



Norges miljø- og
biovitenskapelige
universitet

Masteroppgave 2021 30 stp
Fakultetet for realfag og teknologi

Deteksjon av oversvømt jordbruksareal basert på data fra Sentinel-1

Detection of flooded agricultural land based on data
from Sentinel-1

Olivia Ragna Magnorsdotter Rypdal
Institutt for geomatikk

Sammendrag

I takt med en økende nedbørsmengde i Norge og store deler av verden, øker også behovet for å kartlegge landområder hvor store nedbørsmengder potensielt kan gjøre stor skade. Jordbruksareal er områder som er sårbare for klimaendringene, da oversvømmelser på matjorda kan redusere og ødelegge avlinger. Fjernmålingsdata egner seg godt til en slik kartlegging, og ved å benytte billedata fra radarsatellitten Sentinel-1 er man sikret tilgang til opptak med høy nøyaktighet fra alle slags vær- og lysforhold.

I denne gradsoppgaven undersøkes det i hvilket grad oversvømt jordsbruksareal kan detekteres i SAR-data fra Sentinel-1. Analysen er gjennomført for et studieområde i den sør-østlige delen av Viken, i to kraftige nedbørsperioder rundt juli og oktober 2020. Basert på radarsignalenes tilbakespredningsintensitet, som i SAR-bildene er representert ved intensitetsverdi i pikslene, forsøkes det å ekstrahere områdene som kan være oversvømte. Teknikkene som benyttes til dette er terskling og Random forest klassifisering. I analysen undersøkes også utviklingen av pikslenes intensitetsverdi over tid, for å se om det er sammenhenger innad i og mellom tidsperiodene.

Resultatene viser at det er stort potensiale i anvendelse av data fra Sentinel-1 til å kartlegge oversvømte jordbruksområder. Resultatene varierer noe mellom de ulike teknikkene, men har mange likheter som gjør det mulig å anta forekomster av overvann i bildene. Funnene er noe sprikende for bilder med ulikt opptakstidspunkt, og av de to tidsperiodene virket det enklest å detektere overvann i bildene fra oktober. Tidsserieanalysen viste også en klar tendens til synkende intensitetsverdier i pikslene klassifisert som vann i dette opptaket. Uten valideringsdata kan det ikke fastslås noen nøyaktighet for resultatene, men ved å sammenligne med verdier i NDWI-raster basert på data fra Sentinel-1 ble en del av antakelsene om oversvømte områder styrket.

Abstract

With an increasing amount of precipitation in Norway and large parts of the world, the need to map areas where large amounts of precipitation can potentially cause great damage is also increasing. Agricultural land are areas that are vulnerable to climate change, as floods on topsoil can reduce and destroy crops. Remote sensing data is well suited for such a mapping, and by using image data from the radar satellite Sentinel-1, recordings with high accuracy from all kinds of weather and light conditions are ensured.

In this master's thesis, it is investigated to what extent flooded agricultural area can be detected in SAR data from Sentinel-1. The analysis was carried out for a study area in the south-eastern part of Viken County in Norway, during two heavy rainfall periods around July and October 2020. Based on the backscatter intensity of the radar signals, which in the SAR images is represented by intensity value in the pixels, attempts are made to extract the flooded areas. The techniques used for this are thresholding and Random forest classification. The analysis also examines the development of the pixels' intensity value over time, to see if there are connections within and between the time periods.

The results show that there is great potential in the use of data from Sentinel-1 for mapping flooded agricultural land. The results vary somewhat between the different techniques, but have many similarities that make it possible to assume occurrences of surface water in the images. The results are somewhat divergent for images with different sensing date, and of the two time periods, it seemed easiest to detect surface water in the image from October. The time series analysis also showed a clear tendency for the development of intensity values in the pixels classified as water in this image. Without validation data, no accuracy of the results can be determined, but by comparing with values in a NDWI raster based on data from Sentinel-1, some of the assumptions about flooded areas were strengthened.

Forord

Denne masteroppgaven markerer slutten på min tilværelse som geomatikkstudent ved Norges miljø- og biovitenskapelige universitet. Oppgaven er skrevet høsten 2021 og utgjør 30 studiepoeng.

Studiet har gitt meg bred kompetanse innenfor fagområdet geomatikk, og fordypningsemner i fjernmåling og GIS har lagt et godt grunnlag for arbeidet med tematikken i denne oppgaven. SAR-fagfeltet har imidlertid vært ganske ukjent for meg frem til nå. Det har vært tidvis utfordrende, men aller mest spennende, å sette seg inn i dette nye emnet på tampen av studiet.

Jeg vil rette en stor takk til mine dyktige veiledere som har bidratt med kompetanse og råd underveis; Bjørn Borschenius og Jonathan Rizzi ved NIBIO, Dagrun Aarsten ved 3D Analyse og Ivar Maalen-Johansen ved NMBU. En ekstra takk til Bjørn som på vegne av NIBIO kom med oppgaveforslaget, og til Dagrun som lot meg holde forelesning om SAR i emnet GMBB300 i høst.

Takk også til andre som har bidratt med smått og stort underveis i arbeidet - det være seg fagkunnskap, Latex-hjelp, gjennomlesing og motiverende ord.

Til slutt vil jeg takke alle medstudenter, forelesere, kollegaer og andre som har gjort årene på Ås til en lærerik og uforglemmelig tid.

Ås, 15. desember 2021

Olivia Rypdal

Forkortelser

ASC: Ascending

DTM: Digital terrengmodell

ESA: European Space Agency

ESRI: Environmental Systems Research Institute

EW: Extra-Wide swath

FKB: Felles kartdatabase

GIS: Geografiske informasjonssystemer

GRD: Ground Range Detected

IW: Interferometric Wide swath

MSI: MultiSpectral Instrument

NDH: Nasjonal detaljert høydemodell

NDWI: Normalized difference water index

NIBIO: Norsk institutt for bioøkonomi

NIR: Nærinfrarød

QTM: Quick Terrain Modeler

RAR: Real aperture radar

RFC: Random forest classifier

SAR: Synthetic aperture radar

SM: Stripmap

SNAP: Sentinel Application Platform

S-1A: Sentinel-1A

S-1B: Sentinel-1B

TIFF: Tagged Image File Format

VH: Vertikal-horisontal

VV: Vertikal-vertikal

WV: Wave

Innhold

Figurer	viii
Tabeller	x
1 Innledning	1
1.1 Bakgrunn	2
1.2 Problemstilling	2
1.3 Relatert arbeid	2
1.4 Struktur av oppgaven	4
2 Teori	5
2.1 Sentinel-1	5
2.2 Radar	6
2.3 SAR	7
2.3.1 Grunnleggende prinsipp	7
2.3.2 Tilbakespredning	8
2.3.3 Polarisering	9
2.3.4 Bølgelengdeområde	9
2.3.5 SAR-bilder	10
2.4 Sentinel-2	11
2.5 NDWI	11
2.6 Terskling	12
2.7 Styrt klassifisering	12
2.7.1 Random forest classifier	13
3 Material	14
3.1 Studieområde	14
3.2 Tidsrom	16
3.3 Filformater	18

3.3.1	TIFF	18
3.3.2	GeoTIFF	18
3.3.3	Shapefil	18
3.4	Datagrunnlag	18
3.4.1	Sentinel-1 GRD	18
3.4.2	FKB-AR5	19
3.4.3	FKB-Vann	19
3.4.4	NDH-DTM	19
3.5	Programvare	19
3.5.1	SNAP	19
3.5.2	ArcGIS Pro	19
3.5.3	QGIS	20
4	Metode	21
4.1	Preprosessering av Sentinel-1-data	21
4.2	Monotemporær analyse av Sentinel-1-data	22
4.2.1	Terskling	22
4.2.2	Klassifisering	24
4.3	Multitemporær analyse av Sentinel-1-data	25
4.3.1	Tidsserieanalyse	25
4.3.2	Rasterkalkulering	25
4.4	NDWI fra Sentinel-2	26
5	Resultater	27
5.1	Terskling	27
5.2	Klassifisering	28
5.3	Tidsserieanalyse	29
5.4	Rasterkalkulering	32
5.5	Sammenligning av resultater	33
6	Diskusjon	38
6.1	Resultater	38
6.2	Funnesnes pålitelighet	40
6.3	Usikkerhet i metode	40
6.3.1	Preprosessering	40
6.3.2	Polarisering	40

6.3.3	Terskling og klassifisering	40
6.4	Andre usikkerheter	41
6.4.1	Skygger	41
6.4.2	Vegetasjon	41
7	Konklusjon	42
7.1	Videre arbeid	42
	Litteraturliste	44
A	Klassifiseringsrapporter	47
A.1	Klassifiseringsrapport juli	47
A.2	Klassifiseringsrapport oktober	47

Figurer

1.1	Dårlig drenert jord fører til overvannsproblematikk ved kraftig nedbør. Foto: Erling Fløistad, NIBIO [Oldertrøen, 2021]	1
1.2	Klassifisering av jordbruksareal som ble øversvømt under orkanen Harvey [Boryan et al., 2018]	3
1.3	VH-polarisert bilde (til høyre), VV-polarisert bilde (i midten) og oversvømt areal utledet fra disse. [Clement et al., 2018]	3
2.1	Opptaksmodus for Sentinel-1 over land- og havområder [ESA, d]	6
2.2	Lengden den fysiske antennen beveger seg mens den gjør opptak av samme område på bakken utgjør den syntetiske antennelengden [Hanssen, 2001]	7
2.3	Ulike returtyper [Meyer, 2019]	8
2.4	Ulike spredningstyper [FFI, 2019]	8
2.5	De fire mulige kombinasjonene av utsendt og mottatt lineær polarisering [ESA, e]	9
2.6	Sentinel-1-bilde over Oslo og omegn	10
2.7	Gråtonebilde tersklet med to ulike terskelverdier [Caubalejo, 2021]	12
2.8	Eksempel på valgtre. I en klassifiseringsmodell vil valgtreet bestå av tallverdier. [Hoare]	13
2.9	Predikasjonen til seks separate valgtreer (A,B,A,B,B,B) legges sammen til en felles predikasjon (B) [Gujral, 2020]	13
3.1	Sentinel-1 tile (blå) og definert studieområde (gul)	14
3.2	Studieområdets utstrekning, presentert med <i>Norges grunnkart WMS</i> [Kartverket]	15
3.3	Nedbørsmengde i forhold til normalen for målestasjon Ås, januar-desember 2020	16
3.4	Nedbørsmengde i forhold til normalen for utvalgte målestasjoner i studieområdet, juli og oktober 2020	16
3.5	Nedbørsmengde per døgn innad i tidsrom 1, målestasjon Ås	17
3.6	Nedbørsmengde per døgn innad i tidsrom 2, målestasjon Ås	17
3.7	Nedbørsmengde per døgn for målestasjon Ås, juli 2018	17
4.1	Preprosesseringsløype satt opp i <i>Graph builder</i> i SNAP.	21
4.2	Utsnitt fra originalt GRD-produkt (til venstre) og preprosessert bilde (til høyre)	22
4.3	Utsnitt polarisert i VH (til venstre) og VV (til høyre), med tilhørende histogrammer	23

4.4	Binærisert bilde med terskelverdier -22 (VH) og -17 (VV)	23
4.5	Binærisert bilde med terskelverdier -28 (VH) og -20 (VV)	24
4.6	Eksempel på treningsdata for vann (rød) definert innenfor FKB-Vann innsjø (blå)	24
4.7	Eksempel på linjeplott som viser utviklingen av tilbakespredning over tid, i en tilfeldig valgt enkelt piksel.	25
4.8	Utsnitt fra NDWI	26
5.1	Eksempel på resultater fra terskling i opptak fra oktober	27
5.2	Eksempel på resultater fra terskling i opptak fra oktober	27
5.3	Eksempel på resultat fra klassifisering i opptak fra oktober	28
5.4	Eksempel på resultat fra klassifisering i opptak fra oktober	28
5.5	Tidsserieplottene til seks piksler som har blitt klassifisert som vann i ulike områder, for VH- og VV-polarisering	29
5.6	Typisk tidsserieplott for de tre typene piksler	30
5.7	Utviklingen i intensitetsveri mellom de to tidsrommene for piksler klassifisert som vann	30
5.8	Tidsserieplottene til seks piksler som har blitt klassifisert som vann i ulike områder, for VH- og VV-polarisering	31
5.9	Fra venstre til høyre: Bilde fra 2018, bilde fra 2020 og resultatraster	32
5.10	Ekstempel på resultat fra rasterkalkulering visualisert sammen med resultat fra klassifiseringen i bildet fra juli	32
5.11	Sammenligning av resultat fra klassifisering (gul) og terskling (rød) for opptak fra oktober	33
5.12	Sammenligning av resultat fra klassifisering (gul) og terskling (rød) for opptak fra oktober	33
5.13	Eksempel på resultat fra klassifisering (gul) og terskling (rød)	34
5.14	Sammenligning av resultat fra klassifisering i opptak fra juli (rød) og oktober (gul)	34
5.15	Eksempler på resultat fra klassifisering i opptak fra juli (rød) og oktober (gul)) . .	35
5.16	Sammenligning av resultat fra klassifisering (gul) og terskling (rød) av bilde fra juli	35
5.17	Sammenligning av resultat fra klassifisering (venstre) og NDWI (høyre)	36
5.18	Sammenligning av resultat fra klassifisering (venstre) og NDWI (høyre)	37
5.19	Sammenligning av resultat fra klassifisering (venstre) og NDWI (høyre)	37
6.1	Hølandselva, juli (til venstre) og oktober (til høyre)	38
6.2	GRD-produkt fra juli (til venstre) og oktober (til høyre)	39
6.3	GRD-produkt fra juli (til venstre) og oktober (til høyre)	39
6.4	Eksempler på piksler som kan være feilklassifisert på grunn av radarskygge	41

Tabeller

2.1	Opptaksmodus for Sentinel-1 med egenskaper [ESA, d]	5
2.2	Egenskaper og bruksområder til bånd brukt i SAR [Meyer, 2019]	10
2.3	Spektrale bånd i Sentinel-2A MSI [EOS]	11
3.1	Sentinel-1 GRD-produktene som er brukt i oppgaven	18
5.1	Prosent oversvømt jordbruksareal detektert i bildene med de to metodene	36

1 Innledning

De siste årene er det observert en økende nedbørsmengde i Norge og store deler av verden. Dette kan sees i sammenheng med en økt lufttemperatur – når temperaturen øker, kan lufta holde på mer vann, og da vil det også falle ned mer vann. Nedbørsøkningen skyldes både flere dager med nedbør enn tidligere, men også mer nedbør de respektive dagene. Størst økning har det vært i korttidsnedbør, det vil si nedbør som kommer på kortere tidsrom enn ett døgn. [Met.inst., 2021]

Landbruket, med sine verdifulle jordbruksarealer, er et av områdene som er spesielt sårbare for disse klimaendringene. Særlig er økningen i de intense nedbørsperiodene med store mengder nedbør på kort tid en fare for sektoren. Resultatet av slike kraftige regnfall er i mange tilfeller reduserte eller ødelagte avlinger, forårsaket av store mengder overvann som blir liggende på matjorda.

For å redusere skadeomfanget på jordbruksarealene ved økende nedbørsmengde, kan det settes inn avbøtende tiltak i form av drenering. I Norge ble det gjort mye dreneringsarbeid for 30-40 år siden, et arbeid som nå er i ferd med å gå ut på dato. Grøfter faller sammen og drensrør tettes, og mange områder har nå behov for en ny runde med dreneringsarbeid for å kunne takle klimaendringene. I tillegg til å redusere tapet av verdifulle avlinger, vil drenering også bidra til redusert klimagassutslipp i form av redusert utslipp av lystgass (N_2O) fra jorda. [Oldertrøen, 2021].



Figur 1.1: Dårlig drenert jord fører til overvannsproblematikk ved kraftig nedbør.
Foto: Erling Fløistad, NIBIO [Oldertrøen, 2021]

1.1 Bakgrunn

Bakgrunnen for denne oppgaven er at NIBIO ønsker å kartlegge jordbruksareal som er utsatt for oversvømmelse ved kraftige nedbørsperioder, og dermed få bedre oversikt over i hvilke områder det er behov for å sette inn avbøtende tiltak.

Behov for tilgang til kontinuerlig oppdaterte opptak av jordoverflaten, gjør fjernmålingsdata fra satellitt egnet til formålet. For å oppdage oversvømt areal bør man bruke benytte opptak fra under eller rett etter kraftige nedbørsperioder. I slike perioder vil værforholdene kunne gjøre det utfordrende å benytte opptak fra optiske sensorer, som blir påvirket av nedbør og skydekke. Muligheten til å kunne benytte data fra alle slags værforhold gjør det ønskelig å undersøke om data fra radarsatellittene Sentinel-1 egner seg til en slik kartlegging av oversvømt areal for norske forhold.

Dersom resultatet av analysen er tilfredsstillende, er det tenkbart at NIBIO vil ta i bruk data fra Sentinel-1 og bruke elementer av oppgaven som bakgrunn for en tilsvarende kartlegging på nasjonalt nivå.

1.2 Problemstilling

Det er kjent at radarsignaler i liten, eller ingen, grad reflekteres tilbake til avsenderen når det treffer vannoverflater som elver, innsjøer og hav. På grunn av sin lave tilbakespredningsintensitet kommer slike områder derfor tydelig frem i radarbilder. Oversvømt jordbruksareal som ønskes kartlagt vil i mange tilfeller være forholdsvis små i utstrekning, og heller ikke veldig dype. Det er dermed usikkert hvor godt disse områdene vil la seg detektere i radarbildene. Målet med oppgaven er som nevnt i 1.1 å undersøke hvor godt data fra Sentinel-1 egner seg til en kartlegging av oversvømt jordbruksareal. Første forskningsspørsmål kan formuleres slik:

- *I hvilken grad kan oversvømt jordbruksareal detekteres i opptak fra radarsatellitten Sentinel-1?*

Siden det ikke eksisterer noe fasit-datasett som kan brukes til å verifisere nøyaktigheten av resultatene fra analysen, blir alle resultater å regne som antagelser. På grunnlag av dette tilkommer andre forskningsspørsmål:

- *Kan resultatene fra analysen basert på data fra Sentinel-1 styrkes av resultat fra en annen datakilde?*

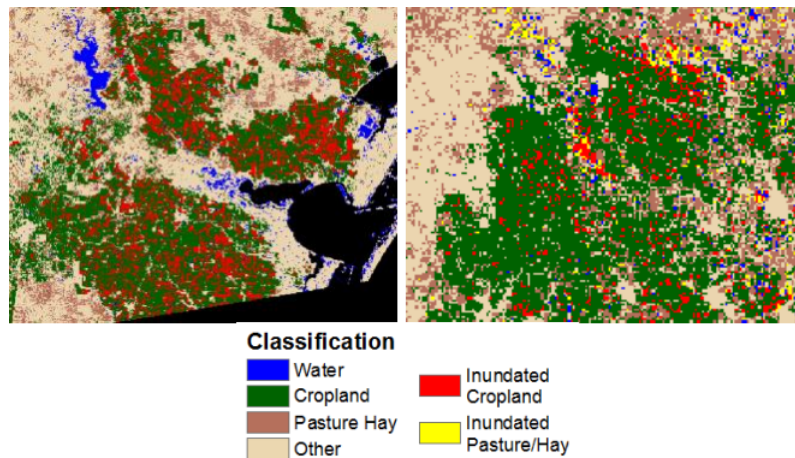
1.3 Relatert arbeid

Tradisjonelt har data fra optiske sensorer og visuell tolkning av slike data blitt benyttet mye til endringsanalyse og deteksjon av fenomener som oversvømmelser på jordoverflata. Utviklingen av SAR-teknologien, samt moderne programvare og teknikker for prosessering og analysing av slike data, har ført til økende anvendelse av SAR til tilsvarende formål. Den enkle og almenne tilgjengeligheten til Sentinel-1-data fra hele verden gjør denne plattformen godt brukt og dokumentert som datakilde til tolkning av SAR-data.

