



Norges miljø- og
biovitenskapelige
universitet

Masteroppgave 2017 30 stp
Institutt for matematiske realfag og teknologi

Automatisk klassifisering av vannflater med LiDAR-data

Automatic detection of waterbodies in LiDAR

Joakim Tveit Husefest
Geomatikk

Forord

Denne masteroppgaven avslutter min mastergrad i Geomatikk på NMBU - Norges Miljø og Biovitenskapelige Universitet. Oppgaven er skrevet ved IMT - Institutt for matematiske realfag og teknologi.

Problemstillingen ble valgt med et eget ønske om å bidra til en aktuell utfordring i dagens geomatikkbransje. Jeg retter derfor en takk til Håkon Dåsnes i Statens Kartverk for initialisering av oppgaven og biveileder Christian Malmquist for veiledning av datagrunnlag. Det er ekstra givende å skrive masteroppgave når produktet er etterspurt.

Takk til veileder Håvard Tveite, som har vært en god støtte i faglige utfordringer som har oppstått underveis.

Til mine medstudenter på geomatikklinjen: Takk for 5 fine år. Det blir stas å bli kollegaer i geomatikkbransjen etter endte studier.

Til slutt, en spesiell takk til min far for oppmuntrende ord og uvurderlig støtte. Du har alltid troen på meg og at jeg kommer til å klare det jeg setter meg for. Det betyr mye.

Sammendrag

Flybåren laserskanning for å generere topografiske 3D terrengmodeller har for tiden økende popularitet. Dette fordi det er rimeligere og raskere enn tradisjonelle metoder basert på stereoskopiske bilder.

Utfordringen med laserskanning er imidlertid at vannflater gir dårlige reflekser til skan-neren, og man må derfor i ettertid legge på vannflatene manuelt etter å ha hentet dem fra eksisterende datasett, som FKB-vann. Utfordringen er at disse er til dels gamle og av varierende kvalitet, og å legge FKB-vann inn i terrengmodellen krever også manuell innsats.

Denne masteroppgaven presenterer to metoder for automatisk klassifisering av vann-flater, basert på data fra flybåren laserskanning. Den ene er en implementasjon av en eksisterende metode og den andre er egenutviklet. Metodene er basert på mange av de samme prinsippene og har til hensikt å klassifisere stille vann og innsjøer.

For klassifisering benyttes egenskapene intensitet og høyde fra flybåren laserskanning, eller LiDAR. Det velges en tilnærming der man filtrerer bort karakteristikkene som ikke tilfredsstillende kriterier, for eksempel neglisjeres områder som ikke er flate. Vannflater gir dessuten lav intensitetsverdi eller ingen retur i det hele tatt, og man neglisjerer derfor områder med høye intensitetsverdier. Potensielle vannflater er til slutt en kombinasjon av intensitets- og høydeinformasjon. De potensielle vannknekklinjene fremkommer som konturer av potensiell vannflate.

Denne masteroppgaven har kommet frem til at automatiske metoder kan gi like god eller bedre nøyaktighet enn FKB-vann med begge metoder, da vi kun har hatt tilgang til eldre datasett. Den egenutviklede metoden lyktes med å redusere krav til regnekapa-sitet betraktelig sammenlignet med den eksisterende eksisterende metoden. Her ligger det altså store potensielle gevinster. Men for å verifisere omfanget av et eventuelt bruks-område, kreves testing av ulike datasett, ulike topografi, ulike parametere i metoden og mer forskning på bruk av intensitetsegenskapene i LiDAR.

Abstract

Airborne laser-scanning to create topographic 3D terrain-models is becoming increasingly popular. It's a more efficient method and has lower cost compared to classic methods based on stereoscopic imaging.

A known challenge is that water surface is returning weak signals to the scanner, low intensity or no signal at all. Until now, it has been necessary to use map-data for lakes from existing datasets, like the Norwegian standard "FKB-vann". These seems to be old and unprecise, and its a manual processes.

This thesis presents two methods for automatic classification of water surface based on data from airborne LiDAR (Light Detection and Ranging). One method is an implementation of an existing method, and the other is developed by the author and described in this document. Both are using many of the same principles and aims to distinguish lakes and ponds from other terrain.

Both methods use elevation and intensity data from LiDAR as attributes for classification. They exclude bins in a raster not fulfilling certain criteria required to be classified as water surface. For instance, a filter-based method removes areas that are not flat, and then removes bins with intensity values that are considered too high to possibly be returned from water surface. Potential lake contours are in the end determined by surrounding regions of bins in a raster satisfying both intensity and elevation requirements.

Furthermore, the new method described herein is drastically reducing the need for computing capacity compared to the existing method presented. But to verify the findings, we need more testing on different datasets, topography, setting of parameters in the method and more information on how to use intensity values to determine water surface with higher certainty.

Innhold

Forord	III
Sammendrag	V
Abstract	VII
Figurer	XI
1 Innledning	1
1.1 Introduksjon	1
1.1.1 Vannflater, vannknekklinjer og Hydro Flattening	1
1.1.2 FKB-vann og bruk i topografiske terrengmodeller	2
1.1.3 Detaljert Norsk Høydemodell	3
1.1.4 Utfordringer med klassifisering i vann	4
1.1.5 Tidligere arbeid for klassifisering av vann	5
1.1.6 Introduksjon til metode	6
2 Teknologi og Datagrunnlag	8
2.1 Flybåren Laserskanning	9
2.1.1 Bestemmelse av 3D punkter	9
2.1.2 Egenskaper i laserskanner	12
2.2 Datagrunnlag	17
2.2.1 Riegl LMS-Q1560	17
2.2.2 Aktuelt datasett	19
3 Teori	22
3.1 Digitale Terrengmodeller	22
3.1.1 Interpolasjonsmetoder	23
3.1.2 Inverse Distance Weighting	23
4 Avgrensninger	26
4.0.1 Avgrensninger i Hydro Flattening	26
4.0.2 Andre avgrensninger	26
5 Metode	28
5.1 Forberedelser og valg	28
5.1.1 Generelle valg for metoder	28

5.2	Toscano 2013	31
5.2.1	Detektering av flate områder	32
5.2.2	Mediansjekk med intensitet	39
5.2.3	Vannflater fra intensitetsraster - Toscana	41
5.2.4	Oppsummering Toscana et al. (2013)	43
5.3	Egenutviklet metode	44
5.3.1	Valg av rutestørrelse i egenutviklet metode	46
5.3.2	Filterbasert deteksjon av flate områder	46
5.3.3	Vannflater fra intensitetsraster - Egenutviklet metode	53
5.3.4	Forbedring av vannknekklinjene med buffer og høyde	55
6	Resultater	58
6.0.1	Egenutviklet metode og FKB-vann	59
6.0.2	Egenutviklet metode sammenlignet med Toscana	60
6.0.3	2 punkt per kvadratmeter datasett	61
6.0.4	Redusert høydeoppløsning	62
6.0.5	Radius i IDW	63
6.0.6	Datasett med flere vann	64
7	Drøfting	66
7.1	Utfordringer for klassifisering i LiDAR-data	66
7.2	Hydro Flattening	66
7.3	Feilkilder i datasett og målemetode	67
7.3.1	Intensitet	67
7.4	Usikkerhet i metode	68
7.4.1	Filterbasert metode vs. histogramanalyse	68
7.5	Klassifisering av Frøvik tjenn	68
7.6	Valg av parametere	69
7.6.1	Interpolator	69
7.6.2	Høydeoppløsning	69
7.6.3	Filtrere bort veier	70
7.6.4	Skannevinkelkorreksjon	70
7.7	Ulik punkttetthet og uregistrerte områder	71
7.7.1	2pkt per kvadratmeter datasett	71
7.8	Videre arbeid	72
7.8.1	Strandsone og kystkontur	72
7.8.2	Videre tester og kvalitetssikring	72
7.8.3	Toscana 2015	73
8	Konklusjon	74

Figurer

1.1	Ved å bruke et vannpolygon på LiDAR-generert terrengmodell blir det omkringliggende landskapet fremhevet og støy (hull og spiker) fjernet . . .	2
1.2	Status NDH per 1. Mars 2017 Kilde:	3
2.1	Laserskanningsprinsippet. Kilde: www.gpsworld.com/surveymappingsparse-network/	9
2.2	Komponenter i flyet. (Vosselman and Maas, 2010)	10
2.3	Tidsendring av pulser i faste intervaller. (Vosselman and Maas, 2010) . . .	10
2.4	Lokal orientering av flyet. (Vosselman and Maas, 2010)	11
2.5	Swath Width (Vosselman and Maas, 2010)	12
2.6	Returer fra LIDAR. (Kartverket, 2010)	13
2.7	Returer fra LIDAR. (Vosselman and Maas, 2010)	13
2.8	Roterende polygon og svingende speil. Figur: (Rasmussen, 2000)	14
2.9	Puls fra utsendt signal og mottatt signal. Figur (Hug, 1996)	15
2.10	Refleksjon fra ulike objekter ved ulike bølgelengder. Kilde: (Hug, 1996) . .	16
2.11	Refleksjonsmønsteret i stille vann. (Toscano, 2015)	16
2.12	Riegl LMS-Q1560	17
2.13	Riegl Skannemønster	18
2.14	Flystriper Kragerø Drangedal 5 punkt per kvadratmeter datasett(Terratec, 2016)	19
2.15	Punkttetthet Kragerø Drangedal 5 pkt datasett (Terratec, 2016)	20
3.1	Et eksempel på en digital terrengmodell i et rasterformat, her representert ved gråskalaverdier fra 17moh-174moh	23
3.2	Et rutenett med $1m^2$ rutestørrelse lagt på punktsky av kontrollpunkter. Kontrollpunktene her er tatt fra ett av $5 \frac{Pkt}{m^2}$ datasett fra Kragerø-Drangedal	24
3.3	Formler IDW (O'Sullivan and Unwin, 2014)	24
3.4	IDW eksempel fra (O'Sullivan and Unwin, 2014)	25
5.1	Frøvikjtjenn Kilde: Topografisk Norgeskart 2 gråtone	29
5.2	Terrengmodell etter IDW interpolasjon	30
5.3	Intensitetsraster	30
5.4	Toscano et al. (2013) Høydedata og intensitetsdata kombineres for å klassifisere vann	32
5.5	Histogram av Frøvikjtjenn	33
5.6	Skalerte histogram	34
5.7	Topper detektert i Frøvikjtjenn	36

5.8	Verdi 1 representerer uregistrerte områder i et binært bilde	37
5.9	Topper i histogrammet og uregistrerte områder som representerer 1 i binært bilde	37
5.10	Connected Component Labeling. Rutene som sammenstøter og har samme verdi får en region-ID. Antall regioner illustreres i skala	38
5.11	Arealer mindre enn $1000m^2$ filtrert bort	38
5.12	Median av intensitetsverdier med og uten uregistrerte områder	39
5.13	Intensitetsverdier under 20 og uregistrerte områder	41
5.14	sammensetting av uregistrerte punkter, intensitetsverdier under 20 og dekket vannflate	42
5.15	Egenutviklet metode	45
5.16	GRID over terrengmodell med $20*20$ ruter blokker	47
5.17	Resultat etter filterbasert metode trinn 1 med 12.5cm høydeoppløsning og 5m grunnrissnøyaktighet	48
5.18	Connected Component Labeling på 5.17	49
5.19	Resultat av høydedeteksjon med grunnrissnøyaktighet 1m og 3cm høydeoppløsning	50
5.20	Skannevinkel mellom -2 og 2 gir i vårt datasett en belte under flyet som er omtrent 100meter bredt	51
5.21	Foreløbig polygon av potensiell vannflate ved filterbasert metode	52
5.22	Figuren viser intensitetsraster korrigert for skannevinkel fra 5.3.2 og ruter med intensitetsverdi under 20	53
5.23	Sammensatt intensitetsraster og høyderaster	54
5.24	$0.0625m^2$ ruter som tilfredsstillter 12.5cm høydeoppløsning i en 5 meters buffer er representert med 1. Mørke områder er enten selve vannflaten eller områder utenfor 5 meters buffer	56
5.25	Vannflate etter korreksjon av høyde og intensitet	57
6.1	Egenutviklet metode mot FKB-Vann	59
6.2	Toscano mot egenutviklet metode Toscano (2015)	60
6.3	$2 \frac{Pkt}{m^2}$ datasett	61
6.4	Redusert høydeoppløsning øker størrelsen på vannflaten	62
6.5	2 meters radius og 4 meters radius	63
6.6	Høydemodell med flere vann	64
6.7	Datasett med mindre tydelige vannknekklinjer	64
7.1	Uregistrerte områder representeres som svarte felter	71

1 Innledning

1.1 Introduksjon

Statens Kartverk har et ønske om å få en automatisk metode for klassifisering av vannflater i LiDAR. Bakgrunnen for dette, er utfordringen om å detektere korrekt grense mellom vann og land i data fra flybåren laserskanning. Utfordringene er knyttet til målemetoden selv, varierende topografi og valg av metoder. En vellykket klassifisering av vann vil bidra til at digitale terrengmodeller bedre beskriver topografien til virkelig terrengoverflate. Til nå er FKB-vann Geovekst (2016) det mest brukte kartproduktet for dette formålet, men datasettet er til dels gammelt og laget med manuelle metoder. Derfor eksisterer behovet for en automatisk klassifisering.

Med dette som bakgrunn vil denne masteroppgaven fokusere på følgende problemstilling:

Hvordan klassifisere vannflater med data fra flybåren laserskanning med automatisk metode?

Første del av innledningen introduserer relevante begreper, noen bruksområder for vannflater i LiDAR-genererte terrengmodeller og annen bakgrunn for arbeidet. Innledningen avsluttes med utfordringer knyttet til klassifisering av vann i LiDAR, tidligere arbeid og introduksjon til metode.

1.1.1 Vannflater, vannknekklinjer og Hydro Flattening

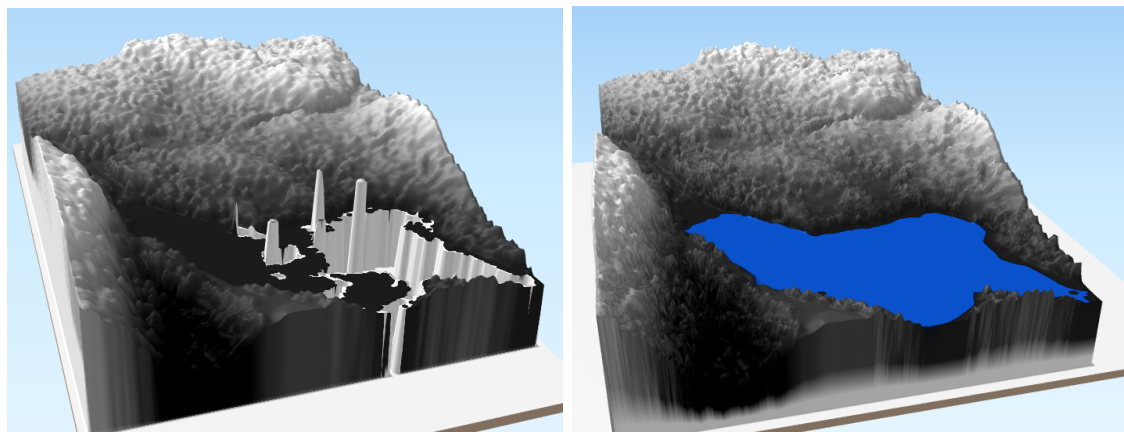
En *knekklinje* blir beskrevet av Bjørke (2010) som *linjer som definerer markante endringer i terrengoverflatens hellning*. Knekklinjene går under kategorien *diskretiseringsmønstre*, som i tillegg til interpolasjonsmetoder og punktdistribusjon er avgjørende for hvordan terrengmodellen beskriver virkelig topografi. I Bjørkes eksempler brukes knekklinjene i genereringen av TIN-nettverk (Triangular Irregular Network). En vannknekklinje defineres i denne oppgaven som grensen mellom vann og land.

Et annet begrep relevant for problemstillingen er "Hydro Flattening". "Hydro Flattening" er en teknikk for LiDAR-genererte terrengmodeller der vann er flate, uten påvirkning av vanndatasett laget ved fotogrammetriske metoder. "Hydro Flattening" er formulert av U.S. Geological Survey Standards (2014a), som er organisasjonen ansvarlig for ASPRS LAS File formatet.

Samme kilde påpeker at LiDAR-modeller laget med "Hydro Flattening" er en topografisk terrengmodell med hensyn på å fremheve omkringliggende topografi og fenomener. Derfor skal en slik modell ikke forveksles med andre typer terrengmodeller med annet formål, selv om modellen inkluderer vann. "Hydro Flattening" omfatter ikke endringer i selve punktskyen eller filformatet, men inngår i modelleringen av endelig terrengmodell. U.S. Geological Survey Standards (2014a) har satt kriterier for hvordan vannflater skal behandles i teknikken. Kriteriene er sentrale hvilke resultater vi ønsker å oppnå med de automatiske metodene som blir presentert. Kriteriene blir derfor nøyere gjennomgått i kapittel 4.

1.1.2 FKB-vann og bruk i topografiske terrengmodeller

Figur 1.1 illustrerer hvordan vannflater fra FKB-vann kan benyttes for å forbedre en LiDAR-generert terrengmodell. Figur 1.1a visualiserer også noen vanlige situasjoner som ofte oppstår ved LiDAR målinger i vann, nemlig hull, spiker og uregistrerte områder.



(a) LiDAR-generert terrengmodell der vann ikke er klassifisert. Vi ser hull, spiker og uregistrerte områder
(b) LiDAR-generert terrengmodell der FKB-vann (eksisterende vanndatasett) har blitt lagt på modellen i 1.1a

Figur 1.1: Ved å bruke et vannpolygon på LiDAR-generert terrengmodell blir det omkringliggende landskapet fremhevet og støy (hull og spiker) fjernet

Fra figur 1.1 ser vi hvordan eventuelle feil i vannpolygonets grunnrisskoordinater og høyde ville ført til ukorrekt topografisk terrengmodell.

Utfordringen ved å benytte FKB-vann er at datasettet er varierende kvalitet. Noen data er til dels gamle. Eldre datafangstmetoder er benyttet, deriblant stereoinstrument, digitalisering fra frihåndstegning og overføring fra papirkart ved et digitaliseringsbord

Geovekst (2016). I tillegg mangler flere vannflater høydeinformasjon, som betyr at data settet ikke kan brukes i sammenheng med eksempler som 1.1. Årstid, oppdemming, tidevann, landheving eller andre endringer i landskapet tas ikke hensyn til i geometrien (Geovekst, 2016).

For å møte disse utfordringene, er det behov for oppdatere disse dataene, men samtidig redusere manuelt arbeid. Siden flybåren laserskanning regnes som en kostnadseffektiv og høyoppløselig kilde til terrengmodeller, kan målemetoden tenkes å være en ny kilde for vannflater i fremtiden.

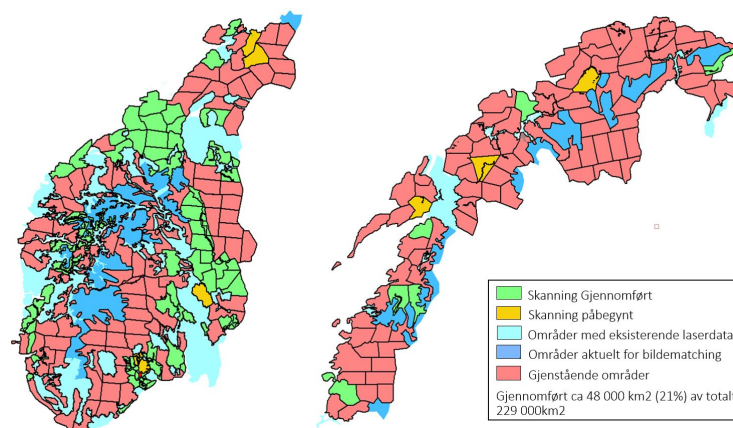
I denne oppgaven skal vi se på mulighetene utfordringene med automatiske metoder, der bare LiDAR brukes som målekilde.

I neste del skal målemetoden flybåren laserskanning introduseres i sammenheng med det nasjonale kartleggingsprosjektet "Detaljert Norsk Høydemodell".

1.1.3 Detaljert Norsk Høydemodell

"Detaljert Norsk Høydemodell" er et nasjonalt kartleggingsprosjekt med hensikt å lage en høyoppløselig topografisk høydemodell av Norge Kartverket (2014a).

Med "høyoppløselig" menes det et overordnet oppløsning med minst 2 3D-punkt per kvadratmeter. Noen steder er det målt 5 punkt per kvadratmeter (Kartverket, 2014b). Statens Kartverk står ansvarlig for gjennomføring og forvaltning av høydemodellen. Prosjektet begynte i 2015 og skal etter planen være ferdigstilt i 2020. Det brukes hovedsaklig to målemetoder for generering av 3D data. I 2017 er det bestilt 92000 kvadratkilometer LiDAR data og 10000 kvadratkilometer med bildematchingKartverk (2014). LiDAR er med andre ord den største kilden til terrengmodeller i NDH. I denne oppgaven benyttes LiDAR-data generert i 'Norsk Detaljert Høydemodell', heretter omtalt som NDH (Kartverket, 2014a).



Figur 1.2: Status NDH per 1. Mars 2017 Kilde:

Fotogrammetrisk bildematching går ut på å studere samme koordinat i to flybilder med en bestemt overlapp mellom bildene. Resultatet fra bildematching er alltid en 3D koordinat per piksel, med tilhørende RGB verdi som representerer overflaten. Fra 2012 har hvert flybilde i omløpsfotograferingen en oppløsning på 0.25m i grunnriss. Oppløsningen fører til opptil 16 høydepunkter pr m^2 og en kan forvente +/- 0.5 m i høydenøyaktighet. Kartverket (2014a).

Flybåren laserskanning er en optisk fjernmålingsmetode, som regnes som en kostnads-effektiv og tidsbesvarende målemetode i områder som ikke krever høyeste punktetthet. Bildematching brukes for det meste i bratte fjellområder, da LiDAR dårlig fanger opp detaljer i slikt landskap. (Kartverket, 2014a)

En viktig fordel med LiDAR er at 3D-punkter er et direkte resultat av målemetoden og ikke krever etterprosessering. I tillegg kan laserskanneren bidra til registrere andre egenskaper ved hvert 3D punkt, for eksempel skannevinkel, GPS tid og intensitet. Siden denne oppgaven baserer seg klassifisering i LiDAR-data, er målemetoden videre beskrevet i kapittel 2.

1.1.4 utfordringer med klassifisering i vann

"ASPRS LAS file format" ASPRS (2013) er en standard for håndtering og prosessering av LiDAR-data. Standarden legger blant annet føringer for klassifisering. "Terreng", "skog", "bygning" og "vann" er eksempler på klasser som skal representere respektive fysiske objekter i LiDAR-punktsky.

Mest kritisk for terrengmodeller er klassifiseringen av klassen "terreng", som består av punkter som tilhører bakken. Punkter som ligger på eller under vannoverflaten skal ifølge ASPRS (2013) ikke være i terrengklassen, men i "vann"-klassen. På grunn av opplevd feilaktig klassifisering av vann, leveres ikke denne klassen i standardleveransene i NDH. Med andre ord omfatter terrengklassen i dagens leveranser også punkter på vannoverflater, enten det er innsjø, hav eller elv.

Intensitet er en egenskap med LiDAR som kan brukes til klassifisering. Siden ulike objekter på bakken reflekterer forskjellig styrke i retursignalet, kan denne målingen diskriminere ulike objekter fra hverandre.

Vann har i denne sammenheng noen generelle karakteristikk. De fleste LiDAR-systemer benytter nærinfrarødt lys (800-1400nm) som bølgelengde. Det første man kan forvente er at nærinfrarødt lys ikke penetrerer vann, men reflekterer stort sett fra vannoverflaten Rasmussen (2000). For det andre reflekteres laserstrålene vekk fra sensor i større grad i vann enn andre fysiske objekter. Dermed kan man generelt forvente lavere intensitet fra vann enn andre objekter. Hug (1996)

Likevel viser det seg problematisk å klassifisere vann i LiDAR. Det er nyanser rundt målemetoden som setter krav til en klassifiseringsmetode. Spesielt i intensitetsverdier gir LiDAR-målingene varierende resultater.

