



Technical Note

Fast Fusion of Sentinel-2 and Sentinel-3 Time Series over Rangelands

Paul Senty^{1,*}, Radoslaw Guzinski¹, Kenneth Grogan¹, Robert Buitenwerf², Jonas Ardö³, Lars Eklundh³, Alkiviadis Koukos¹, Torbern Tagesson^{3,4} and Michael Munk¹

¹ DHI Water & Environment, 2970 Hørsholm, Denmark

² Center for Ecological Dynamics in a Novel Biosphere (ECONOVO), Department of Biology, Aarhus University, Ny Munkegade 114, 8000 Aarhus, Denmark

³ Department of Physical Geography and Ecosystem Science, Lund University, S-223 62 Lund, Sweden

⁴ Department of Geosciences and Natural Resource Management, University of Copenhagen, 1350 København, Denmark

* Correspondence: pase@dhigroup.com

Abstract: Monitoring ecosystems at regional or continental scales is paramount for biodiversity conservation, climate change mitigation, and sustainable land management. Effective monitoring requires satellite imagery with both high spatial resolution and high temporal resolution. However, there is currently no single, freely available data source that fulfills these needs. A seamless fusion of data from the Sentinel-3 and Sentinel-2 optical sensors could meet these monitoring requirements as Sentinel-2 observes at the required spatial resolution (10 m) while Sentinel-3 observes at the required temporal resolution (daily). We introduce the Efficient Fusion Algorithm across Spatio-Temporal scales (EFAST), which interpolates Sentinel-2 data into smooth time series (both spatially and temporally). This interpolation is informed by Sentinel-3's temporal profile such that the phenological changes occurring between two Sentinel-2 acquisitions at a 10 m resolution are assumed to mirror those observed at Sentinel-3's resolution. The EFAST consists of a weighted sum of Sentinel-2 images (weighted by a distance-to-clouds score) coupled with a phenological correction derived from Sentinel-3. We validate the capacity of our method to reconstruct the phenological profile at a 10 m resolution over one rangeland area and one irrigated cropland area. The EFAST outperforms classical interpolation techniques over both rangeland (−72% in the mean absolute error, MAE) and agricultural areas (−43% MAE); it presents a performance comparable to the Spatial and Temporal Adaptive Reflectance Fusion Model (STARFM) (+5% MAE in both test areas) while being 140 times faster. The computational efficiency of our approach and its temporal smoothing enable the creation of seamless and high-resolution phenology products on a regional to continental scale.

Keywords: data fusion; spatiotemporal fusion; Sentinel-2; Sentinel-3; interpolation; phenology; time series; STARFM; rangelands



Citation: Senty, P.; Guzinski, R.; Grogan, K.; Buitenwerf, R.; Ardö, J.; Eklundh, L.; Koukos, A.; Tagesson, T.; Munk, M. Fast Fusion of Sentinel-2 and Sentinel-3 Time Series over Rangelands. *Remote Sens.* **2024**, *16*, 1833. <https://doi.org/10.3390/rs16111833>

Academic Editors: Wei Yang and Pamela L. Nagler

Received: 27 February 2024

Revised: 16 April 2024

Accepted: 17 May 2024

Published: 21 May 2024



Copyright: © 2024 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

Monitoring changes in vegetation attributes and phenology is critical for understanding the impacts of climate change and human activities on ecosystems [1]. Remote sensing techniques have become essential tools for studying vegetation dynamics over large spatial extents [2]. In the most recent decade, large-scale and high-resolution vegetation datasets emerged, ranging from a global map of forest cover change at a 30 m resolution [3] to a sub-continental map of carbon stocks from individual trees in African drylands [4]. This rise in large-scale products has been facilitated by the availability of free and open-access satellite data. Among these, Sentinel-2 satellites, forming part of the European Union's Copernicus satellite constellation, are among the most frequently used Earth observation satellites for monitoring vegetation due to their high spatial and temporal resolutions [5].

However, in seasonally dry ecosystems such as the vast savanna rangelands of Africa, vegetation growth typically coincides with periods of frequent precipitation and therefore cloud cover, leading to prolonged periods without Sentinel-2 data (Figure A1). With long data gaps, traditional temporal interpolation methods can fail to capture key phenological information, e.g., the timing of the green-up stage, the vegetation peak, or the maximum value of the observed vegetation index, which are important for estimating an ecosystem's primary production or to understand interannual vegetation dynamics [6].

Sentinel-3, another Copernicus satellite constellation, acquires daily observations at the expense of a coarser spatial resolution (300 m). Because of its higher temporal resolution, Sentinel-3 is more likely to predict the aforementioned variables. Conversely, its coarser resolution can restrict the acquisition of adequate spatial detail. Various fusion algorithms, including unmixing [7–9] or machine learning [10–13] approaches, successfully fused fine and coarse images to generate synthetic data with a high resolution both temporally and spatially. The most widely used fusion algorithm is the Spatial and Temporal Adaptive Reflectance Fusion Model (STARFM) [14], which has been successfully applied to various ecosystems [15–17]. In the context of the fusion of Sentinel-2 and Sentinel-3 data, the STARFM corrects a cloud-free Sentinel-2 image based on Sentinel-3's temporal change to produce a synthetic high resolution for a time when no Sentinel-2 data are available.

The STARFM and other spatio-temporal fusion algorithms, such as the ESTARFM [18] or Fit-FC [19], rely on a spatial-averaging step that aggregates multiple carefully selected neighboring pixels to estimate the temporal change correction. This step reduces the impacts of geometry misalignment between Sentinel-2 and Sentinel-3 and tends to improve performance over heterogeneous areas. However, despite efforts to accelerate the STARFM [20], this spatial-averaging step remains computationally expensive, which hampers the feasibility of fusing Sentinel-2 and Sentinel-3 data over large scales (regional or continental) [21]. Additionally, the choice of Sentinel-2 data used as input has a high impact on the prediction [22]. When interpolating long time series, different Sentinel-2 images must be used as inputs for different sections of the time series. This leads to abrupt changes in the fused time series at the transition between one input Sentinel-2 image and another. Likewise, if partially clouded images are used (which would be necessary for a large-scale analysis), spatial discontinuities appear along the cloud mask of the Sentinel-2 input.

Our aim is to create a fusion algorithm that is scalable, outperforms single-source interpolation methods, and is suitable for large-scale analyses (i.e., continental) while also mitigating spatial or temporal discontinuities associated with cloud cover. To achieve this, we introduce a method, the Efficient Fusion Algorithm across Spatio-Temporal scales (EFAST), which replaces the spatial-averaging step of the STARFM with temporal averaging. Removing the spatial-averaging step makes the predictions much faster at the expense of reducing the quality of the prediction at the boundary between two land cover types and in heterogeneous areas. Additionally, temporal averaging makes the predictions more resilient to atmospheric effects and leads to smooth time series. The temporal average is a weighted sum modulated by a distance-to-cloud score, which has two advantages: it assigns a higher importance to completely cloud-free images (less likely to contain remnant clouds or cloud shadows) and also leads to smooth transitions in the resulting fusion products around the cloud mask.

Fusing Sentinel-2 and Sentinel-3 over African savannas spanning more than 2000 Sentinel-2 tiles and long time periods would be resource-intensive as most spatiotemporal fusion methods have focused on accuracy, especially with respect to resolving sub-pixel features [18], at the expense of computational efficiency [21]. Moreover, they sometimes require the manual selection of input images [16]. Our highly scalable and fully automated methodology aims to streamline the production of high-resolution phenological products at a continental scale.

2. Materials and Methods

2.1. Study Area

This paper focuses on two 16 km² areas, both located in the Senegalese Louga region (Figure 1). Though our method is designed to predict high-resolution phenological products at a continental scale for all African rangelands, focusing our analysis on those two areas allows us to better visualize the performance of our approach over diverse, small-scale features.

