SMHI

EUMETNET DATA MANAGEMENT WORKSHOP 2025, OSLO

GAP FILLING OF SWEDISH SNOW DEPTH OBSERVATIONS FOR CLIMATE NORMAL VALUE CALCULATIONS

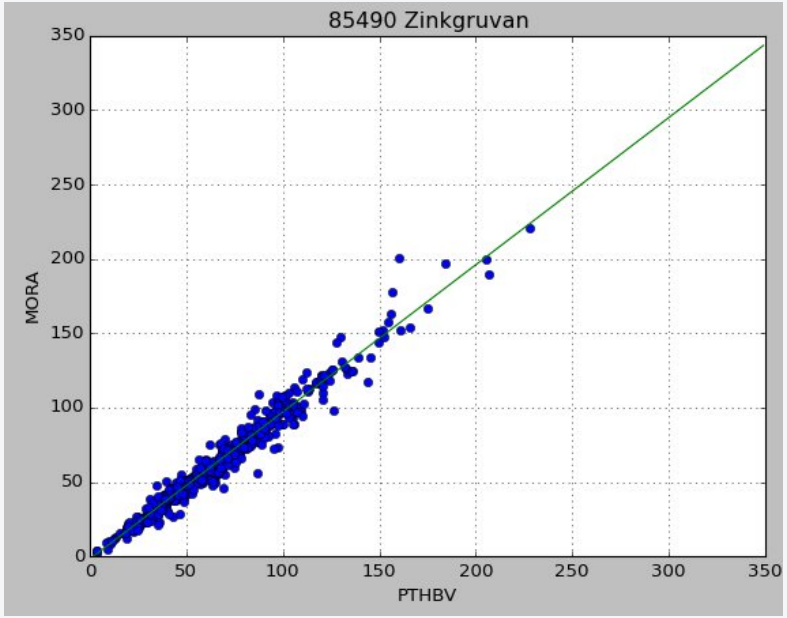
Background

Snow depth normal value

- Before 2021, new climate normal values would be calculated
- Among them snow depth normal values, such as:
  - Maximum snow depth (annual and monthly)
  - Number of days with snow depth
    - ≥ 1 cm (snow cover)
    - ≥ 10 cm
    - ≥ 50 cm
  - First and last day with snow cover, length of snow season (difference between the two)
  - Length of longest period with snow cover
  - Snow depth the 15th of each month
  - White Christmas (snow depth ≥ 1 cm the 25 December)

Other normal value calculations

- Normal value calculations require complete data series, in this case between 1961–2020, covering the reference normal period 1961–1990 and the current standard normal period 1991–2020.
- Gaps need to be filled and short series need to be extended.
- Temperature and precipitation series were filled and extended by a linear or cubic regression model between observations (MORA) and the gridded dataset ptHBV.



