



Norges miljø- og  
biovitenskapelige  
universitet

**Masteroppgave 2024 30 stp**  
Fakultet for realtek og teknologi

# **Kartlegging av lille istids utbredelse på Hardangerjøkulen ved bruk av Sentinel-2 data**

Mapping the extent of the Little Ice Age on  
Hardangerjøkulen using Sentinel-2 data

**Julie Løkeberg**  
Geomatikk





# Forord

Denne oppgaven indikerer slutten på geomatikkstudiet og fem fine år ved Norges miljø- og biovitenskapelige universitet (NMBU). Gjennom disse årene på Ås har jeg fått utviklet meg både faglig og som person. Jeg er veldig taknemlig for alle de fantastiske menneskene jeg har møtt og som har motivert meg gjennom studietiden.

Jeg vil rette en stor takk til medveilederne mine Liss Marie Andreassen og Paul Weber for god teoretisk hjelp og veiledning gjennom hele oppgaven. Deres kompetanse er enorm og jeg har lært utrolig mye av dere i løpet av dette semesteret. I tillegg vil jeg gi en stor takk til hovedveileder Ivar Maalen-Johansen for hans gode veiledning, tilgjengelighet og fine oppmuntrende samtaler. Oppgaven hadde ikke blitt den samme uten deres støtte og hjelp.

Oppgaven er skrevet våren 2024 ved Fakultet for realfag og teknologi ved NMBU og utgjør 30 studiepoeng.

Ås, mai 2024  
Julie Løkeberg

# Sammendrag

Klimaendringer har en betydelig innvirkning på smeltingen av norske isbreer. Regelmessige masseberegninger og lengdemålinger av isbreene er essensielle for å overvåke breenes utvikling. Denne studien ser på Hardangerjøkulen under den lille istid (LIA), rundt 1750, da isbreen var på sitt største. Fra den gang og frem til 2019, har Hardangerjøkulens areal minnet med omtrent  $45\text{km}^2$  (Andreassen, 2022; Weber mfl., 2019).

Tidligere studier har vist at satellittbilder er en effektiv metode for å kartlegge isbreer. Alle norske breer har blitt kartlagt med Sentinel-2 bilder fra 2018 og 2019 (Andreassen, 2022). Satellittbilder sammen med ortofoto, muliggjør videre studier av LIA utbredelsen. Denne studien undersøker bruken av Sentinel-2 for å kartlegge LIA utbredelse på Hardangerjøkulen ved hjelp av ortofoto og kartlagte breomkretser fra tre forskjellige årstall. Satellittdatasettet er levert av NVE med opptaksdato fra 2019, mens ortofoto er tatt av Terratec AS sensommeren 2019. Breomkretsene er levert av Weber mfl. (2019) og ble manuelt kartlagt i 2016 og 2017. Maskinlæringsalgoritmene Support Vector Machine (SVM) og Random Trees (RT) brukes i en styrt klassifisering, hvor brearmene Midtdalsbreen og Blåisen fungerer som treningsområder. De trente modellene testes deretter på Rembesdalskåka, som ligger på motsatt side av breen mot vest. Videre undersøkes de karakteristiske trekkene innenfor og utenfor LIA med en pikselbasert statistisk analyse for hele Hardangerjøkulen.

Metodene gir gode visuelle resultater og jevnt gode klassifiseringsresultater, med en total nøyaktighet på henholdsvis 80% for SVM og 77% for RT. De statistiske analysene viser til et betydelig skille innenfor og utenfor LIA, med vegetasjon som den mest signifikante variabelen. Bånd 8 (NIR), 11 (SWIR) og indeksene NDVI, GNDVI og BSI viser seg å være de mest betydningsfulle i studien.

# Abstract

Climate change has a significant impact on the melting of Norwegian glaciers. Regular mass calculations and length measurements of the glaciers are essential to monitor their development. This study is focused on the glacier Hardangerjøkulen during the Little Ice Age (LIA), ca. 1750, when the glacier was at its largest. From then until 2019, the area of Hardangerjøkulen has decreased by approximately  $45\text{km}^2$  (Andreassen, 2022; Weber et al., 2019).

Previous studies have shown that satellite imagery is an effective method for mapping ice glaciers. All Norwegian glaciers have been mapped with Sentinel-2 imagery from 2018 and 2019 (Andreassen, 2022). Satellite images, together with orthophotos, enable further studies of the LIA distribution. This study investigates the use of Sentinel-2 to map LIA distribution on Hardangerjøkulen using orthophotos and mapped glacier circles from three different years. The satellite dataset is provided by NVE with a recording date from 2019, while the orthophoto was taken by Terratec AS in late summer 2019. The glacier perimeters are provided by Weber et al (2019) and were manually mapped in 2016 and 2017. The machine learning algorithms Support Vector Machine (SVM) and Random Trees (RT) are used in a supervised classification, where the glacier arms Midtdalsbreen and Blåisen serve as training areas. The trained models are then tested on Rembesdalskåka, which is located on the other side of the glacier to the west. Furthermore, the characteristic features inside and outside the LIA with a pixel-based statistical analysis for the entire Hardangerjøkulen.

The methods provide good visual results and consistently good classification results, with an overall accuracy of 80% for SVM and 77% for RT. The statistical analyses show a significant difference inside and outside the LIA, with vegetation being the most significant variable. Band 8 (NIR), 11 (SWIR) and the indices NDVI, GNDVI and BSI proved to be the most significant in the study.

# Innhold

<b>Forord</b>	<b>i</b>
<b>Sammendrag</b>	<b>ii</b>
<b>Abstract</b>	<b>ii</b>
<b>Figurer</b>	<b>vii</b>
<b>Tabeller</b>	<b>xi</b>
<b>Forkortelser</b>	<b>xii</b>
<b>1 Innledning</b>	<b>1</b>
1.1 Bakgrunn . . . . .	1
1.2 Problemstilling . . . . .	3
1.3 Tidligere arbeid . . . . .	3
1.4 Struktur . . . . .	4
<b>2 Teori</b>	<b>5</b>
2.1 Sentinel-2 . . . . .	5
2.2 NDVI . . . . .	6
2.3 GNDVI . . . . .	7
2.4 BSI . . . . .	7
2.5 Support Vector Machine . . . . .	8
2.6 Random Trees . . . . .	9
2.7 Iso Cluster Unsupervised Classification . . . . .	9
2.8 Valideringsparametere . . . . .	10

<b>3</b>	<b>Studieområde og material</b>	<b>12</b>
3.1	Studieområde . . . . .	12
3.2	Datagrunnlag . . . . .	13
3.2.1	Ortofoto . . . . .	14
3.2.2	Sentinel-2 satellittbilder . . . . .	14
3.2.3	Breomkrets og rebasseng . . . . .	14
3.3	Utvalgte Sentinel-2 bånd . . . . .	15
3.4	Båndkombinasjoner og indekser . . . . .	17
3.5	Programvarer . . . . .	20
<b>4</b>	<b>Metode</b>	<b>21</b>
4.1	Styrt klassifisering . . . . .	21
4.1.1	Utklipp av trening- og testområdet . . . . .	22
4.1.2	Treningsdata . . . . .	23
4.1.3	Support Vector Machine . . . . .	24
4.1.4	Random Trees . . . . .	24
4.1.5	Testdata . . . . .	25
4.1.6	Kvalitetsvurdering . . . . .	25
4.2	Pikselbasert analyse . . . . .	26
4.2.1	Iso Cluster Unsupervised Classification . . . . .	26
4.2.2	Buffer . . . . .	27
4.2.3	Utklipp av brearmer . . . . .	27
4.2.4	Feature Vertices To Point . . . . .	28
4.2.5	Extract Multi Values To Point . . . . .	29
4.2.6	Sentinel-2 pikselverdier . . . . .	29
4.3	Feilklassifisering . . . . .	30
<b>5</b>	<b>Resultater</b>	<b>31</b>
5.1	Klassifisering av LIA utbredelse . . . . .	31
5.1.1	Statistisk test . . . . .	33
5.2	Sentinel-2 pikselverdier statistikk . . . . .	34
5.2.1	Breomkretser . . . . .	34
5.2.2	LIA utbredelse . . . . .	37
5.2.3	Midtdalsbreen, Blåisen og Rembesdalskåka . . . . .	39
5.2.4	Båndkombinasjoner over Midtdalsbreen og Blåisen . . . . .	42
5.2.5	Båndkombinasjoner over Rembesdalskåka . . . . .	44
<b>6</b>	<b>Diskusjon</b>	<b>46</b>

6.1	Usikkerhet i datagrunnlag . . . . .	46
6.2	Usikkerheter knyttet til metode . . . . .	47
6.2.1	Klassifiseringsmetoder . . . . .	47
6.2.2	Trening- og testområde . . . . .	47
6.2.3	Effekten av treningsdata . . . . .	47
6.2.4	Valideringspunkter . . . . .	48
6.2.5	Andre usikkerheter . . . . .	48
6.3	Resultater . . . . .	49
6.3.1	Klassifisering . . . . .	49
6.3.2	Pikselbasert analyse . . . . .	52
6.3.3	Visuell analyse av brearmer . . . . .	53
6.4	Bruk av kunstig intelligens . . . . .	55
6.5	Mulige feilkilder . . . . .	55
6.6	Videre arbeid . . . . .	55
<b>7</b>	<b>Konklusjon</b>	<b>57</b>
	<b>Litteraturliste</b>	<b>61</b>
	<b>A Effekten av treningsdata</b>	<b>62</b>
	<b>B Minimum- og maksimumsverdier</b>	<b>68</b>
	<b>C Korrelasjon</b>	<b>72</b>

# Figurer

1.1.1	Morenerygger mot Midtdalsbreen tatt september 2023. (Foto: Julie Løkeberg) . . . . .	2
1.1.2	Utsnitt av terrenget ved Midtdalsbreen og Blåisen tatt september 2023. (Foto: Julie Løkeberg) . . . . .	2
2.1.1	Sentinel-2 spektralbånd og romlig oppløsning. Bilde hentet fra:(Gascon mfl., 2017) . . . . .	5
2.5.1	SVM finner det optimale hyperplanet som fungerer som en skillelinje mellom klassene. Hyperplanet er parallelt med de begrensende hyperplanene der støttevektorene ligger. Marginen er definert som avstanden mellom disse to begrensende hyperplanene. Illustrasjon hentet fra: (Cardoso-Fernandes mfl., 2020) . . . . .	8
2.6.1	RT bruker de ulike beslutningstrær til å bestemme det endelige resultatet. Illustrasjon hentet fra: (M. Y. Khan mfl., 2021) . . . . .	9
2.8.1	Eksempel på en flerdimensjonal forvirringsmatrise. . . . .	10
2.8.2	Den generelle strukturen i en forvirringsmatrise. Illustrasjon inspirert av (Ahmed, 2023) . . . . .	11
3.1.1	Oversikt over plasseringen til Hardangerjøkulen med brearmene Midtdalsbreen (MB), Blåisen (BI) og Rembesdalskåka (RS). . . . .	13
3.2.1	Rekonstruksjon av LIA fra 2016 og 2017 gjort av Paul Weber, (Weber mfl., 2019). Her ser man også brearmene Midtdalsbreen (MB), Blåisen (BI), Torsteinsfonna (TF), Austra Leirebottsskåka (AL), Vestra Leirebottsskåka (VL), Isdøleskåka (IS) og Rembesdalskåka (RS) . . . . .	15
3.3.1	Bånd 2: blå, 492.7nm, GSD 10m. . . . .	16
3.3.2	Bånd 3: grønn, 559.8nm, GSD 10m. . . . .	16

3.3.3	Bånd 4: rød, 664.6nm, GSD 10m. . . . .	16
3.3.4	Bånd 8: NIR, 832.8nm, GSD 10m. . . . .	16
3.3.5	Bånd 11: SWIR, 1613.7nm, GSD 20m. . . . .	16
3.3.6	Bånd 12: SWIR, 2202.4nm, GSD 20m. . . . .	16
3.4.1	Naturlig bilde fremstilt i rød (4), grønn (3) og blå (2). . . . .	18
3.4.2	Falsk bilde fremstilt med båndene NIR (8), rød (4) og grønn (3). . . . .	18
3.4.3	Falsk bilde fremstilt i SWIR (11), NIR (8) og rød (4). . . . .	18
3.4.4	NDVI . . . . .	19
3.4.5	GNDVI . . . . .	19
3.4.6	BSI . . . . .	19
4.0.1	Flytskjema for metoden. . . . .	21
4.1.1	Flytskjema for klassifiseringen. . . . .	22
4.1.2	Treningsområdet over Midtdalsbreen og Blåisen fremstilt i båndkombinasjonen NIR(8), rød(4) og grønn(3). . . . .	22
4.1.3	Testområdet over Rembesdalskåka fremstilt i båndkombinasjonen NIR(8), rød(4) og grønn(3). . . . .	22
4.1.4	Fordeling av treningsdata representert i prosent og antall polygoner. . . . .	23
4.1.5	Oversikt over plassering av treningsdata med bånd 8 (NIR) i bakgrunn. . . . .	24
4.1.6	Utsnitt av klassifisert bildet med tilhørende valideringspunkt. Bildet viser den tilfeldige fordelingen til valideringspunktene for evaluering av klassifiseringen. . . . .	25
4.2.1	Flytskjema for pikselbasert analyse. . . . .	26
4.2.2	Tre polygoner delt inn fra 1km utenfor LIA og frem til LIA (1750), fra LIA (1750) til 1920 og innerste polygon fra 1920 til 2013. . . . .	27
4.2.3	Tre polygoner delt inn fra 1km utenfor LIA og frem til LIA (1750), fra LIA (1750) til 1920 og innerste polygon fra 1920 til 2013, for hver av brearmene Midtdalsbreen (MB), Blåisen (BI) og Rembesdalskåka (RS). . . . .	28
4.2.4	Utklipp av punkt som ligger 20m og 100m på innsiden og utsiden av den sikre kartleggingen av LIAs morenerygg. . . . .	29
5.1.1	Resultatet av klassifisering med SVM på testområdet Rembesdalskåka. . . . .	32
5.1.2	Resultatet av klassifisering med RT på testområdet Rembesdalskåka. . . . .	32



5.1.3	SVM forvirringsmatrise. . . . .	33
5.1.4	RT forvirringsmatrise. . . . .	33
5.2.1	Gjennomsnittsverdier for S-2A båndene for hvert polygon. . . . .	35
5.2.2	Gjennomsnittsverdier for indeksene NDVI, GNDVI og BSI for hvert polygon. . . . .	35
5.2.3	(a), (b) og (c) viser gjennomsnittsverdien til hvert av polygonene rundt Hardangerjøkulen med tilhørende standardavvik. . . . .	36
5.2.4	Gjennomsnittsverdier for S-2A båndene for hver punktlinje. . . . .	37
5.2.5	Gjennomsnittsverdier for NDVI, GNDVI og BSI for hver punktlinje. . . . .	37
5.2.6	(a), (b), (c) og (d) viser gjennomsnittsverdien til hver av punktlinjene rundt Hardangerjøkulen med tilhørende standardavvik. . . . .	39
5.2.7	(a), (b) og (c) viser S-2A gjennomsnittsverdiene for polygonene tilhørende hver av brearmene. . . . .	40
5.2.8	(a), (b) og (c) viser indeksene sin gjennomsnittsverdien for polygonene tilhørende hver av brearmene. . . . .	41
5.2.9	(a), (b) og (c) Visuelt resultat av Midtdalsbreen og Blåisen gitt i tre ulike båndkombinasjoner. . . . .	43
5.2.10	(a), (b) og (c) Visuelt resultat av Rembesdalskåka gitt i tre ulike båndkombinasjoner. . . . .	45
6.3.1	SVM klassifiseringsresultat for Rembesdalskåka. . . . .	50
6.3.2	RT klassifiseringsresultat for Rembesdalskåka. . . . .	50
6.3.3	NDVI over testområdet Rembesdalskåka. . . . .	50
6.3.4	LIA fasit rundt Rembesdalskåka. . . . .	50
6.3.5	SVM klassifiseringsresultat på utsiden av LIA ved Rembesdalskåka. . . . .	51
6.3.6	RT klassifiseringsresultat på utsiden av LIA ved Rembesdalskåka. . . . .	51
6.3.7	NDVI ved Rembesdalskåka utenfor moreneryggen til LIA. . . . .	51
6.3.8	Ortofoto over samme området fra (Norgebilder.no, 2024). . . . .	51
6.3.9	(a), (b) og (c) Visuelt resultat over Rembesdalskåka, viser markante overganger innenfor og utenfor LIA . . . . .	54
A.0.1	Resultatet etter klassifisering med SVM på treningsområdet Midtdalsbreen og Blåisen med treningsdatasett 1. . . . .	64
A.0.2	Resultatet etter klassifisering med SVM på treningsområdet Midtdalsbreen og Blåisen med treningsdatasett 2. . . . .	64

A.0.3	SVM forvirringsmatrise for treningsdata 1. . . . .	65
A.0.4	SVM forvirringsmatrise for treningsdata 2. . . . .	65
A.0.5	Resultatet etter klassifisering med RT på treningsområdet Midtdalsbreen og Blåisen med treningsdatasett 1. . . . .	66
A.0.6	Resultatet etter klassifisering med RT på treningsområdet Midtdalsbreen og Blåisen med treningsdatasett 2. . . . .	66
A.0.7	RT forvirringsmatrise for treningsdata 1. . . . .	67
A.0.8	RT forvirringsmatrise for treningsdata 2. . . . .	67
B.0.1	(a) og (b) representerer minimum- og maksimumsverdier for de tre polygonene. . . . .	68
B.0.2	(a) og (b) representerer minimum- og maksimumsverdier for de fire punktlinjene. . . . .	69
B.0.3	(a), (b) og (c) representerer minimumsverdiene for brearmene Midtdalsbreen (MB), Blåisen (BI) og Rembesdalskåka (RS). . . . .	70
B.0.4	(a), (b) og (c) representerer maksimumsverdiene for brearmene Midtdalsbreen (MB), Blåisen (BI) og Rembesdalskåka (RS). . . . .	71

# Tabeller

2.1.1	Sentral oppløsning og bølgelengde for Sentinel-2A og Sentinel-2B (ESA, 2024d; Sentinel-hub, 2024) . . . . .	6
3.1.1	Senterlinje lengde og areal for brearmene Midtdalsbreen, Blåisen og Rembesdalskåka under den lille istids maksimale utbredelse. Kilde: (Weber mfl., 2019). . . . .	13
5.1.1	Den totale nøyaktigheten for SVM og RT. . . . .	34
5.1.2	Presisjon, gjenkall og F1-verdi for SVM og RT innad i hver klasse. . . . .	34
5.2.1	Variansen for S-2A og indeksene mellom polygonene. . . . .	35
5.2.2	Varians for S-2A og indeksene mellom linjene. . . . .	37
5.2.3	Variansen mellom brearmene for hver av polygonene. . . . .	39
A.0.1	Den totale nøyaktigheten for SVM og RT for treningsdata 1 og 2. . . . .	62
A.0.2	Presisjon, gjenkall og F1-verdi for SVM innad i hver klasse for treningsdata 1 og 2. . . . .	63
A.0.3	Presisjon, gjenkall og F1-verdi for RT innad i hver klasse for treningsdata 1 og 2. . . . .	63
C.0.1	Korrelasjon mellom polygonene for de utvalgte Sentinel-2 båndene. . . . .	73
C.0.2	Korrelasjon mellom punktene langs linjene for de utvalgte Sentinel-2 båndene. . . . .	74

# Forkortelser

<b>AL</b>	Austra Leirebottsskåka
<b>BI</b>	Blåisen
<b>BSI</b>	Bare Soil Index
<b>DN</b>	Digital Numbers
<b>DTM</b>	Digital terrengmodell
<b>GNDVI</b>	Green Normalized Difference Vegetation Index
<b>IS</b>	Isdøleskåka
<b>KI</b>	Kunstig intelligens
<b>LIA</b>	Little Ice Age
<b>MB</b>	Midtdalsbreen
<b>MSI</b>	MuliSpectral Instrument
<b>NDVI</b>	Normalized Difference Vegetation Index
<b>NIR</b>	Nær infrarød
<b>RGB</b>	Rød, grønn og blå
<b>RS</b>	Rembesdalskåka
<b>RT</b>	Random Trees
<b>S-2A</b>	Sentinel-2A
<b>S-2B</b>	Sentinel-2B
<b>SVM</b>	Support Vector Machine
<b>SWIR</b>	Short-wave infrared
<b>TF</b>	Torsteinsfonna
<b>VL</b>	Vestra Leirebottsskåka

# Innledning

Oppgaven er gitt av seksjon for bre, is og snø ved NVE, med et ønske om å se hvor godt satellittbilder fra Sentinel-2 kan brukes til å kartlegge lille istid (LIA) utbredelse på Hardangerjøkulen.

Totalt 0.7% av Norge er dekket av isbreer, og er en stor ressurs innenfor vannkraft (Andreassen mfl., 2020). Brekontoret ved NVE ble innført på 1960-tallet som en respons på et økende behov for måling av massebalanse. Behovet etter å finne bidraget isbreene hadde til Norges energi, var grunnlaget for etableringen. I tillegg har klimaendringenes innvirkning på isbreer og lengdemålinger av disse blitt en vesentlig komponent i overvåkningen av klimaendringene (Andreassen mfl., 2005). Brekartlegging med Sentinel-2 er en del av Copernicus prosjektet som NVE, Norsk Polarinstitutt og Institutt for Geofag ved UiO samarbeider om og som videreføres i prosjektet NVE Copernicustjenester (Andreassen, Moholdt mfl., 2021; Andreassen, Engeset mfl., 2021).

## 1.1 Bakgrunn

Vår nåværende geologiske periode, Holocene, begynte etter den siste store istiden. I løpet av denne tiden har det vært flere mindre klimaendringer, med den lille istid som en av de mer fremtredende. Historisk data relatert til havisen i Nord-Atlanteren indikerer at det var en reduksjon i isdekket rundt Island og i Barentshavet i første halvdel av 1700-tallet sammenlignet med midten av 1600-tallet og slutten av 1700-tallet. Etter flere år med overvåking av havisen, ble det observert et mønster i vinterhavisen, som korrelerte med betydelige svingninger i lufttemperatur og havnivåtrykk. En gjennomsnittlig nedgang i temperatur, kombinert med en økning i vinternedbør, resulterte i de mest omfattende brefremstøtene på Vestlandet i løpet av 1900-tallet, og muligens siden tidlig på 1700-tallet (Nesje mfl., 2008).

LIA strekker seg over en lenger periode og man kan lese flere definisjoner rundt LIA. Det er usikkerhet knyttet til det nøyaktige tidspunktet for LIA, men det er antatt at i

Skandinavia kan denne perioden ha vart fra 1400-tallet til tidlig på 1900-tallet (Palacios mfl., 2024). Det eksakte tidspunktet for breenes maksimale utbredelse varierer fra isbre til isbre. Variasjonen kan i tillegg variere mellom breutløpere tilhørende samme isbre (Weber, 2020). For eksempel viser historien om Folgefonna at dens brearmer nådde sine største punkter i tre forskjellige tiår: rundt 1750, 1890, og en gang på 1930-tallet (Tvede, 1973). For Hardangerjøkulen er det antatt at den lille istids kulminasjon var rundt 1750 (Andersen & Sollid, 1971; Weber mfl., 2019). Når isen ekspanderer og gjør et fremstøt, dannes det morenerygger (Liestøl, 1963), vist i figurene 1.1.1 og 1.1.2.



**Figur 1.1.1:** Morenerygger mot Midtdalsbreen tatt september 2023. (Foto: Julie Løkeberg)



**Figur 1.1.2:** Utsnitt av terrenget ved Midtdalsbreen og Blåisen tatt september 2023. (Foto: Julie Løkeberg)

Målinger av brelegder og masseberegninger strekker seg tilbake til slutten av 1800-tallet. Metodene å måle på har blitt betydelig forbedret ved innføring av nyere teknologier, som satellittfjernmåling (Weber mfl., 2019; Zemp mfl., 2015). Tidligere rekonstruksjoner og kartlegginger har blitt gjort basert på historiske kart, topografiske bilder og fotografering fra fly (Andreassen, 2022; Weber mfl., 2019). Brelegdene ble målt med målebånd, fra brekant til et kjent fastmerke. Nøyaktigheten lå da på rundt 2m. Laser tok over jobben til målebåndet og nøyaktigheten økte til 1m (Andreassen mfl., 2005). Ved bruk av satellittdata åpnes mulighetene til å måle et betydelig større antall breer, selv i tilfeller der tilgjengeligheten er begrenset (Andreassen mfl., 2023).

Optiske bilder fra Landsat viste seg å være svært effektivt for å kartlegge isbreenes utbredelse (Winsvold mfl., 2014). Landsat 1 ble skutt opp i 1972 og hadde en romlig oppløsning på 68 x 83m. Utviklingen innen Landsat-satellittene økte den spektakle og romlige opp-

løsningen til 30m (Andreassen mfl., 2012). I 2015 ble Sentinel-2 satellitten satt i drift og reduserte to utfordringer knyttet til isbrekartlegging. Med en hyppigere omløpstid, bedre romlig og temoral oppløsning (10m) og med flere spektralbånd, kunne ruskete breer avgrensnes mer presist og sjansen for skyfrie bilder ble større (Paul mfl., 2016). Fra perioden 2018/2019 er alle norske breer klartlagt med Sentinel-2 (Andreassen, 2022).

I denne oppgaven vil Sentinel-2 bilder bli benyttet i en styrt pikselklassifisering, samt en dypere analyse innenfor bestemte områder for å undersøke brukervennligheten Sentinel-2 har i en kartlegging av LIA utbredelse på Hardangerjøkulen.

## 1.2 Problemstilling

Hvordan kan utbredelsen av LIA detekteres fra Sentinel-2 bilder? Og hvilke mulige motarbeidene faktorer finnes som kan gjøre denne deteksjonen vanskelig? Dette er to spørsmål som skal besvares i denne oppgaven. Med Sentinel-2 skal det undersøkes hvor godt satellittbilder egnet seg til å detektere LIA sin utbredelse på Hardangerjøkulen.

Problemstilling:

- Hvordan kan Sentinel-2 bilder brukes til å identifisere og kartlegge utbredelsen på Hardangerjøkulen under den lille istid?

## 1.3 Tidligere arbeid

Artikkelen *Mapping glaciers in Jotunheimen, South-Norway, during the “Little Ice Age” maximum* (Baumann mfl., 2009) hadde som mål å kartlegge isbreområdet i Jotunheimen under LIA maksimum og å utarbeide en isbreinventar. For Jotunheimen ble breutbredelsen under LIA kartlagt ved hjelp av fjernmålingsteknikker. Kartleggingen ble validert gjennom tolkning av eksisterende glasiokronologiske studier, analyse av geomorfologiske kart og egne GPS-feltemålinger. Datagrunnlag som ble benyttet for å kartlegge LIA-isbreene inkluderte Landsat 5 TM, flyfoto, geomorfologiske kart med LIA-konturer, GPS-punkter for LIA-morenerygger, digitale topografiske kart, DTM, digitale isbrekonturer og grenser for hydrologiske bassenger.

Studien viste at satellittbilder og flyfoto kan brukes til manuell kartlegging av isbrekonturer ved LIA-maksimum i Jotunheimen på en regional skala. Isbrestrømningslinjer ble manuelt digitalisert for alle isbreene for å beregne LIA-isbrelelengder. Resultatene var tilfredsstillende og det ble observert kun et fåtall isbreer som forsvant helt fra LIA til 2003. Generelt var isbreene større ved LIA-maksimum og den relative reduksjonen i isbreområdet var på 35%.