Höfle et al. (2009) beskriver her noen situasjoner som kan oppstå ved LiDAR-målinger i vann og forhold å ta hensyn til i en klassifiseringsmetode:

- I stille vann kan det oppstå en urealistisk høye intensitetsverdier dersom målingen er målt med liten skannevinkel.
- Uregistrerte områder kan skyldes:
 1. Nærinfrarødt lys reflekteres i ”flere retninger”, og lyset oppfattes ikke av sensor.
 2. Høy skannevinkel.
 3. Stor flyhøyde og stor flyfart.
 4. Absorpsjon av nærinfrarødt lys.
- Andre påvirkninger på intensitetsmålinger kan være:
 1. Vegetasjon i vannet
 2. Myrområder
 3. Veier har mye av de samme refleksjonene som vann, for eksempel enkel-retur og lav intensitet
 4. Nærinfrarødt lys kan penetrere vannet i grunner og strender
 5. Variasjoner i topografi kan påvirke klassifiseringsalgoritme. Grunner, strender og bratte stup.
 6. Bølgeeffekt.

1.1.5 Tidligere arbeid for klassifisering av vann

Forskning på klassifisering i LiDAR har økt i takt med bruken av målemetoden. Ulike former for vann blitt analysert med ulike representasjoner og ved ulike metoder.

Höfle et al. (2009) demonstrerer en ”3D growth” metode som leter etter store rektangler i et TIN (Triangular Irregular Network). Store triangler i nettverket gir indikasjon på potensielt vann. Prosjektet brukte data flybåren laserskanning og kombinerer intensitetsverdier og høydeverdier. Området som ble valgt består av to ulike typer elver. Det ene datasettet en 20km lang elv som karakteriseres av sand, grunne områder og vidstrakt landskap. Det andre datasettet med en smalere 3 kilometers elv med gjennomsnittshelling på 3%. Resultatet er en konturlinje som definerer knekklinjer med 97% klassifiseringsrate med 0.45meters grunnrissnøyaktighet. Valideringen baserer seg på GPS-måling og ortofoto. Höfles undersøkelser viser at punkttetthet og variasjon i intensitetsverdier avhenger av skannevinkel. Med skannevinkel menes vinkelen mellom laserstrålens utgangsretning fra flyet og normalen fra bakken. Ifølge kilden fører målinger i vann med skannevinkel høyere enn 9 grader til betydelig økning av antall uregistrerte

områder. I skannevinkler under 2 grader oppstår det i tillegg urealistisk høye intensitetsverdier. En viktig del av Höfle et al. (2009) er derfor en radiometrisk korreksjon som påføres intensitetsverdiene i et preprosesseringssteg.

I denne oppgaven har en tatt lærdom av observasjonene i Höfle et al. (2009) når det gjelder korreksjon av høye intensitetsverdier forårsaket av lave skannevinkler. I oppgavens datasett er dette problemet oppsvert, og dette regnes som en nødvendig observasjon.

Brzank and Heipke (2007) viser til detektering av knekklinjer ved bruk av statistiske beregninger per punkt i punktskyen. Datasettet består av et havområde på 7300 kvadratkilometer. "Vann" og "grunne strender" ('mudflat') blir forsøkt klassifisert og egenskapene som blir analysert er høydeverdier, intensitet og punkttetthet. Klassifiseringen skjer punkt for punkt, uten naboanalyse. Fuzzy-logic metoden går ut på å tillegge nevnte egenskaper forskjellig vekt basert på gitte kriterier. For eksempel avhenger intensitet av skannevinkel. På samme måte øker sannsynligheten for et vannpunkt ved lav intensitetsverdi og lav høydeverdi.

Ulempen ved Brzank and Heipke (2007) er at lave høydeverdier ikke nødvendigvis er kjennetegn på vann, da det kan finnes flere vann i samme område ved ulike høyder. Dette har også blitt tatt hensyn til i metodene som blir presentert i denne oppgaven.

Klassifisering av vannknekklinjer har også blitt utviklet i rasterformatet. Liu (2008) viser til en metode som benytter hellingsraster. Vannknekklinjene detekteres ved ruter i grenselinjen mellom flate områder og signifikant stigning. Brugelmann (2000) analyserte et raster av andrederiverte med hypotesetesting.

Smeeckaert et al. (2013) tar også sikte på å klassifisere vann i rasterformatet. Metoden baserer seg på "Support Vector Machine" teori (også kjent som "Machine Learning"). Analyserte egenskaper er høyde, lokal punkttetthet i nabolag, høyest punkttetthet i nabolag, forhold mellom to punkttettheter i nabolag, egenverdier og spredning. Punkttettheten varierer med ulike datasett. For eksempel tar metoden hensyn til flystriper som overlapper og som gir høyere punkttetthet i noen områder. Resultatene viser til 95% klassifisering i kystkontur og 89% i elver i testområdet.

Smeeckaert et al. (2013) poengterer at punktdistribusjon er viktig for klassifisering i LiDAR og at den i tillegg varierer innenfor hvert datasett. Punkttetthet er i denne oppgaven nøyere beskrevet i kapittel 2.

1.1.6 Introduksjon til metode

Som vist i 1.1.5, er det ulike metoder for å klassifisere vann og hver metode har sine styrker og svakheter. Felles for metodene nevnt er at det benyttes utelukkende data fra flybåren laserskanning som kilde, uten satellittbilder eller alternative målemetoder involvert. Dette er et sentralt poeng fordi det reduserer manuelt arbeid.

Metoden som er presentert i denne masteroppgaven er sterkt inspirert av Toscano et al.

(2013). Toscano er kjent for sin metode for klassifisering av stille vannflater. Det senere arbeidet, Toscano (2015), er nok mer omfattende enn Toscano et al. (2013), som vi har tatt utgangspunkt i. Metoden har i rasterformatet fokus på å klassifisere innsjøer og gjør det med følgende utgangspunkt:

1. Vann har stort sett samme høyde
2. Vannet ligger lavere enn det omkringliggende landskapet
3. Intensitetsverdier er generelt lavere i vann enn i det omkringliggende landskapet
4. Nærinfrarødt lys har særegne reflekterende egenskaper.

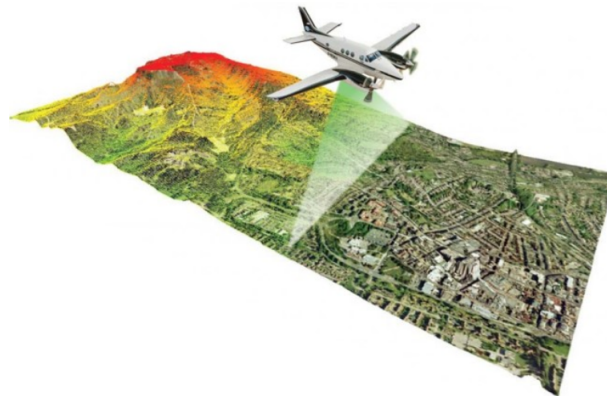
Toscano (2015)

Istedenfor å klassifisere i punktskyen, benyttes det en terrengmodell i rasterformatet. Metoden klassifiserer vannflater ved arealberegning, der vannknekklinjene fremkommer som konturer av detektert vannflateareal.

Metodedelen i kapittel 5 består av en beskrivelse av Toscano et al. (2013) og presentasjon av den egenutviklede metoden. Denne metoden er basert på mange av de samme prinsippene som Toscano et al. (2013). Metodene har blitt testet med motivasjon om å oppfylle kravene til "Hydro Flattening", som blir gjort rede for i kapittel 4.

Et teknologikapittel om flybåren laserskanning er avsatt for å forstå det teoretiske grunnlaget for klassifisering i LiDAR. Metodedelen innledes deretter av et avgrensningskapittel som fastsetter hvilke kriterier som settes for metodene presentert. I kapittel 7 diskuteres resultatene av vannflatene og vannknekklinjene som fremkommer av metodene som presenteres.

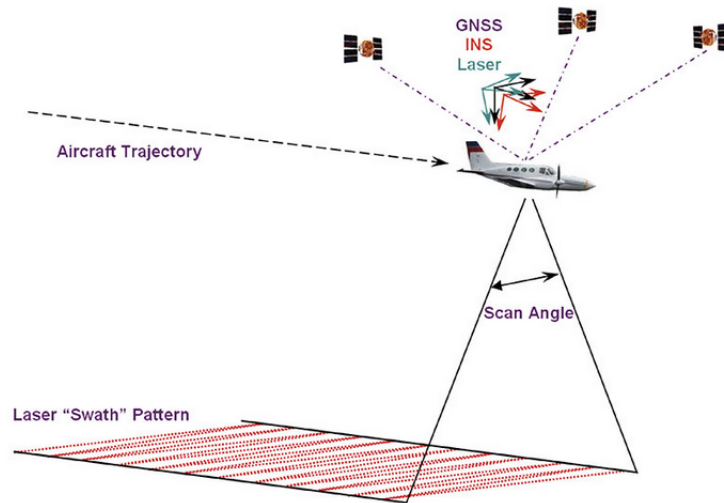
2 Teknologi og Datagrunnlag



LIDAR er en optisk fjernmålingsteknikk for måling av fysiske objekters posisjon i 3 dimensjoner. Selv om denne oppgaven omhandler data fra flybåren laserskanning, benyttes måleteknikken også blant landbaserte kjøretøy, teodolitter og droner.

Dette kapitlet tar først for seg teknologi i flybåren laserskanning før det avsluttes med detaljer rundt aktuell laserskanner og resulterende datasett. Kapitlet blir satt i sammenheng med hva som er relevant for klassifisering i data fra flybåren laserskanning.

2.1 Flybåren Laserskanning



Figur 2.1: Laserskanningsprinsippet. Kilde: www.gpsworld.com/surveymappingparse-network/

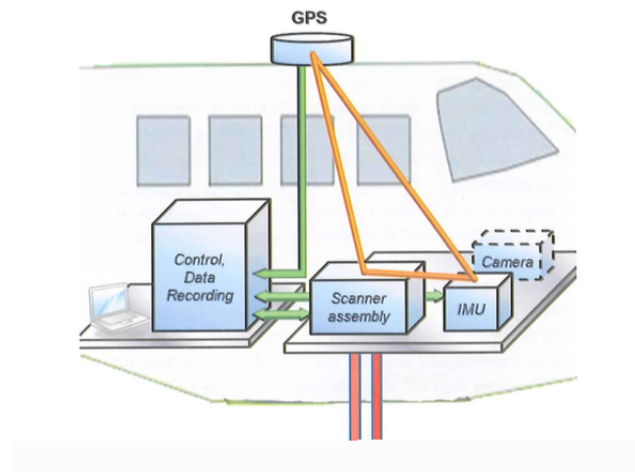
Flybåren laserskanning skjer ved at et fly beveger seg over landskapet og måler terrengoverflaten automatisk. Figur 2.1 viser laserskanningsprinsippet. Vi ser at flybanen er en lineær bane der en laserskanner sveiper landskapet i en skannevinkel. I flyet er det montert tre systemer; GNSS, INS og en laserskanner som i kombinasjon genererer 3D punktsky.

Resulterende datasett er påvirket av valg av skannevinkel, flyhøyde og topografi. I tillegg påvirkes målingene av parametre i laserskanneren, som for eksempel bølgelengde og repetisjonsfrekvens.

2.1.1 Bestemmelse av 3D punkter

3D punkter genereres i en kombinasjon av INS (Inertial Navigation System), GPS (Global Positioning System) og en laserskanner.

I tillegg til komponentene i flyet, viser figur 2.2 et montert multispektralt kamera i flyet. Flybilder har ofte høyere oppløsning enn LIDAR målinger og støtter LIDAR målinger på objekter som trenger høyere punkttetthet.



Figur 2.2: Komponenter i flyet. (Vosselman and Maas, 2010)

Avstandsmåling

Laserskanneren måler avstanden mellom flyet og overflaten.

Det er to måter å måle tidsendringen på; pulsmåling og "Continuous Wave"(CW). I pulsmåling blir signalene utsendt i faste intervaller, mens i CW sendes signalene kontinuerlig ("Continuous"). I pulsmåling måles tidsendring mellom utsendt og reflektert signal av samme puls, men CW måler dette i form av bølgelengdeforskyvning. (Rasmussen, 2000)

I denne oppgavens datagrunnlag er skannere med pulsmåling benyttet. Grunnprinsippet i pulsmåling er å måle tidsendring av utsendt og reflektert signal av samme puls:

Figur 2.3: Tidsendring av pulser i faste intervaller. (Vosselman and Maas, 2010)

$$\rho = \frac{c * \tau}{n * 2} \quad (2.1)$$

ρ er avstand mellom laserskanner og reflekterende objekt og c er konstant lyshastighet. τ er tiden lyset bruker fra laserskanner, gjennom reflekterende objekt og tilbake til

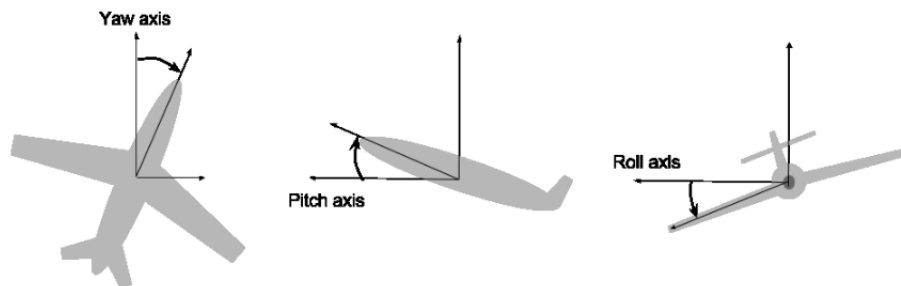
skanneren. n betraktes å være omtrent 1.00025 og er en refraksjonsfaktor som beskriver lysets påvirkning av temperatur, trykk og luftfuktighet. (Vosselman and Maas, 2010)

De fleste systemer opererer med en repetisjonsfrekvens mellom 2 og 80 kHz, som tilsvarer mellom 2000 og 8000 skudd i sekundet. Frekvensen er avgjørende for punkttettheten på bakken. Punkttettheten er nærmere utredet i 2.1.2.

Posisjonering i rommet

En kombinasjon av INS (Inertial Navigation System) og GPS beregner hvor flyet befinner seg i rommet.

INS består av bevegelsessensorer som akselerometer (hastighetsendring) og gyroskop (rotasjon). Sensorene opererer vanligvis med 200 målinger per sekund. Systemets oppgave er å måle den lokale rotasjonen til flyet i 3 dimensjoner, herunder *roll*, *pitch* og *yaw*. (Vosselman and Maas, 2010)



Figur 2.4: Lokal orientering av flyet. (Vosselman and Maas, 2010)

I tillegg til rotasjonsdata kan INS også beregne posisjon basert på akselerasjonsdata (akselerometer). Men på grunn av naturlig drift i INS systemet vil posisjonen miste nøyaktighet over tid. GPS tilbyr derimot koordinater med høyere nøyaktighet enn INS, men med lavere oppdateringsfrekvens (2 Hz). Derfor brukes GPS og INS i kombinasjon for å oppnå koordinater. Kombinasjonen er fordelaktig fordi systemene utligner hverandre i form av målefrekvens og posisjonsoppdatering. (Rasmussen, 2000)

I tillegg mottar GPS-antennen i flyet korreksjoner fra differensielle GNSS stasjoner på bakken med høyere posisjonsnøyaktighet. Korreksjonene som mottas kompenserer for atmosfæriske effekter og annet støy som kan påvirke posisjonsoppdatering fra GPS-antenne. (Vosselman and Maas, 2010)

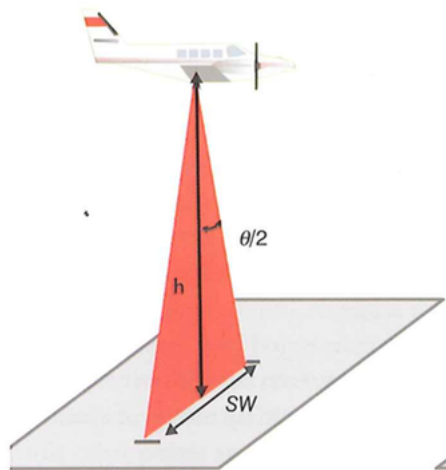
Et kalmanfilter er ofte benyttet for sammenkobling av GPS og INS. Filteret er en rekursiv algoritme som statistisk beregner posisjon basert på INS og GPS. Filteret kan brukes i sanntid og i tillegg gi statistiske beregninger hvor flyet befinner seg inntil neste posisjonsoppdatering fra GPS (postprosessering). (Vosselman and Maas, 2010)

2.1.2 Egenskaper i laserskanner

Skannevinkel og skannebredde

En av laserskannerens egenskaper er at den kan sveipe i en vinkel fra nadir som kalles skannevinkel θ . Med "swath width" menes nedslagsfeltet på bakken som vist i figur 2.5 og avhenger av flyhøyde og skannevinkel.

$$SW = 2h \tan\left(\frac{\theta}{2}\right) \quad (2.2)$$

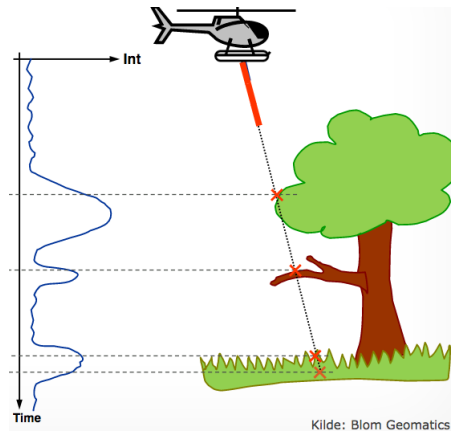


Figur 2.5: Swath Width (Vosselman and Maas, 2010)

Valg av skannevinkel er et spørsmål om økonomi og presisjon. Større skannevinkel fører til større målt areal, men går på bekostning av punkttettheten. For eksempel vil et fly med flyfart 66,6 m/s og 2Hz repetetisjonsfrekvens avlegge 30 punkter ved skannebredde 30m og 0.6 punkter/m ved skannebredde 50.

Ekko

Laserskanneren kan i tillegg til tidsmåling registrere flere returer ved samme puls. Større skannevinkel forsterker denne effekten, som vist i 2.6.

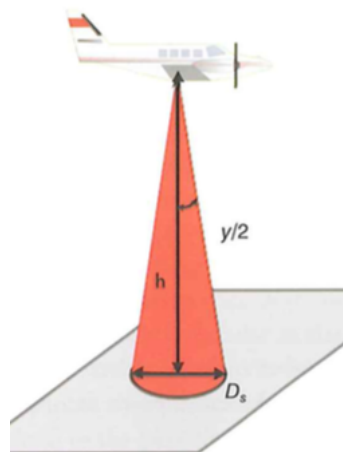


Figur 2.6: Returer fra LIDAR. (Kartverket, 2010)

Lyset kan penetrere blader i skogområder og returnere trekronen og terrengpunkt i samme retur. Antall returer avhenger også av objekt, siden skog reflekterer flere returer enn absorberende objekt som flatt vann eller asfaltvei. Denne funksjonen er derfor mye brukt for å beregne volum og areal i skogområder.

Fotavtrykk

Med fotavtrykket mener vi divergensen til laserstrålen eller nedslagsfeltet til laserstrålen på bakken.



Figur 2.7: Returer fra LIDAR. (Vosselman and Maas, 2010)

$$D_s = 2h * \tan\left(\frac{\gamma}{2}\right) \quad (2.3)$$

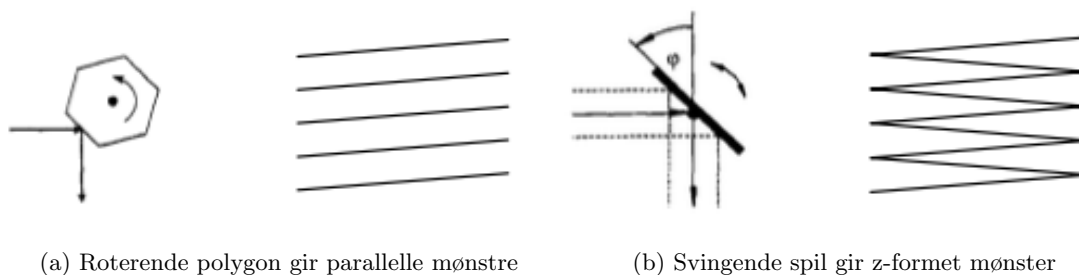
γ er divergensvinkel og h er flyhøyde. En γ på 0.25 mrad ved 1000 meters flyhøyde tilsvarer et fotavtrykk på 0.5m.

En for stor D_s kan gi negative konsekvenser for klassifisering. I tilfelle skogmodellering kan fotavtrykket bli så stort at flere ekkoer registreres uten at strålen egentlig penetrerer. En avstandsfeil oppstår dersom laserstrålen treffer et ulendt terreng og fotavtrykket utbreder seg. (Rasmussen, 2000)

Skannemønstre

Ulike skannemønstre er benyttet for å oppnå ulike målingsmønstre på bakken. Skannemønstre kan være spiralformede, parallelle og z-formede.

I Riegl produktspesifikasjon ¹ poengteres det at de fleste kommersielle laserskannere enten benytter *svingende speil* eller *roterende polygon*.



Figur 2.8: Roterende polygon og svingende speil. Figur: (Rasmussen, 2000)

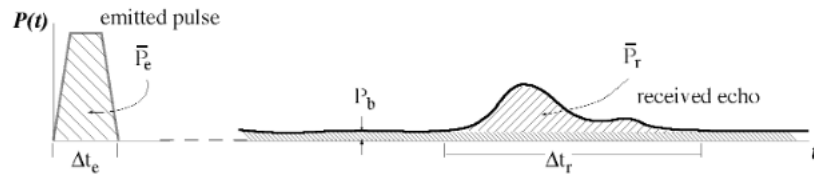
Fra figur 2.8a ser vi at skannemønsteret består av parallelle linjer. Roterende polygon speil beskrives av Riegl² som et godt alternativ for høy repetisjonsrate og jevn punkt-distribusjon på bakken.

Intensitet

¹How to read LIDAR specifications http://www.riegl.com/uploads/tx_pxpriegldownloads/Airborne_Laser_Scanning-what_to_expect_from_your_sensor_2013-03-08_03.pdf

²http://www.riegl.com/uploads/tx_pxpriegldownloads/Airborne_Laser_Scanning-what_to_expect_from_your_sensor_2013-03-08_03.pdf

En viktig egenskap når det gjelder klassifisering i LiDAR er intensitet. Når laserskanneren registrerer retursignalet av en puls, har retursignalet en annen form enn pulsen som var utsendt, som vist i figur 2.9.



Figur 2.9: Puls fra utsendt signal og mottatt signal. Figur (Hug, 1996)

Figur 2.9 viser at amplituden (kraften) er mindre i retursignalet enn i utsendt signal. I hvilken grad styrken har forandret seg mellom utsendt og reflektert puls kalles for intensitet.

Intensitet er derfor definert som styrkeforholdet mellom reflektert og utsendt puls:

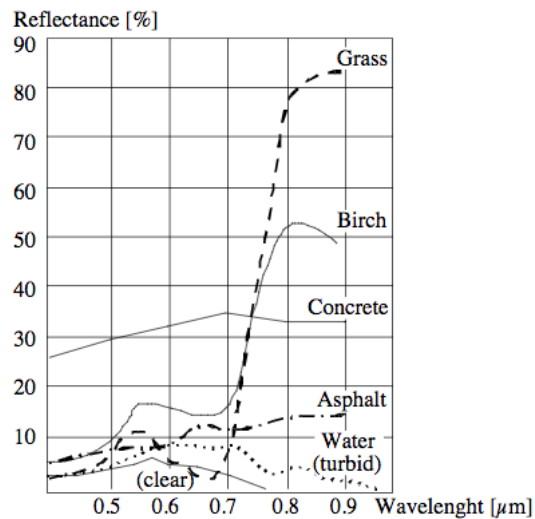
$$\rho = \frac{P_{reflektert}}{P_{utsendt} * \tau_{atm}} \quad (2.4)$$

der τ_{atm} regnes som tap av energi i atmosfæren.

