

Prosjekt:

Slik endres Norge



Innsendt

Data-SKUP 2023

Innsendere

Runa Victoria Engen

runa.engen@vg.no

+47 41147181

Jari Bakken

jari@vg.no

+47 48185271

Ingeborg Huse Amundsen

Ingeborg.Huse.Amundsen@vg.no

+47 92276929

Espen Rasmussen

espen.rasmussen@vg.no

+47 90549004

Simen Grytøyr

simen.grytoyr@vg.no

+47 41541450

Einar Otto Stangvik

einar.otto.stangvik@vg.no

+47 92233601

1. Innledning	2
2. Ideen	2
2.1 Få oversikt over tilgjengelig data	2
3. Veivalg	3
3.1 Målestasjoner: Alle målinger eller der folk flest bor?	3
3.2 Regn: Millimeter eller prosent?	3
3.3 Snø: Dager med snø eller skivær?	4
3.4 Tidsserier: Hva sammenligner vi?	4
4. Datagrunnlaget	6
5. Visuell presentasjon	7
6. Åpenhet og deling	9
7. Merverdi og etterliv	9
8. Resultat	10

URL hovedsak

<https://www.vg.no/spesial/klima/slik-endres-norge/>

Tilknyttede artikler:

Hvor fort gror Norge igjen?

<https://www.vg.no/spesial/klima/slik-endres-norge/artikler/etne/>

Havets bunn

<https://www.vg.no/spesial/klima/slik-endres-norge/artikler/havets-bunn/>

Torskens vandring

<https://www.vg.no/spesial/klima/slik-endres-norge/artikler/fiske-flukten/>

Kloakken flommer over

<https://www.vg.no/spesial/klima/slik-endres-norge/artikler/oslo-overlop/>

Ekstremtørken

<https://www.vg.no/spesial/klima/slik-endres-norge/artikler/ekstrem-torke/>

1. Innledning

Vi snakker oftere om klima og global oppvarming, men hva er egentlig klimaendringer? Hvordan ser de ut der folk bor?

Høsten 2022 skrev VG flere saker om konsekvensene av global oppvarming, blant annet om hetebølge og tørke på flere kontinenter. De mest dramatiske effektene av oppvarmingen er ofte i land langt unna, eller langt fram i tid. Samtidig vet vi at klimaendringer er synlige her og nå, også i Norge. Ikke minst fordi oppvarmingen er høyere jo lenger nord man kommer, så når temperaturen i verden har økt med én grad i gjennomsnitt, er økningen rundt to grader i Norge.

VG ønsket å gjøre klimaendringene konkret og synlig. Ved å analysere og bearbeide et av de desidert største datasettene VG noensinne har jobbet med, viste vi frem endringene som som skjer der vi bor - akkurat nå.

2. Ideen

VGs journalist hadde over tid ønsket å bringe temaet klima tettere på folk. Ideen startet med et ønske om å vise fram den oppvarmingen som allerede har skjedd, og at det går raskere i Norge enn globalt. Det er også stor forskjell i klimaendringenes effekt internt i landet. Med VGs kapasitet til å behandle store datasett var det en naturlig tanke at vi kunne gi lesere muligheten til å søke opp eget hjemsted og lage kart som viste de store kontrastene i landet. Reportasjer om enkelte utslag av oppvarmingen skulle ta et overveldende komplekst tema og gjøre det håndgripelig.

At leserne skulle få se endringene med egne øyne var en bærende tanke. Det var viktig at tjenesten skulle ha robust kode som kontinuerlig oppdateres med tilgjengelige data, slik at den er aktiv og kan leve i mange år. Spesialen skulle danne grunnlaget for et helt univers med klimastoff.

2.1. Få oversikt over tilgjengelig data

Det er Meteorologisk institutt (Met) som eier værdata og temperaturmålinger i Norge. De har også spisskompetansen innen behandling av værdata for å illustrere klimaendringer. Det var derfor naturlig å først kontakte dem. De var svært positive til ideen, og i januar 2023 hadde vi et første møte for å se nærmere på muligheter og begrensninger. Ingen andre mediehus hadde behandlet disse dataene på egen hånd og på denne måten før.