Paul og Kääb diskuterer i sin artikkel *Perspectives on the production of a glacier inventory from multispectral satellite data in Arctic Canada: Cumberland Peninsula, Baffin Island* (Paul & Kääb, 2005) effektiviteten av satellitt- og DOM-data for isbreinventering i arktiske områder. Ved å bruke Landsat ETM+ og Terra ASTER data, samt en ASTER-avledet DOM, demonstrerer studien hvordan GIS-basert prosessering kan kartlegge isbreer på det avsidesliggende stedet Cumberland Peninsula, som tilhører Baffin Island i det arktiske Canada. Resultatene viser et gjennomsnittlig arealetap på 11% for 225 isbreer. Tidsbruken for prosesseringen er også vurdert, med et gjennomsnitt på 5 minutter per isbre, som tilsvarer 7 år for alle anslåtte isbreene (160 000) over hele verden. Dette indikerer at selv om kartleggingen fra satellitt er rask, krever etterbehandlingen betydelig tid. Til slutt fremheves det at den ASTER-avledelede DOM-en gir flere muligheter for automatisert 3D-isbreinventering.

I artikkelen *Machine-Learning Algorithms for Mapping Debris-Covered Glaciers: The Hunza Basin Case Study* (A. A. Khan mfl., 2020), ble det utført en grundig undersøkelse av anvendelsen av tre forskjellige maskinlæringsalgoritmer for å klassifisere isbreer, ruskdekkede isbreer og utenfor isbreområdet. Disse algoritmene inkluderer Support Vector Machine (SVM), Artificial Neural Network (ANN) og Random Forest (RF). Dataene som ble brukt i denne studien, ble hentet fra Sentinel-2. Resultatene fra studien viste at RF-algoritmen utmerket seg ved å oppnå en F1-verdi på 95% for alle de tre klassene. Til sammenligning presterte ANN og SVM med en F1-verdi på rundt 92%.

## 1.4 Struktur

Resten av strukturen i oppgaven er organisert som følger. Først presenteres studieområdet og de brukte datasettene. Deretter beskrives klassifiseringsmetode samt metode for analyse over angitte områder. Til slutt gjennomføres en vurdering og sammenligning av resultatene, som avsluttes med en konklusjon av oppgaven og problemstillingen.

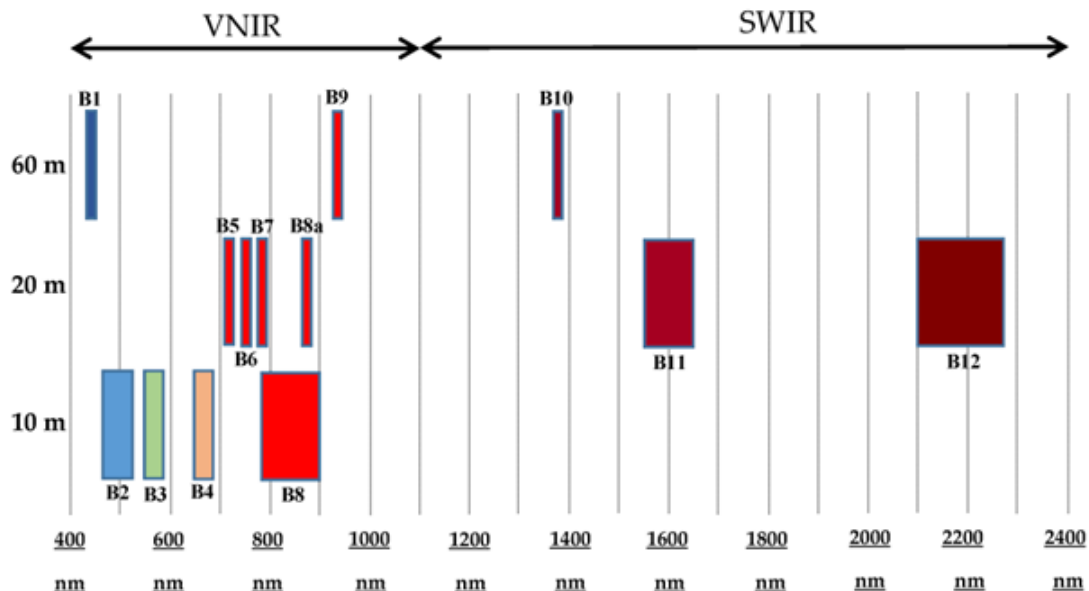


# Kapittel 2

## Teori

### 2.1 Sentinel-2

Sentinel-2 består av to satellitter, Sentinel-2A og Sentinel-2B. Disse satellittene opererer i identiske baner som er forskjøvet med 180 grader i forhold til hverandre. De beveger seg i en solsynkron bane med en inklinasjon på 98,62 grader, som sikrer optimal solbelysning av Jorden med minimal skyggelegging. Omløpstiden for satellittene er 10 dager ved ekvator, og de opererer i en høyde av  $786\text{ km}$ . Satellittene er en del av Copernicus programmet. Dataene som samles inn av disse satellittene er fritt tilgjengelige for publikum. Sentinel-2A ble skutt opp 23. juni 2015 og Sentinel-2B 7. mars 2017 (ESA, 2024c; SIC, 2024).



**Figur 2.1.1:** Sentinel-2 spektralbånd og romlig oppløsning. Bilde hentet fra: (Gascon mfl., 2017)

Sentinel-2 innhenter data ved hjelp av et multispektralt instrument (MSI), som er konstruert for å oppfylle strenge krav til nøyaktighet i både geometriske og spektrale målinger. MSI er utstyrt med tretten spektralbånd som opererer over diverse bølgelengder, en fullstendig oversikt kan ses i tabell 2.1.1. Båndene inneholder kanaler for synlig lys, nær-infrarødt (NIR) og kortbølget infrarødt (SWIR) lys og er designet for å registrere jordens reflekterte stråling. Verdiene fra Sentinel-2 er gitt i digitale nummer (DN).

**Tabell 2.1.1:** Sentral oppløsning og bølgelengde for Sentinel-2A og Sentinel-2B (ESA, 2024d; Sentinel-hub, 2024)

Båndnr.	Båndnavn	Bølgelengde S-2A (nm)	Bølgelengde S-2B (nm)	Oppløsning (m)
1	Coastal aerosol	442.7	442.3	60
2	Blue	492.7	492.3	10
3	Green	559.8	558.9	10
4	Red	664.6	664.9	10
5	Vegetation Red Edge	704.1	703.8	20
6	Vegetation Red Edge	740.5	739.1	20
7	Vegetation Red Edge	782.8	779.7	20
8	NIR	832.8	832.9	10
8a	Vegetation Red Edge	864.7	864.0	20
9	Water vapour	945.1	943.2	60
10	SWIR - Cirrus	1373.5	1376.9	60
11	SWIR	1613.7	1610.4	20
12	SWIR	2202.4	2185.7	20

Satellittbilder har et bredt spekter av bruksområder og dataene egner seg godt til blant annet skog- og vegetasjonsovervåking, isbrekartlegging, naturkatastrofer og areal planlegging (ESA, 2024b). I følge Yan mfl. (2021) har Sentinel-2 bilder vist potensialet i kartlegging av isbreer på grunn av radiometriske egenskaper, lavt støynivå og høy romlig oppløsning.

## 2.2 NDVI

NDVI står for Normalized Difference Vegetation Index og er en av de mest utbredte indeksene for vegetasjonsanalyse. NDVI genererer verdier i intervallet mellom -1 og 1, hvor tall nærmere -1 tyder på fravær av vegetasjon, mens tall nærmere 1 indikerer robust og sunn vegetasjon. Formelen for NDVI, vist i formel 2.1, involverer en kombinasjon av bånd 4 (rød) og bånd 8 (NIR). Sunn vegetasjon inneholder klorofyll som absorberes

sterkt i det synlige lyset blå og rød, samtidig som det reflekteres godt i det nær-infrarøde spekteret og grønt synlig lys (GISGeography, 2017).

$$NDVI = \frac{NIR - Red}{NIR + Red} \quad (2.1)$$

## 2.3 GNDVI

GNDVI er en anerkjent vegetasjonsindeks som benyttes for å estimere fotosyntetisk aktivitet og for å vurdere vann- og nitrogenopptaket i plantenes krone. GNDVI står for Green Normalized Difference Vegetation Index og ble utviklet av Gitelson mfl. (1996).

I følge utvikleren viste GNDVI seg å være mer sensitiv enn NDVI for å identifisere variasjoner i klorofyllkonsentrasjoner, noe som er tett knyttet til nitrogeninnholdet. Bruken av det grønne spektralbåndet har vist seg å være mer effektivt enn det røde spektralbåndet for å skille mellom ulike nivåer av nitrogen (ESA, 2024a; Esri, 2024b).

$$GNDVI = \frac{NIR - Green}{NIR + Green} \quad (2.2)$$

## 2.4 BSI

Bare Surface Index (BSI) er en indeks som kombinerer spektralbåndene blå, rød, nær-infrarød (NIR) og kortbølget infrarød (SWIR) for å detektere variasjoner i jordoverflaten. BSI brukes til å forstå mengden av bart jordsmonn i et gitt område og er spesielt nyttig i fjernmålingsanalyse. Formelen for BSI er beskrevet i formelen 2.3.

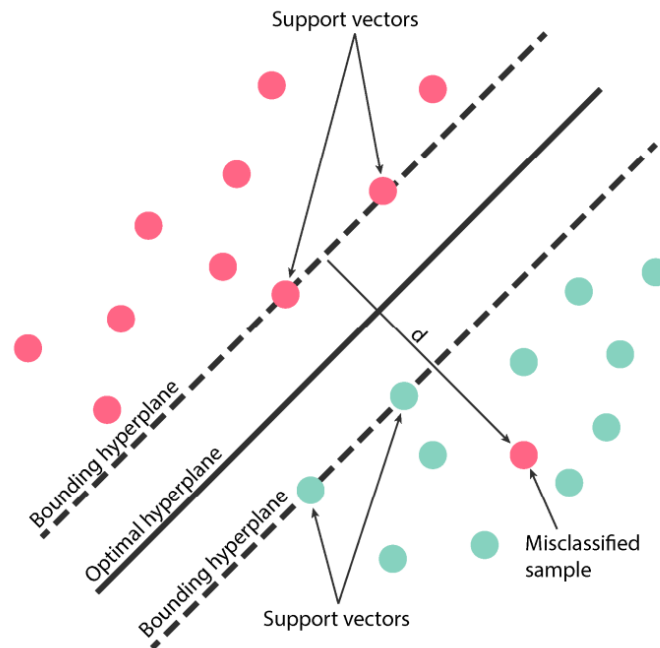
$$BSI = \frac{(SWIR + Red) - (NIR + Blue)}{(SWIR + Red) + (NIR + Blue)} \quad (2.3)$$

Høyere BSI-verdier indikerer en større forekomst av bart jordsmonn, mens lavere verdier tyder på mindre eksponering av jordsmonn. Anvendelsesområder for BSI er blant annet jordsmonn kartlegging og avlingsidentifikasjon, ofte i kombinasjon med NDVI (Sykas, 2024).

## 2.5 Support Vector Machine

Support Vector Machine (SVM) er en kraftig maskinlæringsalgoritme som brukes i klassifisering. Den er spesielt kjent for sin evne til å håndtere høydimensjonale data og for sin effektivitet i situasjoner hvor antall egenskaper overstiger antall datapunkt. SVM-algoritmen er designet for å håndtere slike situasjoner effektivt, og den er i stand til å finne en beslutningsgrense, også kaldt hyperplan, som best skiller de forskjellige klassene. Dette er en av grunnene til at SVM er en populær metode i maskinlærning for både klassifisering og regresjon.

SVM fungerer ved å finne et hyperplan i et flerdimensjonalt rom som best skiller de forskjellige klassene. I to dimensjoner kan dette hyperplanet visualiseres som en linje som deler et plan i to deler, hvor hver klasse ligger på sin side av linjen. SVM prøver å optimalisere marginen mellom de nærmeste punktene i hver klasse og hyperplanet (Patel, 2017).

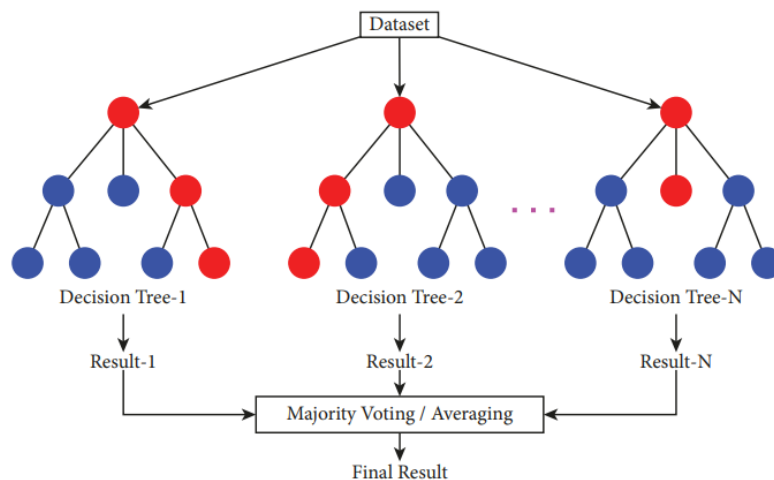


**Figur 2.5.1:** SVM finner det optimale hyperplanet som fungerer som en skillelinje mellom klassene. Hyperplanet er parallelt med de begrensende hyperplanene der støttevektorene ligger. Marginen er definert som avstanden mellom disse to begrensende hyperplanene. Illustrasjon hentet fra: (Cardoso-Fernandes mfl., 2020)

## 2.6 Random Trees

Random Trees (RT) er en maskinlæringsalgoritme som er kjent for sin robusthet mot overtilpasning. I tillegg har denne algoritmen evnen til å behandle segmenterte bilder og andre supplerende rasterdatasett. RT består av individuelle beslutningstrær, som er generert fra varierte prøver fra treningsdataene. Hvert tre utfører en rekke av beslutninger i en rangert rekkefølge av betydning for hver piksel som skal klassifiseres.

Navnet RT kommer fra mønsteret som klassifiseringen etterlater. Stien til en piksel ligner en gren, og når alle pikslene er klassifisert, danner grenene et komplett tre (Esri, 2024e).



**Figur 2.6.1:** RT bruker de ulike beslutningstrær til å bestemme det endelige resultatet. Illustrasjon hentet fra: (M. Y. Khan mfl., 2021)

## 2.7 Iso Cluster Unsupervised Classification

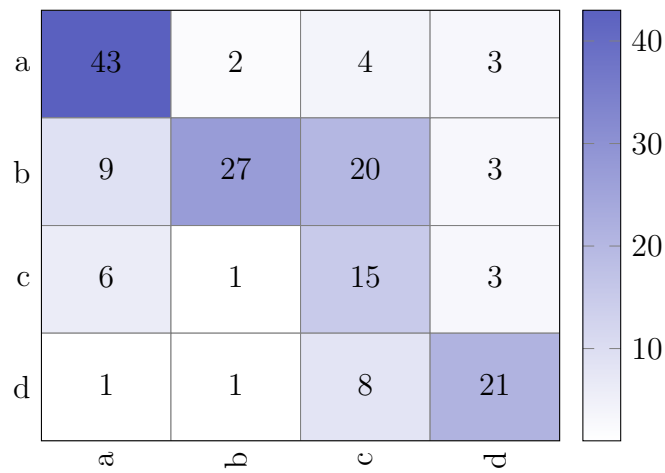
Iso Cluster Unsupervised Classification er en algoritme innen fjernmåling og bildeanalyse som brukes til å klassifisere data uten forhåndsbestemte klasser. Denne metoden benytter seg av en iterativ prosess for å gruppere data i klasser basert på deres spektrale signatur.

I denne metoden blir datasettets elementer systematisk tildelt til eksisterende klyngesentre gjennom hver iterasjon, hvor nye gjennomsnitt beregnes på nytt for hver klasse. Det ideelle antall klasser som skal spesifiseres, er ofte ikke forhåndsbestemt (Esri, 2024c).

For å oppnå en uspesifisert klassifisering, kombinerer algoritmen funksjonalitetene til Iso Cluster og Maximum Likelihood Classification. Resultatet er et klassifisert raster som representerer de forskjellige klassene som er identifisert. Med Iso Cluster Unsupervised Classification, kan man klassifisere landskap som jordområder, skog, vann og urbaniserte områder (Esri, 2024d).

## 2.8 Valideringsparametere

I en evalueringen av klassifiseringsmodeller er anvendelsen av en forvirringsmatrise mye benyttet. Denne matrisen er et sentralt verktøy innen maskinlæring for å bedømme effektiviteten til en klassifiseringsalgoritme. Forvirringsmatrisen illustrerer antallet korrekte og feilaktige prediksjoner, innenfor hver klasse, som tydelig visualiserer algoritmens prestasjon. Matrisen består av rader og kolonner tilsvarende antallet klasser i datasettet. Radene indikerer de predikerte verdiene, mens kolonnene representerer de faktiske sanne verdiene. En illustrasjon av en forvirringsmatrise for et datasett med fire klasser er presentert i figur 2.8.1.



**Figur 2.8.1:** Eksempel på en flerdimensjonal forvirringsmatrise.

Resultatene fra forvirringsmatrisen kan benyttes til å evaluere klassifiseringsmodellen ved å utføre diverse beregninger. De mest brukte beregningene inkluderer total nøyaktighet, presisjon, gjenkall og F1-verdi, som alle gir et mål på metodens ytelse (Hoang mfl., 2009).

Forvirringsmatrisen inneholder følgende punkter:

- Sann positiv (SP): Antall riktige positive tilfeller.
- Falsk positiv (FP): Antall feilaktig positive tilfeller.
- Sann negativ (SN): Antall riktige negative tilfeller.
- Falsk negativ (FN): Antall feilaktig negative tilfeller.

		Forventet	
		Positiv	Negativ
Predikert	Positiv	SP	FP
	Negativ	SN	FN

**Figur 2.8.2:** Den generelle strukturen i en forvirringsmatrise. Illustrasjon inspirert av (Ahmed, 2023)

Nøyaktigheten beregnes ved å dele antall sanne positive prediksjoner med det totale antall observasjoner. Dette resultatet gir en overordnet indikasjon på klassifiseringsmodellens ytelse.

$$N\acute{o}yaktighet = \frac{SP}{SP + FP + SN + FN} \quad (2.4)$$

Presisjon sier noe om modellens evne til å korrekt klassifisere sanne verdier. Denne verdien beregnes ved å dele antall sanne positive verdier med summen av alle predikerte verdier. Dette gir et mål på modellens nøyaktighet når det gjelder å forutsi positive resultater. Gjenkall representerer andelen av de sanne positive verdiene som modellen korrekt har identifisert som positive. Gjenkall beregnes ved å dele antall sanne positive verdier med summen av alle forventede positive verdier. Dette gir et mål på modellens evne til å identifisere og klassifisere positive resultater korrekt (Hoang mfl., 2009).

$$Presisjon = \frac{SP}{SP + FP} \quad (2.5)$$

$$Gjenkall = \frac{SP}{SP + FN} \quad (2.6)$$

F1-verdien beregnes med hensyn til presisjon og gjenkall og gir en god helhetsvurdering av klassifiseringen. En høy F1-verdi indikerer at klassifiseringsmodellen har oppnådd en lav forekomst av både falske positive og falske negative resultater, noe som betyr at modellen har en høy grad av nøyaktighet.

$$F1 - verdi = 2 * \frac{presisjon * gjenkall}{presisjon + gjenkall} \quad (2.7)$$

Alle tre beregningene resulterer i en verdi mellom 0 og 1, hvor 1 indikerer best ytelse (Raschka & Mirjalili, 2019).

## Studieområde og material

### 3.1 Studieområde

Det valgte studieområdet for denne oppgaven er Hardangerjøkulen, som er den sjette største isbreen i Norge. Hardangerjøkulen ligger i Vestland fylke, i nærheten av Finse. Isbreen har for tiden et areal på omtrent  $64\text{km}^2$ , en maks høyde på  $1856\text{ moh}$  og består av syv sentrale brearmer som strekker seg i alle himmelretninger (Andreassen, 2022; Weber mfl., 2019).

Fra nordvest og med klokka finner vi følgende brearmer: Midtdalsbreen, Blåisen, Torsteinsfonna, Austra Leirebottsskåka, Vestra Leirebottsskåka, Isdøleskåka og Rembesdalskåka. Terrenget rundt breen er åpent og med lite eller ingen vegetasjon.





**Figur 3.1.1:** Oversikt over plasseringen til Hardangerjøkulen med brearmene Midtdalsbreen (MB), Blåisen (BI) og Rembesdalskåka (RS).

Denne studien undersøker hele Hardangerjøkulen, men gjennomfører også analyser kun på tre av brearmene: Midtdalsbreen, Blåisen og Rembesdalskåka, hvor Rembesdalskåka er den største utløperen (Liestøl, 1963). Lengden på disse tre brearmene, samt deres areal under lille istid maksimale utbredelse, kan ses i tabell 3.1.1. Data for de resterende brearmene på Hardangerjøkulen kan studeres i artikkelen til Paul Weber (Weber et al., 2019).

**Tabell 3.1.1:** Senterlinje lengde og areal for brearmene Midtdalsbreen, Blåisen og Rembesdalskåka under den lille istids maksimale utbredelse. Kilde: (Weber mfl., 2019).

BreID	Brenavn	Lengde (m)	Areal (km <sup>2</sup> )
2964	Midtdalsbreen	5971	10.9
2966	Blåisen	5384	10.9
2968	Rembesdalskåka	11027	20.7

## 3.2 Datagrunnlag

Oppgaven benytter fire typer datasett: satellittbilder fra Sentinel-2, ortofoto, breomkrets for LIA, 1920 og 2013, samt rebasseng for Midtdalsbreen, Blåisen og Rembesdalskåka.

### 3.2.1 Ortofoto

Ortofotoet ble brukt til å generere treningsdata for klassifiseringen og som støtte for validering av resultatene, noe som bidrar til å sikre at klassifiseringen er nøyaktig og pålitelig. Ortofotoet er hentet fra Norgebilder.no (2024) og ble fløyet sensommeren 2019 av Terratec AS, på oppdrag fra Statens kartverk. Referansesystemet er gitt i Euref89, UTM sone 32N med høydegrunnlag NN2000. Oppløsningen er på  $0.25m$ .

### 3.2.2 Sentinel-2 satellittbilder

Satellittbildene fra Sentinel-2 ble brukt for å analysere reflektansverdiene i hvert enkelt bånd, samt for å studere båndkombinasjoner, indekser og klassifisering. Sentinel-2 datasettet for denne oppgaven ble tatt den 27.08.2019. Dette datasettet er levert av NVE, med fem utvalgte bånd allerede inkludert. Oppløsningen er på  $10m$  for båndene rød (4), grønn (3), blå (2) og NIR (8), og  $20m$  for båndet SWIR (11). Utvalget av bånd ble gjort basert på båndenes egenskaper, som opererer innenfor det synlige lyset, nær infrarødt og kortbølget nærinfrarødt. Sentinel-2 data er gitt i referansesystemet WGS84.

I følge en artikkel på Geodata sin nettside, (Esri, 2024a), vil en konvertering til EUREF89 ikke ha noen innvirkning på datasettet. Esri har bestemt at transformasjonen skal settes til 0, da transformasjonen mellom disse to datumene vil være mindre nøyaktig enn utgangspunktet. Forskjellen mellom EUREF89 og WGS84 ligger på  $\pm 40cm$ , sammenlignet med transformasjonen som har en nøyaktighet på  $\pm 1m$ . Derfor ble det bevisst besluttet å ikke konvertere satellittdatasettet.

### 3.2.3 Breomkrets og brebasseng

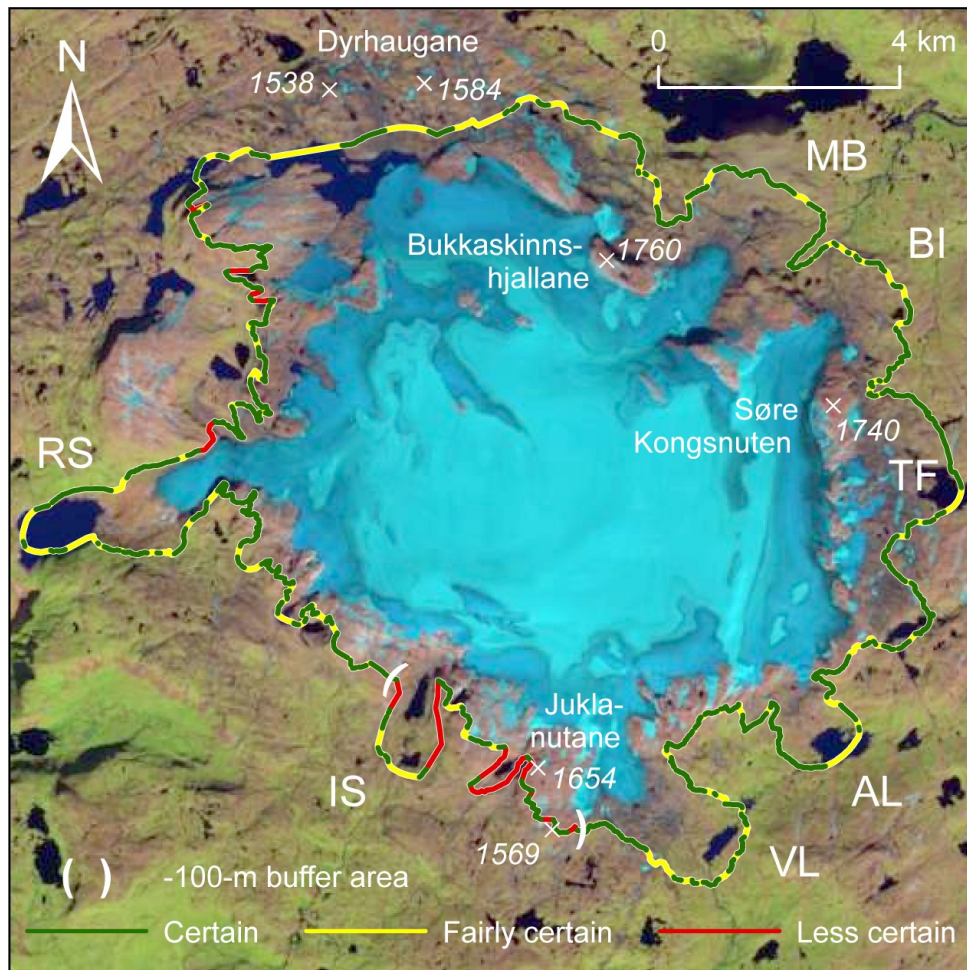
Breomkretsen til LIA er et sentralt datasett for oppgaven og har gjennom hele oppgaven blitt brukt som fasit. Sammen med breomkretsen til LIA, bidro breomkretsene fra 1920 og 2013 inn i analysene som skiller for ulike tidsepoker. Omkretsene ble benyttet til å dele opp datasettet inn i polygoner og linjer, samt til evaluering av resultatene. Videre ble brebassengene for Midtdalsbreen, Blåisen og Rembesdalskåka benyttet til utklipp av de gitte brearmene for individuell analysering.

I 2016 og 2017 ble det utført en manuell kartlegging av utbredelsen til LIA på Hardangerjøkulen. Resultatet er illustrert i figur 3.2.1, som viser den kartlagte omkretsen av LIA. Omkretsen er representert med fargene grønn, gul og rød, som indikerer graden av sikkerhet i kartleggingen.

Den totale omkretsen av LIA er på  $82,4km$ , hvorav  $59,7\%$  er klassifisert som sikker (grønn),  $31,5\%$  som ganske sikker (gul) og de resterende  $8,8\%$  som mindre sikker (rød). Den sikre

delen av kartleggingen vil bli brukt videre i metoden.