I artikkelen *SAR-based detection of flooded vegetation – a review of characteristics and approaches* [Tsyganskaya et al., 2018] presenteres en sammenstilling av 83 studier om deteksjon av oversvømt vegetasjon basert på SAR. Det ble funnet trender angående hvilke bølgelengder, polariseringer og innfallsvinkler som ble benyttet i analysene, men det var stor variasjon i klassifiseringsteknikker og bruk av informasjon avhengig av datasettene og studieområdet. Imidlertid benyttet bare én av disse studiene Sentinel-1 som datakilde.

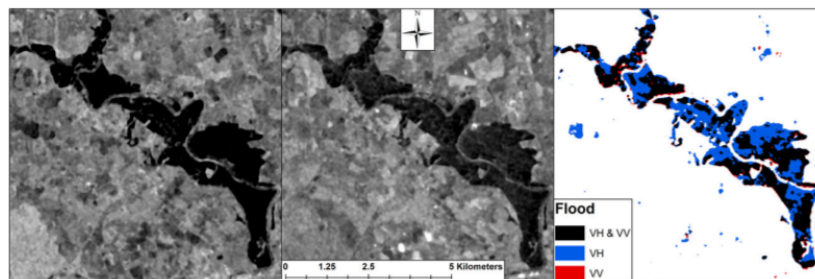
Sentinel-1 er dokumentert benyttet til flomanalyse ved store naturkatastrofer, som ved orkanen Harvey i 2017 [Boryan et al., 2018]. I studien undersøkes det om Sentinel-1 egner seg til overvåking

av dyrka mark ved slike hendelser i USA. Det har blitt utført en terskling på bilder fra før og etter orkanen, og oversvømte områder har blitt bestemt basert på forskjellen mellom disse. Datalag med oversikt over avlingstype har blitt inkludert for å se hvilke avlingstyper som ble berørt av oversvømmelsene. Resultatene er tilfredsstillende, og studien viser at SAR-data fra Sentinel-1 er en effektiv og verdifull datakilde for katastrofevurdering i landbruket.



Figur 1.2: Klassifisering av jordbruksareal som ble øversvømt under orkanen Harvey [Boryan et al., 2018]

I artikkelen *Multi-temporal synthetic aperture radar flood mapping using change detection* [Clement et al., 2017] presenteres en endringsanalyse basert på Sentinel-1-bilder fra vinterflom i Yorkshire, hvor det blir utført terskling for å detektere flomområder. Både VH- og VV-polarisering blir benyttet, og disse gir lik nøyaktighet ved validering av resultatene. I mangel på fasitdata blir *Modified normalised difference water index (MNDWI)* fra Sentinel-2-data brukt til validering.



Figur 1.3: VH-polarisert bilde (til høyre), VV-polarisert bilde (i midten) og oversvømt areal utledet fra disse. [Clement et al., 2018]

I Norge er bruk av Sentinel-1 til flomanalyse mindre dokumentert, men en meget relevant analyse er utført av Norsk Regnesentral på vegne av NVE [Reksten et al., 2019]. I analysen utføres en Random Forest-klassifisering på Sentinel-1-data fra før og under flomhendelser, i områder med kjente oversvømmelser. Det blir også tatt hensyn til andre faktorer som endres mellom bildene, som høstede avlinger og smeltet is. Resultatene er gode, 168 av 179 merkede oversvømmelser blir oppdaget.

Programvaren som hovedsakelig benyttes i analysen, SNAP, er godt dokumentert gjennom publiserte veiledninger hvor den har blitt benyttet på analyser av data fra Sentinel-1. Preprosseseringen og tidsserieanalysen er basert på Brauns *Time-series analysis with Sentinel-1* [Braun, 2021], hvor det blir benyttet multitemporære data til å studere forandring i vegetasjon på jordbruksområder over tid. McVitties *Flood mapping tutorial* [McVittie, 2019] beskriver også en tilsvarende preprosseseringsløype, samt tersklingsprosessen ved en flomanalyse.

1.4 Struktur av oppgaven

Oppgaven er delt inn i sju hovedkapitler med følgende struktur og innhold:

Kapittel 1 - Innledning:

Presentasjon av bakgrunn og motivasjon for oppgaven, med valgt problemstilling og relevant litteratur fra tidligere forskning.

Kapittel 2 - Teori:

Teorien som danner det teoretiske grunnlaget for oppgaven, med hovedvekt på prinsippene bak SAR-teknologien.

Kapittel 3 - Material:

Presentasjon av oppgavens datagrunnlag i form av beskrivelse av filformater, datasett og programvare som er brukt, samt beskrivelse av valgt tidsrom og studieområde.

Kapittel 4 - Metode:

Beskrivelse av fremgangsmåtene som har ført frem til resultatene. Analysen av data fra Sentinel-1 er delt opp i en monotemporær og en multitemporær del.

Kapittel 5 - Resultat:

Presentasjon av resultatene fra analysen basert på fremgangsmåten i metodedelen. Resultatene presenteres og sammenlignes hovedsakelig visuelt.

Kapittel 6 - Diskusjon:

Vurdering av resultatene fra resultatdelen, samt diskusjon rundt usikkerheter, valg av metode og funnenes pålitelighet.

Kapittel 7 - Konklusjon:

Oppsummering av oppgaven som besvarer forskningsspørsmålene ved å sammenstille problemstilling og resultater. Avsluttes med anbefalinger og forslag til videre arbeid.

2 Teori

2.1 Sentinel-1

Sentinel-1 er det første av fem oppdrag i Copernicus-programmet utviklet av Den europeiske romfartsorganisasjonen, ESA. Oppdraget omfatter en konstellasjon av tvillingsatellittene Sentinel-1A og Sentinel-1B, i drift fra henholdsvis 2014 og 2016. Sentinel-1-satellittene bærer en c-bånds syntetisk apertur-radar, et bildedannende radarinstrument, for å innhente data til oppdragets formål. [ESA, a] Radar- og SAR- teknologien blir forklart i påfølgende underkapitler.

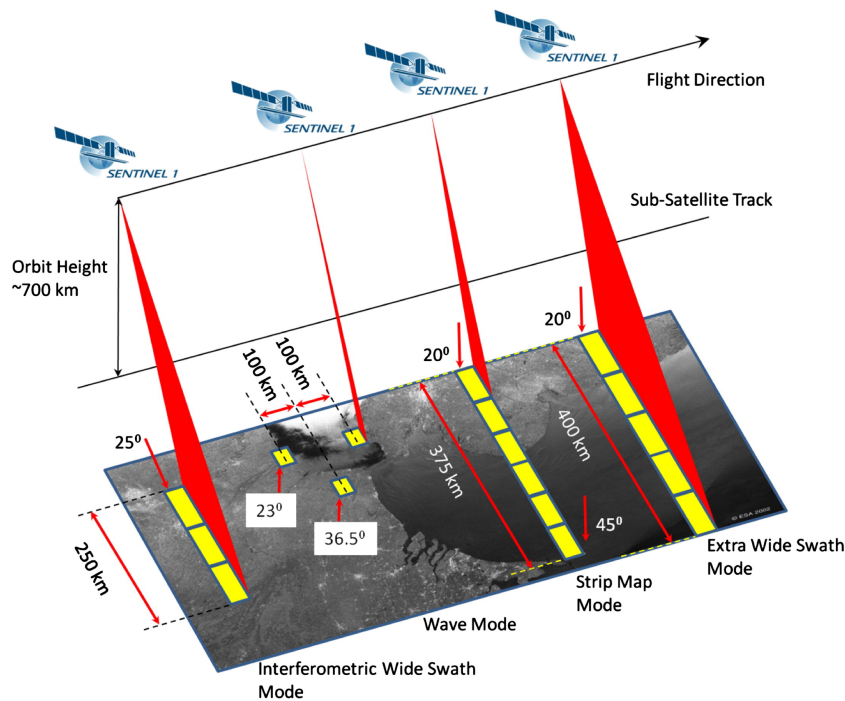
Hovedformålet med Sentinel-1-oppdraget er todelt. I tillegg til å skulle utføre en systematisk og kontinuerlig overvåking av jordoverflaten, er formålet å sikre systemrespons ved katastrofer og nødsituasjoner. Noen konkrete eksempler på bruksområder for Sentinel-1 er overvåking av isbreer og marine områder, deteksjon av flomområder og oljesøl, kontroll av landbevegelser, og kartlegging av arealressurser som skog-, vann- og jordbruksområder. [ESA, b]

Med en høy temporær oppløsning er Sentinel-1 godt egnet til begge sine hovedformål. Hver av Sentinel-1 satellittene beveger seg i identisk nærpolar solsynkron bane på rundt 700 kilometers høyde, med 180 graders faseforskyvning og 12-dagers repetisjonssyklus. Det vil si at de to satellittene i konstellasjonen til sammen har en temporær oppløsning på seks dager, ved ekvator. Siden baneavstanden er kortere ved høyere breddegrader, vil nordligere områder ha en enda høyere temporær oppløsning. [ESA, c]

Sentinel-1 kan operere i fire ulike opptaksmodus med ulik skårbredde (dekningsområde), insidensvinkel, polariseringsmulighet og romlig oppløsning [ESA, d]. Teorien bak polarisering blir utdypet i kapittelet om SAR. Tabell 2.1 gir en oversikt over de ulike opptaksmodusene. Over landområder opererer Sentinel-1 hovedsakelig i opptaksmodus IW, og i WV over havområder.

Opptaksmodus	Skårbredde	Insidensvinkel	Polarisering	Bakkeoppløsning
Stripmap (SM)	80 km	18.3° - 46.8°	HH, VV, HH+HV, VV+VH	5 x 5 m
Interferometric Wide swath (IW)	250 km	29.1° - 46.0°	HH, VV, HH+HV, VV+VH	5 x 20 m
Extra-Wide swath (EW)	410 km	18.9° - 47.0°	HH, VV, HH+HV, VV+VH	20 x 40 m
Wave (WV)	20 x 20 km	21.6° - 25.1° 24.8° - 38.0°	HH, VV	5 x 5 m

Tabell 2.1: Opptaksmodus for Sentinel-1 med egenskaper [ESA, d]



Figur 2.1: Opptaksmodus for Sentinel-1 over land- og havområder [ESA, d]

2.2 Radar

Radar, akronym for *R*Adio *D*etection *A*nd *R*anging, er en anordning som benytter elektromagnetisk stråling til å detektere og lokalisere reflekterende objekter [Skolnik, 1990].

Et radarsystem består i sin enkleste form av en transmitter (sender), en mottaker og en antenne. Transmitteren generer elektromagnetiske bølger med passende bølgelengde, som via antennen forplanter i rommet. Noe av den elektromagnetiske strålingen vil treffe måleobjekter på bakken eller i luften, og vil reflekteres med ulik styrke i ulike retninger, avhengig av objektene egenskaper. Noe av denne reflekterte strålingen, kalt ekko, vil returnere til radarsystemet og samles opp av antennen. Ekkosignalet vil så bli forsterket og behandlet i mottakeren, for å kunne brukes til å bestemme måleobjektets posisjon. [Skolnik, 2008] Avstanden til objektet bestemmes ved å beregne tiden det tar fra signalet forlater antennen til det returnerer i form av ekko. Retningen, eller vinkelen, til objektet bestemmes ut fra ankomstretningen til det reflekterte signalet. [Skolnik, 1980]

På grunn av at radar bruker sin egen transmitter til å sende ut signal, kalles det en aktiv sensor. Det vil si at sensoren ikke er avhengig av eksterne strålingsskilder, slik som de fleste optiske og infrarøde sensorer er. Radarpulsenes bølgelengde gjør at signalene kan penetrere meteorologiske fenomener som regn, snø og skyer. Dette vil i praksis si at et radarinstrument kan detektere og bestemme posisjonen til relativt små objekter på lang avstand, uavhengig av lys- og værforhold. [Skolnik, 1990]

De første komplette radarsystemene fra midten av 1930-tallet var bakkebaserte, og ble brukt til å lokalisere luftbårne objekter, hovedsakelig fly. Gjennom andre verdenskrig utviklet radarteknologien seg raskt, og siden starten av 1940-tallet har radarinstrumentene vært små nok til å implementeres i fly. Dette har gitt radar et mye større bruksområde, og har siden da vært mye brukt til jordobservasjon. [Meyer, 2019] Fra midten av 1960-tallet har satellitt vært en viktig plattform for radarinstrumenter, og SAR-teknologien har gjort det mulig å kunne detektere små objekter på jordoverflaten fra hundrevis av kilometers avstand.

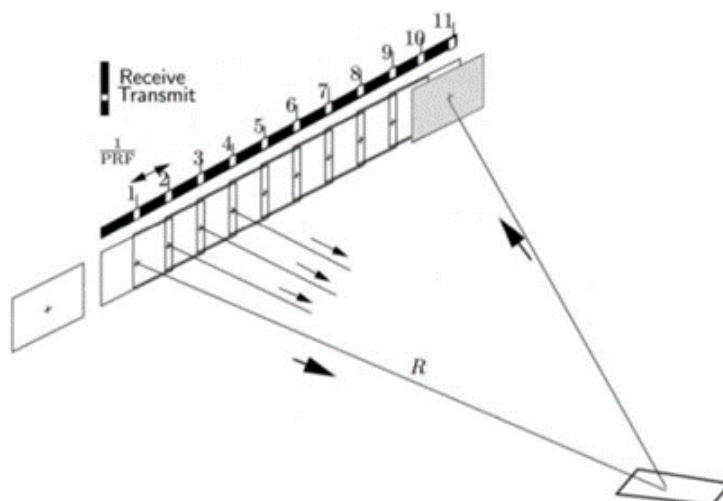
2.3 SAR

2.3.1 Grunnleggende prinsipp

SAR står for *Synthetic Aperture Radar*, eller *syntetisk apertur-radar* på norsk, og er en klasse radarsystemer kalt bildedannende radar [Hanssen, 2001]. Prinsippet bak SAR ble oppdaget av Carl Wiley i 1951, med ideen om at hvis radarpulser er koherente, kan pulsene behandles i sammenheng med hverandre og man kan ut fra dette konstruere todimensjonale bilder med kraftig forbedret oppløsning [Lasswell, 2005].

SAR-prinsippet kan lettest forklares ved å sammenlignes med det enkleste form for radarsystem, *Real Aperture Radar* (RAR). Hos RAR er oppløsningen proporsjonal med den fysiske antennestørrelsen til instrumentet og avstanden til måleobjektet. RAR-målinger fra satellitt med bakkeoppløsning ned mot noen meters nivå ville krevd en usannsynlig stor antennestørrelse. Fordi satellittbanehøyden er konstant høy, ville man måtte redusert nøyaktigheten på oppløsningen kraftig for å redusere antennestørrelsen til et realistisk nivå.

Den fysiske antennestørrelsen til SAR-instrumenter som brukes til jordovervåkning er liten, eksempelvis 12,5 meter hos Sentinel-1 [ESA, b]. Det som skiller SAR fra RAR er at antennestørrelsen som har størst innvirkning på oppløsningen i realiteten ikke er størrelsen til den fysiske antennen, men en syntetisk («kunstig») antenne. Den syntetiske antennen blir konstruert ved at den fysiske antennen forflytter seg over en avstand mens den gjør målinger av samme punkt på bakken. Den syntetiske antennelengden er lik avstanden fra første til siste måling. [Skolnik, 1990]



Figur 2.2: Lengden den fysiske antennen beveger seg mens den gjør opptak av samme område på bakken utgjør den syntetiske antennelengden [Hanssen, 2001]

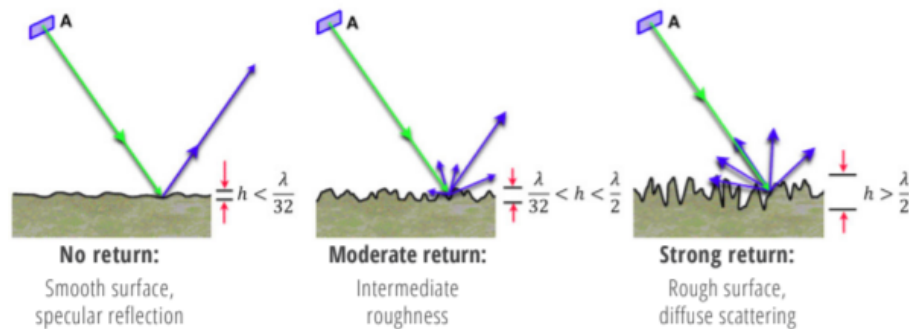
Lengden til den syntetiske antennen, L_{SA} , gis ved formel 2.1, hvor λ er bølgelengde, L er den fysiske antennelengden, R_0 er avstand til måleobjekt, og β er strålebredden $\frac{\lambda}{L}$ [Meyer, 2019].

$$L_{SA} = \frac{\lambda}{L} \cdot R_0 \approx \beta \cdot R_0 \quad (2.1)$$

2.3.2 Tilbakespredning

Når radarbølgen treffer en overflate på bakken blir den spredt i en eller flere retninger før den helt eller delvis finner veien tilbake til SAR-antennen som ekko. Tilbakespredningsintensiteten er et mål på hvor sterke disse retursignalene er. Tilbakespredningsintensiteten fra et mål avhenger av overflatens geometriske struktur, elektriske egenskaper, ruhet og fuktighetsinnhold, samt polarisering og innfallsvinkel til radarpulsen [FFI, 2020].

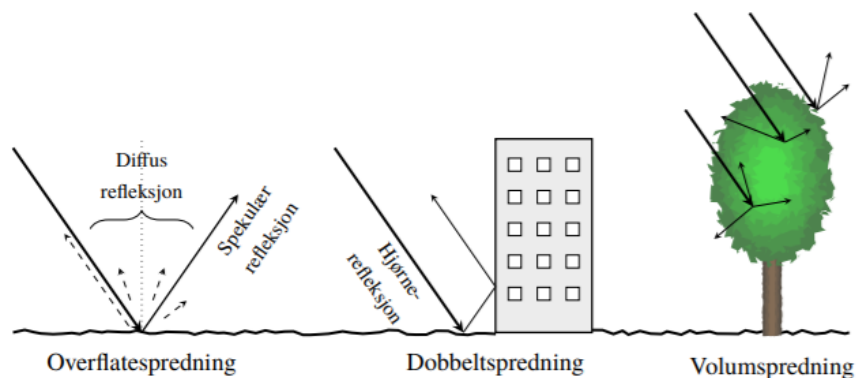
En tommelfingelregel er at jo ruere overflaten er, jo høyere er tilbakespredningsintensiteten. En glatt overflate som vann eller asfalt har lav tilbakespredning. Helt stillestående vann kan ses på som en perfekt glatt overflate, og vil gi null tilbakespredning. Dette kalles spekulær refleksjon. Middels ru overflater, som de fleste typer vegetasjon, vil gi en moderat tilbakespredning, eller diffus refleksjon. Svært ujevne overflater som skrenter og byområder gir høy grad av tilbakespredning. [Meyer, 2019]



Figur 2.3: Ulike returtyper [Meyer, 2019]

Spredningen av radarsignalene kan deles inn i tre kategorier [Meyer, 2019];

- *Overflatespredning*, forekommer når signalet treffer en glatt eller middels ru overflate, som lav vegetasjon, vann og veier. Signalet reflekteres direkte til rommet.
- *Dobbeltspredning*, forekommer når signalet treffer vertikale strukturer som høye bygninger, tre-stammer og skrenter. Signalet gjør et sprett før det reflekteres til rommet.
- *Volumspredning*, forekommer når signalet treffer trekroner og annen høy vegetasjon. Signalet spretter gjentatte ganger mens det forplanter seg gjennom vegetasjonen, før det reflekteres til rommet.



Figur 2.4: Ulike spredningstyper [FFI, 2019]

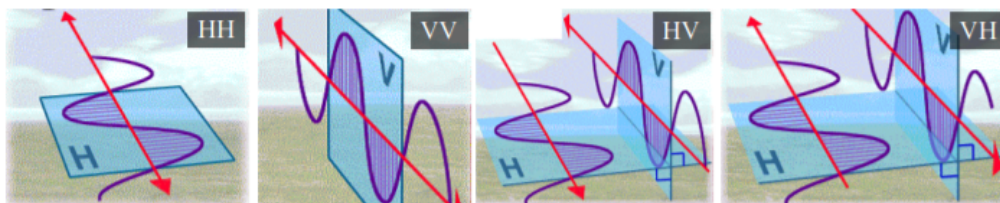
2.3.3 Polarisering

Fordi SAR er en aktiv sensor, kan man kontrollere og utnytte polariseringen av både sendt og mottatt signal. Polarisering er en egenskap ved elektromagnetiske bølger, som beskriver orienteringen til svingningsplanet til bølgen. [Meyer, 2019] Opptaksinnstillinger på sensoren avgjør typen polarisering.

