Supplement of Geosci. Model Dev., 19, 10669–10686, 2026
https://doi.org/10.5194/gmd-19-10669-2026-supplement
© Author(s) 2026. CC BY 4.0 License.



Geoscientific Model Development
Open Access
EGU

Supplement of

MSDT v1.0: an open-source multi-method snow drought toolbox

Xiangfei Li et al.

Correspondence to: Jinzhi Tian (tianjzh@outlook.com)

The copyright of individual parts of the supplement might differ from the article licence.

Supplement

Figures and Tables

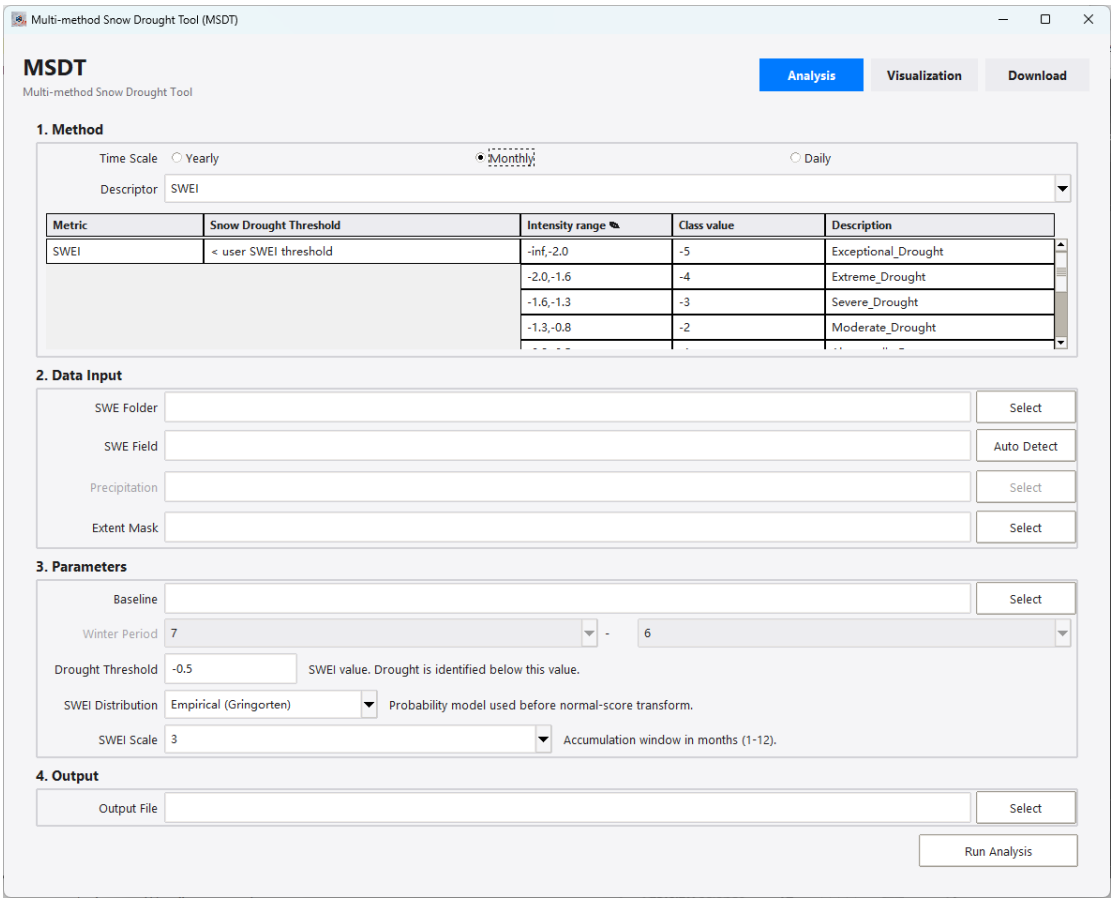


Figure S1. Graphical user interface of MSDT for snow drought identification.

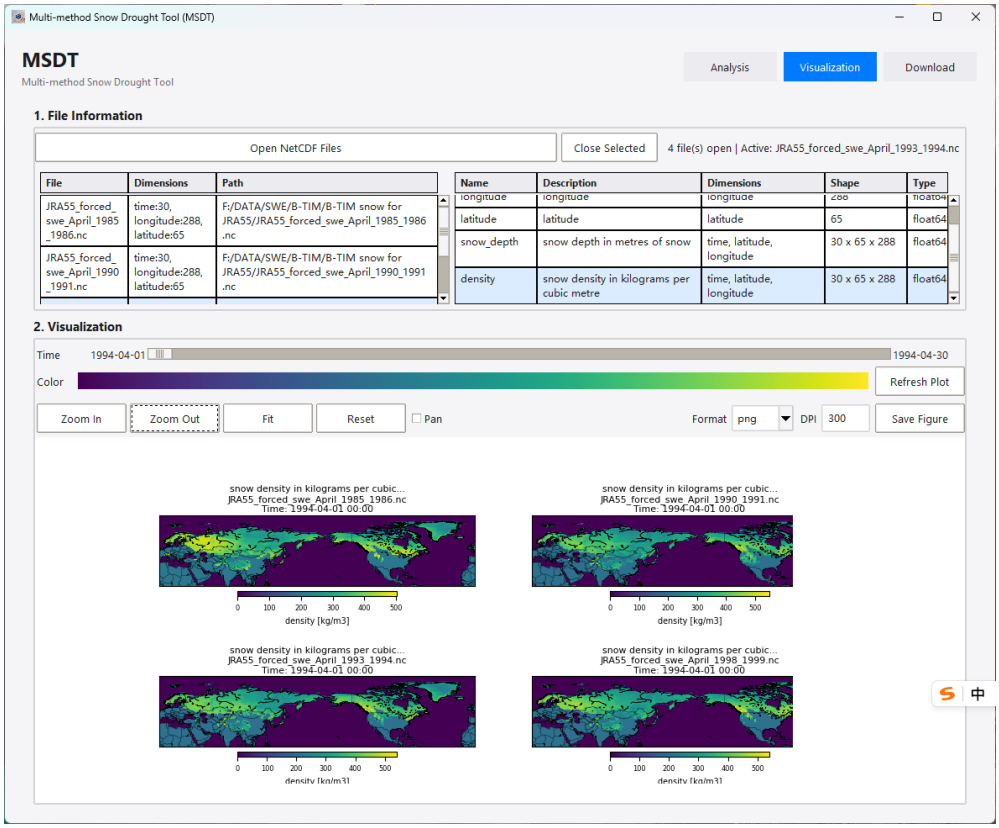


Figure S2. Graphical user interface of MSDT for snow data visualization.

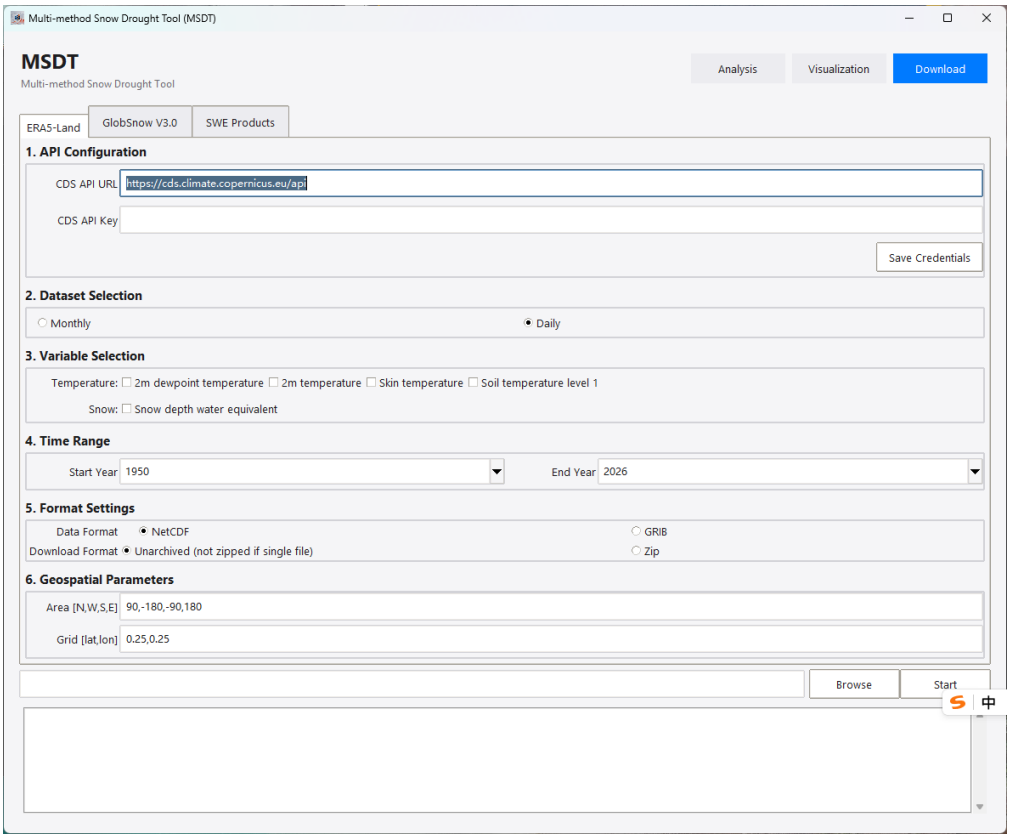


Figure S3. Graphical user interface of MSDT for data download.

Table S1. Classification schemes and flag values of snow drought intensity

Categories	Metric	Classification schemes and flag values
Point-in-time method (Yearly)	Accumulated SWE	$\begin{cases} Light(1), & 50\%-25\% \\ Medium(2), & 25\%-15\% \\ Heavy(3), & 15\%-5\% \\ Extreme(4), & <5\% \end{cases}$
Standardized index (SI) method (Monthly)	n-month accumulated SWEI	$\begin{cases} Abnormally\ Dry(1), & -0.8--0.5 \\ Moderate\ Drought(2), & -1.3--0.8 \\ Severe\ Drought(3), & -1.6--1.3 \\ Extreme\ Drought, & -2.0--1.6 \\ Exceptional\ Drought, & <-2.0 \end{cases}$
Phase diagram method (Daily)	SWE	$\begin{cases} Abnormal(1), & 30\%-20\% \\ Moderate(2), & 20\%-10\% \\ Severe(3), & 15\%-5\% \\ Extreme(4), & 5\%-2\% \\ Exceptional(5), & <2\% \end{cases}$

Table S2. Data type, units, and notes of Snow drought characteristics

Variable	Data type	Unit	Notes
sd_flag	integer, preferably int8	unitless / 1	Binary snow drought record: 1 = snow drought, 0 = non-drought. In current code, SWEI and PDM are explicitly int8; some yearly methods may be written as default integer unless further standardized.
sd_type_flag	int8	unitless / category	Only used by PDM. 0 = non-drought, 1 = warm snow drought, 2 = dry but warmer, 3 = dry but snowier.
intensity	floating point, usually float32 or float64	method-dependent	SWE anomaly for SWE-based methods, days for PeakSWEDate, and standardized index value for SWEI.
intensity_flag	integer, preferably int8	unitless / category	Severity class of snow drought. Flag values follow Table S1.
swe	floating point, usually float32 or float64	same as input SWE, commonly m or mm	SWE-related metric used by the selected method. Written mainly for yearly methods; SWEI and current PDM output do not write raw swe.
threshold	floating point, usually float32 or float64	method-dependent	Threshold used to identify snow drought. Unit follows the compared variable: SWE unit for SWE-based methods, days for PeakSWEDate, and not written as a separate variable for SWEI.

Table S3. SWE-related data products listed in the MSDT data-catalog panel.

