



Norges miljø- og  
biovitenskapelige  
universitet

**Masteroppgave 2023 30 stp**  
Fakultet for landskap og samfunn

# **Sammenheng mellom urbane grønne områder og Covid-19 smitte i Norge - en spatiotemporal Bayes' SIR modell for Oslo og Bergen**

Association between urban green space in Norway  
and transmission of Covid-19- A Bayesian SIR  
modeling approach

Halvor Kjellesvig  
Folkehelsevitenskap

# 1 Forord

Året 2020 ble et annerledes år på svært mange måter. For meg innebar pandemien blant annet at jeg tidlig fikk bidra med smittesporing. Dette ga et spennende innblikk i hvordan Covid-19 spredte seg i befolkningen. Dette var nok med på å inspirere til å søke masterstudiet i Folkehelsevitenskap i en kort rolig stund den våren, mens jeg planla for hvordan jeg skulle komme meg gjennom nedstengninger den kommende vinteren. Det følte også naturlig å vurdere temaer for masteroppgaven som omhandlet Covid-19 når den tiden kom.

Denne masteroppgaven består av en kappe og en artikkel. Artikkelen er mitt forskningsarbeid.

I dag vil jeg takke min arbeidsgiver som alltid har vært støttende og tilrettelagt for at jeg har kunnet kombinere jobb og studier. Og ikke minst min veileder professor Geir Aamodt for hans entusiasme, kunnskap og dyktige veiledning.

Halvor Kjellesvig

15.05.2023

## 2 Sammendrag

**Bakgrunn:** Tilgang til grønne omgivelser kan gi muligheter for fysisk aktivitet og sosial interaksjon i urbane områder i perioder med strenge retningslinjer for sosial distansering, fordi Covid-19 smitte reduseres i ventilerte områder. Formålet med denne studien er å utforske sammenhengen mellom tilgang til grønne områder og utviklingen av Covid-19 smitte på bydelsnivå i Oslo og Bergen gjennom flere bølger av pandemien.

**Metode:** Vi benyttet daglig publiserte antall positive PCR Covid-19 laborietester på bydelsnivå gjennom fire bølger av pandemien, fra 15. oktober 2020 til 15. april 2022, i Oslo og Bergen. Vi inkluderte befolkningens tilgang til grønne omgivelser i deres nærområde ved to objektive beregninger: prosentandel av grønt areal (%Ga) og gjennomsnitt mengde grønn vegetasjon Normalized Difference Vegetation Index (NDVI). Andelen av befolkningen med lavinntekt og forventet levealder ble inkludert som sosiodemografiske variabler. En Bayes' *Susceptible-Infected-Removed* (SIR) modell ble benyttet for å utnytte tilgangen til de daglig oppdaterte dataene av Covid-19 smitte og samtidig ta hensyn til romlig og temporære avhengigheter i den statistiske analysen.

**Resultater:** Vi fant ingen signifikant forskjell i sammenhengen mellom tilgang til grønne omgivelser og Covid-19 smitte for %Ga eller NDVI. Vi fant bare signifikant sammenheng mellom tilgang til grønne omgivelser og Covid-19 smitte i den 4. bølgen i Oslo. Her ble økt tilgang til grønne omgivelser sett i sammenheng med økt smitte av Covid-19. I Oslo økte Covid-19 smitte signifikant med andelen av befolkningen med lavinntekt under bølge to og tre. I Bergen ble det ikke funnet sammenheng mellom tilgang til grønne omgivelser og Covid-19 smitte ved noen av bølgene.

**Konklusjon:** Vi fant at sammenheng mellom tilgang til grønne omgivelser og Covid-19 smitte var liten sammenlignet med faktorer som sosioøkonomi og underliggende smitte.

## Innholdsfortegnelse

<b>1</b>	<b>Forord</b> .....	<b>1</b>
<b>2</b>	<b>Sammendrag</b> .....	<b>2</b>
<b>3</b>	<b>Innledning</b> .....	<b>5</b>
<b>4</b>	<b>Problemstilling</b> .....	<b>6</b>
<b>5</b>	<b>Bakgrunn</b> .....	<b>6</b>
5.1	<i>Helse og urbane grønne omgivelser</i> .....	6
5.2	<i>Hvordan tilgang til natur kan påvirke helse</i> .....	7
5.3	<i>Egenskaper ved grønne omgivelser</i> .....	8
5.4	<i>Sammenheng mellom utvikling av Covid-19 og tilgang til urbane grønne omgivelser i internasjonale studier</i> 8	
5.5	<i>Sammenheng mellom sosioøkonomiske forhold Covid-19 smitte og tilgang til grønne omgivelser</i> .....	9
5.6	<i>Utvikling av Covid-19 og tilgang til urbane grønne omgivelser i Norge</i> .....	9
5.7	<i>Oppgavens aktualitet i et folkehelseperspektiv</i> .....	10
<b>6</b>	<b>Metode</b> .....	<b>10</b>
6.1	<i>Studiedesign</i> .....	10
6.2	<i>Populasjon</i> .....	10
6.2.1	<i>Befolkningsdata</i> .....	10
6.2.2	<i>Sosio demografiske variabler</i> .....	11
6.3	<i>Covid-19 insidens</i> .....	11
6.4	<i>Tilgang til grønne urbane omgivelser</i> .....	11
6.4.1	<i>Urbane grønne områder (green area %Ga)</i> .....	11
6.4.2	<i>Vegetasjonsdekke (NDVI)</i> .....	12
6.5	<i>Statistisk analyse</i> .....	13
6.5.1	<i>SIR modell</i> .....	13
6.5.2	<i>Bayesiansk hierarkisk regresjonsmodell</i> .....	14
6.5.3	<i>Spesifikasjon av priors (a priori fordelinger)</i> .....	15
6.5.4	<i>Modell beregning</i> .....	16
6.5.5	<i>Model seleksjon</i> .....	16
6.6	<i>Pandemiens bølger</i> .....	16
6.7	<i>Etikk</i> .....	17
<b>7</b>	<b>Resultater</b> .....	<b>17</b>
7.1	<i>Beskrivende statistikk</i> .....	17
7.2	<i>Asymtomatisk smitte og evaluering av <math>\alpha</math> verdien</i> .....	18
7.3	<i>Forskjell mellom beregning av tilgang til grønne områder ved NDVI og %Ga og Covid-19 smitte.</i> ...	18
7.4	<i>Forskjell mellom byer og bølger</i> .....	18

<b>8</b>	<b>Diskusjon.....</b>	<b>19</b>
8.1	<i>Hovedresultater.....</i>	<i>19</i>
8.2	<i>Styrker og svakheter.....</i>	<i>19</i>
8.3	<i>Konklusjon.....</i>	<i>20</i>
<b>9</b>	<b>Litteratur:.....</b>	<b>21</b>

**Vedlegg:**

- 1. Artikkel:** *Association between urban green space in Norway and transmission of Covid-19- A Bayesian SIR modeling approach*
- 2. Appendix**

### 3 Innledning

Den 11. mars 2020 erklærte Verdens Helseorganisasjon (WHO) at verden var utsatt for en pandemi pga. den globale spredningen av viruset SARS-CoV-2 som forårsaker Covid-19 (WHO, 2020). Dagen etter innførte regjeringen i Norge de mest inngripende tiltakene i fredstid fokusert på å redusere spredning av viruset for å redusere antall dødsfall av Covid-19 og redusere byrden på helsetjenestene (NOU 2021:6). Fram til utsatte grupper av befolkningen fikk utviklet immunitet gjennom vaksinerings innebar slike tiltak, som i andre land, forskjellige grader av sosial distansering og perioder med stengte grenser, hjemmeskole og hjemmekontor (NOU 2022:5, 2022). Slike tiltak kan føre til sosial isolasjon og redusert fysisk aktivitet og til slutt påvirke den mentale og fysiske helsen i befolkningen (Bonati et al., 2022; Park et al., 2022).

Tilgang til grønne områder kan gi mulighet for fysisk aktivitet og sosiale møter i urbane områder, under perioder med strenge sosiale distanseringstiltak, fordi Covid-19 smitte er mindre i ventilerte områder (Bulfone et al., 2021). I urbane områder har tilgang til grønne arealer flere helsefordeler gjennom å bidra til økt fysisk aktivitet, styrke sosial tilhørighet, redusere stress og motvirke luftforurensing (Hartig et al., 2014). Det er også forskning som peker på at tilgang til grønne områder har hatt positiv effekt på mental helse gjennom pandemien (Pouso et al., 2021; Reid et al., 2022).

I tillegg til de positive effektene tilgang til grønne omgivelser har hatt for folkehelsen gjennom pandemien, har flere studier pekt mot at tilgang til grønne omgivelser kan bidra til redusert Covid-19 smitte (Chen et al., 2023; Klompaker et al., 2021; Venter, Z. S. et al., 2021; You & Pan, 2020). To økologiske studier fra to europeiske byer fant motsettende resultater (Pan et al., 2021; Schmiedege et al., 2023). Det er også studier som finner at tilgang til grønne omgivelser ikke er likt fordelt mellom sosioøkonomiske grupper og at dette kan føre til forskjeller mellom grupper i utvikling av smitte Covid-19 og evne til å tilpasse seg smittereduserende tiltak (Gao et al., 2023; Lu et al., 2021; Spotswood et al., 2021).

Covid-19 pandemien har ført til betydelige belastninger på den urbane befolkningen under perioder med strenge smittereduserende tiltak. Dette har tydeliggjort viktigheten av miljøkvaliteter i byer i en stadig mer urbanisert verden (UN-Habitat, 2021). I en slik sammenheng kan oppmerksomhet på hvordan urbane kvaliteter, som tilgang til grønne områder, påvirker motstandsdyktighet i befolkningen bli verdifullt for fremtidig urban planlegging og byutvikling. Denne studien utforsker sammenhengen mellom tilgang til urbane områder og Covid-19 smitte på bydelsnivå i Oslo og Bergen gjennom flere bølger av

pandemien. Tilgang til grønne områder i befolkningens nærområde er beregnet ved å se på to forskjellige kvaliteter: prosent av grønt areal (%Ga) og gjennomsnittlig mengde grønn vegetasjon etter Normalized Difference Vegetation Index (NDVI). En Bayes' Susceptible-Infected-Removed (SIR) modell er benyttet for å utnytte tilgangen til de daglig oppdaterte dataene av Covid-19 smitte og samtidig ta hensyn til romlig og temporære avhengigheter i den statistiske analysen.

## 4 Problemstilling

Forskningsspørsmål:

- *Ved å benytte en Bayes' SIR-modell og kontrollere for sosioøkonomiske karakteristikk vil denne studien utforske om det er sammenheng mellom tilgang til urbane grønne omgivelser og utvikling av Covid-19.*
- *Videre vil jeg undersøke om det er forskjell på sammenheng mellom tilgang til urbane grønne omgivelser og utvikling av covid-19 i forskjellige bølger av pandemien?*
- *Til slutt vil jeg undersøke om det er forskjell på sammenheng mellom tilgang til urbane grønne omgivelser og utvikling av covid-19 om vi måler grønne omgivelser gjennom %Ga eller NDVI?*

## 5 Bakgrunn

I dette kapittelet vil jeg først kort presentere perspektiver på hvordan tilgang til grønne omgivelser blir sett i sammenheng med helse generelt. Så vil jeg presentere forskning som ser på sammenheng mellom tilgang til grønne omgivelser og smitteutvikling av Covid-19. Deretter vil jeg kort peke på enkelte forhold ved utviklingen av Covid-19 i Norge. Til slutt vil jeg peke på oppgavens aktualitet i et folkehelseperspektiv.

### 5.1 Helse og urbane grønne omgivelser

Ideer om hvordan urbane grønne omgivelser kan påvirke helse og sosiale ulikheter er diskutert av landskapsarkitekt Frederic Law Olmsted da han utviklet Central Park i New York. I Norge utviklet byplanlegger og arkitekten Harald Hals Ullevål Hageby etter tanker fra Ebenezer Howard (Butenschøn, 2013; Szczygiel & Hewitt, 2000). Howard utviklet ideer om hagebyer som et tiltak mot de vanskelige levekårene i London under den industrielle revolusjonen (Butenschøn, 2013).

I dag har livsstilsykdommer utviklet seg til å bli et stort folkehelseproblem og forskning på urbane grønne omgivelser sin påvirkning på fysisk aktivitet, overvekt og stress er blitt utbredt (van den Berg et al., 2015). Van den Berg et al. (2015) gjennomførte en systematisk litteraturgjennomgang av studier som så på sammenhengen mellom grønne omgivelser og helseutfall. Litteraturgjennomgangen fant en signifikant positiv sammenheng mellom tilgang til grønne omgivelser og egenopplevd mental helse og generell dødelighet. De fant også en moderat signifikant positiv sammenheng mellom tilgang til grønne omgivelser og egenopplevd generell helse.

## 5.2 Hvordan tilgang til natur kan påvirke helse

Hartig et al. (2014) pekte i en kunnskapsoppsummering om natur og helse i urbane strøk på luftkvalitet, fysisk aktivitet, sosialt samhold og reduksjon av stress som viktige områder der natur kan påvirke helse. Enkelte av disse er vist i *Figur 1*.

Kunnskapsoppsummeringen peker på at trær og grønne vekster kan påvirke luftkvaliteten og bidra til helse ved å redusere enkelte former for forurensing som ozon, nitrogenoksid og luftpartikler, samt bidra til å regulere temperatur gjennom skygge. Samtidig kan grønn vegetasjon, gjennom enkelte arters utslipp av pollen, være en helsebelastning for allergikere (Hartig et al., 2014).

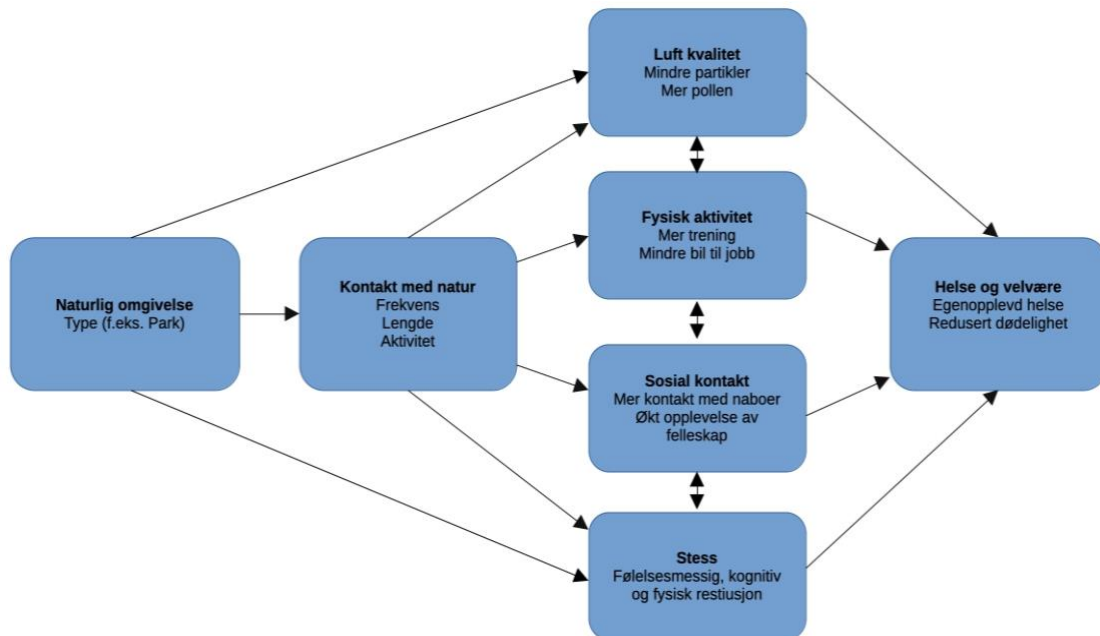
Grønne omgivelser kan også bidra til å øke grad av fysisk aktivitet i befolkningen (Fong et al., 2018). Hartig et al. (2014) peker på at helseutbyttet ved økt aktivitet er størst hos den gruppen som i utgangspunktet er minst aktiv. Grønne omgivelser kan bidra til å øke fysisk aktivitet på flere måter. Først gjennom å tilby områder der fysiske aktiviteter som idrett kan gjennomføres, men også ved at grønne omgivelser i seg selv er attraktive å bevege seg i på grunn av opplevelsen det gir. For å øke fysisk aktivitet som sykling og gange i transportøyemed, er transportinfrastruktur i form sykkel og gangstier langt viktigere enn grad av grønne omgivelser (Hartig et al., 2014).

Hartig et al. (2014) påpeker også at mye forskning dokumenterer en positiv sammenheng mellom sosiale relasjoner og helse og velvære. Egenskaper ved nærmiljø som tilrettelegger for sosiale arenaer og felleskap kan bidra til å styrke opplevelse av sosial tilhørighet. Sosial tilhørighet preges av en opplevelse av å dele normer og verdier og følelse av tilhørighet og aksept i en gruppe. Det pekes derfor på at grønne omgivelser er en egenskap i nærmiljø som kan bidra til å øke grad av sosial tilhørighet.

Den positive betydningen som kontakt med natur har på helseproblemer, relateres også til stress, og er ifølge Hartig et al. (2014) også godt dokumentert. Det pekes på at natur kan



redusere eksponering av utfordrende/stressende egenskaper ved omgivelsene. Dette ved at f.eks. trær kan redusere støypenger fra en vei, skjule lite tiltalende bygningsmasse eller redusere for innsyn og dermed øke opplevelse av privatliv. Tilgang til grønne omgivelser kan også gi rom for mennesker til å trekke seg ut av stressende omgivelser for en periode og på denne måten bidra til å samle krefter.



**Figur 1.** Figur over sammenhenger mellom naturlige omgivelser og helse og velvære (fra Hartig et al. 2014).

### 5.3 Egenskaper ved grønne omgivelser

Forskjellige egenskaper ved grønne omgivelser kan påvirke hvordan vi oppsøker og benytter områder med mye grønt, og dette kan igjen påvirke helseutfall. I en rapport fra WHO (2016) pekes det på at bruken av grønne områder blant annet påvirkes av hvor nære de fra hjemmet, hvor sammenhengende de er og hvordan de er tilrettelagt med fasiliteter som stinett og benker. Tilsvarende funn er observert i Norge (Fongar et al., 2019).

### 5.4 Sammenheng mellom utvikling av Covid-19 og tilgang til urbane grønne omgivelser i internasjonale studier

Internasjonalt er det publisert studier som har undersøkt sammenhenger mellom utvikling av Covid-19 og tilgang til urbane grønne omgivelser. I en studie med globale data finner Venter, Z. S. et al. (2021) at restriksjoner i mobilitet er effektivt mot utvikling av smitte, men at restriksjonene ikke trenger å begrense rekreasjon i grønne omgivelser. You og Pan (2020) fant i en studie av Covid-19 smittetilfeller i 989 fylker i USA en negativ sammenheng mellom grønn urban vegetasjon og Covid-19 smitteutvikling. Klompaker et al. (2021) fant negativ sammenheng mellom vegetasjonstetthet i nærmiljø og covid-19 smitte på fylkesnivå i USA. I

denne studien påpekes det at sammenhengen er sterkere i fylker med høy populasjonstetthet og i fylker med sterke begrensinger på mobilitet. I en studie på bydelsnivå i London for den første bølgen av pandemien, fant Pan et al. (2021) at høyere tilgang til sammenhengende grønne områder var forbundet med økt risiko for Covid-19 smitte. En annen studie på nabolagsnivå i Essen i Tyskland for det første året av pandemien, fant forfatterne en negativ sammenheng mellom tilgang til grønne omgivelser og Covid-19 smitte (Schmiege et al., 2023). I en studie med 2,8 millioner deltakere (Covid Symptom Studien) fant Chen og hans medarbeidere en negativ sammenheng mellom tilgang til grønne omgivelser og Covid-19 smitte (Chen et al., 2023).

### 5.5 Sammenheng mellom sosioøkonomiske forhold Covid-19 smitte og tilgang til grønne omgivelser

Mulighetene til å tilpasse eget liv til de store endringene de smittereduserende tiltakene førte til, er ikke likt fordelt i befolkningen. Forhold som hjemmekontor og størrelse på bolig er eksempler på faktorer som ikke er fordelt likt mellom sosioøkonomiske grupper. Slike faktorer kan derfor ha stor påvirkning på hvordan smitte av Covid-19 utvikler seg. Studier i USA (Gao et al., 2023; Lu et al., 2021; Spotswood et al., 2021) peker på at tilgang til urbane grønne omgivelser heller ikke er likt fordelt i befolkningen og at dette også kan ha påvirkning på hvordan Covid-19 smitte utvikler seg og hvilke forutsetninger forskjellige grupper har for å tilpasse seg belastningen ved smittereduserende tiltak.

### 5.6 Utvikling av Covid-19 og tilgang til urbane grønne omgivelser i Norge

Norske myndigheter valgte tidlig en strategi, med fokus på å redusere utviklingen av covid-19 i befolkningen, som er vurdert vellykket nasjonalt og internasjonalt (Ursin et al., 2020). De smittebegrensende tiltakene i Norge var aldri så inngripende at en ikke kunne tilbringe tid utendørs, så lenge en forholdt seg til gjeldende råd og regler om sosial distansering. I Oslo økte aktivitet i urbane grønne omgivelser betydelig under den første nedstengingen og vedvarte utover i pandemien. (Venter et al., 2020; Venter, Zander S. et al., 2021).

Selv om smittespredningen og dødeligheten av Covid-19 i Norge har vært mindre enn i mange andre land, er utviklingen ujevnt fordelt i befolkningen. En har sett at spredningen har vært særlig stor i de store byene og at enkelte innvandringsgrupper ble særlig rammet (Kjøllesdal et al., 2022).

