Testing remote sensing estimates of snow water equivalent in the framework of the European Drought Observatory

Carmelo Cammalleri^{✉,*}, Paulo Barbosa, and Jürgen V. Vogt

European Commission, Joint Research Centre, Ispra, Italy

Abstract. We evaluated the feasibility for operational snow drought monitoring over Europe based on the near-real-time snow water equivalent (SWE) satellite product from the EUMETSAT Satellite Application Facility on Support to Operational Hydrology and Water Management (H-SAF). To do so, the consistency of this dataset with the consolidated dataset of the Canadian Meteorological Centre (CMC), as well as with the ERA5 reanalysis dataset from the European Centre for Medium-Range Weather Forecasts, was tested in terms of both spatial snow coverage and detection of anomalies from the long-term climatology. The analysis confirms a general good agreement among the three products as well as substantial differences over mountainous terrains, with the H-SAF product capturing only about 30% of the areas identified by CMC as snow-covered in those areas, while a better match between the ERA5 and the CMC spatial coverage is observed. However, significant inconsistencies in the correlation between all three SWE anomalies are observed over mountain areas. Due to the lack of a reliable reference dataset, the observed inconsistencies and the coarse spatial resolution (0.25 deg) of all three products limit the possibility for snow drought monitoring over key European regions such as the Alps. © European Union. Published by SPIE under a Creative Commons Attribution 4.0 International License. Distribution or reproduction of this work in whole or in part requires full attribution of the original publication, including its DOI. [DOI: [10.1117/1.JRS.16.014509](https://doi.org/10.1117/1.JRS.16.014509)]

Keywords: snow drought; H-SAF; ERA5; Copernicus Emergency Management Service.

Paper 210352 received Jun. 3, 2021; accepted for publication Jan. 10, 2022; published online Feb. 1, 2022.

1 Introduction

The concept of snow drought, as the manifestation of reduced water storage in the form of snow, has appeared in the scientific literature since the late seventies and early eighties. For instance, the term was used to describe the reduced snow cover observed over North America during the winter of 1980 to 1981.¹ Around the same period, effects of snow drought on avian communities were discussed in ecological studies,² and in the early 1990s snow drought was discussed when analyzing the lack of snow in both North America and Eurasia during the period of 1970 to 1990.³

Climate change studies on past trends^{4,5} and future projections^{6,7} of snow storage emphasize the increasing importance of snow for the hydrological balance of several snow-dominated water basins.⁸ Following the US West coast winter drought of 2015, a new classification of snow drought was proposed, which differentiates between “dry snow drought,” caused by a lack of winter precipitation as snow, and “warm snow drought,” associated with a lack of snow accumulation even during normal winter precipitation due to high temperatures.⁹ Similar mechanisms were observed for snow droughts over Austria and Norway,¹⁰ where unusual high temperatures and/or low precipitation conditions may play different roles in the lack of snow.

This distinction has several practical implications for water managers, who have to deal with reduced streamflow over the entire year in the case of dry snow drought or early melting in the case of warm snow drought that can lead to winter floods followed by spring/summer droughts. In snow-dominated regions, severe winter droughts can directly impact local tourism, but they also impact subsequent spring and summer water availability and drought conditions. In addition

*Address all correspondence to Carmelo Cammalleri, carmelo.cammalleri@ec.europa.eu

to its hydrological relevance, snow drought also has a major influence on the health and function of numerous regional ecosystems and their response to climate change.¹¹

Independent of the underlying drought mechanism, the snow water equivalent (SWE)—a measure of the depth of water that would result from the melting of the snowpack—represents a key quantity for snow drought monitoring since it can be used to detect both dry and warm snow droughts, as both result in lower than average SWE values¹² (low SWE values are obtained if either less precipitation has fallen or an unusually large amount of melt has occurred). For this reason, in this study, we focus on the use of SWE, and more specifically its deviations from the long-term climatology, as a general proxy variable for snow drought conditions.

Direct SWE observations can be derived on the ground via the snow pillow method, as in the case of the SNOTEL network.¹³ Indirect estimates are based on snow depth (SD) measurements from ultrasonic or GPS sensors,^{14,15} which require time-consuming bulk density measurements with snow samples¹⁶ or tubes¹⁷ for the conversion of SD to SWE. Overall, SWE ground measurements remain unaffordable for large scale monitoring and are rarely deployed in high-elevation terrains. Consequently, land surface modeling and airborne/satellite remote sensing techniques have been extensively used to estimate snow-related quantities with high spatiotemporal continuity.^{18,19}

Satellite estimates of SWE are commonly obtained from spaceborne passive microwave sensors as a function of the brightness temperature.^{20,21} Several properties of the snowpack influence the microwave signal, including the snow temperature, grain size, the number of grains, and the volume to air fraction, as well as the underlying surface conditions and *in situ* vegetation characteristics.²⁰ The strong interconnection among these factors makes remote estimates of SWE challenging, as exemplified by the numerous algorithms proposed in the literature,²¹ ranging from simple empirical relationships to complex modeling of the physical processes. However, despite the abovementioned advantages that satellite estimations have over ground measurements, their coarse spatial resolution²² limits the direct usage of passive microwave SWE products over complex and rough terrain in mountain regions,²³ where they may be characterized by high uncertainty.^{24,25}

Moreover, due to the underlying physics of the retrieval process, passive microwave estimates of SWE tend to saturate at high SD levels²⁶ and are often unreliable in the case of wet snow.²⁷ These well-known limitations have indeed been highlighted in several intercomparison studies,^{28–30} in which pure remote sensing-based products showed large uncertainties. Generally, satellite-only estimates underperform compared with the more reliable ground-satellite blended products, which are becoming increasingly available in a timely manner for operational monitoring systems.

By incorporating ground-based observations into the retrieval algorithms, the abovementioned major limitations of remote sensing-based products have been partially reduced.²⁴ These blending procedures allow for improving the overall accuracy of SWE retrievals, as well as reducing the systematic errors, even if saturation effects persist at high SWE values.³¹ In addition, blending may introduce some artifacts²⁸ that influence the results of snow drought monitoring and still need to be carefully investigated.

