallet af nødvendige realisationer og hvor ofte modellerne skal opdateres med ny information. Oprettelse af varianskort for tolkningsusikkerhed tager nogle timer på en standard laptop, selv på national skala. Dette er derfor ubetydeligt i sammenligning med den nuværende GDM-N-kode implementeret i MATLAB, der tager omkring 3-4 dage for at generere 100 realisationer af hydrostratigrafien ved brug af en kraftig arbejdsstation (13th Gen Intel(R) Core(TM) i9-13900K 3.00 GHz med 24 kerner og 64 GB DDR4 RAM). Det største tidsforbrug ligger i løsningen af mindste-kvadraters problemet i ligningerne for hvert cluster, som tidligere nævnt. Dette er vanskeligt at optimere, medmindre tærskelværdien for cluster-størrelsen reduceres yderligere fra de 5000 celler, da standardalgoritmen til matrixinversion har en tidskompleksitet på O(n³). Selvom 3 dage kan virke omfattende, vil vi argumentere for at tidsrammen er acceptabel, så længe GDM-N-arbejdsgangen leverer værdifulde og pålidelige resultater.

4.3 Konsistens og usikkerheder

Vores resultater viser, at de overordnede strukturer bevares, idet den oprindelige hydrostratigrafi ligger inden for realisationernes opløsning. Samtidig synes usikkerhederne at have et rimeligt udfaldsrum. Helt generelt vil mængden af mulige kilder til usikkerhed i forbindelse med 3D geologisk modellering altid overstige vores evne til at identificere, oversætte og beskrive dem kvantitativt (Høyer el al. 2024). Derfor vil der i enhver model være usikkerheder, som ikke er håndteret, og man vil aldrig med sikkerhed kunne sige, om disse er korrekt håndteret. I den forbindelse er det altid en abstraktion at tale om, hvorvidt en metode kan udspænde det fulde udfaldsrum uden at præcisere, hvilket udfaldsrum der er tale om.

De to relevante spørgsmål i denne sammenhæng er:

- 1. Udspænder modelresultatet det udfaldsrum, der forventes givet de antagelser, vi har om systemet?
- 2. Udspænder modelresultatet det udfaldsrum, der er relevant at modellere?

Da realisationerne ser ud til at flugte med de antagelser, der er gjort om systemet, er det rimeligt at antage, at de udspænder det udfaldsrum, der relaterer sig til spørgsmål 1. For så vidt angår spørgsmål 2 vil GDM-N kun kunne udspænde den del af udfaldsrummet, der er forenelig med de antagelser, vi har til systemet, dvs. at vi kun tilskriver usikkerhed til tolkningen og derfor kun udspænder den usikkerhed, der er relateret til denne. GDM-N håndterer ikke usikkerheder, der stammer fra den konceptuelle model eller modelleringsparadigmet, da udgangspunktet er, at FOHM og DK-modellens hydrostratigrafi er en tilstrækkelig beskrivelse af undergrunden.

4.4 Samlet vurdering

Gennem opfyldelsen af de tre kriterier har vi forbedret modelgrundlaget for den videre hydrologiske modellering. Dette er opnået ved at opdatere fladerne med relevant ny information og designe et robust hydrostratigrafisk modelleringsværktøj. Værktøjet tilføjer ekstra værdi til de eksisterende hydrostratigrafiske modeller ved ikke blot at præcisere modelgrundlaget, men også ved at gøre usikkerheden kvantificerbar og tilgængelig for videre analyser og beslutningsstøtte. Samtidig er metoderne designet med høj fleksibilitet, hvilket tillader at opdaterede flader og datagrundlag kan integreres uden omfattende merarbejde.

5. Perspektiver for metoden (GDM-N)

Det anvendte geostatistiske ensemble er baseret på en første version af konceptet, der forventes at blive videreudviklet i fremtidige projekter. Arbejdet har været underlagt stramme tidsrammer, og de beskrevne koncepter skal derfor betragtes som værende under udvikling. I løbet af processen er der blevet identificeret flere områder, som kan optimeres, forfines og/eller videreudvikles. I det følgende beskrives nogle af de modelleringstekniske dele, der kunne være relevante at videreudvikle i fremtidige projekter, samt de overordnede perspektiver for anvendelsen af den geostatistiske modellering fremadrettet.

5.1 Usikkerhedstemaer

I det nuværende setup betragtes en lang række kilder til usikkerhed og omsættes til en usikkerhed på lagfladernes placering. Der er brugt både information fra boringer og geofysik, samt de individuelle hydrostratigrafiske modellers kvalitet. Derudover er der taget hensyn til afstanden til datapunkterne og gridpunkterne. I det nuværende setup bruges informationen fra geofysikken dog ikke fuldt ud, da usikkerheden er sat til at skalere med dybden, og der tages ikke højde for forskelligheden i opløsning af konduktive og ikke-konduktive lag. Oplagte områder for forbedringer og videreudvikling af settuppet er opridset herunder:

- Det vil være relevant at skelne mellem konduktive lag og resistive lag i de geofysiske inversionsresultater. For fålagsmodeller burde den primære laggrænse være til et konduktivt lag. Ved grænsen til dette lag skal usikkerheden falde markant og tilsvarende være større i modellen nedenunder, da meget af energien er absorberet i det konduktive lag. Dette er en kendt svaghed/styrke ved TEM sonderinger (se evt. Christiansen et al., 2009).
- Det vil også være relevant at undersøge overensstemmelsen mellem lag tolket i FOHM, og de resistiviteter, der kommer ud af inversionen. Dette vil kræve en statistisk tilgang, da der oftest er stort overlap mellem resistivitetsfordelinger for ler og sand, der tilmed også varierer lokalt såvel som regionalt (Barfod et al. 2016).
- Tilsvarende vil det være relevant at indarbejde en analyse af, hvordan laggrænser og litologier i boringer passer med laggrænser og hydrostratigrafiske lag i FOHM-modellen.
- Slutteligt vil det også være oplagt at undersøge den sensitivitet og udbredelse, der skal benyttes, når man bevæger sig væk fra datapunkterne. I det nuværende setup er den geologiske kompleksitet brugt som kriterie, og der er truffet et valg om afstandsfunktionen. En yderligere undersøgelse af hvordan dette valg påvirker usikkerhedstemaerne og alternative afstandsfunktioner, der er mere forankrede i den lokale geologi, vil være et oplagt sted at forsætte arbejdet.

5.2 Opsætning af simuleringsalgoritme

For at køre simuleringerne træffes der ligesom for usikkerhedstemaerne en række valg, der har indflydelse på det endelige resultat. Under udarbejdelsen af GDM-N er der eksperimenteret med en lang række parameteriseringer af algoritmen og enkeltkomponenterne. Da algoritmen bag GDM-N er omfattende, er der dog stadig en lang række forbedringsmuligheder og sensitivitetsanalyser at udføre. Her nævnes nogle stykker:

- En af de vigtigste parametre for slutresultatet er brugen af masker. Maskers indflydelse på resultatet er også visuelt nemt at detektere og sætte i en geologisk kontekst. I Jyllands og Sjællandsmodellerne er der brugt masker til at tvinge nogle lag til at opføre sig mere som lag i FOHM-modellen for Jylland og DKmodellen for Sjælland. Det vil i den forbindelse være naturligt at afdække i hvor stort et omfang dette er med til at forbedre den geostatistiske model ud fra den geologiske viden, der er til stede.
- I det nuværende setup bliver enkeltområder sammensmeltet med en simpel afstandsvægtning. Dette kan i nogle tilfælde medføre, at der tabes information om usikkerhed i randzonen mellem to simuleringsområder med forskellige statistiske modeller. Det vil være naturligt at kigge på, hvordan dette tab kan kompenseres gennem alternative vægtningskriterier.
- Lagkorrektion spiller også en meget vigtig rolle i enhver modellering, hvor der ikke er opsat en statistisk model til at beskrive korrelationen mellem de enkelte lag. Dette er også tilfældet i GDM-N. En undersøgelse af påvirkningen af lagkorrektionen er derfor nærliggende.
- Antallet af punkter, der udtages fra modellen, påvirker også modelresultatet og i hvor høj grad, at den geostatistiske model bindes til udgangspunktsmodellen. Under udviklingen af GDM-N er disse blevet testet på mindre skala og/eller kvalitativt på ad-hoc basis på national skala. En mere systematisk og kvantitativ afdækning af punkternes betydning for den samlede usikkerhed vil være oplagt.

6. Referenceliste

Aarhus University 2005. Landskabselementer (Landscape elements). Online: https://dca.au.dk/forskning/den-danske-jordklassificering.

Barfod, A. A. S., Møller, I., & Christiansen, A. V., 2016. Compiling a national resistivity atlas of Denmark based on airborne and ground-based transient electromagnetic data. Journal of Applied Geophysics, 134, 199–209. https://doi.org/10.1016/j.jappgeo.2016.09.017.

Christiansen, A. V., & Auken, E., 2012. A global measure for depth of investigation. Geophysics, 77(4), WB171–WB177. https://doi.org/10.1190/geo2011-0393.1.

Christiansen, A. V., Auken, E., & Sørensen, K., 2009. The transient electromagnetic method. In R. Kirsch (Ed.), Groundwater geophysics, 179–226. Springer Berlin Heidelberg. http://link.springer.com/10.1007/978-3-540-88405-7.

Damsgaard, A. & Makwarth, S., 2020. Quality assessment of boreholes within the Jupiter database. <u>https://gitlab.com/mst-gko/borehole_quality</u>.

Ditlefsen, C., Sørensen, J., Pallesen, T.M., Pedersen, D., Nielsen, O.B., Christiansen, C., Hansen, B., Gravesen, P., 2008. Jordprøver Fra Grundvandsboringer: Vejledning i Udtagning, Beskrivelse Og Geologisk Tolkning i Felten. Geovejledning 1, særudgivelse. https://www.geus.dk/media/7301/vejledning-jordproever.pdf.