Vi ønsket data så langt tilbake i tid som mulig. Met har detaljerte data over temperatur, regn og snødybde i et rutenett med 1x1 kilometer oppløsning for hele landet, tilbake til 1957. Vi kom altså ikke til førindustriell tid, men det er 65 sammenhengende år. Mets forskningsleder forklarte at i Norge har klimaendringene vært særlig synlig i statistikken siden tidlig 80-tallet. Derfor var det viktigst at vi hadde data før dette.

3 Veivalg

3.1 Målestasjoner: Alle målinger eller der folk flest bor?

Met ga oss et råd i det første møtet, basert på tidligere erfaringer. De hadde tidligere regnet ut de forventede klimaendringene i framtiden, men som et gjennomsnitt av alle målestasjoner i kommunen. En del målestasjoner står høyt på fjellet, eller i øde områder, mens befolkningen oftest bor sentralisert og i lavere strøk eller tett på vann. Høydeforskjell har store utslag på vær, og beregningene ble mindre representative. De rådet oss derfor til å velge ut de målestasjonene som ligger i de mest folkerike delene av hver kommune.

Dette ble diskutert i VG, og vi kom fram til at det viktigste for oss var å vise været der folk flest bor. Dette innebar samtidig at arbeidet VG nå gjorde, ikke hadde blitt gjort før. Det ville gi et resultat heller ikke instituttet hadde sett tidligere.

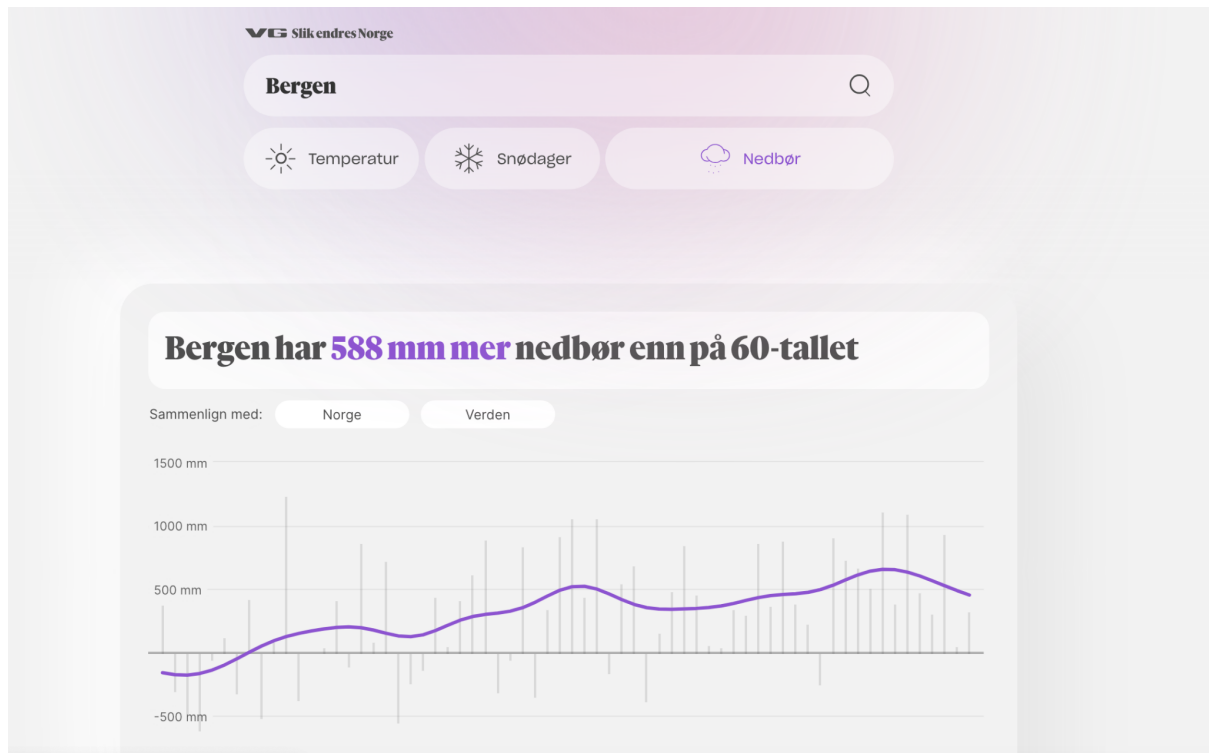
For å kunne hente ut meteorologiske data for den mest folkerike delen av landets kommuner, trengte vi en oversikt over hvor folk bor med like høy oppløsning som Met sine klimadata. Dette fant vi i [SSBs geodata](#), som inneholder befolkningstall i et fast geografisk rutenett på 1x1 kilometer fra 2019. Rutenettet består av omtrent 55 000 slike datapunkter over hele landet.

