https://doi.org/10.5194/egusphere-2026-5105
Preprint. Discussion started: 3 November 2026
© Author(s) 2026. CC BY 4.0 License.

Assessing the vertical representativeness of ERA5 humidity for interpreting satellite upper-tropospheric humidity trends

Running title: Vertical representativeness of satellite humidity trends

Author: Zhuolun Li

Affiliation: College of New Energy and Environment, Jilin University, Changchun, China 130021

Corresponding author: Zhuolun Li, College of New Energy and Environment, Jilin University, Changchun, China 130021. Email: lizl9924@mails.jlu.edu.cn

Abstract

Long-term upper-tropospheric humidity (UTH) trends are sensitive not only to sampling and observing-system changes but also to the vertical definition used for cross-product comparison. We analysed the CM SAF version-2 microwave UTH climate record over 60°S–60°N and tested whether ERA5 relative humidity at 200–500 hPa, individually or as non-negative vertical mixtures, could represent its 1996–2018 change. Area-weighted CM SAF UTH declined by 1.238 percentage points decade^{-1} (95% moving-block-bootstrap interval: −1.472 to −1.008; AR(1)-adjusted P = 0.003), with negative slopes in 97.3% of valid grid cells. The sign persisted across seasons, latitude bands, coverage thresholds, later start years, and a platform-transition audit. ERA5 showed weak and uncertain trends at 200–300 hPa but significant moistening at 400–500 hPa. The CM SAF trend was 1.017 percentage points decade^{-1} more negative than the most negative ERA5 level and therefore lay outside the envelope of every non-negative mixture of the five pressure levels. A mixture optimized for spatial trends retained only moderate correspondence (r = 0.474), a +0.991 percentage-points-decade^{-1} bias, and 65.8% sign agreement. A separately optimized mixture failed to predict detrended annual variability under leave-one-year-out validation (r = −0.098). Thus, the satellite decline is internally robust but is not representable by

https://doi.org/10.5194/egusphere-2026-5105
Preprint. Discussion started: 3 November 2026
© Author(s) 2026. CC BY 4.0 License.

fixed-pressure ERA5 humidity or simple vertical averaging. Cross-product climate assessments require an observation-operator-aware comparison rather than selection or averaging of pressure levels.

Keywords: upper-tropospheric humidity; vertical representativeness; satellite climate data record; ERA5; trend uncertainty; observation operator

1. Introduction

Water vapour is the atmosphere's dominant natural greenhouse gas, and changes in upper-tropospheric moisture have a disproportionate effect on outgoing long-wave radiation because they occur above much of the lower-tropospheric water vapour burden (Held and Soden 2000; Soden and Held 2006). Upper-tropospheric humidity is also dynamically informative: it reflects the interaction of convective detrainment, subsidence, large-scale advection, and microphysical processes (Soden and Fu 1995; Sherwood et al. 2010). Detection studies have identified human influence on atmospheric moisture and satellite evidence of upper-tropospheric moistening, but those results do not imply that every product-defined relative-humidity record must increase (Santer et al. 2007; Chung et al. 2014). Long, spatially extensive UTH observations are therefore needed to evaluate climate models and to characterize the background against which radiative feedbacks operate.

Satellite humidity records provide the required spatial coverage but present several interpretation challenges. UTH products derived from infrared or microwave channels represent broad, humidity-dependent vertical weighting rather than a point measurement at a fixed pressure level (Buehler and John 2005; Houshangpour et al. 2005; Gettelman et al. 2006). Operational microwave records have enabled long-term UTH monitoring, but their interpretation depends on inter-satellite consistency, retrieval definition, and comparison with radiosondes and other products (Buehler et al. 2008; Moradi et al. 2010; Lang et al. 2020). Observations near the 183-GHz water-vapour absorption line are additionally affected by spectroscopy, radiative-transfer, calibration, cloud and precipitation, and representativeness uncertainties (Brogniez et al. 2016). Experience with multi-satellite microwave climate records also shows that inter-satellite offsets, orbital drift, diurnal sampling, and cross-platform recalibration must be treated explicitly during homogenization

https://doi.org/10.5194/egusphere-2026-5105
Preprint. Discussion started: 3 November 2026
© Author(s) 2026. CC BY 4.0 License.

(Zou et al. 2006; Mears and Wentz 2017). Clear-sky sampling, changes in observing systems, calibration, and retrieval assumptions can therefore affect apparent long-term behaviour (Bates and Jackson 2001; John et al. 2011). Trend estimates cannot be compared reliably unless the product definition, quality control, sampling support, spatial weighting, and statistical treatment are explicit.

The EUMETSAT Satellite Application Facility on Climate Monitoring (CM SAF) provides climate data records designed for consistent satellite-based climate analysis (Schulz et al. 2009; Hollmann et al. 2013). The version-2 UTH thematic climate data record combines microwave sounder observations into a long daily record and documents its retrieval and homogenization framework (Lang et al. 2020). Climate-data-record uncertainty should nevertheless be communicated at the level of the measured quantity and its processing chain, rather than reduced to a single generic error term (Merchant et al. 2017). A reproducible analysis is therefore needed to establish how the CM SAF trend depends on latitude, season, spatial significance, data-retention thresholds, and the selected time window. These issues are scientifically consequential: unequal grid-cell area can bias a domain mean; serial correlation can make a nominal trend test too permissive; and thousands of uncorrected grid-cell tests can produce an extensive but spurious pattern.

Here we use the CM SAF version-2 microwave UTH record to separate two questions that are often conflated in climate-data-record assessment: whether a trend is internally robust and whether it is vertically representable in another product. We first quantify the 1996–2018 domain, latitude-band, seasonal, and gridded trends under observation-count weighting, serial-correlation adjustment, false-discovery-rate control, block-bootstrap uncertainty, breakpoint selection, and a contributing-platform audit. We then compare the record with ERA5 relative humidity at 200, 250, 300, 400, and 500 hPa and introduce a constrained vertical-representability stress test. This test asks whether any non-negative mixture of the five fixed pressure levels can reproduce the CM SAF domain trend, spatial trend field, or detrended annual variability. The analysis is deliberately observational and diagnostic: it tests representability and robustness but does not attribute the trend to circulation, temperature, radiative forcing, or a specific instrumental transition.

2. Data and Methods

2.1 CM SAF upper-tropospheric humidity record

We used the Copernicus Climate Data Store collection satellite-upper-troposphere-humidity,

https://doi.org/10.5194/egusphere-2026-5105
Preprint. Discussion started: 3 November 2026
© Author(s) 2026. CC BY 4.0 License.

variable mean_uth, version 2. The retrieved archive contains 8,944 daily NetCDF files from 6 July 1994 to 31 December 2018 on a native 1° × 1° latitude–longitude grid. The product is the CM SAF Upper Tropospheric Humidity Thematic Climate Data Record based on microwave sounders (Courcoux and Schröder 2015; Lang et al. 2020), archived under DOI 10.5676/EUM_SAF_CM/UTH/V002. Although the original request covered 90°S–90°N, the primary analysis was restricted a priori to 60°S–60°N because the product metadata caution against interpreting the high-latitude values as representative UTH. This restriction yielded 120 latitude rows, 360 longitude columns, and 43,200 grid cells. The record contains 294 calendar months, beginning with a partial July 1994.

2.2 Quality control and monthly aggregation

Daily files were processed without modifying the downloaded archive. A file was retained only when its record_status indicated a valid record (status 0). At each grid cell, valid UTH values were required to be finite, to have a positive number_of_observations, and to fall within the product-provided valid range (0–109%). Daily UTH values were then aggregated to calendar-month means using number_of_observations as weights:

```
x_(m,i,j) = sum_(d in m)[n_(d,i,j) x_(d,i,j)] / sum_(d in m)[n_(d,i,j)]
```

where x is daily mean UTH, n is the corresponding observation count, d indexes the observation day, and (i, j) identifies the grid cell. A monthly grid cell was retained when at least three distinct days contributed valid observations. Across the full domain and record, 92.44% of grid-months met this criterion. Tests using thresholds of one and five days were performed to quantify sensitivity to this choice.