Påvirkninger av intensitet

Ulike fysiske fenomener som vann, vegetasjon og bygninger blir registrert ved forskjellig intensitet. Refleksjonen av ulike objekt varierer i tillegg ved ulik bølgelengde. For eksempel vil en bølgelengde på 532nm (grønt/gult spekter) kunne penetrere vannedybder, men det nærinfrarøde båndet reflekterer vannoverflaten (Hug, 1996).

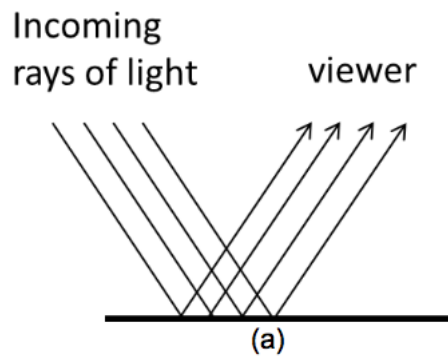
Figur 2.10 viser intensiteten til ulike fysiske objekter ved synlig og nærinfrarødt lys:



Figur 2.10: Refleksjon fra ulike objekter ved ulike bølgelengder. Kilde: (Hug, 1996)

I figuren ser vi at vann reflekteres relativt lite og enda mindre ved økende bølgelengde (nærinfrarødt; 800-1500nm). Asphalt reflekterer noe mer enn vann mens vegetasjon reflekteres mest i det nærinfrarøde spekteret.

Som følge av at nærinfrarødt lys ikke penetrerer vann, har vann i tillegg en karakteristisk spredning. Figur 2.11 viser hvordan laserstrålene kan reflekteres bort fra laserstrålene i et stille vann.



Figur 2.11: Refleksjonsmønsteret i stille vann. (Toscano, 2015)

Fra figur 2.11 kan vi også se hvordan økt skannevinkel fører til mindre refleksjon og uregistrerte områder. Som nevnt i innledningen, er unntaket når skannevinkel er rundt 0 og en oppnår unaturlig høye intensitetsverdier.

I tillegg til reflekterende objekt og skannevinkel er det en annen faktor som kan senke

amplituden i $P_{reflektert}$. Det er økt diameter i fotavtrykket, som oppstår ved økt flyhøyde
 2.2.1. Økt diameter fører til at mindre energi blir reflektert tilbake til laserskanner. (Hug, 1996)

Punkttetthet

Punkttettheten regnes ut med PRF repetisjonsfrekvens, SW skannevinkel, v flyfart og θ skannevinkel ved tiden t .

$$F = v * t * SW \quad (2.5)$$

$$P = \frac{PRF * t}{F} = \frac{PRF}{v * SW} = \frac{PRF}{2vh * \tan \frac{\theta}{2}} \quad (2.6)$$

(Vosselman and Maas, 2010)

Forklart med ord blir punkttettheten lavere med høyere skannevinkel, større flyfart og mindre repetisjonsfrekvens.

I NDH er det fastsatt en punkttetthet på minst 2 punkt per kvadratmeter. Hvis en skal beregne repetisjonsfrekvensen til ønsket punkttetthet $2 \text{ pkt}/\text{m}^2$ (fastsatt i NDH), kan vi snu på formel 2.6. Gitt at et fly har skannevinkel 20° , flyhøyde 500m og flyfart 70 m/s blir repetisjonsfrekvensen 2468p per sekund. (Vosselman and Maas, 2010)

Formel 2.6 sier bare noe om punkttettheten sendt ut fra laserskanneren og i forhold til flyhøyde og skannevinkel. På bakken vil derimot punktdistribusjonen av variere med skannemønster, reflekterende objekt, skannevinkel, fotavtrykk, topografi, flyhøyde.

2.2 Datagrunnlag

2.2.1 Riegl LMS-Q1560



Figur 2.12: Riegl LMS-Q1560

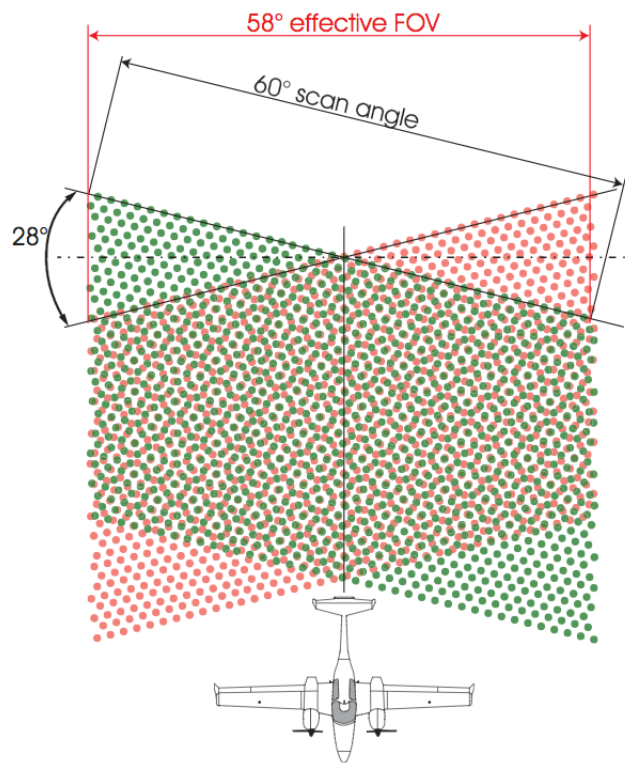
Laserskanneren benyttet for datagrunnlaget i denne masteroppgaven heter Riegl LMS-Q1560.³

Skanneren tilbyr opptil 60°skannevinkel og 800Khz repetisjonsrate. Skanneren har kapasitet til å registrere opptil 10 returer per puls og opererer i det nærinfrarøde båndet (800-1400nm). Divergensvinkel er 25mrad. Se .

Skanneren regnes som et godt alternativ når ønsker høy punkttetthet og måling av store arealer, dvs høye skannevinkler.

Komponentene i laserskanningsprosessen beskrevet i 2.1.1 er ikke separert med i Riegl Q1560, men integrert i laserskannerens maskinvare. Nøyaktigheten i dette systemet beskrives av produsenten til å være 0.05m - 0.3m i grunnriss og 0.2m i avstand.

Roterende polygon 2.1.2 er benyttet som skanningsmekanisme. Skanneren karakteriseres ved å ha to kanaler (to laserskannere) som opererer i luften samtidig. Kanalene er tiltet mot hverandre 28°og muliggjør høyere oppløsning i overlappsområdet, som illustrert i figur 2.13.



Figur 2.13: Riegl Skannemønster

Ifølge produktspesifikasjonen kompenserer dette skanningsmønsteret for sviktende dis-

³http://www.riegl.com/uploads/tx_pxpriegl/downloads/LMS-Q1560_at_a_glance_2015-08-24.pdf

tribusjon i områder ved ujevn topografi.

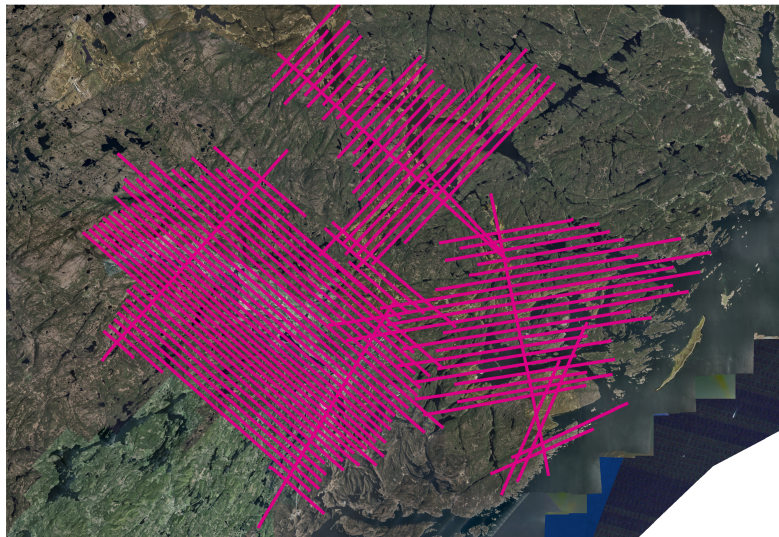
I tillegg til skannevinkel i bredden, kan Riegl Q1560 skanne bakover og framover i 8°. Dette muliggjør 3D modellering av bygninger og andre fysiske objekter. Det er ikke klart om denne funksjonen benyttes i vårt datasett.

Produktspesifikasjonen⁴ oppsummerer målinger av innsjøkanter og elver med Riegl Q1560 som "excellent" i sammenligningen av produsentens ulike laserskannere.

2.2.2 Aktuelt datasett

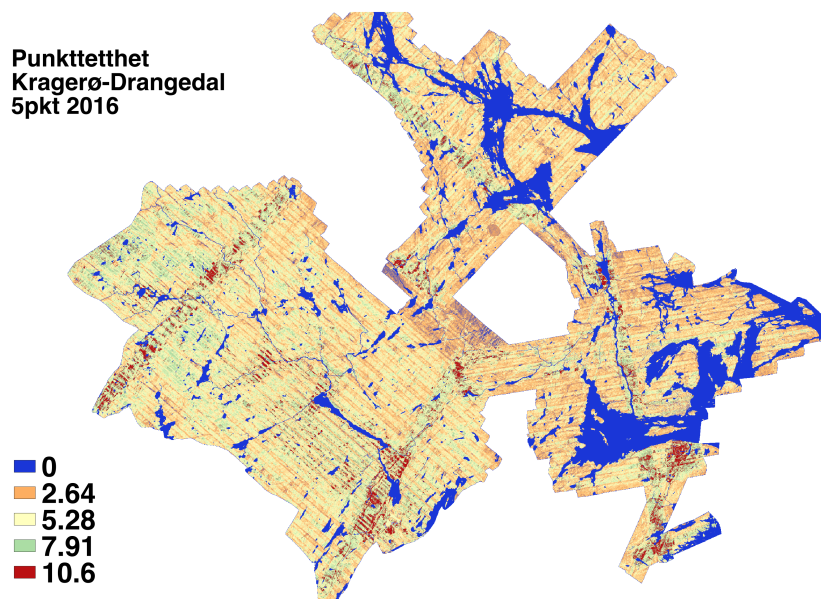
Datasettet benyttet i denne masteroppgaven er tatt fra Kragerø-Drangedal. Datasettet ble valgt på grunn av Riegl laserskanner, høy punkttetthet og nylig datafangst. I 2017 er det planlagt flere målinger med Riegl Q1560. Figurene nedenfor viser flystripene og punkttettheten til datasettet (Terratec, 2016).

Flystriper Kragerø-Drangedal 5pkt



Figur 2.14: Flystriper Kragerø Drangedal 5 punkt per kvadratmeter datasett(Terratec, 2016)

⁴http://www.riegl.com/uploads/tx_pxpriegldownloads/DataSheet_LMS-Q1560_2015-03-19.pdf



Figur 2.15: Punkttetthet Kragerø Drangedal 5 pkt datasett (Terratec, 2016)

De blå områdene i 2.15 med verdi 0 er vann i Kragerø-Drangedal. Mer spesifikt er det FKB-vann som er lagt på punktskyen for å dekke over sviktende punktdistribusjon i vannflatene. I aktuelle lasfiler brukt i denne oppgaven finner vi likevel vannpunkter som forventet.

Leveransene

I leveransene blir lasfilene avgrenset til områder på 800m*600m, som er et areal på 480000m². Koordinatsystemet er Euref89 UTM 32, geoidmodell Href2016A og vertikalt høydesystem NN2000. ifølge prosjektrapporten, var været skyfritt og det var lite vind på måletidspunktet.

Produsenten har utført justeringer og kontroller på datasettet som forbedrer datasettets nøyaktighet både i grunnriss og høyde. Ifølge statistikk i prosjektrapport (Terratec, 2016) kan vi forvente i gjennomsnitt 1.1cm avvik i høyde. Grunnrissnøyaktigheten varierer ulike steder på kartet, men ligger rundt 1cm (Terratec, 2016).

Dette kapitlet avrundes med en oversikt over de spesifikke egenskapene som finnes per 3D punkt i gjeldende datasett:

LAZ Kragerø-Drangedal	
Classification	Klassifiseringskode
EdgeOfFlightLine	Indikerer enden av flystripe før retningsforandring ved en verdi 0-1
GpsTime	GPS tid
Intensity	Intensitetsverdi
NumberOfReturns	Antall returer i punkt
PointId	Punktnummer i datasettet
PointSourceId	ID som sier hvilket datasett som er punktets opphav
ReturnNumber	Returnnummer for en gitt puls
ScanAngleRank	Skannvinkel
ScanDirectionFlag	Retningen speilet var vendt ved tidspunkt av utsendt puls
UserData	Uspesifisert
X	Øst Euref89 UTM 32
Y	Nord Euref89 UTM 32
Z	Høyde NN2000

I neste kapittel skal teori som brukes på disse lasfilene fra flybåren laserskanning for klassifisering av vannflater.

3 Teori

Mens forrige kapittel handlet om teknologien bak flybåren laserskanning og datasettets egenskaper, skal vi i dette kapitlet se på kjent teori som anvendes på datasettet for å identifisere vann i terrenget. Teorien bruker vi i senere kapittel der metodene beskrives.

3.1 Digitale Terrengmodeller

Digitale terrengmodeller beskriver formen til en terrengoverflate og karakteriseres ved at:

- Datagrunnlaget er et utvalg av diskrete datapunkter representert ved (x,y,z) .
- En interpolasjonsmetode benyttes for å beregne z (høyde) i et vilkårlig punkt når (x,y) er gitt.

(Bjørke, 2010)

I denne oppgaven benyttes 2.5 dimensjonale terrengmodeller. I disse modellene er det kun én høydeverdi per (x,y) koordinat. Naturligvis er det flere (x,y) koordinater per høydeverdi, disse representerer like høyder i terrenget.

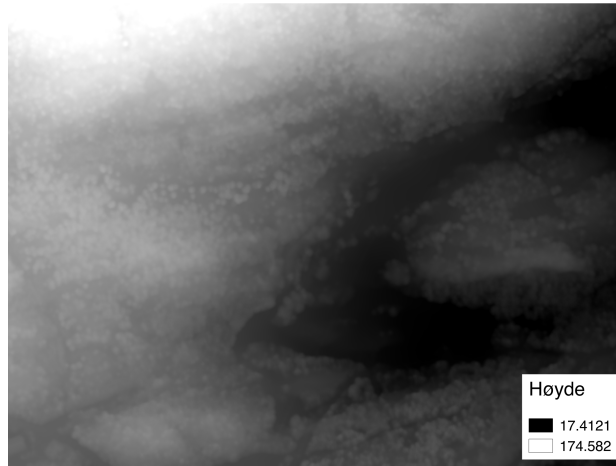
$$f(x, y) : (x, y) \rightarrow h \quad (3.1)$$

(Bjørke, 2010)

Slike digitale terrengmodeller kan representeres ved et rasterformat eller vektorformat. Med en vektorrepresentasjon menes et datasett bestående av TIN (Triangular Irregular Network), et trekantnett av triangler hvor hvert kontrollpunkt blir representert som et hjørne. (Bjørke, 2010)

For klassifisering av vannflater i denne oppgaven benyttes digitale terrengmodeller i rasterformat. En rasterrepresentasjon er et rutenett bestående av like store kvadratiske ruter, der hver rute representeres ved en høydeverdi. Oppløsningen i en rasterrepresentasjon blir bestemt som lengden av en rutekant. (Bjørke, 2010)

Siden metodene som presenteres detekterer vannflater ved arealberegning, omtaler vi i denne oppgaven ruteoppløsning (lengden på sidekant) som rutestørrelse med benevnning m^2 .



Figur 3.1: Et eksempel på en digital terrengmodell i et rasterformat, her representert ved gråskalaverdier fra 17moh-174moh

3.1.1 Interpolasjonsmetoder

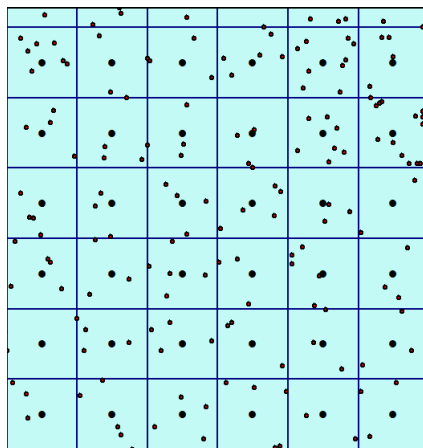
En interpolasjonsmetode er en beregningsprosedyre som regner ukjent z i punkter hvor (x,y) er gitt og z er ukjent. Interpolatorens oppgave er å definere en kontinuerlig terrengoverflate basert på en mengde målte kontrollpunkter (x,y,z) , der z er en kontinuerlig variabel. Den kontinuerlige variabelen kan være høyde, skannevinkel, intensitet eller en annen verdi som er målt som z i et kontrollpunkt. (Bjørke, 2010)

I valg av interpolator tar en hensyn til punkttetthet, prosesseringstid og interpolatorens nøyaktighet. I datasett med stor punkttetthet og stort antall kontrollpunkter kan en for eksempel benytte mindre nøyaktige interpolatorer. Interpolasjonsmetoder med kort prosesseringstid går generelt på bekostning av nøyaktighet. (Bjørke, 2010)

(O’Sullivan and Unwin, 2014) skiller mellom deterministiske og geostatistiske interpolasjonsmetoder. Felles er at begge tar i betraktning distanse beregnet fra (x,y) med ukjent z . Med geostatistiske metoder tar en hensyn til både distanse og romlig autokorrelasjon, samtidig som justering av z er tillatt. Et statistisk estimat gir en indikasjon på feil begått i interpolasjonen. En geostatistisk metode, som for eksempel kriging, regnes derfor som mer nøyaktig interpolator enn deterministiske metoder (Burger and Burge, 2016). I *deterministiske* metoder tar en bare i betraktning distanse fra kontrollpunkt, i tillegg til å forutsette at samtlige kontrollpunkter (x,y,z) er korrekte.

3.1.2 Inverse Distance Weighting

”Inverse Distance Weighting” er deterministisk interpolasjonsmetode med hensikt å finne lokalt gjennomsnitt av kontrollpunkter i et søkefelt r . Først legges raster med valgt oppløsning over et nett av (x,y,z) kontrollpunkter. (O’Sullivan and Unwin, 2014)



Figur 3.2: Et rutenett med $1m^2$ rutestørrelse lagt på punktsky av kontrollpunkter. Kontrollpunktene her er tatt fra ett av $5 \frac{Pkt}{m^2}$ datasett fra Kragerø-Drangedal

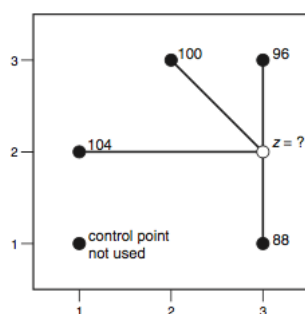
Man slår en sirkel rundt sentrum i hver rute, og bestemmer z verdier basert på kontrollpunkt som faller innenfor en radius r . Ut fra målte distanser tillegges kontrollpunktene ulik vekt. Med andre ord får punkter som ligger langt unna, mindre vekt enn nære. Formlene for IDW illustreres i 3.3 og prinsippet for IDW i figur 3.4.

$$\hat{z}_j = \frac{1}{m} \sum_{i=1}^m *z_i \quad (3.2)$$

$$\omega_{ij} = \frac{1}{d_{ij}} \quad (3.3)$$

$$\hat{z}_j = \frac{1}{m} \sum_{i=1}^m \omega_{ij} * z_i \quad (3.4)$$

Figur 3.3: Formler IDW (O'Sullivan and Unwin, 2014)



Figur 3.4: IDW eksempel fra (O’Sullivan and Unwin, 2014)

Input i IDW er radius og rutestørrelse/oppløsning. Fordi flere punkter tas med i beregningen ved økt radius, må større prosesseringstid påberegnes. På en annen side vil dette bidra til å glatte avledede verdier av z . Radiusen bør øke dersom punkttettheten går ned, slik at en får nok kontrollpunkter innenfor sirkelen.

Økt rutestørrelse vil også gi glattere overflate, som følge av det opptar ”mer plass” i modellen. Valg av rutestørrelse velges generelt ut fra hvor mye detaljer en ønsker å oppnå. Liten rutestørrelse vil føre til mer detaljert terrengmodell, gitt stor nok punkttetthet, men stiller større krav til regnekapasitet.

IDW kalles en *eksakt* interpolator fordi at den garanterer at alle kontrollpunkter er inkludert (dersom interpolasjonssirkelen omslutter ruten). Hvis det oppstår 0 under brøkstreken i formel 3.3, sørger algoritmen for at nærmeste kontrollpunkt blir interpolert verdi.

IDW har likevel svakheter. Siden det regnes gjennomsnittsverdi, vil interpolert verdi aldri være høyere enn maksimums- eller lavere enn minimumsverdien av kontrollpunktene i sirkelen. I tillegg kommer IDW dårlig ut i datasett med clustring og stor variasjon i punkttetthet. Da vil utenforliggende punkter få mindre vekt og gjerne gi urealistiske måleresultat. (Burger and Burge, 2016)

I datasett med høy punkttetthet og jevn punktdistribusjon er regnes derimot IDW et godt alternativ av mange, for eksempel Liu (2008).

4 Avgrensninger

Til nå har vi gjennomgått teori for flybåren laserskanning som ligger til grunn for klassifisering av vannflater med LiDAR-data. Dette kapitlet handler om avgrensninger, som gjenspeiler hvilke resultater vi ønsker å oppnå med presenterte metoder.

4.0.1 Avgrensninger i Hydro Flattening

(U.S. Geological Survey Standards, 2014b) setter følgende kriterier for teknikken "Hydro Flattening":

- Vannflater med areal lik eller større enn $8000 m^2$ skal flates ut.
- Vannflater som flates ut skal ha en høydeverdi som representerer hele overflaten
- Vannknekklinjene skal være på eller under det omkringliggende landskapet
- Lange vannforekomster - Elver, fjorder, vassdrag eller andre vannforekomster der vann har varierende høydeforskjeller skal behandles som elver

Med hensyn til kriteriene satt i 4.0.1 klassifiserer ikke presenterte metoder hav, bekker, elver, vassdrag eller noen form for rennende vann.

4.0.2 Andre avgrensninger

Selv om nedre arealgrense for vann som flates ut er $8000 m^2$ i 4.0.1, senkes dette kriteriet til $1000m^2$ for å tilpasse oss FKB-vann. I praksis betyr det at metoden detekterer vannflater som er større enn $1000m^2$.

I tillegg til kravene i 4.0.1 har metoden disse avgrensningene:

- Eventuelle feil og unøyaktigheter i målemetoden
- Bruker datasettet som det er, uten bearbeiding
- Forutsetter at dataprodusentens metoder som ekskluderer skog, bygninger og andre klasser som ikke tilhører terreng, er korrekte.
- Vi bruker ikke radiometrisk korreksjon
- Klassifiserer ikke vegetasjon i vann

- Tar ikke hensyn til konsekvenser av overlappende flystriper
- Forutsetter at eventuelle returer fra vann gir samme grunnrissnøyaktighet som ellers i terrenget.

5 Metode

Metodekapittelet beskriver metodene for klassifisering av vannflater og vannknekklinjer.

Toscano et al. (2013) er den beste metoden funnet for klassifisering av stille vann og innsjøer. Metoden regnes som et utgangspunkt for egenutviklet metode og beskrives derfor grundig i første del av dette kapittelet.

Deretter følger trinnvis gjennomgang av egenutviklet metode. Målet med denne metoden, er å få til en like god eller bedre nøyaktighet, men samtidig være raskere enn Toscano et al. (2013).