Polariseringsprinsippet kan visualiseres ved at man ser for seg at man holder i enden av et tau. Dersom man beveger hånda fram og tilbake sidelengs, vil bølgen som oppstår forplante seg i et horisontalt plan, bølgen er horisontalt polarisert (H). Dersom man beveger hånda opp og ned, vil bølgen forplante seg i et vertikalt plan, bølgen er vertikalt polarisert (V). Ved horisontal polarisering står det elektriske feltet til den utsendte bølgen fra sensoren omtrent normalt på innfallsplanet. Ved vertikal polarisering er det elektriske feltet parallelt med innfallsplanet. [FFI, 2019]. Slike polariseringer hvor svingningsplanet er konstant langs forplantningsbanen til bølgen, kalles lineær polarisering. [Meyer, 2019].

De fleste SAR-instrumenter er lineært polariserte, og de fleste nyere instrumenter støtter både kopolarisering og krysspolarisering. Kopolarisering betyr at polariseringen til utsendt og mottatt bølge er lik – polariseringen er da enten VV (vertikal utsendt, vertikal mottatt) eller HH (horisontal utsendt, horisontal mottatt). Ved krysspolarisering kan utsendt og mottatt bølge ha ulik polarisering, og polariseringen kan da være VH (vertikal utsendt, horisontal mottatt) eller HV (horisontal utsendt, vertikal mottatt). [FFI, 2019]

Sensoren på Sentinel-1 støtter dobbelpolarisering, som betyr at den kan samle forskjellige bilder med ulike polarisering fra samme serie med pulser [Meyer, 2019]. I tabell 2.1 i kapittel 2.1 kan man se opptaksmodus SM, IW og EW gjør det mulig å få radarbilder fra samme opptak med polariseringene HH og HV eller VV og VH. Opptaksmodus WV støtter bare enkeltpolarisering, altså enten HH eller VV.



Figur 2.5: De fire mulige kombinasjonene av utsendt og mottatt lineær polarisering [ESA, e]

Det er viktig å kjenne til polariseringen når man skal benytte et radaropptak, fordi ulike polariseringskanaler påvirker tilbakespredningen ulikt. Spredningstypene som er forklart i kapittel 2.3.2 sprer radarsignalet med ulik styrke avhengig av polarisering. Blant annet gir VV-polarisering sterkest overflatespredning, HH-polarisering sterkest dobbeltspredning og HV- og VH-polarisering gir sterkest volumspredning. [Meyer, 2019]

2.3.4 Bølgelengdeområde

Som nevnt i kapittel 2.1, er Sentinel-1 bærer av et C-bånds SAR-instrument. Begrepet *bånd* er i denne sammenhengen det samme som bølgelengdeområde. SAR-instrumenter kan sende og motta radiobølger med bølgelengder på mellom 2.4 og 100 cm. C-båndet omfatter bølgelengder fra 3.8 til 7.5 cm, og er hovedbåndet som blir brukt i de fleste SAR-instrumenter. Andre bånd som blir benyttet i SAR-instrumenter er X-, S-, L- og P-båndene. [Meyer, 2019]

Båndene har ulike bruksområder, fordi ulike bølgelengder har ulik evne til å penetrere overflatetyper. Tabell 2.2 beskriver egenskaper og bruksområder til bånd benyttet i SAR-instrumenter.

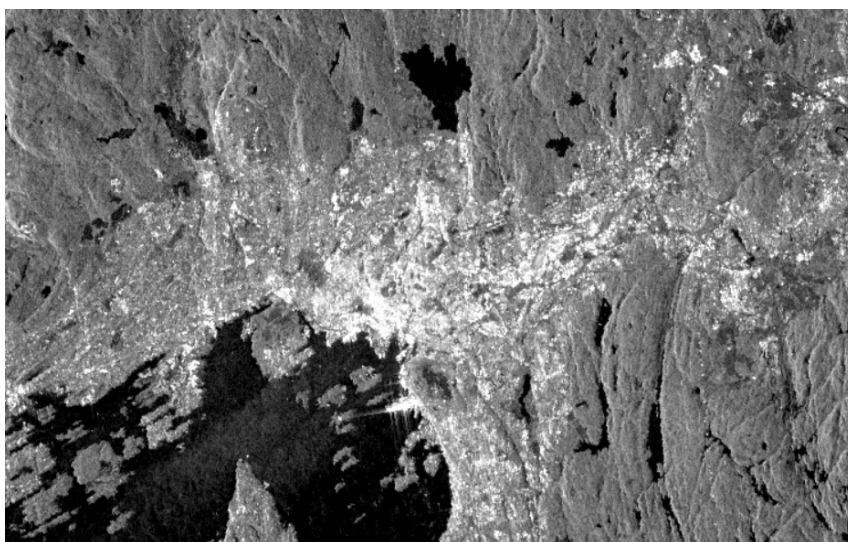
Bånd	Frekvens	Bølgelengde	Bruksområde
X	8-12 GHz	2.4-3.8 cm	Høyoppløselig SAR - Urban overvåking, is, snø. Lite penetrasjon i vegetasjonsdekke.
C	4-8 GHz	3.8-7.5 cm	Hovedbåndet benyttet i SAR - Global kartlegging, endringsdeteksjon, is, maritime områder, overvåking av områder med lav til middels høy vegetasjon. Medium penetrasjonsevne.
S	2-4 GHz	7.5-15 cm	Liten, men økt bruk i jordobservasjon - Overvåking av jordbruksområder. Medium penetrasjonsevne.
L	1-2 GHz	15-30 cm	Medium-oppløselig SAR - Geofysisk overvåking, biomasse- og vegetasjonskartlegging, InSAR. Høy penetrasjonsevne.
P	0.3-1 GHz	30-100 cm	Nytt eksperimentielt bånd i SAR - Vegetasjonskartlegging og overvåking av biomasser. Høy penetrasjonsevne

Tabell 2.2: Egenskaper og bruksområder til bånd brukt i SAR [Meyer, 2019]

2.3.5 SAR-bilder

Et SAR-bilde er et gråtoneraster hvor intensitetsverdien i pikslene representerer tilbakespredningsintensiteten til radarsignalet i dette området. Innenfor utstrekningen til hver piksel er det ikke bare én, men mange reflektorer. Det betyr at radarsignalet som returneres som ekko til mottakeren egentlig er en sum av refleksjoner fra mange punkter som ikke kan skilles fra hverandre. [FFI, 2019]

Overflatetyper som gir sterk tilbakespredning vil fremstå som hvite eller lys grå i SAR-bildet. Typisk vil byområder med mange høye bygninger som gir dobbeltspredning fremstå veldig lyse. Periodiske strukturer på bakken, for eksempel pløyde jorder, kan fremstå svært lyse dersom lengden mellom bølgetoppene er omtrent lik bølgelengden til radarsignalet. Overflatetyper med middels tilbakespredning vil fremstå som grå i bildet. Skogsområder kan fremstå som både lyse og mørke på grunn av volumspredning. Overflatetyper med liten eller ingen tilbakespredning vil fremstå som mørk grå og svarte - jo glattere overflaten er, jo mørkere i bildet. Radarsignalet penetrerer lettere tørr jord enn våt jord, så jord med mye fuktighet vil dermed i mange tilfeller fremstå lysere i bildet enn tørr jord. [FFI, 2019]



Figur 2.6: Sentinel-1-bilde over Oslo og omegn

2.4 Sentinel-2

Sentinel-2 er, i likhet med Sentinel-1, en konstellasjon av to tvillingsatellitter i Copernicus-programmet. Satellittene beveger seg i en identisk, nærpolar og solsynchron bane med 180 graders faseforskyving. Hver satellitt har en repetisjonssyklus på 10 dager ved ekvator, som vil si fem dager totalt for konstellasjonen. Skårbredden til sensoren er 290 km. [ESA, f]

Sentinel-2-satellittene bærer et multispektralt kamera, kalt *MultiSpectral Instrument (MSI)*. MSI er en optisk, passiv sensor som samler sollyst som er reflektert fra jordoverflaten [ESA, g]. Sensoren støtter datainnsamling i 13 spektrale bånd med ulike bølgelengder og romlig oppløsning, presentert i tabell 2.3.

Båndnummer	Båndnavn	Sentral bølgelengde (nm)	Oppløsning (m)
1	Coastal aerosol	443.9	60
2	Blue	496.6	10
3	Green	560.0	10
4	Red	664.5	10
5	Vegetation red edge	703.9	20
6	Vegetation red edge	740.2	20
7	Vegetation red edge	782.5	20
8	NIR	835.1	10
8a	Narrow NIR	864.8	20
9	Water vapour	945.0	60
10	SWIR - Cirrus	1374.5	60
11	SWIR	1613.7	20
12	SWIR	2202.4	20

Tabell 2.3: Spektrale bånd i Sentinel-2A MSI [EOS]

Spennet fra blått til kortbølget infrarødt bølglengdeområde gjør sensoren egnet til overvåking av mange slags overflater og fenomener til lands, vanns og i lufta. Viktige bruksområder for Sentinel-2 er skog- og vegetasjonskartlegging, overvåking av is og vann, karbonmasseestimering, og kartlegging ved naturkatastrofer og humanitære kriser. [ESA, f]

2.5 NDWI

NDWI står for *Normalized difference water index*, og er en spektral indeks som er følsom for vannoverflater og vegetasjon med høyt vanninnhold. NDWI-indeksen opptrer i forskjellige varianter, her forklares indeksen som er definert av McFeeters [McFeeters1995].

NDWI-indeksen til McFeeters er designet for å maksimere reflektansen fra vannforekomster i grønt bånd, minimere den lave reflektansen til vannforekomster i nærinfrarødt bånd og ta fordel av den høye reflektansen til terrestrisk vegetasjon og jordegenskaper i nærinfrarødt bånd. [Sun et al, 2012]

Båndkombinasjonen i NDWI er gitt ved formel 2.2, hvor *GREEN* er grønt bånd og *NIR* er nærinfrarødt bånd. I tabell 2.3 kan man se at i Sentinel-2-sensoren tilsvarer *GREEN* bånd 3 og *NIR* bånd 8.

$$NDWI = \frac{GREEN - NIR}{GREEN + NIR} \quad (2.2)$$

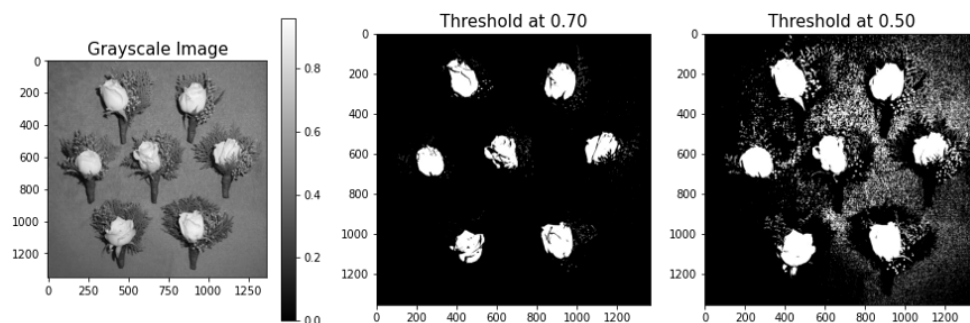
NDWI kan i de fleste tilfeller effektivt detektere vannflater, men indeksen er følsom for bebyggelse, noe som kan resultere i en overvurdering av vannforekomster [Abazaj, 2020]. Indeksen går fra -1 til 1, hvor vannforekomster hovedsakelig har verdier over 0.5, bebygde områder har verdier mellom 0 og 0.2, og vegetasjon og jord har negative verdier. [Sun et al, 2012]

2.6 Terskling

Terskling er den enkleste metoden for bildesegmentering, og kan brukes til å konstruere binære bilder av gråtonebilder [Guruprasad, 2020].

De enkleste tersklingsmetodene erstatter alle pikselverdier i rasteret med enten 0 (svart) eller 1 (hvit), bestemt ut fra en manuelt eller automatisk satt terskelverdi. Metoden kan også beskrives ved formel 2.3, hvor (x, y) er pikselkoordinat, f er pikselverdi og T er terskelverdi.

$$g(x, y) = \begin{cases} 1, & \text{if } f(x, y) > T \\ 0, & \text{if } f(x, y) \leq T \end{cases} \quad (2.3)$$



Figur 2.7: Gråtonebilde tersklet med to ulike terskelverdier [Caubalejo, 2021]

Tersklingsmetoder kan bli kategorisert i seks grupper, etter hva slags informasjon de utnytter for å komme frem til resultatet. En av disse gruppene kalles *histogramform-baserte metoder*, hvor rasterets histogram, med toppe, daler og formasjoner studeres for å sette en terskelverdi. [Sezgin et al., 2004].

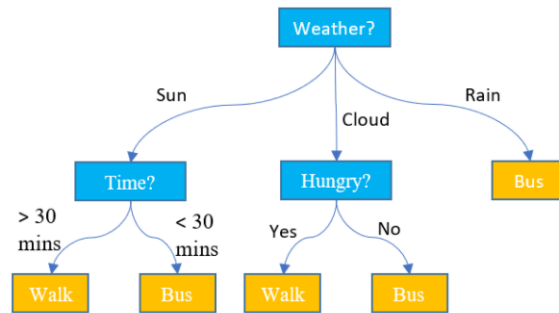
2.7 Styrt klassifisering

Styrt klassifisering er en metode for kvantitativ analyse av bildedata. I likhet med ikke-styrt klassifisering er formålet med klassifiseringen å kategorisere pikslene i bildet i ulike klasser, basert på landdekketyper til korresponderende område på bakken. Ideen med styrt klassifisering er at brukeren definerer treningsdata, i form av å velge ut eksempelpikslers som er representative for valgte klasser. Deretter utfører en styrt klassifiserings-algoritme en klassifisering av alle pikslene i bildet basert på disse treningsdataene. [Punia]

2.7.1 Random forest classifier

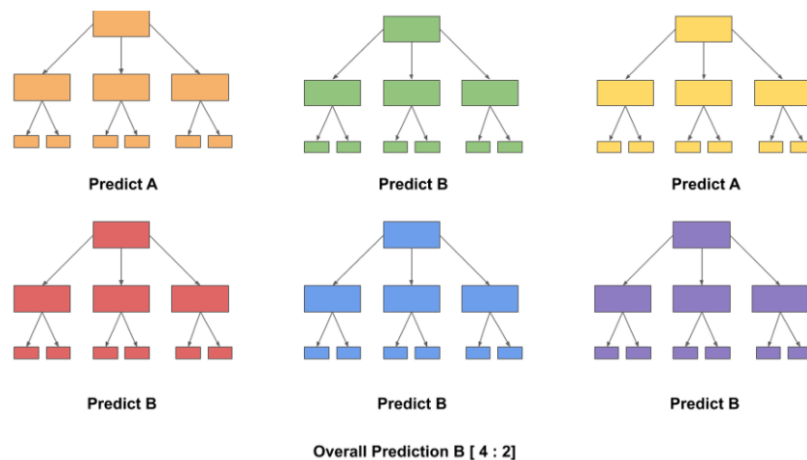
En av mange algoritmer benyttet til styrt klassifisering er *Random forest classifier (RFC)*. Bakgrunnen til navnet *Random forest* er at algoritmen består av en mengde separate, tilfeldig genererte valgtrær, som i fellesskap utgjør en «skog» av valgtrær.

Et valgtre er en klassifiseringsmodell med en trestruktur. Modellen består av beslutninger som må tas for å nå et bestemt mål, i denne sammenheng en klasse. Hver node i valgtreet representerer et spørsmål eller en kriterie, som bryter inputdataene i mindre og mindre deler.



Figur 2.8: Eksempel på valgtre. I en klassifiseringsmodell vil valgtreet bestå av tallverdier. [Hoare]

RFC består av en mengde individuelle valgtrær som samlet utgjør en kraftig klassifiseringsmodell. Hvert tre gjør en predikasjon av klasse, og klassen som predikeres av flest trær blir modellens samlede predikasjon. [Gujral, 2020]



Figur 2.9: Predikasjonen til seks separate valgtrær (A,B,A,B,B,B) legges sammen til en felles predikasjon (B) [Gujral, 2020]

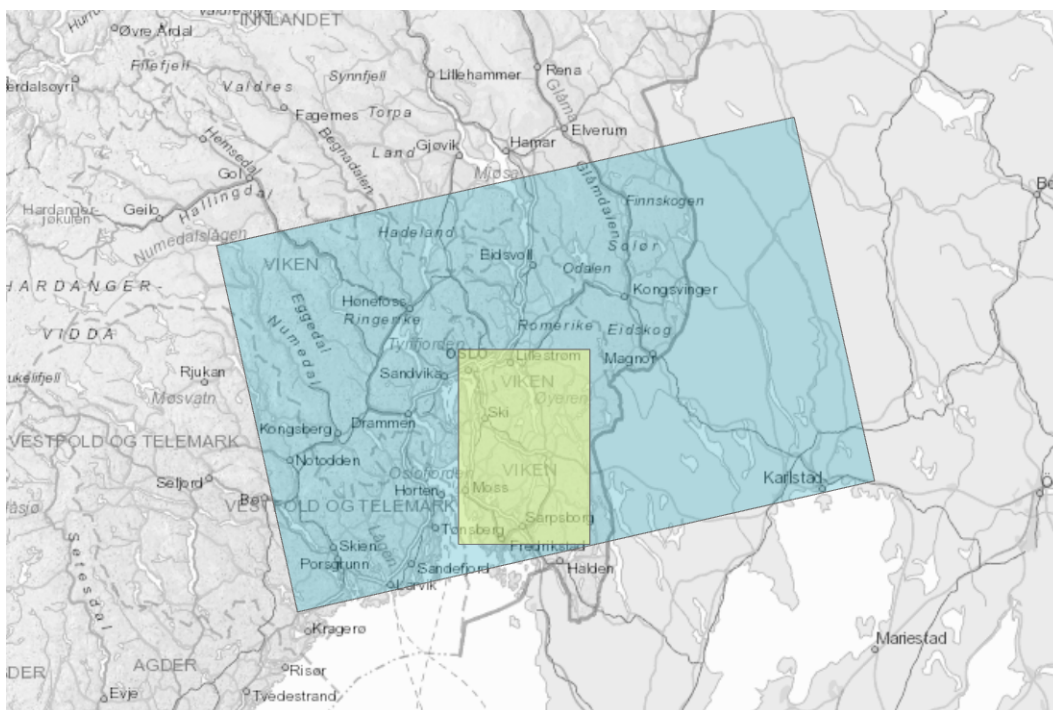
Brukerdefinerte parametere i RFC-algoritmen er valgtrærnes dybde og antall noder. For mange oppdelinger og for høyt antall av iterasjoner kan føre til overtilpasning i valgtreet, som kan gi en feil predikasjon. [Raschka et al, 2019]. Styrken med RFC er at skogen av valgtrær utkonkurrerer de separate trærne. Ved et flertall av riktige predikasjoner vil valgtrærne som predikerer feil klasse ikke ha noe å si for modellens endelige predikasjon. En forutsetning for at RFC skal fungere bra er tilstrekkelig og korrekt treningsdata, slik at modellene ikke bare gjetter seg frem til en predikasjon. I tillegg må predikasjonene til de separate valgtrærne ha lav korrelasjon med hverandre. [Gujral, 2020]

3 Material

3.1 Studieområde

NIBIO ønsket at analysen skulle gjennomføres for et område på østlandet. Studieområdet ble delvis definert basert på plasseringen til Sentinel-1 *tiles*, eller fotavtrykk. Ved å velge et område innenfor samme tile er man sikret at det finnes opptak som dekker hele studieområdet, og opptakene er mindre påvirket av geometriske parametre som insidensvinkel og forskyvning.