2.3 Domain and latitude-band time series

Spatial averages were weighted by the cosine of latitude to account for unequal grid-cell area:

```
x-bar_t = sum_(i,j)[x_(t,i,j) cos(phi_i)] / sum_(i,j)[cos(phi_i)]
```

where missing cells were omitted from the numerator and denominator. Annual means were calculated from monthly domain means and retained when at least ten months were valid. The complete annual trend sequence was 1996–2018 (23 years); 1994 and 1995 were excluded from

https://doi.org/10.5194/egusphere-2026-5105
Preprint. Discussion started: 3 November 2026
© Author(s) 2026. CC BY 4.0 License.

trend estimation because their annual coverage was incomplete.
We additionally calculated area-weighted series for four pre-defined latitude bands: 30°S–60°S, 0°–30°S, 0°–30°N, and 30°N–60°N. Monthly climatologies used all quality-controlled monthly data from July 1994 through December 2018. Seasonal uncertainty in Fig. 3B is described by interannual 10th–90th and 25th–75th percentile envelopes for each calendar month.

2.4 Trend estimation and temporal uncertainty

OLS trends were fitted to annual mean UTH as a function of calendar year. A Theil–Sen median slope was calculated as a robust alternative (Sen 1968). For the OLS hypothesis test, lag-1 correlation (r_1) of the regression residuals was used to calculate an effective sample size:

```
n_eff = n (1 - r_1) / (1 + r_1)
```

bounded between 3 and n. The two-sided trend P value was recalculated from the regression correlation using the effective sample size minus two degrees of freedom, following effective-sample-size practice for autocorrelated climate trends (Santer et al. 2000). When residual autocorrelation was negative, the effective sample size was capped at n, preventing a negative correlation from producing a more permissive test.
Trend uncertainty shown in Figs. 2–4 and 6–7 was estimated with a circular residual moving-block bootstrap, which preserves short-range dependence within resampled blocks (Künsch 1989). Residual blocks of three consecutive years were sampled with replacement, wrapped at the ends of the sequence, and added to the OLS fitted values. We used 4,000 replicates for the domain series and 3,000 replicates for latitude-band and sensitivity estimates. The 2.5th and 97.5th percentiles of the bootstrap distributions define the displayed 95% confidence intervals. These intervals describe sensitivity to temporally clustered residual variability; the reported P values remain those from the AR(1)-adjusted test.

2.5 Gridded trends and multiple testing

Calendar-month climatologies were removed at each grid cell before annual gridded anomalies were calculated. A grid-cell annual value required at least ten valid monthly anomalies. OLS trends and AR(1)-adjusted P values were computed independently at all cells with at least eight annual observations. To control the expected proportion of false discoveries across the spatial field, all valid grid-cell P values were subjected to the Benjamini–Hochberg procedure at $\alpha = 0.05$ (Benjamini and Hochberg 1995), following its

https://doi.org/10.5194/egusphere-2026-5105
Preprint. Discussion started: 3 November 2026
© Author(s) 2026. CC BY 4.0 License.

application to field significance in spatial climate analyses (Wilks 2006). Map stippling denotes only cells passing this FDR criterion.

2.6 Sensitivity analyses

Two sets of sensitivity tests were specified. First, the minimum number of valid observation days per grid-month was set to one, three, or five while all other steps were unchanged. Second, the trend start year was set to 1996, 2000, or 2005 with the end year fixed at 2018. Each sensitivity trend was evaluated using the same OLS, residual autocorrelation adjustment, and moving-block-bootstrap interval procedure as the primary domain trend.

2.7 Seasonal trend atlas and supplementary mode analysis

Seasonal means were calculated separately for December–February (DJF), March–May (MAM), June–August (JJA), and September–November (SON). December was assigned to the following seasonal year. A seasonal grid cell was retained only when all three constituent months were valid, and trends were estimated for seasonal years 1996–2018. Each season received an independent AR(1)-adjusted grid-cell analysis and Benjamini–Hochberg FDR correction. Calendar-month domain trends and season-by-latitude-band trends were also calculated as supplementary temporal diagnostics; their uncertainty used the same three-year circular moving-block bootstrap as the main analysis.

As a supplementary, non-inferential description of spatial covariance, empirical orthogonal functions (EOFs) were calculated from 1996–2018 annual calendar-month anomaly fields. Grid cells required at least 20 of 23 valid annual anomalies, retaining 40,016 cells (92.6% of the domain). Missing annual values within retained cells represented 5.9% of the EOF input matrix and were linearly interpolated in time. The anomaly matrix was weighted by the square root of cosine latitude before singular-value decomposition. Because EOF sampling and mode separation are sensitive to record length (North et al. 1982), and the principal components can have strong serial persistence, their trends were evaluated with the same AR(1) adjustment and were not treated as independent evidence of the domain trend.

2.8 Change-point and observing-platform audit

We tested whether the domain and latitude-band records were better represented by a structural change than by one linear trend. Candidate breakpoints were scanned across all admissible years while retaining at least five annual observations on each side. Three

https://doi.org/10.5194/egusphere-2026-5105
Preprint. Discussion started: 3 November 2026
© Author(s) 2026. CC BY 4.0 License.

alternatives were fitted: a level shift with a common slope, a continuous slope change, and a combined level-and-slope change. Model support was evaluated relative to the single-linear-trend null using corrected Akaike information criterion (AICc) and Bayesian information criterion (BIC). Because selection of the best breakpoint inflates nominal significance, improvement in residual sum of squares was tested against 5,000 null simulations generated by circular three-year moving-block resampling of residuals from the single linear model. Monthly calendar-month anomalies were analysed analogously, with at least 60 months on each side and 12-month residual blocks. Pettitt tests (Pettitt 1979) were additionally applied to residuals after removal of the single linear trend to screen specifically for abrupt distributional shifts. The interpretation of candidate segmented models follows the broader structural-change framework of Bai and Perron (1998).
The daily files contain a bit-encoded satellites_contributing_to_daily_mean variable. We decoded all 18 documented satellite bits and calculated, for each calendar month, the fraction of valid 60°S–60°N satellite-code grid-cell-days containing each platform. Fractions can sum to more than one because multiple satellites may contribute to the same grid-cell daily mean. The resulting timeline was compared descriptively with statistical breakpoint candidates; temporal coincidence was not treated as proof of instrumental causation.

2.9 ERA5 cross-product comparison

We compared the satellite record with monthly averaged ERA5 relative humidity at 200, 250, 300, 400, and 500 hPa for January 1996–December 2018 (Hersbach et al. 2020). ERA5 was obtained on its 0.25° latitude–longitude grid and sampled at the 1° cell centres of the CM SAF domain. For each pressure level, monthly fields were averaged to annual values and then processed using the same 60°S–60°N domain, cosine-latitude area weights, linear trend estimator, AR(1) adjustment, and three-year moving-block-bootstrap interval as the satellite domain series. We reported every selected level rather than choosing a level after inspecting agreement.
Cross-product agreement was evaluated at three scales. Domain trends and their uncertainty tested agreement in the sign and magnitude of long-term change. Pearson correlations were calculated for both raw annual domain series and residuals after independently removing each series' linear trend; the latter isolate year-to-year co-variability from the opposing influence of long-term slopes. Spatial agreement was assessed from grid-cell trend

https://doi.org/10.5194/egusphere-2026-5105
Preprint. Discussion started: 3 November 2026
© Author(s) 2026. CC BY 4.0 License.

correlations, trend-sign agreement, and zonal-mean profiles on the common 1° grid. A Taylor diagram summarized detrended temporal correlation and the ratio of residual standard deviations (Taylor 2001). Because satellite UTH has a broad, humidity-dependent weighting function, ERA5 fixed-pressure relative humidity was treated as a set of vertical proxies, not as an observation-operator-equivalent reference or ground truth.

2.10 Vertical-representability stress test

Fixed-pressure ERA5 fields cannot reproduce the broad, scene-dependent microwave weighting function, but they can test whether the satellite change is contained within the span of plausible pressure-level behaviour. We therefore fitted non-negative weights for the five ERA5 levels under the constraint that the weights summed to one. This convex constraint preserves the physical interpretation of an average of supplied pressure-level relative humidities and prevents cancellation through negative coefficients. Three properties were tested. First, the domain-trend envelope was defined by the minimum and maximum trends among the five levels; every convex mixture must lie within that interval. Second, weights were optimized by least squares against the coincident CM SAF grid-cell trend field, after which spatial correlation, root-mean-square error, mean bias, and trend-sign agreement were calculated. Third, independent weights were fitted to linearly detrended annual domain series. Temporal generalization was evaluated by leave-one-year-out fitting, so each predicted year was excluded when its weights were estimated. The spatial and temporal optimizations were kept separate to test whether one vertical representation could satisfy both objectives. This procedure is a representativeness stress test rather than a microwave radiative-transfer observation operator; failure indicates that fixed-level selection or simple vertical averaging is inadequate, not which product is closer to the unobserved atmospheric truth.

