S1 Details about reanalysis products

All three global reanalyses assimilate conventional atmospheric measurements, with ERA5 and JRA-55 additionally assimilating surface snow depth observations and satellite-derived snow extent information. In these two reanalyses, in-situ snow depth measurements are assimilated during the land surface analysis step using 2-D optimal interpolation (2D-OI) schemes. 2D-OI primarily impacts observation-dense regions and nudges the first guess field towards the observed values. In regions with few or no in-situ snow depth observations, the land surface/snow model plays a more significant role in generating the reanalysis SWE. Further, the available historical snow data tends to be biased to open terrain, low-elevations, and the mid-latitudes.(Dyer & Mote, 2006; Mortimer et al., 2020). ERA5 and JRA-55 assimilate SYNOP snow depth observations, and JRA-55 additionally assimilates station data from Russia, USA, and Mongolia.

Apart from data assimilation, differences exist in data resolution and snow model complexity. Gridded ERA5 data are at a finer resolution of 0.25°x0.25°, while JRA-55 gridded data are coarser (1.25°x1.25°), and MERRA-2 falls in between (0.5°x0.625°). The ERA5 land model allows a single layer of snow on each sub-grid scale land tile. This snow layer has a temperature, mass, density, and albedo. Terrain albedo and snow-covered fraction, which is determined based on physical snow depth (diagnosed from mass and density), are used for other calculations at the snow-atmosphere interface. In the JRA-55 land model, there can be only one snow layer with evolving SWE and temperature, but with a constant density of 200 kg m^{-3} of snow assumed [source: https://www.jma.go.jp/jma/jma-eng/jma-center/nwp/outline2007-nwp/index.htm, last accessed 22-Jan-2024]. There is no snow albedo evolution. Finally, the MERRA-2 land model allows up to three snow layers (Reichle et al., 2017b). Snow depth, snow heat content, and snow water equivalent are all modeled.

S2 Bias correction methodology

Due to the differences in temperature and precipitation forcing between reanalyses, we implement a simple bias correction to test the impact of driving data biases on resulting SWE. For a chosen dataset, monthly multiplicative scaling factors are calculated with respect to a target dataset at each grid cell. The B-TIM is then run using driving data with scaling factors applied at each time step. Our method is based on climatologies and is intended to yield matching monthly climatologies (for precipitation or temperature) between the chosen dataset and a target dataset.

For a variable X representing precipitation or 2-meter temperature, the bias corrected version would be:

$$X_b(i,j,s,m,y) = X(i,j,s,m,y) \times SF(i,j,m), \tag{10}$$

where i and j represent the spatial dimensions. The variables s, m, and y indicate sub-monthly timestep, calendar month, and year. The scaling factors, $SF(i,j,m) = \frac{\overline{X_T(i,j,m)}}{\overline{X(i,j,m)}}$, depend on location and month and X_T refers to a target dataset. The overline indicates temporal averaging over all years for a particular calendar month, m. This method preserves the pattern and number of precipitation-free days and retains the fractional interannual variability of a given dataset, all while matching it to the target monthly climatology. Swapping different target datasets X_T provides a simple way to test the effect of reduced forcing differences – in particular, if it reduces the SWE biases that result. See for example Fig. 5.