Vi lastet rutenettet fra SSB inn i databasesystemet PostgreSQL med utvidelsen PostGIS, som muliggjør geografiske spørringer mot databasen. Vi lastet også inn kommunegrensene. Til sammen kunne vi da skrive en SQL-spørring som fant den mest folkerike kvadratkilometeren i hver kommune. Koden er tilgjengelig som [åpen kildekode](#). Vi kvalitetssjekk resultatet i geoverktøyet QGIS.

3.2 Regn: Millimeter eller prosent?

Et annet valg var spørsmålet om hvordan vi ønsket å gjengi endringer i regn. For noen gir det mest mening med prosent, men da får man ikke fram at 20 prosent økning i nedbør i Eigersund er noe helt annet enn i Bergen. Samtidig er det umiddelbart satt i perspektiv.

Her er det ikke ett svar som virket mest riktig eller galt. Nordmenn er generelt værinteresserte og gjengir ofte nedbør i millimeter i dagligtale. Derfor endte vi til slutt med mm, mens vi omtaler den forventede endringen i framtiden i prosent.



3.3 Snø: Dager med snø eller skivær?

Mens snøen forsvinner lavt i landet og langs kysten, blir det mer snø i fjellene. Vi ville vise denne forskjellen med kart, men skrive om hvordan snøen forsvinner i de områdene der det bor flest folk.

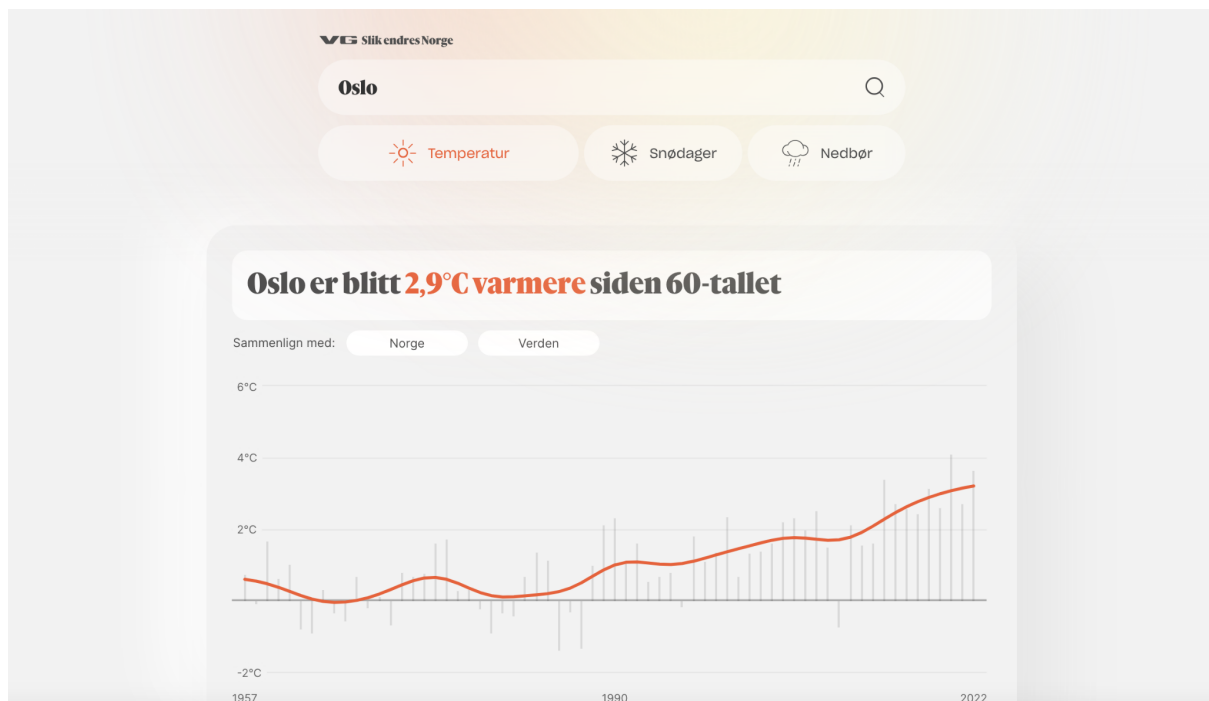
Det kan man beskrive med tap av dager med skiføre (minst 25 cm snø) eller tap av dager med snødekke på bakken (5 eller 10 cm med snø). Igjen var det sentralt for oss at det vi omtalte skulle lesere kjenne igjen. Mange steder har det aldri vært regelmessig dager med skiføre, men de har hatt dager med snø. Vi valgte derfor den laveste mengden man ville ansett som en "hvit snødag", med fem centimeter dekke på bakken.

