

Article

# A Near Real-Time Method for Forest Change Detection Based on a Structural Time Series Model and the Kalman Filter

Martin Puhm <sup>\*</sup>, Janik Deutscher, Manuela Hirschmugl , Andreas Wimmer, Ursula Schmitt and Mathias Schardt

Joanneum Research, Institute for Information and Communication Technologies, Steyrergasse 17, 8010 Graz, Austria; janik.deutscher@joanneum.at (J.D.); manuela.hirschmugl@joanneum.at (M.H.); andreas.wimmer@joanneum.at (A.W.); ursula.schmitt@joanneum.at (U.S.); mathias.schardt@joanneum.at (M.S.)

<sup>\*</sup> Correspondence: martin.puhm@joanneum.at; Tel.: +43-316-876-5107

Received: 16 July 2020; Accepted: 22 September 2020; Published: 24 September 2020



**Abstract:** The increasing availability of dense time series of earth observation data has incited a growing interest in time series analysis for vegetation monitoring and change detection. Vegetation monitoring algorithms need to deal with several time series characteristics such as seasonality, irregular sampling intervals, and signal artefacts. While common algorithms based on deterministic harmonic regression models account for intra-annual seasonality, inter-annual variations of the seasonal pattern related to shifts in vegetation phenology due to different temperature and rainfall are usually not accounted for. We propose a transition to stochastic modelling and present a near real-time change detection method that combines a structural time series model with the Kalman filter. The model continuously adapts to new observations and allows to better separate phenology-related deviations from vegetation anomalies or land cover changes. The method is tested in a forest change detection application aiming at the assessment of damages caused by storm events and insect calamities. Forest changes are detected based on the cumulative sum control chart (CUSUM) which is used to decide if new observations deviate from model-based forecasts. The performance is evaluated in two test sites, one in Malawi (dry tropical forest) and one in Austria (temperate deciduous, coniferous and mixed forests) based on Sentinel-2 time series. Both forest areas are characterized by a distinct, but temporally varying leaf-off season. The presented change detection method shows overall accuracies above 99%, users' accuracies of 76.8% to 88.6%, and producers' accuracies of 68.2% to 80.4% for the forest change stratum (minimum mapping unit: 0.1 ha). Results are based on visually interpreted points derived by stratified random sampling. A further analysis revealed that increasing the time series density by merging data from two Sentinel-2 orbits yields better forest change detection accuracies in comparison to using data from one orbit only. The resulting increase in users' accuracy amounts to 7.6%. The presented method is capable of near real-time processing and could be used for a variety of automated forest monitoring applications.

**Keywords:** state space models; forest disturbance mapping; near real-time monitoring; Sentinel-2; CUSUM

## 1. Introduction

Current Earth observation (EO) missions employing optical sensors such as Sentinel-2 acquire a vast volume of data: a new image every 5 days of almost every place on earth. By taking orbit overlaps into account, the time between consecutive images of the same region is reduced even further and the chance of acquiring cloud-free observations is further increased. Through high-quality georeferencing and atmospheric correction of the satellite images, it is possible to create consistent time series of

measured reflectance values for any given spectral band. The vast availability of high-resolution optical data allows—for the first time—to also map small changes in near real-time. Here, “small” may apply to both spatial extent as well as spectral change magnitude. However, dense time series of high-resolution optical data have a number of characteristics that pose a challenge to change detection applications. In addition to noise effects remaining after atmospheric correction and uncertainties in the geometric registration, these challenging characteristics include:

- **Seasonality:** Recurring seasonal patterns can be attributed to plant phenology and/or varying illumination conditions due to topography and solar angle. Considering that annual variations of temperature and rainfall also affect phenology, the seasonal patterns can vary between years.
- **Irregular sampling interval:** Satellites with a regular nadir acquisition scheme usually have a constant revisit cycle, for example, 5 days for the Sentinel-2A and B constellation. If images from overlapping orbits are integrated, though, the sampling interval becomes irregular. Data gaps due to masked clouds, cloud shadows, and snow also add to this irregularity.
- **Presence of signal artefacts:** Despite using state-of-the-art screening algorithms like Fmask [1], un-masked clouds, cloud shadows, and snow-covered areas remain in the pre-processed imagery. Corresponding observations have to be treated as invalid since the measured reflectance values neither represent the undisturbed land cover state nor a persistent change of it.

Algorithms can be divided based on how they deal with the time series characteristics described above. The different approaches used to handle these characteristics strongly affect the algorithms' suitability for monitoring changes in near real-time. Approaches that do not account for seasonality form a first group of algorithms. A review of change detection studies using Landsat time series concludes that many older studies focused on mapping changes only at annual or biannual time scales based on series of cloud-free composite images, which always represent the same season [2]. In this context, both the Vegetation Change Tracker (VCT) [3] as well as the LandTrendr approach [4] represent widely used algorithms, but they are not designed for near real-time mapping.

A second group of algorithms explicitly accounts for seasonality by using regression models based on trigonometric functions to capture the intra-annual variations (variations within one year, i.e., seasonality) of the spectral signatures independently for each pixel. With this approach, also frequently referred to as harmonic regression, periods of stable land cover are modelled as a deterministic, continuous function of time. Irregular sampling intervals and data gaps are therefore not a problematic issue, but the deterministic nature of the model does not allow inter-annual variations of the seasonal pattern (variations between different years, e.g., shifts in seasonality). The seasonal model represents an average of different conditions occurring within a stable period, e.g., dry and wet years, late and early leave outbreak. A widely used algorithm belonging to this group is the Breaks for Additive Season and Trend (BFAST) algorithm [5] and its evolution BFAST Monitor [6]. While the latter is tailored to near real-time mapping of new changes, the original version is intended for the analysis of historic time series. Both versions have been used in a variety of studies and can be applied to detect both abrupt and gradual changes. Concerning the robustness of BFAST to invalid observations, it has been stated that occasional signal artefacts are well handled, but temporally aggregated occurrences such as several consecutively un-masked clouds can be a source of error [7]. Also, additional pre-processing to eliminate artefacts was applied [8–10]. The second widely used implementation of the harmonic regression approach is the Continuous Change Detection and Classification algorithm (CCDC), where the original concept [11] is extended to include more types of land cover besides forest, as well as a classification framework [12]. From the beginning, CCDC was designed to work with dense Landsat time series and can handle seasonality, irregularly spaced observations, and signal artefacts to some extent. Both abrupt and gradual changes can be detected. Some further updates to the algorithm include (i) a mechanism to automatically adjust the complexity of the time series model based on the number of available clear observations, as well as (ii) a different method to estimate the model parameters which reduces overfitting [13]. A third algorithm employing harmonic regression

utilizes residuals from the regression together with statistical quality control charts [14]. This approach comprises signal artefact detection with Shewhart X-bar charts. After the elimination of artefacts, both abrupt and gradual changes are indicated by exponentially weighted moving average (EWMA) charts in a near real-time manner. All of the described algorithms of the second group share certain basic concepts, but the individual implementations vary. They are designed to process large amounts of data in a highly automated way and therefore rely on data-driven statistical boundaries for detecting change, although the distinct nature and computation of these boundaries is quite different.

The third group of algorithms stands out by also taking inter-annual variations of the seasonal pattern into account. Structural time series models are set up in terms of components, such as trends and cycles, which have a direct interpretation. Their statistical treatment is based on the state space form and the Kalman filter and first described for time series analysis in econometrics [15]. Compared to the harmonic regression approach, the model is no longer deterministic. Kalman filtering denotes a versatile parameter estimation technique which yields optimal estimates in a statistical sense [16]. It is well established in many application fields and has been applied to numerous signal tracking problems [17]. The combination of structural time series models with the Kalman filter and the concept's suitability for remote sensing purposes has been investigated in a proof of concept study, where Landsat time series are used to detect storm damages in a small forest test site in Germany [18]. The Kalman filter has also been applied to time series of MODIS 8-day composites in order to detect insect-induced defoliation in near real-time at a forest test area in northern Sweden [19]. This study makes use of the CUSUM control chart [20] to indicate changes, but it does not combine structural time series models with the Kalman filter. Instead, the filter is used to derive a smoothed time series of the Normalized Difference Vegetation Index (NDVI) based on a global model trajectory.

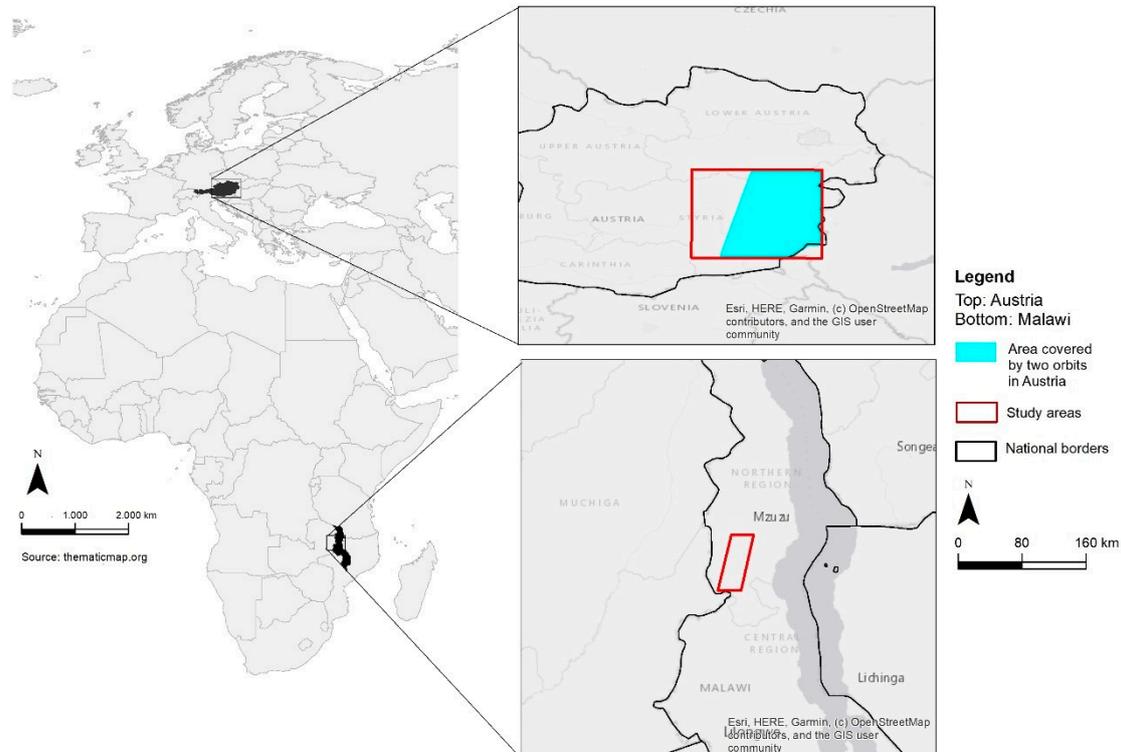
With the advantages and limitations of existing algorithms in mind, this work combines the pixel-by-pixel modelling typical for existing harmonic regression algorithms with the Kalman filter's capability to dynamically adjust the model based on new observations. The main aim of this study is to present an innovative change detection algorithm for optical EO data which is based on a structural time series model and the Kalman filter. It is largely data-driven and designed especially for near real-time mapping in web- or cloud-based monitoring services. The algorithm presented in this paper accounts for seasonality and also allows inter-annual variations of the seasonal pattern, e.g., vegetation phenology. Furthermore, strategies for handling irregular sampling intervals and signal artefacts are presented. The algorithm is tested in a forest change detection application using time series of Sentinel-2 data (S-2). Forest disturbances are detected at two complex forest test sites in Austria and Malawi. The first test site is an alpine area in Austria characterized by frequent cloud cover, snow cover, strong topographic effects, and pronounced forest seasonality and phenology. The second test site is located in the dry tropical forests of Malawi, where forests show a strong and varying seasonality between dry and rainy seasons. In the Malawi test site, we also analyze and compare accuracy results for two different data scenarios: first, using S-2 images from only one orbit, and second, using all available S-2 imagery from two orbits. The aim of this analysis is to investigate if different viewing angle and inconsistent geo-location of the pixels resulting from the combination of two orbits decrease the overall change detection accuracies despite the boosted time series density.