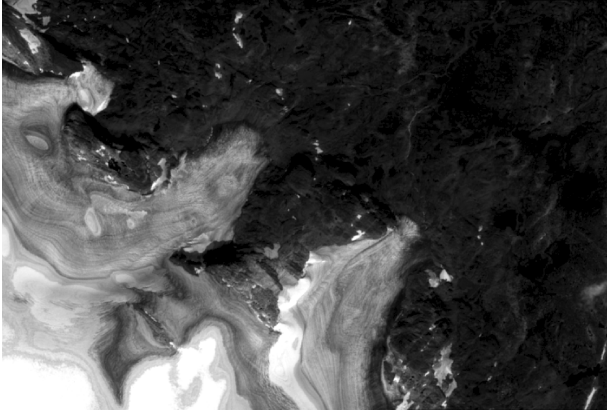
Rekonstruksjonen av LIA og brebassengene ble utført av Paul Weber og er publisert i 2019 (Weber et al., 2019).



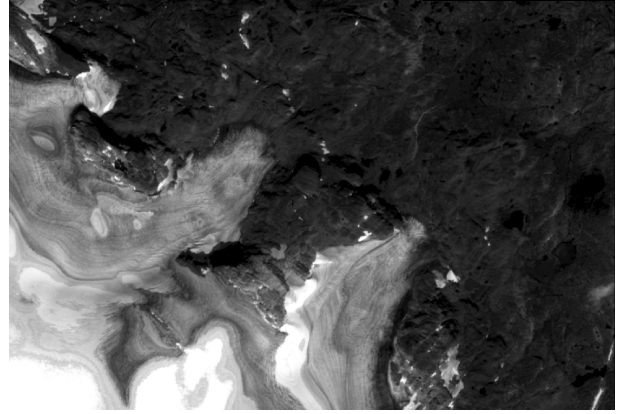
**Figur 3.2.1:** Rekonstruksjon av LIA fra 2016 og 2017 gjort av Paul Weber, (Weber mfl., 2019). Her ser man også brearmene Midtdalsbreen (MB), Blåisen (BI), Torsteinsfonna (TF), Austra Leirebottsskåka (AL), Vestra Leirebottsskåka (VL), Isdøleskåka (IS) og Rembedalskåka (RS)

### 3.3 Utvalgte Sentinel-2 bånd

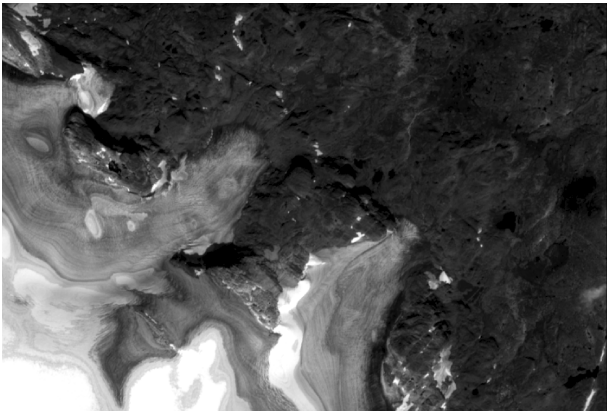
De valgte båndene fra Sentinel-2 er illustrert nedenfor over brearmene Midtdalsbreen og Blåisen. Bølgelengdene som er vist i de følgende figurene er hentet fra (ESA, 2024d). Hvert av de spektrale båndene i Sentinel-2 inneholder unike egenskaper som et direkte resultat av deres individuelle bølgelengder (GISGeography, 2014; Sinergise, 2024b).



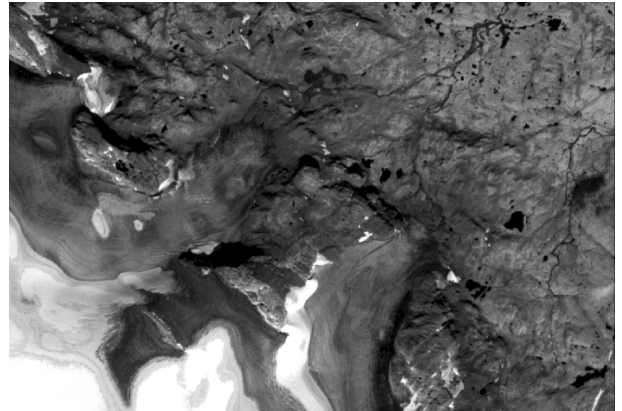
**Figur 3.3.1:** Bånd 2: blå, 492.7nm, GSD 10m.



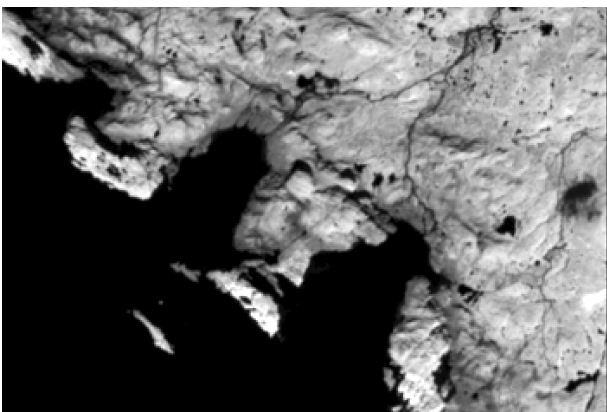
**Figur 3.3.2:** Bånd 3: grønn, 559.8nm, GSD 10m.



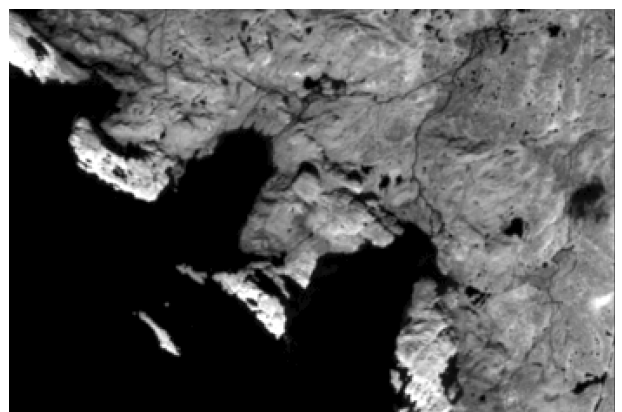
**Figur 3.3.3:** Bånd 4: rød, 664.6nm, GSD 10m.



**Figur 3.3.4:** Bånd 8: NIR, 832.8nm, GSD 10m.



**Figur 3.3.5:** Bånd 11: SWIR, 1613.7nm, GSD 20m.



**Figur 3.3.6:** Bånd 12: SWIR, 2202.4nm, GSD 20m.

Bånd 2, illustrert i figur 3.3.1, ligger innenfor det synlige lysspekteret og reflekterer blått lys. Dette kortbølgede båndet har en unik evne til å fremheve objekter som befinner seg i skyggen. I tillegg har dette båndet gunstige egenskaper for kartlegging av jordtyper og identifisering av menneskeskapte objekter. Det er også i stand til å trenge gjennom klart vann effektivt. Denne bølgelengden absorberes av klorofyll, noe som resulterer i at planter fremstår som mørke.

Bånd 3, illustrert i figur 3.3.2, opererer også innenfor det synlige lysspekteret og reflekterer grønt lys. Dette båndet er spesielt nyttig for å skille ut kontrastene i klart og grumsete vann. Likt Bånd 2, kan dette båndet også identifisere menneskeskapte objekter.

Bånd 4, vist i figur 3.3.3, reflekterer rødt lys og er et av de mest brukte båndene i kartlegging. Dette båndet er viktig for å identifisere ulike vegetasjonstyper, jordsmonn og byområder. Vegetasjon absorberer dette lyset, og død vegetasjon kommer tydelig frem.

Bånd 8, vist i figur 3.3.4, opererer innenfor det nær-infrarøde spekteret og er mye brukt til klassifisering av vegetasjon. I dette spekteret reflekteres klorofyllet i plantene godt, og det er dermed et nyttig bånd for å analysere sunnheten i vegetasjonen.

Bånd 11 og 12, illustrert i figur 3.3.5 og 3.3.6, ligger innenfor det kortbølgede infrarøde spekteret og benyttes mye til kartlegging av isbreer. I dette spekteret blir vann og snø absorbert og blir derfor fremvist som mørk (GISGeography, 2014; Sinergise, 2024b).

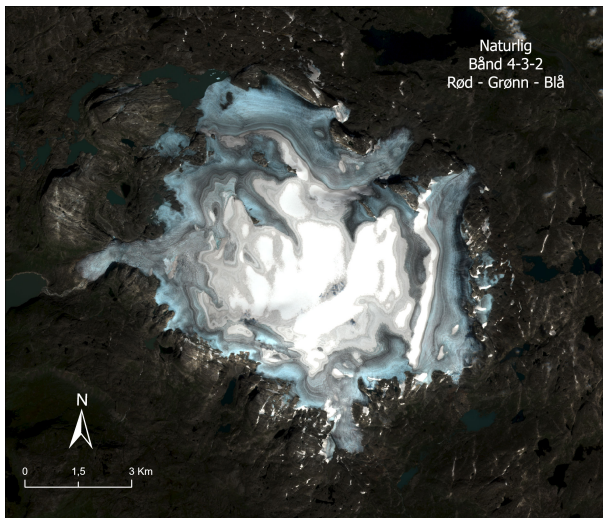
## 3.4 Båndkombinasjoner og indekser

Satellittbåndene er en god kilde til informasjon som kan kombineres på forskjellige måter for å utføre visuell analyse. I figurene 3.4.1, 3.4.2 og 3.4.3 presenteres tre mye brukte kombinasjoner for en slik analyse.

I tillegg til båndkombinasjoner, blir indekser også mye brukt i analyser. I figurene 3.4.4, 3.4.5 og 3.4.6 vises tre forskjellige indekser over Hardangerjøkulen: NDVI, GNDVI og BSI. Disse indeksene gir verdifull innsikt i forskjellige aspekter av landskapet.

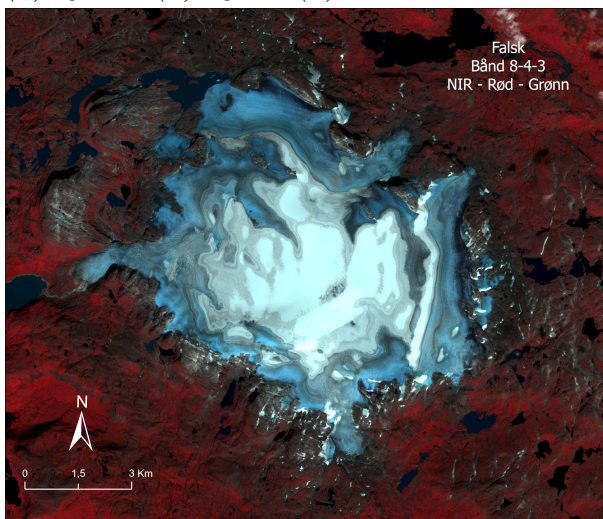
Både båndkombinasjonene og indeksene vil bli brukt videre i denne studien for å analysere Hardangerjøkulen og utbredelsen av LIA.





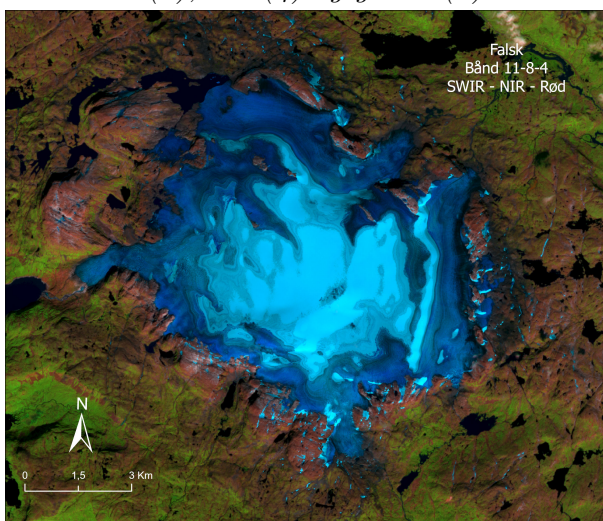
I denne båndkombinasjonen er det benyttet tre bånd som tilhører det synlige spekteret: rød (4), grønn (3) og blå (2). Denne kombinasjonen gir en visuell fremstilling av terrenget slik det ville sett ut for det menneskelige øyet. Terrengområdene fremstår i nyanser av brun og grønn, mens isbreen vises i lys blå. Snøen er hvit og vannet fremstår i mørke toner. Dette skyldes at snø reflekterer lys i det synlige spekteret og absorberer lys i det kortbølgede infrarøde (SWIR) spekteret.

**Figur 3.4.1:** Naturlig bilde fremstilt i rød (4), grønn (3) og blå (2).



Falsk nær-infrarød båndkombinasjon er en effektiv metode for å analysere vegetasjonen i et gitt område. Frisk vegetasjon, som inneholder mye klorofyll, vil fremstå i en tydelig rød farge. På den andre siden vil mindre frisk vegetasjon fremstå i en dypere rød farge. Allerede her er det mulig å få et inntrykk av hvor LIA sin utbredelse har vært.

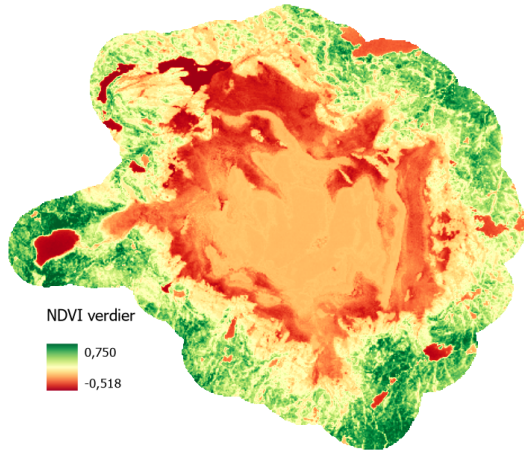
**Figur 3.4.2:** Falsk bilde fremstilt med båndene NIR (8), rød (4) og grønn (3).



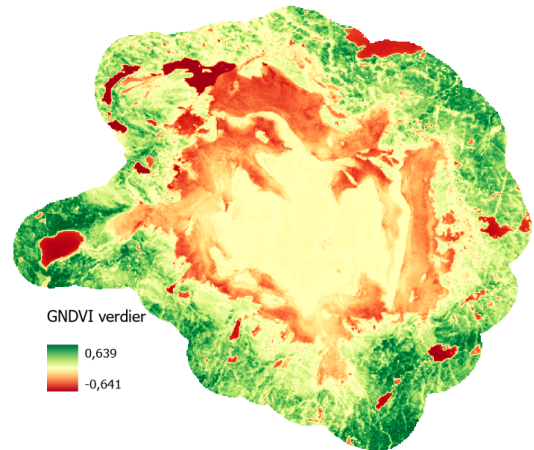
Denne båndkombinasjonen, som inkluderer SWIR (11), er også en effektiv metode for å analysere vegetasjon. Inkluderingen av SWIR gir et bredere informasjonsspekter fra fjell, stein og berg. Sunn vegetasjon fremstår tydelig i lyse grøntoner, mens områder med mindre vegetasjon, hvor berggrunnen er synlig, blir representert i bruntoner.

**Figur 3.4.3:** Falsk bilde fremstilt i SWIR (11), NIR (8) og rød (4).

NDVI og GNDVI er begge vegetasjonsindekser. Disse sier noe om hvor sunn vegetasjonen i området er. Den yngste vegetasjonen vil være nærmest breen og det vil derfor være forventet lavere innhold av klorofyll og dermed lavere indeks verdi. Som det kommer frem i figurene 3.4.4 og 3.4.5, ser man tydelig mindre sunn vegetasjon nærmest breen.

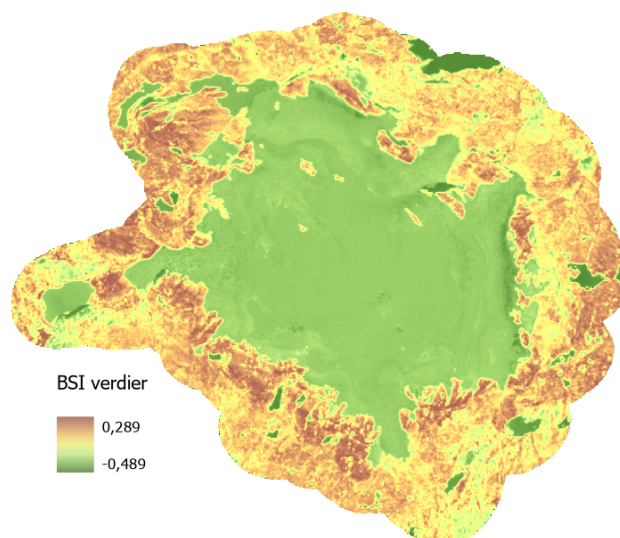


**Figur 3.4.4:** *NDVI*



**Figur 3.4.5:** *GNDVI*

BSI er en jordsmonnsindeks, som brukes til jordsmonnskartlegging og avlingsidentifikasjon (Sykas, 2024). Bart fjell og steiner kommer tydelig frem i en typ rød/brun farge. Vegetasjonen varierer fra gul til grønn utfra hvor mye som er tilstedet. Snø, isbre og vann blir også fremvist i mørkere grøntoner.



**Figur 3.4.6:** *BSI*

BSI blir ofte anvendt innen fjernmåling sammen med indekser som NDVI og GNDVI. Det er viktig å understreke at resultatene fra indeksene kan variere med årstiden bildene

blir tatt. Dette betyr at sesongvariasjoner kan ha innvirkning på indeksverdiene, ettersom vegetasjonens tilstand endre seg over tid (Sinergise, 2024a).

## 3.5 Programvarer

ArcGIS Pro er en GIS-programvare som er utviklet av Esri. Programvaren er kan brukes til å visualisere og analysere i både 2D og 3D, med mulighet for å dele arbeidet offentlig via ArcGIS Online (ESA, 2024b).

I denne oppgaven brukes ArcGIS Pro til klassifisering og analyse av Hardangerjøkulen.

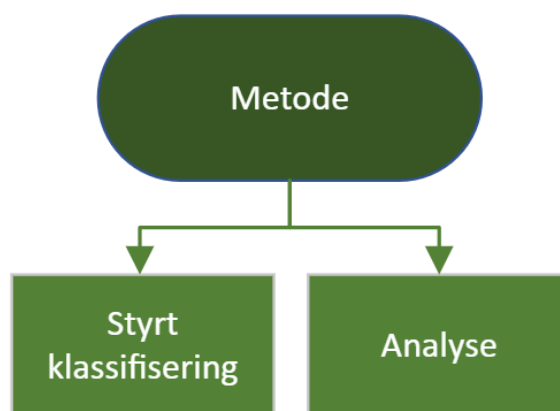
Analyseresultatene fra ArcGIS Pro ble overført til Excel for videre analysering. Dette inkluderte beregninger som forvirringsmatrise, varians, korrelasjon og deskriptiv statistikk, som gir en oppsummering av datasettets sentrale trender, spredning og form. I tillegg ble det laget histogrammer for å visualisere datafordelingen.



## Metode

Kapittel 4 presenterer en gjennomgang av metoden som er benyttet i denne oppgaven. Metoden er delt inn i to hoveddeler, samt en del som ikke gikk helt etter planen. Beskrivelse av hvilke programvarer, funksjoner og parametere som er anvendt i prosessene, kommer frem i påfølgende avsnitt.

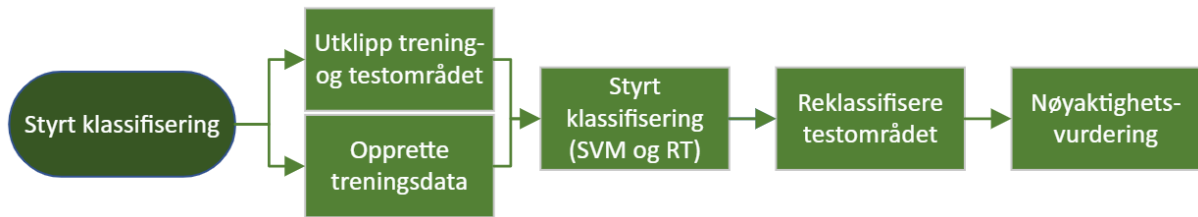
Den overordnede oversikten over metoden vises i flytskjema under.



**Figur 4.0.1:** *Flytskjema for metoden.*

### 4.1 Styrt klassifisering

For å klassifisere brearmene Midtdalsbreen og Blåisen gjennomføres en styrt klassifisering. Den styrte klassifiseringen sin input er datasettet fra Sentinel-2, som består av fem bånd: RGB (4, 3, 2), NIR (8) og SWIR (11). I tillegg ble indeksene NDVI, GNDVI og BSI inkludert. Rembesdalskåka blir brukt som testområde for validering av klassifiseringen.



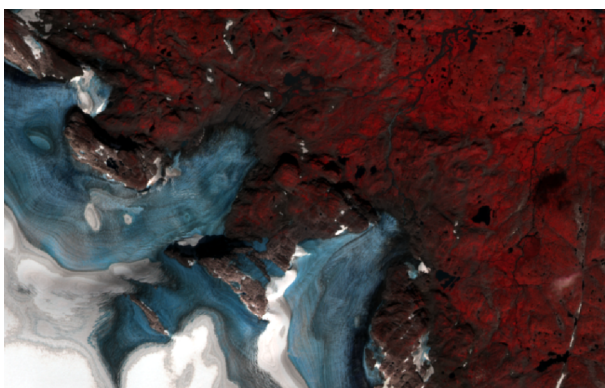
**Figur 4.1.1:** Flytskjema for klassifiseringen.

Det ble gjennomførte to separate klassifiseringer. Den første brukte en Support Vector Machine (SVM), mens den andre benyttet Random Trees (RT). Klassifiseringsmetodene blir sammenlignet under kapittel 6 Diskusjon. Stegene videre forklares nedenfor i samme rekkefølge gitt i flytskjemaet.

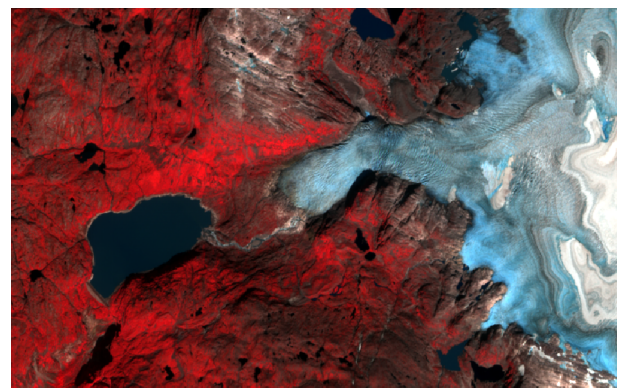
#### 4.1.1 Utklipp av trening- og testområdet

Første trinn i klassifiseringen er å dele datasettet til to mindre områder. Hvert av områdene fungerer henholdsvis som treningsområdet og testområdet. Formålet med denne delingen er todelt: For det første gir det muligheten til å evaluere metoden ved hjelp av testdatasettet og vurdere hvor godt den fungerer på områder utenfor treningsfeltet. For det andre reduserer det størrelsen på datasettene, noe som gjør dem mer håndterbare.

Utklippingen gjennomføres i ArcGIS Pro med funksjonen *Extract by Mask*. Satellittbildet klippes etter et polygon tilpasset Midtdalsbreen og Blåisen (figur 4.1.2) og et polygon tilpasset Rembesdalskåka (figur 4.1.3). Begge utsnittene inneholder godt med data både innenfor og utenfor LIA.



**Figur 4.1.2:** Treningsområdet over Midtdalsbreen og Blåisen fremstilt i båndkombinasjonen NIR(8), rød(4) og grønn(3).



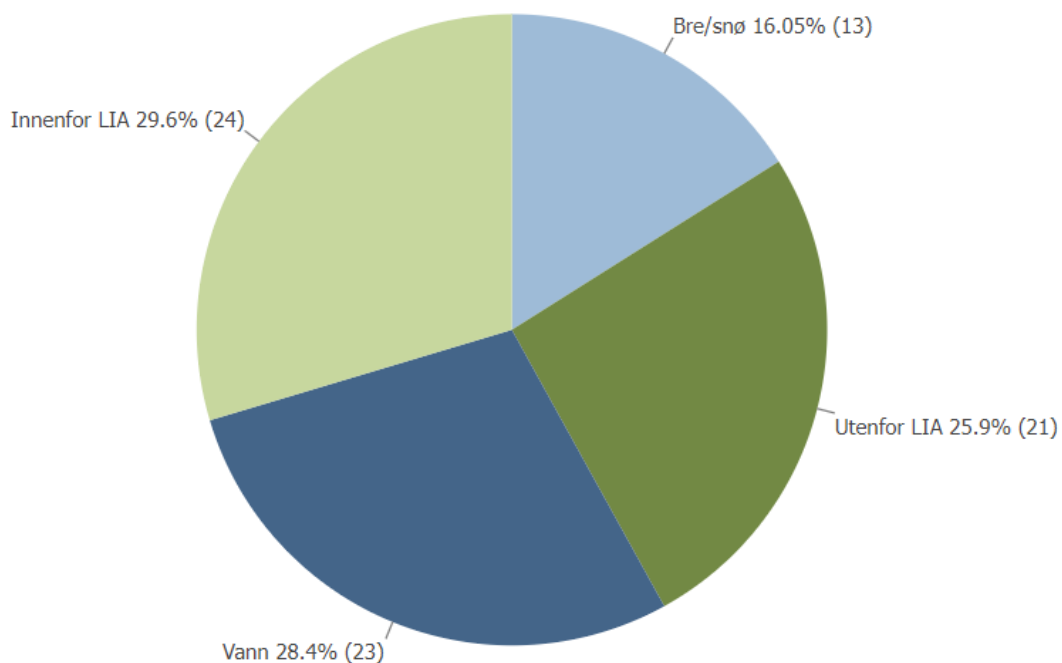
**Figur 4.1.3:** Testområdet over Rembesdalskåka fremstilt i båndkombinasjonen NIR(8), rød(4) og grønn(3).

## 4.1.2 Treningsdata

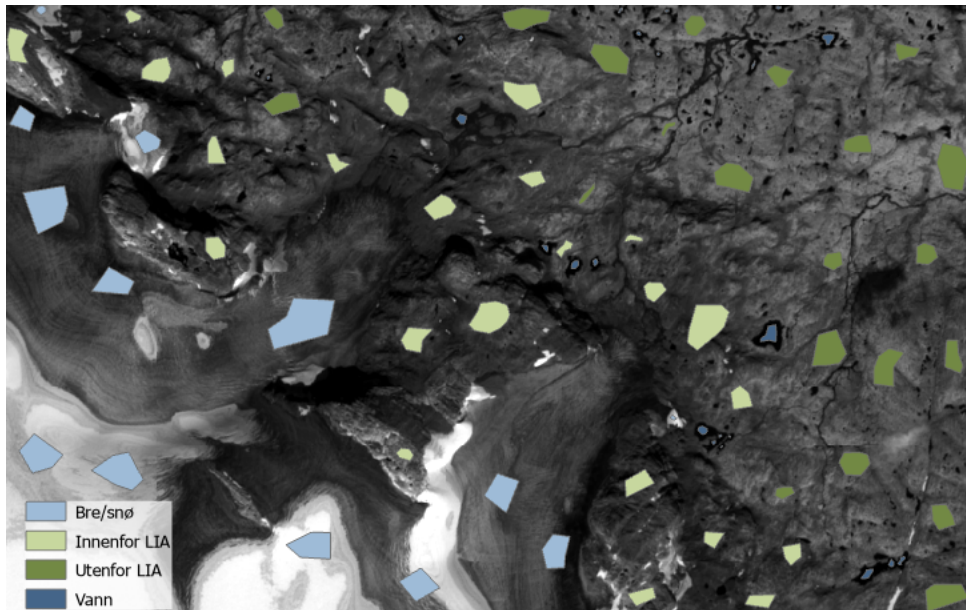
For å trene klassifiseringsmetodene benyttes spektralsignaturen fra områdene i treningsdatasette. Treningsdataen ble utarbeidet i ArcGIS Pro med funksjonen *Training Samples Manager*. Her opprettes et nytt skjema med de klassene man ønsker og deretter lages polygonene. Treningsdataen ble visuelt plukket ut ved hjelp av ortofoto fra Norgebilder.no. I tillegg ble polygonet som representerer LIA brukt som mal for å plassere polygonene i riktig klasse. De fire klassene i treningsdatasettet består av:

- Bre og snø
- Vann
- Innenfor LIA
- Utenfor LIA

Polygonene ble valgt spredt utover området for å sikre variasjon i klassene. Dette bidrar til å lære modellen å håndtere variert terreng og dermed oppnå mer nøyaktig klassifisering. Den totale fordelingen av polygoner er jevnt fordelt i hver klasse, som illustrert i kakediagrammet i figur 4.1.4. Plasseringen av treningsdataen fremgår i figur 4.1.5.