Category	Product	Time	Resolution	Data source	Access link
In-situ	NRCS SNOTEL / Snow Course	long records to present	Point stations	USDA NRCS	https://www.nrcs.usda.gov/resources/data-and-reports/snow-and-water-interactive-map
	NOAA GHCN Daily	historical to present	Point stations	NOAA NCEI	https://www.ncei.noaa.gov/products/land-based-station/global-historical-climatology-network-daily
Reanalysis	ERA5-Land	1950-present	0.1 deg; hourly/monthly/daily derivatives	ECMWF / Copernicus CDS	https://cds.climate.copernicus.eu/datasets/reanalysis-era5-land
	MERRA-2 Land Diagnostics	1980-present	0.5 x 0.625 deg; hourly/time-averaged	NASA GMAO / GES DISC	https://disc.gsfc.nasa.gov/datasets/M2T1NXLND_5.12.4/summary
	GLDAS Noah	2000-present for v2.1	0.25 deg; 3-hourly/monthly	NASA GES DISC	https://disc.gsfc.nasa.gov/datasets/GLDAS_NOAH025_3H_2.1/summary
	SNODAS	30 September 2003-present	1 km; daily; CONUS	NOAA NOHRSC / NSIDC	https://nsidc.org/data/g02158/versions/1
Remote sensing	GlobSnow SWE v3.0	1979-2022 archive	25 km EASE-Grid; daily/monthly	Finnish Meteorological Institute / ESA heritage	https://www.globsnow.info/swe/archive_v3.0/
	ESA Snow CCI SWE	1979 onwards; product dependent	Passive-microwave grid products; product dependent	ESA Climate Change Initiative	https://climate.esa.int/en/projects/snow/
	AMSR-E / AMSR2 SWE	AMSR-E 2002-2011; AMSR2 2012-present	25 km EASE-Grid; daily	NASA NSIDC	https://nsidc.org/data/ae_dysno/versions/3

Text S1. Snow drought algorithm description

This document provides pseudocode, flowcharts, input and output conventions, and example parameter files for the seven snow drought identification methods implemented in the Multi-method Snow Drought Toolbox (MSDT).

Common input and output conventions

All methods use NetCDF input data on a regular time-latitude-longitude grid. The main snow variable is standardized to `sd`; precipitation for PDM is standardized to `tp`. Coordinate aliases such as `latitude`, `longitude`, `y`, `x`, `valid_time`, `T`, and `t` are standardized internally to `lat`, `lon`, and `time` before calculation. The expected gridded structure is:

```
Input SWE dataset:
  dimensions: time, lat, lon
  variable:   sd(time, lat, lon)

Input precipitation dataset for PDM:
  dimensions: time, lat, lon
  variable:   tp(time, lat, lon)

Common output variables:
  swe or method-specific metric, threshold, intensity, sd_flag
  optional: intensity_flag, percentile_swe, percentile_pre, sd_type_flag
```

For annual methods, SWE values less than or equal to 0.001 are treated as non-snow or invalid values before annual aggregation. For PDM, SWE values less than or equal to 0.001 are set to zero before daily rolling calculations. Hydrological-year methods remove 29 February and retain only complete analysis years containing all months required by the user-defined period. Missing values are ignored in xarray aggregation functions where `skipna=True` is used. In SWEI, a grid cell is calculated only when the full time series is non-negative and finite enough for the selected probability calculation; otherwise, the output is set to missing values for that grid cell.

The binary snow drought flag is consistently defined as:

```
sd_flag = 1 if current metric < threshold
sd_flag = 0 otherwise
```

except SWEI, where the flag is defined by the standardized index threshold:

```
sd_flag = 1 if SWEI < user-defined threshold
sd_flag = 0 otherwise
```

Algorithm S1. 1AprSWE

Pseudocode

```
Input: SWE dataset D, optional baseline B, threshold factor f
1. Sort D by time and retain SWE values greater than 0.001.
2. If B is not provided, set B = D.
3. Select all 1 April records from B.
4. For each grid cell, retain the cell only if at least 75 percent of baseline years are valid.
5. Compute baseline mean SWE on 1 April: M(lat, lon).
6. Compute threshold: T = M * f.
7. Select all 1 April records from D using the same validity logic.
8. Compute intensity: I(time, lat, lon) = current_1AprSWE - T.
9. Set sd_flag = 1 where I < 0, otherwise 0.
10. Write swe, threshold, intensity, sd_flag, and metadata.
```

Flowchart

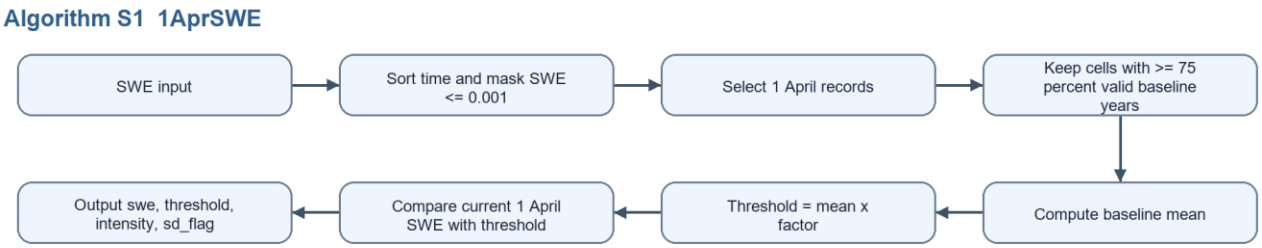


Figure S1-1. Algorithm S1. 1AprSWE.

Algorithm S2. PeakSWEMean

Pseudocode

```
Input: SWE dataset D, optional baseline B, winter period, threshold factor f
1. Sort D by time and retain SWE values greater than 0.001.
2. If B is not provided, set B = D.
3. Select complete hydrological years from B using the user-defined period.
4. For each hydrological year and grid cell, calculate annual peak SWE.
5. Compute long-term mean of baseline annual peak SWE: M(lat, lon).
6. Compute threshold: T = M * f.
7. Calculate annual peak SWE for D using the same hydrological-year definition.
8. Compute intensity: I = current_peak - T.
9. Set sd_flag = 1 where I < 0, otherwise 0.
10. Write swe, threshold, intensity, sd_flag, and metadata.
```

Flowchart

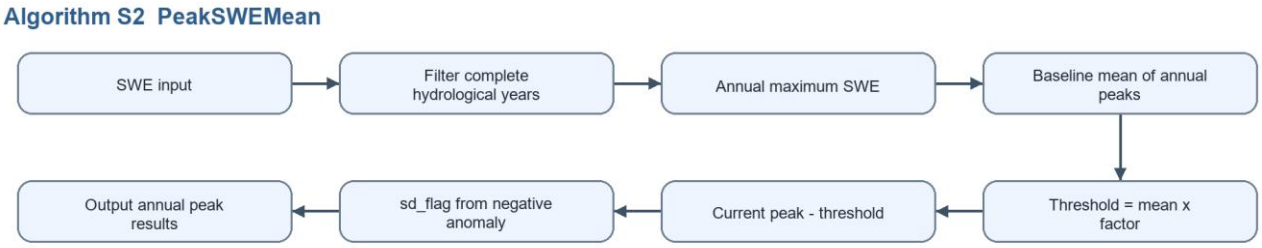


Figure S1-2. Algorithm S2. PeakSWEMean.

Algorithm S3. PeakSWEPctl

Pseudocode

```
Input: SWE dataset D, optional baseline B, winter period, percentile q
1. Sort D by time and retain SWE values greater than 0.001.
2. If B is not provided, set B = D.
3. Select complete hydrological years from B.
4. Calculate annual peak SWE for each baseline year and grid cell.
5. Compute threshold T as the qth percentile of baseline annual peaks using the nearest quantile method.
6. Calculate annual peak SWE for D.
7. Compute intensity: I = current_peak - T.
8. Set sd_flag = 1 where I < 0, otherwise 0.
9. Write swe, threshold, intensity, sd_flag, and metadata.
```

Flowchart

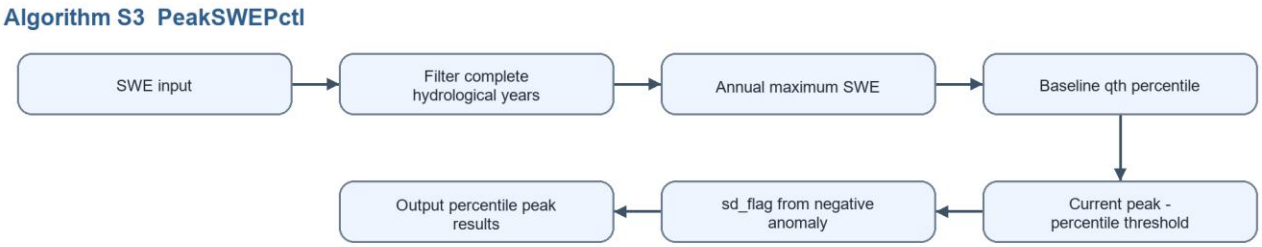


Figure S1-3. Algorithm S3. PeakSWEPctl.

Algorithm S4. PeakSWEDate

Pseudocode

```
Input: SWE dataset D, optional baseline B, winter period, percentile q
1. Sort D by time and retain SWE values greater than 0.001.
2. If B is not provided, set B = D.
3. Select complete hydrological years from B.
4. For each baseline hydrological year and grid cell, find the date of maximum SWE.
5. Convert the peak date to day number since the hydrological-year start.
6. Compute threshold T as the qth percentile of baseline peak dates using the nearest quantile method.
7. Calculate peak dates for D using the same period.
8. Compute intensity: I = current_peak_date - T.
9. Set sd_flag = 1 where I < 0, indicating an earlier-than-threshold peak date.
10. Write peak_swe_date, threshold, intensity, sd_flag, and metadata.
```

Flowchart

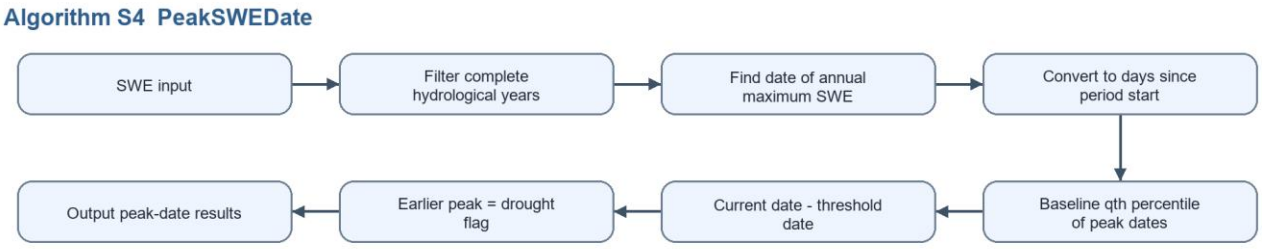


Figure S1-4. Algorithm S4. PeakSWEDate.

Algorithm S5. AccSWE

Pseudocode

```
Input: SWE dataset D, optional baseline B, winter period, drought percentile q, intensity rules R
1. Sort D by time and retain SWE values greater than 0.001.
2. If B is not provided, set B = D.
3. Select complete hydrological years from B.
4. For each hydrological year and grid cell:
   a. Sort records by time.
   b. Calculate positive SWE increments between consecutive time steps.
   c. Sum only positive increments to obtain annual accumulated SWE.
5. Compute all percentile boundaries required by q and R from baseline accumulated SWE.
6. Calculate current accumulated SWE for D.
7. Compute intensity: I = current_accumulation - threshold(q).
8. Set sd_flag = 1 where I < 0, otherwise 0.
9. Classify intensity_flag using validated, contiguous percentile ranges from 0 to q.
10. Write swe, threshold, intensity, sd_flag, intensity_flag, and metadata.
```

Flowchart

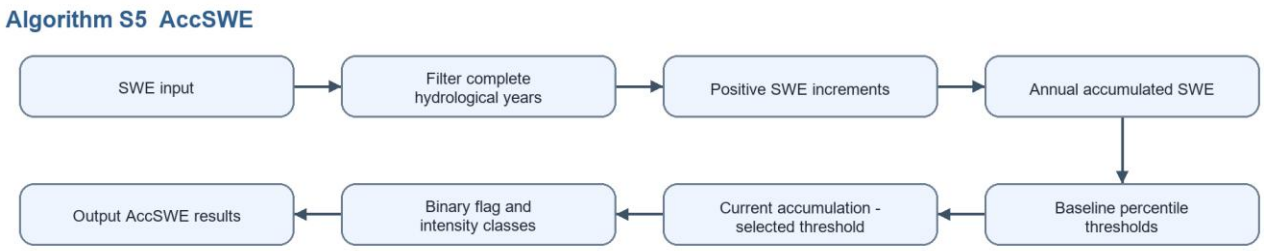


Figure S1-5. Algorithm S5. AccSWE.

