Article

Five-Year Evaluation of Sentinel-2 Cloud-Free Mosaic Generation Under Varied Cloud Cover Conditions in Hawai'i

Francisco Rodríguez-Puerta [1,*] , Ryan L. Perroy [2,3] , Carlos Barrera [4], Jonathan P. Price [2,3] and Borja García-Pascual [4,5]

1   EiFAB-iuFOR, Campus Universitario Duques de Soria s/n, Universidad de Valladolid, 42004 Soria, Spain
2   Department of Geography & Environmental Science, University of Hawai'i at Hilo, Hilo, HI 96720, USA; rperroy@hawaii.edu (R.L.P.); jpprice@hawaii.edu (J.P.P.)
3   Spatial Data Analysis & Visualization Laboratory, University of Hawai'i at Hilo, Hilo, HI 96720, USA
4   Föra Forest Technologies SLL, Campus Universitario Duques de Soria s/n, 42004 Soria, Spain; carlos.barrera@fora.es (C.B.); borja.garciapascual@luke.fi (B.G.-P.)
5   Natural Resources Institute Finland (Luke), Yliopistokatu 6, 80100 Joensuu, Finland
*   Correspondence: francisco.rodriguez.puerta@uva.es; Tel.: +34-645830476

Abstract: The generation of cloud-free satellite mosaics is essential for a range of remote sensing applications, including land use mapping, ecosystem monitoring, and resource management. This study focuses on remote sensing across the climatic diversity of Hawai'i Island, which encompasses ten Köppen climate zones from tropical to Arctic: periglacial. This diversity presents unique challenges for cloud-free image generation. We conducted a comparative analysis of three cloud-masking methods: two Google Earth Engine algorithms (CloudScore+ and s2cloudless) and a new proprietary deep learning-based algorithm (L3) applied to Sentinel-2 imagery. These methods were evaluated against the best monthly composite selected from high-frequency Planet imagery, which acquires daily images. All Sentinel-2 bands were enhanced to a 10 m resolution, and an advanced weather mask was applied to generate monthly mosaics from 2019 to 2023. We stratified the analysis by cloud cover frequency (low, moderate, high, and very high), applying one-way and two-way ANOVAs to assess cloud-free pixel success rates. Results indicate that CloudScore+ achieved the highest success rate at 89.4% cloud-free pixels, followed by L3 and s2cloudless at 79.3% and 80.8%, respectively. Cloud removal effectiveness decreased as cloud cover increased, with clear pixel success rates ranging from 94.6% under low cloud cover to 79.3% under very high cloud cover. Additionally, seasonality effects showed higher cloud removal rates in the wet season (88.6%), while no significant year-to-year differences were observed from 2019 to 2023. This study advances current methodologies for generating reliable cloud-free mosaics in tropical and subtropical regions, with potential applications for remote sensing in other cloud-dense environments.

Keywords: cloud-free mosaics; Sentinel-2; planet imagery; deep learning; Google Earth Engine; cloud-masking algorithms; Hawai'i

1. Introduction

Cloud cover is a significant obstacle to acquiring useful satellite imagery, which is critical for monitoring ecosystems and managing natural resources. Global annual mean cloud cover is estimated at around 66%, according to the International Satellite Cloud Climatology Project [1]. This extensive cloud presence poses substantial challenges for optical satellite observations, as clouds obscure the Earth's surface, limiting the acquisition of clear images required for precise environmental assessments [2]. Satellite remote sensing, using platforms like Sentinel-2, has become an essential tool for environmental assessments, offering global coverage, cost-effectiveness, and standardized data across extensive spatial and temporal scales [3]. Sensors such as Sentinel-2 and Planet are particularly valuable for ecosystem monitoring; however, atmospheric conditions, especially cloud cover, frequently

obstruct data acquisition [4]. The challenge is particularly pronounced on Hawai'i Island, a geographically diverse landmass with significant cloud cover variability, where clouds complicate the capture of high-quality satellite imagery essential for studying land use change, ecosystem health, and environmental responses to disturbances [5,6]. Effective monitoring under such varied climatic conditions requires advanced cloud-masking techniques capable of reliably addressing cloud interference, a factor that critically impacts the usability of satellite data [4,7].

Accurate cloud identification is critical, as errors in cloud detection can affect downstream analyses, thereby impacting the quality of Earth observation products [8,9]. Achieving a balance between accuracy and a conservative approach to cloud masking is essential, as overly conservative algorithms may restrict valid observations and reduce the availability of cloud-free data [7]. Hawai'i Island's diverse climate zones and regions of persistent cloud cover make it an ideal testing ground for cloud-masking techniques, with applications for addressing ecological challenges such as forest pathogens like rapid 'Ōhi'a death [10–13], endangered and invasive species mapping and monitoring [14,15], and identifying disturbance events such as storms and fires [16–18].

The need for reliable cloud-free imagery in environmental monitoring, particularly in cloud-dense tropical regions, has driven advancements in cloud-masking algorithms. These tools range from rule-based models and supervised learning methods to more sophisticated deep learning models designed to address complex cloud patterns and atmospheric conditions [19]. Traditional algorithms, such as FMask and Sen2Cor, have been widely adopted but exhibit notable limitations. FMask, while effective at detecting thick clouds and shadows, often struggles with thin cloud edges, leading to inaccuracies in complex atmospheric conditions [20]. Similarly, Sen2Cor frequently misclassifies bright surfaces, such as snow and ice, as clouds due to its reliance on fixed spectral thresholds [21]. Advancements such as CloudScore+ and the WASP (Weighted Average Synthesis Processor) address some of these shortcomings by offering greater flexibility and multi-temporal synthesis capabilities, respectively, to improve the quality of cloud-free imagery [22,23]. Deep learning techniques, particularly convolutional neural networks, are increasingly preferred for their ability to capture complex spatial patterns directly from the data, often achieving higher precision in cloud detection by minimizing false positives [24,25]. However, challenges such as computational demands and scalability issues persist [26]. Emerging transformer-based models have introduced innovative approaches to cloud detection by leveraging self-attention mechanisms to capture global spatial dependencies and long-range relationships within satellite data. These models, such as U-TILISE and hybrid architectures integrating CNNs and transformers, have demonstrated enhanced accuracy in complex scenarios, particularly in handling thin clouds and dynamic cloud patterns [27–30]. However, transformers are computationally intensive, requiring larger and more diverse datasets for effective training and often facing scalability challenges due to their high resource consumption [31,32]. Furthermore, overfitting remains a concern in cases where training data are insufficiently representative of diverse conditions, potentially limiting generalizability [33]. Despite these drawbacks, their ability to integrate spatial and temporal dependencies offers significant potential for advancing cloud-free mosaic generation, particularly in environments with persistent cloud cover.

Google Earth Engine (GEE) has revolutionized large-scale geospatial data analysis by offering a powerful cloud-based platform that combines extensive repositories of remote sensing data with seamless processing capabilities, eliminating the need for time-consuming data downloads [34,35]. While numerous algorithms have been developed to mask clouds in optical satellite imagery, comparative studies are still limited, especially in regions with high and variable cloud cover, such as Hawai'i [23,36].

This study aims to evaluate the performance of three advanced cloud-masking algorithms applied to Sentinel-2 imagery, CloudScore+ and s2cloudless, commonly used in Google Earth Engine [35], along with a new deep learning model (L3), for generating cloud-free mosaics over Hawai'i Island. By comparing these methods against high-frequency

Planet imagery as a reference dataset, this analysis provides insights into each algorithm's strengths and limitations under a range of cloud cover conditions, from low to very high. We hypothesize that the deep learning-based L3 algorithm will exhibit higher precision in handling complex cloud patterns, while CloudScore+ will maintain stable accuracy across varying levels of cloud cover. This research contributes essential insights for advancing cloud-free mosaic generation in cloud-heavy environments and fills critical gaps in the comparative evaluation of cloud-masking techniques [34,37]. Through a longitudinal, statistically rigorous approach, we assess the performance and consistency of each algorithm under diverse cloud cover frequencies and seasonal conditions, aiming to support future remote sensing applications in tropical and subtropical regions.

2. Materials and Methods

2.1. Study Area

Hawai'i Island, located between 19°N and 20°N, is the largest island in the Hawaiian archipelago, spanning 10,430 km². Its formation by five shield volcanoes contributes to a rich tapestry of Köppen climate groups, encompassing tropical, arid, and tundra climates [38]. Of the 14 recognized climate zones, Hawai'i Island hosts 10, including the cold-summer Mediterranean climate, demonstrating exceptional climatic variability (Figure 1a). Sea-level temperatures oscillate from 26 °C in summer to 22 °C in winter, rarely exceeding 32 °C or dropping below 16 °C [39]. Snowfall may occur at the Mauna Kea and Mauna Loa summits throughout the year. Trade winds create a persistent moisture discontinuity between 1200 and 2400 m, prevailing 50 to 70 percent of the time, inhibiting vertical air movement and constraining cloud formation just below this threshold. The leeward areas to the west typically experience clear skies, while the windward side receives substantial rainfall from trade winds and orographic precipitation. Coastal regions, especially the south and west, exhibit drier conditions. Hawai'i's annual precipitation is highly variable, ranging from just 188 mm at the summit of Mauna Kea to over 7500 mm on the windward slopes, despite being only 30 km apart. Unlike most tropical regions, Hawai'i's wet season occurs primarily in winter, adding a unique seasonal dynamic to its climatic diversity. Cloud cover also varies significantly (Figure 1b), with persistent clouds often enveloping the windward side and the southern Kona coast, while much of the central region and the northern leeward side enjoy clearer skies [39].
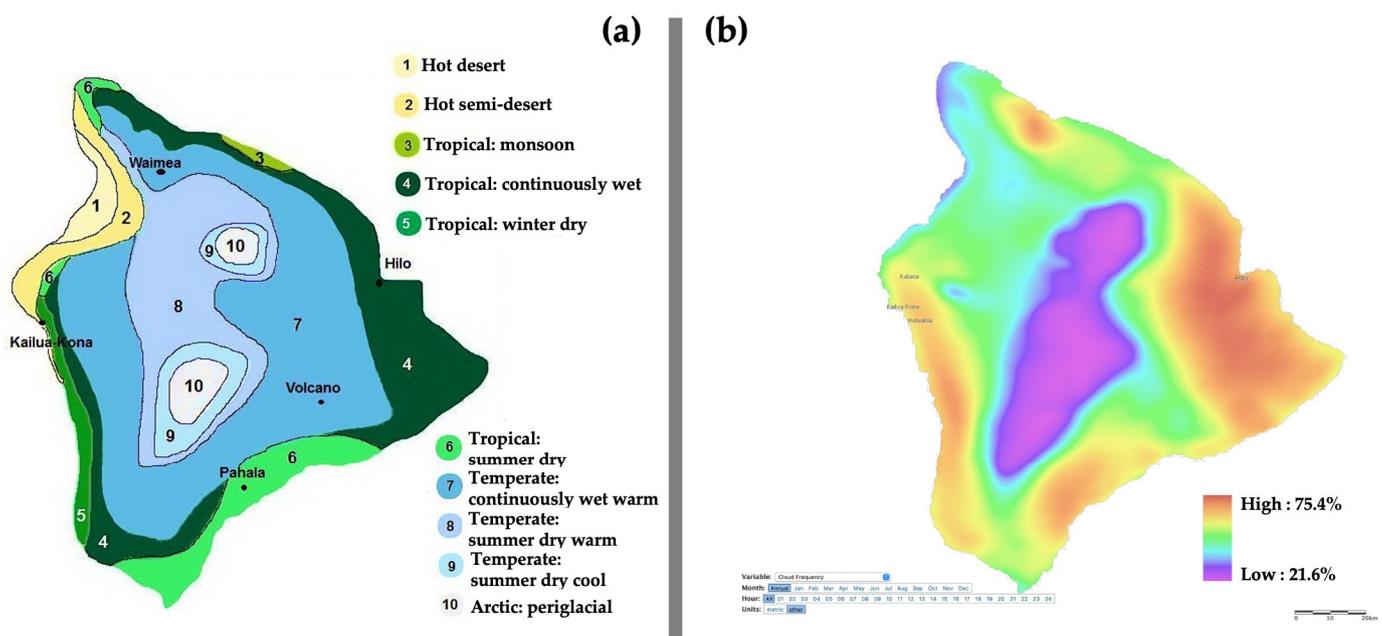
Figure 1. Island of Hawai'i: (**a**) Köppen climate map adapted from Mauna Loa Observatory Report [40] and (**b**) the annual mean cloud cover percentage based on the Hawai'i Climate Atlas [39].

2.2. Availability of Satellite Imagery

In this study, we evaluated three cloud-masking methods to produce cloud-free mosaics from Sentinel-2 imagery over Hawai'i Island, using high-frequency Planet imagery as a reference dataset. Sentinel-2 provides a 10 m spatial resolution in four of its 13 spectral bands and has a revisit cycle of five days, which makes it a valuable tool for monitoring land cover, climate change, and natural disasters [41]. Furthermore, the integration of Sentinel-2 with other global satellite programs like Landsat and SPOT enhances timely data availability, aiding in change detection and progress tracking for Sustainable Development Goals [42]. Planet satellite imagery has also become highly relevant for scientific and commercial applications, especially with the availability of free tropical Basemaps through Norway's International Climate and Forests Initiative (NICFI). Using tools like Google Earth Engine [35], users can access these Basemaps at a resampled resolution of 4.77 m.

2.2.1. Planet Basemap Imagery

Operated by Planet Labs, a constellation of low-Earth orbit satellites provides daily global imaging [43]. These satellites capture images in four spectral bands: blue (455–515 nm), green (500–590 nm), red (590–670 nm), and near-infrared (780–860 nm) [44], offering a high revisit frequency. Access to these data was facilitated through an academic license, allowing researchers to download individual images and monthly Basemaps. The Basemaps were created from the best images of each month, providing a reliable view of the Earth's surface. As noted by Pascual et al. [45], these maps are carefully normalized to reduce atmospheric and sensor-related distortions and are available at a resolution of 4.77 m.

2.2.2. CloudScore+ Algorithm on Google Earth Engine

The CloudScore+ algorithm [37] is an advanced cloud-masking tool that uses a machine learning-based optical image quality assessment (QA) with weak supervision to improve cloud detection and mask generation for Sentinel-2 imagery. Building on the foundational concepts of Landsat TDOM and the original CloudScore algorithm, CloudScore+ enhances detection accuracy and reduces false positives, particularly in challenging conditions such as snowy or highly reflective surfaces [46]. Initially designed for Landsat, CloudScore+ operates at the pixel level, masking clouds and shadows through the combined use of visible, NIR, SWIR, and thermal bands [47]. Integrated as a function in GEE ('*ee.Algorithms.Landsat.simpleCloudScore*'), it was later adapted for Sentinel-2 MSI data, incorporating cirrus and aerosol bands to replace Landsat's thermal infrared band [48]. Leveraging machine learning on labeled datasets, CloudScore+ benefits from video sequence analysis to capture temporal dynamics, which enhances predictions of atmospheric effects with greater precision.