2.11 Spatial machine-learning diagnostic

To test whether a flexible spatial predictor could recover information not captured by fixed-pressure mixtures, we fitted a quadratic ridge-regression model to the 43,200 coincident grid cells. Predictors were the five ERA5 pressure-level trends, latitude, longitude, squared latitude and longitude, and smooth longitude and latitude–longitude interaction terms. The target was the CM SAF grid-cell trend. The regularisation strength was fixed at $\alpha = 10$; no target-derived tuning was performed. Generalisation was evaluated with six contiguous 60° longitude blocks, each held out in turn. This design prevents neighbouring longitudes from appearing in both training and test sets and therefore provides a stricter spatial test than random cross-validation. The model is used only as a

https://doi.org/10.5194/egusphere-2026-5105
Preprint. Discussion started: 3 November 2026
© Author(s) 2026. CC BY 4.0 License.

representativeness diagnostic; predictive success does not establish a physical observation operator or causal mechanism.

2.12 Reproducibility and software

The workflow was implemented in Python using netCDF4 and xarray for data access, NumPy and pandas for array and table operations, SciPy for statistical estimation, Matplotlib for plotting, and Cartopy for map projection. The processing scripts write annual series, latitude-band summaries, gridded trend fields, breakpoint scans, platform contributions, ERA5 comparison metrics, coverage diagnostics, and sensitivity results to machine-readable CSV, NetCDF, and JSON files. Publication figures are exported as 600-dpi PNG and TIFF files and as vector PDF and SVG files. No spatial interpolation was applied to the CM SAF trend field; ERA5 was sampled only at coincident 1° cell centres, and no missing statistics were imputed. OpenAI Codex was used to assist with code drafting, debugging, figure-layout implementation, manuscript-language revision, and submission-file preparation. All analysis choices, source data, numerical results, figures, citations, and final text were checked and approved by the author, who takes full responsibility for the work.

3. Results

3.1 The decline is spatially coherent across the analysis domain

The primary spatial result was a domain-wide tendency toward lower UTH. Negative slopes occurred in 42,038 of 43,200 valid cells (97.3%). After residual-autocorrelation adjustment and Benjamini–Hochberg FDR control, 18,918 cells (43.8% of the analysed grid) remained significant, and every significant cell had a negative slope. The zonal profile remained negative throughout 60°S–60°N. The composition plot in Fig. 1C further shows that the latitude bands differed mainly in the severity distribution of drying, not in the sign of change. The agreement between the map, zonal summary, and severity composition indicates that the domain trend was not produced by a small number of spatially extreme regions.

https://doi.org/10.5194/egusphere-2026-5105
Preprint. Discussion started: 3 November 2026
© Author(s) 2026. CC BY 4.0 License.

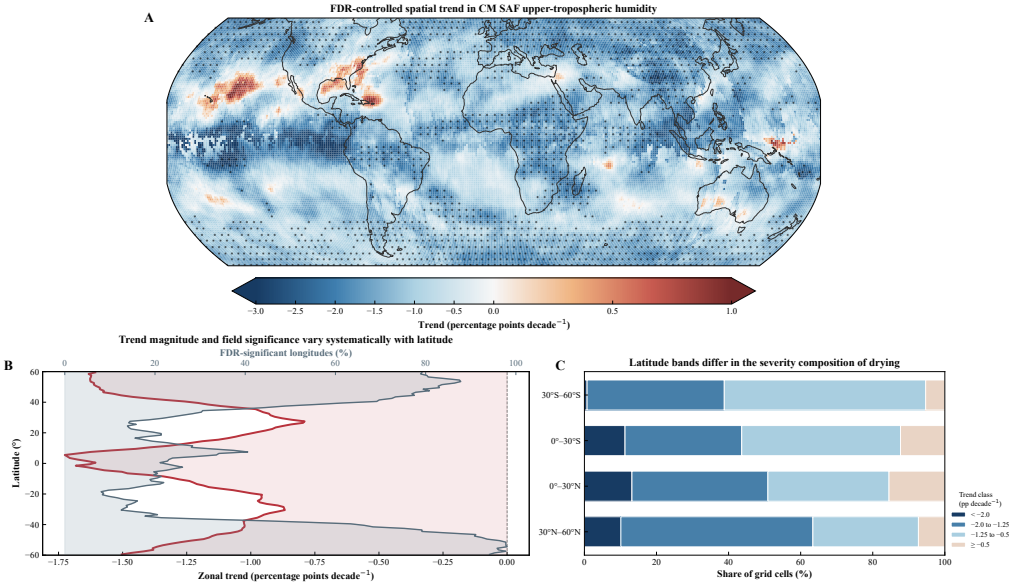


Fig. 1 Spatial coherence and severity composition of CM SAF upper-tropospheric humidity trends during 1996–2018. a Grid-cell OLS trends in annual UTH anomalies over 60°S–60°N; black stippling marks cells passing the Benjamini–Hochberg false-discovery-rate threshold at α = 0.05 after residual lag-1-autocorrelation adjustment. b Zonal-mean trend, shown with the percentage of longitudes passing the FDR threshold on the upper axis. c Composition of grid-cell trend classes within four latitude bands; the stacked bars give the percentages of cells in four trend intervals. CM SAF data credit: © EUMETSAT 2023.

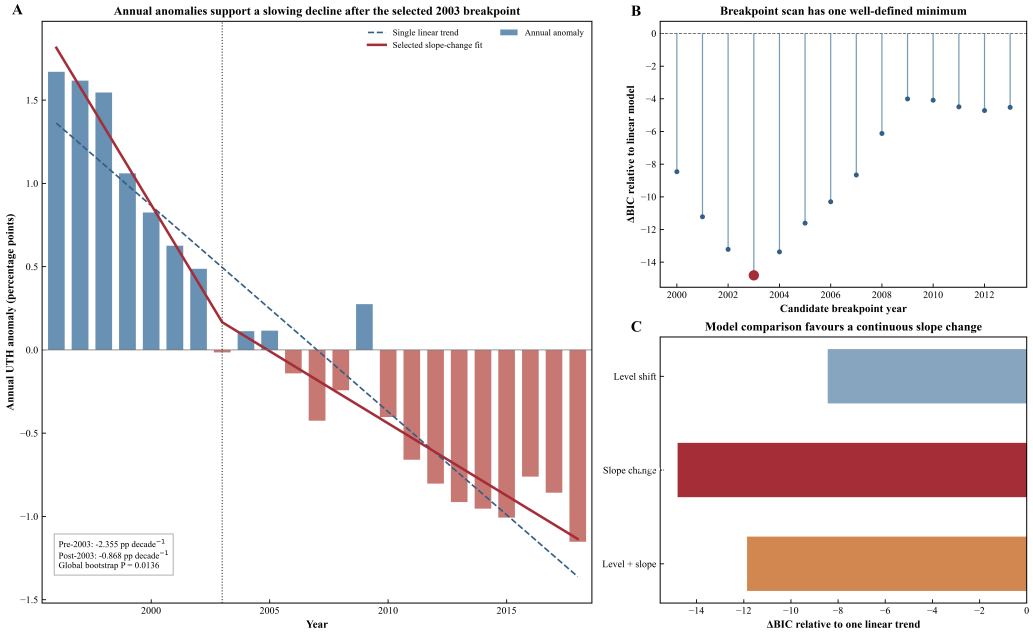
3.2 The long-term decline slowed after a selected 2003 slope change

Area-weighted annual mean UTH decreased from 37.750% in 1996 to 34.926% in 2018, a net change of −2.824 percentage points. The OLS slope was −1.238 percentage points decade^{-1} (95% moving-block-bootstrap CI: −1.472 to −1.008), with residual lag-1 correlation of 0.574, AR(1)-adjusted P = 0.003, and R^2 = 0.898. The Theil–Sen estimate was also −1.238 percentage points decade^{-1}, indicating that the result was not controlled by one or two extreme years.

