GERMANYS FIRST CLOUD-BASED WEB SERVICE FOR LAND MONITORING USING COPERNICUS SENTINEL-2 DATA

Patrick Knoefel[1, *], David Herrmann[2], Marcus Sindram[2], Michael Hovenbitzer[1]

1 Federal Agency for Cartography and Geodesy, Frankfurt, Germany; patrick.knoefel@bkg.bund.de

2 GAF AG, Munich, Germany

KEY WORDS: Copernicus, Sentinel-2, change detection, monitoring, spatio-temporal features, web application, fully automatic.

ABSTRACT:

The research and development project named Landscape Change Detection Service (German abbreviation: LaVerDi) was initiated by the German Federal Agency for Cartography and Geodesy (BKG). Within the scope of the project a monitoring service for landscape changes was developed and implemented using free Copernicus satellite data for an automated derivation of potential land cover change. This change indication is meant to be used to update or continue BKG in-house products, such as the Digital Land Cover Model Germany (LBM-DE), in a comprehensive and uniform quality. The results can be further used for numerous applications or as change information for administration and planning, and for the compilation of spatial statistics. It satisfies the users' need for a national service for open data on land cover changes and thus represents the first automatic and verified national satellite product for land cover changes in Germany. As input data the service uses pre-processed Sentinel-2 data from the European Copernicus satellite program, as well as an image segmentation approach to extract change objects. Using an improved cloud mask algorithm, Sentinel-2 tiles with up to maximum cloud coverage of 60% can be used for analysis. The service (data processing, change detection, visualisation) runs on the German "Copernicus Data and Utilization Platform" (CODE-DE). As of December 2020, the INSPIRE-compliant LaVerDi web service is operational. The thematic accuracy of the generated change layers is above the given requirements (minimum of 80%), considering the 95% confidence interval for all relevant land cover classes in certain test areas. The transferability of the methodology has been successfully shown by a prototypic nationwide demonstrator in early 2020 and is therefore expected to reliably detect both long-term and seasonal changes.

1. INTRODUCTION

Determining changes of the Earth's surface is a common application of satellite-based Earth observation (Townshend et al. 1991). First change analysis methods were already presented in the 1980s (Singh 1989). Since then a variety of increasingly improved methods and algorithms have been developed. Correspondingly, the availability of potentially suitable satellite image data has vastly increased. In addition to the number of available sensors with high spatial resolution (e.g. Landsat, SPOT, Sentinel), a large portion of this data is now available free of charge, thus enabling large-scale evaluation and application in various disciplines. In order to be able to process these large amounts of data (cost-) efficiently, a high degree of automation is necessary.

The research project named Landscape Change Detection Service (German abbreviation: LaVerDi) was initiated in spring 2017. Development and implementation of the project were contracted to a consortium comprising GAF AG (project coordination) and the German Aerospace Center (DLR) as scientific project partners in the initial phase. The aim of the project was to develop and implement a regular monitoring of landscape changes in Germany, using free Copernicus data to support the updating and validation of BKG internal products with a focus on the LBM-DE (Land Cover Model for Germany). The detected changes were to be derived with an accuracy of ≥80% and primarily serve as update information for the LBM-DE. Besides the development of suitable methodologies for land cover change analysis, the developed procedures were to be embedded on the CODE-DE platform (Copernicus Data and Exploitation Platform - Germany) and made available via an online viewing service.

Within the framework of LaVerDi, update information should be automatically derived for the following land covers, among others:

- Forest
- Agricultural use (fallow land, arable land)
- Water areas / wetlands

The final goal of the project was the development of a performant web application for a regular and automatic cloud-based provision of change indications for Germany. Specifically, the project is divided into two phases, the pilot phase and the rollout phase. In the first project phase, an algorithm for change analysis based on time series of optical Sentinel-2 data with a spatial resolution of 10 meters was developed and demonstrated in defined test areas using the LBM-DE. The test areas comprised four Sentinel-2 tiles (each approx. 100 x 100 km) in the "Test Area South" between Munich and Stuttgart, as well as two tiles in the "Test Area North" nearby Berlin. An assessment of appropriate update frequency of the change indications and a cloud cover assessment were also carried out during this project stage.

The second phase focused on the integration of the developed algorithms on the CODE-DE platform with the aim of complete automation. To this end, the relevant processing chains and interfaces were already developed in a standardised manner during the pilot phase. The implementation was fully rolled out on CODE-DE at the end of November 2020 and handed over to the BKG in the first half of December 2020.

---

* Corresponding author

2. DATA AND PROCESSING

2.1 Land Cover Reference Model

The Digital Land Cover Model for Germany "LBM-DE" was developed in coordination with the German Federal Environment Agency and has been made available by the BKG since 2009 (at that time known as DLM-DE) (Hovenbitzer 2012). It describes topographic objects of the landscape in vector format under the aspect of land cover (LC) and land use (LU). It provides information on 31 LC classes with a minimum mapping unit (MMU) of 1 ha. In this way, the state of the environment at a given point in time is recorded, thus enabling analyses according to different viewpoints. The LBM-DE has so far been updated for the reference years 2012, 2015 and 2018 covers the entire extent of Germany. Based on the boundaries of the area, objects of the base landscape model (base DLM), information on land cover and land use in terms of the European CORINE Land Cover (CLC) nomenclature is derived. This is achieved by evaluating multispectral satellite image time series from the respective reference year.

