POLITECNICO MILANO 1863

SCUOLA DI INGEGNERIA INDUSTRIALE E DELL'INFORMAZIONE

Implementation of a Multi-Sensor Algorithm for Time Series Snow Cover Mapping in Lombardy, Italy

Geoinformatics Project

Authors:
Hadi Kheiri Gharajeh - 10946669
Ola Elwasila Abdelrahman Yousif – 10963045

Advisor: Prof. Giovanna Venuti and Prof. Daniela Stroppiana

Academic Year: 2025

Contents

Contents 2
Introduction 4
Content and Motivation 4
Objectives: 5
1.1. Dataset Summary: 5
2 Related Work 7
3 Comparative View 8
4 Innovation and Contribution 9
5 Project Development 10
5.1. Overview and workflow 10
5.2. Preprocessing: 11
5.2.1. Resampling and Reprojecting: 11
5.2.2. Weekly Aggregation and Alignment: 12
5.3. Comparison and Agreement Assessment: 15
5.3.1. Pairwise Agreement: 15
5.3.2. Snow Consensus Mapping: 16
5.3.3. Weekly Snow Area statistics: 18
5.3.4. Area Bias 19
5.3.5. Spatial Discrepancy Maps 20
5.3.6. Sentinel-2 vs Sentinel-3 Agreement 23
5.4. Algorithm Implementation 25
5.4.1. Random Forest 25
5.4.2. The "Let-It-Snow" (LIS) Approach: 27
5.4.3. Regression-Based Comparison with MODIS 29
6 Conclusion 31
7 Reference 32
Project Link in GitHub 33

Figure 1: Pairwise Agreement Over Time 15
Figure 2: Snow Consensus 17
Figure 3: Weekly Area Bias 19
Figure 4:MODIS vs. GFSC W11-2022 21
Figure 5: S2 vs. GFSC W11-2022 21
Figure 6: MODIS vs. GFSC W08-2023 22
Figure 7: S2 vs. GFSC W08- 2023 23
Figure 8: S2 vs S3 24

Table 1: Dataset Description 6
Table 2: Weekly Snow Statistics (Sample) 18
Table 3: Area Bias 20
Table 4: Performance Metrics of the Random Forest Classifier Using Combined Predictors 26
Table 5: Performance Metrics of RF for Single-Product Classification 27
Table 6: Performance Metrics of LIS 29
Table 7: Regression Output 30

Introduction

Content and Motivation

Snow cover plays a key role in hydrological processes, mountain ecology, and global energy balance, due to its high albedo and seasonal dynamics. Monitoring its spatio-temporal evolution is essential for understanding runoff, plant and animal phenology, soil erosion, and water management strategies in alpine regions (Revuelto et al., 2021).

MODIS imagery has provided global-scale daily snow cover observations since 2001 with a spatial resolution of 500 meters, enabling the development of long-term snow climatologies (Revuelto et al., 2021). However, this resolution is insufficient for local-scale hydrological or ecological analysis, especially in areas with complex terrain and variable snow distribution patterns below 100 meters (Gascoin et al., 2019) .

The Sentinel-2 mission, launched in 2015, addresses this limitation by offering high-resolution optical imagery (up to 20 meters) across visible and near-infrared bands. Despite its finer resolution and 5-day revisit cycle, Sentinel-2 has a relatively short historical record and is susceptible to cloud cover, which limits its direct use for long-term trend analysis (Revuelto et al., 2021).

To bridge the gap between MODIS's temporal continuity and Sentinel-2's spatial detail, this project investigates methods for integrating and downscaling MODIS observations using Sentinel-2 products. The aim is to generate high-resolution snow cover maps that retain MODIS's temporal depth while leveraging Sentinel-2's spatial granularity.

Specifically, this work:

- Compares multiple snow cover products (MODIS, GFSC, S2, S3) over the Lombardy Alps.
- Develops a regression-based approach to relate MODIS fractional snow cover to high-resolution Sentinel-2 snow masks (Revuelto et al., 2021; Alonso-González et al., 2022) .
- Implements and evaluates logic-based and machine-learning-based classifiers to enhance snow cover mapping performance (Fugazza et al., 2019).

Objectives:

Harmonize, resample, and align snow products from MODIS, Sentinel-2, Sentinel-3, and GFSC over Lombardy.

Train machine learning models (Random Forest) to estimate snow presence using EO inputs.

Evaluate product agreement, accuracy, and correlation through regression, logic-based fusion, and model prediction.

1.1. Dataset Summary:

The following datasets were analyzed:

MODIS Snow Mask (500m): Thematic classification by EURAC with irregular frequency.

ARPA GFSC (60m): Weekly gap-filled product with fractional snow cover values (0–205), requiring quality filtering.

Sentinel-2 (20m): Biweekly binary snow masks with associated date maps for snow detection chronology.

Sentinel-3 (300m): Weekly binary snow masks with a date map index (0–3).

Note: All datasets used in this project were selected and processed strictly in accordance with the technical requirements defined by the **National Research Council of Italy (CNR)**. These requirements dictated the temporal resolution, spatial extent, and preprocessing protocols. The data themselves were officially provided by **ARPA Lombardia**, the regional environmental protection agency. Every step in our dataset preparation and usage — from product selection (MODIS, GFSC, Sentinel-2, Sentinel-3) to alignment and weekly aggregation — was carried out to fully meet the expectations and operational constraints communicated by the CNR and validated against ARPA-supplied ground truth references. As such, the project is fully aligned with institutional standards and intended downstream applications.

Table 1: Dataset Description

Product	Sensor	Resolution	Value Range	Temporal Res.	Notes
MODIS Snow Mask	MODIS	500 m	1–2	~5–10 days (irregular)	Thematic classes (e.g., 1 = snow)
GFSC – GF Layer	MODIS/S2	60 m	0–205	Weekly (7-day)	Gap-filled snow index (scaled)
GFSC – QC Layer	MODIS/S2	60 m	0–205	Weekly (7-day)	Pixel-level quality mask
Sentinel-2 Snow Mask	Sentinel-2	20 m	1 (binary)	Bi-weekly (14-day)	Binary snow presence
Sentinel-2 Date Map	Sentinel-2	20 m	1–14	Bi-weekly (14-day)	Acquisition index (ordinal)
Sentinel-3 Snow mask	Sentinel-3	300 m	1 (binary)	Weekly (7-day)	Binary snow presence
Sentinel-3 Date Map	Sentinel-3	300 m	0–3	Weekly (7-day)	Indicates which acquisition saw snow

2 Related Work

Monitoring snow cover through remote sensing has been extensively studied in the past two decades. MODIS (Moderate Resolution Imaging Spectroradiometer) has been one of the most widely used instruments for this purpose, thanks to its high temporal frequency and global coverage. However, several studies have highlighted its limitations in mountainous regions due to its coarse spatial resolution and cloud contamination issues (Gascoin et al., 2019).

To overcome these challenges, hybrid snow mapping approaches have been developed. For instance, the **GlobSnow (GFSC)** product combines passive microwave and optical measurements to improve snow cover estimates, especially in forested or cloudy regions. The GFSC dataset has been evaluated and validated across various climate zones and topographic conditions (Gascoin et al., 2019).

On the high-resolution end, **Sentinel-2** has emerged as a promising alternative for snow detection due to its 10–20 m spatial resolution and multispectral capabilities. Nonetheless, its low temporal revisit rate and cloud dependency make standalone applications problematic. As a result, several studies propose using Sentinel-2 in combination with MODIS to improve spatial detail while maintaining consistent temporal coverage. A key example is the downscaling of MODIS snow cover to Sentinel-2 scale using **Random Forest regression models**, which showed strong improvements in fractional snow estimation over complex terrain (Fiddes et al., 2019).

A recent trend is the use of **data fusion and ensemble learning** techniques to merge outputs from different snow products. Revuelto et al. (2021) introduced a **"Let-It-Snow" logic-based voting method**, where snow presence is determined based on agreement among several products. This method reduces the impact of sensor-specific biases and helps fill gaps caused by cloud cover or missing data (Revuelto et al., 2021).