**Figur 4.1.4:** Fordeling av treningsdata representert i prosent og antall polygoner.



**Figur 4.1.5:** *Oversikt over plassering av treningsdata med bånd 8 (NIR) i bakgrunn.*

### 4.1.3 Support Vector Machine

Support Vector Machine (SVM) er en innbygd klassifiseringsmetode i ArcGIS Pro. Denne klassifiseringsmetoden henter inn treningsdata som input, sammen med datasettet man ønsker å klassifisere - i dette tilfellet satellittbildet med 5 bånd og indeksene NDVI, GNDVI og BSI. Videre oppgis det høyeste antallet av treningspunkt innad i en klasse. Som vist i kakediagrammet i figur 4.1.4, har klasse *Innenfor LIA* 24 treningsområdet. Resultatet er et klassifisert bilde av området.

### 4.1.4 Random Trees

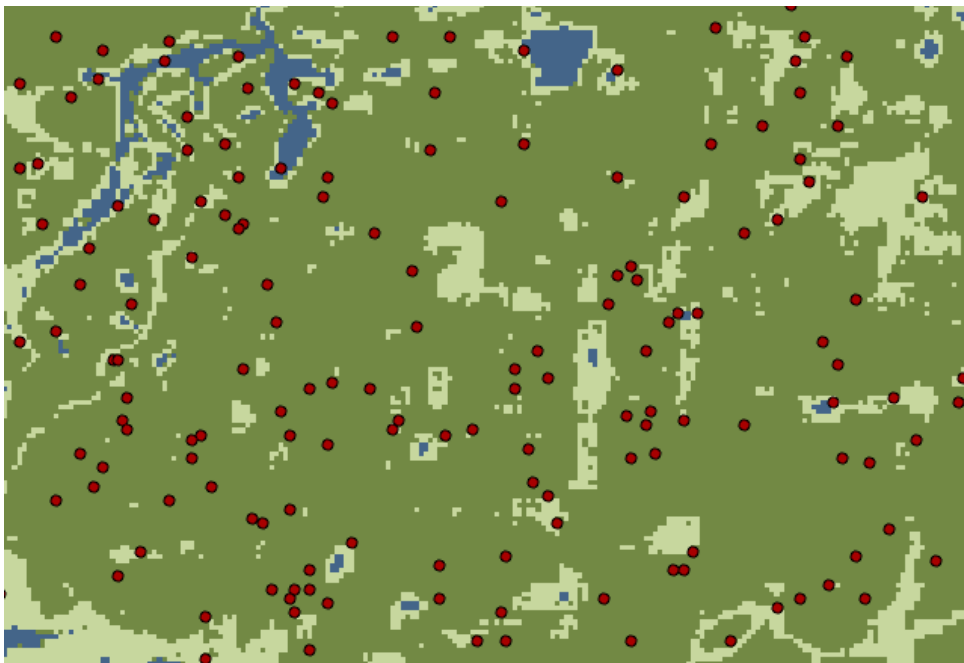
Random Trees (RT) er en annen innebygd klassifiseringsmetode i ArcGIS Pro. Metoden henter inn treningsdata som input og en rekke variabler må velges. I denne klassifiseringen ble maksimum av trær satt til 50, maksimum tredybde til 20 og maksimum antall treningspolygoner innenfor en klasse ble satt til 24. Standardverdiene for antall trær og tredybde ble benyttet, da dette ga best resultat under opptreningen av modellen.

### 4.1.5 Testdata

Etter evaluering av klassifiseringsmetodene ble SVM valgt som den mest effektive modellen for denne type klassifisering. SVM-modellen ble derfor anvendt på testområdet, Rembesdalskåka. For å utføre klassifiseringen på testområdet ble funksjonen *Classify raster* benyttet. Funksjonen tar imot testområdet og det klassifiserte treningsområdet som input. Resultatet er et klassifisert bildet av Rembesdalskåka, basert på opptreningen.

### 4.1.6 Kvalitetsvurdering

Evaluering av klassifiseringer er viktig for å kunne vurderer hvor godt resultatet er. Totalt 2000 tilfeldige punkt fordelt over treningsområdet og 2000 tilfeldige punkt fordelt over testområdet ble opprettet. Basert på ortofoto fikk punktene en sann verdi, som ble vurdert opp mot den klassifiserte verdien.

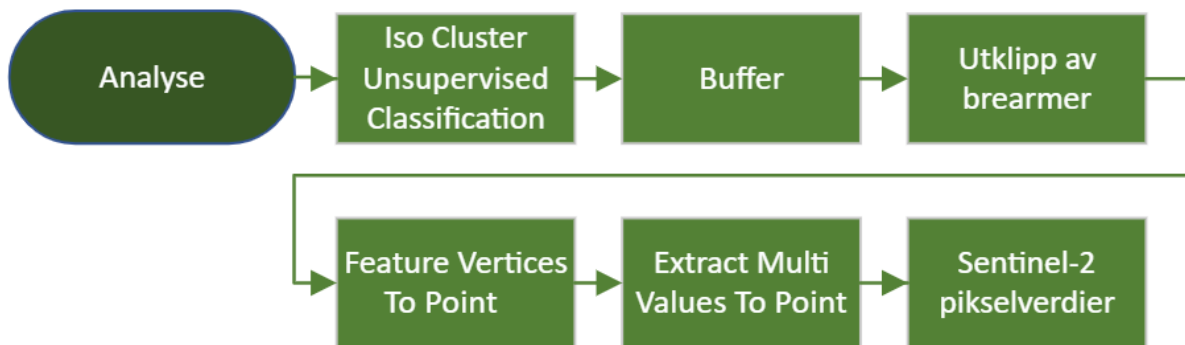


**Figur 4.1.6:** Utsnitt av klassifisert bildet med tilhørende valideringspunkt. Bildet viser den tilfeldige fordelingen til valideringspunktene for evaluering av klassifiseringen.

Basert på punktenes nøyaktighet ble det generert en forvirringsmatrise. Denne matrisen gir innsikt i hvor godt klassifiseringen stemmer overens med virkeligheten. Videre utføres statistisk analyse av resultatene - presisjon, gjenkall og F1-verdi. I tillegg ble de klassifiserte bildene visuelt analysert opp mot ortofoto og breomkretsen til LIA.

## 4.2 Pikkelsbasert analyse

En del av problemstillingen er å undersøke om satellittbilder kan brukes til å detektere lille istids utbredelse på Hardangerjøkulen. For å studere dette et steg nærmere, ble Hardangerjøkulen delt opp i tre ulike polygoner basert på tidligere kartlegginger og fire linjer basert på de sikre områdene til LIA sin morenerygg (se konstruert morenerygg i figur 3.2.1). Innenfor polygonene ble det generert tilfeldige punkt, som videre ble analysert opp mot hverandre. Det samme ble gjort for linjene, samt for tre av brearmene.



**Figur 4.2.1:** Flytskjema for pikkelsbasert analyse.

En nøyere beskrivelse av stegene i flytskjemaet følger under.

### 4.2.1 Iso Cluster Unsupervised Classification

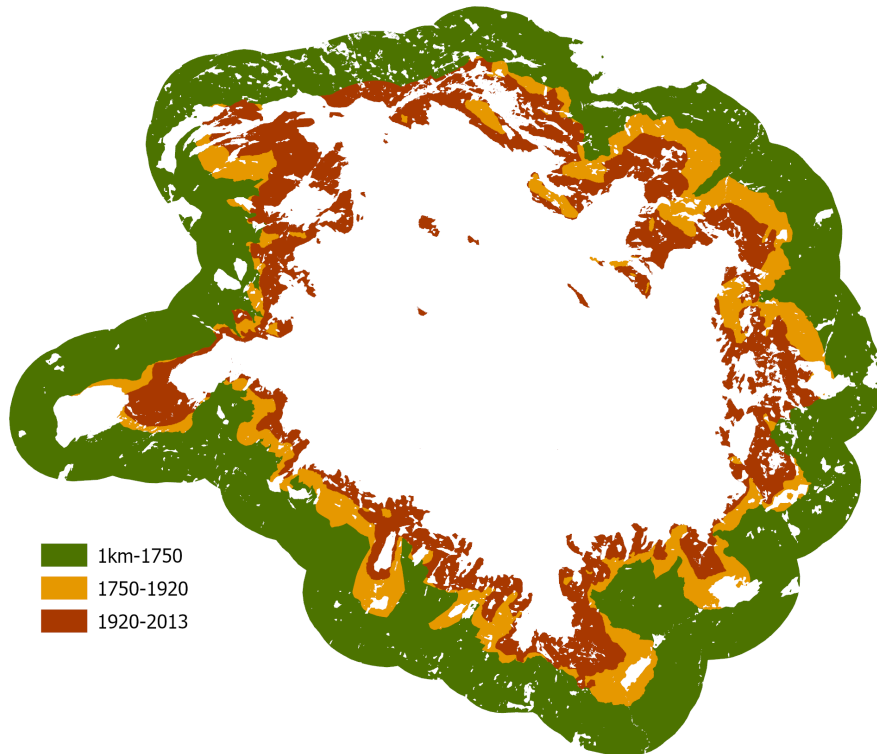
For å kun få verdier basert på fjell og terreng ble klassifiseringsmetoden *Iso Cluster Unsupervised Classification* brukt for å filtrere bort bre, snø og vann. NDVI-datasett ble satt som input og 15 klasser ble valgt. Dette ble gjort for å sikre at algoritmen effektivt skille ut de ulike terrenstypene. Videre ble klassene slått sammen og deretter reklassifisert slik at det ble stående igjen to klasser:

- Klasse 0: Bre, snø og vann
- Klasse 1: Terreng og fjell

Klasse 1 ble lagret til et nytt polygon og benyttet videre til oppdeling av områder og generering av punkter, som beskrevet i påfølgende avsnitt.

## 4.2.2 Buffer

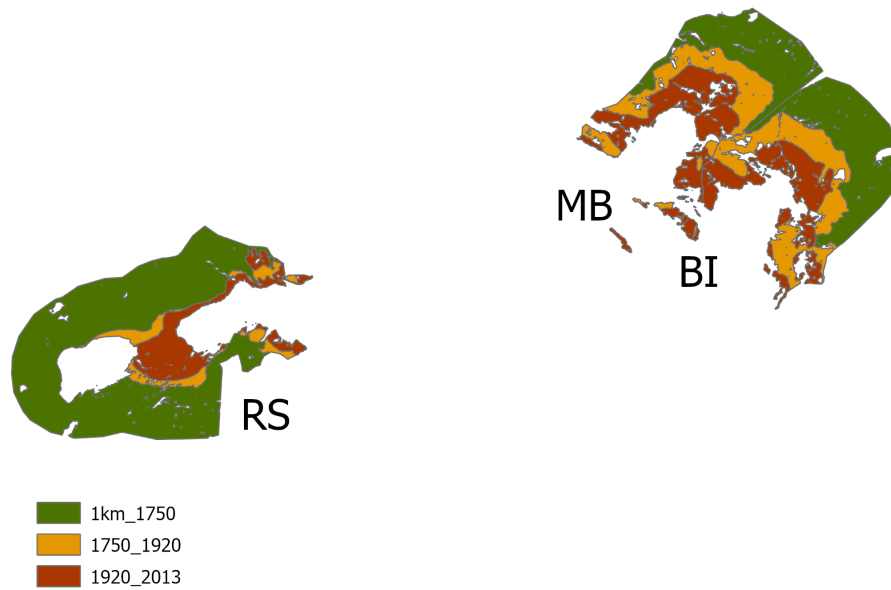
Buffer ble benyttet for å lage linjer basert på de sikre områdene rundt LIA, samt polygoner. Det ble generert linjer på 20m og 100m både på innsiden og utsiden av LIA. I tillegg ble det generert tre buffer-polygoner over hele Hardangerjøkulen i årsskillene 1km utenfor LIA og frem til 1750 (LIA), fra 1750 til 1920 og fra 1920 til 2013. Polygonene er representert nedenfor i figur 4.2.2, samt et utsnitt av linjene mellom Midtdalsbreen og Blåisen i figur 4.2.4.



**Figur 4.2.2:** Tre polygoner delt inn fra 1km utenfor LIA og frem til LIA (1750), fra LIA (1750) til 1920 og innerste polygon fra 1920 til 2013.

## 4.2.3 Utklipp av brearmer

Brebassengene til Midtdalsbreen, Blåisen og Rembesdalskåka ble klippet ut og lagret som nye polygoner. Brearmene vil videre bli sammenlignet med hverandre i den kommende analysen. Utklippingen samt oppdelingen av polygoner kan ses under i figur 4.2.3.

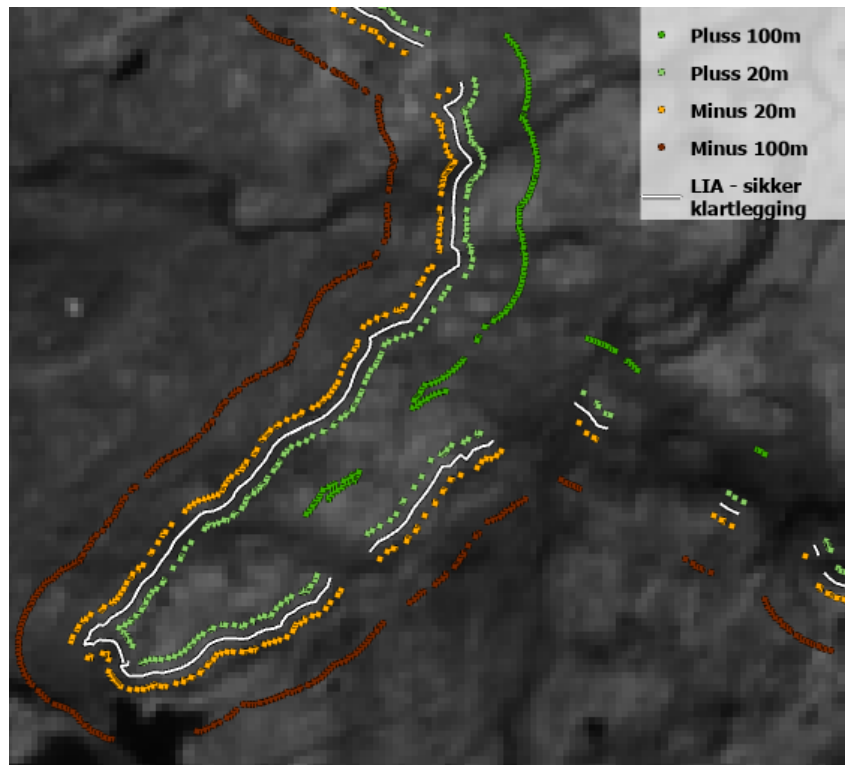


**Figur 4.2.3:** Tre polygoner delt inn fra 1km utenfor LIA og frem til LIA (1750), fra LIA (1750) til 1920 og innerste polygon fra 1920 til 2013, for hver av brearmene Middalsbreen (MB), Blåisen (BI) og Rembesdalskåka (RS).

#### 4.2.4 Feature Vertices To Point

*Feature vertices to point* ble brukt for å opprette punkt i knekklinjene i hver av linjene. I tillegg ble det opprettet 5000 tilfeldige punkt innefor hvert polygon. Noen av punktene ble fjernet manuelt grunnet feil plassering. Punktene som ble fjernet havnet enten rett på utsiden av polygonet eller på toppen av breen hvor det er et par området med fjell. Av den grunn vil et par av polygonene inneholde litt færre punkt.





**Figur 4.2.4:** Utklipp av punkt som ligger 20m og 100m på innsiden og utsiden av den sikre kartleggingen av LIAs morenerygg.

#### 4.2.5 Extract Multi Values To Point

For å gi punktene fra det forrige trinnet en funksjonell betydning, var det nødvendig å tildele verdier til hvert enkelt punkt. Funksjonen *Extract Multi Values To Points* ble benyttet til dette formålet. Denne funksjonen gir hvert punkt verdier basert på de valgte datasettene - satellittbilder med fem bånd og indeksene NDVI, GNDVI og BSI.

#### 4.2.6 Sentinel-2 pikselverdier

For å utføre en grundigere analyse og sammenligning mellom polygonene og mellom linjene, ble de genererte verdiene eksportert og deretter importerte inn i Excel. Denne prosessen muliggjorde utførelsen av mer detaljerte beregninger og statistiske analyser.

Følgende analyser ble utført over hele Hardangerjøkulen, brearmene og de kartlagte sikre punktlinjene:

- **Gjennomsnitt, median, minimum og maksimum:** Disse analysene gir en forståelse av den sentrale trenden og spredningen av verdiene.

- **Variansanalyse:** Analyse for å vurdere variasjonen i attributtverdiene. Denne analysen er med på å identifisere eventuelle signifikante forskjeller.
- **Korrelasjon:** Korrelasjonen gir en innsikt i eventuelle sammenhenger mellom attributtene.

Verdiene fra satellittdatasettene er gitt i digitale nummer (DN). Dette formatet representerer pikselens intensitetsverdi. Hver piksel i et satellittbilde har en tilhørende DN-verdi som gir informasjon om gjennomsnittlig lysstyrke eller reflektans. Disse verdiene er oppskalert fra de opprinnelige reflektansverdiene fra hver piksel. Det er viktig å merke seg at DN-verdiene er enhetsløse og gir oss et mål på pikselens intensitet (Sentinel-hub, 2024). Resultatene av disse beregningene og analysene presenteres i tabellene under kapittel 5 Resultater.

### 4.3 Feilklassifisering

I analysen av Hardangerjøkulen er det viktig å ta hensyn til terrenget rundt breen. Terrenget er preget av åpne flater og minimal vegetasjon. For å styrke signalene fra fjell og stein, var det ønskelig å inkludere bånd 12 (SWIR). Bånd 12 fanger opp refleksjon fra kortbølget infrarødt lys og gir verdifull informasjon om geologiske strukturer og mineraler i området. For å oppnå dette, ble satellittdatasettet nedlastet på nytt. Satellittdata er tilgjengelige for alle gjennom Copernicus sin nettside ([link til nedlasting her](#)). Det nedlastede datasettet har samme tidstempel som det opprinnelige datasettet levert fra NVE.

Klassifiseringen ble gjennomført på lik måte som i avsnitt 4.1 - Styrk klassifisering.

Utklipp av brearmene Midtdalsbreen og Blåisen, samt generere treningsdata var de første stegene. Videre kjørt de to klassifiseringsalgoritmene: Support Vector Machine (SVM) og Random Trees (RT). Algoritmene ble anvendt på Sentinel-2 datasettet bestående av 6 bånd: RGB (4, 3, 2), NIR (8) og SWIR (11, 12) og indeksen NDVI. Resultatet ble et klassifisert bildet basert på satellittdata.

Ved en grundig undersøkelse av resultatene, ble det observert at det klassifiserte datasettet var lagret som en 16-bit usignert datatype. Denne datatypen tillater kun positive verdier, noe som kan påvirke klassifiseringens nøyaktighet ved å overse manglende data eller skydekke. Videre, etter at satellittdatasettet ble slått sammen med NDVI, ble det observert at indeksverdien endret seg til enten 1 eller 0. Dette resulterte i at indeksen ikke bidro i klassifiseringen.

Etter betydelig tid og innsats i denne retningen, ble det besluttet å avslutte arbeidet med det selvnedlastede datasettet. I stedet ble det valgt å starte på nytt med det forhåndsprosesserte datasettet levert fra NVE.

# Kapittel 5

## Resultater

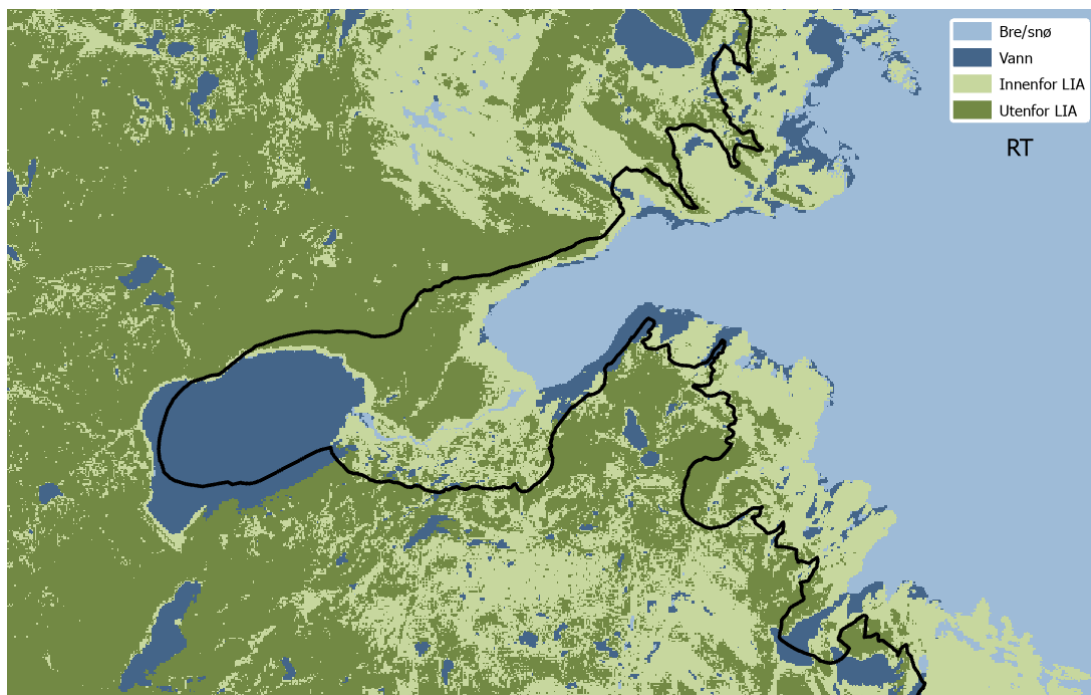
Resultatene i dette kapitlet presenterer analysene fra kapittel 4 Metode. Første del av resultatene representerer den styrte klassifiseringen gjennom SVM og RT med tilhørende statistikk. Andre del av resultatene representerer tilfeldig utvalgte punkt forelt i polygoner (1km-1750, 1750-1920 og 1920-2013) og langs fire linjer ( $\pm 20m$  og  $100m$  fra LIA) og dens tilhørende statistikk.

### 5.1 Klassifisering av LIA utbredelse

Figurene 5.1.1 og 5.1.2 viser det visuelle resultatet av klassifisering basert på SVM og RT. Klassifiseringene er presentert sammen med den kartlagte moreneryggen til LIA.



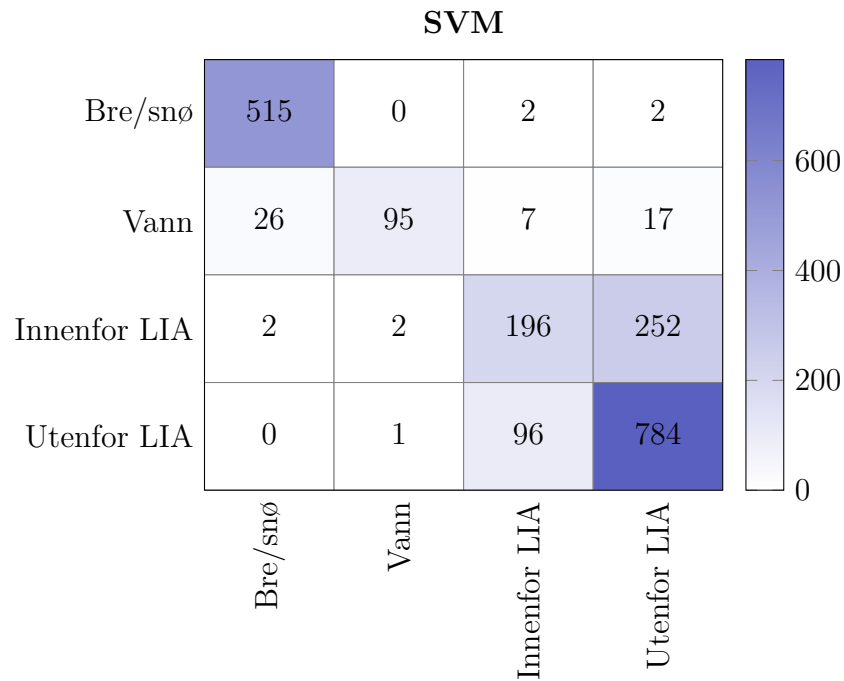
**Figur 5.1.1:** Resultatet av klassifisering med SVM på testområdet Rembesdalskåka.



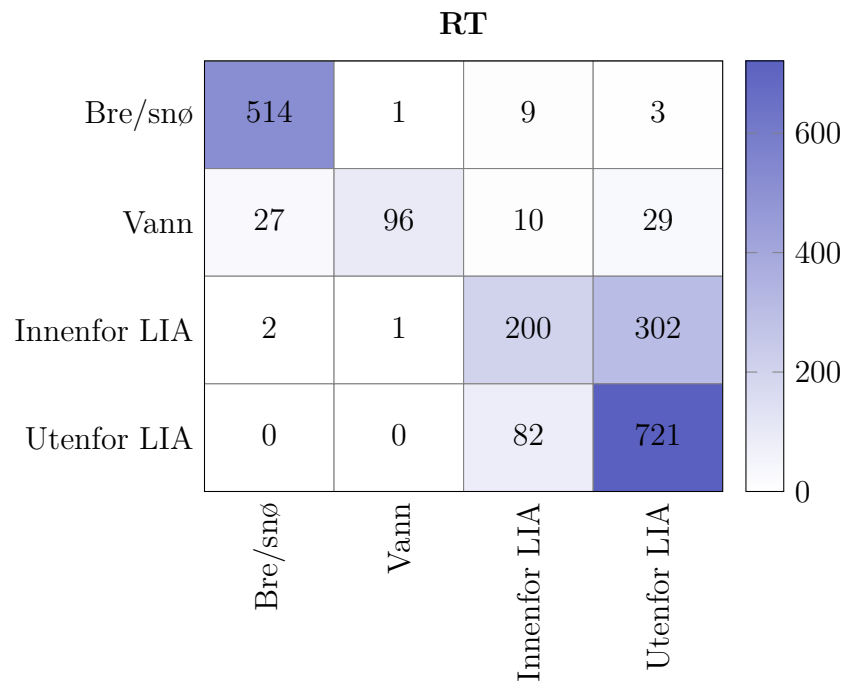
**Figur 5.1.2:** Resultatet av klassifisering med RT på testområdet Rembesdalskåka.

### 5.1.1 Statistisk test

Nedenfor presenteres forvirringsmatrisene fra klassifiseringen fra henholdsvis SVM og RT.



**Figur 5.1.3:** SVM forvirringsmatrise.



**Figur 5.1.4:** RT forvirringsmatrise.

En forrvirringsmatrise per modell ble generert ut fra klassifiserte verdier og sanne verdier, figur 5.1.3 og 5.1.4. Ut fra matrisene ble presisjon, gjenkall, F1-verdi og total nøyaktighet regnet ut. De statistiske resultatene vises under i tabellene 5.1.1 og 5.1.2.