## 5.7 Oppgavens aktualitet i et folkehelseperspektiv

Covid-19 har gjennom sin store smittespredning ført til mange alvorlige sykdomsforløp og dødsfall som også har ført til stor belastning på helsevesenet. Tiltak iverksatt for å begrense smitte i befolkningen har også ført til store belastninger for samfunnet, både sosialt og økonomisk. Totalt sett har dette hatt store konsekvenser for folkehelsen i Norge og ellers i verden. Det er også sannsynlig at vi må regne med at konsekvensene av pandemien vil ha betydning på folkehelsen i årene som kommer. Koronakommisjonen (NOU 2021:6) peker blant annet på at hjemmeundervisning har ført til læringstap på barne- og mellomtrinnet og at særlig utsatte grupper av ungdom har økt forekomst psykiske belastninger.

Internasjonal forskning peker på at urbane grønne omgivelser gjennom flere virkningsmekanismer kan ha helsebringende egenskaper og dermed styrker folkehelse. Det er også forskning som peker mot at urbane grønne omgivelser har reduserende effekt på utvikling av smitte. Tilgang til urbane grønne omgivelser kan dermed være et virkemiddel for å redusere smitte som også har andre positive egenskaper for folkehelse i seg selv. Resultater fra denne studien vil derfor kunne benyttes til å gi råd til helsemyndigheter ved en eventuell ny pandemi.

## 6 Metode

### 6.1 Studiedesign

I studien benyttes data for antall Covid-19 smittede og sosio-demografiske data på bydelsnivå i Oslo og Bergen. Disse dataene kombineres med geografiske data for arealdekke (%Ga) og vegetasjonsdekke (NDVI) til å gjennomføre en økologisk studie. Økologiske studier benytter data på gruppenivå, noe som ofte er lettere tilgjengelig enn data på individnivå. Økologiske studier er ikke egnet til å trekke kausale slutninger. Dette fordi studiene er sårbare for blant annet konfundere og ved at sammenhenger en observerer på gruppenivå ikke nødvendigvis representerer sammenhenger på individnivå (Webb et al., 2019).

### 6.2 Populasjon

#### 6.2.1 Befolkningsdata

I denne studien vil vi begrense oss til byene Oslo og Bergen og vi vil benytte data fra de siste bølgene. Den første bølgen utelates fordi registreringer av sykdomstilfeller er beheftet med mer usikkerhet enn de siste bølgene i pandemien. Oslo og Bergen er Norges to største byer med henholdsvis 697 010 og 285 601 innbyggere i 2021 (FHI, 2022). Byene er igjen oppdelt i bydeler, Oslo 15 og Bergen 8, som har betydelig lokalforvaltning (Thorsnæs, 2023; Thorsnæs & Thune, 2023). I denne oppgaven benyttes data over befolkningsfordelingen, i form av bosatte

på rutenett a (250m x 250), hentet fra den offentlige kart-katalogen GeoNorge (SSB, 2019). Det geografiske datasettet er produsert av Statistisk sentralbyrå (SSB) på bakgrunn av populasjonsdata i folkeregisteret knyttet opp mot posisjoner i det offentlige eiendomsregisteret (SSB, 2019).

### 6.2.2 Sosio demografiske variabler

Sosiodemografiske og sosioøkonomiske data på bydelsnivå i Oslo og Bergen i 2020 ble hentet fra Folkehelseinstituttet (FHI) (FHI, 2022). I utgangspunktet ble data på prosentandel innvandrere og personer født i Norge av innvandrerforeldre, prosentandel av befolkning med lavinntekt etter EU skala, prosentandel av befolkning med fullført videregående skole og forventet levealder inkludert i modellen. Dessverre ble det nødvendig å kun inkludere andel av befolkning med lavinntekt og forventet levealder. Dette pga. høy korrelasjon mellom de sosioøkonomiske variablene og tilsvarende høy kompleksitet i modellen.

### 6.3 Covid-19 insidens

31. januar blir koronaviruset (SARS-CoV-2) definert som allmennfarlig smittsom sykdom og meldepliktig etter smittevernloven (Helse- og omsorgsdepartementet, 2020). FHI er ansvarlig for MSIS (*MSIS-regulation*, 2003) og har gjennom pandemien løpende tilgjengeliggjort tall for daglige totale positive PCR laboratorietester aggregert til bydelsnivå for Oslo og Bergen og kommunenivå for resterende kommuner. I starten av pandemien var testkapasiteten begrenset, blant annet pga. lav kapasitet på laboratoriene (NOU 2021:6, 2021), det benyttes derfor ikke data fra den første bølgen.

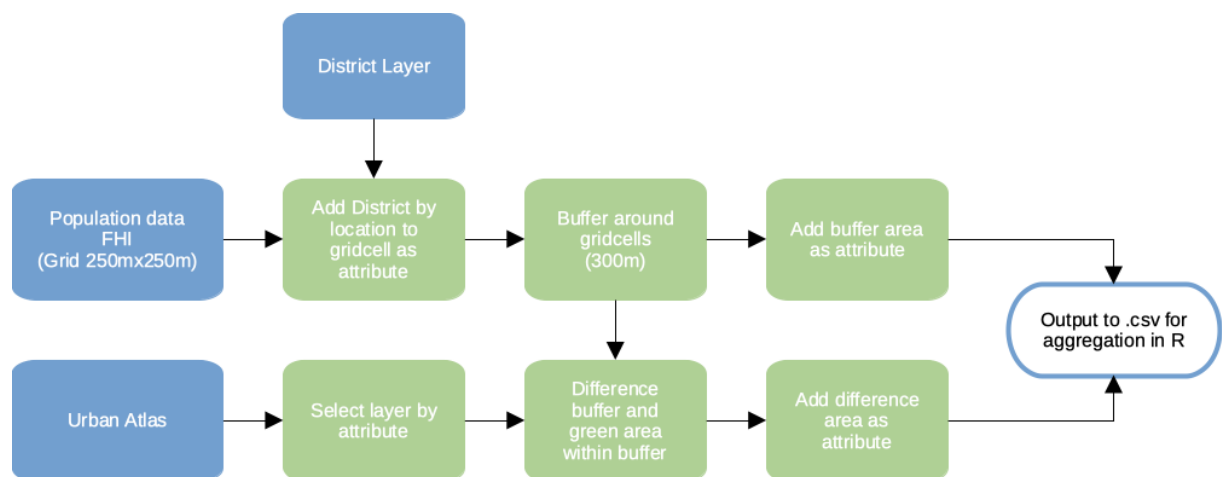
### 6.4 Tilgang til grønne urbane omgivelser

Denne studien benytter to beregninger for å tallfeste befolkningens tilgang til grønne urbane omgivelser; gjennomsnittlig mengde grønn vegetasjon (NDVI) i nærområdet og andel grønne arealer (%Ga) i nærområdet. Fremgangsmåten for å beregne dataene er basert på arbeidet til Barboza et al. (2021), som i sin studie ser på sammenheng mellom tilgang til grønne omgivelser og dødelighet i en rekke europeiske byer, deriblant Oslo og Bergen. Deres studie beskriver hvordan NDVI og %Ga beregnes. Studien benytter seg også av åpent tilgjengelige og godt etablerte datakilder. Ved å velge en tilsvarende fremgangsmåte etterstrebes det at dataene som estimeres for å anslå studiepopulasjonens tilgang til grønne omgivelser vil ha god validitet.

#### 6.4.1 Urbane grønne områder (green area %Ga)

Arealdekkedata fra Urban Atlas kombineres med populasjonsdata fra SSB, fordelt på rutenett, til å beregne en indikator for hvor stor andel (%Ga) av befolkningens nærområde som

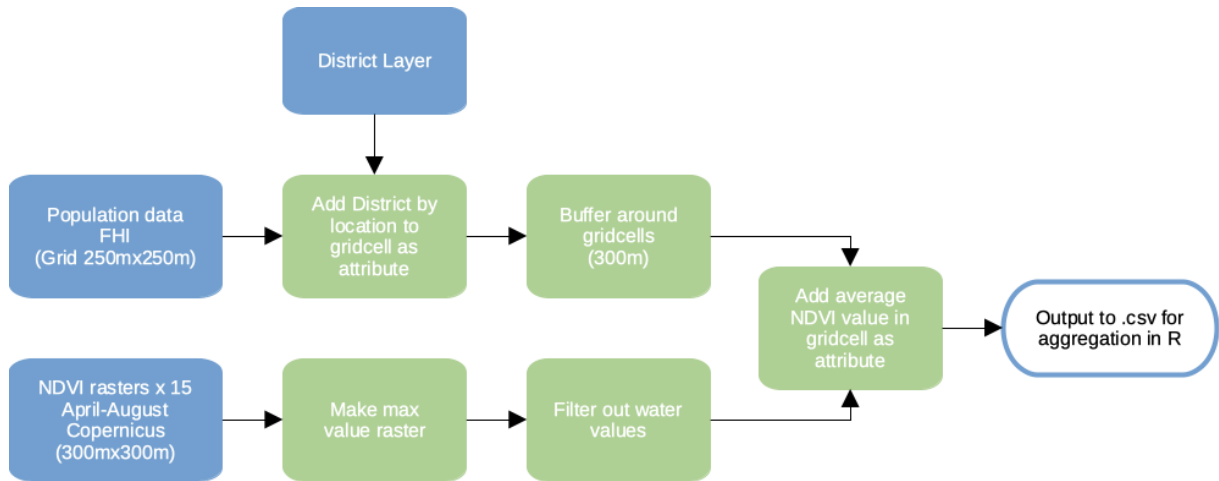
er dekket av grønne omgivelser. Urban Atlas er et prosjekt i det europeiske jordobservasjonsprogrammet Copernicus som leverer sammenlignbare data for urbane områder i Europa. Det nyeste datasettet er fra 2018 og dekker 788 urbane områder i Europa med en oppløsning på 0.25 hektar (Copernicus, 2018). Urban Atlas definerer en rekke kategorier av arealdekke ut fra arealers egenskap og funksjon. I denne studien inkluderes samme kategorier som studien til Barboza et al. (2021) , se seksjon 3.3 i artikkelen. Som enhet for urbant tilgjengelige nærområde brukes en buffersone på 300 meter rundt hver populasjonscelle a 250m x 250m. Denne enheten benyttes også av Barboza et al. (2021) og baserer seg på anbefalinger fra WHO (2016). De geografiske analysene ble utført i QGIS (versjon 3.28.2) og er beskrevet med i et flytskjema i **Figur 2** nedenfor.



**Figur 2.** Flytskjema for beregning av tilgang til grønne arealer i QGIS.

#### 6.4.2 Vegetasjonsdekke (NDVI)

Vi fanger opp vegetasjonsdekke ved å benytte variabelen Normalized Difference Vegetation Index (NDVI). NDVI er basert på satellittbilder og ble samlet av den europeiske jordobservasjon satellitten Sentinel 3. Dataene blir bearbeidet og distribuert via nettportalen Copernicus (2022). Copernicus leverer tre ganger hver måned bilder av NDVI-verdier med en oppløsning på 300m x 300m (raster). Hvert raster er basert på flere satellittmålinger slik at det blir tatt hensyn til feilkilder, som f.eks. skydekke. I denne studien ble NDVI bilder fra perioden april 2021 til august 2021 (n = 15) lastet ned og bearbeidet i QGIS for å finne maks NDVI verdi gjennom året for hvert pixel-området. Verdier som representerer vann, ble så filtrert ut av datasettet. Det ferdig genererte raster-laget ble så brukt til å beregne gjennomsnittlig NDVI verdi i 300m buffersoner rundt populasjonsceller fra SSB. Fremgangsmåten er beskrevet med et flytskjema i **Figur 3**.



Figur 3. Flytskjema for beregning av gjennomsnittlig NDVI verdier i QGIS.

## 6.5 Statistisk analyse

For å utforske sammenhengen mellom tilgang til urbane grønne omgivelser og covid-19 smitte, med utgangspunkt i dataene presentert over, tar denne oppgaven utgangspunkt i en Bayes' hierarkisk SIR (*Susceptible-Infected-Removed*) modell skissert og utarbeidet av Lawson og Kim (2021).

### 6.5.1 SIR modell

SIR-modeller har lenge blitt brukt i epidemiologisk arbeid for å beskrive utbrudd av infeksjonssykdommer og er først beskrevet i 1927 av Kermack og McKendrick (1991). SIR modeller ser på balansen, stegvis gjennom en tidsserie, mellom delen av en populasjon som er mottakelige for smitte (*Susceptible*), delen som er smittet (*Infected*) og delen som ikke lenger er smitteførende (*Removed*). I studien modellerer vi separate bølger i Oslo eller Bergen som et utbrudd. De romlige og temporale enhetene defineres som:  $i = 1, \dots, M$  bydeler og  $j = 1, \dots, T$  dager. Til en tidspunkt  $i$  er den mottakelige delen av populasjonen i bydel  $j$  gitt ved:

$$S_{ij} = S_{ij-1} - I_{ij-1} - R_{ij-1}$$

Likningen sier at for bydel  $i$  og for tid (dag)  $j$  så er antall sårbare de som var sårbare i går minus de som ble smittet og som døde i går. I denne modellen antas det at den totale populasjonen i hver bydel ikke endrer seg. Populasjonsdata for bydeler er aggregert fra FHI sitt populasjonsgrid. Data fra MSIS over positive tester i populasjonen benyttes til å modellere antall daglig nye smittebærende etter formelen:

$$I_{ij-1} = I_{ij-1}^{Sym} + I_{ij-1}^{Asym}$$

$I_{ij-1}^{Sym}$  er antall positivt testede personer.  $I_{ij-1}^{Asym}$  er antall asymptomatiske smitteførende personer gitt ved:

$$I_{ij-1}^{Asym} = \alpha I_{ij-1}^{Sym}$$

Her er  $\alpha$  en konstant som beskriver hvor mange asymptomatisk smittede vi må beregne pr. positivt testet smittede. Vi evaluerte den beste verdien for  $\alpha$  ved å beregne Deviance Information Criteria (DIC). Vi evaluerte verdiene 0,2 0,33 og 0,5  $\alpha$  og utførte analysen kun på bølge 2 og 5. I denne studien er *Removed* delen av SIR-modellen gitt ved en enkel formel for tilfriskning:

$$R_{ij-1} = \beta I_{ij-1}^{Sym}$$

$R_{ij-1}$  har ikke ledd for antall døde av Covid-19 fordi prosjektet ikke har tilgang til slike data på bydelsnivå. Antall døde av Covid-19 i Norge er svært lavt sammenlignet med antall smittede. Det vurderes derfor at en slik variabel vil ha liten påvirkning på modellen.

### 6.5.2 Bayesiansk hierarkisk regresjonsmodell

Dataene som benyttes i SIR-modellen, beskrevet over, har en ukjent grad av både romlig og temporal avhengighet. Det vil si at smitte eller andre faktorer som forventet levealder i en bydel kan påvirkes av omliggende bydeler og at smitten en dag påvirkes av smitte de foregående dager. Slike avhengigheter kan beskrives ved å inkludere dem som latente tilfeldige variabler i en Bayesianske hierarkisk modell (Storvik, 2012). Bayesianske hierarkiske metoder har ofte stor fleksibilitet når en skal bergene data med kompliserte lag av avhengighet, sammenlignet med klassiske frekventistiske metoder (Storvik, 2012). Bayesianske hierarkiske modeller er samtidig svært beregningskrevende. De er med tiden allikevel stadig mer brukt, ettersom beregningsmetoder er blitt mer effektive og beregningshastighet har økt med bedre datakraft (de Valpine et al., 2017; Storvik, 2012).

I Bayesianske modeller behandles parameterne i modellen som tilfeldige variable. Disse parameterne kan vi beregne (*posterior*) oppdaterte sannsynlighetsfordelinger av, ved hjelp av dataene vi putter inn i modellen og en (*prior*) forhåndsantagelse av parameter fordelingene (Storvik, 2012). Modellen som benyttes i denne oppgaven benytter seg av svakt informative prior fordelinger og numeriske MCMC (Markov chain Monte Carlo) beregningsmetoder for å beregne posterior parameter fordelingene. Oppgavens omfang gir ikke rom for nærmere utledning av matematikken bak disse metodene.

Antall mottagelige for smitte er betydelig større enn antall positive testede ved en gitt dag og en Poisson fordeling for antall positivt testede kan dermed antas:

$$I_{ij}^{sym} \sim \text{Poisson}(\mu_{ij})$$

En Poisson fordeling er gitt ved en parameter  $\mu$  som er lik gjennomsnittet og variansen for fordelingen (Samuelsen & Veierød, 2012).

$$\mu_{ij} = E(I_{ij}^{sym}) = Var(I_{ij}^{sym})$$

Det Bayesianiske rammeverket hvor vi inkluderer latente tilfeldige variable gjør at vi kan anta betinget uavhengige data og ta høyde for økt variasjon ved økt antall smittede (Lawson & Kim, 2021). I denne modellen beregnes  $\mu_{ij}$  ved antall mulig smittbare og en funksjon som beskriver hvor stor andel av mulig smittbare som tester positiv for Covid-19.

$$\mu_{ij} = S_{ij} f(I_{ij-1}^{sym} \dots)$$

I denne oppgaven definerer vi en modell etter følgende funksjon:

$$\log f(I_{ij-1}^{sym} \dots) = b_{0j} + b_1 \log(y_{ij-1}^{sym} + y_{ij-1}^{asym}) + b_2 [x_i^{Ga/NDVI}] + b_3 [x_i^{Socio}] + b_4 [x_i^{Socio}] + B_i$$

Her er  $b_{0j}$  en temporal trend for tidsavhengig smittenivå over alle bydeler gjennom en bølge.  $b_1$  er en parameter som fanger opp sammenhengen mellom smittede ved dag  $j$  og dag  $j-1$  basert både på symptomatiske og asymptomatiske smittede tilfeller over bydeler gjennom en bølge.  $b_2$ ,  $b_3$  og  $b_4$  er parametere for bydel-populasjoners tilgang til grønne omgivelser og sosiodemografiske sammensetninger.  $B_i$  er en romlig avhengig effekt mellom bydelene.

### 6.5.3 Spesifikasjon av priors (a priori fordelinger)

Bayesianske statistiske metoder er avhengige av at vi gjør antagelser av sannsynlighetsfordelingene for de frie parameterne (prior fordelinger) før vi gjør beregninger med data (Storvik, 2012). Valg av prior fordelinger påvirker også egenskapene til de numeriske beregningsmetode som modellen beregnes med. Dette ved at fullstendig eller svært uinformerte prior fordelinger kan føre til ustabile modeller som kan bli svært beregningskrevende. For å sikre at modellen konverger uten at den preges av en forforståelse av hvordan Covid-19 smitte påvirkes av de frie parameterne i modellen benyttes svakt informative prior fordelinger. Modellen benytter samme prior distribusjoner som Lawson og Kim (2021). For regresjonsparametrene  $b_1$ ,  $b_2$ ,  $b_3$  og  $b_4$  samt  $b_{0j}$  antar vi at disse følger en normalfordeling:

$$b_n \sim N(0, \tau_n)$$

Her er presisjonsparameter  $\tau_n = 1/\sigma_n^2$  definert ved:

$$\tau_n \sim \text{Gamma}(2, 0,5)$$

For den romlig avhengige effekten  $B_i$  antas en ICAR fordeling (Besag et al., 1991) hvor:

$$B_{i|j \neq i} \sim N(M_i, \tau_B^{-1}/n_{\delta_i})$$



$$M_i = \sum_{l \in \delta_i} B_l / n_{\delta_i}$$

I distribusjonen har bydel  $B_i$  avhengighet til de bydelene den deler grense med ( $\delta_i$ ), antall bydeler dette innebærer er gitt ved  $n_{\delta_i}$ . Presisjonsparameteren  $\tau_B$  er også her gitt ved:

$$\tau_n \sim \text{Gamma}(2, 0,5)$$

I Appendix seksjon 2 og 3 er koden for bearbeiding av data lagt ved sammen med koden for å sette opp og beregne den definerte modellen. Dataprogrammet R (versjon 4.2.1) er brukt til å analysere data og R-pakken NIMBLE (de Valpine et al., 2022) til å sette opp og beregne den statistiske modellen.

#### 6.5.4 Modell beregning

De numeriske metodene for å beregne posterior distribusjoner vil føre til svake resultater om modellen er satt opp på måte som fører til at MCMC prøvetakeren ikke konvergerer. For å evaluere dette er benyttet R-pakken ggmcmc (Fernandes i Marin, 2021). Det ble gjennomført visuell vurdering av MCMC kjedene, vurdering av autokorrelasjon mellom iterasjonene og Geweke diagnostikk (Geweke, 1991), se appendix 1 seksjon 4 for MCMC diagnostikk på et utvalg av kjøringar.

#### 6.5.5 Model seleksjon

For å sammenligne modeller med forskjellige verdier av asymptomatisk andel ( $\alpha$ ) og forskjellige forklaringvariabler ble Deviance Information Criterion (DIC) benyttet (Spiegelhalter et al., 2014).