Our work draws inspiration from these advances and implements a complete multi-source processing and validation pipeline using MODIS as reference and Sentinel-2, Sentinel-3, and GFSC as predictors. The methodology integrates regression, classification, and logical fusion, contributing to the growing body of literature on high-resolution snow cover mapping in alpine environments.

3 Comparative View

This project evaluates and compares multiple Earth Observation (EO) snow products over the Lombardy region, using **MODIS** as a reference due to its extensive coverage and public availability. The three predictor datasets considered — **GFSC**, **Sentinel-2**, and **Sentinel-3** — were chosen based on their complementary spatial and temporal characteristics.

MODIS Snow Mask (500 m): A binary thematic product provided by EURAC, based on snow/no-snow classification with irregular temporal frequency (~5–10 days). It is used as the **reference layer** in both classification and regression analyses.

GFSC (60 m): The ARPA GFSC product combines **Sentinel-1 SAR** and **Sentinel-2 optical** data to generate weekly **fractional snow cover** maps (values: 0–205). The integration improves performance under cloudy conditions and in forested terrain. A **quality control (QC) layer** is used to mask unreliable pixels. GFSC is gap-filled but not directly binary, so additional binarization and QC-based filtering are required before comparison.

Sentinel-2 (20 m): A binary snow mask (SnowMask_latest) built from a 14-day observation window. An additional datemap layer encodes the acquisition date. It offers **high spatial resolution** for detailed mapping but is limited by **cloud cover** and **lower revisit rate**.

Sentinel-3 (300 m): Similar in structure to S2, but based on a 7-day period. It provides moderate resolution with more frequent updates (~3–4 images per week), making it a **balance** between S2's detail and MODIS's frequency.

4 Innovation and Contribution

This project delivers a unified and modular pipeline for harmonizing, comparing, and evaluating multi-sensor snow cover datasets over the Lombardy Alps. Its innovation stems from the integration of heterogeneous Earth Observation (EO) products using a reproducible and scalable workflow — combining machine learning, logic-based fusion, and statistical regression.

Key contributions include:

Multi-sensor spatial-temporal harmonization: All four snow products (MODIS, GFSC, Sentinel-2, Sentinel-3) were resampled to a 120 m grid, temporally aligned over 23 winter weeks, and clipped to a common alpine region of interest — ensuring direct pixel-level comparison. This step required careful preprocessing (resampling, reprojection, filtering) and was implemented in a fully automated batch-processing pipeline `snow_processing.py, main.ipynb.`

Dual snow classification strategies:

A **Random Forest model** was trained to classify snow/no-snow using GFSC, S2, and S3 as predictors and MODIS as ground truth. This supervised model achieved **94% accuracy** and **F1 = 0.83** for snow detection.

A **Let-It-Snow-style fusion approach** used majority voting among the binary products, yielding an alternative logic-based snow mask. Though simpler, it achieved comparable results with slightly lower accuracy.

Pixel-level regression over 23 weeks: A large-scale, memory-safe regression routine was used to compute correlation and slope values between each product and MODIS for 500,000 sampled pixels per week. This enabled **quantitative comparison of agreement trends** over time

Interpretability and reproducibility: The modular design (with run_regression_batch, compare_classifiers, and harmonize_week functions) allows the pipeline to be reused for any region, any time period, and any set of products — extending its utility beyond this study

Transparency of data limitations: The project does not rely on raw spectral bands, and both ML and logic-based methods were shown to produce **equivalent results** due to the binary nature of inputs — a finding that reinforces the role of product-level uncertainty in snow classification.

5 Project Development

5.1. Overview and workflow

The project establishes a comprehensive processing framework for harmonizing, analyzing, and classifying multi-source snow cover data over the Lombardy region. The methodology is organized into two primary phases: (i) preprocessing and comparison of existing snow cover products, and (ii) implementation and evaluation of custom classification algorithms.

The preprocessing phase ensures both spatial and temporal harmonization of four key datasets: MODIS, GFSC, Sentinel-2, and Sentinel-3. All datasets are reprojected to a common spatial reference system (EPSG:32632) and resampled to a unified 60 m grid. Temporal alignment is achieved by aggregating snow masks to the ISO-week level, facilitating inter-product comparisons. Each dataset undergoes a tailored preparation process to address its unique format and characteristics:

MODIS: Daily snow masks are processed using `reproject_resample_visualize()` followed by `aggregate_weekly()`, which applies a temporal maximum to create weekly binary composites.

GFSC: Snow products derived from Sentinel-1 and Sentinel-2 inputs are preprocessed with `resample_reproject_gfsc()` and grouped weekly via `aggregate_weekly_gfsc()`, preserving quality flags and masking invalid data.

Sentinel-2: Individual masks are associated with 14-day composite identifiers and processed with `process_s2_weekly()` to split and aggregate them into weekly layers.

Sentinel-3: Daily masks are grouped by ISO week based on acquisition dates and processed with `reproject_s3_weekly()` using a float-to-binary conversion and maximum value selection, while filtering out invalid values.

Following the harmonization step, the second phase implements and evaluates two snow classification strategies:

A **machine learning model** based on the Random Forest algorithm is trained to reproduce MODIS binary snow cover using GFSC, Sentinel-2, and Sentinel-3 as explanatory variables. Performance is assessed using standard classification metrics.

A **logic-based fusion approach**, inspired by the Let-It-Snow methodology, combines the three inputs using a majority voting scheme to estimate weekly snow presence.

Finally, a **regression module** performs consistency analysis across 23 selected weeks, comparing each product against MODIS using linear fitting and correlation metrics. This dual-phase approach results in a robust and modular pipeline for multi-source snow monitoring, enabling both operational deployment and scientific validation.

5.2. Preprocessing:

5.2.1. Resampling and Reprojecting:

The initial stage of preprocessing was dedicated to spatial harmonization across all input datasets to enable pixel-wise comparison and subsequent integration within the classification pipeline. To this end, all products were reprojected to a common coordinate reference system—**EPSG:32632 (UTM Zone 32N)**—and resampled to a target resolution of **60 meters**, selected to match the native spatial scale of the GFSC product.

This reprojection and resampling process was tailored to the specific characteristics of each dataset, ensuring preservation of their semantic and numeric integrity while enabling downstream alignment:

MODIS: Originally provided at a 500 m resolution, MODIS snow cover masks were downsampled to 60 m using *nearest-neighbor interpolation* to retain binary integrity (0: no snow, 1: snow, 255: nodata). This transformation was executed via the `reproject_resample_visualize()` function in the `snow_processing.py` module.

GFSC: As a fusion product combining Sentinel-1 and Sentinel-2 inputs, GFSC is natively at 60 m but includes associated quality control layers. The `resample_reproject_gfsc()` function was employed to:

Filter out pixels not meeting acceptable quality thresholds (acceptable_qc = {0,1,2,3}),

Resample fractional snow values using *bilinear interpolation*, and

Resample QC flags using *nearest-neighbor*. Invalid or low-quality pixels were set to NaN or recoded as 255 to indicate nodata.

Sentinel-2: Provided as binary masks accompanied by date identifiers, Sentinel-2 scenes were grouped into ISO-week periods and processed using `process_s2_weekly()`. This function extracted and resampled each mask to the unified 60 m grid using *nearest-neighbor interpolation*, preserving binary class boundaries. All invalid or missing values were uniformly set to nodata = 255.

Sentinel-3: Delivered as daily snow masks in float32 format, Sentinel-3 data were first grouped by ISO-week and then aggregated temporally using a max-composite operation. The aggregated weekly outputs were reprojected to the 60 m grid via *nearest-neighbor* resampling using the `reproject_s3_weekly()` function, while values exceeding valid thresholds were masked.

This standardized reprojection-resampling protocol ensured that all input products shared a common spatial footprint, projection, resolution, and nodata handling—forming a foundational prerequisite for reliable temporal aggregation, classification, and regression tasks in the subsequent phases of the workflow.

5.2.2. Weekly Aggregation and Alignment:

After harmonizing spatial resolution and projection, the second stage of preprocessing addressed the temporal dimension by aggregating all datasets into consistent ISO-weekly snow masks. This step was essential for enabling direct comparisons across products with varying revisit periods, acquisition formats, and data conventions. Each product required a tailored weekly aggregation procedure, followed by alignment to a unified grid for inter-product compatibility.