## 2. Materials and Methods

### 2.1. Test Sites, Data, and Pre-Processing

For the forest change detection demonstration, we selected two test sites, one in Austria and one in Malawi. The location of the two sites is shown in Figure 1. The Austrian test site is located in the south-eastern part of the country. The test site is characterized by strong topography in the northern part (Alpine area), where coniferous forests dominate. The southern part of the test site is located in the foothills of the Alps with moderate topography and the forests are predominantly mixed forests composed of coniferous and deciduous trees. The annual temperature amplitude in the Alpine

area is stronger than in the foothills; however, the Alpine coniferous forests have less pronounced phenological dynamics in comparison to the mixed and deciduous forests that dominate in the foothills. The eastern half of the study area is covered by two orbits, and the western half only by one orbit (see Figure 1). This is a typical data scenario encountered in practical applications.

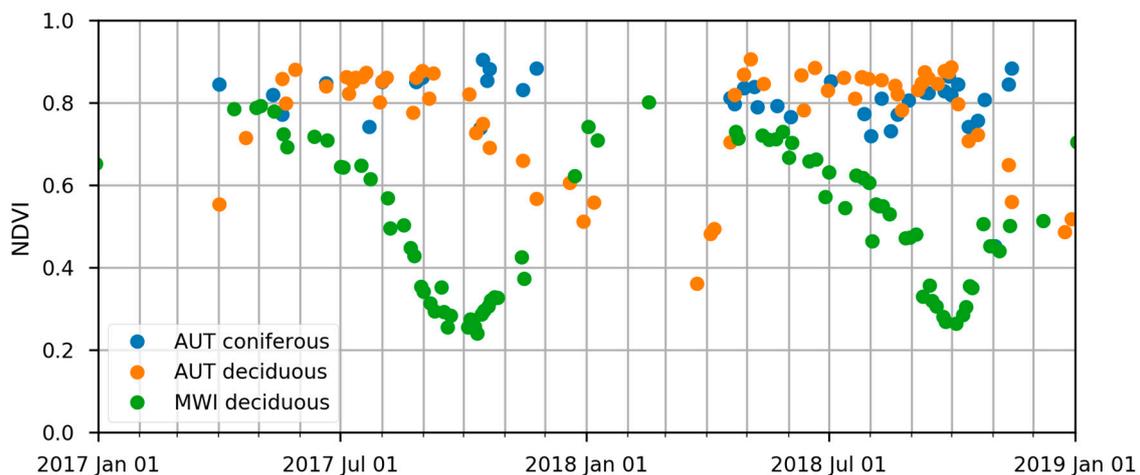


**Figure 1.** Location of the test sites.

The Malawi test site is characterized by flat to slightly hilly topography. The land cover is very heterogeneous and is subject to a precipitation gradient from east to west. As a result, the area is characterized by a vegetation phenology gradient from east to west. The forested areas differ in tree-cover density and tree-type composition and therefore, show very different spectral behavior. Dry tropical forests and the surrounding land use classes are more difficult to classify and monitor than humid evergreen forests, as they show a typical phenological development from highly vital in the rainy season to dry and leafless in the dry season [21]. Understory and grassland fires beneath the forest canopy can further complicate forest classification. Two S-2 orbits (relative orbits 92 and 135) cover the Malawi study area. The size of the test site was clipped to the overlap area of two orbits to investigate the effect of separately processing data from one and two orbits.

Both test sites are located in areas that show distinct seasonal patterns due to phenology. Austria has the typical European summer growing season with leaf-off time in winter for deciduous species due to low temperatures. The forests of Malawi also show strong phenological variation as water scarcity during the dry season (typically May–October) causes leaf-fall. The typical temporal NDVI signatures of different forest types are shown in Figure 2. Each time series represents two years of NDVI observations for a single pixel corresponding to a specific forest type and test site. In all cases, data from two orbits is used and cloud/snow masking has been applied as described below. Aside from the different seasonal patterns, several other characteristic properties of the data can be observed:

- The time series density achieved with the Sentinel-2 constellation and overlapping orbits has a high potential for near real-time monitoring applications.
- Larger data gaps occur during winter (snow cover in Austria) or the rainy season (frequent cloud cover in Malawi).
- The illustration also gives an impression of the smoothness of the time series. Significant short-term variance of consecutive observations caused by limited multi-temporal geometric registration accuracy [22], different viewing angles, and remaining atmospheric effects has to be expected.



**Figure 2.** Typical temporal NDVI (Normalized Difference Vegetation Index) signatures over two years for single pixels corresponding to specific forest types and test sites. Larger data gaps occur in the winter months due to snow cover in Austria (AUT) or the rainy season in Malawi (MWI). Outside of these periods, the observation density is high enough to capture the seasonal patterns.

The current data quality report issued by the Sentinel-2 Mission Performance Centre (S2 MPC) gives statistics for the multi-temporal geometric registration performance [22]. For about half of the images, the co-registration error is larger than 0.5 pixels at 10 m resolution. Within homogenous land-cover areas, this can be treated as an additional noise component. The high geometric uncertainty becomes a larger problem at the border regions between land-cover classes, especially if a given pixel jumps between forested and non-forested states.

All available Sentinel-2 scenes with a nominal cloud cover below 90% were downloaded for the test sites (Table 1) at Level-1C and atmospherically corrected to surface reflectance (SR) values using the Sen2Cor processor version 2.5.5 [23]. We then resample the 20 m bands to 10 m spatial resolution and stack all bands to a 10-band output image. We calculate a combined cloud, cloud shadow, and snow mask with the FMask algorithm [24,25]. This mask is slightly altered by morphological operations (erode, expand) to fill cloud holes and the masked pixels are then removed from the pre-processed S-2 imagery by assigning no-data values to them. We also perform a topographic correction based on a modified Minnaert correction [26] using the Shuttle Radar Topography Mission (SRTM) model at 30 m spatial resolution as digital elevation model (DEM).

**Table 1.** Earth ObservationData Information.

Test Site	Tile ID	No. of Images Used	Time Windows	
			Initialization	Change Detection
Malawi—one orbit (R135)	36LWM	86	2016–2018	2019
Malawi—two orbits (R135 and R092)	36LWM	134	2016–2018	2019
Austria—two orbits (R122 and R079)	33TWN	160	2016–2018	2019

## 2.2. Change Detection Method Using a Structural Time Series Model and the Kalman Filter

The underlying assumption of the monitoring approach presented in this section is that the normal temporal trajectory of a given spectral band over the course of the year can be captured by a univariate structural time series model. Within the structural model, we can further distinguish between an observation sub-model and a dynamic sub-model. The observation sub-model on the one hand defines the relationship of the measurements to a set of state variables which cannot be observed directly. In a structural time series model, the state variables usually represent the series' additive decomposition into trend, seasonal, and long-term cyclical components. The dynamic sub-model on the other hand describes the expected temporal evolution of the state variables. By formulating the dynamic sub-model in continuous time, the problem of irregular sampling intervals is addressed. The Kalman filter is used to fit the model to the data and operates recursively from one point in time to the next. Each recursion may be divided into two steps. In the time-update step, the states' temporal evolution is predicted based on the dynamic sub-model. It is followed by the measurement update step, where the predicted state estimate is enhanced by incorporating newly available observations. Abrupt changes of the spectral signature, possibly linked to a forest disturbance, are indicated by statistically significant deviations between new observations and the Kalman filter predictions. The recursive operation of the filter further implies that prior knowledge about the initial state is required. Because the ability to distinguish anomalies from normal seasonal changes depends on the quality of the state estimates, a proper initialization is of high importance. Therefore, a robust least squares method is used to estimate the initial state from a historic time series covering at least one full year prior to the beginning of the monitoring period (see initialization time window in Table 1). The following sub-sections describe the implementation in more detail.

### 2.2.1. Time Series Model

Structural time series models are mathematically formulated using the discrete-time state space representation [15]. This concept assumes that a linear, time-variant system can be described by a set of state variables. Because these variables can usually not be observed directly, an observation sub-model linking the system state to a set of measurements is required. In case of univariate time series, the measurement equation is

$$z_k = \mathbf{h}_k \mathbf{x}_k + r_k, \quad (1)$$

where  $\mathbf{h}_k$  is a row vector,  $\mathbf{x}_k$  denotes the state vector, and  $z_k$  is a scalar observation made at time  $t_k$ . In addition,  $r_k$  represents serially uncorrelated observation noise with mean zero and variance  $R_k$ . We will refer to discrete points in time as epochs and define an epoch index  $k \in \mathbb{N}$  to indicate a possible time-dependency of variables. The temporal evolution of the state vector is described by a dynamic sub-model using the transition equation:

$$\mathbf{x}_k = \Phi_k \mathbf{x}_{k-1} + \mathbf{q}_k, \quad (2)$$

where  $\Phi_k$  denotes the transition matrix and  $\mathbf{q}_k$  is a vector of serially uncorrelated process noise with mean zero and covariance matrix  $\mathbf{Q}_k$ . Note that no assumptions regarding the distributions of the observation and process noises are made at this point, but they are supposed to be uncorrelated with each other in all epochs. It is further assumed that the initial state  $\mathbf{x}_0$  is known with a level of uncertainty characterized by the state error covariance matrix  $\mathbf{P}_0$ .