A full candidate search favoured a continuous slope change near 2003 over a single linear trend (ΔBIC = −14.80; global three-year block-bootstrap P = 0.0136). The fitted decline was −2.355 percentage points decade^{-1} before 2003 and −0.868 percentage points decade^{-1} afterwards (Fig. 2). An abrupt annual level shift was not supported, and a Pettitt test on detrended annual residuals gave P = 0.365. The evidence therefore supports a change in rate

https://doi.org/10.5194/egusphere-2026-5105
Preprint. Discussion started: 3 November 2026
© Author(s) 2026. CC BY 4.0 License.

282 rather than a sudden discontinuity.



283 **Fig. 2 Temporal evolution and change-point evidence for the area-weighted CM SAF UTH**
284 **record. a Annual domain anomalies for 1996–2018, the single linear trend, and the selected**
285 **continuous slope-change fit with a 2003 candidate breakpoint. The annotation reports fitted**
286 **slopes before and after the breakpoint and the global moving-block-bootstrap P value. b BIC**
287 **scan across all admissible breakpoint years, expressed relative to the single-linear-trend**
288 **model; the selected 2003 breakpoint is the minimum. CM SAF data credit: © EUMETSAT 2023.**

289 **Table 1. Latitude-band trend estimates for 1996–2018**

Latitude band	OLS trend (percentage points decade⁻¹)	95% block-bootstrap CI	Theil–Sen trend (percentage points decade⁻¹)	95% Theil–Sen CI	AR(1)-adjusted P
30°S–60°S	−1.380	−1.756 to −0.986	−1.210	−1.622 to −0.858	<0.001
0°–30°S	−0.744	−1.197 to −0.310	−0.921	−1.282 to −0.579	0.029
0°–30°N	−0.707	−1.199 to −0.259	−0.983	−1.230 to −0.756	0.025
30°N–60°N	−1.158	−1.400 to −0.928	−1.140	−1.467 to −0.846	<0.001

290 **3.3 Seasonal structure changes the magnitude but not the sign of the result**

291 The phase–amplitude plot in Fig. 3B shows that the latitude bands occupied distinct

https://doi.org/10.5194/egusphere-2026-5105
Preprint. Discussion started: 3 November 2026
© Author(s) 2026. CC BY 4.0 License.

combinations of seasonal-maximum month and annual amplitude, while point size indicates their mean UTH level. Despite this seasonal structure, trends were negative in all latitude bands and all four seasons. Negative grid-cell slopes occurred in 88.0% of valid DJF cells, 92.7% of MAM cells, 93.1% of JJA cells, and 88.4% of SON cells. The latitude-band bootstrap intervals showed that several tropical seasonal estimates were less precise, but the dominant direction of change remained negative (Fig. 3C). Full seasonal trend maps are retained in Online Resource 1 (Fig. S2) because they resolve within-season geography without repeating multiple maps in the main article.

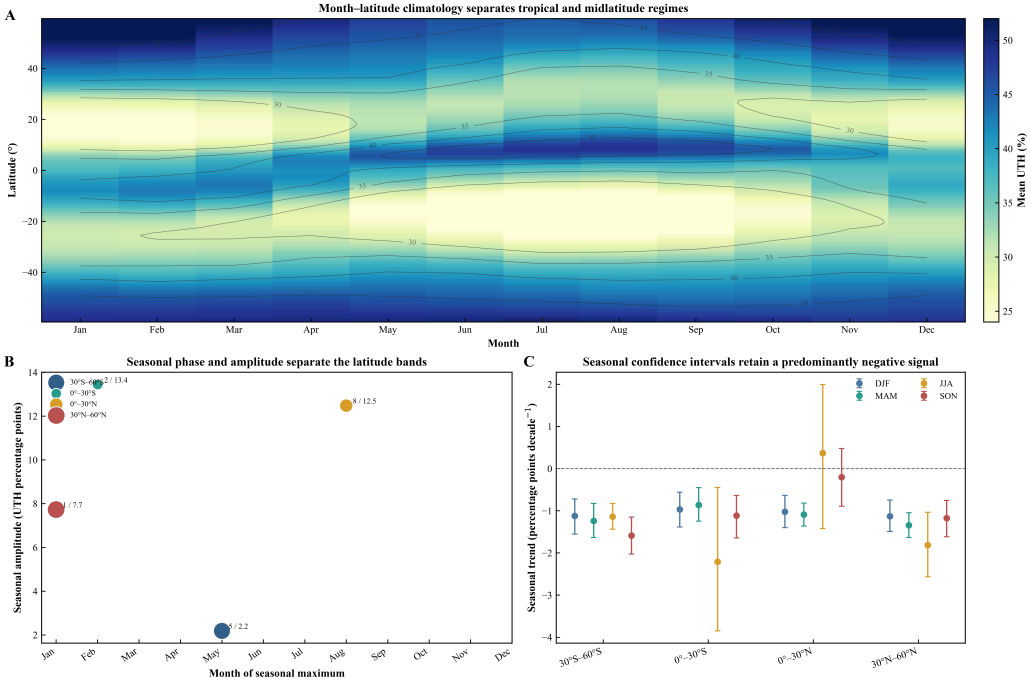


Fig. 3 Seasonal and latitudinal organization of CM SAF UTH. a Month–latitude climatology of zonal-mean UTH. b Phase–amplitude scatter plot for four latitude bands; x is the month of maximum mean UTH, y is seasonal max–min amplitude, point size scales with mean UTH, and colour identifies latitude band. c Seasonal OLS trends with 95% circular moving-block-bootstrap intervals. CM SAF data credit: © EUMETSAT 2023.

3.4 Sampling and analysis choices do not explain the negative sign

The three-day monthly rule retained 11,741,182 of 12,700,800 possible grid-months (92.44%). Coverage was lower in the earliest years but exceeded 95% across nearly the entire domain

https://doi.org/10.5194/egusphere-2026-5105
Preprint. Discussion started: 3 November 2026
© Author(s) 2026. CC BY 4.0 License.

after approximately 1999 (Fig. 4A). The paired-period comparison in Fig. 4B shows that this gain in support occurred across all latitude bands. Changing the minimum number of valid observation days per grid-month from one to three to five preserved negative domain trends of −1.205, −1.238, and −1.780 percentage points decade^{-1}, respectively; all 95% bootstrap intervals remained below zero. Advancing the start year to 2000 or 2005 reduced the fitted magnitude to −0.991 and −0.892 percentage points decade^{-1} but did not change the sign or statistical detectability (Fig. 4C). The primary estimate therefore did not depend on one reasonable coverage threshold or on inclusion of the earliest complete years.

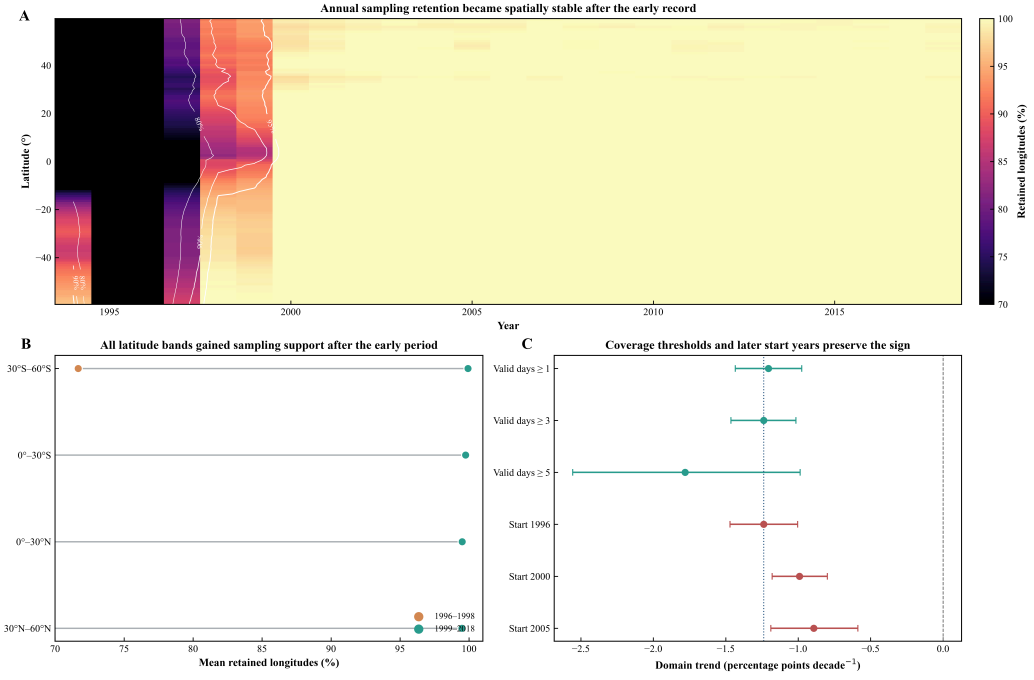


Fig. 4 Sampling support and robustness of the CM SAF domain trend. a Annual percentage of longitudes retained at each latitude under the three-valid-day monthly rule; contours identify 80%, 90%, and 95% retention. b Mean retained-longitude percentages for 1996–1998 and 1999–2018 in each latitude band; connecting segments show the gain in support. c Domain trends under one-, three-, and five-valid-day monthly thresholds and under 1996, 2000, and 2005 start years with 2018 fixed as the end year. Horizontal intervals are 95% circular moving-block-bootstrap intervals; the dashed line marks zero and the dotted line marks the primary estimate. CM SAF data credit: © EUMETSAT 2023.