MODIS

MODIS daily snow masks, originally at 500 m resolution, were first reprojected and down-sampled to the target 60 m grid. The function `aggregate_weekly()` was then applied to stack all daily masks within each ISO week and compute a weekly composite using the temporal maximum operation. A pixel was classified as snow-covered if it exhibited snow on at least one day in the week. The binary convention (0 = no snow, 1 = snow) was preserved, with 255 indicating nodata.

Function: `aggregate_weekly()`

Method: temporal max across binary stack

Key Features: preserves nodata (255); snow presence if any day is snowy

GFSC

GFSC combines Sentinel-1 and Sentinel-2 data with multiple quality levels. After resampling and filtering using `resample_reproject_gfsc()`, valid snow fraction maps were binarized and grouped by ISO week using date information parsed from filenames (format YYYYMMDD). The function `aggregate_weekly_gfsc()` applied a max operation across valid binary snow masks per week. Nodata (255) and invalid values were propagated.

Function: `aggregate_weekly_gfsc()`

Method: weekly max over filtered daily binary layers

Key Features: quality-controlled input; respects nodata masking

Sentinel-2

Sentinel-2 data was provided as bi-weekly composite masks tagged by acquisition date. Using the `process_s2_weekly()` function, each mask was split into ISO-week subsets based on the datemap metadata (values 1–7 for Week A, 8–14 for Week B). A temporal max-stack was computed within each week, and nodata regions (value > 1 or set by source) were preserved.

Function: `process_s2_weekly()`

Method: bi-weekly tag → ISO week split → max-stack

Key Features: binary masks; uses datemap; nodata=255

Sentinel-3

Sentinel-3 snow masks were delivered as daily float32 rasters. Weekly grouping was performed using ISO calendar week parsing from filenames. The reproject_s3_weekly() function stacked all valid float values for a week and computed the maximum value using `np.nanmax()`, after excluding invalid values (>1 or NaN). Masks were then reprojected to the 60 m common grid using nearest-neighbor interpolation.

Function: `reproject_s3_weekly()`

Method: weekly max with NaN-masking, followed by reprojection

Key Features: float32 input; invalid → NaN → max-stack → 255 for nodata

Weekly Availability

Although all products were successfully aligned in both space and time, the temporal overlap across datasets presented a significant constraint. No snow cover data was available from any product until **week 44 of 2021**, which marked the beginning of partial data availability. Full overlap across all four datasets—MODIS, GFSC, Sentinel-2, and Sentinel-3—only started from **week 2 of 2022 (2022_W02)**. In total, **23 common weeks** were identified where all products had valid, aligned data for the Lombardy region.

These common weeks are:

2022_W02, 2022_W04, 2022_W06–2022_W12, 2022_W44, 2022_W47, 2022_W49–2022_W52, 2023_W04, 2023_W06, 2023_W08–2023_W11, 2023_W14–2023_W15, 2023_W17.

This limited temporal intersection—only **23 out of 159 ISO weeks** in the considered period—reduced the dataset available for inter-product comparison, model training, and evaluation. Consequently, this sparsity in common weeks may have impacted the statistical robustness of agreement metrics, classifier accuracy, and regression results. Future improvements could be achieved by extending the temporal window or integrating additional sources to increase overlap.

5.3. Comparison and Agreement Assessment:

To evaluate the coherence and consistency among the snow cover products, a multi-level comparison was performed, including quantitative agreement metrics, weekly snow coverage statistics, and spatial consensus maps. This analysis focused on a subset of **23 common weeks** from 2022 and 2023, during which all four datasets (MODIS, GFSC, Sentinel-2, Sentinel-3) were simultaneously available and spatially aligned at 60 m resolution.

5.3.1. Pairwise Agreement:

We computed the weekly pairwise agreement between all product pairs using the function `compute_pairwise_agreement()` defined in `snow_processing.py`. The metric is based on pixel-wise correspondence across binary masks, where agreement is defined as both products assigning the same class (snow or no-snow) on the same pixel. This was repeated for all weeks and product combinations.

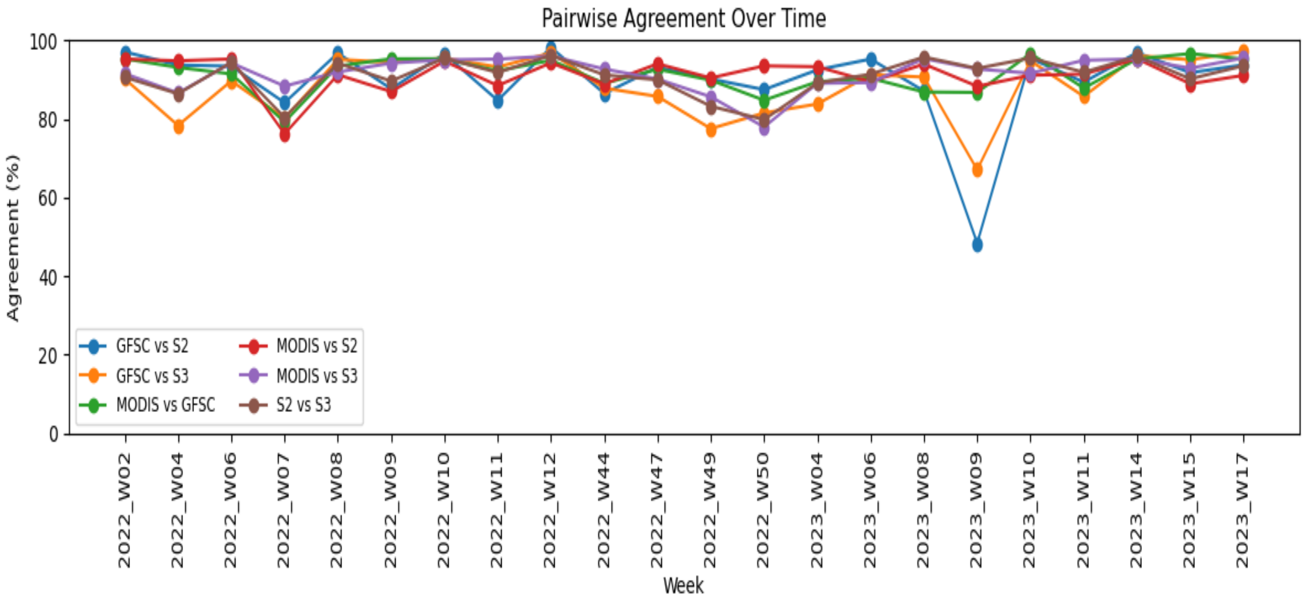

Figure 1: Pairwise Agreement Over Time

The resulting agreement percentages are plotted in **Figure 1**, which illustrates the temporal behavior of inter-product agreement over the 23-week period. Most product pairs maintained agreement rates above 90%, with the **GFSC–S2** and **S2–S3** pairs exhibiting the highest consistency throughout the two snow seasons. This observation aligns with expectations, given that GFSC integrates Sentinel-2 data, and both S2 and S3 are based on higher-resolution optical sensors.

An evident anomaly is observed in **week 2023_W09**, where multiple product pairs, particularly GFSC–S2 and GFSC–S3, experienced a sharp drop in agreement, falling below 50%. This may be attributed to missing acquisitions or increased noise in cloud or snow detection during that specific week. Apart from this outlier, the agreement trajectories remained relatively stable across time.

5.3.2. Snow Consensus Mapping:

In parallel, we generated weekly consensus maps to visually assess multi-product agreement. For each week, we computed the total number of products classifying a given pixel as snow using the function `compute_multisensor_snow_agreement()`.