	Nøyaktighet
SVM	0.80 - 80%
RT	0.77 - 77%

**Tabell 5.1.1:** Den totale nøyaktigheten for SVM og RT.

	Presisjon		Gjenkall		F1-verdi	
	SVM	RT	SVM	RT	SVM	RT
Bre/snø	0.95	0.95	0.99	0.95	0.97	0.96
Vann	0.97	0.98	0.66	0.59	0.78	0.74
Innenfor LIA	0.65	0.66	0.43	0.40	0.52	0.50
Utenfor LIA	0.74	0.68	0.89	0.90	0.81	0.78

**Tabell 5.1.2:** Presisjon, gjenkall og F1-verdi for SVM og RT innad i hver klasse.

## 5.2 Sentinel-2 pikselverdier statistikk

Videre i resultatene vil man finne ulike statistiske analyser gjort for hele Hardangerjøkulen, linjevis og for de tre brearmene, Midtdalsbreen, Blåisen og Rembesdalskåka.

Under analyseringen kom det frem at minimum- og maksimumsverdiene ikke er den beste måten å undersøke denne problemstillingen på. Etter at bre, snø og vann var fjernet, var det fortsatt noen punkt som havnet på nabopikslene, hvor verdien ble preget av de fjernede områdene. Dette førte til at 1920-2013 polygonet fikk høyere ekstremverdier enn de to andre polygonene som ligger lenger ut fra breen. Dette blir også de andre analysene preget av, men i mindre grad. Tabellene over minimum- og maksimumverdier kan studeres nærmere i Tillegg B.

Det ble ikke funnet noen korrelasjon mellom polygonene eller mellom linjene. Derfor er ikke disse tatt med i resultatene. Oversikt over korrelasjonen kan ses i Tillegg C.

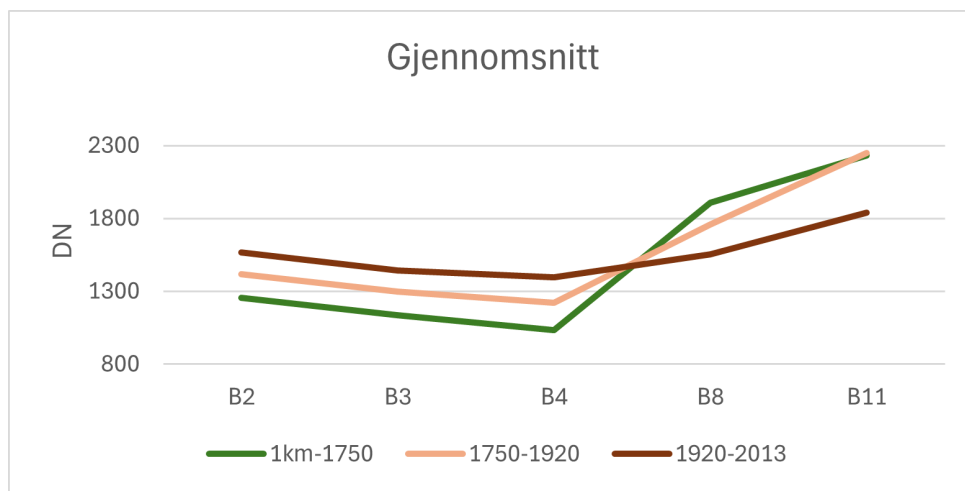
### 5.2.1 Breomkretser

Resultatene i tabell 5.2.1 samt figurene 5.2.1 og 5.2.2 viser variasjonen mellom polygonene for Sentinel-2 og indeksene, samt gjennomsnittsverdiene for polygonene: 1km-1750,

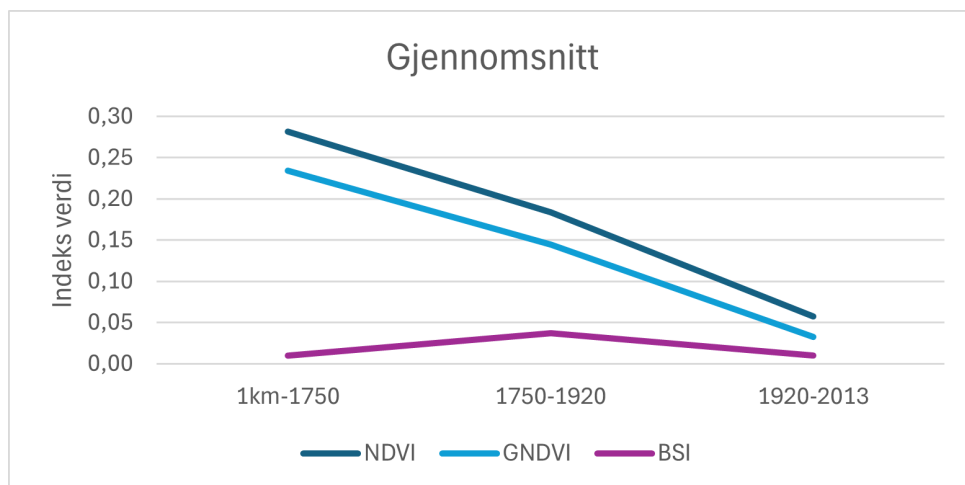
1750-1920 og 1920-2013. Videre presenteres gjennomsnittene per polygon med tilhørende standardavvik ( $\pm 1$  std) i figur 5.2.3.

Varians		
	F	F-kirt
Sentinel-2	144	3
NDVI	2895	3
GNDVI	2893	3
BSI	193	3

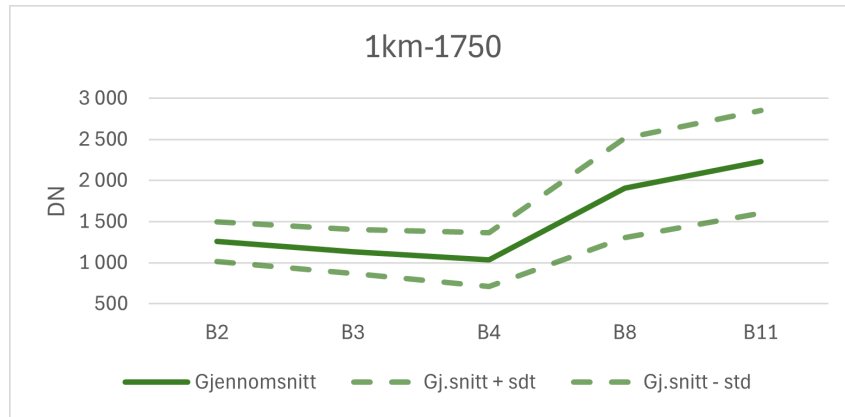
**Tabell 5.2.1:** Variansen for S-2A og indeksene mellom polygonene.



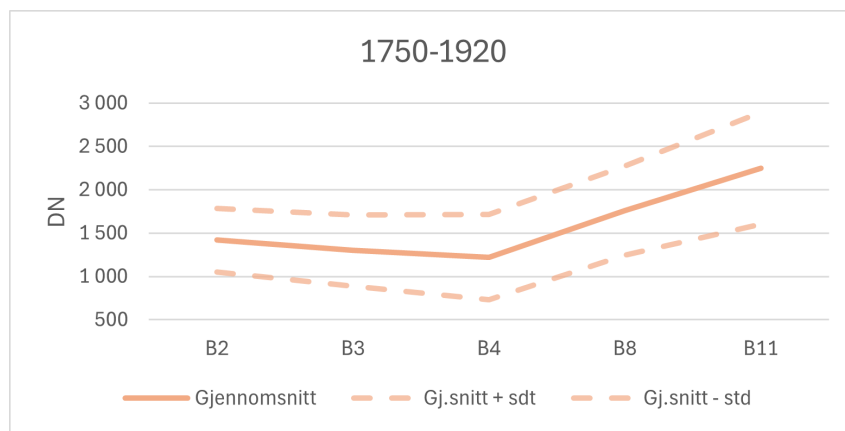
**Figur 5.2.1:** Gjennomsnittsverdier for S-2A båndene for hvert polygon.



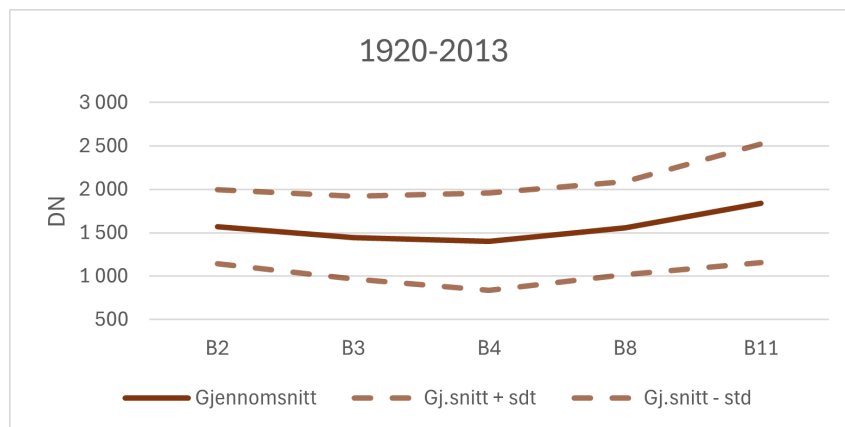
**Figur 5.2.2:** Gjennomsnittsverdier for indeksene NDVI, GNDVI og BSI for hvert polygon.



(a)



(b)



(c)

**Figur 5.2.3:** (a), (b) og (c) viser gjennomsnittsverdien til hvert av polygonene rundt Hardangerjøkulen med tilhørende standardavvik.

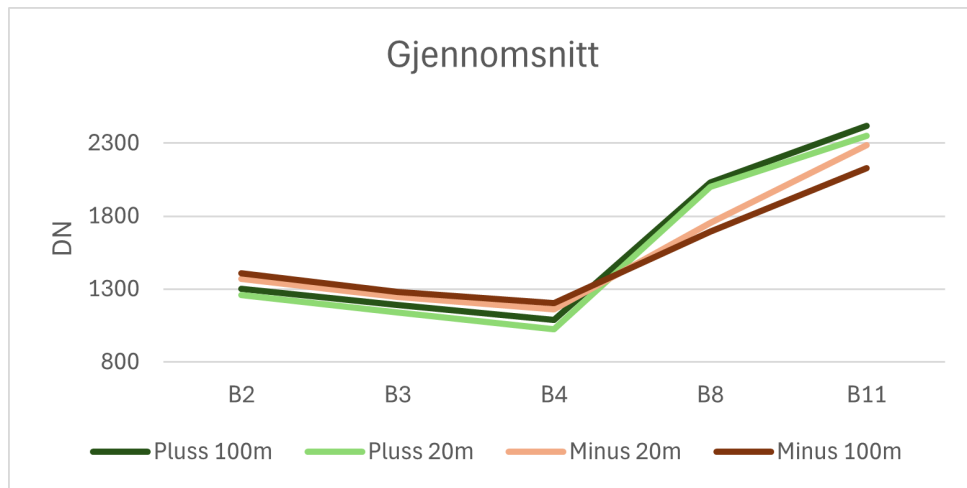


### 5.2.2 LIA utbredelse

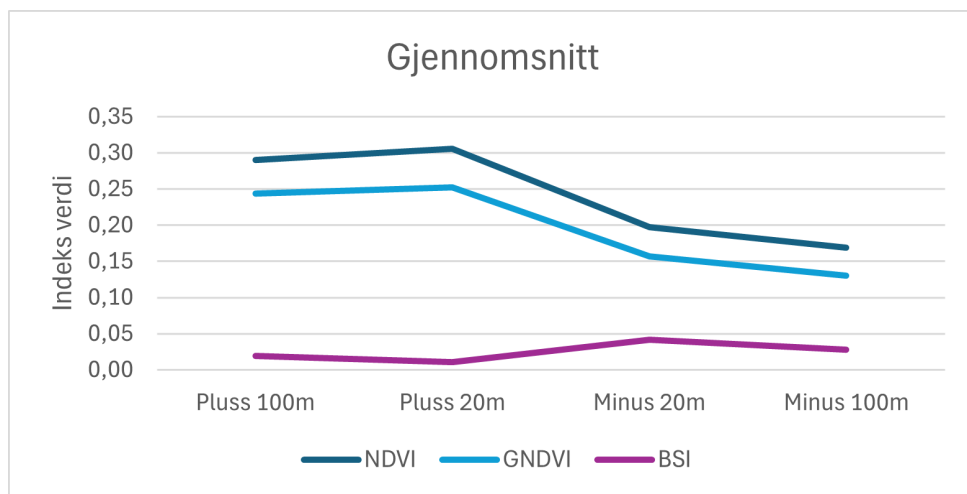
Tilsvarende som for polygonene rundt Hardangerjøkulen, følger variansen mellom linjene i tabell 5.2.2. Dette kun for de sikre kartlagte stedene rundt LIA morenerygg. Videre følger gjennomsnittsverdier for hver av linjene for Sentinel-2 båndene og indeksene i figurene 5.2.4 og 5.2.5, samt gjennomsnitt med tilhørende standardavvik ( $\pm 1$  std) i figur 5.2.6.

	Varians	
	F	F-kirt
Sentinel-2	50	3
NDVI	852	3
GNDVI	876	3
BSI	130	3

**Tabell 5.2.2:** Varians for S-2A og indeksene mellom linjene.



**Figur 5.2.4:** Gjennomsnittsverdier for S-2A båndene for hver punktlinje.



**Figur 5.2.5:** Gjennomsnittsverdier for NDVI, GNDVI og BSI for hver punktlinje.



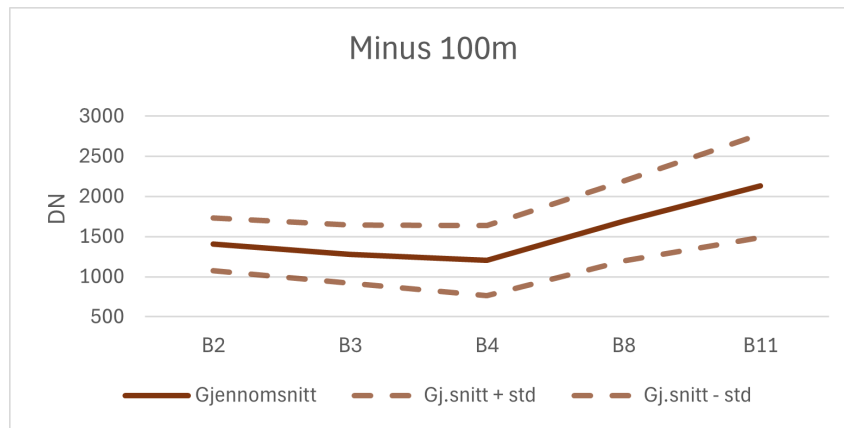
(a)



(b)



(c)



(d)

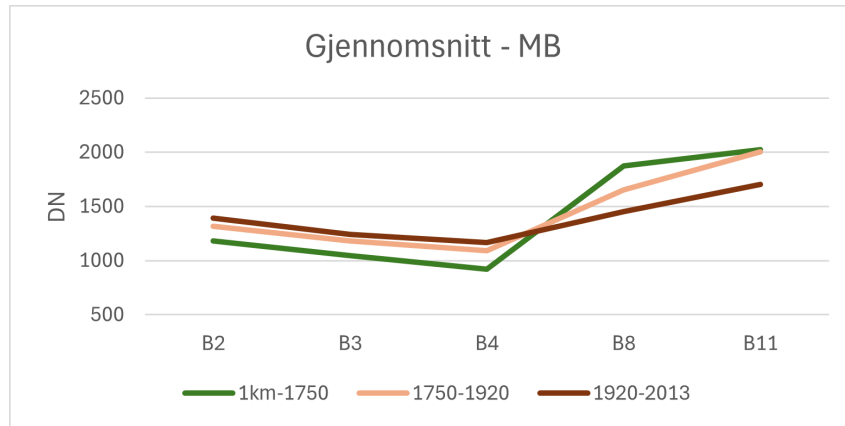
**Figur 5.2.6:** (a), (b), (c) og (d) viser gjennomsnittsverdien til hver av punktlinjene rundt Hardangerjøkulen med tilhørende standardavvik.

### 5.2.3 Midtdalsbreen, Blåisen og Rembesdalskåka

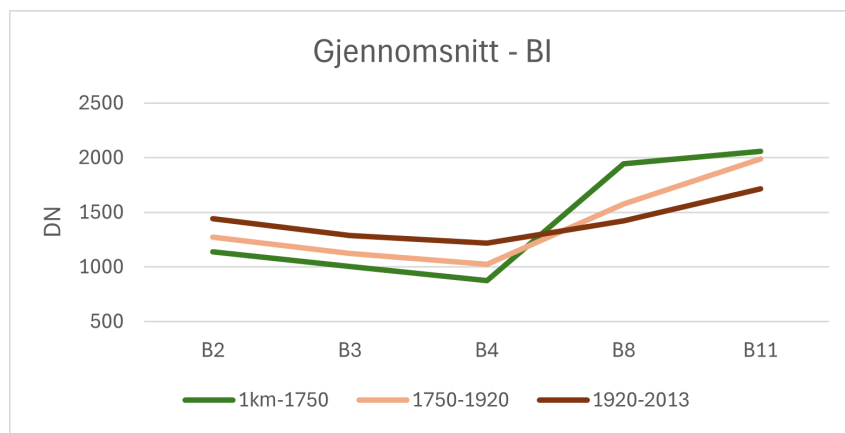
Brearmene Midtdalsbreen, Blåisen og Rembesdalskåka fikk generert nye tilfeldig punkt innad i hver polygon. Nedenfor i tabell 5.2.3 presenteres variansen mellom brearmene. Videre presenteres brearmene sine gjennomsnittsverdier for Sentinel-2 båndene og indek-sene, hver for seg i figurene 5.2.7 og 5.2.8. For gjennomsnittet til satellittbåndene kan man observere at de største forskjellene ligger i det nærinfrarøde (bånd 8).

	Varians	
	F	F-kirt
1km-1750	738	3
1750-1920	255	3
1920-2013	705	3

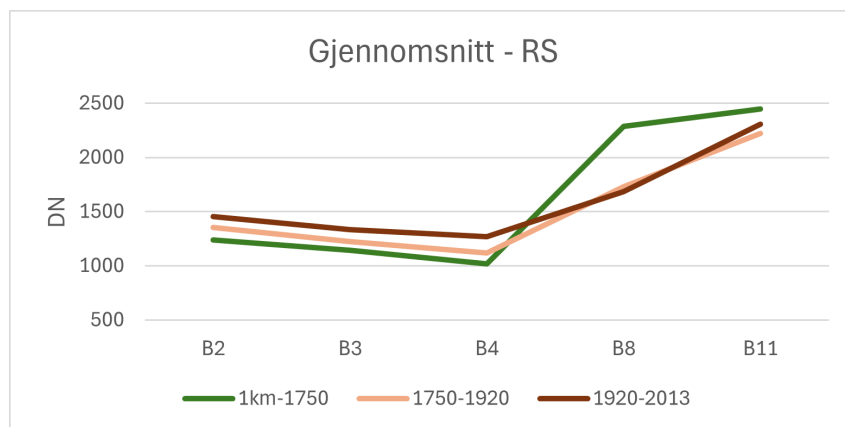
**Tabell 5.2.3:** Variansen mellom brearmene for hver av polygonene.



(a) *Midtdalsbreen (MB)*.

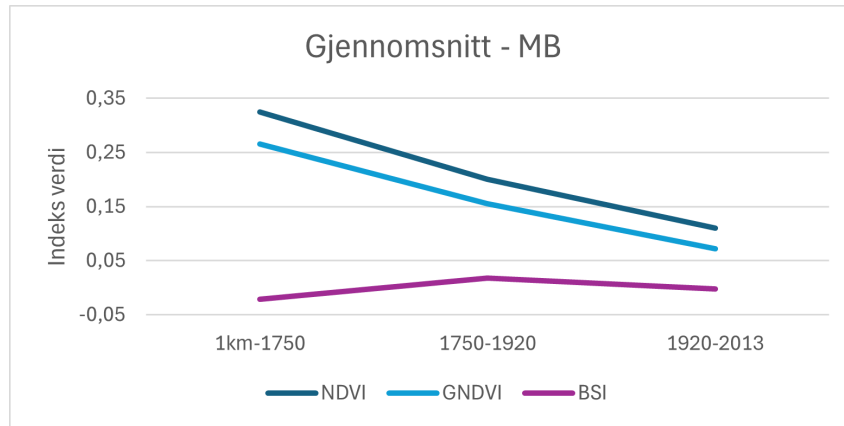


(b) *Blåisen (BI)*.

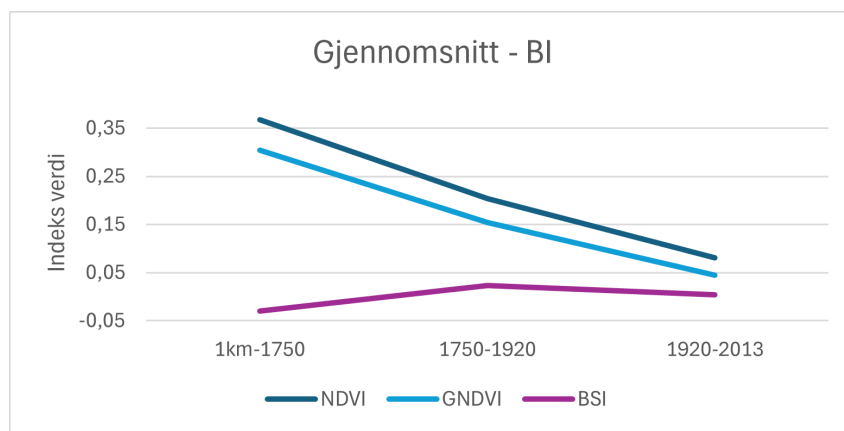


(c) *Rembesdalskåka (RS)*.

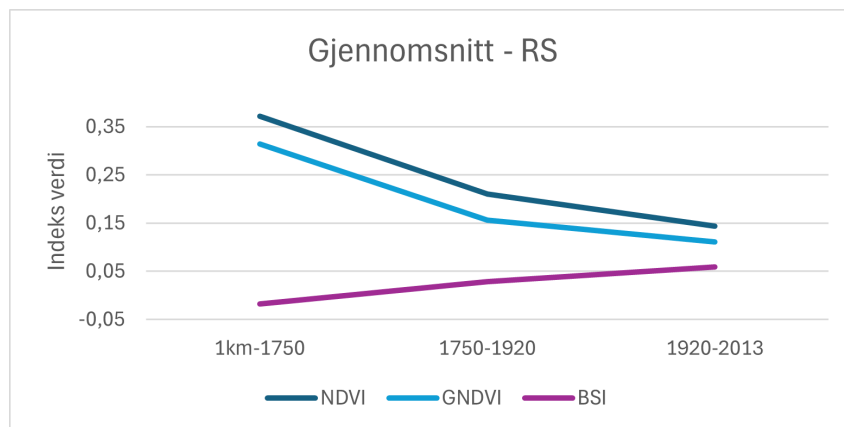
**Figur 5.2.7:** (a), (b) og (c) viser S-2A gjennomsnittsverdiene for polygonene tilhørende hver av brearmene.



(a) *Midtdalsbreen (MB)*.



(b) *Blåisen (BI)*.

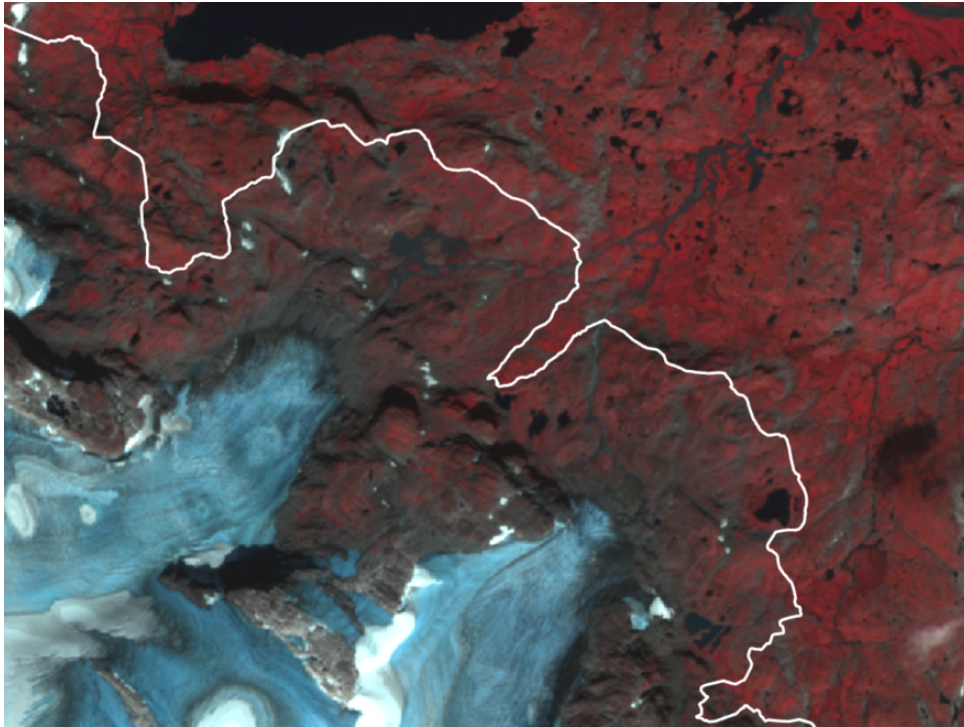


(c) *Rembesdalskåka (RS)*.

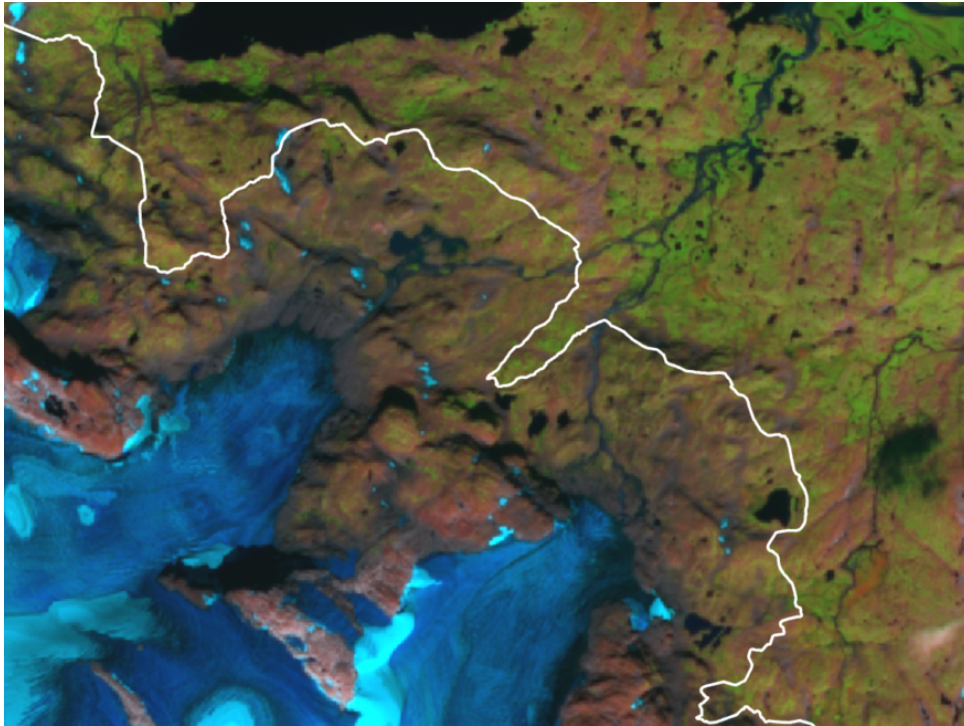
**Figur 5.2.8:** (a), (b) og (c) viser indeksene sin gjennomsnittsverdien for polygonene tilhørende hver av brearmene.