3.5 The selected rate change coincides with observing-system evolution

The composition of the contributing microwave constellation changed substantially during

https://doi.org/10.5194/egusphere-2026-5105
Preprint. Discussion started: 3 November 2026
© Author(s) 2026. CC BY 4.0 License.

the record. The weighted trajectories in Fig. 5A show discrete observing regimes, with line thickness proportional to the monthly contribution fraction, while the phase-average bubble matrix in Fig. 5B identifies the dominant contributors in 1994–2002, 2003–2010, and 2011–2018. Several satellites contributed simultaneously, and platform fractions can therefore sum to more than one. The timing is relevant to the uncertainty audit but is not evidence that a particular platform caused the fitted rate change. Together with the absence of a supported abrupt level shift, the result motivates observation-operator and calibration diagnostics rather than a post hoc instrumental correction.

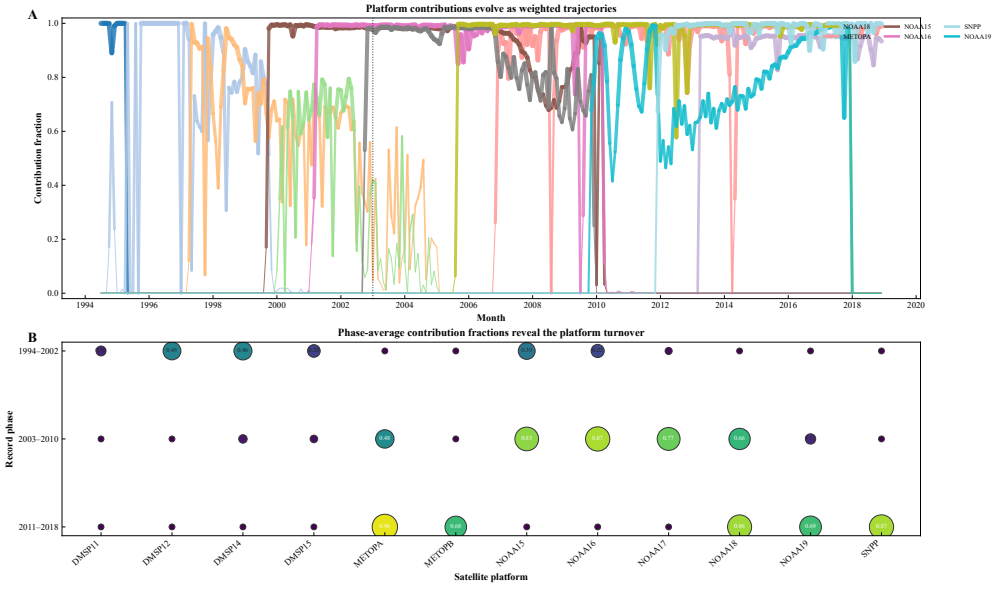


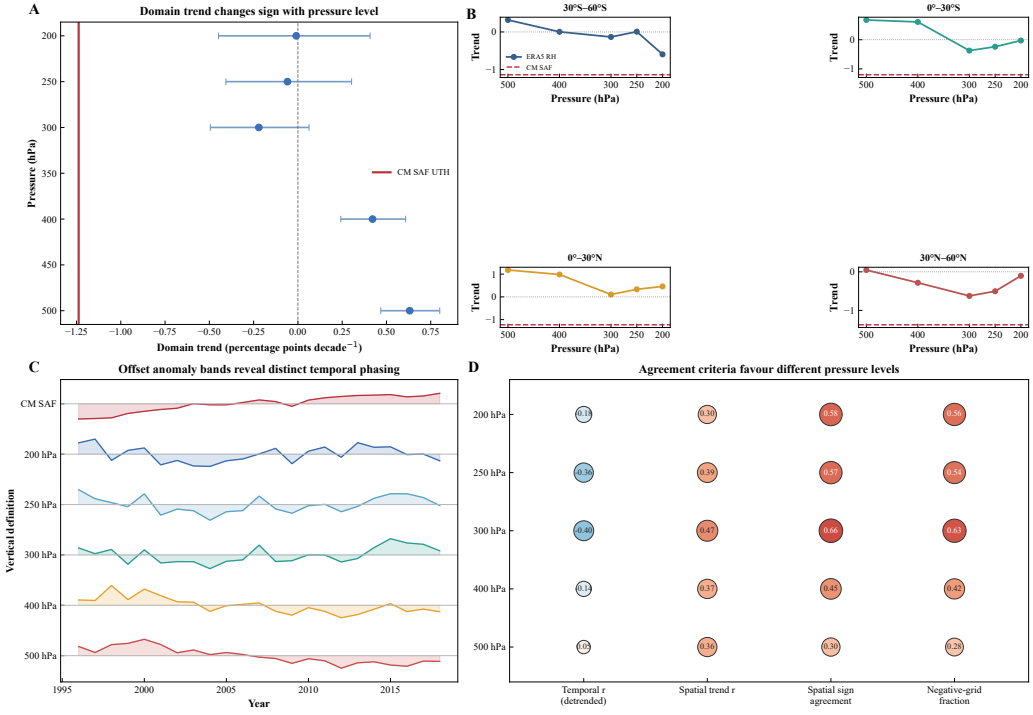
Fig. 5 Observing-platform audit of the selected temporal rate change. a Monthly platform-contribution trajectories; line thickness is proportional to contribution fraction and the vertical line marks 2003. b Phase-average platform contributions for three record phases; bubble area and colour encode mean contribution fraction. Fractions can sum to more than one because platforms may contribute simultaneously. CM SAF data credit: © EUMETSAT 2023.

3.6 ERA5 cross-product behaviour depends strongly on pressure level

ERA5 trends at 200, 250, and 300 hPa were −0.010, −0.059, and −0.221 percentage points decade^{-1}, with intervals spanning zero, whereas 400 and 500 hPa showed significant increases of 0.421 and 0.631 percentage points decade^{-1}. The domain-level sign reversal is shown in Fig. 6A. The regional pressure summaries in Fig. 6B shows that the pressure dependence of the trend varies across latitude bands, while the offset anomaly bands in Fig. 6C demonstrates that annual evolution is not a simple rescaling of CM SAF. The strongest individual-level spatial correspondence occurred at 300 hPa (grid-cell trend r =

https://doi.org/10.5194/egusphere-2026-5105
Preprint. Discussion started: 3 November 2026
© Author(s) 2026. CC BY 4.0 License.

348 0.474; sign agreement = 65.6%), but no pressure level simultaneously matched the domain
349 trend, spatial pattern, and interannual evolution (Fig. 6B).