Algorithm S6. SWEI

Pseudocode

```
Input: SWE dataset D, accumulation scale s, drought threshold z0, probability distribution P, intensity rules R
1. Sort D by time and extract sd(time, lat, lon).
2. For each grid cell:
   a. If the full time series contains invalid or negative values, set all SWEI values to missing.
   b. Set the first s - 1 months to missing because rolling accumulation is incomplete.
   c. Calculate s-month rolling accumulated SWE.
   d. Split accumulated values by calendar month.
   e. For each calendar month, estimate non-exceedance probability using P.
   f. Clip probabilities to [1e-6, 1 - 1e-6].
   g. Transform probabilities to standard normal scores: SWEI = inverse_normal_cdf(probability).
3. Set sd_flag = 1 where SWEI < z0, otherwise 0.
4. Classify intensity_flag using validated contiguous SWEI ranges.
5. Write intensity, sd_flag, intensity_flag, selected distribution, fallback policy, and metadata.
```

Probability distribution selection and rejection criteria

MSDT allows the user to select one of six probability options: empirical Gringorten plotting position, Gamma, Pearson type III, Normal, Lognormal, and Weibull. The empirical option is distribution-free and is used directly on ranked finite values. Gamma, Lognormal, and Weibull are positive-support distributions; they are fitted only to positive accumulated SWE values, while the empirical probability of zero values is retained through a zero-probability adjustment. Normal and Pearson type III are fitted to all finite accumulated values. For each grid cell and calendar month, parametric fitting is rejected and replaced by the empirical Gringorten method if fewer than two finite distinct values are available or if the fitting routine fails. MSDT therefore does not automatically choose a statistically optimal distribution; instead, it implements a user-selected distribution with a conservative grid-level fallback. For reproducible studies, users should report the selected distribution and are encouraged to test sensitivity across several distributions in regions with short records, intermittent snow cover, or strongly skewed accumulated SWE.

Flowchart

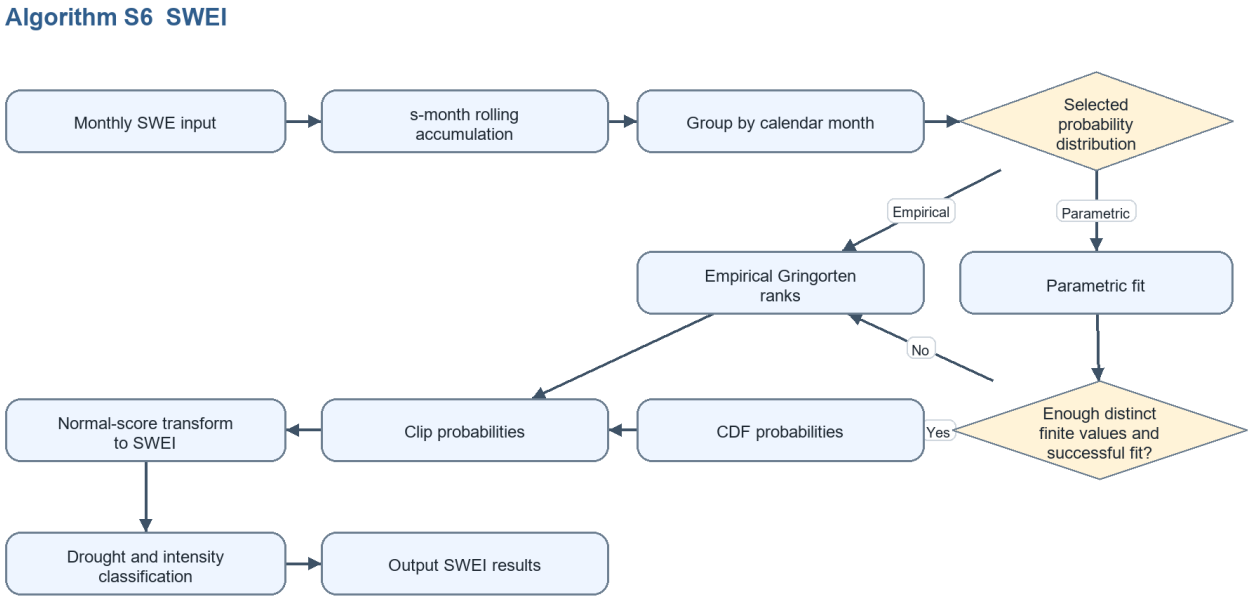


Figure S1-6. Algorithm S6. SWEI.

Algorithm S7. PDM

Pseudocode

```
Input: daily SWE D, optional SWE baseline B, daily precipitation P, winter period, percentile q, rules R, optional
extent mask.
1. Standardize coordinates; crop D, B and P to the active mask bounding box when a mask is supplied.
2. Select winter months from every input, including both calendar years for a cross-year winter.
3. Estimate filtered input size and available memory; choose one in-memory block or latitude blocks (see memory-
planning pseudocode below).
4. Skip fully masked blocks; retain the full reference period within every active spatial block.
5. Schedule at most W block tasks (W <= 2), with a bounded queue of pending/completed block results.
6. For each block, load D, optional B and P as float32, then perform the following complete PDM calculation:
   a. Replace current SWE values not greater than 0.001 with zero; use the cleaned D as baseline if B is absent.
   b. Sort inputs by time and validate identical SWE-precipitation latitude-longitude coordinates.
   c. Remove 29 February and select complete hydrological years for the requested winter period.
   d. Calculate centered 7-day means within each complete year, requiring at least 3 finite observations.
   e. Reuse baseline preprocessing when current and baseline datasets are identical.
   f. Calculate daily threshold climatology T = np.percentile(processed_baseline, q, axis=year).
   g. Calculate intensity I = processed_current_SWE - T; set sd_flag = 1 if I < 0, otherwise 0.
   h. Rank current SWE against the baseline at each winter day and grid cell to obtain percentile_swe.
      Empirical ranks count strictly smaller reference values; ties therefore have the same lower rank.
   i. Rank processed precipitation against its own analysis-period distribution to obtain percentile_pre.
   j. Validate R as contiguous, non-overlapping ranges covering 0 to q with positive class values.
      Calculate current-period SWE quantiles at the range boundaries and assign intensity_flag.
      Intervals are lower-exclusive and upper-inclusive; the lowest interval includes its lower end.
      Severity classes are assigned independently of sd_flag and do not reuse baseline percentile_swe.
   k. Assign sd_type_flag only where sd_flag = 1 (otherwise leave 0):
      percentile_pre > 50: 1, Warm_Snow_Drought;
      percentile_pre <= 50 and percentile_pre > percentile_swe: 2, Dry_but_Warmer;
      percentile_pre <= 50 and percentile_pre <= percentile_swe: 3, Dry_but_Snowier.
   l. Return the block's seven output variables, keeping T as a compact day-latitude-longitude climatology.
7. A single writer appends each completed block to every relevant annual NetCDF file at its latitude slice.
   Expand T only for the current annual write; apply the extent mask and do not create a full-period tiled threshold.
8. Release the written block and its task reference; submit the next block without accumulating full-domain outputs.
9. After all active blocks are written, finalize files as <output_stem>_<hydrological_year>.nc.
   Use NetCDF4 zlib level 4, shuffle, and chunks aligned with the spatial block rows.
Output: threshold, intensity, sd_flag, intensity_flag, percentile_swe, percentile_pre and sd_type_flag.
        All exported variables have dimensions (time, lat, lon); no day, hydro_year or day_in_period is exported.
```

Implementation notes: current-SWE cleaning also replaces NaNs with zero because the condition SWE > 0.001 is false for NaNs. A separately supplied baseline and precipitation retain NaNs. Rolling means require three finite values; thresholds and severity boundaries use np.percentile, not np.nanpercentile, so missing values may propagate. Non-finite percentile ranks are missing, whereas flags remain zero when their comparisons fail; a zero flag alone does not establish that observations were complete. Output values outside an extent mask are missing. The current PDM time coordinate is a sequential step index; hydrological-year labels are carried by annual filenames.

Flowchart

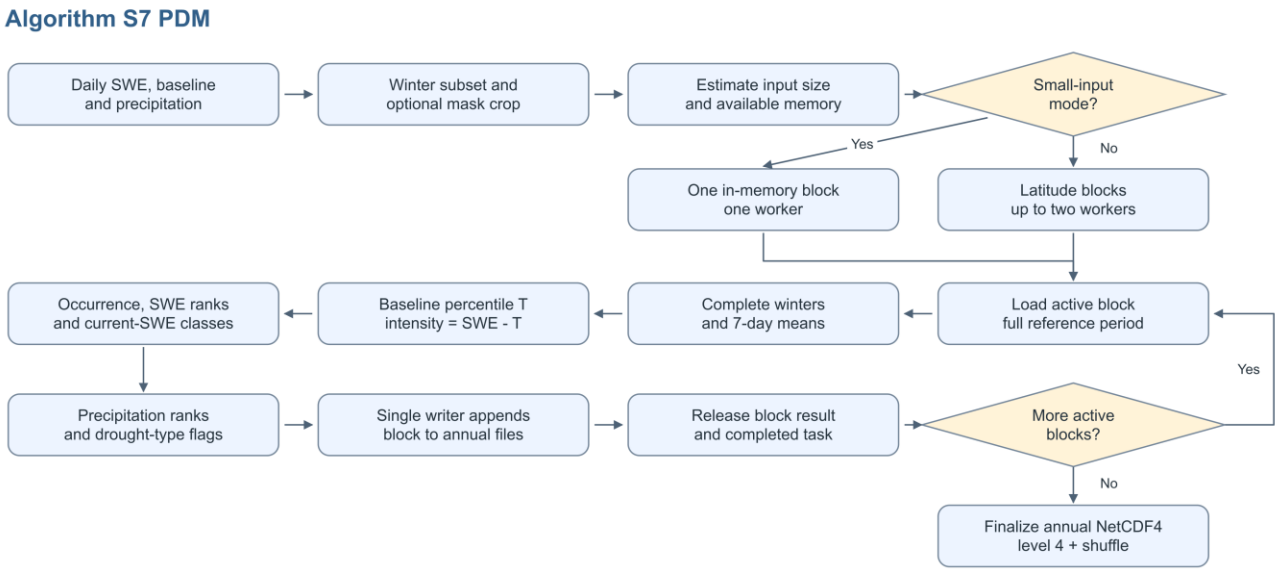


Figure S1-7. Daily PDM with spatial blocking throughout calculation and incremental annual output.

Adaptive parallelism and latitude-block partitioning

The PDM GUI calls calculate_daily_snow_analysis_to_files in DailyStreaming.py. Its memory plan covers the complete calculation and incremental output pipeline, not only daily preprocessing. The planner uses the largest float32 input-field estimate, reserves memory for the operating system, limits concurrent block tasks to two, and releases each block after writing. Available memory is queried using psutil.virtual_memory().available; conservative fallbacks are used if the query is unavailable. The following pseudocode reproduces the implemented planning heuristic.

```
Input: winter-filtered D, optional B, P after extent-mask bounding-box cropping.
Definitions: clamp(x, lo, hi) = min(hi, max(lo, x)); integer divisions below use floor.
Constants: L0 = 2 GiB; Lmin = 512 MiB; T0 = 512 MiB; Tmin = 256 MiB; Tmax = 1 GiB.
           Wmax = 2; F = 16 (whole-pipeline working-memory multiplier).

1. Detect A = psutil.virtual_memory().available; use A = unknown if detection fails.
2. Estimate S_j = n_time_j * n_lat * n_lon * 4 bytes for each supplied input field.
   S = max(S_D, S_P, S_B if supplied); n_time_max = max(n_time_j).
   This estimate precedes removal of 29 February and incomplete hydrological years.
3. Set Wreq = clamp(valid MSDT_DAILY_WORKERS value, 1, 2), or 2 when unset/invalid.
4. Set the in-memory limit L:
   configured MSDT_DAILY_IN_MEMORY_GB value if valid;
   otherwise max(Lmin, min(L0, floor(0.35 * A))) if A is known; otherwise L0.
5. Set initial working-memory budget P0:
   configured MSDT_DAILY_PARALLEL_MEMORY_GB value if valid;
   otherwise max(1 GiB, floor(0.50 * A)) if A is known; otherwise 2 GiB.
6. Apply the operating-system reserve:
   P = min(P0, max(1 byte, A - max(2 GiB, floor(0.20 * A)))) if A is known; otherwise P = P0.
7. Obtain the preliminary block target Tbase from the existing preprocessing helper:
   if MSDT_DAILY_BLOCK_GB is valid: Tbase = clamp(configured value, Tmin, Tmax).
   else if A is known: Tbase = clamp(max(T0, floor(P0 / (Wreq * 3))), Tmin, Tmax).
   else: Tbase = T0.
8. Account for all PDM intermediate and output arrays, not only preprocessing:
   T = clamp(min(Tbase, max(1 byte, floor(P / (Wreq * F)))), Tmin, Tmax).
   W = min(Wreq, max(1, floor(P / (T * F)))).
9. Select the small-input path only if both the input limit and working budget permit:
   if S <= min(L, floor(P / F)): T = S; W = 1.
   This path uses one full spatial block; annual output is still incremental.
10. Partition latitude using all filtered input times:
    bytes_per_lat_row = max(1 byte, n_time_max * n_lon * 4 bytes).
    rows = clamp(floor(T / bytes_per_lat_row), 1, n_lat).
    slices = consecutive latitude slices of at most rows rows.
    Remove slices containing no active mask cells; error if the mask has no active cells.
    W = min(W, number of slices before mask skipping, number of active slices).
11. Execute a bounded compute-write-release pipeline:
    if W == 1: compute active blocks sequentially.
```

```
    else: use ThreadPoolExecutor(max_workers=W), with at most W pending block tasks.
    for each completed block in latitude order:
        load all reference-period times for this block and compute the full PDM result (Algorithm S7).
        retain threshold as a compact day-latitude-longitude climatology within the block.
        use one writer and the shared NetCDF library lock to append the block to annual files.
        use chunks = (min(16, winter_days), min(rows, n_lat), min(360, n_lon)).
        write the current annual threshold slice, not a globally tiled threshold array.
        release the block and its completed task reference, then submit the next active block.
12. Finalize annual files after all active slices are written; clean unfinished temporary files on
cancellation/error.
```