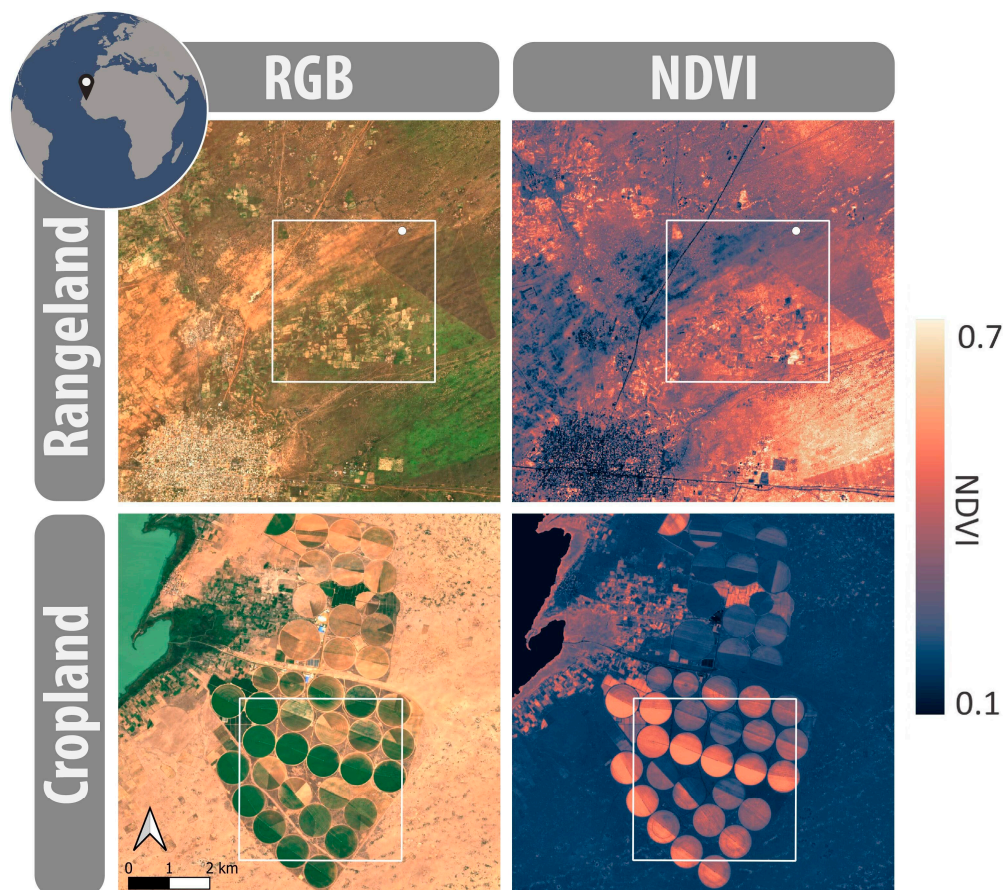


Figure 1. The test areas are shown in white squares. The rangeland true-color image and NDVI data (top) were acquired by Sentinel-2 on 6 October 2021, and the cropland images (bottom) date from 13 December 2022. The cropland area is surrounded by grasslands along Lake Guiers. The rangeland area is situated a few kilometers northeast of the town of Dahra. The Dahra field site (represented by a small white dot on the top-right corner of the rangeland area) includes a hemispherical NDVI sensor that we use to evaluate the reliability of our method.

The rangeland area in the Southern Sahel region, with distinct dry and wet seasons, presents strong vegetation seasonality driven by precipitation. The wet season (lasting from July to September) is characterized by prevalent cloud cover, leading to long time periods without Sentinel-2 observations. Yet these temporal gaps are less pronounced than in more humid tropical areas (Figure A1). This area also contains the Dahra field site, which is equipped with a multispectral sensor measuring the normalized difference vegetation index (NDVI) [23]. This in situ sensor provides an accurate data source for validating the temporal interpolation of the proposed fusion method.

The second area consists of multiple irrigated parcels surrounded by natural herbaceous vegetation. The parcels are harvested once or twice a year, depending on the parcel and the year. The timing of vegetation growth varies from one parcel to another. This heterogeneous area allows us to assess the fusion's ability to interpolate different vegetation dynamics for which the coarse resolution of Sentinel-3 would be restricting.

2.2. Sentinel-2 Processing

We downloaded four years of Sentinel-2 L2A (bottom-of-atmosphere) products covering tile 28PDC (relative orbits 37 and 80) from January 2019 to December 2022. Sentinel-2's normalized difference vegetation index (NDVI) was derived as follows:

$$S_2 = \frac{B08 - B04}{B08 + B04}$$

where B08 is the Sentinel-2 spectral band centered on 842 nm, and B04 is the spectral band centered on 665 nm. Clouds and cloud shadows were masked using the scene classification (SCL) map.

2.3. Sentinel-3 Processing

All Sentinel-3 SYN L2A products (SY_2_SYN) acquired between 2019 and 2022 and overlapping with tile 28PDC were also used in this study. Each was reprojected on to a 300 m resolution grid (EPSG: 32628) and clipped to the extent of tile 28PDC. The following Land Quality Science Flag (LQSF) layers, distributed as part of the SY_2 SYN__ product, were used as a cloud mask: LQSF.CLOUD, LQSF.CLOUD_AMBIGUOUS, and LQSF.CLOUD_MARGIN. Finally, the NDVI was derived from surface directional reflectance (SDR) bands:

$$S_3 = \frac{Oa17 - Oa08}{Oa17 + Oa08}$$

where Oa17 is the Sentinel-3 spectral band centered on 865 nm and Oa08 is the spectral band centered on 665 nm.

Some Sentinel-3 acquisitions contain unflagged clouds and cloud shadows characterized by underestimations of the NDVI; these are particularly apparent around September in Figure 2. These atmospheric effects are reduced and the data are interpolated into a smooth and continuous time series using a moving average.

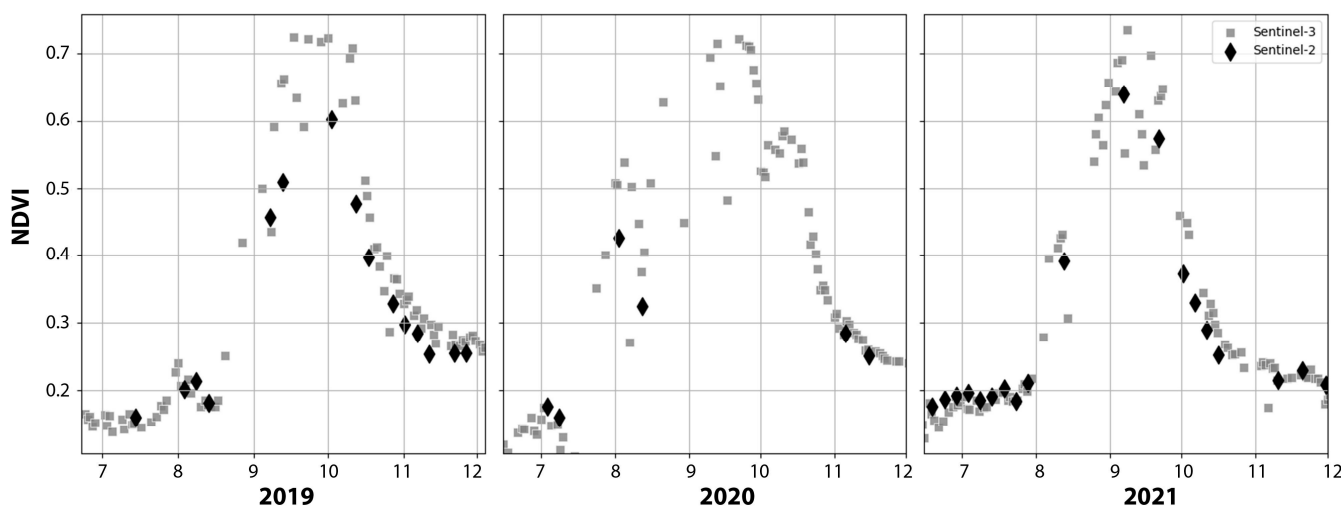


Figure 2. Example showing Sentinel-2 and Sentinel-3 NDVI data around the wet seasons (months 7–12) of 2019, 2020, and 2021 in Senegal. Atmospheric effects lead to underestimations of Sentinel-2 and Sentinel-3 data which are especially apparent around September in 2020 and 2021. The timing of vegetation growth varies from July to August. Higher cloud cover during the wet season leads to fewer acquisitions, with an especially long time without Sentinel-2 data in 2020.

2.4. Fusion Principle

Let us assume that the Sentinel-3 NDVI values are aggregations of Sentinel-2 NDVI values (linear mixing model [24]):

$$S_3(t, p) = \sum_{x \in C(p)} w_p(x) \times S_2(t, x) \quad (1)$$

where $S_3(t, p)$ is the Sentinel-3 NDVI value at time t for a Sentinel-3 coarse pixel p , $C(p)$ is the set of Sentinel-2 pixels belonging to the Sentinel-3 resolution cell corresponding to pixel p , and $S_2(t, x)$ is the Sentinel-2 NDVI value at the same time t . $w_p(x)$ is the contribution of the fine pixel x to the Sentinel-3 resolution cell, and the sum of these partial contributions is equal to 1.

If the temporal change is spatially homogeneous over Sentinel-3's resolution cell $C(p)$ between two timesteps t^* and t , then the value of the temporal change $S_2(t, x) - S_2(t^*, x)$ is the same for every 10 m pixel x belonging to that resolution cell. In particular, it is equal to the value at the central pixel, which coincides with the position of Sentinel-3's pixel center p . We can derive from Equation (1) and the previous statement the following equation:

$$S_2(t, p) - S_2(t^*, p) = S_3(t, p) - S_3(t^*, p)$$

which translates to the following: if the temporal change is homogeneous locally, then the temporal change at a 10 m scale corresponds to the temporal change measured by Sentinel-3. This equation is generalized to every pixel by applying a resampling of Sentinel-3 data down to Sentinel-2's resolution:

$$S_2(t, x) - S_2(t^*, x) = S_3(t, x) - S_3(t^*, x),$$

Or, as represented in Figure 3,

$$S_2(t) = S_2(t^*) + \underbrace{[S_3(t) - S_3(t^*)]}_{\text{temporal change correction}} \quad (2)$$

