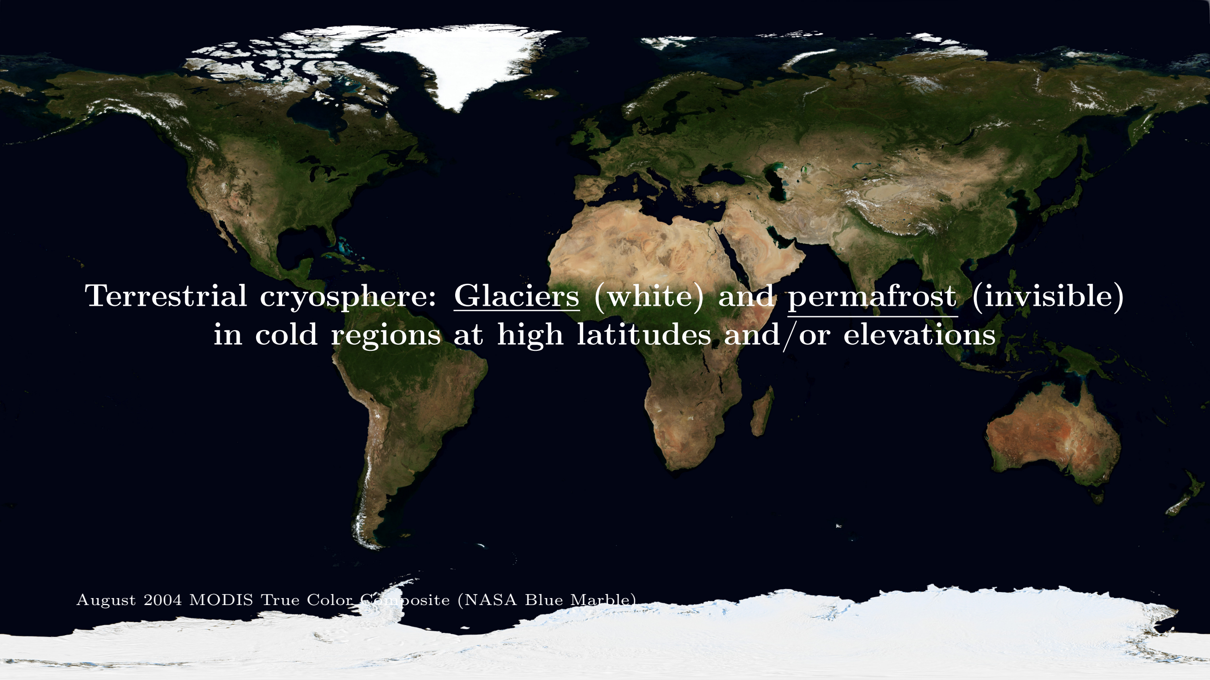
Ensemble-based reanalysis of the terrestrial cryosphere

Kristoffer Aalstad[1], Researcher, Dept. Geosci., Uni. Oslo, kristaal@geo.uio.no
R. Hock[1], M. Mazzolini[1], N. Pirk[1], D. Treichler[1], S. Westermann[1], Y. A. Yılmaz[1] J. Fiddes[2] (SLF), E. Alonso-González[3] (IPE)

UNIVERSITY OF OSLO

SLF

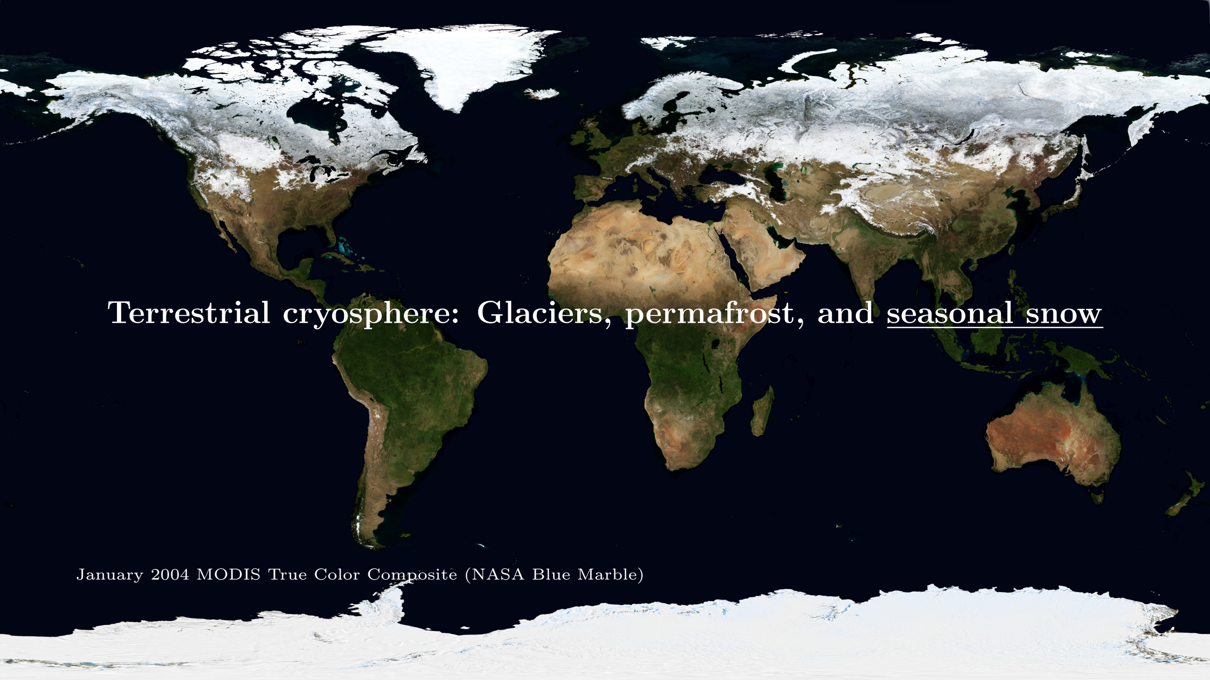
IPE CSIC



Terrestrial cryosphere: Glaciers (white) and permafrost (invisible) in cold regions at high latitudes and/or elevations

August 2004 MODIS True Color Composite (NASA Blue Marble)

Terrestrial cryosphere: Glaciers, permafrost, and seasonal snow



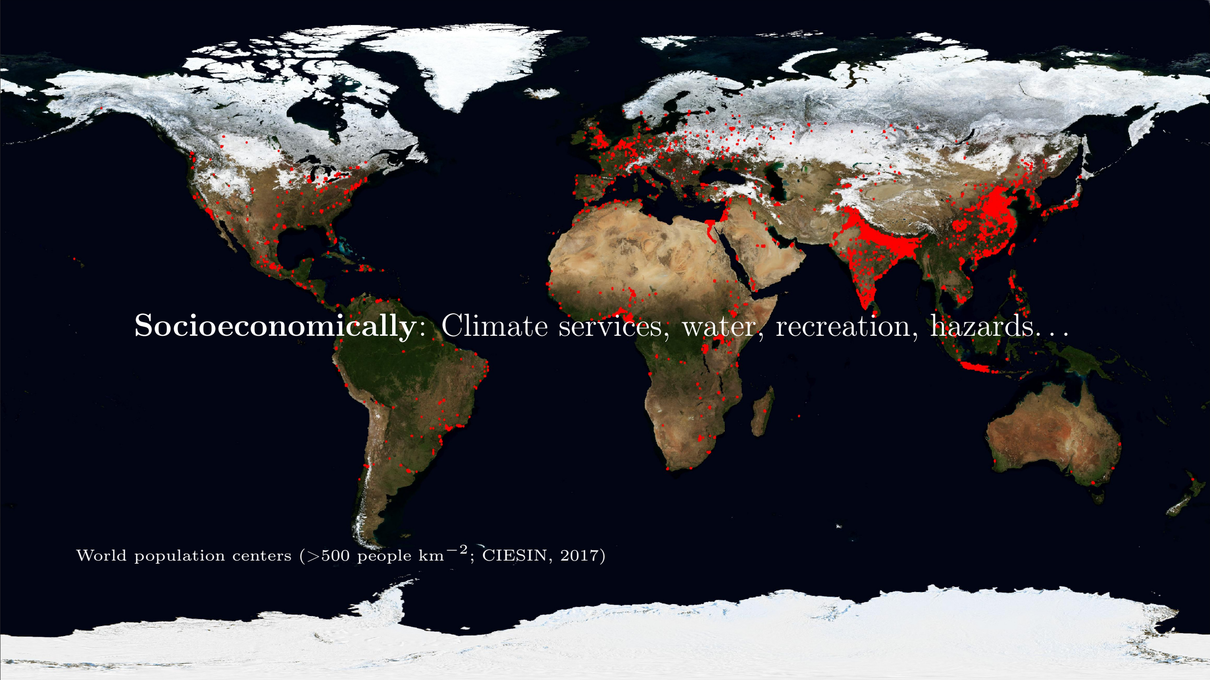
January 2004 MODIS True Color Composite (NASA Blue Marble)

Why we are interested in the cryosphere?