The factor 3 in step 7 belongs to the existing preliminary preprocessing helper; the outer PDM planner additionally applies the full-pipeline factor F = 16 in steps 8-9. Block targets describe one float32 input field, not the total memory occupied by SWE, precipitation, intermediate arrays and outputs. The 256 MiB-1 GiB target range applies to the blocked path; a permitted single-block small-input run can use up to the in-memory limit. These are planning estimates rather than a guarantee of peak resident memory. Numeric environment overrides are interpreted in GiB (despite their _GB names); worker count remains capped at two.

Adaptive parallelism flowchart

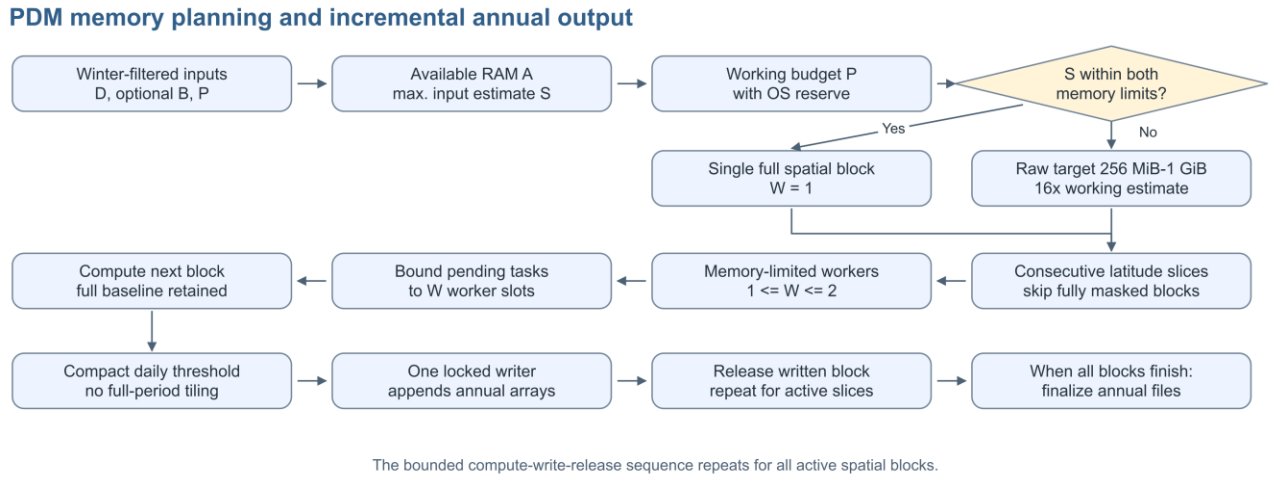


Figure S1-8. Whole-pipeline memory planning and the bounded compute-write-release sequence for daily PDM.

Example parameter files

1AprSWE

Identify snow drought when April 1 SWE is below the baseline mean multiplied by 1.0. This example applies to the Northern Hemisphere.

1apr_swe.yaml

```
method: 1AprSWE
time_scale: yearly
input:
  swe_folder: ./data/swe_daily/
  swe_variable: sd
  baseline_folder: ./data/swe_daily/
  extent_mask: ./data/masks/extent_mask.xlsx
parameters:
  drought_threshold: 1.0
output:
  file: ./output/1AprSWE.nc
```

Output fields: `swe`, `threshold`, `intensity`, `sd_flag`.

PeakSWEMean

Identify snow drought when the July-June peak SWE is below the baseline mean annual peak multiplied by 1.0.

peak_swe_mean.yaml

```
method: PeakSWEMean
time_scale: yearly
input:
  swe_folder: ./data/swe_daily/
  swe_variable: sd
  baseline_folder: ./data/swe_daily/
  extent_mask: ./data/masks/extent_mask.xlsx
parameters:
  start_month: 7
  end_month: 6
  drought_threshold: 1.0
output:
  file: ./output/PeakSWEMean.nc
```

Output fields: `swe`, `threshold`, `intensity`, `sd_flag`. `time_bnds` defines each annual reporting interval.

PeakSWEPctl

Identify snow drought when the July-June peak SWE is below its baseline 25th percentile. Enter 25 in the GUI Drought Threshold field.

peak_swe_percentile.yaml

```
method: PeakSWEPctl
time_scale: yearly
input:
  swe_folder: ./data/swe_daily/
  swe_variable: sd
  baseline_folder: ./data/swe_daily/
  extent_mask: ./data/masks/extent_mask.xlsx
parameters:
  start_month: 7
  end_month: 6
  drought_threshold: 0.25
output:
  file: ./output/PeakSWEPctl.nc
```

Output fields: `swe`, `threshold`, `intensity`, `sd_flag`. `time_bnds` defines each annual reporting interval.

PeakSWEDate

Identify snow drought when the peak SWE occurs earlier within the July-June reporting period than the baseline 25th-percentile occurrence day. Enter 25 in the GUI Drought Threshold field.

peak_swe_date.yaml

```
method: PeakSWEDate
time_scale: yearly
input:
  swe_folder: ./data/swe_daily/
  swe_variable: sd
  baseline_folder: ./data/swe_daily/
  extent_mask: ./data/masks/extent_mask.xlsx
parameters:
  start_month: 7
  end_month: 6
  drought_threshold: 0.25
output:
  file: ./output/PeakSWEDate.nc
```

Output fields: `peak_swe_date`, `threshold`, `intensity`, `sd_flag`. `time_bnds` defines each annual reporting interval.

AccSWE

Identify snow drought when the accumulation of positive SWE increments during July-June is below its baseline median. Enter 50 in the GUI Drought Threshold field.

acc_swe.yaml

```
method: AccSWE
time_scale: yearly
input:
  swe_folder: ./data/swe_daily/
  swe_variable: sd
  baseline_folder: ./data/swe_daily/
  extent_mask: ./data/masks/extent_mask.xlsx
parameters:
  start_month: 7
  end_month: 6
  drought_threshold: 0.5
  intensity_rules:
  - break_value: 25-50
    value: 1
    description: Light
  - break_value: 15-25
    value: 2
    description: Medium
  - break_value: 5-15
    value: 3
    description: Heavy
  - break_value: 0-5
    value: 4
    description: Extreme
output:
  file: ./output/AccSWE.nc
```

Output fields: `swe`, `threshold`, `intensity`, `sd_flag`, `intensity_flag`. `time_bnds` defines each annual reporting interval.

SWEI

Calculate a three-month SWEI using empirical Gringorten probabilities and identify snow drought when SWEI is below -0.5. The complete classification spans dry, near-normal, and wet conditions.

swei.yaml

```
method: SWEI
time_scale: monthly
input:
  swe_folder: ./data/swe_monthly/
  swe_variable: sd
  extent_mask: ./data/masks/extent_mask.xlsx
parameters:
  scale: 3
  distribution: empirical
  drought_threshold: -0.5
  intensity_rules:
  - break_value: -inf,-2.0
    value: -5
    description: Exceptional_Drought
  - break_value: -2.0,-1.6
    value: -4
    description: Extreme_Drought
  - break_value: -1.6,-1.3
    value: -3
    description: Severe_Drought
  - break_value: -1.3,-0.8
    value: -2
    description: Moderate_Drought
  - break_value: -0.8,-0.5
    value: -1
    description: Abnormally_Dry
  - break_value: -0.5,0.5
    value: 0
    description: Near_Normal
  - break_value: 0.5,0.8
    value: 1
    description: Abnormally_Wet
  - break_value: 0.8,1.3
    value: 2
    description: Moderate_Wet_Spell
  - break_value: 1.3,1.6
    value: 3
    description: Severe_Wet_Spell
  - break_value: 1.6,2.0
    value: 4
    description: Extreme_Wet_Spell
  - break_value: 2.0,inf
    value: 5
    description: Exceptional_Wet_Spell
output:
  file: ./output/SWEI.nc
```

Output fields: `intensity`, `sd_flag`, `intensity_flag`.

PDM

Identify daily snow drought below the baseline 30th-percentile SWE threshold during November-March. Daily SWE and precipitation are required for the diagnostic type classification.

pdm.yaml

```
method: PDM
time_scale: daily
input:
  swe_folder: ./data/swe_daily/
  swe_variable: sd
  baseline_folder: ./data/swe_daily/
  precipitation_folder: ./data/precip_daily/
  precipitation_variable: tp
  extent_mask: ./data/masks/extent_mask.xlsx
parameters:
  start_month: 11
  end_month: 3
  drought_threshold: 30.0
  intensity_rules:
  - break_value: 20-30
    value: 1
    description: Abnormally_Dry
  - break_value: 10-20
    value: 2
    description: Moderate_Drought
  - break_value: 5-10
    value: 3
    description: Severe_Drought
  - break_value: 2-5
    value: 4
    description: Extreme_Drought
  - break_value: 0-2
    value: 5
    description: Exceptional_Drought
output:
  file: ./output/PDM.nc
```

Output fields: `threshold`, `sd_flag`, `intensity`, `intensity_flag`, `percentile_swe`, `percentile_pre`, `sd_type_flag`.

The output path `./output/PDM.nc` produces files such as `PDM_1980.nc`, where 1980 denotes the starting year of the November 1980-March 1981 reporting period. Each file contains actual daily dates, the hydrological-period metadata, and `threshold(time, lat, lon)`.

Text S2. NetCDF output structure

File organization

MSDT exports results as NetCDF4 files following the Climate and Forecast (CF) metadata conventions, version 1.8, identified by the global attribute `Conventions = "CF-1.8"`. Each file contains spatial and temporal coordinates, method-specific diagnostic variables, and metadata describing the analysis settings. Annual and monthly results are saved in individual files, and daily PDM results are organized into separate hydrological-period files.

Dimensions and coordinates

Table S2-1 Dimensions and coordinate variables

Name	Dimensions	Stored type	Units and attributes	Interpretation
`time`	`(time)`	float64	Days since 1970-01-01 00:00:00; explicit `calendar`; `standard_name = "time"`; `axis = "T"`	Observation date or start date of an annual reporting period.
`lat`	`(lat)`	Input coordinate type	`degrees_north`; `standard_name = "latitude"`; `axis = "Y"`	Geographic latitude of grid-cell centres.
`lon`	`(lon)`	Input coordinate type	`degrees_east`; `standard_name = "longitude"`; `axis = "X"`	Geographic longitude of grid-cell centres.
`time_bnds`	`(time, nv)`	float64	Reference-time units and calendar associated with `time`	Start and end of each reporting interval for PeakSWEMean, PeakSWEPctl, PeakSWEDate, and AccSWE; `nv = 2`.

For 1AprSWE, `time` records the actual April 1 observation date. For PeakSWEMean, PeakSWEPctl, PeakSWEDate, and AccSWE, it records the first day of the selected reporting period, and `time_bnds` gives the inclusive start and exclusive end. For example, a July-June reporting period beginning in 1980 has bounds `[1980-07-01, 1981-07-01)`.