2.2 Copernicus Data

Since 2014, the European Earth observation programme Copernicus has been providing a modern and powerful infrastructure for Earth observation and geoinformation. Earth observation includes measurements from satellites, aircraft, but also ground- and sea-based observation infrastructures. The Copernicus infrastructure is subdivided into a space component and an in-situ component. The space component covers CORE services in the overarching themes of land, climate and atmosphere, oceans, crisis and disaster management and security with the so-called Sentinel missions that complement each other. The Sentinel-2 satellite in particular was developed for the collection of land cover information and vegetation. Sentinel-2A was launched in June 2015, followed by Sentinel-2B in March 2017. Sentinel-2 carries an optical instrument (MSI - multispectral imager) with 13 spectral channels. The spatial resolution ranges 10 to 60 m, depending on the channel. As a tandem mission, Sentinel-2 has a temporal resolution of 5 days at the equator. The areas of application of the optical data are extremely diverse and range from agricultural and forestry applications to the derivation of water quality parameters.

Sentinel data is provided in different processing levels. These are the levels 0 (raw), 1 (upper level) and 2. LaVerDi uses Sentinel-2 Level 2A data. This is cartographically rectified and atmospherically corrected Sentinel-2 data, which is being generated as bottom of atmosphere (BOA) reflectance products from the Level 1C products using atmospheric and topographic correction algorithms. The products are processed by the European Space Agency (ESA) and since mid-March 2018, the Level-2A became an operational product, starting with coverage of the Euro-Mediterranean region. Level-2A can also be generated by the user from the Level-1C product using the Sentinel-2 toolbox or a standalone version of the Sen2Cor processor.

2.3 Cloud Processing

The developed change detection application itself as well as the relevant data mentioned in section 2.2 will be either provided, stored or implemented in a cloud infrastructure. For this purpose, the national Copernicus platform CODE-DE was chosen. As part of the German strategy on geoinformation, CODE-DE provides access to remote sensing data in a simple and efficient way with

a virtual working environment for processing these data, and extensive information material and training to support users. As a central entry point, CODE-DE is not only to provide data and information on Copernicus, but also to offer online capacities for processing by national users. CODE-DE has been set up by the Federal Ministry of Transport and Digital Infrastructure (BMVI) to ensure faster and better provision of data, products and applications for national users. Its processing and storage capabilities were used to set up the Landscape Change Detection Service. Further information is given in section 4.3.

3. METHOD

In this section the relevant processing steps implemented in the analysis workflow are described. Starting with the pre-processing (section 3.1) followed by the explanation of relevant products in the change detection workflow (section 3.2) to gain the final change layer (3.2.3).

3.1 Pre-processing

The first step of the LaVerDi processing chain is the integration of all Sentinel-2 Level-2A data available on CODE-DE with a specified cloud cover ($\leq 60\%$) for the respective selected reference period (e.g. 01.03. - 31.10.2017/2018). The metadata is extracted and transferred to a relational database. Subsequently, the default 10m, 20m and 60m bands are resampled to a reference resolution of 10m, followed by a cloud and cloud shadow detection. For cloud masking, a modified Scene Classification Layer (SCL) is used, which generates better results in cloud/shadow masking than the native SCL of the Level2A product. Furthermore, some disadvantages of the SCL over urban areas are partially compensated by the overall dense Sentinel-2 time series for Germany.

These modified cloud masks in combination with the Sentinel-2 L2A data build the input for the index calculation and the band-specific multi-temporal, pixel-based statistics. The most suitable index identified was the NDVI (Normalized Difference Vegetation Index). Other indices like NDWI (Normalized Difference Water Index) and brightness were tested in this study but have not been considered in the operational service.
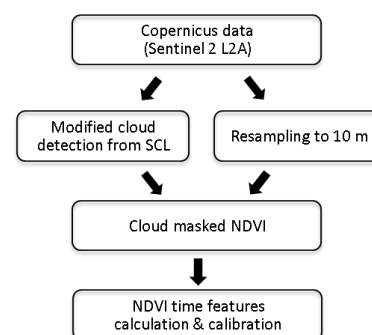
Figure 1. Simplified Sentinel-2 Pre-Processing for LaVerDi.

Based on the scene-specific NDVI, a series of spatiotemporal features are then calculated, which can also be derived from individual spectral bands if required. The features are created as a vertical layer stack according to their expression (minimum, maximum, median, mean, etc.) for the selected reference period (e.g. the vegetation period). First, based on experiences from other research projects (e.g. ECoLaSS - Evolution of Copernicus Land Services based on Sentinel data), a set of 44 spatiotemporal features (also referred to as "time features" in the rest of the

paper) was derived in a standardised way. Finally, only five time features for mean, 10%-percentile, 90%-percentile and the percentile difference of 90%/10%, plus the median features of the Sentinel-2 bands for the observation period specified by the user were used for the application. In a final step these features are calibrated and then directly incorporated into the class-specific change analysis. With the completion of the calculation of the time features, all requirements for change detection are fulfilled and the pre-processing is finished. All relevant information and generated metadata are written to an object-relational database (open source PostgreSQL) during the processing. Directly afterwards, the thematic processing of the data can be initiated.