Issues with snow data

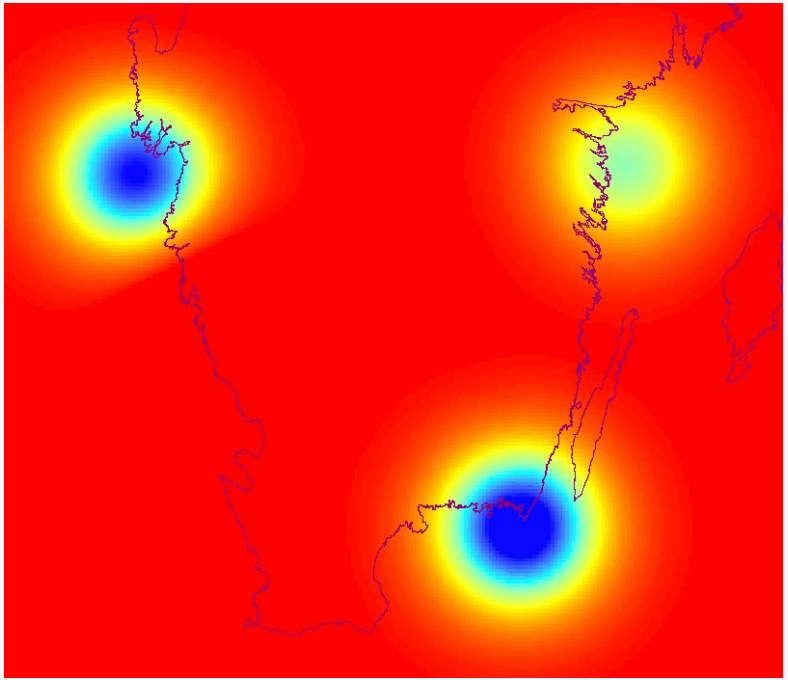
SMHI

Choice of gridded product

- ptHBV does not include snow
- SMHIGridclim (based on UERRA) covers 1961–2018 (currently), due to discontinuation of UERRA
- CERRA, CERRALand is available from 1985
- ERA5, coarse resolution
- S-HYPE, based on drainage basin

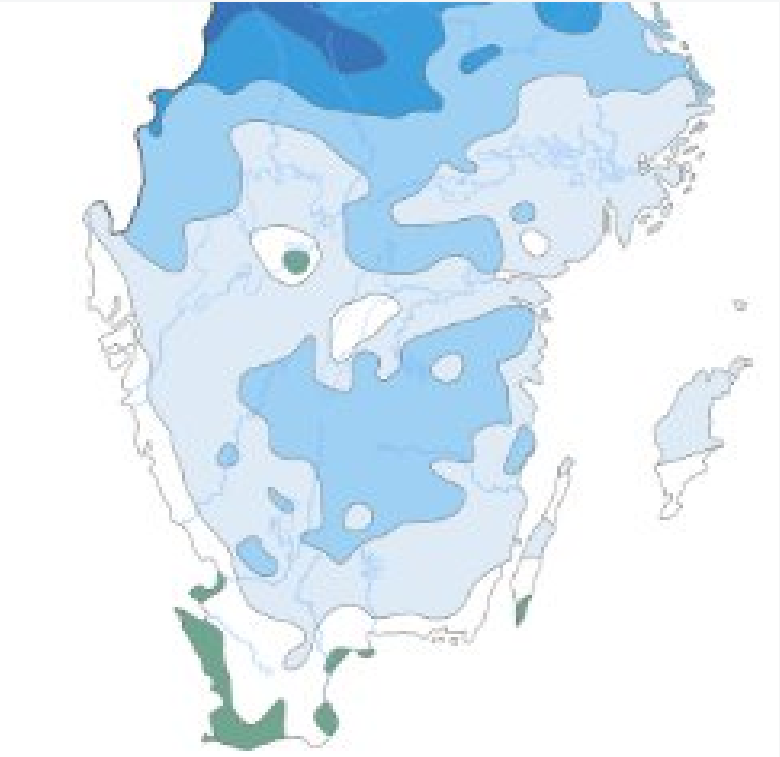
State of snow depth in gridded products

- Gridded products occasionally have quality issues in their snow depth
- There is not necessarily a linear relation between gridded product and observations



Observations

- Numerous missing data
  - Commonly, the observers has not reported “obvious” lack of snow
- Observations sometimes of questionable quality
- A station might not be representable for a larger area