Figur 3.1 viser en Sentinel-1 tile over østlandet hvor satellitten beveger seg i stigende bane (*ascending mode*). Dette fotavtrykket ble brukt som utgangspunkt, og det ble innenfor denne avgrensningen definert et mindre studieområde.



Figur 3.1: Sentinel-1 tile (blå) og definert studieområde (gul)

Valgt studieområde dekker store deler av sør-østlige Viken, øst for Oslofjorden. Studieområdet omfatter blant annet Norges største jordbrukskommune, Indre Østfold, samt andre betydelige landbrukskommuner som Ås, Enebakk, Aurskog-Høland og Rakkestad.



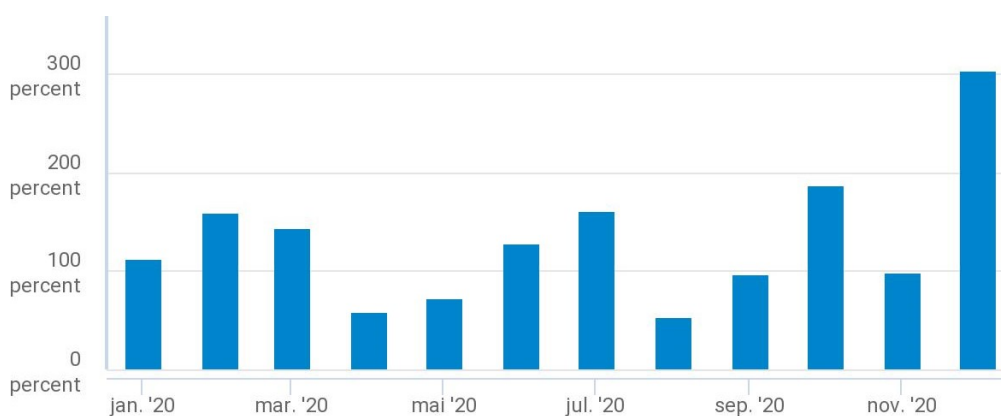
Figur 3.2: Studieområdets utstrekning, presentert med *Norges grunnkart WMS* [Kartverket]

Alle steg i metoddelen ble utført på hele studieområdet, men i resultat- og diskusjonsdelen presenteres resultatene for mindre områder for bedre visualisering.

3.2 Tidsrom

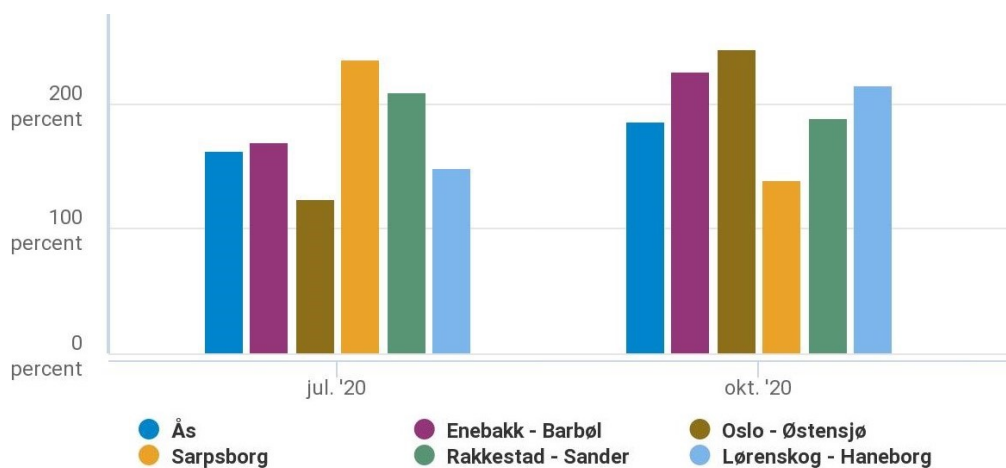
Nettsiden *Norsk Klimaservicesenter* [Nor] ble brukt til å studere nedbørshistorikk for å finne egnede tidsrom for analysen. Alle tall og figurer i dette delkapittelet er hentet fra egendefinerte søk under *Observasjoner og værstatistikk* på denne nettsiden.

Målestasjon Ås (SN17850) ble brukt som utgangspunkt, men andre stasjoner i studieområdet ble også undersøkt. 2020 var et vått år over hele landet, og på Ås var årsnedbøren 31,1% høyere enn normalen (1991-2021). Utenfor vintersesongen var juli og oktober ekstra våte måneder, og det ble valgt å definere et tidsrom for analysen rundt disse månedene.



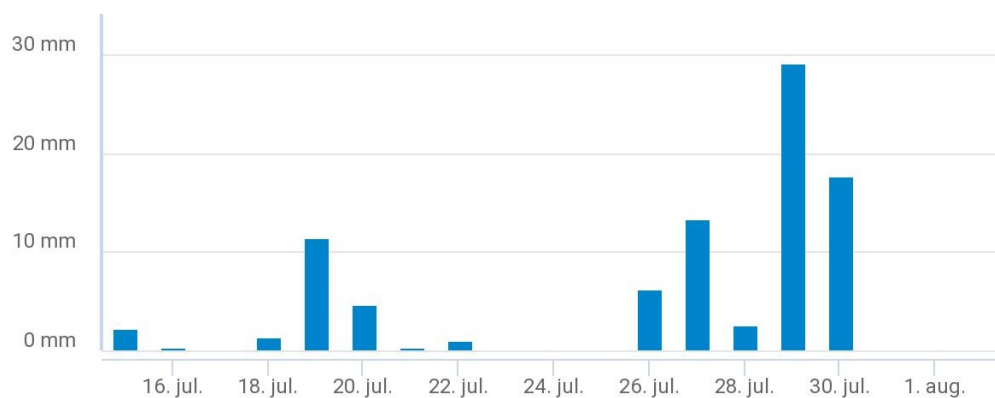
Figur 3.3: Nedbørsmengde i forhold til normalen for målestasjon Ås, januar-desember 2020

Det er observert overraskende store variasjoner i nedbørsmengde innad i studieområdet, både på døgn- og månedsnivå. Det ble konstanterert at nedbørsmengden var større enn normalen i hele studieområdet i begge de valgte tidsrommene.

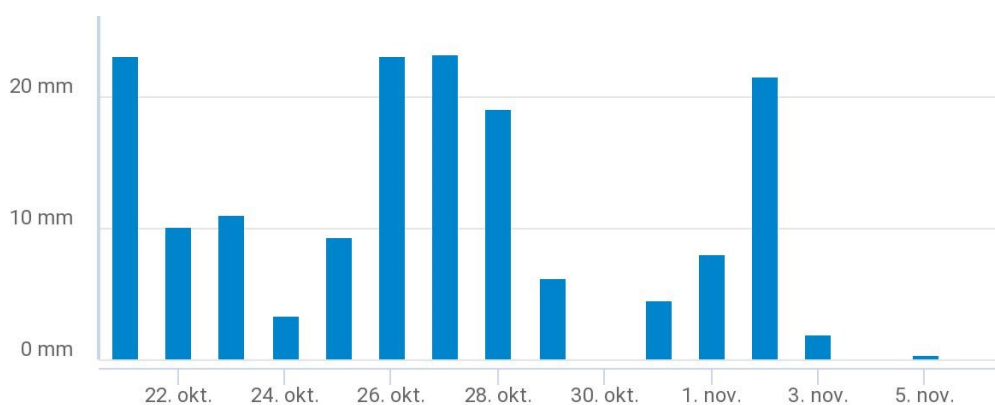


Figur 3.4: Nedbørsmengde i forhold til normalen for utvalgte målestasjoner i studieområdet, juli og oktober 2020

På grunnlag av nedbørsfordeling på døgnnivå innad i månedene, samt tilgang til Sentinel-1-data fra periodene, ble nøyaktig tidsrom for analysen definert til 15. juli-2. august og 21. oktober-6. november.

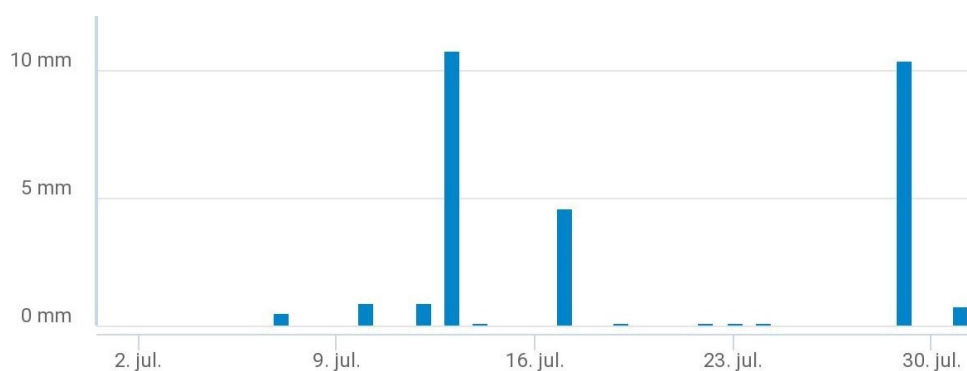


Figur 3.5: Nedbørsmengde per døgn innad i tidsrom 1, målestasjon Ås



Figur 3.6: Nedbørsmengde per døgn innad i tidsrom 2, målestasjon Ås

Sommeren 2018 var, i motsetning til samme periode i 2020, meget tørr i store deler av landet. På Ås var den totale nedbørsmengden i juli bare 29,4 mm, som tilsvarer 35,8% av normalen siden 1991. For å kunne sammenligne de våte periodene i 2020 med en uvanlig tørr periode, ble 26. juli 2018 valgt som sammenligningsgrunnlag.



Figur 3.7: Nedbørsmengde per døgn for målestasjon Ås, juli 2018

3.3 Filformater

3.3.1 TIFF

TIFF står for *Tagged Image File Format*, og er et filformat for beskrivelse og lagring av rasterbilder. TIFF er et lossless (tapsfritt) format, og støtter derfor svært høy bildekvalitet. [Adobe, 1992]

3.3.2 GeoTIFF

GeoTIFF er en utvidelse av TIFF-formatet med mulighet for georeferering av bilder. Kort sagt er en GeoTIFF-fil en TIFF-fil med geografisk metadata. GeoTIFF blir typisk brukt til lagring av fly- og satellittbilder, scannede kart og digitale høydemodeller, og støttes av de fleste GIS.

3.3.3 Shapefil

Shapefil-formatet, utviklet av ESRI, er et filformat for GIS-programvare. I en shapefil lagres ikke-topologisk geometri og attributtinformasjon for romlige objekter. Geometrien til objektene lagres som sett med vektorkoordinater, og representeres som punkter, linjer eller polygoner. [ESRI, 1998]

3.4 Datagrunnlag

3.4.1 Sentinel-1 GRD

GRD-produktene fra Sentinel-1 er bilder som er derivert fra rådata fra SAR-målingene. Nivå 1-GRD-produkter består av SAR-data som har blitt prosessert og projisert til bakkeplanet ved en ellipsoidmodell. [ESA, h]

Måledata fra Sentinel-opdragene er fritt tilgjengelige for offentligheten, og kan lastes ned gratis gjennom portaler fra ESA. Sentinel-1 data over Norge kan blant annet lastes ned gjennom nettsidene *Copernicus Open Access Hub* [Cop, a] eller *MET Norway - Copernicus Data Hub* [Cop, b]. Sistnevnte er en database med norske Sentinel-data som støtter ekstra rask tilgang til produktene, og det er denne som er brukt til datanedlasting i denne oppgaven. I nedlastingsportalen ble valgt studieområde og tidsrom definert i søkefunksjonen, og GRD ble valgt som produkttype. Plattform ble satt til Sentinel-1, som omfatter både Sentinel-1A og Sentinel-1B. Opptaksmodus og polarisering ble ikke spesifisert.

Totalt ble ni GRD-produkter lastet ned og brukt i oppgaven. Tabell 3.1 viser en oversikt over de ulike målingene. *Tidsrom 0* refererer til enkeltmålingen 26. juli 2018, *tidsrom 1* til perioden 15.juli-2.august 2020 og *tidsrom 2* til perioden 21. oktober-6. november 2020.

Måling	Måledato	Plattform	Modus	Orbit	Tidsrom
1	26.07.2018	S-1A	IW	ASC	0
2	15.07.2020	S-1A	IW	ASC	1
3	21.07.2020	S-1B	IW	ASC	1
4	27.07.2020	S-1A	IW	ASC	1
5	02.08.2020	S-1B	IW	ASC	1
6	19.10.2020	S-1A	IW	ASC	2
7	25.10.2020	S-1B	IW	ASC	2
8	31.10.2020	S-1A	IW	ASC	2
9	06.11.2020	S-1B	IW	ASC	2

Tabell 3.1: Sentinel-1 GRD-produktene som er brukt i oppgaven

3.4.2 FKB-AR5

AR5 står for *arealressurskart i målestokk 1:5000*, og er et nasjonalt heldekkende datasett som deler inn landareal etter arealtype, skogbonitet, treslag og grunnforhold. AR5 inngår i felles kartdatabase (FKB), og har NIBIO som fagansvarlig. [NIBIO, 2016]

I denne oppgaven blir AR5 brukt til å avgrense analysen kun til jordbruksareal. Jordbruksareal i AR5 omfatter arealtypene fulldyrka jord, overflatedyrka jord og innmarksbeite [NIBIO, 2016]. Det ble gjort en spørring mot AR5-databasen med de tre arealtypene og definert studieområde som input. Resultatet er en Shapefil med bestående av kun jordbruksområder innenfor studieområdet.

3.4.3 FKB-Vann

FKB-Vann er et landsdekkende datasett som beskriver geografisk beliggenhet, forløp og form til blant annet bekker, elver, grøfter, isbreer og innsjøer [SOSI, 2018]. FKB-Vann er, som FKB-AR5 og andre FKB-data, presentert på vektorform.

For bruk i klassifiseringsdelen i denne oppgaven, ble det fra FKB-Vann ekstrahert flater innenfor studieområdet med objekttype innsjø og størrelse mellom 500 og 100.000 m².

3.4.4 NDH-DTM

NDH står for *Nasjonal detaljert høydemodell*, og er det største landkartleggingsprosjektet som er gjennomført i Norge noensinne. Prosjektet har samlet inn høyde- og dybde data ved laserscanning, bildematching og sjømåling, som er gjort tilgjengelig for allmennheten som punkt skyer og høydemodeller. De digitale terrengmodellene (DTM) har på land en oppløsning på 1, 10 og 50 meter, og er lagret på GeoTIFF-format. [Kartverket]

I denne oppgaven benyttes NDH-DTM med oppløsning 10 meter til terrengkorrigerings av Sentinel-1-bilder.

3.5 Programvare

3.5.1 SNAP

The Sentinel Application Platform (SNAP) er ESAs programvare for prosessering, analysering og visualisering av data fra Sentinel-opdragene. SNAP er en samling av separate open source-moduler, *Sentinel-Toolboxes*, som behandler ulike typer data som radar- og satellittbilder. I tillegg til data fra Sentinel-sensorene støtter verktøyene også en rekke andre sensorer. [ESA, i]

SNAP og Sentinel-1 toolbox brukes i denne oppgaven til prosessering og analyse av GRD-produktene.

3.5.2 ArcGIS Pro

ArcGIS pro er en GIS-programvare utviklet av ESRI for å utforske, visualisere og analysere geografiske data i ulike dimensjoner og formater [ESRI].

ArcGIS PRO brukes i denne oppgaven til visualisering av resultater fra analysen.

3.5.3 QGIS

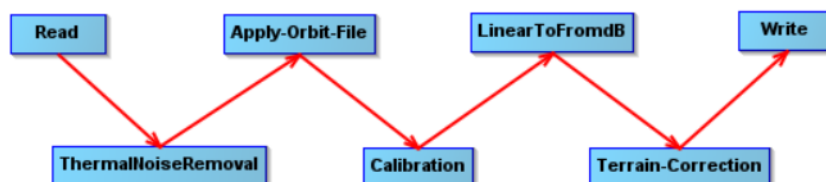
QGIS er en open source GIS-programvare for oppretting, visualisering, behandling, analysering og publisering av geografiske data. QGIS støtter et stort antall raster- og vektorformater. [QGIS, a]

QGIS brukes i denne oppgaven til visualisering av resultater fra analysen og etablering av treningsdata for klassifisering.

4 Metode

4.1 Preprosessering av Sentinel-1-data

GRD-produktene er allerede delvis prosessert ved nedlasting, men har behov for ytterligere forbedring før de tas i bruk i analysen. Preprosesseringen utføres i SNAP, og fremgangsmåten er basert på brukerveiledningene til Braun [2021] og Filipponi [2019]. Videre forklares de ulike stegene i preprosesseringen.



Figur 4.1: Preprosesseringsløype satt opp i *Graph builder* i SNAP.

Thermal noise removal

Intensiteten i SAR-bildene er preget av forstyrrende termisk støy, særlig i de krysspolariserte båndene. I dette steget normaliseres tilbakespredningskoeffisienten, og den termiske støyen fjernes eller reduseres. [Filipponi, 2019]

Apply orbit file

Satellittbanevektorene som ligger i SAR-produktenes metadata er ikke fullstendig presise. En stund etter opptakstidpunktet, typisk et par dager, vil nøyaktig informasjon om satellittens posisjon og hastighet være tilgjengelig i en egen fil kalt *orbit file*. I dette steget blir riktig orbit file automatisk hentet inn, og SAR-produktets metadata blir oppdatert med informasjonen fra denne filen. [Filipponi, 2019]

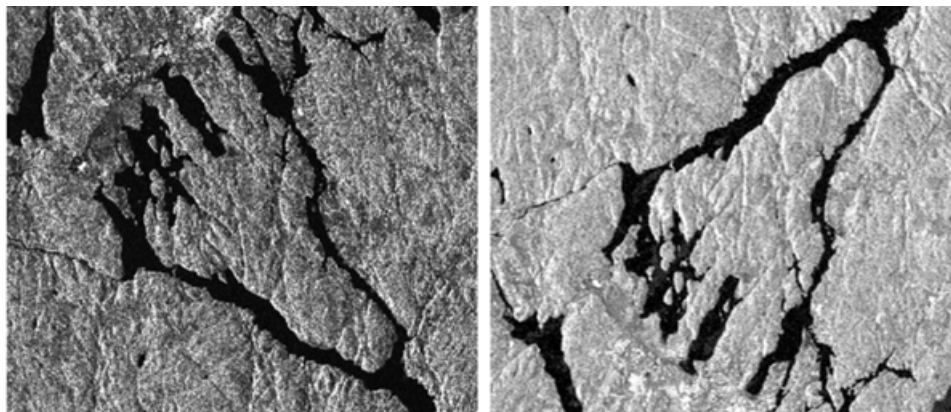
Calibration

I dette steget konverteres digitale pikselverdier, som representerer bildeintensitetsverdi, til radiometrisk kalibrert SAR-tilbakespredning i form av sigma nought-verdier. Koeffisienter som blir benyttet i kalibreringen, som insidensvinkel og polarisering, blir hentet fra produktets metadata. [Filipponi, 2019]

Convert bands to dB Tilbakespredningsintensiteten har til nå opptrådt som en koeffisient uten enhet. I dette steget blir koeffisienten ved en logaritmisk transformasjon konvertert til desibel (dB). [Filipponi, 2019]

Terrain correction

På grunn av topologi vil GRD-produktene inneholder en del geometriske forvrengninger. I dette steget brukes *Range doppler terrain correction* til å korrigere pikslenes posisjon basert på en automatisk innlastet eller ekstern digital høydemodell. [Filipponi, 2019] Her blir NDH-DTM 10m lastet inn. Fordi valgt DTM er klippet etter utstrekningen til studieområdet, blir bildene i dette steget automatisk også klippet til riktig utstrekning, og det er ikke behov for Subset-steget som er inkludert i fremgangsmåten til Braun.