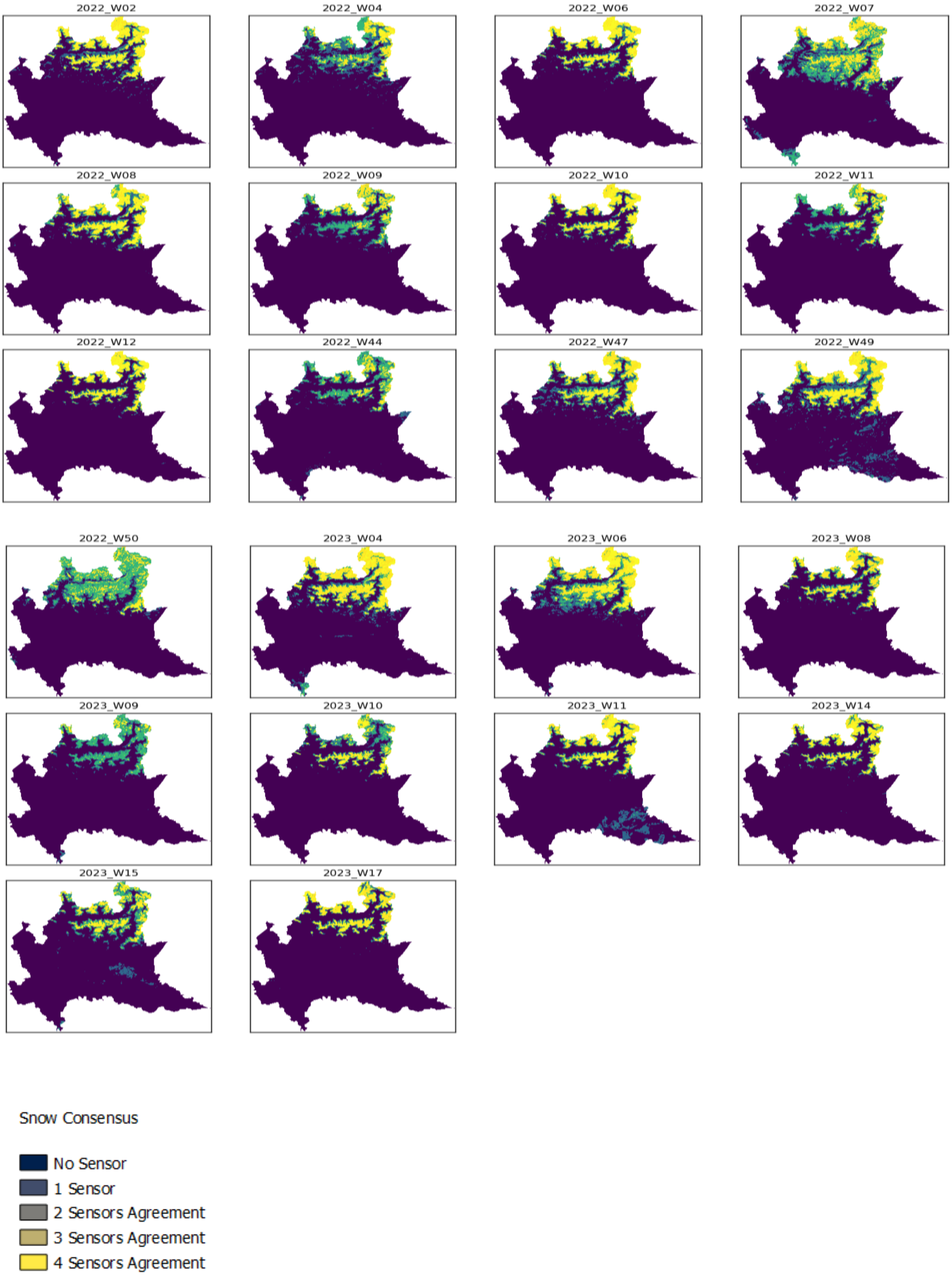

Figure 2: Snow Consensus

The resulting maps contain values from 0 (no product classified snow) to 4 (full agreement across all products). The consensus outputs for all 23 common weeks are presented in Figure 2

These visualizations reveal a clear spatial trend: regions with persistent snow cover in high altitudes, such as northern mountainous areas, frequently achieved full agreement (value = 4, in yellow). In contrast, areas with transient or marginal snow—typically located at lower elevations or along the snowline—showed more variability and lower consensus values.

Furthermore, inter-annual seasonal patterns are evident. For instance, widespread agreement is visible in 2022_W06–W11 and 2023_W09–W10, which correspond to peak accumulation periods. Meanwhile, weeks such as 2023_W15–W17 exhibit limited snow extent and consensus, indicating the onset of spring melt and the retreat of the snowpack.

5.3.3. Weekly Snow Area statistics:

To assess the magnitude of snow cover reported by each product, we extracted weekly statistics using the function `compute_weekly_statistics()`. For each product and week, the snow-covered area (in km²), total pixel count, and percentage of coverage were computed over the aligned grid. As shown in the table excerpt below, MODIS and Sentinel-3 (S3) consistently report larger snow-covered areas compared to Sentinel-2 (S2) and GFSC. For instance, in week 2022_W02, MODIS detected approximately 3,266.68 km² of snow-covered area, while GFSC reported only 1,033.08 km² for the same week.

Table 2: Weekly Snow Statistics (Sample)

product	week	total pixels	snow pixels	missing pixels	snow_area_km2	coverage
GFSC	2022_W02	3212289	286968	10767495	1033.0848	8.933443
MODIS	2022_W02	6628201	907411	7351583	3266.6796	13.690155
S2	2022_W02	6628201	722284	7351583	2600.2224	10.897135
S3	2022_W02	6628201	1016658	7351583	3659.9688	15.33837
GFSC	2022_W04	406417	111634	13573367	401.8824	27.467847

This pattern persists across multiple weeks and can be attributed to the differing spatial resolutions, sensing characteristics, and snow classification strategies of the products. MODIS, with its coarser 500 m resolution and broader cloud/snow discrimination criteria, tends to overestimate snow cover, especially in mixed or heterogeneous terrain. In contrast, GFSC applies stricter quality control and is based on higher-resolution inputs, leading to more conservative snow estimates.

5.3.4. Area Bias

To quantify spatial discrepancies between the snow products, we computed the area bias (in km²) for each pair of sensors across all common weeks. As illustrated in Figure 3, the bias values vary significantly over time, especially in comparisons involving MODIS and S3. In many weeks, MODIS vs S2 and MODIS vs S3 exhibit strong negative or positive biases respectively, often exceeding ±1,000 km². This confirms a systematic tendency of MODIS to overestimate snow extent relative to S2, while S3 often detects even more snow than MODIS, particularly in complex alpine terrain.