The structural time series model implemented in this study enables a decomposition of the time series into additive trend, seasonal, and irregular components. We assume a constant trend and directly introduce the state variable  $\mu_k$ , which represents the trend level at time  $t_k$ . The seasonal component on the other hand is constructed by a sum of  $P$ -periodic cosine waves, with  $P$  denoting the fundamental duration of the seasonal cycle. Considering the nature of the time series investigated here, it is appropriate to measure time in days and thus set  $P$  to 365.25. For each cosine wave added to

the seasonal component, two more variables are appended to the state vector. Taking a wave with frequency  $\omega_i$ , the state variable  $\gamma_{i,k}$  representing the cosine waves' level at time  $t_k$  is added along with the variable  $\gamma_{i,k}^*$  whose interpretation is not particularly important. The structural model for the Malawi test site features a seasonal component with frequencies corresponding to one and two periods per year. In case of the Austrian test site, we use just the fundamental frequency because of the less complex seasonal pattern. The measurement Equation (1) of the Malawi model becomes

$$z_k = [ 1 \quad 1 \quad 0 \quad 1 \quad 0 ] \begin{bmatrix} \mu \\ \gamma_1 \\ \gamma_1^* \\ \gamma_2 \\ \gamma_2^* \end{bmatrix}_k + r_k \tag{3}$$

Note that  $z_k$  represents the observed reflectance value at time  $t_k$  and  $\mathbf{h}$  is time-invariant. Using  $\Delta t_k = t_k - t_{k-1}$  measured in days and  $\omega_i = 2\pi i / 365.25$ , the transition Equation (2) of the Malawi model becomes

$$\begin{bmatrix} \mu \\ \gamma_1 \\ \gamma_1^* \\ \gamma_2 \\ \gamma_2^* \end{bmatrix}_k = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & \cos(\omega_1 \Delta t_k) & \sin(\omega_1 \Delta t_k) & 0 & 0 \\ 0 & -\sin(\omega_1 \Delta t_k) & \cos(\omega_1 \Delta t_k) & 0 & 0 \\ 0 & 0 & 0 & \cos(\omega_2 \Delta t_k) & \sin(\omega_2 \Delta t_k) \\ 0 & 0 & 0 & -\sin(\omega_2 \Delta t_k) & \cos(\omega_2 \Delta t_k) \end{bmatrix} \begin{bmatrix} \mu \\ \gamma_1 \\ \gamma_1^* \\ \gamma_2 \\ \gamma_2^* \end{bmatrix}_{k-1} + \begin{bmatrix} q_\mu \\ q_{\gamma_1} \\ q_{\gamma_1^*} \\ q_{\gamma_2} \\ q_{\gamma_2^*} \end{bmatrix}_k \tag{4}$$

In order to fully specify the state space model, two more quantities need to be defined:  $R_k$  and  $\mathbf{Q}_k$ . The first one will be addressed in the next sub-section. The covariance matrix  $\mathbf{Q}_k$  quantifies the uncertainties within the dynamic sub-model and is defined depending on the nature of the process noise. Here, we assume that the state transition is affected by multivariate, continuous time white noise with constant variances  $Q_c$  independently specified for the trend and seasonal component, that is:

$$\mathbf{Q}_k = \Delta t_k \begin{bmatrix} Q_c^{\text{trend}} & 0 & 0 & 0 & 0 \\ 0 & Q_c^{\text{seas}} & 0 & 0 & 0 \\ 0 & 0 & Q_c^{\text{seas}} & 0 & 0 \\ 0 & 0 & 0 & Q_c^{\text{seas}} & 0 \\ 0 & 0 & 0 & 0 & Q_c^{\text{seas}} \end{bmatrix} \tag{5}$$

Note that all off-diagonal elements of  $\mathbf{Q}_k$  equal zero and we therefore assume that the different process noise components are uncorrelated. Implementing the dynamic model in continuous time means that the time series model is specified for arbitrary values of  $\Delta t_k$  and irregular sampling intervals do not pose a problem. Finding appropriate values for  $Q_c$  is one of the difficulties. Our experiments showed that setting individual values for each pixel and band proportional to the respective observation variance  $R$  works quite well. Since  $R$  is estimated from the data (see the next sub-section), the user has to specify two proportionality factors, one for the trend component and one for the seasonal component. We recommend to let  $Q_c^{\text{seas}}$  be larger than  $Q_c^{\text{trend}}$  in order to make the time series model more responsive to phenological variations.

### 2.2.2. Initial State Estimation

In order to obtain estimates of the initial state vector  $\mathbf{x}_0$ , its associated covariance matrix  $\mathbf{P}_0$ , and the observation variance  $R$ , the state space model outlined above is transformed to a linear regression model with the measurement equation:

$$\begin{bmatrix} z_1 \\ z_2 \\ \vdots \\ z_m \end{bmatrix} = \begin{bmatrix} \mathbf{h}\Phi(t_1, t_0) \\ \mathbf{h}\Phi(t_2, t_0) \\ \vdots \\ \mathbf{h}\Phi(t_m, t_0) \end{bmatrix} \mathbf{x}_0 + \begin{bmatrix} r_1 \\ r_2 \\ \vdots \\ r_m \end{bmatrix} \quad (6)$$

Equation (6) also illustrates the key difference between a structural time series model and a regression model. The latter does not distinguish between observation sub-model and dynamic sub-model and has no notion of process noise. Standard least squares methods may be used to obtain estimates of  $\mathbf{x}_0$  and  $\mathbf{P}_0$  from a batch of  $m$  historic observations acquired before the monitoring period. We applied a robust parameter estimation approach known as iteratively reweighted least squares (IRLS). The technique belongs to the class of  $M$ -estimators [27] and the implementation follows [28]. We further use the mean squared error of the weighted least squares fit as an estimate for the observation variance  $R$ . If the number of valid historic observations is low, for example in areas with extremely high cloud probability, the observation variance can be severely underestimated. Existing studies applying harmonic regression also report this issue and we take up the suggestion to define a certain minimum value for  $R$  that is used if the estimate is lower [12,29].

### 2.2.3. Kalman Filter

Once the time series model and the initial values  $\mathbf{x}_0$  and  $\mathbf{P}_0$  are defined, the discrete-time Kalman filter recursion can be used to obtain estimates for the state and its error covariance matrix in subsequent epochs. The time update step yields the predicted (a-priori) estimates  $\tilde{\mathbf{x}}_k$  and  $\tilde{\mathbf{P}}_k$  based on the dynamic model and the previous estimates at time  $t_{k-1}$ :

$$\tilde{\mathbf{x}}_k = \Phi_k \hat{\mathbf{x}}_{k-1} \quad (7)$$

$$\tilde{\mathbf{P}}_k = \Phi_k \hat{\mathbf{P}}_{k-1} \Phi_k^T + \mathbf{Q}_k \quad (8)$$

Then, the a-priori measurement residual  $y_k$  and its variance  $C_k$  is computed according to Equations (9) and (10). The residual represents the difference of the prediction to the actual measurement and is referred to as innovation, since it contains new information currently not present in the predicted state [17].

$$y_k = z_k - \mathbf{h}\tilde{\mathbf{x}}_k \quad (9)$$

$$C_k = \mathbf{h}\tilde{\mathbf{P}}_k \mathbf{h}^T + R \quad (10)$$

In the measurement update step of each recursion, new information is merged with the predictions to obtain improved (a-posteriori) estimates  $\hat{\mathbf{x}}_k$  and  $\hat{\mathbf{P}}_k$ . The Kalman gain  $\mathbf{k}_k$  (11) determines how much the newly acquired measurement will influence the a-posteriori estimates of the state and its error covariance and appears in both update Equations (12) and (13), where  $\mathbf{I}$  represents an identity matrix.

$$\mathbf{k}_k = \tilde{\mathbf{P}}_k \mathbf{h}^T C_k^{-1} \quad (11)$$

$$\hat{\mathbf{x}}_k = \tilde{\mathbf{x}}_k + \mathbf{k}_k y_k \quad (12)$$

$$\hat{\mathbf{P}}_k = (\mathbf{I} - \mathbf{k}_k \mathbf{h}) \tilde{\mathbf{P}}_k \quad (13)$$

#### 2.2.4. Signal Artefact Handling and Change Detection

When the sequence of measurements processed by the filter may contain artefacts, an additional anomaly detection step should be included before the measurement update. The properties of the innovations can be exploited to detect anomalous measurements by means of a statistical test. Provided that the underlying model assumptions are valid, and the observation noise is Gaussian, the innovations should be normally distributed with mean zero and variance  $C_k$ . The test statistic  $\hat{T}_k$  given in (14) follows the  $\chi^2$ -distribution with a single degree of freedom. The hypotheses to be tested on a significance level  $\alpha$  are stated in (15) and (16), respectively:

$$\hat{T}_k = \frac{y_k^2}{C_k}, \quad \hat{T}_k \sim \chi^2 \quad (14)$$

$$H_0 : y_k = 0 \text{ if } \hat{T}_k \leq \chi_{1, 1-\alpha}^2 \quad (15)$$

$$H_1 : y_k \neq 0 \text{ otherwise} \quad (16)$$

Considering that anomalous observations will cause large innovations, the null hypothesis will be rejected. In that case, the measurement update is not carried out so that the state estimate remains unbiased.

The change detection part of the algorithm is based on the cumulative sum control chart (or CUSUM) [20]. The key quantity here is again the sequence of filter innovations, which should have zero mean if the Kalman filter model assumptions reflect the truth closely enough. We use the CUSUM control chart to monitor if the innovations deviate from mean zero. Of course, the presence of artefacts like un-masked clouds presents a problem in this regard, because a single innovation corresponding to an artefact may shift the mean significantly and hence trigger a (inaccurate) change signal. The previously described anomaly test cannot distinguish between an artefact and an abrupt land cover change. To work around this limitation, a new quantity we call edited innovation  $\check{y}_k$  is introduced. It represents the original innovation divided by its standard deviation, but also limited in magnitude based on the significance level  $\alpha$  specified for the anomaly test, that is:

$$\check{y}_k = \begin{cases} \min\left(\frac{y_k}{\sqrt{C_k}}, \sqrt{\chi_{1, 1-\alpha}^2}\right) & \text{if } y_k \geq 0 \\ \max\left(\frac{y_k}{\sqrt{C_k}}, -\sqrt{\chi_{1, 1-\alpha}^2}\right) & \text{if } y_k < 0 \end{cases} \quad (17)$$

Dividing by the standard deviation means that sequences of edited innovations should be close to having unit variance across different bands and pixels. The limit operation on the other hand ensures that the CUSUM control chart with  $\check{y}_k$  as input is less sensitive to single statistical outliers. A temporal aggregation of edited innovations with the same sign is required to shift the mean of the sequence significantly. The CUSUM test statistic for a positive mean shift with respect to spectral band  $i$  is implemented as:

$$\begin{aligned} S_i^+(0) &= 0 \\ S_i^+(k) &= \max(0, S_i^+(k-1) + \check{y}_i(k) - d) \end{aligned} \quad (18)$$

where  $d$  is a drift parameter which generally compensates small deviations and also ensures that effects of occasional signal artefacts on the test statistic fade away over time. A mean shift is signaled if the test statistic crosses a predefined threshold. Instead of evaluating all processed bands separately, we decided to aggregate the test statistics of several bands by simple summation and then use a global threshold. We look for anomalous reflectance increases in the red, red edge, and short wave infrared

(SWIR) bands, as these have proven to be sensitive to vegetation changes [30]. The implemented criterion for detecting a forest disturbance at time  $t_k$  is given in Equation (19) using S-2 band numbers.

$$\sum_i S_i^+(k) > \text{change threshold} \text{ where } i \in \{B04, B05, B11, B12\} \tag{19}$$

Both the change threshold in (19) and the drift parameter in (18) are user-defined tuning parameters. Appropriate values need to be determined empirically. Because of the limit operation described by (17), the maximum of a single edited innovation is a known constant and amounts to  $\sim 2.57$  if  $\alpha = 1\%$ . This knowledge provides a helpful yardstick for setting both threshold and drift parameters. The statistical normalization applied to the edited innovations means that the same values can be used globally for all pixels. The flowchart depicted in Figure 3 illustrates how the methods discussed in the preceding sections are joined together in order to create a data-driven algorithm capable of detecting abrupt changes on the pixel level. A summary and some additional explanatory comments are given below.