CloudScore+ generates two main products: the *cs_band*, an atmospheric similarity score sensitive to haze and cloud edges, and the *cs_cdf band*, a cumulative distribution score less affected by minor spectral shifts or terrain shadows [49,50]. The *clear_threshold* parameter, which sets the detection threshold for atmospheric phenomena, typically ranges from 0.5 to 0.65 but can be adapted to meet specific conditions. Given Hawai'i's significant cloud cover variability, we optimized this parameter, testing thresholds of 0.5, 0.65, 0.75, and 0.85 to strike an ideal balance between excluding clouds and retaining data. This customization aimed to address frequent and variable cloud cover by adjusting the threshold as cloud presence increased. Finally, we modified the foundational CloudScore+ script available on GEE (https://bit.ly/3UHAORF, accessed on 12 February 2024) to fit our study's parameter requirements, expanding its applicability across diverse research contexts.

2.2.3. s2cloudless Algorithm on Google Earth Engine

The s2cloudless algorithm, based on the Light Gradient Boost Machine (Light GBM) framework [51], is an automated tool for detecting clouds in Sentinel-2 imagery. Developed by Sinergise's EO Research team, it is available under the MIT License on GitHub. This model, trained on extensive global datasets, excels in identifying clouds for specific time

frames, though it does not incorporate spatial context, which enhances its adaptability across various resolutions. Using Sentinel-2 Level-1C TOA reflectance bands, it generates a cloud probability map that users can convert into a cloud mask by applying a recommended threshold (typically 0.4) to mitigate cloud omission errors. The algorithm provides several adjustable parameters, including *cloud_filter*, which sets the maximum allowable cloud cover in images; *cld_prb_thresh*, which determines the cloud probability threshold; *nir_drk_thresh*, which specifies the near-infrared reflectance threshold; and *cld_prj_dist*, which sets the maximum search distance (in kilometers) for cloud shadows from cloud edges. In our study, we tested three values for *cld_prb_thresh* (0.35, 0.4, and 0.45) and three values for *nir_drk_thresh* (0.10, 0.15, and 0.20), selecting optimal values for comparative analysis. Both Sentinel Hub and Google Earth Engine provided pre-processed s2cloudless cloud probability maps for the full Sentinel-2 archive. For our implementation in GEE, we utilized Sentinel-2 Surface Reflectance (SR) and Sentinel Cloud Probability (SCP) data. While the foundational script is available on GEE's website (https://bit.ly/3y03JaB, accessed on 12 February 2024), we customized it to meet our parameterization requirements.

2.2.4. A New Masking Algorithm Based on Deep Learning (L3)

The L3 algorithm, adapted by föra forest technologies, employs deep learning to generate cloud-free mosaics from Sentinel-2 imagery by analyzing temporally adjacent images in a time series. Integrating the WASP methodology [52], L3 applies super-resolution, cloud masking, and weighted averaging to create a seamless, high-resolution mosaic specifically optimized for Sentinel-2's 10 m resolution. This approach effectively manages cloud cover variability through a multi-stage weighting system based on cloud presence and temporal proximity.

L3 was trained using a manually curated dataset derived from Baetens-Hagolle's Sentinel-2 cloud mask dataset [53]. The training workflow consisted of the following: (1) generating Level-2 (L2) products from Sentinel-2 imagery, (2) applying super-resolution to enhance spatial detail, (3) curating high-quality ground truth labels for cloud masks, and (4) training the deep learning model. This dataset, enriched with super-resolved imagery, provided a robust foundation for the model to learn complex cloud patterns and atmospheric conditions, enhancing its ability to generalize across diverse climatic settings.

Building on this foundation, the algorithm begins with the super-resolution of all bands in the Sentinel-2 time series using DSen2, a deep learning model based on the Enhanced Deep Super-Resolution (EDSR) architecture [54,55]. This step improves spatial detail, especially in the 20 and 60 m bands, by analyzing features from the 10 m bands, allowing for the detection of thin clouds and semi-transparent atmospheric phenomena.

Following super-resolution, L3 applied cloud masking with HRNet+OCR, a CNN segmentation model [56] that classifies each pixel into six categories: clouds, haze, cloud shadows, land, water, and snow. The HRNet+OCR mask was combined with Sen2Cor masks [57,58], produced during atmospheric correction, to improve the identification of cloud-affected pixels and reduce misclassifications on bright surfaces. Morphological operations, particularly erosion and dilation, are then employed to refine the edges of the mask. Erosion, using a 9×9 pixel matrix, eliminates small, isolated artifacts and sharpens the boundaries of invalid pixels. On the other hand, dilation, also with a 9×9 pixel matrix, ensures smoother transitions and connectivity between neighboring regions. This two-step process enhances the precision of the binary mask, effectively distinguishing valid (land and water) from invalid pixels (clouds, haze, cloud shadows, and snow).

Once the binary mask was created, cloud weights were assigned to each pixel based on their proximity to cloud-affected areas. This step utilizes Gaussian filters at two scales (10 m and 240 m) to avoid abrupt transitions in the final mosaic. For the 10 m mask, two Gaussian filters with standard deviations of 8 and 20 were applied independently, producing weight matrices G_{10} for cloud proximity. The resulting weight matrix W_{10}^{cloud} is calculated as follows:

$$W_{10}^{cloud} = (1 - G_{10,\,\sigma=8} * M_{10}) * (1 - G_{10,\,\sigma=20} * M_{10}) \tag{1}$$

where M_{10} is the binary cloud mask at 10 m resolution.

For additional smoothing, the binary mask was down-sampled to a 240 m resolution and processed with Gaussian filters (standard deviations of 2 and 10). This 240 m weight matrix was then rescaled to 10 m and combined with the 10 m matrix, yielding the final cloud weight matrix W^{cloud} for the scene. This multi-scale approach, inherited from the WASP methodology, balances computational efficiency and the seamless integration of cloud-free pixels in the final composite (Figure 2).
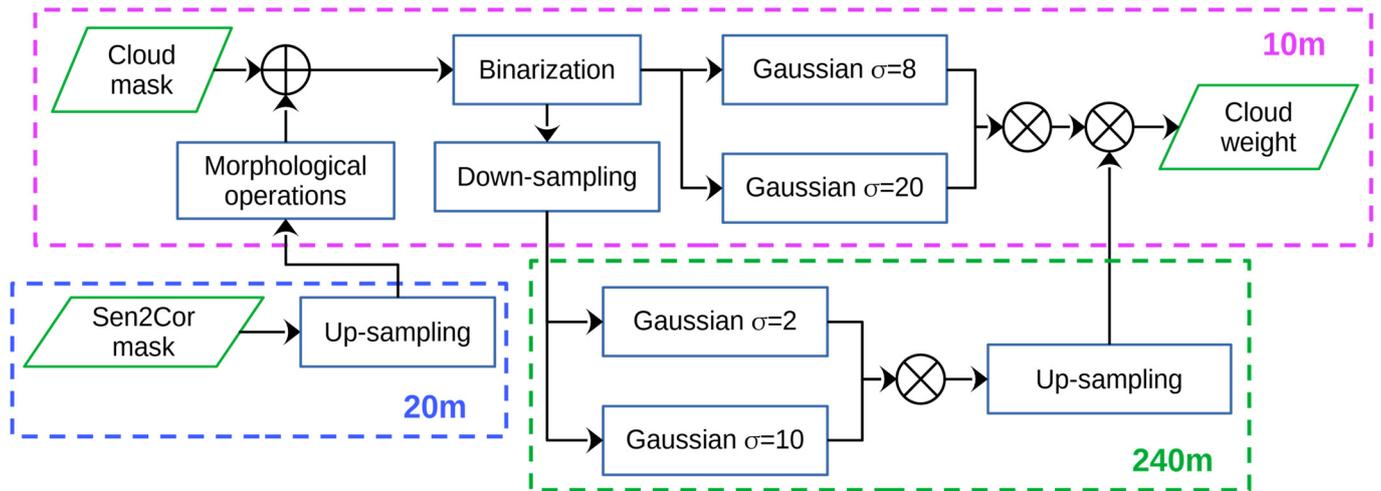
Figure 2. Cloud weight computation for a single Sentinel-2 image, showing different spatial resolutions used during the process.

Each pixel weight is adjusted by a temporal weight W^{date}, calculated as the relative difference in days between the acquisition date of each image and the target date:

$$W^{date} = e^{-\frac{|t-t_0|}{\tau}} * (1 - W_{min}) \tag{2}$$

where t is the acquisition date, t_0 is the target date, τ is the time from which Sentinel 2 images are sampled, and W_{min} is the minimum magnitude of W^{date}.

The combined weight for each pixel at time t is given by the following:

$$W_t^{final} = W_t^{cloud} \cdot W_t^{date} \tag{3}$$

Finally, a weighted average of all super-resolved images was computed to produce the L3 mosaic:

$$L3 = \frac{\sum_{t=1}^{T} W_t^{final} \cdot Image_t}{\sum_{t=1}^{T} W_t^{final}} \tag{4}$$

where L3 is the final cloud-free product, W_t^{final} is the weight associated with the super-resolved Sentinel-2 image ($Image_t$) at time t, and T is the total number of images in the Sentinel 2 time series.

This three-step process allows the L3 algorithm to generate a high-quality, cloud-free composite that effectively handles areas with persistent cloud cover. The combination of super-resolution, multi-scale cloud weighting, and temporal weighting in L3 addresses common challenges in remote sensing, producing a final product that minimizes abrupt boundaries and provides a temporally cohesive representation of the landscape.

2.3. Experimental Design

Numerous studies have employed cloud reference datasets from Sentinel-2 and Landsat 8, such as L8Biome [59], CESBIO [23], and GSFC [60]. However, these datasets have limitations, particularly the lack of temporal information and incomplete labeling of thin clouds or cloud shadows. Despite efforts to address these limitations with datasets like CloudSEN12 [61], we opted to develop our own dataset tailored specifically for this study. This new dataset uses only Sentinel-2 and Planet imagery, taking advantage of Hawai'i's diverse cloud cover and the precise cloud mapping capabilities provided by the Hawai'i Climate Atlas [39].

To define areas with distinct cloud cover (CC) frequencies, we used raster data from the Hawai'i Climate Atlas and stratified the island into four cloud cover categories based on the annual mean cloud cover recorded at 11 am HST, which corresponds to the Sentinel-2 acquisition time. The four strata are as follows: low (<30% cloud cover), moderate (30–50%), high (50–70%), and very high (>70%). These strata cover 7.9%, 29.9%, 32.4%, and 29.7% of the island area, respectively (Figure 3).
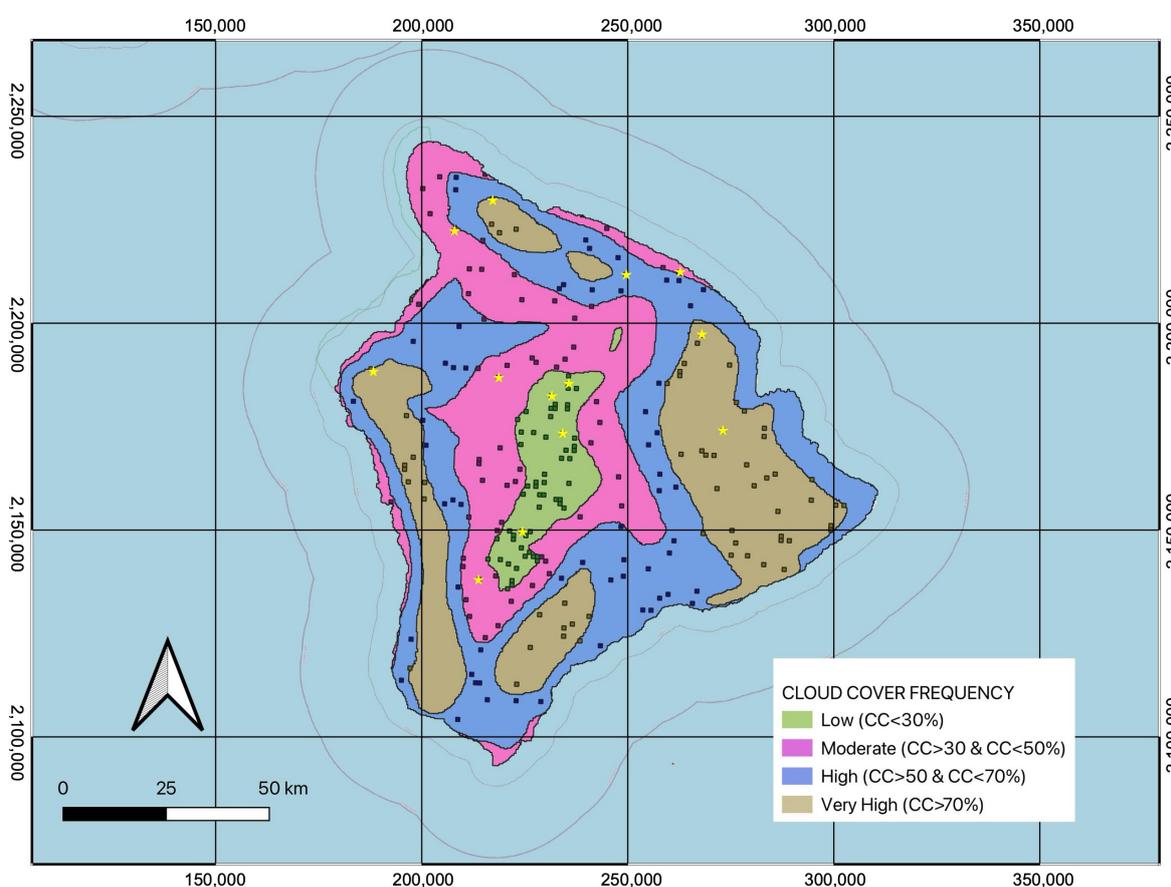
Figure 3. Cloud cover stratification across the island of Hawai'i, derived from data provided by the Hawaiian Climate Atlas [39]. The black squares represent the locations of the evaluated blocks (a total of 240 blocks, each covering 100 hectares) within each cloud cover stratum (60 blocks allocated to each cloud cover stratum). Yellow stars indicate the spatial distribution of the sample blocks presented in Results.

Using a stratified random sampling approach within each cloud cover stratum, we selected 60 blocks, each representing a unique combination of month and year, covering the period from January 2019 to December 2023. Given the four cloud cover strata, this resulted in a total of 240 blocks. In each block, we applied three Sentinel-2 cloud masks and used Planet imagery as a reference, generating a total of 960 analyzed images (Figure 3). This

systematic approach provided comprehensive coverage of the island while minimizing redundant sampling across distinct cloud cover zones.