5.1 Forberedelser og valg

5.1.1 Generelle valg for metoder

Håndtering av støy og terrengpunkter

Både Toscano et al. (2013) og egenutviklet metode benytter LiDAR-genererte terrengmodeller bestående av utelukkende terrengpunkter. Produsenten som utgir gjeldende datasett har fra før utført filtreringsalgoritmer på datasettet som klassifiserer punkter som ligger tilhører terrenget. (Terratec, 2016)

I prosjektrapport Terratec (2016) står det: ”hull og spiker aksepteres kun dersom de er forårsaket av permanente vannoverflater eller med lav refleksjon”. For ordens skyld ble det i denne oppgaven utført en vasking som klassifiserer spikere og hull som ”støy” der de ellers ville vært klassifisert i terrengklassen. Terskelen for denne operasjonen var punkter som ligger 100 meter over terrenget. Eventuelle spikere og hull vil dermed representeres som uregistrerte områder etter denne operasjonen.

Eksempeldatasett

For metodedelen velges ut et terreng som er representativ for eksempler. Terrengmodellen inneholder kun ett vann.

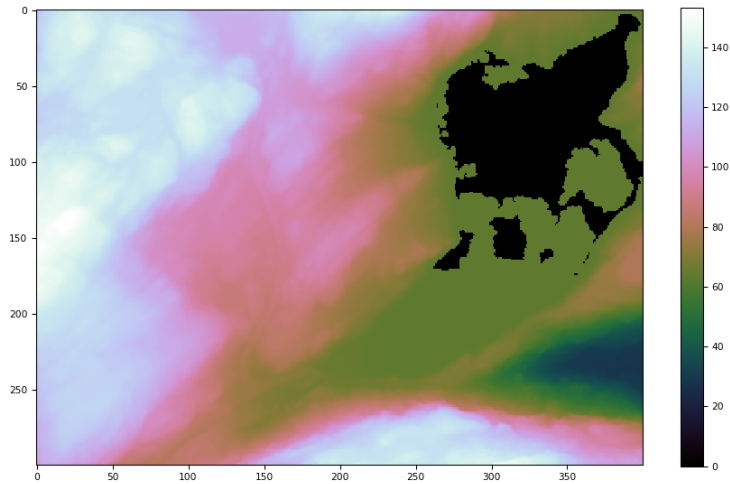


Figur 5.1: Frøvikjønn Kilde: Topografisk Norgeskart 2 gråtone

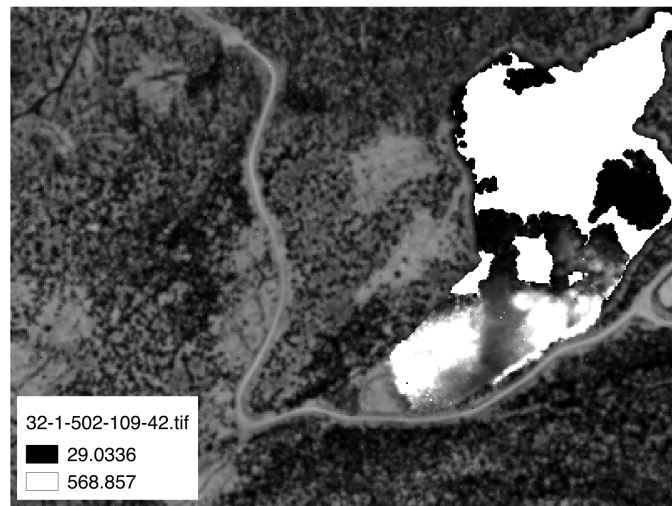
Vannet heter Frøvikjønn og ligger vest for Kilsfjorden i Kragerø. Manuelle beregninger i QGIS viser at vannet ligger omtrent 700m fra kysten. Vårt datasett omspinner omtrent hele vannet og terrenget ligger mellom 40-153 moh. I FKB-vann har vannet en høyde på 63.2moh og areal $82637m^2$.

Datasettet ble også valgt på grunn av relativt liten skannevinkel og en myraktig overflate i overgangen mellom vann og land. Langs vannet er det en vei som kan påvirke klassifiseringsresultatet. I tillegg er det en bekk som renner ut av vannet. Datafangstdato er ifølge FKB-vann 01.06.1981 og er produsert ved manuell tegning på et digitaliseringsbord Geovekst (2016).

Figur 5.2 og 5.3 illustrerer gjeldende modeller for henholdsvis høyde og intensitet.



Figur 5.2: Terrengmodell etter IDW interpolasjon



Figur 5.3: Intensitetsraster

$5 \frac{pkt}{m^2}$ datasett er benyttet for eksempler i metodedelen. For beskrivelsen av Toscano et al. (2013) er rutestørrelse $4m^2$. I henhold til punkttettheten i $5 \frac{pkt}{m^2}$ datasett tilsvarer dette 20 punkter per rute, regnet utfra laserstråler sendt fra laserskanner. I egenutviklet metode benyttes en mindre rutestørrelse for å oppnå mer detaljer i og rundt vannflatene.

I kapittel 6 skal vi teste $2 \frac{pkt}{m^2}$ datasett, samt ulike parametre i interpolator og metode. I tillegg tester vi metodene på datasett der vannknekklinjene kan være ekstra vanskelig å detektere.

Interpolator

”Inverse Distance Weighting” er benyttet interpolasjonsmetode på både intensitets - og høyderaster.

IDW (3.1.1) anbefales i datasett med høy punkttetthet og jevn punktdistribusjon Liu (2008). Siden metoden benytter returer fra vannet, velges det en *eksakt* interpolator som tar i bruk alle punkter. I tillegg regnes spikere og hull som filtrert bort før utført interpolasjonsmetode.

Benyttet radius i interpolator er 4 meter for å glatte registrerte returer på vannflater og samtidig opprettholde holdbar prosesseringstid. En glatting av vannkantene vil også forekomme, samtidig som uregistrerte områder forventes å bli mindre betydningsfull. En fordel med større radius er at det glatter ut små uregistrerte områder som ikke er vann.

5.2 Toscano 2013

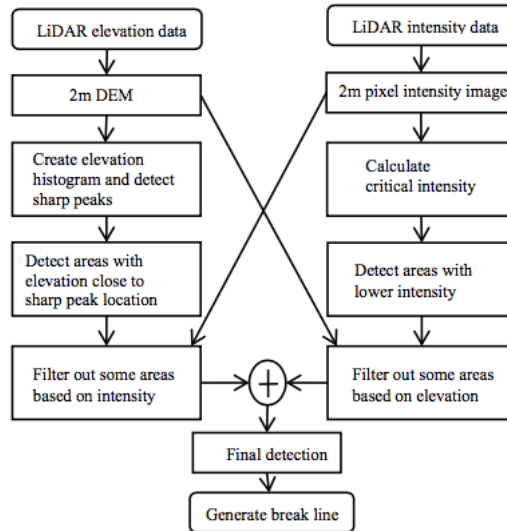
Mens vi i 5.1.1 beskrev forberedelse for begge metoder, følger nå en beskrivelse av forhold som angår Toscano et al. (2013) spesielt.

Resultatene fra Toscano et al. (2013) er basert på egen implementasjon og tolkning. På slutten av underkapittel 5.2 refereres det til avvik fra original metode.

Grunnprinsippet i både Toscano et al. (2013) og egenutviklet metode er å sammenkoble høydedata og intensitetsdata for klassifisering av vannflater. Det lages ett raster for høydeverdier (figur 5.2) og ett raster for intensitetsverdier (figur 5.3) med $4m^2$ rutestørrelser. Toscano et al. (2013) detekterer vannflater i høyderaster og intensitetsraster uavhengig av hverandre. Hvis enten høyderaster eller intensitetsraster indikerer vann, konkluderer testene at området er en potensielt vannflate.

I løpet av klassifiseringsprosessen utføres ulike tester som forkaster eller godkjenner potensielle vannflater.

Til slutt fremkommer vannknekklinjene som konturer av potensielle vannflater.



Figur 5.4: Toscano et al. (2013) Høydedata og intensitetsdata kombineres for å klassifisere vann

5.2.1 Detektering av flate områder

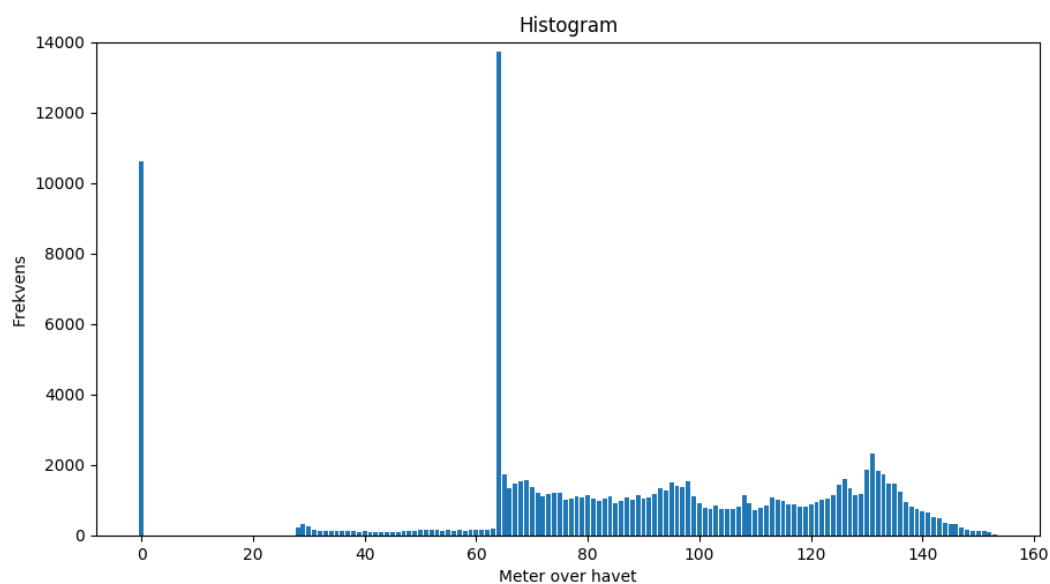
Histogram

En histogramanalyse i Toscano et al. (2013) detekterer flate områder. Histogrammet viser frekvensen av de forskjellige høydene i terrengmodellen. Ideen er at en topp i histogrammet indikerer potensiell vannflate.

Romlig indeksering i implementasjonen skjer ved at hver høydeverdi får tilført tilhørende pikselkoordinater i en liste. Hver unik høydeverdi har dermed flere pikselkoordinater (Digitale terrengmodeller 3.1).

Uregistrerte områder har etter interpolering verdien -9999 og var i figur 5.2 visualisert som de mørkeste områdene. I henhold til kriteriene satt for vannklassifisering, settes områdene til 0 før generering av histogram. Dette resulterer i en topp for 0.

Histogrammet av hele terrengmodellen vises i figur 5.5. X-aksen representerer meter over havet og y-aksen er frekvens av en gitt høydemeter. I dette histogrammet er terskelen for å havne innenfor en søyle avrundet til nærmeste høydemeter.



Figur 5.5: Histogram av Frøviktjenn

I histogram 5.5 ser en to tydelige topper i søyle 0 og i søyle 64. 0 representerer uregistrerte områder og søyle 64 representerer 64 meter over havet.

Arealet til vannet Frøviktjenn kan estimeres direkte fra histogrammet ved å multiplisere frekvensen med rutestørrelse. Søyle 0 representerer uregistrerte områder. Legger vi sammen søyle 0 og søyle 64 får vi et areal på 97444 m^2 . Sammenlignet med FKB-vann, som er 82637 m^2 , er foreløpig estimert areal for vannet 97444 m^2 .

Skalere histogrammet

Toscano et al. (2013) velger videre å øke oppløsningen i histogrammet ved å øke antall søyler langs x-aksen til 1023 søyler. Dette gjøres ved å multiplisere hver høydeverdi i høydemodellen med følgende skalafaktor:

$$Skalafaktor = \frac{1023}{\text{Høyeste verdi DTM}} \quad (5.1)$$

$$(5.2)$$

$$Skalafaktor = \quad (5.3)$$

$$x, \text{ hvis } 8 < x < 25 \quad (5.4)$$

$$8, \text{ hvis } x \leq 8 \quad (5.5)$$

$$25, \text{ hvis } x \geq 25 \quad (5.6)$$

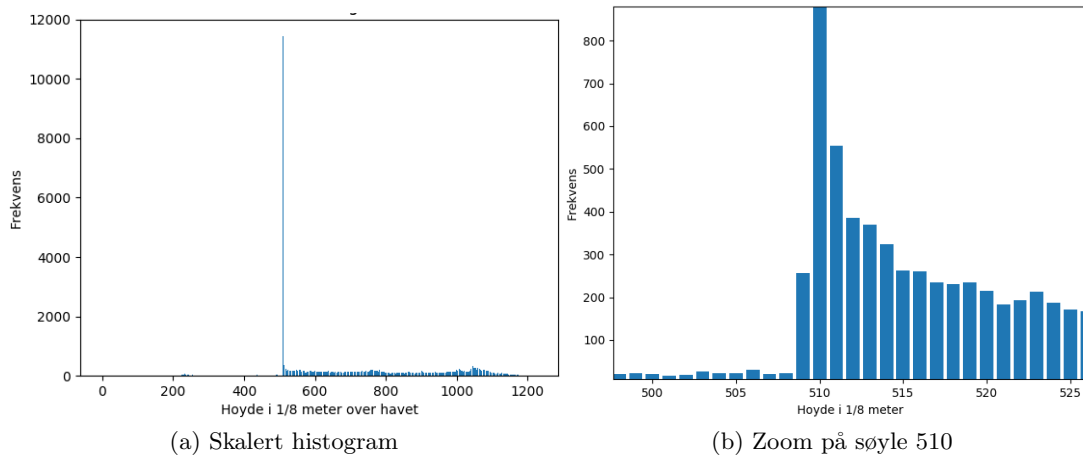
$$(5.7)$$

$$\text{For gjeldende datasett:} \quad (5.8)$$

$$Skalafaktor = \frac{1023}{153.14m} = 6.68 \frac{1}{m} \quad (5.9)$$

I henhold til likning 3.5 velges skalaverdi for gjeldende datasett til 8.

Etter skalering er hver søylebredde $\frac{1}{8}m$. Figur 5.6a nedenfor viser skalert histogram i samme zoom som i 5.5. For å illustrere effekten av skalering er det i figur 5.6b zoomet til en markant topp 510 og omkringliggende søyler. Denne søylen tilsvarer 64 i tidligere eksempel.



Figur 5.6: Skalerte histogram

Frekvensen til søyle 510 er nå redusert til 11440 ruter fordi frekvensen har fordelt seg på omkringliggende søyler. Ved å dele 510 på skalafaktor finner vi at denne søylen representerer 63.75 moh i uskalert histogram.

En eventuell forenkling av Toscano et al. (2013) metode er å avrunde hver meter til desimeter. Søylebredden reduseres da med $\frac{1}{10}m$.

En ny arealberegning av vannet kan beregnes dra skalert histogram. Legger vi sammen søyle 0 og søyle 510 finner vi et areal på $88272m^2$. Sammenlignet med beregningen fra det uskalerte histogrammet på $97444m^2$ er vi på dette punkt nærmere FKB-vann med $82637m^2$. Hvorfor arealet har vi blitt mindre, forklares med høydeoppløsning i neste avsnitt.

Høydeoppløsning

I sammenheng med det skalerte histogrammet i 5.2.1, er nødvendig å forklare en viktig parameter i denne oppgaven, som vi kaller for høydeoppløsning. Høydeoppløsningen er spennet av høydeverdier som kommer innenfor hver søyle i histogrammet 5.6a. En lavere oppløsning (større tall) gir vannflaten større areal, mens en høyere høydeoppløsning gir mindre areal.

Høydeoppløsningen i Toscano et al. (2013) kan beskrives i følgende uttrykk:

$$\text{Høydeoppløsning} = \text{Topp } m \pm \frac{1}{\text{Skalafaktor}} m \quad (5.10)$$

I gjeldende datasett med skalert histogram ser vi at høydeoppløsningen er 0.125m eller 12.5cm gitt en skalafaktor på 8 (5.1).

Kriteriene satt i 5.1 viser at valgt høydeoppløsning i histogrammet avhenger av skalafaktor. Skalafaktor økes ved liten maksimumsverdi i terrengmodellen. Største høydeoppløsning tilgjengelig i 5.1 er $1/25$ m eller 4 cm.

Høydeoppløsningen brukes også som begrep i egenutviklet metode i 5.3.

Detektering av topper

Detektering av topper i histogrammet skjer ved iterering (for-løkke) av frekvensene i histogrammet, i rekkefølgen av stigende høydeverdier langs x-aksen i 5.6a.

I løpet av iterasjonen, sjekkes følgende kriterier for å skille ”skarpe” og ”vanlige” topper:

1. Skarp topp dersom: $Topp_{frekvens_{(n-1)}} < \frac{1}{\sqrt{2}} * Topp_{frekvens_n}$
2. Skarp topp dersom: $Topp_{søylebredde_n} < \sqrt{2} * skalafaktor$

(Toscano et al., 2013)(side 101)

Forklaring av kriteriene

Kriterium 1 betyr at en krever en frekvensverdi som er 30% høyere enn frekvensverdien til foregående søyle for å bli konstatert som ”skarp”. Med andre ord, foregående søyle må være 70% av potensiell topp eller mindre for at søylen skal godkjennes eventuell topp.

Kriterium 2 betyr at vi neglisjerer søyler som ikke har søyler umiddelbart omkring seg. Hvis ikke det finnes ruter med lignende høyde, er det ikke et vann. Søylene kan derfor ikke brukes for å oppfylle kravet i kapittel 4, som sier at vannflater skal ligge på eller under det omkringliggende landskapet.

Hull og spiker representeres ved uregistrerte områder i vårt datasett (5.1.1), er disse områdene søyle 0 i vårt histogram 5.6a. Kriteriene overfor gjelder derfor ikke søyle 0.

Følgende topper ble detektert i Frøvikjtjenn:

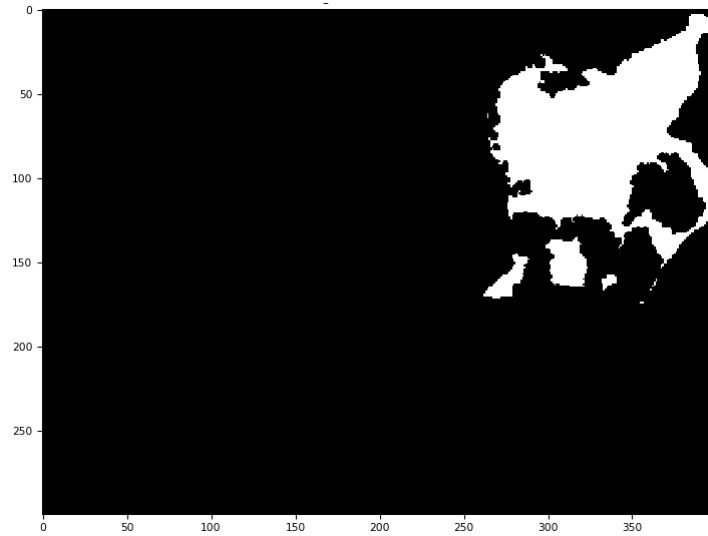
Datasett	Topper i skalert histogram
Frøvikjtjenn	[225, 226, 509, 510, 780, 563, 1153]

Figur 5.7: Topper detektert i Frøvikjtjenn

Markering av flate områder

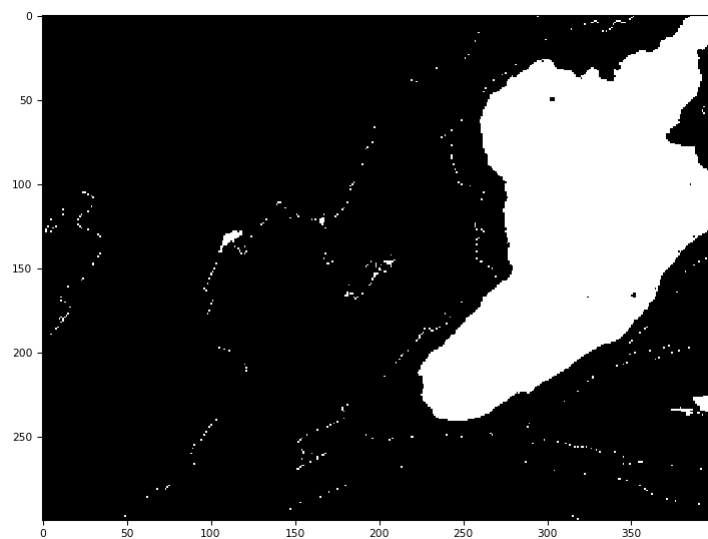
I forrige avsnitt ble det gjort rede for deteksjon av flate områder ved hjelp av histogram-analyse. I tillegg ble det drøftet hvilken høydeoppløsning en kan forvente med Toscano et al. (2013) gitt gjeldende skalafaktor. Nå følger fremgangsmåte som resulterer i potensielle vannflater med Toscano et al. (2013).

Målet for deteksjon av vannflater er et binært bilde der vann representeres ved verdi 1. Først settes uregistrerte områder til verdien 1 i et ellers tomt binært bilde, som vist i figur 5.8.



Figur 5.8: Verdi 1 representerer uregistrerte områder i et binært bilde

Uregistrerte områder fra figur 5.8 blir sammensatt med detekterte topper i histogrammet i 5.7, som vist i figur 5.9.

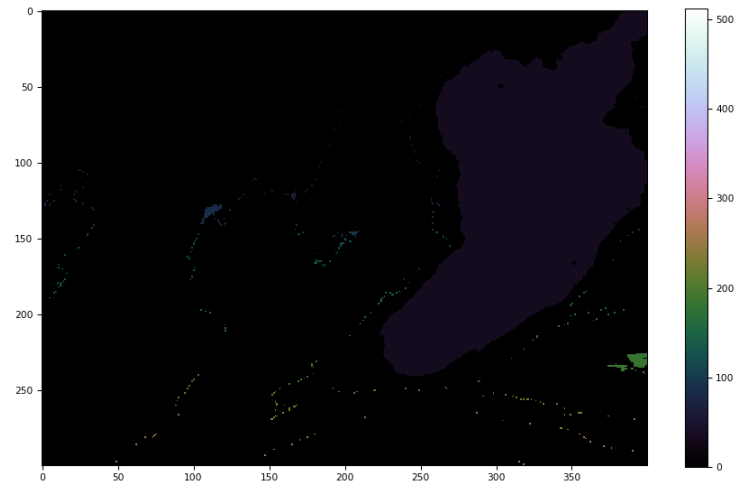


Figur 5.9: Topper i histogrammet og uregistrerte områder som representerer 1 i binært bilde

Deretter ble det utført Connected Component Labeling¹ på binært bildet. Hvis detekterte ruter (minst to) sammenstøter og har samme verdi lages det en region med tilhørende

¹<http://scikit-image.org/docs/dev/api/skimage.measure.html#skimage.measure.label>

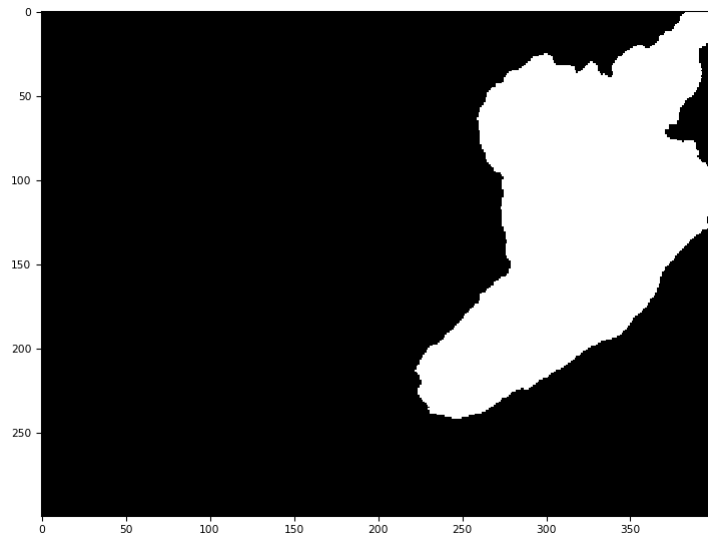
ID. Hver region-ID sees langs skalaen til høyre i figur 5.10.