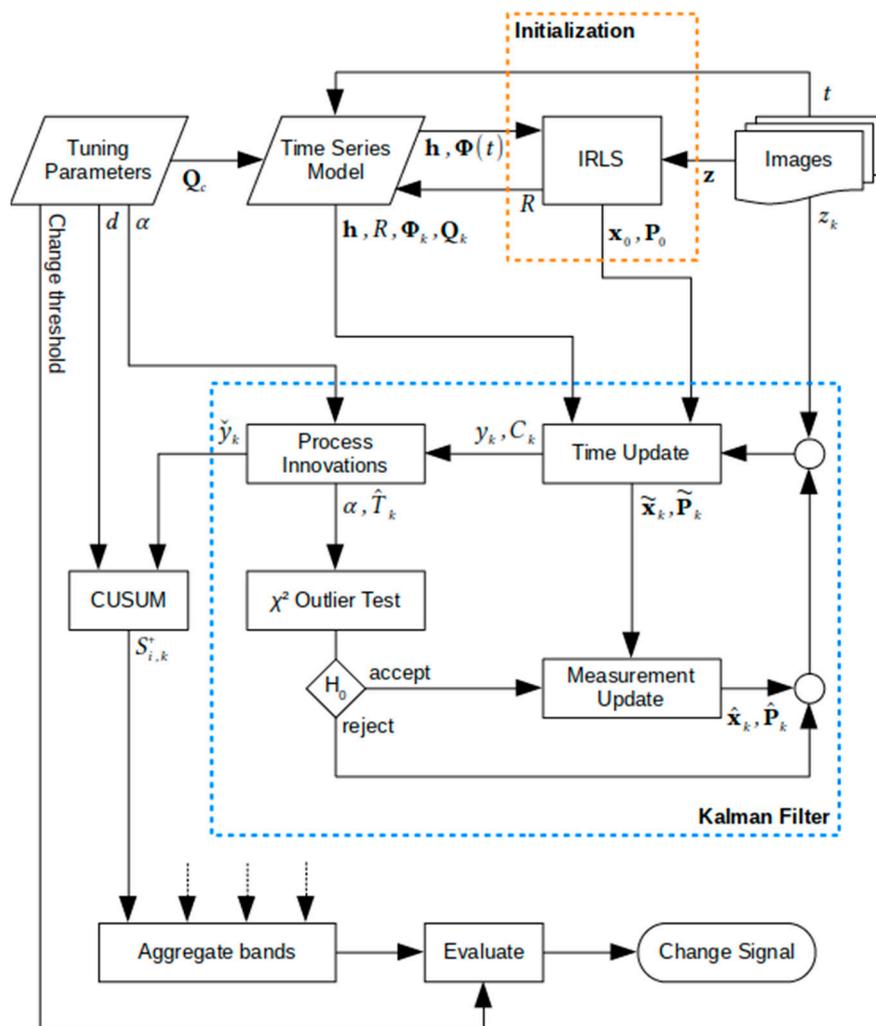
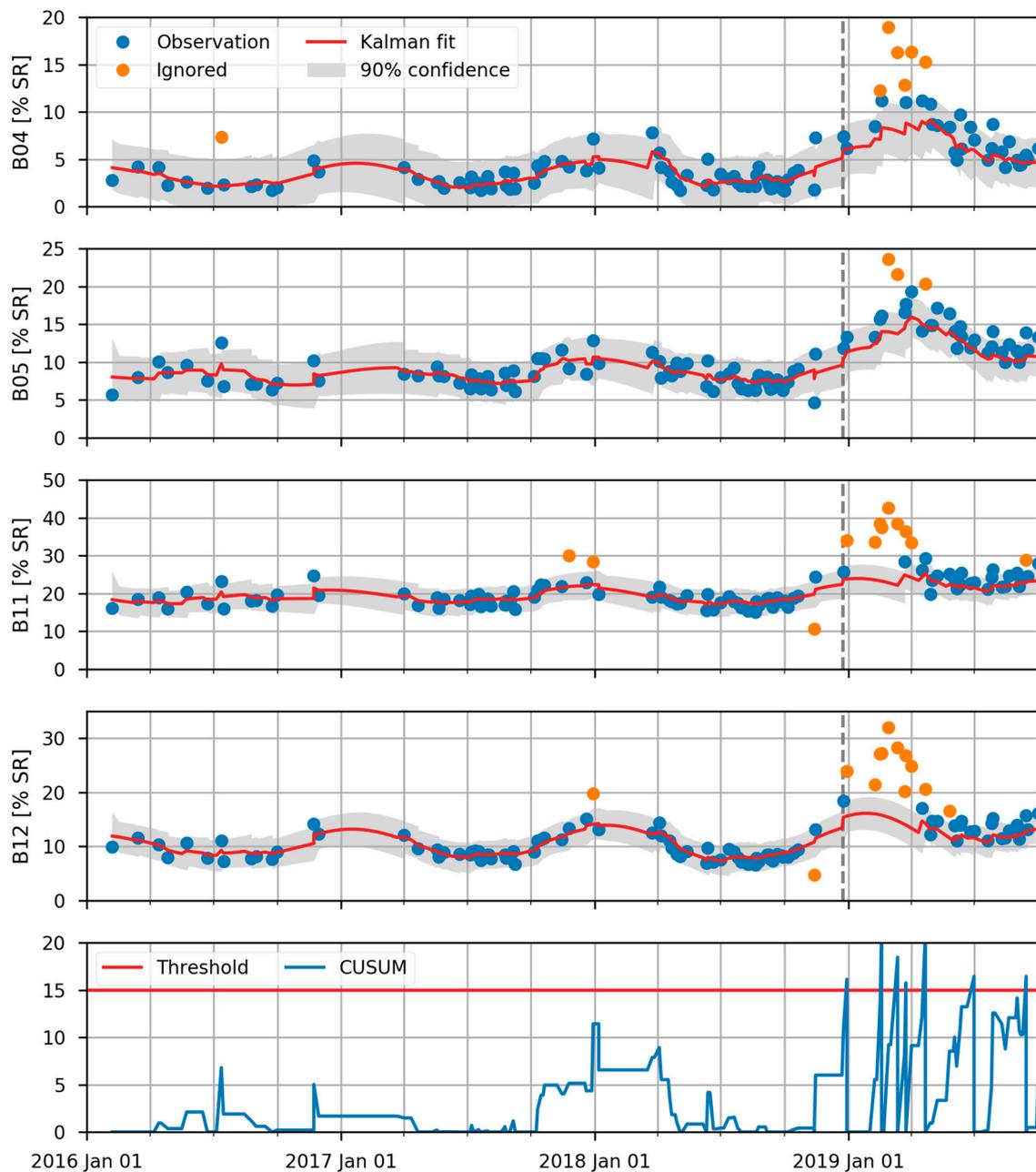


Figure 3. Flow chart of the Kalman filter approach.

1. The initial state as well as the observation noise are estimated on the pixel level using the iteratively reweighted least squares (IRLS) method. Therefore, the user has to supply a stack of historic images as a training dataset. At least one year of observations is required, but, especially for models using two seasonal frequencies, we recommend two or three years to ensure a stable initial model.
2. New images acquired in the monitoring period are processed one-at-a-time in a Kalman filter loop. A hypothesis test is used to identify anomalous observations showing significant deviations to the prediction.
3. Whenever an observation is marked as anomalous, the measurement update step is bypassed in order to avoid adverse influence of signal artefacts on the state estimates.
4. Spectral bands are processed in parallel and the aggregated CUSUM test statistic used for change detection is evaluated after processing each new image.

Figure 4 visualizes the method for a forest change pixel (10 by 10 m) in a deciduous forest in the Austrian test site. The first four sub-plots show the surface reflectance of Sentinel-2 bands 4 (red), 5 (red edge), 11 (SWIR1), and 12 (SWIR2) over time. The sub-plot at the bottom shows the value of the CUSUM test statistic defined in Section 2.2.4 over time. A phenological cycle that is typical for deciduous tree species can be observed. Higher reflectance values correspond to the leaf-off season during winter. The red line represents the Kalman fit and the grey area corresponds to the 90% confidence interval of the model forecast. While blue observations are considered in the model and measurement update step (compare Figure 3), the orange observations are flagged as anomalous and therefore ignored during the measurement update step. Please note that the level of significance for the anomaly test is  $\alpha = 1\%$ , hence blue observations may also appear outside of the plotted confidence interval. Occasional positive anomalies will cause a short-lived increase of the related CUSUM, but in most cases, it should not exceed the threshold (e.g., middle of 2016 or end of 2017 in Figure 4). Such anomalies occur if for example small cloud artefacts remain in the pre-processed data. A prolonged increase of the CUSUM is triggered by a persistent signal shift. At some point, the threshold is exceeded, and the pixel is flagged as changed, in this case, at the end of 2018. The vertical dashed line in Figure 4 marks the date on which the change is visible for the first time in the S-2 imagery. Note that anomalies with significant magnitude are first detected only in the SWIR bands, while the signal shift in the red and red edge band initially occurs within the 90% confidence interval; however, the aggregated CUSUM test statistic allows a timely detection of the change. After a threshold crossing, the test statistic is reset to zero. Repeated change alerts, as shown in this example, may thus occur. This could be used to increase the confidence about a detected change at the cost of delaying the detection; however, this aspect has not been investigated in detail in this work. For the resulting forest change maps at the two test sites, we first include all Sentinel-2 single pixel changes (10 by 10 m) that were detected by the described approach within the change detection time window. For the final forest change maps, a minimum mapping unit of 0.1 ha is applied to the forest change stratum, which relates to change areas represented by at least ten connected 10 m pixels. Detected forest change areas smaller than 10 pixels are removed from the final forest change maps.