Each selected block covered an area of 100 hectares (1 km^2), further divided into 100 pixels of 1 hectare each. This scale allowed for detailed visual inspection and precise assessment within each block. For each block, we recorded the number of pixels classified as "clear" (indicating successful cloud masking), "cloudy" (mask failure), and "no data" (mask inability). Additionally, we incorporated two seasonal variables: the dry season (May 1 to September 30) and the wet season (October 1 to April 30), as well as the study year (2019 to 2023). These factors, combined with the cloud cover strata and different mask performances, provided a nuanced understanding of cloud cover dynamics across various conditions.

2.4. Benchmarking

2.4.1. Visual Inspection

To assess the effectiveness of the cloud-masking algorithms in generating cloud-free mosaics (comprising three Sentinel-2 masks and the best monthly image from Planet), we conducted a detailed visual inspection across multiple conditions, including cloud cover categories (low, moderate, high, and very high), seasonal variations (dry and wet), and years (2019–2023). This analysis encompassed all 960 images, covering 240 blocks of 100 hectares each and applying four different masks.

The visual inspection was carried out on each 100-hectare block, further subdivided into one-hectare pixels (as shown in Figure 4 by the yellow grid). Figure 4 displays a representative 100-hectare block with moderate cloudiness (30–50% cloud cover), illustrating each mask's performance: (1) CloudScore+, (2) s2cloudless, (3) L3, and (4) Planet as the reference. Panel (A) shows the composite images generated by the masks, while Panel (B) presents the visual inspection results, categorizing pixels as "clear" (purple), "cloudy" (orange), or "no data" (gray), indicating effective cloud removal, cloud presence, or mask failure, respectively.
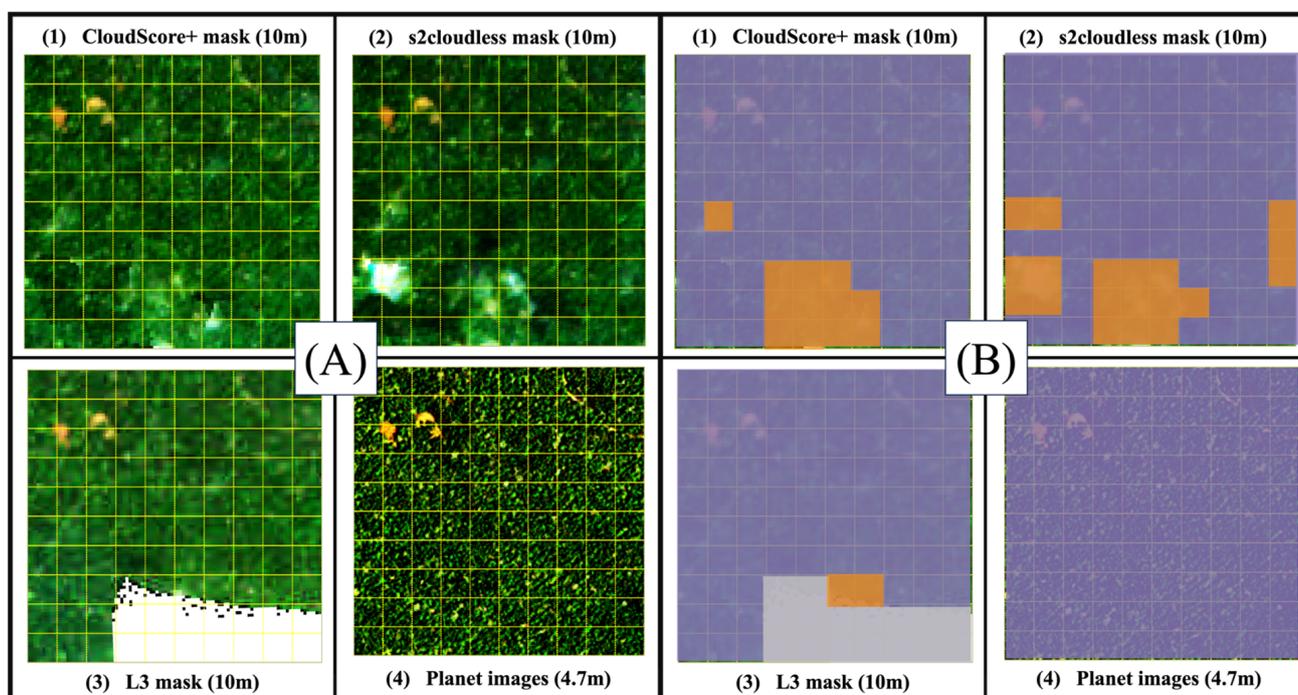
Figure 4. Example of a visual inspection of a moderate cloud cover block (ID = 13,020): Panel (**A**) shows mosaics of the evaluated masks: (**1**) CloudScore+, (**2**) s2cloudness, (**3**) L3, and (**4**) Planet reference image for comparison. Panel (**B**) shows the results after visual inspection, with "clear" pixels colored purple, "cloudy" pixels colored orange, and "no data" pixels colored gray.

For each 100-hectare block, we recorded the count of pixels classified as "clear" (effective cloud removal), "cloudy" (cloud presence), and "no data" (mask failure). These categories were assessed for each of the three Sentinel-2 cloud masks (CloudScore+, s2cloudless, and L3). In the Planet imagery, we recorded only "clear" and "cloudy" pixels, as Planet images consistently provided data without no-data areas. For the masks that allow parameter adjustment (CloudScore+ and s2cloudless), we also documented the selected parameter values.

This visual inspection process encompassed a total of 960 images, yielding 10,560 records—7680 values representing pixel classifications ("clear", "cloudy", or "no data") and 2880 for the parameter settings. Notably, Figure 4 reflects the differing spatial resolutions: 4.7 m for Planet and 10 m for Sentinel-2, which explains the smoother appearance of the Planet images compared to Sentinel-2.

2.4.2. Graphical and Statistical Analysis

All graphical and statistical analyses were conducted using R software [62]. Box plots were employed to visually represent the distribution of continuous variables, summarizing statistics such as the median, hinges, and whiskers, along with any outlying data points [63]. In these plots, successful cloud removal ("clear" pixels) is depicted in purple, cloud detection errors ("cloudy" pixels) are shown in orange, and missing data ("no data" pixels) are represented in gray.

One-way ANOVA tests were conducted to compare the proportions of "clear", "cloudy", and "no data" pixels across various factors, including mask type, year, season, and cloud cover category. Tukey's honestly significant difference (HSD) test was then applied to explore significant differences within each factor, using pairwise comparisons [64]. A two-way ANOVA model was also implemented to examine interactions between the most influential factors—mask type and cloud cover. Based on these interactions, adjusted means and pairwise comparisons were calculated using Tukey's method. The R package *emmeans* [65] was employed to compute adjusted marginal means for each combination of mask and cloud cover level [66]. In these comparisons, "estimated difference" refers to the difference between group means, where positive values indicate higher means in the first group, negative values indicate the opposite, and values near zero suggest no significant difference.

3. Results

During the visual inspection, pixels were categorized into three groups: those identified as "clear" (cloud-free), those marked as "cloudy" (clouds or other meteorological phenomena detected), and those labeled as "no data" (where the masks provided no output). In total, 10,560 values were recorded and analyzed from the 960 images included in this study, with 7680 representing pixel classifications ("clear", "cloudy", or "no data") and 2880 corresponding to parameter values used.

To better illustrate the performance of the evaluated cloud detection algorithms under varying cloud cover conditions, Figure 5 presents a series of representative 1 × 1 km blocks stratified into four cloud cover categories: low, moderate, high, and very high. Each quadrant in the figure corresponds to one cloud cover level, with individual scenes ordered from left to right based on the accuracy quartiles of the L3 mask. Specifically, the leftmost images within each quadrant represent the highest accuracy quartile (Q1), while the rightmost images depict the lowest accuracy quartile (Q4). The spatial distribution of the evaluated blocks shown in Figure 5 is further contextualized in Figure 3, where yellow stars mark the locations of the displayed examples.
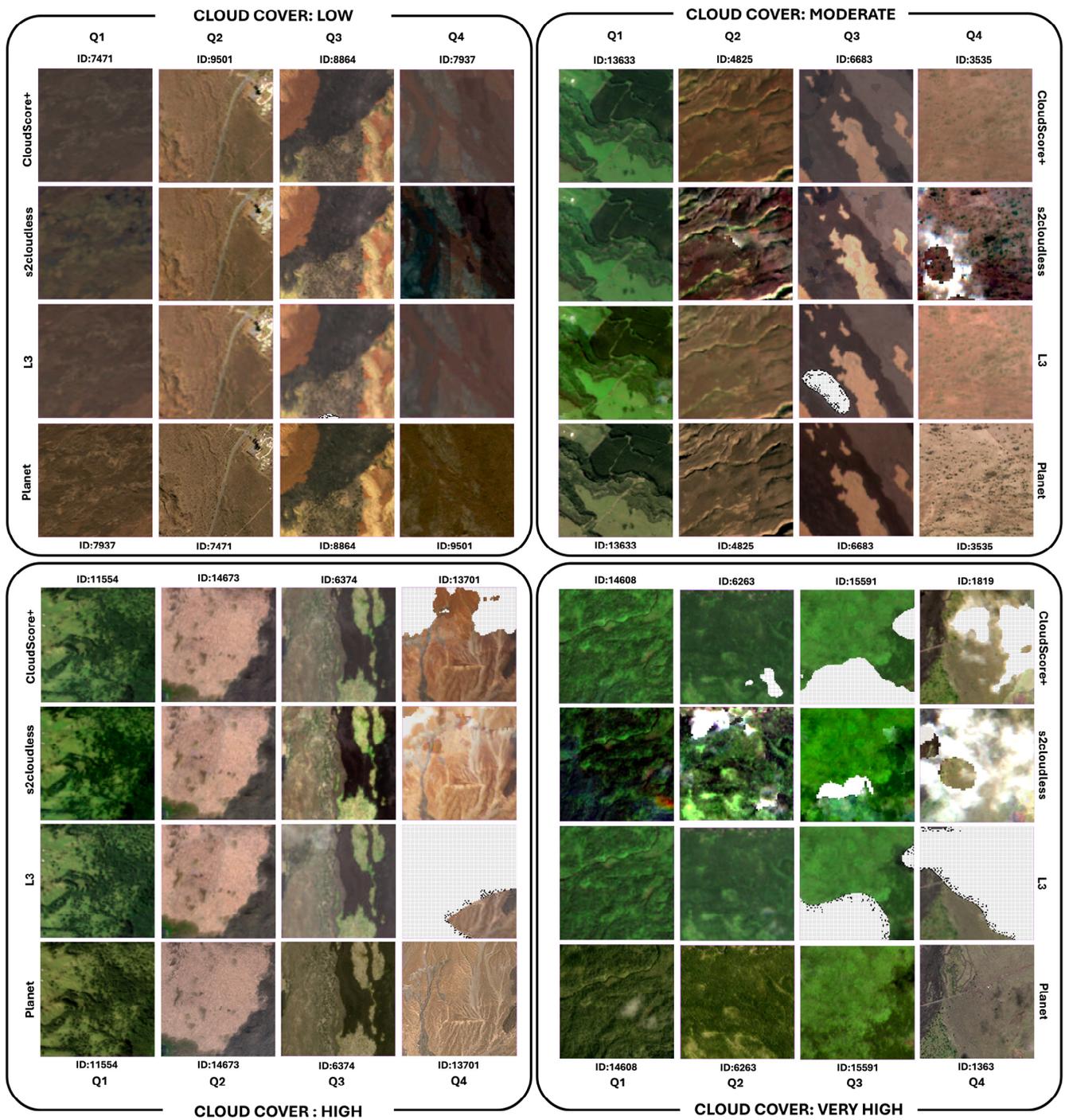
Figure 5. Visual examples of cloud detection results under varying cloud cover conditions (low, moderate, high, and very high) for CloudScore+, s2cloudless, and L3 algorithms. This figure is divided into four quadrants, each representing a specific cloud cover level. Within each quadrant, individual 1×1 km blocks are ordered from left to right according to the accuracy quartiles of the L3 mask, with the leftmost column showing the highest accuracy quartile (Q1) and the rightmost column representing the lowest accuracy quartile (Q4). Rows correspond to the evaluated masks (top to bottom: CloudScore+, s2cloudless, L3) and PlanetScope imagery as ground truth. Errors (clouds) are displayed as white areas, while "no data" pixels appear as large gray grid patterns. The spatial locations of these blocks are indicated in Figure 3 by yellow stars.

Under low and moderate cloud cover conditions, all evaluated algorithms perform consistently well across all quartiles, exhibiting minimal errors or "no data" pixels. In the

case of high cloud cover, results remain nearly perfect in the highest accuracy quartiles (Q1 and Q2), with minimal errors across all algorithms. In the third quartile (Q3), small errors begin to appear, particularly for the s2cloudless mask, where undetected clouds are represented as white areas. In the lowest accuracy quartile (Q4), s2cloudless shows more noticeable errors, while CloudScore+ and L3 primarily exhibit "no data" regions, represented as large gray grid patterns, reflecting their tendency to mask uncertain areas rather than misclassify pixels. For very high cloud cover, the challenges of cloud detection become more pronounced, with accuracy declining progressively from Q1 to Q4. In the highest accuracy quartile (Q1), results remain satisfactory for all algorithms, but the second quartile (Q2) begins to reveal errors in s2cloudless, while CloudScore+ and L3 maintain more robust results. In the third (Q3) and fourth quartiles (Q4), s2cloudless displays significant cloud-related errors, while CloudScore+ and L3 prioritize masking uncertain areas, leading to increased occurrences of "no data" pixels. This highlights the balance achieved by CloudScore+ and L3 in minimizing cloud detection errors at the cost of introducing "no data" regions, particularly under extreme cloud cover conditions.

Table 1 presents the results from a one-way ANOVA conducted for each variable across the factors considered: mask, year, season, and cloud cover. Additionally, Table 1 includes the groupings determined by the Tukey HSD mean separation test. It is important to note that each one-way ANOVA (Table 1) analyzed a single factor independently (mask, cloud cover, season, or year) without interaction with other factors.

Table 1. One-way ANOVA results for each analyzed factor (mask, cloud cover, season, and year). "Clear" and "Cloudy" pixels indicate successful and failed cloud removal, respectively. Letters "a", "b", and "c" denote statistically distinct groups according to Tukey's HSD test. Statistically significant differences are indicated in bold and by an asterisk.