### 6.6 Pandemiens bølger

Forskjellige bølger av Covid-19 smitte er definert etter en tidslinje av pandemien i Norge beskrevet av den Norske Koronakommisjonen (NOU 2022:5, 2022). Koronakommisjonen beskriver 5 bølger av pandemien basert på nivå av Covid-19 smitte, antall sykehusinnleggelser pga. Covid-19 og alvorlighet og Covid-19 retningslinjer. I vår studie er start og slutt datoer for bølgene valgt på bakgrunn av endringer i Covid-19 smitte og nasjonale Covid-19 retningslinjer. I starten av pandemien hadde Norge redusert Covid-19 testkapasitet noe som førte til restriktive testkriterier (NOU 2021:6, 2021). Derfor ble bare bølge 2-5 inkludert i studien, se figur 2 i artikkelen. I bølge 5 ble registrering av testresultater akkumulert gjennom julen samt at Covid-19 retningslinjene ble betydelig endret etter nyttår. Vi valgte derfor kun å inkludere andre halvdel av bølge 5 i studien.

## 6.7 Etikk

Studien har et økologisk studiedesign der all data er anonym og åpent tilgjengelig. Det er derfor ikke behov for å søke Sikt Kunnskapssektorens tjenesteleverandør om godkjenning.

## 7 Resultater

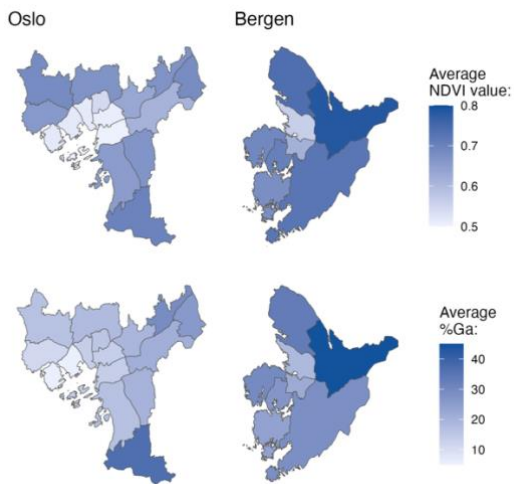
En oppsummering av resultatene i studien vil presenteres nedenfor. Først gjennom en kort gjennomgang av beskrivende statistikk for de to byene. Jeg vil så trekke fram resultater fra analysen utført med Bayes' SIR modell.

### 7.1 Beskrivende statistikk

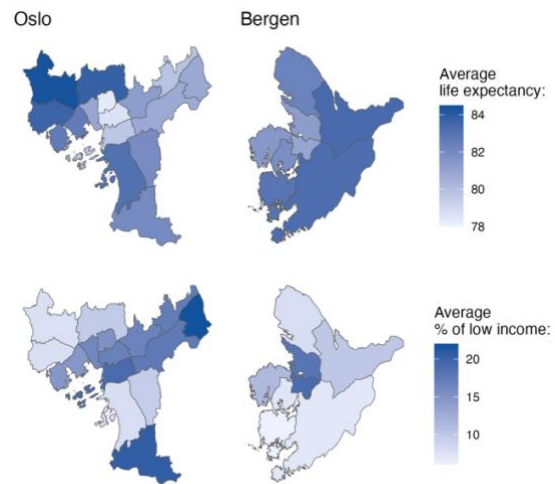
I **Figur 4** vises kartutsnitt med bydelenes gjennomsnittlige verdi for NDVI og %Ga i Oslo og Bergen. I de sentrumsnære områdene i Oslo og Bergen ser vi de laveste verdiene av tilgjengelighet til grønne omgivelser. Befolkningen i Bergen har høyere tilgang til grønne omgivelser sammenlignet med befolkningen i Oslo (gjennomsnittlig NDVI = 0.61, gjennomsnittlig %Ga = 18,3), se tabell 1 artikkelen.

Kartutsnitt med bydelenes gjennomsnittlige forventede levealder og prosentandel av befolkning med lavinntekt er vist i **Figur 5**. Vi ser her at befolkningen i Bergen har en høyere forventet levealder (gjennomsnittlig levealder = 82.2 år) enn befolkningen i Oslo (gjennomsnittlig levealder = 81.4 år). Bergen hadde også en lavere andel av befolkningen med lavinntekt (gjennomsnitt = 10.83%) enn Oslo (gjennomsnitt = 14,5%),.

Vi så også på om det var romlige avhengighet ved å berogene Morans I (Moran, 1950). Her fant vi romlig avhengighet for forventet levealder (p-verdi = 0,01) og %Ga (p-verdi = 0,02) i Oslo, se tabell 1 i artikkelen.



**Figur 4.** Kartfigur av gjennomsnittlige verdier av NDVI og %Ga i bydelen i Oslo og Bergen. Data: Copernicus.



**Figur 5.** Kartfigur av gjennomsnittlig forventet levealder og prosentandel med lavinntekt i bydelene i Oslo og Bergen. Data: FHI.

## 7.2 Asymtomatisk smitte og evaluering av $\alpha$ verdien

Vi benyttet DIC for å evaluere hvilken verdi av  $\alpha$  som skulle benyttes i modellen. Her kom det fram at en  $\alpha$  verdi på 0,2 ga best resultat for bølge 2 og en  $\alpha$  verdi på 0,5 ga best resultat for bølge 5, se tabell 2 i artikkelen. Vi valgte å benytte en  $\alpha$  verdi på 0,2 for alle modellsimuleringene.

## 7.3 Forskjell mellom beregning av tilgang til grønne områder ved NDVI og %Ga og Covid-19 smitte.

To parallelle modeller ble beregnet for de fire bølgene i Oslo og Bergen. I den ene modellen ble tilgang til grønne omgivelser beregnet med %Ga i den andre modellen tilsvarende for NDVI. Alle de andre faktorer (variablene) var identiske for de to modellene. I figur 5 i artikkelen er resultatet av disse beregningene vist. Vi fant ingen betydelig forskjell i resultatene om tilgang til grønne omgivelser ble beregnet med %Ga eller NDVI.

## 7.4 Forskjell mellom byer og bølger

I Bergen var det bare smittenivået dagen før som viste signifikant sammenheng mellom tilgang til grønne områder og Covid-19 smitte. Vi så heller ikke betydelig endring i resultatene mellom de 4 bølgene, se figur 5 og tabell 3 i artikkelen.

I Oslo var også smittenivået dagen før signifikant forbundet til smittenivået dagen etter. Her så vi også signifikant sammenheng mellom andelen av befolkningen som hadde lavinntekt og Covid-19 smitte for bølge 2 ( $b_4(\%Ga) = 0,064$  95% CI: 0,029-0,099,  $b_4(NDVI) = 0,075$  95% CI: 0,041-0,107) og bølge 3 ( $b_4(\%Ga) = 0,065$  95% CI: 0,023-0,106,  $b_4(NDVI) = 0,079$  95% CI:

0,038-0,118). Vi fant også at tilgang til grønne omgivelser hadde en signifikant positiv sammenheng med utvikling av Covid-19 smitte i bølge 4 ( $b_2 (\%Ga) = 0,022$  95% CI: 0,003-0,040  $b_2 (NDVI) = 0,025$  95% CI: 0,005-0,45), se figur 5 og tabell 3 i artikkelen.

## 8 Diskusjon

I dette kapittelet vil jeg kort presentere resultatene fra studien i lys av forskningsspørsmålene som er stilt. Jeg vil kort peke på svakheter i studien som kan ha påvirket disse resultatene. Til slutt vil jeg drøfte hvordan erfaringer fra studien kan bidra til videre kunnskapsutvikling.

### 8.1 Hovedresultater

Vi fant bare en signifikant sammenheng mellom tilgang til grønne omgivelser og utvikling av Covid-19 smitte under den 4. bølgen i Oslo. Her var høyere tilgang til grønne omgivelser sett i sammenheng med en høyere Covid-19 smitte. I Oslo økte også utviklingen av Covid-19 smitte signifikant med andelen av befolkningen med lavinntekt gjennom de første to bølgene. I Bergen fant vi ikke sammenheng mellom tilgang til grønne omgivelser og utvikling av Covid-19 smitte i noen av bølgene.

I studien ble to vanlige beregninger, NDVI og %Ga, benyttet for å utforske sammenhengen mellom tilgang til grønne omgivelser og utvikling av Covid-19 smitte. Det ble ikke funnet noen forskjell i sammenhengen mellom tilgang til grønne omgivelser og Covid-19 smitte om vi benyttet beregninger for NDVI eller %Ga.

### 8.2 Styrker og svakheter

Ved å benytte en Bayes' SIR modell til å utføre den statistisk analysen tok vi utbytte av de daglig oppdaterte tallene for Covid-19 smitte samtidig som vi kan ta hensyn til romlige og temporale avhengigheter i dataene.

Studien har flere metodiske svakheter. Vi hadde kun tilgang på Covid-19 testresultater og sosiodemografiske data på bydelsnivå. Vi må anta at tilgang til grønne områder, Covid-19 smitte og sosiodemografiske karakteristikk varierer innenfor bydelene. Dette kan føre til økologiske feilslutninger. Vi har også et lavt antall bydeler i hver modell, noe som kan føre til upresise resultater.

Dataene på positive Covid-19 tester trenger ikke gjenspeile befolkningens smittetrykk perfekt. Enkelte grupper i befolkningen kan være mindre testet enn andre.

Vi benyttet ikke metoder som tar hensyn til variasjon på kvaliteten til de grønne områdene i befolkningens nærområde. Det lykkes ikke å finne en standardisert måte til å bestemme når en bølge starter og slutter. Vi har heller ikke inkludert forskjeller i Covid-19 retningslinjer i de forskjellige bølgene. Resultatene fra denne studien er ikke overførbare til andre byer.

### 8.3 Konklusjon

Gjennom historien har pandemier som svartedøden i middelalderen og spanskesyken på starten av 1900-tallet ført til store tap av menneskeliv (Webb et al., 2019). Covid-19 er derfor ikke enestående, og mye tyder på at globale endringer som økt urbanisering og utvikling av klimakrisen kan øke risikoen for at verden igjen blir utsatt for en ny pandemi (Baker et al., 2022). Erfaringer som er gjort gjennom Covid-19 pandemien om hva slags effekt smittebegrensende tiltak har hatt og hvordan tiltakene har påvirket folkehelsen kan derfor bli sentrale for myndighetene ved et senere tidspunkt.

Tilgang til grønne omgivelser har positive effekter for befolkningens fysiske og mentale helse. Vi har også sett at bruken av grønne omgivelser til rekreasjon økte betydelig under starten av pandemien. Samtidig peker forskning på at tilgangen til grønne omgivelser ikke er likt fordelt i befolkningen mange steder. Resultatene fra denne studien finner ikke at tilgang til grønne omgivelser har påvirket utviklingen av smitte i stor grad i Oslo og Bergen. Forskning på hvordan grønne omgivelser har vært tilgjengelig og brukt av forskjellige grupper av befolkningen gjennom pandemien kan gi viktig informasjon for hvordan vi kan tilrettelegge til at de positive egenskapene grønne omgivelser har for folkehelsen ikke preges av sosial ulikhet.

## 9 Litteratur:

- Baker, R. E., Mahmud, A. S., Miller, I. F., Rajeev, M., Rasambainarivo, F., Rice, B. L., Takahashi, S., Tatem, A. J., Wagner, C. E., Wang, L. F., et al. (2022). Infectious disease in an era of global change. *Nat Rev Microbiol*, 20 (4): 193-205. doi: 10.1038/s41579-021-00639-z.
- Barboza, E. P., Cirach, M., Khomenko, S., lungman, T., Mueller, N., Barrera-Gómez, J., Rojas-Rueda, D., Kondo, M. & Nieuwenhuijsen, M. (2021). Green space and mortality in European cities: a health impact assessment study. *The Lancet Planetary Health*, 5 (10): e718-e730. doi: 10.1016/s2542-5196(21)00229-1.
- Besag, J., York, J. & Mollié, A. (1991). Bayesian image restoration, with two applications in spatial statistics. *Annals of the Institute of Statistical Mathematics*, 43 (1): 1-20. doi: 10.1007/BF00116466.
- Bonati, M., Campi, R. & Segre, G. (2022). Psychological impact of the quarantine during the COVID-19 pandemic on the general European adult population: a systematic review of the evidence. *Epidemiol Psychiatr Sci*, 31: e27. doi: 10.1017/S2045796022000051.
- Bulfone, T. C., Malekinejad, M., Rutherford, G. W. & Razani, N. (2021). Outdoor Transmission of SARS-CoV-2 and Other Respiratory Viruses: A Systematic Review. *J Infect Dis*, 223 (4): 550-561. doi: 10.1093/infdis/jiaa742.
- Butenschøn, P. (2013). *Oslo. Steder i Byen*. Oslo: Forlaget Press.
- Chen, K., Klompmaker, J. O., Roscoe, C. J., Nguyen, L. H., Drew, D. A., James, P., Laden, F., Fecht, D., Wang, W., Gulliver, J., et al. (2023). Associations between greenness and predicted COVID-19-like illness incidence in the United States and the United Kingdom. *Environ Epidemiol*, 7 (1): e244. doi: 10.1097/EE9.0000000000000244.
- Copernicus. (2018). Urban Atlas. I: *European Commission*. Tilgjengelig fra: <https://land.copernicus.eu/local/urban-atlas/urban-atlas-2018?tab=download> (lest 05.07.2022).
- Copernicus. (2022). *Normalized Difference Vegetation Index*. Copernicus. Tilgjengelig fra: <https://land.copernicus.eu/global/products/ndvi> (lest 25.09.2022).
- de Valpine, P., Turek, D., Paciorek, C. J., Anderson-Bergman, C., Lang, D. T. & Bodik, R. (2017). Programming With Models: Writing Statistical Algorithms for General Model Structures With NIMBLE. *Journal of Computational and Graphical Statistics*, 26 (2): 403-413. doi: 10.1080/10618600.2016.1172487.
- de Valpine, P., Paciorek, C., Turek, D., Michaud, N., Anderson-Bergman, C., Obermeyer, F., Wehrhahn Cortes, C., Rodriguez, A., Temple Lang, D., Paganin, S., et al. (2022). *NIMBLE: MCMC, Particle Filtering, and Programmable Hierarchical Modeling* (Versjon 0.13.1). R-package. Tilgjengelig fra: <https://cran.r-project.org/package=nimble>.
- Fernandes i Marin, X. (2021). *Tools for Analyzing MCMC Simulations from Bayesian Inference* (Versjon 1.5.1.1). Tilgjengelig fra: [https://cran.r-project.org/web/packages/ggmcmc/vignettes/using\\_ggmcmc.html](https://cran.r-project.org/web/packages/ggmcmc/vignettes/using_ggmcmc.html).
- FHI. (2022). *Kommunehelse statistikkbank*: Folkehelseinstituttet. Tilgjengelig fra: <https://khs.fhi.no/webview/> (lest 25.09.2022).
- Fong, K. C., Hart, J. E. & James, P. (2018). A Review of Epidemiologic Studies on Greenness and Health: Updated Literature Through 2017. *Curr Environ Health Rep*, 5 (1): 77-87. doi: 10.1007/s40572-018-0179-y.
- Fongar, C., Aamodt, G., Randrup, T. B. & Solfeld, I. (2019). Does Perceived Green Space Quality Matter? Linking Norwegian Adult Perspectives on Perceived Quality to

- Motivation and Frequency of Visits. *Int J Environ Res Public Health*, 16 (13). doi: 10.3390/ijerph16132327.
- Forskrift om Meldingssystem for smittsomme sykdommer (MSIS-forskriften). (2003). I 2003 hefte 9: Helse- og omsorgsdepartementet.
- Gao, S., Zhai, W. & Fu, X. (2023). Green space justice amid COVID-19: Unequal access to public green space across American neighborhoods. *Front Public Health*, 11: 1055720. doi: 10.3389/fpubh.2023.1055720.
- Geweke, J. (1991). *Evaluating the accuracy of sampling-based approaches to the calculation of posterior moments*.
- Hartig, T., Mitchell, R., de Vries, S. & Frumkin, H. (2014). Nature and health. *Annu Rev Public Health*, 35: 207-28. doi: 10.1146/annurev-publhealth-032013-182443.
- Helse- og omsorgsdepartementet. (2020). *Koronavirus blir meldepliktig sykdom og defineres som en allmennfarlig smittsom sykdom*. Regjeringen.no: Helse- og omsorgsdepartementet. Tilgjengelig fra: <https://www.regjeringen.no/no/dokumentarkiv/regjeringen-solberg/aktuelt-regjeringen-solberg/hod/nyheter/2020ny/koronavirus-blir-meldepliktig-sykdom/id2688415/> (lest 24.09.2022).
- Kermack, W. O. & McKendrick, A. G. (1991). Contributions to the mathematical theory of epidemics--I. 1927. *Bull Math Biol*, 53 (1-2): 33-55. doi: 10.1007/BF02464423.
- Kjøllesdal, M., Skyrud, K., Gele, A., Arnesen, T., Klovstad, H., Diaz, E. & Indseth, T. (2022). The correlation between socioeconomic factors and COVID-19 among immigrants in Norway: a register-based study. *Scand J Public Health*, 50 (1): 52-60. doi: 10.1177/14034948211015860.
- Klomp maker, J. O., Hart, J. E., Holland, I., Sabath, M. B., Wu, X., Laden, F., Dominici, F. & James, P. (2021). County-level exposures to greenness and associations with COVID-19 incidence and mortality in the United States. *Environ Res*, 199: 111331. doi: 10.1016/j.envres.2021.111331.
- Lawson, A. B. & Kim, J. (2021). Space-time covid-19 Bayesian SIR modeling in South Carolina. *PLoS One*, 16 (3): e0242777. doi: 10.1371/journal.pone.0242777.
- Lu, Y., Chen, L., Liu, X., Yang, Y., Sullivan, W. C., Xu, W., Webster, C. & Jiang, B. (2021). Green spaces mitigate racial disparity of health: A higher ratio of green spaces indicates a lower racial disparity in SARS-CoV-2 infection rates in the USA. *Environ Int*, 152: 106465. doi: 10.1016/j.envint.2021.106465.
- Moran, P. A. (1950). Notes on continuous stochastic phenomena. *Biometrika*, 37 (1-2): 17-23. doi: <https://doi.org/10.2307/2F2332142>.
- NOU 2021:6. (2021). *Myndighetenes håndtering av koronapandemien — Rapport fra Koronakommisjonen*. Regjeringen.
- NOU 2022:5. (2022). *Myndighetenes håndtering av koronapandemien-del 2*. Regjeringen.
- Pan, J., Bardhan, R. & Jin, Y. (2021). Spatial distributive effects of public green space and COVID-19 infection in London. *Urban For Urban Green*, 62: 127182. doi: 10.1016/j.ufug.2021.127182.
- Park, A. H., Zhong, S., Yang, H., Jeong, J. & Lee, C. (2022). Impact of COVID-19 on physical activity: A rapid review. *J Glob Health*, 12: 05003. doi: 10.7189/jogh.12.05003.
- Pouso, S., Borja, A., Fleming, L. E., Gomez-Baggethun, E., White, M. P. & Uyarra, M. C. (2021). Contact with blue-green spaces during the COVID-19 pandemic lockdown beneficial for mental health. *Sci Total Environ*, 756: 143984. doi: 10.1016/j.scitotenv.2020.143984.

- Reid, C. E., Rieves, E. S. & Carlson, K. (2022). Perceptions of green space usage, abundance, and quality of green space were associated with better mental health during the COVID-19 pandemic among residents of Denver. *PLoS One*, 17 (3): e0263779. doi: 10.1371/journal.pone.0263779.
- Samuelsen, S. O. & Veierød, M. B. (2012). Poisson regression. I: Veierød, M. B., Lydersen, S. & Laake, P. (red.) *Medical statistics in clinical and epidemiological research*. Oslo: Gyldendal Akademisk.
- Schmiege, D., Haselhoff, T., Ahmed, S., Anastasiou, O. E. & Moebus, S. (2023). Associations Between Built Environment Factors and SARS-CoV-2 Infections at the Neighbourhood Level in a Metropolitan Area in Germany. *J Urban Health*: 1-11. doi: 10.1007/s11524-022-00708-5.
- Spiegelhalter, D. J., Best, N. G., Carlin, B. P. & van der Linde, A. (2014). The deviance information criterion: 12 years on. *Journal of the Royal Statistical Society. Series B (Statistical Methodology)*, 76 (3): 485-493.
- Spotswood, E. N., Benjamin, M., Stoneburner, L., Wheeler, M. M., Beller, E. E., Balk, D., McPhearson, T., Kuo, M. & McDonald, R. I. (2021). Nature inequity and higher COVID-19 case rates in less-green neighbourhoods in the United States. *Nature Sustainability*, 4 (12): 1092-1098. doi: 10.1038/s41893-021-00781-9.
- SSB. (2019). Befolkning på rutenett. Tilgjengelig fra: <https://kartkatalog.geonorge.no/metadatas/befolkning-paa-rutenett-250-m/0c0ad0ce-55e8-4d73-9c12-0eb0e2454acb> (lest 06.02.2023).
- Storvik, G. (2012). Bayesian Methods. I: Veierød, M. B., Lydersen, S. & Laake, P. (red.) *Medical statistics in clinical and epidemiological research*. Oslo: Gyldendal Akademisk.
- Szczygiel, B. & Hewitt, R. (2000). Nineteenth-Century Medical Landscapes: John H. Rauch, Frederick Law Olmsted, and the Search for Salubrity. *Bulletin of the History of Medicine*, 74: 708-734.
- Thorsnæs, G. (2023). *Oslo*. Store norske leksikon. snl.no. Tilgjengelig fra: <http://snl.no/Oslo> (lest 05.02.2023).
- Thorsnæs, G. & Thune, N. A. (2023). *Bergen*. Store norske leksikon. snl.no. Tilgjengelig fra: <http://snl.no/Bergen> (lest 05.02.2023).
- UN-Habitat. (2021). *Cities And Pandemics: Towards a More Just, Green and Healthy Future*. United Nations Human Settlements Programme: UN HABITAT.
- Ursin, G., Skjesol, I. & Tritter, J. (2020). The COVID-19 pandemic in Norway: The dominance of social implications in framing the policy response. *Health Policy Technol*, 9 (4): 663-672. doi: 10.1016/j.hlpt.2020.08.004.
- van den Berg, M., Wendel-Vos, W., van Poppel, M., Kemper, H., van Mechelen, W. & Maas, J. (2015). Health benefits of green spaces in the living environment: A systematic review of epidemiological studies. *Urban Forestry & Urban Greening*, 14 (4): 806-816. doi: 10.1016/j.ufug.2015.07.008.
- Venter, Z. S., Barton, D. N., Gundersen, V., Figari, H. & Nowell, M. (2020). Urban nature in a time of crisis: recreational use of green space increases during the COVID-19 outbreak in Oslo, Norway. *Environmental Research Letters*, 15 (10). doi: 10.1088/1748-9326/abb396.
- Venter, Z. S., Barton, D. N., Gundersen, V., Figari, H. & Nowell, M. S. (2021). Back to nature: Norwegians sustain increased recreational use of urban green space months after the



- COVID-19 outbreak. *Landscape and Urban Planning*, 214. doi: 10.1016/j.landurbplan.2021.104175.
- Venter, Z. S., Sadilek, A., Stanton, C., Barton, D. N., Aunan, K., Chowdhury, S., Schneider, A. & Iacus, S. M. (2021). Mobility in Blue-Green Spaces Does Not Predict COVID-19 Transmission: A Global Analysis. *Int J Environ Res Public Health*, 18 (23). doi: 10.3390/ijerph182312567.
- Webb, P., Bain, C. & Page, A. (2019). *Essential Epidemiology: An Introduction for Students and Health Professionals*. 4 utg. Cambridge: Cambridge University Press.
- WHO. (2016). *Urban green spaces and health* Europe, W. R. O. f. Copenhagen.
- WHO. (2020). *WHO Director-General's opening remarks at the media briefing on COVID-19 - 11 March 2020*: WHO. Tilgjengelig fra: <https://www.who.int/director-general/speeches/detail/who-director-general-s-opening-remarks-at-the-media-briefing-on-covid-19---11-march-2020> (lest 26.03.2023).
- You, Y. & Pan, S. (2020). Urban Vegetation Slows Down the Spread of Coronavirus Disease (COVID-19) in the United States. *Geophysical Research Letters*, 47 (18). doi: 10.1029/2020gl089286.