A dynamic modeling of SWE can also be achieved from the snowpack components of hydrological and land surface models.²² These models solve the energy balance problem of snow formation and melting by adopting different degrees of complexity in the representation of the main physical processes, such as absorption of incoming radiation, advection, and convection, as well as by making different assumptions for the representation of the internal structure of the snowpack.³² Since various degrees of abstraction are involved in the modeling of each process, the accuracy of SWE estimates is inherently related to both the model complexity and parameterization. As for any of the outputs of hydrological models, the quality of the modeled SWE is also strongly related to the quality of the meteorological forcing data.

For an SWE product to be adopted in the framework of an operational snow drought monitoring as would be the case in the frame of the European Drought Observatory (EDO),³³ two main characteristics are key: (i) the timeliness of data delivery and (ii) the temporal consistency of the time series as a basis for the detection of anomalies from the long-term climatology. These two features are not a common focus of intercomparison studies of SWE products, which are often centered on absolute accuracy and bias of the products; hence specific analyses are needed

to evaluate the feasibility of introducing a reliable SWE-based snow drought product in EDO.

In searching for a near-real time and operational product for snow drought monitoring over Europe, we investigated the suitability of the SWE dataset provided by the EUMETSAT Satellite Application Facility on Support to Operational Hydrology and Water Management (H-SAF) and more specifically of derived anomalies for inclusion in EDO. Given the lack of an independent dataset of ground truth for a direct validation of the anomalies derived from the H-SAF product, this product was compared against the similar near-real-time product from the European Centre for Medium-Range Weather Forecasts (ECMWF) ERA5 reanalysis, as well as against the consolidated dataset produced annually (i.e., not in near-real time) by the Canadian Meteorological Centre (CMC). These three products are selected to cover the main means to detect SWE at a continental scale, namely ground measurements, modeling, and satellite retrievals. However, none of them can currently be considered an absolute ground truth due to various imitations. As a consequence, the main goals of the intercomparison are (i) to quantify the spatial agreement of the three products in detecting the areas potentially affected by snow drought and (ii) to assess the intermodel consistency in the estimated SWE anomalies. Special emphasis is given to the results obtained over mountain areas, given the well-known uncertainty of SWE products over complex terrains and the relevance of such areas in the hydrology of many European river basins.

2 Materials and Methods

This section describes the near-real-time remote sensing-based SWE product identified as candidate for an operational snow drought monitoring, as well as the two additional datasets used in the analyses, one based on a reanalysis (ERA5) and one on the interpolation of *in situ* data (CMC). The strategy adopted to intercompare the three products and the criteria adopted to assess the spatiotemporal consistency of the products for snow drought monitoring are also outlined.

2.1 H-SAF Dataset

As part of the EUMETSAT H-SAF, a near-real-time SWE dataset is produced and freely distributed.³⁴ The H-SAF SWE product has some similarities with the GlobSnow dataset³⁵ produced by the Finnish Meteorological Institute (FMI). However, a key feature of the H-SAF procedure is the separate treatment of flat/forest areas and mountain areas following the algorithms developed by the FMI and by the Turkish State Meteorological Services, respectively.³⁶ In summary, both institutes provide estimates of SWE for the entire European continent, but the optimized estimate of H-SAF is used for each of the two regions (flat/forest or mountain) according to a mountain map as determined using simple rules on mean altitude and standard deviation of the slope derived from the GTOPO30³⁷ digital elevation model.³⁸

The flat/forest areas component combines the sparse synoptic SD ground observations with passive microwave brightness temperature maps retrieved from the Special Sensor Microwave/Imager sensor on board the US Air Force Defense Meteorological Satellite Program series of satellites. The underlying assimilation procedure³⁶ estimates the SWE by inverting a semiempirical snow emission model³⁹ and by combining these estimates with a weather station-based Kriging-interpolated SD map using a Bayesian spatial assimilation approach. A snow density value is applied to each grid cell to convert SD to SWE. Given that spaceborne microwave-based SWE estimates are reliable only for dry snow, the assimilation is not performed under wet snow conditions.⁴⁰

Over mountain areas, SWE estimates are based on spaceborne data only,³⁶ using an algorithm analogous to the one used for flat terrain. In addition, empirical estimates of grain size and density are considered more suitable over these terrains rather than ancillary information due to the common data scarcity. The mountain algorithm is designed for dry snow only; hence no SWE calculation is done if wet snow status is detected by the dedicated H-SAF product.

The merged product comprises daily maps in near-real-time (12 h time lag) covering Europe (25–75°N lat, 25°W–45°E lon) at a 0.25 deg spatial resolution, starting in 2013.³⁶ When ground

observations are not available in time for a near-real-time delivery or when wet snow conditions are detected in mountain areas, the daily maps are masked over the corresponding pixels, resulting in partially missing data in the daily dataset. To maximize the data coverage at the monthly scale, monthly maps were obtained as maximum value composites for those months with at least three valid daily datasets available.

The dataset has been previously validated over Finland and Turkey.⁴⁰ Some key outcomes of the validation exercise of this product are an overall accuracy of monthly estimates [in terms of root-mean-square error (RMSE)] on the order of 30 to 40 mm and a clear tendency to underestimate SWE for values >150 mm due to signal saturation. Also it is worth pointing out that most of the validations were performed against ground data collected between January and March to focus on dry snow only.

2.2 CMC and ECMWF Datasets

The CMC provides a dataset of estimated monthly SWE⁴¹ starting in 1998 and regularly updated to the last full year (i.e., 2019 in this case).⁴² Since this dataset is produced only once per year, it can be considered a consolidated dataset over the northern hemisphere based on *in situ* observations that are not constrained by near-real-time requirements.

The monthly mean SWE values provided by CMC are used in this study. Those maps were obtained by applying to monthly mean SD maps the snow density values stored in a look-up table and derived from the Canadian snow-climate classification for different snow-climate classes.^{43,44} The SD analysis is performed using a combination of *in situ* daily observations and optimal interpolation with a first-guess field generated from a simple snow accumulation and melt model based on temperature and forecasted precipitation.⁴⁵ *In situ* observations include data from synoptical stations and aviation reports.

Even if the look-up table values were derived from information acquired over Canada, the dataset includes monthly average maps for the entire canonical northern hemisphere snow season (namely October to June, the months when SD observations are available) on a standard 24-km polar stereographic grid. For this study, data were resampled over Europe on a lat/lon regular grid at a 0.25 deg spatial resolution using the nearest neighbor algorithm to obtain a dataset co-registered to the H-SAF grid.