**Figure 4.** Example of the time series model for a deciduous forest pixel at the Austrian test site, where a forest change occurred at the end of 2018. The plots show the surface reflectance of Sentinel-2 bands 4, 5, 11, and 12. The vertical dashed line marks the date on which the change is visible for the first time in the S-2 image time series. The plot at the bottom shows the cumulative sum control chart (CUSUM) used for detecting the change event (coordinates: X 559045 m, Y 5235095 m, in EPSG 32633).

### 2.3. Evaluation Method

In order to test the operational application of the presented forest change detection method, a full area-based validation of the resulting change maps is performed in both test sites based on stratified random sampling points that are located within forest areas outlined by benchmark forest masks (status 2016 for Malawi and summer 2018 for Austria). The overall number of samples and samples per stratum are based on recommendations for land cover accuracy estimation [31,32]. Points that were found to be already deforested before the beginning of the change detection time window were excluded from the analysis since these cases correspond to errors of the initial forest masks.

For Malawi, 849 reference sample points are used for statistical analysis, with 735 sample points belonging to stratum forest and 114 sample points belonging to stratum change. During blind interpretation, we flagged all sample points that were located at or very close (~10 m) to the border of a forest change patch. This allows us to treat these points as correctly classified in a subsequent plausibility analysis. This approach is usually termed “plausibility analysis”, because it is deemed plausible that the border point can also belong to the other stratum. This is especially true for the small-scale 10 by 10 m change assessment we use in this study considering that some of the spectral bands of Sentinel-2 used to derive the changes only have a nominal spatial resolution of 20 m. For Malawi, assessments are based on the following validation approaches and input data options:

- Blind versus plausibility validation approach.
- Input data from only one orbit versus combined input data from two orbits.

The plausibility analysis is only performed for the two-orbit input data. Thus, the comparison of the two input data options in Malawi is based only on the blind assessment approach.

For Austria, we interpreted a total of 1585 reference points, of which 21 points were removed as they were found to be non-forest already before the beginning of the change detection window. From the remaining 1564 points, 1212 belong to stratum forest and 352 to stratum change. The plausibility approach was carried out in the same manner as in Malawi. Please note that the Austrian test site is only partly covered by two orbits, which is a typical data scenario when working with Sentinel-2 data. Combined input data from two orbits is used where possible. For all assessments, we provide unbiased estimates of the mapped area proportions as well as the products’, users’, and producers’ accuracies by applying Equations (1) and (6)–(8) from Reference [31].

### 3. Results

Table 2 gives a summary of key validation results for both test sites, different assessment approaches, and input data scenarios. The detailed results of the forest disturbance detection are shown in Tables 3–7, where the upper part presents the sample counts and the lower part presents the unbiased estimates of area proportion and accuracy measures. Overall accuracies are very high (96.4–99.3%). This is not surprising since the unchanged forest class accounts for 98.7% of the validation area at the Malawi test site and for 98.6% at the Austrian test site. It is also evidence for a low rate of false-positive change detections. For better comparison, Table 2 also lists the users’ accuracies and producers’ accuracies of the different validation approaches for the change class. The plausibility analysis increases the users’ accuracy of the change class by 17.5% in Malawi and by 4.1% in Austria. Producers’ accuracies show a strong increase of 31.1% in Malawi and 29.1% in Austria. Overall accuracies after plausibility analysis reach 99.3% at both test sites. Results show that combining data from two orbits leads to an increase in users’ accuracy of 7.6% (Malawi—blind validation approach) compared to using data from only one orbit, while producers’ accuracies remain the same. The detailed accuracy metrics for Malawi are shown in Tables 3–5, and for Austria, they are listed in Tables 6 and 7. For comparison between one and two orbits, please compare Tables 3 and 4. For comparison of blind and plausibility results, please compare Tables 4 and 5 for Malawi and Tables 6 and 7 for Austria.

**Table 2.** Summary of the accuracy measures for the forest change maps at both test sites for different assessment approaches (blind interpretation and plausibility analysis).

Country	Validation Approach	Change Class		Overall
		Users	Producers	
Malawi	Blind—1 orbit	63.5%	37.2%	96.4%
Malawi	Blind—2 orbits	71.1%	37.1%	98.0%
Malawi	Plausibility—2 orbits	88.6%	68.2%	99.3%
Austria	Blind—2 orbits	72.7%	51.5%	98.6%
Austria	Plausibility—2 orbits	76.8%	80.4%	99.3%

**Table 3.** Error matrix for Malawi, blind approach—1 orbit.

		Reference			$A_i$ (ha)	Users	Producers	Overall
		Forest	Change	Total				
Map	Forest	714	20	734	55,258			
	Change	42	73	115	1407			
	Total	756	93	849	56,665			
	Forest	0.949	0.027	0.975	54,266	97.3 ± 1.2%	99.1 ± 0.2%	
	Change	0.009	0.016	0.025	2399	63.5 ± 8.8%	37.2 ± 10.7%	
	Total	0.958	0.043	1.000	56,665			96.4 ± 1.2%

Confidence interval of accuracy measures: 95%.

**Table 4.** Error matrix for Malawi, blind approach—2 orbits.

		Reference			$A_i$ (ha)	Users	Producers	Overall
		Forest	Change	Total				
Map	Forest	723	12	735	55,909			
	Change	33	81	114	756			
	Total	756	93	849	56,665			
	Forest	0.971	0.016	0.987	55,215	98.4 ± 0.9%	99.6 ± 0.1%	
	Change	0.004	0.009	0.013	1450	71.1 ± 8.4%	37.1 ± 13.4%	
	Total	0.974	0.026	1.000	56,665			98.0 ± 0.9%

Confidence interval of accuracy measures: 95%.

**Table 5.** Error matrix for Malawi, plausibility approach—2 orbits.

		Reference			$A_i$ (ha)	Users	Producers	Overall
		Forest	Change	Total				
Map	Forest	731	4	734	55,909			
	Change	13	101	115	756			
	Total	744	105	849	56,665			
	Forest	0.981	0.005	0.987	55,691	99.5 ± 0.5%	99.9 ± 0.1%	
	Change	0.002	0.012	0.013	974	88.6 ± 5.9%	68.2 ± 21.0%	
	Total	0.983	0.017	1.000	56,665			99.3 ± 0.5%

Confidence interval of accuracy measures: 95%.

**Table 6.** Error matrix for Austria, blind approach.

		Reference			$A_i$ (ha)	Users	Producers	Overall
		Forest	Change	Total				
Map	Forest	1084	11	1095	426,192			
	Change	128	341	469	6253			
	Total	1212	352	1564	432,445			
	Forest	0.976	0.010	0.986	423,617	99.0 ± 0.6%	99.6 ± 0.1%	
	Change	0.004	0.010	0.014	8828	72.7 ± 4.0%	51.5 ± 14.8%	
	Total	0.980	0.020	1.000	432,445			98.6 ± 0.6%

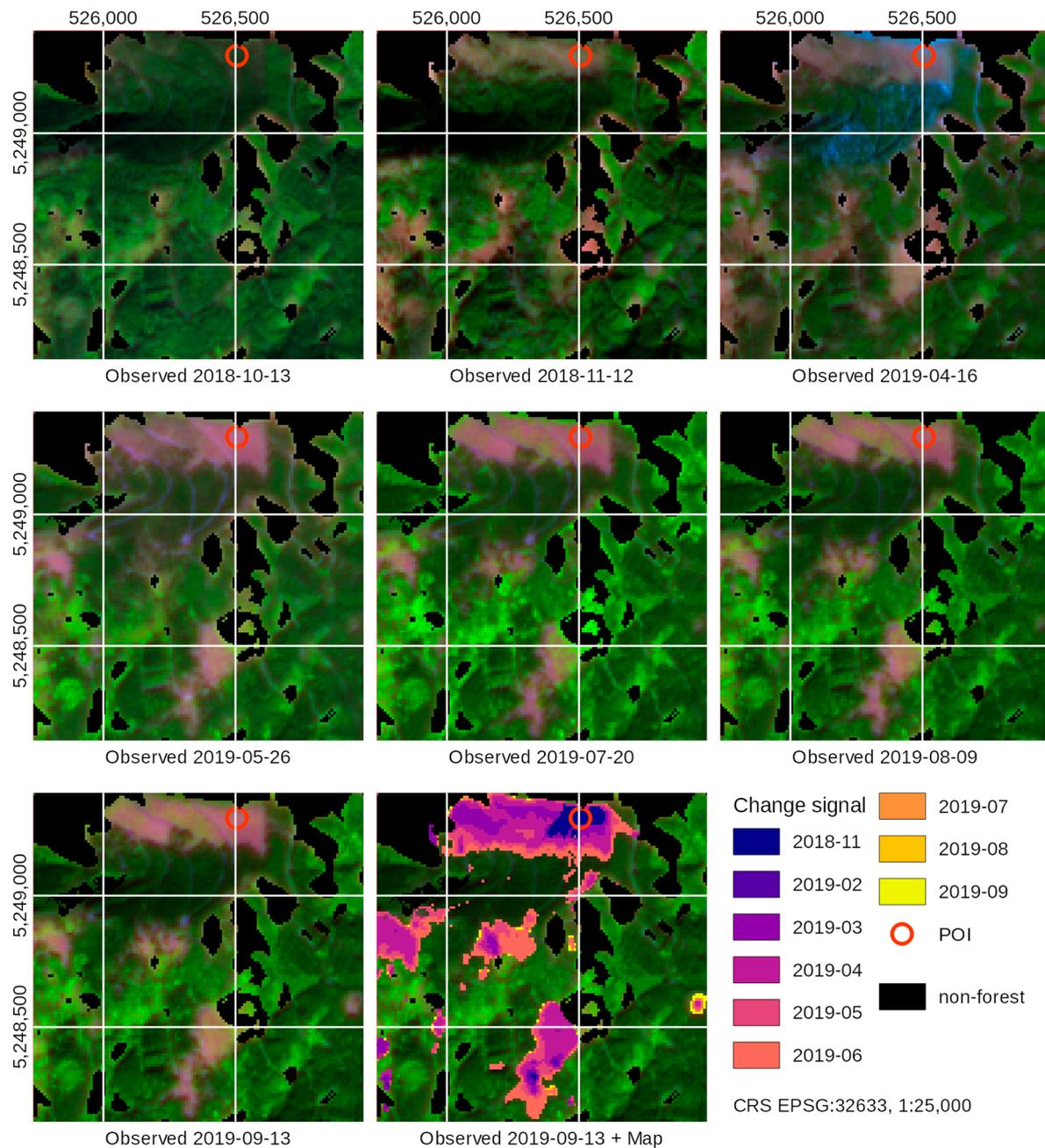
Confidence interval of accuracy measures: 95%.