350 **Fig. 6 Pressure-level dependence of the CM SAF–ERA5 comparison. a Domain trends and**
351 **bootstrap intervals. b Regional pressure-level summaries for the four**
352 **latitude bands, with dashed CM SAF band means. c Offset anomaly bands for CM SAF and ERA5**
353 **levels. d Bubble comparison of temporal correlation, spatial trend correlation, spatial**
354 **sign agreement, and negative-grid fraction. CM SAF data credit: © EUMETSAT 2023; ERA5 data**
355 **credit: Copernicus Climate Change Service/ECMWF.**

356 **3.7 No non-negative ERA5 mixture represents all dimensions of CM SAF change**

357 The convex-mixture test strengthened the individual-level comparison. The most negative
358 ERA5 trend was −0.221 percentage points decade⁻¹ at 300 hPa, so the CM SAF trend of
−1.238
359 percentage points decade⁻¹ lay 1.017 percentage points decade⁻¹ below the lower bound
360 attainable by any non-negative mixture of the five levels (Fig. 7A). Vertical averaging
361 therefore cannot recover the domain magnitude without extrapolating beyond the supplied
362 pressure-level trends.
363 Optimizing weights for the spatial field assigned 93.4% to 300 hPa and 6.6% to 200 hPa.
364 This mixture retained a correlation of 0.474 but overestimated the CM SAF trend by 0.991
365 percentage points decade⁻¹ on average, had an RMSE of 1.324 percentage points decade⁻¹,
and

https://doi.org/10.5194/egusphere-2026-5105
Preprint. Discussion started: 3 November 2026
© Author(s) 2026. CC BY 4.0 License.

matched the trend sign in only 65.8% of coincident cells (Fig. 7B). The paired weight comparison in Fig. 7C shows that the spatial and temporal objectives selected incompatible vertical representations: the spatial fit concentrated at 300 hPa, whereas the leave-one-year-out temporal fit concentrated at 500 hPa. The year-specific residual bars in Fig. 7D show the year-by-year disagreement directly; the cross-validated correlation was −0.098 and sign agreement was 43.5%. Thus, spatial pattern, domain trend, and interannual variability cannot be reconciled by selecting or averaging the tested ERA5 pressure levels.

https://doi.org/10.5194/egusphere-2026-5105
Preprint. Discussion started: 3 November 2026
© Author(s) 2026. CC BY 4.0 License.

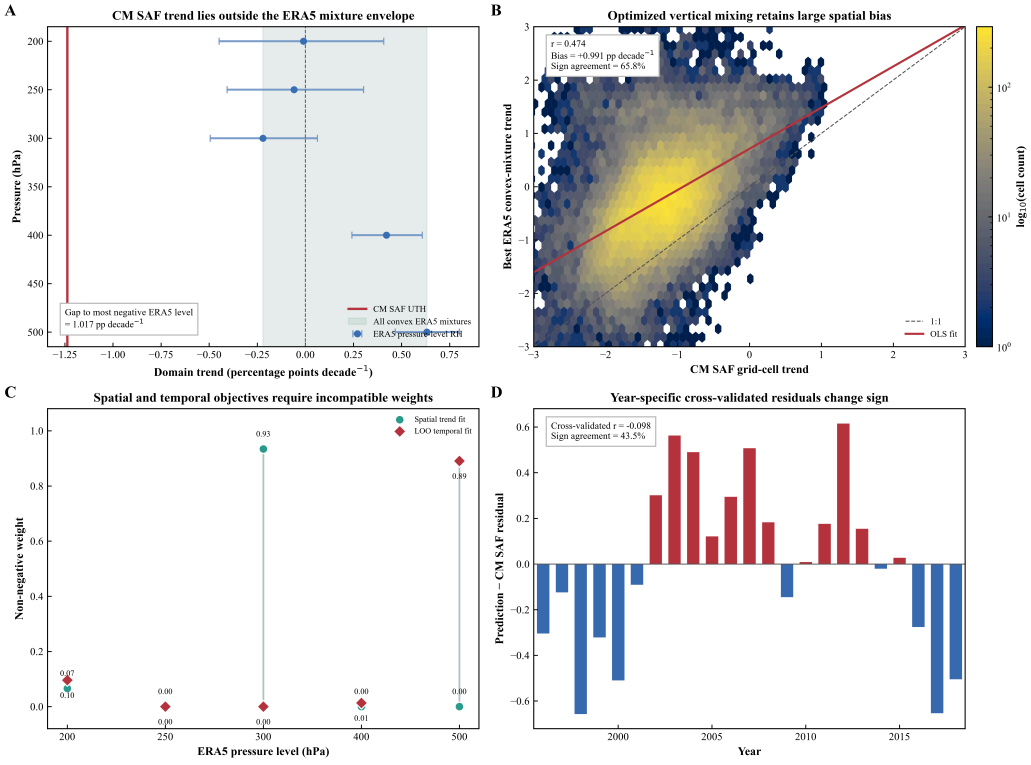


Fig. 7 Vertical-representativeness stress test. a ERA5 domain trends and attainable convex-mixture envelope. b CM SAF grid-cell trends versus the spatially optimized ERA5 mixture. c Paired spatial-fit and leave-one-year-out temporal-fit weights connected by vertical dumbbells. d Year-specific prediction-minus-CM SAF residual bars for the leave-one-year-out temporal mixture. The procedure is a representativeness stress test rather than a microwave radiative-transfer observation operator.

3.8 A spatial machine-learning diagnostic improves pattern prediction but not physical representability

The spatial-block machine-learning diagnostic provided a deliberately flexible benchmark. A quadratic ridge model using the five ERA5 pressure-level trends and geographic terms achieved r = 0.355, R^2 = 0.042, RMSE = 0.591 percentage points $decade^{-1}$, and 97.2% sign agreement across six held-out longitude blocks. The low explained variance shows that allowing smooth nonlinear spatial structure did not recover the CM SAF trend magnitude or detailed pattern with high fidelity. The result complements the convex-mixture analysis: a flexible predictor can reproduce the dominant negative sign in most cells, but it does not

https://doi.org/10.5194/egusphere-2026-5105
Preprint. Discussion started: 3 November 2026
© Author(s) 2026. CC BY 4.0 License.

constitute a validated microwave observation operator and does not remove the cross-product representativeness gap (Online Resource 1, Fig. S6).

4. Discussion

4.1 A coherent product-level decline

The central result is a spatially extensive decline in CM SAF mean UTH between 60°S and 60°N during 1996–2018. The evidence is internally consistent across four levels of analysis: the area-weighted domain mean, robust regression, fixed latitude bands, and the FDR-controlled gridded field. The decline was not limited to a small number of grid cells or one latitude zone: 97.3% of valid cells had negative slopes, every FDR-significant cell declined, and all four latitude-band estimates were negative. Coverage-threshold and trend-window tests further show that the result is not an artefact of a single defensible processing choice.

Earlier satellite and reanalysis studies have reported both long-term UTH variability and substantial sensitivity to observing system and analysis method (Bates and Jackson 2001; Dessler and Davis 2010; John et al. 2011). Satellite analyses have also detected increases in upper-tropospheric water vapour, particularly when expressed as specific or absolute moisture rather than the product-defined relative-humidity quantity analysed here (Chung et al. 2014). That literature is not contradicted directly by the present decline because records differ in humidity variable, spectral channel, cloud screening, homogenization, vertical sensitivity, temporal span, and spatial weighting. The present result should therefore be interpreted first as a well-defined change within one microwave-based climate data record. Its value is the explicit linkage between product-level quality flags, observation-count weighting, area weighting, autocorrelation adjustment, FDR control, and robustness tests.

4.2 Latitude and season constrain interpretation

The climatological and seasonal results reveal strong hemispheric asymmetry. The tropical bands had the largest seasonal amplitudes, but the strongest long-term OLS decline occurred in the southern midlatitudes. This separation between seasonal amplitude and trend magnitude indicates that the long-term domain change cannot be inferred from the seasonal cycle alone. It also argues against interpreting the area-weighted decline as a uniform multiplicative shift across all regimes.

https://doi.org/10.5194/egusphere-2026-5105
Preprint. Discussion started: 3 November 2026
© Author(s) 2026. CC BY 4.0 License.

The seasonal trend atlas adds a second distinction: the decline was present throughout the calendar year, yet FDR-significant spatial coverage was approximately 45% greater in JJA and SON than in DJF and MAM. This pattern shows that an annual trend map combines seasonally different centres of change. The season-by-latitude matrix further indicates that the tropical contribution is not uniform; for example, the northern tropical seasonal estimates in JJA and SON were weak and uncertain even though the annual 0°–30°N trend was negative. These features provide testable targets for future comparisons with convection and circulation diagnostics.

UTH responds to convection, detrainment, subsidence, and horizontal transport (Soden and Fu 1995; Sherwood et al. 2010). The spatial pattern is compatible with changes in one or more of these processes, but the present analysis contains no direct circulation, cloud, temperature, or radiation diagnostics. Accordingly, the midlatitude concentration of significant trends identifies where follow-up attribution should focus; it does not establish a dynamical cause.

4.3 Statistical robustness and remaining structural uncertainty

The agreement between OLS and Theil–Sen slopes reduces concern that isolated annual anomalies dominate the main estimate. Accounting for positive residual autocorrelation makes the significance test more conservative, while FDR correction prevents the spatial field from being interpreted through thousands of unadjusted P values. The moving-block-bootstrap intervals provide a separate uncertainty view that allows residual variability to remain temporally clustered. Together, these procedures support the conclusion that the observed product-level decline is statistically robust.