**Association between urban green space in Norway and transmission of Covid-19- A Bayesian SIR modeling approach**

Halvor Kjellesvig, Department of Public Health Sciences, LANDSAM, NMBU, Ås, Norway

# 1 Abstract

**Background:** Access to green spaces can provide opportunities for physical activities and social interactions in urban areas during times with strict social distancing guidelines because Covid-19 transmission is reduced in ventilated areas. The aim of this study is to explore the association between access to urban green spaces and Covid-19 transmission at the district level in Norway's two largest cities during several waves of the pandemic.

**Methods:** We used daily published numbers of confirmed laboratory PCR tests on district levels during four waves of the Covid-19 pandemic, from October 15, 2020 to April 15, 2022 in Oslo and Bergen. We included the population's access to urban green spaces using two objective measurements: percentage of green area (%Ga) and vegetation cover using the Normalized Difference Vegetation Index (NDVI). The percentages of the population with low income and average life expectancy were included as sociodemographic variables. A Bayesian Susceptible-Infected-Removed (SIR) model was used to take advantage of the daily updated data on Covid-19 incidence and account for spatial and temporal dependencies in the statistical analysis.

**Results:** No significant differences in the association between access to urban green spaces and Covid-19 transmission were found when we measured by %Ga or NDVI. Furthermore, access to urban green spaces was only significantly associated with Covid-19 transmission during the fourth wave of the pandemic in Oslo, where it contributed to a higher incidence rate of Covid-19. In Oslo, Covid-19 transmission increased significantly with the proportion of low-income population during waves second and third. In Bergen, no association between access to urban green spaces and Covid-19 transmission was found in any wave.

**Conclusions:** We found that access to green space showed minor association to Covid-19 transmission in comparison with factors such as socio-economy and the underlying transmission.

**Keywords:** *Natural environment; Green space; COVID-19; Bayesian; NIMBLE*

## 2 Introduction

On March 11, 2020, the World Health Organization (WHO) declared that the spread of the SARS-Co-V2 virus, which causes Covid-19, had led to a global pandemic (WHO, 2020a). The following day, the Norwegian government introduced the most interventional national measures in peacetime aimed to curtail the spread of the virus to reduce the number of Covid-19 deaths and the burden on health services (NOU 2021:6). Until large parts of the vulnerable population had obtained immunization from vaccination, such measures, as in other countries, included different degree of social distancing and periods with closed borders, mandatory homeschool, and home office (NOU 2022:5). The measures can lead to social isolation and reduced physical activity and eventually impact the mental and physical health of the population (Bonati et al., 2022; Park et al., 2022).

Access to green spaces can provide opportunities for physical activities and social interactions in urban areas during times with strict social distancing guidelines because Covid-19 transmission is reduced in ventilated areas (Bulfone et al., 2021). Access to green spaces in urban areas has many health benefits. These include increased physical activity, improved social cohesion, reduced stress, and less exposure to air pollution (Hartig et al., 2014). Evidence also suggests that access to green spaces has positively affected mental health during the pandemic (Pouso et al., 2021; Reid et al., 2022).

In addition to the positive benefits access to green spaces in urban areas has had for public health during the pandemic, several ecological studies on the county level suggest that access to green spaces can contribute to decreased Covid-19 transmission (Klompmaker et al., 2021; Venter et al., 2021; You & Pan, 2020). At lower administrative levels, two studies found opposite results. One study at borough level in London during the first wave, found that higher access to connected green spaces was associated with an increased risk of Covid-19 transmission (Pan et al., 2021). Another study from Essen, Germany at neighborhood level during the first year of the pandemic, found a negative association between access to urban green spaces and Covid-19 transmission (Schmiege et al., 2023). The results from the county-level studies are supported by a study with data from 2.8 million participants in the Covid Symptom Study (Chen et al., 2023). Some studies also suggest that access to green spaces in urban areas is not equally distributed among socioeconomic groups and that this can lead to differences between groups on how levels of Covid-19 transmission develop and the group's ability to adapt to contagion-reducing measures (Gao et al., 2023; Lu et al., 2021; Spotswood et al., 2021)

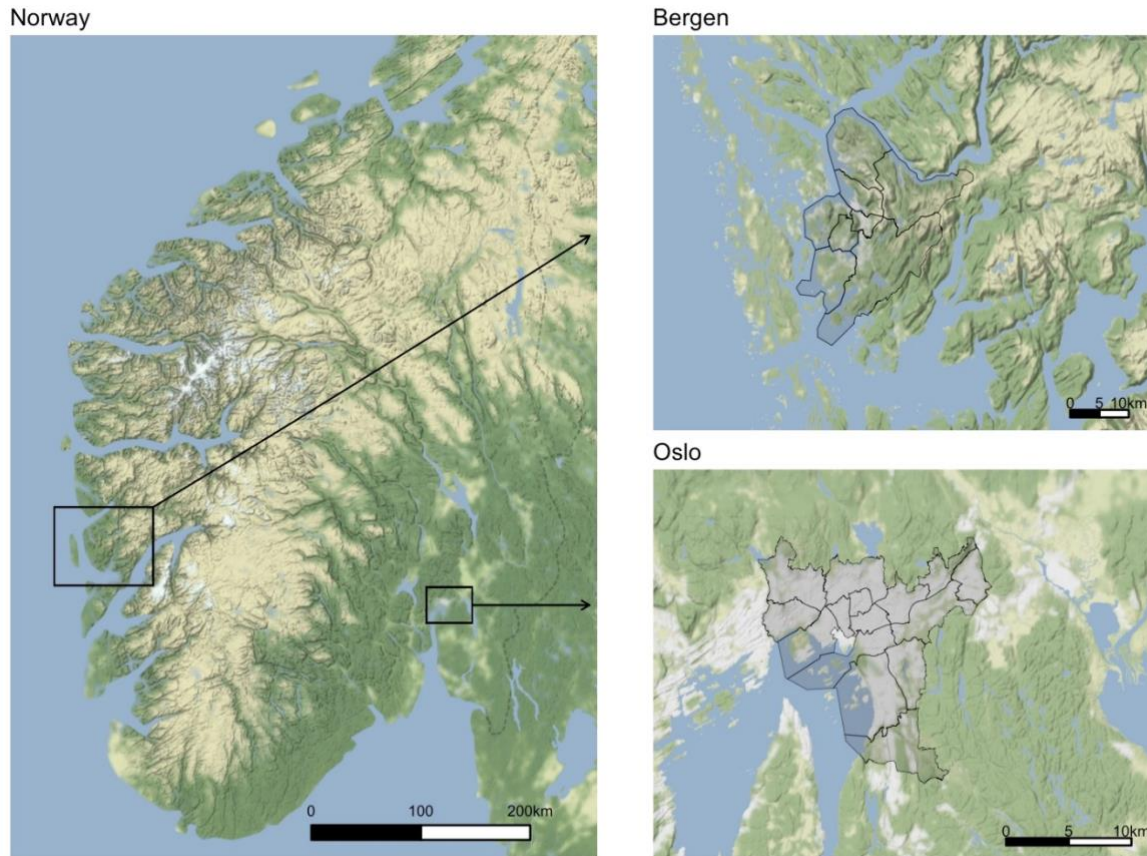
The Covid-19 pandemic has led to severe strain on urban populations during lockdowns. This has demonstrated the importance of environmental qualities in cities in an increasingly urbanized world (UN-Habitat, 2021). Global changes such as urbanization and climate change have also heightened the risk of new outbreaks of infectious diseases (Baker et al., 2022). In this context, awareness of how urban qualities such as access to green spaces impacts resilience may be valuable for future urban planning and city development. The aim of this study is to explore the association between access to urban green spaces and Covid-19 transmission at the district level in Norway's two largest cities during several waves of the pandemic. Access to urban green spaces in nearby residential areas is measured with two commonly used proxies; Normalized Difference Vegetation Index (NDVI) and percentage of green area (%GA). NDVI detects live vegetation, and %Ga represents the green area using land cover and land use maps. A Bayesian Susceptible-Infected-Removed (SIR) model is used to take advantage of the daily updated data on Covid-19 incidence and account for spatial and temporal dependencies in the statistical analysis. This study uses data from four different waves, controlled for selected socioeconomic variables, to explore the following research questions:

- *By using a Bayesian SIR-model adjusting for socio-economic characteristics is there a association between access to green spaces and incidence of Covid-19?*
- *Is there a difference in the association between access to urban green spaces and Covid-19 transmission during different waves of the pandemic?*
- *Is there a difference in the association between access to urban green spaces and Covid-19 transmission if we measure green spaces by NDVI or %Ga?*

## 3 Methods

### 3.1 Study area and study population

Oslo is the capital of Norway, with 697 010 inhabitants, and Bergen is the second largest city, with 285 602 inhabitants in 2021 (FHI, 2022b). The cities, shown in Figure 1, comprise 15 districts in Oslo and 8 in Bergen. The districts have some degree of local political governance, but Covid-19 guidelines are made at city and national level (Thorsnæs, 2023; Thorsnæs & Thune, 2023). This study retrieves data on the number of inhabitants on a grid cell with a resolution of 250m x 250 in 2019 (SSB, 2019). Statistics Norway generates the population grid from the national population register and the national property register (SSB, 2019).



**Figure 1.** Bergen is the center of the Norwegian west coast. Oslo is the center of the southeast part of Norway. Map tiles by [Stamen Design](#), under CC BY 3.0. Data by [OpenStreetMap](#), under ODbL.

### 3.2 Covid-19 incidence

On January 31, 2020, Covid-19 was declared a dangerous infectious disease by national authorities in Norway. This required incidents of Covid-19 to be reported in accordance with the Norwegian infection control act (Helse- og omsorgsdepartementet, 2020). The Norwegian Institute of Public Health (NIPH) is responsible for the national surveillance system for communicable diseases (MSIS ) (*MSIS-regulation*, 2003). NIPH has published daily numbers of confirmed laboratory PCR tests on the district level in Oslo and Bergen from March 26, 2020 until December 7, 2022. This data was retrieved from an open-source initiative that publishes public Norwegian Covid-19 data (FHI, 2022a).

### 3.3 Access to urban green space

We included the population’s access to urban green spaces using two objective measurements: percentage of green area (%GA) and vegetation cover using the normalized difference vegetation index (NDVI). A residential proximity area is defined as a 300-meter buffer around each grid cell in the population dataset for both measures. The 300-meter buffer, shown in Figure 2, represents a 5 min walk on walkable pathways and is chosen on

recommendation from WHO (2016). In the proximity area, the %GA and average NDVI value are calculated by an approach described by Barboza et al. (2021).

%GA is measured based on land cover and land use maps from Urban Atlas (Copernicus, 2018). Urban Atlas delivers harmonized land cover and land use data for over 700 urban areas in Europe with a spatial resolution of 0.25 ha (50 m). Urban Atlas is produced every third year; the latest data set is from 2018. Although Urban Atlas has a separate category for urban green areas, this study uses a broader classification, see Barboza et al. (2021), with the following Urban Atlas categories: Green urban areas, sport and leisure facilities, arable land, permanent crops, pastures, complex and mixed cultivation patterns, forests, herbaceous vegetation associations and wetlands. See Appendix section 1.2 for a calculation flowchart.

NDVI values are captured by the European earth observation satellite Sentinel 3. The satellite measures the spectral reflectance of near-infrared (NIR) and red (Red) wavebands to obtain the indicator of the greenness of biomes by the simple formulation (Copernicus, 2022):

$$NDVI = \frac{NIR - Red}{NIR + Red}.$$

The European earth observation program Copernicus processes and distributes the NDVI values as rasters with a spatial resolution of 300m x 300m three times monthly. In this study, NDVI rasters from April to August 2021 were downloaded and processed to make a maximum NDVI value raster for that year, representing the maximum photosynthetic level in the canopy (Brown, 2018). NDVI pixels with values representing water are filtered out. Finally, the max NDVI raster is combined with the population buffer layer to calculate the average NDVI value in each residential proximity area. We used QGIS (version 3.28.2) for geospatial operations. See Appendix section 1.1 for the calculation flowchart.

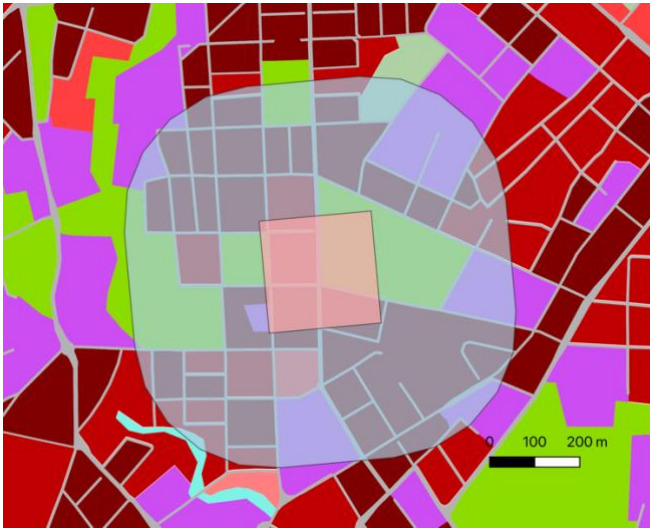


Figure 2 Example of one population grid cell with a 300-meter buffer in Oslo. Map Urban Atlas 2018.

### 3.4 Sociodemographic data

Socio-economic and sociodemographic data at the district level in 2020 were obtained from the Norwegian Institute of Public Health (FHI, 2022b). Initially, data on the percentage of immigrants and Norwegian born with immigrant parents in the population, the percentage of the population with low income by EU-income scale, the percentage of the population with high school education or higher and life expectancy were included in the model. However, due to the relatively strong correlation coefficients between socioeconomic variables and increasing model complexity with the number of included variables, only data on the percentage of the population with low income and life expectancy are included in the final analyses.

### 3.5 Statistical analyses

A Bayesian Susceptible-Infected-Removed (SIR) model, demonstrated by Lawson and Kim (2022), was used to investigate the association between Covid-19 incidence and access to urban green spaces. The model combines the SIR model technique, for outbreak modeling of infectious diseases, with a Bayesian hierarchical framework that considers spatial and temporal dependencies in the data.

This study considers a separate Covid-19 wave in Oslo or Bergen as an outbreak. The spatial and temporal units in the model are defined as:  $i = 1, \dots, M$  districts and  $j = 1, \dots, T$  days. The susceptible portion of the population in a district  $i$  at time  $j$  is given by:

$$S_{ij} = S_{ij-1} - I_{ij-1} - R_{ij-1}$$

The model assumes that the total population in each district is constant since the population size is relatively stable over the timespan of this study. Data from MSIS, on positive laboratory tests, is used to model the daily infected population given by:

$$I_{ij-1} = I_{ij-1}^{Sym} + I_{ij-1}^{Asym}$$

Here  $I_{ij-1}^{Sym}$  is the number of positively tested persons at day  $j - 1$ , and the asymptomatic part of the population is given by:

$$I_{ij-1}^{Asym} = \alpha I_{ij-1}^{Sym}$$

The positive constant  $\alpha$  describes how many asymptomatic infected persons we have per confirmed positive test. The “optimal” value of  $\alpha$  was evaluated using Deviance Information Criteria (DIC). The following values were included in this evaluation (0.2, 0.33 and 0.5) and we performed the analysis for waves two and five only, see section *Asymptomatic infection and evaluation of the  $\alpha$  value*.



In the SIR-model, the Removed part ( $R_{ij-1}$ ) only considers a recovery rate as a function of the positive tested:

$$R_{ij-1} = \beta I_{ij-1}^{sym}$$

and the  $\beta$  value is set to 0.1 following Lawson and Kim (2021). In our analyses, we did not include the number of deaths caused by Covid-19 since this number was negligibly low compared to the number of cases.

The SIR function estimates the susceptible population provided by our data on positive tests the day before. The susceptible population is considerably larger than the number of positive tests, and a Poisson model can be assumed to make Bayesian statistical inferences:

$$I_{ij}^{sym} \sim \text{Poisson}(\mu_{ij})$$

$$\mu_{ij} = S_{ij} f(I_{ij-1}^{sym} \dots)$$

The Bayesian hierarchical framework with latent random variables makes it possible to assume conditional independence and account for possible overdispersion (Lawson & Kim, 2022). The fraction of the susceptible population that tests positive for Covid-19 is in this model given by the function:

$$\log f(I_{ij-1}^{sym} \dots) = b_{0j} + b_1 \log(y_{ij-1}^{sym} + y_{ij-1}^{asym}) + b_2 [x_i^{\%Ga/NDVI}] + b_3 [x_i^{Life\ eceptensy}] + b_4 [x_i^{\% Low\ income}] + B_i$$

In this equation:  $b_{0j}$  is a temporally dependent trend, the regression parameter  $b_1$  is associated with the level of infection (tested and asymptomatic) the day before,  $b_2$ ,  $b_3$  og  $b_4$  is covariates for access to urban green spaces and sociodemographic characteristics for the district and  $B_i$  is a spatial structured effect.

For the spatial structured effect  $B_i$ , the model assumes a prior ICAR distribution by (Besag et al., 1991) with a precision parameter given by a prior Gamma (2, 0.5) distribution. For the rest of the model parameters ( $b_{0j}$ ,  $b_1$ ,  $b_2$ ,  $b_3$ , and  $b_4$ ), a zero mean normal prior distribution is assumed with a precision parameter also given by a prior Gamma (2, 0.5) distribution. The precision parameter is considered weakly informative, which means that the prior distribution should contain enough information to rule out most of the unreasonable parameter values but be able to include those that could make sense. In addition, we assumed the temporal trend  $b_{0j}$  to be uncorrelated following Lawson and Kim (2022).

### 3.5.1 Model computation

The computational program R (version 4.2.1) is used for data processing and statistical analyses. The R-package NIMBLE by de Valpine et al. (2022) implements the Bayesian hierarchical model and runs MCMC (Markov chain Monte Carlo) samplers to obtain the

posterior distribution of the unknown parameters. See Appendix section 3 for the model simulation code. MCMC convergence was evaluated with the R-package ggMCMC by Fernandes i Marin (2021). Considering visual assessment of MCMC-chains, autocorrelations between iterations, Geweke diagnostics (Geweke, 1991) and limitations due to computational time, each outbreak simulation was run with two MCMC chains with 2 million iterations, a burn-in of 500 000 and a thinning interval of 300. See appendix section 4 for a selection of plots from the MCMC diagnostics.