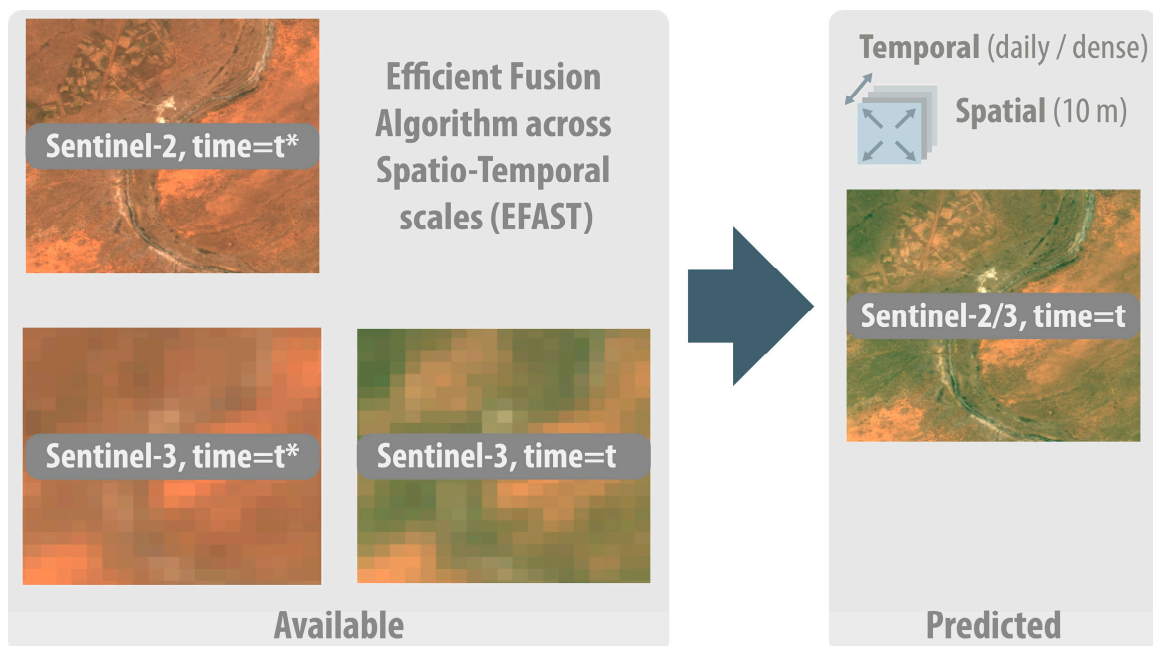


Figure 3. The fusion principle—the EFAST creates synthetic high-resolution data through a simple transformation of Sentinel-2 and Sentinel-3 images as follows: $S_2(t^*) + S_3(t) - S_3(t^*)$.

Looking at the right-hand side of the equation, the reference image $S_2(t^*)$ is the high-resolution component, while $S_3(t) - S_3(t^*)$ is a coarse-scale correction accounting for the temporal change measured by Sentinel-3. This equation is the principle behind our approach which, in theory, only applies in areas presenting relatively uniform temporal profiles; the temporal change at Sentinel-3's 300 m resolution should closely reflect the temporal change at a 10 m resolution.

2.5. Spatial and Temporal Smoothing

A limitation of the previous approach is its reliance on one single reference Sentinel-2 image, $S_2(t^*)$. If we interpolate the NDVI profile between two cloud-free acquisitions at times t_1^* and t_2^* , using the closest cloud-free Sentinel-2 image as a reference, we can use $S_2(t_1^*)$ as a reference for all timesteps before the transition time $\frac{1}{2}(t_1^* + t_2^*)$, which is halfway between t_1^* and t_2^* , and $S_2(t_2^*)$ after the transition, resulting in a discontinuity in the interpolated time series. Our approach produces smooth time series by considering a weighted sum of all the corrected Sentinel-2 images instead of relying on a single reference image:

$$S_2(t) = \sum_{t^*} w_t(t^*) \times [S_2(t^*) + S_3(t) - S_3(t^*)] \quad (3)$$

where $w_t(t^*)$ represents normalized weights which depend on two components: the distance-to-clouds score of the Sentinel-2 image $S_2(t^*)$ and the time delta between t and t^* . The normalized scores are derived as follows:

$$w_t(t^*) \propto \min\left(\frac{d_{t^*}}{D}, 1\right) \times \exp\left[-\frac{(t - t^*)^2}{2s^2}\right]$$

where d_{t^*} is the distance between each pixel and the nearest cloud [25] in the Sentinel-2 image $S_2(t^*)$, D is the size of the transition region, and s is the smoothing parameter (Figure A2). For this paper, we set s and D to 20 days and 5 km, respectively.

2.6. Validation Strategy

We assess the performance of the EFAST over the two areas presented in Figure 1, as well as at the position of the Dahra field site. Our approach was compared to two other methods:

- The Whittaker filter, which smooths and interpolates time series while being resilient to missing data, making it a commonly used method in remote sensing [26,27]. The Whittaker filter also contains a smoothing parameter that we set to $400 \text{ days}^2 = (20 \text{ days})^2$, which appears consistent with the EFAST smoothing parameter $s = 20$ days. This method only uses Sentinel-2 data, so the comparison of the EFAST and the Whittaker filter aims to demonstrate the value of adding Sentinel-3 to the equation.
- The STARFM, a spatio-temporal fusion algorithm [14], with the following parameters: four classes and a window size of 31 pixels. We use Mileva's 2018 open-source implementation in Python [28] to compare its speed with our approach in the same environment. We use the single-pair version of the STARFM and choose the closest cloud-free Sentinel-2 image as input data. A comparison of the performance of our method with STARFM allows us to verify whether the increase in the computational efficiency of the EFAST over the STARFM translates into a reduction in performance and to quantify this reduction.

We compare the three interpolation methods (the EFAST, STARFM, and Whittaker filter) in two experiments:

1. A comparison using in situ data at the Dahra field site (experiment 1). The interpolated time series are produced using all Sentinel-2 observations that do not contain clouds within a radius of 1 km from the site (to avoid undetected clouds and cloud shadows). For the STARFM and EFAST, we also use the entire smoothed Sentinel-3 time series. The Sentinel-2 input data and the predictions are displayed at the position of the Dahra field site (as the average value over a 3-by-3-pixel box to account for misalignment

between the Sentinel-2 resolution cell and the multispectral sensor). We compare these time series to in situ data obtained over four years from 2019 until the end of the year 2022.

- Across the two study areas highlighted in Figure 1 (experiment 2), to assess performance on a larger scale and at a high resolution, we use the Sentinel-2 data itself for validation. We keep the Sentinel-2 images acquired in July, August, or September for validation (Figure 4), leading to temporal gaps of three months. Discarding three months' worth of data emulates plausible conditions in these semi-arid ecosystems (Figure A1). To avoid contaminating the errors with clouds and cloud shadows, we only consider Sentinel-2 images that are cloud-free over the extent of the study area. The absolute difference between the Sentinel-2 images kept for validation (12 images for the rangeland area and 17 for the cropland area) and the corresponding predictions are aggregated and displayed as error maps.

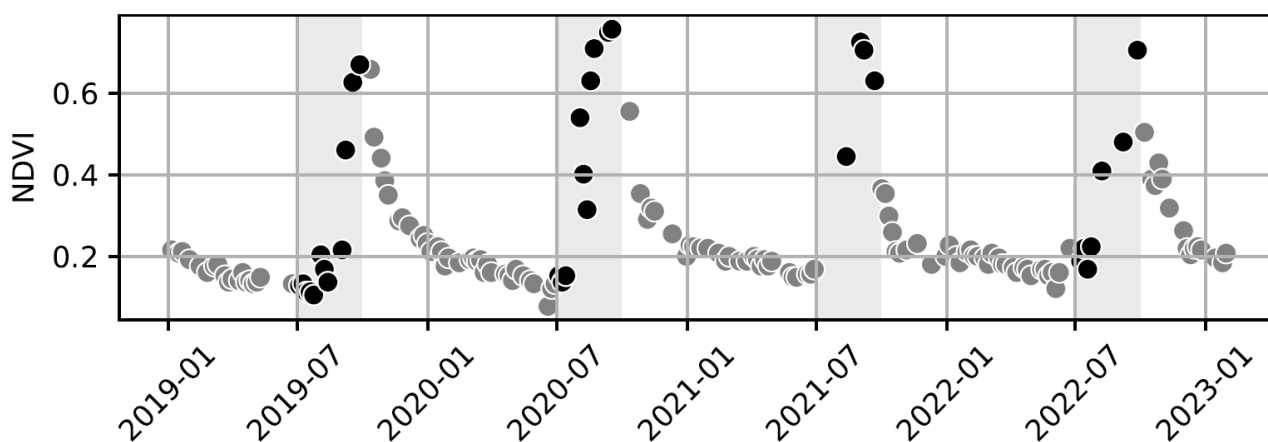


Figure 4. Validation was performed on all cloud-free Sentinel-2 images acquired during the wet season (black points); the rest of the Sentinel-2 cloud-free acquisitions (gray points) and all the Sentinel-3 observations were used for interpolation.

Additionally, the computation times of the three methods are compared. All tests are implemented in Python and tested on a desktop computer with Windows 10 OS, an Intel i7-10850H CPU (2.7 GHz, 6 Cores), and 16 GB RAM.

3. Results

3.1. Field Site Evaluation

Both the STARFM and EFAST are able to capture full phenological cycles over the 4 years of data (Figure 5). The Whittaker filter is able to interpolate the wet season of 2020 despite the lack of observation around the vegetation peak in September but fails to capture the maximal value of the NDVI in 2022. The in situ data present a positive bias relative to Sentinel-2 data, but the timings of vegetation growth and senescence are consistent.