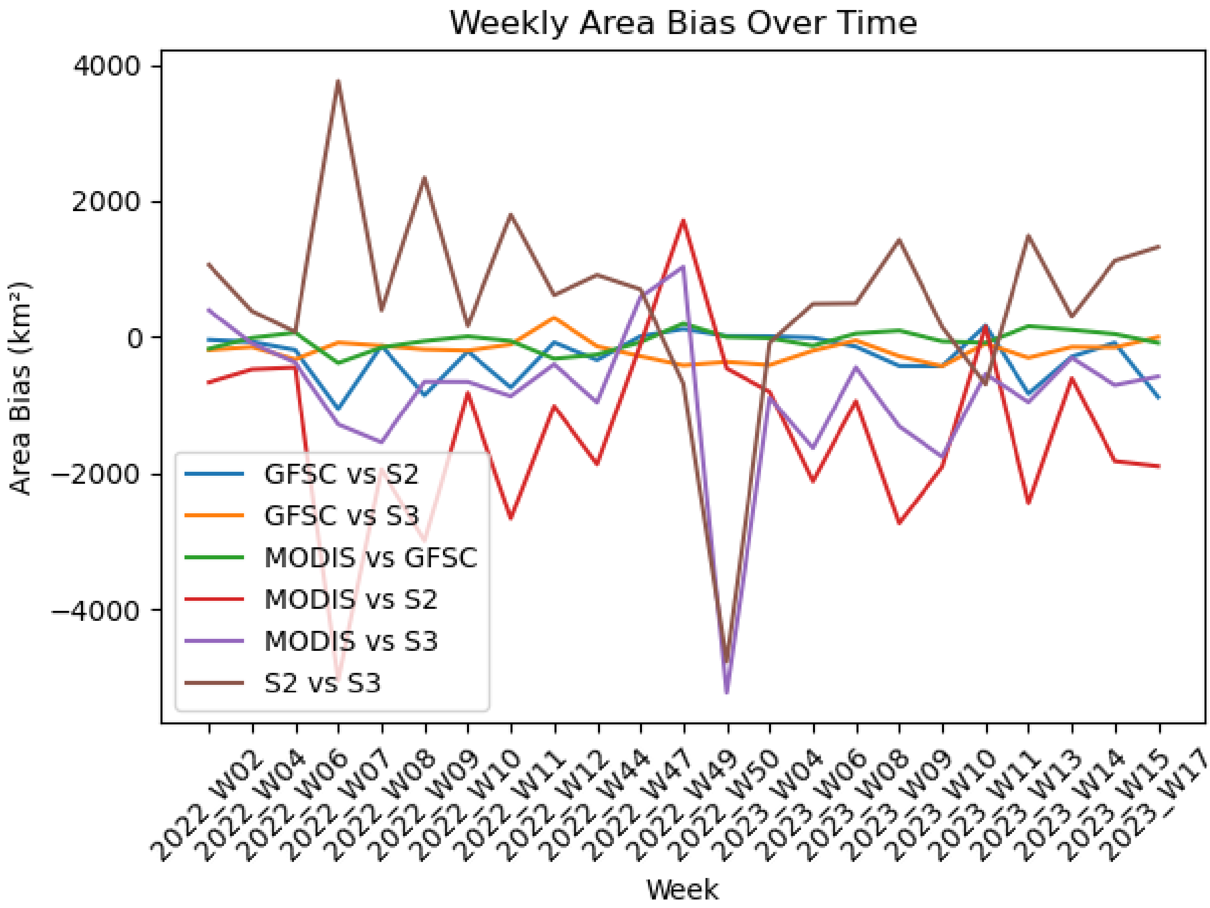

Figure 3: Weekly Area Bias

A detailed example is reported in the table below for week 2022_W02, showing that MODIS vs S2 had a negative bias of −666.46 km², while MODIS vs S3 had a positive bias of +393.29 km². Meanwhile, the GFSC vs S2 pair showed only −43.28 km², indicating high consistency. This reinforces the role of resolution and detection strategy: coarser products like MODIS and S3 are more prone to aggregating mixed or ambiguous pixels into snow, while higher-resolution products (S2 and GFSC) apply stricter detection criteria, leading to more conservative estimates.

Table 3: Area Bias

Week	Product 1	Product 2	Agreement (%)	Area Bias (km²)
2022_W02	GFSC	S2	98.35	−43.28
2022_W02	GFSC	S3	94.27	−192.92
2022_W02	MODIS	GFSC	97.76	−174.93
2022_W02	MODIS	S2	95.24	−666.46
2022_W02	MODIS	S3	91.54	+393.29

5.3.5. Spatial Discrepancy Maps

The maps in **Figure 4** refer to week **2022_W11**, a period characterized by moderate snow coverage. Notably, the MODIS vs GFSC comparison reveals substantial overestimation by MODIS, especially in the southern and central parts of the region. This discrepancy can be attributed to MODIS's coarser resolution and more permissive cloud-masking criteria, which lead to snow detection in mixed or cloudy pixels.

In contrast, the **S2 vs GFSC** difference map for the same week shows more localized disagreement, primarily along elevation gradients where partial snow coverage is more ambiguous. This reinforces the idea that both S2 and GFSC benefit from higher-resolution inputs and stricter cloud handling, resulting in closer alignment.

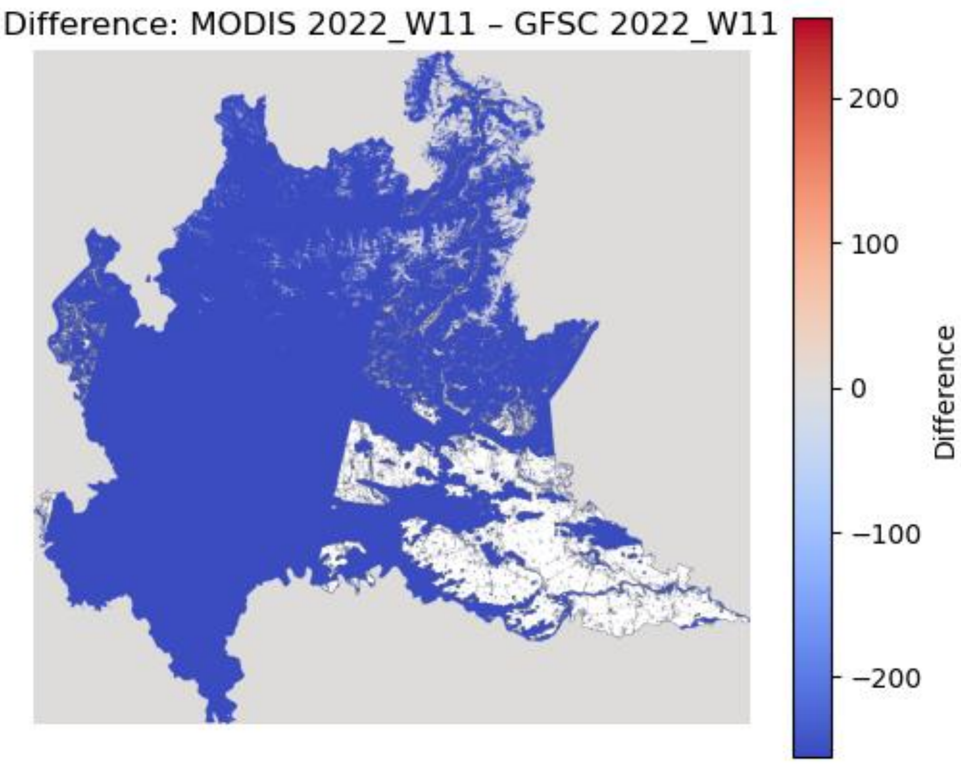

Figure 4:MODIS vs. GFSC W11-2022

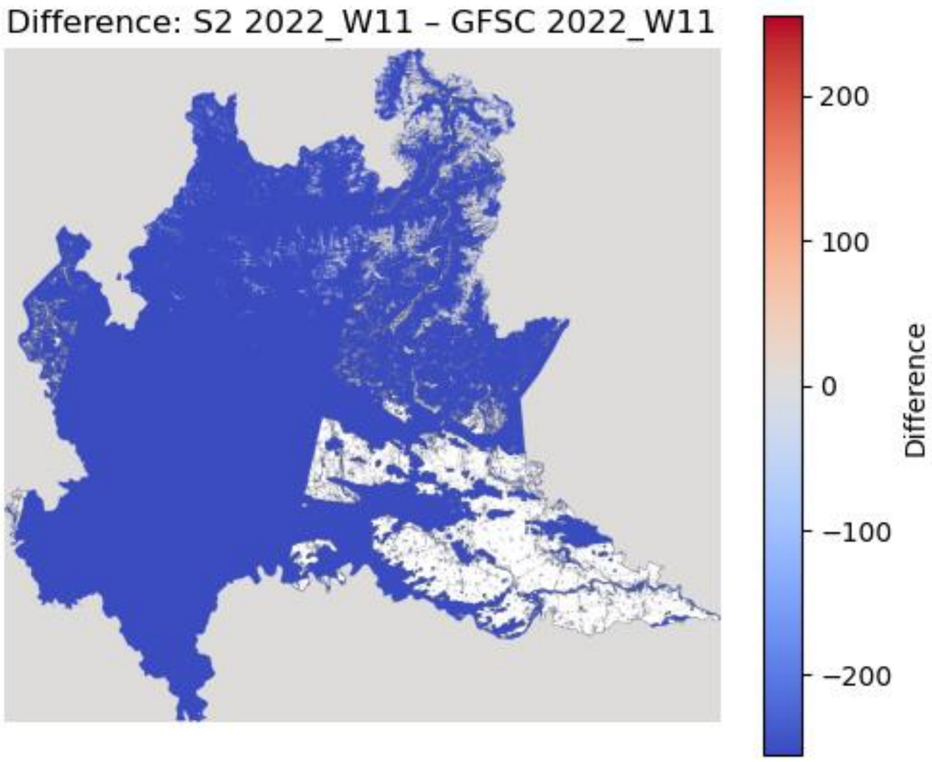

Figure 5: S2 vs. GFSC W11-2022

To further explore temporal consistency, we generated difference maps for week **2023_W08**, as shown in **Figure 6 and Figuer 7**. During this later snow season week, differences between MODIS and GFSC were considerably reduced, indicating a convergence in snow detection due to clearer atmospheric conditions and more extensive coverage. The **S2 vs GFSC** map also shows minimal discrepancies, highlighting a strong agreement between the high-resolution products under favorable conditions.

