Martin Puhm, BSc

Analysis of Landsat Time Series Using State Space Models and Kalman Filtering

Application to change detection in forest areas

Master's Thesis

to achieve the university degree of

Diplom-Ingenieur

Master's Degree Programme: Geomatics Science

submitted to

Graz University of Technology

Supervisor

Univ.-Prof. Dipl.-Forstw. Dr. Mathias Schardt
Institute of Geodesy

Graz, November 2018

Statutory Declaration

I declare that I have authored this thesis independently, that I have not used other than the declared sources/resources, and that I have explicitly indicated all material which has been quoted either literally or by content from the sources used. The document uploaded to TUGRAZonline is identical to the present thesis.

Eidesstattliche Erklärung

Ich erkläre an Eides statt, dass ich die vorliegende Arbeit selbstständig verfasst, andere als die angegebenen Quellen/Hilfsmittel nicht benutzt, und die den benutzten Quellen wörtlich und inhaltlich entnommenen Stellen als solche kenntlich gemacht habe. Das in TUGRAZonline hochgeladene Dokument ist mit der vorliegenden Arbeit identisch.

Date/Datum	Signature/Unterschrift

Abstract

Current Earth observation missions employing spaceborne optical sensors acquire a vast data volume never available before. The American Landsat and the European Sentinel-2 mission are of special interest. Their satellite design represents a good compromise between relatively high geometric resolution and short time intervals between images of the same surface region. Through high-quality georeferencing of the satellite images it is possible to create a time series of measured gray values for a given spectral band at pixel-level. This thesis explores a novel approach to detect forest disturbances based on these time series. Similar to some existing methods, the normal temporal signature of a given spectral band over the course of one year is captured by a time series model. For example, the phenological cycle typically encountered in forests can be approximated by a sum of trigonometric functions of different frequencies and a trend component. Consequently, newly available observations may be compared to a model-based forecast. Abrupt changes of the spectral signature, possibly linked to a forest disturbance, are indicated by statistically significant deviations between new observations and the forecast. Regression models together with robust least-squares techniques to estimate their parameters are widely used. In contrast, this thesis investigates the applicability of time series models formulated in state space form in conjunction with the Kalman filter. Two distinct advantages of the Kalman filter approach include the fact that more recent observations have a larger weight on the forecast and that the time series model is dynamically updated. Aside from a few user-defined tuning parameters, the proposed change detection algorithm is data-driven. A multi-temporal stack of Landsat surface reflectance data is used to implement a test case regarding storm damage detection. Three features obtained by applying the tasseled cap transformation, namely brightness, greenness, and wetness, have been chosen as observables. The test site is located in Baden-Württemberg, Germany. Three different change maps, each based on one of the observables, have been produced. A limited quantitative evaluation of the change maps based on a ground-truth data set describing windthrow areas after a storm in 2012 has been carried out. The best results are yielded by the wetness change map, where changed pixels were classified with a producer accuracy of up to 80.6% and a user accuracy of up to 86.1%.

Kurzfassung

Aktuelle Missionen zur Erdbeobachtung mittels satellitengestützter optischer Sensoren, insbesondere das amerikanische Landsat- und das europäische Sentinel-2-Programm, liefern eine noch nie da gewesene Datenmenge. Hinsichtlich ihrer Auslegung bieten sie einen guten Kompromiss zwischen relativ hoher geometrischer Auflösung und möglichst geringem zeitlichen Abstand zwischen aufeinanderfolgenden Aufnahmen desselben Gebietes. Eine genaue Georeferenzierung der Satellitenbilder ermöglicht es, die in verschiedenen Spektralbereichen gemessenen Grauwerte auf Pixelebene als Zeitreihen darzustellen. Die vorliegende Diplomarbeit beschäftigt sich mit der Detektion von Waldschäden mithilfe dieser Zeitreihen. Wie bei bereits existierenden Algorithmen wird der Ansatz verfolgt, den normalen Jahresverlauf einer Spektralsignatur mit einem Zeitreihenmodell zu erfassen. Die für Wald typischen Phänologiekurven können z.B. mit einer Summe aus trigonometrischen Funktionen verschiedener Frequenz sowie einer Trendkomponente approximiert werden. In weiterer Folge können neu verfügbare Beobachtungen mit einer Vorhersage auf Basis des Modells verglichen werden. Abrupte, z.B. von Sturmschäden verursachte Änderungen der Spektralsignatur werden durch statistisch signifikante Abweichungen zwischen Modellvorhersage und Beobachtung signalisiert. Weit verbreitet ist dabei die Verwendung von Regressionsmodellen, deren Parameter mithilfe von robusten Least-Squares-Verfahren geschätzt werden. Im Gegensatz dazu werden in dieser Arbeit State-Space-Modelle und Kalman-Filterung eingesetzt. Der Kalman-Filter bietet den Vorteil, dass neuere Beobachtungen in der Vorhersage stärker gewichtet und das Zeitreihenmodell dynamisch angepasst wird. Der entwickelte Algorithmus arbeitet nach der Definition von wenigen Grundparametern vollautomatisch. Er wurde auf eine historische Landsat-Zeitreihe für eine Sturmwurfdetektion im Schwarzwald angewendet, wobei die Komponenten der Tasseled-Cap-Transformation (Brightness, Greenness, Wetness) als Beobachtungen dienten. Als Ergebnis wird für jedes prozessierte Band eine Karte ausgegeben, in der detektierte Änderungen und das Datum der Signalisierung verzeichnet sind. Ein Vergleich mit Referenzdaten eines Sturmwurfs im Jahr 2012 ergab die höchste Übereinstimmung mit der Karte der Wetness-Änderungen, wobei die Sturmwurfflächen mit einer Produzentengenauigkeit von bis zu 80,6% sowie einer Anwendergenauigkeit von bis zu 86,1% erfasst wurden.

Contents

Acknowledgements

First I would like to express my appreciation to my thesis supervisor Professor Mathias Schardt, whose encouragement, advice, and valuable feedback helped a great deal to finish this work. I am especially thankful to him for giving me the opportunity to be a part of the remote sensing team at Joanneum Research. Of course my sincere thanks go to all my colleagues who pointed me in the right direction whenever I got lost in the “data jungle” and gave me the feeling that my work is actually useful. Finally I would like to extend my deepest gratitude to my parents who gave me the support to pursue this goal. Without their patience I would not be where I am today.

Martin Puhm
Graz, November 2018

Credits

This thesis has received funding from the European Union's Horizon 2020 research and innovation program under grant agreement no. 633 464 (DIABOLO).

Landsat imagery courtesy of the U.S. Geological Survey.

A reference map of damages in the Black Forest caused by a storm in 2012 has been used with kind permission of the Chair of Remote Sensing and Landscape Information Systems, University of Freiburg.

1 Introduction

Within the field of remote sensing, the topic of this thesis can be categorized into the area of forest monitoring and change detection with the help of spaceborne optical sensors. The presence of the keywords *monitoring* and *change detection* almost automatically implies that multi-temporal data and time series analysis methods are involved. The first of the following sections highlights how this thesis is embedded in the remote sensing and time series analysis context. Its objectives are outlined in the second section and the introduction concludes with an overview of the thesis' structure.

1.1 Background

High-resolution images of the Earth's surface acquired by satellites may be considered as a vast historical archive. The various satellites of the Landsat program have added valuable data to this archive since 1972, and since 2008 all Landsat data is publicly available (Wulder et al., 2016). The long history and relatively high geometric resolution of 30 meters makes Landsat images very attractive for the purpose of land-cover monitoring, because many man-made changes of the Earth's surface can be captured with this kind of resolution, for example agricultural units like fields. With the launch of Landsat 8, the newest of the currently operational Landsat satellites, the rate at which new images of the same surface region can be acquired has also increased. Single images can be joined to a sequence, and if this sequence is sorted by time it is called a time series.

The extraction of as many information as possible out of these time series is an active field of research within the remote sensing community. Another incentive for further development has been offered when Sentinel-2A was launched successfully in June 2015. The Sentinel-2 satellite mission provides data with a spatial resolution of either 10, 20, or 60 meters depending on the spectral band. Once fully operational, the mission features two identical satellites which acquire images from opposite sides of the same orbit in order to reduce the time interval between consecutive images of the same surface area. Indeed Sentinel-2B was launched in March 2017 (Wikipedia, 2018).

There already are several operational change-detection algorithms based on time series analysis. However, the concept of Kalman filtering has been rarely applied in remote sensing contexts so far, but the technique is well established in many other fields with no apparent connection. Kalman filtering denotes a parameter estimation technique which yields optimal estimates in a statistical sense. In general the Kalman filter approach requires the definition of a dynamic model and an observation model. The observation model on the one hand defines the relationship of the measurements to a set of state variables which cannot be observed directly. For a time series, the state variables usually represent the series' additive decomposition into trend, seasonal, and long-term cyclical components. The dynamic model on the other hand describes the

expected temporal evolution of the state variables. The Kalman filter framework has some interesting properties regarding its application to remote sensing time series:

- The filter operates recursively in two steps:

  a) In the time-update step, the states' evolution from one point in time to the next is *predicted* based on the dynamic model.

  b) In the measurement-update step, the predicted state estimate is enhanced by incorporating newly available observations.

  Therefore, all the information of past observations is present in the current state estimate and does not have to be kept in memory.

- The filter can handle uneven temporal intervals between measurements if the dynamic model is formulated in continuous time. This way, the problem of gaps in the time series due to masked clouds and their shadows is addressed. For example, the seasonal phenology cycle typically encountered in forest can be modeled using trigonometric functions.

- The noise in the filtered time series is reduced.

- The filter predictions can be used to identify abrupt structural change in a time series.

1.2 Objectives and structure of the thesis

The first objective is to give an account of the current state of the art approaches to change-detection based on time series analysis. Along with an emphasis on applications using Landsat data, methods capable of modeling normal intra-annual variations are of special interest. The primary objective of this thesis was to link these current approaches to Kalman filtering techniques commonly found in other disciplines and investigate their applicability for change-detection in forests. Key requirements for the intended Kalman filter algorithm include robustness against un-masked clouds and shadows, sensitivity to forest damages like windthrow as well as a high level of automation. The third and final objective is to test the proposed algorithm using a historic Landsat time series and evaluate its performance.

The thesis is organized in 6 Chapters. Following the introductory Chapter 1, Chapter 2 reviews some of the complexities encountered when working with remotely-sensed time series and summarizes the working principles of several existing and operational algorithms. Chapter 3 deals with the mathematical background and implementation details of a Kalman-filter-based change-detection algorithm. Chapter 4 describes the used data set and the results of the field test are presented and discussed in Chapter 5. Concluding remarks and suggestions for future work are given in Chapter 6.

2 State of the Art

This chapter presents a review of change detection methods based on per-pixel time series analysis that have been published in the recent years. The focus lies especially on algorithms capable of processing *dense* time series at spatial resolutions of *optical* sensors like the Landsat family or Sentinel-2. In this context, a time series is considered dense if several valid observations are available within a specific time period. Before discussing most common existing methods, a summary of the challenges which a change detection algorithm has to address is given.

2.1 Properties of dense time series acquired with optical sensors

Dense time series acquired with high-resolution optical sensors have a number of properties which make them challenging to work with. Additional to noise caused by different atmospheric conditions and uncertainties in the geometric registration process, they include:

- Seasonality
- Unequal temporal intervals between observations
- Missing observations
- Presence of invalid observations

Each property and its source as well as how it influences the ability to detect changes are discussed in the respective subsections below. Note that the distinction between irregularly spaced and missing observations is important. Consider a sequence of measurements sampled at a constant interval, but with occasionally missing values. The time interval between consecutive observations will always be an integer multiple of the basic sampling interval. This restriction does not apply to truly irregular time series.

2.1.1 Seasonality

Seasonal patterns in remotely sensed time series are primarily caused by the annual variations of temperature and rainfall that influence plant phenology. Effects of the Bidirectional Reflectance Distribution Function (BRDF) can also add to the seasonality (Zhu et al., 2012). An analysis of different seasonal patterns for a number of land-cover/land-use classes based on MODIS NDVI time series is given by Geerken (2009). Fourier analysis was applied to the data in order to find the dominant frequencies in the annual cycle. The results indicate that most of the seasonal variation of many land cover types can be modeled by using 3 harmonics corresponding to a frequency of 1, 2, and 3 periods per year, without overfitting the noise. Dynamic land cover types like cropland show larger amplitudes at the higher frequencies whereas forest types have the highest amplitude at a frequency of 1 period per year. Hence, time series with increased density and well distributed observations over all seasons are required to describe the spectral dynamics of different land cover types over time.

An algorithm for change detection has to be able to distinguish between normal seasonal changes and abnormal behavior.

2.1.2 Unequal temporal interval between observations

Satellites with a regular nadir acquisition scheme usually have a constant revisit cycle, for example 16 days for Landsat 8, 10 days for Sentinel-2A or 5 days if data from Sentinel-2A and -2B are combined. However, with increasing latitude adjacent paths overlap and this can be used to increase the density of the time series in certain areas at the cost of irregularly spaced observations. The revisit time could also be increased by integrating data from Landsat and Sentinel-2, which is deemed possible by Wulder et al. (2016). Because the Sentinel-2 satellites have a different repeat cycle, this would also lead to irregularly spaced observations. However, many methods used in time series analysis require a constant sampling interval. In order to use these methods, an additional pre-processing step like filtering or interpolation becomes necessary in order to create a regularly-spaced time series. A review and comparison of some existing interpolation methods is given by Kandasamy et al. (2013). However, this pre-processing may introduce other undesired effects like smoothing abrupt signal jumps possibly corresponding to a change event. At spatial resolutions of Landsat or higher, the additional computational effort might also be significant. Depending on the application, the use of more sophisticated time series analysis methods which can handle irregularly spaced observations would be preferable.

2.1.3 Missing observations

Clouds and cloud shadows as well as snow greatly influence the reflectance of different spectral bands (Dozier, 1989; Irish et al., 2006). Therefore they have to be considered as noise in the data and must be masked, resulting in missing observations in the time series. Not all methods used in time series analysis are designed to account for missing values. Analogous to the reasons pointed out in Section 2.1.2, a decision between additional pre-processing to fill the gaps or more powerful time series analysis methods capable of handling missing observations must be made. Another aspect of the problem regards the amount of cloud and snow cover that has to be expected for certain areas on Earth. In the tropical climate zone or northern parts of Europe for example, the cloud cover probability is very high (Wylie and Menzel, 1999) and it is therefore difficult to obtain dense time series in the first place.

2.1.4 Presence of invalid observations

Although there are some powerful algorithms for cloud, cloud shadow and snow screening, none of them is absolutely accurate. As a consequence, the presence of outliers in the time series must be considered. A reliable change detection method has to incorporate a mechanism to avoid confusion between outliers and real changes.

2.2 Existing methods and algorithms

A number of studies published recently aimed at assessing the current state of change detection methods based on high resolution optical earth observation data. A review by Hirschmugl et al. (2017) focused on the mapping of forest disturbances and degradation. They concluded "*that there are already many methods available for bi-temporal change detection from high-resolution data on the one hand and for time series analysis from coarse resolution data on the other. The current main challenge and research development focus is transferring these approaches to high resolution time series.*"

Without a limitation to forest applications, Zhu (2017) reviewed change detection studies based on Landsat time series. Although the number of studies is quite high, the author observed that "*most of the time series studies were only interested in producing annual or biannual change maps*" and therefore "*tended to select multiple images* (*partly cloudy images*) *acquired in the same season and the same year to produce cloud-free composite images as their inputs. In this way, seasonal differences caused by solar angle differences and vegetation phenological changes were minimized, and the data volume was reduced substantially*". The author also noted that newer studies are more likely to use a higher amount of images per year acquired in different seasons.

The conclusions from both reviews as well as an independently conducted literature search show that there are few algorithms for change detection which have been applied to Landsat time series with a high temporal resolution. Many older, established methods like the vegetation change tracker (VCT) proposed by Huang et al. (2010) or LandTrendr (Kennedy et al., 2010) circumvent the problems outlined in Section 2.1 by using annual time series of cloud-free composites. Algorithms capable of modeling the intra-annual seasonal changes as well as commonly applied pre-processing steps for Landsat data are presented in the following sub-sections.

2.2.1 Pre-processing

The literature search revealed that there are many pre-processing steps commonly applied to Landsat data in order to use them as input to time series analysis methods. Only images conforming to high-level requirements with respect to radiometric and atmospheric calibration as well as geometric registration are suitable for applications on a pixel-level. These requirements are met by Landsat Surface Reflectance (SR) products with processing level L1T. In case of Landsat 4-7 data, the conversion to SR is based on the Landsat Ecosystem Disturbance Adaptive Processing System (LEDAPS) (Masek et al., 2006). Landsat 8 data is processed to SR using the Provisional Landsat 8 Surface Reflectance Code (LaSRC) (Vermote et al., 2016). To mask clouds, cloud shadows or snow, the Fmask algorithm (Zhu and Woodcock, 2012) is commonly used. High-quality geometric registration is achieved by using ground control points and a digital elevation model (USGS, 2018). The temporal consistency of the geometric registration can be increased even further by employing matching algorithms to a stack of images.

2.2.2 BFAST

The Breaks for Additive Season and Trend (BFAST) algorithm was originally designed to detect trend and seasonal changes in MODIS 16-day NDVI composites (Verbesselt et al., 2010a). It is based on the decomposition of a time series into a trend, seasonal, and remainder component. Change detection is performed by finding breakpoints in either of the estimated components. Through this approach the number, timing, and type of changes in historic time series can be determined. A second version employed a different seasonal model based on harmonic functions which the authors deemed to be "*more suitable and robust for phenological change detection with satellite image time series*" (Verbesselt et al., 2010b). The next stage in development was BFAST Monitor, "*a multi-purpose time-series-based disturbance detection approach that identifies and models stable historical variation to enable change detection within newly acquired data*" (Verbesselt et al., 2012). Initially it was also used with MODIS data, but a variety of more recent studies listed in Table 1 demonstrate that BFAST and BFAST Monitor can also be applied to Landsat time series to detect both abrupt and gradual change.

Depending on the individual application and test site, the authors adapted the algorithm to their needs. In conclusion, the BFAST algorithms present rather a framework than an "out-of-the-box" solution, but the components are flexible and can be adjusted to different requirements. It can handle irregularly spaced and missing observations. According to DeVries et al. (2015b), it is also robust to occasional outliers, but temporally aggregated occurrences such as several consecutively un-masked clouds can be a source of error. Additional pre-processing to eliminate the effects of un-masked clouds or cloud shadows was applied by DeVries et al. (2016, 2015a) and Hamunyela et al. (2016).

Table 1: Studies using BFAST or BFAST Monitor with Landsat time series

Goal and observables	Test sites	Reference
Monitoring of forest cover loss, NDVI, combination with MODIS or rainfall data	Bolivia, tropical forest	(Dutrieux et al., 2015)
Forest monitoring in regions with persistent cloud cover, fusion of Landsat NDVI and SAR	Fiji, tropical forest	(Reiche et al., 2015)
Tracking of disturbance-regrowth dynamics using all available Landsat data, NDMI	Southern Peru, tropical forest	(DeVries et al., 2015a)
Monitoring of small-scale forest disturbances, NDVI	Southern Ethiopia	(DeVries et al., 2015b)
Deforestation mapping, sNDVI (spatially normalized NDVI, reduced seasonality by using spatial context)	Humid/dry forest, Brazil/Bolivia	(Hamunyela et al., 2016)
Mapping of deforestation and degradation, various spectral bands and indices, focus on correct change classification	Southern Ethiopia	(DeVries et al., 2016)
Evaluation of how specific effects of site and radiometric correction affect the accuracy of deforestation monitoring when using BFAST Monitor	Brazil, Ethiopia, Vietnam	(Schultz et al., 2016)
Forest disturbance detection and change agent attribution (windthrow, cleared windthrow, bark beetles, and other harvest), tasseled cap wetness	Bohemian forest, Kalkalpen, Tatra	(Oeser et al., 2017)

2.2.3 CCDC

Originally, Zhu et al. (2012) developed an algorithm called Continuous Monitoring of Forest Disturbance Algorithm (CMFDA). "*Using all the available Landsat ETM+ images in two years, models using sines and cosines are fit for each pixel and each spectral band. These models can predict Landsat images at any date assuming there is not any land cover change.*" Change detection is performed by differencing the predicted and a newly acquired image on pixel-level. If the calculated difference of a change index sensitive to forest disturbance crosses a certain threshold for 3 consecutive times, a pixel is flagged as changed. The CMFDA algorithm also incorporates two-stage cloud screening where previously un-masked clouds, shadows, and snow can be detected by using multi-temporal data. This approach was later refined and published separately as "*a new algorithm ... for automated masking of cloud, cloud shadow, and snow for multi-temporal Landsat images*" (Zhu and Woodcock, 2014a).

The change detection part of CMFDA has been further developed to the Continuous Change Detection and Classification (CCDC) algorithm where the concept is extended to include more types of land cover beside forest as well as a classification framework (Zhu and Woodcock, 2014b). From the beginning, CCDC was designed to work with dense Landsat time series and can therefore handle seasonality, missing or irreg-

ularly spaced observations, and outliers to some extent. Both abrupt and gradual changes can be detected. Some further updates to the algorithm are described by Zhu et al. (2015). They include a mechanism to automatically adjust the complexity of the time series model based on the number of available clear observations as well as a different method to estimate the model parameters that reduces overfitting. Some recently published studies employing the CCDC algorithm are listed in Table 2.

Table 2: Studies using CCDC with Landsat time series

Topic	Test sites	Reference
Investigation of the possibilities for monitoring gradual changes using dense Landsat time series	Several locations, USA	(Vogelmann et al., 2016)
Analysis of vegetation greenness trends considering effects due to land cover change	Guangzhou, China	(Zhu et al., 2016)
Analysis of urbanization induced land use and land cover change	Atlanta metropolitan area, USA	(Fu and Weng, 2016)
Evaluation of CCDC for use within the USGS Land Change Monitoring, Assessment, and Projection program (LCMAP)	Several locations, USA	(Pengra et al., 2016)
Mapping of forest degradation	Lam Dong Province, Vietnam	(Vogelmann et al., 2017)
Near-real-time monitoring of insect defoliation	Southern New England, USA	(Pasquarella et al., 2017)

2.2.4 Other algorithms

Beside the more widely used algorithms discussed so far, there are also others which have been designed for analyses of dense Landsat data. To monitor and map forest disturbances, Brooks et al. (2014) presented "*a method that utilizes residuals from harmonic regression over years of Landsat data, in conjunction with statistical quality control charts, to signal subtle disturbances in vegetative cover. These charts are able to detect changes from both deforestation and subtler forest degradation and thinning. First, harmonic regression residuals are computed after fitting models to interannual training data. These residual time series are then subjected to Shewhart X-bar control charts and exponentially weighted moving average charts. The Shewhart X-bar charts are also utilized in the algorithm to generate a data-driven cloud filter, effectively removing clouds and cloud shadows on a location-specific basis. Disturbed pixels are indicated when the charts signal a deviation from data-driven control limits.*" Test sites were located in Alabama and Mississippi, USA, and the used change index was tasseled cap angle. This study deviates from others with respect to the pre-processing. Additional to the conversion to surface reflectance, a dark object subtraction using band minima was applied. "*The dark object subtraction had a significant effect in reducing time series noise, since post-LEDAPS evaluation of the reflectance data revealed that some scenes were uniformly brighter or darker than the remainder of the time series.*" The authors also did not use a

cloud mask but filtered out images with a nominal cloud cover higher than 10%. The remaining clouds and shadows were eliminated using the built-in data-driven cloud filter.

Another approach was proposed by Thonfeld et al. (2015) in a study dedicated to compare bi-temporal-, change-trajectory- and time-series-analysis-based methods for change detection. The test site was located in southern Vancouver Island, Canada. All available Landsat images with a nominal cloud-cover below 80% were used to create time series of the Normalized Difference Moisture Index (NDMI) with a temporal coverage of multiple years. Linear interpolation was performed to create regularly-spaced observations. To reduce the influence of un-masked clouds and cloud shadows, Savitzky-Golay filtering was applied. Change detection is based on the computation of a statistical measure for a moving window of 365 days. The authors hypothesize that any seasonal effects are leveled out by setting the window size to a full year. The statistical measure is used to find the most significant breakpoint in the time series, thus multiple abrupt or gradual changes are not considered in this study.

2.2.5 Limitations

All algorithms 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. Most of the algorithms handle seasonality through harmonic models based on sines and cosines, but the methods used to fit them to the data differ. Regarding the limitations of time-series-analysis-based change detection algorithms, the influence of the following interdependent parameters has to be considered:

(1) *Properties of the input data.* In order to detect a certain type of change, it has to be reflected by the input data and spectral bands or indices have to be chosen accordingly. Recently, Cohen et al. (2017) have compared forest disturbance maps produced from the output of seven different change detection algorithms. Among them were CCDC and the method proposed by Brooks et al. (2014) the remaining five only considered annual time series. One of the conclusions by Cohen et al. (2017) was that "*Spectral change magnitudes associated with forest disturbance are highly variable, with a population likely to be skewed towards lower-magnitude occurrences. Such disturbances are challenging to map because they are often difficult to distinguish from spectral noise common in temporal trajectories of spectral signals.*"

(2) *The number of frequencies used in the harmonic model.* The complexity of the seasonal pattern is dependent on the climate zone, the land cover class and the spectral band or index. To include higher frequencies, a higher number of clear observations distributed over all seasons is required. This may be a major limitation for areas with persistent cloud or snow cover. More complex models are also more sensitive to noise and outliers, hence the risk of overfitting may increase in spite of using robust estimation methods.