The relatively dense Sentinel-2 time series over this semi-arid region in Senegal limits the need for fusion; a simple interpolation method (such as the Whittaker filter) would usually capture most of the phenological cycle relying on Sentinel-2 data only, as it did for the years 2019, 2020, and 2021 (Figure 5). The next subsection documents the ability of the Whittaker filter, STARFM, and EFAST to reconstruct the phenological profile when no data are available for multiple months, which is common in tropical savannas (Figure A1).

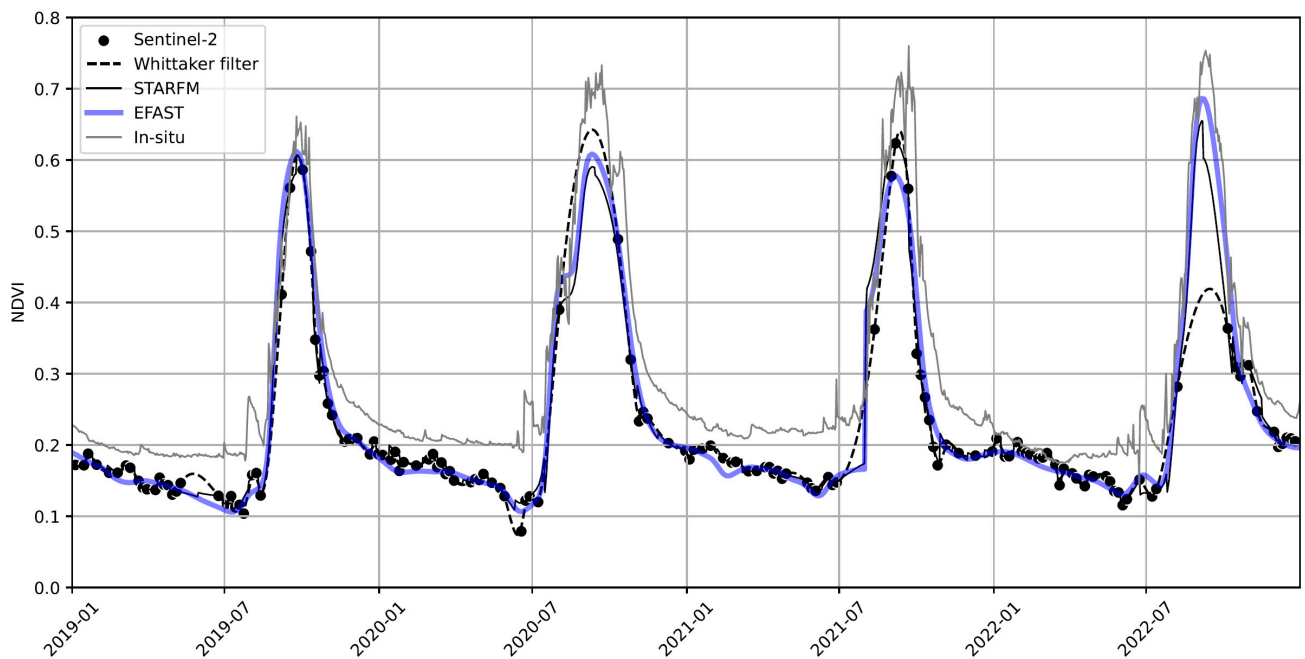
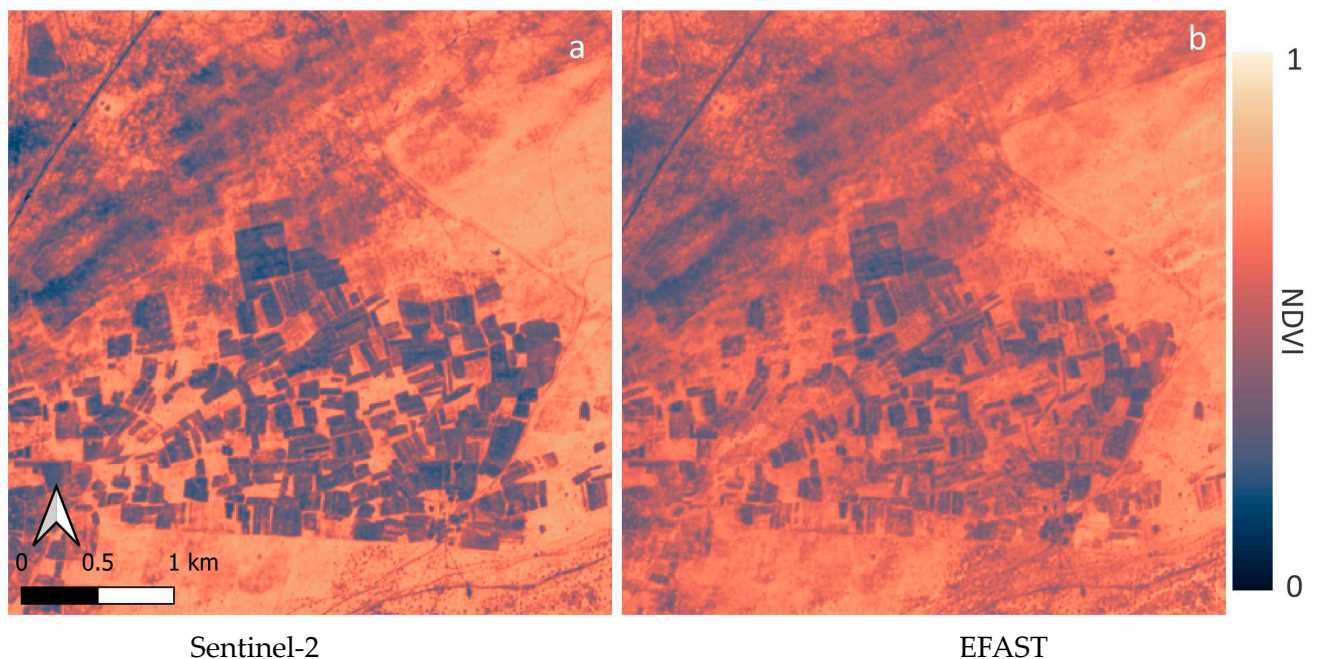


Figure 5. STARFM, EFAST, and Whittaker filter compared to in situ data at Dahra field site.

3.2. Reconstruction of the Wet Season

3.2.1. Rangeland Area

The EFAST accurately identifies regions of high and low values, but the resulting image does not present the same contrast as the Sentinel-2 image (Figure 6a); the NDVI values of bare parcels are overestimated and, conversely, the vegetation index over the surrounding grasslands is underestimated.



Sentinel-2

EFAST

Figure 6. Example of a Sentinel-2 image used for validation, dating from 17 September 2019 (a), and the corresponding prediction by the EFAST (b). The absolute difference between these two images is one of the 12 terms (one for each validation image) of the mean absolute error map (Figure 7c).

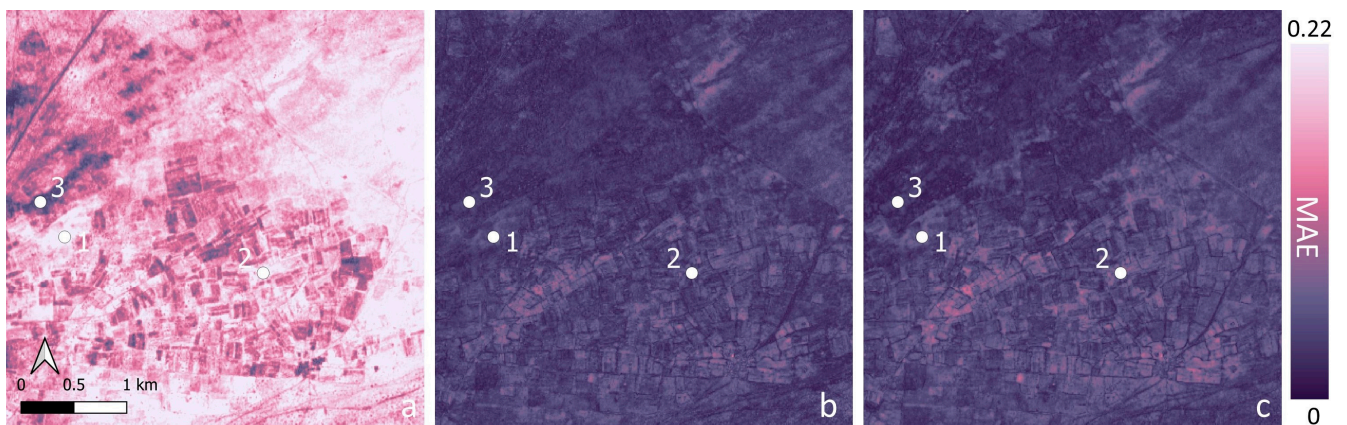


Figure 7. Mean absolute error maps of the reconstructed NDVI profiles using the Whittaker filter (a), STARFM (b), and EFAST (c). The depicted dots represent specific points for which the corresponding time series are illustrated in Figure 8.