Overall, these visual analyses align with the trends observed in the area bias time series (Section 5.3.4), confirming that MODIS tends to overestimate snow presence in complex terrain or cloudy conditions, while GFSC and S2 show stronger spatial coherence.

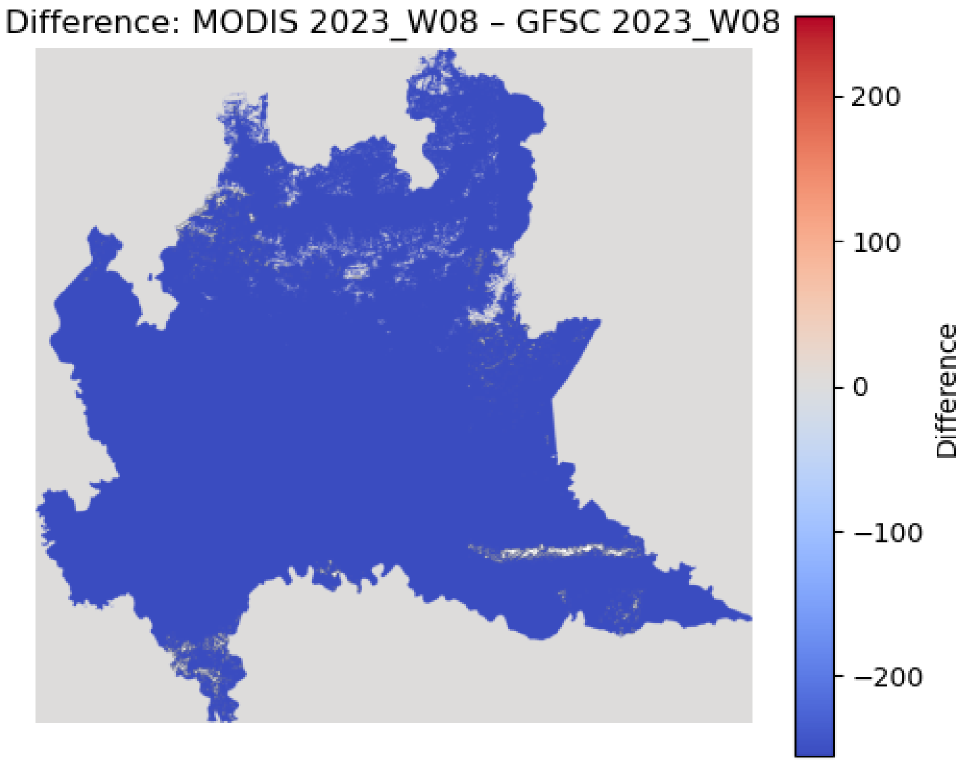

Figure 6: MODIS vs. GFSC W08-2023

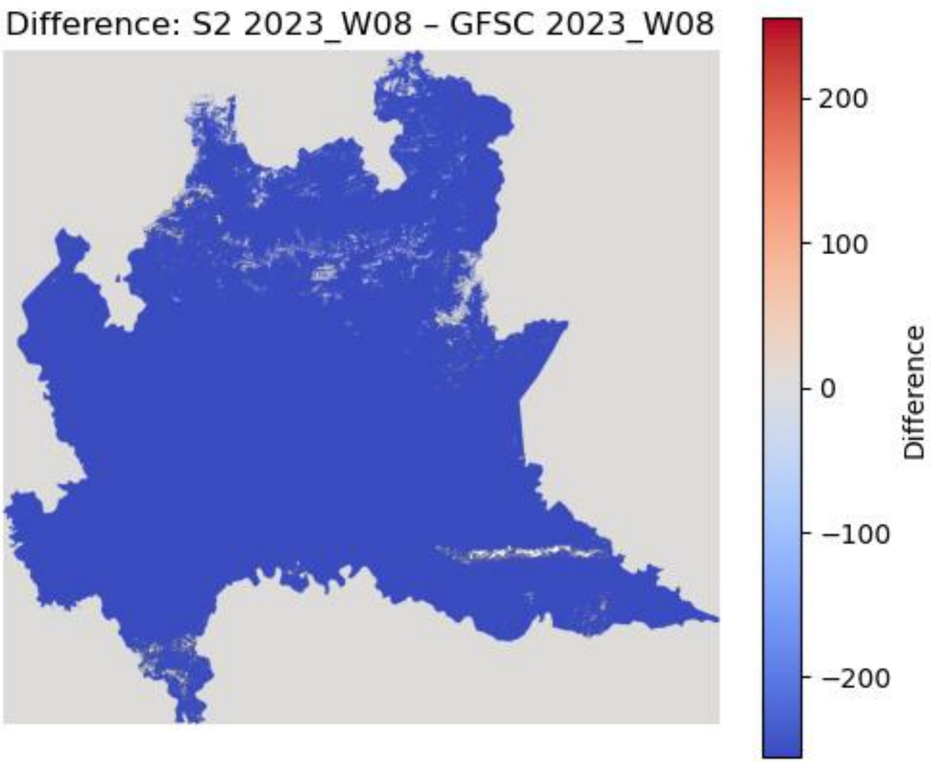

Figure 7: S2 vs. GFSC W08- 2023

5.3.6. Sentinel-2 vs Sentinel-3 Agreement

To further investigate agreement levels between the two high-resolution sensors, we computed the weekly agreement percentage between Sentinel-2 (S2) and Sentinel-3 (S3) using the pairwise comparison function. The output, summarized in the figure below, shows a consistently high level of spatial agreement across most weeks, with values typically above 90%. Weeks such as 2022_W10, 2023_W08, and 2023_W14 exhibit agreement percentages above 95%, suggesting strong consistency in snow detection between the two products.