### 5.2.4 Båndkombinasjoner over Midtdalsbreen og Blåisen

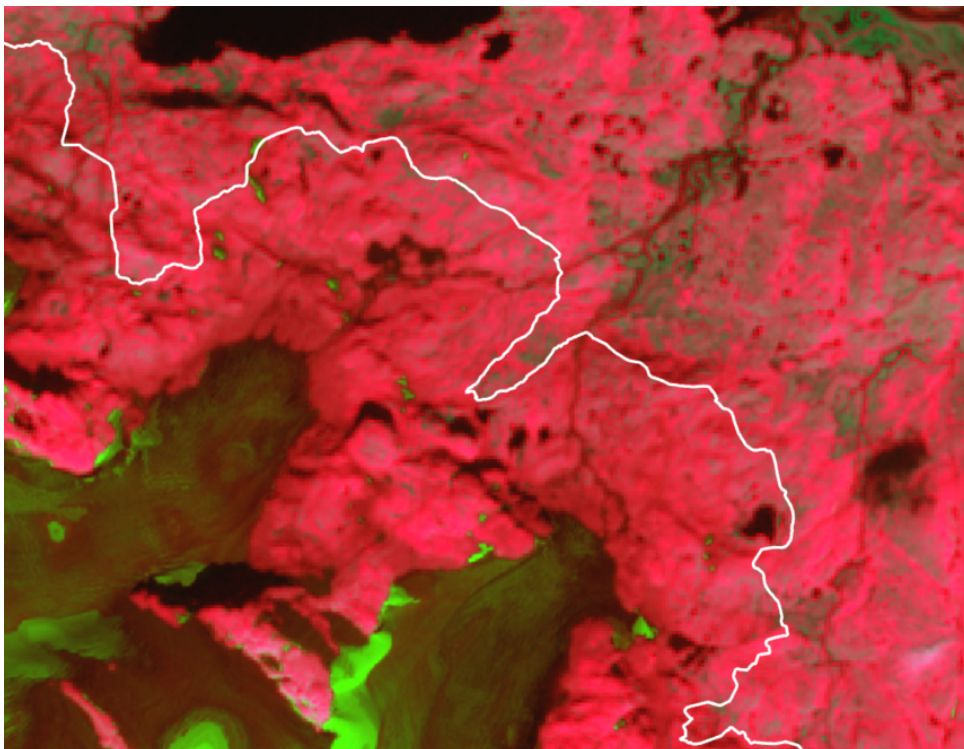
Figurene 5.2.9 (a,b,c) presenterer Midtdalsbreen og Blåisen visuelt i tre forskjellige båndkombinasjoner, sammen med fasitlinjen til LIA. For området i denne oppgaven er det visuelt mulig å antyde hvor LIA sin morenerygg går, kun basert på de ulike båndkombinasjonene. I alle tre kombinasjonene ser man tydelig forskjell på områder med og uten vegetasjon, dette danner et mønster i terrenget.



(a) Båndkombinasjon:  $B8 - B4 - B3$



(b) Båndkombinasjon: B11 - B8 - B4



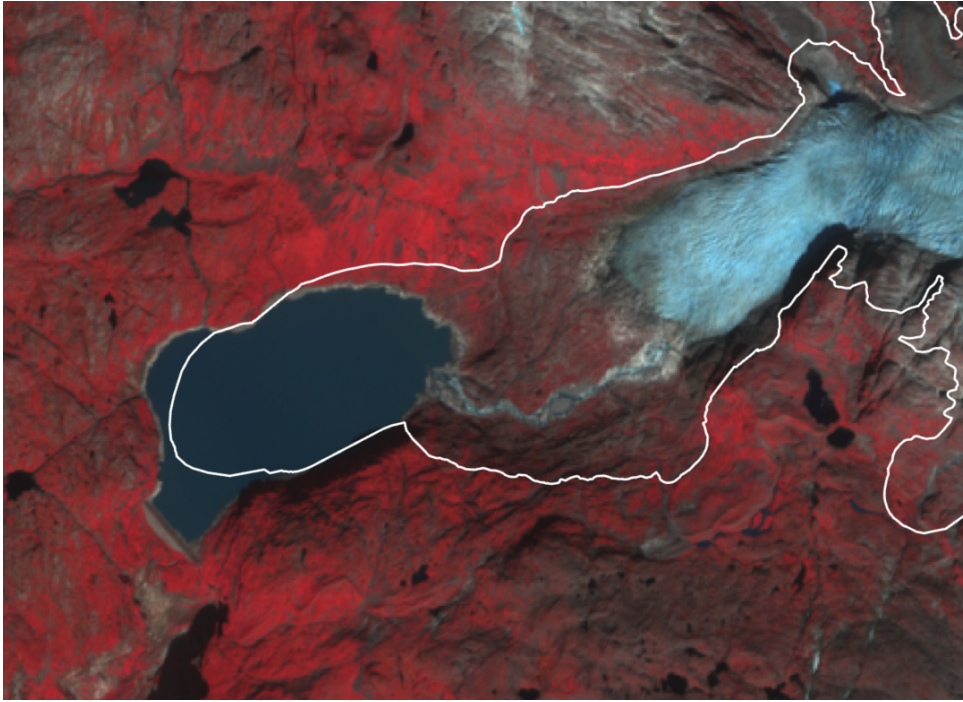
(c) Båndkombinasjon: BSI\*6.25 - B8 - B11

**Figur 5.2.9:** (a), (b) og (c) Visuelt resultat av Midtdalsbreen og Blåisen gitt i tre ulike båndkombinasjoner.

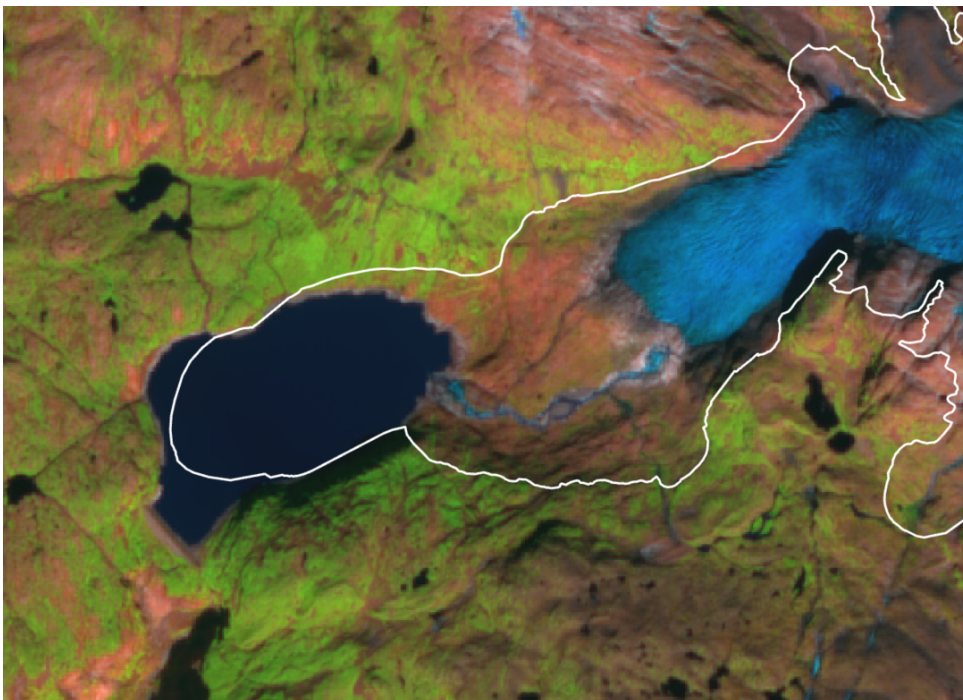


### 5.2.5 Båndkombinasjoner over Rembesdalskåka

På samme måte som for Midtdalsbreen og Blåisen, presenterer figurene 5.2.10 (a,b,c) de visuelle resultatene over Rembesdalskåka. Mønsteret i terrenget går igjen her og muligheten for å antyde LIA er tilstede, spesielt mellom tuppen av breen og det store vannet.

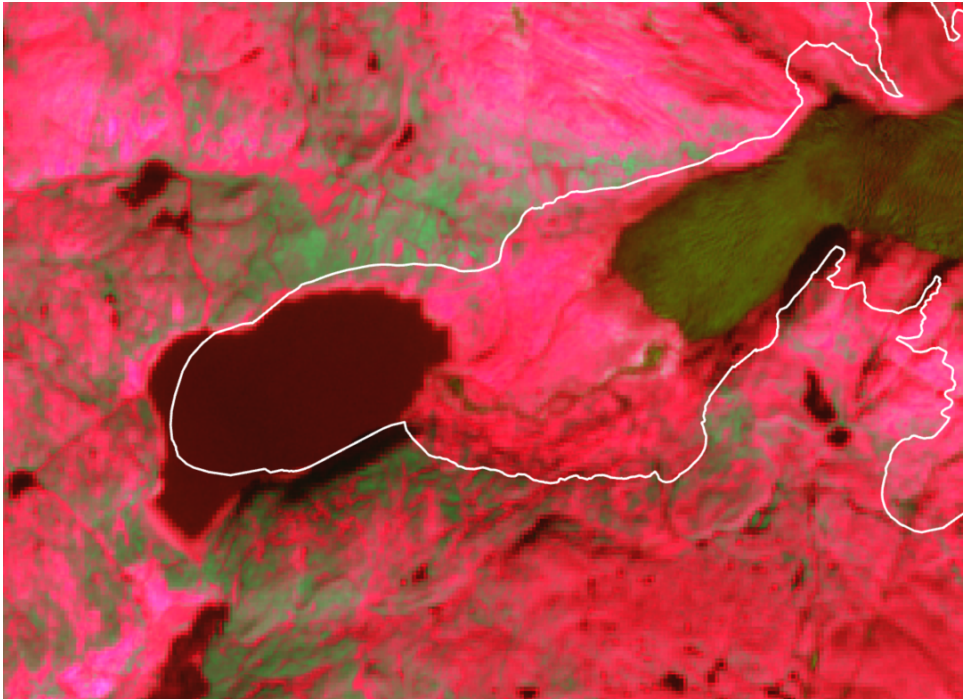


(a) Båndkombinasjon: B8 - B4 - B3



(b) Båndkombinasjon: B11 - B8 - B4





(c) Båndkombinasjon:  $BSI*6.25 - B8 - B11$

**Figur 5.2.10:** (a), (b) og (c) Visuelt resultat av Rembesdalskåka gitt i tre ulike båndkombinasjoner.

## Diskusjon

### 6.1 Usikkerhet i datagrunnlag

I denne studien ble Sentinel-2 datasett med 10m oppløsning anvendt for klassifisering og pikselanalyse. Datasettet viste seg i denne oppgaven å ha god nok oppløsning for kartlegging av LIA utbredelse ved Hardangerjøkulen. I delkapittel 6.3 diskuteres funnene i oppgaven.

En potensiell kilde til usikkerhet i datasettet kan være de kartlagte breomkretsene rundt Hardangerjøkulen. Disse kartene ble utarbeidet manuelt, noe som innebærer en viss grad av usikkerhet knyttet til resultatene. Omtrent 60% av moreneryggen fra LIA ble fastslått som helt sikker, basert på israndmorener, grenser for breelvavsetninger, trimlinjer og identifiserbare erosjons- og forvitringsgrenser. For kortere avstander ble litt over 30% interpolert mellom sikre linjer. De resterende 8,8% ble bestemt basert på terrengets topografi. Breomkretsen fra 1920-tallet ble bestemt gjennom feltarbeid og et topografisk kart basert på bakkeundersøkelser utført på 1920-tallet (Weber mfl., 2019).

Som nevnt i kapittel 3, er Hardangerjøkulen et åpent område med minimal vegetasjon. I slike terreng kan det være nødvendig med en mer detaljert undersøkelse av bergtyper og geologiske strukturer. Bånd 12 (SWIR) er et forslag som et tillegg i analysen. Dette båndet, sammen med bånd 11 (SWIR), vil begge bidra til en mer omfattende analyse av områdets geologi.

Problemstillingen for oppgaven var å undersøke om hvor godt Sentinel-2 bilder kan brukes til å detektere LIA utbredelse på Hardangerjøkulen. Morenerygger er karakteristiske og har potensialet til å bidra i en klassifisering. I tillegg til et satellittdatasett, kan det være fordelaktig å inkludere en digital terrenngmodell (DTM) for en mer robust kartlegging av terrenget.

## 6.2 Usikkerheter knyttet til metode

Metoden utført i oppgaven, er gjennomført veldig stegvis. Det er imidlertid fullt mulig å forestille seg at en mer effektiv tilnærming er tilstede, enten gjennom bruk av GIS-programvare eller ved hjelp av programmering. En av fordelene med den anvendte metoden er dens evne til å tydeliggjøre eventuelle feil som kan oppstå mellom de forskjellige trinnene i prosessen.

### 6.2.1 Klassifiseringsmetoder

Valget av klassifiseringsmetodene Support Vector Machine (SVM) og Random Trees (RT) ble gjort basert på tidligere positive resultater fra lignende områder, f.eks (A. A. Khan mfl., 2020).

Generelt sett var klassifiseringene tilfredsstillende. Som det fremgår av resultatene, presterer SVM litt bedre enn RT. Det er ikke store forskjeller mellom metodene, og eventuelle avvik er systematiske. Gjenkall for vann viser det største avviket mellom metodene, med en forskjell på 7%.

### 6.2.2 Trening- og testområde

Treningsområdet for denne studien ble valgt som et utsnitt av Midtdalsbreen og Blåisen, mens testområdet var et utsnitt av Rembesdalskåka. Disse brearmene ligger på hver sin side av Hardangerjøkulen og kan ha ulik påvirkning av vær og vind, noe som potensielt kan påvirke resultatene. For å forbedre opplæringen av klassifiseringsmetodene, kunne det vært en strategi å gradvis utvide treningsområdet. Dette ville tillate algoritmene å bli kjent med et bredere spekter av terrengetyper. I tillegg kunne det ha vært fordelaktig å bruke et testområde som ikke er direkte knyttet til treningsområdet. I denne studien kan trening- og testområdene ha vært for like områder og derav gode resultater med 77% og 80% totalt nøyaktighet. Treningsdataen fra Hardangerjøkulen kan mest sannsynlig ikke benyttes på andre breområdet med mer vegetasjon.

### 6.2.3 Effekten av treningsdata

For å trene algoritmene ble det generert treningsdata bestående av polygoner. Totalt ble det produsert 81 treningspunkter, som utgjør omtrent 5% av treningsområdet. Det ble besluttet å etablere fire klasser: snø/bre, vann, innenfor LIA og utenfor LIA. Disse fire

klassene ble valgt basert på problemstillingen for oppgaven. Klassene *snø/bre* og *vann* ble opprettet for å kunne skilles fra resten av datasettet, da disse områdene ikke var av interesse for oppgaven. Klassene *innenfor LIA* og *utenfor LIA* var av primær interesse. Disse klassene ble relativt store og inneholdt et bredt spekter av spektralsignaturer. Sett i ettetid kunne det vært hensiktsmessig å dele klassene ytterligere. Klasser som vegetasjon på innsiden og utsiden av LIA, samt fjell på innsiden og utsiden av LIA, er potensielle kategorier som kan benyttes for å oppnå mer nøyaktige resultater.

Å finne den rette balansen for et treningsdatasett er en utfordrende oppgave. Man ønsker verken et overklassifisert bilde, men heller ikke for lite treningsdata. Effekten av mengden treningsdata ble testet for å undersøke dens innvirkning på klassifiseringen. To treningsdatasett med en prosentandel på henholdsvis litt over 5% og 1% ble testet opp mot hverandre. Resultatene ga en visuell forskjell, men en minimal statistisk forskjell. Differansen i den totale nøyaktigheten havnet på under 1%. Dette kan indikere at området rundt Midtdalsbreen og Blåisen er karakteristisk og et enkelt området å klassifisere. Resultatene av denne undersøkelsen kan studeres nærmere i Tillegg A.

#### 6.2.4 Valideringspunkter

Validering av klassifiseringene ble utført ved å generere 2000 tilfeldige punkt. Antallet ble valgt for å sikre robust statistikk. Punktene ble manuelt tildelt en sann verdi basert på ortofoto og fasitlinjen til LIA. Det kan imidlertid stilles spørsmål ved påliteligheten og nøyaktigheten av disse valideringspunktene.

#### 6.2.5 Andre usikkerheter

I analysene ble snø, vann og bre ekskludert, slik at fokuset utelukkende var på terrenget rundt Hardangerjøkulen. Piksler som befant seg i nærheten av vann eller snø/bre ble påvirket av dette og resulterte i datasettets ekstremverdier. Dette kan observeres i histogrammene for minimum- og maksimumverdier i Tillegg B. På grunn av denne påvirkningen, vil ikke minimum- og maksimumverdiene, samt BSI for polygonet fra perioden 1920-2013, gi et pålitelig resultat.

En annen betydelig faktor kan være årstiden da bildene ble tatt. I løpet av sommeren vil naturen blomstre, og trærne vil bli grønnere og mer frodige. Mengden av vegetasjon vil variere gjennom året. Som det fremgår av resultatene, har bånd 8 (NIR) en betydelig innvirkning på resultatene. Det nærinfrarøde båndet er spesielt effektivt for å detektere vegetasjon og vil derfor bli påvirket i de forskjellige årstidene. Følgelig vil indeksene også variere avhengig av når på året bildene blir tatt. Dette understreker viktigheten av å ta hensyn til sesongvariasjoner ved analyse med satellittbilder.

## 6.3 Resultater

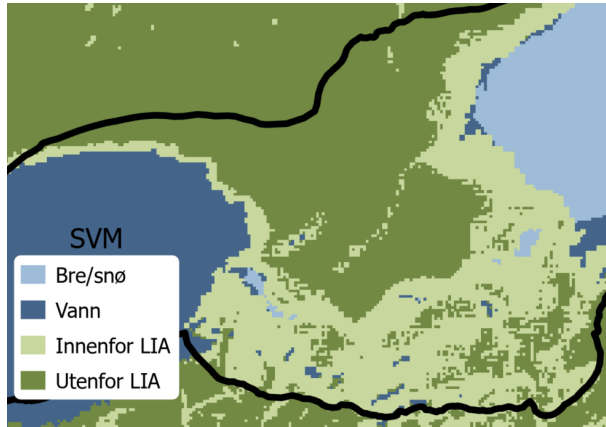
### 6.3.1 Klassifisering

Resultatene indikerer at både SVM og RT oppnår en relativ høy total nøyaktighet, med henholdsvis 80% og 77%. Ved å vurdere klassene individuelt, observeres det at *innenfor LIA* presterer dårligst med en presisjon på 65%, gjenkall på 43% og F1-verdi på 52%, for SVM. I toppen oppnår klassen *bre/snø* de høyeste prosentverdiene, med verdier mellom 95%-99%, for begge algoritmene. Dette var forventet gitt klassens karakteristiske spektralsignatur sammenlignet med de andre klassene. *Vann* og *utenfor LIA* oppnår omtrent like F1-verdi, med henholdsvis 78% og 81% for SVM. Presisjonen ligger endel høyere for *vann*, med en score på 97% og 98%, mens gjenkall er høyere for *utenfor LIA* med en score på 89% og 90%.

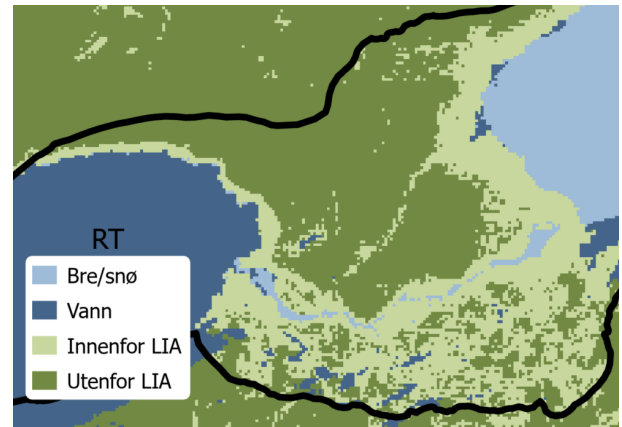
Til tross for at analysene for hver av klassifiseringsmetodene er relativt like, kan det observeres visuelle forskjeller mellom dem. Visuelt sett ser SVM ut til å være mer treffsikker enn RT. En gjentakende utfordring for RT er feilklassifisering av flere større områder, spesielt områder som utelukkende består av fjell.

Det er en merkbar forskjell mellom den forventede og kartlagte plasseringen av moreneryggen til LIA maksimum og hvor den ble klassifisert. I figur 5.2.10 ser det ut til at den kartlagte moreneryggen samsvarer godt med landskapet. Når man ser på de klassifiserte bildene i figur 6.3.1 og 6.3.2, kan det virke som om moreneryggen burde vært flyttet lenger inn. Dette gapet i klassifiseringen, til tross for klarheten i de visuelle bildene, kan være et resultat av begrensninger i treningsområdet. Det er en mulighet for at treningsområdet var for snevret, med utilstrekkelig variasjon i terrenget. To interessante eksempler følger under.

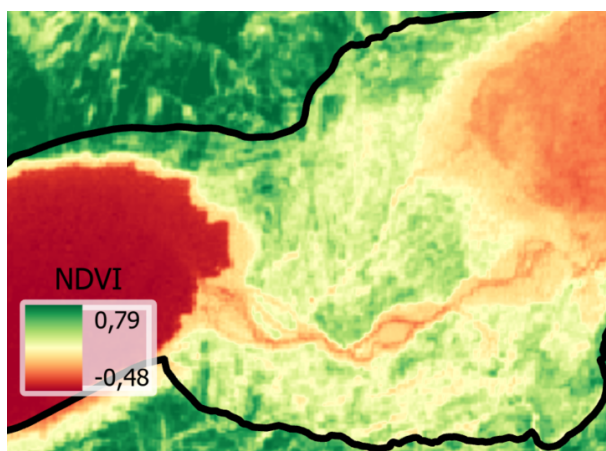
Første eksempel er en visuell fremstilling av området beskrevet over. Her er det observert at en del av områdets indre er feilaktig klassifisert som ytre. Basert på ortofoto og NDVI, er det minimal eller ingen vegetasjon til stede, noe som vanligvis ville ført til en klassifisering som *innenfor LIA*. Figur 6.3.4 viser at klassifiseringen av LIAs morenerygg har blitt tildelt kategorien gul - ganske sikker over deler av området. Fasongen på feilklassifiseringen i både SVM og RT samstemmer godt, samtidig som de følger fasongen til fasitlinjen. Dette området inneholder interessante faktorer som har potensialet for ytteligere undersøkelser.



**Figur 6.3.1:** SVM klassifiseringsresultat for Rembedalskåka.



**Figur 6.3.2:** RT klassifiseringsresultat for Rembedalskåka.

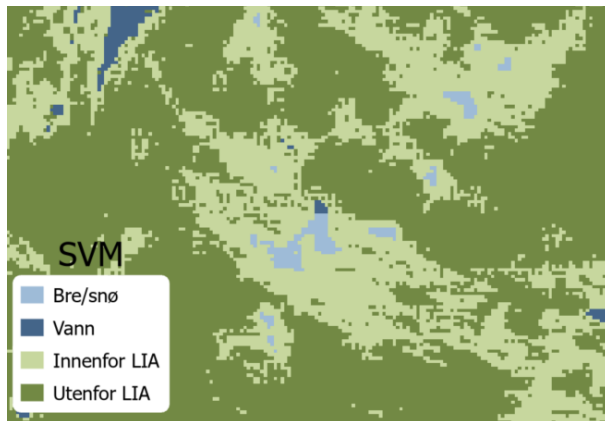


**Figur 6.3.3:** NDVI over testområdet Rembedalskåka.

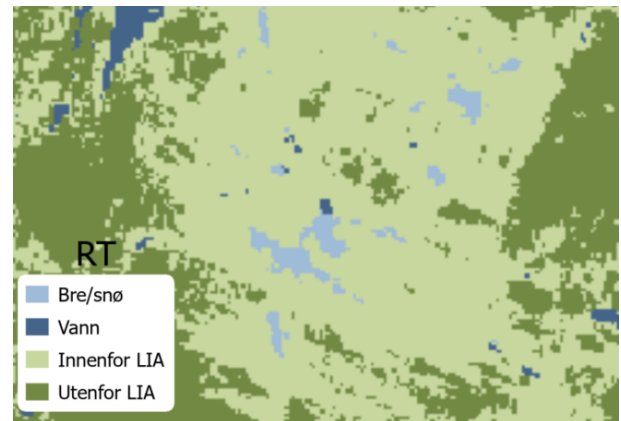


**Figur 6.3.4:** LIA fasit rundt Rembedalskåka.

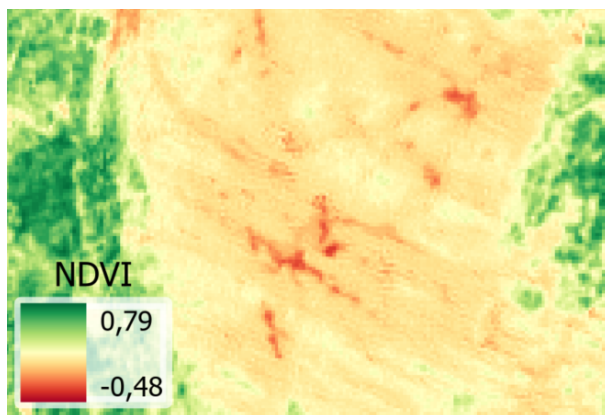
Et annet eksempel illustrerer klassifiseringsresultatene for et område utenfor LIA. Dette området består hovedsaklig av fjell og stein, noe som kan observeres i NDVI bildet. Det er tydelig at RT modellen viser en mindre presis klassifisering ved å identifisere nesten hele området som *innenfor LIA*. Dette mønsteret er gjennomgående i hele klassifiseringsprosessen. Spesielt i regioner med overveiende fjell og stein, viser RT en tendens til å kategorisere områdene som *innenfor LIA*. I kontrast oppnår SVM en mer treffsikker klassifisering i dette området, selv om det fortsatt er rom for forbedringer.



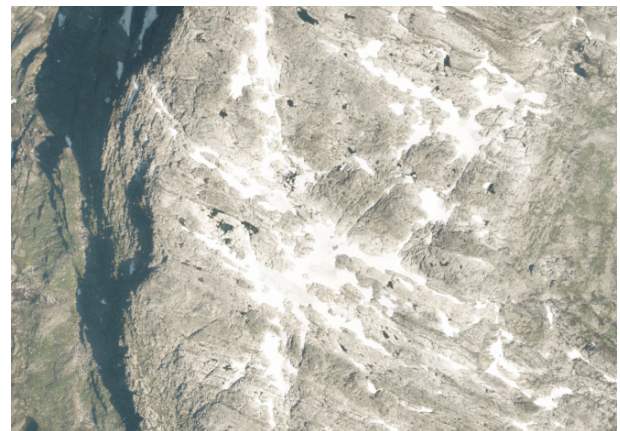
**Figur 6.3.5:** SVM klassifiseringsresultat på utsiden av LIA ved Rembesdalskåka.



**Figur 6.3.6:** RT klassifiseringsresultat på utsiden av LIA ved Rembesdalskåka.



**Figur 6.3.7:** NDVI ved Rembesdalskåka utenfor moreneryggen til LIA.



**Figur 6.3.8:** Ortofoto over samme området fra (Norgebilder.no, 2024).

### 6.3.2 Pikselbasert analyse

For å utforske de karakteristiske trekkene i terrenget, bidraget fra de ulike spektralbåndene og effekten av indeksene, ble området analysert på tre forskjellige måter. Isbreer og vann ble ekskludert fra datasettet slik at analysen baserte seg på LIA sin utbredelse og området opptil  $1km$  utenfor.

I metode 1 ble hele Hardangerjøkulen segmentert i tre polygoner basert på tidligere kartlegginger av isbreens bevegelse. Mer om arealendringene kan leses her: (Weber mfl., 2019). Polygonenes segmentering er gjengitt under i rekkefølgen ytterst til innerst:

- Fra  $1km$  på utsiden av LIA til 1750, som markerer LIAs periode for Hardangerjøkulen.
- Fra 1750 (LIA) til 1920.
- Fra 1920 til 2013, som representerer isbreens nåværende posisjon.