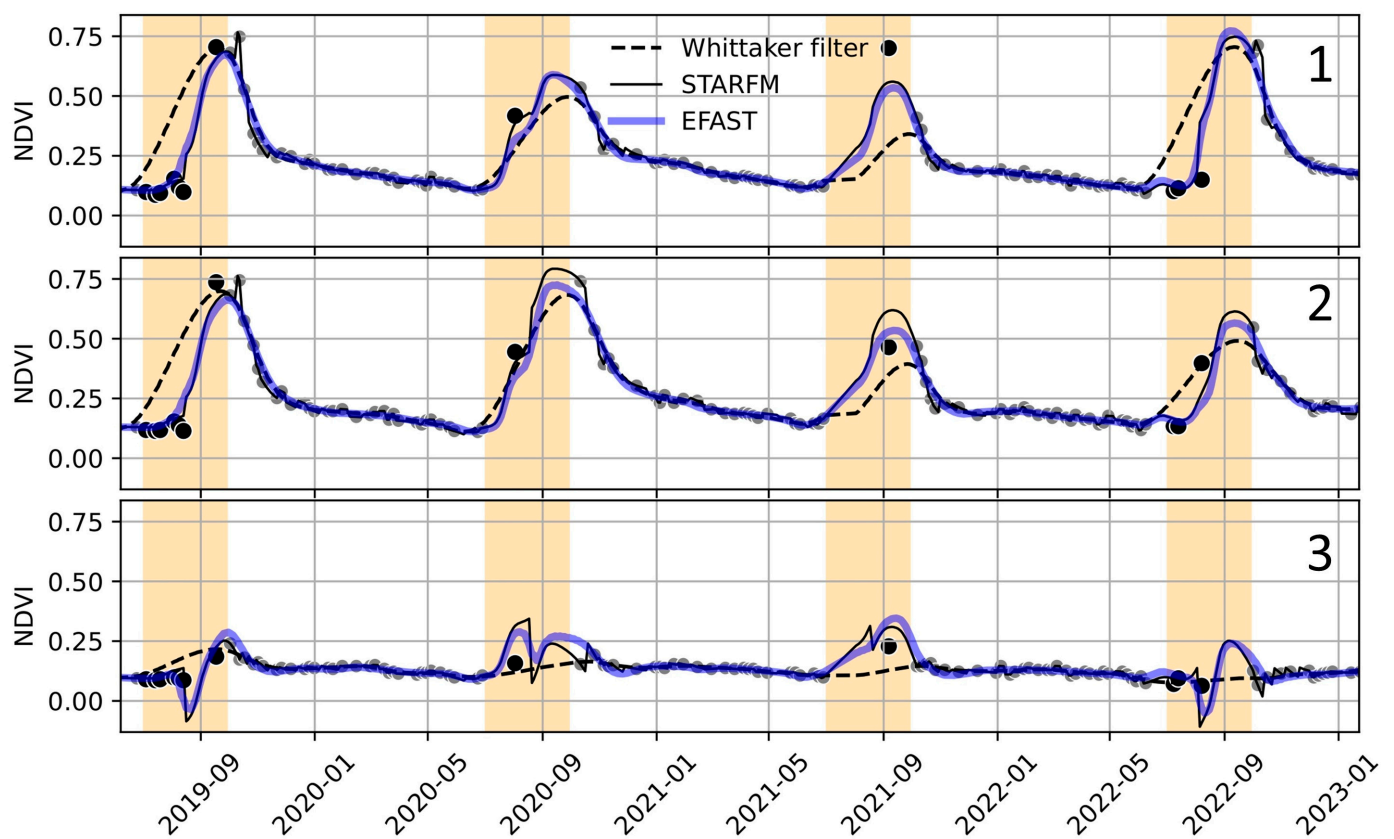


Figure 8. The predicted time series of the three interpolation methods (the Whittaker filter, STARFM, and EFAST) for the three points of the rangeland area (Figure 7). Black points represent Sentinel-2 validation points, and gray points represent the Sentinel-2 data used for interpolation. Orange areas correspond to the time frames in which the reconstruction of the NDVI profile is assessed.

In the area surrounding the Dahra flux site, which is mainly composed of grasslands and a few trees (~3% canopy cover [23]), both the STARFM and EFAST considerably outperform the Whittaker filter (Figure 7). The Whittaker filter tends to underestimate both the peak of vegetation and the timing of vegetation growth (Figure 8), while both the STARFM and the EFAST successfully reconstruct the NDVI values' phenology. The first time series in Figure 8 present a considerable underestimation of the third peak compared to the validation for all three methods. Overall, the STARFM performs marginally better

than the EFAST, with a less than 5% decrease in the mean absolute error (Table 1). The more noticeable difference between the EFAST and STARFM time series (Figure 8) is the presence of temporal discontinuities for the STARFM, while the temporal averaging of the EFAST leads to smoother transitions.

Table 1. The mean absolute errors of the reconstructed NDVI profiles using the Whittaker filter, STARFM, and EFAST. These correspond to the average values of the mean absolute maps (Figure 7).

Area	Whittaker	STARFM	EFAST
Rangeland	0.172	0.042	0.044
Cropland	0.075	0.040	0.042

It is apparent from Figure 8(3) that the EFAST and STARFM overestimate the phenological variations in point number 3 (Figure 7), sometimes leading to negative NDVI values (2019 and 2022). The Sentinel-3 signal over this bare area is corrupted by the surrounding grasslands. Conversely, the Whittaker filter underestimates its variations, failing to predict the small increase in September 2021.

3.2.2. Cropland Area

In the more heterogeneous landscape of the cropland area, the EFAST and STARFM also outperform the Whittaker filter for most crop parcels, and, similarly to the rangeland area, the EFAST and STARFM present similar performances overall (Table 1).

The STARFM significantly outperforms the EFAST over the grass between the crop parcels (e.g., point 5 in Figure 9). This is probably the result of a combination of two factors:

- The spatial averaging of the STARFM makes use of the lower part of the study area, where the Sentinel-3 pixels are more homogenous and mainly composed of grass.
- The temporal averaging of the EFAST gives a lower weight to individual cloud-free pixels, leading to the corruption of the phenological signal, even in periods of low cloud cover. This is particularly apparent in Figure 10(5), where the vegetation growth of the irrigated croplands around December in 2020, 2021, and 2022 leads to a predicted bi-seasonality of the grass.

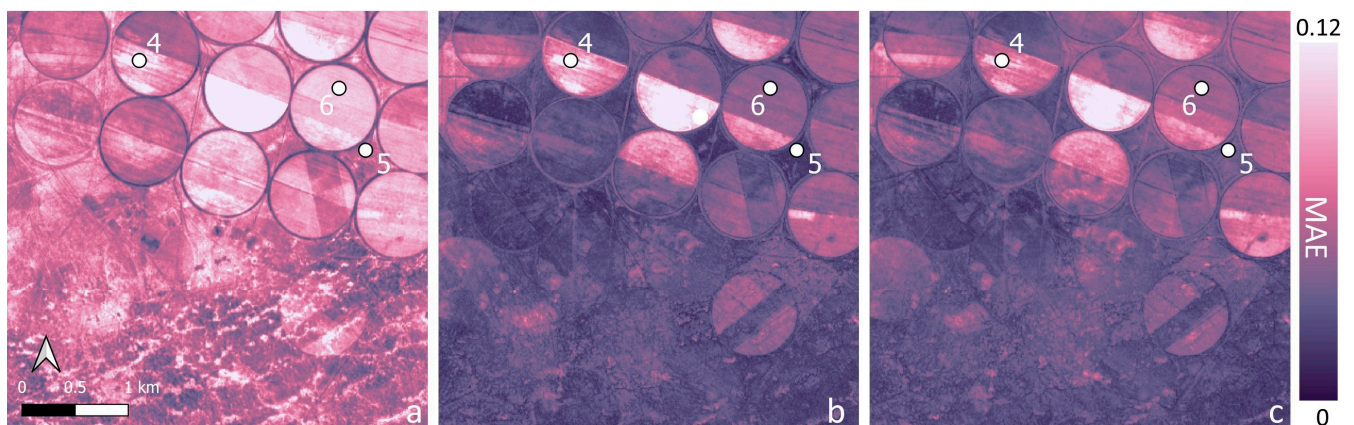


Figure 9. The mean absolute error using the Whittaker filter (a), STARFM (b) and EFAST (c). The dots correspond to the points for which time series are displayed in Figure 10.

The STARFM is also slightly better at predicting large crop parcels (point 6 of Figure 9), though the difference is barely noticeable in the time series (Figure 10(6)).