3.2 Change Detection Workflow

The detection of changes in LaVerDi is based on the concept of spatio-temporal features (Probeck et al. 2019). The basic idea is to compare two Sentinel-2 time series from a pre-defined reference period (T0/T1) (see section 3.3). Time features are calculated for the historical period T0 and the current period T1, which are both incorporated into the change detection workflow. The usage of time features (depending on the quality of the cloud masks used) thereby largely "solves" the well-known data gap and cloud cover issues in multi-temporal satellite data.
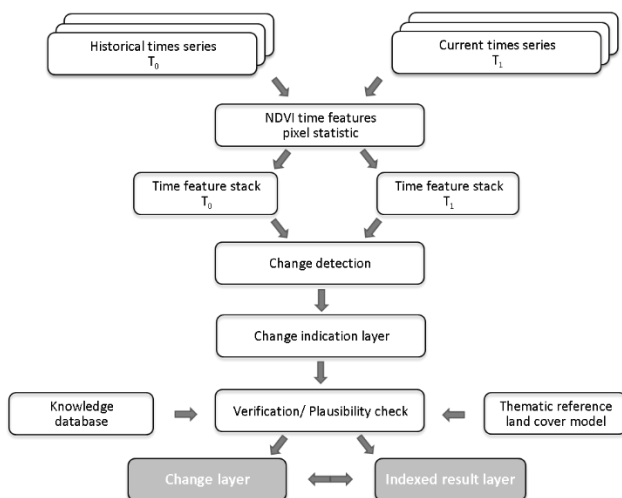
Figure 2. LaVerDi Workflow for automatic change detection.

Furthermore, the consideration of suitable statistical key indicators (e.g. maximum, minimum, mean, percentiles, etc.) enables comparability of the data. In addition, the temporal comparison of periods ensures the avoidance of dynamic seasonal effects (phenology) to a great extent. However, changes in land cover that only appear towards the end of the reference period are problematic. These may not have sufficient relevance in the time series to be adequately represented by the selected feature.

The change indications are mapped separately for each relevant land cover category in the form of a pixel-based Change Indication Layer (CIL, see section 3.2.1) and then subjected to a plausibility check before being filtered and attributed into the final products, the change layer and the indexed reference dataset. The changes provided by the CIL are subjected to land-cover-specific MMU filtering and plausibility or verification using the LBM-DE as a thematic reference in combination with a knowledge database, which contains empirical information on the a priori transition probabilities. The resulting change layer is blended with the LBM-DE for final attribution.

In addition, object geometries from the LBM-DE that overlap with the change layer are selected and indexed with regard to a change. The result is unchanged LBM-DE geometries that provide information on the change in land cover (change: yes/no) in the form of the indexed reference dataset. Change layer and indexed reference dataset are the final products of LaVerDi.

3.2.1 Change Indication Layer

The Change Indication Layer is the centrepiece of the Landscape Change Service. It is derived for five land cover categories (Build-up, cropland, grassland, forest, and water) of the LBM-DE (see Table 2) and represents land cover changes within 12 LBM-DE classes at pixel level without evaluating them. The land cover categories open areas with/without vegetation and specific wetlands are currently not mapped by LaVerDi.
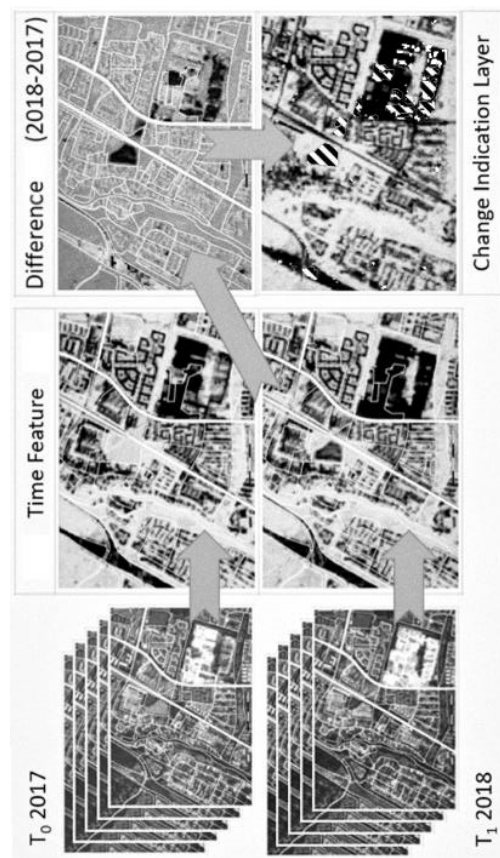
Figure 3. Production of the change indication layer.

Figure 3 illustrates the principle for deriving the change indication layer from the compared time periods. Specific time features are derived from the Sentinel-2 time series. The actual change detection is then carried out via calculating the difference between the relevant time feature data sets. Using thresholding (x-times standard deviation), probability of change indications is then extracted from the difference image in the form of pixel geometries. The change polygons derived from the CIL are checked for plausibility in a further step with the use of the LBM-DE and an empirical knowledge database (next section) and subjected to post-processing (MMU filtering) to generate the final change layer (3.2.3).