A full quality assessment of the SD dataset at the base of this product is described in Ref. 45. Over Europe, the authors report an accuracy in the modeled SD of about 13 cm (RMSE), with a slightly negative bias of 5 cm.⁴⁵ Overall, the major errors of this product were found in Arctic regions (where no observations are available) and land areas with mean annual maximum snow accumulation exceeding 300 mm (e.g., Canadian arctic lands and Tibetan plateau).

In addition to the H-SAF and CMC datasets, the snow map provided by the ECMWF as part of the ERA5 reanalysis dataset was analyzed. Currently, the ERA5 dataset includes a large set of atmospheric, land, and oceanic climate variables covering the full globe with a grid of about 30 km on a reduced Gaussian grid (T1639) and freely distributed within 5 days after real time. ERA5T data may still undergo a correction in a second step but no later than 3 months beyond real time.

In the ERA5 modeling framework, SWE is first simulated according to a physically based energy and mass-balance snow scheme.⁴⁶ The outputs of the snow analysis are then updated based on a two-dimensional optimal interpolation of station observations of SD and snow cover.^{47,48}

In this study, the monthly product distributed by ECMWF through the Copernicus Climate Data Store⁴⁹ is used. The raw data were cropped over Europe and regridded on the same 0.25 lat/lon grid used by H-SAF. Monthly data between 1998 and 2019 were used in the analysis, constituted only by consolidated maps (i.e., no ERA5T data were used).

2.3 Evaluation Strategy

Droughts are generally associated with anomalous conditions compared with the long-term climatology; hence all analyses in this study are focused on standardized *z* scores (i.e., anomalies), defined as

$$z_{i,j} = \frac{\text{SWE}_{i,j} - \mu_i}{\sigma_i}, \quad (1)$$

where the subscripts i and j represent the month and the year, respectively, and μ_i and σ_i are the average and standard deviation values for the given month, respectively, computed on a baseline period.

The agreement of the z score products derived from the H-SAF, CMC, and ERA5 data is evaluated by two analyses focused on: (1) snow detection over recurring snow-covered regions as defined by the CMC dataset and (2) consistency of the anomaly values of the three datasets. To assess the suitability of the H-SAF dataset for an implementation within the EDO operational system, threshold criteria are defined based on the indices derived during the two analyses.

The first analysis is made to detect the area that may be affected by snow drought events, i.e., areas with a systematic presence of at least a minimum amount of snow. To do that, a “snow month” is defined as when the monthly SWE is greater than the minimum threshold of 10 mm (analogous to the value used in H-SAF as an optimal accuracy⁴⁰). This minimum threshold allows for excluding regions with modeled SWE values that are >0 but still too small to be considered potentially affected by snow drought.

For each month, the number of past years with snow (N_s) are computed, and the grid cells where snow drought conditions can occur are obtained by defining a minimum number of years ($N_s > 10$) in the full 1998 to 2019 CMC dataset. The CMC dataset is used here as baseline since it is the only one based on consolidated *in situ* data. The same procedure is then applied to both H-SAF and ERA5 datasets to perform a confusion matrix analysis, the outcomes of which are synthesized by the statistical indicators probability of detection (POD), probability of false detection (POFD), accuracy (ACC), and Heidke skill score (HSS), defined as

$$\text{POD} = \frac{a}{a + c}, \quad (2)$$

$$\text{POFD} = \frac{b}{b + d}, \quad (3)$$

$$\text{ACC} = \frac{a + d}{a + b + c + d}, \quad (4)$$

$$\text{HSS} = 2 \frac{ad - bc}{(a + c)(c + d) + (a + b)(b + d)}, \quad (5)$$

where a , b , c , and d represent the true positive, false negative, false positive, and true negative, respectively.

The POD and POFD indices have opposite behaviors, with the former being best at 1 and the latter at 0. For this study, we assume a POD of 0.8 as a threshold value to define a target consistency, representing an overlap on at least 80% of the areas detected by CMC. Analogously, a value of 0.2 is used as an acceptable threshold for POFD. In terms of accuracy and skill, target values of 0.8 and 0.5 are used as criteria for assessing the suitability of the product for inclusion in EDO.

In the second step, the consistency among the SWE anomalies modeled by H-SAF, CMC, and ERA5 is analyzed by means of a linear correlation analysis. In the absence of a clear SWE reference dataset for validation and given the uncertainty underlying all three datasets, the correlation analysis is used to identify areas where significant inconsistencies among anomalies in the three datasets are detected.

Two main metrics are extracted from the correlation analysis: the Pearson correlation coefficient (r) and its statistical significance (p). It is worth mentioning that for standardized anomalies the r coefficient represents both the correlation and the slope of the regression (the intercept is always zero); hence it also quantifies the accuracy and the bias. In particular, the percentage of the domain with positive correlation values and $p < 0.05$ is used to identify the overall consistency between two models, and a value of 80% is assumed to be the minimum threshold for a satisfactory agreement.

In addition, the fraction of monthly values that fall outside a buffer region around the regression line is computed for each cell (F_o), assuming an amplitude of the buffer equal to ± 0.5 . The value of the buffer corresponds to the width of the main drought classes in commonly used standardized precipitation index classifications.⁵⁰ The criterion for suitability of the dataset is set to a maximum of 1/3 of the data that can fall outside the two buffer lines.

3 Results and Discussion

3.1 Analysis of the Spatial Coverage

The need for snow drought monitoring is limited to areas that have a temporal-consistent snow presence, i.e., where a deficit compared with the usual snow amount may constitute a concern for water managers. The temporal consistency of the CMC SWE grid data is summarized in Fig. 1, where the number of years (N_s) with SWE > 10 mm is depicted for each month in the CMC snow season (between October and June). These maps show how the snow cover reaches its maximum around January/February and is negligible in October and June.

By defining a consistent snow presence when at least 50% of the years have SWE > 10 mm ($N_s > 10$, out of the total 22 years in the 1998 to 2019 CMC dataset), the temporal evolution of

Fig. 1 Spatial distribution of the number of years (N_s) with SWE > 10 mm for each month in the CMC snow season. The lower panel shows the temporal evolution of the percentage of the domain with at least 10 years with SWE > 10 mm (out of a total of 22 years).

the percentage of the domain with consistent snow presence is reported in lower panel of Fig. 1, with values up to about 50% of Europe. The data show how two main areas are affected by consistent snow presence during the whole November to May period, namely the Alps and the Scandinavian peninsula, whereas the rest of the domain, mostly eastern Europe, have consistent snow only between December and April. Other threshold values (between 0 and 20 mm) were also tested (not reported) and confirm the general key outcome to focus the successive analyses on the period of November to May only (7 months).