3.4 Tidsserier: Hva sammenligner vi?

I det første møtet med Met ble det også diskutert hvilken tidsperiode vi setter som et utgangspunkt for å sammenligne med nåtid. I klimatologi er det vanlig å si at "klima er vær i serier på minst 30 år". Det utjevner naturlige variasjoner i vær med svingende varme og kalde perioder. Samtidig øker tempoet i oppvarmingen, temperaturen stiger mer per tiår nå enn før. En periode på 30 år vil kunne underdrive

de faktiske endringene, mens kortere perioder gjør det vanskelig å skille klimaendringer fra naturlige variasjoner i vær.

Klimaendringene har blitt enda tydeligere det siste tiåret, til tross for en periode med unaturlig kaldt vær. VG valgte derfor å sammenligne perioder på ti år, for å vise et mest mulig oppdatert bilde av hvordan temperatur, nedbør og snømengde har endret seg. Det innebar at vi måtte være tydelige på hvilke perioder vi sammenligner, og at dette er *endringer i vær* påvirket av oppvarmingen. Som en del av vår research og kontroll av dataene, diskuterte vi dette valget med klimaforskere fra flere institusjoner, som NMBU, UIB, Cicero senter for klimaforskning og stiftelsen Zero. Vi tok også med oss den ferdige søkbare tjenesten til Meteorologisk institutt før publisering, for å kontrollere at metoden vår ville gi et representativt bilde av klimaendringenes effekt på været.



4. Datagrunnlaget

For å kunne lage tjenesten vi ønsket, trengte vi data med høy geografisk oppløsning og lengst mulig historikk. Vi fikk råd fra Meteorologisk institutt om den beste kilden i norsk sammenheng er de observasjonsbaserte rutenettdataene (se [Lussana et al. 2019](#) og [Saloranta, T. M. 2016](#)) i prosjektet [SeNorge](#).

Datasettet oppdateres daglig med data om temperatur og nedbør over hele landet på et 1x1 kilometer rutenett, og finnes med samme detaljeringsgrad helt tilbake til 1957.

SeNorge bruker en statistisk modell for å beregne verdier for hver kvadratkilometer i Norge, basert på observasjoner fra målestasjoner over hele landet. Siden målestasjoner kan komme og gå over tid, gjøres det en statistisk interpolering mellom de til enhver tid tilgjengelige målingene for hvert område og tidspunkt. Dette er det nærmeste vi kommer et konsistent datasett for værobservasjoner som dekker hele landet over en lang tidsperiode.

Vi lastet ned de daglige datasettene fra METs datatjeneste [Thredds](#) for alle årene mellom 1957 og 2022 for både *seNorge* (temperatur og nedbør) og *seNorge_snow* (snø). Komprimert utgjør datasettet 150 GB med data, ukomprimert er det snakk om over 900 GB. Totalt er det over 220 milliarder datapunkter i materialet vi har analysert.

Størrelsen på dette rådatasettet kommer av høy oppløsning i flere dimensjoner. Det er daglige data i 65 år, som dekker hver kvadratkilometer i landet for fem ulike variabler. Hver dag inneholder data om 1,8 millioner geografiske ruter. Et år med data utgjør dermed over 670 millioner datapunkter.

Dette er klart et av de største datasettene vi noensinne har jobbet med i VG.

Datamengden gjorde at våre første naive forsøk på å analysere dem måtte forkastes, og vi brukte en del tid på å finne ut hvordan vi kunne håndtere størrelsene uten at prosesseringen brukte altfor lang tid eller at maskinene vi jobbet på gikk tom for minne underveis.