January 2004 MODIS True Color Composite (NASA Blue Marble)

Geoscientifically: Energy (albedo, latent heat, insulation), water (storage, runoff, sea level), and carbon (permafrost, vegetation) cycle

January 2004 MODIS True Color Composite (NASA Blue Marble)



World population centers (>500 people km^{-2}; CIESIN, 2017)

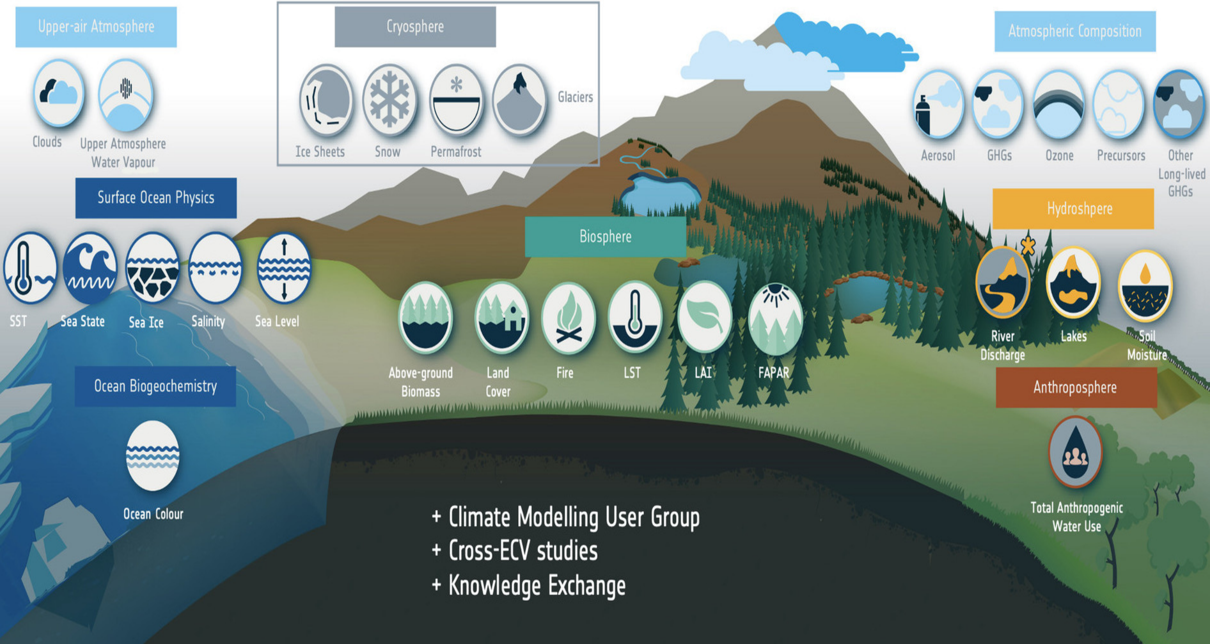


Figure: Snow, permafrost, and glaciers are **essential climate variables** (ESA Climate Change Initiative)

Working towards the **long-term goal** of generating hillslope-scale (100 m), daily, and multi-decadal estimates of the state of the global cryosphere

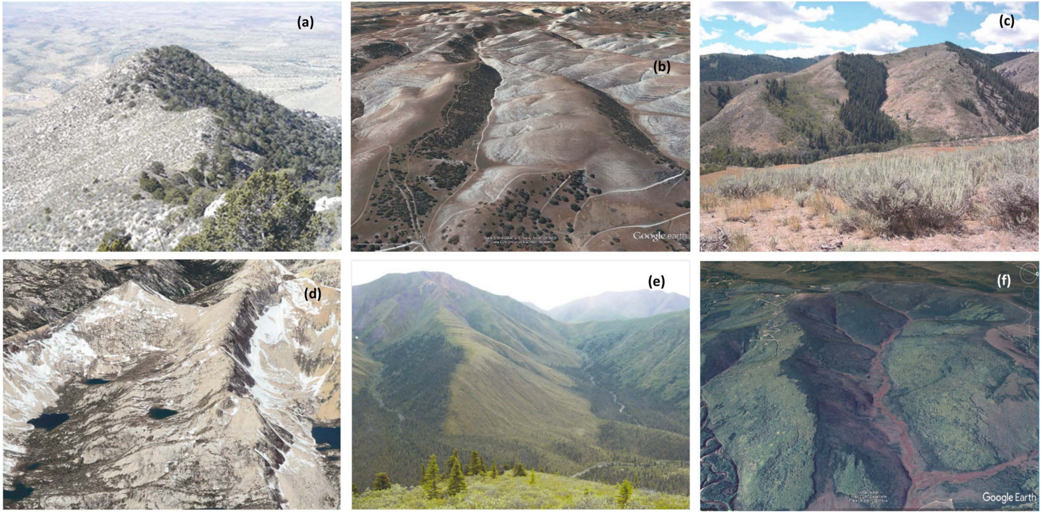
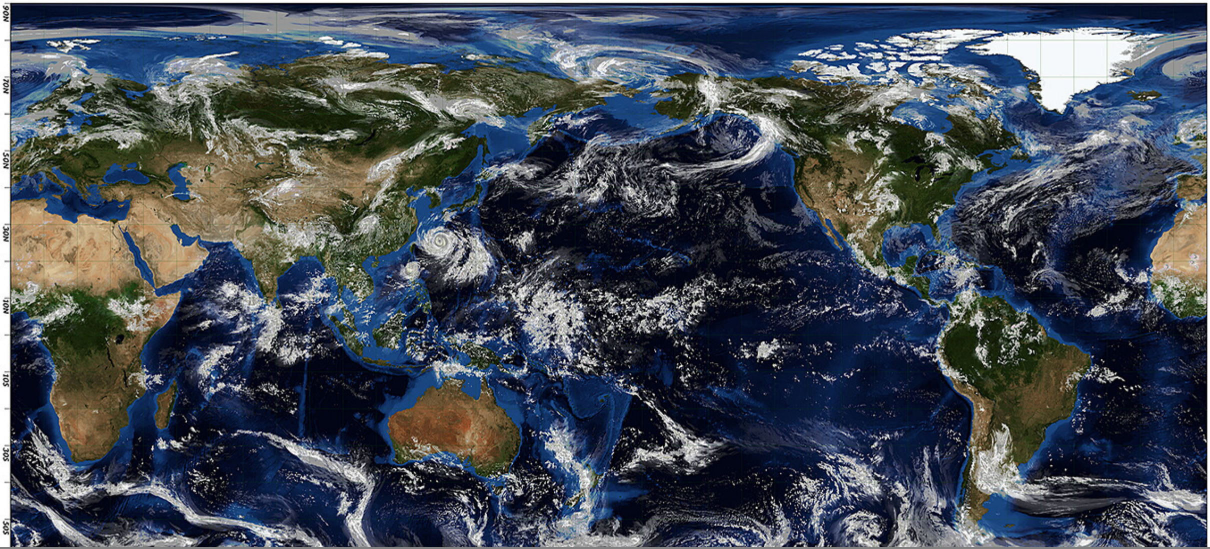


Figure: Example images showing the importance of the hillslope-scale in organizing the storage and fluxes of water, energy, and carbon, across the terrain. Adapted from Fan et al. (2019)



Earth system modeling, especially NWP (see Miyamoto et al., 2013), has come a long way, but multi-decadal subkilomter simulations are beyond the state-of-the-art. Even if we could afford this, **chaos needs to be tamed by observations**.

A burgeoning zoo of **emerging Earth observations. . .**

Earth observation **data can't speak for itself**

- Usually **indirect**.
- **Gaps** in space and time.
- Limited temporal **coverage**.
- Measurement **uncertainty**.
- Mismatch between measurement **scale** and process scale of interest.
- Often these data are already some form of **retrieval**.
- The future and most of the past is currently **unobservable**.