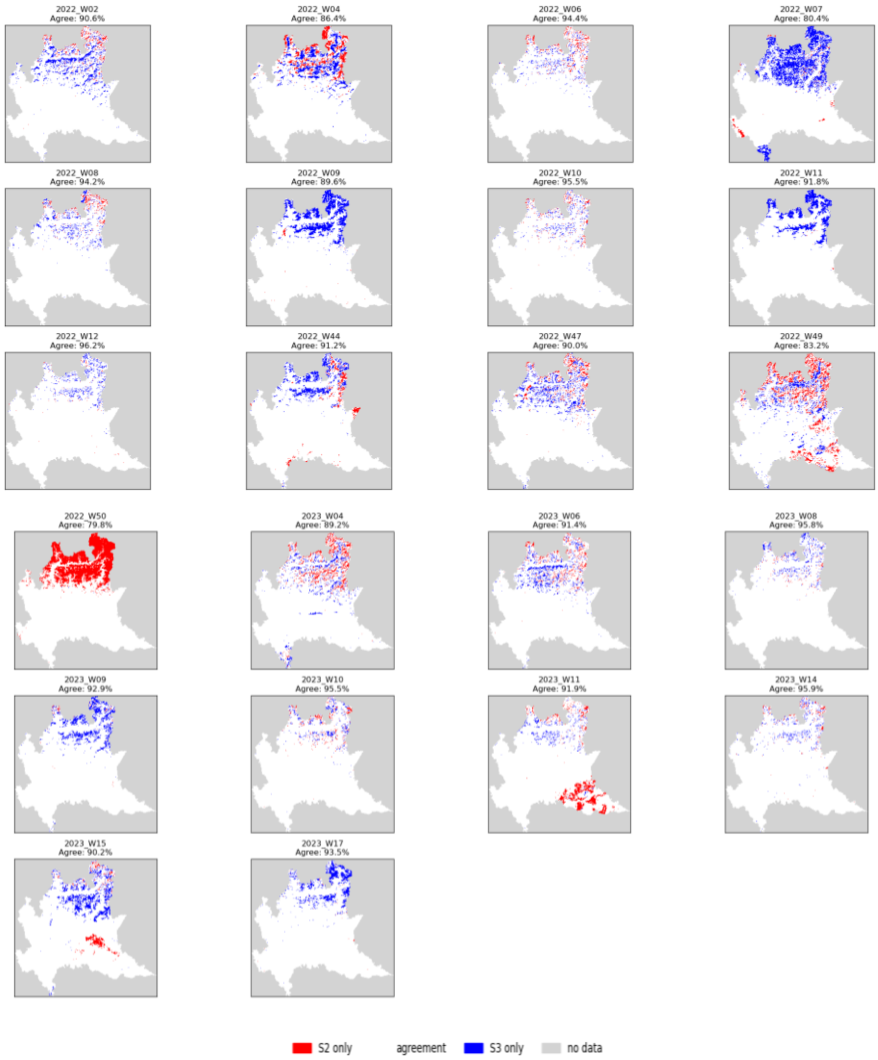

Figure 8: S2 vs S3

This high degree of alignment is expected, as both S2 and S3 provide finer spatial resolution compared to MODIS and adopt stricter cloud and snow masking strategies. However, notable drops in agreement (e.g., weeks 2022_W07 and 2022_W50) indicate occasional divergence, potentially due to residual cloud cover, acquisition gaps, or differences in observation timing.

5.4. Algorithm Implementation

This section presents the practical implementation of three distinct snow classification approaches, each applied to generate weekly snow cover masks over the Lombardy region. The goal was to assess their reliability and alignment with the MODIS reference product. The implemented methods include:

Random Forest classification, applied in two configurations — one trained directly on MODIS masks, and another trained using aligned products (GFSC, S2, S3);

A **logic-based classifier** inspired by the structure of the *Let-It-Snow* (LIS) algorithm but adapted to the available data and simplified for operational use;

A **regression-based evaluation**, focusing on the quantitative agreement between the snow cover estimates and the MODIS reference across multiple weeks.

Each method was carefully designed to process large-scale raster data efficiently, leveraging cloud-free, pre-aligned datasets. Performance metrics including accuracy, precision, recall, F1-score, and correlation were computed to validate the outputs. The following subsections describe the implementation steps, configuration details, and results for each of the three approaches.

5.4.1. Random Forest

The Random Forest (RF) model was selected as the first classification approach due to its robustness, non-parametric nature, and ability to handle heterogeneous feature spaces. In this study, RF was used to predict MODIS-derived snow cover by exploiting both individual predictors and a fusion of products. Two configurations were tested: (i) using GFSC, Sentinel-2, and Sentinel-3 together as a combined feature set, and (ii) using each product separately to evaluate their individual predictive performance. This dual setup allowed us to assess the added value of multi-source data fusion compared to single-product classification baselines.

Multi-Product Fusion Classification

In the first implementation, a Random Forest classifier was trained using combined inputs from GFSC, Sentinel-2 (S2), and Sentinel-3 (S3) products to predict the snow cover label provided by MODIS for week 2023_W06. Before training, all raster datasets were clipped to a common extent and aligned to a uniform 60 m grid. Pixels with missing values (coded as 255) were excluded to ensure consistency across inputs.

Each valid pixel was represented by a three-dimensional feature vector — one value from each of the GFSC, S2, and S3 products — while the MODIS pixel served as the binary label (0: No Snow, 1: Snow). The final training set contained over six million valid samples. A balanced Random Forest model with 100 trees was trained and evaluated on this dataset.

The model achieved an overall accuracy of **93.7%**, with an F1-score of **0.96** for the *No Snow* class and **0.83** for the *Snow* class. The confusion matrix revealed strong agreement for *No Snow* predictions but a slightly lower recall for *Snow*, likely due to underrepresentation or disagreement in complex terrain or cloud-affected areas. These results confirm the benefit of integrating multiple products to improve predictive performance, especially in regions where single-product outputs may be uncertain or noisy.

Table 4: Performance Metrics of the Random Forest Classifier Using Combined Predictors

Class	**Precision**	**Recall**	**F1-score**	**Support**
No Snow (0)	0.95	0.97	0.96	4,897,358
Snow (1)	0.87	0.80	0.83	1,161,704
Accuracy			**0.9373**	6,059,062
Macro avg	0.91	0.88	0.90	6,059,062
Weighted avg	0.94	0.94	0.94	6,059,062

Random Forest: Single-Product Classification

To further explore the predictive capabilities of each individual product, we trained three separate Random Forest classifiers using a single predictor at a time (i.e., GFSC, S2, or S3), with MODIS snow labels as ground truth. This approach isolates the contribution of each source, highlighting its standalone performance in replicating MODIS snow cover detection.

All classifiers were trained on the same set of valid pixels for week 2023_W06, excluding any locations where either the label or the predictor was missing. Each model was evaluated based on overall accuracy, precision, recall, and F1-score, with a detailed breakdown by class (snow vs. no snow).

The results, summarized in the tables below, indicate that:

GFSC achieved the highest accuracy (0.9036), largely due to strong performance in detecting *no snow*, but it underperformed in identifying snow pixels (recall = 0.52).

S2 offered more balanced metrics, with higher recall for the snow class (0.58), though slightly lower accuracy (0.8947).

S3 achieved the best *precision* for the snow class (0.88), but also suffered from reduced recall (0.62), leading to a final accuracy of 0.8926.

Table 5: Performance Metrics of RF for Single-Product Classification

Product	Class	Precision	Recall	F1-Score	Support	Accuracy
GFSC	No Snow	0.9	0.99	0.94	4897358	0.9036
	Snow	0.96	0.52	0.67	1161704	0.9036
S2	No Snow	0.89	0.99	0.94	5079743	0.8947
	Snow	0.94	0.58	0.72	1548458	0.8947
S3	No Snow	0.89	0.97	0.93	5079743	0.8926
	Snow	0.88	0.62	0.73	1548458	0.8926

5.4.2. The "Let-It-Snow" (LIS) Approach:

Developed by the Theia Snow Team, LIS is an automated framework for snow cover detection from optical multispectral images (e.g., Sentinel-2, Landsat-8). Designed to operate at 10–30 m resolution, the algorithm targets mountainous terrain and large-scale operational monitoring gitlab.orfeo-toolbox.org+9github.com+9researchgate.net+9.

Key components of the LIS methodology include:

Input Requirements:

Level-2A optical products with visible and short-wave-infrared (SWIR) bands.

A Digital Terrain Model (e.g., SRTM) for slope correction and snow-line analysis gitlab.orfeo-toolbox.org+2zenodo.org+2gitlab.orfeo-toolbox.org+2data.code.gouv.fr+6github.com+6researchgate.net+6.

NDSI-Based Thresholding:

Snow detection leverages the Normalized Difference Snow Index (NDSI), calculated from green and SWIR bands.

A two-pass approach applies progressively relaxed thresholds depending on scene characteristics.

Configurable band reflectance limits (e.g., SWIR thresholds) are used to remove cloud, shadow, or mixed-pixel effects researchgate.net.

Cloud Exclusion and Slope Masking:

Multi-stage cloud masking removes bright clouds and cloud shadows.

A slope-derived Snow-Line Elevation (zs) dynamically adjusts detection sensitivity across elevation bands av.tib.eu+15essd.copernicus.org+15github.com+15gitlab.orfeo-toolbox.org+5researchgate.net+5gitlab.orfeo-toolbox.org+5.

Output Masks:

Binary snow/no-snow maps, with additional layers for clouds and slopes where detection is uncertain.

Fully integrated into Theia Snow collection, with comprehensive quality controls github.com+2researchgate.net+2essd.copernicus.org+2.

Operational Excellence:

Efficient implementation via C++/Python hybrid streamlines processing of large tile sets.

LIS has been validated across Western Europe. Evaluation shows **94% overall accuracy** and **0.91 F1-score**, outperforming ESA Sen2Cor products, with caveats in cloud-prone areas researchgate.net.