To perform the same analysis on the H-SAF shorter time series, it is necessary to define a threshold analogous to 10 years (used for 1998 to 2019) for the period of 2013 to 2019 covered by this dataset. With this aim, threshold values between 2 and 5 years were tested on the CMC (not shown), selecting the value that returns the minimum difference in terms of year-average coverage. This optimal value, corresponding to $N_s > 3$ years, was used to estimate the overlap between CMC and the H-SAF and ERA5 datasets in the period of 2013 to 2019, as reported in Fig. 2(a).

These data show that both datasets detect about 85% of the grids defined by CMC as consistently snow-covered (7-month average of 84.9% and 85.3% for H-SAF and ERA5, respectively), with slightly more temporal-consistent results for ERA5 compared with H-SAF (7-month standard deviation of 5.8% and 4.0% for H-SAF and ERA5, respectively). The high overlap among all three datasets suggests an overall good agreement on the detection of the areas potentially affected by snow drought at the European scale, while the temporal consistency in the results further suggests a minimal impact of the selected threshold since lower SWE values may be expected early and late in the season due to SWE values being closer to the threshold.

One of the main features of the H-SAF algorithm is the coverage of mountain areas through a dedicated subalgorithm, the lack of which is a major limitation of other remotely sensed based products,²³ including the similar GlobSnow product. For this reason, the data in Fig. 2(b) report the overlaps in the case of mountain areas only, defined in this study as the common areas in the H-SAF and GlobSnow mountain masks³⁶ [see the mask map reported in Fig. 4(c)]. These data show a negligible deterioration of the overlap for the ERA5 dataset (83.1 ± 4.8) compared with the full domain, which is in contrast to the noticeable reduction observed for H-SAF (36.6 ± 8.1). These results seem to suggest that both CMC and ERA5 have quite consistent

Fig. 2 Overlap between CMC and both H-SAF (gray bars) and ERA5 (black bars) snow areas for the (a) whole domain and (b) mountain areas only.

Fig. 3 Indicators derived from the confusion matrix between CMC and both H-SAF (gray bars) and ERA5 (black bars) snow areas. Probability of detection: (a) POD, the closer to 1 the better; (b) POFD, the closer to 0 the better; (c) ACC, the closer to 1 the better; and (d) HSS, positive values are skillful.

outcomes in terms of coverage, whereas the H-SAF product seems to be characterized by a noticeable incompleteness over snowy mountain regions, even if major snow areas are captured by the dataset in high-elevation regions. Part of the reason for the mismatch in the extension of the snow area in mountain regions may be ascribed to missing or low-quality satellite acquisitions (i.e., wet snow conditions, not modeled in mountain areas) rather than poorly detected snow conditions. This explanation is further supported by the worst performing months, November and May, which are located early and late in the season, respectively, when freshly accumulated wet snow is expected. Instead, it excludes a major effect of the selected 10 mm threshold since areas close to this limit are not predominantly located in mountain regions.

A more in-depth analysis of these SWE coverage maps was performed by means of a confusion matrix, the results of which are reported in Fig. 3 and summarized for the full study area in Table 1. Although the POD results are analogous to the overlap analysis reported in Fig. 2, the POFD highlights how a higher number of false detection is observed in the H-SAF product (0.21 ± 0.05) compared with ERA5 (0.05 ± 0.02), which seems to contrast with the previously reported underestimation of snow coverage over mountain regions, but it is in agreement with the small positive bias of SWE observed in H-SAF validation studies for low SWE values (<150 mm).⁴⁰ These results combine in a general better matching between CMC and ERA5 compared with H-SAF, with the latter showing fewer areas with consistent snow presence in mountain regions and a slightly higher extension of the area affected by consistent snow beyond the ones detected by CMC and ERA5.

The main effect of these behaviors is nicely summarized by both ACC and HSS data, which are both higher for ERA5 (0.92 ± 0.02 and 0.78 ± 0.07 , respectively) compared with H-SAF (0.82 ± 0.02 and 0.55 ± 0.16 , respectively). Overall, all of the statistical metrics for the H-SAF product at the European scale are in line with the values set for a suitable use for operational drought monitoring, most notably the high values of POD and ACC (0.85 and 0.82,

Table 1 Summary of the results of the confusion matrix analysis for the full study domain. See Fig. 3 for a synthetic description of the different indicators.

Dataset	POD	POFD	ACC	HSS
H-SAF	0.85 ± 0.06	0.21 ± 0.05	0.82 ± 0.02	0.55 ± 0.16
ERA5	0.85 ± 0.04	0.05 ± 0.02	0.92 ± 0.02	0.78 ± 0.07

respectively), and they are generally comparable with the values obtained for ERA5. On average, the three models seem to be overall consistent in terms of the areas that can be considered consistently snow-covered, with the only exception being the still present underrepresentation of snow coverage in mountain areas in the H-SAF dataset.

3.2 Analysis of the Consistency Among SWE Anomalies

Once the overall agreement between the snow-covered areas retrieved by CMC, H-SAF, and ERA5 products was established, the consistency of the three products in terms of SWE anomalies was evaluated by means of a correlation analysis on the period of 2013 to 2019 (up to 49 monthly values). Figure 4 reports the temporal Pearson correlation coefficient (r) maps for CMC versus H-SAF (a) and ERA5 (b), as well as ERA5 versus H-SAF (d). Each r value was computed independently for every grid cell with at least 25 monthly values available.

The maps in Fig. 4 highlight an overall good correspondence among the SWE anomaly time series of the three datasets, which is particularly true over flat areas and eastern Europe. Sparse r values lower than 0.2 can be observed over mountain regions [see image in Fig. 4(d)], including some negative values. On average, r values are positive and statistically significant ($p < 0.05$) in at least 90% of the domain in all cases (Table 2), highlighting an overall satisfactory consistency among the three SWE anomaly estimations over the majority of the domain. The lower average correlation observed between H-SAF and both CMC and ERA5 (compared with the correlation between CMC and ERA5) can be ascribed primarily to the reduced performances over mountain areas, as summarized by the average statistics over these regions (Table 3). The results over the mountain areas highlight a general low consistency among the outcomes of all three models, exemplified by both low average r values and the large variability in the percentage of the

Fig. 4 Spatial distribution of the Pearson correlation coefficients (r) between CMC and both (a) H-SAF and (b) ERA5 SWE anomalies, as well as between ERA5 and H-SAF (d). (c) The map depicts in blue the areas considered mountain terrain in the H-SAF algorithm.