We always use **models** to represent and predict regularities in nature.
A **separation of data and models is misleading** since measurement always involves some form of modeling (Parker, 2017).

A unifying perspective: Bayesian **data assimilation** (inversion)

$$\mathbf{y} = \mathcal{G}\left(\mathbf{x}\right) + \boldsymbol{\varepsilon}\,, \tag{1}$$

- $\mathbf{y}$: Observation vector
- $\mathbf{x}$: Model state and/or parameter vector
- $\mathcal{G}$: **Data generating model** of state dynamics and observables.
- $\boldsymbol{\varepsilon}$: Observation error, also known as noise.

The inverse problem (solving for $\mathbf{x}$) is always ill-posed, so we adopt a probabilistic (Bayesian) perspective and infer the posterior distribution

$$\underbrace{p(\mathbf{x}|\mathbf{y})}_{\text{Posterior}} \propto \overbrace{p(\mathbf{y}|\mathbf{x})}^{\text{Likelihood}} \underbrace{p(\mathbf{x})}_{\text{Prior}} \tag{2}$$

by updating the prior (what we knew before conditioning on data) with the likelihood (what the data tell us).

Algorithms for Bayesian DA (MacKay, 2003; Wikle et al., 2007; Murphy, 2023)

(A loose classification)

- Grid approximation†.
- Quadratic approximation (Laplace's method)†.
- Variational methods (4D-Var)†.
- Importance sampling (particle filters/smoothers, GLUE).*
- Ensemble-based methods (EnKF and smoothers)*.
- Approximate Bayesian Computation (ABC)*.
- Markov chain Monte Carlo (MCMC)*.
- Hybrid methods .

†: Deterministic methods.
*: Monte Carlo (ensemble) methods using (pseudo) random number generators to sample distributions.

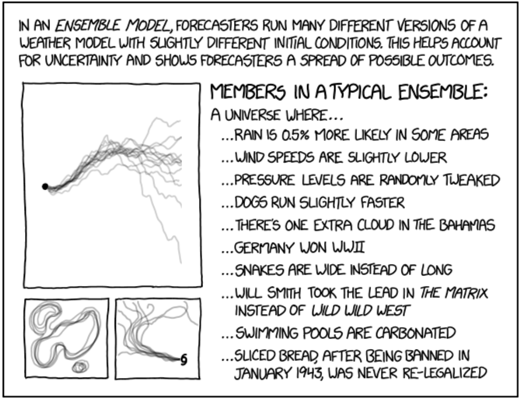


Figure: Ensemble model https://xkcd.com/1885/.

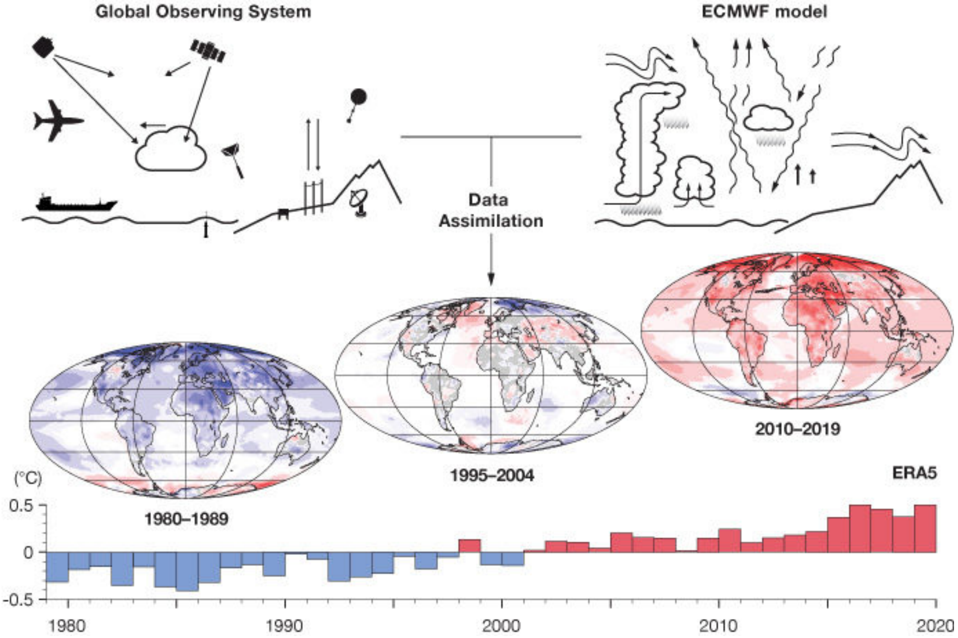
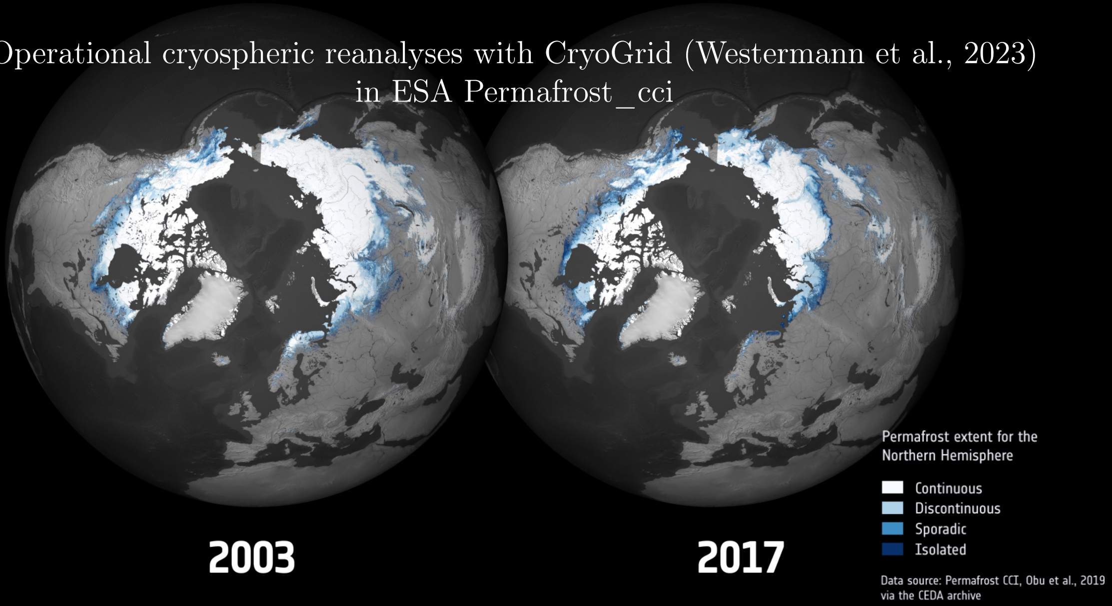


Figure: Data assimilation is used operationally for **climate reanalysis** and NWP (source: ECMWF)

Operational cryospheric reanalyses with CryoGrid (Westermann et al., 2023) in ESA Permafrost_cci



Seasonal snow reanalysis

$$\underbrace{\textbf{Satellite retrievals}}_{\text{snow-covered area (fSCA), depth}} = \mathcal{G}(\underbrace{\textbf{Snowpack state}}_{\text{SWE, fSCA, depth}}) + \varepsilon$$