Figur 5.10: Connected Component Labeling. Rutene som sammenstøter og har samme verdi får en region-ID. Antall regioner illustreres i skala

Operasjonen gjør det mulig å filtrere bort små arealer. Regioner mindre enn $1000m^2$ blir filtrert bort i henhold til avgrensningene i kapittel 4. Øyer og vegetasjon blir ikke filtrert bort, da disse områdene var 0 fra før i det binære bildet.

Etter arealfiltrering ser vannet slik ut:



Figur 5.11: Arealer mindre enn $1000m^2$ filtrert bort

I neste underkapittel skal vi se hvordan implementasjonen filtrerer bort eventuelle regio-

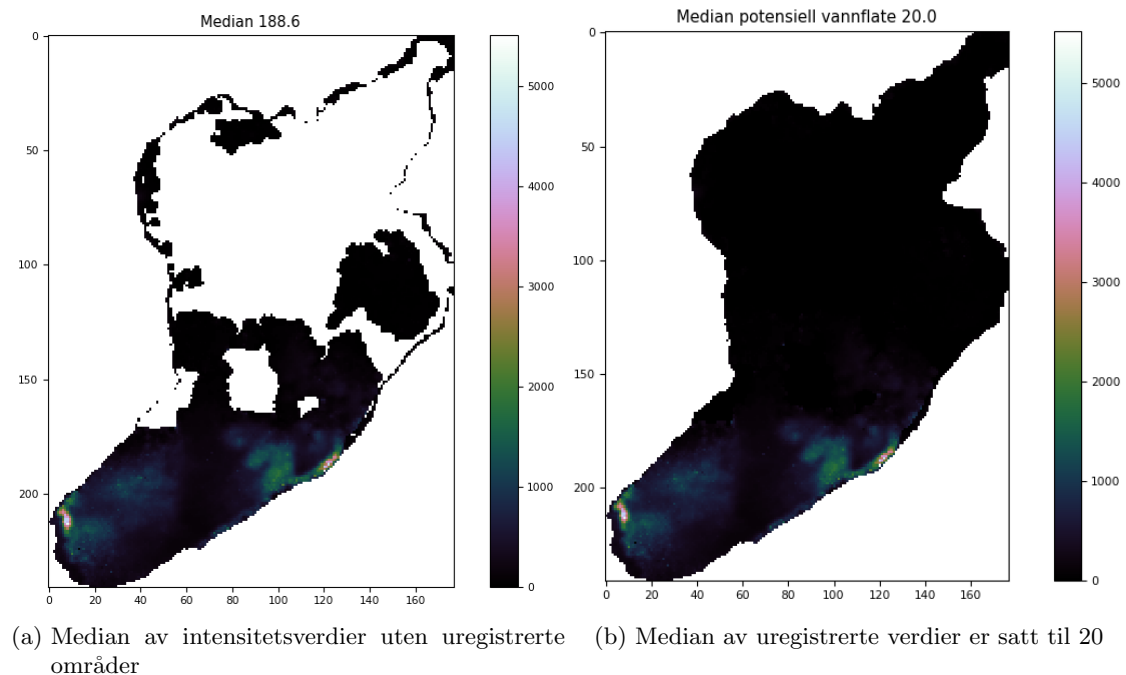
ner fra 5.10 med intensitet.

5.2.2 Mediansjekk med intensitet

Vann har generelt lavere intensitetsverdier enn vegetasjon, bygninger og omkringliggende landskaper. Etter deteksjon av flate områder i histogramanalysen, brukes denne informasjonen for å filtrere bort flate områder som ikke er vann.

Testene i Toscano et al. (2013) gjøres ved å regne medianen av intensitetsverdier i regioner (flate områder) som er detektert fra "Connected Component Labeling" fra figur 5.10. Testene er nødvendig for å sikre at flate områder kan regnes som potensielle vannflater. Dette underkapittelet forklarer hvordan disse testene er implementert, og viser noe problematikk knyttet til lav skannevinkel.

Utgangspunktet for testen er medianen av hele intensitetsrasteret, som ble regnet til verdi 149.5. Dersom medianen av intensitetsverdier i vannflaten er over denne terskelen, fjernes potensiell vannflate. Figur 5.12a viser utregning av medianen i gjeldende vannflate uten å inkludere uregistrerte områder i medianen, mens figur 5.12b viser hva som skjer med inkludering av uregistrerte områder.



Figur 5.12: Median av intensitetsverdier med og uten uregistrerte områder

Figur 5.12 viser at mediantesten i dette tilfelle forkaster Frøvikjtjenn. Årsaken er urea-

listisk høye intensitetsverdier som oppstår i stille vann ved lave skannevinkler, som er en kjent utfordring nevnt i 1.1.4. Testen godkjenner derimot vannflaten ved å sette uregistrerte verdier til 20. Denne versjonen benyttes i implementasjonen og i egenutviklet metode (figur 5.12b).

Dette underkapittelet illustrerer problemet med håndtering av intensitet i områder med lav skannevinkel i Toscano et al. (2013). En terskel som er bestemt uavhengig av skannevinkel kan føre til feilaktig filtrering av potensielle vannflater med Toscano et al. (2013). I tillegg tar ikke testen hensyn til hvor stor del av potensiell vannflate som består av speilrefleks effekt.

I denne masteroppgaven har mediansjekken blitt videreutviklet for sikrere detektering av potensielle vannflater. Med dette menes at en korrigerer intensitetsrasteret for skannevinkel før den brukes i mediansjekk. I seksjon for egenutviklet metode presenteres korreksjonen av intensitetsverdier.

Avvik fra Toscano et al. (2013) og andre feilkilder

Implementasjonen av metoden har følgende avvik:

- **Gaussianfilter** for glatting av histogrammet. En glattefunksjon på histogrammet ville detektert færre topper og gjøre histogrammet mer pålitelig.
- **Bufferanalyse** er ikke del av implementasjonen

Av andre feilkilder finner vi:

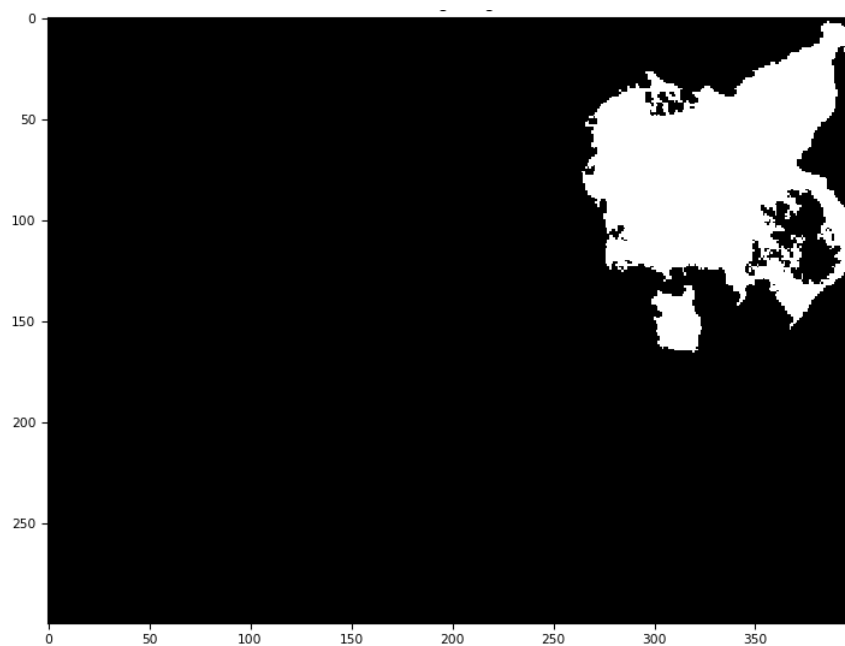
- **Radius i IDW** er ikke spesifisert i Toscano et al. (2013).
- **Området** til Toscano et al. (2013) er 500m*500m. Vårt datasett benytter 600m*800m.

5.2.3 Vannflater fra intensitetsraster - Toscana

Som nevnt i innledningen av kapitlet klassifiseres vannflater i høyderaster og i intensitetsraster hver for seg før modellene kombineres. Forrige underkapittel handlet om histogramanalyse for flate områder og til slutt intensitetssjekker som filtrerer flate områder som ikke er vann.

Siden søyler med liten frekvens ikke fremstår som topp i histogrammet, brukes intensitetsrasteret for å detektere små vannflater (Toscano et al., 2013).

Figur 5.13 viser regioner av uregistrerte områder og intensitetsverdier under eller lik 20.



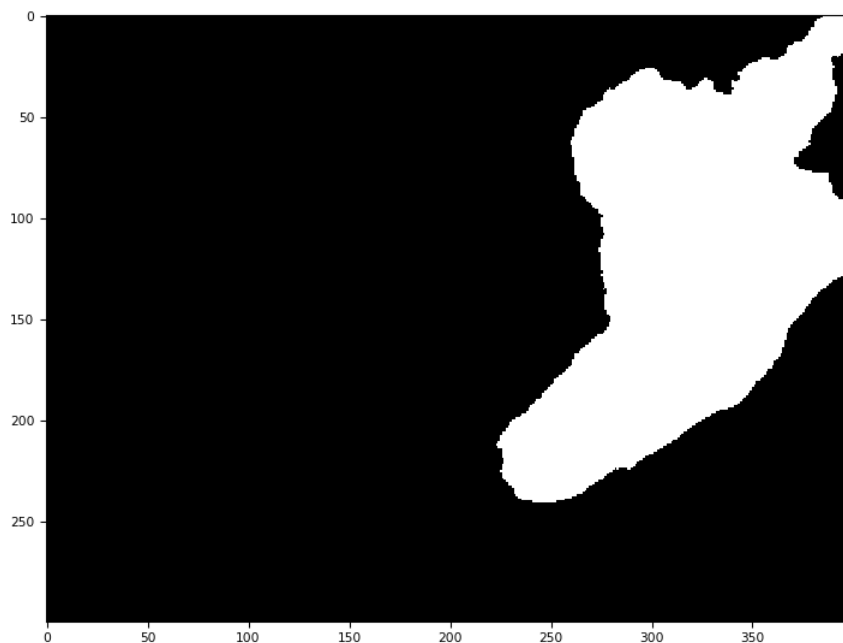
Figur 5.13: Intensitetsverdier under 20 og uregistrerte områder

I egenutviklet metode vil intensitetsrasteret se noe anderledes ut med samme parametre, fordi vi har korrigert for skannevinkel (1.1.5).

Sammensetting av høyderaster og intensitetsraster

Før generering av potensielle knekklinjer sammensettes intensitetsrasteret med potensielle vannflater fra histogramanalysen.

Dersom mediansjekken i implementasjonen blir bestått, kan vi beregne potensielle vannknekklinjer som konturer av den detekterte vannflaten. Potensiell vannflate er vist i figur 5.14.



Figur 5.14: sammensetting av uregistrerte punkter, intensitetsverdier under 20 og detektert vannflate

I dette punktet av implementasjonen av Toscano et al. (2013) har vi detektert endelig potensiell vannflate med Toscano et al. (2013). Fra dette kan vi gjøre en ny arealberegning. Antall ruter i regionen multiplisert med $4m^2$ viser et areal på $89248 m^2$ sammenlignet med $88272 m^2$ i skalert histogram i 5.2.1. Arealet til vannet har økt med $976 m^2$ ved sammensetting av lave intensitetsverdier.

Avvik fra Toscano et al. (2013) og andre feilkilder

- **Interpolasjonsmetode** for intensitet er ikke IDW, men medianen av intensitetsverdier av 3D punkter hvert $4m^2$ ruteareal. Argumentet er at en unngår variasjon i høye refleksjonsverdier. Denne interpolasjonsmetoden ble ikke testet og har derfor usikker effekt.
- **Kontinuitetstest** er ikke utført på intensitetsraster. Denne testen i vår implementering er erstattet av en soliditetstest.
- **En 18 meters** utsidebuffer og innsidebuffer er ikke implementert. Toscano et al. (2013) tar differansen mellom medianen i en innsidebuffer og medianen en utsidebuffer for en enten-eller filtrering.

5.2.4 Oppsummering Toscano et al. (2013)

Toscano et al. (2013) er til nå den beste metoden funnet for automatisk klassifisering av stille vannflater i rasterformatet. Derfor har metoden for blitt grundig beskrevet. Det har blitt gjort rede for histogramanalyse og deretter håndtering av intensitetsdata. Avhengig av skalafaktor, muliggjør histogramanalysen deteksjon av flate områder med 4 - 12.5cm høydeoppløsning.

En median av intensitetsverdier settes som kritisk terskel for enten-eller klassifikasjon av potensielle vannflater. I 5.2.2 ble det illustrert problematikk rundt skannevinkel og feilaktige filtreringer.

Toscano et al. (2013) er arealavhengig i histogramanalysen og setter en nedre arealgrense for å fjerne støy og små vannflater. Små vannflater finnes ved bruk av intensitetsdata, men på grunn av manglende korreksjon med vil denne antagelsen være noe usikker.

Toscano et al. (2013) velger å tette hull i rasteret vegetasjon som for eksempel kan indikere vegetasjon.

Metoden benytter seg av $4m^2$ ruter og gir derfor, uavhengig av punkttetthet, en usikkerhet på 2m i grunnriss. Ved valg av mindre rutestørrelse virker implementasjonen å kreve betydelig regnekapasitet. Implementasjonens indekseringsmetode (lister) regnes som en lite tidseffektiv indekseringsmetode. Et økt antall ruter vil øke kravene til regnekapasitet betraktelig.

5.3 Egenutviklet metode

Bakgrunnen for masteroppgaven er et ønske om raskere, sikrere og mer nøyaktig klassifisering av vannflater i rasterformatet. Metoden er inspirert av Toscano et al. (2013) i flere deler, blant annet mediansjekker, regionbasert klassifisering og bufferanalyse. Metoden er utviklet med hensyn på utfordringene oppsummert i 5.2.4.

I dette underkapittelet skal vi se på en filterbasert metode som et alternativ til histogramanalyse for deteksjon av flate områder. En videreutviklet versjon av Toscano et al. (2013) mediansjekk skal presenteres, der en korrigerer for skannevinkel.

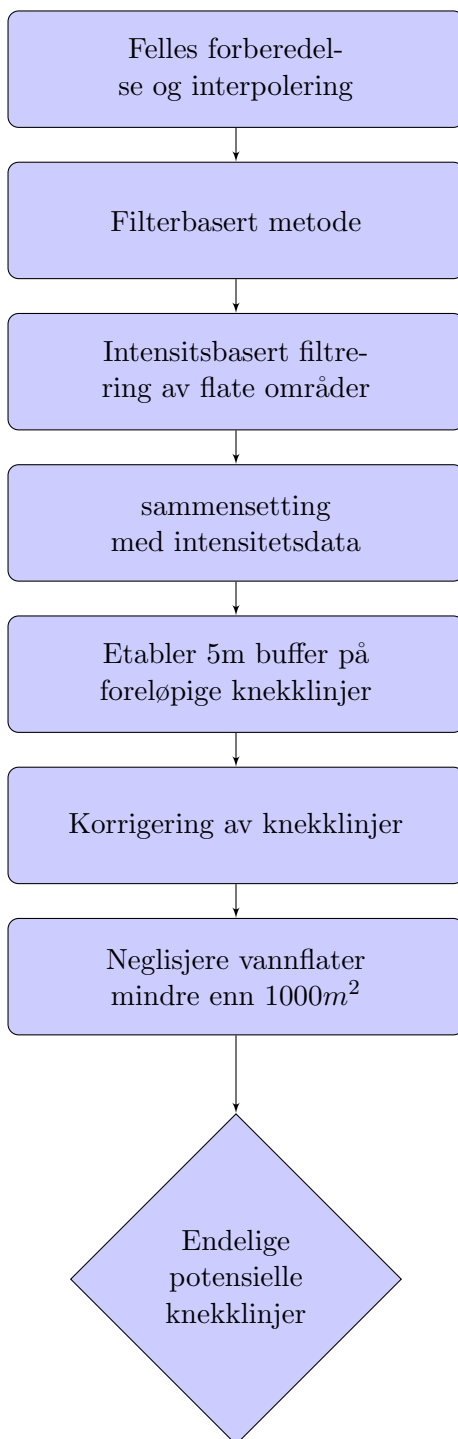
Til slutt utføres en bufferanalyse på potensielle vannknekklinjer som ved hjelp av høyderaster og intensitetsraster gir endelige potensielle vannknekklinjer og vannflater.

For gjennomføring av egenutviklet metode trengs det følgende tre typer modeller i rasterformatet:

- Terrengmodell for detektering av flate områder
- Intensitetsmodell for klassifisering
- Skannevinkelmodell

Før trinnvis gjennomgang av egenutviklet metode skal vi se på hvilke vurderinger som gjøres i forkant av denne metoden.

Klassifisering av vannflater



Figur 5.15: Egenutviklet metode

5.3.1 Valg av rutestørrelse i egenutviklet metode

I tillegg til å utvikle en raskere metode enn Toscano et al. (2013), er det ønskelig å utnytte punkttettheten som er tilgjengelig i et $5 \frac{Pkt}{m^2}$ datasett.

I bestemmelse av rutestørrelse tar en gjerne utgangspunkt i forhold som punkttetthet, areal i datasett og antall terrengpunkter (3.1.1). For eksempel vil en rutestørrelse på $0.25 m^2$ tilsvare $1.25 \frac{Pkt}{rute}$ i et $5 \frac{Pkt}{m^2}$ datasett. På bakken vil punktdistribusjonen variere, spesielt i vannområder.

I egenutviklet metode velges likevel $0.0625 m^2$ ($0.25m * 0.25m$) som rutestørrelse for å oppnå mest detaljer i knekklinjene. Mange ruter vil som følge av dette ikke inneholde kontrollpunkter. Men den filterbaserte metoden krever likevel liten rutestørrelse for å detektere potensielle vannflater med grunnrissnøyaktighet på 1m. Den trinnvise gjennomgangen vil vise hvordan dette henger sammen.

5.3.2 Filterbasert deteksjon av flate områder

Den filterbaserte metoden for deteksjon av flate områder følger en hierarkisk struktur med to trinn:

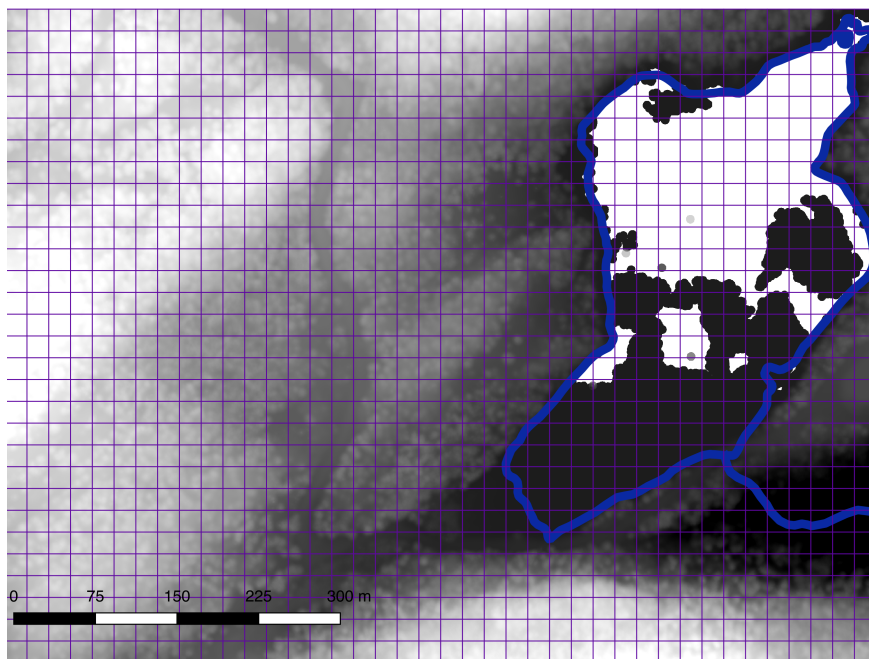
1. Terrengmodellen deles inn i blokker, der hver blokk har areal $25m^2$. Hver blokk som tilfredstiller en høydeoppløsning 12.5cm, blir tillagt verdien 1 i et binært bilde. Connected Component Labeling² blir utført på binært bilde og hver flat blokk blir identifisert ved et omliggende rektangel.
2. For hvert omliggende rektangel detektert i trinn 1, reduseres blokkstørrelse til et areal på $1m^2$. Hver blokk som tilfredstiller høydeoppløsning på 3cm, blir indikasjon på endelige flate områder.

Trinn 1: Filtrere bort store arealer som ikke er flate

Som i Toscano et al. (2013) er målet å ende opp med et binært bilde hvor flate områder representeres ved 1. Uregistrerte områder settes til 1 i henhold til antakelsen om at disse områdene er vann.

Terrengmodellen blir først delt opp i blokker på $20 * 20$ ruter. Siden hver rute i blokkene er $0.0625 m^2$ ($0.25m * 0.25m$) tilsvarer blokken et areal på $25 m^2$. Figur 5.16 viser rutenettet i trinn 1. De blå vannknekklinjene til FKB-vann er lagt på for å visualisere av den grove tilnærmingen vi ønsker i trinn 1.

²<http://scikit-image.org/docs/dev/api/skimage.measure.html#skimage.measure.label>

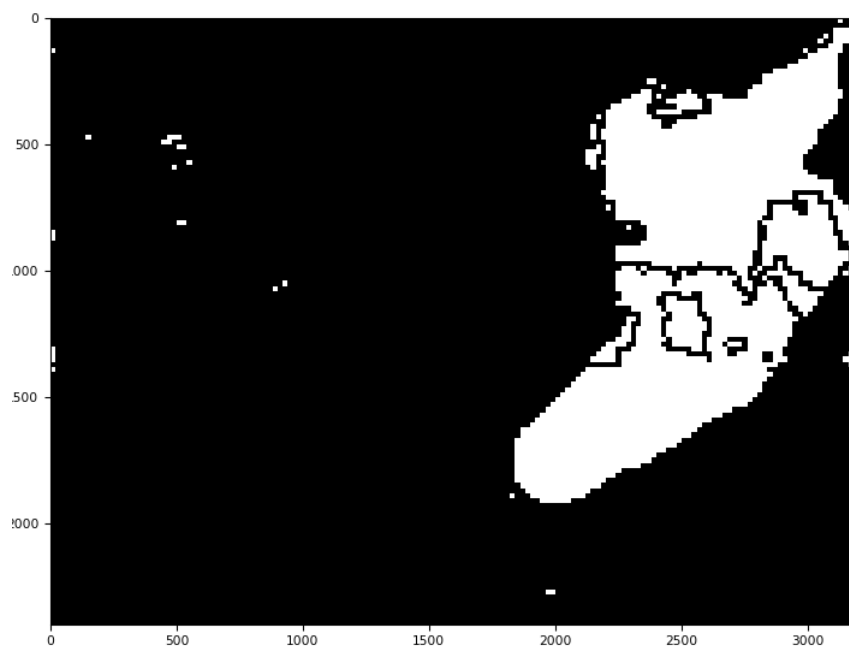


Figur 5.16: GRID over terrengmodell med 20*20 ruter blokker

For hver blokk regnes differansen mellom største og minste høydeverdi. Hvis differansen er mindre enn kravet for høydeoppløsning (5.2.1), defineres blokken på $25m^2$ som et flatt område. I praksis kan man med denne blokkstørrelsen detektere et mindre vann (et tjern), men også en åker eller en fotballbane. For hver blokk sjekkes altså følgende betingelse:

$$\min(blokk) - \max(blokk) < \text{Høydeoppløsning} \quad (5.11)$$

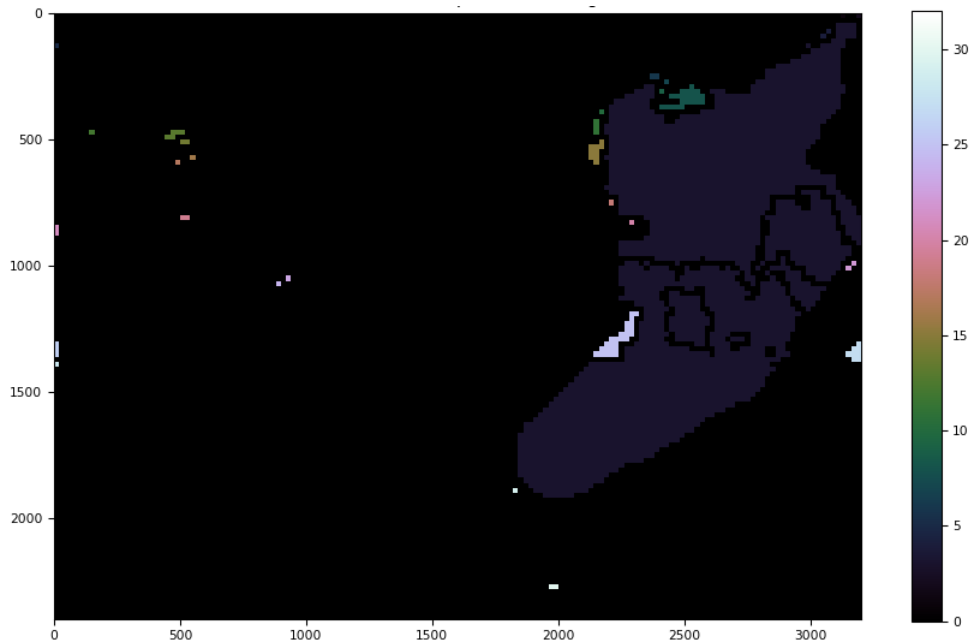
Hvis kriteriet oppfylles legges blokken med verdien 1 i et ellers tomt binært bilde. Valgt høydeoppløsning ble valgt til å være 12.5cm, som er Toscano et al. (2013) oppløsning i samme datasett. Høydeoppløsningen regnes i denne metoden som lav nok til vannflaten skaper regioner med "Connected Component Labeling".