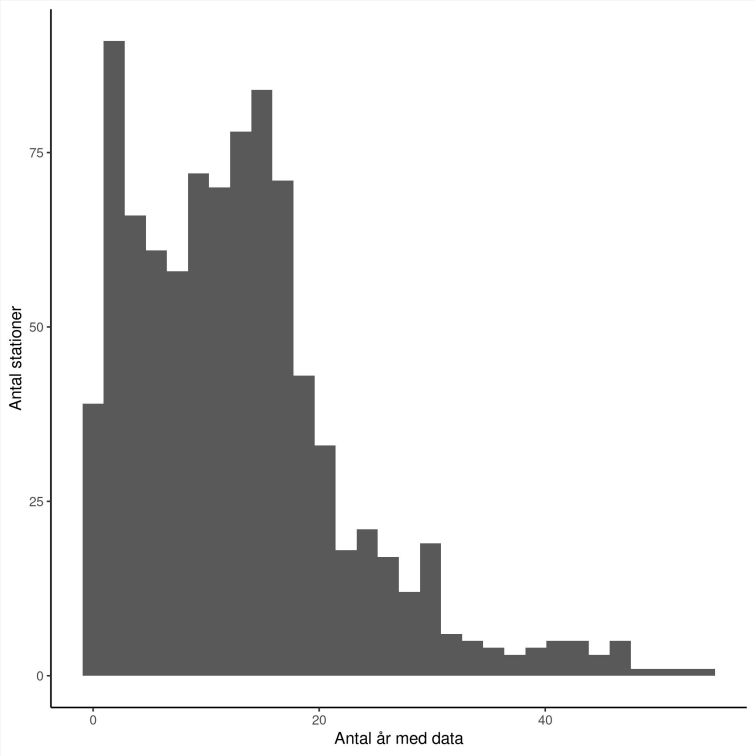


Number of days with > 10 cm,
average 1961–1990

Data description

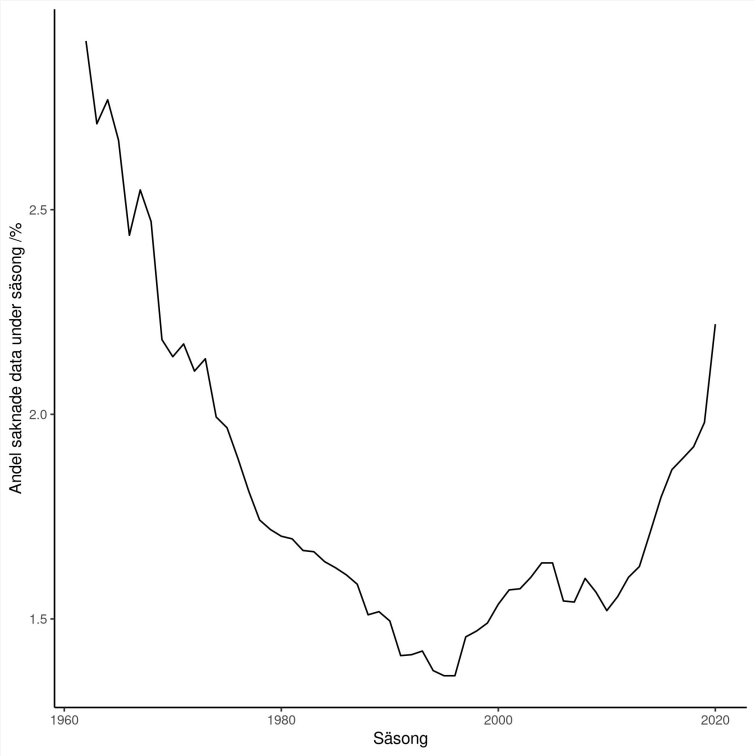
Length of series, daily snow depth

- 897 stations (848 with at least 400 days)
- 1961–2020 (60 years)



Missing data

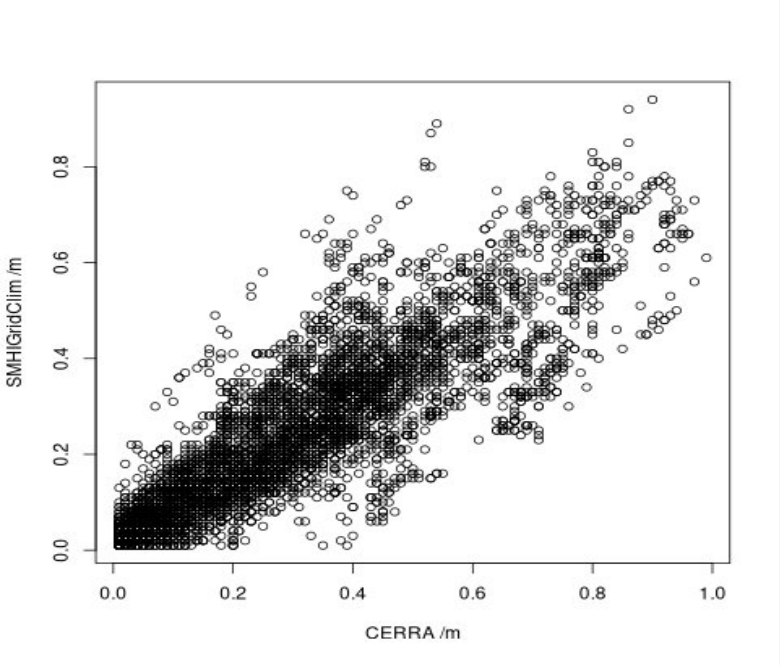
- Missing values:
  - 1961–1990: 84 %
  - 1991–2020: 73 %
  - Most series does not cover the entire period
  - Data commonly missing off-season
- During entire period, during snow season: 200 000 gaps
  - 25 %: 1–2 days
  - 50 %: up to 7 days
  - 20 %: 14–30 days
  - 5 %: > 30 days