(3) *Settings regarding the computation of statistical boundaries.* The discussed algorithms have certain tuning parameters which control the thresholds for detecting change. To make the algorithm more robust to noise and outliers, these thresholds have to be increased. As a consequence, changes with a low magnitude cannot be detected.

3 Methods

This chapter covers the mathematical basics of the time series analysis method which has been applied in this thesis. The first section contains remarks on the nature of the underlying data and introduces some basic terms and notation. The idea of decomposing a time series into several components is discussed in the second section, followed by an introduction to stochastic processes. Based on these concepts, the structural time series models outlined in Section 3.4 present the core of this chapter. The penultimate section deals with regression models and a technique for robust parameter estimation with respect to the presence of outliers in the data. The chapter concludes with a description of the implemented change detection algorithm.

3.1 Preliminary remarks

All time series investigated within this thesis are generated by using multi-temporal stacks of Landsat images as data source. Every image of the stack is geometrically registered to a common spatial grid and provides a set of observations including multiple spectral bands as well as any indices derived from them. Therefore, the available observation data can be organized in a four-dimensional array defined according to Table 3.

Table 3: Array representation of observations

Dimension	Interpretation	Length
1	Time	m
2	Spectral bands and indices	n
3	Spatial grid rows, extent north to south	n_{rows}
4	Spatial grid columns, extent east to west	n_{cols}

For every combination of band, row and column, a sequence of m observations can be extracted, meaning that the observation array represents a total of $n \cdot n_{rows} \cdot n_{cols}$ time series. Every one of these sequences could be treated separately as *univariate* time series. However, the nature of the data also suggests a *multivariate* approach where several series, for example all bands for a given row and column, are modeled jointly in order to take correlations between them into account. With the acquisition date of an image referred to as *epoch*, the notation used throughout the following sections is as follows:

k … epoch index, $k = 1,2,\ldots m$
z_k … scalar observation made in epoch k (univariate case)
$\mathbf{z}_k$ … vector of observations made in epoch k (multivariate case)
t_k … epoch time
Δt … time interval between consecutive epochs, $\Delta t = t_k - t_{k-1}$

The software modules implemented in this thesis are limited to univariate time series models, but several bands or indices may be processed in parallel to make use of the different information content. Therefore most of the content in this chapter is centered on univariate time series. In theory though, the presented methods are also capable of supporting true multivariate models, but the complexity of the implementation would have been beyond the scope of this thesis.

3.2 Additive decomposition of a time series

A widely used starting point for the development of time series models is to consider each observation z_k as the realization of a random variable Z_k, which is itself a sum of four components:

$$Z_k = T_k + C_k + S_k + R_k, \quad k = 1,2,\dots m \tag{3.1}$$

The first two terms on the right hand side of equation (3.1) are called *trend* and *cycle* respectively. Both reflect non-random long term movements in the series and while the trend may take a variety of shapes, the cycle component is explicitly defined as a recurring pattern of growth and decline. Furthermore, S_k is called the *seasonal* component and reflects non-random short term cyclical patterns which repeat themselves more or less every year. The remaining non-systematic deviations left in the series are captured by the *irregular* component R_k, which is a random variable used to model all kinds of random influences (Falk et al., 2012, sec. 1.1).

Additive decomposition models like (3.1) *describe* the series in terms of components which cannot be observed directly. One of the advantages of this strategy is that each component of interest can be analyzed and interpreted separately. Additionally, the complexity of the model can be scaled simply by adding or removing components in order to find a model which is consistent with the data, without being unnecessarily over-specified. Once an appropriate model for a given time series has been found, it can be used to *predict* future observations (Harvey, 1989, chap. 1).

3.3 Stochastic processes

Stochastic processes are a basic element of time series analysis. This section is based on Gibbs (2011, sec. B.3) and Harvey (1989, sec. 2.3.1, 2.4.1, 9.1.1) and discusses the fundamentals needed to later define the properties of structural time series models in Section 3.4.

Let the sequence of values y_1, y_2, … y_m be a sample of the output of a univariate stochastic process. Using the expectation operator E(·), some well-known statistical quantities describing the sequence include the mean

$$\mathrm{E}(y_k) = \mu_k, \tag{3.2}$$

as well as the autocovariances

$$\mathrm{E}\left[(y_k - \mu_k)(y_{k-\ell} - \mu_k)\right] = \gamma_k(\ell), \quad \ell = 0,1,\dots \tag{3.3}$$

where $\gamma_k(0)$ equals the variance

$$\mathrm{E}\left[(y_k - \mu_k)^2\right] = \gamma_k(0) = \mathrm{Var}(y_k). \tag{3.4}$$

The subscript k indicates that a variable may vary with time. When the current value y_k only depends on the previous value y_{k-1} and a disturbance term q_k, it is an autoregressive process characterized by the equation

$$y_k = \varphi y_{k-1} + q_k, \quad k = 1,2,\ldots m, \tag{3.5}$$

where φ is a parameter. The statistical properties of the process are determined by the definition of the disturbance term and the value of φ. Here it is assumed that the disturbances are white noise, which is a sequence of serially uncorrelated random variables with a constant mean of zero and constant variance. The value of the parameter φ determines whether the process is stationary or not. If $|\varphi| < 1$, it is said to be *wide-sense stationary* and its mean and autocovariances are independent of time. Consequently, the said quantities can be estimated from a single sample (or realization) of the process. The estimators are given by the sample mean

$$\hat{\mu} = \bar{y} = \frac{1}{m}\sum_{i=k}^{m} y_k, \tag{3.6}$$

and the sample autocovariances

$$\hat{\gamma}(\ell) = c(\ell) = \frac{1}{m}\sum_{k=\ell+1}^{m} (y_k - \bar{y})(y_{k-\ell} - \bar{y}), \quad \ell = 0,1,\ldots \tag{3.7}$$

where $c(0)$ is the sample variance

$$\hat{\gamma}(0) = c(0) = \frac{1}{m}\sum_{k=1}^{m} (y_k - \bar{y})^2. \tag{3.8}$$

The definition of *strict stationarity* requires additionally that the joint probability distribution of two samples taken at different time periods remains identical.

The only non-stationary stochastic process with practical importance for the time series models discussed in this thesis is the *random walk,* designated by $\varphi = 1$. Assuming that the random walk has started somewhere in the past, its current value may also be written as

$$y_k = \sum_{j=0}^{\infty} q_{k-j}, \tag{3.9}$$

thus it can be interpreted as the cumulative sum of white noise. Figure 3.1 shows three realizations of autoregressive processes with varying values of φ, including a random walk. In all cases the value at $k = 0$ is set to zero and the same realization of normally distributed white noise with mean zero and unit variance is used as disturbance. The different evolutions of the processes are therefore only caused by the value of φ.

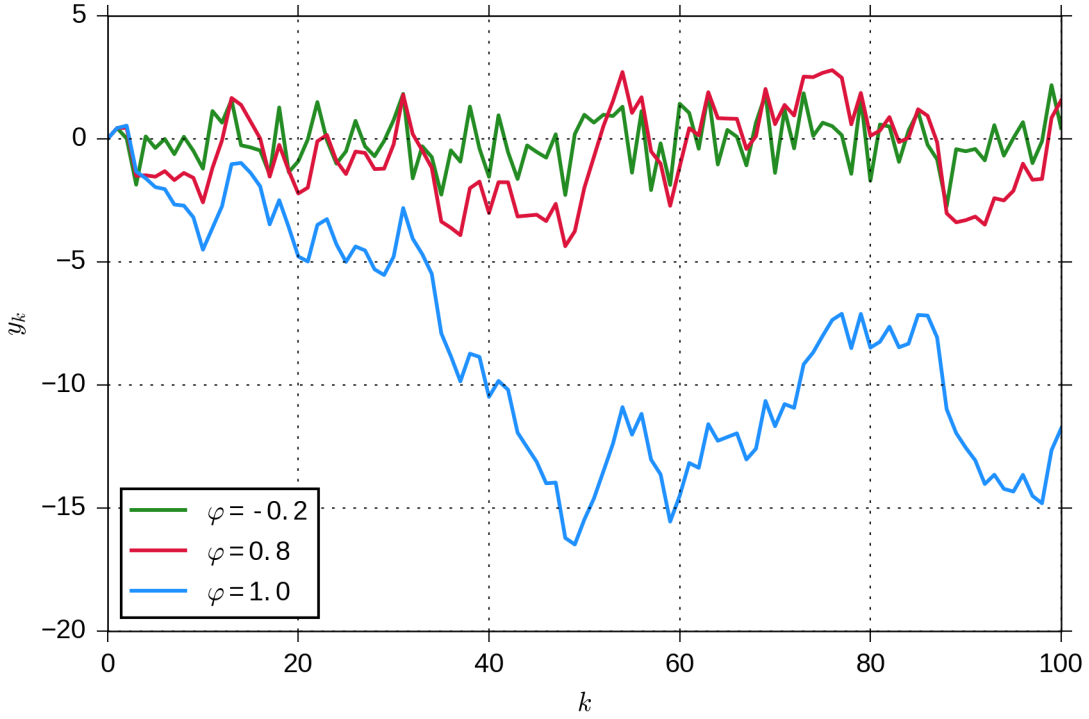

Figure 3.1: Three realizations of autoregressive stochastic processes. If $|\varphi| > 1$, the process is stationary. The non-stationary process designated by $\varphi = 1$ is called random walk.

The drawback of (3.5) and (3.9) is that they are only valid if the sampling interval $\Delta t = t_k - t_{k-1}$ is constant. In order to handle unequal temporal intervals between observations, the process can also be modeled in continuous time. In case of the random walk, the corresponding continuous-time equivalent is known as *Wiener* or *Brownian motion* process characterized by the differential equation

$$\frac{d}{dt} y(t) = q_c(t), \tag{3.10}$$

where $q_c(t)$ is continuous-time white noise with mean zero and variance Q_c. White noise inputs in differential equations have to be treated specially because the definition of the integral or derivative of a function $f[q_c(t), t]$ is not unique. Both Harvey (1989) and Gibbs (2011) point to Jazwinski (1970) for further information. The consequence is that Equation (3.10) is not valid in a strict sense, but it is still used because of its demonstrative form, which indicates that the output of a Wiener process $W(t)$ is integrated white noise. The increment $W(t_2) - W(t_1)$ has the properties

$$\mathrm{E}\left[\int_{t_1}^{t_2} q_c(\tau)\, d\tau\right] = 0, \tag{3.11}$$

$$\mathrm{Var}\left[\int_{t_1}^{t_2} q_c(\tau)\, d\tau\right] = Q_c(t_2 - t_1) \quad \text{and} \tag{3.12}$$

$$\mathrm{E}\left[\int_{t_1}^{t_2} q_c(\tau)\,d\tau \int_{t_3}^{t_4} q_c(\tau)\,d\tau\right] = 0, \quad t_1 < t_2 < t_3 < t_4. \tag{3.13}$$

While (3.12) says that the variance of the continuous random walk grows linear with time, (3.13) expresses that the increments of different time periods are uncorrelated.

3.4 State space models and the Kalman filter

Structural time series models are formulated using the discrete-time state space representation (Harvey, 1989, chap. 3). This concept assumes that a linear, time-variant system can be described by a set of state variables. Due to the fact that these variables can usually not be observed directly, a measurement model linking the system state to a set of observables is required. The measurement equation is

$$\mathbf{z}_k = \mathbf{H}_k\,\mathbf{x}_k + \mathbf{r}_k, \tag{3.14}$$

where $\mathbf{z}_k$ is an n_k-element vector of observations, $\mathbf{x}_k$ is the p-element state vector, $\mathbf{H}_k$ is an $n_k \times p$ matrix, and $\mathbf{r}_k$ is an n_k-element vector of serially uncorrelated observation noise with mean zero and $n_k \times n_k$ covariance matrix $\mathbf{R}_k$:

$$\mathrm{E}(\mathbf{r}_k) = \mathbf{0} \quad \text{and} \quad \mathrm{Var}(\mathbf{r}_k) = \mathbf{R}_k. \tag{3.15}$$

The temporal evolution of the state vector is described by a dynamic model using the transition equation

$$\mathbf{x}_k = \mathbf{\Phi}_k\,\mathbf{x}_{k-1} + \mathbf{q}_k, \tag{3.16}$$

where $\mathbf{\Phi}_k$ denotes the $p \times p$ transition matrix and $\mathbf{q}_k$ is a p-element vector of serially uncorrelated process noise with mean zero and $p \times p$ covariance matrix $\mathbf{Q}_k$:

$$\mathrm{E}(\mathbf{q}_k) = \mathbf{0} \quad \text{and} \quad \mathrm{Var}(\mathbf{q}_k) = \mathbf{Q}_k. \tag{3.17}$$

The subscript k indicates that a variable may vary with time. 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 variables $\mathbf{H}_k$, $\mathbf{R}_k$, $\mathbf{\Phi}_k$, and $\mathbf{Q}_k$ are known as system matrices. To clarify the notation it should be mentioned that an index k on $\mathbf{\Phi}$ always means that the respective transition matrix describes the temporal change of the state vector with respect to the previous epoch, that is

$$\mathbf{\Phi}_k = \mathbf{\Phi}(t_k, t_{k-1}). \tag{3.18}$$

The process noise vector $\mathbf{q}_k$ in Equation (3.16) is a key element of the dynamic model. While the matrix $\mathbf{\Phi}_k$ represents a purely deterministic transition of the state, the added process noise allows a stochastic evolution. The technical term is that the process noise *drives* the state variables. When the time interval between consecutive epochs is not constant, structural time series models are built on the assumption that

the noise in the dynamic model is generated by continuous-time white noise processes. Each state variable may be driven by a separate noise process. Since $\mathrm{E}(\mathbf{q}_k) = \mathbf{0}$, the noise vector is not directly modeled, but its associated covariance matrix is given by the integral

$$\mathbf{Q}_k = \int_{t_{k-1}}^{t_k} \mathbf{\Phi}(t_k, \tau)\, \mathbf{Q}_c\, \mathbf{\Phi}^{\mathrm{T}}(t_k, \tau)\, d\tau\,, \tag{3.19}$$

where $\mathbf{Q}_c$ denotes the covariance matrix of multivariate continuous-time white noise. Hence $\mathbf{Q}_k$ represents the variance build-up from one epoch to the next. The process noise driving one state variable may also influence others, including $\mathbf{\Phi}_k$ in (3.19) accounts for these effects (Gibbs, 2011, chap. 2; Harvey, 1989, chap. 9).

When the state space concept is applied to a time series model, the system state represents the various components like trend and seasonality. Note that the measurement equation given in (3.14) applies to a multivariate time series. However, Section 3.1 indicated that the mathematical and computational complexity of multivariate time series analysis quickly grows beyond the scope of this thesis. In the univariate case, n_k is equal to unity in all epochs and the dimensions of the involved variables are reduced accordingly. Using the p-element *row* vector $\mathbf{h}_k$, the measurement model can be simplified to

$$z_k = \mathbf{h}_k \mathbf{x}_k + r_k, \quad \text{where} \quad \mathrm{E}(r_k) = 0 \quad \text{and} \quad \mathrm{Var}(r_k) = R_k\,. \tag{3.20}$$

The following sub-sections first discuss the derivation of the system matrices needed to implement univariate structural time series models. Secondly, the Kalman filter algorithm and its application are described.

3.4.1 Trend

In structural models, the formulation of the trend component is based on the current level μ_k of the trend at time t_k, that is

$$T(t = t_k) = \mu_k\,. \tag{3.21}$$

The level is directly introduced as a state variable. In order to model a linear trend, the level's rate of change $\dot{\mu}_k$ is added to the state vector. The dot notation represents the derivatives of a variable with respect to time. Trends of a higher order may be implemented by including derivatives up to the corresponding order, like $\ddot{\mu}_k$ for a quadratic trend. However, this section focuses on the linear case. With the time interval between consecutive epochs denoted as Δt, the transition equation according to (3.16) is

$$\begin{pmatrix} \mu_k \\ \dot{\mu}_k \end{pmatrix} = \begin{pmatrix} 1 & \Delta t \\ 0 & 1 \end{pmatrix} \begin{pmatrix} \mu_{k-1} \\ \dot{\mu}_{k-1} \end{pmatrix} + \mathbf{q}_k\,. \tag{3.22}$$

The state space representation reflects the idea of a *local* linear trend, where the slope adapts gradually as new measurements are coming in. Hence the dynamic model is made stochastic based on the assumption that $\dot{\mu}_k$ follows a random walk. Considering

that the value of Δt may not be constant, the process noise is modeled in continuous time. With known transition matrix of the system, the discrete-time covariance matrix corresponding to a given (positive) time interval can be computed by evaluating the integral

$$\mathbf{Q}_k = \int_{\tau=0}^{\Delta t} \begin{pmatrix} 1 & \tau \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 0 & 0 \\ 0 & Q_c \end{pmatrix} \begin{pmatrix} 1 & 0 \\ \tau & 1 \end{pmatrix} d\tau\,, \tag{3.23}$$

where Q_c is the variance of continuous-time white noise. The local linear trend component for a univariate time series is completely specified by the state vector and system matrices summarized in (3.24). Assuming a constant measurement variance $\sigma_z{}^2$ the measurement model following (3.20) is time-invariant.

$$\mathbf{x}_k = \begin{pmatrix} \mu_k \\ \dot{\mu}_k \end{pmatrix},\ \mathbf{h} = \begin{pmatrix} 1 & 0 \end{pmatrix},\ R = \sigma_z^2,\ \mathbf{\Phi}_k = \begin{pmatrix} 1 & \Delta t \\ 0 & 1 \end{pmatrix},\ \mathbf{Q}_k = Q_c \begin{pmatrix} \Delta t^3/3 & \Delta t^2/2 \\ \Delta t^2/2 & \Delta t \end{pmatrix} \tag{3.24}$$

3.4.2 Seasonality

The dynamic model of the seasonal component follows the assumption that the seasonal variations in the spectral signatures because of plant phenology effects can be approximated by a sum of trigonometric functions of different frequencies, similar to the principle of Fourier synthesis. This section outlines the state space representation of a single trigonometric function, while the next sub-section discusses how to combine several functions of different frequencies. The underlying modeling assumption is appropriate especially for forest, where seasonal change events like the green-up in spring occur gradually and thus have a continuous characteristic. Other land cover types like grassland and agriculture exhibit discontinuities due to mowing and harvesting events. Larger model errors have to be expected for these land cover types, because the maximum number of frequencies to be included in the model is limited by the observation density.

Consider a cosine wave γ of frequency ω modulated by the parameters amplitude and phase, denoted by A and ϕ respectively:

$$\gamma(t) = A\cos(\omega t - \phi)\,. \tag{3.25}$$

In this form, the phase parameter is an argument to the cosine function. A trigonometric identity can be exploited to obtain the alternative form

$$\gamma(t) = \alpha\cos(\omega t) + \beta\sin(\omega t)\,, \tag{3.26}$$

where the original amplitude and phase parameters are replaced by α and β. The relationships of the new parameters to A and ϕ are given in (3.27).

$$\begin{aligned} A &= \sqrt{\alpha^2 + \beta^2} \\ \phi &= \arccos(\alpha/A) = \arcsin(\beta/A) \end{aligned} \tag{3.27}$$

Similar to the formulation of the trend, the current level γ_k of the wave at time t_k is introduced as a state variable. Furthermore, the variable γ_k^*, whose interpretation will be

clarified later, is added to the state vector. At time $t = 0$, the state variables correspond to the parameters α and β of Equation (3.26), hence the initial state vector is given by

$$\mathbf{x}_0 = \begin{pmatrix} \gamma_0 \\ \gamma_0^* \end{pmatrix} = \begin{pmatrix} \alpha \\ \beta \end{pmatrix}. \tag{3.28}$$

This relationship is further illustrated in Figure 3.2, where γ_k and γ_k^* are variables on the x_1- and x_2-axis of a Cartesian coordinate system representing the state space. The drawing shows that the transition matrix of the state space model is a rotation matrix of the form

$$\hat{\mathbf{R}}(\theta) = \begin{pmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{pmatrix}, \tag{3.29}$$

which rotates the state vector in the x_1x_2-plane clockwise through an angle θ about the origin of the coordinate system. The rotation angle corresponding to a certain time interval Δt between consecutive epochs is given by the product $\omega \cdot \Delta t$, thus

$$\mathbf{\Phi}_k = \hat{\mathbf{R}}(\omega \Delta t). \tag{3.30}$$

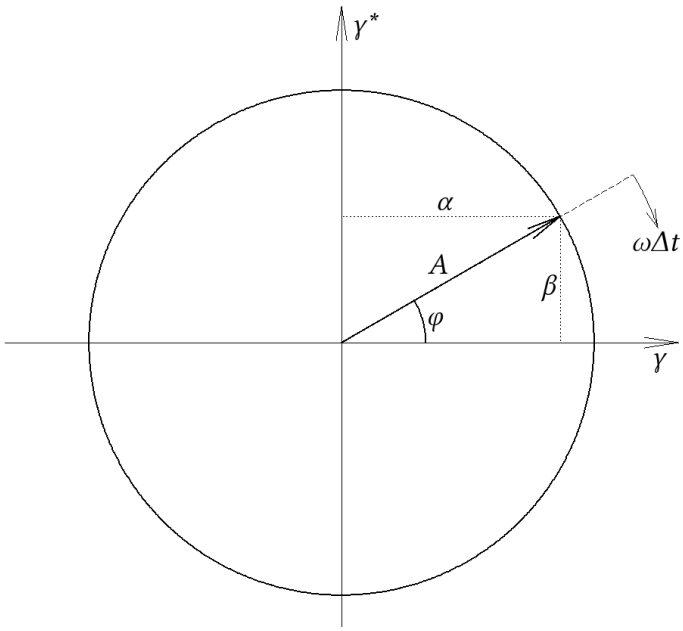

Figure 3.2: Initial state vector representing a cosine wave of frequency ω, modulated by amplitude A and phase ϕ. The transition matrix of the model rotates the state vector clockwise.

The benefit of the state-space approach is that the seasonal component may evolve over time as new measurements are obtained. Hence the dynamic model is made stochastic based on the assumption that both amplitude and phase of the cosine wave follow a random walk. This is implemented by including white process noise on both state variables. Taking into account that the value of Δt may not be constant, the influence of the process noise is modeled in continuous time. With known transition matrix of the system, the discrete-time covariance matrix corresponding to a given (positive) time interval can be computed by evaluating the integral

$$\mathbf{Q}_k = \int_{\tau=0}^{\Delta t} \begin{pmatrix} \cos(\omega\tau) & \sin(\omega\tau) \\ -\sin(\omega\tau) & \cos(\omega\tau) \end{pmatrix} \begin{pmatrix} Q_c & 0 \\ 0 & Q_c \end{pmatrix} \begin{pmatrix} \cos(\omega\tau) & -\sin(\omega\tau) \\ \sin(\omega\tau) & \cos(\omega\tau) \end{pmatrix} d\tau\,, \tag{3.31}$$

where Q_c is the variance of continuous-time white noise. The state vector and system matrices required to model a cosine wave are summarized in (3.32), where $\mathbf{I}$ represents a 2×2 identity matrix. Assuming a constant measurement variance $\sigma_z{}^2$, the measurement model following (3.20) is time-invariant.

$$\mathbf{x}_k = \begin{pmatrix} \gamma_k \\ \gamma_k^* \end{pmatrix},\quad \mathbf{h} = \begin{pmatrix} 1 & 0 \end{pmatrix},\quad R = \sigma_z^2,\quad \mathbf{\Phi}_k = \begin{pmatrix} \cos(\omega\Delta t) & \sin(\omega\Delta t) \\ -\sin(\omega\Delta t) & \cos(\omega\Delta t) \end{pmatrix},\quad \mathbf{Q}_k = Q_c\,\Delta t\,\mathbf{I} \tag{3.32}$$

3.4.3 Combination of model components

The seasonal component is constructed by a sum of P-periodic cosine waves, with P denoting the fundamental duration of the seasonal cycle. Therefore, the set Ω containing the seasonal (angular) frequencies is

$$\Omega = \{\omega_1, \omega_2, \ldots\} = \left\{ \frac{2\pi j}{P} \;\middle|\; j = 1, 2, \ldots, \infty \right\}, \tag{3.33}$$

with j as the corresponding number of periods per cycle. Considering the nature of the time series investigated in this thesis, it is appropriate to measure time between epochs in days and thus set P to 365.25.