### 3.5.2 Model selection

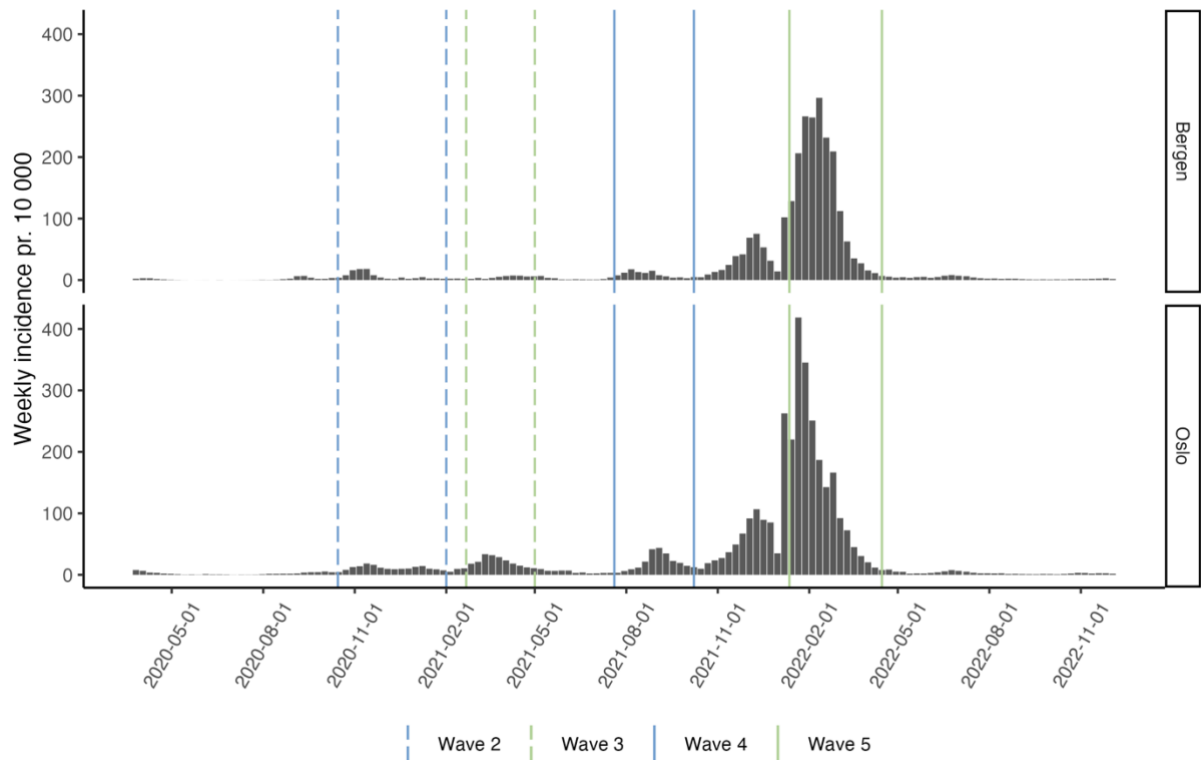
In our study we used the Deviance Information Criterion (DIC) following Lawson and Kim (2022) to compare models with different selections of fixed variables and asymptomatic rates ( $\alpha$ ). The DIC value is based on the principle of measuring the “goodness of fit” penalized by model complexity (Spiegelhalter et al., 2014). Lower DIC values suggest a better model fit to the data.

## 3.6 Waves

Different waves of Covid-19 transmission are defined based on a timeline of the pandemic in Norway described by the Norwegian Covid-19 Commission (NOU 2022:5). The Commission describes five Covid-19 waves based on levels of Covid-19 transmission, levels of Covid-19 induced hospitalizations and severity of Covid-19 guidelines. In the present study we selected start and end dates based on changes in national Covid-19 transmission and national Covid-19 guidelines. During the initial stages of the pandemic, Norway faced a limited laboratory test capacity for Covid-19, resulting in strict test criteria (NOU 2021:6, 2021). Therefore, only waves 2-5 were selected for analyses (see Figure 3). Furthermore, only the second part of wave five is included due to changes in Covid-19 guidelines and an accumulation of reported test results during Christmas, which fell in the middle of the wave.

## 3.7 Ethics

The study used an ecological design where all data are anonymous and approval from the Norwegian Agency for Shared Services in Education and Research is not required.



**Figure 3.** Weekly incidence of positive laboratory PCR-tests of Covid-19 and intervals of selected Covid-19 waves. Wave 2: October 15, 2020 – February 01, 2021. Wave 3: February 21, 2021 – May 01, 2021. Wave 4: July 20, 2021 – October 08, 2021. Wave 5: January 12, 2022 – April 15, 2022. Data: Norwegian Institute of Public Health.

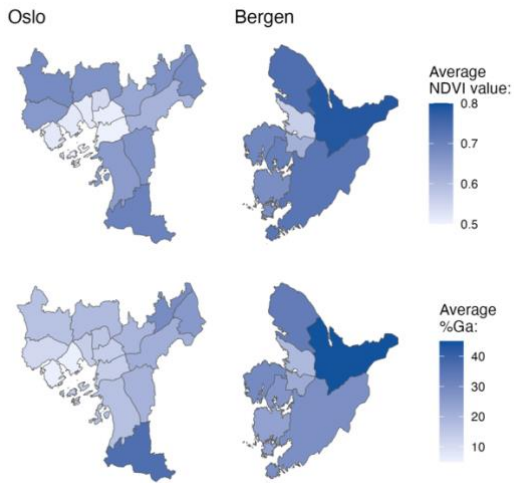
## 4 Results

### 4.1 Descriptive statistics

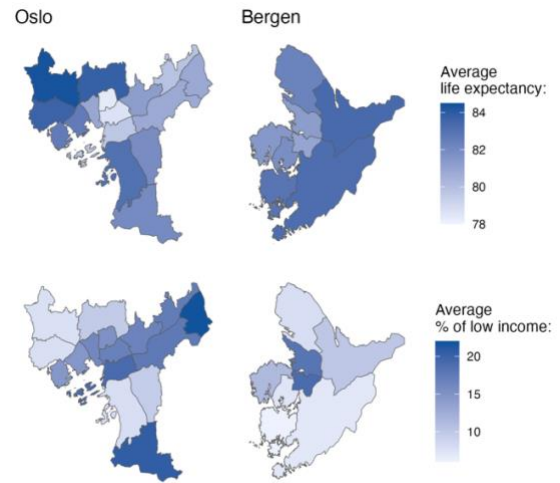
Figure 4 shows the distribution of vegetation cover (NDVI) and percentage of green area (%GA) for both Oslo and Bergen. The central districts in both Bergen and Oslo, which correspond to the city centers, have the lowest values of access to urban green spaces measured in NDVI and %Ga. Moran's I tests, which examine spatial autocorrelation (Moran, 1950), suggest a non-significant, but an interesting spatial clustering for NDVI ( $p$ -value = 0.08) and a significant ( $p$ -value = 0.02) spatial clustering of %Ga in Oslo (see Table 1). No such indications of spatial autocorrelation were found for NDVI and %Ga in Bergen. The population in Bergen has higher access to urban green spaces (mean NDVI = 0.69, mean %Ga = 28.3) compared to Oslo (mean NDVI = 0.61, mean %Ga = 18.3), see Table 1.

We also show the distribution of expected life years and low income among the districts for both cities in Figure 5. The population in Bergen has a higher life expectancy (mean age = 82.2 yrs) than in Oslo (mean age = 81.4 yrs). We also observed that Bergen had a lower percentage of individuals with low-income (mean = 10.83%) in their districts compared to Oslo (mean = 14.5%). Moran's I test suggested spatial clustering for life expectancy in Oslo, only

( $p$ -value = 0.01), with a higher life expectancy in the western districts than eastern districts. No significant values were found for low income. But low-income districts are seen from the city center and north-east along the Groruddalen. See Table 1 for additional information.



**Figure 4.** Plot of average NDVI values and %Ga in residential areas on the district level in Oslo and Bergen. Data: Copernicus.



**Figure 5.** Plot of average life expectancy and % of low income on the district level in Oslo and Bergen. Data: Norwegian Institute of Public Health.

*Table 1 Descriptive statistics of the study variables and population on district level. Data: (FHI, 2022b)*

Variable	Mean	Std.Dev	Min	Max	Moran's I	p-value
<b>Bergen</b>						
Life expectancy	82.16	0.92	80.95	83.35	0.12	0.16
Low income	10.83	5.06	6.10	19.00	-0.01	0.23
NDVI	0.69	0.07	0.56	0.78	-0.25	0.66
%Ga	28.30	8.00	18.47	44.42	-0.18	0.53
Population	35 072	10 208	13 800	42 389		
<b>Oslo</b>						
Life expectancy	81.42	1.86	78.30	84.35	0.36	0.01
Low income	14.53	4.83	7.60	21.80	0.11	0.16
NDVI	0.61	0.07	0.50	0.70	0.19	0.08
%Ga	18.30	7.68	5.75	36.82	0.30	0.02
Population	44 789	10 335	27 336	60 663		

## 4.2 Asymptomatic infection and evaluation of the $\alpha$ value

Model fit using the DIC value was performed to evaluate a set of values for the proportion of asymptomatic cases compared to symptomatic cases ( $\alpha$ ) as described above. We evaluated the values 0.2, 0.33 and 0.5 of  $\alpha$  for waves two and five. Table 2 shows the results, which indicate that an  $\alpha$  value of 0.2 gives the best model fit for wave two and an  $\alpha$  value of

0.5 for wave five. Nevertheless, an  $\alpha$  value of 0.2 was chosen for all model simulations to ensure consistent analyses between waves.

Table 2. Model fit by DIC value for different values of  $\alpha$  on waves two and five in Oslo and Bergen. The lowest DIC value is marked in gray.

	Oslo		Bergen	
	%Ga DIC	NDVI DIC	%Ga DIC	NDVI DIC
<b>Wave 2</b>				
$\alpha = 0.2$	9 473.57	9 473.29	3 361.20	3 361.37
$\alpha = 0.33$	9 474.52	9 474.69	3 361.34	3 361.34
$\alpha = 0.5$	9 475.64	9 476.01	3 361.37	3 361.38
<b>Wave 5</b>				
$\alpha = 0.2$	15 573.06	15 573.26	5 520.56	5 520.76
$\alpha = 0.33$	15 563.92	15 564.11	5 518.99	5 518.80
$\alpha = 0.5$	15 551.03	15 551.03	5 516.17	5 516.50

### 4.3 Difference between measurement of NDVI and %Ga and Covid-19 transmission

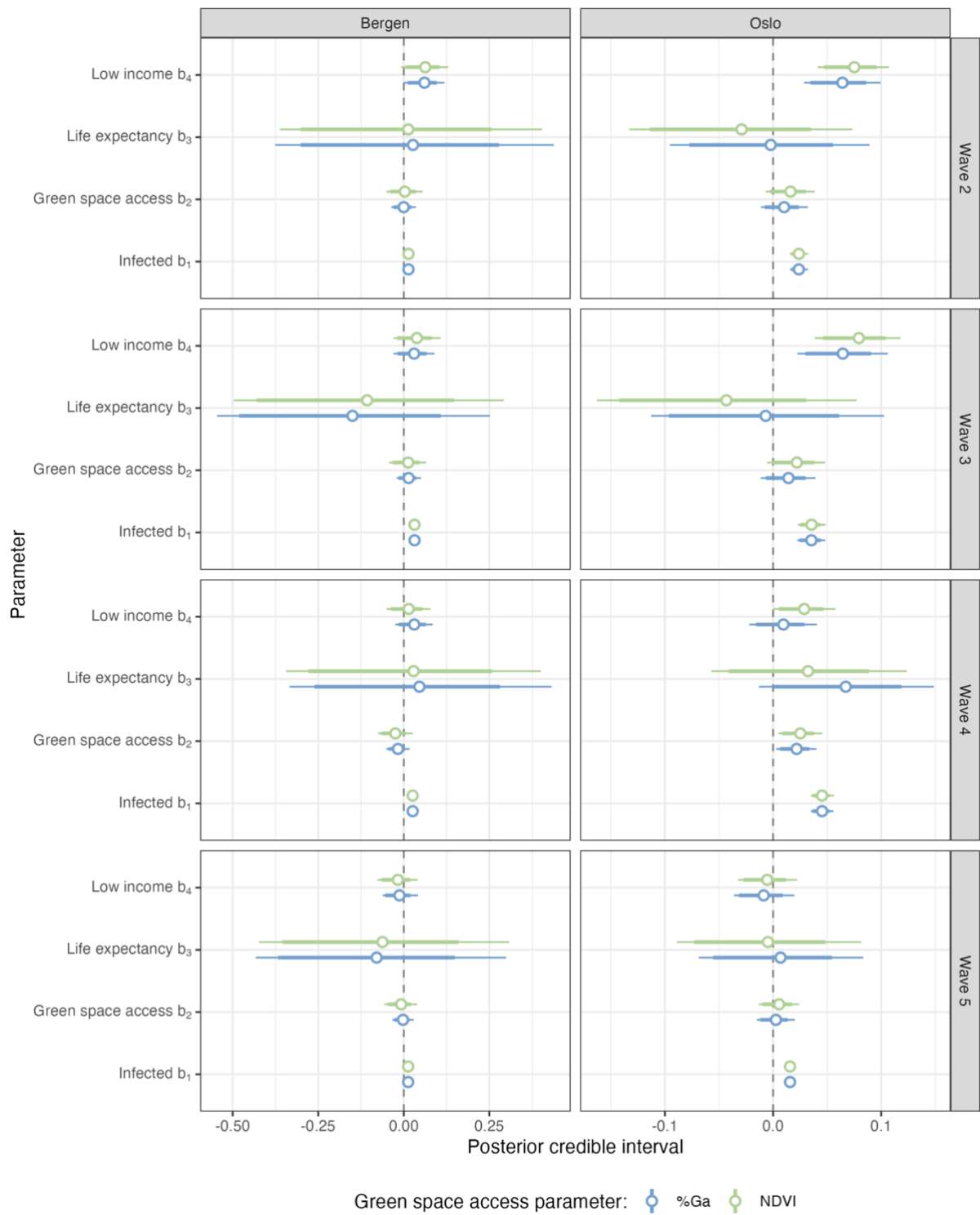
For the four waves two parallel models were run, the first model was run with the %Ga and the second model with vegetation cover (NDVI) as the variable representing access to urban green space. Everything else was kept similar between these two models. This procedure was followed for both Oslo and Bergen. In Figure 6, the outcome of this process is visualized. Each city and wave are represented by a plot, with circles that denote the posterior mean for the parameters  $b_1$ ,  $b_2$ ,  $b_3$  and  $b_4$ . The thick lines indicate the 90% posterior credibility interval, while the thin lines indicate the 95% credibility interval. The NDVI model is represented by green, and the %Ga model is represented by blue. For all waves in Oslo and Bergen, no statistically significant differences in the association between NDVI or %Ga and Covid-19 were found.

### 4.4 Differences between cities and waves

In Bergen, only the infection level the day before ( $b_1$ ) showed a significant association with Covid-19 incidence rates for all the waves. With the highest association in wave three ( $b_1(\%Ga) = 0.031$ , 95% CI: 0.017, 0.046) and lowest in wave five ( $b_1(\%Ga) = 0.013$ , 95% CI: 0.006, 0.019), see Table 3. No other variables had a significant effect. A non-significant but interesting association between Covid-19 transmission and percentage of low-income population when measuring access to urban green space with %Ga ( $b_4(\%Ga) = 0.060$ , 95% CI: 0.000, 0.119) was found in wave two. None of the two green access variables were significantly

related to Covid-19 transmission and we did not observe considerable differences between the four waves.

In Oslo, the effect of the infection level the day before ( $b_1$ ) significantly affects the incidence level during all the waves, see Figure 6. In addition, the percentage of the low-income population significantly affects the incidence level during waves two ( $b_4 (\%Ga) = 0.064$ , 95% CI: 0.029, 0.099,  $b_4 (NDVI) = 0.075$ , 95% CI: 0.041, 0.107) and wave three ( $b_4 (\%Ga) = 0.065$ , 95% CI: 0.023, 0.106,  $b_4 (NDVI) = 0.079$ , 95% CI: 0.038, 0.118). We observed that green access was associated with incidence rates of Covid-19 for the fourth wave only ( $b_2 (\%Ga) = 0.022$ , 95% CI: 0.003, 0.040,  $b_2 (NDVI) = 0.025$ , 95% CI: 0.005, 0.045). Here incidence rates of Covid-19 increased with increasing access to green space giving evidence for differences in the association between different waves.



**Figure 6.** Posterior mean parameter estimates with 95% and 90% credible interval for components  $b_1$ ,  $b_2$ ,  $b_3$ , and  $b_4$  with %Ga and NDVI parameter for  $b_2$ . Posterior mean is represented with a circle, the 95% credible interval with a thin line and the 90% credible interval with a thick line.

Table 3. Posterior mean parameter estimates with 95% and 90% credible interval for components  $b_1$ ,  $b_2$ ,  $b_3$  and  $b_4$  with %Ga and NDVI parameter for  $b_2$ .

Parameter	Coefficient	Upper95	Low95	Upper90	Low90	Coefficient	Upper95	Low95	Upper90	Low90
	Oslo					Bergen				
<b>Wave 2 NDVI</b>										
Infected $b_1$	0.024	0.032	0.016	0.030	0.017	0.014	0.026	0.001	0.022	0.003
Green access $b_2$	0.016	0.039	-0.007	0.030	-0.003	0.003	0.055	-0.051	0.035	-0.040
Life expectancy $b_3$	-0.029	0.073	-0.133	0.035	-0.114	0.013	0.404	-0.362	0.256	-0.301
Low income $b_4$	0.075	0.107	0.041	0.096	0.047	0.063	0.129	-0.008	0.105	0.006
<b>Wave 2 %Ga</b>										
Infected $b_1$	0.024	0.032	0.016	0.029	0.017	0.014	0.026	0.001	0.022	0.003
Green $b_2$	0.010	0.032	-0.011	0.024	-0.008	0.000	0.035	-0.037	0.022	-0.030
Life expectancy $b_3$	-0.002	0.089	-0.095	0.056	-0.077	0.027	0.438	-0.376	0.278	-0.301
Low income $b_4$	0.064	0.099	0.029	0.086	0.035	0.060	0.119	0.000	0.096	0.012
<b>Wave 3 NDVI</b>										
Infected $b_1$	0.036	0.049	0.023	0.044	0.025	0.031	0.046	0.017	0.041	0.019
Green access $b_2$	0.022	0.048	-0.006	0.038	0.000	0.013	0.065	-0.042	0.046	-0.032
Life expectancy $b_3$	-0.043	0.077	-0.163	0.031	-0.142	-0.107	0.292	-0.496	0.146	-0.430
Low income $b_4$	0.079	0.118	0.038	0.104	0.046	0.038	0.108	-0.030	0.082	-0.018
<b>Wave 3 %Ga</b>										
Infected $b_1$	0.035	0.048	0.022	0.044	0.024	0.031	0.046	0.017	0.041	0.019
Green access $b_2$	0.014	0.039	-0.011	0.030	-0.007	0.014	0.050	-0.021	0.037	-0.015
Life expectancy $b_3$	-0.007	0.103	-0.113	0.061	-0.096	-0.150	0.251	-0.545	0.108	-0.481
Low income $b_4$	0.065	0.106	0.023	0.091	0.030	0.030	0.089	-0.030	0.067	-0.018
<b>Wave 4 NDVI</b>										
Infected $b_1$	0.045	0.056	0.035	0.052	0.037	0.026	0.038	0.014	0.034	0.016
Green access $b_2$	0.025	0.045	0.005	0.038	0.009	-0.025	0.026	-0.075	0.006	-0.065
Life expectancy $b_3$	0.032	0.124	-0.057	0.089	-0.041	0.029	0.400	-0.344	0.258	-0.279
Low income $b_4$	0.029	0.058	0.000	0.047	0.005	0.014	0.078	-0.050	0.054	-0.038
<b>Wave 4 %Ga</b>										
Infected $b_1$	0.045	0.056	0.035	0.052	0.037	0.026	0.038	0.014	0.034	0.016
Green access $b_2$	0.022	0.040	0.003	0.034	0.006	-0.017	0.017	-0.051	0.004	-0.044
Life expectancy $b_3$	0.067	0.149	-0.013	0.119	0.001	0.046	0.432	-0.334	0.281	-0.261
Low income $b_4$	0.010	0.040	-0.022	0.029	-0.016	0.031	0.084	-0.025	0.065	-0.015
<b>Wave 5 NDVI</b>										
Infected $b_1$	0.016	0.019	0.012	0.018	0.013	0.013	0.019	0.006	0.017	0.007
Green access $b_2$	0.006	0.024	-0.013	0.018	-0.010	-0.008	0.039	-0.057	0.021	-0.046
Life expectancy $b_3$	-0.005	0.082	-0.089	0.049	-0.073	-0.062	0.308	-0.423	0.160	-0.355
Low income $b_4$	-0.005	0.022	-0.032	0.012	-0.027	-0.017	0.041	-0.077	0.019	-0.065
<b>Wave 5 %Ga</b>										
Infected $b_1$	0.016	0.019	0.012	0.018	0.013	0.013	0.019	0.006	0.017	0.007
Green access $b_2$	0.003	0.020	-0.015	0.014	-0.012	-0.002	0.029	-0.033	0.017	-0.027
Life expectancy $b_3$	0.007	0.083	-0.069	0.055	-0.055	-0.079	0.299	-0.432	0.149	-0.368
Low income $b_4$	-0.009	0.020	-0.036	0.009	-0.031	-0.013	0.041	-0.061	0.019	-0.054



## 5 Discussion

This study has used two commonly used measurements, NDVI and %Ga, to explore the relationship between access to urban green spaces and Covid-19 in Norway's two biggest cities during four waves of the pandemic. No significant differences in the association between access to urban green spaces and Covid-19 transmission is found if we measure by %Ga or NDVI. Furthermore, access to urban green spaces was only significantly associated with Covid-19 transmission during the fourth wave of the pandemic in Oslo, where it contributed to a higher incidence rate of Covid-19. In Oslo, Covid-19 transmission increased significantly with the proportion of low-income population during waves two and three. In Bergen, no association between access to urban green spaces and Covid-19 transmission was found in any wave.

The similarities between the %Ga and NDVI results indicate that the broad selections of land cover and land use categories included in the %Ga measurement, following Barboza et al. (2021), corresponds to the maximum "greenness" assessed from the NDVI value. Such a broad definition is commonly used to define green space (Nguyen et al., 2021), contributing to this study's generalizability.