Factor	Class	"Clear" (Success)		"Cloudy" (Error)		"No Data" (NA)	
		Pr (>F) [1]	Value ± SD (%)	Pr (>F) [1]	Value ± SD (%)	Pr (>F) [1]	Value ± SD (%)
mask	Planet	**<3.46 × 10^{-10} *****	94.57 ± 19.30 a	**<2 × 10^{-16} *****	5.43 ± 19.30 b	**<2 × 10^{-16} *****	--
	L3		79.26 ± 37.25 b		0.88 ± 6.38 c		20.24 ± 37.28 a
	s2cloudless		80.79 ± 28.31 b		15.39 ± 22.88 a		3.83 ± 14.92 c
	CloudScore+		89.42 ± 23.63 a		0.63 ± 1.93 c		9.94 ± 23.35 b
cloud cover	low	**2 × 10^{-16} *****	99.03 ± 4.57 a	**<2 × 10^{-16} *****	0.65 ± 3.49 c	**<2 × 10^{-16} *****	0.43 ± 3.49 c
	moderate		95.14 ± 16.34 a		2.72 ± 10.65 bc		2.86 ± 14.56 c
	high		87.01 ± 26.06 b		6.41 ± 18.23 b		9.85 ± 24.16 b
	very high		62.47 ± 38.97 c		12.69 ± 23.42 a		32.55 ± 40.38 a
season	dry	**0.0008 ***	82.35 ± 31.58 b	0.063	6.75 ± 17.92 a	**0.009 ***	14.49 ± 30.58 a
	wet		88.61 ± 25.96 a		4.75 ± 15.25 a		9.09 ± 25.12 b
year	2019	0.468	85.95 ± 26.99 a	0.206	6.97 ± 17.48 a	0.654	9.43 ± 25.15 a
	2020		88.72 ± 26.91 a		3.41 ± 11.33 a		10.49 ± 27.88 a
	2021		86.73 ± 27.18 a		6.00 ± 16.69 a		10.41 ± 26.33 a
	2022		85.27 ± 29.56 a		4.92 ± 15.96 a		12.38 ± 28.13 a
	2023		83.40 ± 31.91 a		6.62 ± 19.52 a		13.94 ± 30.45 a

[1] Significance codes: '***' for $p < 0.0001$; '*' for $p < 0.01$

Regarding the mask variable, significant statistical differences were observed in success rates ("clear"), error rates ("cloudy"), and occurrences of "no data" pixels. Planet, used as the ground truth reference, exhibited an average of approximately 95% "clear" pixels, indicating that, even with the optimal daily images selected for each month, around 5% of pixels remained cloud covered. Among the Sentinel-2 masks, CloudScore+ achieved a notable accuracy of 89.4% for "clear" pixels, closely matching Planet's results, with no significant statistical differences from the ground truth. By comparison, L3 and s2cloudless masks yielded similar and statistically indistinguishable success rates, with averages of 79.3% and 80.8% "clear" pixels, respectively.

In terms of error rates ("cloudy"), CloudScore+ and L3 had low error rates, at 0.6% and 0.9%, respectively, while s2cloudless showed a considerably higher error rate of 15.4%. Since Planet imagery does not provide "no data" values, it recorded a 5.4% error rate,

representing the residual cloud-covered pixels. The Sentinel-2 masks, capable of assigning "no data" values to avoid classification errors, yielded average "no data" percentages of 20.2% for L3, 9.9% for CloudScore+, and 3.8% for s2cloudless.

The cloud cover factor (Table 1), alongside the mask type, was identified as one of the most influential variables affecting the effectiveness of cloud-free mosaic generation. Results from the one-way ANOVA indicated strong statistical significance ($p < 2 \times 10^{-16}$) across all three outcomes: success rates ("clear" pixels), error rates ("cloudy" pixels), and "no data" occurrences. Areas with low cloud cover (<30%) displayed notably high success rates, reaching 99% accuracy, with minimal error rates and "no data" occurrences (both <1%). However, as cloud cover frequency increased, success rates declined, dropping to 62.47% in regions with very high cloud cover.

Error rates ("cloudy" pixels) and "no data" occurrences rose in tandem with cloud cover. Under low cloud cover conditions, error rates were only 0.65%, with "no data" pixels accounting for 0.43%. Conversely, in high cloud cover areas, error rates increased to 12.69%, with "no data" pixels reaching 32.55%. These findings underscore the difficulty of producing quality cloud-free mosaics in areas with very high cloud cover (over 70% year-round), a condition affecting over 30% of Hawai'i's area.

The additional factors of season and year (Table 1) showed relatively lower levels of statistical significance. Generally, the wet season yielded more favorable results, with statistically significant differences in success rates (Pr > F = 0.0008) and "no data" occurrences (Pr > F = 0.009). However, no significant seasonal differences were detected in error rates (Pr > F = 0.063).

Regarding the year factor, no statistically significant differences were found across any of the three variables analyzed ("clear", "cloudy", and "no data"). The mean success rate ("clear") ranged from 83.4% to 88.7%, error rates ("cloudy") varied from 3.4% to 7.0%, and the incidence of "no data" ranged from 10.4% to 13.9%.

Box plots were created to visually highlight the key results (Figures 6 and 7). Figure 6 presents one-way box plots showing the percentage of successful pixels ("clear") in purple, pixel errors ("cloudy") in orange, and "no data" occurrences (available only for Sentinel-2 masks) in gray. In each box plot, the boxes span from the 25th to the 75th percentile of each group's value distribution, with vertical lines extending from the boxes to indicate adjacent values, representing the most extreme points within 1.5 interquartile ranges of the 25th and 75th percentiles. Points beyond these lines are shown as outliers. This graphical representation complements the data in Table 1, providing insights into the mean values and variability across factors. For instance, the interquartile range for error rates with the s2cloudless mask indicates that most values fall between 0% and slightly above 20%, offering a quick visual interpretation of dispersion patterns across the different factors.

We further analyzed how each cloud mask interacts with factors such as cloud cover, season, and year. Table 2 displays the mean and standard deviation for the percentages of "clear", "cloudy", and "no data" pixels across these factors. Figure 7 specifically illustrates the interaction between mask type and cloud cover levels. The box plots in Figure 7 are organized according to cloud cover categories: low (upper left), moderate (upper right), high (bottom left), and very high (bottom right). As expected, Planet imagery, used as the ground truth, generally outperforms the other masks, achieving a higher percentage of cloud-free pixels across all conditions.

Under low cloud cover conditions, all masks demonstrated high performance, with nearly 100% "clear" pixels and minimal errors. Specifically, the success rates for "clear" pixels were 99.5% for Planet, 98.0% for s2cloudless, 99.2% for CloudScore+, and 99.4% for L3. Under moderate cloud cover, no statistically significant differences were observed among the masks; however, s2cloudless showed the lowest performance, with a reduced success rate and higher error percentage. The "clear" pixel success rates for each mask were 97.6% for Planet, 92.2% for s2cloudless, 97.5% for CloudScore+, and 93.2% for L3. Notably, in both low and moderate cloud cover scenarios, the percentage of "no data" pixels remained close to zero, except for L3, which exhibited a slightly higher "no data" rate of approximately 6%.
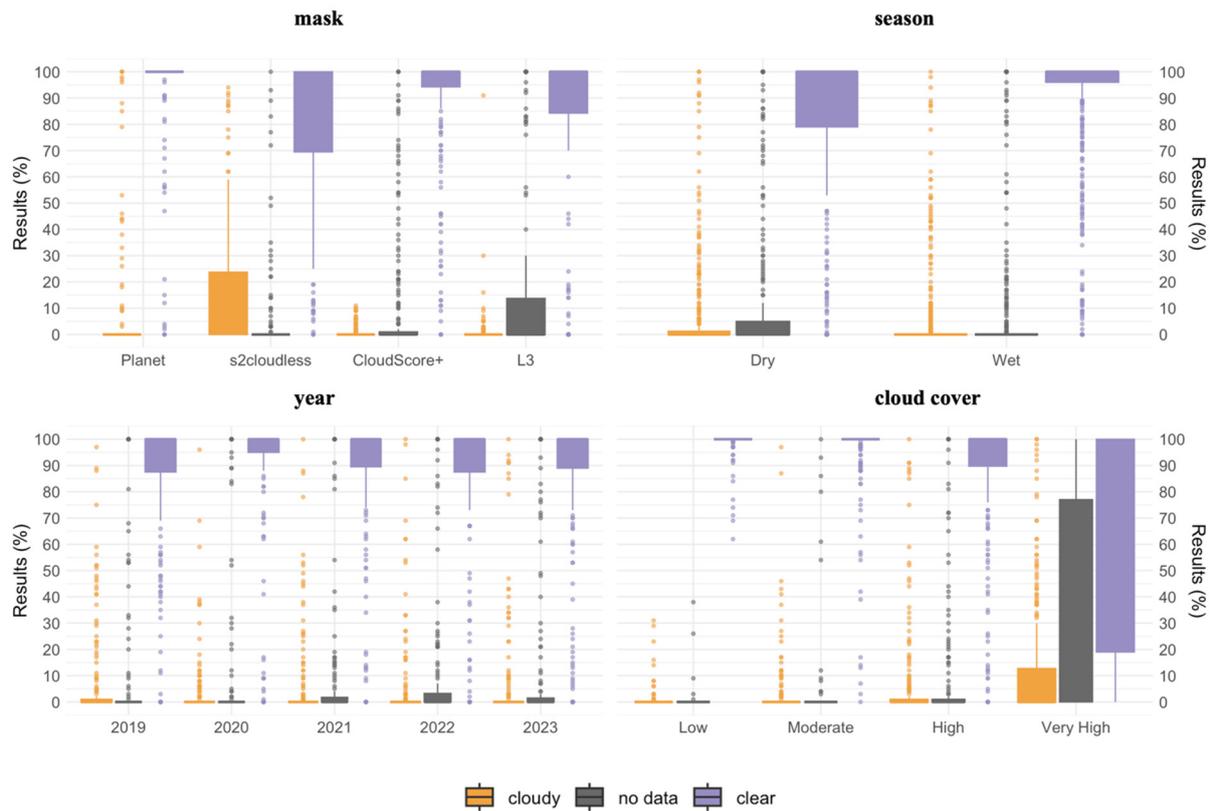
Figure 6. Box plots showing percentages of success ("clear") in purple, errors ("cloudy") in orange, and "no data" in gray for the one-way factors analyzed: (**top left**)—mask, (**top right**)—season, (**bottom left**)—year, and (**bottom right**)—cloud cover.
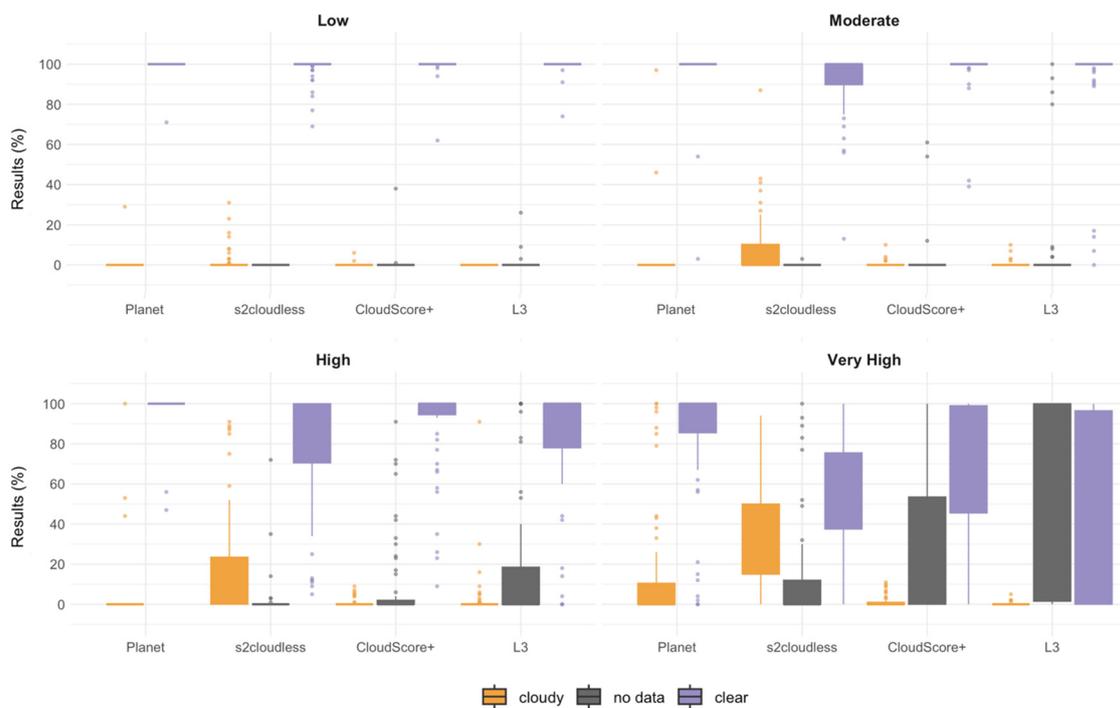
Figure 7. Box plots showing the percentages of success ("clear") in purple, errors ("cloudy") in orange, and "no data" in gray for each mask and cloud cover level combination (low, moderate, high, and very high).

Table 2. Mean and standard deviation (SD) of the observed percentage of pixels for each combination of the mask factor with other factors (cloud cover, year, and season). "Clear" indicates successful cases where no clouds are visible after visual inspection. "Cloudy" represents failures where clouds are detected, and "No Data" applies only to cases where no pixels are available for visualization.

2nd Factor	Class	"Clear" (Success) (%)								"Cloudy" (Error) (%)								No Data (%)					
		Planet		CloudScore+		s2cloudless		L3		Planet		CloudScore+		s2cloudless		L3		CloudScore+		s2cloudless		L3	
		Mean	SD	Mean	SD	Mean	SD	Mean	SD	Mean	SD	Mean	SD	Mean	SD	Mean	SD	Mean	SD	Mean	SD	Mean	SD
cloud cover	low	99.52	3.74	99.22	4.95	98.03	5.68	99.37	3.55	0.48	3.74	0.13	0.81	1.97	5.68	0.00	0.00	0.65	4.91	0.00	0.00	0.63	3.55
	moderate	97.62	13.77	97.53	10.88	92.23	15.23	93.18	22.75	2.38	13.77	0.35	1.46	7.72	15.11	0.42	1.65	2.12	10.51	0.05	0.39	6.40	22.61
	high	98.36	8.89	90.12	21.28	79.14	29.54	80.44	33.17	3.34	8.89	0.71	1.95	18.69	27.08	2.90	12.54	9.17	20.76	2.17	10.46	18.20	33.15
	very high	82.66	32.29	70.53	33.94	53.25	30.28	43.46	45.90	15.64	32.29	1.34	2.78	33.53	23.39	0.24	0.82	28.14	34.21	13.22	26.03	56.31	46.11
year	2019	93.42	17.84	89.83	22.72	78.21	26.87	82.38	35.47	6.58	17.84	0..35	1..18	20..29	25.51	0.65	2.07	9.81	22.51	1.50	6.17	16.98	35.46
	2020	97.71	13.90	90.60	23.18	84.42	26.71	82.15	36.74	2.29	13.90	0.58	1.61	9.94	15.71	0.83	4.40	8.81	23.32	5.65	19.76	17.02	36.87
	2021	94.37	17.16	91.74	19.86	82.37	25.73	78.43	38.44	7.80	22.06	0.72	1.87	15.33	22.05	0.17	0.77	7.54	19.58	2.30	7.04	21.39	38.46
	2022	92.92	24.57	87.56	26.45	82.81	27.17	77.79	37.17	5.00	20.41	0.23	0.88	13.94	22.05	0.52	2.51	12.21	25.99	3.25	12.99	21.69	37.28
	2023	94.44	21.63	87.48	25.88	76.19	34.44	75.50	39.46	5.56	21.63	1.27	3.17	17.44	26.99	2.21	13.13	11.25	25.34	6.38	21.43	24.19	39.32
season	dry	90.72	25.34	85.62	27.58	77.12	30.94	75.95	38.84	8.27	23.62	0.60	1.74	16.47	22.00	1.68	9.37	13.79	27.38	6.40	20.25	23.29	38.98
	wet	97.32	12.88	92.14	20.02	83.40	26.08	81.61	36.03	3.40	15.29	0.65	2.06	14.61	23.53	0.32	2.60	7.21	19.65	1.99	9.11	18.07	36.01

Under high cloud cover conditions, statistically significant differences between masks emerged, affecting both "clear" and "cloudy" pixels, particularly for Sentinel-2 masks, where the proportion of "no data" is also relevant. In terms of "clear" pixel success rates, no statistically significant differences were observed between Planet and CloudScore+ under both high and very high cloud cover conditions. Planet achieved success rates of 98.35% and 82.66%, respectively, while CloudScore+ reached 90.12% and 70.52%. Although Planet displayed slightly better performance, these differences were not statistically significant. L3 showed intermediate performance, with "clear" pixel success rates of 80.44% under high cloud cover and 43.45% under very high cloud cover. Conversely, s2cloudless had the lowest performance, with "clear" pixel success rates of 70.14% under high cloud cover and 53.25% under very high cloud cover.