The system matrices of a univariate structural time series model incorporating a trend and a seasonal component consisting of several frequencies are obtained by combining the elements specified in (3.24) and (3.32). Concerning the vectors $\mathbf{x}$ and $\mathbf{h}$, this is done by concatenating the individual trend and seasonal elements in the manner of (3.34). An additional numerical subscript indicates the seasonal frequency associated with the respective element.

$$\mathbf{x}_k = \begin{pmatrix} \mathbf{x}_k^{\text{trend}} \\ \mathbf{x}_{1,k}^{\text{seas}} \\ \mathbf{x}_{2,k}^{\text{seas}} \\ \vdots \end{pmatrix},\quad \mathbf{h} = \left(\mathbf{h}^{\text{trend}}, \mathbf{h}_1^{\text{seas}}, \mathbf{h}_2^{\text{seas}}, \ldots \right) \tag{3.34}$$

The individual transition matrices are combined to a single matrix with block-diagonal structure illustrated in (3.35). The same holds true for the process noise covariance matrices.

$$\mathbf{\Phi}_k = \begin{pmatrix} \mathbf{\Phi}_k^{\text{trend}} & \mathbf{0} & \mathbf{0} & \\ \mathbf{0} & \mathbf{\Phi}_{1,k}^{\text{seas}} & \mathbf{0} & \cdots \\ \mathbf{0} & \mathbf{0} & \mathbf{\Phi}_{2,k}^{\text{seas}} & \\ & \vdots & & \ddots \end{pmatrix},\quad \mathbf{Q}_k = \begin{pmatrix} \mathbf{Q}_k^{\text{trend}} & \mathbf{0} & \mathbf{0} & \\ \mathbf{0} & \mathbf{Q}_{1,k}^{\text{seas}} & \mathbf{0} & \cdots \\ \mathbf{0} & \mathbf{0} & \mathbf{Q}_{2,k}^{\text{seas}} & \\ & \vdots & & \ddots \end{pmatrix} \tag{3.35}$$

3.4.4 Kalman filter

Once the system matrices and the initial values $\mathbf{x}_0$ and $\mathbf{P}_0$ are defined, the discrete-time Kalman filter algorithm can be used to obtain estimates for the state and its error covariance matrix in subsequent epochs $k = 1, 2, \ldots m$. The Kalman filter processes observations recursively, one epoch at a time, and each recursion may be divided into several steps. The first one is known as time update step and yields the predicted (*a-priori*) estimates $\widetilde{\mathbf{x}}_k$ and $\widetilde{\mathbf{P}}_k$ based on the dynamic model and the previous estimates at time t_{k-1}:

$$\begin{aligned} \widetilde{\mathbf{x}}_k &= \boldsymbol{\Phi}_k \hat{\mathbf{x}}_{k-1} \\ \widetilde{\mathbf{P}}_k &= \boldsymbol{\Phi}_k \hat{\mathbf{P}}_{k-1} \boldsymbol{\Phi}_k^{\mathrm{T}} + \mathbf{Q}_k \end{aligned} \tag{3.36}$$

Step two is to compute the *a-priori* measurement residual $\mathbf{y}_k$ and its associated covariance matrix $\mathbf{C}_k$ using (3.37). The residual represents the difference of the prediction to the actual measurements and is referred to as *innovation,* since it contains new information currently not present in the predicted state.

$$\begin{aligned} \mathbf{y}_k &= \mathbf{z}_k - \mathbf{H}_k \widetilde{\mathbf{x}}_k \\ \mathbf{C}_k &= \mathbf{H}_k \widetilde{\mathbf{P}}_k \mathbf{H}_k^{\mathrm{T}} + \mathbf{R}_k \end{aligned} \tag{3.37}$$

In the final step of each recursion, the new information is merged with the predictions to obtain improved (*a-posteriori*) estimates $\hat{\mathbf{x}}_k$ and $\hat{\mathbf{P}}_k$. Therefore it is known as measurement update step. The Kalman gain matrix $\mathbf{K}_k$ given in (3.38) determines how much the newly acquired measurements will influence the *a-posteriori* estimates of the state and its error covariance.

$$\mathbf{K}_k = \widetilde{\mathbf{P}}_k \mathbf{H}_k^{\mathrm{T}} \mathbf{C}_k^{-1} \tag{3.38}$$

The elements of $\mathbf{K}_k$ are ranging from 0 to 1, thus it can be interpreted as a weighting matrix. With the update equations stated in (3.39), the basic Kalman filter recursion is complete.

$$\begin{aligned} \hat{\mathbf{x}}_k &= \widetilde{\mathbf{x}}_k + \mathbf{K}_k \mathbf{y}_k \\ \hat{\mathbf{P}}_k &= \left(\mathbf{I} - \mathbf{K}_k \mathbf{H}_k\right) \widetilde{\mathbf{P}}_k \end{aligned} \tag{3.39}$$

When the sequence of measurements processed by the filter contains outliers, an additional outlier detection step should be included *before* the measurement update step. 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 correct and the observation noise is Gaussian, the innovations will be normally distributed with mean zero and covariance matrix $\mathbf{C}_k$, that is

$$\mathbf{y}_k \sim N\left(0, \mathbf{C}_k\right). \tag{3.40}$$

The test statistic $\hat{T}_k$ given in (3.41) follows the χ^2-distribution with n_k degrees of freedom, where n_k is the number of observations in epoch k. The hypotheses to be tested on a significance level α are stated in (3.42).

$$\hat{T}_k = \mathbf{y}_k^{\mathrm{T}} \mathbf{C}_k^{-1} \mathbf{y}_k, \quad \hat{T}_k \sim \chi^2(n_k). \tag{3.41}$$

$$\begin{aligned} H_0: &\quad \boldsymbol{y}_k = 0 \quad \text{if} \quad \hat{T} \le \chi^2_{n_k,1-\alpha} \\ H_1: &\quad \boldsymbol{y}_k \neq 0 \quad \text{otherwise} \end{aligned} \tag{3.42}$$

Considering that anomalous observations will cause large innovations, the null hypothesis will be rejected if outliers are present in the current epoch. In order to avoid a negative influence on the state estimate, the measurement update step should not be carried out (Gibbs, 2011, chap. 8).

3.5 Linear regression models and robust parameter estimation

A general linear regression model links an m-element observation vector $\mathbf{z}$ to a p-element parameter vector $\mathbf{x}$ through the measurement equation

$$\mathbf{z} = \mathbf{A}\,\mathbf{x} + \mathbf{r}, \tag{3.43}$$

where $\mathbf{A}$ is the $m \times p$ design matrix of the model and $\mathbf{r}$ is an m-element vector of normally distributed observation noise with mean zero and covariance matrix $\mathbf{R}$. The observation noise represents the differences between the data and the model and therefore $\mathbf{r}$ is also known as vector of residuals.

Regarding a univariate time series, there is a clear relationship between the structural models outlined in the previous section and linear regression models. The same modeling principles can be applied to obtain the design matrix and consequently (3.43) may be written as

$$\begin{pmatrix} z_1 \\ z_2 \\ \vdots \\ z_m \end{pmatrix} = \begin{pmatrix} \mathbf{h}\,\boldsymbol{\Phi}(t_1, t_0) \\ \mathbf{h}\,\boldsymbol{\Phi}(t_2, t_0) \\ \vdots \\ \mathbf{h}\,\boldsymbol{\Phi}(t_m, t_0) \end{pmatrix} \mathbf{x}_0 + \begin{pmatrix} r_1 \\ r_2 \\ \vdots \\ r_m \end{pmatrix}. \tag{3.44}$$

Equation (3.44) illustrates the key difference between a structural model and a regression model: The latter does not include the concept of process noise and thus the parameter vector is time-invariant and referenced to a fixed epoch (Gibbs, 2011, sec. 4.1).

Still, (3.44) can be used to model a time series within a certain training period in order to obtain an estimate for the initial state of a structural model. Assuming $m > p$ and normally distributed and uncorrelated observation noise with mean zero and constant variance σ^2, the parameter vector can be estimated using the method of ordinary least squares (OLS). Considering the presence of outliers in the data, this assumption is violated and a robust parameter estimation approach following the implementation of Heiberger and Becker (1992) is employed. The technique is known as iteratively reweighted least squares (IRLS) and belongs to the class of M-estimators (Huber, 1964). The parameter estimate $\hat{\mathbf{x}}$ is computed based on the cost function

$$F = \left(\mathbf{z} - \mathbf{A}\,\hat{\mathbf{x}}\right)^{\mathrm{T}} \mathbf{W} \left(\mathbf{z} - \mathbf{A}\,\hat{\mathbf{x}}\right) \;\longrightarrow\; \min. \tag{3.45}$$

where $\mathbf{W}$ is an $m \times m$ diagonal matrix of observation weights ranging from 0 to 1. Consequently, $\mathbf{W}$ can also be expressed in terms of an m-element vector $\mathbf{w}$, that is

$$\mathbf{W} = \mathrm{diag}(\mathbf{w}). \tag{3.46}$$

The parameter estimate minimizing F is given by

$$\hat{\mathbf{x}} = (\mathbf{A}^{\mathrm{T}} \mathbf{W} \mathbf{A})^{-1} \mathbf{A}^{\mathrm{T}} \mathbf{W} \mathbf{z}. \tag{3.47}$$

The initial solution of the IRLS procedure is obtained by setting $\mathbf{W}$ equal to an identity matrix, thus it is equal to the OLS solution. In subsequent iterations, the weights are recalculated based on the residuals $r_1, r_2, \ldots r_m$. Observations with large residual values are down-weighted. Each residual is divided by the scale s to obtain the normalized residual u_i:

$$u_i = r_i / s \quad \text{using} \quad s = \mathrm{median}(|r_1|, |r_2|, \ldots) / 0.6745 \tag{3.48}$$

The normalized residuals are then used as input to a cost function which determines the weight of the corresponding observation. Heiberger and Becker (1992) suggest a two-stage procedure employing the Huber cost function (Huber, 1964) until convergence followed by (up to) two additional iterations applying the Bisquare cost function (Beaton and Tukey, 1974). Convergence is achieved when the difference of the Euclidean norms of $\hat{\mathbf{x}}$ between consecutive iterations is below a given threshold. The respective elements of $\mathbf{w}$ are determined by using (3.49) or Error: Reference source not found, where c is a tuning parameter. An illustration of the cost functions is given in Figure 3.3. The Huber cost function is more conservative regarding the down-weighting of suspicious observations.

$$w_{i,\mathrm{Huber}} = \begin{cases} 1 & |u_i| \le c \\ \dfrac{c}{|u_i|} & |u_i| > c \end{cases} \quad \text{default } c = 1.345 \tag{3.49}$$

$$w_{i,\mathrm{Bisquare}} = \begin{cases} \left(1 - \left(\dfrac{u_i}{c}\right)^2\right)^2 & 0 < \left|\dfrac{u_i}{c}\right| \le 1 \\ 0 & \left|\dfrac{u_i}{c}\right| > 1 \end{cases} \quad \text{default } c = 4.685 \tag{3.50}$$

Additional to the parameter estimate $\hat{\mathbf{x}}$, several other interesting quantities can be derived within the weighted least squares framework. An unbiased estimator for the observation noise variance σ^2 is given by

$$\hat{\sigma}^2 = \frac{\mathbf{r}^{\mathrm{T}} \mathbf{W} \mathbf{r}}{m - p} \quad \text{where} \quad \mathbf{r} = \mathbf{z} - \mathbf{A}\hat{\mathbf{x}}. \tag{3.51}$$

This quantity is also known as the *mean squared error* (MSE) of the fit. The uncertainty of the parameter estimate is reflected by the covariance matrix

$$\hat{\mathbf{P}} = \hat{\sigma}^2 (\mathbf{A}^{\mathrm{T}} \mathbf{W} \mathbf{A})^{-1}. \tag{3.52}$$

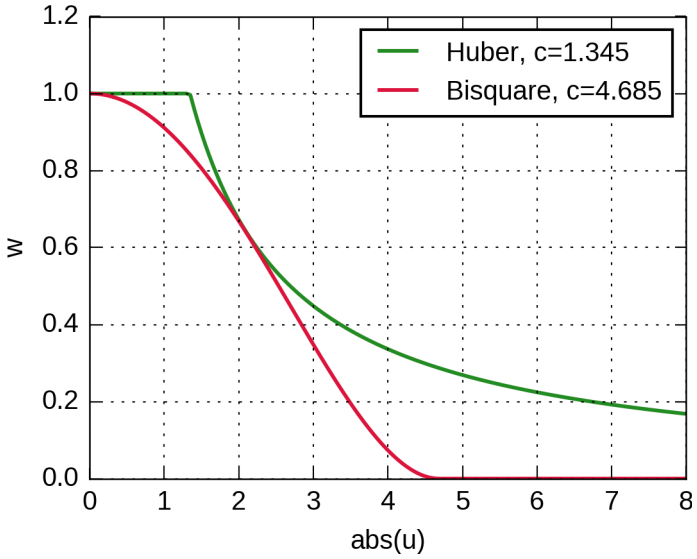

Figure 3.3: Huber and Bisquare weighting functions with default tuning parameters. The curves depict the assigned observation weight w *dependent on the absolute value of the normalized residual* u.

Another simple measure used to assess how well a regression model fits the underlying data is given by the coefficient of determination R^2. A robust version of R^2 which takes any previously computed weights into account can be written as

$$R^2 = 1 - \frac{\sum_{i=1}^{m} w_i \left(z_i - \hat{z}\right)^2}{\sum_{i=1}^{m} w_i \left(z_i - \bar{z}_w\right)^2} \quad \text{where} \quad \bar{z}_w = \left(1/\sum w_i\right) \sum w_i z_i. \tag{3.53}$$

The denominator in the ratio above represents the total variation of the data about its weighted mean. The numerator on the other hand is equal to the weighted residual sum of squares $\mathbf{r}^{\mathrm{T}} \mathbf{W} \mathbf{r}$. As a result, the coefficient of determination will be close to unity if most of the total variation is captured by the fitted model (Renaud and Victoria-Feser, 2010).

3.6 Change detection algorithm

The flowchart depicted in Figure 3.4 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 pixel-level. Some further explanatory comments are given below.

(1) The tuning parameters and the time series model components specified by the user are applied globally, hence they are the same for all pixels.

(2) The initial state as well as the observation noise are estimated on pixel-level using the IRLS method. Therefore the user has to supply a stack of historic images as training dataset.

(3) New images covering 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.

(4) Each pixel features an “anomaly counter”. Each time an observation is marked as anomalous, the counter is *incremented* by 1. In contrast, the counter is also *decremented* by 1 if the current observation is not statistically conspicuous, although it can never become lower than zero. This principle ensures that a certain user-defined “change threshold” of the counter can only be reached when outliers occur temporally aggregated.

(5) Whenever an observation is marked as anomalous, the measurement update step is bypassed.

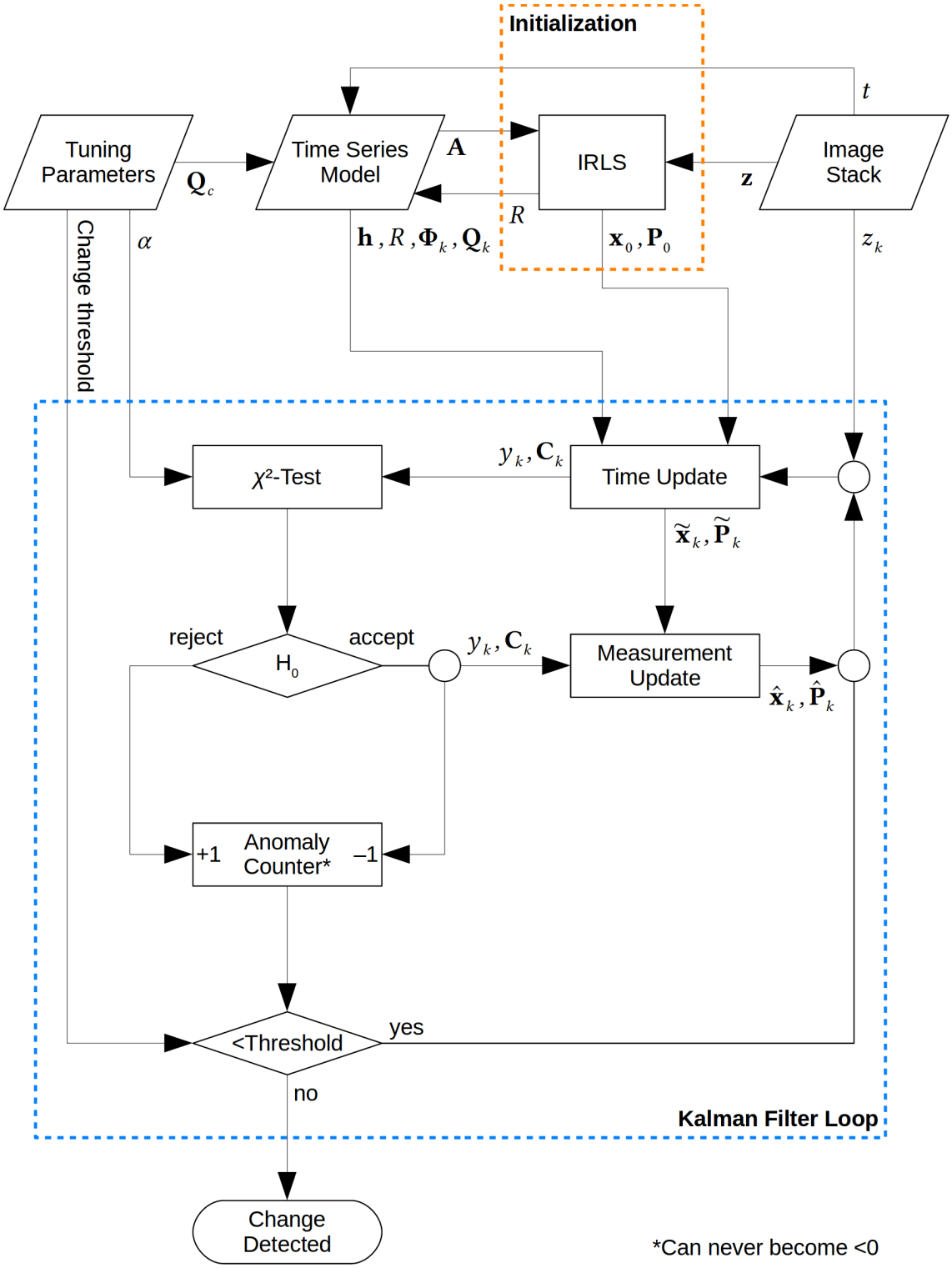

Figure 3.4: Change detection algorithm flowchart

4 Data Pre-Processing and Test Site

This chapter is divided into two parts. The first section describes the properties of the used Landsat satellite imagery and discusses the applied pre-processing. Section 4.2 presents specific information about the selected test site located in the federal state of Baden-Württemberg, Germany.

4.1 Data and pre-processing

In order to carry out experimental testing of the methods outlined in the preceding chapter, Landsat Surface Reflectance (SR) products provided by the U.S. Geological Survey Earth Resources Observation and Science Center (USGS EROS) are used. Since Landsat 4, all Landsat data are organized using the Worldwide Reference System 2 (WRS-2) path and row notation (Wulder et al., 2016). All Landsat 5, 7, and 8 images of path 195 and row 96 available in the time period from 2009-01-01 to 2016-12-31 have been downloaded using the USGS EarthExplorer service (earthexplorer.usgs.gov). The various Landsat satellites carry different instruments to acquire data, see Table 4. Each instrument measures different ranges of frequencies along the electromagnetic spectrum. These spectral ranges are illustrated in Figure 4.1, along with the associated band number of the corresponding satellite.

Table 4: Instruments carried by Landsat satellites

Satellite	Instruments
Landsat 1-5	Multispectral Scanner System (MSS)
Landsat 4-5	Thematic Mapper (TM)
Landsat 7	Enhanced Thematic Mapper (ETM+)
Landsat 8	Operational Land Imager (OLI), Thermal Infrared Sensor (TIRS)
Landsat 9*	Operational Land Imager 2 (OLI-2), Thermal Infrared Sensor 2 (TIRS-2)

*not yet operational

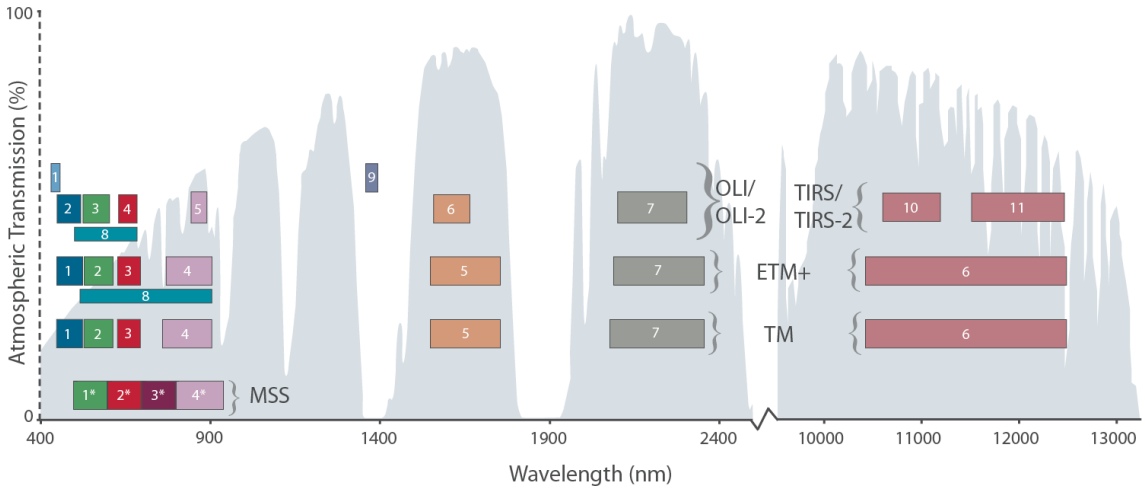

Figure 4.1: Spectral bands of various Landsat instruments (Source: NASA/USGS Landsat)

Landsat SR products are considered as a high-level data set (USGS, 2016a, 2016b). There are three processing levels with respect to the quality of radiometric calibration and geometric registration. The processing level of an image can be retrieved from the supplied metadata. Only the highest-level products, designated as L1T, are suitable for time series analysis applications on a pixel-level (USGS, 2018). In case of Landsat 5 or 7 data, the conversion to SR is based on the Landsat Ecosystem Disturbance Adaptive Processing System (LEDAPS, Masek et al., 2006). Landsat 8 data is processed to SR using the Provisional Landsat 8 Surface Reflectance Code (LaSRC, Vermote et al., 2016). A cloud, cloud shadow, snow, and water identification is provided based on the results of the Fmask algorithm (Zhu and Woodcock, 2012). The product specifications of spectral and mask bands which have been used in this thesis are listed in Table 5.

Table 5: Landsat Surface Reflectance product band specifications (USGS, 2016a, 2016b)

Band name	**Band number**		**Data type**	**Units**	**Valid Range**	**Fill Value**	**Scale Factor**
	TM, ETM+	**OLI**					
Blue	1	2	INT16	Reflectance	0 – 10 000	–9999	0.0001
Green	2	3	INT16	Reflectance	0 – 10 000	–9999	0.0001
Red	3	4	INT16	Reflectance	0 – 10 000	–9999	0.0001
NIR	4	5	INT16	Reflectance	0 – 10 000	–9999	0.0001
SWIR1	5	6	INT16	Reflectance	0 – 10 000	–9999	0.0001
SWIR2	7	7	INT16	Reflectance	0 – 10 000	–9999	0.0001
Fmask	NA	NA	UINT8	Flag	0 Clear 1 Water 2 Cld. Shadow 3 Snow 4 Cloud	255	NA

Abbreviations: NIR near infrared, SWIR short wave infrared, INT16 16-bit signed integer, UINT8 8-bit unsigned integer, NA not applicable

The first pre-processing step was to select the highest-quality images from all 241 downloaded products based on metadata such as the processing level, acquisition date, and cloud-cover properties. A detailed account of the selection criteria is given in Table 6. In order to reduce the data volume further, the 6 SR bands have been transformed to the tasseled cap (TC) features brightness, greenness, and wetness. The transformation coefficients given in Table 7 were published by Crist (1985). Additional to the reduced data volume, several existing studies reported that TC components also have desirable properties for change or disturbance detection in forested areas, see Oeser et al. (2017), Pasquarella et al. (2017) or Brooks (2014).

Table 6: Image selection steps

Step	Description	Remaining images
0	All available Landsat 5, 7, and 8 images, WRS-2 Path 195 / Row 26, date range from 2009-01-01 to 2016-12-31	241
1	Exclude images which are not designated to have L1T processing level	180
2	Exclude images taken in winter months (Dec, Jan, Feb)	150
3	Exclude images where more than 90% of the pixels in the test site are flagged as “cloud”, “cloud shadow” or “snow” according to the Fmask layer	111

Table 7: Tasseled cap transformation coefficients (Crist, 1985)

Feature	Blue	Green	Red	NIR	SWIR1	SWIR2
Brightness	0.2043	0.4158	0.5524	0.5741	0.3124	0.2303
Greenness	−0.1603	−0.2819	−0.4934	0.7940	−0.0002	−0.1446
Wetness	0.0315	0.2021	0.3102	0.1594	−0.6806	−0.6109

Using data from multiple satellites and therefore different measurement instruments raises questions concerning the spectral consistency between them. Figure 4.1 shows that the spectral bands of both TM and ETM+ are defined almost identically, but there are larger differences to the OLI bands. Furthermore, the conversion of OLI data to SR is based on a different algorithm. In the time series analysis context, the problem was discussed by Zhu et al. (2016). The authors concluded that data from Landsat 5 and 7 are fairly consistent, but there is a bias with respect to Landsat 8. They observed that the reflectance values of all 6 bands listed in Table 5 are lower when measured with the OLI instrument, with the largest differences occurring in the visible bands. The authors also noted that the bias is magnified in normalized-difference vegetation indices.