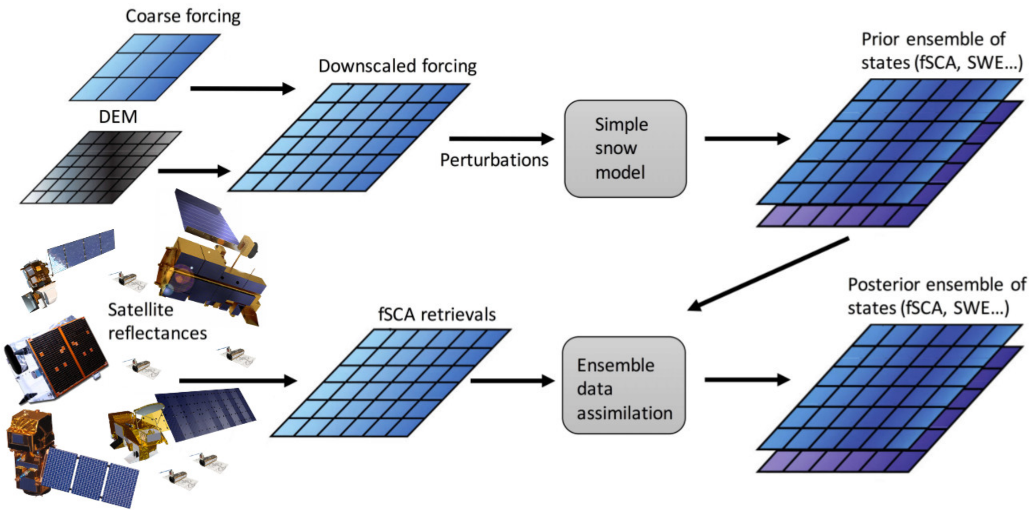
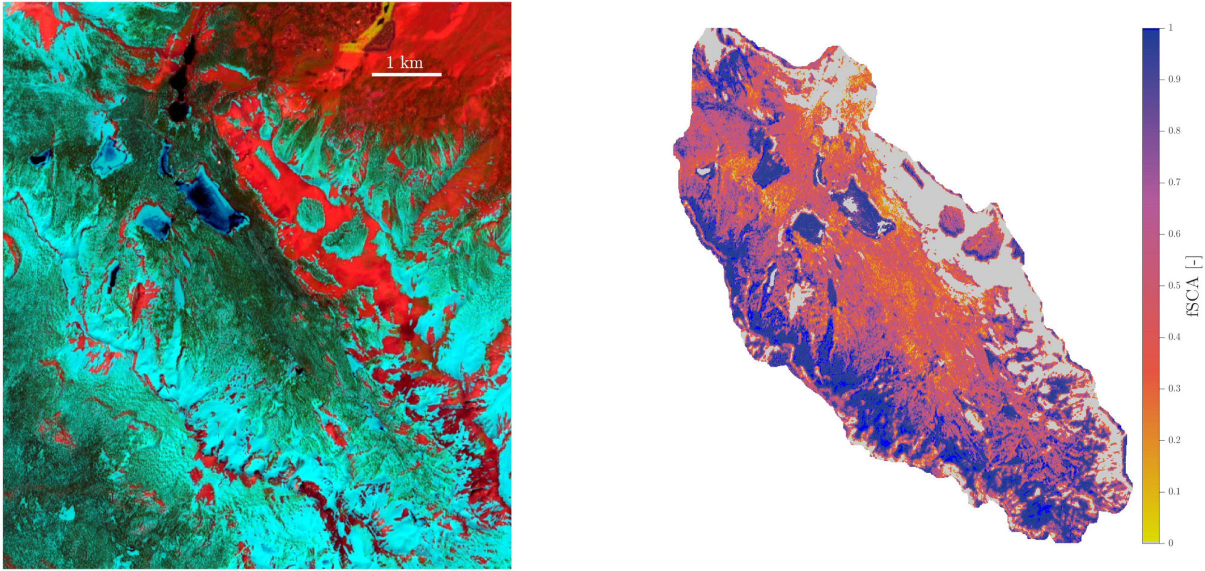


Figure: Workflow for the ensemble-based snow reanalysis method (Martinec and Rango, 1981; Kolberg and Gottschalk, 2006; Margulis et al., 2015; Aalstad et al., 2018; Alonso-González et al., 2022).

Lakes basin snow reanalysis (30 km^2, US Sierra Nevada)

- Assimilating fSCA retrievals from myriad optical satellites (Landsat 7-8, Sentinel-2, Planet cubesats)
- Simple energy & mass balance snow model (Aalstad et al., 2018) for 1 layer snowpack at 100 m res.
- Hybrid particle-ensemble smoother (Pirk et al., 2022).
- Forced by topographically-downscaled NLDAS.
- Validation with airborne snow observatory lidar.
- Comparing performance to existing regional (UCLA, Spires-ParBal, UA), national (CONUS404, SNODAS, NWM, NLDAS), and global (ERA5L) products.

S2B MSI 11-Jun-2019 18:54:09Z



RGB=(SWIR1,NIR,R)

Figure: Sentinel-2 10 m false color (left, snow is turquoise) and fSCA retrievals (right) over the Lakes basin.

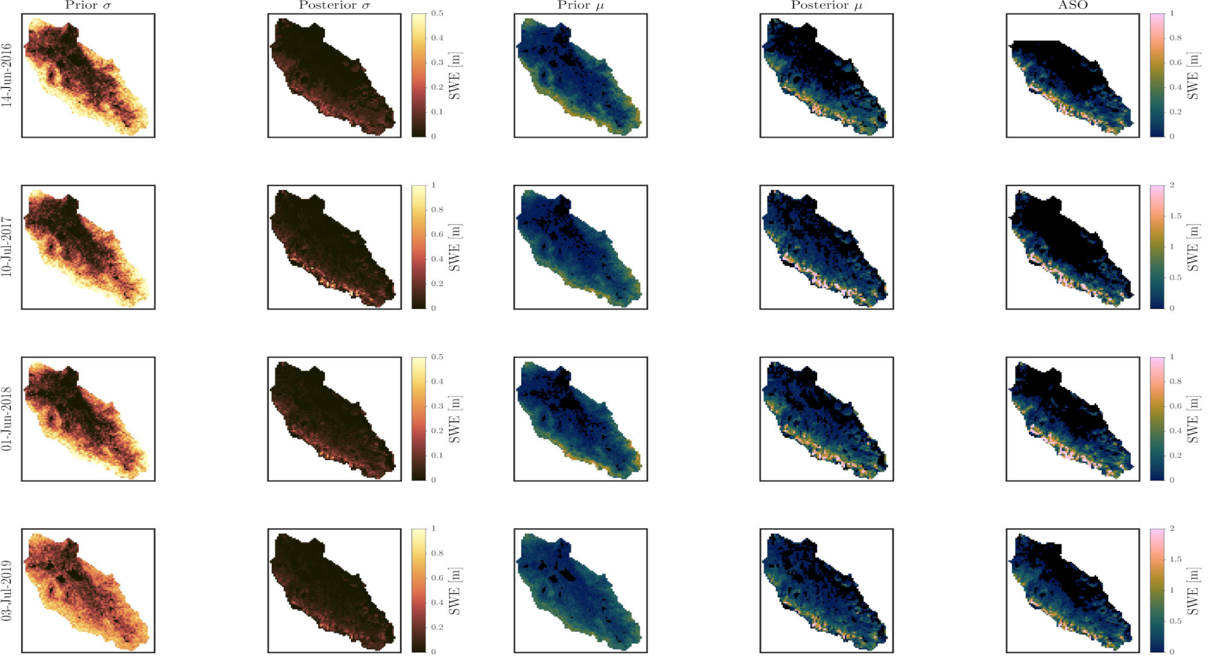


Figure: Prior & posterior spread σ (col 1&2) and mean μ (col 3&4) SWE and ASO SWE (col 5) for 4 melt seasons (rows).

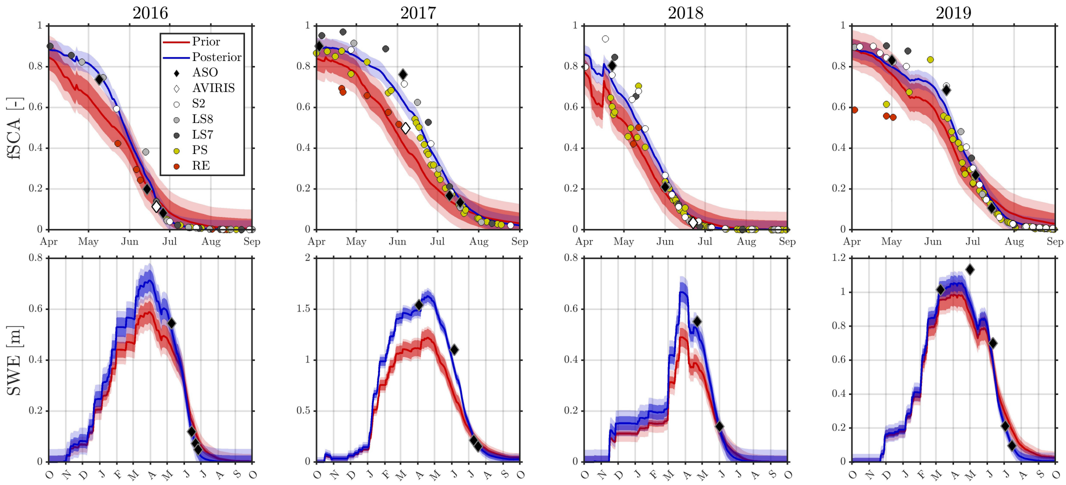


Figure: Basin-averaged time series of fractional snow-covered area (top, melt season only) and snow water equivalent (bottom) showing the prior (red) and posterior (blue) mean (line) and spread (shading, $\pm 5\sigma$) along with assimilated satellite data (dots) and independent airborne observations (triangles). Each column is for one water year (WY), i.e. 1-Oct-[WY-1] to 30-Sep-WY, in the range 2016-2019.