Figur 4.2: Utsnitt fra originalt GRD-produkt (til venstre) og preprosessert bilde (til høyre)

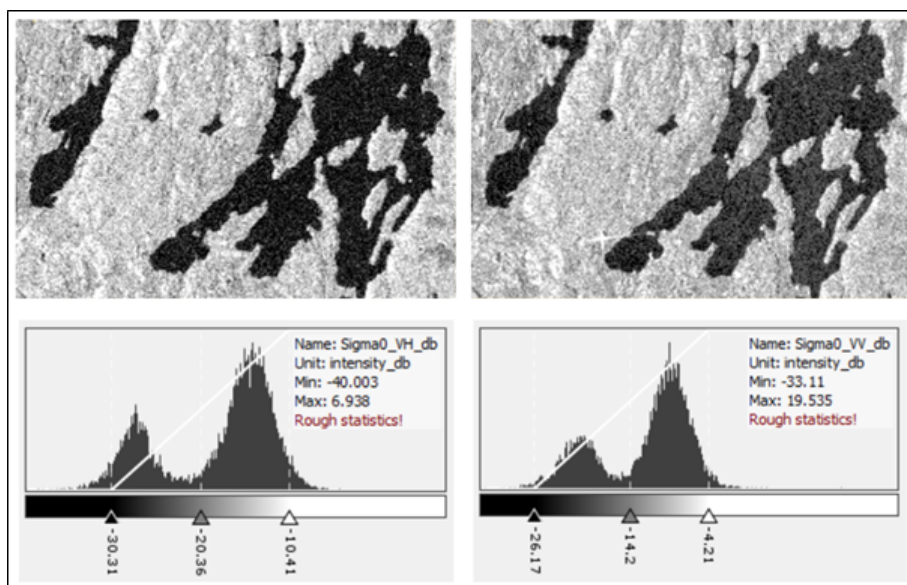
Det ble vurdert å tilføye *Speckle filtering* som et steg i preprosesseringen. På grunn av at områdene som ønskes å dekkes kan være så små som enkeltpikslers, kan en slik filtrering føre til tap av viktig informasjon, og det ble utelatt.

4.2 Monotemporær analyse av Sentinel-1-data

I den monotemporære analysen brukes pikselverdien i enkeltbilder til å detektere mulige oversvømte områder. Det ble gjennomført to monotemporære analyser, en for hver av tidsrommene. Fra tidsrom 1 ble 27. juli valgt som dato for analysen, og fra tidsrom 2 ble 31. oktober valgt. Metodene i kapittel 4.2.1 og 4.2.2 blir altså utført på opptakene fra begge disse datoene. Fremgangsmåten er vist på bildet fra oktober.

4.2.1 Terskling

For å bestemme en passende terskelverdi ble histogrammene til SAR-bildet studert i begge polariseringskanaler. Ved å undersøke et lite utsnitt av bildet hvor antall pikslers vann og ikke-vann er rimelig likt fordelt, vil man lettere se sammenhengen mellom bilde og histogram. Figur 4.3 viser utsnitt av bildet som ble undersøkt, med tilhørende histogrammer.



Figur 4.3: Utsnitt polarisert i VH (til venstre) og VH (til høyre), med tilhørende histogrammer

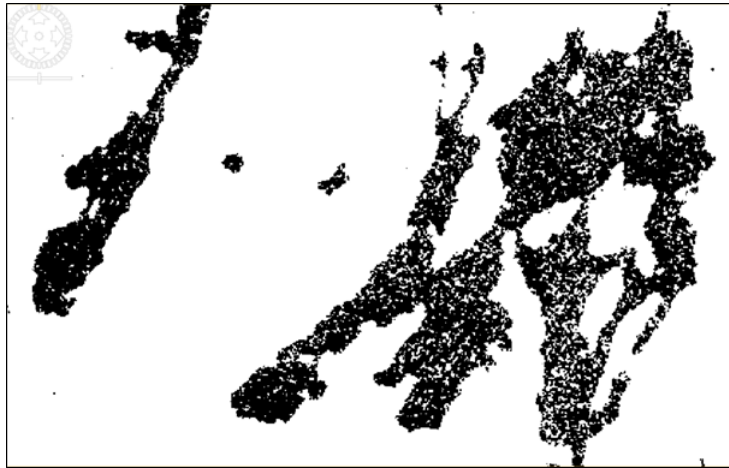
De to trekantformasjonene i histogrammene kan sees på som to klasser av piksler; vann og ikke vann. Ved å sette en terskelverdi i lokalt minimumspunkt, altså i bunnen mellom de to toppene, vil man få et binært bilde hvor de store vannene er tydelig avgrenset og kun består av svarte piksler. Dette tilsvarer terskelverdi -22 for VH og -17 for VV. Tersklingen i SNAP utføres i kalkulatorverktøyet *Band Maths*, hvor følgende uttrykk ble angitt:

if Sigma0_VH_db > -22 and Sigma0_VV_db > -17 then 1 else 0



Figur 4.4: Binærisert bilde med terskelverdier -22 (VH) og -17 (VV)

Fordi vannene i SAR-bildene fra figur 4.2 også består av en del lysere piksler, vil mange av ikke-vann-områdene også bli kategorisert som vann ved å sette terskelverdien på denne måten. Ytterligere studering av det binære bildet tilsier at det er lite sannsynlig at alle disse pikslene er oversvømte områder. Det bør derfor settes en lavere terskelverdi, slik at færre piksler blir kategorisert som vann. Terskelverdien ble satt i det venstre toppunktet i histogrammet, som vil si -28 for VH og -20 for VV. Som man ser i figur 4.5 vil man ved å velge en lavere terskelverdi gå glipp av piksler som faktisk er vann, men det er større sannsynlighet for at pikslene som har blitt kategorisert som vann faktisk er det.



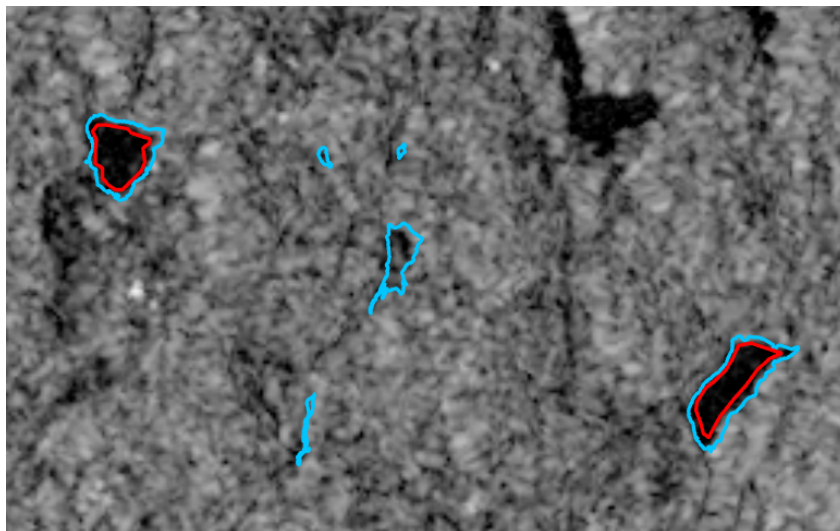
Figur 4.5: Binærisert bilde med terskelverdier -28 (VH) og -20 (VV)

Siden det kun er jordbruksareal som skal studeres videre, må andre arealtyper maskeres ut fra de binære bildene. Til dette brukes verktøyet *Land/Sea Mask* i SNAP, hvor AR5-datalaget med jordbruksområder blir brukt som maske. Områder utenfor masken vil få verdi *nodata*.

4.2.2 Klassifisering

Den ble utført en styrt klassifisering med Random forest-algoritmen i SNAP. Klassifiseringen skal dele området inn i to klasser; vann og ikke vann.

Som treningsdata for vann-klassen ble det definert 80 polygoner med utgangspunkt i FKB-vann datasettet med innsjø mellom 500 og 100.000 m². Tanken med denne størrelsesbegrensningen er at små innsjøer vil opptre mer likt som overvann på jordbruksareal enn det større innsjøer gjør. Det ble forsøkt å velge mørkest mulig piksler i vannet som treningsdata. Som treningsdata for ikke-vann ble FKB-AR5-datalaget benyttet.



Figur 4.6: Eksempel på treningsdata for vann (rød) definert innenfor FKB-Vann innsjø (blå)

Variabler som kan justeres i RFC-algoritmen i SNAP er *Number of training samples* og *Number of trees*. Førstnevnte vil si antall piksler innenfor treningspolygonene som blir brukt til trening av modellen, som ble satt til 5000. Antall trær ble satt til 25.

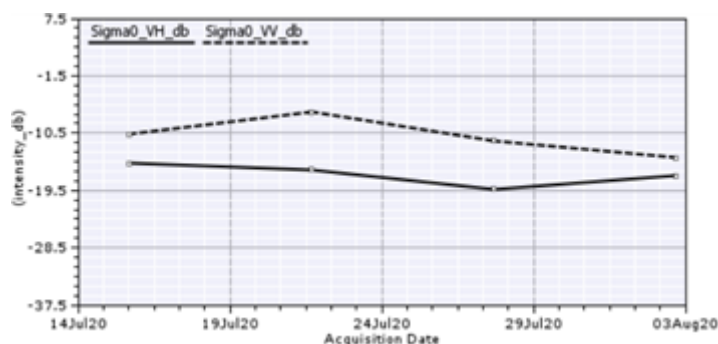
4.3 Multitemporær analyse av Sentinel-1-data

I den multitemporære delen av analysen brukes endring i pikselverdi mellom to eller flere tidspunkter til å se på sammenhengen mellom pikselverdi og nedbørperiode.

4.3.1 Tidsserieanalyse

Verktøyet *Time series* i SNAP brukes til å visualisere pikselverdier over en gitt periode i et linjeplott, og kan dermed brukes til å visualisere variasjonen av tilbakespredningsintensiteten til overflater over tid [Braun, 2021].

Det ble utført tidsserieanalyser innenfor hver av tidsrommene, samt en sammenligning av tidsrommene. Opptak som skulle inngå i tidsserien ble lagt til i analyseverktøyet. Ved å bevege musepekeren eller klikke på bestemte piksler vil man få opp linjeplottet til respektive piksel, med plott for hver av polariseringskanalene.



Figur 4.7: Eksempel på linjeplott som viser utviklingen av tilbakespredning over tid, i en tilfeldig valgt enkelt piksel.

Hovedsakelig studeres signaturen til piksler og områder som har blitt definert som oversvømte områder tidligere i analysen, men det undersøkes også om det er en trend i signaturen til andre piksler. Basert på framgangsmåten til Braun [Braun, 2021] blir tidsseriene studert ved sammenligning av tidsseriene til enkelt piksler.

4.3.2 Rasterkalkulering

Verktøyet *Raster Calculator* i QGIS gjør det mulig å utføre matematiske beregninger på grunnlag av eksisterende pikselverdier i valgte raster. Resultatene fra kalkuleringen skrives til et nytt rasterdatalag. [QGIS, b]

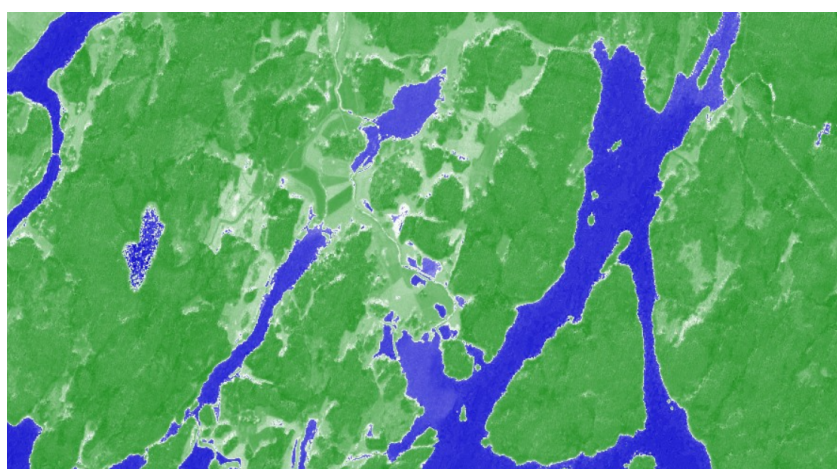
En rasterkalkulering vil kunne vise forskjellen i pikselverdi i bilder fra to tidspunkter. Det er her ønskelig å studere forskjellen mellom en veldig tørr og en veldig våt periode. Opptak fra samme tid på året vil gi best mulig sammenligningsgrunnlag, derfor benyttes opptak fra 26. juli 2018 og 27. juli 2020.

I rasterkalkulatoren subtraheres 2020-rasteret fra 2018-rasteret. I resultat-rasteret vil pikslene med høyest verdi være pikslene hvor tilbakespredningsintensiteten har størst negativ endring mellom 2018 og 2020. Det undersøkes om oversvømte områder kan oppdages i et slikt differanse-datalag.

4.4 NDWI fra Sentinel-2

De fleste Sentinel-2 opptak fra begge tidsrommene er preget av kraftig skydekke i hele studieområdet, men fra 3. november 2020 finnes det faktisk opptak som er tilnærmet skyfrie i store deler av studieområdet. Dermed kan data fra Sentinel-2 brukes som en kilde for sammenligning med resultatene fra Sentinel-1.

Prosessert Sentinel-2-data ble hentet fra nettsiden *Sentinel Hub EO Browser* [ESA, j], en portal hvor man kan studere og laste ned satellittdata fra ulike sensorer. Her ble det valgt datakilde Sentinel-2, dato 3. november 2020, og det ble navigert til relevant område. Bruker kan velge mellom gitte eller egendefinerte båndkombinasjoner, ettersom hvordan man ønsker å visualisere opptakene. Her ble NDWI valgt. I studien til Clement et al. [Clement et al., 2017] benyttes MNDWI til verifikasjon av resultatene. MNDWI er en modifisert versjon av NDWI som redserer effekten som gir bygninger en høy indeks. Fordi bebygde områder ikke inngår i analysen, blir vanlig NDWI benyttet. Høy indeks representeres ved blå områder i rasteret, middels høy ved lyse, og lav indeks ved grønne.



Figur 4.8: Utsnitt fra NDWI

For å kunne benytte Sentinel-2 NDWI til en visuell sammenligning med resultatene fra Sentinel-1 ble valgt område og båndkombinasjon lastet ned som en geoTIFF-fil. Det ble lastet opp en vektor av studieområdets utstrekning for å avgrense området.

5 Resultater

5.1 Terskling

Resultatet av tersklingen indikerer mange små oversvømte områder på på størrelse med én eller noen få piksler (figur 5.1), altså ned mot en størrelse på 10 x 10 meter. Resultatene viser også større sammenhengende områder med antatt oversvømmelse (figur 5.2) og områder hvor enkeltpiksler og små gupper av piksler opptrer hyppigere. Det finnes indikasjoner på forekomster av overvann spredt utover hele studieområdet.

For bedre visualisering og lettere forståelse av størrelse og område er pikslene som er definert som vann ekstrahert og lagt over *Norge i Bilder WMS Ortofoto*.



Figur 5.1: Eksempel på resultater fra terskling i opptak fra oktober



Figur 5.2: Eksempel på resultater fra terskling i opptak fra oktober

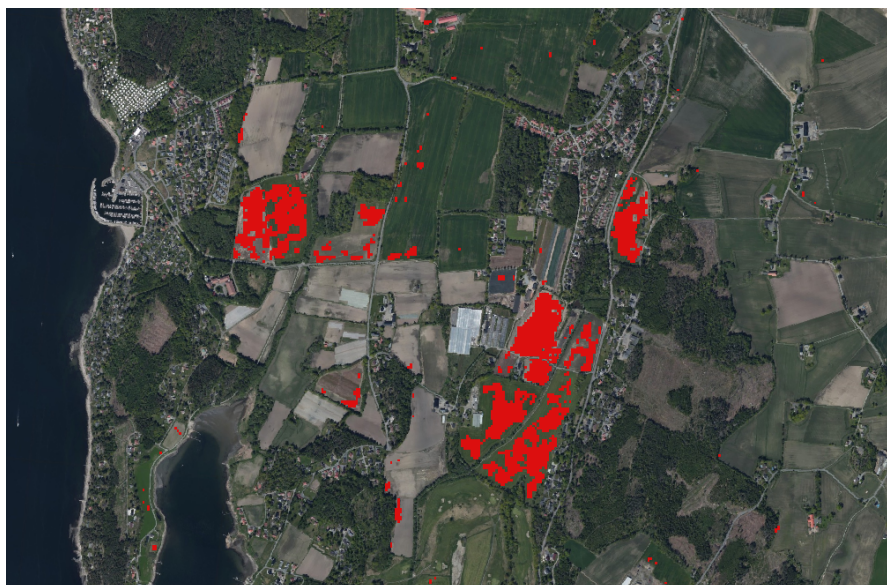
5.2 Klassifisering

Klassifiseringsrapportene fra de to random forest-klassifiseringene finnes i appendix A1 og A2. I rapportene finnes verdier for blant annet *Root Mean Squared Errors (RMSE)*, som er et mål på den gjennomsnittlige prediksjonsfeilen til modellen. I klassifiseringene fra juli og oktober viser rapporten RMSE på henholdsvis 0.19 og 0.05.

Som for tesklingen indikerer resultatene fra klassifiseringen mange små oversvømte områder på størrelse med én eller noen få piksler (figur 5.3), men også større sammenhengende områder (figur 5.4). Forekomstene er spredt utover studieområdet, og visualiseres på samme måte som tersklingen.



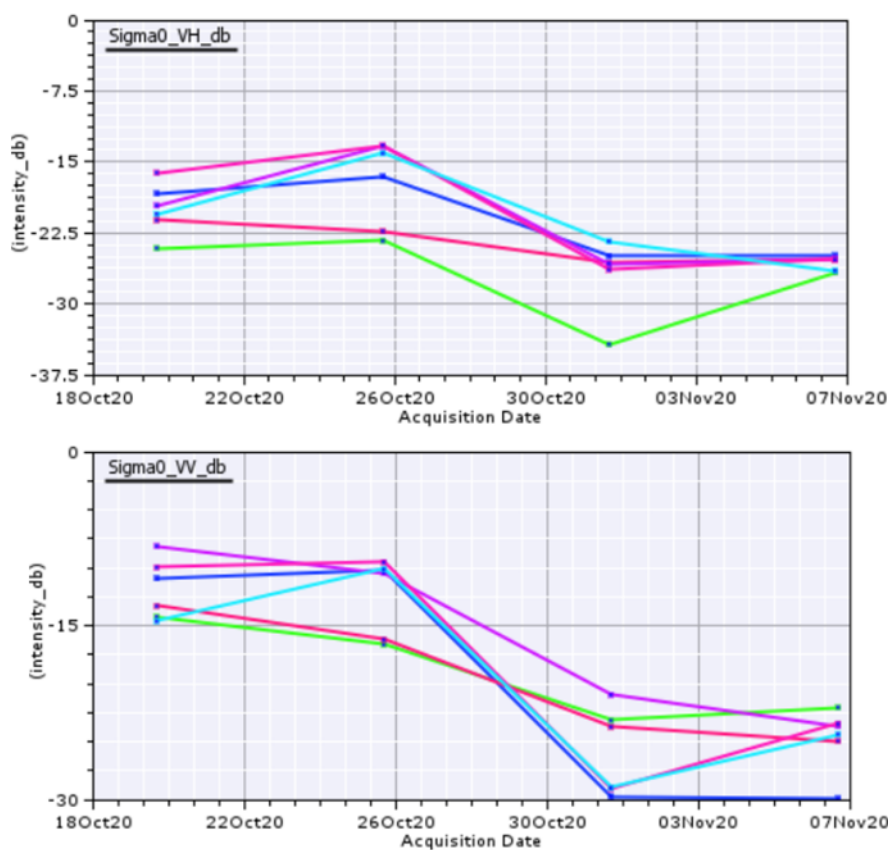
Figur 5.3: Eksempel på resultat fra klassifisering i opptak fra oktober



Figur 5.4: Eksempel på resultat fra klassifisering i opptak fra oktober

5.3 Tidsserieanalyse

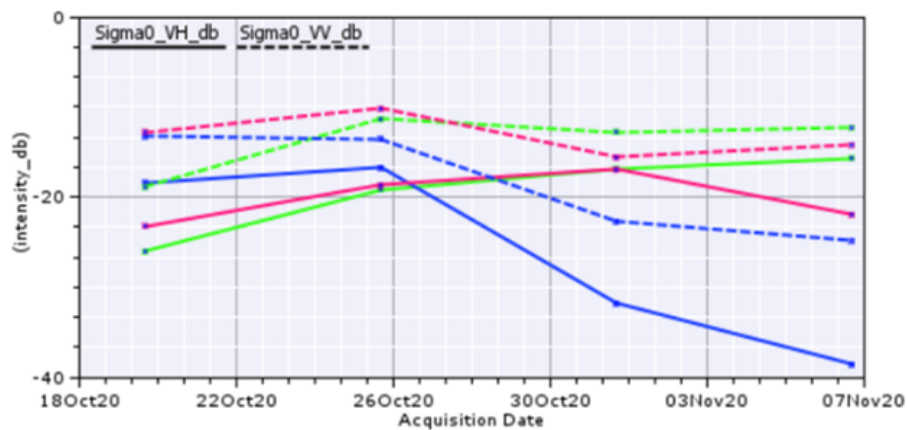
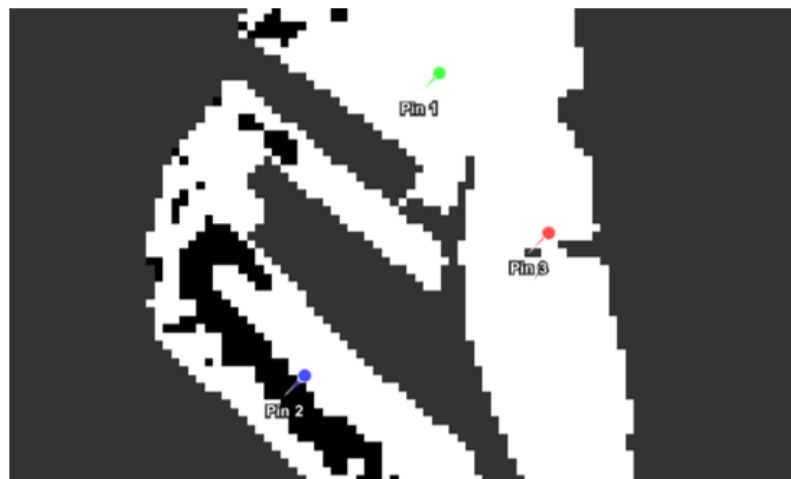
Figur 5.5 viser linjeplottene i VH- og VV-polarisering til seks valgte piksler fra seks forskjellige områder som antas å være oversvømte, i tidsrom 2. Plottene viser en synkende tendens i utviklingen av tilbakespredningsintensitet innad i tidsrommet. Tendensen er et stort fall i kurven mellom opptak 2 og 3. Mellom opptak 1 og 3 er kurven flatere, for noen av pikslene også stigede, og mellom opptak 3 og 4 er kurven svakt stigende eller flatere.