Another investigation of the spectral consistency problem was carried out by Oeser et al. (2017). Since the study employed TC wetness as primary observable, the authors analyzed the differences in wetness calculated from ETM+ and OLI data and found that values derived from the latter tended to be lower. They concluded “*As the observed differences generally were small* (*average difference in TC wetness <0.01 for forested areas*) *compared to changes associated with forest disturbance and OLI data made up only a small part of our analyzed time series, we did not apply any normalization to the OLI imagery.*” Following this conclusion, no additional pre-processing has been applied to the Landsat 8 data used in this thesis.

4.2 Test site

The test site is located in the federal state of Baden-Württemberg, Germany. Figure 4.2 shows its location relative to the state borders and the footprint of WRS-2 path 195,

row 26. The square-shaped test site covers an area of 900 km², which corresponds to 1000×1000 Landsat pixels. A cloud-free Landsat 8 image acquired on June 8 2014 is contrasted with a forest type classification based on the Copernicus High Resolution Layer (HRL) 2015 in Figure 4.3, showing that a large part of the test site is forested and both broadleaved and coniferous types occur. Figure 4.4 depicts a digital elevation model (DEM) of the test site, showing that the terrain elevation is ranging from about 130 to 1260 meters. Possible systematic influences of the relief on the results will be addressed in Chapter 5.

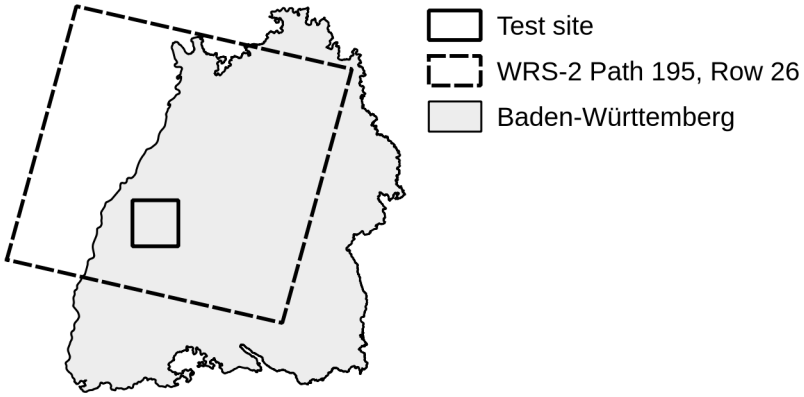

Figure 4.2: Location of the test site (*WRS-2 data courtesy of the U.S. Geological Survey. Baden-Württemberg federal state border available as open data, www.lgl-bw.de*)

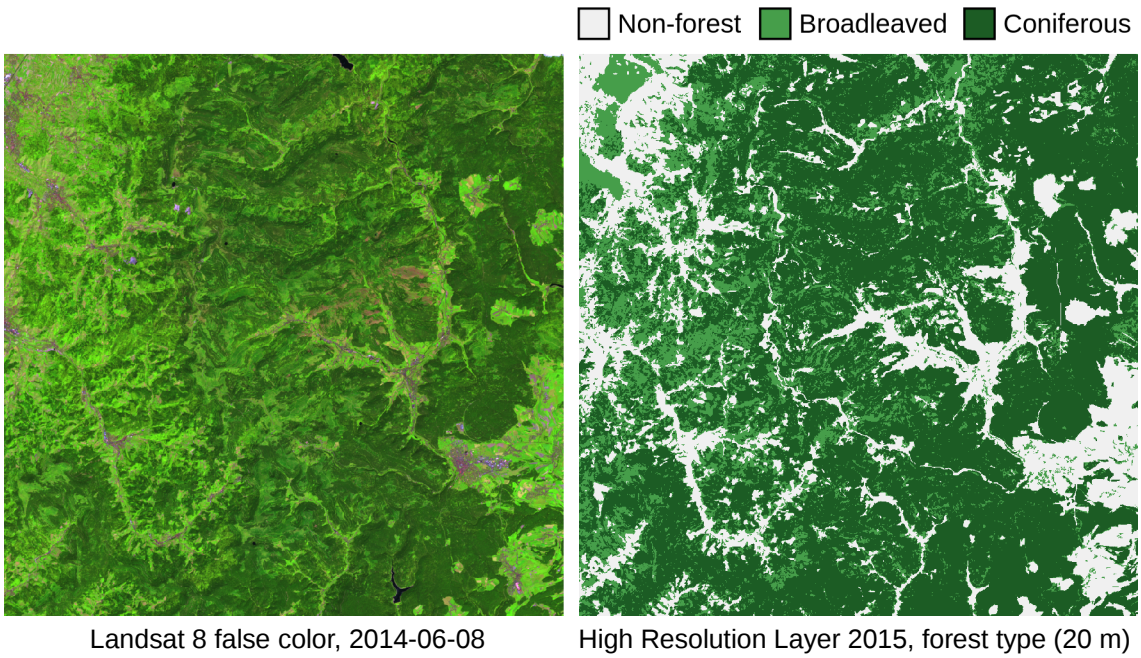

Figure 4.3: Distribution of forested areas within the test site (*Copernicus HRL 2015 accessible open and free at land.copernicus.eu*)

The number of available valid observations per year has been analyzed for each pixel of the test site. The results are summarized in Figure 4.5. Note that the last Landsat 5 imagery covering the test site was acquired at the end of 2011, resulting in a reduced observation count until the first Landsat 8 data became available in April 2013. Low minimum values correspond to very bright objects which frequently trigger false posi-

tives in the cloud masking process. Unfortunately not all of these bright objects are masked out if a forest mask derived from the HRL 2015 is applied, hence the minimum values reported in Figure 4.5 do not correspond to forest pixels. The median numbers of observations per year can be considered as a good estimate of the true observation count for a given forest pixel. The numbers indicate that the phenological cycle associated with forest is captured by the available data. An exemplary time series plot of the TC components for a coniferous forest pixel is given in Figure 4.6. Recurring seasonal patterns are clearly visible in the signals.

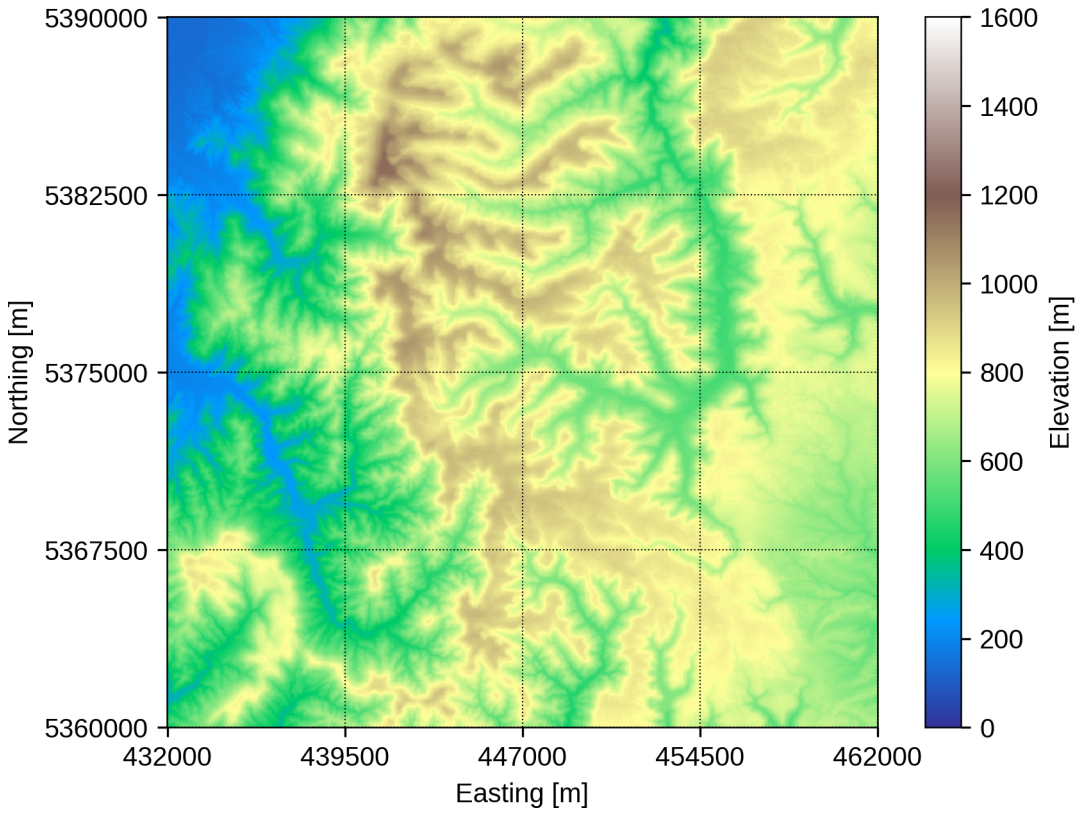

Figure 4.4: Digital elevation model (DEM) of the test site (Coordinate reference frame EPSG 32 632. Data source: Advanced Land Observing Satellite (ALOS) Global Digital Surface Model "ALOS World 3D – 30 m", © Japan Aerospace Exploration Agency)

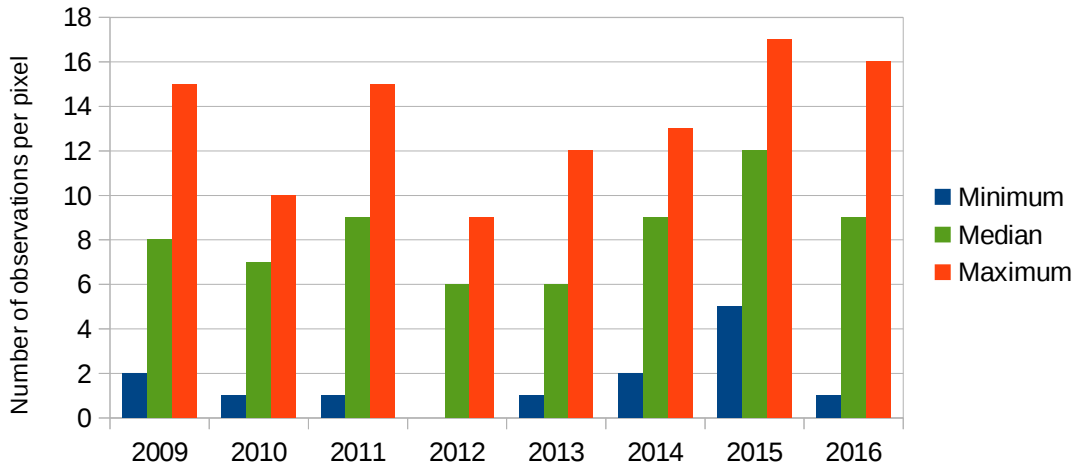

Figure 4.5: Analysis of the time series density per-pixel

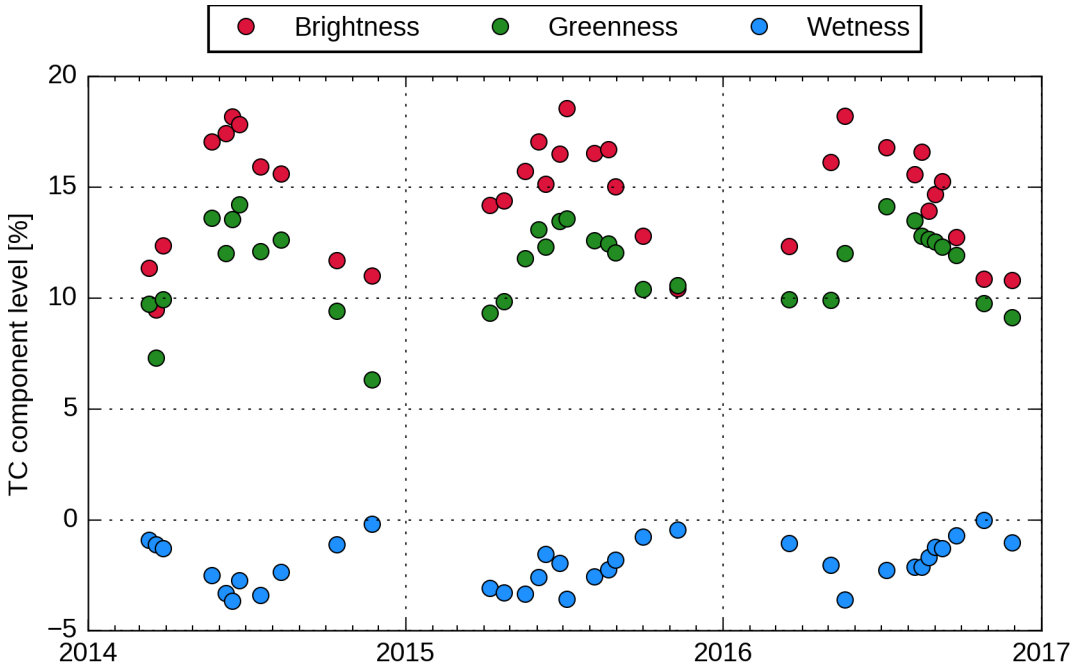

Figure 4.6: Typical time series of TC components for a coniferous forest pixel

5 Results and Discussion

The first part of this chapter presents and discusses the results obtained by applying the regression model approach outlined in Section 3.5. These results are required to initialize the Kalman filter, whose application and output are covered in the second section. The temporal domain of the results is illustrated in detail by time series plots of single, representatively selected sample pixels within the test site. Each of the pixels listed in Table 8 has been selected to show and discuss the algorithms behavior regarding different properties like the presence of un-masked clouds and cloud shadows, abrupt change events as well as the underlying forest type and terrain-dependent effects. In the spatio-temporal domain, the results are discussed using maps and sequences of images.

Table 8: Sample pixels

ID	x [m]	y [m]	Forest type	Property
1	441 060	5 362 320	coniferous	Un-masked cloud and cloud shadows occur
2	444 000	5 363 220	broadleaved	None
3	441 030	5 366 670	coniferous	Southward hillside
4	440 910	5 366 760	coniferous	Northward hillside
5	452 970	5 378 460	coniferous	Storm damage 2012
6	451 980	5 376 360	coniferous	Storm damage 2012, omission error

Coordinate reference frame: EPSG 32 632 (European Petroleum Survey Group Geodesy)

5.1 Initialization of the state space model

Initial values for the Kalman filter state are obtained by fitting a regression model to a given training period using the robust IRLS method. The time series model used here features a constant trend as well as a seasonal component using the frequencies ω_1 and ω_2 as defined in Equation (3.33). Thus, there are 5 model parameters (μ, α_1, β_1, α_2, β_2) to estimate. The appropriate length of the training period is mostly governed by the number of parameters, p, and the desired minimum number of available observations n to estimate them from. Tests showed that the IRLS procedure yields good and stable results when the degree of over-determination is set to about 3, this means that $n \approx 3p$. Another requirement concerning the length of the training period is that all seasons (except winter) should be represented equally, which is accomplished easily by allowing only an integer number of complete years. Following the considerations above and taking into account the median time series density reported in Figure 4.5, a 3-year training period is deemed appropriate when using a 5-parameter model. For a 3-parameter model featuring only the fundamental seasonal frequency ω_1, the length may be reduced to 2 years.

5.1.1 Down-weighting of anomalous observations

Exemplary results illustrating the methods capability to separate outliers are discussed using sample pixel 1, where three different types of anomalous observations (un-masked cloud, fog, and cloud shadow) as shown in Figure 5.1 occur within the training period. Figure 5.2 below shows time series of the tasseled cap components brightness, greenness and wetness in a 3-year training period ranging from the beginning of 2009 to the end of 2011. Additional to the discrete observations, different stages of the IRLS curve fit are plotted.

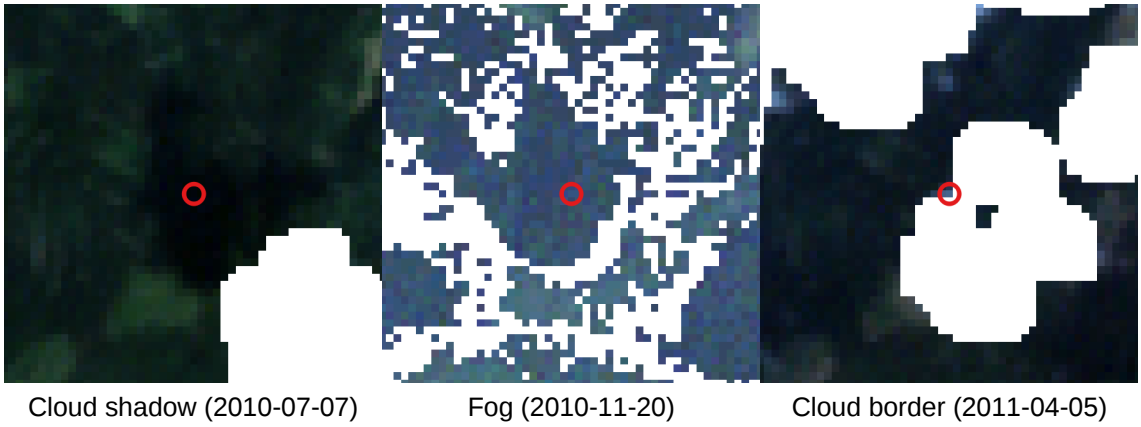

Figure 5.1: True-color images showing different kinds of outliers occurring in the time series of sample pixel 1, marked with a red circle.

The obtained curve fits show that the model captures the seasonal dynamics quite well, especially the greenness cycle. The anomalous observations corresponding to the images of Figure 5.1 are annotated accordingly. All 3 of them are clearly identifiable as outliers in the brightness signal. The cloud shadow is also visible in the greenness time series and the un-masked cloud border also presents itself in the wetness signal. Figure 5.3 depicts the final weights assigned to each observation by the Huber and Bisquare weighting functions. It can be verified that the outliers are effectively down-weighted. The results indicate that the influence of occasionally un-masked clouds and cloud shadows on the parameter estimate can be effectively reduced by employing the IRLS procedure, given that the available number of observations is large enough to allow a fair degree of over-determination. However, high-quality cloud masks are still required to reduce the overall number of outliers in the first place.

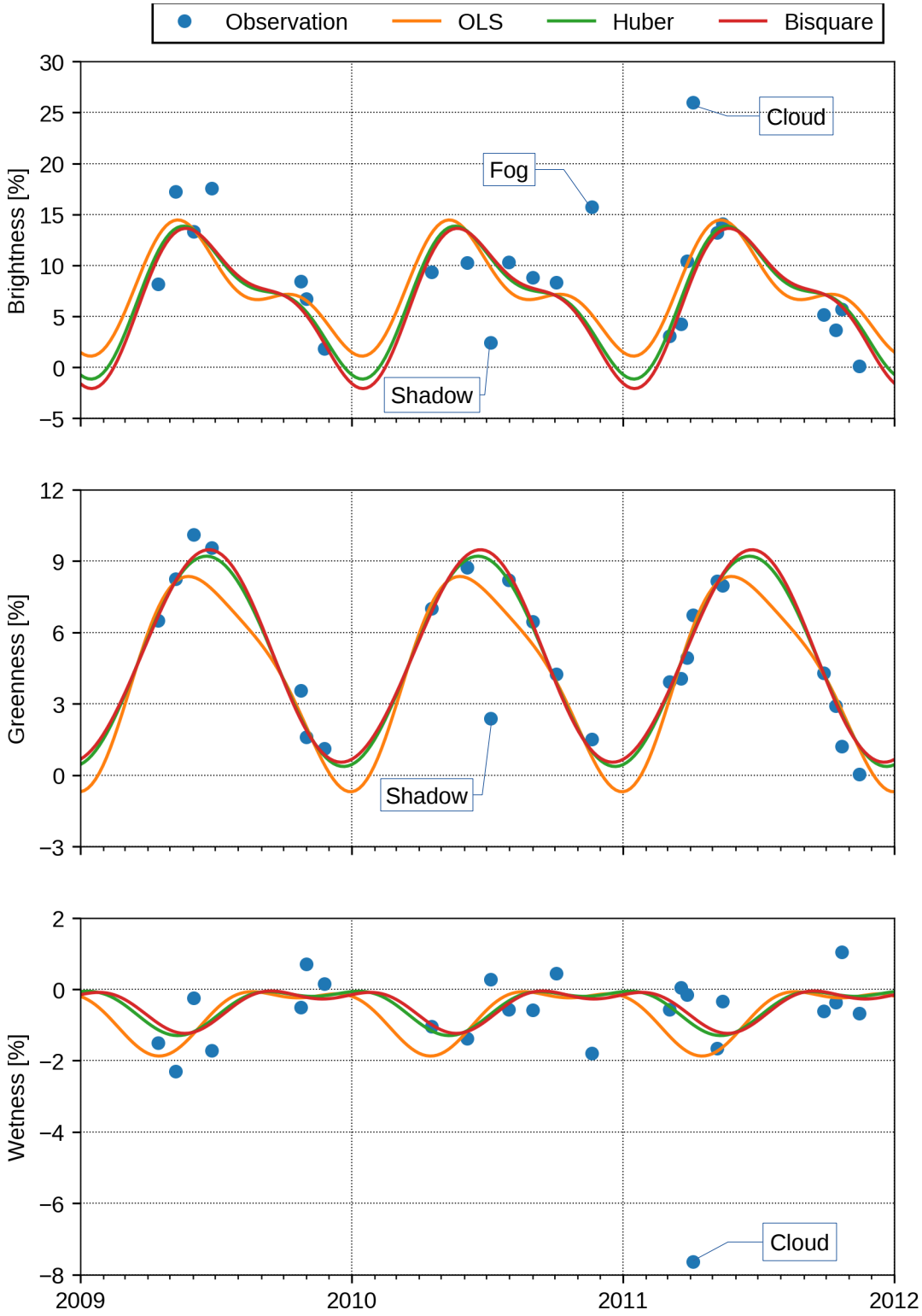

Figure 5.2: Different stages of the IRLS procedure, sample pixel 1.

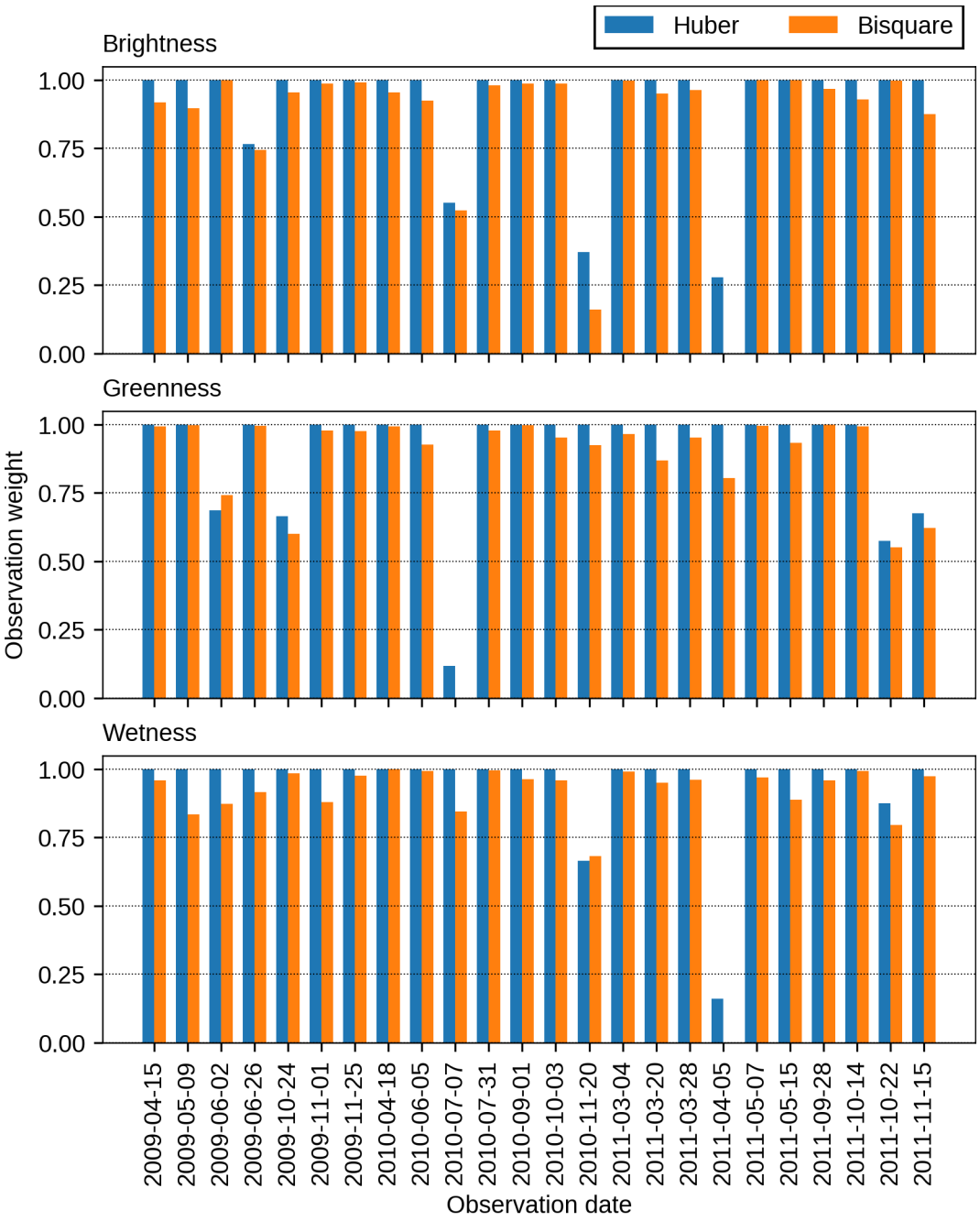

Figure 5.3: Resulting IRLS observation weights, sample pixel 1.