Dataene kom i filformatet NetCDF, som fra før var helt ukjent for oss. Etter noen tidlige forsøk på å behandle filene i programmeringsspråket R, landet vi på at Python ville være bedre egnet. Biblioteket [xarray](#) var til stor hjelp for å sikre at vi kunne jobbe med filene på en effektiv måte. Vi fikk her også god hjelp av OpenAIs språkmodell ChatGPT for å forstå hvordan NetCDF og xarray fungerte og hvordan oppgavene kunne løses.

All koden som er brukt for nedlasting, prosessering og analyse av dataene er lagt ut som [åpen kildekode på Github](#).

For å redusere rådataene til noe vi kunne presentere for leserne, måtte vi aggregere de daglige verdiene til årlige verdier. Dette var variablene vi var interessert i:

- Gjennomsnittlig av daglig temperatur
- Sum av daglig nedbør

- Gjennomsnittlig snødybde per dag
- Antall skidager (mer enn 25 cm snø)
- Antall snødager (mer enn 5 cm snø)

Koden beregner dette for hver av de 1,8 millioner rutene i det geografiske rutenettet, og dette gjentas for hvert av de 65 årene. Vi brukte så den tidligere omtalte rutenettstatistikken fra SSB for å trekke ut de arealutsnittene vi var interessert i – ett isolert datasett per kommune. De ferdige filene som ble brukt til å bygge tjenesten inneholdt:

- **Tidslinje per kommune:** Årlig tidslinje for alle de fem variablene for den mest befolkningsrike kvadratkilometeren i kommunen
- **Nasjonal tidslinje:** Et gjennomsnitt av kommunenes tidslinjer
- **Kartdata:** Endringen fra første til siste tiår for temperatur, nedbør og antall snødager for hver rute i rutenettet.

For å kunne tilby sammenligninger i tjenesten hentet vi også inn data om global temperatur og nedbør fra National Oceanic & Atmospheric Administration (NOAA) og Environmental Protection Agency (EPA).

5. Visuell presentasjon

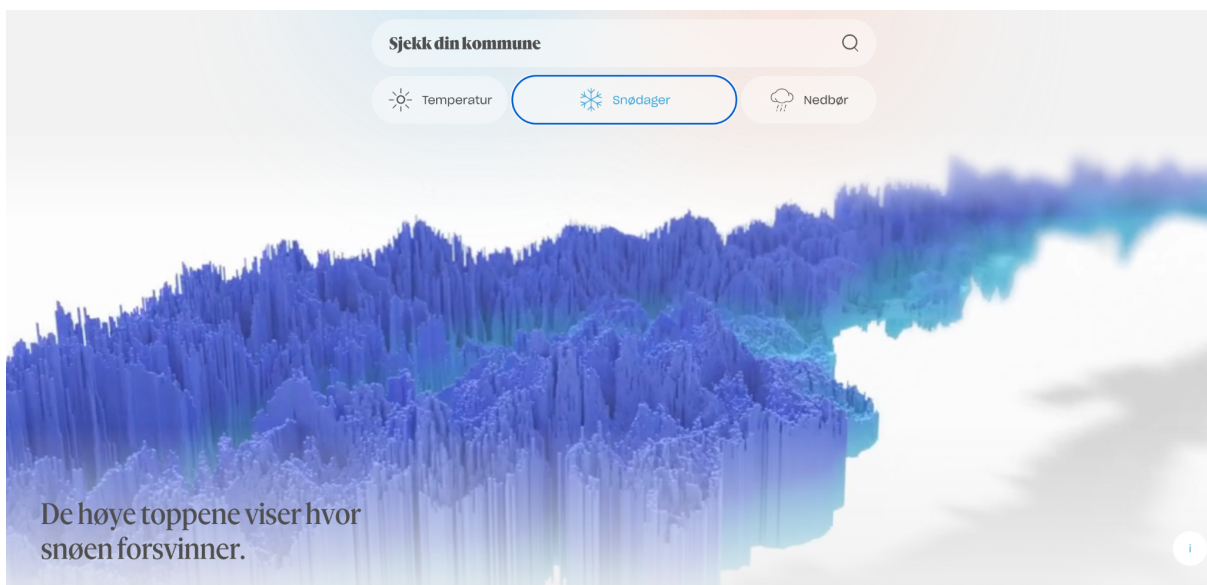
Siden premisset handlet om å vise frem tallmateriale på lokalt nivå, satte vi oss noen visuelle regler og rammer for hva prosjektet skulle oppleves som.