3.2.2 Knowledge Database

The evaluation of the detected change indications is subject to consideration in the plausibility check. For this purpose, the change areas are linked with additional information from the LBM-DE (attributes and topological properties) and the feature statistics from the reference periods (T0/T1). In combination with the a-priori transition probabilities of specific land cover types, a set of rules for determining the direction of change for detected change polygons can thus be built. This knowledge-based set of rules is called knowledge database in the project context. With regard to the LBM-DE reference class key, it was filled with empirical directions of land cover change by analysing the LBM-DE products 2015 and 2018 on the frequency of certain types of land cover change (class transitions). Furthermore, a class called "tree loss" was introduced in order not to have to differentiate between man-made and natural changes in tree and forest cover. Temporary states or changes in the transformation process are always given by their own class code.

3.2.3 Resulting Change Layers

Change layer and indexed result layer represent the final vector products provided to the user via the application. First, the unfiltered change indications provided in the 10m raster geometry are calculated from the CIL. These pixel-based indications of change are extracted and converted into a binary mask (0/1) in order to be able to carry out performant raster operations. First, all gaps within the change geometries up to the specified MMU (≥0.5 or 1ha) are filled using a 4x4 connectivity filter. Then, all pixels and pixel groups smaller than the specified minimum mapping unit (MMU) are eliminated. The final Change Layer is generated by intersecting the MMU-filtered CIL with the most recent version of the LCR reference data, considering the previous plausibility step. The change layer is attributed using the knowledge database, the most recent LCR and a background classification of the LCR for the two reference periods (T0/T1) using a machine learning algorithm in order to determine the most probable direction of change in combination with a change probability in 25% steps for each polygon. The change layer and indexed reference dataset allow the land cover model to be updated with target and object accuracy.

3.2.4 Class Transitions

The sampling design for the classification consists of a stratified random sample in the relevant reference classes with more than 1.6 million samples. The following process steps are performed:

- Rasterization of the LBM-DE in vector format and transfer to the 10m pixel geometry of Sentinel-2.
- Application of an 8x8 filter for each relevant LBM-DE class to minimise transition zones and mixed pixel effects at the edges.
- Performing a stratified random sampling considering class weighting in the LBM-DE (minimum of 50,000 samples per class) followed by extraction of the necessary statistical parameters.
- Performing a statistical K-Nearest Neighbour outlier analysis to optimise the sample data set.
- Training of the Random Forest algorithm.

For this process/analysis, the use of machine learning to determine change probability and direction using the knowledge database was found to be very effective. The process, however,

proved to be very memory intensive. For this reason, LightGBM (Guolin et al., 2017) was chosen as the algorithm for the operational service, as it works much more RAM-efficiently without having to accept a loss of accuracy.

4. RESULTS

4.1 Update Frequency

As already described in the sections above, the use of cloud-free data sets is not a mandatory requirement for the change detection in LaVerDi. This is of advantage as cloud-free data sets are generally rather rare and play a subordinate role, compared to the abundance of data. With regard to the reference period, the shorter it is, the higher the probability of not having enough cloud-free pixels for a valid change detection. Consequently, the update frequency should not run in too short intervals in order not to avoid the risk of receiving too little information for a reliable detection. Based on the results of the project pilot phase, the shortest possible interval was set at 3 months. For such an interval, there is a sufficiently high probability that enough cloud-free observations are available for a reliable change detection. The minimum number of observations for an efficient change analysis based on a time series is therefore set to two observations per month with low cloud cover (≤10%). Such dense time series enable the detection of changes in dynamic land cover classes, such as the change in phenology of agricultural crop species (which is used to determine these classes). In addition, it allows the determination of the time of change in more stable classes, such as the mowing time on extensive grassland or the time of new construction in urban areas within the time series. Long-term changes in land cover (e.g. as a result of construction measures) can already be detected in short observation periods, while seasonal changes (arable/fallow or differentiation of arable/grassland) require a significantly longer observation period.

With the tandem constellation of the two Sentinel-2A+B satellites, a satellite image can be expected every five days under good acquisition conditions. Theoretically, a maximum of six images per month can be available. Due to the wide-swath recordings of the satellites with approx. 290 x 290 km coverage, the repetition rate in the overlap areas (partly <10%) is increased (every 2-5 days), so that a maximum of up to 15 (partial) recordings per tile and month can be available. With four full coverages per month, therefore, a good and with six full coverages a quasi-optimal data availability for the time series analysis can be achieved. This density of time series is consistently achieved for the first time in the year 2018, starting with the operational tandem constellation (commissioning phase of Sentinel-2B starts from spring 2017). Here, four or more acquisitions per month were available for change detection. The approach chosen in LaVerDi for change detection provides a comparison of two selected periods of the same duration within two consecutive years. Project-independent analyses by GAF AG have shown that the inclusion of the winter months in the reference period does not add any significant value for change detection. Reasons for this are, among others:

- snow/ice cover in the winter months,
- higher commission errors in cloud detection due to snow and ice,
- low sun angles and resulting shadow issues,
- a fundamentally lower variability of vegetation characteristics (e.g. NDVI) in the winter months.