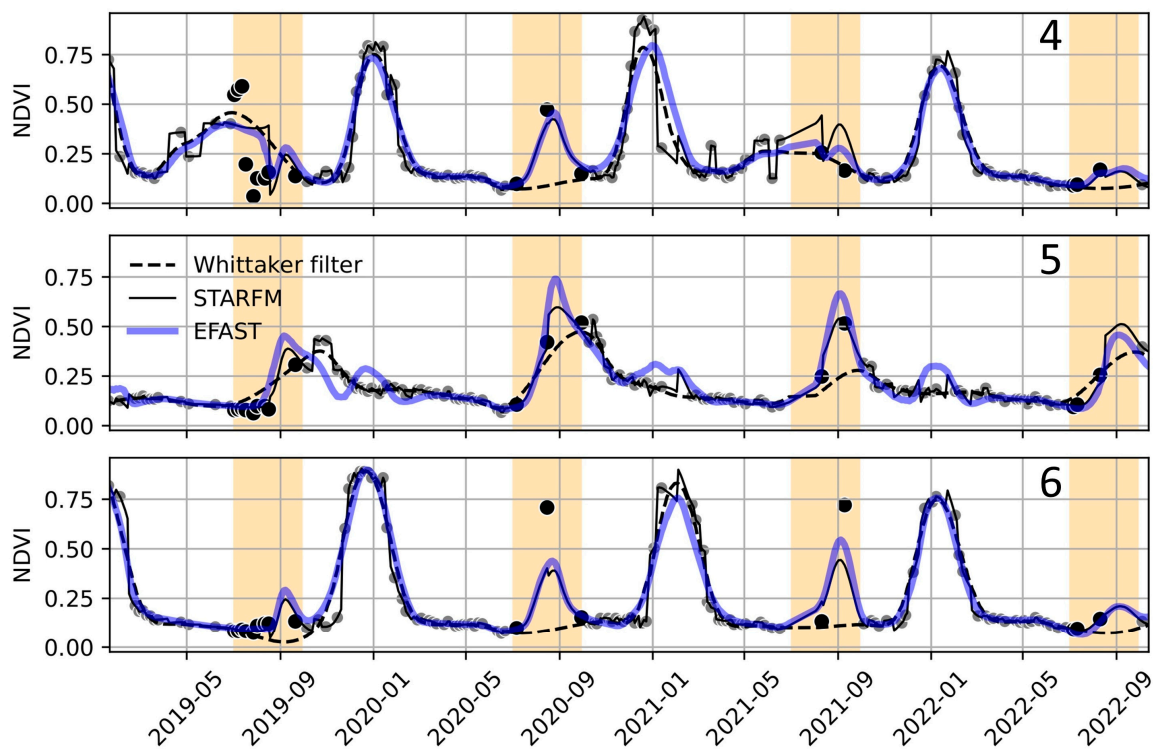


Figure 10. The predicted time series of the three interpolation methods (the Whittaker filter, STARFM, and EFAST) for three points in the cropland area (Figure 9). Black points represent Sentinel-2 validation points, and gray points represent the Sentinel-2 data used for interpolation. Orange areas correspond to time frames in which the reconstruction of the NDVI profile is assessed.

3.3. Computation Time

The EFAST is about 140 times faster than the STARFM (Table 2). As the EFAST treats fine pixels independently, one can think of the EFAST as a version of the STARFM with a window size of one single pixel instead of 31 in the STARFM. Thus, removing the spatial-averaging step decreases the number operation by $31^2 \approx 1000$. But the EFAST averages multiple Sentinel-2 images, therefore increasing the number of operations by one order of magnitude (around 10 Sentinel-2 images are typically used in the weighted average when $s = 20$ days). These considerations are consistent with the measured decrease in computation time of two orders of magnitude ($1000/10$) between the STARFM and EFAST.

Table 2. Computation time, in seconds, taken by the Whittaker filter, STARFM, and EFAST to produce a single 400 pixel by 400 pixel NDVI prediction.

Whittaker	STARFM	EFAST
0.1 *	85	0.6

* with a constant weight matrix [26].

4. Discussion

4.1. Efficiency over Large Scales

The EFAST demonstrates substantial potential for large-scale applications, which can primarily be attributed to its significantly enhanced processing speed compared to the STARFM achieved by removing spatial complexity. While acknowledging a marginal increase in error compared to spatiotemporal fusion methods designed to resolve heterogeneous landscapes, the EFAST's efficiency makes it particularly suitable for extensive analyses. The independence of pixels within the EFAST allows for efficient parallelization and seamless integration into cloud computing environments, enhancing scalability for large-scale assessments.

The very limited number of parameters (only the smoothing parameter s has an important impact on results) contributes to the method's applicability for large-scale analyses without the need for extensive parameter tweaking. Additionally, the EFAST's automated image weighting mechanism eliminates the necessity of the manual selection of individual Sentinel-2 input images as the weights inherently prioritize closer images with lower cloud cover. These attributes collectively position the EFAST as a pragmatic and efficient tool for large-scale remote sensing endeavors, emphasizing its suitability for such applications without indulging in landscape-specific complexities.

4.2. Consequences for Rangeland Monitoring

Harmonized and uninterrupted time series on vegetation phenology are crucial to ecosystem monitoring. In seasonally dry ecosystems such as African rangelands, phenological changes can proceed rapidly, especially at the start of the growing season when herbaceous vegetation responds strongly to the first rains of the wet season [29]. A high observation frequency is therefore key, but this is complicated by cloudy conditions during the growing season which can result in gaps in NDVI time series that result in inaccurate estimates of the start of the growing season. Simultaneously, African savannas and rangelands vary in vegetation composition and structure, driven by local environmental gradients such as catenal sequences and the localized impacts of herbivory and fire [30]. This implies that capturing ecosystem dynamics requires high-spatiotemporal-resolution data.

By enabling uninterrupted time series at a high resolution, the EFAST is therefore expected to have significant potential for improving ecosystem monitoring. Since any ecosystem monitoring system relies on fast and efficient implementation, we expect that the EFAST will be able to provide baseline data for various monitoring platforms based on spectral methods. Further investigation across diverse ecoregions and climatic zones is ongoing within the framework of the ESA-funded Rangeland Monitoring for Africa Using Earth Observation (RAMONA, <https://app.ramona.earth/>, accessed on 16 May 2024), aiming to assess the utility of the EFAST's fusion method for different types of rangelands (over 2000 Sentinel-2 tiles will be produced).

4.3. Limitations over Heterogeneous Areas

Over homogeneous areas, all Sentinel-2 pixels present a common phenology locally. The Sentinel-3 and Sentinel-2 values show a similar temporal profile. In this case, Sentinel-3 adds valuable information to the Sentinel-2 time series.

However, in highly heterogeneous areas, Sentinel-2 and Sentinel-3 time series can present different patterns. For example, in the second study area, the Sentinel-3 signal over the grass between the agricultural parcels (Figure 9, dot 5) is corrupted by the surrounding parcels. The EFAST's informed interpolation is therefore badly informed, and the homogeneity hypothesis (described in Section 2.4) leads to the supposal of a crop-like temporal change over this grass. Under these conditions, the EFAST can perform worse than traditional interpolation methods.

By calculating the Pearson correlation coefficient between the Sentinel-2 and Sentinel-3 time series (Figure 11), we can determine regions in which the homogeneity hypothesis is likely to hold true. This correlation coefficient offers insights into whether data fusion with Sentinel-3 would potentially worsen or enhance the interpolation. We are currently extending our approach to automatically determining which Sentinel-2 pixels should be corrected by Sentinel-3's temporal change and which ones should not.

Alternatively, more sophisticated spatio-temporal fusion methods that derive conversion coefficients, such as the ESTARFM [18], Fit-FC [19], or NDVI-LMGM [31], are more suitable for heterogeneous landscapes. Indeed, the conversion coefficient identifies regions associated with low or high variability, thus reducing the overestimation of the change in bare areas or grasslands during the dry season. These algorithms are recommended for use over heterogeneous landscapes, provided that computational constraints are not a limiting factor.

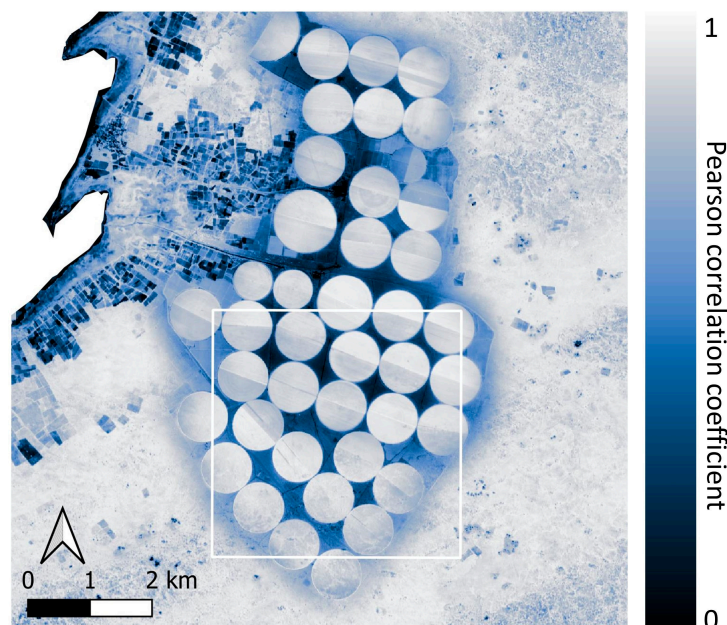


Figure 11. Pearson correlation coefficient between Sentinel-2 and Sentinel-3 time series. Small-scale features stand out as having a low correlation because of Sentinel-3's limited spatial resolution. Conversely, large crop parcels and homogeneous grasslands present a high correlation. The white box corresponds to the area in Figure 9.

4.4. Land-Cover Change

A transition in land cover, such as a change caused by fire, flood, or deforestation, induces abrupt changes in vegetation indices. Given its emphasis on generating smooth time series data, the EFAST is not well-suited to characterizing disturbances. Indeed, the utilization of multiple reference images in EFAST results in the blending of the Sentinel-2 images before and after the disturbance, leading to an overestimation of change before the event and an underestimation during the event. Algorithms like STAARCH [32] are more appropriate for analyzing land-cover changes and for applications in ecosystems prone to sudden disturbances. Indeed, STAARCH, for each date, determines the best reference image, either the last available Sentinel-2 image before the disturbance or the next available Sentinel-2 image after the disturbance, depending on the prediction time's relation to the disturbance event.