Til å svare på problemstillingen, er de to første polygonene av spesiell interesse da de representerer perioden isbreen var på sitt største. Det indre polygonet inneholder flere ekstreme verdier som påvirket resultatene noe, samt inneholder mindre vegetasjon, noe som synes å være en viktig faktor i analysen.

For metode 2 ble fire linjer trukket ut fra LIA sine sikre kartlagte områder (se figur 3.2.1, grønne linjer), for å undersøke om det er signifikante forskjeller nært moreneryggen. Avstander på  $20m$  og  $100m$  ble valgt for denne undersøkelsen.

Til slutt ble metode 3 gjennomført på samme måte som metode 1, men ble anvendt separat for tre av brearmene: Midtdalsbreen, Blåisen og Rembesdalskåka. Hensikten var å se om de forskjellige fremstøtene og plasseringene av brearmene kunne ha en innvirkning på analysen.

En gjennomgående observasjon fra alle tre metodene er variansen som gir et tydelig signal om signifikante forskjeller mellom breomkretsene rundt Hardangerjøkulen og mellom linjene tilknyttet LIAs morenerygg. Variansen var størst for indeksene NDVI og GNDVI, med verdier på omtrent 2890 for hele Hardangerjøkulen og 852-576 for linjene. En større forskjell i disse indeksene var forventet, da teorien om økt vegetasjon utenfor LIA var tilstede. Dette understreker at vegetasjon bruker tid på å etablere seg robust. Histogrammet i figur 5.2.4 illustrerer dette poenget klart, hvor det er et merkbart gap i det nærinfrarøde båndet (bånd 8) mellom linjene utenfor og på innsiden av LIA. Ser man på variansen for brearmene oppdager man at det er minst variasjon i årstallet 1750-1920 enn for de andre årstallene, med en verdi på 255 kontra litt over 700.

Variansen mellom breomkretsene for hele Hardangerjøkulen kan potensielt være noe misvisende, ettersom alle tre ble analysert sammen. Det er mulig at det innerste polygonet (1920-2013) har hatt for stor innvirkning på resultatene. Standardavviket for polygonene øker proporsjonalt med nærheten til isbreen som kan observeres i histogrammene 5.2.3 (a,c,b).



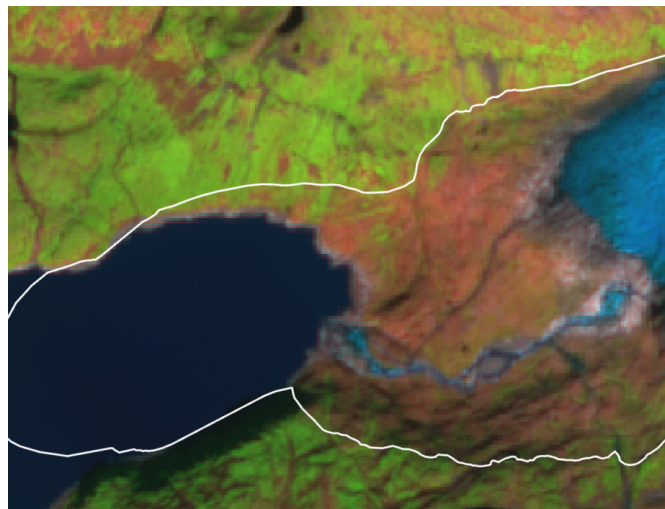
Observasjoner tilknyttet linjene 20m og 100m på innsiden av LIA, viser en jevn trend på tvers av alle spektralbånd. Dette mønsteret gjenspeiler seg også for de to linjene utenfor LIA. I de synlige spektralbåndene 2, 3 og 4 (RGB), er refleksjonen relativt stabil, mens i det nærinfrarøde båndet (bånd 8) opplever begge linjene utenfor LIA en økning i refleksjonsverdiene. Dette korrelerer med en gjennomsnittlig økning på 0.1 i NDVI og GNDVI indeksene, samtidig med en liten reduksjon i BSI.

### 6.3.3 Visuell analyse av brearmer

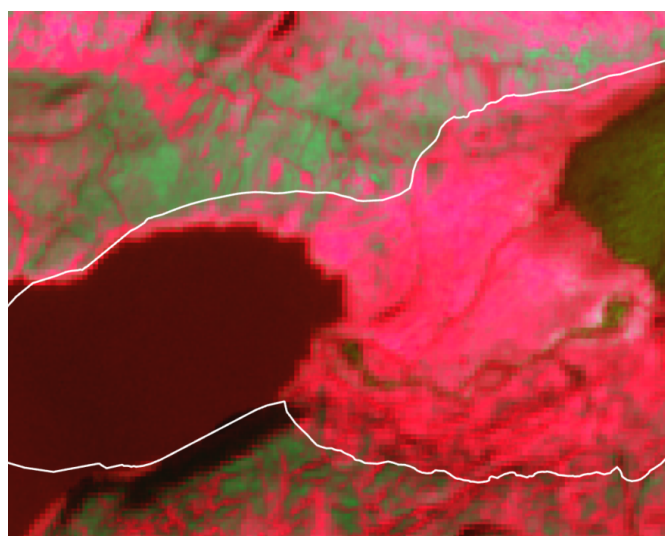
De visuelle satellittbildene av brearmene viser at det er mulig å identifisere signifikante forskjeller mellom områdene innenfor og utenfor LIA ved hjelp av de tre båndkombinasjonene. De mest fremtredende forskjellene observeres ved Rembesdalskåka, hvor overgangene er ganske markante. Disse forskjellene kan undersøkes nærmere i figurene 6.3.9 (a,b,c). Vegetasjonen er tydelig i alle båndkombinasjonene, noe som bekrefter tidligere resultater om at spektralbåndene inneholder rikelig med informasjon som reflekterer vegetasjon. Bånd 11 (SWIR) har også en god innvirkning på resultatene, som fremhever bar bakke.



(a) Båndkombinasjon: B8 - B4 - B3



(b) Båndkombinasjon: B11 - B8 - B4



(c) Båndkombinasjon: BSI\*6.25 - B8 - B11

**Figur 6.3.9:** (a), (b) og (c) Visuelt resultat over Rembesdalskåka, viser markante overganger innenfor og utenfor LIA

## 6.4 Bruk av kunstig intelligens

Kunstig intelligens (KI) har opplevd betydelig vekst og utvikling i de senere år, og har etablert seg som en verdifull ressurs. I denne oppgaven har KI blitt anvendt for å forbedre og omformulere setninger og avsnitt, med mål om å oppnå et mer akademisk språk. Videre har KI periodisk bistått med å foreslå referanser relatert til teori. Alle referanser foreslått av KI-generatoren har blitt nøye vurdert.

## 6.5 Mulige feilkilder

I enhver vitenskapelig undersøkelse som denne, er det uunngåelig at det oppstår feil. Noen av de mulige feilkildene er oppsummert nedenfor, og det er sannsynlig at det finnes ytterligere:

- Menneskelige feil, slik som feil tastetrykk eller feil i kopiering og liming.
- Manuell bestemmelse av sanne verdier. De sanne verdiene for valideringspunkter i klassifiseringene ble fastsatt manuelt.
- Variabilitet i treningsdataen. Forskjellig størrelse på polygonene, samt noe varierende antall.

## 6.6 Videre arbeid

Sentinel-2 viser stort potensial i kartleggingen av morenerygger. Selv om resultatene er positive, finnes det rom for forbedringer. Nedenfor er forslag til videre arbeid:

- Implementering av DTM. Utforske ulike statistikker som for eksempel helning, høyder og 2.derivert for å forbedre nøyaktigheten.
- Indekser har vist seg å gi mye informasjon. I åpne områder som dette studieområdet, kan inkludering av flere indekser forbedre analysen, spesielt de som er sensitive for stein og fjell.
- Videre kan inkludering av bånd 12 støtte analyser relatert til stein og fjell.
- Flere klasser i klassifiseringen. Et forslag til oppdeling kan være:
  - Vegetasjon innefor LIA
  - Fjell innefor LIA

- Vegetasjon utenfor LIA
- Fjell utenfor LIA
- Sammenligning med andre isbreer, eksempelvis i Jotunheimen hvor breene har mer vegetasjon, for å validere resultatene.
- Bruke mer pålitelig valideringsdata for å sikre en mer robust validering av metoden.
- Testing av flere klassifiseringsalgoritmer.
- Vurdere fargevalget for å sikre at figurene kan tolkes av alle, inkludert personer med fargeblindhet. Fargene brukt i denne oppgaven vil oppfattes for like og kan derfor bli vanskelig å tolke for en som er fargeblind (Uutilsynet, 2024).

## Konklusjon

Problemstillingen for oppgaven var å se hvordan Sentinel-2 bilder kan brukes for å kartlegge LIA utbredelse på Hardangerjøkulen. For å svare på oppgaven ble pikselbasert klassifisering og statistiske analyser basert på pikselverdier testet ut.

Satellittbilder viser utvilsomt et betydelig potensial for kartlegging av LIAs maksimale utbredelse på Hardangerjøkulen. Valget av satellittbåndene 2 (B), 3 (G), 4 (R), 8 (NIR) og 11 (SWIR), samt indeksene NDVI, GNDVI og BSI viste seg å være gode bidragsytere. Resultatene fra klassifiseringen var tilfredstillende og presterte med en totalt nøyaktighet på 77% og 80% for klassifiseringsmetodene RT og SVM. Begge metodene har et potensial for videre opplæring og forbedring, med mål om å oppnå en høyere total nøyaktighet.

I denne studien ble det observert at moreneryggen til LIA er tydelig og merkbar i store deler av de kartlagte områdene. Resultatene fra linjen 20m på innsiden og utsiden av LIA styrker oppunder dette, da det var tydelig skille i den gjennomsnittlige verdien i bånd 8 (NIR). De tre båndkombinasjonene støtter visuelt disse resultatene.

En potensiell motarbeidende faktor i en slik kartlegging virker å være bart fjell. I et slikt område som Hardangerjøkulen er et nøyaktig og godt trent datasett en viktig faktor. Gitt analysens sterke avhengighet av vegetasjon, kan det antas at slike analyser er mest effektive i sommerhalvåret, med mindre andre datasett og sensorer blir inkludert.

Det kan konkluderes med at Sentinel-2 er godt egnet til kartlegging av LIAs maksimale utbredelse på Hardangerjøkulen. Klassifiseringen og analysene basert på vegetasjon viste seg å gi størst utnytte, både analytisk og visuelt. Med noe vegetasjon i området gir båndkombinasjonen og indekser en svært god indikasjon på LIA. Ved en klassifisering er en mer avhengig av å ha en anelse om hvor LIA utbredelse er for å kunne lage gode treningsdata.

# Litteraturliste

- Ahmed, N. A. (2023). What is A Confusion Matrix in Machine Learning? The Model Evaluation Tool Explained. Hentet 9. mai 2024, fra <https://www.datacamp.com/tutorial/what-is-a-confusion-matrix-in-machine-learning>
- Andersen, J. L., & Sollid, J. L. (1971). Glacial Chronology and Glacial Geomorphology in the Marginal Zones of the Glaciers, Midtdalsbreen and Nigardsbreen, South Norway. *Norsk Geografisk Tidsskrift - Norwegian Journal of Geography*, 25(1), 1–38. <https://doi.org/10.1080/00291957108551908>
- Andreassen, L. M. (2022). Breer og fonner i Norge. *Norges vassdrags- og energidirektorat*.
- Andreassen, L. M., Elvehøy, H., Kjøllmoen, B., & Belart, J. M. C. (2020). Glacier change in Norway since the 1960s – an overview of mass balance, area, length and surface elevation changes. *Journal of Glaciology*, 66(256), 313–328. <https://doi.org/10.1017/jog.2020.10>
- Andreassen, L. M., Elvehøy, H., Kjøllmoen, B., Engeset, R. V., & Haakensen, N. (2005). Glacier mass-balance and length variation in Norway. *Annals of Glaciology*, 42, 317–325. <https://doi.org/10.3189/172756405781812826>
- Andreassen, L. M., Moholdt, G., Kääb, A., Messeril, A., Nagy, T., & Winsvold, S. H. (2021). Monitoring glaciers in mainland Norway and Svalbard using Sentinel. *NVE Rapport 3-2021*, 94 s. [https://publikasjoner.nve.no/rapport/2021/rapport2021\\_03.pdf](https://publikasjoner.nve.no/rapport/2021/rapport2021_03.pdf)
- Andreassen, L. M., Robson, B. A., Sjørusen, K. H., Elvehøy, H., Kjøllmoen, B., & Carrivick, J. L. (2023). Spatio-temporal variability in geometry and geodetic mass balance of Jostedalbreen ice cap, Norway. *Annals of Glaciology*, 64(90), 26–43. <https://doi.org/10.1017/aog.2023.70>
- Andreassen, L. M., Winsvold, S. H., Paul, F., & Hausberg, J. E. (2012). Inventory of Norwegian glaciers. <https://doi.org/10.5167/UZH-73855>
- Andreassen, L. M., Engeset, R. V., Melvold, K., Müller, K., Orthe, N. K., Sund, M., & Winsvold, S. H. (2021). Developing Satellite Services for Hydrology, Glaciers and Avalanches in Norway. *NVE Rapport 25-2021*, 43 s. [https://publikasjoner.nve.no/rapport/2021/rapport2021\\_25.pdf](https://publikasjoner.nve.no/rapport/2021/rapport2021_25.pdf)
- Baumann, S., Winkler, S., & Andreassen, L. M. (2009). Mapping glaciers in Jotunheimen, South-Norway, during the "Little Ice Age" maximum. *The Cryosphere*, 3(2), 231–243. <https://doi.org/10.5194/tc-3-231-2009>

- Cardoso-Fernandes, J., Teodoro, A. C., Lima, A., & Roda-Robles, E. (2020). Semi-Automatization of Support Vector Machines to Map Lithium (Li) Bearing Pegmatites. *Remote Sensing*, 12(14), 2319. <https://doi.org/10.3390/rs12142319>
- ESA. (2024a). GNDVI Algorithm Specification. Hentet 3. mai 2024, fra <https://step.esa.int/main/wp-content/help/versions/10.0.0/snap-toolboxes/eu.esa.opt.opttbx.radiometric.indices.ui/gndvi/GndviAlgorithmSpecification.html>
- ESA. (2024b). Introducing Sentinel-2. Hentet 26. april 2024, fra [https://www.esa.int/Applications/Observing\\_the\\_Earth/Copernicus/Sentinel-2/Introducing\\_Sentinel-2](https://www.esa.int/Applications/Observing_the_Earth/Copernicus/Sentinel-2/Introducing_Sentinel-2)
- ESA. (2024c). Orbit - Sentinel 2 - Mission - Sentinel Online. Hentet 3. mai 2024, fra <https://sentinels.copernicus.eu/web/sentinel/missions/sentinel-2/satellite-description/orbit>
- ESA. (2024d). Spatial - Resolutions - Sentinel-2 MSI - User Guides - Sentinel Online. Hentet 25. april 2024, fra <https://sentinels.copernicus.eu/web/sentinel/user-guides/sentinel-2-msi/resolutions/spectral>
- Esri. (2024a). Datumtransformasjon mellom WGS84 og ETRS89 (Euref89) i ArcGIS for Desktop. Hentet 25. april 2024, fra <https://www.geodata.no/guider/datumtransformasjon-mellom-wgs84-og-etrs89-euref89-i-arcgis-for-desktop>
- Esri. (2024b). GNDVI—ArcGIS Pro | Documentation. Hentet 3. mai 2024, fra <https://pro.arcgis.com/en/pro-app/latest/arcpy/image-analyst/gndvi.htm>
- Esri. (2024c). How Iso Cluster works—ArcGIS Pro | Documentation. Hentet 3. mai 2024, fra <https://pro.arcgis.com/en/pro-app/latest/tool-reference/spatial-analyst/how-iso-cluster-works.htm>
- Esri. (2024d). Iso Cluster Unsupervised Classification (Spatial Analyst)—ArcGIS Pro | Documentation. Hentet 3. mai 2024, fra <https://pro.arcgis.com/en/pro-app/latest/tool-reference/spatial-analyst/iso-cluster-unsupervised-classification.htm>
- Esri. (2024e). Train Random Trees Classifier (Spatial Analyst)—ArcGIS Pro | Documentation. Hentet 7. mai 2024, fra <https://pro.arcgis.com/en/pro-app/latest/tool-reference/spatial-analyst/train-random-trees-classifier.htm>
- Gascon, F., Bouzinac, C., Thépaut, O., Jung, M., Francesconi, B., Louis, J., Lonjou, V., Lafrance, B., Massera, S., Gaudel-Vacaresse, A., Languille, F., Alhammoud, B., Viallefont, F., Pflug, B., Bieniarz, J., Clerc, S., Pessiot, L., Trémas, T., Cadau, E., ... Fernandez, V. (2017). Copernicus Sentinel-2A Calibration and Products Validation Status. *Remote Sensing*, 9(6), 584. <https://doi.org/10.3390/rs9060584>
- GISGeography. (2014 mars). Spectral Signature Cheatsheet in Remote Sensing. Hentet 2. mai 2024, fra <https://gisgeography.com/spectral-signature/>
- GISGeography. (2017 mai). What is NDVI (Normalized Difference Vegetation Index). Hentet 25. april 2024, fra <https://gisgeography.com/ndvi-normalized-difference-vegetation-index/>
- Gitelson, A., Merzlyak, M., & Grits, Y. (1996). Novel algorithms for remote sensing of chlorophyll content in higher plant leaves. *IGARSS '96. 1996 International Geoscience and Remote Sensing Symposium*, 4, 2355–2357. <https://doi.org/10.1109/IGARSS.1996.516985>
- Hoang, G., Bouzerdoum, A., & Lam, S. (2009 oktober). Learning Pattern Classification Tasks with Imbalanced Data Sets. I P.-Y. Yin (Red.), *Pattern Recognition*. InTech. <https://doi.org/10.5772/7544>
- Khan, A. A., Jamil, A., Hussain, D., Taj, M., Jabeen, G., & Malik, M. K. (2020). Machine-Learning Algorithms for Mapping Debris-Covered Glaciers: The Hunza Basin Case

- Study. *IEEE Access*, 8, 12725–12734. <https://doi.org/10.1109/ACCESS.2020.2965768>
- Khan, M. Y., Qayoom, A., Nizami, M. S., Siddiqui, M. S., Wasi, S., & Raazi, S. M. K.-R. (2021). Automated Prediction of Good Dictionary EXamples (GDEX): A Comprehensive Experiment with Distant Supervision, Machine Learning, and Word Embedding-Based Deep Learning Techniques (S. Sarfraz, Red.). *Complexity*, 2021, 1–18. <https://doi.org/10.1155/2021/2553199>
- Liestøl, O. (1963). Et senglacialt brefremstøt ved Hardangerjøkulen. *Norsk Polarinstitutt Årbok 1962*, 132–139.
- Nesje, A., Dahl, S. O., Thun, T., & Nordli, Ø. (2008). The ‘Little Ice Age’ glacial expansion in western Scandinavia: summer temperature or winter precipitation? *Climate Dynamics*, 30(7-8), 789–801. <https://doi.org/10.1007/s00382-007-0324-z>
- Norgebilder.no. (2024). Norge i Bilder. Hentet 29. april 2024, fra <https://www.norgebilder.no/>
- Palacios, D., Hughes, P. D., Jomelli, V., & Tanarro, L. M. (Red.). (2024). *European glacial landscapes: the holocene*. Elsevier.
- Patel, S. (2017 mai). Chapter 2 : SVM (Support Vector Machine) — Theory. Hentet 3. mai 2024, fra <https://medium.com/machine-learning-101/chapter-2-svm-support-vector-machine-theory-f0812effc72>
- Paul, F., & Kääb, A. (2005). Perspectives on the production of a glacier inventory from multispectral satellite data in Arctic Canada: Cumberland Peninsula, Baffin Island. *Annals of Glaciology*, 42, 59–66. <https://doi.org/10.3189/172756405781813087>
- Paul, F., Winsvold, S., Kääb, A., Nagler, T., & Schwaizer, G. (2016). Glacier Remote Sensing Using Sentinel-2. Part II: Mapping Glacier Extents and Surface Facies, and Comparison to Landsat 8. *Remote Sensing*, 8(7), 575. <https://doi.org/10.3390/rs8070575>
- Raschka, S., & Mirjalili, V. (2019). *Python machine learning: machine learning and deep learning with Python, scikit-learn, and TensorFlow 2* (Third edition). Packt.
- Sentinel-hub. (2024). Sentinel-2 L2A. Hentet 18. april 2024, fra <https://docs.sentinel-hub.com/api/latest/data/sentinel-2-l2a/>
- SIC. (2024). Sentinel-2A SatelliteSensor | Satellite Imaging Corp. Hentet 26. april 2024, fra <https://www.satimagingcorp.com/satellite-sensors/other-satellite-sensors/sentinel-2a/>
- Sinergise, S.-H. b. (2024a). Barren Soil Script. Hentet 2. mai 2024, fra [https://custom-scripts.sentinel-hub.com/custom-scripts/sentinel-2/barren\\_soil/](https://custom-scripts.sentinel-hub.com/custom-scripts/sentinel-2/barren_soil/)
- Sinergise, S.-H. b. (2024b). Sentinel-2 Bands. Hentet 2. mai 2024, fra <https://custom-scripts.sentinel-hub.com/custom-scripts/sentinel-2/bands/>
- Sykas, D. (2024). Spectral Indices with multispectral satellite data. Hentet 2. mai 2024, fra <https://www.geo.university/pages/blog?p=spectral-indices-with-multispectral-satellite-data>
- Tvede, A. M. (1973). Folgefonni - en glasiologisk avvikler. *Naturen*, 97(1), 11–16.
- Uutisynet. (2024 april). Bruk av farger | Tilsynet for universell utforming av ikt. Hentet 29. april 2024, fra <https://www.uutisynet.no/veiledning/bruk-av-farger/206>
- Weber, P. (2020). Ice-marginal processes and retreat dynamics of Norwegian plateau ice-fields . [Unpublished doctoral thesis] *University of Portsmouth*.
- Weber, P., Boston, C. M., Lovell, H., & Andreassen, L. M. (2019). Evolution of the Norwegian plateau icefield Hardangerjøkulen since the ‘Little Ice Age’. *The Holocene*, 29(12), 1885–1905. <https://doi.org/10.1177/0959683619865601>



- Winsvold, S. H., Andreassen, L. M., & Kienholz, C. (2014). Glacier area and length changes in Norway from repeat inventories. *The Cryosphere*, 8(5), 1885–1903. <https://doi.org/10.5194/tc-8-1885-2014>
- Yan, S., Xu, L., Yu, G., Yang, L., Yun, W., Zhu, D., Ye, S., & Yao, X. (2021). Glacier classification from Sentinel-2 imagery using spatial-spectral attention convolutional model. *International Journal of Applied Earth Observation and Geoinformation*, 102, 102445. <https://doi.org/10.1016/j.jag.2021.102445>
- Zemp, M., Frey, H., Gärtner-Roer, I., Nussbaumer, S. U., Hoelzle, M., Paul, F., Haeberli, W., Denzinger, F., Ahlstrøm, A. P., Anderson, B., Bajracharya, S., Baroni, C., Braun, L. N., Cáceres, B. E., Casassa, G., Cobos, G., Dávila, L. R., Delgado Granados, H., Demuth, M. N., ... Vincent, C. (2015). Historically unprecedented global glacier decline in the early 21st century. *Journal of Glaciology*, 61(228), 745–762. <https://doi.org/10.3189/2015JoG15J017>

## Effekten av treningsdata

Nedenfor følger visuelle og analytiske resultater etter test av effekten til et treningsdatasett. Først fremvises resultatene for SVM for to forskjellige treningsdatasett, før resultatene fra RT følger etter. Treningsdata 1 er datasettet som ble benyttet i oppgaven.

	Nøyaktighet
SVM 1	0.79 - 79%
SVM 2	0.78 - 78%
RT 1	0.77 - 77%
RT 2	0.77 - 77%

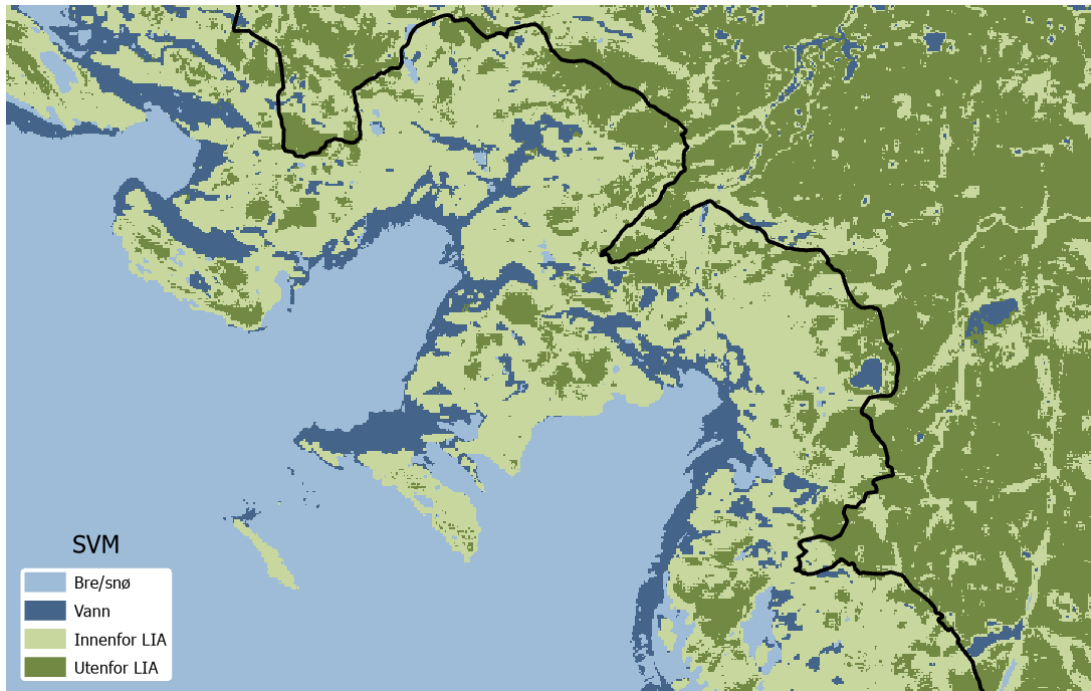
**Tabell A.0.1:** *Den totale nøyaktigheten for SVM og RT for treningsdata 1 og 2.*

	Presisjon		Gjenkall		F1-verdi	
	SVM 1	SVM 2	SVM 1	SVM 2	SVM 1	SVM 2
Bre/snø	0.90	0.90	1.00	1.00	0.95	0.95
Vann	0.61	0.57	0.12	0.11	0.20	0.19
Innenfor LIA	0.70	0.76	0.73	0.69	0.73	0.72
Utenfor LIA	0.74	0.64	0.75	0.78	0.75	0.70

**Tabell A.0.2:** Presisjon, gjenkall og F1-verdi for SVM innad i hver klasse for treningsdata 1 og 2.

	Presisjon		Gjenkall		F1-verdi	
	RT 1	RT 2	RT 1	RT 2	RT 1	RT 2
Bre/snø	0.91	0.89	0.99	1.00	0.94	0.94
Vann	0.64	0.61	0.11	0.11	0.19	0.19
Innenfor LIA	0.70	0.81	0.70	0.68	0.70	0.74
Utenfor LIA	0.68	0.59	0.77	0.80	0.72	0.68

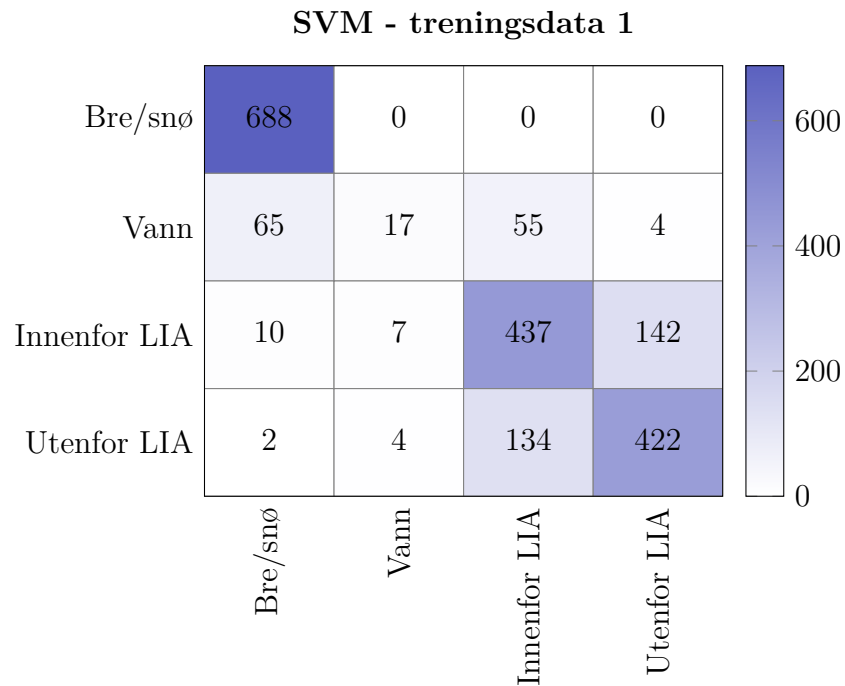
**Tabell A.0.3:** Presisjon, gjenkall og F1-verdi for RT innad i hver klasse for treningsdata 1 og 2.