Figur 5.17: Resultat etter filterbasert metode trinn 1 med 12.5cm høydeoppløsning og 5m grunnrissnøyaktighet

Fra 5.17 ser vi grove skisser av potensielle vannflater. De svarte konturene i vannet kommer av at betingelsen 5.11 ikke blir oppfylt i grensen mellom registrerte og uregistrerte områder.

”Connected Component Labeling” ble igjen benyttet på filtrert binært raster, som vist i figur 5.18.



Figur 5.18: Connected Component Labeling på 5.17

Som i Toscano et al. (2013), filtrerer vi bort arealer mindre enn $1000m^2$.

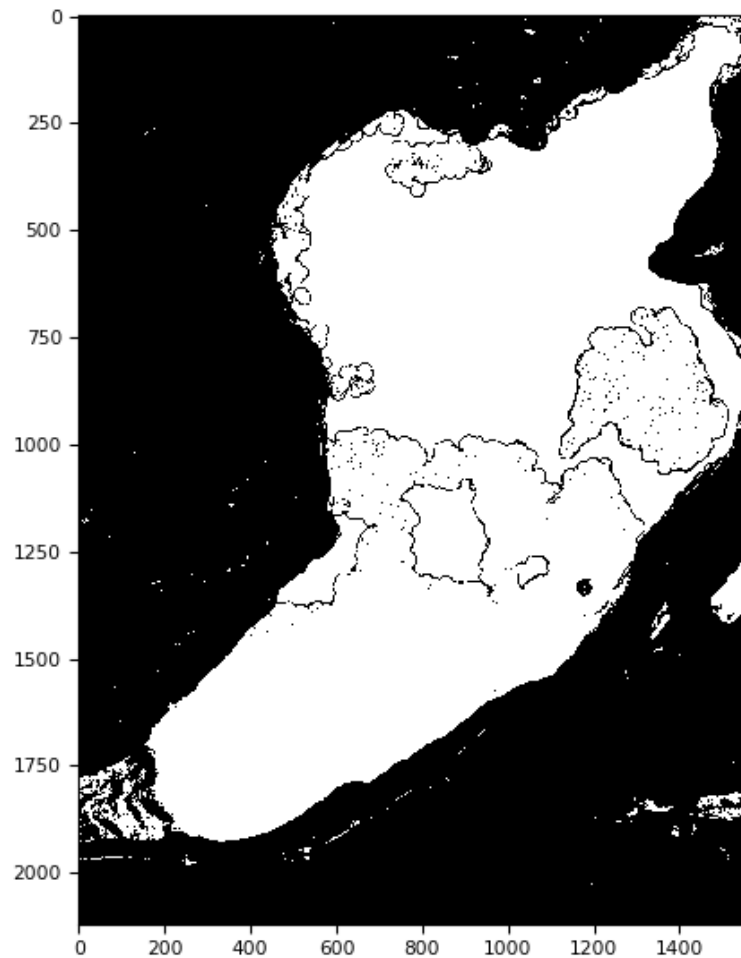
Trinn 2: Omkringliggende rektangel og større oppløsning

Med grov oppløsning har man til nå filtrert bort deler av terrenngmodellen som ikke inneholder flate områder. Fra ”Connected Component Labeling” sitter vi igjen med omkringliggende rektangel som omspinner potensielle vannflater.

Som følge av blokkstørrelse på $25m^2$, ansees endelige vannknekklinjer å ligge maksimum 5 meter unna foreløbig detekterte vannknekklinjer. Rektangelet forstørres derfor med 10 meter utover i alle retninger.

Samme prosedyre beskrevet i forrige paragraf utføres deretter innenfor omliggende rektangel til hver region som er større enn arealparameter. Her benyttes derimot en blokkstørrelse $1m^2$ ($4*4$ ruter) og høydeoppløsning på 3cm. En liten høydeoppløsning er valgt for sikre seg at hver blokk er flat og at en ekskluderer blokker med litt hellning, for eksempel en grunne eller strand. Å øke høydeoppløsningen fra 12.5cm til 3cm er en naturlig konsekvens av at færre ruter blir tatt med i beregningen (sammenlignet med Toscano et al. (2013) på 12.5cm) og at høydeforskjellene i hver blokk på $1m^2$ er mindre. Høydeoppløsningen drøftes mer utdypende i 7.6.2.

Figur 5.19 nedenfor viser resultatet av klassifiseringen med 1m grunnrissnøyaktighet og 3cm høydeoppløsning.

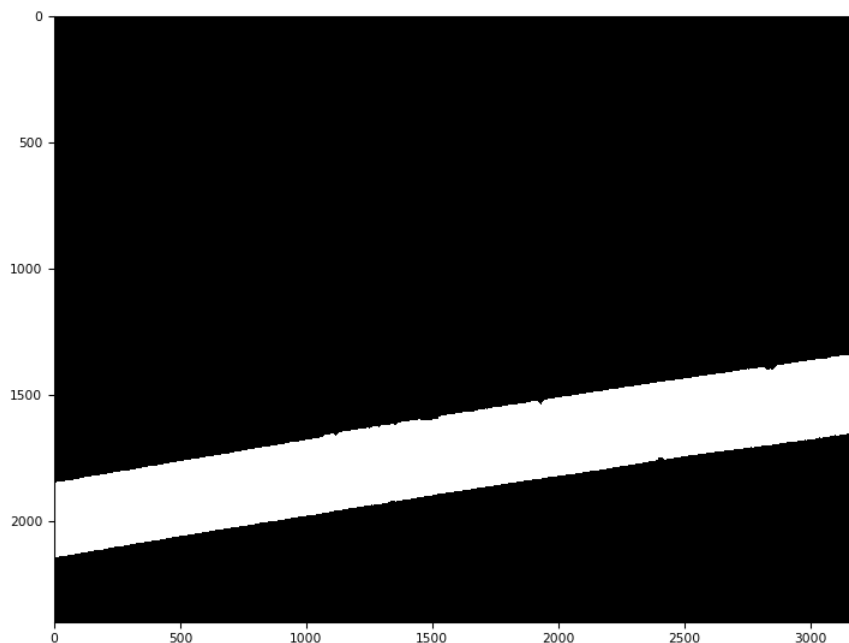


Figur 5.19: Resultat av høydedeteksjon med grunnrissnøyaktighet 1m og 3cm høydeoppløsning

Mediansjekk med intensitetsraster korrigert for skannevinkel

Som i 5.2.2 skal vi gjøre mediansjekken til Toscano et al. (2013) i egenutviklet metode. Forskjellen er at egenutviklet metode korrigerer for skannevinkel med formål å oppnå en sikrere mediansjekk.

Undersøkelser i Höfle et al. (2009) viser at skannevinkler mindre enn 2 grader ut fra normalen fører til speilrefleks i vann. Derfor er 2 grader satt som terskel for korreksjonen av intensitetsverdier.



Figur 5.20: Skannevinkel mellom -2 og 2 gir i vårt datasett en belte under flyet som er omtrent 100meter bredt

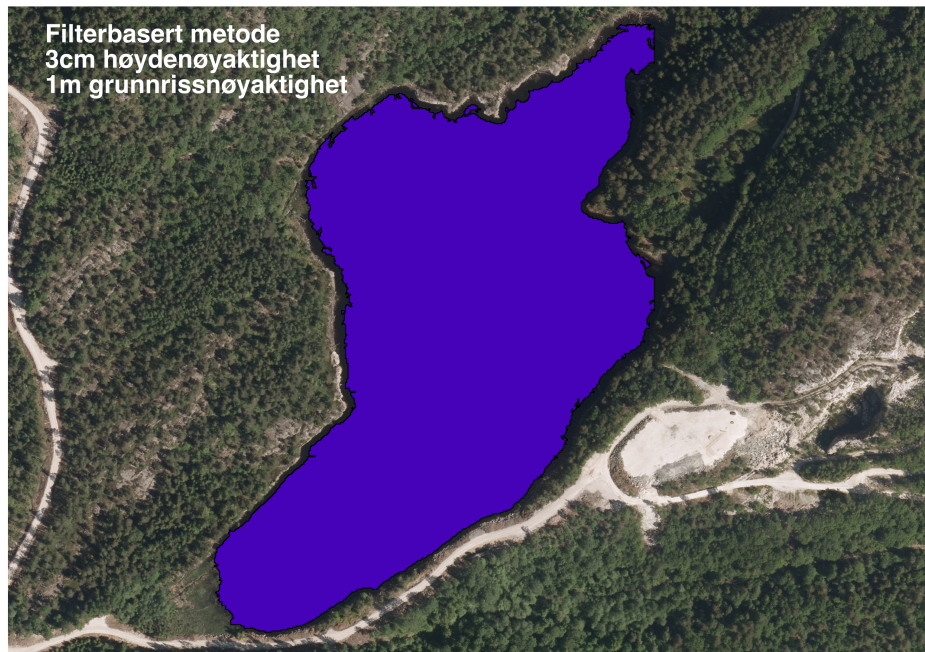
Korreksjon med skannevinkel gir bare mening hvis det brukes i potensielle vannflater. Derfor utføres korreksjon bare i vannflaten som er detektert fra figur 5.19.

Samtidig vil en sørge for terskelen for intensitetsverdier er "høy nok" til at intensitetsverdien karakteriseres som speilrefleks. Dette er for å filtrere bort regioner innenfor skannebredde (swath width: 2.1.2) i figur 5.20 som har blitt detektert av filterbasert metode, men som ikke har speilreflekseffekt.

Av erfaring i gjeldende datasett regnes 400 som en høy nok intensitetsverdi. Alle intensitetsverdier over 400 innenfor et område med skannevinkel mellom -2 og 2 blir altså omgjort til intensitetsverdi 10, som indikerer vann.

Med andre ord snur vi om på den generelle antakelsen om lav refleksjon i vann. Nå klassifiseres også høye intensitetsverdier som vann, så lenge høye intensitetsverdier finnes innenfor høy skannevinkel og detektert vannflate.

Dersom mediansjekken består (som den gjør etter korreksjonen av skannevinkel i dette tilfelle), kan vi tegne polygon av det flate området, som er resultatet av filterbasert metode. som vist i figur 5.21.



Figur 5.21: Foreløbig polygon av potensiell vannflate ved filterbasert metode

Filterbasert metode gir et foreløbig areal på $86140m^2$ på Frøvikjtjenn. Toscano et al. (2013) ble testet med $0.25m^2$ ruterstørrelse og gav et areal på $88355 m^2$ mot $82637m^2$ i FKB-vann.

Vi er nå ved et punkt i egenutviklet metode der vi er ferdig med beskrivelse av filterbasert deteksjon av flate områder.

5.3.3 Vannflater fra intensitetsraster - Egenutviklet metode

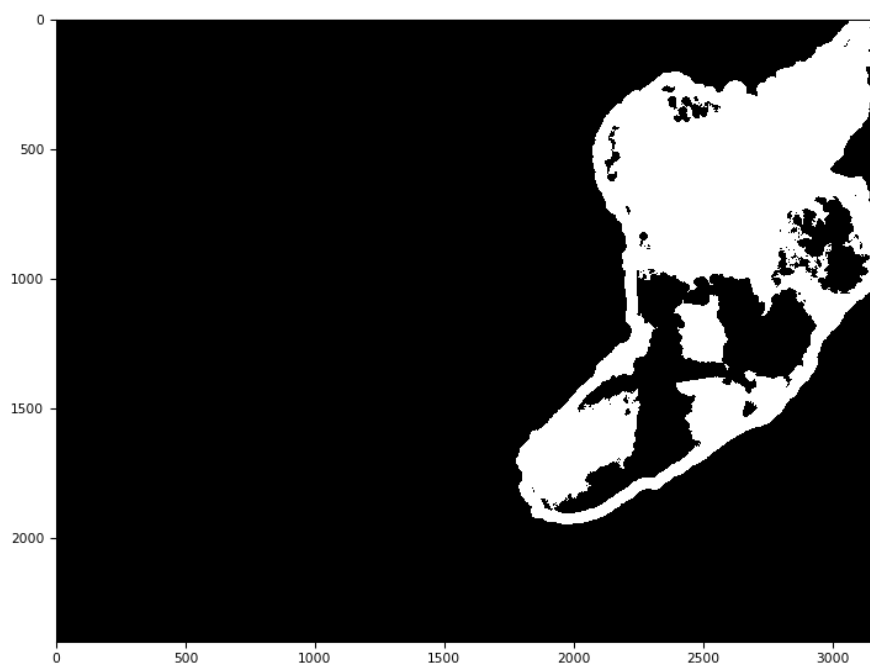
Etter gjennomføring av filterbasert metode kan vi være rimelig sikker på at hver $1m^2$ blokk har høydeforskjell under 3cm (sett bort fra feil i laserskanner) og hver region har et areal større enn $1000m^2$. Dette medfører at endelige potensielle vannknekklinjer på dette punkt ansees å ligge utenfor detekterte flate blokker.

Dette underkapittelet tilsvarer underkapittelet 5.2.3, der vi sammensatte intensitetsrasteret med flate områder fra histogrammet med Toscano et al. (2013). I det etterfølgende skal vi gjøre samme prosess med egenutviklet metode. Forskjellen er at vi bruker flate områder fra filterbasert metode istedenfor histogramanalysen, og et intensitetsraster som er korrigert for skannevinkel (forklart i 5.3.2).

Intensitetsraster korrigert for skannevinkel

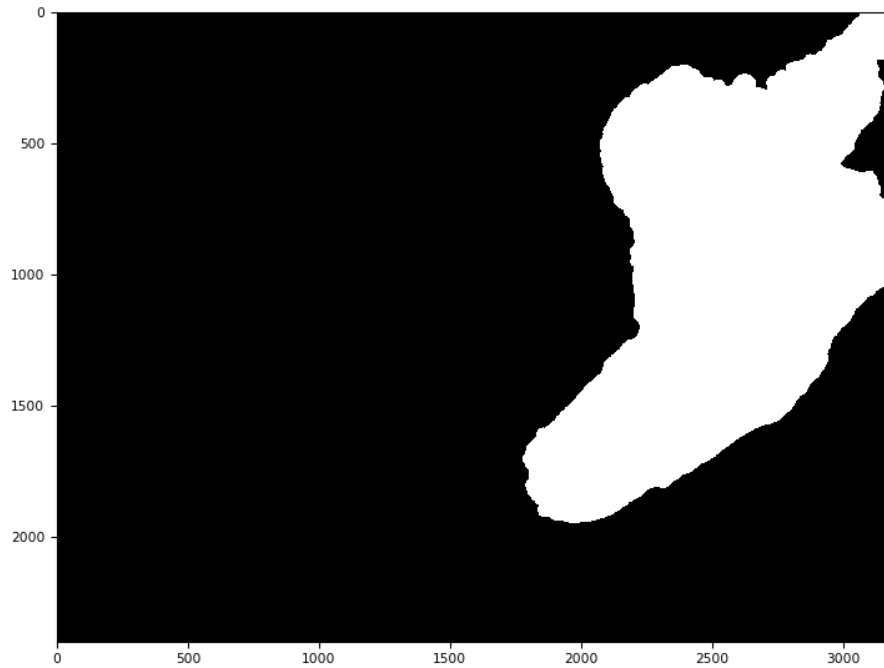
Inspirert av Toscano et al. (2013), ble 20 definert som den høyeste intensitetsverdien et vann kan ha i intensitetsrasteret.

Det endelige intensitetsrasteret i egenutviklet metode vises i figur 5.22.



Figur 5.22: Figuren viser intensitetsraster korrigert for skannevinkel fra 5.3.2 og ruter med intensitetsverdi under 20

Deretter sammennesettes intensitetsrasteret i figur 5.22 med flate områder fra filterbasert metode i 5.21. Resultatet vises i 5.22:



Figur 5.23: Sammensatt intensitetsraster og høyderaster

Verdt å merke seg er at flate områder i filterbasert metode fortsatt er blokker på $1m^2$. Intensiteten representeres derimot ved $0.0625m^2$ ruter. Ser vi nøye på figur 5.23 ser vi at sammensettingen gir vannflaten glattere konturer. Dette poenget kommer fram i en ny arealberegning, som viser nytt areal på $87044 m^2$. Arealet har økt med $904 m^2$ ved å inkludere lave intensitetsverdier (under 20) i figur 5.23.

Fra figur 5.23 kunne vi konstruert endelige potensielle vannknekklinjer som konturer av vannflaten, men vi skal i neste avsnitt forsøke å forbedre nøyaktigheten med knekklinjene ved å inkludere høyde.

5.3.4 Forbedring av vannknekklinjene med buffer og høyde

Vi forventer på generelt grunnlag at intensitetsverdier forholder seg lave i vann så lenge det ikke er speilrefleks. Men det kan tenkes at intensitetsverdiene øker ved for eksempel langs grunnene, strender og ved vegetasjon i vannet. Derfor kan ikke vår definisjon av intensitet (mindre enn 20) brukes til å klassifisere videre inn mot endelige vannknekklinjer.

I dette underkapittelet presenteres en metode som kan øke vannflatens areal ytterligere med høydeverdier. Metoden bygger på antakelsen om at vannet ligger lavere enn omkringliggende landskaper 4. I denne operasjonen bestemmer vi også den endelige høyden på vannflaten i egenutviklet metode.

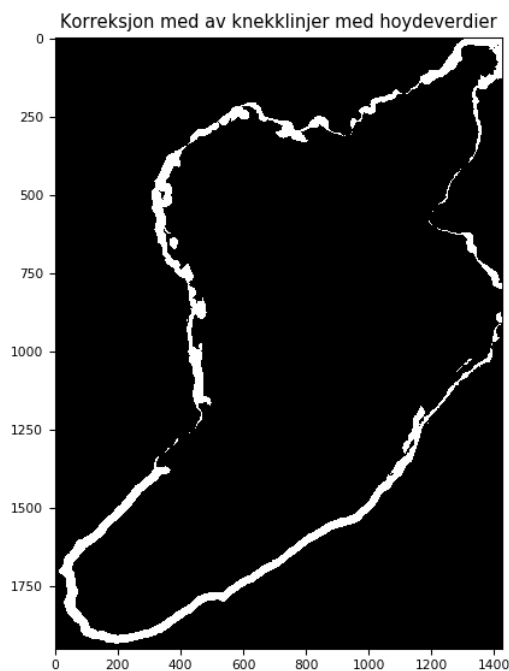
Bufferanalyse og høyde på vannet

Vi lager en buffer rundt foreløpige konturer av potensiell vannflate. Med "Connected Component Labeling" var det mulig å representere hver rute i buffer som høydeverdier fra terrengmodellen. Med andre ord brukes terrengmodellen som "bakgrunn" i en buffer. Bufferanalysen skal i prinsippet virke uavhengig av bufferstørrelse, men siden filterbasert metode antar at vannkanten ligger maksimum 5 meter utenfor detekterte ruter i trinn 1 (5.3.2), velges 5 meter om radius.

Høyden på vannflaten finner vi som medianen av høydeverdiene i buffer. Vi valgte medianutregning med inspirasjon fra Toscano et al. (2013).

Medianen brukes deretter som terskel mot de andre rutene i buffer, der vi godkjenner alle $0.0625m^2$ ruter som tilfredsstillende høydeoppløsning 12.5cm fra medianverdien i bufferen. Med denne høydeoppløsningen skal egenutviklet metode etter resonnementet oppnå samme resultat i vannknekklinjene som det skalerte histogrammet i 5.6a som Toscano et al. (2013) (gitt samme rutestørrelse).

Figur 5.25 viser hvilke ruter som tilfredsstillende 12.5cm høydeoppløsning i 5 meters buffer. Disse rutene er representert ved 1.



Figur 5.24: $0.0625m^2$ ruter som tilfredsstillter 12.5cm høydeoppløsning i en 5 meters buffer er representert med 1. Mørke områder er enten selve vannflaten eller områder utenfor 5 meters buffer

Til tross for at interpolering gir glattere overflate enn virkeligheten kunne tilsi, kan vi se ulik topografi forskjellige steder i figur 5.24. Noen steder på vannknekklinjene finner vi en bratt vannkant, mens andre steder en myk overgang. Meningen med denne operasjonen er altså å detektere disse rutene ved hjelp av høyde.

Til slutt sammensettes regionene fra filterbasert metode, ruter med lave intensitetsverdier (under 20) og rutene som tilfredsstillter høydenøyaktighet på 12.5cm i en 5 meters buffer. Nedenfor ser vi resultatet av korreksjonen av intensitet og bufferanalysen.