- Tallene er stjernen. Den visuelle fremstillingen skal fremheve funnene, og det er grafikken som skal bære presentasjonen.
- Det skal være enkelt å søke opp nye kommuner og bytte variabel (mellom temperatur, snødager og nedbør). Derfor er søket alltid til stede i skjermbildet sammen med meny der leseren kan velge hvilken variabel den vil se. Det skal være sømløst å bytte kommune eller variabel. Det skal ikke virke som om brukeren forlater en side og kommer til en annen.
- En konsekvens av at tallene er stjernen, er at vi fjernet farger fra alt, bortsett fra elementene som viser tallene og det som handler om navigering mellom de tre variablene (temperatur, snødager og nedbør). Hver av variablene fikk sitt eget fargesett, så når brukeren ser på en graf, skal den vite hva den handler om. Grafene skifter fargen når man navigerer.
- Titlene i grafene skal være fortellende, slik at brukeren intuitivt får hjelp til å forstå hva grafen og tallene viser. Dette bidrar også til at grafene blir enklere delbare. En isolert graf forteller dens egen historie.

- Holde alle reportasjeelementer utenfor i tjenesten som handler om tallene. Det er også en konsekvens av at tallene skal bære presentasjonen, samtidig som det vil skape en god dynamikk mot reportasjene vi lagde rundt tjenesten.

Dette kom raskt på plass, og vi designet opp en visuell profil for prosjektet. Samtidig visste vi tidlig at vi på nasjonalt nivå ville fremstille dette i form av kart – men uten en konkret ide om hvordan.

Vi hadde lagt merke til interessante visualiseringer av høyoppløselige befolkningsdata fra Twitter-brukeren [researchremora](#), og forsøkte å gjenskape metoden med værdataene vi hadde samlet. Det ble først gjort forsøk med pakken [Rayshader](#) i programmeringsspråket R, men det ble for tungvint for animasjonen vi ønsket oss. Vi koblet på redaksjonell utvikler Einar Otto Stangvik, som er den med størst kompetanse på 3D i VG. Ved hjelp av softwaren [Blender](#) kunne han importere datasettene vi hadde forberedt og lage kartanimasjoner som ble et bærende visuelt element i hovedsaken.



Det var et ønske helt fra starten av prosjektet at vi skulle skape et univers, der tjenesten lever hånd i hånd med reportasjene rundt, der alt henger sammen. Leserne skulle aldri «sette seg fast», men heller bli oppfordret til videre lesning.

Når det kommer til reportasjene, hadde vi to tydelige premisser: Vise tilstedeværelse, samtidig som vi fikk frem forskningen og tallene. Det førte til at vi landet på en todelt løsning som skaper en trekkspill-dynamikk avhengig av hvor visuell saken var. Når vi er til stede i reportasjen er saken bildedrevet, med korte tekster og store bilder. Når saken går inn i forskningen og tallene fjernes det meste av bilder og tekstuell tilstedeværelse. På denne måten ønsket vi å få frem en intern dynamikk i saken. Det

ble da enkelt å strukturere reportasjene i idé-, planlegging- og produksjonsfasen, samtidig som det gir en stram leserkontrakt.

6. Åpenhet og deling

Vi ønsket å være åpne om hvordan vi hadde jobbet med dataene, og hvilke valg vi hadde tatt underveis. Vi laget derfor en faktaboks der vi forklarte hva funnene våre var basert på og hvordan vi hadde regnet.

I VG har vi en åpenhetsportal for å gi leserne innsikt i hvordan vi jobber og våre avveininger. Vi skrev en egen [redaksjonell vurdering](#) for prosjektet der vi redegjorde for de viktigste avveiningene og la den inn i samtlige reportasjer.