Enemark, T., Madsen, R. B., Sonnenborg, T. O., Andersen, L. T., Sandersen, P. B. E., Kidmose, J., Møller, I., Hansen, T. M., Jensen, K. H., & Høyer, A.-S., 2024. Incorporating interpretation uncertainties from deterministic 3D hydrostratigraphic models in groundwater models. Hydrology and Earth System Sciences, 28(3), 505-523. https://doi.org/10.5194/hess-28-505-2024.

Falk, F. A., & Madsen, R. B., 2023. Machine learning-based estimation and clustering of statistics within stratigraphic models as exemplified in Denmark. GEUS Bulletin, 53, Artikel 8353. https://doi.org/10.34194/geusb.v53.8353.

Higham, N. J., 2009. Cholesky factorization. Wiley Interdisciplinary Reviews Computational Statistics 1, 251–254. <u>https://doi.org/10.1002/wics.018</u>.

Højberg, A. L., Thodsen, H., Børgesen, C. D. (red.) Andersen, A. H., Andersen, L. T., Audet, J., Bach, E. O., Christiansen, D., T., Christiansen, D. A., Conde, J., T., Falk, F. A., Frederiksen, R. R., Giannini-Kurina, F., Gudbjerg, J., Hansen, B., Henri, C. V., Henriksen, E. S., Hermansen, N., Hoffmann, C. C., Høyer, A. -S., Iversen, B. V., Jakobsen, R., Jørgensen, M. S., Kim, H., Kjeldgaard, A., Koch, J., Kronvang, B., Larsen, S. E., Liu, J., Madsen, R. B., Martin, N. L., Molis, M., Mortensen, M. H., Motevalli, A., Muff, E., Møller, I., Ondracek, M., Petersen, R. J., Pugliese, L., Rosenkrantz, A., Sandersen, P. B. E., Schneider, R. J. M., Sonnenborg, T. O., Stisen, S., Sørensen, P. B., Thorling, L., Tornbjerg, H., Troldborg, L., Uldall-Jessen, L., Voutchkova, D., Aamand, J. 2025. National kvælstofmodel - version 2025.

Udvikling af nye kvælstofretentionskort. Metode rapport. De Nationale Geologiske Undersøgelser for Danmark og Grønland. GEUS særudgivelse

Høyer, A. S., Sandersen, P. B. E., Andersen, L. T., Madsen, R. B., Mortensen, M. H., & Møller, I., 2024. Evaluating the chain of uncertainties in the 3D geological modelling workflow. Engineering Geology, 343, Artikel 107792. <u>https://doi.org/10.1016/j.enggeo.2024.107792</u>.

Jakobsen, P.R. & Tougaard, L., 2020. Danmarks digitale jordartskort 1:25000. Version 5.0. Geological Survey of Denmark and Greenland Report 2020/18.

Kristensen, M., Vangkilde-Pedersen, T., Rasmussen, E.S., Dybkjær, K., Møller, I., Andersen, L.T., 2015. Miocæn 3D, opdateret 2015. Den rumlige geologiske model. Danmarks og Grønlands geologiske Undersøgelse Rapport 2015/90. https://doi.org/10.22008/gpub/30740.

Madsen, R. B., Falk, F. A., Møller, I., & Høyer, A.-S., 2024. Introducing INPOX: a method for informed point extraction from geological 2D surfaces exemplified on the Danish national hydrostratigraphic model. GEUS Bulletin, 57, Artikel 8364. https://doi.org/10.34194/geusb.v57.8364.

Madsen, R. B., Høyer, A.-S., Andersen, L. T., Møller, I., & Hansen, T. M., 2022. Geologydriven modeling: A new probabilistic approach for incorporating uncertain geological interpretations in 3D geological modeling. Engineering Geology, 309, 106833. https://doi.org/10.1016/j.enggeo.2022.106833.

Miljøstyrelsen 2020. Samling af geologiske modeller i Jylland. FOHM – fælles offentlig hydrologisk model. Rapport, Miljøstyrelsen, Miljø- og Fødevareministeriet.

Mortensen, M.H, Andersen L., T., Møller, I., Sandersen, P.B.E & Høyer, A.-S. 2025. Lokale opdateringer af FOHM-modellens hydrostratigrafi til Kvælstofretentionskortet version 2025. Danmark og Grønlands Geologiske Undersøgelse Rapport 2025/19.

Orbicon, 2018. Hydrostratigrafisk model for Fyn, Odense Kommune. Rapport til Miljøstyrelsen. Orbicon.

Sandersen, P.B.E, 2021. A basic geological complexity map for use in the implementation of the MapField concept. GEUS. Danmarks og Grønlands Geologiske Undersøgelse Rapport 2021/37. https://doi.org/10.22008/gpub/34597.

Smed. P., 1978. Landskabskort over Danmark. Geografforlaget 1987.

Stisen, S., Ondracek, M., Troldborg, L., Schneider, R.J.M., van Til, M.J., 2020. National vandressource model. Modelopstilling og kalibrering af DK-model 2019. Danmarks og Grønlands geologiske Undersøgelse Rapport 2019/31. https://doi.org/10.22008/gpub/32631.