SWEI uses the input monthly timestamps; each index corresponds to the accumulation ending in the labelled month. PDM uses the actual daily dates within the selected winter period. A November-March period beginning in 1980 covers November-December 1980 and January-March 1981. Hydrological-period processing excludes 29 February, as documented in `leap_day_policy`. The `calendar` attribute specifies the calendar used to interpret time.

Diagnostic variables

Most diagnostic variables have dimensions `(time, lat, lon)`. Annual thresholds have dimensions `(lat, lon)`, and PDM thresholds have dimensions `(time, lat, lon)`. SWEI records its scalar identification threshold in the global `drought_threshold` attribute.

In Table S2-2, `U` denotes the physical SWE units specified by the input data, such as `m`, `mm`, or `kg m-2`. SWE quantities and departures use these units, peak-date departures use `days`, and SWEI is dimensionless. Input SWE units are recorded in `input_swe_units`.

Table S2-2 Diagnostic variables

Variable	Dimensions	Stored type	Units	Definition and availability
`sd_flag`	`(time, lat, lon)`	int8	`1`	Binary snow drought classification for all methods: `0 = non_drought`, `1 = snow_drought`.
`sd_type_flag`	`(time, lat, lon)`	int8	`1`	PDM snow drought type, determined from SWE and precipitation percentiles.
`intensity`	`(time, lat, lon)`	float32 or float64	`U`, days, or `1`	SWE-metric departure from the identification threshold, peak-date departure, standardized SWEI value, or processed daily SWE departure.
`intensity_flag`	`(time, lat, lon)`	Signed integer; int8 for default classes	`1`	Classification of drought severity for AccSWE and PDM, and dry-to-wet conditions for SWEI.
`swe`	`(time, lat, lon)`	float32 or float64	`U`	April 1 SWE for 1AprSWE, annual peak SWE for PeakSWEMean and PeakSWEPctl, or accumulated positive SWE increments for AccSWE. Includes all valid annual records.
`peak_swe_date`	`(time, lat, lon)`	float64	days	One-based elapsed day of peak SWE relative to the reporting-period start; PeakSWEDate only.
`threshold`	`(lat, lon)` for annual methods; `(time, lat, lon)` for PDM	float32 or float64	`U`; days for PeakSWEDate	Numeric SWE or peak-date threshold used for identification.
`percentile_swe`	`(time, lat, lon)`	float32	percent	Percentile rank of processed daily SWE; PDM only.
`percentile_pre`	`(time, lat, lon)`	float32	percent	Percentile rank of processed daily precipitation; PDM only.

SWEI and PeakSWEDate diagnostics use float64, and PDM diagnostics use float32. The precision of annual SWE diagnostics depends on the input data and calculation. Variables are described by `long_name`, `units`, and method-specific descriptive attributes. Annual peak SWE carries `cell_methods = "time: maximum"`, and `ancillary_variables` links `intensity` to its classification fields.

Classification and missing values

Categorical variables carry `flag_values`, `flag_meanings`, and `valid_range`. Attribute arrays have the same integer type as their associated variable. `flag_meanings` contains space-separated category tokens, while `flag_descriptions` stores the corresponding descriptive labels. Editable intensity intervals, codes, and descriptions are recorded in `classification_rules`.

Table S2-3 Default classification codes

Variable	Default values	Default meanings
`sd_flag`	`0, 1`	Non-drought; snow drought.

Variable	Default values	Default meanings
AccSWE `intensity_flag`	`0, 1, 2, 3, 4`	Non-drought; light; medium; heavy; extreme.
SWEI `intensity_flag`	`-5, -4, -3, -2, -1, 0, 1, 2, 3, 4, 5`	Negative values indicate increasing dryness with decreasing code; zero indicates near-normal conditions; positive values indicate increasing wetness.
PDM `intensity_flag`	`0, 1, 2, 3, 4, 5`	Non-drought; abnormally dry; moderate, severe, extreme, and exceptional drought.
PDM `sd_type_flag`	`0, 1, 2, 3`	Non-drought; warm snow drought; dry but warmer; dry but snowier.

The output attributes specify the classification rules used for each analysis, including user-defined intervals and labels. Binary identification follows the strict comparison recorded in `drought_comparison`; intensity-class intervals use their documented endpoint conventions.

Floating-point diagnostics use a NaN `_FillValue`. Default int8 classifications use `_FillValue = -128`. Custom classification codes are stored as int8, int16, int32, or int64 according to their range, with the minimum representable integer reserved for missing data. Missing diagnostics, unavailable accumulation periods, and cells outside the extent mask are represented by fill values. A classification code of zero denotes a valid non-drought or near-normal record, according to the variable.

Global metadata and compression

Global attributes identify the file contents and provenance through `title`, `institution`, `source`, `history`, `software`, `output_schema_version`, and `identification_method`. Analysis metadata specify the baseline period, physical units, temporal interpretation, reporting-period months, and identification thresholds. SWEI metadata also record the accumulation length, selected probability distribution, and fitting fallback policy. Classification metadata describe the severity intervals, codes, and labels.

PDM files include `hydrological_year_start`, `reporting_period_start`, and `reporting_period_end_exclusive`, identifying the hydrological period represented by each file.

Numeric diagnostic variables use NetCDF4 zlib level-4 lossless compression with shuffle and dimension-based storage chunks. Coordinates and time bounds provide the spatial and temporal reference for interpreting the diagnostic fields.

Text S3. Performance benchmarks

Benchmark configuration

All benchmarks used an Intel Core i9-12900K workstation (16 cores, 24 logical processors), 128 GB installed RAM, Windows 11 Pro, and Python 3.11.7. Inputs and final outputs were stored on a SATA hard drive. Each configuration was run three times in an independent process. Wall time covers input opening, calculation, compression and completion of output files, excluding interpreter startup and validation. Peak RAM is the maximum process resident memory, using both sampled RSS and the Windows high-water mark. Operating-system caches were not flushed.

The full-domain inputs cover 20–84°N at 0.25° over 1980–2024. Annual methods use daily SWE (1.52 GiB); SWEI uses monthly SWE (0.37 GiB); PDM uses daily SWE and precipitation (8.79 GiB combined). The same extent mask gives a 234 × 1440 bounding grid with 116,867 active cells. Each selected input period also serves as its baseline.Default thresholds and classifications are used. Annual analyses use July–June; SWEI uses a 3-month accumulation and empirical Gringorten plotting positions (threshold −0.5); PDM uses November–March and the 30th-percentile threshold.Full-domain computational performance

Table S3-1. Full-domain performance of all seven methods.

Method	Scale	Wall time (s) Median (range)	Peak RAM (GiB) Median (maximum)
1AprSWE	Yearly	124.6 (124.0–130.9)	1.76 (1.80)
PeakSWEMean	Yearly	232.3 (230.9–236.5)	27.11 (27.23)
PeakSWEPctl	Yearly	213.2 (212.2–216.0)	3.88 (3.92)
PeakSWEDate	Yearly	319.3 (317.5–324.7)	20.54 (20.84)
AccSWE	Yearly	314.2 (310.8–317.6)	19.79 (19.79)
SWEI	Monthly	117.3 (115.3–120.2)	5.28 (5.28)
PDM	Daily	1,644.0 (1,643.3–1,663.6)	28.45 (28.50)

Main findings. The five annual methods completed in 124.6–319.3 s, while monthly SWEI required 117.3 s. Daily PDM had the highest computational cost (1,644.0 s, or 27.4 min), with median peak RAM of 28.45 GiB. Memory demand also varied substantially among annual methods (1.76–27.11 GiB), showing that temporal scale alone does not determine resource requirements.

Lossless output compression

Table S3-2. Lossless NetCDF4 storage using identical decoded outputs.

Method	Uncompressed (MiB)	Compressed (MiB)	Reduction (%)
1AprSWE	573.5	50.0	91.3
PeakSWEMean	573.5	67.6	88.2
PeakSWEPctl	573.5	58.7	89.8
PeakSWEDate	573.5	10.6	98.2
AccSWE	668.5	62.6	90.6
SWEI	4,303.2	274.9	93.6

Method	Uncompressed (MiB)	Compressed (MiB)	Reduction (%)
PDM	1,495.3	171.4	88.5

Each pair re-encodes identical values, dimensions and data types with the same chunk layout; compression uses zlib level 4 with shuffle. Decoded values are exactly equal in all tests. Rows cover complete output sets except PDM, which uses one representative 151-day winter (year label 2000). MiB = 2^{20} bytes; GiB = 2^{30} bytes.

Main findings. Lossless NetCDF4 compression reduced file sizes by 88.2–98.2% without changing decoded values. PeakSWEDate showed the largest reduction (98.2%); reductions for SWEI and the representative PDM winter were 93.6% and 88.5%, respectively.

Temporal and spatial scaling

Scaling tests evaluate AccSWE, SWEI and PDM as representative annual, monthly and daily workflows. The full-domain tests vary input duration and spatial sampling; the regional tests vary grid spacing over the same 4° × 4° area. All use the platform, timing definition, defaults and three-run protocol described above.

Table S3-3. Temporal coverage and spatial grid scaling of annual, monthly and daily workflows. Each method cell gives median wall time (s) / peak RAM (GiB).

Domain, spacing and input period	Grid size (lat × lon)	AccSWEs / GiB	SWEIs / GiB	PDMs / GiB
A Temporal coverage at 0.25°				
20–84°N, 0.25° 1980–1989	234 × 1440	83.2 / 5.31	89.8 / 1.32	206.3 / 28.97
20–84°N, 0.25° 1980–1999	234 × 1440	152.1 / 10.83	96.9 / 2.45	504.5 / 24.73
20–84°N, 0.25° 1980–2024	234 × 1440	314.2 / 19.79	117.3 / 5.28	1,644.0 / 28.45
B Full-domain spatial coarsening from the 0.25° reference				
20–84°N, 0.5° 1980–2024	117 × 720	186.5 / 19.68	35.5 / 1.86	537.7 / 24.22
20–84°N, 1° 1980–2024	56 × 360	149.0 / 19.62	12.5 / 1.67	205.0 / 7.98
C Regional grid refinement over a fixed 4° × 4° domain				
Regional, 0.25° 1980–2024	16 × 16	2.7 / 0.24	1.9 / 0.20	3.4 / 0.36
Regional, 0.1° 1980–2024	40 × 40	3.5 / 0.39	2.7 / 0.22	7.1 / 0.87
Regional, 0.01° 1980–2024	400 × 400	252.0 / 5.85	101.9 / 2.62	969.2 / 28.15

Grid size denotes the allocated bounding grid after masking. Coarser full-domain grids use spatial subsampling. Regional tests cover 44–48°N and 84–88°E; 0.1° and 0.01° inputs are nearest-neighbour replications for computational scaling, not native fine-resolution observations or evidence of full-domain 0.01° feasibility. Adaptive blocking means PDM peak memory need not increase monotonically with input size.

Main findings. Extending the full-domain record from 10 to 45 years increased runtime by factors of 3.8 (AccSWE),

1.3 (SWEI) and 8.0 (PDM). Coarser full-domain sampling reduced runtime, although AccSWE peak RAM remained close to 20 GiB. Regional grid refinement increased computational cost markedly: the 0.01° PDM test required 969.2 s (16.2 min) and 28.15 GiB. PDM peak memory was non-monotonic across record lengths under adaptive blocking.

Practical hardware guidance

Memory requirements depend on the method, record length and analysis extent. For the larger workloads represented here, 64 GB RAM is a conservative planning recommendation, not a validated minimum. All measurements were obtained on the 128 GB reference platform. Users should reserve memory for the operating system and other applications, apply a suitable extent mask, and restrict the region or period when resources are limited.

Text S4. MSDT method comparison

Methods

To quantify the dependence of snow-drought assessments on the identification method, we compared binary drought classifications over 44 common hydrological-year start years (1980-2023). The Northern Hemisphere comparison included all seven methods. Because 1AprSWE is restricted to the Northern Hemisphere, the comparison across the global study domain included the remaining six methods. All calculations retained the original identification thresholds and used the existing 0.25-degree output grids without spatial interpolation.

Native annual classifications were retained for 1AprSWE, PeakSWEMean, PeakSWEPctl, PeakSWEDate, and AccSWE. For SWEI and PDM, an additional annual label was defined from the fraction of valid winter time classified as drought. The winter window was November-March in the Northern Hemisphere and May-September in the Southern Hemisphere. Monthly SWEI labels were weighted by the number of days in each month, whereas daily PDM labels were weighted equally. February 29 was excluded, consistently with the PDM output calendar. A year was labelled as drought when at least 50% of valid winter time was affected; thresholds of 25% and 75% were evaluated separately. This cumulative duration criterion did not require consecutive drought days or months and was used only for cross-scale comparison, not as a replacement for the native definitions.