In terms of "cloudy" pixels, the results showed distinct patterns. Under high cloud cover conditions, significant differences were found only between s2cloudless and the other masks. The s2cloudless mask produced the highest percentage of "cloudy" pixels at 18.69%, while Planet, CloudScore+, and L3 recorded lower rates at 3.34%, 0.71%, and 2.89%, respectively. Additionally, CloudScore+ and L3 exhibited substantial "no data" rates, at 9.17% and 18.20%, respectively.

Under very high cloud cover conditions, only L3 and CloudScore+ exhibited similar behavior, with no statistically significant differences between them. Both achieved minimal error rates of 0.24% and 1.34%, respectively. Notably, these low error rates are due to the tendency of both masks to classify a significant proportion of pixels as "no data" under very high cloud cover, with "no data" percentages reaching 56.30% for L3 and 28.15% for CloudScore+.

As shown in Table 2, no significant interaction was found between the mask and year factors nor between the mask and season factors. Figure 8 demonstrates the consistency in mask performance patterns across the seasons and years analyzed. Generally, the wet season shows a higher percentage of cloud-free ("clear") pixels, while inter-annual differences are minimal. Although there is a slight improvement in cloud-free percentages in 2020, this difference is not statistically significant when compared to other years. These findings were further confirmed by the results of the two-way ANOVA.

In contrast, a significant interaction was observed between the mask and cloud cover factors, suggesting that the effectiveness of each mask varies according to cloud cover levels. To assess this interaction's impact on success percentage ("clear"), error percentage ("cloudy"), and "no data" occurrences, we performed a two-way ANOVA using mask and cloud cover as categorical variables. The F-values and *p*-values indicated the statistical significance of each factor and their interaction. Table 3 presents pairwise comparisons, showing estimated differences between each combination of cloud cover level and mask algorithm. Statistically significant differences are highlighted in bold, marking instances where mask selection significantly influences outcomes based on cloud cover levels.

Table 3. Pairwise comparison of success percentage ("clear" pixels) and error percentage ("cloudy" pixels) across all combinations of cloud cover levels and mask algorithms, using Tukey's HSD Test. Statistically significant differences are highlighted in bold and marked with an asterisk.

Cloud Cover	Mask Comparison	Success ("Clear" Pixels)		Error ("Cloudy" Pixels)	
		Pr (>F) [1]	Value (%) [2]	Pr (>F) [1]	Value (%) [2]
	s2cloudless-Planet	1.0000	−1.48	0.9999	1.48
	CloudScore+-Planet	1.0000	−0.30	1.0000	−0.35
	L3-Planet	1.0000	−0.15	1.0000	−0.48
low	CloudScore+-s2cloudless	1.0000	1.18	0.9999	−1.83
	L3-s2cloudless	1.0000	1.33	0.9999	−1.97
	L3-CloudScore+	0.9997	0.15	1.0000	−0.13

Table 3. *Cont.*

Cloud Cover	Mask Comparison	Success ("Clear" Pixels)		Error ("Cloudy" Pixels)	
		Pr (>F) [1]	Value (%) [2]	Pr (>F) [1]	Value (%) [2]
moderate	s2cloudless-Planet	0.9971	−5.38	0.7596	5.33
	CloudScore+-Planet	1.0000	−0.08	0.9999	−2.03
	L3-Planet	0.9997	−4.43	0.9999	−1.97
	CloudScore+-s2cloudless	0.9975	5.30	0.2109	−7.37
	L3-s2cloudless	1.0000	0.95	0.2241	−7.30
	L3-CloudScore+	0.9998	−4.35	0.9405	−4.38
high	s2cloudless-Planet	**0.0010 *****	**−19.22**	**0.0000 *****	**15.36**
	CloudScore+-Planet	0.8725	−8.24	0.9998	−2.63
	L3-Planet	**0.0038 *****	**−17.91**	1.0000	−0.44
	CloudScore+-s2cloudless	0.4410	10.98	**0.0000 *****	**−17.98**
	L3-s2cloudless	1.0000	1.30	**0.0000 *****	**−15.80**
	L3-CloudScore+	0.6680	−9.68	0.9999	2.19
very high	s2cloudless-Planet	**0.0000 *****	**−29.41**	**0.0000 *****	**17.88**
	CloudScore+-Planet	0.2653	−12.14	**0.0000 *****	**−14.30**
	L3-Planet	**0.0000 *****	**−39.20**	**0.0000 *****	**−15.41**
	CloudScore+-s2cloudless	**0.0068 *****	**17.27**	**0.0000 *****	**−32.19**
	L3-s2cloudless	0.6478	−9.80	**0.0000 *****	**−33.29**
	L3-CloudScore+	**0.0000 *****	**−27.07**	1.0000	−1.10

[1] Significance code: '***' for $p < 0.0001$. [2] The sign of the value (%) indicates which of the two compared masks has a higher percentage of successes or errors, depending on the variable ("clear" or "cloudy"). Positive values indicate higher means in the first group, negative values indicate higher means in the second group, and values close to zero indicate no significant difference between the group means.
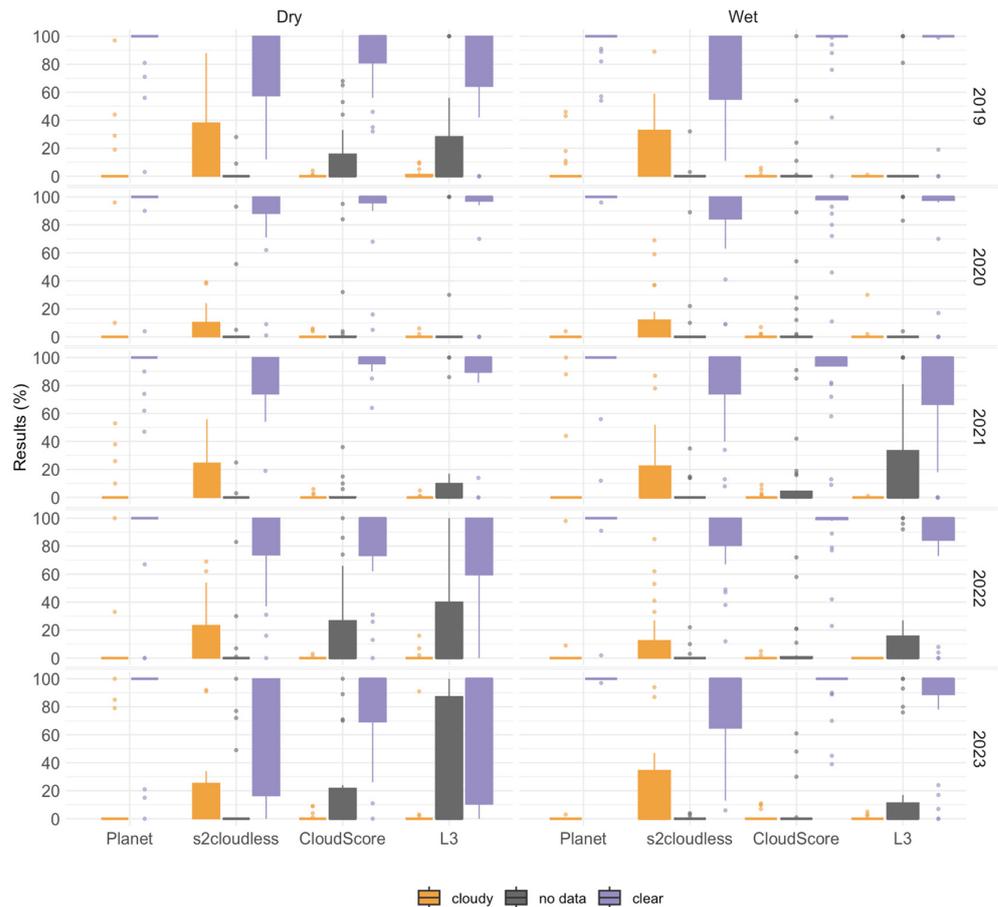

Figure 8. Box plots showing the percentages of success ("clear") in purple, errors ("cloudy") in orange, and "no data" in gray for each combination of mask, year, and season.

For the success rate of the mask ("clear" pixels), no statistically significant differences were found between the masking methods (CloudScore+, s2cloudless, L3, and Planet) under low and moderate cloud cover conditions, indicating that all methods effectively generate cloud-free mosaics in these scenarios. As detailed in Table 3, the sign of the variable value (%) reflects the difference in magnitude between the first and second masks in each pairwise comparison. For example, when comparing s2cloudless with Planet, a result of -1.48 in success rates ("clear" pixels) indicates that Planet achieves a success rate 1.48% higher than s2cloudless. Overall, Planet consistently yields the highest percentage of cloud-free pixels across all cloud cover levels. Notably, no statistically significant differences were observed between Planet and CloudScore+ under high and very high cloud cover conditions (Figure 9).
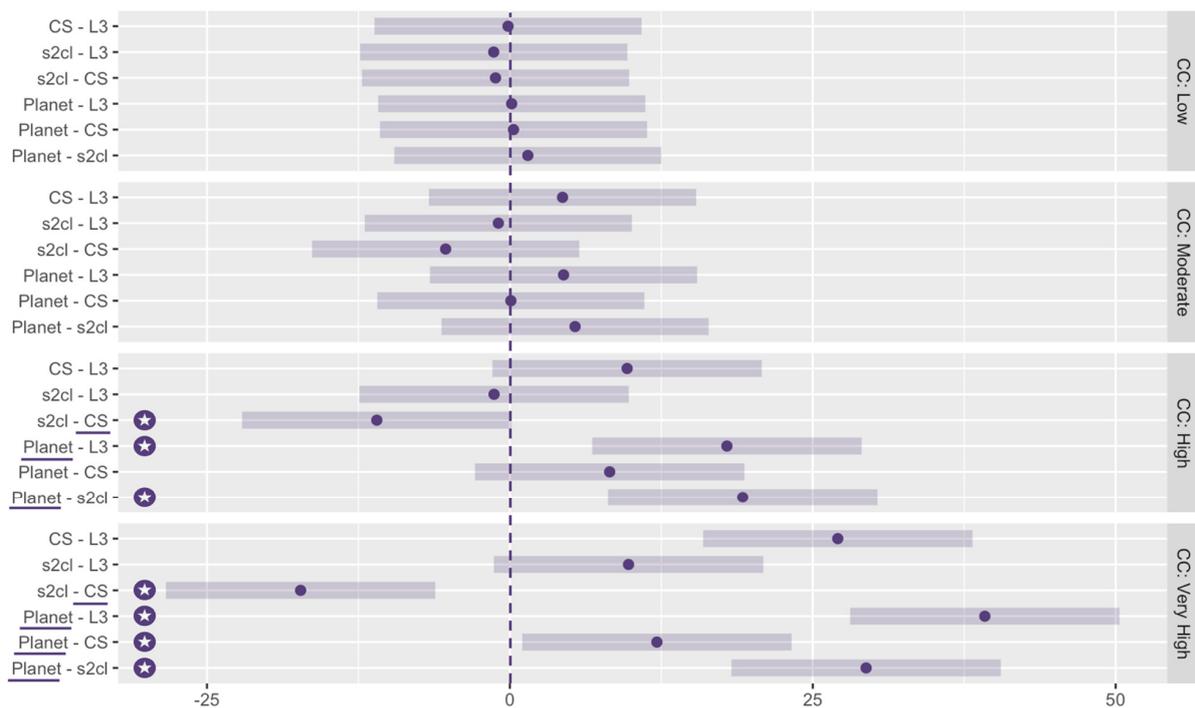
Figure 9. Pairwise comparison of success rates ("clear" pixels) across different levels of cloud cover and mask algorithms, using Tukey's HSD Test. From top to bottom, the cloud cover levels are ordered as low, moderate, high, and very high. Statistically significant differences are marked with an asterisk, and the superior mask (indicated by a higher percentage of "clear" pixels) is underlined. Abbreviations: s2cl = s2cloudless mask, CS = CloudScore+.

For error rates, representing the inability to achieve cloud-free mosaics (percentage of "cloudy" pixels), no differences were observed among the masks under low cloud cover conditions. At moderate and high cloud cover levels, CloudScore+ and L3 produced results comparable to Planet, with no significant differences, while s2cloudless showed poorer performance. Under very high cloud cover, CloudScore+ and L3 yielded similar results, both prioritizing error minimization at the expense of a higher percentage of "no data" pixels (Figure 10).