5.1.2 Spectral signatures of different forest types

The following Figures 5.4, 5.5, and 5.6 show comparisons of the spectral signatures of coniferous (sample pixel 1) and broadleaved (sample pixel 2) forest for each tasseled cap component. Additional to the discrete observations, the regression model fits including 99% confidence intervals are plotted. The coefficient of determination R^2 computed using (3.53) as well as the estimated observation standard deviation σ according to Equation (3.51) are given in the top left corner of each sub-plot. All three tasseled cap component time series behave quite differently depending on the forest type. The differences can be observed in several qualities including the overall level, the seasonal amplitude, and the "temporal smoothness" of the signal. The latter quality is most

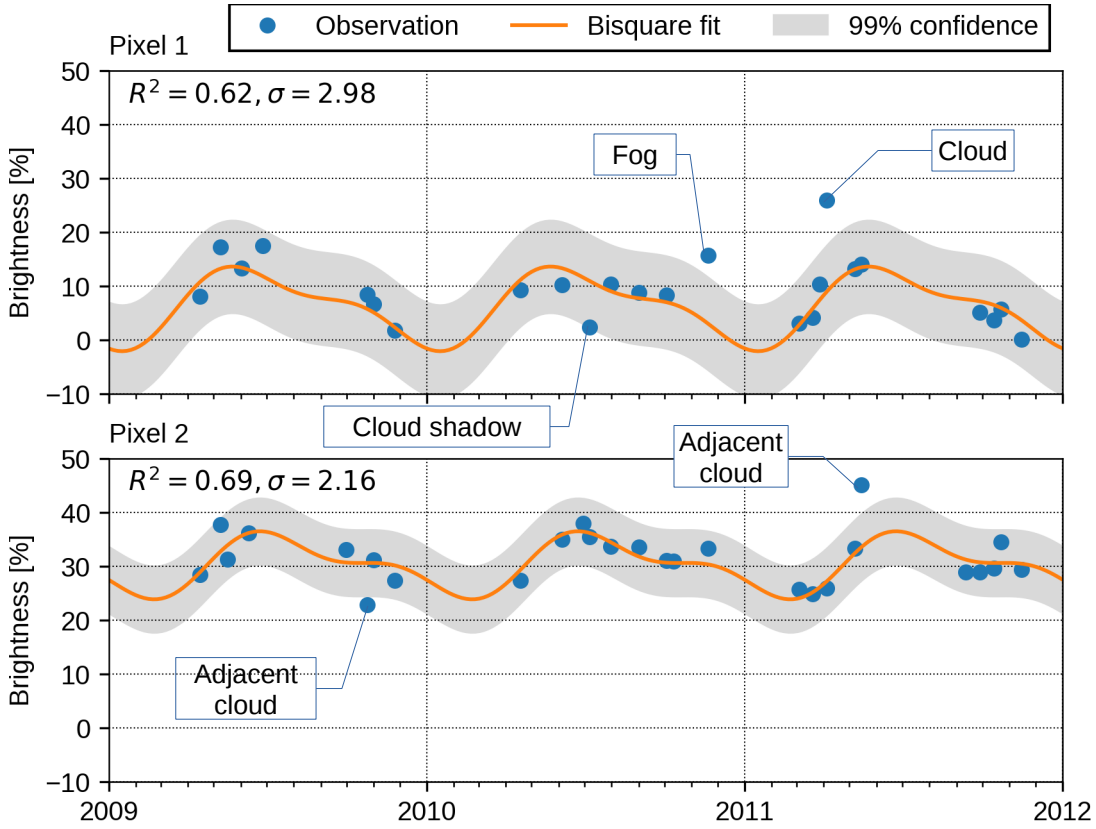

Figure 5.4: Brightness signatures of coniferous (Pixel 1) and broadleaved forest (Pixel 2)

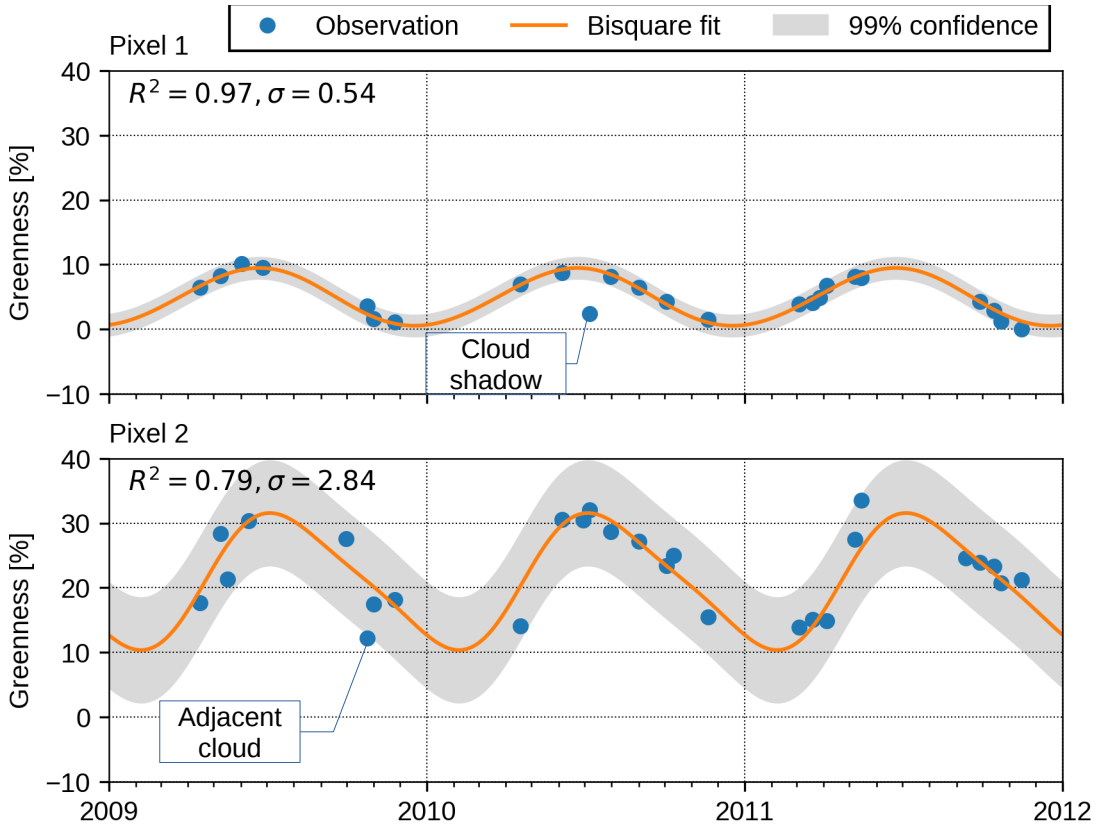

Figure 5.5: Greenness signatures of coniferous (Pixel 1) and broadleaved forest (Pixel 2)

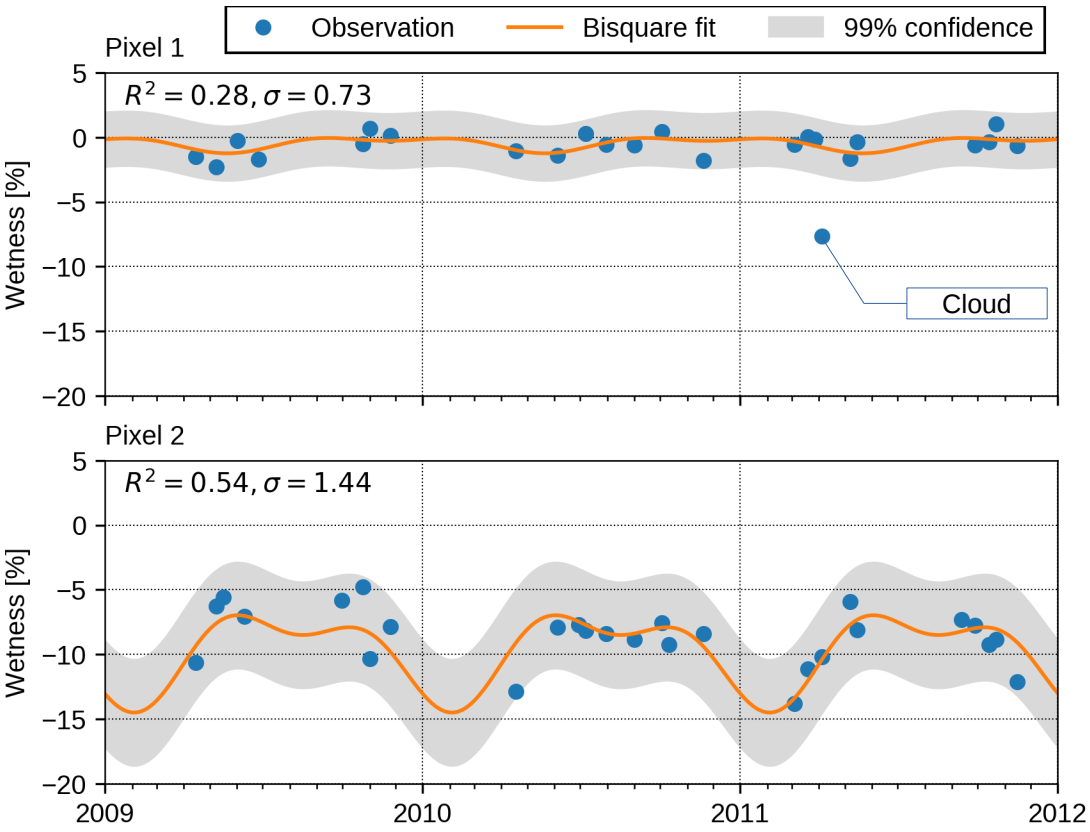

Figure 5.6: Wetness signatures of coniferous (Pixel 1) and broadleaved forest (Pixel 2)

distinct in the greenness time series of sample pixel 1, leading to a high coefficient of determination and a low estimated observation variance. The seasonal variations of the greenness and wetness signatures are more pronounced in deciduous forest (pixel 2). Foliation in spring is accompanied by steep increases of both signals, whereas leaf senescence and finally defoliation lead to larger decreases in comparison to evergreen pixels.

5.1.3 Comparison of models by means of a deciduous pixel

So far, the shown results were based on a 5-parameter regression model featuring a constant trend and two seasonal frequencies. In the next Figures 5.7, 5.8, and 5.9, the 5-parameter model is contrasted with a 3-parameter model featuring only a constant trend and the fundamental frequency. Considering that the seasonal variation is more pronounced in broadleaved forest (see previous section), sample pixel 2 is used. The results show that the 5-parameter model produces higher values of R^2 and lower estimations of σ. Due to the fact that there are additional degrees of freedom compared to the 3-parameter model, this is not surprising and cannot be used as sole evidence to claim that it is better. However, it can be seen that the steep increase of all TC components in spring can be modeled more accurately by including the frequency ω_2 into the model. For this example, the largest difference in the coefficient of determination can be observed in the wetness signal. Based on a qualitative assessment, the seasonal wetness variation is indeed captured more closely by the 5-parameter model, especially for the year 2010 where many observations are available in summer.

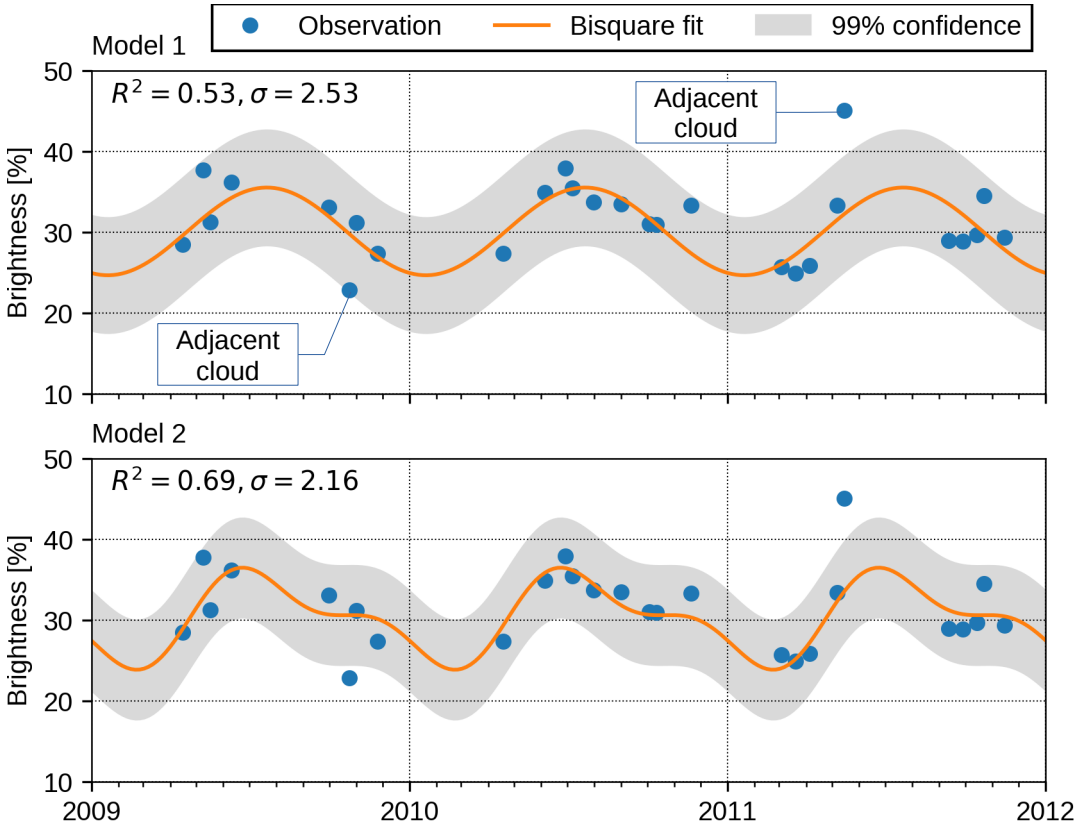

Figure 5.7: Brightness time series of sample pixel 2 (broadleaved forest) and fitted regression models. Model 1: 3 parameters, Model 2: 5 parameters.

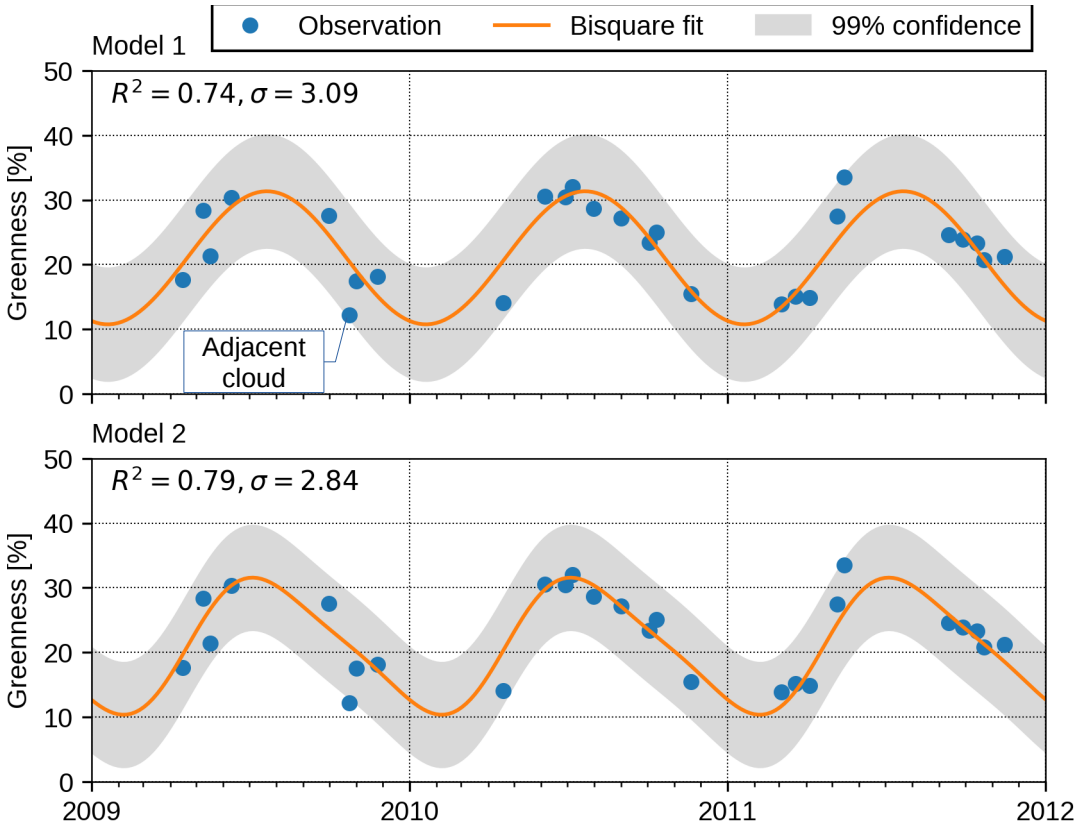

Figure 5.8: Greenness time series of sample pixel 2 (broadleaved forest) and fitted regression models. Model 1: 3 parameters, Model 2: 5 parameters.

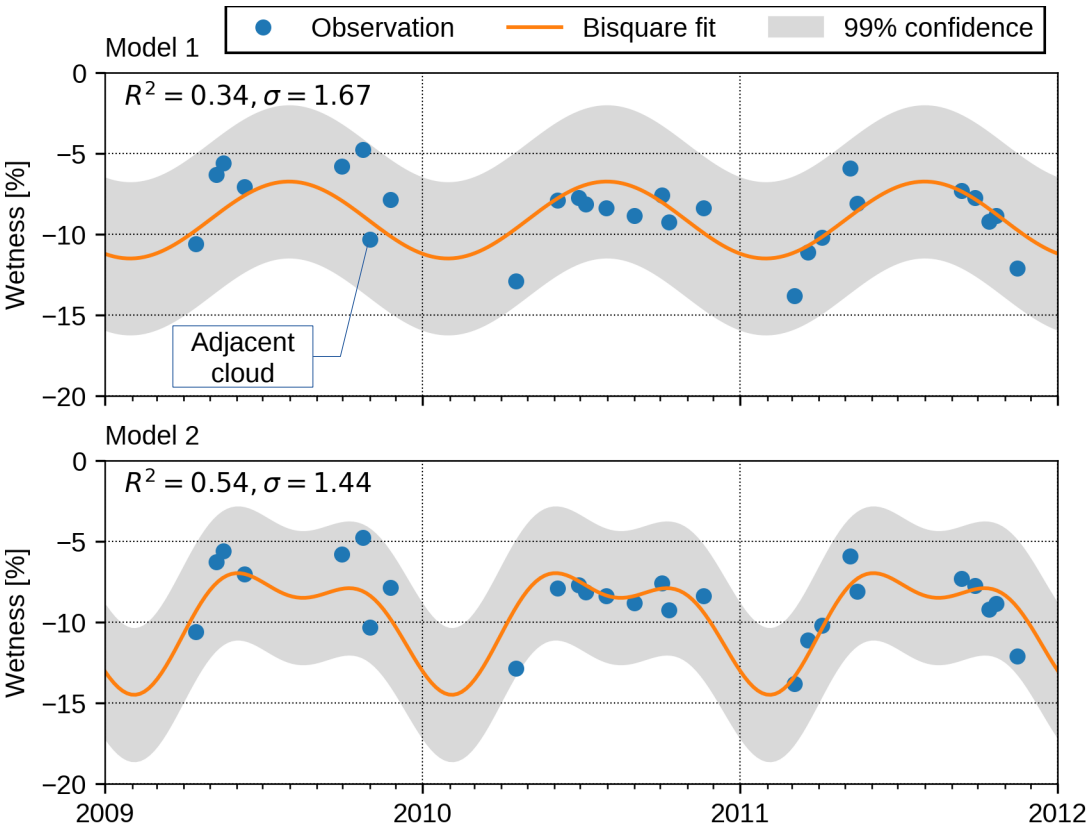

Figure 5.9: Wetness time series of sample pixel 2 (broadleaved forest) and fitted regression models. Model 1: 3 parameters, Model 2: 5 parameters.

5.1.4 Terrain influence

It was mentioned in Section 4.2 that the terrain of the test site is not flat and therefore systematic effects caused by the local topography may occur. Figure 5.10 shows a relief map of the test site on the left and a map of the R^2-values computed for the greenness

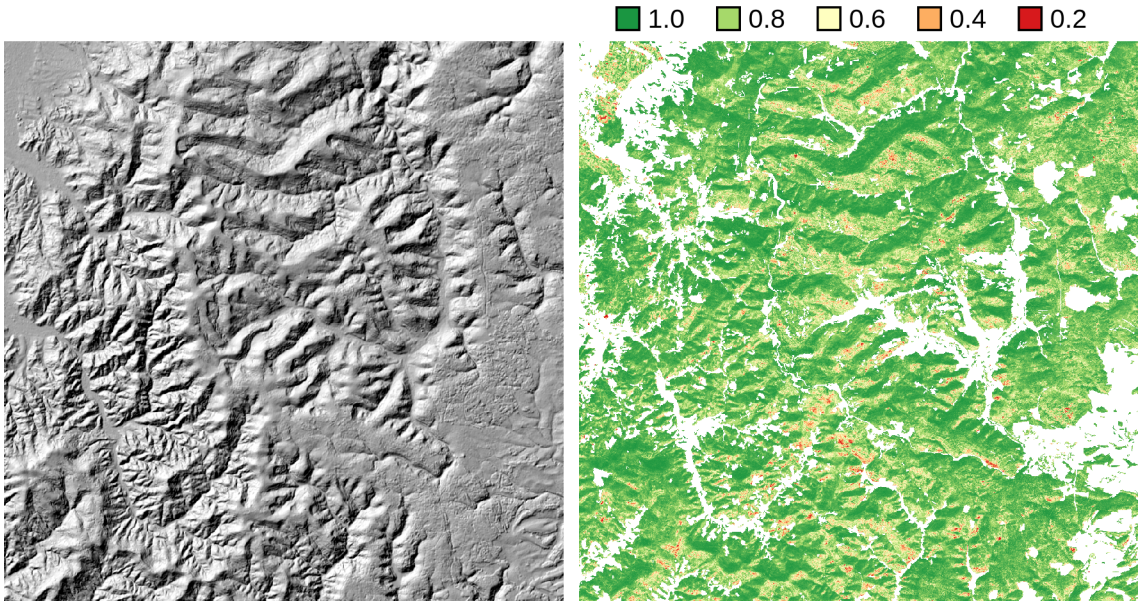

Figure 5.10: Terrain dependency of the coefficient of determination

signal on the right side. Again a 5-parameter model and 3-year training period were used. The illustration shows that the coefficient of determination is systematically lower in areas on southward-facing hillsides compared to northward-facing hillsides. A comparison on pixel-level is carried out using sample pixels 3 and 4, which are located 150 m apart on southward- and northward-facing hillsides respectively. The brightness and greenness time series depicted in the Figures 5.11 and 5.12 show that the investigated pixels have different seasonal patterns. Looking at pixel 4, both brightness and greenness are close to zero from late autumn to early spring because the northward-facing hillside lies in the shadow. The effect of the strong intra-annual illumination variance is superimposed on the phenology, resulting in a higher overall amplitude of the seasonal pattern compared to pixel 3. This effect is captured by the model, hence it is unlikely that shadows caused by the local topography will be a source of error in the change detection process. However, at this point an assessment whether this behavior is desired or not cannot be made. Further investigations including comparisons to topographically normalized data are necessary. Finally, Figure 5.13 shows that the dynamic range of the wetness signals is low regardless of the hillside, which is also in agreement with plots of coniferous pixels shown in previous sections.

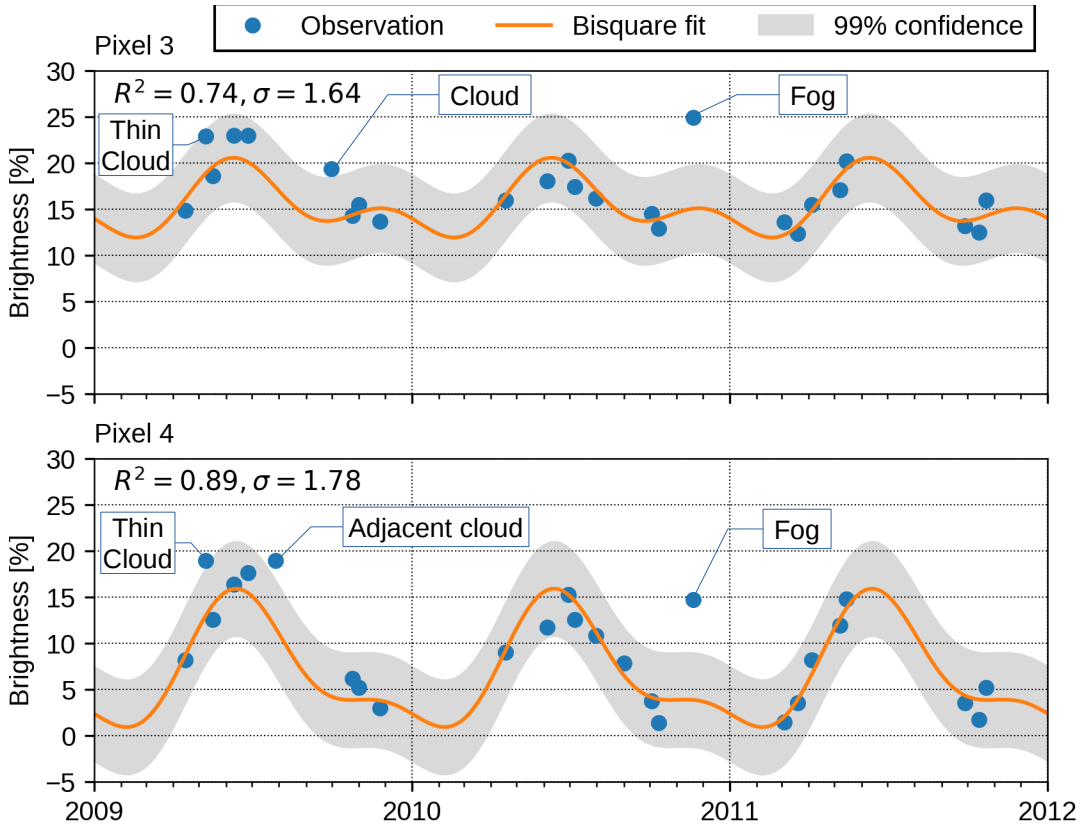

Figure 5.11: Terrain dependency of the brightness signal. Pixel 3: southward slope. Pixel 4: northward slope.