The supplied SWEI output used three-month accumulation, empirical Gringorten probabilities, and a drought threshold of -0.5. The PDM output used a 30th-percentile SWE threshold. Mean-based annual methods used a factor of 1.0; PeakSWEPctl and PeakSWEDate used the 25th percentile, and AccSWE used the median. These native settings were not equalized to a common drought frequency. Annualization harmonized the reporting unit, not the original diagnostic windows: in particular, a three-month SWEI value for a winter month can contain antecedent months outside the selected winter.

A binary flag was considered valid only when `sd_flag` was zero or one and `intensity` was finite and unmasked. This prevented undefined index values that had been encoded as zero flags from being treated as no drought. Annualized SWEI and PDM records required at least 90% of the expected winter duration to be valid. This implies all five winter months for SWEI, and at least 136 of 151 winter days in the north or 138 of 153 in the south for PDM. Within each comparison, a fixed spatial intersection was then retained only where every method was valid in every common year. The same spatial mask was used for all three annualization thresholds.

All years were indexed by hydrological-year start. In the Northern Hemisphere, 1 April of calendar year y+1 was paired with start year y; thus, the comparison covered 1980/81-2023/24. PeakSWEDate ordinal time records were mapped to starts from 1980, following the documented interpretation of these files, and PDM years were read from the filenames. Southern Hemisphere seasons were indexed by their own start year. A global start-year label therefore combined the northern winter y/y+1 with the southern winter within y, rather than representing one concurrent calendar interval.

At each grid cell and year, F was the fraction of methods identifying drought, and classification consensus was C = max(F, 1 - F). F = 0 and F = 1 represent unanimous no-drought and drought classifications, respectively; high C can therefore describe either label. In a six-method ensemble, C = 0.5 represents a three-against-three tie. Representative map years were chosen objectively: one had an area-weighted mean minority fraction (1 - C) closest to its temporal median, and the other had the largest mean minority fraction. Pairwise agreement was evaluated using area-weighted Cohen's kappa (Cohen, 1960), ordinary agreement, and the drought Jaccard index, pooling grid-cell-years with equal weights for years and spherical grid-cell area weights in space.

Annual drought area was expressed as a percentage of the fixed valid domain. Between-method spread was summarized by the maximum-minus-minimum range and the interquartile range (IQR); neither was interpreted as a confidence interval. Sen slopes were calculated from the continuous annual area-fraction series and expressed in percentage points per decade. These slopes were descriptive, with no significance claims or treatment of the spatial grid cells as independent replicates.

Results

The fixed common domain comprised 67,752 cells (24.855 million km2) for the seven-method Northern Hemisphere comparison and 111,620 cells (48.531 million km2) for the six-method global-study-domain comparison. The smaller northern domain resulted from requiring complete valid records across all methods, particularly 1AprSWE. It retained 48.8% of the area with any valid northern output, whereas the six-method global comparison retained 93.6% of the area with any valid output. The Southern Hemisphere contributed 0.805 million km2 (1.66%) of the latter domain, so the global aggregate was strongly weighted towards the Northern Hemisphere. Results for the two comparisons cannot therefore be interpreted as a controlled hemispheric contrast.

Under the primary 50% annualization rule, pairwise kappa ranged from 0.054 to 0.852 in the Northern Hemisphere (median 0.459) and from 0.043 to 0.725 across the global study domain (median 0.377; Fig. S4-2). SWEI and PDM had the highest kappa in both comparisons, with values of 0.852 and 0.725 and drought Jaccard indices of 0.811 and 0.658, respectively. PeakSWEDate and SWEI had the lowest kappa (0.054 and 0.043). The low agreement of the peak-timing definition with SWE-amount definitions was consistent with the distinction between early snowpack timing and a deficit in snow amount, although the comparisons alone did not establish the physical causes of individual disagreements.

Spatial consensus varied within individual years. The Northern Hemisphere map years selected by the objective disagreement criterion were 2012/13 (typical) and 2002/03 (largest disagreement; Fig. S4-1). All seven methods gave the same classification over 36.3% and 22.5% of the fixed domain in these years, respectively. The corresponding global-study-domain start years were 1997 and 1994 (Fig. S4-4). These spatial-disagreement selections were based on minority votes and need not coincide with the year of the largest aggregate area spread, because geographically compensating disagreements can produce similar total areas.

The mean annual maximum-minus-minimum drought-area range was 32.6 percentage points in the Northern Hemisphere and 35.1 percentage points across the global study domain (Fig. S4-3). The corresponding mean IQRs were 23.0 and 20.6 percentage points. Both area ranges were largest for start year 2002, reaching 48.3 and 43.8 percentage points. Thus, the spatial extent inferred for an individual year depended substantially on the definition. These differences included the effects of the original thresholds: for example, the northern mean drought-area fractions were approximately 52% for mean-based SWE definitions, 50% for AccSWE, and 24-25% for the percentile-based peak amount and peak date definitions.

For the primary rule, all methods had positive descriptive Sen slopes, but their magnitudes ranged from 0.42 to 2.60 percentage points per decade in the Northern Hemisphere and from 2.13 to 3.47 across the global study domain (Fig. S4-7). Positive slope signs also persisted at both alternative annualization thresholds. The direction of change was therefore consistent within the tested definitions, retained domains, and period, whereas the estimated magnitude was method dependent. This finding was not a test of trend significance.

Annualization sensitivity was non-negligible (Fig. S4-6). At the 25%, 50%, and 75% thresholds, median pairwise kappa was 0.468, 0.459, and 0.360 for the Northern Hemisphere and 0.455, 0.377, and 0.288 for the global study domain. Mean annual area ranges were 32.5, 32.6, and 35.9 percentage points in the north and 33.1, 35.1, and 43.4 percentage points globally. The 50% rule was retained as the stated primary convention, not selected to maximize agreement. Restricting the comparison to native annual methods on the same common domains also left appreciable disagreement: median pairwise kappa was 0.462 for the five northern annual methods and 0.314 for the four global annual methods.

Discussion

These analyses show that method choice affects both the location and reported extent of snow drought. The close agreement between annualized SWEI and PDM should not be interpreted as independent validation: both depend on SWE variability, and the shared winter-duration convention also promotes comparability. Conversely, low agreement involving PeakSWEDate need not indicate poor performance, because an early seasonal maximum

is not equivalent to a low snowpack amount. Users should choose a definition that matches the timing, storage, or persistence aspect relevant to the application, and report sensitivity to plausible alternatives when drawing conclusions about drought extent.

The reported spread represents dependence on the selected definitions, their original thresholds, and the explicit temporal harmonization rule. It does not isolate algorithmic uncertainty at matched event frequencies, quantify observational uncertainty, or establish which method is most accurate. Kappa is also affected by marginal drought frequencies; the paired Jaccard indices, prevalence estimates, and complete contingency tables are therefore provided alongside it. The ensemble members are not independent or exhaustive, so vote fractions must not be interpreted as calibrated drought probabilities and ensemble spread must not be interpreted as a probabilistic uncertainty interval. The restricted common domains, particularly the 1AprSWE-limited northern domain, further constrain geographical generalization. Physical attribution of disagreements would require additional meteorological and snow-process evidence.

Reference

Cohen, J.: A coefficient of agreement for nominal scales, Educ. Psychol. Meas., 20, 37-46, https://doi.org/10.1177/001316446002000104, 1960.

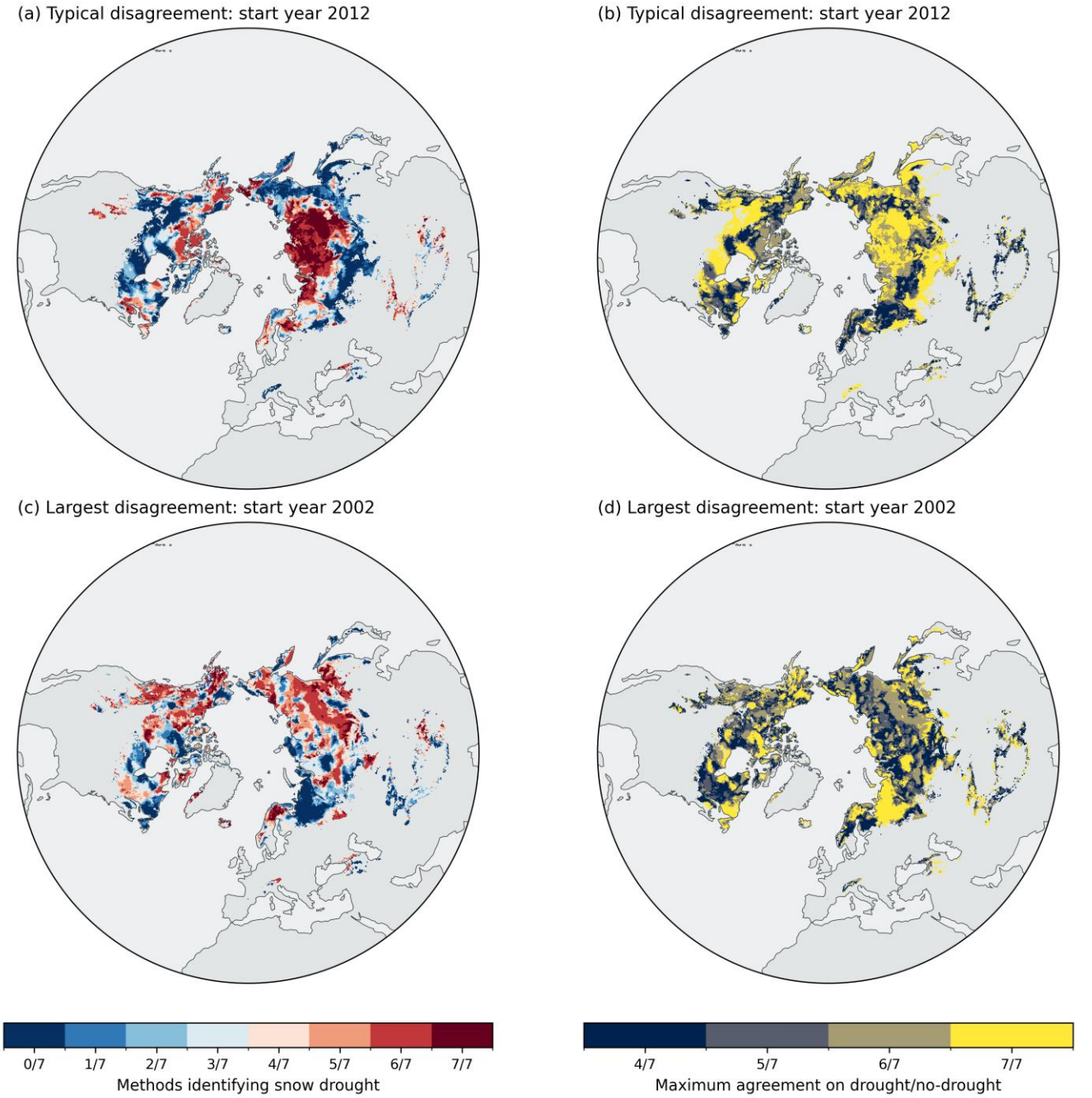


Figure S4-1. Spatial agreement among seven snow-drought methods in the Northern Hemisphere. Panels (a, b) show hydrological year 2012/13, selected as the year closest to the median area-weighted minority fraction; panels (c, d) show 2002/03, with the largest minority fraction. Left panels show F, the number of methods identifying drought divided by seven. Right panels show C = max(F, 1 - F), the maximum share assigning the same drought/no-drought label. Thus, 7/7 consensus can denote either unanimous drought or unanimous no drought; the left panels distinguish these cases. Annualized SWEI and PDM require drought over at least 50% of valid winter time. Gray areas are outside the fixed common valid domain. The comparison period is 1980/81-2023/24.