Figure 11 illustrates Tukey's mean comparisons of the percentage of "no data" pixel occurrences, highlighting both the direction and magnitude of differences between the groups compared. Statistically significant differences are marked with an asterisk, and the mask yielding fewer "no data" pixels is underlined. Under low and moderate cloud cover, no statistically significant differences were observed in "no data" occurrences across the three masks analyzed. When cloud cover is high, significant differences appear only between L3 and s2cloudless, with L3 producing 16.0% more "no data" pixels. In conditions of very high cloud cover, all masks show significant differences, with the largest observed

difference being between L3 and s2cloudless, where L3 produces 43.1% more "no data" pixels. Overall, L3 and CloudScore+ are the masks that yield the highest percentages of "no data" pixels under very high cloud cover.
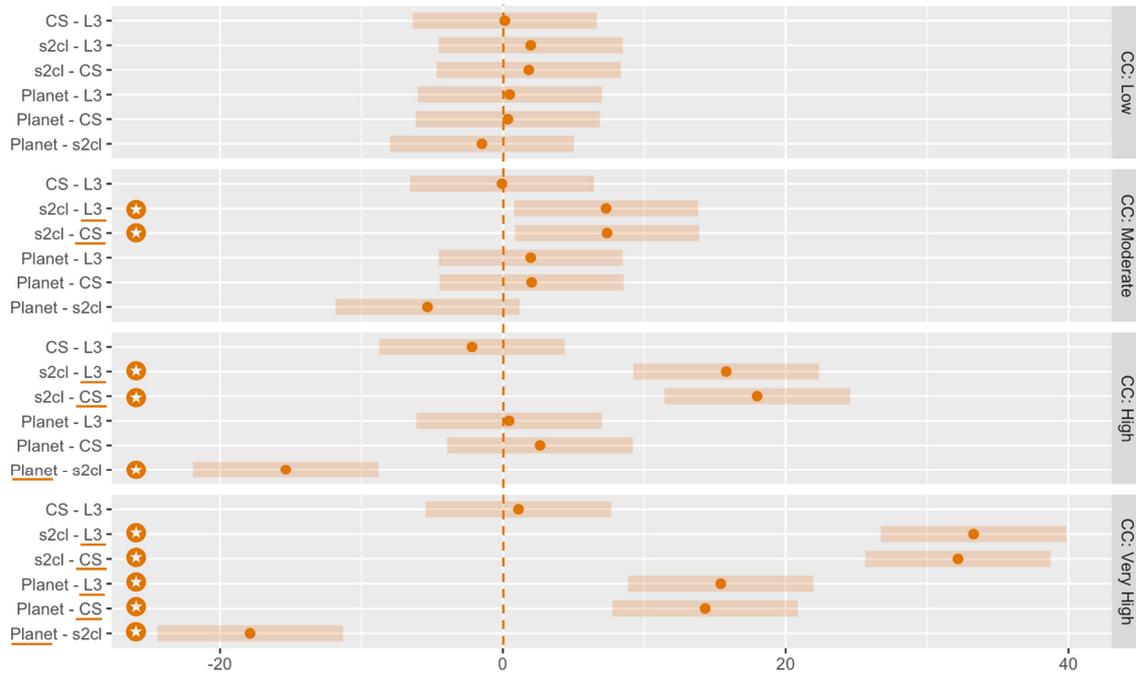
Figure 10. Pairwise comparison of error rates ("cloudy" pixels) across different levels of cloud cover and mask algorithms, using Tukey's HSD Test. From top to bottom, cloud cover levels are ordered as low, moderate, high, and very high. Statistically significant differences are marked with an asterisk, and the superior mask (indicated by a lower percentage of "cloudy" pixels) is underlined. Abbreviations: s2cl = s2cloudless mask, CS = CloudScore+.
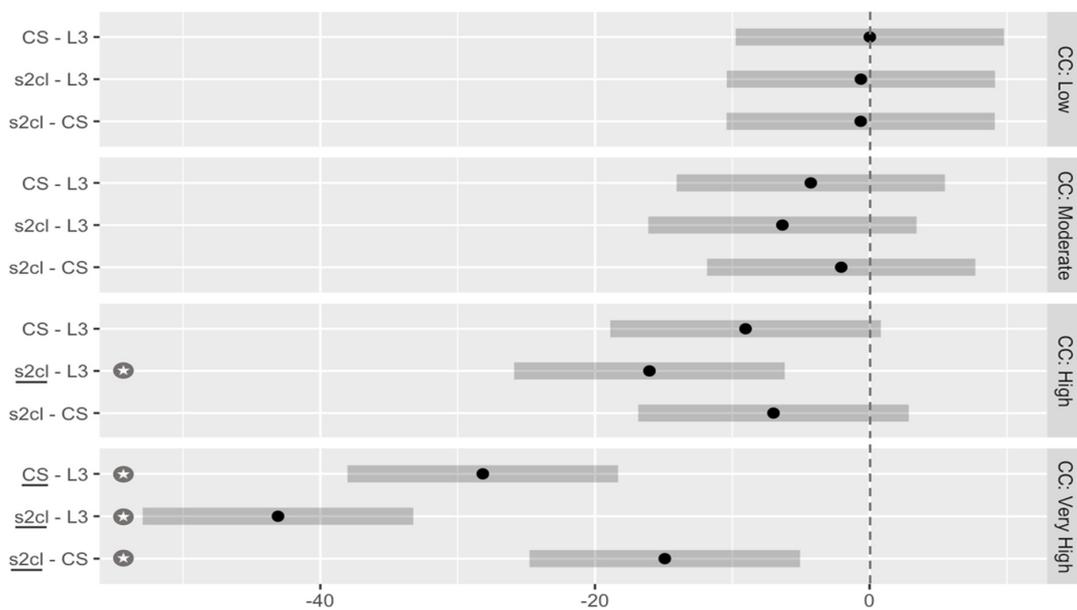
Figure 11. Pairwise comparison of "no data" percentage across different levels of cloud cover and mask algorithms, using Tukey's HSD Test. From top to bottom, cloud cover levels are ordered as low, moderate, high, and very high. Statistically significant differences are marked with an asterisk, and the mask with fewer "no data" pixels is underlined. Abbreviations: s2cl = s2cloudless mask, CS = CloudScore+.

4. Discussion

Generating cloud-free mosaics remains a substantial challenge in remote sensing, especially in regions with complex cloud dynamics like Hawai'i. This study addresses this challenge by applying and evaluating several cloud detection algorithms over a five-year period. Specifically, we compared the performance of three cloud-masking algorithms for Sentinel-2 imagery—CloudScore+, s2cloudless, and L3—against daily PlanetScope imagery as a reference.

Our findings demonstrate that all masks achieved satisfactory performance under low and moderate cloud coverage (mean annual cloud cover <50%). However, as cloud cover increased, differences in mask performance became more evident. Specifically, CloudScore+ maintained robust accuracy even under high cloud cover conditions, whereas s2cloudless and L3 displayed greater sensitivity to cloud cover variability, resulting in higher error rates and an increase in 'no data' pixels under these challenging conditions. These observations emphasize the necessity of selecting cloud detection algorithms that align with the predominant cloud conditions. Each algorithm exhibits distinct strengths and limitations depending on cloud frequency, with CloudScore+ showing greater stability, while L3's deep learning model sacrifices some accuracy to minimize errors in data masking, a trade-off that can affect long-term ecological monitoring in cloud-dense environments.

Among the Sentinel-2 masks evaluated, CloudScore+ distinguishes itself due to its demonstrated adaptability and efficiency in processing data across varied geographic contexts and application areas. Previous applications of CloudScore [67–69] highlight its foundational reliability, which has been further enhanced in the updated CloudScore+ [49,50,70,71]. Studies comparing CloudScore+ to other cloud-masking techniques, including QA60, S2cloudless, and CDI, consistently showcase its robustness and accuracy in detecting clouds, affirming its status as a competitive choice for reliable cloud masking [49,50].

Of the additional masks evaluated, L3 and s2cloudless, L3 consistently produced more robust results, particularly in areas with moderate to high cloud cover. However, under very high cloud cover, both masks exhibited noticeable performance reductions. The L3 mask demonstrated statistically lower error rates even in high cloud conditions, but this accuracy comes at a cost, with a substantial increase in 'no data' pixel values. This trade-off highlights a critical consideration in ecological monitoring, where maintaining data completeness is often as important as accuracy. Selecting an algorithm must therefore account for this balance, with L3 offering an option that minimizes errors at the expense of data coverage. This balance is essential for applications in cloud-dense regions where complete cloud removal remains challenging. Both L3 and CloudScore+ prioritize data accuracy by minimizing false detections, a critical aspect in contexts such as ecological monitoring and resource management, where erroneous data can significantly hinder decision-making processes [72,73], underscoring the importance of reliable tools for environmental analysis [74].

Effective cloud masking not only enhances the quality of environmental assessments but also strengthens the reliability of data-driven decisions, which is essential for ecosystem management and conservation [75]. This focus on data precision is particularly critical in cloud-heavy regions such as the Hawaiian Islands, where frequent cloud interference poses a major challenge to obtaining accurate satellite imagery [76]. In remote sensing, accurate data are indispensable; errors can lead to suboptimal decisions that negatively impact ecosystem health and conservation efforts.

Regions with persistently high cloud cover pose substantial challenges for optical remote sensing, especially in ecosystems of significant conservation value, like Hawai'i's rainforests and coffee-growing areas. The leeward side of the island, characterized by frequent cloud cover, includes critical native forest habitats currently facing multiple environmental threats, as well as coffee plantations essential to the state's agriculture [77,78]. The limited availability of reliable imagery from these areas hinders large-scale efforts to monitor forest pathogens, invasive species, and other threats to native biodiversity. While radar satellite imagery can penetrate clouds and holds promise for permanently clouded

tropical regions [79–81], it falls short in detecting some forest health indicators critical to ecological management. Aerial platforms, both crewed and uncrewed, which can operate below the cloud deck, provide a viable alternative; however, the high costs and regulatory restrictions on aerial operations limit their practicality for extensive monitoring across large areas. Based on this study's findings, CloudScore+ offers a partial solution for these persistently cloudy regions, and Planet imagery also shows promise in providing frequent, cloud-free data. An integrated approach combining multiple masks, daily Planet imagery, and radar data could offer a viable strategy for areas with frequent cloud cover.

Google Earth Engine (GEE) has established itself as a transformative platform for large-scale Earth observation data processing, enabling the integration of deep learning models for cloud detection, such as CloudScore+ and s2cloudless [82], and supporting algorithms like BFAST for extensive environmental monitoring [83]. This study underscores GEE's robust capabilities for generating large-scale, cloud-free mosaics, which prove essential for addressing complex environmental monitoring challenges, such as tracking Amazon deforestation [84]. The platform's compatibility with advanced convolutional neural networks (CNNs), including Faster R-CNN and DeepMask, has markedly improved the accuracy and availability of cloud-free mosaics for a wide range of applications [85].

While this study demonstrates the effectiveness of CNN-based cloud-masking approaches, future advancements could benefit from integrating transformer-based architectures. Transformers excel in capturing complex spatial and temporal relationships, offering the potential to enhance the detection of clouds in challenging environments [27,28,30]. However, their high computational demands and reliance on extensive, diverse training datasets currently present significant barriers to operational scalability [31,33].

Our findings indicate that PlanetScope imagery generally outperformed Sentinel-2 in producing cloud-free mosaics, primarily due to its high revisit frequency, which is beneficial for regions with fluctuating cloud cover. However, in areas classified as having very high cloud cover, even PlanetScope images may occasionally lack a completely cloud-free scene within specific timeframes. Selecting the optimal satellite system is therefore essential and should consider unique environmental and climatic conditions. Sentinel-2 provides valuable data for land cover and land use monitoring, as highlighted by Phiri et al. [41]; however, PlanetScope's combination of high spatial resolution and frequent imaging presents distinct advantages. For instance, integrating high-resolution multispectral imagery, such as that from PlanetScope, has been shown to enhance forest monitoring accuracy and classification, capturing detailed canopy information [86]. High-resolution satellite data have also proven valuable in monitoring agricultural parameters, such as stem water potential in vineyards (Helman et al., 2018) and estimating canopy features and management zones in almond orchards [87]. Additionally, this detailed imagery supports predictive analyses in forestry, such as assessing tree mortality risk in tropical eucalyptus plantations [45].

The Hawaiian Islands' diverse climatic landscape poses substantial challenges for remote sensing, particularly due to variations in cloud cover associated with phenomena like El Niño, which modulates cloud patterns and environmental conditions [88]. Analyses by Barnes et al. [89] illustrate the complexity of Hawaiian cloud cover patterns, highlighting the islands' distinct diurnal and seasonal shifts in cloud dynamics. Further studies by Zhang et al. [90] examined the relationships between cloud cover, sea surface temperature, and ENSO (El Niño–Southern Oscillation) events, providing critical insights for environmental monitoring and disaster response in areas sensitive to such variations. This correlation between precipitation and cloud frequency [91] underscores the need for adaptable remote sensing approaches, especially considering how cloud cover variability impacts ecosystem productivity [92]. The subtle seasonal differences observed in this five-year study may reflect the influence of the Oceanic Niño Index (ONI) on cloud distribution, emphasizing the necessity of seasonally responsive and adaptable remote sensing techniques.

Optimizing remote sensing strategies requires a nuanced understanding of how climate patterns and phenomena influence cloud cover, especially in regions like the Hawaiian Islands where persistent cloud cover poses considerable challenges for acquiring accu-

rate satellite imagery. This issue is particularly pronounced in humid ecosystems, where precipitation critically influences ecosystem productivity and growth [76]. Implementing tailored cloud-masking approaches across diverse ecological regions is thus essential for reliable data acquisition [33]. In tropical forests, cloud cover substantially impacts net CO2 uptake and rainforest tree growth during rainy seasons [92]. Similarly, improved cloud-masking techniques are crucial in boreal forest regions to address challenges like cloud–snow confusion, which can affect the accuracy of snow mapping products [93]. Our findings support those of Robinson et al. [2] and Park [94], who documented the limitations imposed by cloud cover on environmental monitoring and disaster response, emphasizing the critical role of cloud-free satellite imagery in understanding and managing terrestrial ecosystems.

The limited performance of cloud-masking algorithms under high and very high cloud coverage represents a significant challenge in optical remote sensing. A promising approach to mitigate this limitation is the integration of optical imagery with synthetic aperture radar (SAR) data, such as Sentinel-1. Unlike optical sensors, SAR can penetrate cloud cover, providing complementary information that enhances land cover mapping and change detection [95,96]. Moreover, the application of advanced machine learning techniques to combine these multimodal datasets could further improve the reconstruction of cloud-obscured areas, offering a pathway towards more accurate and reliable environmental monitoring [97,98].

5. Conclusions

This study evaluated the performance of three Sentinel-2 cloud detection algorithms—CloudScore+, s2cloudless, and the CNN-based L3—alongside PlanetScope imagery for environmental remote sensing on Hawai'i Island. Among the Sentinel-2 masks, Cloud-Score+ demonstrated the highest accuracy, closely matching the performance of PlanetScope imagery. Both s2cloudless and L3 also showed effective results, underscoring the potential of machine learning in enhancing cloud detection. Despite its lower radiometric resolution, PlanetScope imagery provided superior spatial detail and achieved cloud-free images in over 95% of cases, highlighting the advantages of high-frequency imaging for constructing reliable cloud-free mosaics.

These findings reinforce the importance of pairing advanced cloud detection algorithms with carefully selected satellite systems to optimize remote sensing for environmental monitoring. While the current L3 algorithm effectively balances accuracy and data coverage, future advancements could explore a hybrid approach that combines the computational efficiency of CNNs with the spatial-temporal modeling capabilities of transformers. Such an approach has the potential to significantly enhance cloud detection in persistently cloudy regions, though challenges related to computational demands and the need for diverse training datasets must be addressed.

Future research should focus on refining these cloud detection algorithms and incorporating complementary data sources, such as radar, to mitigate the limitations posed by persistent cloud cover and seasonal climate variability. By advancing these methods, remote sensing can continue to serve as a vital tool for accurate ecological assessments and long-term climate studies, especially in cloud-heavy regions like Hawai'i.