Figur 5.25: Vannflate etter korreksjon av høyde og intensitet

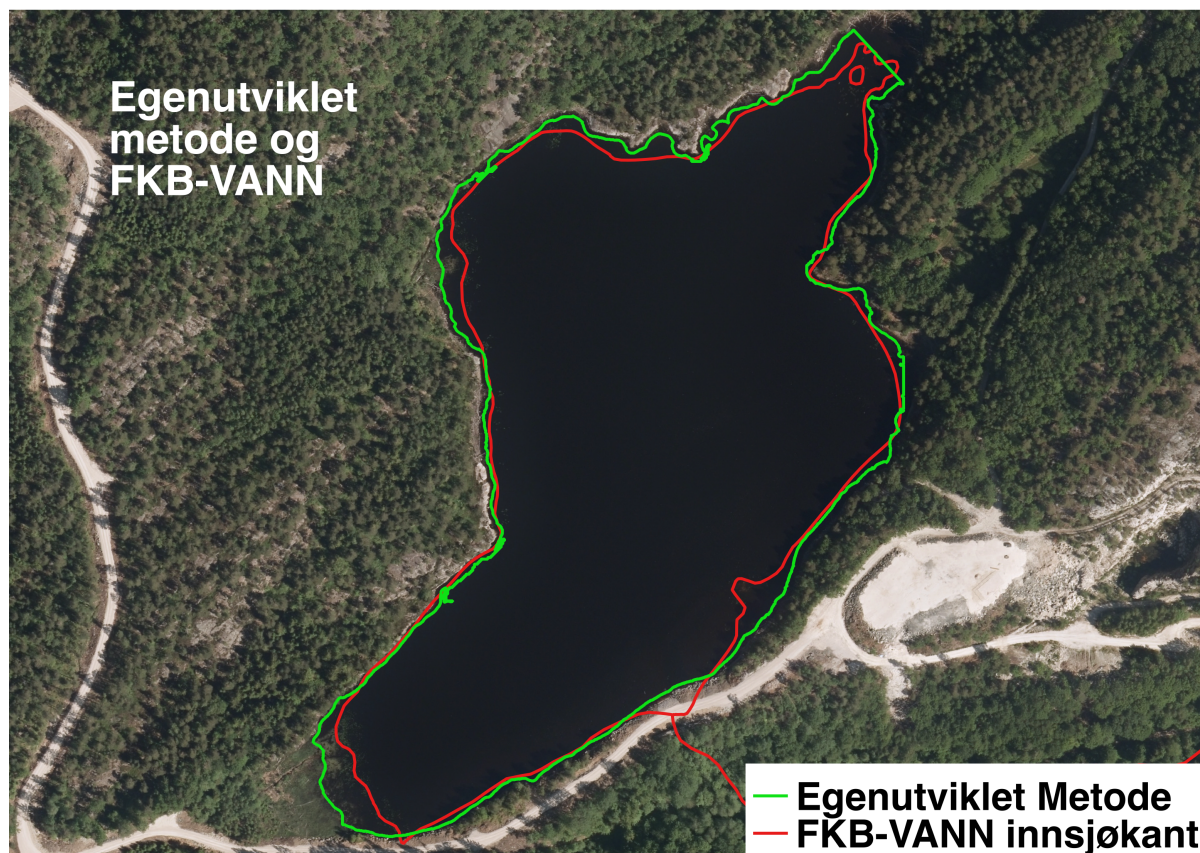
Arealet til endelig potensiell vannflate blir nå beregnet til å være 88865 m^2 mot 82637 m^2 i FKB-vann. Endelig potensielle vannknekklinjer fremkommer som konturer av vannflaten i figur 5.25, som presenteres i kapittel 6.

6 Resultater

I dette kapitlet presenteres resultatene i denne masteroppgaven. Toscano et al. (2013) og egenutviklet metode testes på datasett med ulik oppløsning, radius i interpolator og parametere i metode.

Vi bruker Norge i Bilder for verifisering av vannknekklinjene, som masteroppgaven har fokusert på. Flybildene som brukes er tatt i 2016, samme år som aktuelt datasett. Hvis ikke annet er spesifisert, gjelder resultatene $5 \frac{Pkt}{m^2}$ datasett. Resultatene analyseres videre i kapittel 7.

6.0.1 Egenutviklet metode og FKB-vann



Figur 6.1: Egenutviklet metode mot FKB-Vann

Vi ser at egenutviklet metode stort sett presterer bedre nøyaktighet enn FKB-vann. Resultatet viser i dette tilfelle at egenutviklet metode presterer bedre i grunne områder. I metodedelen fant vi ut at høyden er 63.7moh mot 63.2moh i FKB-vann.

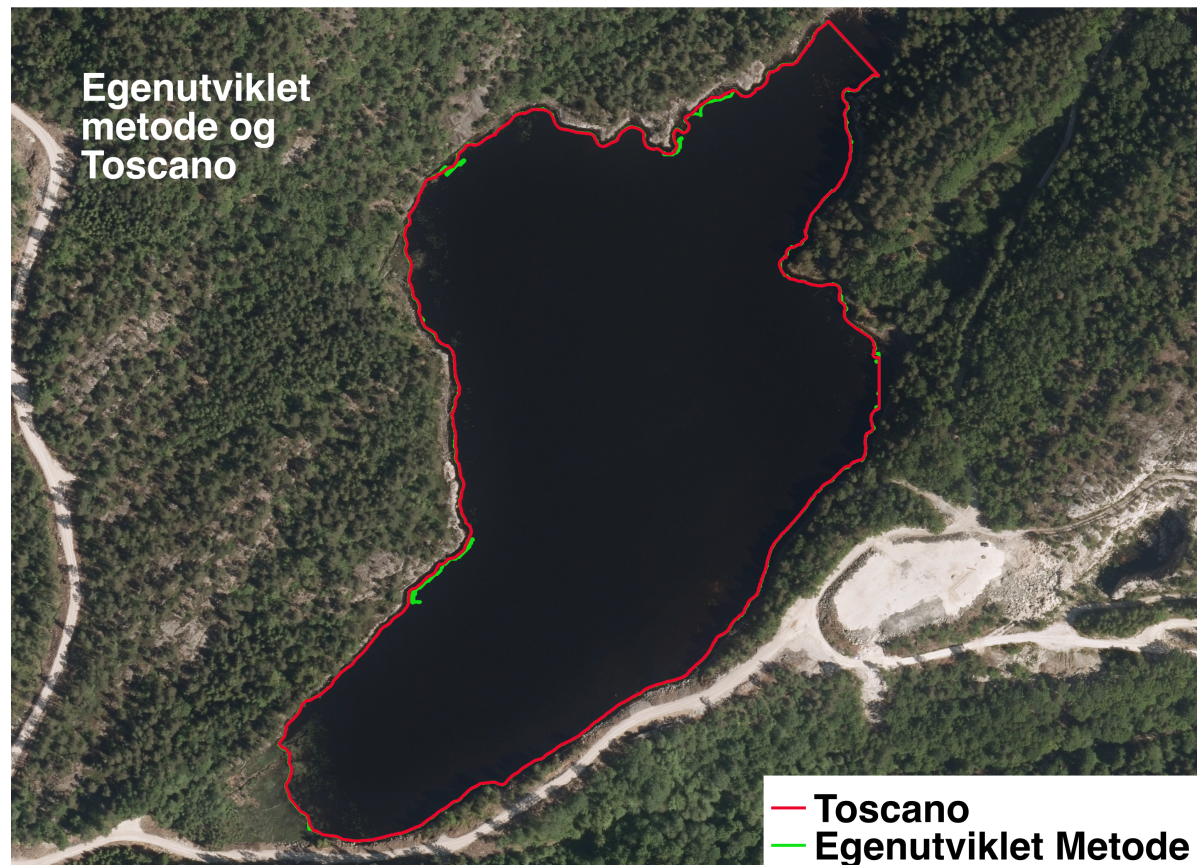
Nede på bildet ser ut om som om FKB-vann indikerer en liten vik og en liten bekk som renner ut av vannet. En årsak til avviket kan være at veien er ny, og det er fylling i rør under veien.

Oppe til høyre i vannet har vi kommet til datasettets yttergrense, så dette skyldes ikke feil i metoden.

Resultatene drøfting videre i kapittel 7.

6.0.2 Egenutviklet metode sammenlignet med Toscano

For sammenligning med Toscano et al. (2013) ble rutestørrelse $0.0625m^2$ benyttet for begge metoder. Avvikene mellom metodene visualiseres i figur 6.2. Her er Toscano et al. (2013) representert ved rød farge, mens egenutviklet metode er grønn farge.



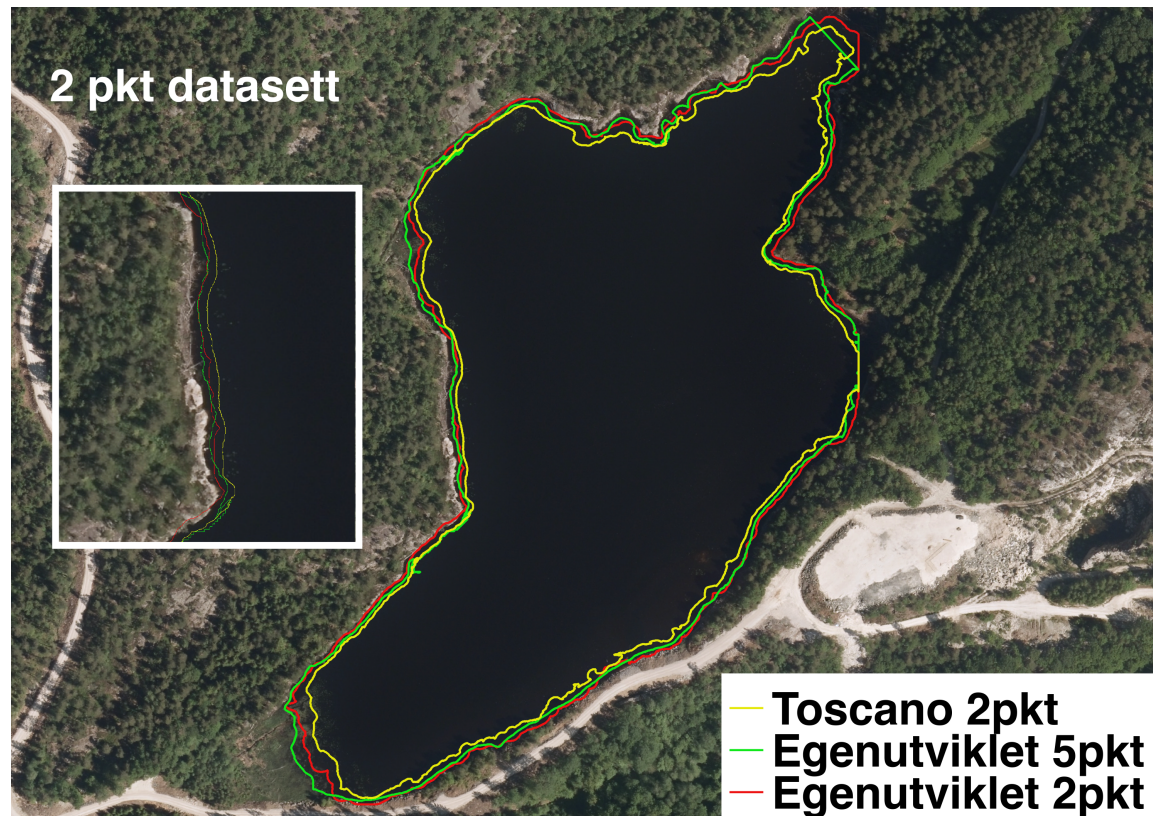
Figur 6.2: Toscano mot egenutviklet metode Toscano (2015)

Med høydeoppløsning 12.5cm får både Toscano et al. (2013) og egenutviklet metode omtrent samme resultat bortsett fra sporadiske avvik.

Med en Macbook Pro 2011 bærbar datamaskin brukte Toscano et al. (2013) 6 minutter og 31 sekunder på å klassifisere Frøviktjenn. Egenutviklet metoden brukte 1 minutt og 22 sekunder. Dersom dette er representativt for alle typer vann, så er ambisjonen oppnådd; vi har laget en metode med samme nøyaktighet, men klassifiserer raskere enn implementasjonen av Toscano et al. (2013). Hva som skal til for å bekrefte denne observasjonen, kommer vi tilbake til i drøftingskapittelet.

6.0.3 2 punkt per kvadratmeter datasett

Både Toscano et al. (2013) og egenutviklet metode ble testet med $2 \frac{Punkt}{m^2}$ datasett. Siden Toscano et al. (2013) og egenutviklet metode er praktisk talt like for $5 \frac{Punkt}{m^2}$ datasett, har vi i figur 6.2 bare sammenlignet $2 \frac{Punkt}{m^2}$ datasett (som gjelder begge metoder) med $5 \frac{Punkt}{m^2}$ i egenutviklet metode.



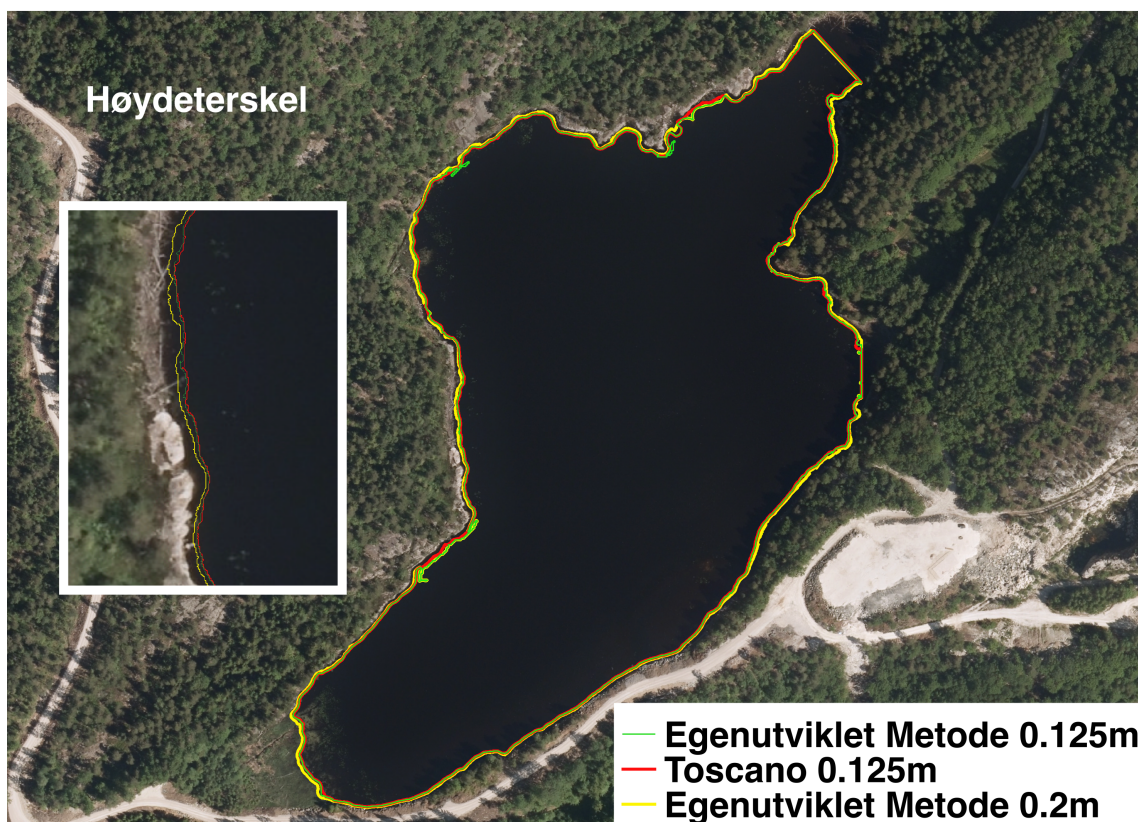
Figur 6.3: $2 \frac{Pkt}{m^2}$ datasett

For dette vannet ser vi at egenutviklet metode presterer tilsynelatende med bedre nøyaktighet enn Toscano et al. (2013) med $2 \frac{Pkt}{m^2}$ datasett. Denne situasjonen drøftes i kapittel 7.

I tillegg ser vi at vannknekklinjene med $5 \frac{Pkt}{m^2}$ datasett gir generelt bedre resultater enn $2 \frac{Pkt}{m^2}$ datasett, som naturlig er.

6.0.4 Redusert høydeoppløsning

I figur 6.4 ser vi hva som skjer om vi reduserer høydeoppløsningen fra 12.5cm til 20cm.



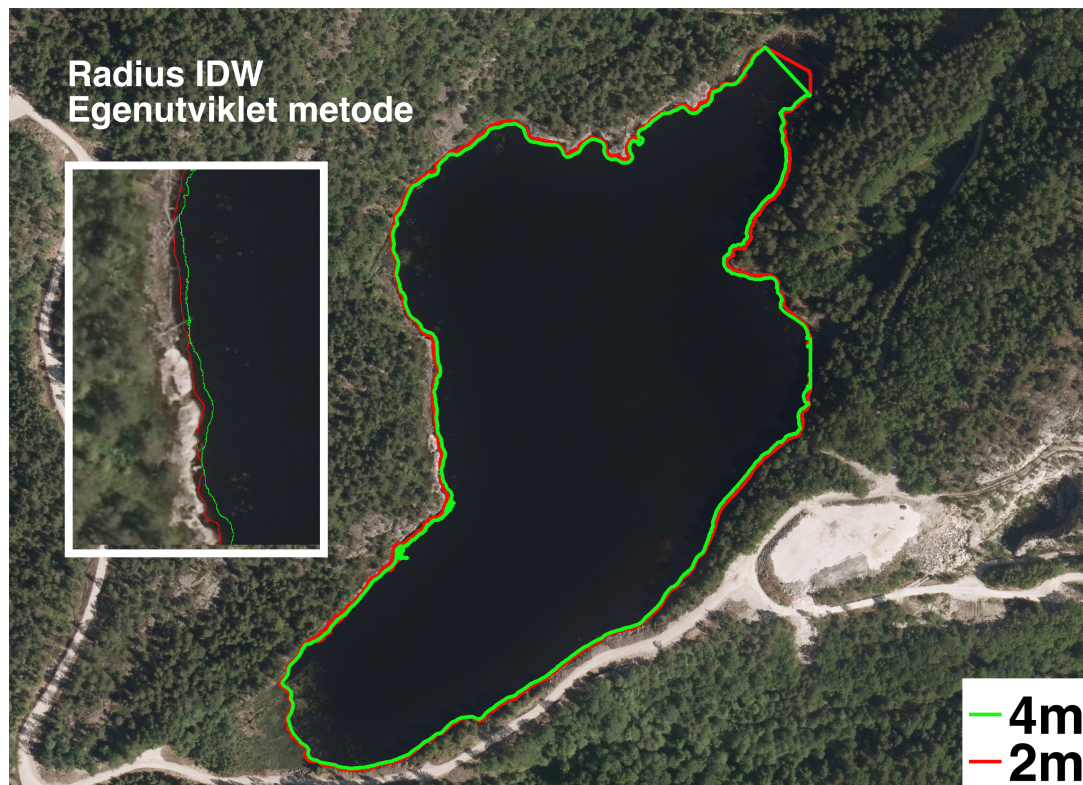
Figur 6.4: Redusert høydeoppløsning øker størrelsen på vannflaten

Figur 6.4 viser at redusert høydeoppløsning øker størrelsen på vannflaten. Hva som er den ideelle høydeoppløsningen i forhold til nøyaktighet drøftes i kapittel 7.

6.0.5 Radius i IDW

Når vi reduserer radiusen fra 4 meter til 2 meter forventer vi at overgangen mellom vann og land blir skarpere (IDW 3.1.1).

Figur 6.5 viser en sammenligning mellom opprinnelig 4 meters radius og 2 meters radius i IDW.

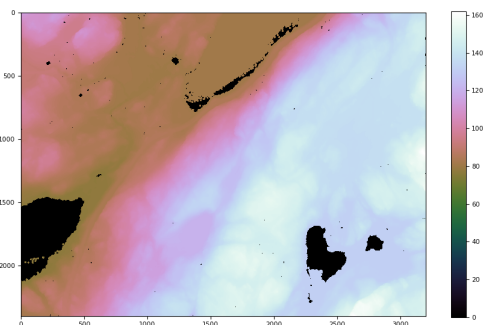


Figur 6.5: 2 meters radius og 4 meters radius

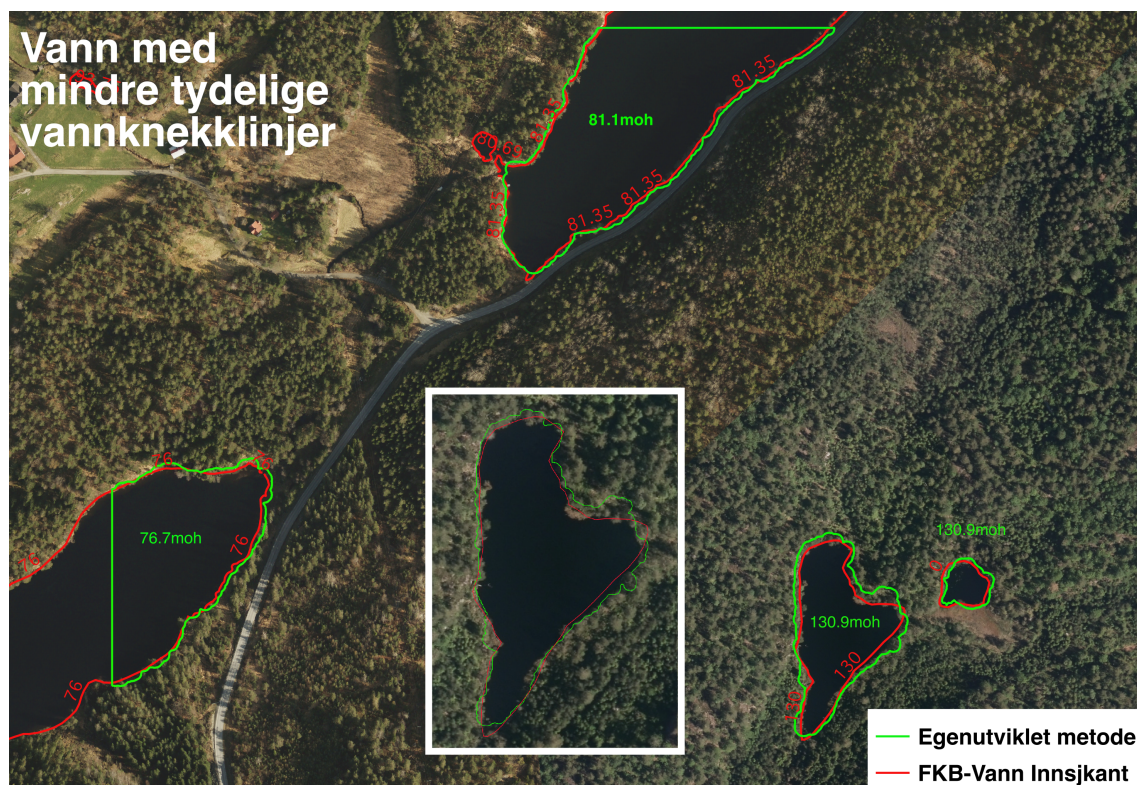
Figur 6.5 viser hvordan vannknekklinjene beveger seg inn mot land ved mindre glatting. Dette er forventet, med tanke på teorien beskrevet i 3.1.1.

6.0.6 Datasett med flere vann

Både egenutviklet metode og Toscano et al. (2013) er laget for å detektere flere vann i samme datasett. Men eksempelet nedenfor regnes som et problematisk på grunn av myr, stor andel uregistrerte punkter og små vann. Figur 6.6 viser terrengmodellen som er utgangspunktet for resultatet i figur 6.7.



Figur 6.6: Høydemodell med flere vann



Figur 6.7: Datasett med mindre tydelige vannknekklinjer

Figur 6.7 illustrerer utfordringene når en skal klassifisere vann med stor andel uregistrerte områder som gjerne intreffer med høy skannevinkel. Vannet uthevet i den hvite rammen har ikke skarpe kanter, slik som Frøvikjtjenn, men har en mer på myraktig overgang. Figuren illustrerer prisen å betale ved å øke skannevinkel, der man mangler høyde,- og intensitetsinformasjon i uregistrerte områder. Dermed kan man være mindre sikker på høyden på vannet i disse tilfellene. Disse forholdene gjøres mer rede for i kapittel 7.

En annen observasjon er at arealkravet på 1000 m^2 har fjernet det lille vannet på vestsiden av det nordligste vannet på 81.1moh.

7 Drøfting

Problemstillingen for masteroppgaven var å utvikle en metode som automatisk klassifiserer vannknekklinjer med data fra flybåren laserskanning.

Metoden skal kreve mindre regnekapasitet enn andre metoder, som for eksempel Toscano et al. (2013), men ikke ha vesentlig dårligere nøyaktighet.

Metoden skal etter beste evne tilfredstille kravene til "Hydro Flattening" fra kapittel 4 og generere vannknekklinjer og vannflater som gir bedre topografiske 3D-modeller.

Resultatene fra kapittel 6 kan tyde på at en automatisk metode kan brukes for klassifisering av vannflater i LiDAR. I hvilken grad disse resultatene er til å stole på, skal vi drøfte i dette kapittelet.