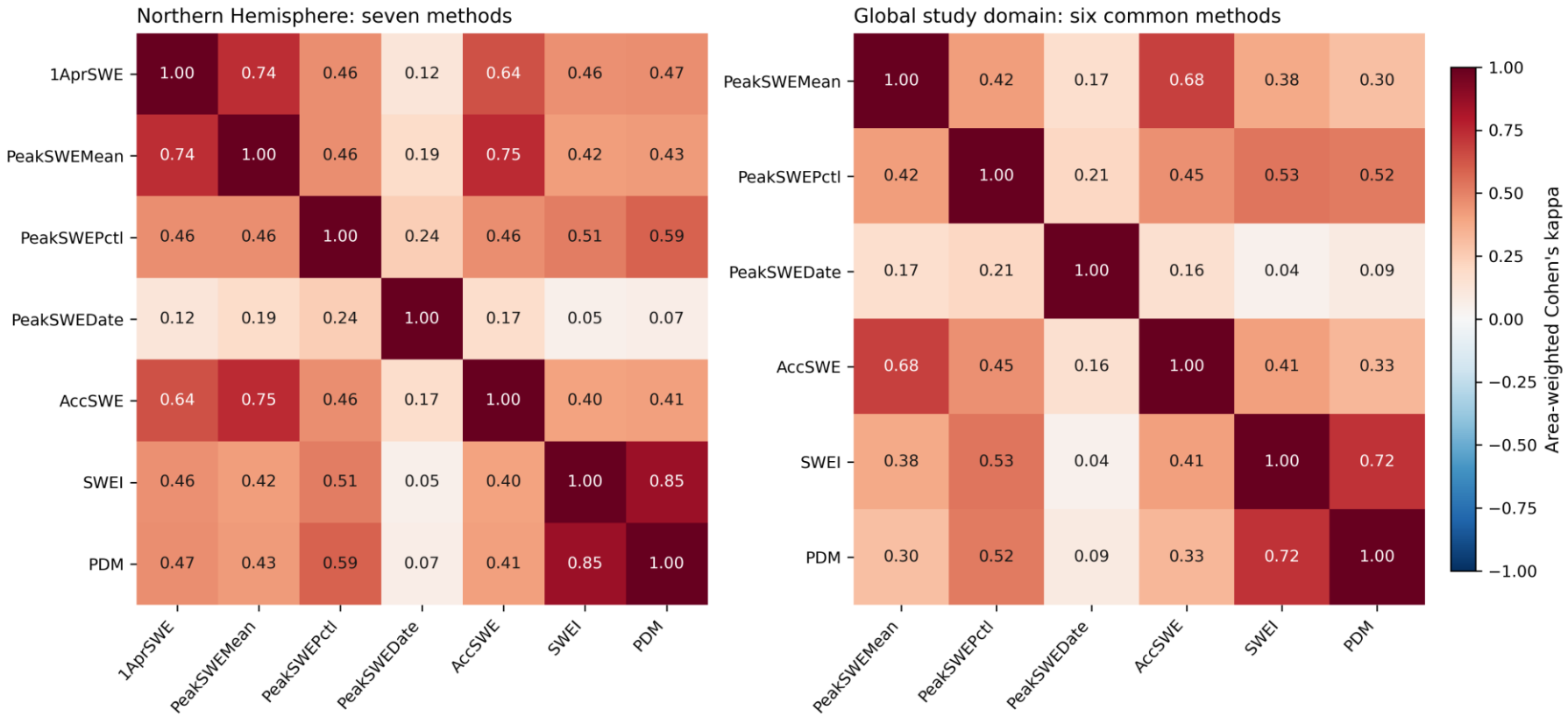


Figure S4-2. Pairwise area-weighted Cohen kappa for drought/no-drought classification in the seven-method Northern Hemisphere comparison (left) and six-method global-study-domain comparison (right). Values pool 44 hydrological-year start years (1980-2023) using spherical grid-cell area weights and equal weights for years. SWEI and PDM use the 50% winter-duration annualization criterion. Each comparison uses its own fixed common valid domain. PeakSWEPctl uses the stored 25th-percentile output. Kappa measures agreement beyond that expected from the marginal label frequencies; it does not establish accuracy against observations.

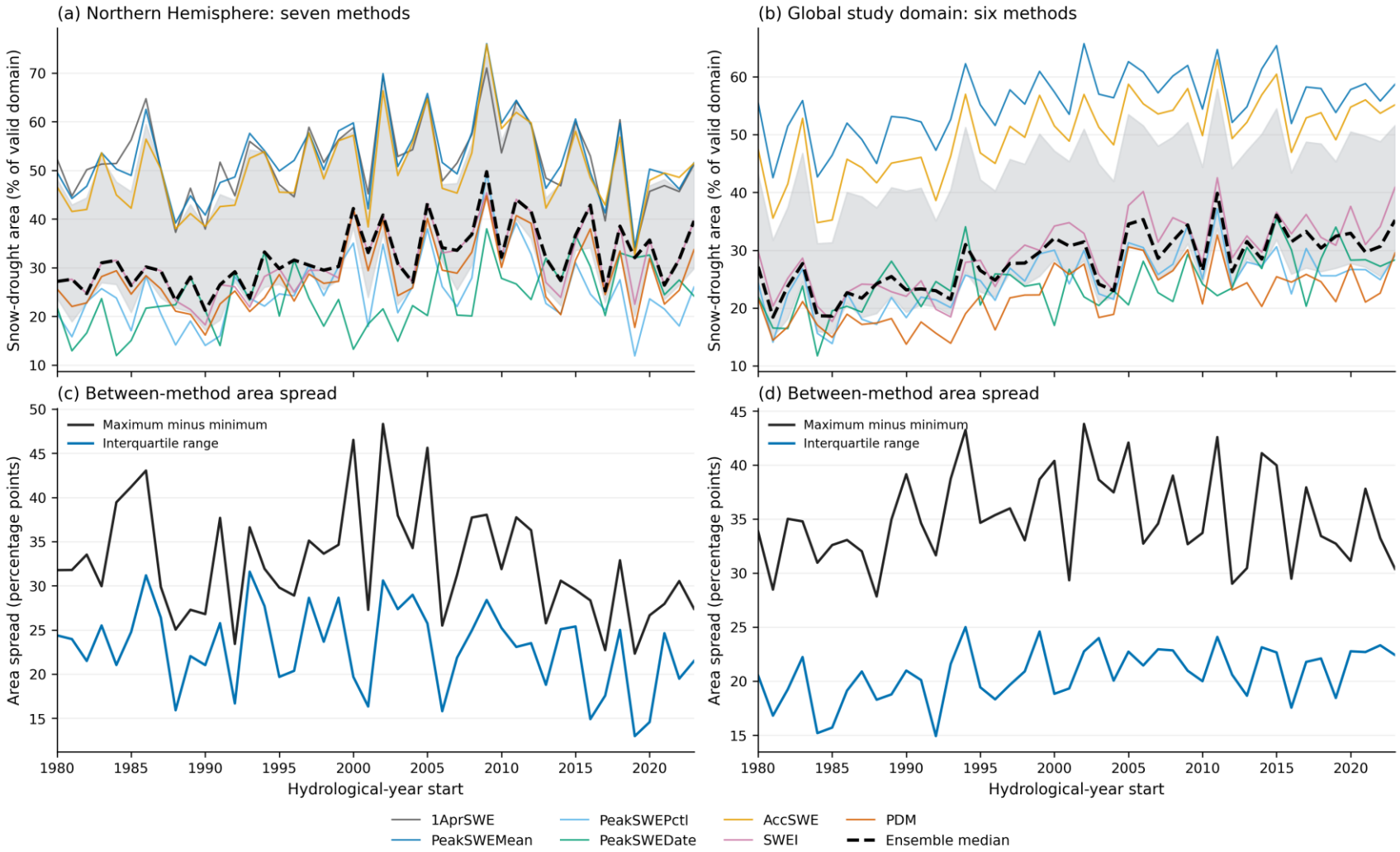


Figure S4-3. Dependence of annual drought area on the identification method. Panels (a, b) show method-specific area fractions for the Northern Hemisphere and global study domain. The dashed line and gray shading indicate the across-method median and IQR, respectively. Panels (c, d) show the maximum-minus-minimum range and IQR of area fractions in percentage points. All area fractions use a fixed valid-domain denominator. The 50% winter-duration criterion is used for SWEI and PDM. Start year y denotes northern winter y/y+1 and southern winter within y; the global values combine these hemisphere-specific seasons. Spread describes the selected definitions, not a probabilistic confidence interval.

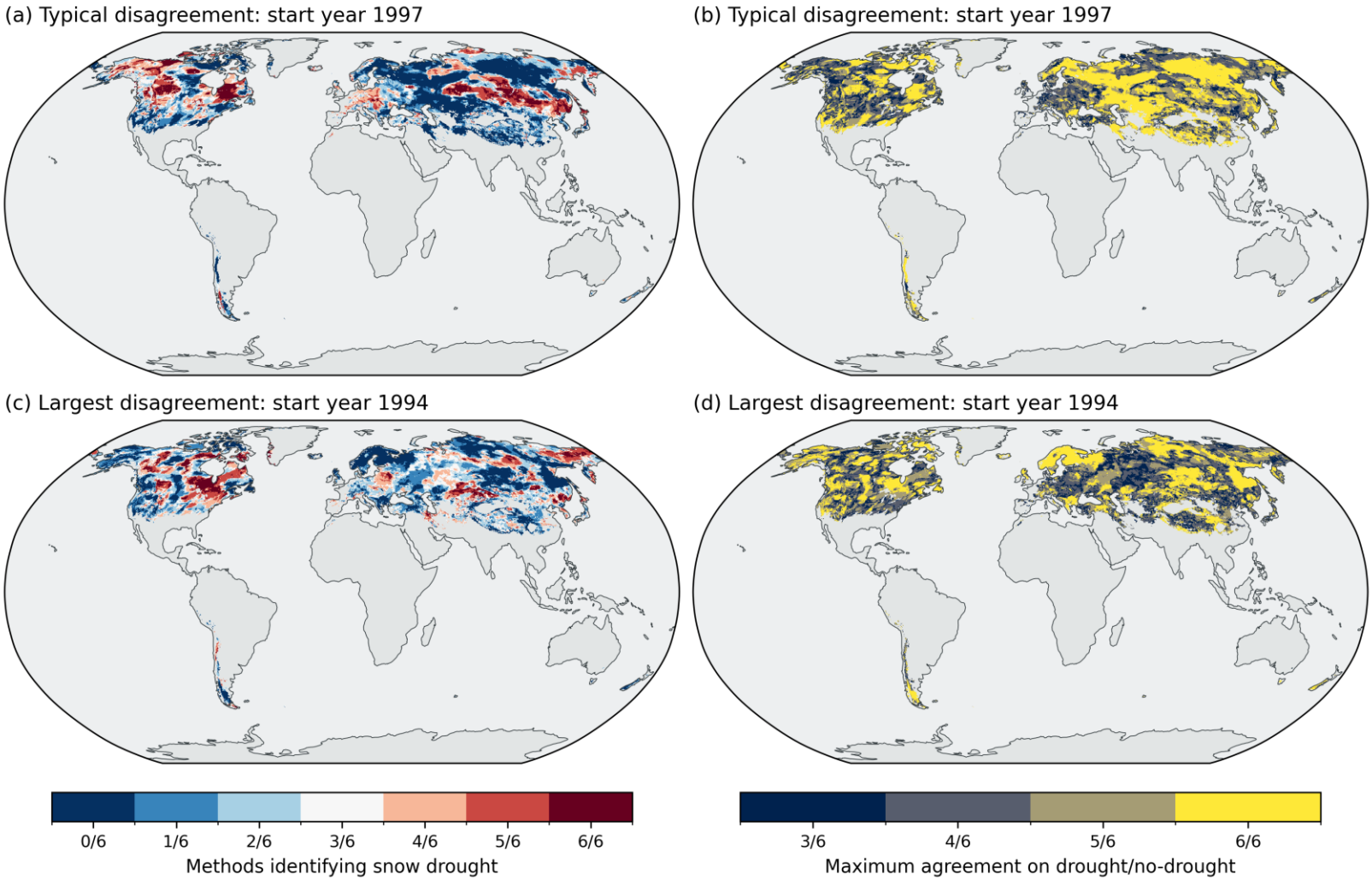


Figure S4-4. Spatial agreement among the six methods applicable across the global study domain. The typical-disagreement start year is 1997 (a, b), and the largest-disagreement start year is 1994 (c, d), selected from the area-weighted minority fraction. Left panels show the fraction identifying drought; right panels show C = max(F, 1 - F). A value of 3/6 is a tie, not a strict majority. Gray areas are outside the common valid domain, including Antarctica and regions outside the supplied masks. The annualization, temporal alignment, and data-validity rules are identical to those used in the main analysis. The domain is strongly dominated by northern snow-covered land.