References

1. Zou, X.; Yi, Z.; Tao, P.; Cui, Y. Research on Cloud Removal Based on Fusing Multi-Temporal Remote Sensing Images. In Proceedings of the 2022 IEEE 10th Joint International Information Technology and Artificial Intelligence Conference (ITAIC), Chongqing, China, 17–19 June 2022. [CrossRef]

2. Robinson, T.; Rosser, N.; Walters, R.J. The Spatial and Temporal Influence of Cloud Cover on Satellite-Based Emergency Mapping of Earthquake Disasters. *Sci. Rep.* **2019**, *9*, 12455. [CrossRef]

3. Meyer, S.T.; Koch, C.; Weisser, W.W. Towards a Standardized Rapid Ecosystem Function Assessment (REFA). *Trends Ecol. Evol.* **2015**, *30*, 390–397. [CrossRef] [PubMed]

4. Kustas, W.P.; Norman, J.M. Use of Remote Sensing for Evapotranspiration Monitoring over Land Surfaces. *Hydrol. Sci. J.* **1996**, *41*, 495–516. [CrossRef]

5. Pettorelli, N.; Laurance, W.F.; O'Brien, T.G.; Wegmann, M.; Nagendra, H.; Turner, W. Satellite Remote Sensing for Applied Ecologists: Opportunities and Challenges. *J. Appl. Ecol.* **2014**, *51*, 839–848. [CrossRef]

6. Turner, M.G.; Calder, W.J.; Cumming, G.S.; Hughes, T.P.; Jentsch, A.; LaDeau, S.L.; Lenton, T.M.; Shuman, B.N.; Turetsky, M.R.; Ratajczak, Z. Climate Change, Ecosystems and Abrupt Change: Science Priorities. *Philos. Trans. R. Soc. B* **2020**, *375*, 20190105. [CrossRef] [PubMed]

7. Rossow, W.B.; Garder, L.C. Cloud Detection Using Satellite Measurements of Infrared and Visible Radiances for ISCCP. *J. Clim.* **1993**, *6*, 2341–2369. [CrossRef]

8. Sishodia, R.P.; Ray, R.L.; Singh, S.K. Applications of Remote Sensing in Precision Agriculture: A Review. *Remote Sens.* **2020**, *12*, 3136. [CrossRef]

9. Burgan, R.E. *Monitoring Vegetation Greenness with Satellite Data*; US Department of Agriculture, Forest Service, Intermountain Research Station: Ogden, UT, USA, 1993; Volume 297.

10. Cannon, P.; Friday, J.B.; Harrington, T.; Keith, L.; Hughes, M.; Hauff, R.; Hughes, F.; Perroy, R.; Benitez, D.; Roy, K. Rapid 'Ōhi 'a Death in Hawai 'i. In *Forest Microbiology*; Elsevier: Amsterdam, The Netherlands, 2022; pp. 267–289.

11. Roy, K.; Granthon, C.; Peck, R.W.; Atkinson, C.T. *Effectiveness of Rapid Ohia Death Management Strategies at a Focal Disease Outbreak on Hawaii Island*; HCSU Technical Report Series 99; University of Hawai'i at Manoa: Honolulu, HI, USA, 2021.

12. Perroy, R.L.; Sullivan, T.; Benitez, D.; Hughes, R.F.; Keith, L.M.; Brill, E.; Kissinger, K.; Duda, D. Spatial Patterns of 'ōhi 'a Mortality Associated with Rapid 'ōhi 'a Death and Ungulate Presence. *Forests* **2021**, *12*, 1035. [CrossRef]

13. Luiz, B.C.; Giardina, C.P.; Keith, L.M.; Jacobs, D.F.; Sniezko, R.A.; Hughes, M.A.; Friday, J.B.; Cannon, P.; Hauff, R.; Francisco, K. A Framework for Establishing a Rapid 'Ōhi 'a Death Resistance Program. *New For.* **2023**, *54*, 637–660. [CrossRef]

14. Perroy, R.L.; Sullivan, T.; Stephenson, N. Assessing the Impacts of Canopy Openness and Flight Parameters on Detecting a Sub-Canopy Tropical Invasive Plant Using a Small Unmanned Aerial System. *ISPRS J. Photogramm. Remote Sens.* **2017**, *125*, 174–183. [CrossRef]

15. Pejchar, L.; Lepczyk, C.A.; Fantle-Lepczyk, J.E.; Hess, S.C.; Johnson, M.T.; Leopold, C.R.; Marchetti, M.; McClure, K.M.; Shiels, A.B. Hawaii as a Microcosm: Advancing the Science and Practice of Managing Introduced and Invasive Species. *Bioscience* **2020**, *70*, 184–193. [CrossRef]

16. Kendall, S.J.; Rounds, R.A.; Camp, R.J.; Genz, A.S.; Cady, T.; Ball, D.L. Forest Bird Populations at the Big Island National Wildlife Refuge Complex, Hawai'i. *J. Fish Wildl. Manag.* **2024**, *14*, 410–432. [CrossRef]

17. Schörghofer, N.; Leopold, M.; Klasner, F.L. A Third Pond on the Mauna Kea Summit Plateau. *Pac. Sci.* **2023**, *76*, 277–287. [CrossRef]

18. Trauernicht, C. Vegetation—Rainfall Interactions Reveal How Climate Variability and Climate Change Alter Spatial Patterns of Wildland Fire Probability on Big Island, Hawaii. *Sci. Total Environ.* **2019**, *650*, 459–469. [CrossRef] [PubMed]

19. López-Puigdollers, D.; Mateo-García, G.; Gómez-Chova, L. Benchmarking Deep Learning Models for Cloud Detection in Landsat-8 and Sentinel-2 Images. *Remote Sens.* **2021**, *13*, 992. [CrossRef]

20. Han, L.; Wu, T.; Liu, Q.; Liu, Z. A Novel Approach for Cloud Detection in Scenes With Snow/Ice Using High Resolution Sentinel-2 Images. *Atmosphere* **2019**, *10*, 44. [CrossRef]

21. Sánchez, A.; Picoli, M.C.A.; Câmara, G.; Andrade, P.R.; Chaves, M.E.D.; Lechler, S.; Soares, A.R.; Marujo, R.d.F.B.; Simões, R.; Ferreira, K.R.; et al. Comparison of Cloud Cover Detection Algorithms on Sentinel–2 Images of the Amazon Tropical Forest. *Remote Sens.* **2020**, *12*, 1284. [CrossRef]

22. Hagolle, O.; Colin, J.; Coustance, S.; Kettig, P.; d'Angelo, P.; Auer, S.; Doxani, G.; Desjardins, C. Sentinel-2 Surface Reflectance Products Generated by CNES and DLR: Methods, Validation and Applications. *ISPRS Ann. Photogramm. Remote Sens. Spat. Inf. Sci.* **2021**, *51*, 9–15. [CrossRef]

23. Baetens, L.; Desjardins, C.; Hagolle, O. Validation of Copernicus Sentinel-2 Cloud Masks Obtained from MAJA, Sen2Cor, and FMask Processors Using Reference Cloud Masks Generated with a Supervised Active Learning Procedure. *Remote Sens.* **2019**, *11*, 433. [CrossRef]

24. Romero, A.; Gatta, C.; Camps-Valls, G. Unsupervised Deep Feature Extraction for Remote Sensing Image Classification. *IEEE Trans. Geosci. Remote Sens.* **2015**, *54*, 1349–1362. [CrossRef]

25. Mateo-García, G.; Gómez-Chova, L.; Camps-Valls, G. Convolutional Neural Networks for Multispectral Image Cloud Masking. In Proceedings of the 2017 IEEE International Geoscience and Remote Sensing Symposium (IGARSS), Fort Worth, TX, USA, 23–28 July 2017; pp. 2255–2258.

26. Wu, W.; Gao, X.; Fan, J.; Xia, L.; Luo, J.; Zhou, Y. Improved Mask R-CNN-Based Cloud Masking Method for Remote Sensing Images. *Int. J. Remote Sens.* **2020**, *41*, 8910–8933. [CrossRef]

27. Gong, C.; Long, T.; Yin, R.; Jiao, W.; Wang, G. A Hybrid Algorithm With Swin Transformer and Convolution for Cloud Detection. *Remote Sens.* **2023**, *15*, 5264. [CrossRef]

28. Wu, C.; Chen, X.; Li, X. Mask Transformer: Unpaired Text Style Transfer Based on Masked Language. *Appl. Sci.* **2020**, *10*, 6196. [CrossRef]

29. Dosovitskiy, A. An Image Is Worth 16x16 Words: Transformers for Image Recognition at Scale. *arXiv* **2020**, arXiv:2010.11929.

30. Stucker, C.; Garnot, V.S.F.; Schindler, K. U-TILISE: A Sequence-to-Sequence Model for Cloud Removal in Optical Satellite Time Series. *IEEE Trans. Geosci. Remote Sens.* **2023**, *61*, 1–16. [CrossRef]

31. Dimyati, M.; Husna, A.; Handayani, P.T.; Annisa, D.N. Cloud Removal on Satellite Imagery Using Blended Model: Case Study Using Quick Look of High-Resolution Image of Indonesia. *Telkomnika (Telecommun. Comput. Electron. Control)* **2022**, *20*, 373. [CrossRef]

32. Zhao, Y.; Wang, G.; Tang, C.; Luo, C.; Zeng, W.; Zha, Z.-J. A Battle of Network Structures: An Empirical Study of Cnn, Transformer, and Mlp. *arXiv* **2021**, arXiv:2108.13002.

33. Candra, D.S.; Phinn, S.R.; Scarth, P. Automated Cloud and Cloud-Shadow Masking for Landsat 8 Using Multitemporal Images in a Variety of Environments. *Remote Sens.* **2019**, *11*, 2060. [CrossRef]

34. Amani, M.; Ghorbanian, A.; Ahmadi, S.A.; Kakooei, M.; Moghimi, A.; Mirmazloumi, S.M.; Moghaddam, S.H.A.; Mahdavi, S.; Ghahremanloo, M.; Parsian, S. Google Earth Engine Cloud Computing Platform for Remote Sensing Big Data Applications: A Comprehensive Review. *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.* **2020**, *13*, 5326–5350. [CrossRef]

35. Gorelick, N.; Hancher, M.; Dixon, M.; Ilyushchenko, S.; Thau, D.; Moore, R. Google Earth Engine: Planetary-Scale Geospatial Analysis for Everyone. *Remote Sens. Environ.* **2017**, *202*, 18–27. [CrossRef]

36. Singh, R.; Biswas, M.; Pal, M. Cloud Detection Using Sentinel 2 Imageries: A Comparison of XGBoost, RF, SVM, and CNN Algorithms. *Geocarto Int.* **2022**, *38*, 1–32. [CrossRef]

37. Pasquarella, V.J.; Brown, C.F.; Czerwinski, W.; Rucklidge, W.J. Comprehensive Quality Assessment of Optical Satellite Imagery Using Weakly Supervised Video Learning. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, Vancouver, BC, Canada, 17–24 June 2023; pp. 2124–2134.

38. Kottek, M.; Grieser, J.; Beck, C.; Rudolf, B.; Rubel, F. World Map of the Köppen-Geiger Climate Classification Updated. *Meteorol. Z.* **2006**, *15*, 259–263. [CrossRef]

39. Giambelluca, T.W.; Chen, Q.; Frazier, A.G.; Price, J.P.; Chen, Y.-L.; Chu, P.-S.; Eischeid, J.K.; Delparte, D.M. Online Rainfall Atlas of Hawai 'i. *Bull. Am. Meteorol. Soc.* **2013**, *94*, 313–316. [CrossRef]

40. Miller, J. *Mauna Loa Observatory: A 20th Anniversary Report*; US Department of Commerce, National Oceanic and Atmospheric Administration: Silver Spring, MD, USA, 1978.

41. Phiri, D.; Simwanda, M.; Salekin, S.; Nyirenda, V.R.; Murayama, Y.; Ranagalage, M. Sentinel-2 Data for Land Cover/Use Mapping: A Review. *Remote Sens.* **2020**, *12*, 2291. [CrossRef]

42. Helber, P.; Bischke, B.; Hees, J.; Dengel, A. Towards a Sentinel-2 Based Human Settlement Layer. In Proceedings of the IGARSS 2019—2019 IEEE International Geoscience and Remote Sensing Symposium, Yokohama, Japan, 28 July–2 August 2019; pp. 5936–5939.

43. Marta, S. Planet Imagery Product Specifications. *Planet Labs San Fr. CA USA* **2018**, *91*, 170.

44. Frazier, A.E.; Hemingway, B.L. A Technical Review of Planet Smallsat Data: Practical Considerations for Processing and Using Planetscope Imagery. *Remote Sens.* **2021**, *13*, 3930. [CrossRef]

45. Pascual, A.; Tupinambá-Simões, F.; Guerra-Hernández, J.; Bravo, F. High-Resolution Planet Satellite Imagery and Multi-Temporal Surveys to Predict Risk of Tree Mortality in Tropical Eucalypt Forestry. *J. Environ. Manag.* **2022**, *310*, 114804. [CrossRef] [PubMed]

46. Nguyen, H.T.T.; Doan, T.M.; Tomppo, E.; McRoberts, R.E. Land Use/Land Cover Mapping Using Multitemporal Sentinel-2 Imagery and Four Classification Methods—A Case Study From Dak Nong, Vietnam. *Remote Sens.* **2020**, *12*, 1367. [CrossRef]

47. Housman, I.W.; Chastain, R.A.; Finco, M. V An Evaluation of Forest Health Insect and Disease Survey Data and Satellite-Based Remote Sensing Forest Change Detection Methods: Case Studies in the United States. *Remote Sens.* **2018**, *10*, 1184. [CrossRef]

48. Chastain, R.; Housman, I.; Goldstein, J.; Finco, M.; Tenneson, K. Empirical Cross Sensor Comparison of Sentinel-2A and 2B MSI, Landsat-8 OLI, and Landsat-7 ETM+ Top of Atmosphere Spectral Characteristics over the Conterminous United States. *Remote Sens. Environ.* **2019**, *221*, 274–285. [CrossRef]

49. Gong, C.; Yin, R.; Long, T.; Jiao, W.; He, G.; Wang, G. Spatial–Temporal Approach and Dataset for Enhancing Cloud Detection in Sentinel-2 Imagery: A Case Study in China. *Remote Sens.* **2024**, *16*, 973. [CrossRef]

50. Liang, K.; Yang, G.; Zuo, Y.; Chen, J.; Sun, W.; Meng, X.; Chen, B. A Novel Method for Cloud and Cloud Shadow Detection Based on the Maximum and Minimum Values of Sentinel-2 Time Series Images. *Remote Sens.* **2024**, *16*, 1392. [CrossRef]

51. Zupanc, A. Improving Cloud Detection with Machine Learning [WWW Document]. Sentin. Hub Blog. 2017. Available online: https://medium.com/sentinel-hub/improvingcloud-detection-with-machine-learning-c09dc5d7cf13 (accessed on 26 October 2021).