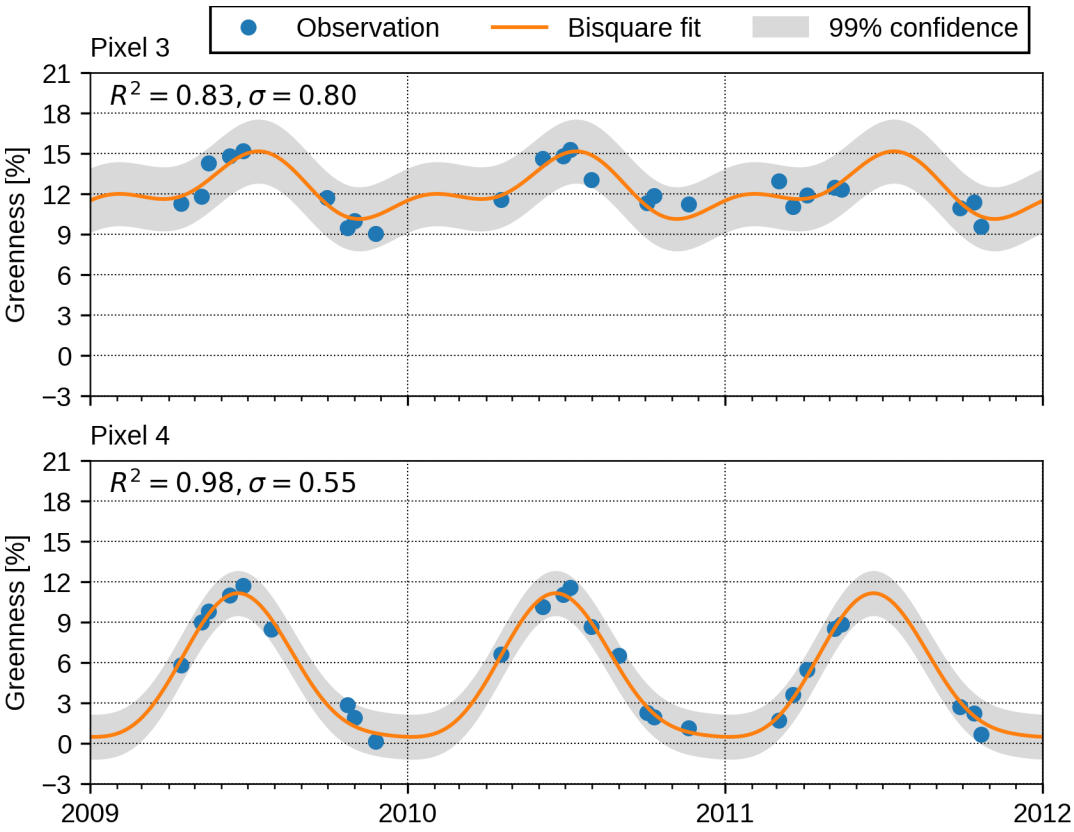

Figure 5.12: Terrain dependency of the greenness signal. Pixel 3: southward slope. Pixel 4: northward slope.

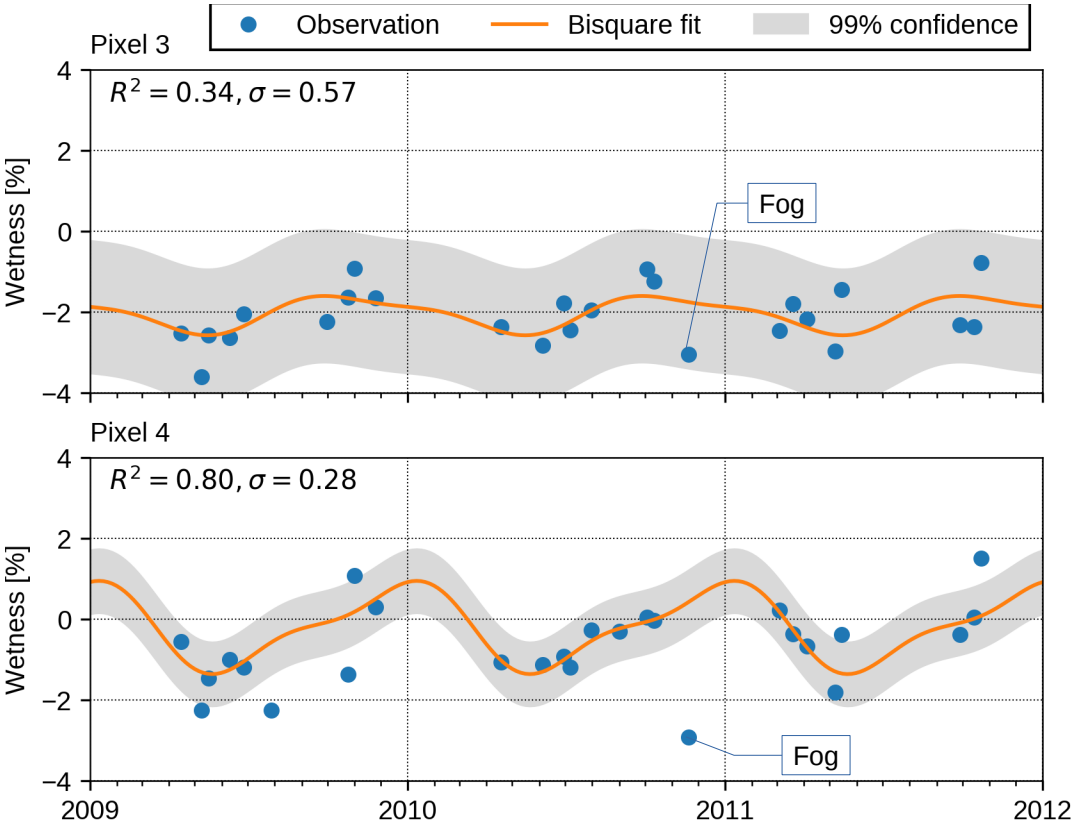

Figure 5.13: Terrain dependency of the wetness signal. Pixel 3: southward slope. Pixel 4: northward slope.

5.1.5 Conclusions concerning the Kalman filter setup

Even though Section 5.1 mostly revolves around the analysis of a few pixels, the results indicate that the proposed approach for estimating the initial filter state can handle various circumstances. A proper initialization is of high importance, because the ability to predict future observations based on a statistical model is an integral part of the Kalman filter change-detection algorithm (see Section 5.2) outlined in Figure 3.4. The reader may recall from Section 3.4.4 that the differences between predicted and actually observed measurements are called innovations. Large absolute innovation values mean that measurements strongly deviate from the prediction. The change detection algorithm uses a χ^2 hypothesis test on the filter innovations to identify statistically "suspicious" observations and thus accomplish two things:

1. The update of the filter state with information from invalid observations due to cloud cover etc. is avoided.
2. The temporally aggregated occurrence of large innovation values indicates an abrupt structural change in the time series.

The variables which tune the hypothesis test are reviewed here. According to Equation (3.41), the outcome of the statistical test is determined by the value y_k of the innovation itself, but also by its associated variance C_k. The innovation variance depends on the uncertainty of the predicted state, reflected by $\widetilde{\mathbf{P}}_k$, as well as the observation variance R, see (3.39). Since the state prediction needs to be known with low levels of uncertainty in order to be able to identify anomalous observations in the first place, the value of C_k, and therefore the sensitivity of the statistical test, is indeed to a large part governed by R alone.

The proposed change-detection algorithm assumes that R is set independently for each time series based on the observation variance $\hat{\sigma}^2$ estimated from the training period. Note that the computation follows Equation (3.51), hence any calculated observation weights also influence the value of $\hat{\sigma}^2$. By looking at the Bisquare weighting function illustrated in Figure 3.3 it can be seen that even observations with low residuals receive weights lower than 1, thus the resulting estimate of the observation variance is more optimistic compared to an OLS equivalent. Some of the plots presented in the preceding sub-sections, namely the Figures 5.5, 5.12, and 5.13, suggest that the resulting values of R can be too low. Low values of R lead to a high sensitivity of the statistical test. In order to reduce the false alarm rate, it is proposed to limit the lower boundary of R to 1, that is

$$R = \begin{cases} \hat{\sigma}^2 & \text{if } \hat{\sigma}^2 > 1 \\ 1 & \text{otherwise.} \end{cases} \tag{5.1}$$

Note that (5.1) is only valid if the time series observations are scaled to a range between −100 and 100 to be interpreted as percent.

Another implementation detail which needs to be addressed is given by the fact that the regression models used throughout Section 5.1 defined a constant *deterministic*

trend component, whereas the state-space definition of sub-section 3.4.1 mentioned a *local linear* trend. The explanation of this inconsistency lies in the length of the training period. For a stable, slowly evolving land cover type like forest, the inclusion of a linear trend component in a regression model which is fitted to only 3 years of data does not make sense and could even lead to distortions. On the other hand a linear trend may emerge over a longer time scale, thus it is advantageous to include an additional slope parameter in the Kalman filter model. In the initial state, the value of the slope is zero. The corresponding main-diagonal element in $\mathbf{P}_0$ can be set to a low value, but always >0 to allow a stochastic evolution of the slope parameter. A value of 0.005^2 proved to be applicable.

5.2 Kalman filter application for change detection

Before the Kalman filter results are discussed, a summary of all the tasks preceding the change detection is given:

1. Definition of the time series model components
2. Definition of the training period
3. Computation of $\mathbf{x}_0$, $\mathbf{P}_0$ and R, for each time series independently
4. Definition of the global change-detection tuning parameters

The first three tasks have been discussed in the previous section. Regarding task 4, Table 9 below lists the set of tuning parameters used to produce the results presented in this section. Note that the given values are not optimized by a benchmarking process due to a lack of appropriate ground truth data. However, the set of reasonable values for each tuning parameter is limited and the given values were derived from empirical tests.

Table 9: Change-detection tuning parameter settings

Parameter	Description	Value
α	Significance level of the outlier hypothesis test	0.01
Change threshold	Threshold value of the anomaly counter for a change to be signaled	3
Q_c trend	Process noise, trend component	$(2.5\times10^{-4})^2$
Q_c seasonal	Process noise, seasonal component(s)	$(2.5\times10^{-2})^2$

While the interpretation of α and the change threshold is quite obvious, the values for the process noise parameters are more abstract. In order to find reasonable values, Gibbs (2011, chap. 2) suggests to apply the equations for the computation of $\mathbf{Q}_k$ given in (3.24) and (3.32) in an inverse fashion, thus set the desired level of process noise related to a certain time interval Δt and solve for Q_c. For example, the equation to compute the process noise of the trend level is

$$q_k = \frac{\Delta t^3}{3} Q_c. \tag{5.2}$$

Note that both q_k and Q_c are variances. If the model error of the trend level is expected to be 1% per year, rearranging (5.2) and substituting yields

$$Q_c = \frac{3 \times 1^2}{365.25^3} \approx 0.00025^2. \tag{5.3}$$

The process noise level of the seasonal components given in Table 9 corresponds to about 3% model error per year, for each frequency.

5.2.1 Undisturbed pixel

At first, an exemplary result of the Kalman filter application to an undisturbed coniferous forest pixel (sample pixel 1) will be discussed. The Figures 5.14, 5.15, and 5.16 depict all available un-masked observations as well as the filtered time series for a given tasseled cap component in the upper subplot, while the lower subplot illustrates the result of the outlier test for each observation. The same pixel has been discussed in Section 5.1.1, therefore the three anomalous observations shown in Figure 5.1 also appear here and can be seen in the brightness signal plotted in Figure 5.14. They are correctly identified by the outlier test integrated in the Kalman filter.

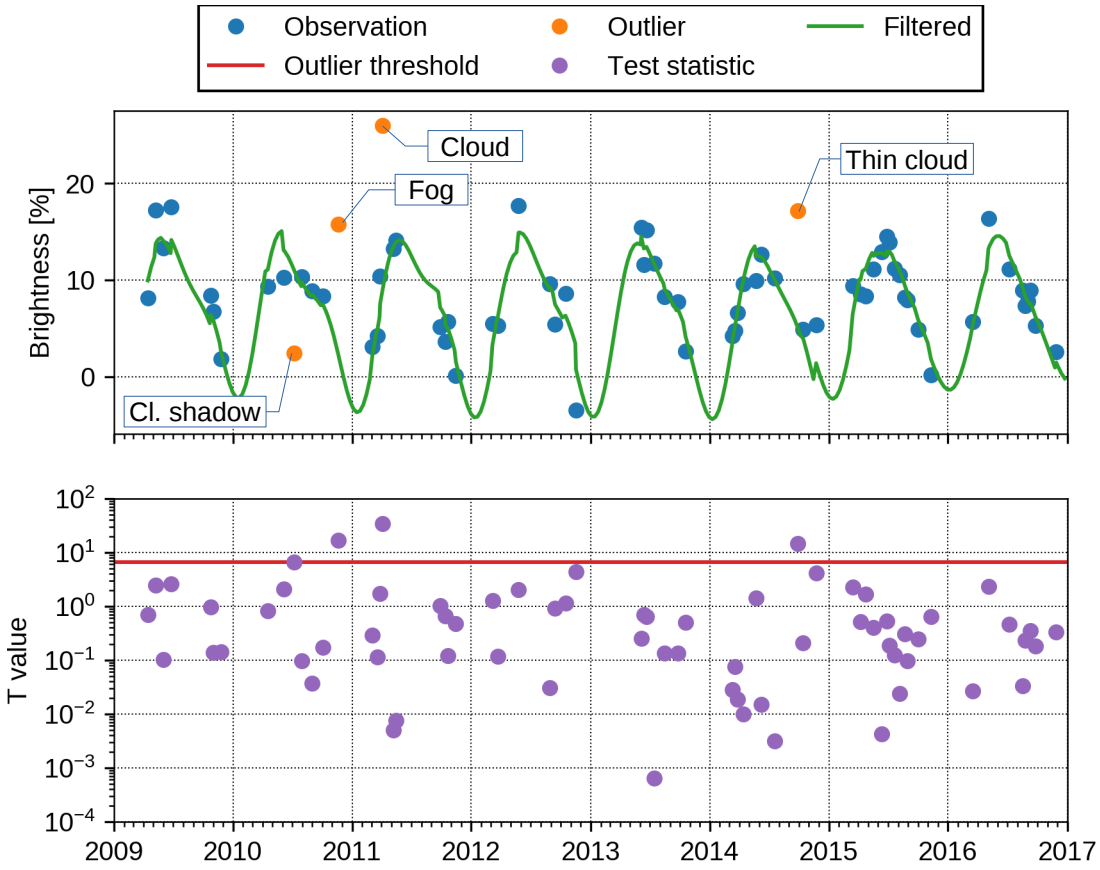

Figure 5.14: Filtered brightness signal of sample pixel 1 (undisturbed, coniferous)

Variations of the yearly peak level can be observed in the greenness signal plotted in Figure 5.15, for example from 2011 to 2012. The filter state is updated heavily in 2012 when that year's peak observation is processed. Because of this update, the prediction

for spring 2013 overshoots and the first observation of 2013 is incorrectly marked as outlier (labeled as “False alarm” in the figure). However, the second observation passes the test and the state is updated accordingly. It has also to be kept in mind that the observation density is exceptionally low in the year 2012 and the first quarter of 2013, because the only operational satellite at that time was Landsat 7.

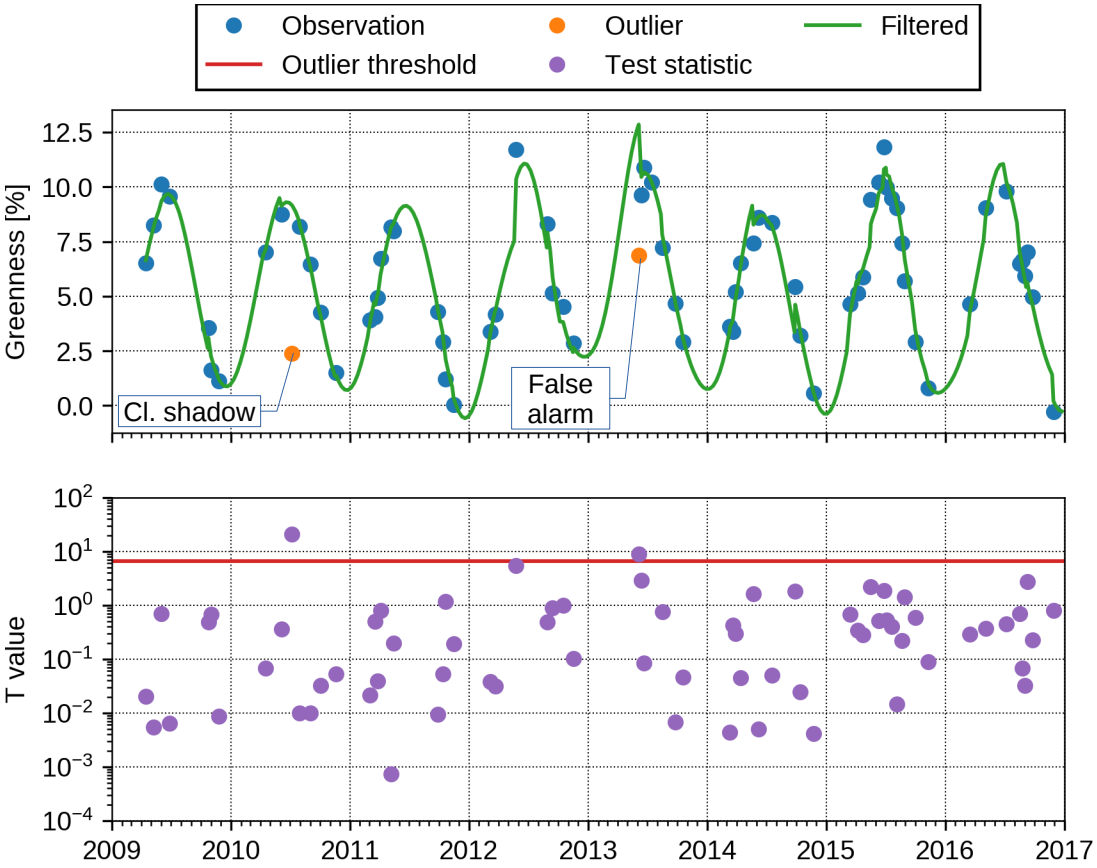

Figure 5.15: Filtered greenness signal of sample pixel 1 (undisturbed, coniferous)

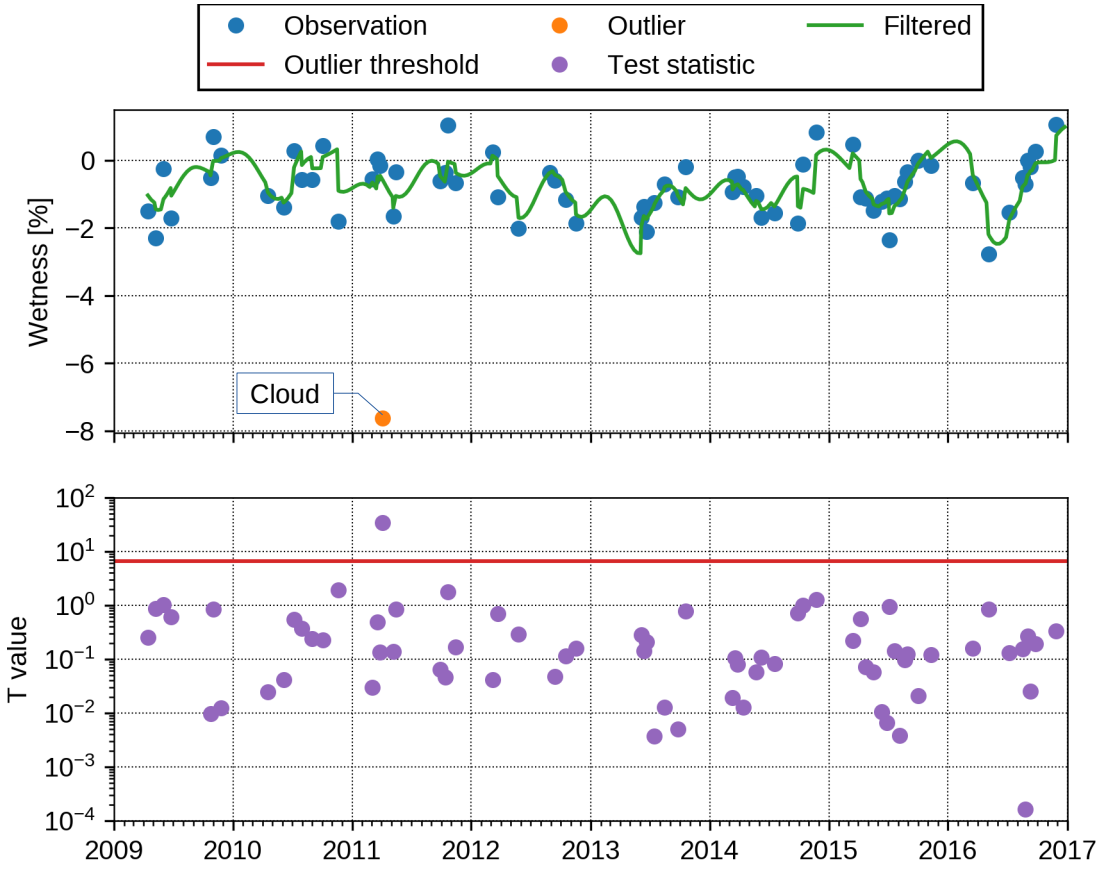

Figure 5.16: Filtered wetness signal of sample pixel 1 (undisturbed, coniferous)

5.2.2 Abrupt change – single pixel

Sample pixel 5 (coniferous, storm damage 2012) is used to demonstrate the Kalman filter result for time series exhibiting abrupt change. The change event resulting from storm damage occurred in July 2012. Unfortunately the observation density is very low in 2012 up to the first quarter of 2013 due to the failure of Landsat 5. Furthermore, the pixel is located on a northward-facing hillside which means that the signal is already very weak in November due to the low sun elevation angle. As a result, a time gap of more than a year lies between the change event and the change signal date. Nonetheless this pixel has been selected because many aspects of the filter operation can be discussed using a single example.

Figure 5.17 shows the brightness signal. The change event is followed by an abrupt increase of the brightness level in July 2012. While the first available observation after the change event is identified as outlier, the second is too close to the prediction. The last available observation of 2012 was acquired in November, where the weak signal is also in agreement with the prediction. What follows is an unusually large time interval to the first observation of 2013. Due to the cumulative effect of the process noise during this observation-free interval, the prediction error increases and the sensitivity of the outlier test decreases, hence the new observations of 2013 are also not identified as outliers.

The greenness signal illustrated in Figure 5.18 shows no abrupt change after the change event, thus the filter can follow the observations and no outliers are signaled. Figure 5.19 shows the wetness signal, where the change event is followed by an abrupt decrease of the wetness level. Two outliers are identified after the change event, but again, due to the terrain influence, the last observation of 2012 is not marked as outlier. As a result two more observations in 2013 are required for the anomaly counter to reach a value of 3 and flag the pixel as changed.