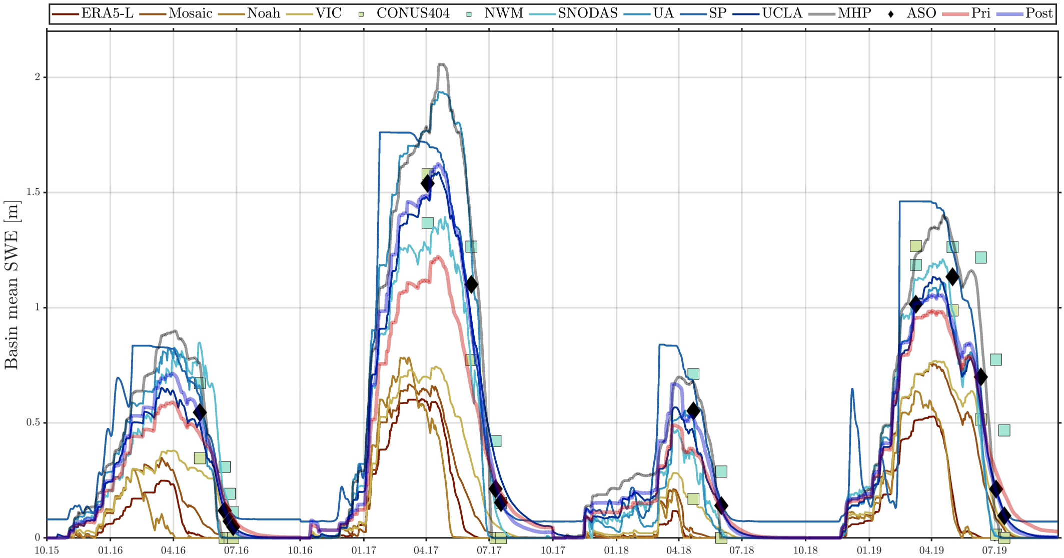


Figure: Basin-average time series of the prior (thick red) and posterior (thick blue) mean SWE for 4 water years in the range 2016-2019 as well as the corresponding estimates from several other regional, national, and global products as well as the local station (MHP, gray) and the ASO ground truth (black diamond).

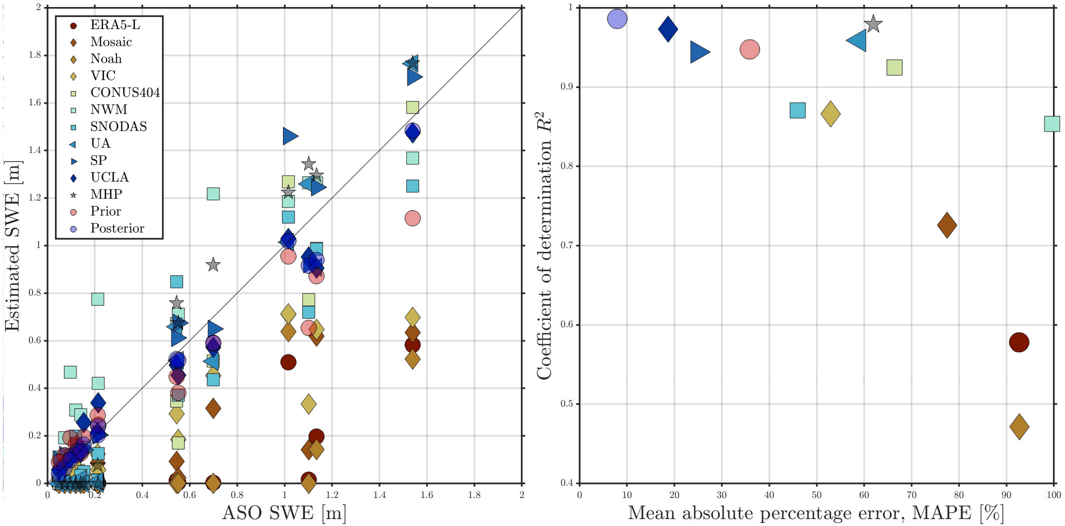


Figure: Scatter plot of all the basin-average SWE estimates from various products vs. the ASO SWE ground truth (left) and the corresponding performance in terms of MAPE (0 is best) and R^2 (1 is best) for all products (right). Naive estimates from the local MHP station (gray star) with SWE MAPE of 60% are worse than the prior. Three snow reanalysis methods (posterior, UCLA, SP) perform best.

https://doi.org/10.5194/egusphere-2024-1404
Preprint. Discussion started: 24 May 2024

Spatio-temporal snow data assimilation with the ICESat-2 laser altimeter

Marco Mazzolini[1], Kristoffer Aalstad[1], Esteban Alonso-González[2], Sebastian Westermann[1], and Désirée Treichler[1]

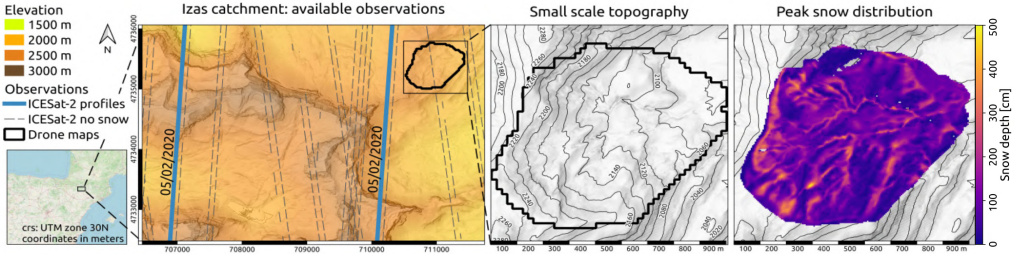


Figure: Assimilating ICESat-2 snow depth & Sentinel-2 fSCA with **MuSA** (Mazzolini et al., 2024)

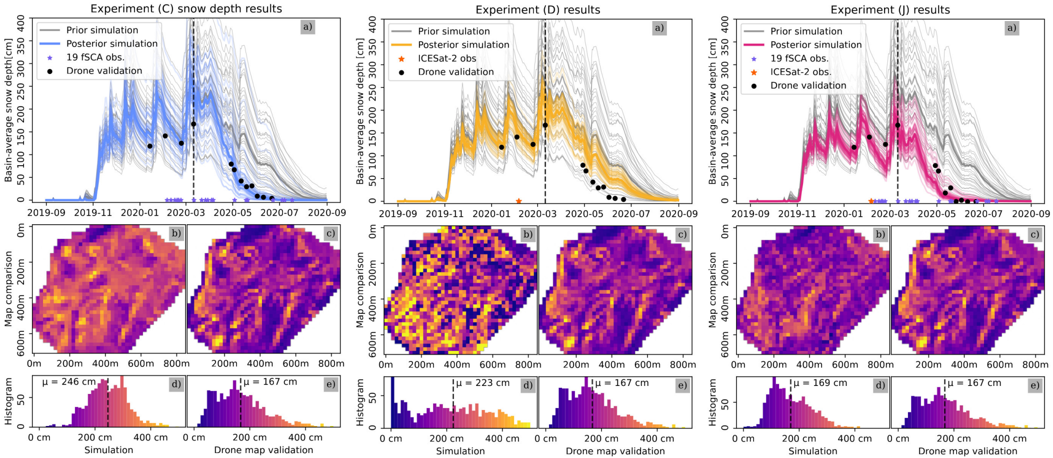


Figure: Snow depth results assimilating snow cover only (**left**) snow depth only (**middle**) and both snow depth and snow cover (**right**). For each column the panels show the: catchment-average (0.5 km^2) time series (a), peak snow depth distribution posterior median estimates (b) and independent drone observations (c), histograms of peak snow depth posterior median estimates (d) and drone observations (e).

Check out github.com/ealonsogzl/MuSA (Alonso-González et al., 2022), we are continuously developing this toolbox and welcome new contributors!