Table 2 Summary of the correlation analysis between the three models' SWE z scores.

Datasets		Start year	r	% $p < 0.05$ and $r > 0$
CMC	H-SAF	2013	0.68 ± 0.25	91.8
	ERA5	2013	0.80 ± 0.22	96.6
ERA5	H-SAF	2013	0.66 ± 0.32	89.9
CMC	ERA5	1998	0.79 ± 0.21	98.4

Table 3 Summary of the correlation analysis between the three models' SWE z scores over mountain areas.

Datasets		Start year	r	% $p < 0.05$ and $r > 0$
CMC	H-SAF	2013	0.08 ± 0.25	15.4
	ERA5	2013	0.43 ± 0.25	64.6
ERA5	H-SAF	2013	0.16 ± 0.23	36.9

domain with a positive and statistically significant correlation (ranging between 15% and 65%). It is worth highlighting that the impact of these inconsistencies on the overall comparison is downplayed by the limited extension of mountain areas.

Overall, a diminished accuracy over mountain regions may be expected for all three approaches. Although the difficulties that satellite products (e.g., H-SAF) have in reproducing SWE dynamics in mountain regions were been observed and discussed in Sec. 3.1, the accuracy of estimates from global reanalysis products (e.g., ERA5) also proved to be limited, with some studies demonstrating improvement only when regional nested approaches are implemented over specific regions.⁵¹ The reliability of the products based on the interpolation of *in situ* measurements (e.g., CMC), finally, is strongly linked to the availability of observations, which is generally completely lacking in high-mountain regions or only available once a year at the end of the accumulation period.⁵² Additionally, it has been shown that, in the case of high-resolution products, even interpolated data from snow pillows may not perform as well as energy-balance-based estimates over some regions.⁵³

The coarse spatial resolution of all three products may be another limiting factor over mountain terrains as it may not be suitable to capturing the complex spatial dynamic observed in such areas. Differences between point scale observation and grid-cell average SWE values have been shown to be up to 200% in regions with complex topography,⁵⁴ suggesting that the locations of *in situ* stations may not always be optimal for representing local average conditions.

Since CMC z scores are computed on a longer baseline period (1998 to 2019) compared with H-SAF and ERA5 (2013 to 2019), another potential source of discrepancy can be associated with this difference, so its impact was evaluated by computing the SWE anomalies for ERA5 over the period of 1998 to 2019 and by performing an additional correlation analysis on the full period (154 monthly values). The correlation map obtained in this case (not shown) is very similar to the one reported in Fig. 4(b), and the negligible differences in the results between the two periods are summarized by the synthetic data reported at the bottom of Table 2. Generally, the ERA5 average r value is only marginally smaller when the full dataset is used, likely due to the increase in the sample size, but with a slightly larger percentage of the domain with the correlation being positive and statistically significant ($p < 0.05$). This result seems to suggest a limited impact of the mismatch in the baseline periods.

An additional set of correlation analyses was performed to test the spatial consistency of the SWE z scores, by computing the r value for all grid cells in each of the 49 months in the period of 2013 to 2019. The average r value for a specific month (from the r values in the 7 different years) is reported in Fig. 5 for CMC versus H-SAF and ERA5 and for ERA5 versus H-SAF. These

Fig. 5 Average spatial correlation for each month between CMC and both H-SAF (dark gray bars) and ERA5 (black bars) SWE anomalies, as well as between ERA5 and H-SAF (light gray).

results confirm many similarities in the consistency between the three anomaly datasets. Even if the correlation between H-SAF and both ERA5 and CMC is lower than the r value for CMC versus ERA5 (about 0.4 on average, compared with 0.6), the overall temporal dynamic is similar, with slightly lower values early/late in the snow season (i.e., December or May) compared with the central months. These low correlation values can be explained by the higher fraction of snow-covered cells in mountain areas during the early/late season compared with the middle season, as well as also by the difficulty of modeling wet snow for microwave-based approaches, which should be predominant during these periods. Expected lower performances for H-SAF during these months due to wet snow are also suggested by the product validation procedure, which focused on data collected between January and March only.⁴⁰

The ending of the snow season, with a prevalence of snow melting, seems to be a particularly problematic period since minimum correlation values are observed in May in all three correlation analyses. During this stage, in addition to the already mentioned issue of wet snow modeling by satellite data, reanalysis approaches also are often limited by the large number of factors involved in the modeling of melting. Even if an improved scheme is implemented in ERA5 for a better timing of melting,⁴⁶ the improvement observed in the offline runs do not necessarily translate to the online outcomes.

A final test on the consistency of the three datasets was performed on the frequency, in which the SWE anomalies of H-SAF and ERA5 differ from the ones from CMC for more than the

Fig. 6 Frequency analysis of data that fall outside the buffer region ± 0.5 (F_o). (a) An example for a cell in Sweden, with the black line representing the linear regression and the dotted lines delimiting the buffer zone. (b) The frequency distribution for all cells in the domain, with the vertical dotted line highlighting the suitability threshold of 33%.

defined threshold of 0.5. Specifically, this fraction F_o was computed independently for each cell [see example on Fig. 6(a)], and the frequency distribution across the domain is reported in Fig. 6(b) for both H-SAF and ERA5.

The analysis shows how high values of F_o are relatively common, with about half of the domain with F_o values greater than the threshold of 0.33 for H-SAF (i.e., only for half of the domain are the differences smaller than 0.5 for $<1/3$ of the monthly values). The agreement between CMC and ERA5 is slightly better (70% for $F_o = 0.33$), and acceptable F_o values can be observed over some mountain areas. By overlapping the areas with F_o values below 0.33 for all three pairs of comparison, only about 30% of the domain has consistent outcomes in all three datasets according to the set criterion, mostly comprised of flat areas in eastern Europe. This result reinforces the difficulties in systematically achieving SWE anomalies that are consistent among all three products for the entire domain and further stresses the impossibility of achieving optimal results everywhere with a single dataset.⁵⁵

4 Summary and Conclusions

The H-SAF SWE satellite product has been tested for its potential use for operational snow drought monitoring over Europe by comparing its near-real time, monthly aggregated SWE anomaly maps against the CMC consolidated datasets and the outputs of the similarly near-real-time ECMWF ERA5 reanalysis.