Therefore, the winter months have been excluded from further consideration. This, in turn has the positive side effect that less input data is needed and the calculation of the required cloud masks, indices and time features can be calculated much faster. The results of this investigation for the update frequency depending on land cover are shown in Table 1.

Land cover category	Possible update frequency	Optimal Reference period
Build-up	3 months	01.03.2017/18
Cropland	12 months	–
Grassland	12 months	31.10.2017/18
Forest	3 months	(8 months)
Water	4-6 months	

Table 1. Recommended update frequency by land cover type.

4.2 Accuracy Assessment

In the pilot phase of the project 9,362 change polygons were detected within the test areas (see section 1) for the reference period 2017 to 2018. These changes were analysed for thematic accuracy to determine the omission error. For this purpose, the detected change areas from the change layers as well as the relevant classes from the thematic reference were sampled by means of a stratified random sample. The results are presented for each change layer in the form of a confusion matrix and subsequently discussed.

The thematic accuracy of the final change layer was determined by stratified random sampling at polygon level with two strata:

- Stratum 1: Stratum with the sum of all detected change polygons (≥0.5 or 1ha).
- Stratum 2: Stratum containing all other LBM-DE object geometries (≥0.5 or 1ha) of the specific LBM-DE land cover category or class.

This approach results in six separate stratifications and precision analyses for six land cover categories (see Table 2). The minimum sample size n is calculated according to formula (1) below (Snedecor and Cochrane 1967) from the expected classification accuracy p (here 0.8), the allowable sampling error E (here 0.05) and the z-value (here z= 1.96, read from standard normal distribution table).

$$n = \frac{p \cdot q}{(E/z)^2} \qquad (1)$$

where
n = sample size
p = desired accuracy
q = 1 - p
E = sampling error
z = z-value from the 95% confidence interval

Taking into account the 95% confidence interval, this results in a minimum sample size of 246 samples (polygons) for stratum 1. In order to absorb possible uncertainties in the interpretation of the multi-temporal satellite image data, a 10% oversampling was applied, so that the sample size comprises a maximum of 270 samples per change layer. The sample for stratum 2 has been increased by a factor of 2, considering the population N to determine the omission error.

Land cover category	Population N	Sample size (Strata 1 / 2)
Build-up area	144,825	270 / 540
Crop land	58,065	270 / 540
Homogenous Grassland	102,873	270 / 540
Grassland	187,455	270 / 540
Forest and semi-natural areas	187,832	270 / 540
Water surfaces	8,186	12* / 540
* Only 12 change polygons were detected in land cover category Water surfaces.		

Table 2. Sample size n for strata 1 and 2.

The change polygons from the change layer as well as the reference polygons were manually evaluated by visual interpretation regarding a real change (change: yes/no) on multi-temporal 10m Sentinel-2 data, considering the applicable MMU into account. Historical Very High Resolution (VHR) satellite data images were used as auxiliary data. The results were then transferred to a confusion matrix to calculate the accuracy measures. Historical images were used as auxiliary data. An investigation into the correct assignment of the direction of change was not part of the accuracy analysis. It was only checked whether the change indication also resulted in a real change of land cover (independent of the land cover type).

A total of 9,362 change polygons with a total area of 19,639 ha were detected in the test areas of the pilot phase, which corresponds to 0.33% of the cumulative LaVerDi test area. Of these, 2,373 areas (25.35%) are in the North test area and 6,989 (74.65%) in the South test area. On average, approximately 1,560 change areas have been detected per Sentinel-2 tile. Overall, the percent change areas correspond strongly with the LBM-DE land cover percentages. As expected, based on the LBM-DE land cover percentages, most changes occur in agricultural land, forests and semi-natural land, and grassland/grassland. These are followed by a wide margin by changes within built-up areas. Changes within water areas, on the other hand, can hardly be observed.

Table 2 shows the confusion matrix for the change layer for LBM-DE category "Built-up areas". The change layer has a total of 329 change areas with an average size of 0.81 ha. The overall accuracy achieved is 97.90%. The producer accuracy for the change is 99.22%, while the user accuracy is 94.44%. The omission error is less than 1%.

		Reference data 2017/2018				
		no change	change	Sum	User	95% CI
Build-up area	no change	538	2	540	99.63	±0,60
	change	15	255	270	94.44	±2.92
	Sum	553	257	810		
	Producer	97.29	99.22			
	95% CI	±1.44	±1.27		Overall accuracy: 97.90%	

Table 2. Confusion matrix for change layer Built-up areas.

Table 3 shows the confusion matrix for the change layer for LBM-DE category "Crop land". In total, the change layer has 4,969 change areas with an average size of 2.62 ha. The overall accuracy achieved is 82.35%. The producer accuracy for the change is 95.68%, while the user accuracy is only 49.26% due to

the high number of commission errors. The omission error is less than 5%. The change layer for crop land reliably shows land cover changes towards urban land.

Cropland		Reference data 2017/2018				
		no change	change	Sum	User	95% CI
	no change	534	6	540	98.89	±0.98
	change	137	133	270	49.26	±6.15
	Sum	671	139	810		
	Producer	79.58	95.68			
	95% CI	±3.12	±3.74		Overall accuracy: 82.35%	

Table 3. Confusion matrix for change layer Crop land.