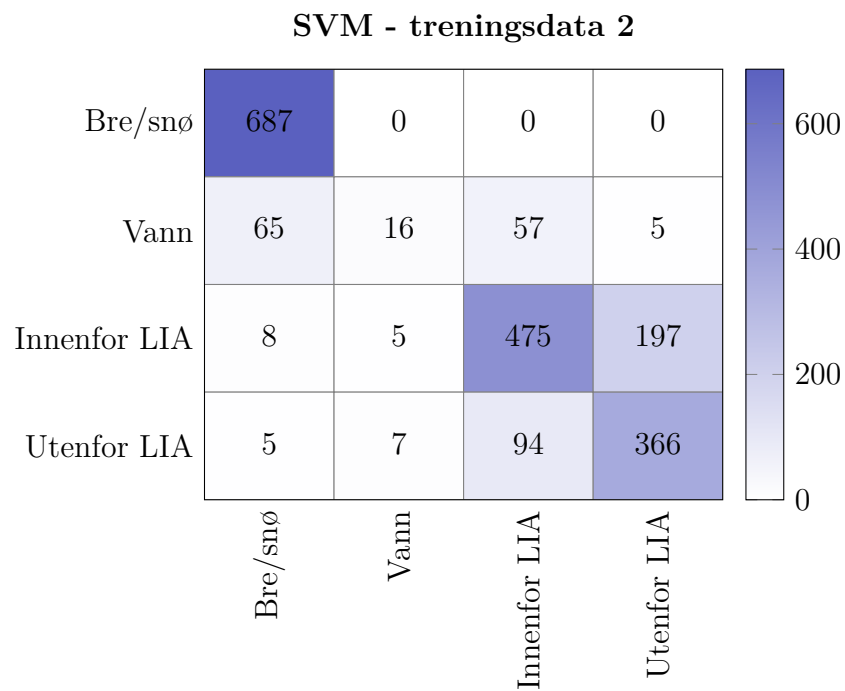
**Figur A.0.1:** Resultatet etter klassifisering med SVM på treningsområdet Midtdalsbreen og Blåisen med treningsdatasett 1.



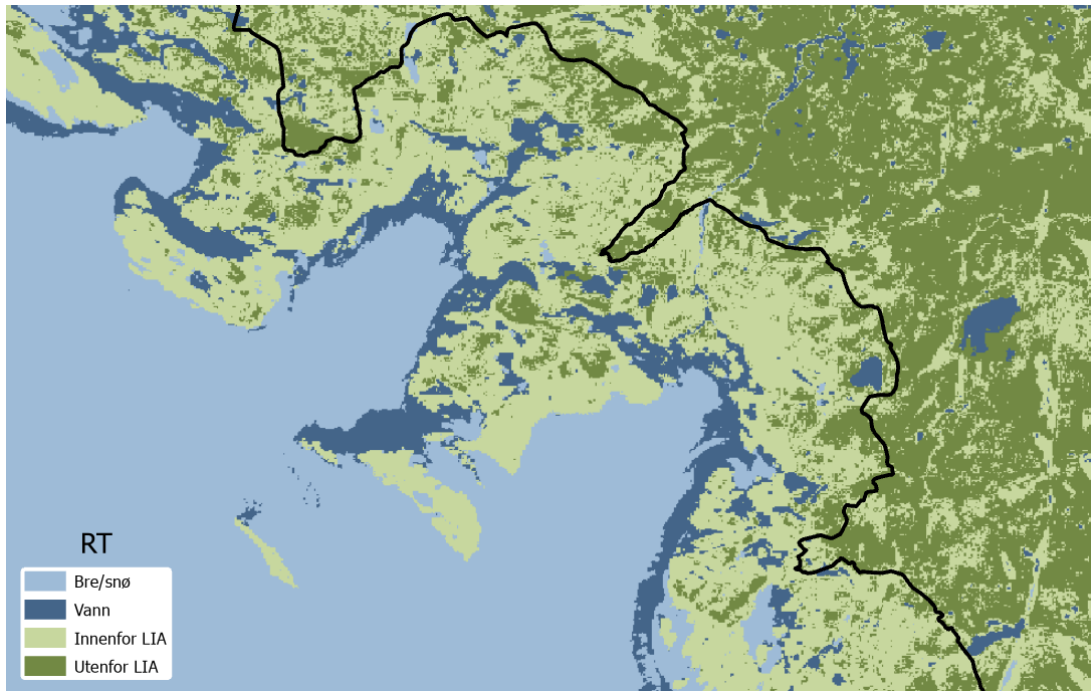
**Figur A.0.2:** Resultatet etter klassifisering med SVM på treningsområdet Midtdalsbreen og Blåisen med treningsdatasett 2.



**Figur A.0.3:** SVM forvirringsmatrise for treningsdata 1.



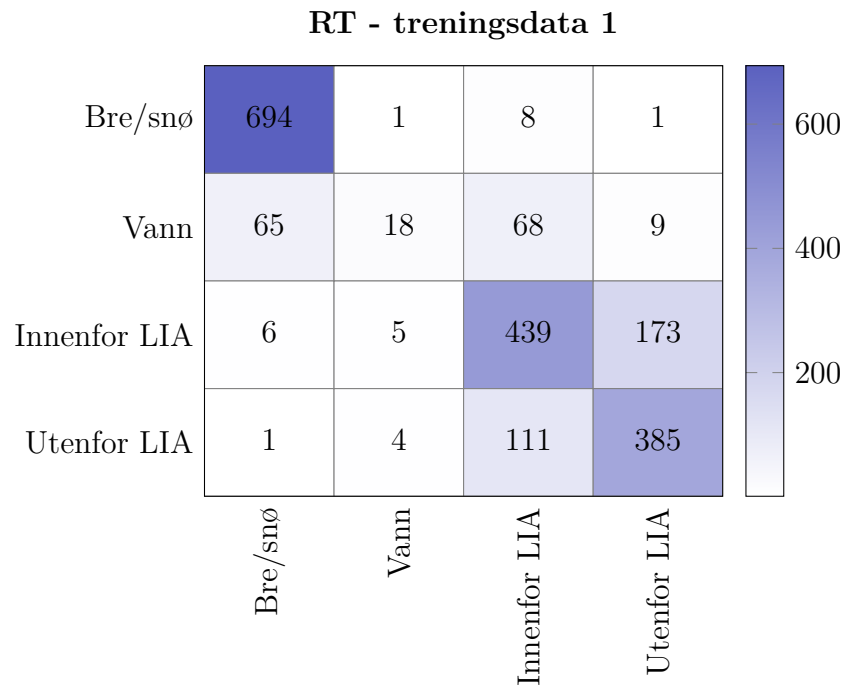
**Figur A.0.4:** SVM forvirringsmatrise for treningsdata 2.



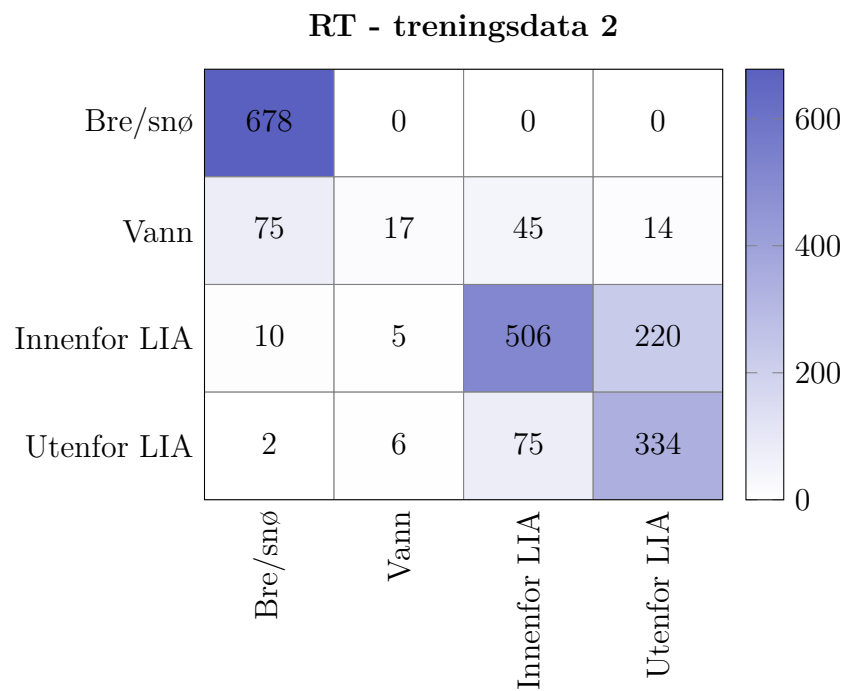
**Figur A.0.5:** Resultatet etter klassifisering med RT på treningsområdet Midtdalsbreen og Blåisen med treningsdatasett 1.



**Figur A.0.6:** Resultatet etter klassifisering med RT på treningsområdet Midtdalsbreen og Blåisen med treningsdatasett 2.



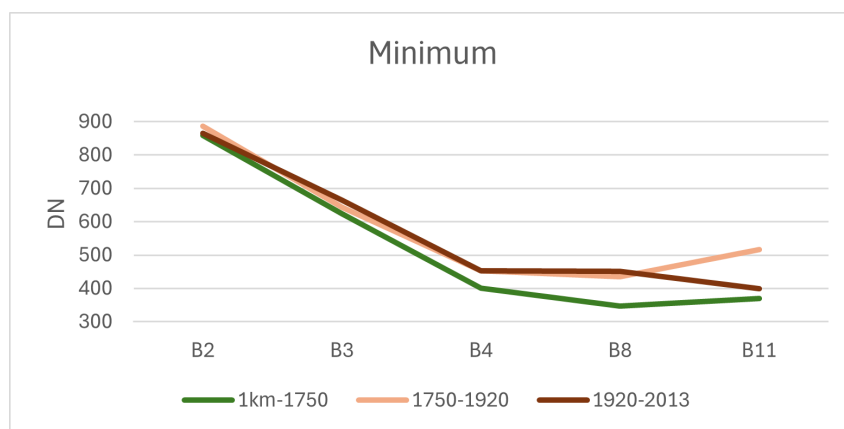
**Figur A.0.7:** RT forvirringsmatrise for treningsdata 1.



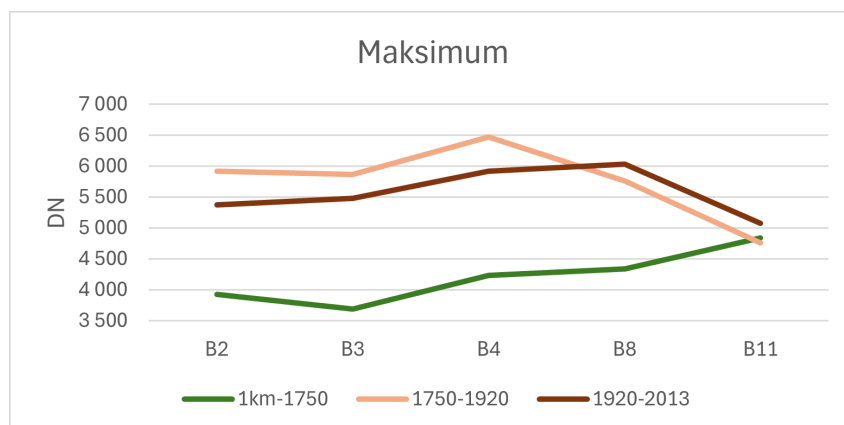
**Figur A.0.8:** RT forvirringsmatrise for treningsdata 2.

## Minimum- og maksimumsverdier

I dette tillegget kan nærmere undersøkelse av minimum- og maksimumsverdier for polygoner, punktlinje og brearmer studeres.



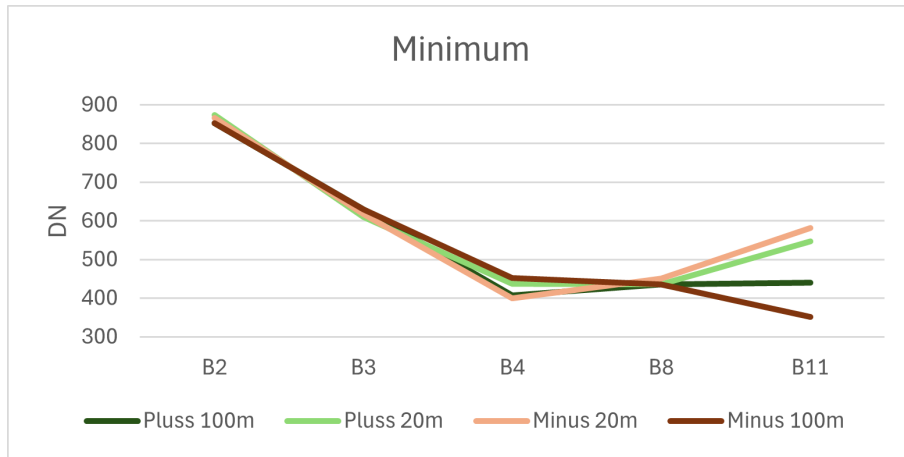
(a)



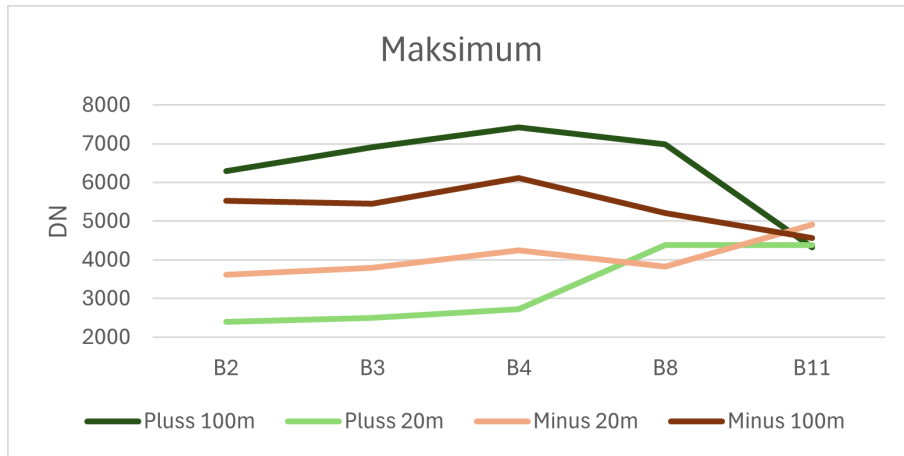
(b)

**Figur B.0.1:** (a) og (b) representerer minimum- og maksimumsverdier for de tre polygonene.



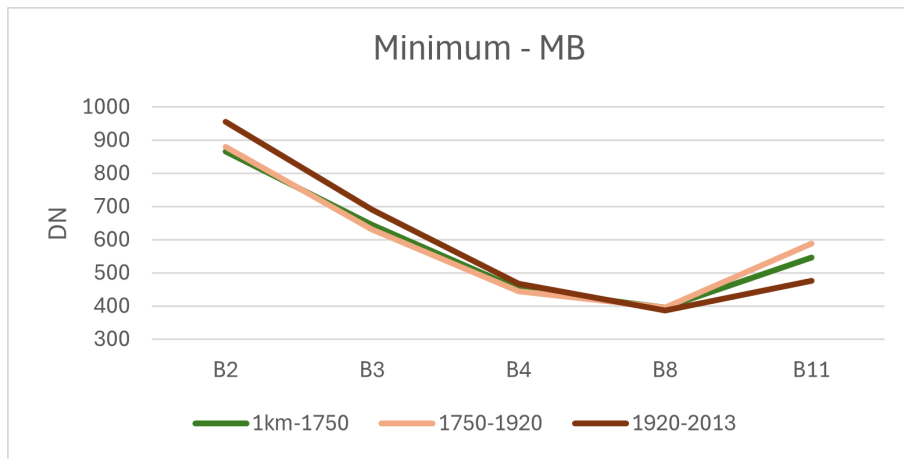


(a)

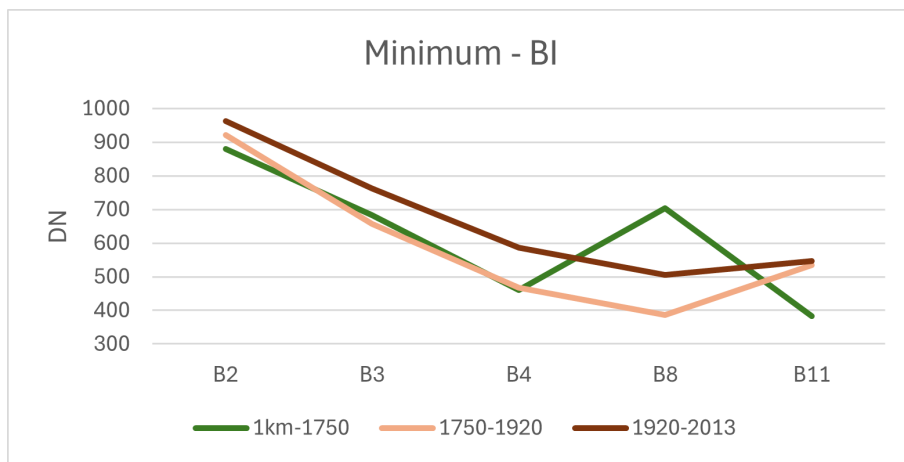


(b)

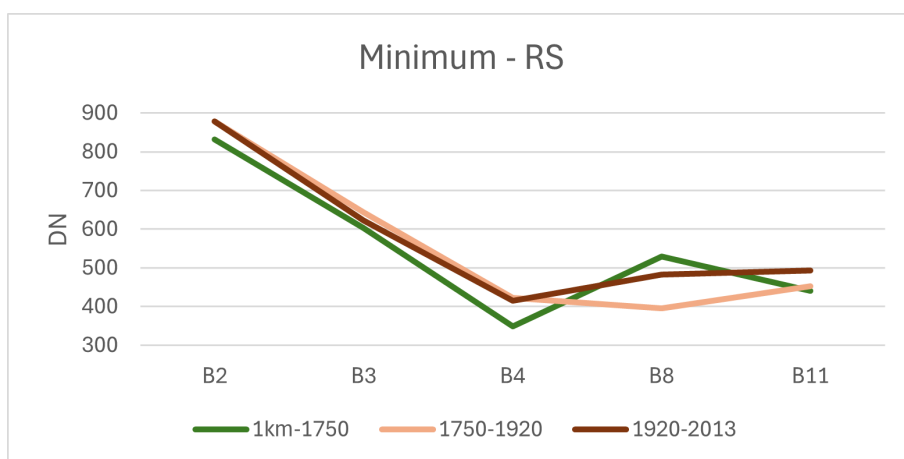
**Figur B.0.2:** (a) og (b) representerer minimum- og maksimumsverdier for de fire punktlinjene.



(a) *Midtdalsbreen (MB)*

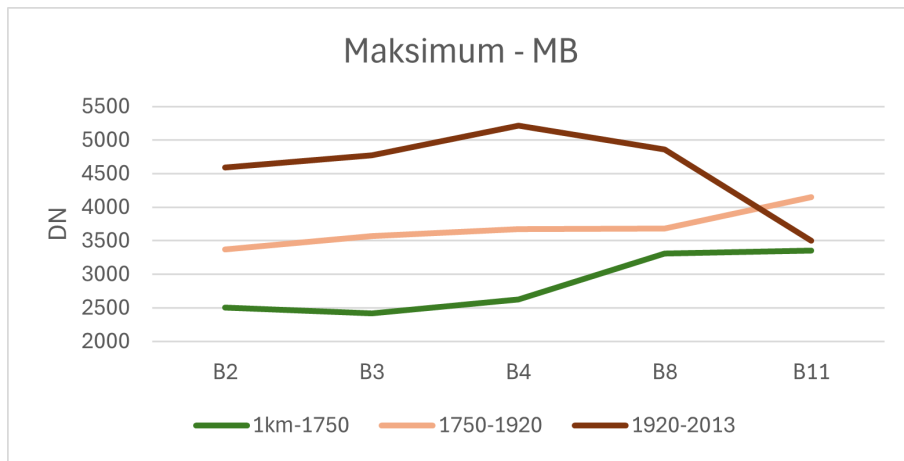


(b) *Blåisen (BI)*

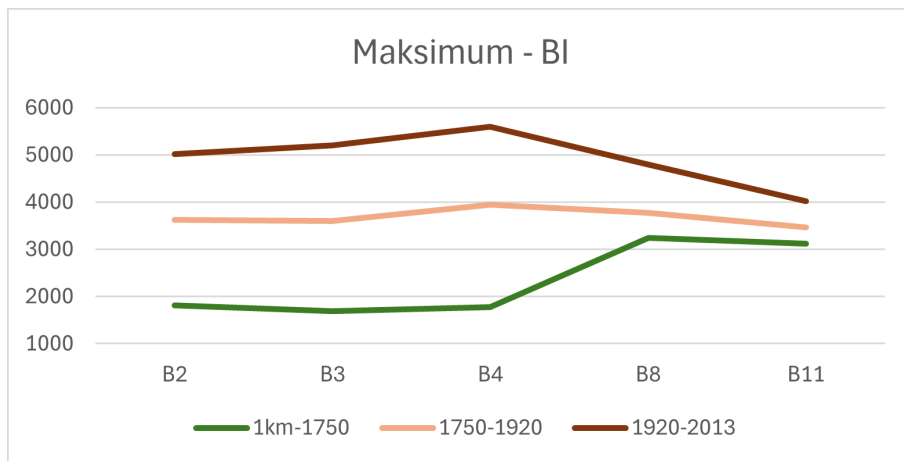


(c) *Rembedalskåka (RS)*

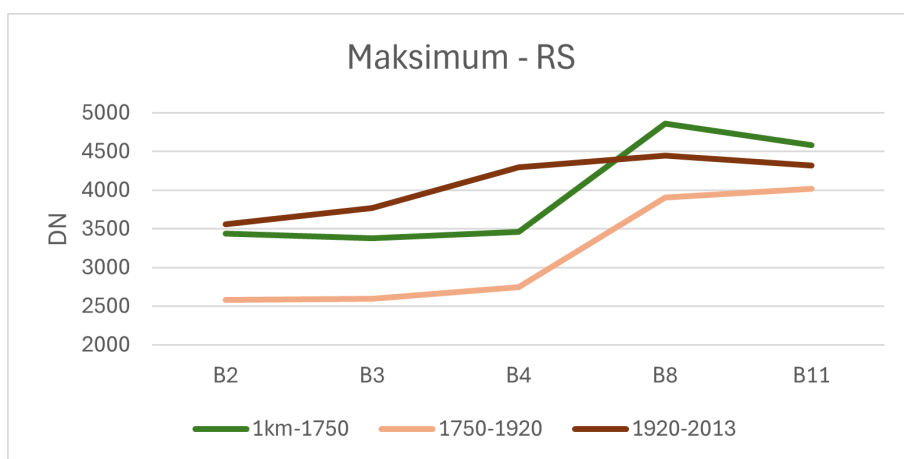
**Figur B.0.3:** (a), (b) og (c) representerer minimumsverdiene for brearmene Midtdalsbreen (MB), Blåisen (BI) og Rembedalskåka (RS).



(a) *Middalsbreen (MB)*



(b) *Blåisen (BI)*



(c) *Rembesdalskåka (RS)*

**Figur B.0.4:** (a), (b) og (c) representerer maksimumsverdiene for brearmene *Middalsbreen (MB)*, *Blåisen (BI)* og *Rembesdalskåka (RS)*.

## Korrelasjon

Nedenfor presenteres korrelasjonsanalysen for hvert individuelle bånd innenfor polygonene og langs punktlinjene. Resultatene fra brearmene viser en likhet med korrelasjonene for polygonene, og er derfor utelatt fra denne presentasjonen.

**Korrelasjon - Polygon**

<b>B2</b>	1km-1750	1750-1920	1920-2013
1km-1750	1		
1750-1920	-0.011	1	
1920-2013	-0.010	0.009	1

<b>B3</b>	1km-1750	1750-1920	1920-2013
1km-1750	1		
1750-1920	-0.017	1	
1920-2013	-0.009	0.010	1

<b>B4</b>	1km-1750	1750-1920	1920-2013
1km-1750	1		
1750-1920	-0.018	1	
1920-2013	-0.010	0.014	1

<b>B8</b>	1km-1750	1750-1920	1920-2013
1km-1750	1		
1750-1920	-0.018	1	
1920-2013	-0.022	0.013	1

<b>B11</b>	1km-1750	1750-1920	1920-2013
1km-1750	1		
1750-1920	-0.020	1	
1920-2013	-0.026	0.012	1

**Tabell C.0.1:** Korrelasjon mellom polygonene for de utvalgte Sentinel-2 båndene.

### Korrelasjon - Linje

<b>B2</b>	Pluss 100m	Pluss 20m	Minus 20m	Minus 100m
Pluss 100m	1			
Pluss 20m	-0.115	1		
Minus 20m	-0.085	0.137	1	
Minus 100m	-0.050	0.048	0.065	1

<b>B3</b>	Pluss 100m	Pluss 20m	Minus 20m	Minus 100m
Pluss 100m	1			
Pluss 20m	-0.105	1		
Minus 20m	-0.091	0.172	1	
Minus 100m	-0.053	0.012	0.059	1

<b>B4</b>	Pluss 100m	Pluss 20m	Minus 20m	Minus 100m
Pluss 100m	1			
Pluss 20m	-0.133	1		
Minus 20m	-0.101	0.126	1	
Minus 100m	-0.052	0.017	0.048	1

<b>B8</b>	Pluss 100m	Pluss 20m	Minus 20m	Minus 100m
Pluss 100m	1			
Pluss 20m	0.090	1		
Minus 20m	0.085	0.166	1	
Minus 100m	-0.005	0.024	0.050	1

<b>B11</b>	Pluss 100m	Pluss 20m	Minus 20m	Minus 100m
Pluss 100m	1			
Pluss 20m	-0.107	1		
Minus 20m	-0.070	0.260	1	
Minus 100m	-0.015	0.035	0.076	1

**Tabell C.0.2:** *Korrelasjon mellom punktene langs linjene for de utvalgte Sentinel-2 båndene.*





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

Postboks 5003  
NO-1432 Ås  
Norway