The breakpoint analysis qualifies, rather than overturns, that conclusion. Both annual and monthly records favoured a continuous slope change near 2003, with a faster decline before and a slower decline afterwards. The nonsignificant annual level-shift and detrended Pettitt tests argue against describing the record as one abrupt discontinuity. Because the breakpoint year was selected from all admissible candidates, the global block-bootstrap probabilities are more appropriate than a nominal test performed at a preselected year. Even so, a 23-year annual record provides limited information for distinguishing low-frequency climate variability, gradual homogenization effects, and genuine changes in the underlying trend.

The platform audit identifies a concrete competing explanation. The early 2000s included changes in the combination of DMSP and NOAA microwave sounders contributing to the merged

https://doi.org/10.5194/egusphere-2026-5105
Preprint. Discussion started: 3 November 2026
© Author(s) 2026. CC BY 4.0 License.

daily field, while later years incorporated Metop and Suomi NPP platforms. Long microwave records commonly require explicit correction of inter-satellite offsets, orbital drift, and changing local sampling time (Zou et al. 2006; Mears and Wentz 2017), while climate-data-record guidance distinguishes measurement, sampling, and structural components of uncertainty (Merchant et al. 2017). Temporal proximity alone cannot establish an instrumental origin: multiple platforms contributed simultaneously, the product was designed as a homogenized climate data record, and no abrupt domain-level residual shift was detected. Conversely, the present analysis cannot rule out smaller calibration or sampling effects distributed over several transitions. The defensible inference is therefore that the decline is robust within CM SAF v2, but its change in rate has both climatic and observing-system explanations that require independent evidence.

However, robustness to statistical choices is not equivalent to invariance across observational definitions. The ERA5 comparison demonstrated this boundary directly: no fixed pressure level reproduced the magnitude of the CM SAF domain decline, and the sign changed from weakly negative at 200–300 hPa to significantly positive at 400–500 hPa. Reanalysis humidity fields can differ because their assimilated observations, forecast models, bias corrections, and observing-system transitions differ, even between successive ECMWF products (Dee et al. 2011; Hersbach et al. 2020). This divergence could therefore reflect real vertical redistribution of relative humidity, the broad and humidity-dependent satellite weighting function, sampling and homogenization differences, or structural uncertainty in either product. Because ERA5 assimilates satellite observations, it is not a fully independent ground truth; agreement would not prove the satellite trend, and disagreement cannot identify which record is closer to the unobserved atmospheric state. The result also ends in 2018 and should not be extrapolated to the present climate state.

4.4 Vertical representativeness, rather than level choice, limits validation

The cross-product analysis separates three criteria that a convincing validation would need to satisfy: domain-trend magnitude, spatial organization, and detrended interannual variability. ERA5 came closest to the spatial pattern at 300 hPa, but neither that level nor any non-negative mixture reached the CM SAF domain trend. The optimized spatial mixture remained biased by almost one percentage point per decade, while an independently optimized temporal mixture failed under leave-one-year-out validation. The discrepancy is therefore not resolved by choosing a more favourable pressure level or by averaging the tested levels.

This outcome is physically and methodologically plausible because satellite UTH is not a

https://doi.org/10.5194/egusphere-2026-5105
Preprint. Discussion started: 3 November 2026
© Author(s) 2026. CC BY 4.0 License.

fixed-pressure variable. Its microwave weighting varies with atmospheric humidity and temperature, and 183-GHz measurements carry radiative-transfer and representativeness uncertainties that depend on the atmospheric scene (Brogniez et al. 2016). Pressure-level ERA5 fields instead follow fixed pressure surfaces whose relation to the upper-tropospheric emission layer changes in space and time. The incompatible spatial and temporal weights also show that a data-fitted vertical average is not a stable surrogate: a mixture optimized for one diagnostic does not generalize to another. The convex-mixture analysis remains a deliberately simple stress test and cannot replace an observation operator. A decisive comparison should apply a validated radiative-transfer model such as RTTOV to reanalysis or model profiles, use the satellite viewing geometry and sampling, and compare synthetic radiances or retrieval-equivalent humidity with the observations (Saunders et al. 2018).

4.5 Implications for model evaluation

A model comparison based on this record should reproduce the satellite operator or otherwise account for the product's broad vertical weighting and sampling characteristics. Comparing the UTH trend directly with relative humidity at a fixed model pressure level could confound physical and observational definitions. A defensible evaluation would collocate model output with the satellite sampling, apply a compatible humidity transformation, and compare the spatial, seasonal, and latitude-band structures shown here.

Under those conditions, the processed record and source-data tables provide a useful benchmark for evaluating whether simulated upper-tropospheric moisture changes have the observed sign, magnitude, and spatial organization.

5. Conclusions

CM SAF version-2 observations show an internally robust decline in upper-tropospheric humidity over 60°S–60°N during 1996–2018. The area-weighted trend was −1.238 percentage points decade^{-1}, negative slopes occupied 97.3% of valid grid cells, and the sign persisted across seasons, latitude bands, coverage thresholds, trend windows, and a platform-transition audit. The central cross-product result is not merely that ERA5 differs: the satellite trend falls outside the range attainable by every non-negative mixture of ERA5 relative humidity at 200–500 hPa. Spatially optimized mixing retained moderate pattern correlation but large trend bias, whereas temporally optimized mixing failed to predict detrended annual variability. Fixed-pressure selection and simple vertical averaging therefore cannot provide an equivalent validation target for this

https://doi.org/10.5194/egusphere-2026-5105
Preprint. Discussion started: 3 November 2026
© Author(s) 2026. CC BY 4.0 License.

microwave UTH record. The CM SAF decrease should be treated as a robust climate-data-record result whose atmospheric interpretation remains observation-operator dependent. Attribution and reconciliation require radiative-transfer-equivalent comparisons together with direct temperature, circulation, cloud, and calibration diagnostics.

Declaration of generative AI and AI-assisted technologies in the writing process

During the preparation of this work, the author used OpenAI Codex to assist with manuscript structuring, language editing, data-analysis code development, and figure preparation. After using this tool, the author reviewed and edited the content as needed, verified the reported analyses and results against the underlying data and code outputs, and takes full responsibility for the content of the publication.

Statements and Declarations

Financial support

No funds, grants, or other support was received for conducting this study or preparing this manuscript.

Competing interests

The author has no relevant financial or non-financial interests to disclose.

Author contribution

Zhuolun Li: Conceptualization; Methodology; Software; Validation; Formal analysis; Investigation; Data curation; Visualization; Writing – original draft; Writing – review and editing; Project administration. The author read and approved the final manuscript and accepts responsibility for the integrity of the work.

Data availability

The primary source data are the CM SAF Upper Tropospheric Humidity Thematic Climate Data Record, version 2.0, distributed through the Copernicus Climate Data Store and archived at https://doi.org/10.5676/EUM_SAF_CM/UTH/V002. ERA5 monthly averaged pressure-level relative humidity was obtained from the Copernicus Climate Data Store dataset reanalysis-era5-pressure-levels-monthly-means for 1996–2018. Derived annual series, latitude-band statistics, gridded trends, uncertainty intervals and figure source tables are supplied in the accompanying supplementary source-data archive. The large original source archives should be accessed through the Copernicus Climate Data Store rather than redistributed with the manuscript.

Code availability

The processing, statistical, breakpoint, platform-audit, ERA5-comparison and plotting

https://doi.org/10.5194/egusphere-2026-5105
Preprint. Discussion started: 3 November 2026
© Author(s) 2026. CC BY 4.0 License.

scripts are supplied in the accompanying supplementary code-and-source-data archive. The exact ERA5 request is recorded in download_era5_validation.py, and package requirements are provided in requirements.txt.

Acknowledgements

The author acknowledges EUMETSAT CM SAF and the Copernicus Climate Change Service/ECMWF for providing the CM SAF UTH and ERA5 data used in this study.

https://doi.org/10.5194/egusphere-2026-5105
Preprint. Discussion started: 3 November 2026
© Author(s) 2026. CC BY 4.0 License.