7.1 utfordringer for klassifisering i LiDAR-data

Klassifiseringsresultatet av vannflater kan påvirkes av følgende faktorer:

1. Nøyaktighet i datasettet levert fra flybåren laserskanning
2. Klassifiseringsmetode og valg av parametere
3. Varierende forekomster i naturen

Utfordringen er å finne metoder og parametre som best mulig klassifiserer innsjøer og tjern. Innsjøer og tjern har varierende forekomster av myr, vegetasjon, strender, grunner og ujevn topografi i vannkantene.

Med kun intensitets,- og høydeinformasjon i LiDAR er det en komplisert oppgave å ta hensyn til alle variasjonene i en og samme automatiske klassifiseringsmetode. Formålet med klassifiseringen må avgjøre hvilken klassifiseringsmetode en velger, samt parametere som punkttetthet og skannevinkel. For metodene presentert i denne oppgaven må en også med omhu velge radius i interpolator og høydeoppløsning.

7.2 Hydro Flattening

Mindre vann (tjern som i figur 6.7) kjennetegnes gjerne med myr og andre myke overganger sammenlignet med et større vann, som for eksempel Frøviktjenn. Dersom en

filtrerer bort små vann etter arealkriteriet i "Hydro Flattening" ((U.S. Geological Survey Standards, 2014b)), slipper en gjerne unna problemer som oppstår som følge av dårlige definerte vannkanter. FKB-vann har flere vann med vesentlig mindre areal enn $8000m^2$, og derfor valgte vi $1000m^2$ arealgrense med våre metoder. Avhengig av formålet med terrengmodellen, er det avveining mellom hvor små vann som skal klassifiseres og hvilken nøyaktighet som kreves.

På generelt grunnlag ser det ut til at metodene blir sikrere jo større vannflater som filtreres bort i arealfiltreringen.

I hvilken grad metodene tilfredsstillkravene for kriteriene for "Hydro Flattening" fra kapittel 4 og (U.S. Geological Survey Standards, 2014b), drøftes i underkapitlene som nå skal gjennomgås.

7.3 Feilkilder i datasett og målemetode

7.3.1 Intensitet

Selv om vann kjennetegnes ved lave intensitetsverdier, er det problematisk å definere "lav nok" intensitetsverdi til å klassifisere vann. Intensiteten vil avhenge av flere forhold beskrevet i kapittel 2. For eksempel vil et større fotavtrykk (2.2.1) føre til en mindre refleksjon fra grunner langs vannkanten enn mindre fotavtrykk. Fotavtrykket i seg selv påvirkes av kontaktflaten laserstrålen treffer, men også av skannevinkel.

Intensiteten påvirkes også av bølgelengde. Vi vet at nærinfrarødt bånd (800-1400nm) benyttes i datagrunnlaget i denne oppgaven, men ikke nøyaktig hvilken bølgelengde. Hvilken bølgelengde i spennet mellom 800 og 1400nm vil ha betydning.

Når det gjelder mediansjekken i 5.2.2 er dette bare ett forslag på å diskriminere potensielle vannflater fra for eksempel en fotballbane, åker eller andre flate områder. Mediansjekken fungerer i mange tilfeller, men det er tvilsomt om den er tilfredsstillende i alle situasjoner. For videreutvikling av testen kunne man prøvd å regne medianen uten områdene som er uregistrerte eller som har urealistisk høye intensitetsverdier ved lav skannevinkel (speilrefleks). Da hadde man muligens fått et mer realistisk mål på intensitetsverdien som burde brukes i en slik test.

Helst skulle man basert mediantesten i 5.2.2 på antagelsen om generelt lave intensitetsverdier (under 20), men kjente problemstillinger som vegetasjon, grunner, skannevinkel vil påvirke intensiteten i et slikt tilfelle. Å unngå bruk av intensitetsverdier fra disse fenomenene i vannet, er vel argumentet for å bruke median i utgangspunktet.

For egenutviklet metode i figur 5.23 kapittel 5 øker arealet med $904m^2$ ved å sammensette intensitetsverdier under 20 og flate områder fra høydefiltrering. Intensitetsverdier (under 20) har mest sannsynlig kompensert for grunner som ikke ble detektert som vann i filterbasert metode trinn 2.

Parameteren ser derfor ut som et vellykket valg for å klassifisere vann i dette datagrunnlaget.

7.4 Usikkerhet i metode

I egenutviklet metode velges en tilnærming der man filtrerer bort karakteristikkene som ikke tilfredsstiller gitte kriterier. For eksempel brukes den filterbaserte delen av metoden til å negligere områder som ikke er flate.

7.4.1 Filterbasert metode vs. histogramanalyse

Man detekterer flate områder i egenutviklet metode med filterbasert tilnærming, i kontrast til den histogrambaserte metoden i Toscano et al. (2013).

En filterbasert metode med hierarkisk struktur ser ut som et steg i riktig retning når det gjelder regnekapasitet og effektivitet i rasterformatet. Egenutviklet metode reduserte tiden med 80% sammenlignet med implementasjonen av Toscano et al. (2013).

Siden små vann ikke fanges opp av histogramanalysen til Toscano et al. (2013), tyder det på at filterbasert metode er bedre egnet, som beskrevet i 5.3.2. Av samme grunn antar vi at filterbasert metode kan detektere vann, der bare en del av vannet er innenfor datasettet. Selv om vi antar disse egenskapene med filterbasert metode, må dette testes nærmere.

7.5 Klassifisering av Frøvikjtjenn

Forskjellige forhold kan føre til at klassifisering av vannflatene kan være problematisk:

- Vannet kan være regulert, dvs. planlagt varierende vannstand.
- Vannet kan blitt utsatt for en lang regnperiode, flom eller tørke.
- Vind gir bølgeeffekt og påvirker klassifikasjonen ved at vannet ikke lenger representeres som "flatt".

Foreløpige resultater basert på flyfoto fra 2016 viser at både Toscano et al. (2013) og egenutviklet metode gir mer nøyaktige vannknekklinjer rundt Frøvikjtjenn enn FKB-vann. Målingene i gjeldende datasett tyder på at vannet har blitt større, med tanke på at overflaten til Frøvikjtjenn ser ut til å være 50cm høyere enn i 1981 da Frøvikjtjenn ble tegnet til FKB-vann. Årsaken kan selvsagt være at vannet er regulert, men ingen spor i flyfotoet tyder på det.

Ifølge prosjektrapporten var det skyfritt og lite vind på måletidspunktet. Fordi det har forekommet en speilrefleks i vannet, kan vi anta at vannet var stille på måletidspunktet. Hvordan målingene hadde vært med bølger på vannet, har vi ikke undersøkt.

En sentral problemstilling rundt høyden til Frøvikjtjenn, er medianberegningen i 5.3.4. Man kunne istedenfor valgt en minimum høyde i buffer for å være enda mer sikker på at vannet ligger under det omkringliggende landskapet (bortsett fra der bekken renner ut av vannet). På en annen side sier medianen mer om ”mengden” av ruter med samme høyde (som indikerer vann), sammenlignet med en minimumsverdi som like godt kan være en avstandsfeil fra laserskanneren eller en grunne.

Tydelige vannkanter, stort areal og relativt stor andel returter fra vannoverflaten er muliggjenn grunnen til at metoden klassifiserte Frøvikjtjenn med godkjent grunnrissnøyaktighet.

7.6 Valg av parametere

Noen parametere som kan påvirke nøyaktighet og krav til regnekapasitet. Disse parametere har vi kort drøftet nedenfor.

7.6.1 Interpolator

2 meters radius viser seg å gi bedre resultat på vannknekklinjene enn 4 meters radius (figur 6.5). Det er som forventet, da mindre glatting fører til skarpere skille mellom vann og land (3.1.1). På en annen side vil det føre til at færre uregistrerte områder blir tillagt en interpolert verdi. Utfordringen knyttet til stor andel uregistrerte områder blir videre drøftet i 7.7.

7.6.2 Høydeoppløsning

Siden metodene presentert beregner vannflater ved arealberegninger, er høydeoppløsning en sentral parameter. Resultatene viser at redusert høydeoppløsning øker størrelsen på vannflaten og fører de beregnede vannknekklinjene lenger inn mot den observerte vannkanten. Med tanke på at IDW gir kontinuerlig overflate, er det altså forventet at vannknekklinjene med redusert høydeoppløsning gir større areal på vannet.

Siden hver rute er $0.0625m^2$, regner en med at rutene i knekklinjene i figur 6.4 har interpolert verdi. På en annen side kan vi anta at kontrollpunktene ligger tettere lenger inn mot land og at hver rute langs knekklinjene gir en mer korrekt høydeverdi.

Toscano et al. (2013) har et poeng ved at høydeoppløsningen må variere med topografien i datasettet. For å finne flate områder i slake terreng trengs det høyere høydeoppløsning enn i bratte landskaper. Utfordringen er at vi finner variasjoner i topografi rundt hvert

vann. Noen steder finnes det skarpe kanter og andre steder myke overganger (grunner, strender).

I egenutviklet metode øker vi høydeoppløsning innenfor hver vannflateregion for å detektere flere detaljer i og rundt vannflaten. Selv om vi i metodedelen valgte å tette eventuelle hull i vannflaten, ville eventuell vegetasjon og eventuelle øyer representeres som nettopp dette; hull og områder som ikke er flate. Med andre ord vil økt høydeoppløsning gi mer informasjon om øyer og vegetasjon i vannet. Videre utvikling av vegetasjon i vannet er ikke behandlet i denne oppgaven.

Selv om økt høydeoppløsning viser til bedre resultater i Frøvik tjenn, betyr det ikke at parameteren bør økes med formål om å forbedre nøyaktigheten i alle tilfeller. Justering på radius i interpolator vurderes som et sikrere valg i de fleste tilfeller, på grunn av færre interpolerte ruter. Det antyder også resultatene dersom en sammenligner vannknekklinjene i figur 6.5 og 6.4.

7.6.3 Filtrere bort veier

Veier har ofte lignende intensitetsverdier som vann, med denne metoden kan de derfor se ut som elver. Vi har i dette arbeidet vært inne på ulike metoder for å filtrere bort veiene, som soliditetsfiltrering/convex hull, måling av høydeforskjell til omkringliggende områder og gjenkjenne menneskeskapte former. Men ingen av disse metodene kom vi i mål med innenfor denne oppgaven.

7.6.4 Skannevinkelkorleksjon

Ifølge resultatene er vannknekklinjene med lav skannevinkel og høye intensitetsverdier identiske med og uten intensitetskorleksjon av skannevinkel. Vi registrerer dermed at korleksjon av skannevinkel ikke har innvirkning på avstandsmålingen, bare intensiteten.

På samme måte som usikkerheten rundt valg av "lav nok" intensitetsverdi for vann (også diskutert i 7.3.1), er det usikkerhet rundt "høy nok" intensitetsverdi for klassifisering av speilrefleks. Vi valgte 400 som "høy nok" terskel i skannevinkelkorleksjonen, men dette er kun basert på egen erfaring i gjeldende datasett.

I våre datasett (800m*600m) dekker en skannevinkel mellom -2 og 2 grader en "swath width" på omtrent 100 meter. Tatt i betraktning lite vind på måletidspunktet og at finnes flere vann i området, antas en form for skannevinkelkorleksjon som nødvendig i gjeldende datasett for sikrere klassifisering.

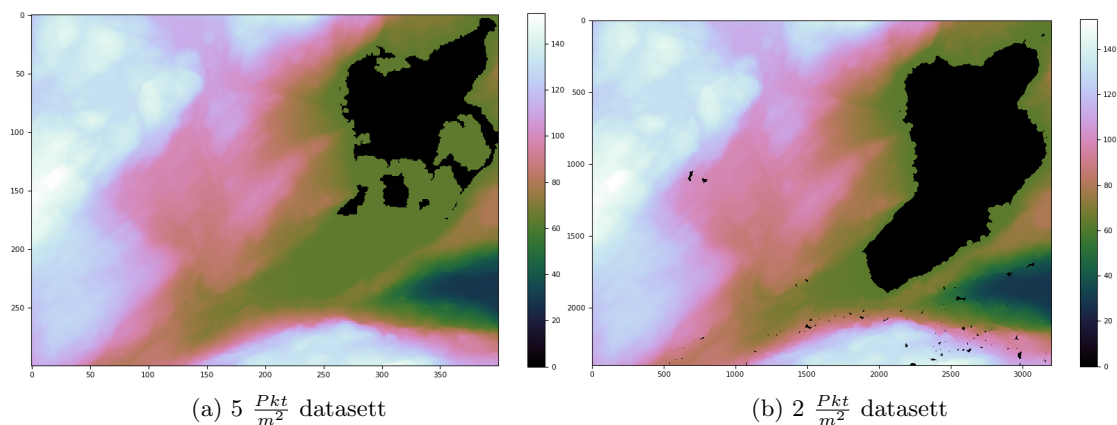
7.7 Ulik punkttetthet og uregistrerte områder

Resultatene i denne oppgaven viser generelt bedre resultater med $5 \frac{Pkt}{m^2}$ datasett enn $2 \frac{Pkt}{m^2}$ datasett. Dette er forventet siden høyere punkttetthet bør gi bedre resultater.

I det etterfølgende skal vi drøfte en mulig årsak til at egenutviklet metode tilsynelatende har bedre nøyaktighet enn Toscano et al. (2013) i $2 \frac{Pkt}{m^2}$ datasett fra figur 6.3 i kapittel 6. Resonnementet beskriver også mulige utfordringer knyttet til uregistrerte områder og det problematiske eksempelet i figur 6.7.

7.7.1 2pkt per kvadratmeter datasett

Fra et $2 \frac{Pkt}{m^2}$ datasett kan vi først og fremst forvente større andel uregistrerte områder og færre kontrollpunkter per rute. Figur 7.1a og figur 7.1b viser terrengmodellene til de ulike datasettene.



Figur 7.1: Uregistrerte områder representeres som svarte felter

Som følge av en stor andel uregistrerte områder i $2 \frac{Pkt}{m^2}$, kan vi tenke oss at histogramanalysen vil detektere topper med mindre frekvensverdier. Siden kontrollpunkter likevel kan finnes i vannkantene på Frøvikjtjenn, er det nærliggende å tro at Frøvikjtjenn likevel gir en topp i histogrammet. Det interessante er hvilken søyle som detekteres med høyest frekvens i datasettet med $2 \frac{Pkt}{m^2}$. Dersom høydeverdien i detektert topp viser seg betydelig lavere enn tilsvarende søyle for $5 \frac{Pkt}{m^2}$ (som var 63.7moh), er det naturlig at arealet blir mindre. Det er fordi metoden Toscano et al. (2013)(histogramanalysen) forutsetter at en detektert topp er under det omkringliggende landskapet.

Egenutviklet metode tilnærmer seg problemstillingen på en annen måte. Mens Toscano et al. (2013) har som kriterium at vannet ligger rett under landskapet ved histogramanalysen, brukes medianen i en 5 meters buffer utenfor potensielle vannknekklinjer (5.3.4)

for å oppfylle dette kriteriet i egenutviklet metode. Med andre ord må bufferoperasjonen finne sted før vi betrakter dette kriteriet som oppfylt. Derfor kan det gi mening hvis vannknekklinjene fra egenutviklet metode ligger opptil 5 meter utenfor vannknekklinjene i Toscano et al. (2013) i figur 6.3.

Det ser ut til at egenutviklet metode gir bedre grunnrissnøyaktighet i vannflater der uregistrerte områder ligger mer enn 5 meter fra vannkanten, som i 2 $\frac{P_{kt}}{m^2}$ datasettet av Frøvik tjenn. Hvis vi derimot havner i en situasjon der uregistrerte områder er mindre enn 5 meter fra vannkanten, har egenutviklet metode en utfordring. Denne utfordringen illustreres nemlig i det problematiske eksempelet i 6.7, der store deler av tjernet dekkes av uregistrert område. Det kan godt tenkes at Toscano et al. (2013) presterer bedre grunnrissnøyaktighet enn egenutviklet metode i disse situasjonene, siden bufferanalysen havner på land.

Generelt er uregistrerte områder en utfordring for begge metoder når det gjelder detektering av vannflatens høyde.

7.8 Videre arbeid

7.8.1 Strandsone og kystkontur

Metodene presentert i denne oppgaven kan potensielt videreføres til å detektere strandsone og kystkontur. For at dette skal være mulig, må resultatene først og fremst vurderes opp mot tidevannstabeller og andre relevante meteorologiske data. I denne sammenheng må det nevnes et LiDAR datasett av havet bør klassifiseres med noe andre kriterier enn stille vann. På havet er det større bølger, mer værutsatt og har flere usikkerhetsmomenter som påvirker målinger fra flybåren laserskanning. Problemstillingene rundt grunner, strender, vegetasjon og andre forhold diskutert i denne oppgaven er likevel relevant i sammenheng med strandsone og kystkontur.

Det er grunn til å anta at kartlegging av strandsonen er spesielt interessant. Man kan med denne metoden raskt registrere tiltak i strandsonen, og sammenholde dette med godkjente søknader.

7.8.2 Videre tester og kvalitetssikring

I tillegg til drøftingstemaene som til nå er diskutert, oppsummeres her videre tester i egenutviklet metode:

- Teste på flere terrengmodeller med ulik topografi
- Teste for vegetasjon og øyer
- Teste for vann med mindre arealer enn $1000 m^2$

- Teste for bølger
- Teste i ulikt vær
- Kvalitetsikring i punktskyen

7.8.3 Toscana 2015

Denne oppgaven tar utgangspunkt i Toscano et al. (2013). Men det finnes et senere arbeid fra samme forfatter, Toscano (2015), som sannsynligvis er mer omfattende enn Toscano et al. (2013). Før man tar egenutviklet metode i bruk, hadde det vært nyttig å holde den opp mot funnene i Toscano (2015).

8 Konklusjon

Denne masteroppgaven har presentert en egenutviklet metode og en implementasjon av den hittil mest kjente metoden for klassifisering av vannflater, (Toscano et al., 2013). Begge metoder klassifiserer vann automatisk, basert på data fra flybåren laserskanning.

Begge metoder har til hensikt å erstatte den manuelle, men hittil nødvendige prosessen ved å legge inn FKB-vann for å markere vannflater i LiDAR genererte terrengmodeller. Her er det altså mye tidsbesparelse å hente.

Siden (Toscano et al., 2013) er den beste metoden hittil kjent, er det denne vi har sammenlignet oss med, i tillegg til den nøyaktigheten vi kjenner fra FKB-vann.

Vi kan trekke følgende konklusjoner av arbeidet:

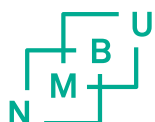
- Egenutviklet metode krever vesentlig mindre regnekapasitet enn vår implementasjon av Toscano et al. (2013). Vi har sett reduksjoner i størrelsesorden 80% mindre regnekapasitet.
- Både (Toscano et al., 2013) og egenutviklet metoder ser ut til å ha bedre nøyaktighet enn testede datasett fra FKB-vann. Siden FKB-vann datasett er produsert på forskjellige tidspunkt, er naturlig å anta nøyaktigheten bedres jo nyere de er.
- På $5 \frac{Pkt}{m^2}$ datasett ser egenutviklet metode ut til å ha tilsvarende nøyaktighet som (Toscano et al., 2013)
- På $2 \frac{Pkt}{m^2}$ datasett ser det ut til at egenutviklet metode har bedre nøyaktighet enn (Toscano et al., 2013)
- Før en trekker endelige konklusjoner, må tester og kvalitetssikring som beskrevet i drøftingskapittelet gjennomføres.
- Dersom testene og kvalitetssikringen gir tilfredsstillende resultat, er det ikke unaturlig at en versjon av denne nye metoden i fremtiden kan brukes til automatisk klassifisering av vannflater basert på LiDAR data.

Bakgrunnen for oppgaven var å redusere det tidkrevende arbeidet med å kombinere FKB-vann og LIDAR-baserte terrengmodeller. Vi har vist at det finnes automatiske metoder for å klassifisere vannflater som kan gi like god eller bedre nøyaktighet. Her ligger det altså store potensielle gevinster.

Bibliografi

- ASPRS (2013). Las spesification version 1.4 - r13. http://www.asprs.org/wp-content/uploads/2010/12/LAS_1_4_r13.pdf.
- Bjørke, J. (2010). Digitale terrengmodeller. *Universitetet for miljø-og biovitenskap*.
- Brzank, A. and Heipke, C. (2007). Supervised classification of water regions from lidar data in the wadden sea using a fuzzy logic concept. pages 12–14.
- Burger, W. and Burge, M. J. (2016). *Digital image processing: an algorithmic introduction using Java*. Springer.
- Geovekst, F.-V. p. (2016). Fkb-vann produktspesifikasjon. <http://www.kartverket.no/globalassets/standard/sosi-kap3-produktspesifikasjoner/fkb-4.02-2011-12-01/2-vann-2011-12-01.pdf>.
- Hug, C. (1996). Combined use of laser scanner geometry and reflectance data to identify surface objects. pages 9–11.
- Höfle, B., Vetter, M., Pfeifer, N., Mandlbürger, G., and Stötter, J. (2009). Water surface mapping from airborne laser scanning using signal intensity and elevation data. *Earth Surface Processes and Landforms*, 34(12):1635–1649.
- Kartverk, S. (2014). Fotogrammetrisk bildematching, nasjonal detaljert høydemodell. <http://www.kartverket.no/globalassets/kart/nasjonal-hoydemodell/nasjonal-detaljert-hoydedatamodell---sluttrapport-20140210.pdf>.
- Kartverket (2010). Høydemodeller fra laserskanning og bildematching. <http://www.statenskartverk.no/globalassets/kart/nasjonal-hoydemodell/fotogrammetrisk-matching.pdf>.
- Kartverket (2014a). Forprosjekt, nasjonal detaljert høydemodell, sluttrapport. <http://www.kartverket.no/globalassets/kart/nasjonal-hoydemodell/nasjonal-detaljert-hoydedatamodell---sluttrapport-20140210.pdf>.
- Kartverket (2014b). Kvalitetskrav ndh, nasjonal detaljert høydemodell. <http://www.kartverket.no/globalassets/kart/nasjonal-hoydemodell/nasjonal-detaljert-hoydedatamodell---sluttrapport-20140210.pdf>.
- Liu, X. (2008). Airborne lidar for dem generation: some critical issues. *Progress in Physical Geography*, 32(1):31–49.

- O’Sullivan, D. and Unwin, D. (2014). *Geographic information analysis*. John Wiley & Sons.
- Rasmussen, J. K. og Flatman, A. C. (2000). *En kvalitetsvurdering, afgangspjekt fra Aalborg Universitet*. Landinspektøruddannelse, Aalborg Universitet.
- Smeeckaert, J., Mallet, C., and David, N. (2013). Classification of water surfaces using airborne topographic lidar data. *ISPRS-International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, (1):321–326.
- Terratec (2016). Prosjektrapport ndh kragerø-drangedal 5pkt 2016. In *NDH Kragerø-Drangedal 5pkt 2016*, pages 2–20. Geovekst.
- Toscano, G., Gopalam, U., and Devarajan, V. (2013). A novel method for automation of 3d hydro break line generation from lidar data using matlab. *ISPRS-International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, 1(2):99–104.
- Toscano, G. J. (2015). A lidar-based auto hydro breakline generation algorithm for standing water bodies. https://uta-ir.tdl.org/uta-ir/bitstream/handle/10106/25368/Toscano_uta_2502D_13295.pdf?sequence=1&isAllowed=y.
- U.S. Geological Survey Standards, Section B, L. B. S. (2014a). Hydro flattening reference, lidar base spesification, techniques and methods. page 30.
- U.S. Geological Survey Standards, Section B, L. B. S. (2014b). Lidar base spesification, hydro flattening requirements, techniques and methods. page 12.
- Vosselman, G. and Maas, H. (2010). *Airborne and Terrestrial Laser Scanning*. Whittles Publishing.



Norges miljø- og biovitenskapelig universitet
Noregs miljø- og biovitenskapelige universitet
Norwegian University of Life Sciences

Postboks 5003
NO-1432 Ås
Norway