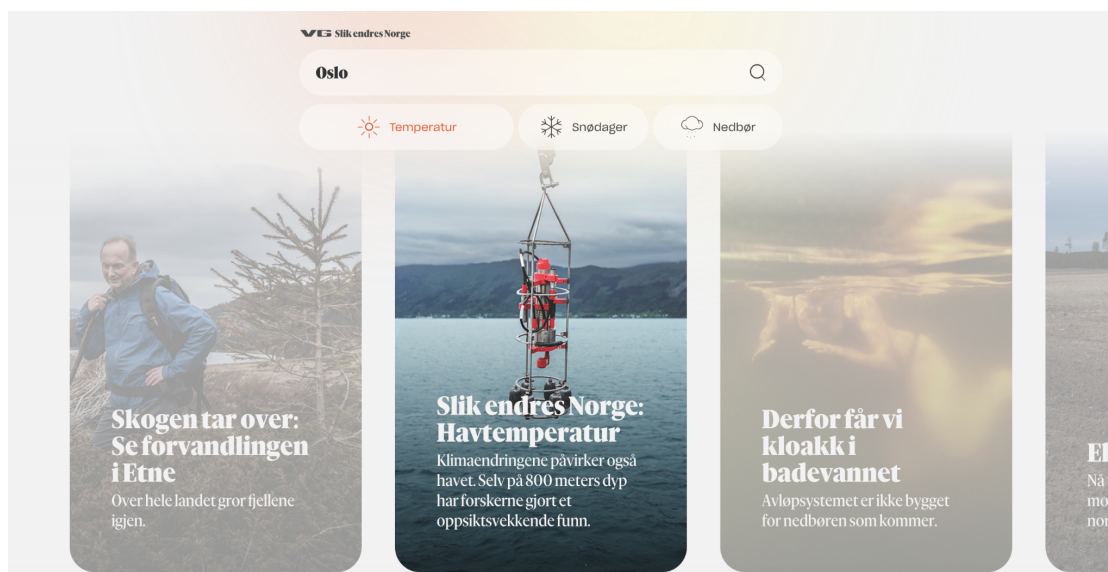
Vi valgte også å dele koden for databehandlingen som [åpen kildekode på Github](#). I tillegg til å bidra til åpenhet om våre metodiske valg og la spesielt interesserte lesere ettergå arbeidet, ønsker vi at det er et bidrag til den voksende delingskulturen innen datajournalistikk. Vi har selv lært og dratt nytte av åpen kildekode fra norske og internasjonale medier, vi ønsker å bidra til at dette blir mer utbredt.

7. Merverdi og etterliv

Før første publisering av hovedspesialen hadde vi startet arbeidet med fem reportasjer. Utgangspunktet var å ta for seg ulike tema, som styrtregn, tørke, endringer i naturen og i havet, og deretter lete etter måter å illustrere det på.

Vi kan for eksempel ikke nødvendigvis planlegge at vi skal klare å vise fram styrtregn, men vi kan vise en av konsekvensene av det. Derfor valgte vi å fokusere på at kloakk slippes ut i fjordene i Oslo når avløpssystemet blir overveldet av de enorme vannmengdene.

Rækkefølgen for publisering av reportasjene ble endret for at de skulle føles relevante for leserne, slik at saken om kloakk i badevannet kom en helg med badetemperaturer. Det var et mål at ikke reportasjene skulle føles som en folkeopplysning, men heller treffe leserne på det de var opptatt av. Det er en styrke for prosjektet at vi kan fortsette å sette fokus på nye temaer etter hvert som de blir relevante og inkludere det i samme univers.



Det er samme fotograf som har tatt bilder til alle disse reportasjene. Det er med på å skape et helhetlig visuelt univers og var uvurderlig i et prosjekt der temaene tross alt spriker mye. Artikkene har fotoelementer som går igjen, som bruk av dronevideo.

Søketjenesten er integrert i artikkene, med værdata fra kommunen reportasjen er laget fra. Denne løsningen er også laget for å være enkel å ta i bruk av andre journalister når de senere lager reportasjer eller nyhetsartikler om klima i Norge.

Fremover er det et mål å videreutvikle spesialen med nye elementer eller bygge nye søketjenester innen klima. Vi kommer til å publisere flere artikler i samme reportasjeserie. Dette er bare begynnelsen.

8. Resultat

Interessen for klimaspecialen og reportasjene har vært svært gode. Hovedspesialen har passert 1.100.000 lesninger. Legger man til kommunesidene fra søketjenesten, har denne delen av prosjektet rundt 2.000.000 ganger. Reportasjene er lest tilsvarende godt.

Det har vært svært gledelig å se at søketjenesten deles aktivt av kommuner til sine innbyggere, og blir brukt av sivilbefolkningen som diskuterer klimaendringer i some. Spesialen har gitt klimadekningen i VG en ny dimensjon. Det er, så vidt vi vet, første gang en norsk redaksjon har behandlet et så stort datasett om vær og klima selv, og presentert en så detaljert og grundig oversikt over utviklingen. I sum viser de hvor gjennomgripende klimaendringene er, som vi nå står midt i.