Geosci. Model Dev., 15, 9127–9155, 2022
https://doi.org/10.5194/gmd-15-9127-2022

Geoscientific Model Development
Open Access

The Multiple Snow Data Assimilation System (MuSA v1.0)

Esteban Alonso-González[1,★], Kristoffer Aalstad[2,★], Mohamed Wassim Baba[3], Jesús Revuelto[4], Juan Ignacio López-Moreno[4], Joel Fiddes[5], Richard Essery[6], and Simon Gascoin[1]

[1]Centre d'Etudes Spatiales de la Biosphère, Université de Toulouse, CNRS–CNES–IRD–INRA–UPS, Toulouse, France
[2]Department of Geosciences, University of Oslo, Oslo, Norway
[3]Center for Remote Sensing Application (CRSA), Mohammed VI Polytechnic University (UM6P), Ben Guerir, Morocco
[4]Instituto Pirenaico de Ecología, CSIC, Zaragoza, Spain
[5]WSL Institute for Snow and Avalanche Research SLF, Davos, Switzerland
[6]School of GeoSciences, University of Edinburgh, Edinburgh, UK
★These authors contributed equally to this work.

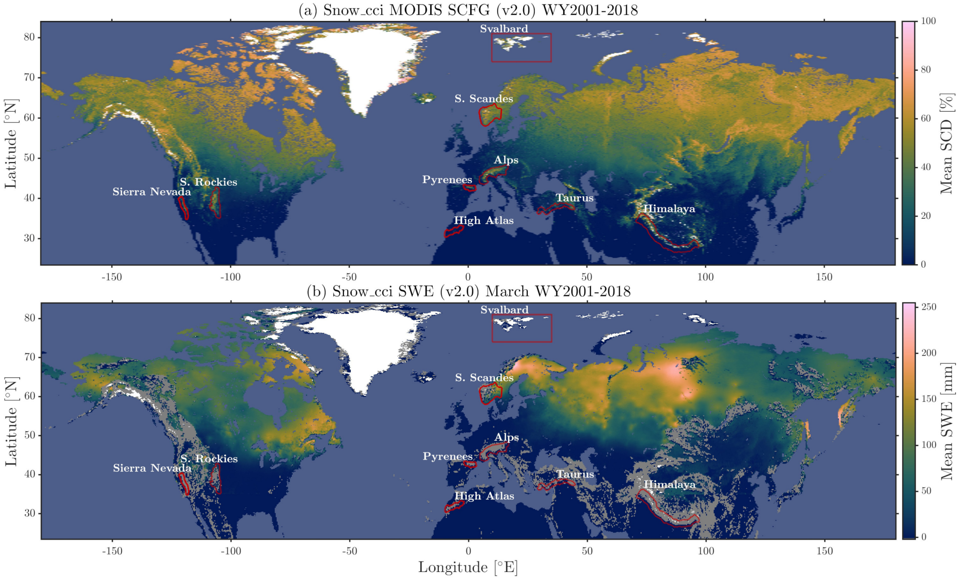


Figure: The ESA PATCHES project seeks to (1) fill gaps in montane SWE (lower) by assimilating snow cover (top) in a snow reanalysis and (2) Extend the kilometer-scale snow cover record back to the 1980s by sharpening AVHRR data using MODIS. More info: mn.uio.no/geo/english/research/projects/patches/

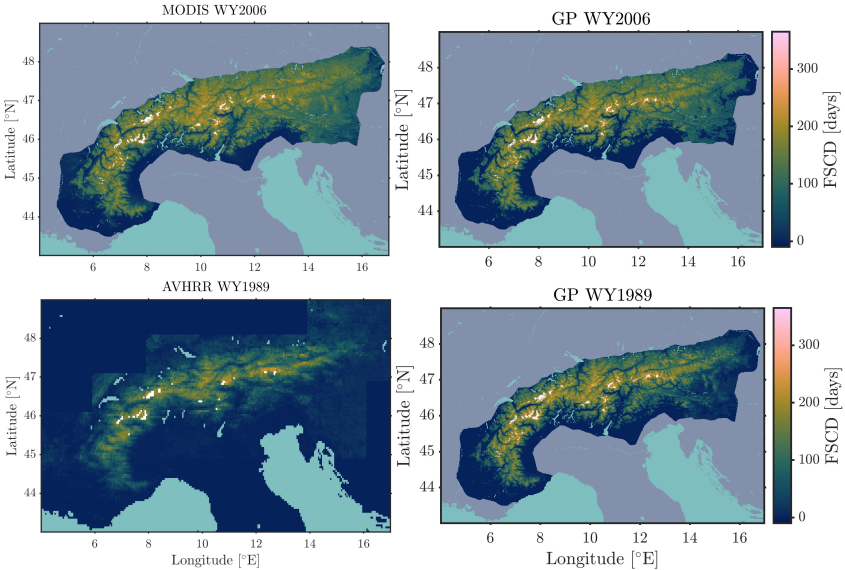


Figure: Examples of snow cover duration for the Alps from MODIS at 0.01°(top left, WY2006) and AVHRR (bottom left, WY1989). Corresponding Gaussian process 0.01° fusion of AVHRR & MODIS in the right column. Only pixels falling within the Alps are shown for the MODIS and GP sharpened imagery.

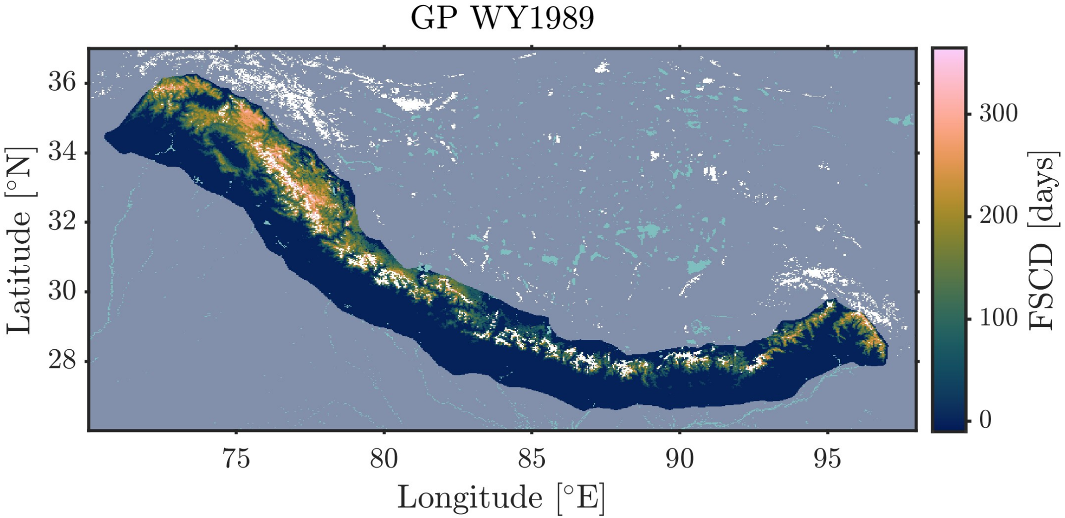


Figure: Ongoing work testing a novel Gaussian process sharpening method over 8 large mountain ranges and an Arctic archipelago, above is an example of sharpened 0.01° snow cover duration across the Himalaya for water year 1989 (prior to the MODIS era). Currently no truly global kilometer scale snow cover data exists prior to water year 2000 when Landsat obtained global coverage and MODIS was launched. In the current first stage of PATCHES we are trying to help fill this gap.

Inspired by **yukigata** (Japanese: 'snow picture')

Repeated patterns in annual snow cover dynamics are key to snow cover sharpening and reanalysis methods.

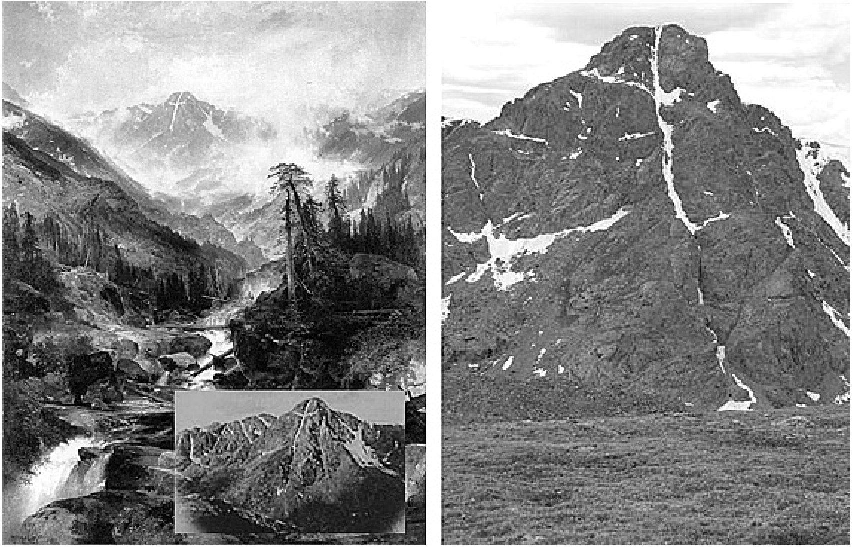


Figure: Mt. of the Holy Cross (CO, USA) in 1873 (left) and 2004 (right) from Sturm and Wagner (2010).