The comparison between the three datasets demonstrated an overall good consistency in the SWE dynamics at the European scale, both in terms of spatial coverage across the continent and agreement in the modeled standardized anomalies. This outcome is consistent with the results in Ref. 55, where the GlobSnow product (very similar to H-SAF in terms of modeling framework over flat terrains) returned a consistent performance with several reanalysis products, including ERA5. In the absence of independent assessments of snow drought, no conclusions can be drawn on the optimal dataset for specific areas, but the high consistency observed among the three datasets over some regions, such as eastern Europe, suggests at least a strong potential for near-real-time drought monitoring with either H-SAF or ERA5 over those regions.

Indeed, if only domain-average statistics are considered, the H-SAF dataset seems to fulfill all of the requirements for a potential operational implementation within the EDO drought monitoring system since no major differences among the three products are observed on average. This is mainly due to the high consistency observed over flat terrains, mostly located in eastern and northeastern Europe, where an overlap of about 85% is obtained for the snow areas derived by the three products, and positive ($r > 0$) and statistically significant correlation values ($p < 0.05$) in more than 90% of the domain between all three SWE anomaly time series are observed. The satisfactory results obtained by comparing H-SAF with ERA5 is in line with the outcome of the study in Ref. 56, in which the consistency between satellite and reanalysis products was observed to be better when the most recent reanalysis (i.e., ERA5 instead of ERA-Interim) was tested. In the same study, a strong consistency between existing SWE data products in terms of anomalies was observed,⁵⁶ even where the results diverge in the absolute values of snow storage. Less consistent results were observed over the Arctic and Alpine regions compared with the Boreal forests, and, indeed, our analysis confirmed the well-known issues related to the characterization of SWE dynamics over mountainous regions, where the agreement among the three datasets is strongly reduced. The reported results highlight the difficulties that remote sensing products have in correctly reproducing SWE anomalies over these regions, especially early and late in the season when wet snow is likely predominant. Only about 30% of the CMC snow coverage (areas with a temporal consistent presence of SWE) over mountain grid cells are also captured by H-SAF data, while CMC and ERA5 seem to be more consistent with each other.

Even if the agreement between the ECMWF reanalysis dataset and CMC in terms of snow coverage in mountain areas is high, the correlation analysis shows noticeable inconsistencies between these two products, with a large fraction of the domain with unacceptable differences (i.e., larger than 0.5) in modeled SWE anomalies for all three datasets. This result confirms that, despite the improvements in both remote sensing and modeling techniques, the characterization of SWE dynamics remains one of the key unresolved problems of snow hydrology.⁵⁷

In particular, the lack of consistency in the anomalies from the climatology over mountain regions underlines the effects of some of the well-known sources of error in all three datasets: (1) the inability of coarse spatial resolution satellite SWE products to match the heterogeneity observed in mountainous environments, (2) the reduction in the accuracy of reanalysis datasets when uncertain meteorological forcing, especially precipitation, are used,⁵⁸ and (3) the failure of interpolation methods to correctly capture the spatial dynamic of SWE when point stations are not representative of the surroundings.⁵⁹

Matching among anomaly products seems to be further reduced during both the early and late seasons, suggesting a possible relationship with the capability of reproducing wet snow dynamics and correctly modeling snow melting, both of which can have different impacts on the capability to capture dry or warm snow droughts. Further validation exercises during those periods are needed to better understand these ramifications.

Unfortunately, a detailed validation of coarse spatial resolution satellite and reanalysis datasets against spatially limited SWE observations seems unfeasible in complex terrain due to the limited representativeness of surface observations in areas with strong elevation gradients.⁵⁵ This situation makes it difficult to understand which approach is potentially better suited for mountain regions. Overall, the difference observed among SWE estimates suggests that more work must be done to characterize water resources in snow-dominated regions, particularly in mountains.⁶⁰

The observed difficulties in reliably modeling the interannual fluctuations in SWE over mountainous terrain and the potential major role of snow drought as a predictor of spring drought for snow-dominated regions,⁶¹ pose strong limitations on the capability to correctly model snow drought over Europe within an operational monitoring framework. Although due to the limited spatial extension of mountain areas the detected inconsistencies seem to have only a small effect on the overall agreement of the three models, the actual impact in an operational drought monitoring system can be much more substantial considering the important role of mountain regions as a temporal hydrological buffer in the form of snow and ice.

The intricate morphology of mountain terrains and the limitation in the spatial resolution of all of the currently available operational SWE products suggest the need for *ad hoc* solutions to correctly characterize the Alpine⁶² or similarly complex environments. This issue has been also highlighted in the recent ESA Climate Change Initiative for the construction of a satellite snow climate data record, which currently lacks coverage of alpine environments.⁶³ Due to the importance of mountain snow coverage and SWE for the hydrology of several big river basins in Europe, the observed major inconsistencies across the three models highlight a major limitation for the use of currently available near-real-time SWE products in the operational set up of EDO, until satisfactory improvements in remote sensing, modeling, and observations are achieved over mountain regions. In this regard, even the combined use of different products, which have demonstrated improved performances over single datasets,⁵⁵ may still be insufficient for operational applications in many cases.

Acknowledgments

The authors would like to thank the EUMETSAT Satellite Application Facility on Support to Operational Hydrology and Water Management, the European Centre for Medium-Range Weather Forecasts, and the Canadian Meteorological Centre for making their data available to the public. All outputs of this study can be requested through the European Drought Observatory.³³

References

1. D. Wiesnet, "Winter show drought," *EOS Trans. Am. Geophys. Union* **62**(14), 137–137 (1981).
2. K. G. Smith, "Drought-induced changes in avian community structure along a Montane Sere," *Ecology* **63**(4), 952–962 (1981).
3. D. A. Robinson, F. T. Keimig, and K. F. Deway, "Recent variation in Northern Hemisphere snow cover," in *Proc. 15th Annu. Clim. Diagn. Workshop*, NOAA, pp. 219–224 (1991).