Approaches

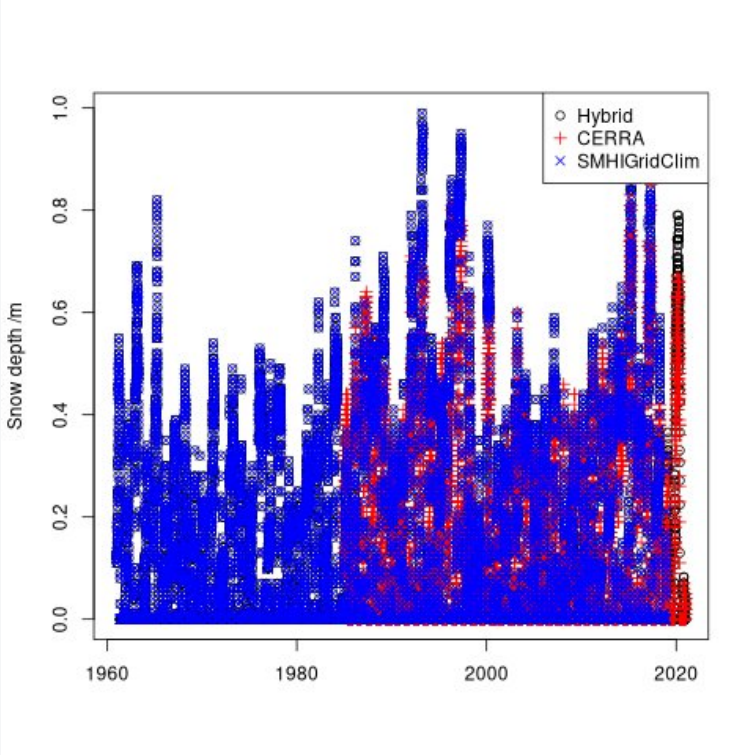
Gridded dataset
Hybrid: SMHIGridClim-CERRA

- SMHIGridClim extented to 2020 by a regression model on CERRA, based on the overlapping period 1985–2018
- Consistency: SMHIGridClim will cover entire period
- Potential homogeneity problem



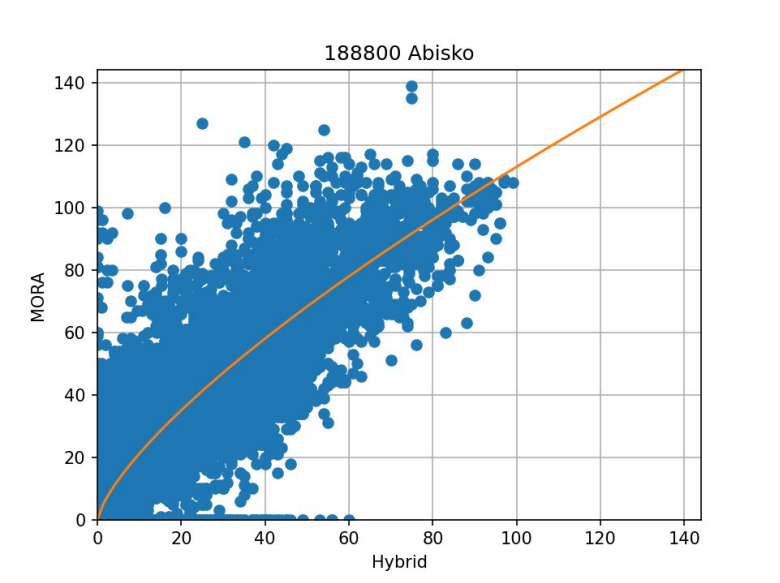
Hybrid: SMHIGridClim-CERRA

- Example time series Abisko
- Hybrid data takes the values from SMHIGridClim 1960–2018
- Hybrid data takes the values from a regression model based on CERRA data 2019–2020