Glacier reanalysis

$$\underbrace{\textbf{Glacier observations}}_{\text{balances, snowlines, albedo}\ldots} = \mathcal{G}(\underbrace{\textbf{Glacier state}}_{\text{mass,volume}\ldots}) + \varepsilon$$

Past and Future High-resolution Global Glacier Mass Changes (**GLACMASS**): mn.uio.no/geo/english/research/projects/glacmass/

- ERC Advanced Grant (Oct 2023 - Oct 2028). PI: Regine Hock (Uni. Oslo, UAF)
- Goal: Reconstruct and project the mass balance of all >200 000 glaciers on Earth.
- Method: Synthesizing emerging glacier observations by combining DA and machine learning with the Python Glacier Evolution Model (PyGEM, Rounce et al., 2023).

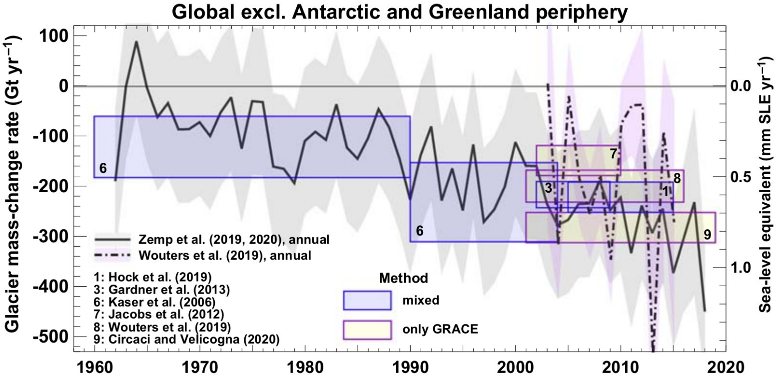
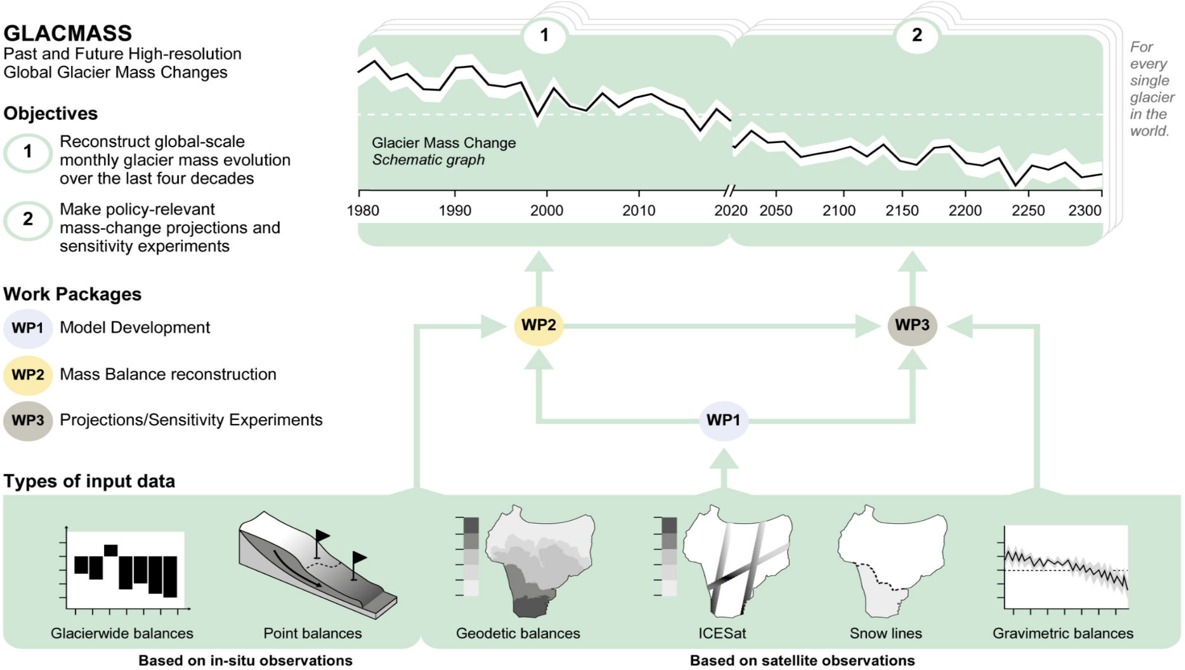


Figure: Global glacier mass-change estimates from Hock and Huss (2021).


GLACMASS
Past and Future High-resolution
Global Glacier Mass Changes
Objectives
1 Reconstruct global-scale monthly glacier mass evolution over the last four decades
2 Make policy-relevant mass-change projections and sensitivity experiments
Work Packages
WP1 Model Development
WP2 Mass Balance reconstruction
WP3 Projections/Sensitivity Experiments
Types of input data
1
2
Glacier Mass Change
Schematic graph
1980
1990
2000
2010
2020
2050
2100
2150
2200
2250
2300
For every single glacier in the world.
WP2
WP3
WP1
Glacierwide balances
Point balances
Geodetic balances
ICESat
Snow lines
Gravimetric balances
Based on in-situ observations
Based on satellite observations

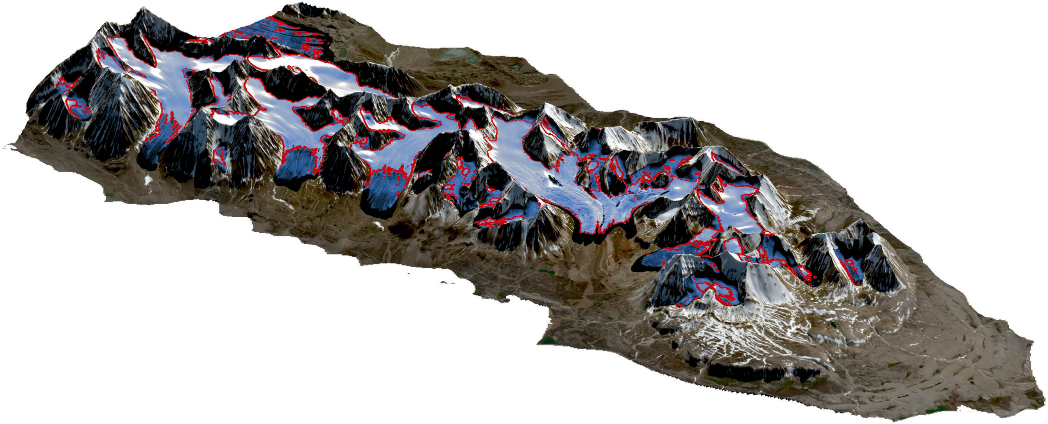


Figure: Combining DA and machine learning, in GLACMASS we will explore the value of emerging Earth observations for constraining glacier evolution models. This figure shows glacier albedo (dark blue: low, light blue: high) over the Brøgger peninsula retrieved using a novel Bayesian spectral unmixing approach on Sentinel-2 data, and the corresponding snow lines (red) estimated from these albedo retrievals.

Proof-of-concept glacier reanalysis

- Assimilating in-situ mass balance in PyGEM-lite.
- Hybrid particle-ensemble smoother.
- Forced by topographically downscaled ERA5 data.
- Rembesdalskåka glacier, Hardangerjøkulen, Norway.