Figur 5.5: Tidsserieplottene til seks piksler som har blitt klassifisert som vann i ulike områder, for VH- og VV-polarisering

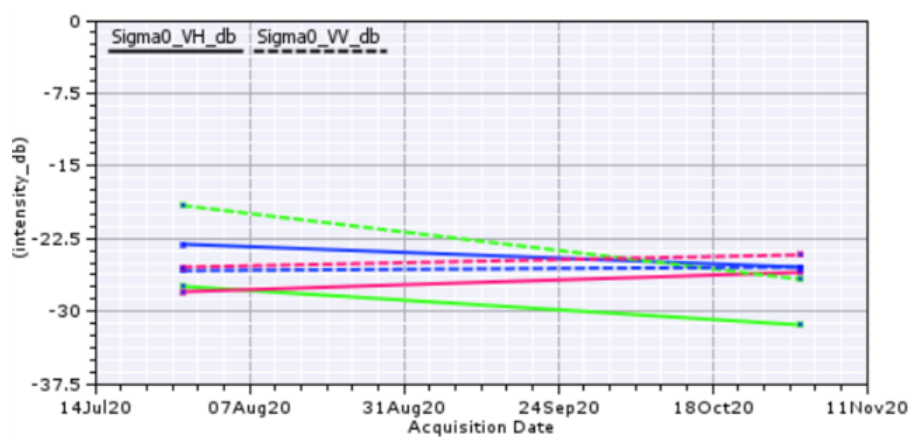
I figur 5.7 presenteres tidsserieplottene til tre typer piksler, markert med nåler i figuren. Nål 1 representerer piksler som ikke har blitt klassifisert som vann. Nål 2 representerer piksler i et lite område som har blitt klassifisert som vann, her på størrelse med to piksler. Nål 3 representerer piksler som har blitt klassifisert som vann og som er en del av en større vannforekomst.

I piksler som er en del av et større antatt oversvømt område synker intensitetsverdier over tid innad i tidsrommet. Dette er gjentakende i flertallet av piksler i områdene som blir undersøkt. I piksler fra områder som ikke er oversvømte har tidsserieplottene ofte en slak stiging eller er ganske flate. I enkeltpiksler som har blitt klassifisert som vann er utviklingen mer vilkårlig, og det er vanskelig å se en trend.



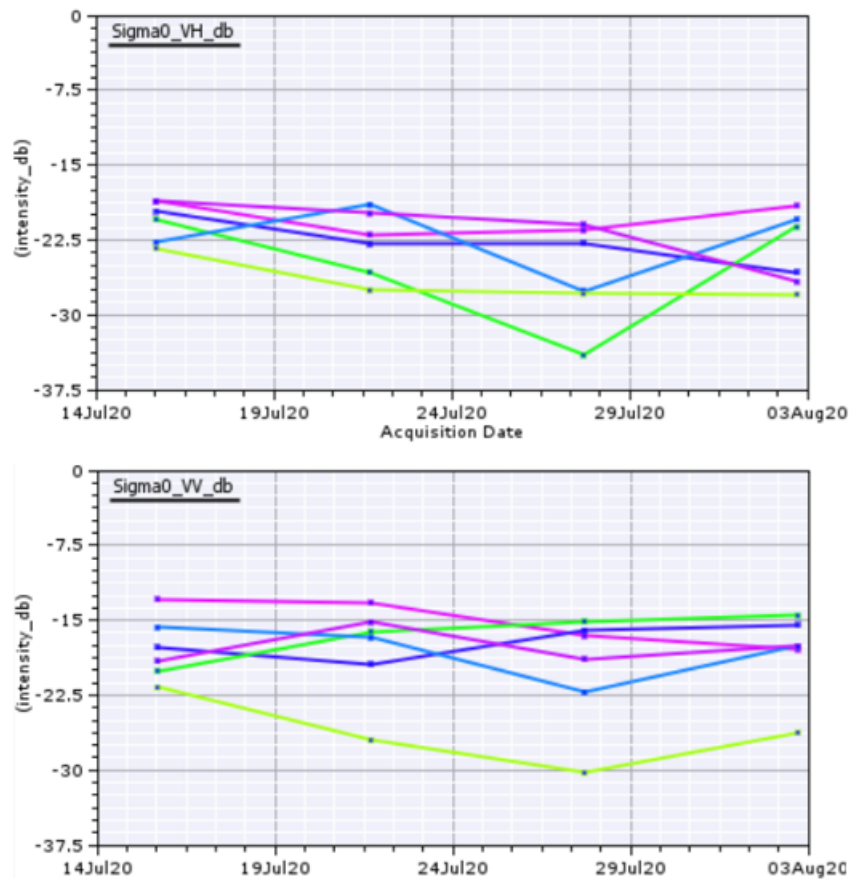
Figur 5.6: Typisk tidsserieplott for tre typer piksler

Det ble også sammenlignet intensitetsverdier i piksler som har blitt klassifisert som oversvømte både i bildet fra juli og oktober. Pikslene har som regel relativt lik verdi, men tendensen for mange av pikslene er en noe lavere verdi i oktober.



Figur 5.7: Utviklingen i intensitetsverdi mellom de to tidsrommene for piksler klassifisert som vann

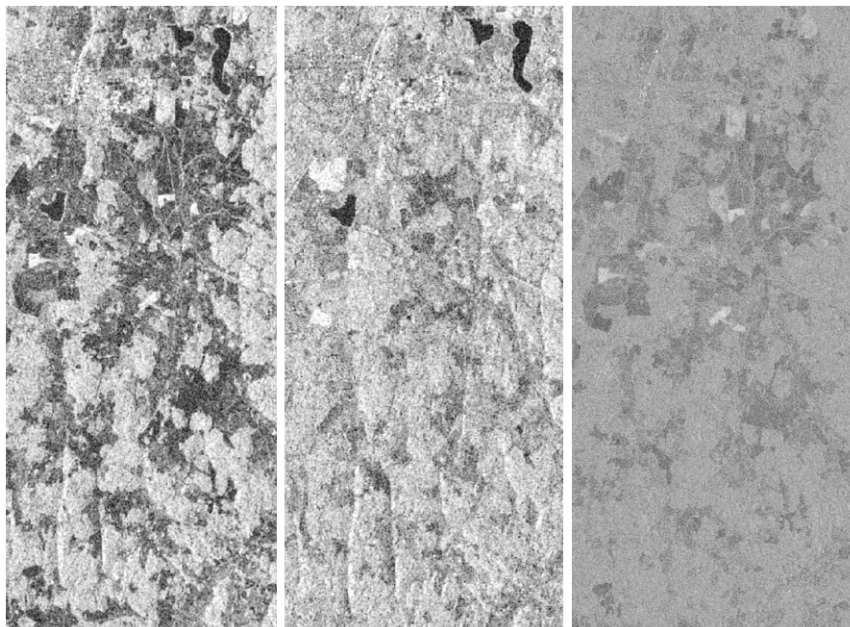
Linjeplottene som viser utviklingen innad i tidrom 1, for piksler klassifisert som vann i bildet fra juli, viser en tendens til synkende verdier utover i perioden. Tendensen er dog ikke veldig markant.



Figur 5.8: Tidsserieplottene til seks piksler som har blitt klassifisert som vann i ulike områder, for VH- og VV-polarisering

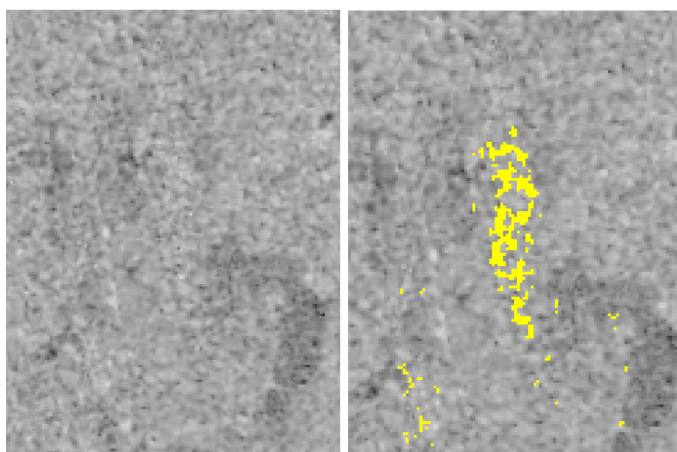
5.4 Rasterkalkulering

Resultatet fra rasterkalkuleringen viser at det har skjedd endringer i mange pikselverdier mellom bildene fra juli 2018 og juli 2020.



Figur 5.9: Fra venstre til høyre: Bilde fra 2018, bilde fra 2020 og resultatraster

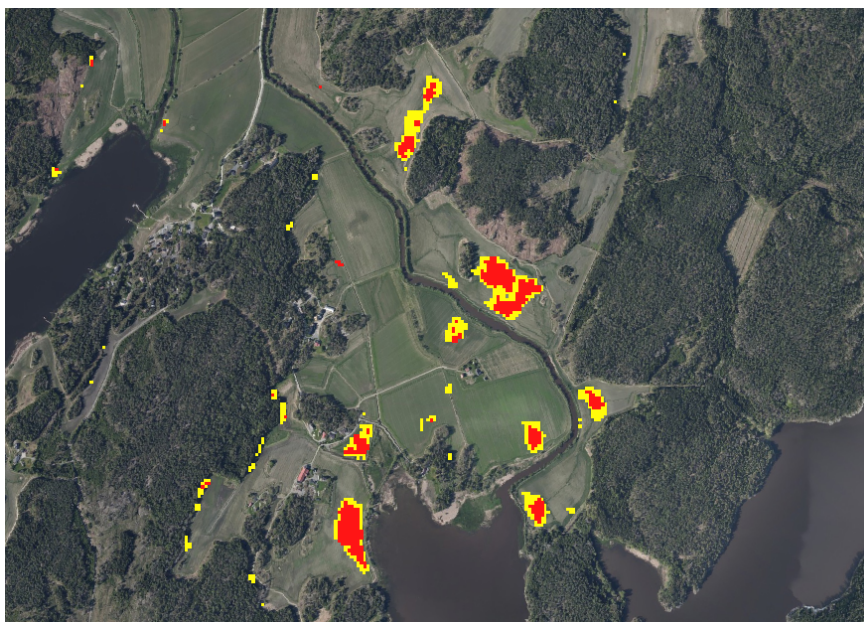
Det viser seg imidlertid at det er vanskelig å bruke endringsrasteret til å detekere områder som kan være oversvømte, basert på sammenligning med områder som ble klassifisert som vann i bildet fra juli 2020.



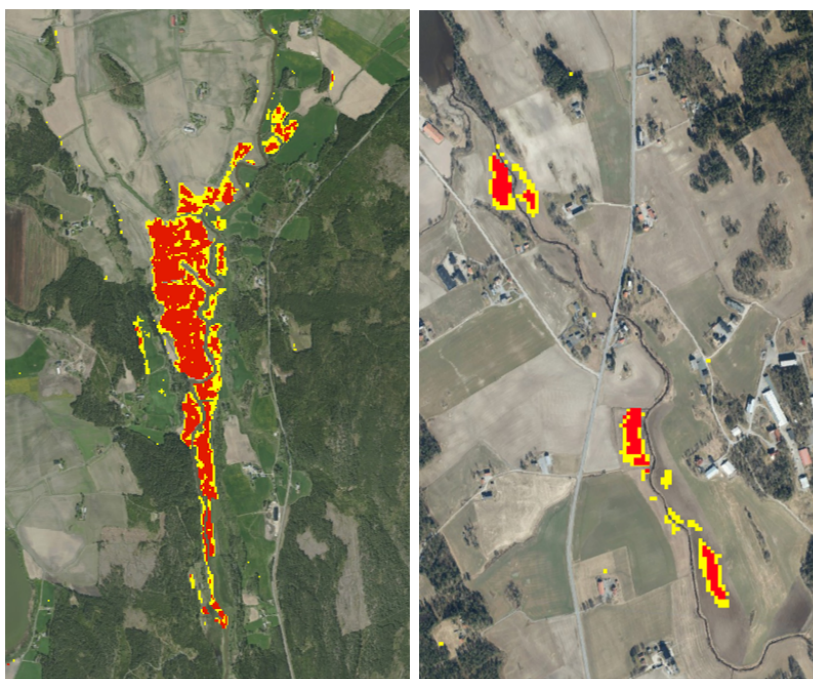
Figur 5.10: Ekstempel på resultat fra rasterkalkulering visualisert sammen med resultat fra klassifiseringen i bildet fra juli

5.5 Sammenligning av resultater

Resultatene viser en klar sammenheng mellom områder hvor det indikeres forekomster av overvann ved metodene terskling og klassifisering i bildet fra oktober. I figur 5.11 og 5.12 visualiseres eksempler på resultater fra de to teknikkene sammen i bildene. Tendensen er at større områder og fler enkeltpixels blir klassifisert som vann ved klassifisering enn ved terskling. I figurene er det klassifisering og terskling av bildet fra oktober som presenteres.



Figur 5.11: Sammenligning av resultat fra klassifisering (gul) og terskling (rød) for opptak fra oktober



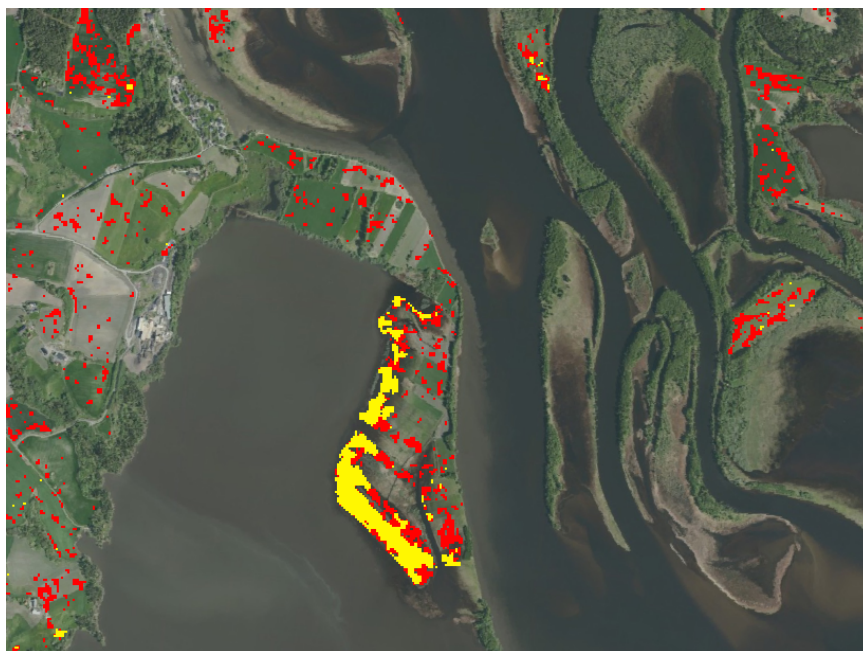
Figur 5.12: Sammenligning av resultat fra klassifisering (gul) og terskling (rød) for opptak fra oktober

De minste områdene som blir detektert som overvann viser seg å ikke sammenfalle ikke i lik grad som de store områdene, vist i figur 5.13.

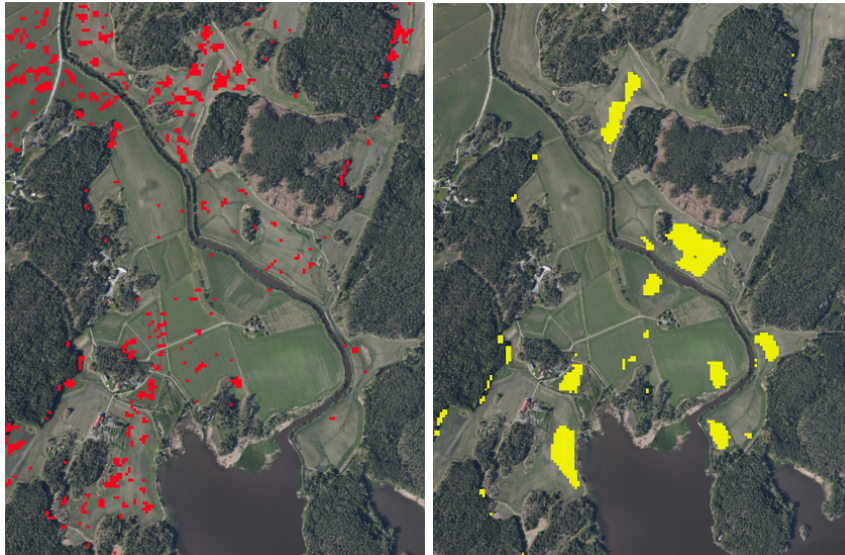


Figur 5.13: Eksempel på resultat fra klassifisering (gul) og terskling (rød)

Det ble benyttet samme metode for terskling og klassifisering på bildene fra 27. juli og 31. oktober. Likevel blir resultatene rimelig ulike for de to tidspunktene, vist i figur 5.14 og 5.15. Resultatene fra klassifiseringen i bildet fra oktober gir flere sammenhengende, større områder, mens resultatene i bildet fra juli er preget av tettere forekomster av mindre områder.