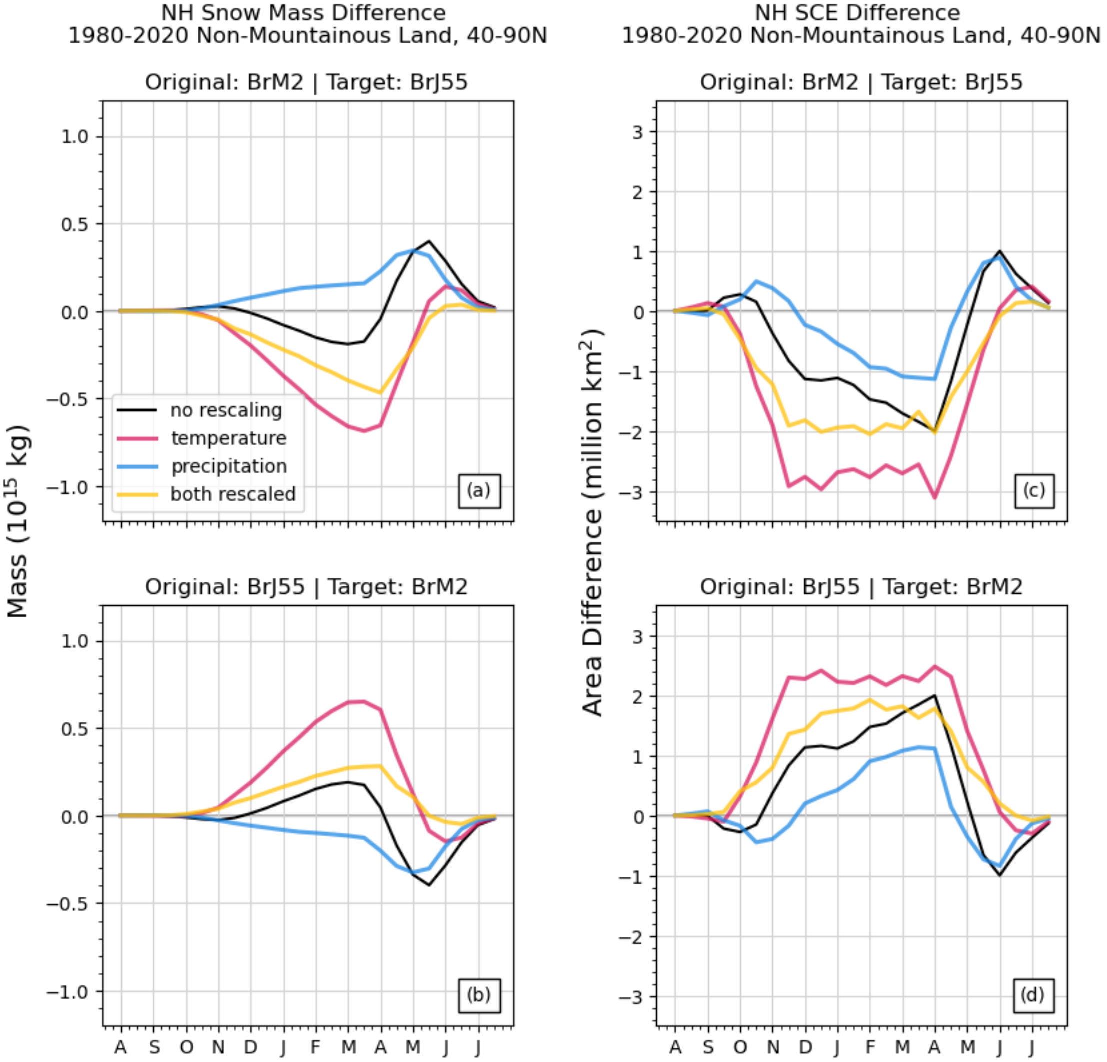

Fig. S1 Same as Fig. 5 for MERRA-2 and JRA-55. (a,b) NH snow mass and (c, d) snow cover extent differences calculated as original minus target. Each panel shows the difference between the original and target snow mass climatologies (black) and the coloured lines represent the versions resulting from adjusting temperature (pink), precipitation (blue), or both (yellow) to the target dataset's climatology.