Photo: H. Elvehøy, (NVE)

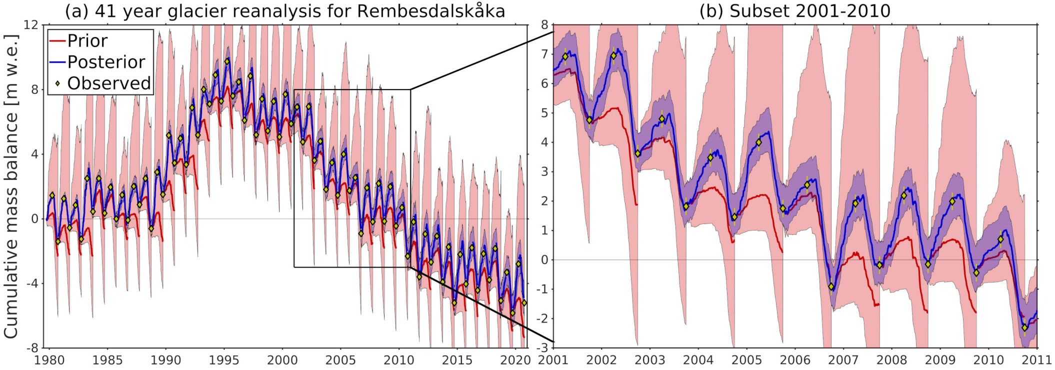


Figure: Glacier mass balance reanalysis for Rembesdalskåka in Norway for (a) the entire 41-year period, and (b) a subset for 2001-2010, showing the prior (red) and the posterior (blue) after having assimilated the mass balance observations (yellow dots). Solid lines indicate the ensemble median while the shading shows the 90$^{\text{th}}$ percentile range of the ensemble. The parameters in each water year are updated independently (**no pooling**), and the prior states are initialized using the last step of the posterior from the previous year. We updated 3 uncertain parameters: degree-day factor, a snowfall correction, and air temperature bias. The observations were provided by the Norwegian Water Resources and Energy Directorate (NVE).

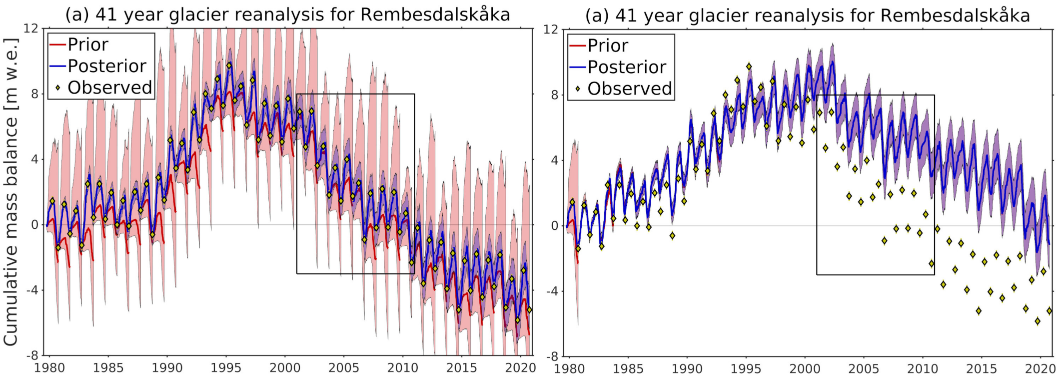


Figure: Panel (a) on the left is the same as the last figure with no pooling, but panel (b) is with **dynamical pooling** (i.e. *filtering*) where the parameter posterior of the current water year becomes the parameter prior for the next water year. In this case, dynamical pooling does not work very well, and neither does batch pooling (not shown) where static parameters are updated using all observations.

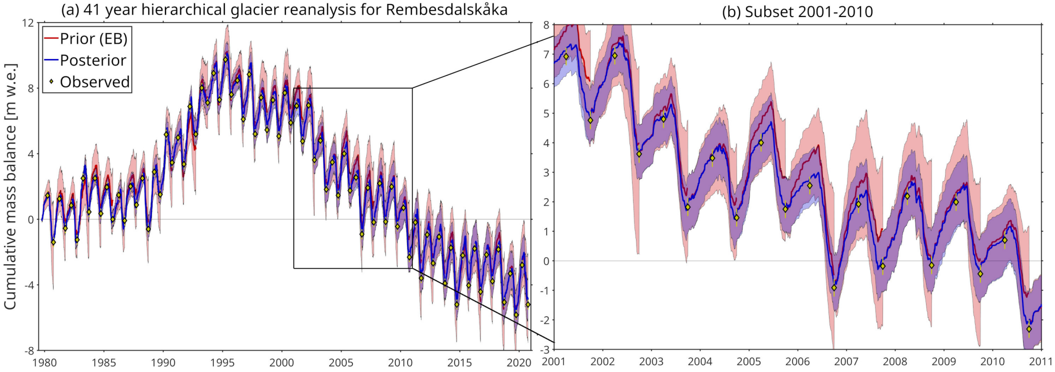


Figure: Between the two extremes of no pooling and complete pooling, we have **partial pooling** using hierarchical Bayesian inference where we define a so-called hyperprior and infer hyperparameters (parameters controlling parameters). Our new hierarchical smoother reanalysis method (**work in progress**) appears to work much better and is less sensitive to temporal (and likely spatial) data gaps.

Some of our recent uses of data assimilation (DA) in GEWEX problems

- Satellite (MODIS & Sentinel-2) snow cover DA (Aalstad et al., 2018).
- High res MODIS snow cover DA with clustering (Fiddes et al., 2019).
- MODIS snow cover DA in FSM-ICAR (Alonso-González et al., 2021).
- **Multiple snow DA system** (MuSA) (Alonso-González et al., 2022).
- High res spatio-temporal snow DA (Alonso-González et al., 2023).
- Spatio-temporal ICESat-2 snow depth DA (Mazzolini et al., 2024).
- Snow reanalysis using DA and deep learning (Guidicelli et al., 2023).
- Drone DA for flux inversion in large eddy simulations (Pirk et al., 2022).
- Guiding drones by flux information gain (van Hove et al., 2023, 2024).
- Inferring carbon fluxes with Bayesian deep learning (Pirk et al., 2024).

Funding: PATCHES (ESA fellowship), Permafrost_cci (ESA), GLACMASS (ERC), Spot-On (RCN), SNOWDEPTH (RCN)

Funded by the European Union

European Research Council

Funded by The Research Council of Norway

Thanks, questions?

Photo: Mount Aka, Daisetsu mountains, Hokkaido (29.06.2017, by K. Aalstad)

Aalstad, K. et al. (2018). doi:10.5194/tc-12-247-2018. *TC*.

Alonso-González, E. et al. (2021). doi:10.5194/hess-25-4455-2021. *HESS*.

Alonso-González, E. et al. (2022). doi10.5194/gmd-15-9127-2022. *GMD*.

Alonso-González, E. et al. (2023). doi:10.5194/hess-27-4637-2023. *HESS*.

Fan, G. et al. (2019). doi:10.1029/2018WR023903. *WRR*.

Fiddes, J. et al. (2019). doi:10.5194/hess-23-4717-2019. *HESS*.

Guidicelli, M. et al. (2023). doi:10.2139/ssrn.4489553. *JH*.

Hock, R. and Huss, M. (2021). doi:10.1016/B978-0-12-821575-3.00009-8. *Elsevier*.

Kolberg, S. A. and Gottschalk, L. (2006). doi:10.1002/hyp.6060. *HP*.

MacKay, D. J. C. (2003). url:inference.org.uk/itprnn/book.pdf. *CU*.

Margulis, S. A. et al. (2015). doi:10.1175/JHM-D-14-0177.1. *JHM*.

Martinec, J. and Rango, A. (1981). doi:10.1029/WR017i005p01480. *WRR*.

Mazzolini, M. et al. (2024). doi:10.5194/egusphere-2024-1404. *TCD*.

Miyamoto, Y. et al. (2013). doi:10.1002/grl.50944. *GRL*.

Murphy, K. P. (2023). url:probml.github.io/book2. *MIT*.

Parker, W. S. (2017). doi:10.1093/bjps/axv037. *BJPS*.

Pirk, N. et al. (2022). doi:10.5194/amt-15-7293-2022. *AMT*.

Pirk, N. et al. (2024). doi:10.1029/2024GL109283. *GRL*.

Rounce, D. et al. (2023). doi:10.1126/science.abo1324. *Science*.

Sturm, M. and Wagner, A. M. (2010). doi:10.1029/2010WR009434. *WRR*.

van Hove, A. et al. (2023). doi:10.5617/nmi.9897. *NMI*.

van Hove, A. et al. (2024). url:proceedings.mlr.press/v233/hove24a. *PMLR*.

Westermann, S. et al. (2023). doi:10.5194/gmd-16-2607-2023. *GMD*.

Wikle, C. K. et al. (2007). doi:10.1016/j.physd.2006.09.017. *PD*.