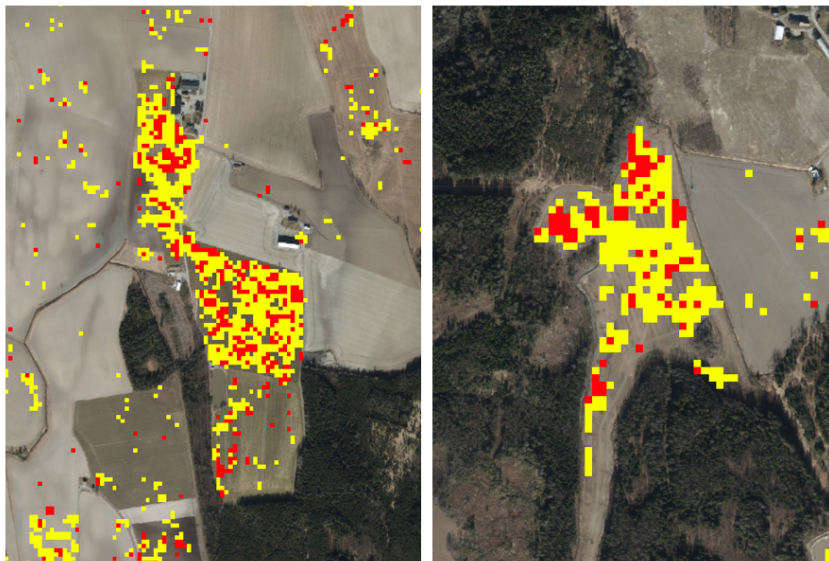


Figur 5.14: Sammenligning av resultat fra klassifisering i opptak fra juli (rød) og oktober (gul)



Figur 5.15: Eksempler på resultat fra klassifisering i opptak fra juli (rød) og oktober (gul)

Det er også til dels sammenfallende resultater mellom tersklingen og klassifiseringen i bildet fra juli, men som vist i figur 5.15 og 5.16 er resultatene i dette bildet mye mer usammenhengende enn i bildet fra oktober.



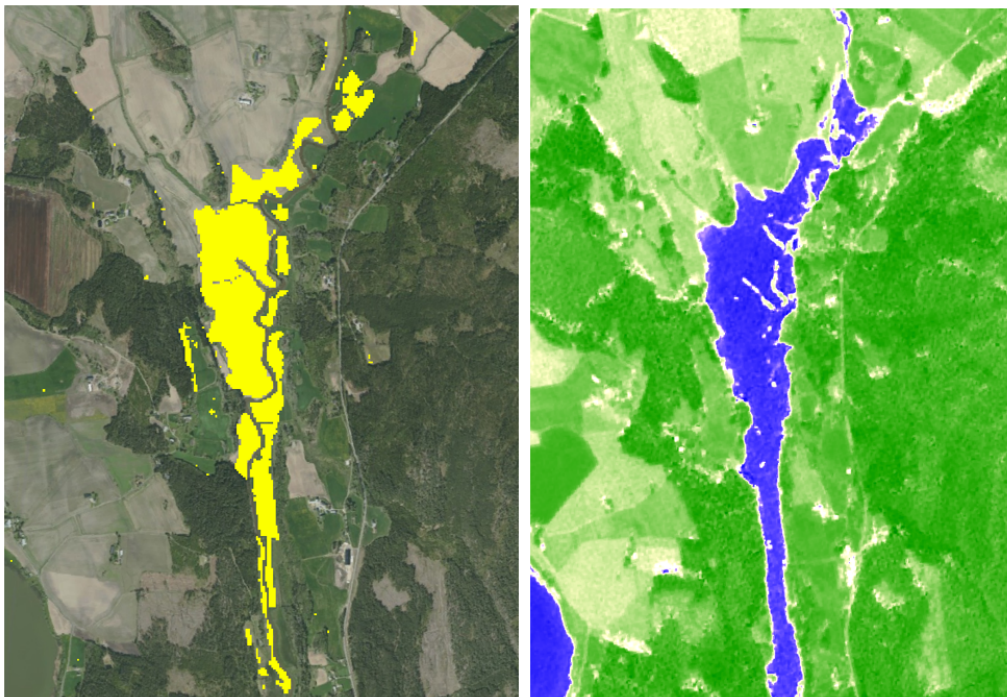
Figur 5.16: Sammenligning av resultat fra klassifisering (gul) og terskling (rød) av bilde fra juli

Tabell 5.1 gir en oversikt over prosentandelen av det totale jordbruksarealet i studieområdet som i analysen ble klassifisert som oversvømt, regnet ut fra pikselfordeling i det binære bildet. Tallene viser at Random forest-klassifiseringen i begge tilfeller klassifiserte flere piksler som vann enn tersklingen gjorde, og at opptaket fra juli indikerte mer oversvømt areal enn opptaket fra oktober.

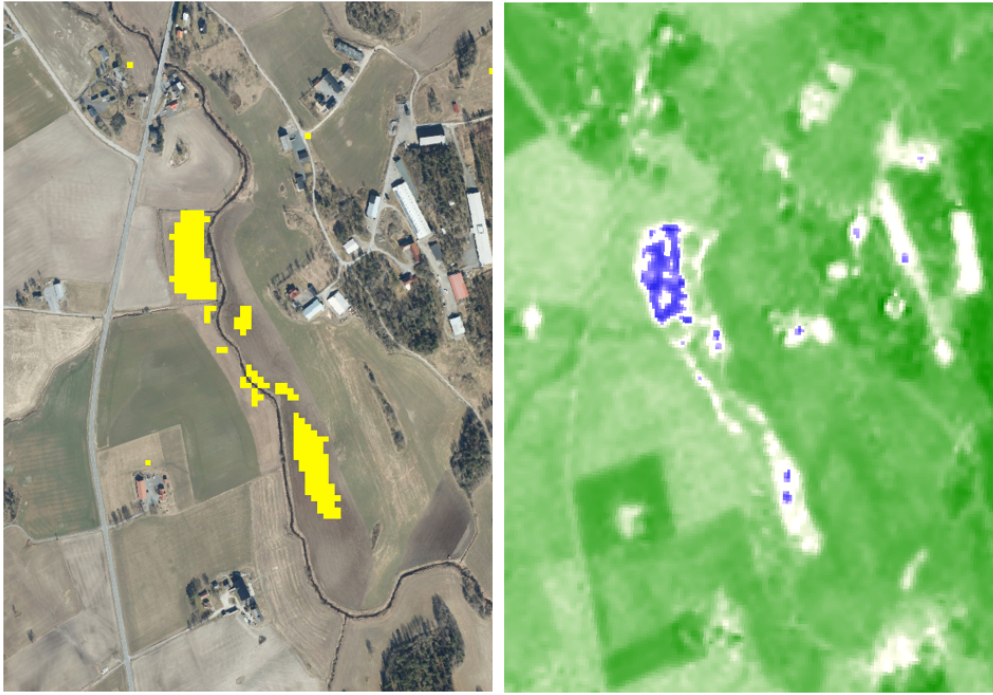
Dato	Metode	Prosent av areal oversvømt
27. juli	Terskling	0.89
27. juli	Klassifisering	3.56
31. oktober	Terskling	0.14
31. oktober	Klassifisering	1.48

Tabell 5.1: Prosent oversvømt jordbruksareal detektert i bildene med de to metodene

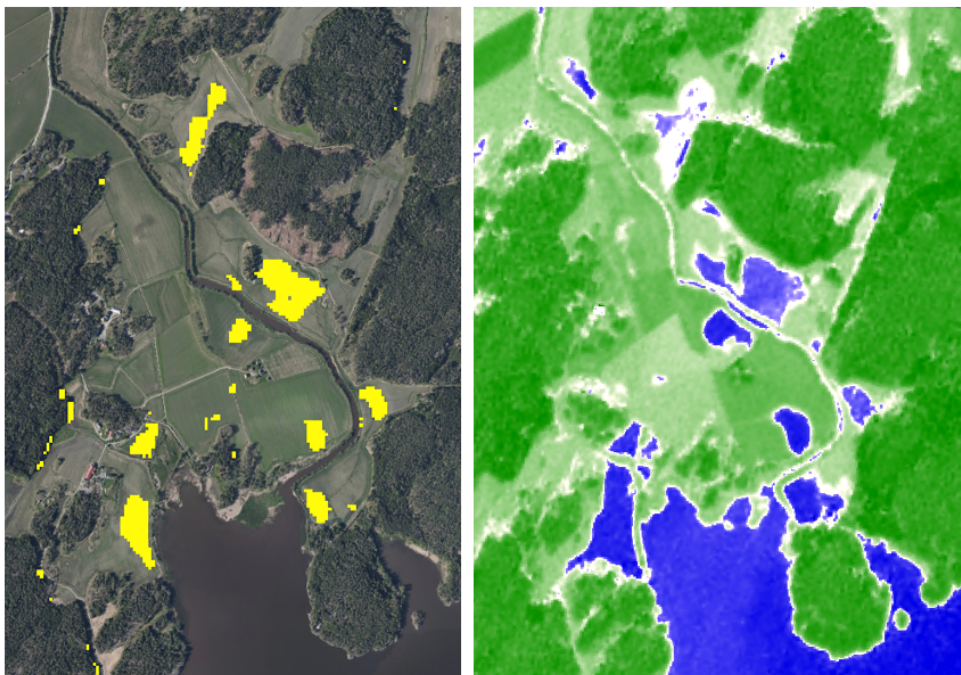
Det ble valgt å benytte resultatene fra klassifiseringen på opptaket fra oktober for å sammenligne med NDWI. Resultater fra tre utvalgte områder vist i figur 5.17, 5.18 og 5.19 viser en tydelig korrelasjon mellom detektere oversvømte områder og områder med høy indeks i NDWI. Merk at NDWI i figurene ikke er klippet etter jordbruksareal-masken, og vil også vise noen forekomster av vann som ikke ligger i jordbruksområder.



Figur 5.17: Sammenligning av resultat fra klassifisering (venstre) og NDWI (høyre)



Figur 5.18: Sammenligning av resultat fra klassifisering (venstre) og NDWI (høyre)



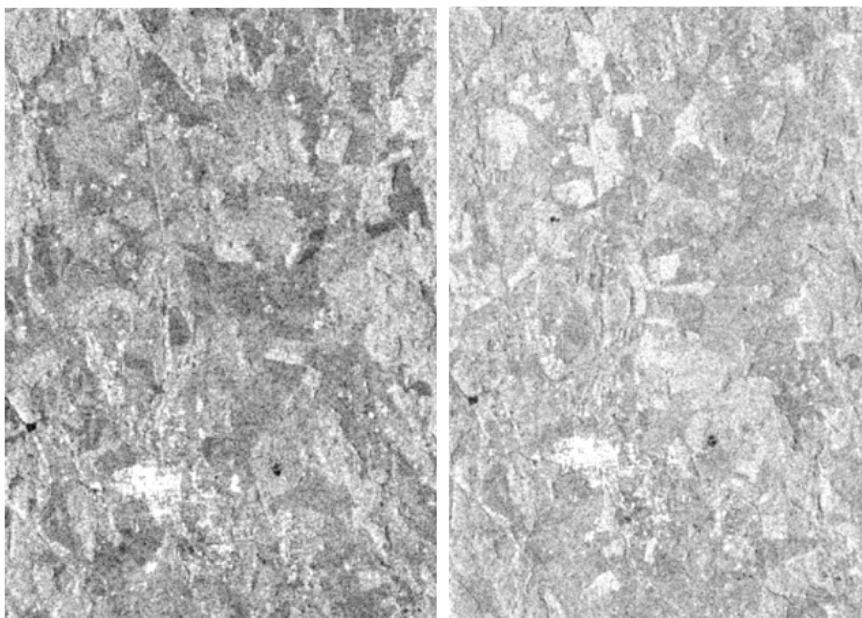
Figur 5.19: Sammenligning av resultat fra klassifisering (venstre) og NDWI (høyre)

6 Diskusjon

6.1 Resultater

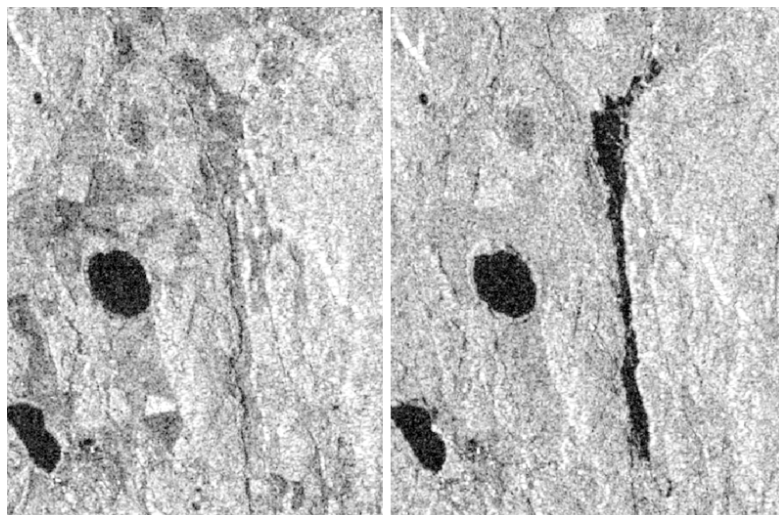
Resultatene fra klassifiseringen og tersklingen utført på samme opptak viser sammenfallende forekomster av potensielle oversvømmelser, i varierende grad. Resultatene av den monotemporære analysen av bildet fra oktober gir indikasjoner på oversvømte områder som er tydelig avgrenset, og mange av indikasjonene virker realistiske. Sammenligningen med NDWI styrker også antagelsene om at dette er noenlunde troverdige resultater. Usikkerheten er større for områder på størrelse med enkelt piksler, da dette også kan være resultat av støy i bildet. Resultatene av den monotemporære analysen av bildet fra juli sammenfaller i mindre grad enn i analysen av bildet fra oktober, og pikslene har en større mye spredning og virker urealistiske som vannforekomster.

Det er ikke forventet at resultatene fra analysen skal bli helt like for opptakene fra juli og oktober, men det er betenkelig at resultatene blir såpass ulike for de to tidsperiodene. Noen svar på hva dette kan skyldes kan kanskje finnes ved en visuell inspeksjon av GRD-produktene, i figur 6.1 representert ved VV-båndet. Bildene viser at opptaket fra juli har en større forekomst av mørke områder enn opptaket fra oktober. Dette kan gjøre det mer krevende å skille mellom vannforekomster og jordtyper som fremstår som mørke i bildet.

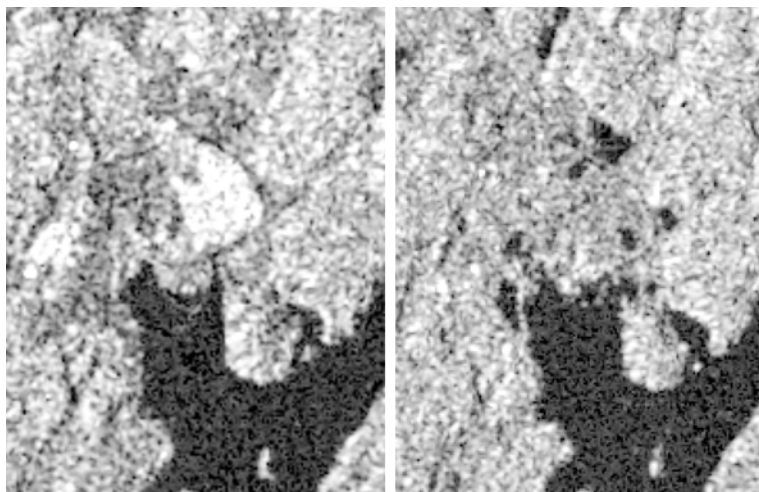


Figur 6.1: GRD-produkt fra 27. juli (venstre) og 31. oktober (høyre))

Dessuten viser bildene at flere av områdene som var oversvømte i oktober rett og slett ikke var oversvømte i juli. Dette gjelder blant annet for området i figur 6.3, som ble brukt til å sammenligne resultater fra klassifiseringene av opptak fra juli og oktober i figur 5.15.



Figur 6.2: GRD-produkt fra juli (til venstre) og oktober (til høyre)



Figur 6.3: GRD-produkt fra juli (til venstre) og oktober (til høyre)

Opptaket fra juli 2018, som ble brukt i rasterkalkuleringen og vises i figur 5.9, har en enda større grad av mørke områder enn opptaket fra juli 2020. Dette kan til dels forklares ved teorien fra kapittel 2.3.5, om at tørr jord kan fremstå mørkere enn våt jord i SAR-bildene på grunn av ulik penetrasjonssevnen til radarstrålen. Det at tendensen til en del av tidsserieplottene for oversvømte områder stiger litt i starten av tidsperioden kan indikere at jorda blir våtere, før den ved neste opptak har blitt oversvømt og får en synkende kurve. Dersom plottet har en synkende kurve helt fra start av tidsperioden, kan det bety at området allerede var oversvømt ved starten av perioden. Økt dybde og størrelse på overvannet kan være grunnen til verdiene fortsetter å synke også etter at områdene har blitt oversvømt. Ved slutten av nedbørsperioden begynner flere av kurvene å stige. Dette kan tyde på at området begynner å tørke opp slik at overvannet avtar i dybde og størrelse. Et bilde fra et opptakstidspunkt med meget høy fuktighet i jorda vil dermed gi større kontrast mellom jord og overvann enn et opptak hvor jorda er mindre fuktig. De diffuse resultatene fra rasterkalkuleringen og analysen basert på bildene fra tidsrom 1 kan skyldes at områdene var mindre våte en antatt.

6.2 Funnenes pålitelighet

I oppgavens problemstilling kommer det fram at det ikke er tilgjengelig fasitdata som kan validere resultatene fra analysen. Påliteligheten til funnene baseres dermed på antagelser, fornuft og hvor godt resultatene fra de ulike metodene samsvarer med hverandre. Dermed blir påliteligheten noe begrenset, selv om noen av resultatene blir styrket ved sammenligning med NDWI. Resultater tyder på at bildet fra oktober egnet seg bedre til å detektere oversvømte områder enn bildet fra juli, og resultatene fra analysen av dette bildet blir betraktet som de mest pålitelige funnene i analysen.

6.3 Usikkerhet i metode

6.3.1 Preprosessering

Løypa som ble benyttet til preprosesseringen er basert på fremgangsmåten og teorien til Braun [Braun, 2021] og Filipponi [Filipponi, 2019]. Andre varianter eller variasjoner av preprosessering ble ikke undersøkt, men ville kunne gitt små forskjeller i SAR-bildene anvendt i analysen. *Speckle filtering* ble utelatt i fare for å miste data fra enkeltpikslar. Siden de mest pålitelige resultatene fra analysen var områder som var større enn dette, kunne muligens dette steget likevel ha blitt inkludert og til og med hatt en positiv effekt på resultatene.

6.3.2 Polarisering

I analysen benyttes begge polariseringskanaler (VH og VV) for å inkludere mest mulig informasjon per piksel i bildet. I tiddserieanalysen analyseres polariseringskanalene hver for seg, mens i tersklingen og klassifiseringen benyttes en kombinasjon av kanalene. Mayer [2019] definerer at VV-kanalen gir sterkere tilbakespredning enn VH-kanalen ved overflatespredning, men i litteraturen finnes det eksempler på at både VV og VH er benyttet ved flomanalyser. Det ble ikke forsøkt å utføre terskling eller klassifisering for bare én av kanalene, men det kunne vært et alternativ å gjøre dette og sammenligne resultatene, slik som Clement [Clement et al., 2017] har gjort i sin studie.

6.3.3 Terskling og klassifisering

Terskelverdiene ble satt basert på en manuell tolkning av histogrammene. Dette gir en god forståelse for hvordan tersklingen fungerer, men er ikke særlig effektivt ved en mer omfattende analyse med bruk av mange bilder med ulik terskelverdi. Dessuten har det vist seg ganske vanskelig å bestemme en passende terskelverdi. Ved å sette en strengere, altså lavere, terskelverdi kunne man unngått å få med en del pikslar som i analysen sannsynligvis er feilkategorisert som vann. Samtidig ville man kunne mistet pikslar som faktisk er vann. Dersom målet er å lokalisere store oversvømte områders plassering kan terskelverdien definitivt settes lavere, men da vil det antakeligvis være vanskelig å fastslå utstrekningen til det oversvømte arealet. Det eksisterer en rekke metoder for automatisk terskling, men i SNAP finnes det inget annet verktøy enn båndkalkulatoren som ble brukt i analysen.