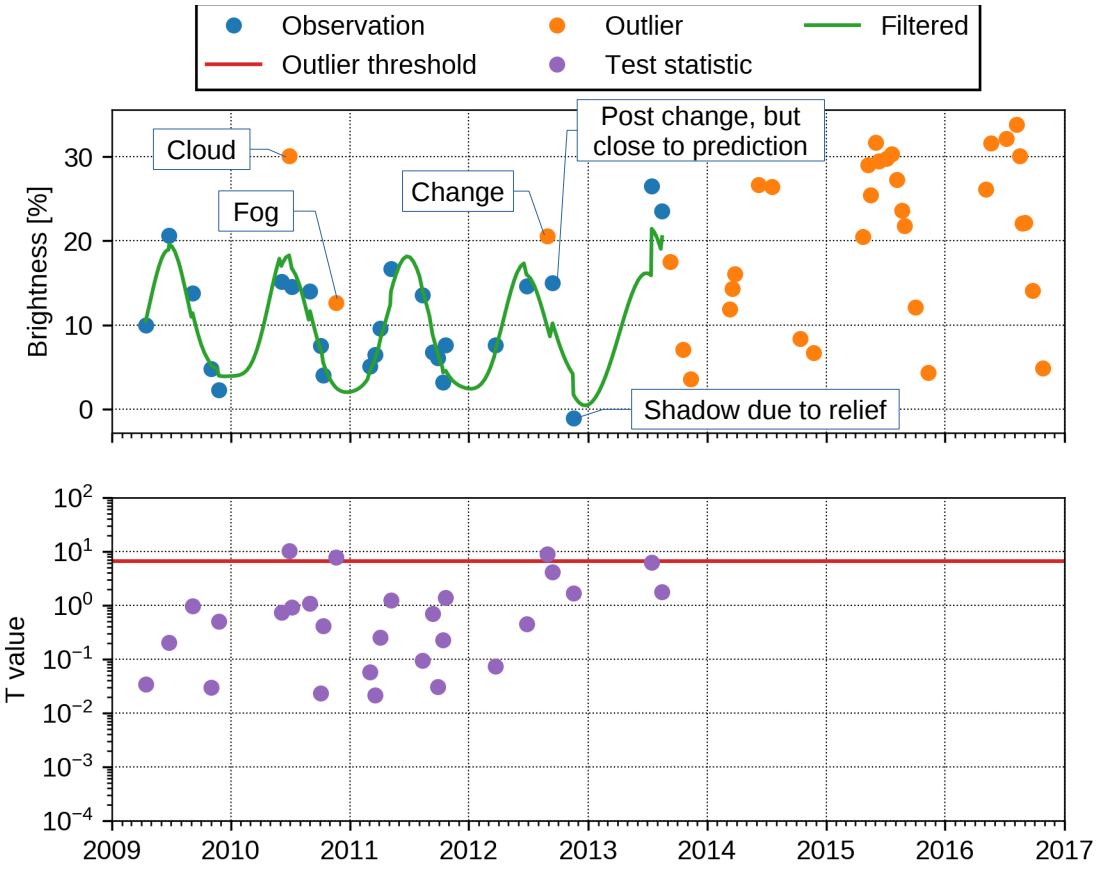

Figure 5.17: Filtered brightness signal of sample pixel 5 (change 2012, coniferous)

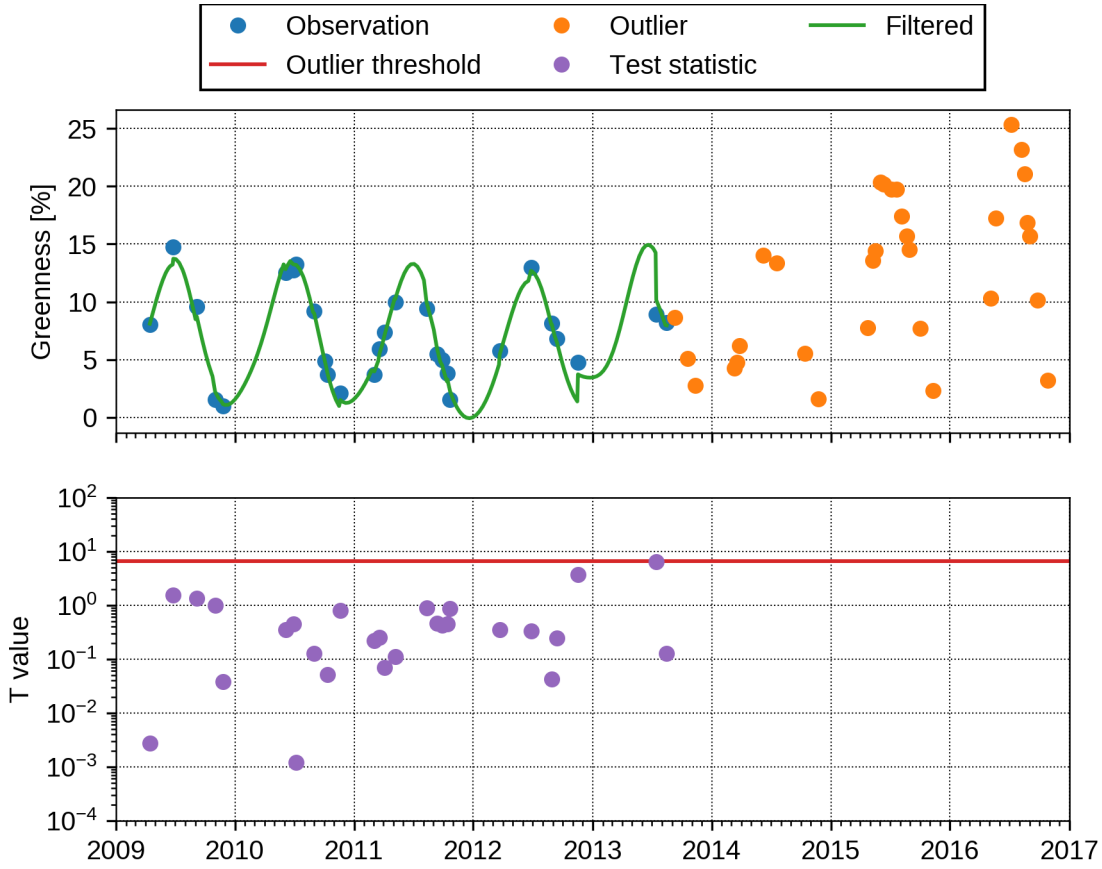

Figure 5.18: Filtered greenness signal of sample pixel 5 (change 2012, coniferous)

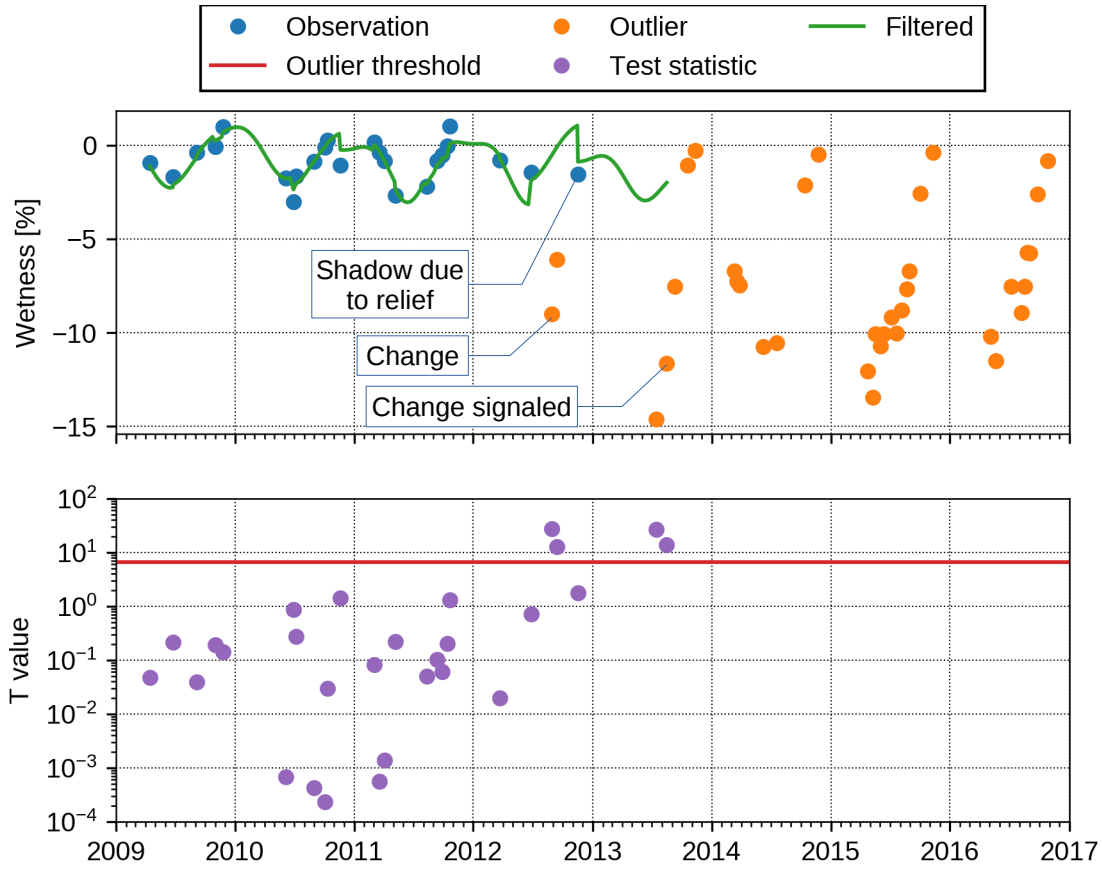

Figure 5.19: Filtered wetness signal of sample pixel 5 (change 2012, coniferous)

5.2.3 Abrupt change – image sequence

Figure 5.20 depicts an image matrix which documents a clipped sub-area of the demonstration site where the storm damage occurred. The first row shows an RGB composite and the last row shows a false-color composite (SWIR1, NIR, red). The time domain is reflected in the columns of the image matrix, where the first column represents the state before the damage. The rows 2, 3, and 4 of the image matrix show the Kalman filter innovations and therefore the difference between predicted and actual observations for each tasseled cap component. The same increase in brightness and decrease in wetness as discussed in the preceding sub-section can be observed.

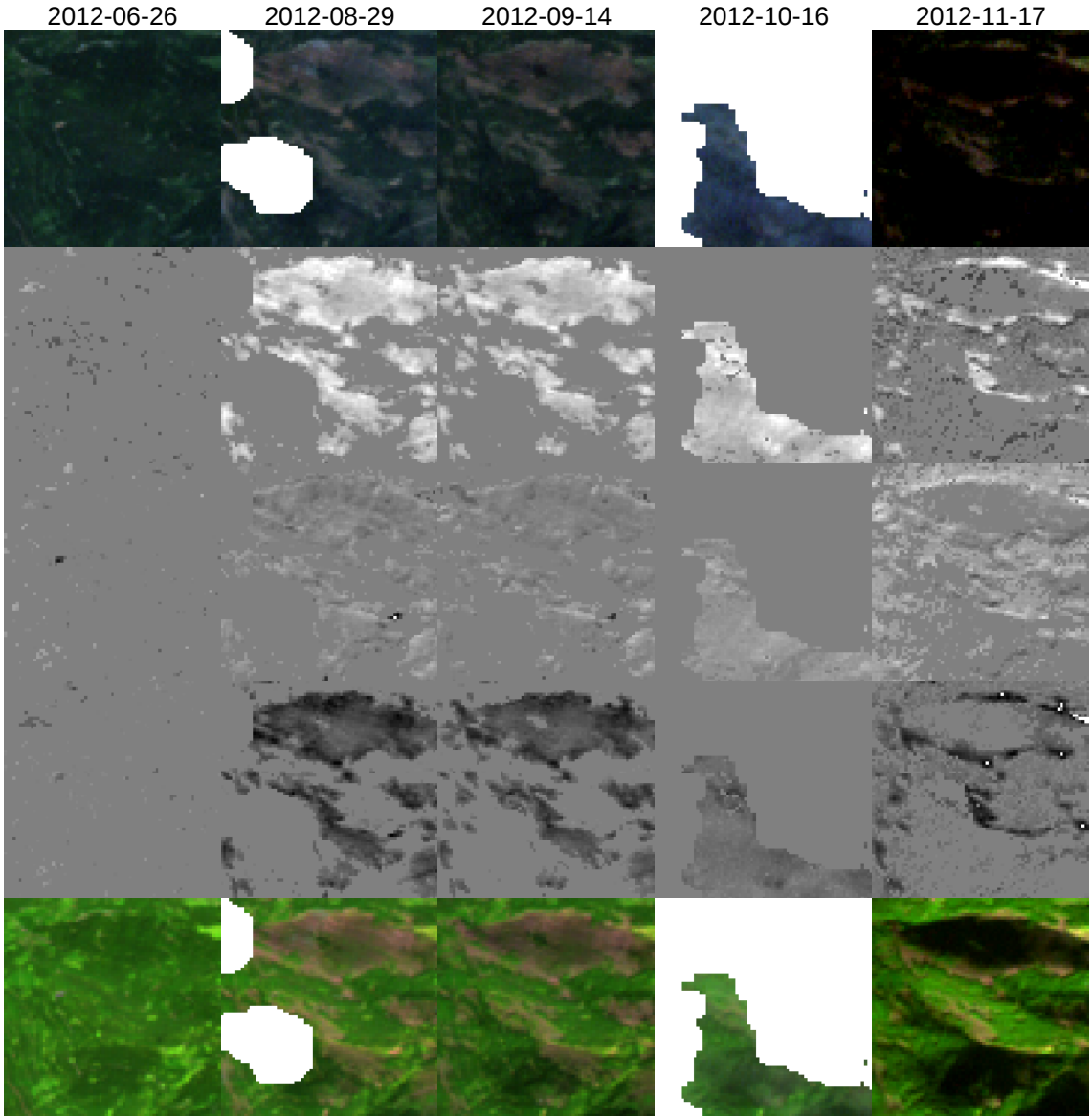

Figure 5.20: Image sequence of a storm damage occurring in 2012. 1st row: RGB image - 2nd row: Brightness filter innovations - 3rd row: Greenness filter innovations - 4th row: Wetness filter innovations - 5th row: False-color image. Light shades of gray mean that an observation is significantly higher than the prediction, whereas dark shades of gray mean that it is lower. A balanced gray value designates masked values or observations close to the prediction.

Column 4 gives an example of the appearance of thin, un-masked clouds. Note that the thin cloud causes spectral signal changes similar to those observed in areas affected by the storm. The influence of shadows caused by the relief can be seen in column 5. A large part of the areas affected by the storm lies in the shadow during the winter months, which leads to a delay of the change signal date if not enough post-change observations are available before the winter period. A detailed discussion of this effect on pixel-level is given in Section 5.2.2.

5.2.4 Change maps

For each processed spectral band or index, the main output of the change-detection algorithm is a map containing information about which pixels were flagged as changed and when. The processed time series of Landsat images covers a time span of 8 years from 2009 to 2016. Subtracting the initial 3 years used as training data leaves a monitoring period ranging from 2012 to the end of 2016. In order to perform a thorough statistical evaluation of the results, a ground-truth data set representative for the whole spatial and temporal extent of the test data would be required. Unfortunately no such data set is available. However, there is a reference map provided by the Chair of Remote Sensing and Landscape Information Systems, University of Freiburg, which documents areas affected by the storm in the summer of 2012. The areas were manually digitized after a visual interpretation of Landsat 7 images. Three different change maps have been created by applying the change-detection algorithm to the time series of the 3 tasseled cap components brightness, greenness, and wetness one at a time. Each of the derived change maps will be evaluated within the spatial and temporal confines of the available reference data set.

At first, the result obtained by processing the brightness signal shall be discussed. Figure 5.21 shows an overlay of the result and the reference map. Each pixel designated to be changed is colored according to the respective date of detection. The provided legend assigns unique colors for the 10 most frequently occurring detection dates. Pixels corresponding to any other detection date are colored in black. After a quick visual assessment, the reader may note that most of the larger, connected areas are roughly in agreement with the reference map, but there are also many small detections scattered across the map. It is unlikely that these small, isolated areas with sizes of a few pixels correspond to actual change events. Therefore they are considered as commission errors and will be filtered in an additional post-processing step. Prior to the discussion of any post-processed results, the confusion matrices based on the unedited results for each tasseled cap feature are given in Table 10, 11, and 12 respectively. The total sample size of $N = 74\,641$ corresponds to the number of forest pixels (according to the forest mask) within the bounding rectangle of the reference map. Both the brightness and wetness based map identify changed pixels with producer and user accuracies over 70%, with a maximum of 80.6% reported for the producer accuracy of the wetness map. The findings of the preceding sub-section already suggested that the greenness signal is less suitable for the detection of storm damages, which is confirmed by the respective confusion matrix. A further discussion of the results will be given after the presentation of the post-processed versions.

Table 10: Confusion matrix of the unedited brightness change map

		Reference			
		Change	No change	Total	User accuracy
Classification	Change	4926	1794	6720	73.3%
	No change	1910	66 011	67 921	97.2%
	Total	6836	67 805	74 641	
	Producer accuracy	72.1%	97.4%		

Table 11: Confusion matrix of the unedited greenness change map

		Reference			
		Change	No change	Total	User accuracy
Classification	Change	3350	3098	6448	52.0%
	No change	3486	64 707	68 193	94.9%
	Total	6836	67 805	74 641	
	Producer accuracy	49.0%	95.4%		

Table 12: Confusion matrix of the unedited wetness change map

		Reference			
		Change	No change	Total	User accuracy
Classification	Change	5511	2212	7723	71.4%
	No change	1325	65 593	66 918	98.0%
	Total	6836	67 805	74 641	
	Producer accuracy	80.6%	96.7%		

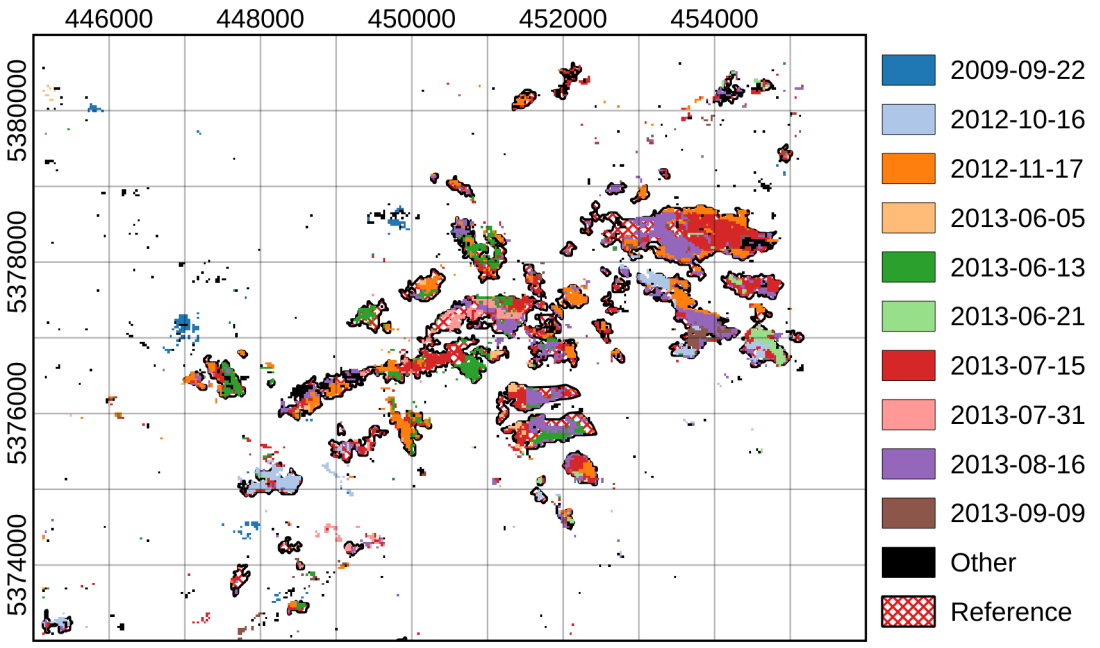

Figure 5.21: Unfiltered map of changes detected in the brightness signal compared to the reference map of storm damages in 2012. The legend indicates the 10 most frequent detection dates.

In order to remove the scattered small-scale detections, the maps have been post-processed by applying a modal-value-filter with a 3×3-pixel window. This filtering operation also closes small gaps in larger areas and smoothes their outlines. Furthermore the homogeneity of detection dates within connected areas is increased. For each tasseled cap feature, the filtered change map is given in Figure 5.22, 5.23, and 5.24 respectively. Comparing Figure 5.22 to Figure 5.21 shows that the selected filter operation effectively removes small objects.

Several observations can be made regarding the reported dates of detection. Although there is a quite large variation overall, they are distributed over many homogeneous patches. One explanation for this effect is given by the local variability in the set of available un-masked observations. As discussed in the preceding sub-sections, the terrain is a second factor of influence. Figure 5.25 shows a Landsat 7 image acquired in the late November of 2009, where all map regions affected by drop-shadows can be identified. By using the coordinate grid, the reader can track the correlation between shadow-areas and delayed change detection.

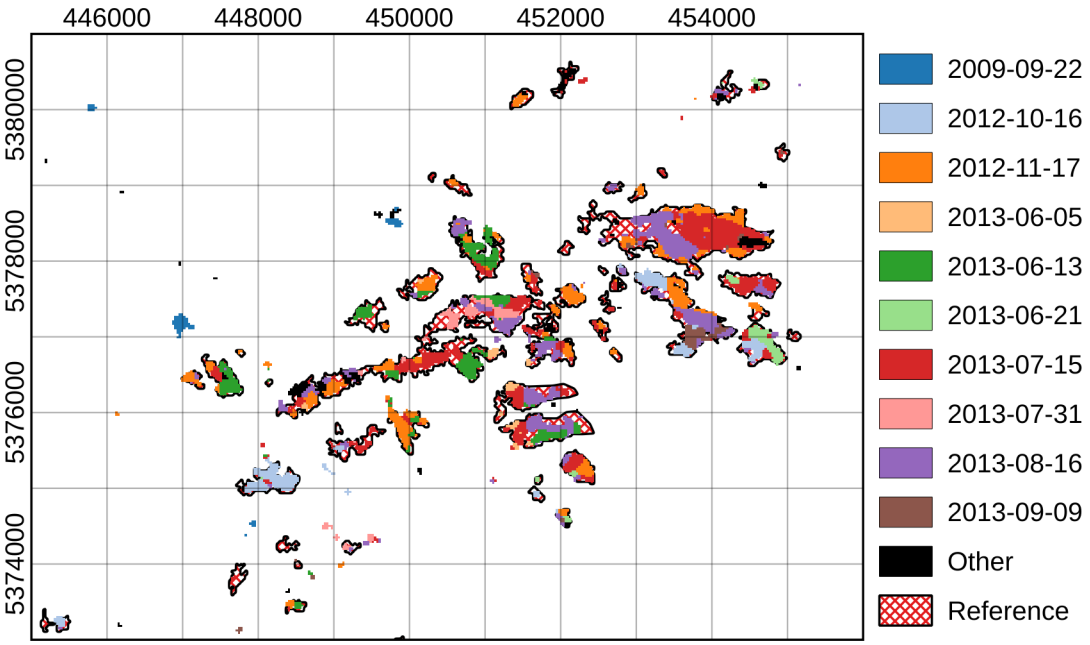

Figure 5.22: Filtered map of changes detected in the brightness signal compared to the reference map of storm damages in 2012. The legend indicates the 10 most frequent detection dates.

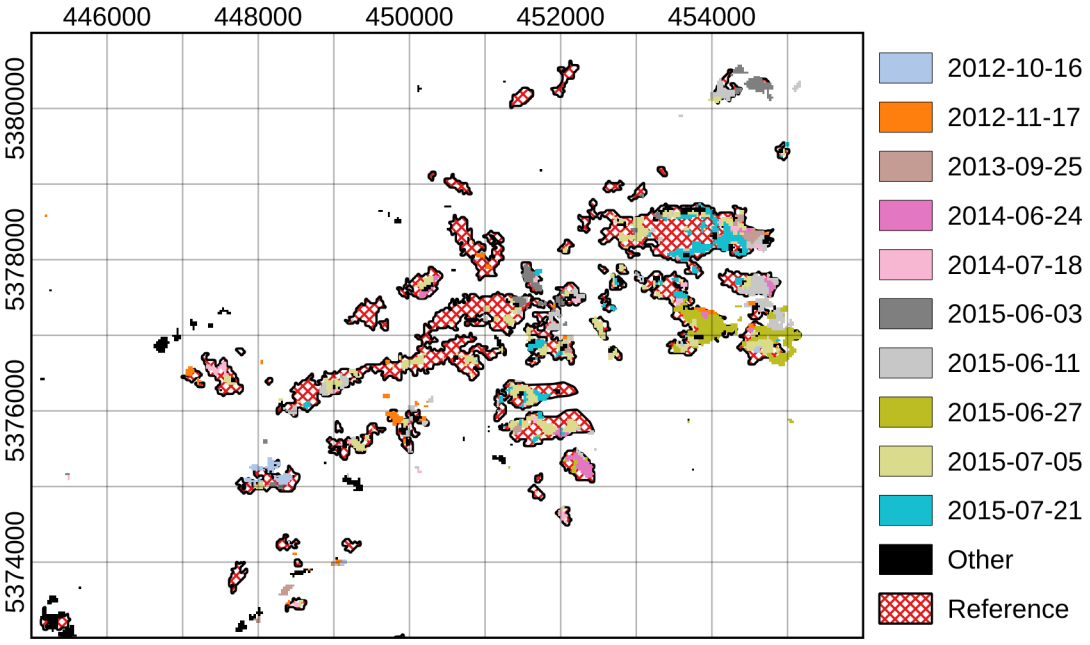

Figure 5.23: Filtered map of changes detected in the greenness signal compared to the reference map of storm damages in 2012. The legend indicates the 10 most frequent detection dates.