A Norwegian study found that self-reported distance from the participants' homes to green space is the most important quality in predicting the use of green spaces (Fongar et al., 2019). This finding is supported by a study on recreational mobility patterns in Oslo during the pandemic outbreak that found a substantial increase in outdoor activity, especially on trails with higher NDVI values and tree canopy cover (Venter et al., 2020). People also traveled shorter distances from their homes when Covid-19 guidelines become stricter (Kamineni et al., 2023). The level of "greenness" in residential proximities may therefore be a predictor for outdoor recreational mobility.

Several causal pathways are suggested between access to urban green spaces and Covid-19 transmission (Spotswood et al., 2021). First, access to green space may boost the immune system by strengthening health and reducing stress. Urban green spaces have less air pollution, which is found to be associated with lower Covid-19 incidence rates (Meo et al., 2021). In addition, access to urban green spaces can provide opportunities for social interactions outside rather than inside, possibly leading to decreased Covid-19 transmission rates. How access to urban green spaces affects social interactions and social distancing may depend on many factors that can develop during the pandemic, such as social distancing rules and weather conditions. We found that access to urban green spaces only had a significant positive association with Covid-19 transmission in wave four in Oslo. During this wave, large parts of the most

vulnerable population were vaccinated, and the government eased up social distancing measures inside while encouraging that social contact preferably should be outside (NOU 2022:5, 2022). The wave developed from the middle of the summer to beginning of autumn, an attractive time for outdoor activity in Oslo. The observational data in this study cannot be used for causal inferences. Still, the differences in the association between access to urban green spaces and Covid-19 during different waves in Oslo suggest further investigations on how access to green space may affect Covid-19 transmission during various waves of the pandemic.

### 5.1 The Bayesian Susceptible-Infected-Removed model

In addition to social distancing and mobility-reducing measurements, WHO emphasized the importance of testing, isolation and contact tracing to reduce Covid-19 incidence from the early start of the pandemic (WHO, 2020b). The Bayesian Susceptible-Infected-Removed model used in this study takes exploits the large daily data set from the massive testing during the pandemic. Furthermore, the spatial structured effect included in the model makes it possible to account for spatial dependencies such as the spatial clustering of high life expectancy in the west side of Oslo.

The SIR model's ability to include asymptomatic infected that found the best fit with an  $\alpha$  value of 0.2 for wave 2 is supported by a systematic review of data from the first part of the pandemic that showed asymptomatic transmission of 24.51% (95% confidence interval 14.38, 36.02) (Ravindra et al., 2022). In our analyses, we found that wave five had an  $\alpha$  value of 0.5, indicating the best fit. This suggests that the model accounts for the impact of new Covid-19 variants that have a higher rate of transmission among the population. In Norway, most Covid-19 infections during the second part of wave five were from the Omicron variant (Whittaker et al., 2022). Studies found that the Omicron variant has more asymptomatic infections than earlier Covid-19 variants (Yu et al., 2022). This could explain the high proportion of asymptomatic cases we observed for the last wave.

We also found a significant effect from the socioeconomic variable, indicating that low income was associated with higher Covid-19 transmission rates in waves two and three in Oslo and an interesting but not significant association in wave two in Bergen. This result is supported by a registry based study with data on confirmed PCR tests on the entire Norwegian population between June 15, 2020 and March 31, 2021 (Labberton et al., 2022).

We found that the results deviated from the different waves in this study. This suggests different characteristics regarding the association between access to green spaces and Covid-19 transmission for the different waves. The Bayesian framework of the Susceptible-Infected-

Removed model is highly flexible and could include more spatiotemporal data, such as mobility data. This could give important information on how Covid-19 social distancing and mobility measures affect the use of urban green spaces and how this is associated with Covid-19 transmission. However, it is important to consider that the Bayesian framework is computationally costly, especially with increased model complexity.

## 5.2 Strengths and limitations

To our knowledge, this is the first study investigating incidence rates of Covid-19 in Norway using a SIR model. We excluded the first wave due to less certain data in this first part of the pandemic, because of limited the test-capacity resulting in restrictive test criteria (NOU 2021:6, 2021). Nevertheless, the Covid-19 cases during the rest of the period was reckoned as satisfactory. The SIR model includes different factors governing transmission of Covid-19 which is also a strength. Other models are focused on reproduction number and do not accommodate the underlying process of transmission.

This study has several limitations. First, the Covid-19 test data and sociodemographic data were only publicly accessible at the district level. However, it is probable that the Covid-19 incidence, access to green space and sociodemographic characteristics vary on a smaller spatial scale within each district leading to heterogeneity in the model's spatial unit. The number of districts in Oslo and Bergen is also low. In Bergen, there are only 8 districts, which challenges the spatial dimension of our model. This could lead to low precision of posterior credible intervals.

We used national Covid-19 guidelines and national Covid-19 transmission levels to determine Covid-19 waves. This approach does not capture individual differences on when the waves developed in Oslo and Bergen or when local Covid-19 guidelines were introduced. Furthermore, the lack of a formal wave definition makes comparing these results to other studies difficult. It is worth mentioning that quantitative definitions of a wave are suggested (Ayala et al., 2021; Zhang et al., 2021), but no common use of these definitions was found.

The buffer area of 300 meters used in this study does not account for differences in the accessibility to green spaces within the buffer. In Bergen, the accessibility of green spaces may be reduced because many residential areas lie between the coastline and steep hills. Residential proximity polygons determined by areas reached with a specific walking time along roads and trails may better capture such geographical characteristics. In addition, the quality of green space in urban areas may be affected by crowding (Ekkel & de Vries, 2017). This could be especially important during times with strict Covid-19 guidelines. A larger residential proximity

area could potentially capture an effect of people traveling longer distances to access less crowded green spaces.

This study does not account for variations in testing capacity and outreach to different populations. There are differences in testing rates and rates of confirmed Covid-19 cases among different socioeconomic groups, particularly in Oslo (Vinjerui et al., 2021). In addition, one study finds that access to urban green spaces differs among socioeconomic groups in Oslo (Suárez et al., 2020). Our model does not capture these socioeconomic differences.

Although we have tried to use widely accepted definitions and methods when gathering data for our model to ensure its generalizability, the results of this study cannot be directly applied to other cities. Our model does not consider differences in Covid-19 restricting measures, and we have not been able to find a commonly used definition of a wave.

### 5.3 Conclusion

In this study we have investigated the relationship between access to urban green spaces and Covid-19 transmission. We found that access to green space was minor in comparison with factors such as socio-economy and the underlying transmission. This knowledge can help authorities plan how to use these spaces for public health benefits while reducing the spread of infectious diseases. This is especially important in a world where global changes such as urbanization and climate change have heightened the risk of new outbreaks of infectious diseases. The Bayesian Susceptible-Infected-Removed model provides a flexible framework that utilizes daily spatial Covid-19 incidence rates to capture differences between waves and has the potential to include additional spatiotemporal data, such as mobility data.

## 6 References

- Ayala, A., Villalobos Dintrans, P., Elorrieta, F., Castillo, C., Vargas, C. & Maddaleno, M. (2021). Identification of COVID-19 Waves: Considerations for Research and Policy. *Int J Environ Res Public Health*, 18 (21). doi: 10.3390/ijerph182111058.
- Baker, R. E., Mahmud, A. S., Miller, I. F., Rajeev, M., Rasambainarivo, F., Rice, B. L., Takahashi, S., Tatem, A. J., Wagner, C. E., Wang, L. F., et al. (2022). Infectious disease in an era of global change. *Nat Rev Microbiol*, 20 (4): 193-205. doi: 10.1038/s41579-021-00639-z.
- Barboza, E. P., Cirach, M., Khomenko, S., lungman, T., Mueller, N., Barrera-Gómez, J., Rojas-Rueda, D., Kondo, M. & Nieuwenhuijsen, M. (2021). Green space and mortality in European cities: a health impact assessment study. *The Lancet Planetary Health*, 5 (10): e718-e730. doi: 10.1016/s2542-5196(21)00229-1.
- Besag, J., York, J. & Mollié, A. (1991). Bayesian image restoration, with two applications in spatial statistics. *Annals of the Institute of Statistical Mathematics*, 43 (1): 1-20. doi: 10.1007/BF00116466.
- Bonati, M., Campi, R. & Segre, G. (2022). Psychological impact of the quarantine during the COVID-19 pandemic on the general European adult population: a systematic review of the evidence. *Epidemiol Psychiatr Sci*, 31: e27. doi: 10.1017/S2045796022000051.
- Brown, J. (2018). *Deriving Phenological Metrics from NDVI*. USGS: Remote Sensing Phenology Available at: <https://www.usgs.gov/special-topics/remote-sensing-phenology/science/deriving-phenological-metrics-ndvi> (accessed: 0.04.2023).
- Bulfone, T. C., Malekinejad, M., Rutherford, G. W. & Razani, N. (2021). Outdoor Transmission of SARS-CoV-2 and Other Respiratory Viruses: A Systematic Review. *J Infect Dis*, 223 (4): 550-561. doi: 10.1093/infdis/jiaa742.
- Chen, K., Klompaker, J. O., Roscoe, C. J., Nguyen, L. H., Drew, D. A., James, P., Laden, F., Fecht, D., Wang, W., Gulliver, J., et al. (2023). Associations between greenness and predicted COVID-19-like illness incidence in the United States and the United Kingdom. *Environ Epidemiol*, 7 (1): e244. doi: 10.1097/EE9.0000000000000244.
- Copernicus. (2018). Urban Atlas. In *European Commission*. Available at: <https://land.copernicus.eu/local/urban-atlas/urban-atlas-2018?tab=download> (accessed: 05.07.2022).
- Copernicus. (2022). *Normalized Difference Vegetation Index*. Copernicus. Available at: <https://land.copernicus.eu/global/products/ndvi> (accessed: 25.09.2022).
- de Valpine, P., Paciorek, C., Turek, D., Michaud, N., Anderson-Bergman, C., Obermeyer, F., Wehrhahn Cortes, C., Rodriguez, A., Temple Lang, D., Paganin, S., et al. (2022). *NIMBLE: MCMC, Particle Filtering, and Programmable Hierarchical Modeling* (Version 0.13.1). R-package. Available at: <https://cran.r-project.org/package=nimble>.
- Ekkel, E. D. & de Vries, S. (2017). Nearby green space and human health: Evaluating accessibility metrics. *Landscape and Urban Planning*, 157: 214-220. doi: 10.1016/j.landurbplan.2016.06.008.
- Fernandes i Marin, X. (2021). *Tools for Analyzing MCMC Simulations from Bayesian Inference* (Version 1.5.1.1). Available at: [https://cran.r-project.org/web/packages/ggmcmc/vignettes/using\\_ggmcmc.html](https://cran.r-project.org/web/packages/ggmcmc/vignettes/using_ggmcmc.html).
- FHI. (2022a). Covid-19 Meldingsystem for smittsomme sykdommer. In *FHI*. Available at: <https://www.covid19data.no/data.html> (accessed: 05.01.2023).
- FHI. (2022b). *Kommunehelse statistikkbank*: Folkehelseinstituttet. Available at: <https://khs.fhi.no/webview/> (accessed: 25.09.2022).

- Fongar, C., Aamodt, G., Randrup, T. B. & Solfeld, I. (2019). Does Perceived Green Space Quality Matter? Linking Norwegian Adult Perspectives on Perceived Quality to Motivation and Frequency of Visits. *Int J Environ Res Public Health*, 16 (13). doi: 10.3390/ijerph16132327.
- Forskrift om Meldingssystem for smittsomme sykdommer (MSIS-forskriften). (2003). I 2003 hefte 9: Helse- og omsorgsdepartementet.
- Gao, S., Zhai, W. & Fu, X. (2023). Green space justice amid COVID-19: Unequal access to public green space across American neighborhoods. *Front Public Health*, 11: 1055720. doi: 10.3389/fpubh.2023.1055720.
- Geweke, J. (1991). *Evaluating the accuracy of sampling-based approaches to the calculation of posterior moments*.
- Hartig, T., Mitchell, R., de Vries, S. & Frumkin, H. (2014). Nature and health. *Annu Rev Public Health*, 35: 207-28. doi: 10.1146/annurev-publhealth-032013-182443.
- Helse- og omsorgsdepartementet. (2020). *Koronavirus blir meldepliktig sykdom og defineres som en allmennfarlig smittsom sykdom*. Regjeringen.no: Helse- og omsorgsdepartementet. Available at: <https://www.regjeringen.no/no/dokumentarkiv/regjeringen-solberg/aktuelt-regjeringen-solberg/hod/nyheter/2020ny/koronavirus-blir-meldepliktig-sykdom/id2688415/> (accessed: 24.09.2022).
- Kamineneni, M., Engo-Monsen, K., Midtbo, J. E., Forland, F., de Blasio, B. F., Frigessi, A. & Engebretsen, S. (2023). Effects of non-compulsory and mandatory COVID-19 interventions on travel distance and time away from home, Norway, 2021. *Euro Surveill*, 28 (17). doi: 10.2807/1560-7917.ES.2023.28.17.2200382.
- Klompemaker, J. O., Hart, J. E., Holland, I., Sabath, M. B., Wu, X., Laden, F., Dominici, F. & James, P. (2021). County-level exposures to greenness and associations with COVID-19 incidence and mortality in the United States. *Environ Res*, 199: 111331. doi: 10.1016/j.envres.2021.111331.
- Labberton, A. S., Godoy, A., Elgersma, I. H., Strand, B. H., Telle, K., Arnesen, T., Nygard, K. M. & Indseth, T. (2022). SARS-CoV-2 infections and hospitalisations among immigrants in Norway-significance of occupation, household crowding, education, household income and medical risk: a nationwide register study. *Scand J Public Health*, 50 (6): 772-781. doi: 10.1177/14034948221075029.
- Lawson, A. B. & Kim, J. (2021). Space-time covid-19 Bayesian SIR modeling in South Carolina. *PLoS One*, 16 (3): e0242777. doi: 10.1371/journal.pone.0242777.
- Lawson, A. B. & Kim, J. (2022). Bayesian Space-time SIR modeling of Covid-19 in two US states during the 2020- 2021 pandemic. doi: 10.1101/2022.02.14.22270941.
- Lu, Y., Chen, L., Liu, X., Yang, Y., Sullivan, W. C., Xu, W., Webster, C. & Jiang, B. (2021). Green spaces mitigate racial disparity of health: A higher ratio of green spaces indicates a lower racial disparity in SARS-CoV-2 infection rates in the USA. *Environ Int*, 152: 106465. doi: 10.1016/j.envint.2021.106465.
- Meo, S. A., Almutairi, F. J., Abukhalaf, A. A. & Usmani, A. M. (2021). Effect of Green Space Environment on Air Pollutants PM2.5, PM10, CO, O(3), and Incidence and Mortality of SARS-CoV-2 in Highly Green and Less-Green Countries. *Int J Environ Res Public Health*, 18 (24). doi: 10.3390/ijerph182413151.
- Moran, P. A. (1950). Notes on continuous stochastic phenomena. *Biometrika*, 37 (1-2): 17-23. doi: <https://doi.org/10.2307%2F2332142>.

- Nguyen, P. Y., Astell-Burt, T., Rahimi-Ardabili, H. & Feng, X. (2021). Green Space Quality and Health: A Systematic Review. *Int J Environ Res Public Health*, 18 (21). doi: 10.3390/ijerph182111028.
- NOU 2021:6. (2021). *Myndighetenes håndtering av koronapandemien — Rapport fra Koronakommisjonen*. Regjeringen.
- NOU 2022:5. (2022). *Myndighetenes håndtering av koronapandemien-del 2*. Regjeringen.
- Pan, J., Bardhan, R. & Jin, Y. (2021). Spatial distributive effects of public green space and COVID-19 infection in London. *Urban For Urban Green*, 62: 127182. doi: 10.1016/j.ufug.2021.127182.
- Park, A. H., Zhong, S., Yang, H., Jeong, J. & Lee, C. (2022). Impact of COVID-19 on physical activity: A rapid review. *J Glob Health*, 12: 05003. doi: 10.7189/jogh.12.05003.
- Pouso, S., Borja, A., Fleming, L. E., Gomez-Baggethun, E., White, M. P. & Uyarra, M. C. (2021). Contact with blue-green spaces during the COVID-19 pandemic lockdown beneficial for mental health. *Sci Total Environ*, 756: 143984. doi: 10.1016/j.scitotenv.2020.143984.
- Ravindra, K., Malik, V. S., Padhi, B. K., Goel, S. & Gupta, M. (2022). Asymptomatic infection and transmission of COVID-19 among clusters: systematic review and meta-analysis. *Public Health*, 203: 100-109. doi: 10.1016/j.puhe.2021.12.003.
- Reid, C. E., Rieves, E. S. & Carlson, K. (2022). Perceptions of green space usage, abundance, and quality of green space were associated with better mental health during the COVID-19 pandemic among residents of Denver. *PLoS One*, 17 (3): e0263779. doi: 10.1371/journal.pone.0263779.
- Schmiege, D., Haselhoff, T., Ahmed, S., Anastasiou, O. E. & Moebus, S. (2023). Associations Between Built Environment Factors and SARS-CoV-2 Infections at the Neighbourhood Level in a Metropolitan Area in Germany. *J Urban Health*: 1-11. doi: 10.1007/s11524-022-00708-5.
- Spiegelhalter, D. J., Best, N. G., Carlin, B. P. & van der Linde, A. (2014). The deviance information criterion: 12 years on. *Journal of the Royal Statistical Society. Series B (Statistical Methodology)*, 76 (3): 485-493.
- Spotswood, E. N., Benjamin, M., Stoneburner, L., Wheeler, M. M., Beller, E. E., Balk, D., McPhearson, T., Kuo, M. & McDonald, R. I. (2021). Nature inequity and higher COVID-19 case rates in less-green neighbourhoods in the United States. *Nature Sustainability*, 4 (12): 1092-1098. doi: 10.1038/s41893-021-00781-9.
- SSB. (2019). Befolkning på rutenett. Available at: <https://kartkatalog.geonorge.no/metadata/befolkning-paa-rutenett-250-m/0c0ad0ce-55e8-4d73-9c12-0eb0e2454acb> (accessed: 06.02.2023).
- Suárez, M., Barton, D. N., Cimburova, Z., Rusch, G. M., Gómez-Baggethun, E. & Onaindia, M. (2020). Environmental justice and outdoor recreation opportunities: A spatially explicit assessment in Oslo metropolitan area, Norway. *Environmental Science & Policy*, 108: 133-143. doi: 10.1016/j.envsci.2020.03.014.
- Thorsnæs, G. (2023). *Oslo*. Store norske leksikon. snl.no. Available at: <http://snl.no/Oslo> (accessed: 05.02.2023).
- Thorsnæs, G. & Thune, N. A. (2023). *Bergen*. Store norske leksikon. snl.no. Available at: <http://snl.no/Bergen> (accessed: 05.02.2023).
- UN-Habitat. (2021). *Cities And Pandemics: Towards a More Just, Green and Healthy Future*. United Nations Human Settlements Programme: UN HABITAT.
- Venter, Z. S., Barton, D. N., Gundersen, V., Figari, H. & Nowell, M. (2020). Urban nature in a time of crisis: recreational use of green space increases during the COVID-19 outbreak

- in Oslo, Norway. *Environmental Research Letters*, 15 (10). doi: 10.1088/1748-9326/abb396.
- Venter, Z. S., Sadilek, A., Stanton, C., Barton, D. N., Aunan, K., Chowdhury, S., Schneider, A. & Iacus, S. M. (2021). Mobility in Blue-Green Spaces Does Not Predict COVID-19 Transmission: A Global Analysis. *Int J Environ Res Public Health*, 18 (23). doi: 10.3390/ijerph182312567.
- Vinjerui, K. H., Elgersma, I. H. & Fretheim, A. (2021). Increased COVID-19 Testing Rates Following Combined Door-to-Door and Mobile Testing Facility Campaigns in Oslo, Norway, a Difference-in-Difference Analysis. *Int J Environ Res Public Health*, 18 (21). doi: 10.3390/ijerph182111078.
- Whittaker, R., Greve-Isdahl, M., Boas, H., Suren, P., Buanes, E. A. & Veneti, L. (2022). COVID-19 Hospitalization Among Children <18 Years by Variant Wave in Norway. *Pediatrics*, 150 (3). doi: 10.1542/peds.2022-057564.
- WHO. (2016). *Urban green spaces and health* Europe, W. R. O. f. Copenhagen.
- WHO. (2020a). *WHO Director-General's opening remarks at the media briefing on COVID-19 - 11 March 2020*: WHO. Available at: <https://www.who.int/director-general/speeches/detail/who-director-general-s-opening-remarks-at-the-media-briefing-on-covid-19---11-march-2020> (accessed: 26.03.2023).
- WHO. (2020b). *WHO Director-General's opening remarks at the media briefing on COVID-19 - 16 March 2020*. WHO: WHO. Available at: <https://www.who.int/director-general/speeches/detail/who-director-general-s-opening-remarks-at-the-media-briefing-on-covid-19---16-march-2020> (accessed: 27.04.2023).
- You, Y. & Pan, S. (2020). Urban Vegetation Slows Down the Spread of Coronavirus Disease (COVID-19) in the United States. *Geophysical Research Letters*, 47 (18). doi: 10.1029/2020gl089286.
- Yu, W., Guo, Y., Zhang, S., Kong, Y., Shen, Z. & Zhang, J. (2022). Proportion of asymptomatic infection and nonsevere disease caused by SARS-CoV-2 Omicron variant: A systematic review and analysis. *J Med Virol*, 94 (12): 5790-5801. doi: 10.1002/jmv.28066.
- Zhang, S. X., Arroyo Marioli, F., Gao, R. & Wang, S. (2021). A Second Wave? What Do People Mean by COVID Waves? - A Working Definition of Epidemic Waves. *Risk Manag Healthc Policy*, 14: 3775-3782. doi: 10.2147/RMHP.S326051.