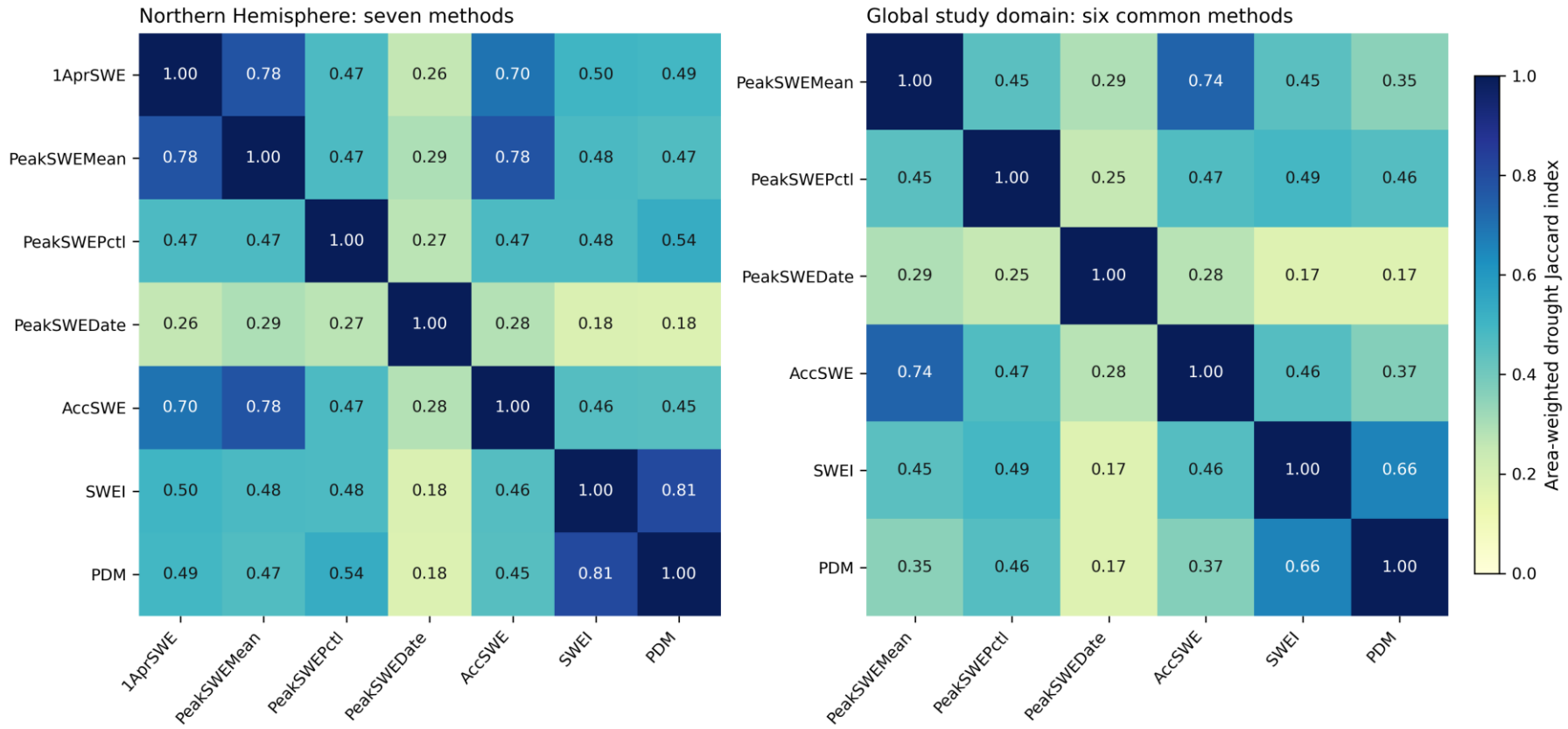


Figure S4-5. Area-weighted drought Jaccard indices for the same pairings, common domains, and primary annualization rule as Fig. F2. The index is the area-time assigned drought by both methods divided by the area-time assigned drought by at least one; joint no-drought records do not contribute. It complements kappa and ordinary agreement when drought frequencies differ. No method is treated as a validation reference.

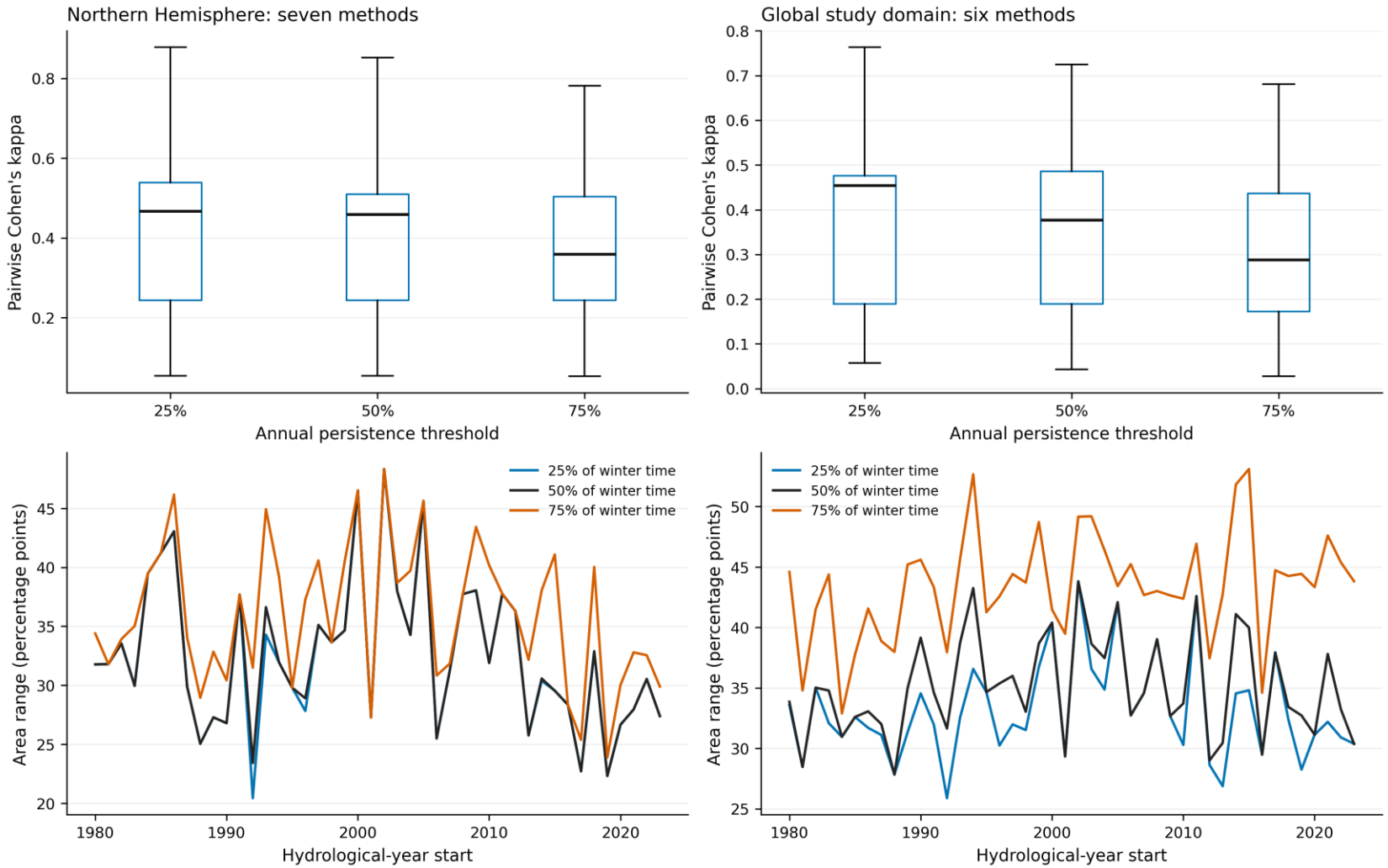


Figure S4-6. Sensitivity to the criterion for converting SWEI and PDM to annual binary labels. Upper panels summarize pooled pairwise kappa across the 21 northern and 15 global method pairs at thresholds of 25%, 50%, and 75% of valid winter time. Boxes show the across-pair IQR and median; whiskers extend to the most extreme values within 1.5 IQR and points beyond are shown individually. These distributions are not confidence intervals. Lower panels show the annual area range across methods. Native annual classifications and fixed spatial domains are unchanged across thresholds.

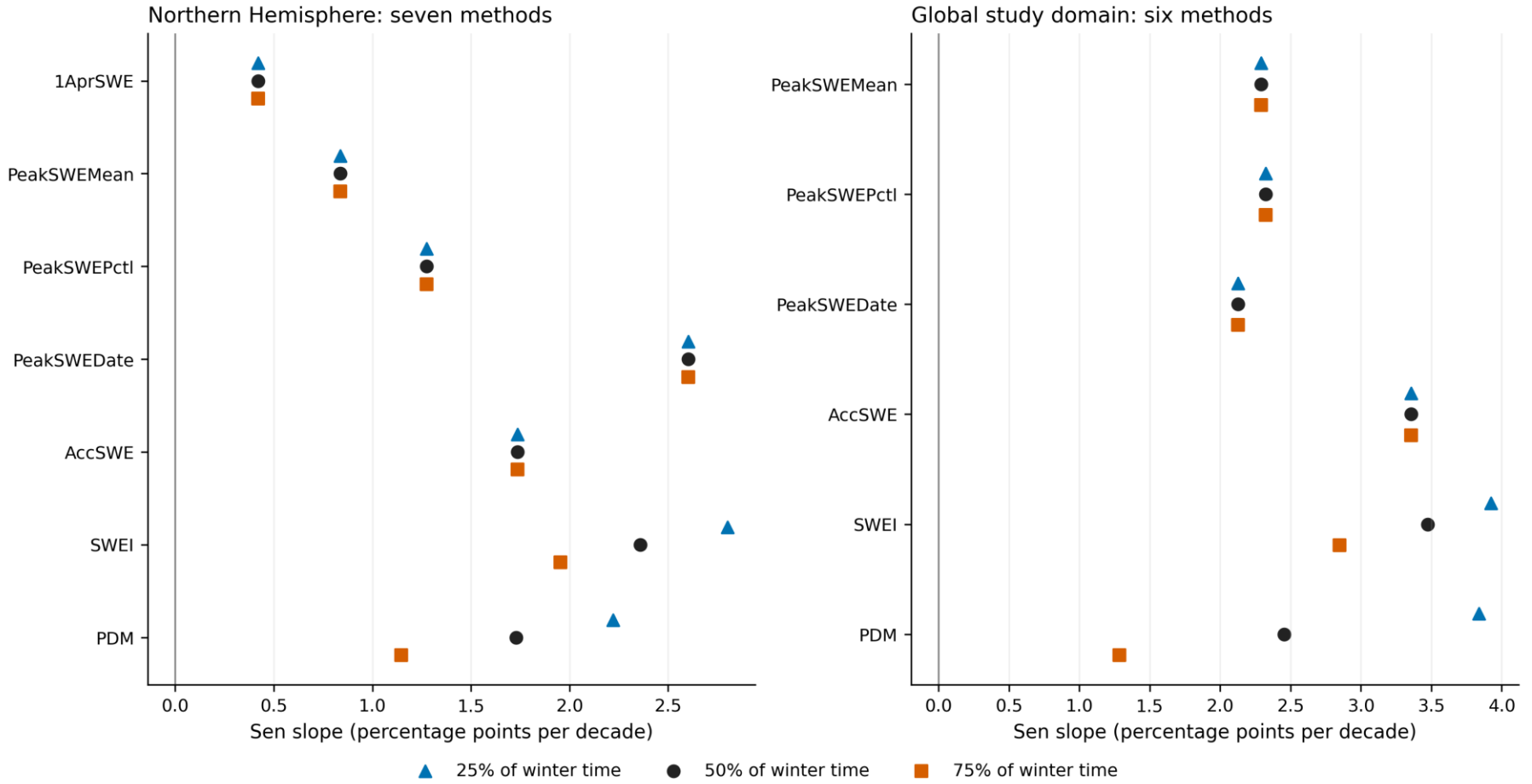


Figure S4-7. Descriptive Sen slopes of annual drought-area fractions for each method under the three SWEI/PDM annualization criteria. Units are percentage points of the fixed valid domain per decade. Native annual methods therefore have coincident estimates across the three criteria. The plotted slopes summarize continuous domain-wide annual area fractions, not individual binary grid-cell flags. No significance test or confidence interval is implied.