The confusion matrix for change layer "Homogeneous grassland" is given in Table 4. In total, the change layer has 1,436 change polygons with an average size of 2.33 ha. In total, the change layer has 1,436 change areas with an average size of 2.33 ha. The overall accuracy achieved is 88.64%. The producer accuracy for the change is 96.84%, while the user accuracy is only 68.15% due to relatively many commission errors. The omission error is less than 5%.

Homogenous Grassland		Reference data 2017/2018				
		no change	change	Sum	User	95% CI
	no change	534	6	540	98.89	±0.98
	change	86	184	270	68.15	±5.75
	Sum	620	190	810		
	Producer	86.13	96.84			
	95% CI	±2.80	±2.75		Overall accuracy 88.64%	

Table 4. Confusion matrix for Homogeneous grassland.

Mainly the changes in "Homogenous Grassland" are detected correctly. Frequent changes can be observed towards crop land and built-up areas, so the layer can be used well to update the LBM-DE. As in the case of class crop land, the extreme year 2018 also causes a large number of commission errors for this layer due to dried-up meadows and pastures. In total, the change layer for the change class grassland has 424 change areas with an average size of 0.92 ha. The overall accuracy achieved is 90.37%. The producer accuracy for the change is 98.48%, the user accuracy is only 72.22% and the omission error is less than 5%.

Grassland		Reference data 2017/2018				
		no change	change	Sum	User	95% CI
	no change	537	3	540	99.44	±0.72
	change	75	195	270	72.22	±5.53
	Sum	612	198	810		
	Producer	87.75	98.48			
	95% CI	±2.68	±1.95		Overall accuracy: 90.37%	

Table 5. Confusion matrix for change layer Grassland

The change layer for the class "Grassland" (see Table 5) is well suited to show changes towards urban areas. In addition, misallocations in the reference (e.g. arable land instead of grassland with trees) can be reliably detected. The majority of commission errors must be attributed to drought-related areas.

The confusion matrix for the change layer category "Forests and semi-natural areas" (see Table 7) has a total of 2,292 change areas with an average size of 1.19 ha. The overall accuracy achieved is 96.54%. The producer accuracy for the change is 99.59%, while the user accuracy is exactly 90.00%. The omission error is less than 1%. Changes in forest land cover are detected very reliably. Changes in the context of forestry measures are detected particularly frequently. In second place comes tree loss as a result of construction activities (primarily infrastructure). The layer also detects misallocations in the LBM-DE (e.g. grassland instead of forest) and is therefore ideally suited for updating the LBM-DE.

Forest/semi-natural areas		Reference data 2017/2018				
		no change	change	Sum	User	95% CI
	no change	539	1	540	99.81	±0.46
	change	27	243	270	90.00	±3.76
	Sum	566	244	810		
	Producer	95.23	99.59			
	95% CI	±1.84	±1.01		Overall accuracy: 96.54%	

Table 6. Confusion matrix for change layer Forest

Table 7 shows the confusion matrix for the change layer for LBM-DE category "water surfaces". Only 12 change polygons with an average size of 0.95 ha are identified. The overall accuracy achieved is 96.01%. The producer accuracy for change, assuming that temporary plant cover does not constitute a change, is only 14.29%; the user accuracy is slightly higher at 16.67%. The omission error is just under 85%.

Water surfaces		Reference data 2017/2018				
		no change	change	Sum	User	95% CI
	no change	528	12	540	96.78	±1.34
	change	10	2	12	16.67	±25.25
	Sum	538	14	552		
	Producer	98.14	14.29			
	95% CI	±1.23	±21.90		Overall accuracy: 96.01%	

Table 7. Confusion matrix for change layer Water surfaces

The change layer shows almost exclusively temporal effects, which in all probability resulted from the extreme drought in 2018. This includes in particular the decrease in the water level of water bodies, which is interpreted as a real change in the context of change detection. In addition, the approach also strikes at algal blooms and vegetation growth (e.g. reeds, water lilies, pondweed), which are, however, evaluated here as Commission errors. The methodology works correctly, but cannot differentiate between vegetation covered waterbodies and land cover. The results show that the methodology delivers predominantly good results.

The required minimum overall accuracy of 80% has been achieved or even exceeded for each change layer, considering temporal or seasonal effects. Under optimal weather and climate conditions, the methodology can provide very good results. Extremes, such as the 2018 European drought, can be considered problematic for the used methodology. As the methodology is based on the NDVI, increased drought and the resulting drought stress can lead to misclassifications in the context of change

detection. This is particularly the case for the class crop land and homogeneous grassland. The effect occurs in both test areas with a focus on the northern test area.

In early 2020, the transferability of the chosen methodology has been proven by set-up of a nationwide demonstrator (hosted in a private cloud), resulting in a Change Layer 2018-2019 with approx. 57,000 change polygons.