**Table 7.** Error matrix for Austria, plausibility approach.

		Reference			$A_i$ (ha)	Users	Producers	Overall
		Forest	Change	Total				
Map	Forest	1092	3	1095	426,192			
	Change	109	360	469	6253			
	Total	1201	363	1564	432,445			
	Forest	0.983	0.003	0.986	426,478	99.7 ± 0.3%	99.6 ± 0.1%	
	Change	0.003	0.011	0.014	5967	76.8 ± 3.8%	80.4 ± 17.8%	
	Total	0.986	0.014	1.000	432,445			99.3 ± 0.3%

Confidence interval of accuracy measures: 95%.

Figure 5 shows some mapping examples of windthrow detection for the Austrian test site. The series of S-2 images illustrates the development of forest disturbances occurring at an alpine subset of the test site that was affected by windthrow late in the year 2018 (storm Vaia on 29/30 October 2018). Please note that the depicted sequence does not represent all available imagery, but a selection of cloud-free images of the area of interest. The first image (top left) shows the undisturbed state two weeks before the storm event. A large windthrow area can be identified in the second image (top middle; surrounding the red circle). Forest change detection is complicated by the fact that the timing of the storm exactly coincides with leaf discoloring and leaf-fall for the broadleaf trees at the site and by subsequent bad weather conditions also leading to snow cover. Remaining snow (blueish colored areas) can still be seen in early April 2019 (top right image). Harvesting of damaged forest areas leads to a continuous increase in deforested areas in the subsequent images. Typically, also areas adjacent to completely thrown forest areas are affected to some degree by wind damage, e.g., single tree throws or broken stems. In Austria, all storm damage affected areas are usually harvested directly after the storm event to prevent bark beetles from spreading. The last image in the series shows which pixels are flagged as changed by the algorithm and in which month the changes were detected. Due to frequent cloud and snow cover at the site from mid-November 2018 to April 2019, only few pixels were detected as changed shortly after the storm as only two usable post-storm observations were available until the end of 2018. Many detections are thus delayed until early spring 2019, when the time series of snow-free Sentinel-2 imagery continues and when the damaged forest areas are being cleared (removal of damaged and broken trees). This example shows both the capabilities, but also the limitations of the method when used in near real-time forest change mapping scenarios for Central European/Alpine forests.



**Figure 5.** Examples of near real-time forest change detection in Austria. Storm “Vaia” windthrows on 29/30 October 2018 at an alpine forest site in Styria (~1400 m above sea level).

The software needed to create the presented mapping example and time series plots has been implemented in Python. This Python implementation is capable of processing smaller test sites (up to 10 MP) for development and demonstration purposes in reasonable time. For large-area processing, a performance-optimized C/C++ implementation is also available. However, both versions rely on certain in-house modules and libraries which are subject to licensing restrictions.

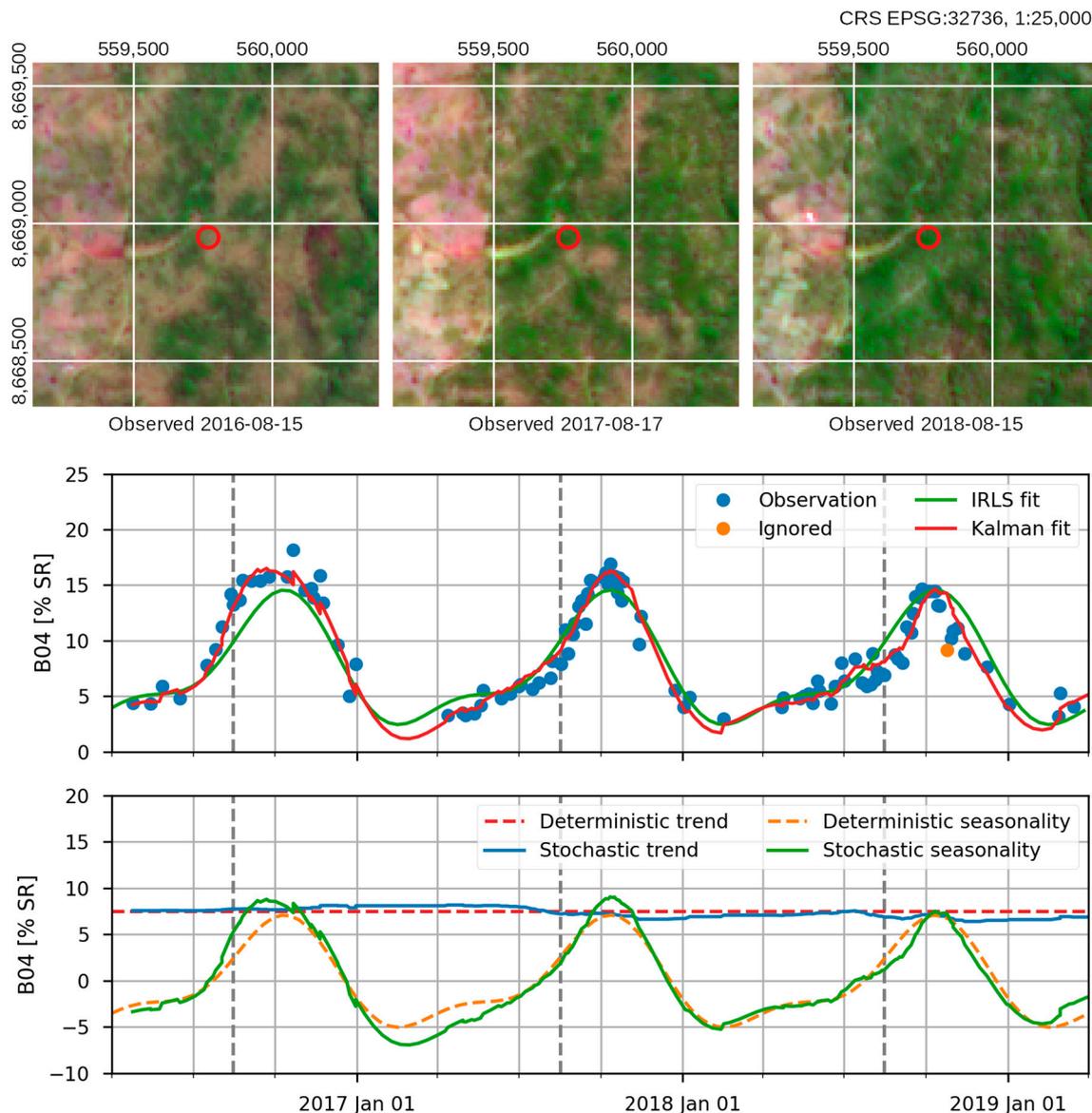
#### 4. Discussion

When comparing our results to those of other studies on near real-time forest change detection, similarly high overall accuracy values can be observed. This is to be expected for a land cover class and test sites that are characterized by a large proportion of unchanged areas. In this case, users’ and producers’ accuracies of the change class are much better measures to determine the suitability and

practical implementation potential of a monitoring approach. Many recent near real-time forest change monitoring approaches detect and evaluate changes at a minimum mapping unit (MMU) of at least one Landsat 8 pixel (~0.1 ha) as most related studies are still primarily based on Landsat 8 imagery with a lower spatial resolution than Sentinel-2 [29,33–35]. For better comparability, we have also applied a 0.1 ha MMU for the change class, but our change polygons can be of any shape representing at least ten connected Sentinel-2 change pixels.

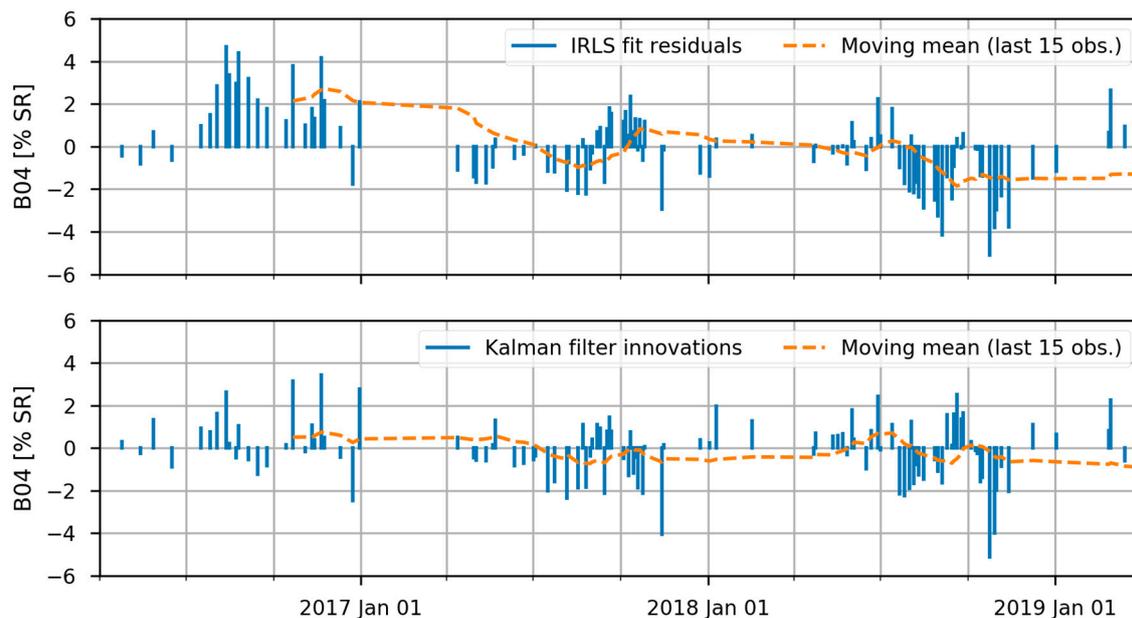
There are a number of near real-time forest change studies and algorithms that provide accuracy statistics similar to those presented here. The near real-time humid tropical forest monitoring approach of Global Forest Watch was evaluated at national scale in Peru with a reported producer's accuracy of 75.4% and a user's accuracy of 92.2% [33]. However, a direct comparison of results is difficult, since 63% of the detected forest changes in this study were at least one hectare in size, while only 4% were the size of one single Landsat pixel (~0.1 ha). In our test sites, the vast majority of change polygons is smaller than 0.5 ha. Another change detection approach has recently been developed and tested for seasonal tropical forests in Myanmar. It uses a harmonic regression model to account for seasonality and a set of time series disturbance probabilities to detect forest changes [35]. The authors report overall disturbance detection accuracies of 78.3% for Landsat 8 data and 83.6% for Landsat 8 data combined with Sentinel-1 data. The reported users' accuracy for the disturbance class at the single Landsat 8 pixel level (~0.1 ha) is 84.1% and producers' accuracy is 78.6%, but disturbance detections are significantly delayed by 65 days on average. Very high users' and producers' accuracies of 88% and 89% were also reported for a forest change detection approach that combines Landsat 8, Sentinel-1, and ALOS-2 PALSAR-2 data at a dry tropical forest site in Bolivia characterized by distinct dry and wet seasons [34]. Instead of adapting the time series model in near real-time to the observed phenology, the authors apply a-priori spatial normalization to reduce the dry forest seasonality in the time series and then apply the change detection analysis on the normalized data [10]. The method is reported to perform well for extreme events, but we would expect that such a combined normalization and change detection approach could fail in years that behave significantly different from the average, e.g., years with extreme dryness (deviation in magnitude) or a very late start (deviation in time) of the wet season.