Det ble ikke testet ut noen andre klassifiseringsalgoritmer enn Random forest i denne oppgaven. Ved å bruke mer tid på tilpasning og utforskning av klassifiseringsmetoder ville kanskje resultatene blitt bedre også for bildet fra juli. I klassifiseringen ble det definert to klasser; vann og ikke-vann. Ved å inkludere flere klasser, som bygninger og veier, kunne muligens klassifiseringen av vann-pikslar blitt mer nøyaktig.

6.4 Andre usikkerheter

6.4.1 Skygger

En del piksler som i analysen indikerer oversvømte områder befinner seg helt på vestkanten av jorder som grenser til tette skogområder. Siden det ser ut som samme fenomen meget sjelder oppstår på østsiden av jorder, kan det tyde på at det er en skyggeeffekt som forårsaker dette. Siden Sentinel-1-satellitten ikke gjør opptakk rett ovenfra og ned, men litt på skrå, kan høy vegetasjon skape et skyggeområde hvor tilbakespredningsintensiteten vil være lavere enn overflatetypen tilsier.



Figur 6.4: Eksempler på piksler som kan være feilklassifisert på grunn av radarskygge

6.4.2 Vegetasjon

Avlinger på jordene kan påvirke deteksjonen av eventuelt overvann i bildene. Dersom plantene stikker opp over vannflata vil de redusere effekten av den glatte overflata som gir den ekstra lave tilbakespredning av radarsignalet. I verste fall kan høye planter med store blader helt eller delvis dekke for overvannet som ligger på jorda. Ulike typer vekster vil også kunne fremstå med ulik intensitet i bildet. Best mulig tidsrom for deteksjon av overvann vil dermed muligens tidlig på våren eller sent på høsten, altså utenfor avlingssesong, i kombinasjon med en kraftig nedbørsperiode.

7 Konklusjon

Det ble i innledningen presentert to forskningsspørsmål:

- *I hvilken grad kan oversvømt jordbruksareal detekteres i opptak fra radarsatellitten Sentinel-1?*
- *Kan resultatene fra analysen basert på data fra Sentinel-1 styrkes av resultat fra en annen datakilde?*

Resultater fra analysen indikerer potensielt oversvømt jordbruksareal i opptakene fra Sentinel-1. Både ved terskling og klassifisering ekstraheres enkeltpiksler og grupper av piksler hvor tilbakespredningsintensiteten tilsier at områdene er dekket av vann. Bruk av begge teknikkene på samme bilde gir noe ulik mengde oversvømt areal i studieområdet, men indikerer oversvømmelse i mange av de samme områdene. Større oversvømte områder synes å detekteres ganske bra, mens områder på størrelse med enkeltpiksler er vanskeligere å slå fast om faktisk er oversvømte områder eller bare støy. Tidsserieplottet viser en tendens til synkende tilbakespredningsintensitet utover i tidsperiodene i områdene som er antatt oversvømte. Forskjellen i resultater mellom opptakstidspunktene viser at man ved deteksjon av oversvømt areal bør velge å bruke data fra en våtest mulig periode - ikke bare på grunn av størrelsen på vannforekomstene, men også kontrasten i tilbakespredningsintensitet mellom våt jord og oversvømt areal.

Resultatene fra deler av analysen av Sentinel-1-data ble i resultatkapittelet visualisert sammen med NDWI-indeks fra Sentinel-2. Sammenligningen viste at store deler av jordbruksområder med en høy indeks i NDWI også var kategorisert som oversvømt areal i tersklingen og/eller i klassifiseringen. Den visuelle sammenligningen av resultatene av analysen og NDWI-rasteret viste dermed at NDWI styrker opp om resultatene fra analysen.

7.1 Videre arbeid

Data fra Sentinel-1 er uten tvil relevant for anvendelse i videre arbeid med deteksjon av oversvømt jordbruksareal. Forbedrede metoder og en eventuell kombinasjon av Sentinel-1-bilder og data fra andre sensorer vil etter all sannsynlighet gi enda bedre indikasjoner på hvilke områder som er flomutsatte enn resultatene fra denne analysen. Det anbefales å studere meteorologiske data for å kunne finne fram til Sentinel-1-bilder fra en våtest mulig periode. Dersom det finnes skyfrie Sentinel-2-opptak fra aktuelt tidspunkt kan den enkle tilgangen NDWI-indeksen i EO Browser være et godt hjelpemiddel til å sammenligne vannforekomstene på ulike tidspunkter.

Det ble i arbeidet med oppgaven også delvis undersøkt hvordan en digital terrengmodell kan brukes til å detektere områder som er utsatt for vannansamling, men det ble ikke gått videre inn på det i denne oppgaven. Å benytte høydedata for å analysere formasjoner i terrenget, som terrengforsenkninger og helningsgrad, vil absolutt være relevant i videre arbeid med en flomanalyse. Avslutningsvis vil jeg komme med noen flere forslag til videre arbeid:

- Innhente pålitelig valideringsdata slik at nøyaktigheten til metodene kan beregnes ved gjennomføring en tilsvarende analyse.

-
- Undersøke effekten av andre fremgangsmåter og variabler ved preprosessering av GRD-produktene.
 - Undersøke om resultatene kan forbedres ved å benytte kun én av polariseringskanalene, og eventuelt sammenligne de to kanalene.
 - Utforske andre metoder for å sette terskelverdi, automatisk eller manuelt.
 - Tilpasse klassifiseringen mer til formålet, og eventuelt teste andre klassifiseringsalgoritmer.
 - Inkludere mer data som er avgjørende for vannets evne til å bli liggende som overvann, for eksempel grunnforhold, jordkvalitet og terrengdata.
 - Kombinere data fra Sentinel-1 og Sentinel-2, på opptak hvor værforholdene tillater det.
 - Utføre en analyse som i tillegg til å detektere oversvømte områder også beregner vanninnhold i vegetasjon og jord.

Litteraturliste

[Abazaj, 2020] F. Abazaj, «SENTINEL-2 Imagery for Mapping and Monitoring Flooding in Buna River Area», Journal of International Environmental Application and Science, vol. 15, no. 2, s. 48-53, 2019. Hentet fra: <http://www.jieas.com/volumes/vol1201-2/abs20-v15-i2-2.pdf>

[Adobe, 1992] Adobe Developers Association. «TIFF. Revision 6.0.» Hentet fra: <https://www.adobe.io/content/dam/udp/en/open/standards/tiff/TIFF6.pdf> (Lastet ned: 20.10.2021).

[Boryan et al., 2018] C.G. Boryan, Z. Yang, A. Sandborn, P. Willis og B. Haack, «Agricultural flood monitoring with Sentinel-1 synthetic aperture radar,» 2018. Hentet fra: doi: 10.1109/I-GARSS.2018.8519458

[Braun 2021] A. Braun, «Time-series analysis with Sentinel-1,» 2021. Hentet fra: <http://step.esa.int/docs/tutorials/S1TBX%20Time-series%20analysis%20with%20Sentinel-1.pdf>

[Caubalejo, 2021] R. Caubalejo «Image Segmentation (Part 1),» 2021. Hentet fra: <https://towardsdatascience.com/image-segmentation-part-1-6e0ddb5f7f8a> (Lastet ned: 02.11.2021)

[Clement et al., 2017] M. Clement, C.G. Kilsby og P. Moore, «Multi-temporal synthetic aperture radar flood mapping using change detection,» 2017. Hentet fra: doi: DOI: 10.1111/jfr3.12303

[Cop, a] Copernicus Open Access Hub. URL: <https://scihub.copernicus.eu/dhus/#/home>

[Cop, b] MET Norway - Copernicus Data Hub. URL: <https://colhub.met.no/#/home>

[EOS] Earth Observing System. «Sentinel-2 Bands.» EOS.com. Hentet fra: <https://eos.com/find-satellite/sentinel-2/> (Lastet ned: 05.11.2021).

[ESA, a] European Space Agency. «Sentinel-1.» ESA.int. Hentet fra: <https://sentinel.esa.int/web/sentinel/missions/sentinel-1/> (Lastet ned: 10.10.2021).

[ESA, b] European Space Agency. «Sentinel-1: ESA's Radar Observatory Mission for GMES Operational Services.» 2012. Hentet fra: https://sentinel.esa.int/documents/247904/349449/S1_SP-1322_1.pdf (Lastet ned: 10.10.2021).

[ESA, c] European Space Agency. «Revisit and Coverage.» ESA.int. Hentet fra: <https://sentinels.copernicus.eu/web/sentinel/user-guides/sentinel-1-sar/revisit-and-coverage/> (Lastet ned: 10.10.2021).

[ESA, d] European Space Agency. «Acquisition Modes.» ESA.int. Hentet fra: <https://sentinels.copernicus.eu/web/sentinel/user-guides/sentinel-1-sar/acquisition-modes/> (Lastet ned: 10.10.2021).

[ESA, e] European Space Agency. «Polarimetry.» ESA.int. Hentet fra: <https://sentinels.copernicus.eu/web/sentinel/user-guides/sentinel-1-sar/product-overview/polarimetry> (Lastet ned: 21.10.2021).

[ESA, f] European Space Agency. «Sentinel-2.» ESA.int. Hentet fra: <https://sentinel.esa.int/web/sentinel/missions/sentinel-2> (Lastet ned: 05.11.2021).

[ESA, g] European Space Agency. «MultiSpectral Instrument (MSI).» ESA.int. Hentet fra: <https://sentinel.esa.int/web/sentinel/missions/sentinel-2/instrument-payload/> (Lastet ned: 05.11.2021).

05.11.2021).

[ESA, h] European Space Agency. «Level-1 GRD Products.» ESA.int. Hentet fra: <https://sentinels.copernicus.eu/web/sentinel/technical-guides/sentinel-1-sar/products-algorithms/level-1-algorithms/ground-range-detected> (Lastet ned: 20.09.2021).

[ESA, i] European Space Agency. «SNAP.» ESA.int. Hentet fra: <https://earth.esa.int/eogateway/tools/snap> (Lastet ned: 10.10.2021).

[ESA, j] Sentinel HUB EO Browser. URL: <https://apps.sentinel-hub.com/eo-browser/?zoom=10&lat=41.9&lng=12.5&themeId=DEFAULT-THEME&toTime=2021-12-14T22%3A31%3A51.318Z>

[ESRI] Environmental Systems Research Institute. «ArcGIS Pro.» Esri.com. Hentet fra: <https://www.esri.com/en-us/arcgis/products/arcgis-pro/overview> (Lastet ned: 10.10.2021).

[ESRI, 1998] Environmental Systems Research Institute, «ESRI Shapefile Technical Description,» 1998. Hentet fra: <https://www.esri.com/content/dam/esrisites/sitecore-archive/Files/Pdfs/library/whitepapers/pdfs/shapefile.pdf>

[FFI, 2019] R.L. Lillestøl, K. Eldhuset, T.N. Hannevik, A.O. Knapskog, R. Olsen og K. Viken, «Anvendelser av SAR-bilder fra satellitter over land,» Forsvarets forskningsinstitutt, 2019. Hentet fra: <https://publications.ffi.no/nb/item/asset/dspace:6420/19-01695.pdf>

[Filipponi, 2019] F. Filliponi, «Sentinel-1 GRD preprocessing workflow,» 2019. Hentet fra: <https://sciforum.net/manuscripts/6201/manuscript.pdf>

[Gujral, 2020] N. Gujral, «Random Forest Classifier — A Forest of Predictions,» 2020. Hentet fra: <https://levelup.gitconnected.com/random-forest-classifier-a-forest-of-predictions-6ee30b59de8c> (Lastet ned: 29.10.2021)

[Guruprasad, 2020] P. Guruprasad, K.S. Mahalingpur og T.N. Manesj, «Overview of different thresholding methods in image processing,» 2020. Hentet fra: https://www.researchgate.net/publication/342038946_overview_of_different_thresholding_methods_in_image_processing#fullTextFileContent

[Hanssen, 2001] R.F. Hanssen, «Radar Interferometry: Data Interpretation and Error Analysis.» Kluwer Academic, 2001. Hentet fra: <https://doi.org/10.1007/0-306-47633-9>

[Hoare, 2019] Hoare, «Example of descision tree» Hentet fra: <https://www.researchgate.net/publication/342624204/figure/fig8/AS:908763234041862@1593677447038/Toy-example-of-decision-tree-Hppm>

[Kartverket] Kartverket. «Høgdedata og djupnedata.» Kartverket.no. Hentet fra: <https://www.kartverket.no/api-og-data/terrengdata> (Lastet ned 11.11.2021).

[Lasswell, 2005] S.W. Lasswell, «History of SAR at Lockheed Martin (formerly Goodyear Aerospace),» 2005. Hentet fra: doi: 10.1117/12.603927

[McFeeters, 1995] S.K. McFeeters, «The use of the Normalized Difference Water Index (NDWI) in the delineation of open water features,» 1995. Hentet fra: doi: 10.1080/01431169608948714

[McVittie, 2019] A. McVittie, «Flood mapping tutorial,» 2019 Hentet fra: http://step.esa.int/docs/tutorials/tutorial_sifloodmapping.pdf

[Met.inst., 2021] Meteorologisk institutt. «Organisasjonen.» Met.no. Hentet fra: <https://www.met.no/vaer-og-klima/det-blir-vatere/> (Lastet ned: 14.10.2021).

[Meyer, 2019] F. Meyer, «Spaceborne Synthetic Aperture Radar: Principles, Data Access, and Basic Processing Techniques,» i The SAR Handbook. USA: SERVIR Global, 2019. Hentet fra: https://gis1.servirglobal.net/TrainingMaterials/SAR/SARHB_FullRes.pdf

[NIBIO, 2016] Norsk institutt for bioøkonomi. «AR5.» Nibio.no. Hentet fra: <https://www.nibio.no/tema/jord/arealressurser/arealressurskart-ar5> (Lastet ned 12.10.2021)

[Nor] Norsk Klimaservicesenter. URL: <https://klimaservicesenter.no/>

[Oldertrøen, 2021] J.O. Oldertrøen. «Drenering gir bedre avling og redusert klimagassutslipp.» Ni-bio.no. Hentet fra: <https://www.nibio.no/nyheter/drenering-gir-bedre-avling-og-reduisert-klimagassutslipp> (Lastet ned: 23.08.2021).

[Punia] M. Punia «Supervised classification,» Hentet fra: <https://www.shivajicollege.ac.in/sPanel/uploads/econtent/6367e01c3436f4564679dd6d45a627b0.pdf>

[QGIS, a] QGIS. «QGIS - The Leading Open Source Desktop GIS.» QGIS.com. Hentet fra: <https://www.qgis.org/en/site/about/index.html> (Lastet ned: 20.10.2021)

[QGIS, b] QGIS. «Raster Calculator.» QGIS.com. Hentet fra: https://docs.qgis.org/2.8/en/docs/user_manual/working_with_raster/raster_calculator.html (Lastet ned: 7.11.2021)

[Raschka et al., 2019] S. Raschka og V. Mirjalili, «Python Machine Learning: Machine Learning and Deep Learning with Python, Scikit-Learn, and TensorFlow, 2nd Edition,» Packt Publishing, 2017. ISBN 1787125939.

[Reksten et al., 2019] J.H. Reksten, A.B. Salberg, og R. Solberg, «Flood detection in Norway based on Sentinel-1 SAR imagery,» Int. Arch. Photogramm. Remote Sens. Spatial Inf. Sci., vol. XLII-3/W8, s. 349–355, 2019. Hentet fra: doi: 10.5194/isprs-archives-XLII-3-W8-349-2019

[Sezgin et al., 2004] M. Sezgin og B. Sankur, «A survey over image thresholding techniques and quantitative performance evaluation,» 2004. Hentet fra: https://www.researchgate.net/publication/309967669_Image_thresholding_techniques_A_survey_over_categories

[Skolnik, 1990] M.I. Skolnik. «Radar handbook Second edition.» McGraw-Hill, 2008. Hentet fra: <http://www.geo.uzh.ch/microsite/rsl-documents/research/SARlab/GMTILiterature/PDF/Skolnik90.pdf>

[Skolnik, 2008] M.I. Skolnik. «Radar handbook.» McGraw-Hill, 2008

[SOSI, 2018] Geonorge, «Produktspesifikasjon FKB-Vann 4.61,» 2018. Hentet fra: http://sosi.geonorge.no/Produktspesifikasjoner/Produktspesifikasjon_FKB-Vann_4.61.pdf

[Sun et al., 2012] F. Sun, W. Sun, J. Chen og P. Gong, «Comparison and improvement of methods for identifying waterbodies in remotely sensed imagery,» International Journal of Remote Sensing, vol. 33 nr. 21, s. 6854-6875, 2012. Hentet fra: doi: 10.1080/01431161.2012.692829

[Tsyganskaya et al., 2018] V. Tsyganskaya, S. Martinis, P. Marzhan og R. Ludwig, «SAR-based detection of flooded vegetation – a review of characteristics and approaches,» International Journal of Remote Sensing, vol. 39, nr. 8, s. 2255-2293, 2018. Hentet fra: doi: 10.1080/01431161.2017.1420938

A Klassifiseringsrapporter

A.1 Klassifiseringsrapport juli

RandomForest classifier newClassifier

Cross Validation

Number of classes = 2

class 0.0: ar5_wetarea_mapping

accuracy = 0.9628 precision = 0.9610 correlation = 0.9284 errorRate = 0.0372

TruePositives = 2412.0000 FalsePositives = 98.0000 TrueNegatives = 2402.0000 FalseNegatives = 88.0000

class 1.0: treningsdata

accuracy = 0.9628 precision = 0.9647 correlation = 0.9284 errorRate = 0.0372

TruePositives = 2402.0000 FalsePositives = 88.0000 TrueNegatives = 2412.0000 FalseNegatives = 98.0000

Using Testing dataset, % correct predictions = 96.2800

Total samples = 10000

RMSE = 0.19287301521985908

Bias = -0.0020000000000000018

Distribution:

class 0.0: ar5_wetarea_mapping 5000 (50.0000%)

class 1.0: treningsdata 5000 (50.0000%)

Testing feature importance score:

Each feature is perturbed 3 times and the % correct predictions are averaged

The importance score is the original % correct prediction - average

rank 1 feature 1 : Sigma0_VH_db score: tp=0.3483 accuracy=0.3483 precision=0.3262

correlation=0.4737 errorRate=-0.3483 cost=-0.5070 GainRatio = 0.2600

rank 2 feature 2 : Sigma0_VV_db score: tp=0.3280 accuracy=0.3280 precision=0.2884

correlation=0.4750 errorRate=-0.3280 cost=-0.4371 GainRatio = 0.2658

A.2 Klassifiseringsrapport oktober

RandomForest classifier randomforest

Cross Validation

Number of classes = 2

class 0.0: ar5_wetarea_mapping

accuracy = 0.9972 precision = 0.9964 correlation = 0.9944 errorRate = 0.0028

TruePositives = 2495.0000 FalsePositives = 9.0000 TrueNegatives = 2491.0000 FalseNegatives = 5.0000

class 1.0: treningsdata

accuracy = 0.9972 precision = 0.9980 correlation = 0.9944 errorRate = 0.0028

TruePositives = 2491.0000 FalsePositives = 5.0000 TrueNegatives = 2495.0000 FalseNegatives = 9.0000

Using Testing dataset, % correct predictions = 99.7200

Total samples = 10000

RMSE = 0.05291502622129181

Bias = -8.000000000000229E-4

Distribution:

class 0.0: ar5_wetarea_mapping 5000 (50.0000%)

class 1.0: treningsdata 5000 (50.0000%)

Testing feature importance score:

Each feature is perturbed 3 times and the % correct predictions are averaged

The importance score is the original % correct prediction - average

rank 1 feature 2 : Sigma0_VV_db score: tp=0.4763 accuracy=0.4763 precision=0.4743

correlation=0.5180 errorRate=-0.4763 cost=-0.9044 GainRatio = 0.4151

rank 2 feature 1 : Sigma0_VH_db score: tp=0.0109 accuracy=0.0109 precision=0.0109

correlation=0.0216 errorRate=-0.0109 cost=-0.0111 GainRatio = 0.3433



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

Postboks 5003
NO-1432 Ås
Norway