Power regression: Obs-gridded data

- $y = x \cdot k^p$
- Dependent variable *x* is from a time line from gridded data
- Independent variable *y* is observations (MORA)
- Values of *k* and *p* are chosen by a optimizing algorithm where the residual is minimized (Nelder-Mead)

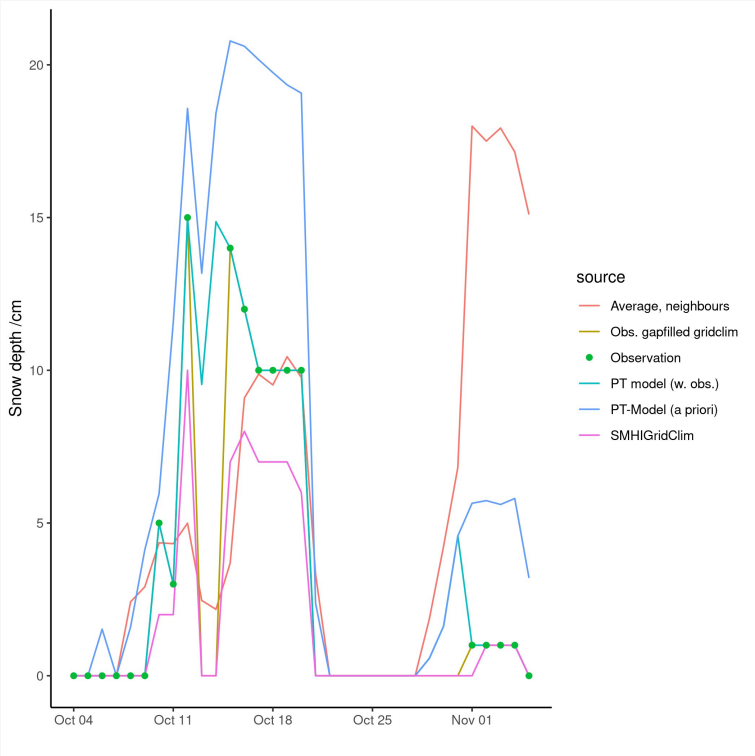


PT-snow model

- Input vectors:
  - air temperature (from ptHBV)
  - precipitation (from ptHBV)
  - initial snow storage (from observation or SMHIGridClim)
- Parameters tuned to observations per station (air temperature from ptHBV might be biased)
  - threshold temperature for snow/rain transition
  - width of snow range temperature
  - threshold temperature for snow melt
  - degree day factor for snow melt
  - liquid water holding capacity of snow
  - potential sublimation/evaporation of snow
- Two time lines constructed per station:
  - Gap-fill, snow model run only in gaps in observational data
  - Snow model run over the entire time period without snow depth observations (apart from t_0)

Use of the different data sets

- The gridded product SMHIGridClim (yellow line) is the basis regression model which is used for
- filling the gaps (green line) between
- observations (purple dots).
- The “a priori” PT-snow model (blue line) is used in quality check, where the
- PT-model with observations (red line) replaces the neglected values