4.3 Web Service

The main objective of the project was the regular provision of automatically derived areas of change in form of a high-performance web application that provides detected land cover changes in polygon form at specified points in time. At the end of 2020, the service was handed over to the BKG by the developer and contractor GAF AG. The functionality of the application is presented in this section. The computing capacity on the CODE-DE cloud was requested in form of a standard processing quota. This quota consists of 128 virtual CPUs and 1 TB GB RAM, 2 TB block storage and a 1 TB file storage, as well as public IP access. These resources were distributed among ten virtual machines, which host the non-monolithic services to grant a balanced data processing and presentation. The processing contingent is designed to be flexible for the duration of LaVerDi and includes support services from the operator.
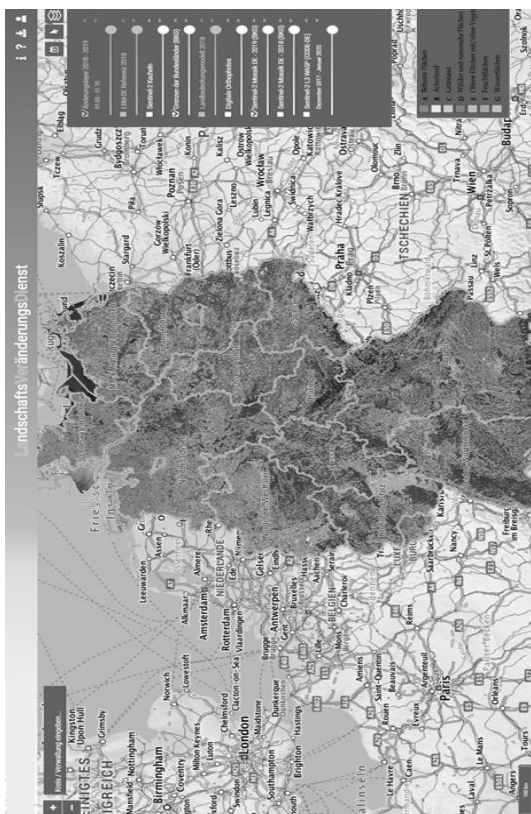

Figure 4. Screenshot of the change detection viewer application.

The developed front-end of the web application is based on the open-source framework Angular Material and the whole back-end programming has been implemented using ASP.NET Core. Communication between the front-end and back-end of the web application has been realised with WebSocket, whereas the communication between web application and the python back-end, representing the thematic processing chain on CODE-DE, is ensured via the open-source message broker RabbitMQ.

Following successful login, the LaVerDi viewer opens automatically, as illustrated in Figure 4. The already calculated and nationwide Change Layer 2018-2019 is activated as the first layer in the layer tree according to the default setting but is greyed out, as the display of the change polygons (approx. 57,000) in full resolution would otherwise lead to considerable performance losses. These are only displayed from a scale of 1:200,000 onwards. The viewer contains a search function for finding change polygons. This is positioned on the left-hand side and uses the Web Map Service (WMS) "Boundaries of the Federal States". The layer tree, the toolbars and the LCR legend are positioned on the right hand of the viewer. The Viewer is deliberately designed intuitively with its components and works only partly with tool tips. Seven WMS services and layers are currently integrated in the viewer (see Table 8), the transparency of which can be changed by the user via a slider.

The user can also control the order or arrangement of the layers as desired in order to ensure overlapping with other information and services. For the 60m Sentinel-2 WASP mosaic, a timeline has also been implemented at the bottom to visualise the timing of potential changes on a monthly basis. From a scale of 1:200,000 onwards, a menu with a result list of the change polygons calculated within the displayed window opens automatically on the left-hand side and the change layer is activated in the layer tree (see Figure 5). Change polygons are always output in the 10m pixel geometry of Sentinel-2. Within the results list, the user can zoom to the desired change polygon using the magnifying glass and view the details of the change detection via the list.

Nr.	Layer Name	Content
1	Change layer 2018-2019	LaVerDi change polygons for Germany
2	LBM-DE Reference 2018	Indexed reference dataset with LBM-DE reference polygons
3	Sentinel-2 tiles	Generalised Sentinel-2 tiles (production units)
4	Boundaries of Federal States	Federal administrative boundaries
5	Sentinel-2 mosaic DE 2019	Low-cloud Sentinel-2 coverage of Germany for 2019 at 10m resolution
6	Sentinel-2 mosaic DE 2018	Low-cloud Sentinel-2 coverage of Germany for 2018 at 10m resolution
7	Sentinel-2 L3 WASP	Monthly low-cloud Sentinel-2 coverage at 60m resolution

Table 8. Description of the layers integrated in the service

For each change polygon, the probability of change in the categories not relevant, very low, low, medium and high is indicated in the upper part of the window next to the area statistics. This is followed by the direction of change (e.g. from crop land in T0 to built-up area in T1) with the NDVI metrics. The latter are always given, but are only relevant for one feature due to the class-specific analysis. As described in section 3.1, the NDVI metrics given are: NDVI mean, NDVI 90% percentile, NDVI 10% percentile and NDVI percentile difference from the 90%/10% percentiles. The metrics are normalised among themselves and therefore only comparable between years, but not within a year.
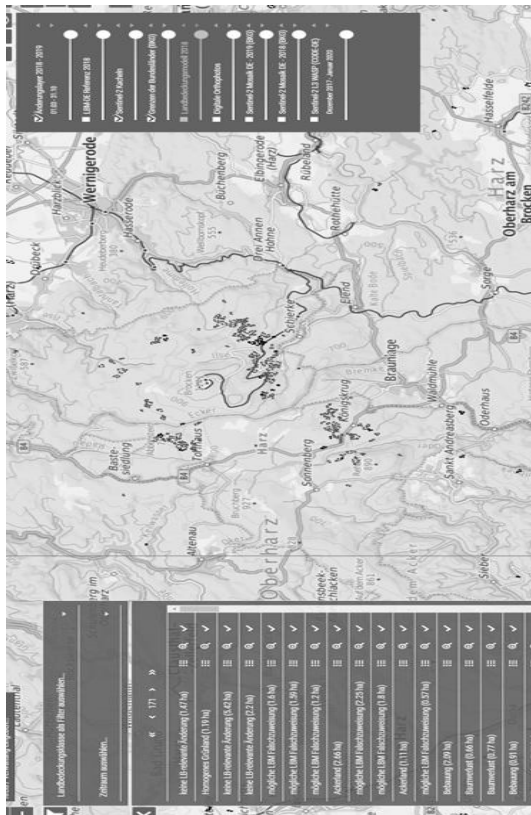
Figure 5. Example of LaVerDi change polygons.