References

Bai J, Perron P (1998) Estimating and testing linear models with multiple structural changes. Econometrica 66:47–78. https://doi.org/10.2307/2998540

Bates JJ, Jackson DL (2001) Trends in upper-tropospheric humidity. Geophys Res Lett 28:1695–1698. https://doi.org/10.1029/2000GL012544

Benjamini Y, Hochberg Y (1995) Controlling the false discovery rate: a practical and powerful approach to multiple testing. J R Stat Soc Series B Stat Methodol 57:289–300. https://doi.org/10.1111/j.2517-6161.1995.tb02031.x

Brogniez H, English S, Mahfouf JF, Behrendt A, Berg W, Boukabara S, Buehler SA, Chambon P, Gambacorta A, Geer A, Ingram W, Kursinski ER, Matricardi M, Odintsova TA, Payne VH, Thorne PW, Tretyakov MY, Wang J (2016) A review of sources of systematic errors and uncertainties in observations and simulations at 183 GHz. Atmos Meas Tech 9:2207–2221. https://doi.org/10.5194/amt-9-2207-2016

Buehler SA, John VO (2005) A simple method to relate microwave radiances to upper tropospheric humidity. J Geophys Res Atmos 110:D02110. https://doi.org/10.1029/2004JD005111

Buehler SA, Kuvatov M, John VO, Milz M, Soden BJ, Jackson DL, Notholt J (2008) An upper tropospheric humidity data set from operational satellite microwave data. J Geophys Res Atmos 113:D14110. https://doi.org/10.1029/2007JD009314

Chung ES, Soden BJ, Sohn BJ, Shi L (2014) Upper-tropospheric moistening in response to anthropogenic warming. Proc Natl Acad Sci USA 111:11636–11641. https://doi.org/10.1073/pnas.1409659111

Courcoux N, Schröder M (2015) The CM SAF ATOVS data record: overview of methodology and evaluation of total column water and profiles of tropospheric humidity. Earth Syst Sci Data 7:397–414. https://doi.org/10.5194/essd-7-397-2015

Dee DP, Uppala SM, Simmons AJ et al (2011) The ERA-Interim reanalysis: configuration and performance of the data assimilation system. Q J R Meteorol Soc 137:553–597. https://doi.org/10.1002/qj.828

Dessler AE, Davis SM (2010) Trends in tropospheric humidity from reanalysis systems. J Geophys Res Atmos 115:D19127. https://doi.org/10.1029/2010JD014192

Gettelman A, Collins WD, Fetzer EJ, Eldering A, Irion FW, Duffy PB, Bala G (2006)

https://doi.org/10.5194/egusphere-2026-5105
Preprint. Discussion started: 3 November 2026
© Author(s) 2026. CC BY 4.0 License.



588 Climatology of upper-tropospheric relative humidity from the Atmospheric Infrared Sounder
589 and implications for climate. J Clim 19:6104–6121. https://doi.org/10.1175/JCLI3956.1
590 Held IM, Soden BJ (2000) Water vapor feedback and global warming. Annu Rev Energy Environ
591 25:441–475. https://doi.org/10.1146/annurev.energy.25.1.441
592 Hersbach H, Bell B, Berrisford P et al (2020) The ERA5 global reanalysis. Q J R Meteorol
593 Soc 146:1999–2049. https://doi.org/10.1002/qj.3803
594 Hollmann R, Merchant CJ, Saunders R et al (2013) The ESA Climate Change Initiative:
595 satellite data records for essential climate variables. Bull Am Meteorol Soc 94:1541–1552.
596 https://doi.org/10.1175/BAMS-D-11-00254.1
597 Houshangpour A, John VO, Buehler SA (2005) Retrieval of upper tropospheric water vapor and
598 upper tropospheric humidity from AMSU radiances. Atmos Chem Phys 5:2019–2028.
599 https://doi.org/10.5194/acp-5-2019-2005
600 John VO, Holl G, Allan RP, Buehler SA, Parker DE, Soden BJ (2011) Clear-sky biases in
601 satellite infrared estimates of upper tropospheric humidity and its trends. J Geophys Res
602 Atmos 116:D14108. https://doi.org/10.1029/2010JD015355
603 Künsch HR (1989) The jackknife and the bootstrap for general stationary observations. Ann
604 Stat 17:1217–1241. https://doi.org/10.1214/aos/1176347265
605 Lang T, Buehler SA, Burgdorf M, Hans I, John VO (2020) A new climate data record of
606 upper-tropospheric humidity from microwave observations. Sci Data 7:218.
607 https://doi.org/10.1038/s41597-020-0560-1
608 Mears CA, Wentz FJ (2017) A satellite-derived lower-tropospheric atmospheric temperature
609 dataset using an optimized adjustment for diurnal effects. J Clim 30:7695–7718.
610 https://doi.org/10.1175/JCLI-D-16-0768.1
611 Merchant CJ, Embury O, Bulgin CE et al (2017) Uncertainty information in climate data
612 records from Earth observation. Earth Syst Sci Data 9:511–527.
613 https://doi.org/10.5194/essd-9-511-2017
614 Moradi I, Buehler SA, John VO, Eliasson S (2010) Comparing upper tropospheric humidity data
615 from microwave satellite instruments and tropical radiosondes. J Geophys Res Atmos
616 115:D24310. https://doi.org/10.1029/2010JD013962
617 North GR, Bell TL, Cahalan RF, Moeng FJ (1982) Sampling errors in the estimation of
618 empirical orthogonal functions. Mon Weather Rev 110:699–706.
619 https://doi.org/10.1175/1520-0493(1982)110<0699:SEITEO>2.0.CO;2
620 Pettitt AN (1979) A non-parametric approach to the change-point problem. Appl Stat
621 28:126–135. https://doi.org/10.2307/2346729

https://doi.org/10.5194/egusphere-2026-5105
Preprint. Discussion started: 3 November 2026
© Author(s) 2026. CC BY 4.0 License.



622 Santer BD, Wigley TML, Boyle JS, Gaffen DJ, Hnilo JJ, Nychka D, Parker DE, Taylor KE (2000)
623 Statistical significance of trends and trend differences in layer-average atmospheric
624 temperature time series. J Geophys Res Atmos 105:7337–7356.
625 https://doi.org/10.1029/1999JD901105
626 Santer BD, Mears C, Wentz FJ et al (2007) Identification of human-induced changes in
627 atmospheric moisture content. Proc Natl Acad Sci USA 104:15248–15253.
628 https://doi.org/10.1073/pnas.0702872104
629 Saunders R, Hocking J, Turner E et al (2018) An update on the RTTOV fast radiative transfer
630 model (currently at version 12). Geosci Model Dev 11:2717–2737.
631 https://doi.org/10.5194/gmd-11-2717-2018
632 Schulz J, Albert P, Behr HD et al (2009) Operational climate monitoring from space: the
633 EUMETSAT Satellite Application Facility on Climate Monitoring (CM-SAF). Atmos Chem Phys
634 9:1687–1709. https://doi.org/10.5194/acp-9-1687-2009
635 Sen PK (1968) Estimates of the regression coefficient based on Kendall's tau. J Am Stat
636 Assoc 63:1379–1389. https://doi.org/10.1080/01621459.1968.10480934
637 Sherwood SC, Roca R, Weckwerth TM, Andronova NG (2010) Tropospheric water vapor,
638 convection, and climate. Rev Geophys 48:RG2001. https://doi.org/10.1029/2009RG000301
639 Soden BJ, Fu R (1995) A satellite analysis of deep convection, upper-tropospheric humidity,
640 and the greenhouse effect. J Clim 8:2333–2351.
641 https://doi.org/10.1175/1520-0442(1995)008<2333:ASAODC>2.0.CO;2
642 Soden BJ, Held IM (2006) An assessment of climate feedbacks in coupled ocean–atmosphere
643 models. J Clim 19:3354–3360. https://doi.org/10.1175/JCLI3799.1
644 Taylor KE (2001) Summarizing multiple aspects of model performance in a single diagram. J
645 Geophys Res Atmos 106:7183–7192. https://doi.org/10.1029/2000JD900719
646 Wilks DS (2006) On field significance and the false discovery rate. J Appl Meteorol
647 Climatol 45:1181–1189. https://doi.org/10.1175/JAM2404.1
648 Zou CZ, Goldberg MD, Cheng Z, Grody NC, Sullivan JT, Cao C, Tarpley D (2006) Recalibration
649 of microwave sounding unit for climate studies using simultaneous nadir overpasses. J
650 Geophys Res Atmos 111:D19114. https://doi.org/10.1029/2005JD006798