4. J. R. Dierauer, D. M. Allen, and P. H. Whitfield, "Snow drought risk and susceptibility in the western United States and southwestern Canada," *Water Resour. Res.* **55**(4), 3076–3091 (2019).
5. W. Xu et al., "Spatial–temporal variability of snow cover and depth in the Qinghai–Tibetan Plateau," *J. Clim.* **30**, 1521–1533 (2017).
6. E. A. Spoles, T. R. Roth, and A. W. Nolin, "Future snow? A spatial-probabilistic assessment of the extraordinarily low snowpacks of 2014 and 2015 in the Oregon Cascades," *Cryosphere* **11**(1), 331–341 (2017).
7. A. J. Collados-Lara, E. Pardo-Igúzquiza, and D. Pulido-Velazquez, "A distributed cellular automata model to simulate potential future impacts of climate change on snow cover area," *Adv. Water Resour.* **124**, 106–119 (2019).
8. B. J. Hatchett and D. J. McEvory, "Exploring the origins of snow drought in the northern Sierra Nevada, California," *Earth Interact.* **22**, 1–13 (2018).
9. A. A. Harpold, M. Dettinger, and S. Rajagopal, "Defining snow drought and why it matters," *EOS Trans. Am. Geophys. Union* **98** (2017).
10. A. F. Van Loon et al., "Hydrological drought types in cold climates: quantitative analysis of causing factors and qualitative survey of impacts," *Hydrol. Earth Syst. Sci.* **19**, 1993–2016 (2015).
11. J. Barichivich et al., "Temperature and snow-mediated moisture controls of summer photosynthetic activity in northern terrestrial ecosystems between 1982 and 2011," *Remote Sens.* **6**(2), 1390–1431 (2014).
12. A. F. Hamlet et al., "Effects of temperature and precipitation variability on snowpack trends in the western United States," *J. Clim.* **18**, 4545–4561 (2005).
13. M. C. Serreze et al., "Characteristics of the western United States snowpack from snowpack telemetry (SNOTEL) data," *Water Resour. Res.* **35**(7), 2145–2160 (1999).
14. K. M. Larson et al., "Can we measure snow depth with GPS receivers?" *Geophys. Res. Lett.* **36**, L17502 (2009).
15. J. L. McCreight, E. E. Small, and K. M. Larson, "Snow depth, density, and SWE estimates derived from GPS reflection data: validation in the western U.S.," *Water Resour. Res.* **50**, 6892–6909 (2014).
16. L. Egli, T. Jonas, and R. Meister, "Comparison of different automatic methods for estimating snow water equivalent," *Cold Regions Sci. Technol.* **57**, 107–115 (2009).
17. D. Dixon and S. Boon, "Comparison of the SnowHydro snow sampler with existing snow tube designs," *Hydrol. Process.* **26**, 2555–2562 (2012).
18. M. P. Clark et al., "Representing spatial variability of snow water equivalent in hydrologic and land-surface models: a review," *Water Resour. Res.* **47**(7), W07539 (2011).
19. T. J. Schmugge et al., "Remote sensing in hydrology," *Adv. Water Resour.* **25**, 1367–1385 (2002).
20. M. Tedesco and J. Jeyaratnam, "A new operational snow retrieval algorithm applied to historical AMSR-E brightness temperatures," *Remote Sens.* **8**, 1037 (2016).
21. N. Saberi et al., "Review of snow water equivalent retrieval methods using spaceborne passive microwave radiometry," *Int. J. Remote Sens.* **41**(3), 996–1018 (2020).
22. L. De Gregorio et al., "Improving SWE estimation by fusion of snow models with topographic and remotely sensed data," *Remote Sens.* **11**(17), 2033 (2019).
23. K. J. Bormann et al., "Estimating snow-cover trends from space," *Nat. Clim. Change* **8**, 924–928 (2018).
24. M. Takala et al., "Estimating northern hemisphere snow water equivalent for climate research through assimilation of space-borne radiometer data and ground-based measurements," *Remote Sens. Environ.* **115**(12), 3517–3529 (2011).
25. R. Kelly et al., "A prototype AMSR-E global snow area and snow depth algorithm," *IEEE Trans. Geosci. Remote Sens.* **41**(2), 230–242 (2003).
26. T. Markus, D. Powell, and J. Wang, "Sensitivity of passive microwave snow depth retrievals to weather effects and snow evolution," *IEEE Trans. Geosci. Remote Sens.* **44**, 68–77 (2006).
27. A. Walker and B. Goodison, "Discrimination of a wet snow cover using passive microwave satellite data," *Ann. Glaciol.* **17**, 307–311 (1993).