# Appendix

## Association between urban green space in Norway and transmission of Covid-19- A Bayesian SIR modeling approach

Halvor Kjellesvig

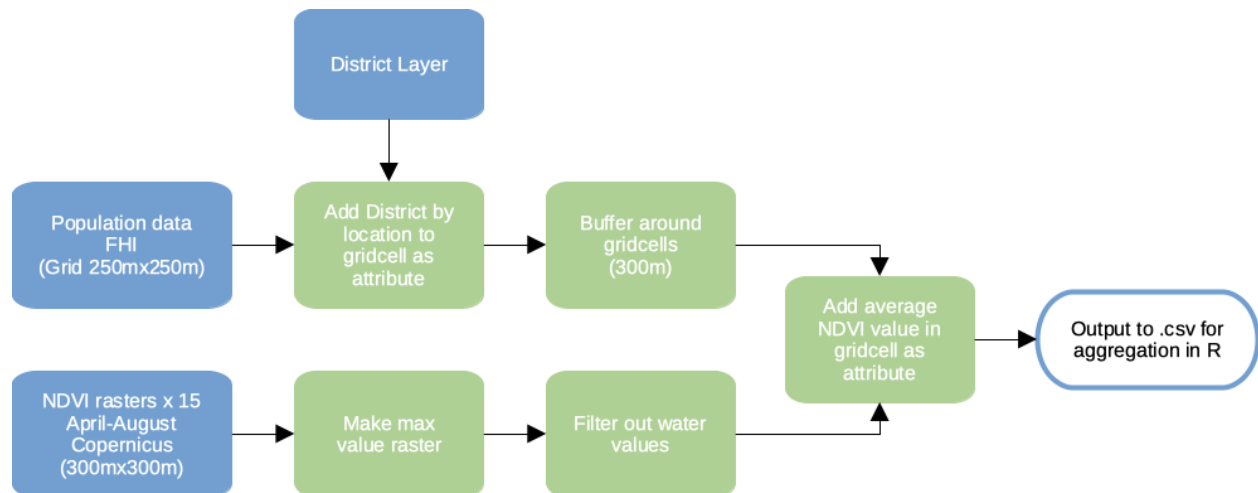
15.05.2023

### Contents

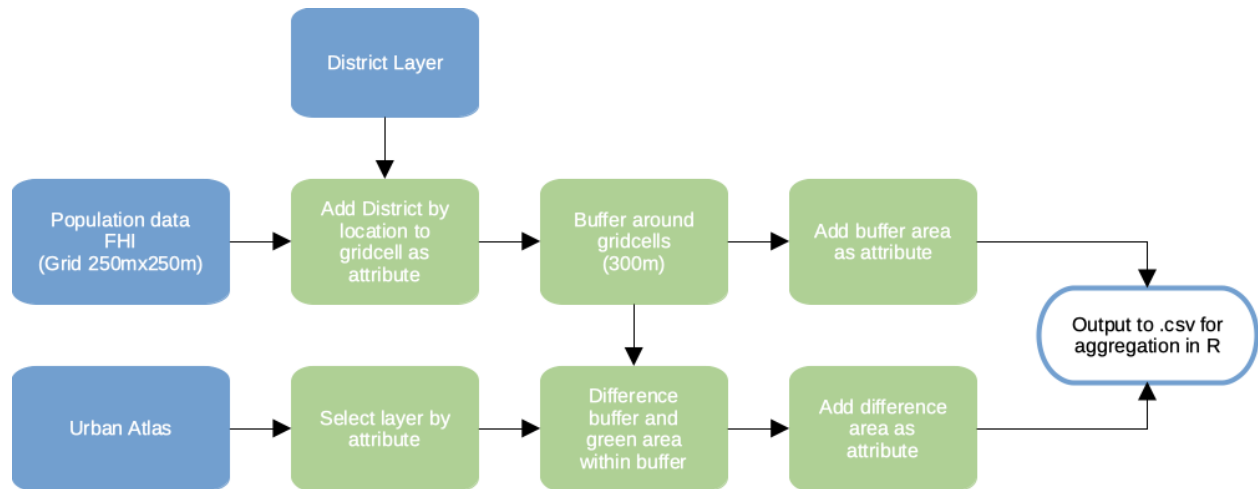
<b>1 Access to urban green spaces</b>	<b>1</b>
1.1 Flowchart for NDVI estimation. . . . .	1
1.2 Flowchart for %Ga estimation. . . . .	2
<b>2 Data generations:</b>	<b>2</b>
<b>3 NIMBLE code:</b>	<b>12</b>
<b>4 MCMC diagnostics</b>	<b>17</b>
<b>References</b>	<b>18</b>

## 1 Access to urban green spaces

### 1.1 Flowchart for NDVI estimation.



## 1.2 Flowchart for %Ga estimation.



*Urban Atlas categories: Green urban areas, sport and leisure facilities, arable land, permanent crops, pastures, complex and mixed cultivation patterns, forests, herbaceous vegetation associations and wetlands.*

## 2 Data generations:

The data for the Bayesian modeling was generated with the following code:

```
## Script for data import and data generation:

### Load packages:

# Data manipulation packages:
library(tidyverse) # Multiple data manipulation and data visualisation packages
library(vctr)     # Package to import data
library(readr)    # Package to import data

# GIS packages:
library(sf)       # Package for importing and manipulate geographical data in R
library(geojsonsf) # Package to import geojson files to sf-objects
library(tmap)     # Package for geographical visualisation
library(tmaptools) # Package for geographical visualisation
library(spdep)    # Package used to calculate adjacency matrix

# Statistical analyses:
library(nimble)   # Package for Bayesian regression
library(coda)     # Package for MCMC output
library(ggcmc)    # Package for MCMC and plotting

# General packages for data presentation:
library(patchwork) # Package to combine plots
library(kableExtra) # Package for table formatting

## Import Covid-data:

# The file is obtained from Geir Aamodt Covid_12.R file that uses data from
# www.covid19data.no
```

```

covid19 <- read_delim("covid19.csv", delim = ";",
                     escape_double = FALSE, trim_ws = TRUE)
# Filter out data that isn't from Oslo or Bergen:
covid19 <- covid19 %>% drop_na(bydel_no) |>
  rename(bydel = kommune_bydel_name)

### Bergen:
##Import data:

# Covid data:
# Filter out data that isn't from Bergen:
covid19_bergen <- covid19 |>
  filter(kommune_name == "Bergen") |>
  filter(kommune_no == 4601) |>
  filter(bydel_name == "Arna" | bydel_name == "Årstad" | bydel_name == "Åsane" |
         bydel_name == "Bergenhush" | bydel_name == "Fana" |
         bydel_name == "Fyllingsdalen" | bydel_name == "Laksevåg" |
         bydel_name == "Ytrebygda")

# Population data
Pop_Bergen <- read.csv("BergenSSBpop.csv")

# Green Area (%Ga) Bergen
Pe_Bergen <- read.csv("GreenareaSSBBergen.csv")

# NDVI Bergen
NDVIBergen <- read.csv("BergenNDVI14.02.23.csv") |>
  drop_na(X_mean) # Remove 1 NA observation in the data.
NDVIBergen$bydel_id <- as.character(NDVIBergen$bydel_id)

# Socio-demographic
sosio_Bergen <- read.csv("sosioall.csv") |>
  filter(bydel_id > 40000) # Chose data from Bergen
sosio_Bergen$bydel_id <- as.character(sosio_Bergen$bydel_id)

# District map
districts <- read_sf(
  "bydeler_bergen_n50_polygoner/bydeler_bergen_n50_polygoner.shp") |>
  drop_na(bydel) |>
  select(bydel_id,bydel,geometry) |>
  st_as_sf()

## Aggregate data:

# Population density:
Pop_Bergen <- Pop_Bergen |>
  group_by(bydel) |>
  summarise(meanpopcell = mean(poptot))

# Percentage of green area in buffer:
Pe_Bergen <- Pe_Bergen |>
  mutate(green_fraction = poptot*((Bufferarea-DiffArea)/Bufferarea))

```

```

Pe_Bergen <- Pe_Bergen |>
  group_by(bydel_id) |>
  summarise(
    green_fraction_bydel = sum(green_fraction), pop_bydel = sum(poptot)) |>
  mutate(green_percent = (green_fraction_bydel*100)/pop_bydel)

Pe_Bergen$bydel_id <- as.character(Pe_Bergen$bydel_id)

# NDVI
NDVIBergen <- NDVIBergen |>
  mutate(popNDVI = poptot*X_mean)

NDVIBergen <- NDVIBergen |>
  group_by(bydel_id) |>
  summarise(NDVI_sum = sum(popNDVI), pop_bydel = sum(poptot)) |>
  mutate(NDVI_bydel =NDVI_sum/(pop_bydel))

## Make a dataframe that contains covariates data for the NIMBLE model:
NDVIBergen <- NDVIBergen |>
  select(-NDVI_sum, -pop_bydel)

sosio_Bergen <- sosio_Bergen |>
  select(-Bydel,-bydel_no)

districts <- left_join(Pop_Bergen,districts) |>
  st_as_sf()

Bergen_data <- left_join(Pe_Bergen,sosio_Bergen) |>
  left_join(NDVIBergen) |>
  left_join(districts) |>
  st_as_sf()

## Adjacency Matrix:

# Make a Adjacency List of district in Bergen.
AdjList <- poly2nb(Bergen_data)

# The NIMBLE package uses a different format on the matrix:
AdjNIMBLE <- nb2WB(AdjList)

## Center and scaling data for the MCMC samplers:
green_per <- Bergen_data$green_percent - mean(Bergen_data$green_percent) # Scaling
NDVI_scale <- (Bergen_data$NDVI_bydel - mean(Bergen_data$NDVI_bydel))*100 # Scaling
lowincome_scale <- Bergen_data$lavinntekt - mean(Bergen_data$lavinntekt) # Scaling
livingage_scale <- Bergen_data$levealder- mean(Bergen_data$levealder) # Scaling
cramped_scale <- Bergen_data$trangbodd - mean(Bergen_data$trangbodd) # Scaling
education_scale <- Bergen_data$utdanning - mean(Bergen_data$utdanning) # Scaling
imigration_scale <- Bergen_data$innvandring - mean(Bergen_data$innvandring) # Scaling
Pop_scale <- (Bergen_data$meanpopcell - mean(Bergen_data$meanpopcell))/10

## Prepare Covid wave data to NIMBLE model format:

```

```

# Wave 1
Wave <- "Wave 1"
wave0_start <- as.Date("2020-10-15")
wave0_end <- as.Date("2021-02-01")

bergen_wave0 <- covid19_bergen |>
  filter(date >= wave0_start & date <= wave0_end)

bydel_wave0 <- bergen_wave0 |>
  select(bydel_name,newdate,ins)|>
  pivot_wider(names_from = newdate, values_from = ins)

NAs <- as.data.frame(table(is.na(bydel_wave0)))
print(paste("Number of NAs in ", Wave, "is:", NAs[2,2],"of", (NAs[2,2]+NAs[1,2])
  , "data points", sep = " "))

bydel_wave0[is.na(bydel_wave0)] <- 0 # simple NA = 0 replacement.
sym_wave1 <-bydel_wave0[,2:(wave0_end-wave0_start+ 2)]
sym_wave1 <- as.matrix(sym_wave1)

ndistict <- nrow(sym_wave1) # M
nwave <- ncol(sym_wave1) # T
nneighbor <- length(AdjNIMBLE$adj) # L

death_dummy <- matrix(0,ndistict,nwave)

Wave1 <- list(Wave = Wave,
  start = wave0_start,
  end = wave0_end,
  sym_wave1 = sym_wave1,
  ndistict = ndistict,
  nwave = nwave,
  nneighbor= nneighbor)

# Wave 2
Wave <- "Wave 2"
wave0_start <- as.Date("2021-02-21")
wave0_end <- as.Date("2021-05-01")

bergen_wave0 <- covid19_bergen |>
  filter(date >= wave0_start & date <= wave0_end)

bydel_wave0 <- bergen_wave0 |>
  select(bydel_name,newdate,ins)|>
  pivot_wider(names_from = newdate, values_from = ins)

NAs <- as.data.frame(table(is.na(bydel_wave0)))
print(paste("Number of NAs in ", Wave, "is:", NAs[2,2],"of", (NAs[2,2]+NAs[1,2])
  , "data points", sep = " "))

bydel_wave0[is.na(bydel_wave0)] <- 0
sym_wave1 <-bydel_wave0[,2:(wave0_end-wave0_start+ 2)]
sym_wave1 <- as.matrix(sym_wave1)

```

```

ndistict <- nrow(sym_wave1) # M
nwave <- ncol(sym_wave1)   # T
nneighbor <- length(AdjNIMBLE$adj) # L

death_dummy <- matrix(0,ndistict,nwave)

Wave2 <- list(Wave = Wave,
              start = wave0_start,
              end = wave0_end,
              sym_wave1 = sym_wave1,
              ndistict = ndistict,
              nwave = nwave,
              nneighbor= nneighbor)

# Wave 3
Wave <- "Wave 3"
wave0_start <- as.Date("2021-07-15")
wave0_end   <- as.Date("2021-10-08")

bergen_wave0 <- covid19_bergen |>
  filter(date >= wave0_start & date <= wave0_end)

bydel_wave0 <- bergen_wave0 |>
  select(bydel_name,newdate,ins)|>
  pivot_wider(names_from = newdate, values_from = ins)

NAs <- as.data.frame(table(is.na(bydel_wave0)))
print(paste("Number of NAs in ", Wave, "is:", NAs[2,2],"of", (NAs[2,2]+NAs[1,2])
            , "data points", sep = " "))

bydel_wave0[is.na(bydel_wave0)] <- 0
sym_wave1 <-bydel_wave0[,2:(wave0_end-wave0_start+ 2)]
sym_wave1 <- as.matrix(sym_wave1)
sym_wave1 <- pmax(sym_wave1,0)# 2 values ar -1 and set to 0.

ndistict <- nrow(sym_wave1) # M
nwave <- ncol(sym_wave1)   # T
nneighbor <- length(AdjNIMBLE$adj) # L

death_dummy <- matrix(0,ndistict,nwave)

Wave3 <- list(Wave = Wave,
              start = wave0_start,
              end = wave0_end,
              sym_wave1 = sym_wave1,
              ndistict = ndistict,
              nwave = nwave,
              nneighbor= nneighbor)

# Wave 4
Wave <- "Wave 4"
wave0_start <- as.Date("2022-01-12") # 2022-01-01

```

```

wave0_end <- as.Date("2022-04-01")

bergen_wave0 <- covid19_bergen |>
  filter(date >= wave0_start & date <= wave0_end)

bydel_wave0 <- bergen_wave0 |>
  select(bydel_name,newdate,ins)|>
  pivot_wider(names_from = newdate, values_from = ins)

NAs <- as.data.frame(table(is.na(bydel_wave0)))
print(paste("Number of NAs in ", Wave, "is:", NAs[2,2],"of", (NAs[2,2]+NAs[1,2])
  , "data points", sep = " "))

bydel_wave0[is.na(bydel_wave0)] <- 0
sym_wave1 <-bydel_wave0[,2:(wave0_end-wave0_start+ 2)]
sym_wave1 <- as.matrix(sym_wave1)

ndistrict <- nrow(sym_wave1) # M
nwave <- ncol(sym_wave1) # T
nneighbor <- length(AdjNIMBLE$adj) # L

death_dummy <- matrix(0,ndistrict,nwave)

Wave4 <- list(Wave = Wave,
  start = wave0_start,
  end = wave0_end,
  sym_wave1 = sym_wave1,
  ndistrict = ndistrict,
  nwave = nwave,
  nneighbor= nneighbor)

# Stack data for Bergen in list:
Bergen <- list(district_data = Bergen_data,
  City = "Bergen",
  AdjNIMBLE = AdjNIMBLE,
  green_per = green_per,
  NDVI_scale = NDVI_scale,
  lowincome_scale = lowincome_scale,
  livingage_scale = livingage_scale,
  cramped_scale = cramped_scale,
  education_scale = education_scale,
  imigration_scale = imigration_scale,
  Pop_scale = Pop_scale,
  Wave1 = Wave1,
  Wave2 = Wave2,
  Wave3 = Wave3,
  Wave4 = Wave4
)

### Oslo
## Import data:

```

```

# Covid data:
# Filter out data that isn't from Oslo:
covid19_oslo <- covid19 |>
  filter(kommune_name == "Oslo") |>
  filter(grepl("Alna|Bjerke|Frogner|GamleOslo|Grorud|Grünerløkka|NordreAker|
              Nordstrand|Østensjø|Sagene|Søndre Norstrand|
              St. Hanshaugen|Stovner|Ullern|
              Vestre Aker", bydel_name))

# Population data
Pop_Oslo <- read_csv("Oslo250bydelSSB.csv")

# Green Area (%Ga) Bergen
Pe_Oslo <- read.csv("AreaBufferDiffSSB0slo.csv")

# NDVI Oslo
NDVI_Oslo <- read.csv("OsloNDVI14.02.2023.csv")

# Socio-demographic
sosio_Oslo <- read.csv("sosioall.csv") |>
  filter(bydel_id < 40000) # Chose data from Oslo
sosio_Oslo$Bydel <- str_sub(sosio_Oslo$Bydel,7,-2) # Select only district Name.
sosio_Oslo <- sosio_Oslo |>
  rename(name= Bydel)
sosio_Oslo$name[4] <- "St. Hanshaugen"

# Import district map
districts <- st_read("OsloOSM32N/OsloUtenHavMarkSentrum.shp")

## Aggregate data

# Population data
Pop_Oslo <- Pop_Oslo |>
  group_by(name) |>
  summarise(meanpopcell = mean(poptot))

# Percentage of green area in buffer:
Pe_Oslo <- Pe_Oslo |>
  mutate(green_fraction = poptot*((Bufferarea-Differenca)/Bufferarea))

Pe_Oslo <- Pe_Oslo |>
  group_by(name) |>
  summarise(green_fraction_bydel = sum(green_fraction),
            pop_bydel = sum(poptot)) |>
  mutate(green_percent = (green_fraction_bydel*100)/pop_bydel)

# NDVI
NDVI_Oslo <- NDVI_Oslo |>
  mutate(popNDVI = poptot*X_mean)

NDVI_Oslo <- NDVI_Oslo |>

```



```

group_by(name) |>
summarise(NDVI_sum = sum(popNDVI), pop_bydel = sum(poptot)) |>
mutate(NDVI_bydel =NDVI_sum/(pop_bydel))

# Make a dataframe that contains covariates data for the NIMBLE model:
Oslo_data <- left_join(Pe_Oslo,sosio_Oslo) |>
  left_join(Pop_Oslo) |>
  left_join(NDVI_Oslo) |>
  left_join(districts) |>
  st_as_sf()

## Adjacency Matrix:

# Make a Adjacency List of distict in Bergen.
AdjList <- poly2nb(Oslo_data)

# The NIMBLE package uses a different format on the matrix:
AdjNIMBLE <- nb2WB(AdjList)

## Center and scaling data for the MCMC samplers:
green_per <- Oslo_data$green_percent - mean(Oslo_data$green_percent) # Scaling
NDVIscale <- (Oslo_data$NDVI_bydel - mean(Oslo_data$NDVI_bydel))*100 #Scaling
lowincome_scale <- Oslo_data$lavinntekt - mean(Oslo_data$lavinntekt) # Scaling
livingage_scale <- Oslo_data$levealder- mean(Oslo_data$levealder) # Scaling
cramped_scale <- Oslo_data$trangbodd - mean(Oslo_data$trangbodd) # Scaling
education_scale <- Oslo_data$utdanning - mean(Oslo_data$utdanning) # Scaling
imigration_scale <- Oslo_data$innvandring - mean(Oslo_data$innvandring) # Scaling
Pop_scale <- (Oslo_data$meanpopcell - mean(Oslo_data$meanpopcell))/10

## Prepare Covid wave data to NIMBLE model format:

## Wave 1
Wave <- "Wave 1"
wave0_start <- as.Date("2020-10-15")
wave0_end <- as.Date("2021-02-01")

oslo_wave0 <- covid19_oslo |>
  filter(date >= wave0_start & date <= wave0_end)

bydel_wave0 <- oslo_wave0 |>
  select(bydel_name,newdate,ins)|>
  pivot_wider(names_from = newdate, values_from = ins) |>
  arrange(bydel_name)

NAs <- as.data.frame(table(is.na(bydel_wave0)))
print(paste("Number of NAs in ", Wave, "is:", NAs[2,2],"of", (NAs[2,2]+NAs[1,2])
  , "data points", sep = " "))

bydel_wave0[is.na(bydel_wave0)] <- 0