In summary, LIS is a robust, terrain-aware NDSI-based snow detection algorithm that employs cloud masking, slope adjustment, and a multi-pass logic to enhance detection performance at high spatial resolution. This serves as a valuable reference for adapting its principles into our simpler implementation for the Lombardy region.

Logic-Based Classification Using Let-It-Snow-Simplified Version

To complement our machine learning experiments, we explored a logic-based approach inspired by the *Let-It-Snow (LIS)* . In our project, we adopted a **simplified version of LIS** to apply consistent logic-based classification to all three input products: GFSC, Sentinel-2, and Sentinel-3. Rather than using raw reflectance bands and multi-criteria decision trees, our logic was applied directly on the binary snow masks produced by each product. By this we assessed the effectiveness of threshold-based classification logic across sensors in a controlled environment.

Implementation Details

The simplified logic is encapsulated in a custom function (logic_based_mask()), which assigns:

1 for pixels flagged as snow,

0 for pixels flagged as no-snow,

NaN for nodata or masked values (i.e., 255 in original data).

Each product's classification output was compared to MODIS as a reference. The evaluation was conducted on the week 2023_W06, using standard classification metrics. Below is a summary of the accuracy and F1-scores obtained:

Table 6: Performance Metrics of LIS

Product	Class	Precision	Recall	F1-Score	Support	Accuracy
GFSC	No Snow	0.9	0.99	0.94	4897358	0.9036
	Snow	0.96	0.52	0.67	1161704	0.9036
S2	No Snow	0.89	0.99	0.94	5079743	0.8947
	Snow	0.94	0.58	0.72	1548458	0.8947
S3	No Snow	0.89	0.97	0.93	5079743	0.8926
	Snow	0.88	0.62	0.73	1548458	0.8926

Interestingly, the logic-based classification approach—though significantly simpler and based only on direct thresholding logic from individual product values—yields results that are remarkably close to those obtained with the more complex Random Forest (RF) classifier. For instance, GFSC achieved an accuracy of **0.9036** using both the logic-based method and the RF model trained on GFSC alone. Similarly, the S2 and S3 products also showed near-identical accuracy scores in both approaches (around **0.89**), with comparable F1-scores for both the *No Snow* and *Snow* classes. This similarity suggests that the snow/no-snow decision boundary is already well-represented in the individual product masks, and that the simplified logic rule effectively captures the essential classification patterns without requiring machine learning. Nonetheless, RF still offers marginal improvements in balancing recall and precision, especially for the snow class.

5.4.3. Regression-Based Comparison with MODIS

In the final stage of our project, we carried out a regression analysis to assess the continuous agreement between each snow cover product and the MODIS reference. The analysis was performed using the run_regression_batch() function, which aligned and compared weekly snow masks from GFSC, Sentinel-2 (S2), and Sentinel-3 (S3) against MODIS over a set of 23 common weeks (from 2022_W02 to 2023_W17). For each product and week, the method sampled up to 500,000 valid pixels and computed a linear regression between the product and MODIS values, extracting the **slope** and **correlation coefficient** as performance metrics.

These two metrics provide complementary insights:

The **slope** indicates how well the product matches the magnitude of MODIS snow cover: a slope of 1 represents perfect agreement.

The **Pearson correlation coefficient** reflects the spatial coherence between the two products: a value close to 1 indicates strong agreement in spatial patterns.

Table 7: Regression Output

Product	**Slope**	**Correlation**
GFSC	0.871	0.847
S2	0.640	0.304
S3	0.626	0.377

As evident from the results, **GFSC shows the highest agreement with MODIS** both in terms of slope and correlation, suggesting that it not only estimates the snow extent with comparable magnitude but also maintains consistent spatial structure. In contrast, both Sentinel-2 and Sentinel-3 exhibit significantly lower correlation values (0.304 and 0.377, respectively), indicating weaker spatial alignment with MODIS observations. The low slope values (0.640 for S2 and 0.626 for S3) also reflect underestimation of snow-covered areas compared to MODIS, possibly due to the effect of cloud contamination, resolution differences, or classification logic.

6 Conclusion

This project presented a full workflow for snow cover detection and evaluation across the Lombardy region, combining both comparative product analysis and algorithmic implementation. The first objective focused on assessing the spatial and temporal agreement between four key snow products: MODIS, GFSC, Sentinel-2, and Sentinel-3. The second objective centered on implementing custom classification algorithms — Random Forest, logic-based (LIS-inspired), and regression-based — and validating them using MODIS as a reference.

Our analysis revealed noticeable discrepancies among the products, particularly in terms of snow extent and consistency across weeks. MODIS and Sentinel-3 showed broader coverage, while GFSC and Sentinel-2 yielded more conservative snow detection. These patterns were supported by agreement metrics and statistical comparisons, forming a reliable basis for algorithm training and validation.

The implementation phase demonstrated that both machine learning and logic-based approaches yielded satisfactory results. The Random Forest classifier trained on MODIS achieved high accuracy (0.94) and solid snow class performance (F1 ≈ 0.83), while the logic-based classifier produced comparable metrics using a simplified version of the Let-It-Snow algorithm. Regression analysis across 23 weeks confirmed that GFSC showed the strongest agreement with MODIS (slope = 0.871, correlation = 0.847), outperforming S2 and S3.

It is important to note that the results were affected by constraints on data access, which were set by the requirements provided by CNR. All data used in the project — including MODIS, Copernicus GFSC, Sentinel-2, and Sentinel-3 — were retrieved through ARPA systems and fully aligned with CNR's expectations. While these limitations influenced the scope of temporal and spatial analysis, we ensured that our processing pipeline and methodological design strictly followed the technical and scientific guidelines shared with us.

In conclusion, the implemented workflow successfully meets the dual goals of product comparison and algorithm development. The modular structure and validated outputs establish a solid foundation for future operational snow monitoring, and the project remains fully compliant with the goals and standards of the involved institutions.

7 Reference

1. **CNR (National Research Council of Italy)**. Technical requirements and guidance provided for snow cover monitoring in Lombardy.
   Website: https://www.cnr.it
2. **ARPA Lombardia (Agenzia Regionale per la Protezione Ambientale)**. Provider of MODIS and Sentinel-aligned datasets for the Lombardy region.
   Website: https://www.arpalombardia.it
3. **Let-It-Snow Algorithm (Orfeo Toolbox)**. Snow detection algorithm for optical imagery.
   GitLab repository: https://gitlab.orfeo-toolbox.org/remote_modules/let-it-snow
4. **ESA Copernicus Programme**. Provider of Sentinel-2 and Sentinel-3 data.
   Website: https://www.copernicus.eu
5. **Copernicus Global Land Service (GFSC)**. GFSC snow product used for comparison.
   Website: https://land.copernicus.eu/global
6. **NASA MODIS (Moderate Resolution Imaging Spectroradiometer)**. Used as a reference dataset for validation.
   Website: https://modis.gsfc.nasa.gov
7. **Pedregosa, F., Varoquaux, G., Gramfort, A., et al. (2011)**. *Scikit-learn: Machine Learning in Python*. *Journal of Machine Learning Research*, 12, 2825–2830.
   https://jmlr.org/papers/v12/pedregosa11a.html
8. **Combal, B., Foga, S., Knaff, J., et al. (2020)**. *Validation of Let-It-Snow snow mask algorithms for Sentinel-2 imagery*. *Remote Sensing of Environment*, 240, 111704.
   https://doi.org/10.1016/j.rse.2020.111704
9. **Gascoin, S., Grizonnet, M., Bouchet, M., et al. (2019)**. *Operational Sentinel-2 snow cover extent mapping using an open-source and cloud-based processing chain*. *Remote Sensing*, 11(5), 516.
   https://doi.org/10.3390/rs11050516
10. **Rasterio**. Python library for geospatial raster I/O operations.
    Website: https://rasterio.readthedocs.io
11. **Seaborn & Matplotlib**. Visualization libraries used for plotting and performance analysis.
    Websites:
    https://seaborn.pydata.org
    https://matplotlib.org

Project Link in GitHub:

Sonw_Cover_Map_GitHub