4.5. Smoothing Parameters

The temporal smoothing parameter s controls temporal smoothing. A low value of s (below 5 days) is equivalent to using only the closest Sentinel-2 data, leading to abrupt transitions in the time series. A higher value of s leads to a smoother time series and reduces the impact of atmospheric effects on individual Sentinel-2 images. However, the smoothing can often reduce the quality of the predictions; indeed, the closest Sentinel-2 acquisition is often the best reference image. The optimal value of s depends on the temporal density of the data and thus on the geographical area. [33] propose a method to automatically deduce the best smoothing parameter for the Whittaker filter. A similar paradigm would make the EFAST more flexible.

The distance-to-clouds scores allow for a smooth transition around the cloud mask while giving a higher weight to observations that do not contain any surrounding clouds. The optimal value of parameter D depends on the quality of the cloud mask as well as the requirements to obtain spatially smooth predictions. A too-high D value would lead one to discard good data from partially clouded images. $D = 5$ km appears suitable for most areas when using the Scene Classification SCL cloud mask (evaluated in Denmark and various areas across Africa).

4.6. Sentinel-3's Temporal Profile

In our second experiment, we emulated higher cloud cover during the wet season by discarding some Sentinel-2 data without affecting the Sentinel-3 time series. We supposed that the coarse-scale temporal profile would still be fully captured by Sentinel-3. This might not hold true in places with very high cloud cover, such as tropical rainforests. Further investigation is needed to assess the applicability of our method in such regions.

5. Conclusions

We introduced the EFAST, the Efficient Fusion Algorithm across Spatial–Temporal scales, which interpolates remote sensing data into smooth time series and is able to reconstruct extended periods without Sentinel-2 data using coarser but more frequent observations from Sentinel-3. We compared its performance and speed to the STARFM and a single-source interpolation method, using 29 validation images over two test areas.

The comparison with single-source interpolation methods underscored the critical role of fusion in accurately capturing the phenological cycle in sub-tropical rangelands. Despite the STARFM showing a slightly lower error (a ~5% difference in the mean absolute error) than the EFAST in both test areas, the EFAST demonstrates a significant advantage in computational efficiency, being more than 100 times faster.

This considerable improvement in computational efficiency and the automated nature of our approach (through the use and scoring of all Sentinel-2 images) allows for its application over large scales both in time and space. Additionally, the pixel-based framework of the EFAST simplifies its integration into cloud computing platforms.

However, the EFAST faces challenges in accurately resolving heterogeneous environments and is not well-suited for abrupt land-cover changes. In cases in which computing time is not a constraint, it is advisable to consider more refined and computationally intensive fusion algorithms that incorporate conversion coefficients. The correlation between phenological profiles at fine and coarse scales helps identify areas in which the EFAST may perform poorly. Further research is needed to pinpoint where fusion is most needed and effective and where the coarser resolution of Sentinel-3 may compromise the Sentinel-2 time series.

With the increasing quantity and availability of satellite data, there is a need for the development and improvement of scalable fusion methods that strike the right balance between accuracy and computational efficiency.

Author Contributions: Conceptualization, P.S., R.G. and M.M.; methodology, P.S., R.G. and K.G.; validation, P.S., J.A., L.E., R.B. and T.T.; formal analysis and investigation, P.S., R.G., K.G. and A.K.; writing—original draft preparation, P.S.; writing—review and editing, R.G., K.G., R.B., J.A., L.E., T.T. and M.M.; visualization, P.S. and M.M.; supervision, M.M., R.G. and K.G.; project administration and funding acquisition, M.M. All authors have read and agreed to the published version of the manuscript.

Funding: This research was conducted in the context of the project RAMONA, Rangeland Monitoring for Africa, funded by the European Space Agency (ESA) under the EO Science for Society, grant number 4000136180/21/I-NB, and partly funded by DHI's research contract with the Danish Ministry of Higher Education and Science. T. Tagesson was additionally funded by the Swedish National Space Agency (SNSA 2021–00144 and 2021–00111) and FORMAS (Dnr. 2021–00644). R. Buitenwerf also considers this work to be a contribution to the Center for Ecological Dynamics in a Novel Biosphere (ECONOVO), funded by Danish National Research Foundation (grant DNRF173).

Data Availability Statement: Publicly available datasets were analyzed in this study (Sentinel-2 and Sentinel-3 level 2 data). These data can be found at the following link: <https://browser.dataspace.copernicus.eu> (accessed on 16 May 2024). The in situ data from Senegal are available upon request from T.T. Information about the resulting products will be communicated through <https://www.ramona.earth> (accessed on 16 May 2024).

Acknowledgments: This work is part of an ESA project, RAMONA, collaboration between DHI, Lund and Aarhus University to produce monthly biomass estimates at 10 m resolution and continental scale and special thanks to partners involved from RAMONA consortium and ESA personnel, particularly Patrick Griffiths for useful feedback and discussions. Special thanks also to Daniel Druce and Rasmus Meyer for supporting analysis (Figure A1).

Conflicts of Interest: Authors Paul Senty, Radoslaw Guzinski, Kenneth Grogan, Alkiviadis Koukos and Michael Munk were employed by the company DHI Water & Environment. The remaining authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Appendix A

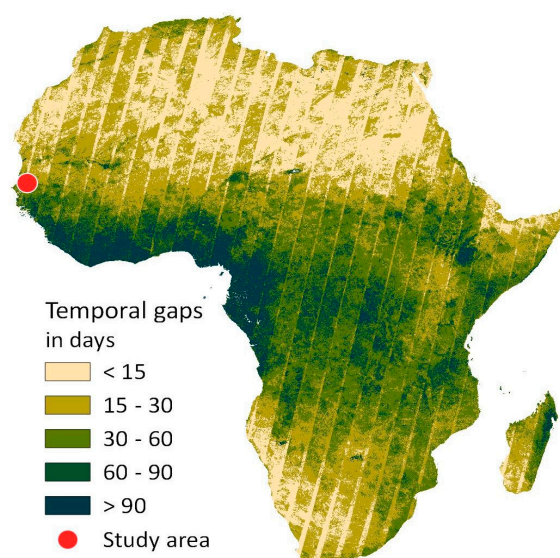


Figure A1. Maximum time, in days, without Sentinel-2 data over the African continent between August 2021 and January 2023. Extracted using Google Earth Engine. The brighter stripes correspond to areas of overlap between two orbits.

Appendix B

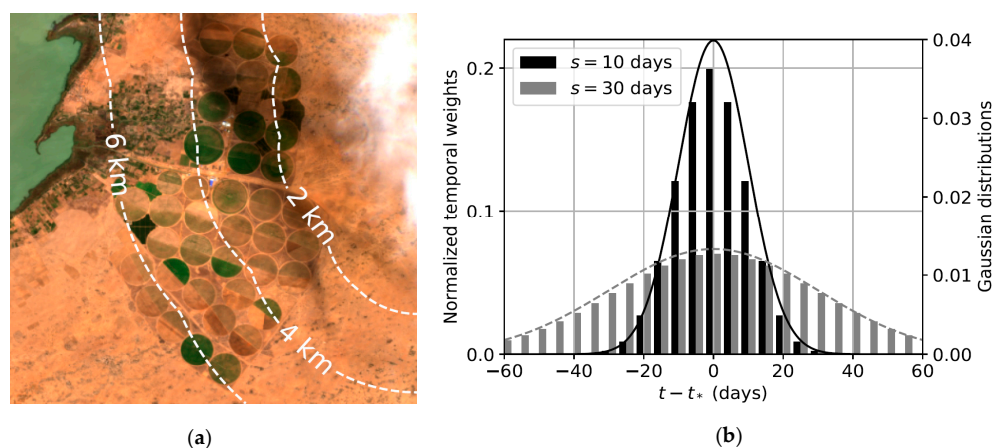


Figure A2. Smoothing parameters s and D . (a) Distance to closest masked cloud in km for the parameter. Distance-to-clouds score is equal to 0.5 two kilometers from the cloud mask and to 1 from $D = 4$ km. (b) Impact of temporal smoothing parameter s (in days) on temporal weights $\exp\left[-\frac{(t-t_*)^2}{2s^2}\right]$, displayed as bars, when there is one cloud-free Sentinel-2 acquisition every five days. Lines represent Gaussian distributions for $s = 10$ and 30 days.