```

```

sym_wave1 <- bydel_wave0[,2:(wave0_end-wave0_start+ 2)]
sym_wave1 <- as.matrix(sym_wave1)

ndistict <- nrow(sym_wave1) # M
nwave <- ncol(sym_wave1) # T
nneighbor <- length(AdjNIMBLE$adj) # L

death_dummy <- matrix(0,ndistict,nwave)

Wave1 <- list(Wave = Wave,
             start = wave0_start,
             end = wave0_end,
             sym_wave1 = sym_wave1,
             ndistict = ndistict,
             nwave = nwave,
             nneighbor= nneighbor)

# Wave 2
Wave <- "Wave 2"
wave0_start <- as.Date("2021-02-21")
wave0_end <- as.Date("2021-05-01")

oslo_wave0 <- covid19_oslo |>
  filter(date >= wave0_start & date <= wave0_end)

bydel_wave0 <- oslo_wave0 |>
  select(bydel_name,newdate,ins)|>
  pivot_wider(names_from = newdate, values_from = ins)|>
  arrange(bydel_name)

NAs <- as.data.frame(table(is.na(bydel_wave0)))
print(paste("Number of NAs in ", Wave, "is:", NAs[2,2],"of", (NAs[2,2]+NAs[1,2])
           , "data points", sep = " "))

bydel_wave0[is.na(bydel_wave0)] <- 0
sym_wave1 <- bydel_wave0[,2:(wave0_end-wave0_start+ 2)]
sym_wave1 <- as.matrix(sym_wave1)

ndistict <- nrow(sym_wave1) # M
nwave <- ncol(sym_wave1) # T
nneighbor <- length(AdjNIMBLE$adj) # L

death_dummy <- matrix(0,ndistict,nwave)

Wave2 <- list(Wave = Wave,
             start = wave0_start,
             end = wave0_end,
             sym_wave1 = sym_wave1,
             ndistict = ndistict,
             nwave = nwave,
             nneighbor= nneighbor)

# Wave 3

```

```

Wave <- "Wave 3"
wave0_start <- as.Date("2021-07-20")
wave0_end <- as.Date("2021-10-08")

oslo_wave0 <- covid19_oslo |>
  filter(date >= wave0_start & date <= wave0_end)

bydel_wave0 <- oslo_wave0 |>
  select(bydel_name,newdate,ins)|>
  pivot_wider(names_from = newdate, values_from = ins) |>
  arrange(bydel_name)

NAs <- as.data.frame(table(is.na(bydel_wave0)))
print(paste("Number of NAs in ", Wave, "is:", NAs[2,2],"of", (NAs[2,2]+NAs[1,2])
  , "data points", sep = " "))

bydel_wave0[is.na(bydel_wave0)] <- 0
sym_wave1 <-bydel_wave0[,2:(wave0_end-wave0_start+ 2)]
sym_wave1 <- as.matrix(sym_wave1)

ndistict <- nrow(sym_wave1) # M
nwave <- ncol(sym_wave1) # T
nneighbor <- length(AdjNIMBLE$adj) # L

death_dummy <- matrix(0,ndistict,nwave)

Wave3 <- list(Wave = Wave,
  start = wave0_start,
  end = wave0_end,
  sym_wave1 = sym_wave1,
  ndistict = ndistict,
  nwave = nwave,
  nneighbor= nneighbor)

# Wave 4
Wave <- "Wave 4"
wave0_start <- as.Date("2022-01-12") #2022-01-01
wave0_end <- as.Date("2022-04-01")

oslo_wave0 <- covid19_oslo |>
  filter(date >= wave0_start & date <= wave0_end)

bydel_wave0 <- oslo_wave0 |>
  select(bydel_name,newdate,ins)|>
  pivot_wider(names_from = newdate, values_from = ins) |>
  arrange(bydel_name)

NAs <- as.data.frame(table(is.na(bydel_wave0)))
print(paste("Number of NAs in ", Wave, "is:", NAs[2,2],"of", (NAs[2,2]+NAs[1,2])
  , "data points", sep = " "))

bydel_wave0[is.na(bydel_wave0)] <- 0 # there are 2 values as NA in the dataset.
sym_wave1 <-bydel_wave0[,2:(wave0_end-wave0_start+ 2)]

```

```

sym_wave1 <- as.matrix(sym_wave1)

ndistict <- nrow(sym_wave1) # M
nwave <- ncol(sym_wave1)    # T
nneighbor <- length(AdjNIMBLE$adj) # L

death_dummy <- matrix(0,ndistict,nwave)

Wave4 <- list(Wave = Wave,
              start = wave0_start,
              end = wave0_end,
              sym_wave1 = sym_wave1,
              ndistict = ndistict,
              nwave = nwave,
              nneighbor= nneighbor)

## Stack data in lists and save for modelsimulation:

Oslo <- list(district_data = Oslo_data,
             City = "Oslo",
             AdjNIMBLE = AdjNIMBLE,
             green_per = green_per,
             NDVI_scale = NDVIscale,
             lowincome_scale = lowincome_scale,
             livingage_scale = livingage_scale,
             cramped_scale = cramped_scale,
             education_scale = education_scale,
             imigration_scale = imigration_scale,
             Pop_scale = Pop_scale,
             Wave1 = Wave1,
             Wave2 = Wave2,
             Wave3 = Wave3,
             Wave4 = Wave4
            )

Masterdata <- list(Oslo = Oslo,
                  Bergen = Bergen)

save(Masterdata, file = "Results/Masterdata.rda")

```

### 3 NIMBLE code:

Based on code supplemented by Lawson and Kim (2021) is the following code used to do Bayesian modeling:

```

# Loading packages:
library(tidyverse) # Multiple data manipulation and data visualisation packages
library(patchwork) # Package to combine plots
# Statistical analyses:
library(nimble) # Package for Bayesian regression
library(coda) # Package for MCMC output
library(ggmcmc) # Package for MCMC and plotting
library(MCMCvis) # Package for MCMC and plotting
library(igraph)

```

```

library(R6)
library(EigenR)
library(foreign)

## Import generated data:
load("Results/Masterdata.rda")

#Define model to run:
place <- "Bergen" # City Oslo or Bergen
wave <- "Wave1" # Wave 1-4 pre-defined
b2 <- "NDVI_scale" # Green acces parameter NDVI_scale or green_per
b3 <- "livingage_scale" # Sosio-demographic parameter
b4 <- "lowincome_scale" # Sosio-demographic parameter
alpha = 0.2 # alpha for asymptomatic infected 0.2, 0.33 or 0.5
mni <- 2 # number for million of iterations

# Simulation settings:
ni <- mni*1000000 # number of iterations
nb <- 500000 # number of iterations as burn-in
nt <- 300 # thinning interval
nc <- 2 # number of chains

# Parameters to monitor during simulation:
my.mon = c("RRbet1", "RRbet2", "RRbet3", "RRbet4",
           "bet0", "bet1", "bet2", "bet3", "bet4", "B", "DevC")

# Load selected data for simulation:
ndistrict <- Masterdata[[place]][[wave]]$ndistrict
nwave <- Masterdata[[place]][[wave]]$nwave
nneighbor <- Masterdata[[place]][[wave]]$nneighbor
sym_wave1 <- Masterdata[[place]][[wave]]$sym_wave1
data2 <- Masterdata[[place]][[b2]]
data3 <- Masterdata[[place]][[b3]]
data4 <- Masterdata[[place]][[b4]]
districtpop <- Masterdata[[place]]$district_data$pop_bydel
AdjNIMBLE <- Masterdata[[place]]$AdjNIMBLE

# Name of model for saving to disk.
modelName <- paste0(place, "_", wave, "_", b2, "_", b3, "_", b4, "_", mni, "M", alpha, "alpha.rda")
print(modelName)

# Initial values:
myInits <- list(
  list(tau0 = 0.5,
       tauB = 0.5,
       B = rep(1.0, ndistrict),
       bet0 = rep(0.1, nwave),
       muc = matrix(2, ndistrict, nwave),
       bet1 = 0.1,
       tau1 = 0.5,
       bet2 = 0.1,
       tau2 = 0.5,
       bet3 = 0.1,

```

```

    tau3 = 0.5,
    bet4 = 0.1,
    tau4 = 0.5),
list(tau0 = 0.5,
    tauB = 0.5,
    B = rep(1.0,ndistict),
    bet0 = rep(-0.1, nwave),
    muc = matrix(2, ndistict, nwave),
    bet1 = -0.1,
    tau1 = 0.5,
    bet2 = -0.1,
    tau2 = 0.5,
    bet3 = -0.1,
    tau3 = 0.5,
    bet4 = -0.1,
    tau4 = 0.5)
)

## Import data and constants to the model:

myData <- list(sym = sym_wave1,
              x2 = data2,
              x3 = data3,
              x4 = data4,
              alpha =alpha)

myConsts <- list(M = ndistict,T = nwave, L = nneighbor,
                susint = districtpop,
                num = AdjNIMBLE$num,
                adj = AdjNIMBLE$adj
)

## The model:
# Model is adapted from code supplemented in Lawson, A and Kim, J (2021)
# Space-time covid-19 Bayesian SIR modeling in South Carolina. PLoS ONE 16 (3)
# https://doi.org/10.1371/journal.pone.0242777.

COV4d <- nimbleCode({
  for(i in 1:M){
    remc[i,1]<-0           # Recovery
    remD[i,1]<-0           # Removed (Death)
    susc[i, 1] <- susint[i] # Susceptible
    muc[i,1] <- 0.001*susc[i,1] # mu
    sym[i,1] ~ dpois(muc[i,1]) # Symptomatic
    asym[i,1] <- alpha*sym[i,1] # Asymptomatic

    LdevC[i,1]<--2*(sym[i,1]*log(muc[i,1]+0.001)
                  -(muc[i,1]+0.001)-lfactorial(sym[i,1]))
  }

  for(i in 1:M){

```

```

for(j in 2:T){
  remc[i,j]<-betaRc*sym[i,j]
  #      remD[i,j]<- deaths[i,j]
  remD[i,j]<- 0
  susc[i,j]<-susc[i,j-1]-sym[i,j-1]-asym[i,j-1]-remc[i,j-1]-remD[i,j-1]
  sym[i,j] ~ dpois(muc[i,j])
  log(muc[i,j]) <- B[i] + bet0[j] + log(susc[i,j]+ 0.001)+
    bet1 * (log(sym[i, j-1] + 0.001) + log(asym[i, j-1] +.001))+
    bet2*x2[i] + bet3 * x3[i] + bet4 * x4[i]

  asym[i,j]<-alpha*sym[i,j]

  LdevC[i,j]<--2*(sym[i,j]*log(muc[i,j]+0.001)
    -(muc[i,j]+0.001)-lfactorial(sym[i,j]))
}
#      v[i] ~ dnorm(0, tauv)
}

for(k in 1:L){wei[k] <- 1}
B[1:M] ~ dcar_normal(adj[1:L], wei[1:L], num[1:M], tauB, zero_mean=1)

bet0[1] ~ dnorm(0, tau0) # Time dependent bet0
for(k in 2:T){
  bet0[k] ~ dnorm(bet0[k-1], tau0)
  #      bet0[k] ~ dnorm(0, tau0)
}

#      bet0 ~ dnorm(0, tau0)
bet1 ~ dnorm(0, tau1)
RRbet1 <- exp(bet1)
bet2 ~ dnorm(0, tau2)
RRbet2 <- exp(bet2)
#      tauv ~ dgamma(2, 0.05)
tauB ~ dgamma(2, 0.5)
tau0 ~ dgamma(2, 0.5)
tau1 ~ dgamma(2, 0.5)
tau2 ~ dgamma(2, 0.5) #0.05
bet3 ~dnorm(0, tau3)
RRbet3 <- exp(bet3)
tau3 ~ dgamma(2, 0.5) #0.05
bet4 ~dnorm(0, tau4)
RRbet4 <- exp(bet4)
tau4 ~ dgamma(2, 0.5) #0.05
#      bet5 ~dnorm(0, tau5)
#      tau5 ~ dgamma(2, 0.05)
#      bet6 ~dnorm(0, tau6)
#      tau6 ~ dgamma(2, 0.05)

DevC<-sum(LdevC[1:M,1:T])

for (j in 1:T){
  muc1[j]<-muc[1,j]
  muc2[j]<-muc[2,j]

```

```

    muc3[j]<-muc[3,j]
  }
  betaRc <- 0.1
})

## MCMC configurations
# I change samplers in some chains to obtain better mixing.

ProgjModel <- nimbleModel(code = COV4d,
                          data = myData,
                          constants = myConsts)
C_ProgjModel <- compileNimble(ProgjModel)

ConfSampler <- configureMCMC(ProgjModel,enableWAIC = TRUE, print = TRUE)
ConfSampler$addMonitors(my.mon)
ConfSampler$printSamplers()
ConfSampler$removeSamplers(c("bet2","bet3","bet4"))
ConfSampler$addSampler(target = c("bet2","bet3","bet4"), type = "AF_slice")
ConfSampler$printSamplers()
SliceSamplerMCMC <- buildMCMC(ConfSampler)
Sliceengine <- compileNimble(SliceSamplerMCMC,
                             project = ProgjModel,
                             resetFunctions = TRUE)

print(modelname)
## Run the model:

t0 <- Sys.time() # For MCMC runtime
mcmc.outd <- runMCMC(mcmc = Sliceengine,
                    niter = ni,
                    nburnin = nb,
                    thin = nt,
                    nchains = nc,
                    inits = myInits,
                    summary = TRUE,
                    samplesAsCodaMCMC = TRUE,
                    WAIC = TRUE)

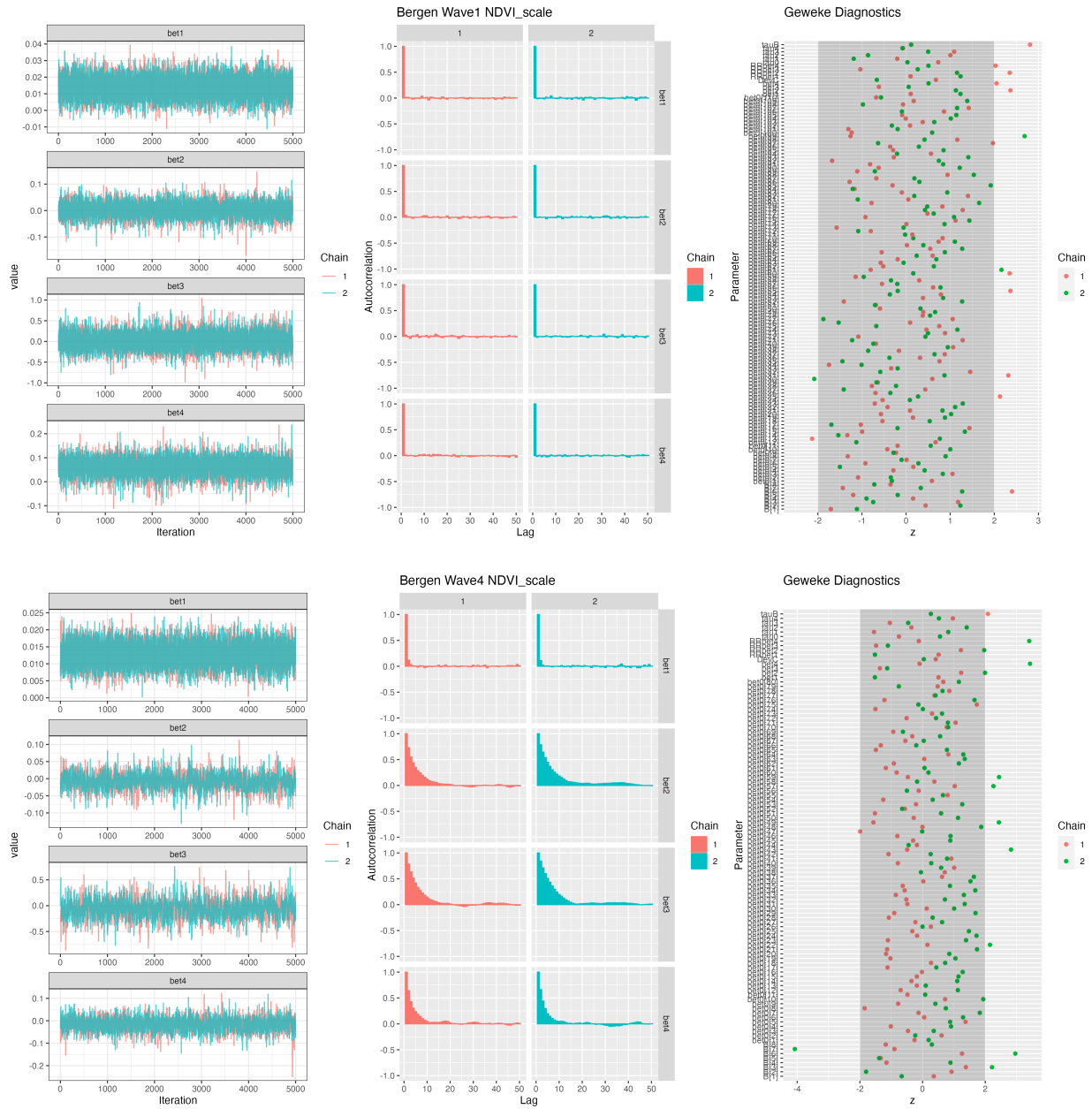
t1 <- Sys.time()
runtime <- round(t1-t0, digits = 1)
print(runtime)
# Save simulation:
simulationlist <- list(mcmc.outd = mcmc.outd,
                       place = place,
                       wave = wave,
                       b2 = b2,
                       b3 = b3,
                       b4 = b4,
                       alpha = alpha,
                       mni = mni,
                       runtime = runtime)
save(simulationlist, file = paste0("Results/Simulations/",modelname))

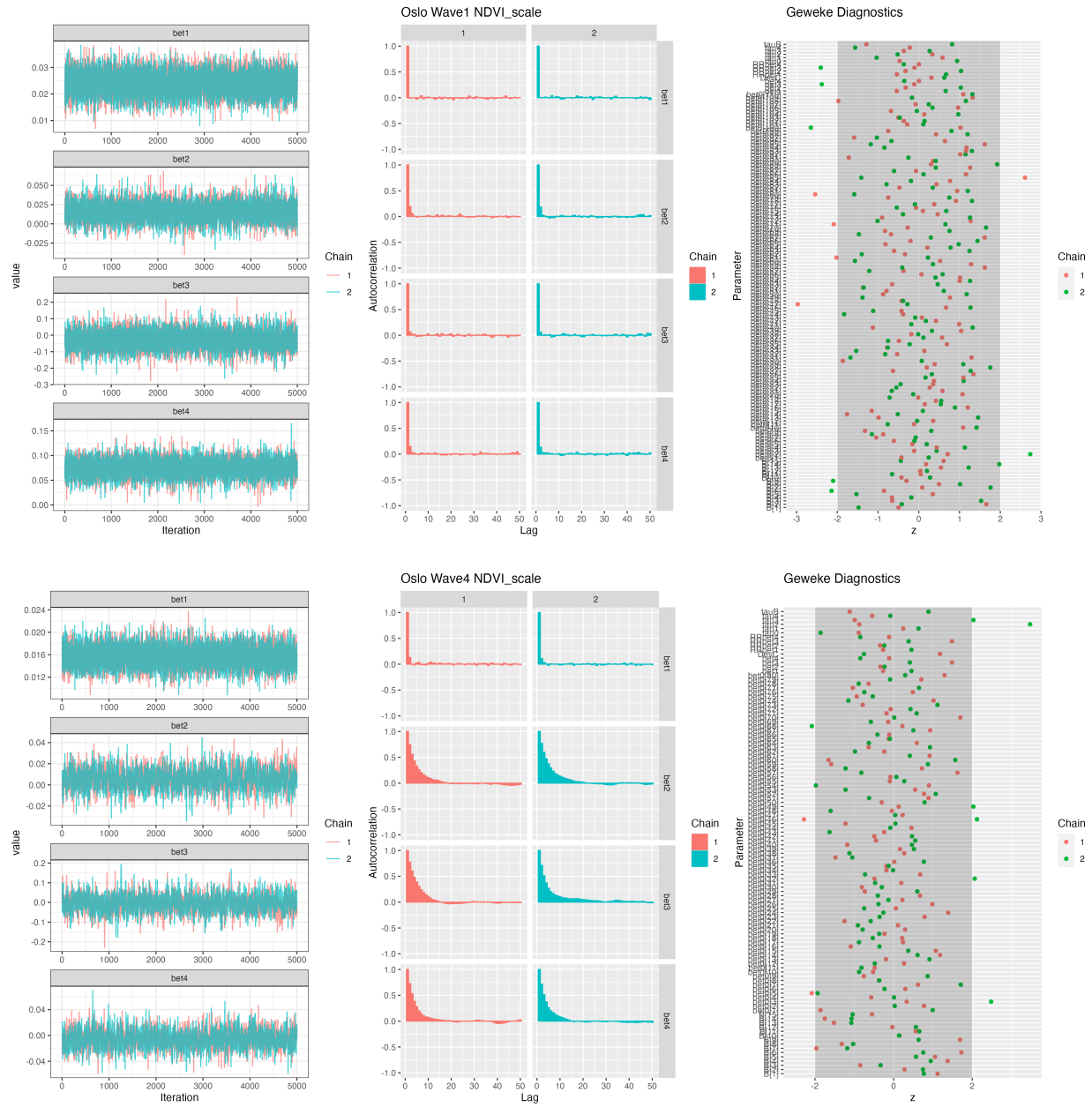
```



# 4 MCMC diagnostics

MCMC diagnostics was used with the package ggmcmc by Xavier Fernández i Marín (n.d.) .





## References

- Lawson, A. B., and J. Kim. 2021. "Space-Time Covid-19 Bayesian SIR Modeling in South Carolina." *Journal Article. PLoS One* 16 (3): e0242777. <https://doi.org/10.1371/journal.pone.0242777>.
- Xavier Fernández i Marín. n.d. "ggmcmc: Graphical Tools for Analyzing Markov Chain Monte Carlo Simulations from Bayesian Inference." *cran-r*. <http://xavier-fim.net/packages/ggmcmc>.



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

Postboks 5003  
NO-1432 Ås  
Norway