For visualisation and analysis of the change polygons, a variety of layers including transparency sliders are available. In addition, the user has the possibility to check the change polygons provided as a change indication by means of a tick box [Set to checked]. This information is written into the database and can be exported for further use, if required.

5. CONCLUSION

The results show that the methodology delivers predominantly good results. The required minimum overall accuracy of 80% has been achieved or even exceeded for each change layer, taking temporal or seasonal effects into account. Under optimal weather and climate conditions, the methodology can provide very good results. The pre-processing chain developed for the Landscape Change Detection Service delivers standardised image data products and was optimised by an improved cloud and cloud shadow detection for cloud masking (up to a maximum cloud cover of 60%) in the operational, cloud-based service. The developed change detection algorithm relies largely on the properties and capabilities of spatio-temporal features and is fully automated. It considers existing object geometries from the reference LC-model to derive change indications and includes elements of object-based post-processing (MMU filter) and thematic plausibility (knowledge database). As an empirical set of rules, the knowledge database enables the assignment of a change direction for detected change indications (e.g. from forest to built-up area). With the introduction of land cover specific MMU filters, the noise in the initial, pixel-based change indications is minimised and the informative value in the direction of an actual (real) change is significantly increased.

Further technological developments and improvements of the currently used processor (Sen2Cor) for the production of the Sentinel-2 Level-2A input data are to be expected in the future. In addition to continuous performance gains, improvements in cloud and cloud shadow detection in particular are likely, so that LaVerDi will be able to benefit from continuous developments. Moreover, a Europe-wide phenology layer based on Sentinel-2 data has been tendered by the European Environment Agency (EEA) and first results are to be expected for summer 2021[1]. These phenology products will be most likely used in the future to address extreme events (such as the European drought of 2018) and to calibrate the data temporally or to provide additional information.

Finally, with regard to the ongoing Sentinel-2 data evolution, it should be noted that the next version of the Global Reference Image (GRI) will provide improved geometric positional accuracy of the Sentinel-2 data. Validation of the GRI has been completed and integration into the new Processing Baseline 3.0 started in March 2021, so that new acquisitions will benefit from improved positional accuracy. A reprocessing of the Sentinel-2 archive has been also announced by ESA and will presumably start in 2021. The aim is to provide a consistent, multi-temporal registration of the Sentinel-2 data, thus enabling an improved starting point for change analyses in a wide range of applications and thematic areas. Depending on the timing and scope of the reprocessing, the Landscape Change Service will be able to fully benefit from improved geometric accuracy of the multi-temporal satellite imagery data from 2022 at the earliest.

REFERENCES

Ke, G., Meng, Q., Finley, T., Wang, T., Chen, W., Ma, W., Ye, Q., & Liu, T. (2017). LightGBM: A Highly Efficient Gradient Boosting Decision Tree. 31st Conference on Neural Information Processing Systems (NIPS 2017), Long Beach, CA, USA.

Hovenbitzer M., Emig F., Wende C., Arnold S., Bock M., Feigenspan S. (2014). Digital Land Cover Model for Germany – DLM-DE. In: Manakos I, Braun M (Eds) Land use and land cover mapping in Europe: practices & trends. Remote Sensing and Digital Image Processing, 18. Springer, Dordrecht, 255-272.

Probeck, M., Schwab, K., Sevillano Marco, E., Herrmann, D., Sandow, C., Richter, R., Moser, L. (2019). Prototypes of Future Copernicus Land Monitoring Products: The ECoLaSS Project. The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences, Volume XLII-2/W16, 2019 PIA19+MRSS19, Munich, Germany. DOI: 10.5194/isprs-archives-XLII-2-W16-201-2019.

Singh, A. (1989). Review article - Digital change detection techniques using remotely-sensed data. International Journal of Remote Sensing, 10, 989-1003.

Snedecor, G.W. & W.F. Cochran (1967). Statistical methods. State University Press, Ames, Iowa, 517.

Townshend, J., Justice, C., Li, W., Gurney, C., & McManus, J. (1991). Global land cover classification by remote sensing: present capabilities and future possibilities. Remote Sensing of Environment, 35, 243-255.

---

[1] https://land.copernicus.eu/pan-european/biophysical-parameters/high-resolution-vegetation-phenology-and-productivity