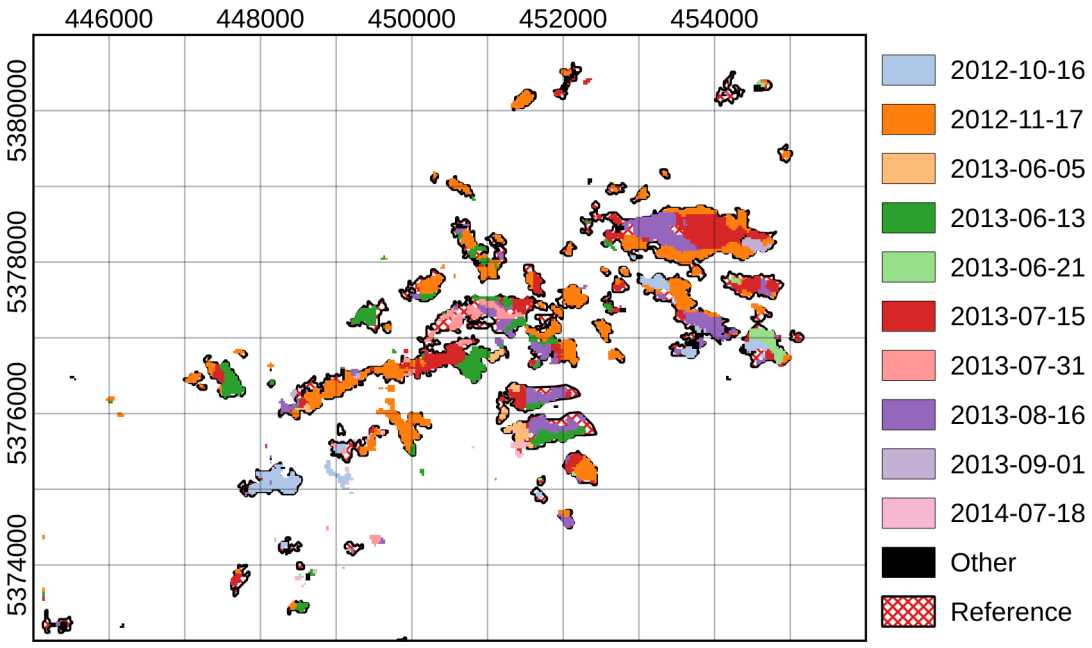

Figure 5.24: Filtered map of changes detected in the wetness signal compared to the reference map of storm damages in 2012. The legend indicates the 10 most frequent detection dates.

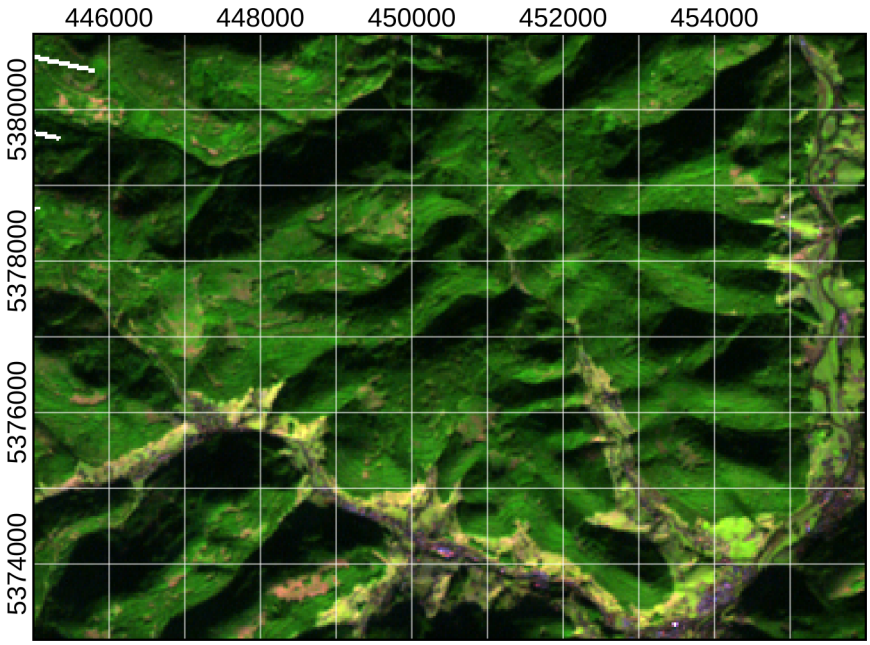

Figure 5.25: False-color composite (SWIR1, NIR, red) of a Landsat 7 image acquired on the November 25th, 2009. The image shows which areas of the change maps are affected by drop-shadows caused by the terrain and low sun elevation.

A visual assessment shows that the brightness and wetness change maps have large similarities. On the other hand, the result based on the greenness signal is very different. If a change is signaled at all, the time-delay is significantly larger in comparison to the other two results. A quantitative evaluation of the post-processed change maps is given by the corresponding confusion matrices in Table 13, 14, and 15. The post-pro-

cessing boosts user accuracies compared to the unfiltered results, which means that the number of commission errors decreased. Unfortunately the filtering also has an adverse effect on the producer accuracies, because originally correctly flagged change-pixels are also filtered out due to the smoothing effect. The best results are yielded by the wetness change map, where changed pixels are classified with a producer accuracy of 74.8% and a user accuracy of 86.1%.

Table 13: Confusion matrix of the filtered brightness change map

		Reference			
		Change	No change	Total	User accuracy
Classification	Change	4471	760	5231	85.5%
	No change	2365	67 045	69 410	96.6%
	Total	6836	67 805	74 641	
	Producer accuracy	65.4%	98.9%		

Table 14: Confusion matrix of the filtered greenness change map

		Reference			
		Change	No change	Total	User accuracy
Classification	Change	2609	1029	3638	71.7%
	No change	4227	66 776	71 003	94.0%
	Total	6836	67 805	74 641	
	Producer accuracy	38.2%	98.5%		

Table 15: Confusion matrix of the filtered wetness change map

		Reference			
		Change	No change	Total	User accuracy
Classification	Change	5115	827	5942	86.1%
	No change	1721	66 978	68 699	97.5%
	Total	6836	67 805	74 641	
	Producer accuracy	74.8%	98.8%		

5.2.5 Sources of error

The reported accuracies are influenced by several error sources. Omission errors may occur if the magnitude of the abrupt spectral shift after the change event is too low. An example is illustrated by Figure 5.26, where the wetness time series of sample pixel 6 is plotted. The pixel in question lies at the border of a region affected by the storm and may in fact be a “mixed” pixel. Unusually large temporal intervals between

consecutive observations also increase the probability of omission errors, because the sensitivity of the underlying hypothesis test used for change-detection decreases linear with time, see Sections 3.3 and 3.4.4.

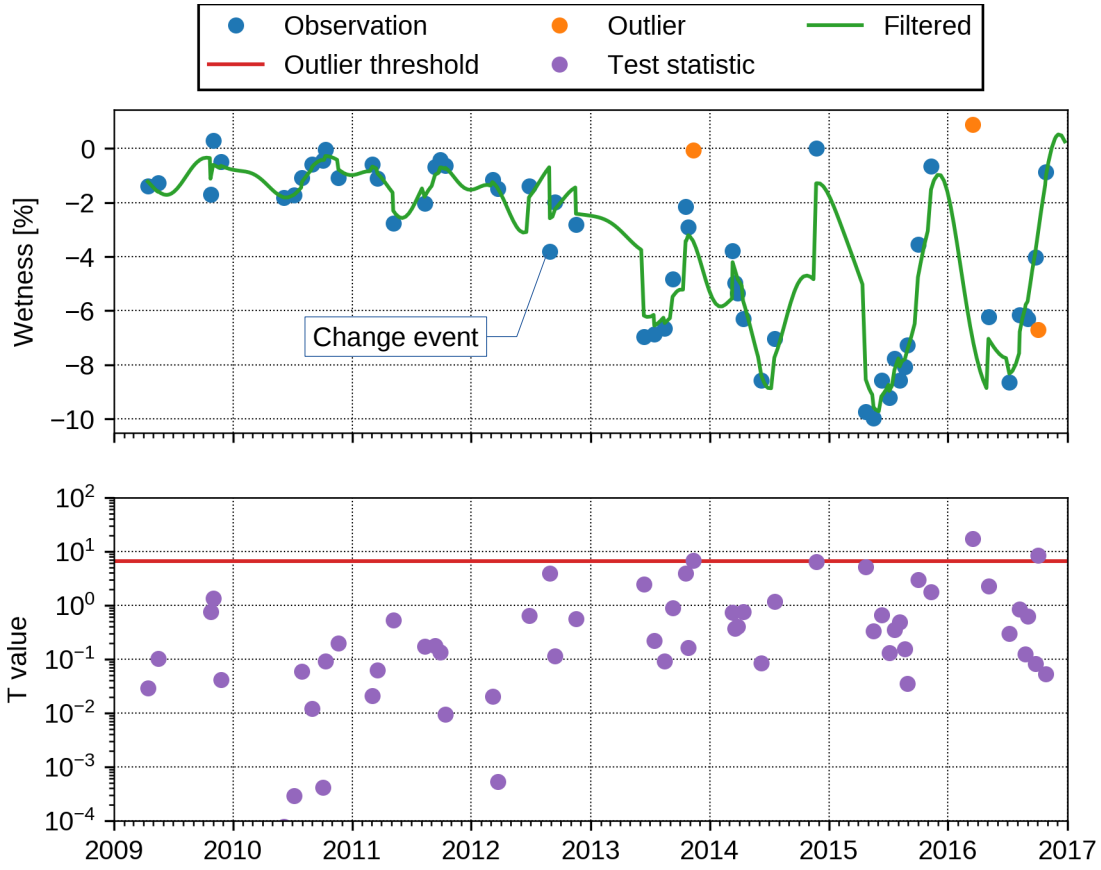

Figure 5.26: Sample pixel 6. Omission error due to an initially low magnitude of spectral change. The Kalman filter updates the state gradually and matches the post-change pattern.

Commission errors may be triggered if a spuriously good fit has been obtained during the initialization process. In this case the observation variance can be underestimated and the hypothesis test used for change-detection becomes too sensitive for the given time series. Another possible commission error source is presented by the fact that the available ground-truth data set is referenced to a single change event, but the applied change-detection algorithm has a continuous character. Correctly identified changes may simply not be reflected by the reference map if the change event occurred at a later time. This especially concerns areas which have been harvested in the course of the clearance of the windthrow areas.

5.2.6 Conclusions concerning the change maps

Considering the limited overall density of the processed historical Landsat time series, the achieved results are promising. The evaluation shows that systematic, large-scale commission errors due to un-masked clouds are avoided by the implemented algorithm. No other major flaws were revealed by the statistical comparison with a reference map. However, the results suggest that the local topography in the area of the

reference map has a major influence on the results. At this point it is probably necessary to put this particular conclusion into perspective. It is important to take into account that the extreme topographical dependency of the results presented in Section 5.2 is to a large part caused by the extraordinary low observation density in 2012 due to the failure of Landsat 5 at that time. If by chance more post-change observations had been available before November 2012, the resulting change maps would appear a lot more homogeneous and the average time-delay between the change event and its detection would also be reduced. On the one hand, an emphasis on the analysis of topographical effects is necessary to understand the results, but on the other hand it should not cloud the overall assessment of the practicability of the algorithm itself. Furthermore, a more detailed quantitative evaluation of the time-delays between the change event and its detection based on the available reference data would not be representative because of the bias introduced by the special satellite constellation prevailing at the time and has therefore been left out.

6 Summary and Outlook

Dense time series acquired with spaceborne high-resolution optical sensors contain a wealth of information. The exploitation of this constantly growing data archive is an active field of research within the remote sensing community. This thesis gives an account of the current state of the art. The use of time series analysis and regression models based on trigonometric functions is identified as central element of several existing and operational algorithms. The primary objective of this thesis was to link these current approaches to Kalman filtering techniques commonly found in other disciplines and investigate their applicability for change-detection in forests.

Central to the intended application of the Kalman filter are time series models formulated in state-space form. Chapter 3 of this thesis reviews how these models are mathematically defined and how they are related to regression models. The Iteratively Reweighted Least Squares (IRLS) technique for robust parameter estimation is introduced to address the problem of un-masked clouds and cloud shadows left in the time series. Furthermore the well-known formulas for the Kalman filter itself are given in conjunction with an extension which can be used to detect anomalous measurements. The chapter concludes with the formulation of a data-driven algorithm capable of detecting abrupt changes on pixel-level.

A multi-temporal stack of Landsat surface reflectance data is used to test the suggested algorithm. Three features obtained by applying the tasseled cap transformation, namely brightness, greenness, and wetness, have been chosen as observables. The test site is located in Baden-Württemberg, Germany.

The results of the field test revealed no major flaws of the implemented algorithm and suggest that a fair degree of robustness against un-masked clouds and cloud shadows has been achieved. Three different change maps, each based on one of the observables, have been produced. A limited quantitative evaluation of the change maps based on a ground-truth data set describing windthrow areas after a storm in 2012 has been carried out. The best results are yielded by the wetness change map, where changed pixels were classified with a producer accuracy of up to 80.6% and a user accuracy of up to 86.1%.

Several starting points for future work are suggested by the results. They may be divided into three categories:

(1) *Further evaluation.* Partly due to the fact that the available reference map only covers storm damages in coniferous forest, the analysis of results for broadleaved forest is underrepresented. A further evaluation of the algorithms' sensitivity to other forest disturbances like fire, drought, or insect attacks based on suitable ground-truth data is necessary.

(2) *Adaptations for other input observables and large-area application.* Landsat data were used to implement the test case due to the availability of long historic time series. However, the algorithm can also be applied to Sentinel-2 data or other dense time

series, for example acquired by Sentinel-1. Another topic is to try other spectral bands and indices and implement a benchmarking framework. Additionally the algorithms' robustness could be further investigated and improved to apply it to larger areas.

(3) *Method improvements.* In the current stage, the algorithm processes only one observable at a time. Improvements could likely be achieved by introducing an integrated processing of multiple observables or the implementation of multivariate time series models. Another limit of the current version is that the change-detection mechanism relies solely on repeatedly large differences between forecast and actual observations. Gradual changes with relatively small initial magnitudes are frequently omitted. By also monitoring significant changes of the estimated time series components, for example shifts of the trend slope, omission errors could be reduced.

References

Beaton, A.E., Tukey, J.W., 1974. The fitting of power series, meaning polynomials, illustrated on band-spectroscopic data. Technometrics 16, 147–185. doi:10.1080/00 401 706.1974.10489 171

Brooks, E.B., Wynne, R.H., Thomas, V.A., Blinn, C.E., Coulston, J.W., 2014. On-the-fly massively multitemporal change detection using statistical quality control charts and Landsat data. IEEE Trans. Geosci. Remote Sens. 52, 3316–3332. doi:10.1109/TGRS.2013.2272545

Cohen, W.B., Healey, S.P., Yang, Z., Stehman, S. V, Brewer, C.K., Brooks, E.B., Gorelick, N., Huang, C., Hughes, M.J., Kennedy, R.E., Loveland, T.R., Moisen, G.G., Schroeder, T.A., Vogelmann, J.E., Woodcock, C.E., Yang, L., Zhu, Z., 2017. How similar are forest disturbance maps derived from different Landsat time series algorithms? Forests 8. doi:10.3390/f8 040 098

Crist, E.P., 1985. A TM Tasseled Cap equivalent transformation for reflectance factor data. Remote Sens. Environ. 17, 301–306. doi:10.1016/0034-4257(85)90 102-6

DeVries, B., Decuyper, M., Verbesselt, J., Zeileis, A., Herold, M., Joseph, S., 2015a. Tracking disturbance-regrowth dynamics in tropical forests using structural change detection and Landsat time series. Remote Sens. Environ. 169, 320–334. doi:10.1016/j.rse.2015.08.020

DeVries, B., Pratihast, A.K., Verbesselt, J., Kooistra, L., Herold, M., 2016. Characterizing forest change using community-based monitoring data and Landsat time series. PLoS One 11, e0 147 121–e0 147 121. doi:10.1371/journal.pone.0147121

DeVries, B., Verbesselt, J., Kooistra, L., Herold, M., 2015b. Robust monitoring of small-scale forest disturbances in a tropical montane forest using Landsat time series. Remote Sens. Environ. 161, 107–121. doi:10.1016/j.rse.2015.02.012

Dozier, J., 1989. Spectral signature of alpine snow cover from the Landsat thematic mapper. Int. J. Remote Sens. 29, 9–22.

Dutrieux, L.P., Verbesselt, J., Kooistra, L., Herold, M., 2015. Monitoring forest cover loss using multiple data streams, a case study of a tropical dry forest in Bolivia. ISPRS J. Photogramm. Remote Sens. 107, 112–125. doi:10.1016/j.isprsjprs.2015.03.015

Falk, M., Marohn, F., Michel, R., Hofmann, D., Macke, M., Spachmann, C., Englert, S., 2012. A first course on time series analysis — Examples with SAS.

Fu, P., Weng, Q., 2016. A time series analysis of urbanization induced land use and land cover change and its impact on land surface temperature with Landsat imagery. Remote Sens. Environ. 175, 205–214. doi:10.1016/j.rse.2015.12.040

Geerken, R.A., 2009. An algorithm to classify and monitor seasonal variations in vegetation phenologies and their inter-annual change. ISPRS J. Photogramm. Remote Sens. 64, 422–431. doi:10.1016/j.isprsjprs.2009.03.001

Gibbs, B.P., 2011. Advanced Kalman filtering, least-squares and modeling: A practical handbook. John Wiley & Sons, Hoboken. doi:10.1002/9780470890042

Hamunyela, E., Verbesselt, J., Herold, M., 2016. Using spatial context to improve early detection of deforestation from Landsat time series. Remote Sens. Environ. 172, 126–138. doi:10.1016/j.rse.2015.11.006

Harvey, A.C., 1989. Forecasting, structural time series models and the Kalman filter. Cambridge University Press, Cambridge.

Heiberger, R.M., Becker, R.A., 1992. Design of an S function for robust regression using iteratively reweighted least squares. J. Comput. Graph. Stat. 1, 181–196. doi:10.1080/10618600.1992.10474580

Hirschmugl, M., Gallaun, H., Dees, M., Datta, P., Deutscher, J., Koutsias, N., Schardt, M., 2017. Methods for mapping forest disturbance and degradation from optical Earth observation data: a review. Curr. For. Reports 3, 32–45. doi:10.1007/s40725-017-0047-2

Huang, C., Goward, S.N., Masek, J.G., Thomas, N., Zhu, Z., Vogelmann, J.E., 2010. An automated approach for reconstructing recent forest disturbance history using dense Landsat time series stacks. Remote Sens. Environ. 114, 183–198. doi:10.1016/j.rse.2009.08.017

Huber, P.J., 1964. Robust estimation of a location parameter. Ann. Math. Stat. 35, 73–101. doi:10.1214/aoms/1177703732

Irish, R.R., Barker, J.L., Goward, S.N., Arvidson, T., 2006. Characterization of the Landsat-7 ETM+ Automated Cloud-Cover Assessment (ACCA) algorithm. Photogramm. Eng. Remote Sens. 72, 1179–1188. doi:10.14358/PERS.72.10.1179

Jazwinski, A.H., 1970. Stochastic processes and filtering theory. Academic Press, New York.

Kandasamy, S., Baret, F., Verger, a., Neveux, P., Weiss, M., 2013. A comparison of methods for smoothing and gap filling time series of remote sensing observations – application to MODIS LAI products. Biogeosciences 10, 4055–4071. doi:10.5194/bg-10-4055-2013

Kennedy, R.E., Yang, Z., Cohen, W.B., 2010. Detecting trends in forest disturbance and recovery using yearly Landsat time series: 1. LandTrendr - Temporal segmentation algorithms. Remote Sens. Environ. 114, 2897–2910. doi:10.1016/j.rse.2010.07.008

Masek, J.G., Vermote, E.F., Saleous, N.E., Wolfe, R., Hall, F.G., Huemmrich, K.F., Gao, F., Kutler, J., Lim, T.-K., 2006. A Landsat surface reflectance dataset for North America, 1990–2000. IEEE Geosci. Remote Sens. Lett. 3, 68–72. doi:10.1109/LGRS.2005.857030

Oeser, J., Pflugmacher, D., Senf, C., Heurich, M., Hostert, P., 2017. Using intra-annual Landsat time series for attributing forest disturbance agents in central Europe. Forests 8, 251. doi:10.3390/f8 070 251

Pasquarella, V., Bradley, B., Woodcock, C., 2017. Near-real-time monitoring of insect defoliation using Landsat time series. Forests 8, 275. doi:10.3390/f8 080 275

Pengra, B., Gallant, A., Zhu, Z., Dahal, D., 2016. Evaluation of the initial thematic output from a continuous change-detection algorithm for use in automated operational land-change mapping by the U.S. Geological Survey. Remote Sens. 8, 811. doi:10.3390/rs8 100 811

Reiche, J., Verbesselt, J., Hoekman, D., Herold, M., 2015. Fusing Landsat and SAR time series to detect deforestation in the tropics. Remote Sens. Environ. 156, 276–293. doi:10.1016/j.rse.2014.10.001

Renaud, O., Victoria-Feser, M.P., 2010. A robust coefficient of determination for regression. J. Stat. Plan. Inference 140, 1852–1862. doi:10.1016/j.jspi.2010.01.008

Schultz, M., Verbesselt, J., Avitabile, V., Souza, C., Herold, M., 2016. Error sources in deforestation detection using BFAST Monitor on Landsat time series across three tropical sites. IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens. 9, 3667–3679. doi:10.1109/JSTARS.2015.2477473

Thonfeld, F., Hecheltjen, A., Menz, G., 2015. Bi-temporal change detection, change trajectories and time series analysis for forest monitoring. Photogramm. - Fernerkundung - Geoinf. 2015, 129–141. doi:10.1127/pfg/2015/0259

USGS, 2018. Landsat Processing details [WWW Document]. URL https://landsat.usgs.gov/landsat-processing-details (accessed 1.18.18).

USGS, 2016a. Product Guide - Landsat 4-7 climate data record (CDR) Surface Reflectance.

USGS, 2016b. Product Guide - Provisional Landsat 8 Surface Reflectance Code (LASRC) product.

Verbesselt, J., Herold, M., Zeileis, A., 2012. Near real-time disturbance detection using satellite image time series. Remote Sens. Environ. 123, 98–108. doi:10.1016/j.rse.2012.02.022

Verbesselt, J., Hyndman, R., Newnham, G., Culvenor, D., 2010a. Detecting trend and seasonal changes in satellite image time series. Remote Sens. Environ. 114, 106–115. doi:10.1016/j.rse.2009.08.014

Verbesselt, J., Hyndman, R., Zeileis, A., Culvenor, D., 2010b. Phenological change detection while accounting for abrupt and gradual trends in satellite image time series. Remote Sens. Environ. 114, 2970–2980. doi:10.1016/j.rse.2010.08.003

Vermote, E., Justice, C., Claverie, M., Franch, B., 2016. Preliminary analysis of the performance of the Landsat 8/OLI land surface reflectance product. Remote Sens. Environ. 185, 46–56. doi:10.1016/j.rse.2016.04.008

Vogelmann, J.E., Gallant, A.L., Shi, H., Zhu, Z., 2016. Perspectives on monitoring gradual change across the continuity of Landsat sensors using time-series data. Remote Sens. Environ. 185, 258–270. doi:10.1016/j.rse.2016.02.060

Vogelmann, J.E., Khoa, P., Lan, D., Shermeyer, J., Shi, H., Wimberly, M., Duong, H., Huong, L., 2017. Assessment of forest degradation in Vietnam using Landsat time series data. Forests 8, 238. doi:10.3390/f8 070 238

Wikipedia, 2018. Sentinel-2 [WWW Document]. URL https://en.wikipedia.org/wiki/Sentinel-2 (accessed 10.15.18).

Wulder, M.A., White, J.C., Loveland, T.R., Woodcock, C.E., Belward, A.S., Cohen, W.B., Fosnight, E.A., Shaw, J., Masek, J.G., Roy, D.P., 2016. The global Landsat archive: Status, consolidation, and direction. Remote Sens. Environ. 185, 271–283. doi:10.1016/j.rse.2015.11.032

Wylie, D.P., Menzel, W.P., 1999. Eight years of cloud statistics using HIRS. J. Clim. 12, 170–184.

Zhu, Z., 2017. Change detection using landsat time series: A review of frequencies, preprocessing, algorithms, and applications. ISPRS J. Photogramm. Remote Sens. 130, 370–384. doi:10.1016/j.isprsjprs.2017.06.013

Zhu, Z., Fu, Y., Woodcock, C.E., Olofsson, P., Vogelmann, J.E., Holden, C., Wang, M., Dai, S., Yu, Y., 2016. Including land cover change in analysis of greenness trends using all available Landsat 5, 7, and 8 images: A case study from Guangzhou, China (2000–2014). Remote Sens. Environ. 185, 243–257. doi:10.1016/j.rse.2016.03.036

Zhu, Z., Woodcock, C.E., 2014a. Automated cloud, cloud shadow, and snow detection in multitemporal Landsat data: An algorithm designed specifically for monitoring land cover change. Remote Sens. Environ. 152, 217–234. doi:10.1016/j.rse.2014.06.012

Zhu, Z., Woodcock, C.E., 2014b. Continuous change detection and classification of land cover using all available Landsat data. Remote Sens. Environ. 144, 152–171. doi:10.1016/j.rse.2014.01.011

Zhu, Z., Woodcock, C.E., 2012. Object-based cloud and cloud shadow detection in Landsat imagery. Remote Sens. Environ. 118, 83–94. doi:10.1016/j.rse.2011.10.028

Zhu, Z., Woodcock, C.E., Holden, C., Yang, Z., 2015. Generating synthetic Landsat images based on all available Landsat data: Predicting Landsat surface reflectance at any given time. Remote Sens. Environ. 162, 67–83. doi:10.1016/j.rse.2015.02.009

Zhu, Z., Woodcock, C.E., Olofsson, P., 2012. Continuous monitoring of forest disturbance using all available Landsat imagery. Remote Sens. Environ. 122, 75–91. doi:10.1016/j.rse.2011.10.030