References

1. Skidmore, A.K.; Coops, N.C.; Neinavaz, E.; Ali, A.; Schaepman, M.E.; Paganini, M.; Kissling, W.D.; Vihervaara, P.; Darvishzadeh, R.; Feilhauer, H.; et al. Priority list of biodiversity metrics to observe from space. *Nat. Ecol. Evol.* **2021**, *5*, 896–906. [[CrossRef](#)] [[PubMed](#)]
2. Senf, C. Seeing the System from Above: The Use and Potential of Remote Sensing for Studying Ecosystem Dynamics. *Ecosystems* **2022**, *25*, 1719–1737. [[CrossRef](#)]
3. Hansen, M.C.; Potapov, P.V.; Moore, R.; Hancher, M.; Turubanova, S.A.; Tyukavina, A.; Thau, D.; Stehman, S.V.; Goetz, S.J.; Loveland, T.R.; et al. High-Resolution Global Maps of 21st-Century Forest Cover Change. *Science* **2013**, *342*, 850–853. [[CrossRef](#)] [[PubMed](#)]
4. Tucker, C.; Brandt, M.; Hiernaux, P.; Kariryaa, A.; Rasmussen, K.; Small, J.; Igel, C.; Reiner, F.; Melocik, K.; Meyer, J.; et al. Sub-continental-scale carbon stocks of individual trees in African drylands. *Nature* **2023**, *615*, 80–86. [[CrossRef](#)] [[PubMed](#)]
5. Misra, G.; Cawkwell, F.; Wingler, A. Status of Phenological Research Using Sentinel-2 Data: A Review. *Remote Sens.* **2020**, *12*, 2760. [[CrossRef](#)]
6. Cleland, E.E.; Chuine, I.; Menzel, A.; Mooney, H.A.; Schwartz, M.D. Shifting plant phenology in response to global change. *Trends Ecol. Evol.* **2007**, *22*, 357–365. [[CrossRef](#)] [[PubMed](#)]
7. Zurita-Milla, R.; Gomez-Chova, L.; Guanter, L.; Clevers, J.G.P.W.; Camps-Valls, G. Multitemporal Unmixing of Medium-Spatial-Resolution Satellite Images: A Case Study Using MERIS Images for Land-Cover Mapping. *IEEE Trans. Geosci. Remote Sens.* **2011**, *49*, 4308–4317. [[CrossRef](#)]
8. Amorós-López, J.; Gómez-Chova, L.; Alonso, L.; Guanter, L.; Zurita-Milla, R.; Moreno, J.; Camps-Valls, G. Multitemporal fusion of Landsat/TM and ENVISAT/MERIS for crop monitoring. *Int. J. Appl. Earth Obs. Geoinf.* **2013**, *23*, 132–141. [[CrossRef](#)]
9. Gevaert, C.M.; García-Haro, F.J. A comparison of STARFM and an unmixing-based algorithm for Landsat and MODIS data fusion. *Remote Sens. Environ.* **2015**, *156*, 34–44. [[CrossRef](#)]
10. Liu, X.; Deng, C.; Wang, S.; Huang, G.-B.; Zhao, B.; Lauren, P. Fast and Accurate Spatiotemporal Fusion Based Upon Extreme Learning Machine. *IEEE Geosci. Remote Sens. Lett.* **2016**, *13*, 2039–2043. [[CrossRef](#)]
11. Song, H.; Liu, Q.; Wang, G.; Hang, R.; Huang, B. Spatiotemporal Satellite Image Fusion Using Deep Convolutional Neural Networks. *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.* **2018**, *11*, 821–829. [[CrossRef](#)]
12. Cai, J.; Huang, B.; Fung, T. Progressive spatiotemporal image fusion with deep neural networks. *Int. J. Appl. Earth Obs. Geoinf.* **2022**, *108*, 102745. [[CrossRef](#)]
13. Wang, Z.; Ma, Y.; Zhang, Y. Review of pixel-level remote sensing image fusion based on deep learning. *Inf. Fusion* **2023**, *90*, 36–58. [[CrossRef](#)]
14. Gao, F.; Masek, J.; Schwaller, M.; Hall, F. On the blending of the Landsat and MODIS surface reflectance: Predicting daily Landsat surface reflectance. *IEEE Trans. Geosci. Remote Sens.* **2006**, *44*, 2207–2218. [[CrossRef](#)]
15. Jia, K.; Liang, S.; Zhang, L.; Wei, X.; Yao, Y.; Xie, X. Forest cover classification using Landsat ETM+ data and time series MODIS NDVI data. *Int. J. Appl. Earth Obs. Geoinf.* **2014**, *33*, 32–38. [[CrossRef](#)]
16. Gao, F.; Anderson, M.C.; Zhang, X.; Yang, Z.; Alfieri, J.G.; Kustas, W.P.; Mueller, R.; Johnson, D.M.; Prueger, J.H. Toward mapping crop progress at field scales through fusion of Landsat and MODIS imagery. *Remote Sens. Environ.* **2017**, *188*, 9–25. [[CrossRef](#)]
17. Olexa, E.M.; Lawrence, R.L. Performance and effects of land cover type on synthetic surface reflectance data and NDVI estimates for assessment and monitoring of semi-arid rangeland. *Int. J. Appl. Earth Obs. Geoinf.* **2014**, *30*, 30–41. [[CrossRef](#)]
18. Zhu, X.; Chen, J.; Gao, F.; Chen, X.; Masek, J.G. An enhanced spatial and temporal adaptive reflectance fusion model for complex heterogeneous regions. *Remote Sens. Environ.* **2010**, *114*, 2610–2623. [[CrossRef](#)]
19. Wang, Q.; Atkinson, P.M. Spatio-temporal fusion for daily Sentinel-2 images. *Remote Sens. Environ.* **2018**, *204*, 31–42. [[CrossRef](#)]
20. Guan, Q.; Peng, X. High-performance Spatio-temporal Fusion Models for Remote Sensing Images with Graphics Processing Units. *AGU Fall Meet. Abstr.* **2018**, *2018*, IN41D-0866.
21. Gao, H.; Zhu, X.; Guan, Q.; Yang, X.; Yao, Y.; Zeng, W.; Peng, X. cuFSDAF: An Enhanced Flexible Spatiotemporal Data Fusion Algorithm Parallelized Using Graphics Processing Units. *IEEE Trans. Geosci. Remote Sens.* **2022**, *60*, 4403016. [[CrossRef](#)]
22. Xie, D.; Gao, F.; Sun, L.; Anderson, M. Improving Spatial-Temporal Data Fusion by Choosing Optimal Input Image Pairs. *Remote Sens.* **2018**, *10*, 1142. [[CrossRef](#)]
23. Tagesson, T.; Fensholt, R.; Guiro, I.; Rasmussen, M.O.; Huber, S.; Mbow, C.; Garcia, M.; Horion, S.; Sandholt, I.; Holm-Rasmussen, B.; et al. Ecosystem properties of semiarid savanna grassland in West Africa and its relationship with environmental variability. *Glob. Change Biol.* **2015**, *21*, 250–264. [[CrossRef](#)] [[PubMed](#)]
24. Zhu, X.; Cai, F.; Tian, J.; Williams, T.K.-A. Spatiotemporal Fusion of Multisource Remote Sensing Data: Literature Survey, Taxonomy, Principles, Applications, and Future Directions. *Remote Sens.* **2018**, *10*, 527. [[CrossRef](#)]
25. Griffiths, P.; van der Linden, S.; Kuemmerle, T.; Hostert, P. A Pixel-Based Landsat Compositing Algorithm for Large Area Land Cover Mapping. *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.* **2013**, *6*, 2088–2101. [[CrossRef](#)]
26. Eilers, P.H.C.; Pesendorfer, V.; Bonifacio, R. Automatic smoothing of remote sensing data. In Proceedings of the 2017 9th International Workshop on the Analysis of Multitemporal Remote Sensing Images (MultiTemp), Brugge, Belgium, 27–29 June 2017; pp. 1–3.
27. Atkinson, P.M.; Jeganathan, C.; Dash, J.; Atzberger, C. Inter-comparison of four models for smoothing satellite sensor time-series data to estimate vegetation phenology. *Remote Sens. Environ.* **2012**, *123*, 400–417. [[CrossRef](#)]

28. Mileva, N.; Mecklenburg, S.; Gascon, F. New tool for spatio-temporal image fusion in remote sensing: A case study approach using Sentinel-2 and Sentinel-3 data. In Proceedings of the Image and Signal Processing for Remote Sensing XXIV, Berlin, Germany, 10–13 September 2018; p. 20.
29. Higgins, S.I.; Delgado-Cartay, M.D.; February, E.C.; Combrink, H.J. Is there a temporal niche separation in the leaf phenology of savanna trees and grasses? *J. Biogeogr.* **2011**, *38*, 2165–2175. [[CrossRef](#)]
30. Scholes, R.J.; Walker, B.H. *An African Savanna: Synthesis of the Nylsvley Study*; Cambridge Studies in Applied Ecology and Resource Management; Cambridge University Press: Cambridge, UK, 1993; ISBN 978-0-521-61210-4.
31. Rao, Y.; Zhu, X.; Chen, J.; Wang, J. An Improved Method for Producing High Spatial-Resolution NDVI Time Series Datasets with Multi-Temporal MODIS NDVI Data and Landsat TM/ETM+ Images. *Remote Sens.* **2015**, *7*, 7865–7891. [[CrossRef](#)]
32. Hilker, T.; Wulder, M.A.; Coops, N.C.; Linke, J.; McDermid, G.; Masek, J.G.; Gao, F.; White, J.C. A new data fusion model for high spatial- and temporal-resolution mapping of forest disturbance based on Landsat and MODIS. *Remote Sens. Environ.* **2009**, *113*, 1613–1627. [[CrossRef](#)]
33. Frasso, G.; Eilers, P.H. L- and V-Curves for Optimal Smoothing. *Stat. Model.* **2015**, *15*, 91–111. [[CrossRef](#)]

Disclaimer/Publisher’s Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.