Our approach to use the Kalman filter is quite different in that it continuously accounts for inter-annual phenology variations and updates the time series model. The method proved to work properly at both test sites even for cases where the phenology curve significantly differs between years. Figure 6 shows three false-color Sentinel-2 images (band combination B11, B04, B03) of an unchanged forest area at the Malawi test site from three consecutive years, acquired almost on the same day of each year (mid-August). Significant differences in the phenological state can be observed both in the images as well as in the corresponding time series plot below (Figure 6). The time series plot shows the Sentinel-2 red band surface reflectance for a pixel at the center of the red circle in the images. Each year shows time periods with strong deviations from the average IRLS fit over all years. Year 2016 shows an early and stronger than average dry season, while the observations of year 2018 indicate a pronounced prolongation of the spring rainy season and thus late start of the dry season. The observed reflectances for August 2016 and August 2018 deviate by more than 7% of total surface reflectance. A widely used deterministic modelling approach based on harmonic regression (see IRLS fit represented by the green line in Figure 6) results in large and prolonged deviations between observations and the model curve which is disadvantageous for change detection. The Kalman filter is able to track the differences in plant vitality much better. In the given example, only one single observation in December 2018 was flagged as anomalous.



**Figure 6.** The illustration shows 3 false-color images (B11, B04, B03) acquired in three consecutive years in Malawi, almost on the same day (gray dashed lines). The upper time series plot shows the red band surface reflectance for a single pixel located in the center of the red circle together with a harmonic regression fit (IRLS—iteratively reweighted least squares) and the Kalman filtered trajectory. The lower plot shows the trend and seasonal components separately.

This seasonal phenological effect is further illustrated in Figure 7, which compares the residuals of the IRLS fit shown in Figure 6 with the corresponding Kalman filter innovations. The IRLS residuals show a much larger degree of unwanted systematic patterns left in the sequence, which is apparent when comparing the moving mean of the last 15 observations (dashed orange line in Figure 7). These systematic patterns represent the difference in phenology (both in time and magnitude) from the “average” fit. In combination with the CUSUM test, these systematic deviations of the IRLS fit would result in a larger amount of erroneous forest change detections. In case of the Kalman filter innovations, remaining non-random patterns appear when the model receives adjustments to the current phenology (especially the second half of 2018 in Figure 7). However, the drift parameter in Equation (18) avoids that small and short-lived deviations from the zero mean assumption accumulate to a change signal.



**Figure 7.** The residuals of the regression fit (IRLS) and the corresponding Kalman filter innovations for the same plot as in Figure 6. The orange dashed line represents the moving mean based on the last 15 observations.

We could show that the combination of Kalman filter innovations and the CUSUM test is suitable for rapid near real-time detection of abrupt forest changes as shown in Figure 4. With the CUSUM test, changes can be detected with a smaller time lag than with traditional change detection methods based on multiple confirmations [21,35–37]. Many existing algorithms require a fixed number of observed anomalies to signal a change, and some enforce the condition that anomalies also have to be detected consecutively. Certain statistical boundaries or change probabilities specific to the method have to be exceeded multiple times for a change to be recorded. When the spectral footprint of a change event is at the border of detectability, these restrictions likely lead to omission errors or a large time lag between change event and detection. Because of its cumulative nature, the proposed CUSUM test statistic is not restricted in the same way. Any post-change observation can add information to the test statistic, regardless of being flagged by the anomaly test or not. Depending on the spectral footprint of the change event, it may take two, three, or more post-change observations to detect the change. At this point, we cannot give a quantitative evaluation of the typical time lag because appropriate reference data were not available.

The reduction of the time lag between change event and change detection is of major importance to near real-time forest monitoring systems. For this reason, several recent studies have combined optical data with Synthetic Aperture Radar (SAR) data in order to reduce this time lag [34,35]. Regarding optical data, persistent cloud cover during rainy seasons represents a major limitation as the time gap between consecutive valid observations may become very large. Change detection methods that require multiple confirmed change detections will therefore show large time lags between the change event and its detection. Two studies carried out in tropical regions reported mean time lags of 63 days and 70 days for Myanmar and Bolivia respectively, if only optical Landsat 8 data is used [34,35].

At the Malawi site, we also tested if exploiting overlaps of two S-2 orbits delivers higher forest change detection accuracies. Using two orbits has the advantage of providing a much denser time series, but a potential disadvantage stems from the geometric shifts related to the different orbit viewing angles. Especially, the forest border is often slightly misplaced in data from two orbits as the DEM that is used by the European Space Agency (ESA) to orthorectify the imagery does not accurately account for tree height. At the 10 m spatial resolution of Sentinel-2, the effect of orbit-related geometric shifts on the spectral reflectance of single forest border pixels is much stronger than that, for example, in 30 m

Landsat 8 pixels. Elevation errors in the DEM may also lead to strong dislocations of up to 20 m at mountain ridges as observed in the alpine test areas in Austria. In Malawi, our test site is characterized by flat to slightly hilly terrain, so errors related to topography are negligible. The results show that for Malawi, the users' accuracy increases from 63.5% to 71.1% when using two orbits instead of only using one orbit. Especially, the commission error is much higher with only one orbit (see mapped areas of 1407 ha versus 756 ha in Tables 3 and 4, respectively). The main reason for this strong improvement may be found by looking at the inter-annual temporal distribution of valid observations. The time series density at the Malawi test site is quite inhomogeneous (see Figure 6). While many clear observations are available during the dry season, the time series is sparse in the rainy season from November to April due to frequent cloud cover. Using two orbits greatly increases the chance of including a few clear observations during the rainy season, which is very important for a proper initialization of the time series model. To our knowledge, such a comparative analysis on using Sentinel-2 data from one or two orbits as input to a change detection approach has not been performed before. Thus, we cannot assess whether our findings are in line with other research results. While our findings suggest that denser time series resulting from orbit overlaps lead to higher forest change mapping accuracies, it is not yet clear if these findings for the dry forests in Malawi are also valid for other forest types and for other regions. Further studies with different data input scenarios and at various forest test sites are still needed to fully answer this question.

The presented approach for detecting forest changes in near real-time using Sentinel-2 imagery has a high potential for operationalization of forest monitoring services, such as improved and automated REDD+ (Reducing Emissions from Deforestation and Forest Degradation) services in the tropics, and windthrow damage assessment or bark beetle monitoring in Central Europe. Future studies to improve the presented forest change detection method and similar near real-time approaches should focus on an integrated and automated separation of different biotic and abiotic forest disturbance agents. Additional developments should include a multivariate analysis of relevant reflectance bands and a spatiotemporal analysis of changes. The near real-time capability should be tested in operational scenarios and the time lag of change detection needs to be assessed with field data or EO data of very high geometric and temporal resolution (e.g., Planet data). For tropical forest monitoring and REDD+ activities, the method should be tested in combination with recent Sentinel-1 SAR forest change detection approaches, such as SAR shadow detection [38], backscatter composite differencing [39], or SAR backscatter thresholding based on the coefficient of variation [40]. With minor adaptations, the presented time series analysis approach might also be directly applicable to SAR data. The approach could also be used for other EO-based applications, such as phenology monitoring in agriculture or grassland monitoring (detection of mowing events).

## 5. Conclusions

In this paper, we presented a new algorithm designed for vegetation monitoring and change detection using optical EO data. The approach is largely data-driven and designed especially for near real-time mapping in web- or cloud-based monitoring services. Compared to existing algorithms employing harmonic regression, this study explored methodological improvements, mainly in two aspects:

- Seasonal patterns related to plant phenology can vary strongly between years because of different climate conditions, such as temperature and rainfall. With widely used harmonic regression models, it can be difficult to separate these normal variations from true disturbances due to the deterministic modelling approach. Our algorithm uses structural time series models in state space form which take into account that the trend and seasonal components in a time series can evolve over time. This class of stochastic models is typically used in conjunction with the Kalman filter, which also enables elegant handling of irregular sampling intervals and signal artefacts like un-masked clouds and cloud shadows. We showed that it is possible to track the phenology-related vegetation dynamics more closely without losing the ability to detect disturbances.

- Many existing algorithms require that specific change probabilities or statistical boundaries have to be exceeded multiple times for a change to be confirmed. In our algorithm, the sequence of Kalman filter innovations (i.e., the differences between the one-step-ahead model forecasts and corresponding observations) opens up alternative possibilities to characterize change. If the time series model assumptions are correct and no changes occur, the innovation sequence should be mostly free of temporal autocorrelation and it should have zero mean. We used the CUSUM control chart to monitor these properties and assumed that significant deviations indicate change. Additionally, a separate anomaly test intended to suppress commission errors due to signal artefacts was implemented. Compared to the multiple confirmation approach, our results indicate that the CUSUM approach has the potential to reduce the time lag between change event and detection and a better chance of detecting subtle and gradual changes because of its cumulative nature. However, a better understanding of different change types and especially reference data for validation are required to support this claim.

Two challenging test sites located in Austria and Malawi were selected to test the algorithm in a forest change detection scenario based on Sentinel-2 data. Both sites show pronounced and dynamic seasonal patterns in the Sentinel-2 time series due to plant phenology. The dominant change types in Malawi are deforestation and forest degradation. In Austria, we were mainly interested in changes caused by storm damages or bark beetle infestations, but of course, deforestation also occurs. The validation of the results was performed based on visually interpreted points derived by a stratified random sampling approach. For the forest change class, we reported users' accuracies of 76.8% (Austria) to 88.6% (Malawi), and producers' accuracies of 68.2% (Malawi) to 80.4% (Austria). Due to the low rate of commission errors and large proportion of stable forest, overall accuracies reached over 99%.

In the Malawi site, we further showed that a denser time series with data from two different orbits results in better change detection results compared to using data from only one orbit. The larger number of input images seems to outweigh the possible negative effects of spectral and geometric differences related to the different viewing angles, which especially occur at forest edges. The observed increase in users' accuracy when using two orbits amounted to 7.6%. However, further studies with different data inputs and at various forest test sites are required to confirm these results.

In summary, it can be stated that the combination of structural time series models and Kalman filtering represents an appropriate method for a variety of automated forest monitoring applications. Beside the possibility of detecting abrupt changes, for example caused by storm damages or deforestation, the method also shows a high potential to detect more subtle and continuous changes, such as bark beetle infestations, forest degradation, or drought stress. Regarding the detection of insect infestations, the interesting research question to be dealt with in future is whether an early detection of bark beetle green attack is feasible.

**Author Contributions:** Conceptualization, M.P., J.D., and M.H.; methodology, M.P.; software, A.W.; formal analysis, M.P., J.D., A.W., and U.S.; investigation, M.P. and A.W.; data curation, M.P. and A.W.; writing—original draft preparation, M.P., J.D., and M.H.; writing—review and editing, M.S.; visualization, M.P. and M.H.; funding acquisition, M.H., J.D., and M.S. All authors have read and agreed to the published version of the manuscript.