Surface code

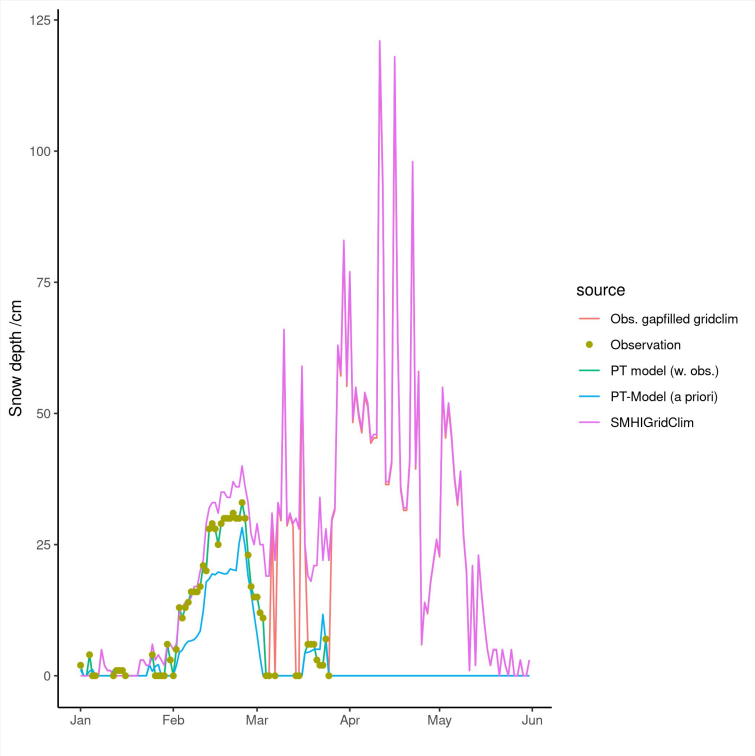
Observations of surface code is used as a quality check. If

- The surface code implies snow and the snow depth < 1 or
- The surface implies no snow and the snow depth ≥ 1

The observation is removed

Quality control

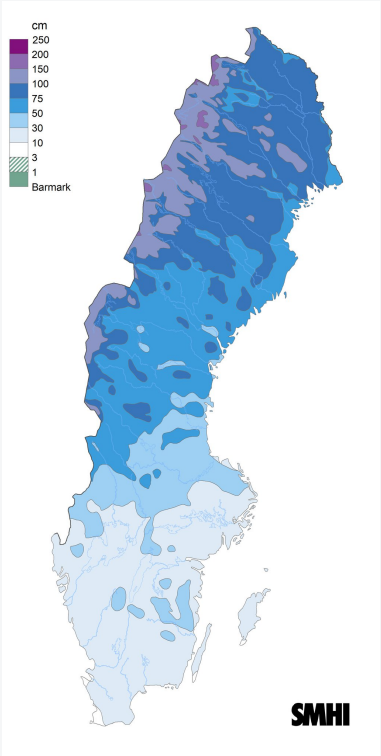
- Segments where the gapfilling data gridded and PT-model are recognised
  - If there is unrealistic sudden growth or decline in the gridded data in a segment, the segment is replaced by the gapfilling PT-model
- Periods of snow in hybrid data, where there is no observations and no snow in the PT-model data is removed

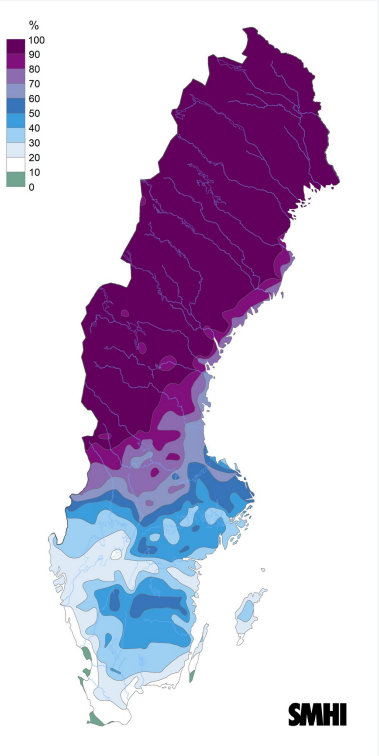


Results and Outlook

Final products

- Normal values (1961–1990, 1971–2000, 1981–2010, and 1991–2020)
  - Annual and monthly maximum snow depth
  - Percentage of years with white Christmas
- Difference between 1961–1990 and 1991–2020





SMHI

Remaining work

- Many of the normal value parameters remain to be published
- Shift to another reanalysis product
  - New SMHIGridClim (under construction)
    - Constructed with gridpp
    - Observation QC with titanlib
    - First guess CERRA (currently being back-extended to 1961)
    - Covering the entire period 1961–2020)
  - S-HYPE (hydrological product, tests have been conducted)
- Hopefully the method can then be simplified

Thank you