28. S. Hancock et al., "Evaluating global snow water equivalent products for testing land surface models," *Remote Sens. Environ.* **128**, 107–117 (2013).
29. N. Dawson, P. Broxton, and X. Zeng, "Evaluation of remotely sensed snow water equivalent and snow cover extent over the contiguous United States," *J. Hydrometeorol.* **19**, 1777–1791 (2018).
30. C. M. Vuyovich, J. M. Jacobs, and S. F. Daly, "Comparison of passive microwave and modeled estimates of total watershed SWE in the continental United States," *Water Resour. Res.* **50**(11), 9088–9102 (2014).
31. J. Pulliainen, "Mapping of snow water equivalent and snow depth in boreal and sub-arctic zones by assimilating space-borne microwave radiometer data and ground-based observations," *Remote Sens. Environ.* **101**(2), 257–269 (2006).
32. N. Rutter et al., "Evaluation of forest snow processes models (SnowMIP2)," *J. Geophys. Res. Space Phys.* **114**, D06111 (2009).
33. The Copernicus Emergency Management Service Drought Observatories, "European drought observatory," <https://edo.jrc.ec.europa.eu> (accessed January 2022).
34. Eumetsat H-SAF, "Snow water equivalent by MW radiometry," <https://hsaf.meteoam.it/Products/Detail?prod=H13> (accessed January 2022).
35. Finnish Meteorological Institute, "GlobSnow," <https://www.globsnow.info> (accessed January 2022).
36. H-SAF, "Product User Manual (PUM) for product H13-SN-OBS-4: snow water equivalent by MW radiometry," Eumetsat, 2018, <1://hsaf.meteoam.it/Products/ProductsList?type=snow> (accessed March 2021).
37. NCAR-UCAR, "USGS 30 ARC-second Global Elevation Data, GTOPO30," <https://rda.ucar.edu/datasets/ds758.0/> (accessed January 2022).
38. P. Lahtinen et al., "Merging flat/forest and mountainous snow products for extended European area," in *IEEE Int. Geosci. Remote Sens. Symp.*, Cape Town, pp. II-563–II-566 (2009).
39. J. Pulliainen, J. Grandell, and M. T. Hallikainen, "HUT snow emission model and its applicability to snow water equivalent retrieval," *IEEE Trans. Geosci. Remote Sens.* **37**, 1378–1390 (1999).
40. H-SAF, "Product Validation Report (PVR) for product H13—SN-OBS-4: snow water equivalent by MW radiometry," Eumetsat, 2012, <http://hsaf.meteoam.it/Products/ProductsList?type=snow> (accessed March 2021).
41. National Snow and Ice Data Center, "Canadian Meteorological Centre (CMC) daily snow depth analysis data, version 1," <https://nsidc.org/data/NSIDC-0447/versions/1> (accessed January 2022).
42. R. D. Brown and B. Brasnett, *Canadian Meteorological Centre (CMC) Daily Snow Depth Analysis Data, Version 1*, NASA National Snow and Ice Data Center Distributed Active Archive Center, Boulder, Colorado (2010).
43. M. Sturm, J. Holmgren, and G. E. Liston, "A seasonal snow cover classification system for local to global applications," *J. Clim.* **8**, 1261–1283 (1995).
44. R. D. Brown and P. W. Mote, "The response of northern hemisphere snow cover to a changing climate," *J. Clim.* **22**, 2124–2145 (2009).
45. B. Brasnett, "A global analysis of snow depth for numerical weather prediction," *J. Appl. Meteorol.* **38**, 726–740 (1999).
46. E. Dutra et al., "An improved snow scheme for the ECMWF land surface model: description and offline validation," *J. Hydrometeorol.* **11**, 899–916 (2010).
47. H. Hersbach et al., "The ERA5 global reanalysis," *Q. J. R. Meteorol. Soc.* **146**(730), 1999–2049 (2020).
48. P. de Rosnay, L. Isaksen, and M. Dahoui, "Snow data assimilation at ECMWF," *ECMWF Newsllett.* **143**, 26–31 (2015).
49. Copernicus Climate Change Service, "ERA5 monthly averaged data on single levels from 1979 to present," <https://cds.climate.copernicus.eu/cdsapp#!/dataset/reanalysis-era5-single-levels-monthly-means?tab=overview> (accessed January 2022).
50. World Meteorological Organization, "Standardized precipitation index user guide," M. Svoboda, M. Hayes and D. Wood, WMO-No. 1090, Geneva (2012).

51. M. L. Wrzesien et al., "A new estimate of North American mountain snow accumulation from regional climate model simulations," *Geophys. Res. Lett.* **45**, 1423–1432 (2108).
52. M. Guidicelli et al., "Continuous spatio-temporal high-resolution estimates of SWE across the Swiss Alps: a statistical two-step approach for high-mountain topography," *Front. Earth Sci.* **9**, 664648 (2021).
53. E. H. Bair et al., "Validating reconstruction of snow water equivalent in California's Sierra Nevada using measurements from the NASA Airborne Snow Observatory," *Water Resour. Res.* **52**(11), 8437–8460 (2016).
54. M. P. Molotch and R. C. Bales, "Scaling snow observations from the point to the grid element: implications for observation network design," *Water Resour. Res.* **41**, W11421 (2005).
55. C. Mortimer et al., "Evaluation of long-term Northern Hemisphere snow water equivalent products," *Cryosphere* **14**, 1579–1594 (2020).
56. L. R. Mudryk et al., "Characterization of Northern Hemisphere snow water equivalent datasets, 1981–2010," *J. Clim.* **28**(20), 8037–8051 (2015).
57. J. Dozier, E. H. Bair, and R. E. Davis, "Estimating the spatial distribution of snow water equivalent in the world's mountains," *WIREs Water* **3**(3), 461–474 (2016).
58. A. Langlois et al., "Simulation of snow water equivalent (SWE) using thermodynamic snow models in Québec, Canada" *J. Hydrometeorol.* **10**(6), 1447–1463 (2009).
59. J. Lundquist et al., "Our skill in modeling mountain rain and snow is bypassing the skill of our observational networks" *Bull. Am. Meteorol. Soc.* **100**(12), 2473–2490 (2019).
60. M. L. Wrzesien et al., "Characterizing biases in mountain snow accumulation from global data sets," *Water Resour. Res.* **55**, 9873–9891 (2019).
61. T. P. Barnett, J. C. Adam, and D. P. Lettenmaier, "Potential impacts of a warming climate on water availability in snow-dominated regions," *Nature* **438**, 303–309 (2005).
62. M. Dall'Amico, S. Endrizzi, and S. Tasin, "MySnowMaps: operative, high-resolution real-time snow mapping," in *Proc. Int. Snow Sci. Workshop*, Innsbruck (2018).
63. K. Luojus et al., *ESA Snow Climate Change Initiative (Snow_cci): Snow Water Equivalent (SWE) level 3C Daily Global Climate Research Data Package (CRDP) (1979–2018), Version 1.0*, Centre for Environmental Data Analysis (2020).

Carmelo Cammalleri received his PhD in hydraulic and environmental engineering from the University of Palermo in 2011. He is a contractual agent at the Joint Research Centre of the European Commission (JRC). He is the author of about 40 papers in peer-reviewed journals covering the topics of hydrology, remote sensing, and risk analysis. His current research interests include the spatial distributed estimation of hydrological variables using both land-surface modeling and remote sensing data, with a focus on drought monitoring and early warning.

Paulo Barbosa received his PhD in forestry from the Technical University of Lisbon in 2000. He is a senior scientist at the JRC. He is currently leading the development of the European and Global Drought Observatories at JRC as well as supporting research on climate change adaptation related to the Covenant of Mayors for Climate and Energy initiative of the European Commission. His current research interests include drought risk and drought impact analysis as well as climate change impacts and adaptation.

Jürgen V. Vogt received his PhD in physical geography, remote sensing, and climatology from Trier University in 1992. He is a senior scientist at the JRC. Among other projects, he led the development of the European and Global Drought Observatories at JRC. His current research interests include terrain and land-use/land cover analysis as well as drought risk analysis and monitoring.