**Funding:** This research was funded by the European Union's Horizon 2020 research and innovation programme, grant numbers 633464 (project DIABOLO) and 685761 (project EOMonDis), and the Austrian Research Promotion Agency (FFG), grant numbers 859764 (project AlpMon) and 878891 (project BEAT IT!).

**Acknowledgments:** We greatly appreciate the free, full, and open data policy adopted by ESA for the COPERNICUS programme. We would also like to thank the reviewers for providing valuable feedback which helped to improve the manuscript.

**Conflicts of Interest:** The authors declare no conflict of interest. The funders had no role in the design of the study; in the collection, analyses, or interpretation of data; in the writing of the manuscript, or in the decision to publish the results.

## References

- Zhu, Z.; Woodcock, C.E. Object-based cloud and cloud shadow detection in Landsat imagery. *Remote Sens. Environ.* **2012**, *118*, 83–94. [[CrossRef](#)]
- Zhu, Z. Change detection using Landsat time series: A review of frequencies, preprocessing, algorithms, and applications. *ISPRS J. Photogramm. Remote Sens.* **2017**, *130*, 370–384. [[CrossRef](#)]
- Huang, C.; Goward, S.N.; Masek, J.G.; Thomas, N.; Zhu, Z.; Vogelmann, J.E. An automated approach for reconstructing recent forest disturbance history using dense Landsat time series stacks. *Remote Sens. Environ.* **2010**, *114*, 183–198. [[CrossRef](#)]
- Kennedy, R.E.; Yang, Z.; Cohen, W.B. Detecting trends in forest disturbance and recovery using yearly Landsat time series: 1. LandTrendr—Temporal segmentation algorithms. *Remote Sens. Environ.* **2010**, *114*, 2897–2910. [[CrossRef](#)]
- Verbesselt, J.; Hyndman, R.; Zeileis, A.; Culvenor, D. Phenological change detection while accounting for abrupt and gradual trends in satellite image time series. *Remote Sens. Environ.* **2010**, *114*, 2970–2980. [[CrossRef](#)]
- Verbesselt, J.; Herold, M.; Zeileis, A. Near real-time disturbance detection using satellite image time series. *Remote Sens. Environ.* **2012**, *123*, 98–108. [[CrossRef](#)]
- DeVries, B.; Verbesselt, J.; Kooistra, L.; Herold, M. Robust monitoring of small-scale forest disturbances in a tropical montane forest using Landsat time series. *Remote Sens. Environ.* **2015**, *161*, 107–121. [[CrossRef](#)]
- DeVries, B.; Pratihast, A.K.; Verbesselt, J.; Kooistra, L.; Herold, M. Characterizing forest change using community-based monitoring data and Landsat time series. *PLoS ONE* **2016**, *11*, e0147121. [[CrossRef](#)]
- DeVries, B.; Decuyper, M.; Verbesselt, J.; Zeileis, A.; Herold, M.; Joseph, S. Tracking disturbance-regrowth dynamics in tropical forests using structural change detection and Landsat time series. *Remote Sens. Environ.* **2015**, *169*, 320–334. [[CrossRef](#)]
- Hamunyela, E.; Verbesselt, J.; Herold, M. Using spatial context to improve early detection of deforestation from Landsat time series. *Remote Sens. Environ.* **2016**, *172*, 126–138. [[CrossRef](#)]
- Zhu, Z.; Woodcock, C.E.; Olofsson, P. Continuous monitoring of forest disturbance using all available Landsat imagery. *Remote Sens. Environ.* **2012**, *122*, 75–91. [[CrossRef](#)]
- Zhu, Z.; Woodcock, C.E. Continuous change detection and classification of land cover using all available Landsat data. *Remote Sens. Environ.* **2014**, *144*, 152–171. [[CrossRef](#)]
- Zhu, Z.; Woodcock, C.E.; Holden, C.; Yang, Z. Generating synthetic Landsat images based on all available Landsat data: Predicting Landsat surface reflectance at any given time. *Remote Sens. Environ.* **2015**, *162*, 67–83. [[CrossRef](#)]
- Brooks, E.B.; Wynne, R.H.; Thomas, V.A.; Blinn, C.E.; Coulston, J.W. On-the-fly massively multitemporal change detection using statistical quality control charts and Landsat data. *IEEE Trans. Geosci. Remote Sens.* **2014**, *52*, 3316–3332. [[CrossRef](#)]
- Harvey, A.C. *Forecasting, Structural Time Series Models and the Kalman Filter*; Cambridge University Press: Cambridge, UK, 1989; ISBN 0-521-40573-4.
- Kalman, R.E. A New Approach to Linear Filtering and Prediction Problems. *J. Basic Eng.* **1960**, *82*, 35–45. [[CrossRef](#)]
- Gibbs, B.P. *Advanced Kalman Filtering, Least-Squares and Modeling*; John Wiley & Sons Inc.: Hoboken, NJ, USA, 2011; ISBN 9780470890042.
- Puhm, M. Analysis of Landsat time series using state space models and Kalman filtering. Master's Thesis, Graz University of Technology, Graz, Austria, 2018. Available online: <https://permalink.obvsg.at/tug/AC15229314> (accessed on 14 July 2020).
- Olsson, P.O.; Lindström, J.; Eklundh, L. Near real-time monitoring of insect induced defoliation in subalpine birch forests with MODIS derived NDVI. *Remote Sens. Environ.* **2016**, *181*, 42–53. [[CrossRef](#)]
- Page, E.S. Continuous Inspection Schemes. *Biometrika* **1954**, *41*, 100. [[CrossRef](#)]
- Hirschmugl, M.; Deutscher, J.; Sobe, C.; Bouvet, A.; Mermoz, S.; Schardt, M. Use of SAR and optical time series for tropical forest disturbance mapping. *Remote Sens.* **2020**, *12*, 727. [[CrossRef](#)]
- S2 MPC Team. L1C Data Quality Report. 2020. Available online: [https://sentinel.esa.int/documents/247904/685211/Sentinel-2\\_L1C\\_Data\\_Quality\\_Report](https://sentinel.esa.int/documents/247904/685211/Sentinel-2_L1C_Data_Quality_Report) (accessed on 18 September 2020).

23. 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, Prague, Czech Republic, 9–13 May 2016; pp. 1–8.
24. Qiu, S.; He, B.; Zhu, Z.; Liao, Z.; Quan, X. Improving Fmask cloud and cloud shadow detection in mountainous area for Landsats 4–8 images. *Remote Sens. Environ.* **2017**, *199*, 107–119. [[CrossRef](#)]
25. Qiu, S.; Zhu, Z.; He, B. Fmask 4.0: Improved cloud and cloud shadow detection in Landsats 4–8 and Sentinel-2 imagery. *Remote Sens. Environ.* **2019**, *231*, 111205. [[CrossRef](#)]
26. Gallaun, H.; Schardt, M.; Linser, S. Remote sensing based forest map of Austria and derived environmental indicators. In Proceedings of the International Conference on Spatial Application Tools in Forestry (ForestSAT 2007), Montpellier, France, 5–7 November 2007.
27. Huber, P.J. Robust estimation of a location parameter. *Ann. Math. Stat.* **1964**, *35*, 73–101. [[CrossRef](#)]
28. Heiberger, R.M.; Becker, R.A. Design of an S Function for Robust Regression Using Iteratively Reweighted Least Squares. *J. Comput. Graph. Stat.* **1992**, *1*, 181–196. [[CrossRef](#)]
29. Zhu, Z.; Zhang, J.; Yang, Z.; Aljaddani, A.H.; Cohen, W.B.; Qiu, S.; Zhou, C. Continuous monitoring of land disturbance based on Landsat time series. *Remote Sens. Environ.* **2020**, *238*, 111116. [[CrossRef](#)]
30. Xue, J.; Su, B. Significant Remote Sensing Vegetation Indices: A Review of Developments and Applications. *J. Sensors* **2017**, *2017*, 1–17. [[CrossRef](#)]
31. Olofsson, P.; Foody, G.M.; Stehman, S.V.; Woodcock, C.E. Making better use of accuracy data in land change studies: Estimating accuracy and area and quantifying uncertainty using stratified estimation. *Remote Sens. Environ.* **2013**, *129*, 122–131. [[CrossRef](#)]
32. Olofsson, P.; Foody, G.M.; Herold, M.; Stehman, S.V.; Woodcock, C.E.; Wulder, M.A. Good practices for estimating area and assessing accuracy of land change. *Remote Sens. Environ.* **2014**, *148*, 42–57. [[CrossRef](#)]
33. Potapov, P.V.; Dempewolf, J.; Talero, Y.; Hansen, M.C.; Stehman, S.V.; Vargas, C.; Rojas, E.J.; Castillo, D.; Mendoza, E.; Calderón, A.; et al. National satellite-based humid tropical forest change assessment in Peru in support of REDD + implementation. *Environ. Res. Lett.* **2014**, *9*. [[CrossRef](#)]
34. Reiche, J.; Hamunyela, E.; Verbesselt, J.; Hoekman, D.; Herold, M. Improving near-real time deforestation monitoring in tropical dry forests by combining dense Sentinel-1 time series with Landsat and ALOS-2 PALSAR-2. *Remote Sens. Environ.* **2018**, *204*, 147–161. [[CrossRef](#)]
35. Shimizu, K.; Ota, T.; Mizoue, N. Detecting forest changes using dense Landsat 8 and Sentinel-1 time series data in tropical seasonal forests. *Remote Sens.* **2019**, *11*, 1899. [[CrossRef](#)]
36. Hansen, M.C.; Krylov, A.; Tyukavina, A.; Potapov, P.V.; Turubanova, S.; Zutta, B.; Ifo, S.; Margono, B.; Stolle, F.; Moore, R. Humid tropical forest disturbance alerts using Landsat data. *Environ. Res. Lett.* **2016**, *11*, 034008. [[CrossRef](#)]
37. 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](#)]
38. Bouvet, A.; Mermoz, S.; Ballère, M.; Koleck, T.; Le Toan, T. Use of the SAR shadowing effect for deforestation detection with Sentinel-1 time series. *Remote Sens.* **2018**, *10*, 1250. [[CrossRef](#)]
39. Rüetschi, M.; Small, D.; Waser, L.T. Rapid detection of windthrows using Sentinel-1 C-band SAR data. *Remote Sens.* **2019**, *11*, 115. [[CrossRef](#)]
40. Deutscher, J.; Gutjahr, K.; Perko, R.; Raggam, H.; Hirschmugl, M.; Schardt, M. Humid tropical forest monitoring with multi-temporal L-C and X-band SAR data. In Proceedings of the 9th International Workshop on the Analysis of Multitemporal Remote Sensing Images (MultiTemp), Brugge, Belgium, 27–29 June 2017; pp. 1–4.