52. Hagolle, O.; Morin, D.; Kadiri, M. Detailed Processing Model for the Weighted Average Synthesis Processor (WASP) for Sentinel-2 (1.4). 2018. Available online: https://zenodo.org/record/1401360 (accessed on 24 March 2021).

53. Baetens, L.; Hagolle, O. Sentinel-2 Reference Cloud Masks Generated by an Active Learning Method. Type: Dataset. 2018. Available online: https://zenodo.org/record/1460961 (accessed on 19 February 2019).

54. Lim, B.; Son, S.; Kim, H.; Nah, S.; Mu Lee, K. Enhanced Deep Residual Networks for Single Image Super-Resolution. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition Workshops, Honolulu, HI, USA, 21–26 July 2017; pp. 136–144.

55. Lanaras, C.; Bioucas-Dias, J.; Galliani, S.; Baltsavias, E.; Schindler, K. Super-Resolution of Sentinel-2 Images: Learning a Globally Applicable Deep Neural Network. *ISPRS J. Photogramm. Remote Sens.* **2018**, *146*, 305–319. [CrossRef]

56. Wang, J.; Sun, K.; Cheng, T.; Jiang, B.; Deng, C.; Zhao, Y.; Liu, D.; Mu, Y.; Tan, M.; Wang, X. Deep High-Resolution Representation Learning for Visual Recognition. *IEEE Trans. Pattern Anal. Mach. Intell.* **2020**, *43*, 3349–3364. [CrossRef] [PubMed]

57. Main-Knorn, M.; Pflug, B.; Louis, J.; Debaecker, V.; Müller-Wilm, U.; Gascon, F. Sen2Cor for Sentinel-2. In Proceedings of the Image and Signal Processing for Remote Sensing XXIII, Warsaw, Poland, 11–13 September 2017; Volume 10427, pp. 37–48.

58. Louis, J.; Debaecker, V.; Pflug, B.; Main-Knorn, M.; Bieniarz, J.; Mueller-Wilm, U.; Cadau, E.; Gascon, F. Sentinel-2 Sen2Cor: L2A Processor for Users. In Proceedings of the Living Planet Symposium 2016, Prague, Czech Republic, 9–13 May 2016; Spacebooks Online. pp. 1–8.

59. Foga, S.; Scaramuzza, P.L.; Guo, S.; Zhu, Z.; Dilley, R.D., Jr.; Beckmann, T.; Schmidt, G.L.; Dwyer, J.L.; Hughes, M.J.; Laue, B. Cloud Detection Algorithm Comparison and Validation for Operational Landsat Data Products. *Remote Sens. Environ.* **2017**, *194*, 379–390. [CrossRef]

60. Skakun, S.; Wevers, J.; Brockmann, C.; Doxani, G.; Aleksandrov, M.; Batič, M.; Frantz, D.; Gascon, F.; Gómez-Chova, L.; Hagolle, O. Cloud Mask Intercomparison EXercise (CMIX): An Evaluation of Cloud Masking Algorithms for Landsat 8 and Sentinel-2. *Remote Sens. Environ.* **2022**, *274*, 112990. [CrossRef]

61. Aybar, C.; Ysuhuaylas, L.; Loja, J.; Gonzales, K.; Herrera, F.; Bautista, L.; Yali, R.; Flores, A.; Diaz, L.; Cuenca, N.; et al. CloudSEN12, a Global Dataset for Semantic Understanding of Cloud and Cloud Shadow in Sentinel-2. *Sci. Data* **2022**, *9*, 782. [CrossRef]

62. R Core Team. *R: A Language and Environment for Statistical Computing*; R Core Team: Vienna, Austria, 2010.

63. McGill, R.; Tukey, J.W.; Larsen, W.A. Variations of Box Plots. *Am. Stat.* **1978**, *32*, 12–16. [CrossRef]

64. Wickham, H.; Stryjewski, L. 40 Years of Boxplots. Technical Report. 2011. Available online: https://vita.had.co.nz/papers/boxplots.html (accessed on 9 January 2024).

65. Lenth, R. V Estimated Marginal Means, Aka Least-Squares Means [R Package Emmeans Version 1.6.0]. *Compr. R Arch. Netw. (CRAN)* **2021**, *34*, 216–221.

66. Searle, S.R.; Speed, F.M.; Milliken, G.A. Population Marginal Means in the Linear Model: An Alternative to Least Squares Means. *Am. Stat.* **1980**, *34*, 216–221. [CrossRef]

67. Castillo, E.B.; Cayo, E.Y.T.; Almeida, C.M.d.; López, R.S.; Briceño, N.B.R.; López, J.O.S.; Gurbillón, M.Á.B.; Oliva, M.; Espinoza-Villar, R. Monitoring Wildfires in the Northeastern Peruvian Amazon Using Landsat-8 and Sentinel-2 Imagery in the GEE Platform. *ISPRS Int. J. Geoinf.* **2020**, *9*, 564. [CrossRef]

68. Khanal, N.; Uddin, K.; Matin, M.A.; Tenneson, K. Automatic Detection of Spatiotemporal Urban Expansion Patterns by Fusing OSM and Landsat Data in Kathmandu. *Remote Sens.* **2019**, *11*, 2296. [CrossRef]

69. Amoakoh, A.O.; Aplin, P.; Awuah, K.T.; Delgado-Fernández, I.; Moses, C.; Alonso, C.P.P.; Kankam, S.; Mensah, J.C. Testing the Contribution of Multi-Source Remote Sensing Features for Random Forest Classification of the Greater Amanzule Tropical Peatland. *Sensors* **2021**, *21*, 3399. [CrossRef] [PubMed]

70. Gao, X.; Chi, H.; Huang, J.; Han, Y.; Li, Y.; Ling, F. Comparison of Cloud-Mask Algorithms and Machine-Learning Methods Using Sentinel-2 Imagery for Mapping Paddy Rice in Jianghan Plain. *Remote Sens.* **2024**, *16*, 1305. [CrossRef]

71. Watt, M.S.; Holdaway, A.; Watt, P.; Pearse, G.D.; Palmer, M.E.; Steer, B.S.C.; Camarretta, N.; McLay, E.; Fraser, S. Early Prediction of Regional Red Needle Cast Outbreaks Using Climatic Data Trends and Satellite-Derived Observations. *Remote Sens.* **2024**, *16*, 1401. [CrossRef]

72. Yu, P.; Zhou, T.; Luo, H.; Liu, X.; Shi, P.; Zhang, Y.; Zhang, J.; Zhou, P.; Xu, Y. Global Pattern of Ecosystem Respiration Tendencies and Its Implications on Terrestrial Carbon Sink Potential. *Earth's Future* **2022**, *10*, e2022EF002703. [CrossRef]

73. Li, S.; Lü, S.; Liu, Y.; Gao, Y.; Ao, Y. Variations and Trends of Terrestrial NPP and Its Relation to Climate Change in the 10 CMIP5 Models. *J. Earth Syst. Sci.* **2015**, *124*, 395–403. [CrossRef]

74. Haara, A.; Kangas, A.; Tuominen, S. Economic Losses Caused by Tree Species Proportions and Site Type Errors in Forest Management Planning. *Silva Fenn.* **2019**, *53*, 10089. [CrossRef]

75. Platnick, S.; King, M.D.; Ackerman, S.A.; Menzel, W.P.; Baum, B.A.; Riédi, J.; Frey, R. The MODIS Cloud Products: Algorithms and Examples From Terra. *IEEE Trans. Geosci. Remote Sens.* **2003**, *41*, 459–473. [CrossRef]

76. Madson, A.; Dimson, M.; Fortini, L.B.; Kawelo, K.; Ticktin, T.; Keir, M.; Dong, C.; Ma, Z.; Beilman, D.W.; Kay, K.; et al. A Near Four-Decade Time Series Shows the Hawaiian Islands Have Been Browning Since the 1980s. *Environ. Manag.* **2022**, *71*, 965–980. [CrossRef] [PubMed]

77. Barton, K.E.; Westerband, A.; Ostertag, R.; Stacy, E.; Winter, K.; Drake, D.R.; Fortini, L.B.; Litton, C.M.; Cordell, S.; Krushelnycky, P.; et al. Hawai'i Forest Review: Synthesizing the Ecology, Evolution, and Conservation of a Model System. *Perspect. Plant Ecol. Evol. Syst.* **2021**, *52*, 125631. [CrossRef]

78. Vaughn, N.R.; Hughes, R.F.; Asner, G.P. Multi-Scale Remote Sensing-Based Landscape Epidemiology of the Spread of Rapid 'Ōhi'a Death in Hawai'i. *For. Ecol. Manag.* **2023**, *538*, 120983. [CrossRef]

79. Espinoza-Mendoza, V.E. Factors Influencing the Estimation of Aboveground Biomass (AGB) in Tropical Forests Using RADAR Remote Sensing. *PeerJ Prepr.* **2018**, *6*, e26534v1. [CrossRef]

80. Caiza-Morales, L.; Gómez, C.; Torres, R.; Nicolau, A.P.; Olano, J.M. MANGLEE: A Tool for Mapping and Monitoring MANgrove Ecosystem on Google Earth Engine—A Case Study in Ecuador. *J. Geovis. Spat. Anal.* **2024**, *8*, 17. [CrossRef]

81. Slagter, B.; Reiche, J.; Marcos, D.; Mullissa, A.; Lossou, E.; Peña-Claros, M.; Herold, M. Monitoring Direct Drivers of Small-Scale Tropical Forest Disturbance in near Real-Time with Sentinel-1 and -2 Data. *Remote Sens. Environ.* **2023**, *295*, 113655. [CrossRef]

82. Yin, Z.; Ling, F.; Foody, G.M.; Li, X.; Du, Y. Cloud Detection in Landsat-8 Imagery in Google Earth Engine Based on a Deep Convolutional Neural Network. *Remote Sens. Lett.* **2020**, *11*, 1181–1190. [CrossRef]

83. Hamunyela, E.; Rosca, S.; Mirt, A.; Engle, E.; Herold, M.; Gieseke, F.; Verbesselt, J. Implementation of BFASTmonitor Algorithm on Google Earth Engine to Support Large-Area and Sub-Annual Change Monitoring Using Earth Observation Data. *Remote Sens.* **2020**, *12*, 2953. [CrossRef]

84. Martinez, J.A.C.; Adarme, M.X.O.; Turnes, J.N.; Costa, G.; De Almeida, C.A.; Feitosa, R.Q. A Comparison of Cloud Removal Methods for Deforestation Monitoring in Amazon Rainforest. *Int. Arch. Photogramm. Remote Sens. Spat. Inf. Sci.* **2022**, *43*, 665–671. [CrossRef]

85. Xu, K.; Guan, K.; Peng, J.; Luo, Y.; Wang, S. DeepMask: An Algorithm for Cloud and Cloud Shadow Detection in Optical Satellite Remote Sensing Images Using Deep Residual Network. *arXiv* **2019**, arXiv:1911.03607.

86. Twumasi, Y.A.; Asare-Ansah, A.B.; Merem, E.C.; Loh, P.M.; Namwamba, J.B.; Ning, Z.; Yeboah, H.B.; Anokye, M.; Armah, R.N.D.; Apraku, C.Y.; et al. Fusion of Landsat 8 OLI and PlanetScope Images for Urban Forest Management in Baton Rouge, Louisiana. *J. Geogr. Inf. Syst.* **2022**, *14*, 444–461. [CrossRef]

87. Sandonís-Pozo, L.; Llorens, J.; Escolà, A.; Arnó, J.; Pascual, M.; Martínez-Casasnovas, J.A. Satellite Multispectral Indices to Estimate Canopy Parameters and Within-Field Management Zones in Super-Intensive Almond Orchards. *Precis. Agric.* **2022**, *23*, 2040–2062. [CrossRef]

88. Warren, S.G.; Eastman, R.; Hahn, C.J. A Survey of Changes in Cloud Cover and Cloud Types Over Land From Surface Observations, 1971–1996. *J. Clim.* **2007**, *20*, 717–738. [CrossRef]

89. Barnes, M.L.; Miura, T.; Giambelluca, T.W. An Assessment of Diurnal and Seasonal Cloud Cover Changes Over the Hawaiian Islands Using Terra and Aqua MODIS*. *J. Clim.* **2015**, *29*, 77–90. [CrossRef]

90. Zhang, C.; Hamilton, K.; Wang, Y. Monitoring and Projecting Snow on Hawaii Island. *Earth's Future* **2017**, *5*, 436–448. [CrossRef]

91. Ray, D.K.; Manoharan, V.S.; Welch, R.M. Cloud Cover Conditions and Stability of the Western Ghats Montane Wet Forests. *J. Geophys. Res. Atmos.* **2011**, *116*, D12104. [CrossRef]

92. Graham, E.; Mulkey, S.S.; Kitajima, K.; Phillips, N.; Wright, S.J. Cloud Cover Limits Net CO_2 Uptake and Growth of a Rainforest Tree During Tropical Rainy Seasons. *Proc. Natl. Acad. Sci. USA* **2003**, *100*, 572–576. [CrossRef] [PubMed]

93. Wang, J.; Wang, Y.; Wang, W.; Shi, L.; Si, H. Transfer-Learning-Based Cloud Detection for Zhuhai-1 Satellite Hyperspectral Imagery. *Front. Environ. Sci.* **2022**, *10*, 1039249. [CrossRef]

94. Park, S. A Dynamic Relationship Between the Leaf Phenology and Rainfall Regimes of Hawaiian Tropical Ecosystems: A Remote Sensing Approach. *Singap. J. Trop. Geogr.* **2010**, *31*, 371–383. [CrossRef]

95. Kpienbaareh, D.; Sun, X.; Wang, J.; Luginaah, I.; Kerr, R.B.; Lupafya, E.; Dakishoni, L. Crop Type and Land Cover Mapping in Northern Malawi Using the Integration of Sentinel-1, Sentinel-2, and PlanetScope Satellite Data. *Remote Sens.* **2021**, *13*, 700. [CrossRef]

96. Hirschmugl, M.; Sobe, C.; Deutscher, J.; Schardt, M. Combined Use of Optical and Synthetic Aperture Radar Data for REDD+ Applications in Malawi. *Land* **2018**, *7*, 116. [CrossRef]

97. Li, L.; Liu, L.; Peng, Y.; Su, Y.; Hu, Y.; Zou, R. Integration of Multimodal Data for Large-Scale Rapid Agricultural Land Evaluation Using Machine Learning and Deep Learning Approaches. *Geoderma* **2023**, *439*, 116696. [CrossRef]

98. Li, J.; Hong, D.; Gao, L.; Yao, J.; Zheng, K.; Zhang, B.; Chanussot, J. Deep Learning in Multimodal Remote Sensing Data Fusion: A Comprehensive Review. *Int. J. Appl. Earth Obs. Geoinf.* **2022**, *112*, 102926. [CrossRef]