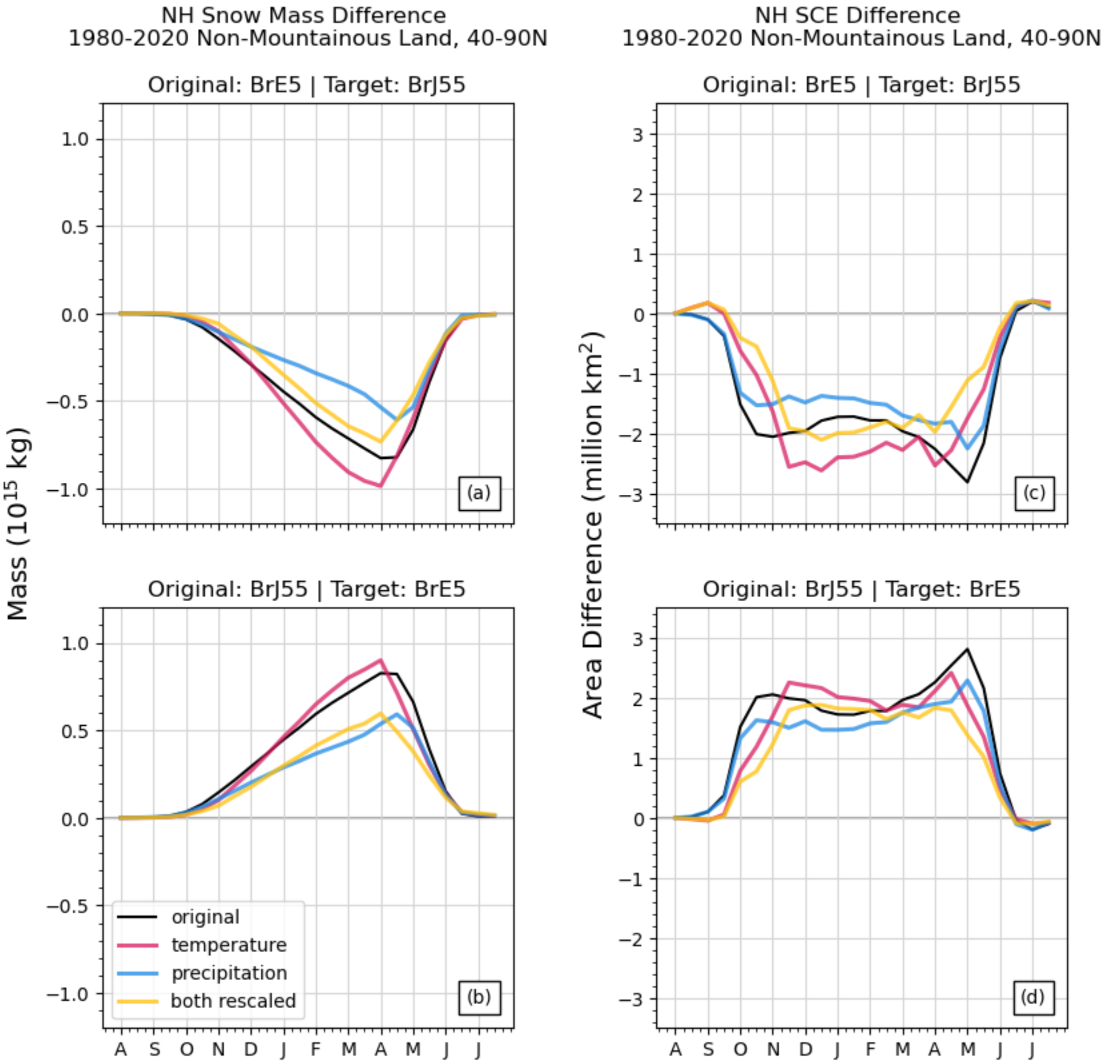

Fig. S2 Same as Fig. 5 for ERA5 and JRA-55. (a,b) NH snow mass and (c, d) snow cover extent differences calculated as original minus target. Each panel shows the difference between the original and target snow mass climatologies (black) and the coloured lines represent the versions resulting from adjusting temperature (pink), precipitation (blue), or both (yellow) to the target dataset's climatology.

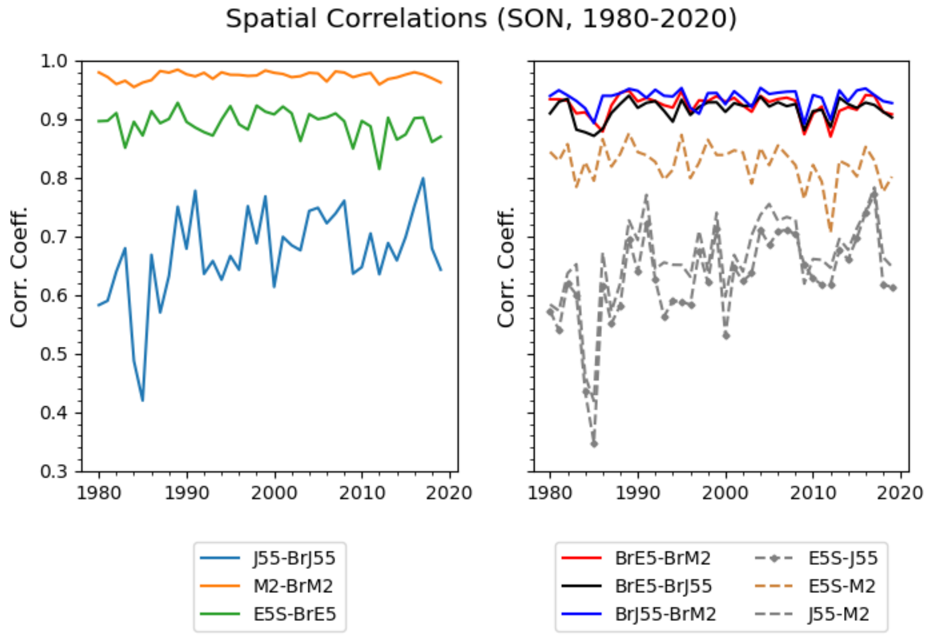

Fig. S3 Spatial correlations for SON calculated between pairs of datasets with the same meteorology (a) and between pairs of similar type (b; either offline-offline or reanalysis-reanalysis). Comparisons involving JRA-55 have the lowest spatial correlations in all cases.