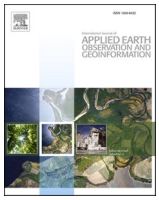
International Journal of Applied Earth Observation and Geoinformation 129 (2024) 103867

Contents lists available at ScienceDirect

International Journal of Applied Earth Observation and Geoinformation

journal homepage: www.elsevier.com/locate/jag



Temporally transferable crop mapping with temporal encoding and deep learning augmentations

Vu-Dong Pham [a,b,*], Gideon Tetteh [c], Fabian Thiel [a], Stefan Erasmi [c], Marcel Schwieder [c,d], David Frantz [e], Sebastian van der Linden [a,b]

[a] Earth Observation and Geoinformation Science Lab, Institute of Geography and Geology, University of Greifswald, Friedrich-Ludwig-Jahn-Str. 16, 17489 Greifswald, Germany
[b] Interdisciplinary Centre for Baltic Sea Region Research (IFZO), University of Greifswald, 17489 Greifswald, Germany
[c] Thünen Institute of Farm Economics, Bundesallee 63, 38116 Braunschweig, Germany
[d] Geography Department, Humboldt-Universitat zu Berlin, Unter den Linden 6, 10099 Berlin, Germany
[e] Geoinformatics – Spatial Data Science, Trier University, Behringstraße 21, Trier 54296, Germany

ARTICLE INFO

Keywords:
Annual crop mapping
Transferability
Temporal encoding
Data augmentations
1D-CNN
Transformer

ABSTRACT

Detailed maps on the spatial and temporal distribution of crops are key for a better understanding of agricultural practices and for food security management. Multi-temporal remote sensing data and deep learning (DL) have been extensively studied for deriving accurate crop maps. However, strategies to solve the problem of transferring crop classification models over time, e.g., training the model with data for a recent year and mapping back to the past, have not been fully explored. This is due to the lack of a generalized method for aggregating optical data with regard to the irregularity in annual clear sky observations and the scarcity of multi-annual crop reference data to support a more generalized DL model. In this study, we tackled these challenges by introducing a method namely Temporal Encoding (TE) to capture the irregular phenological information. Subsequently, we adapted and integrated two methods, i.e., Random Observations Selection (ROS) and Random Day Shifting (RDS) to simulate the variability of temporal sparsity as well as the shifts of crop phenology over different years. We tested this approach with a 1-dimensional Convolutional Neural Network (1D-CNN) and a Transformer Network models. Our results for both classifiers showed that models trained with crop reference data from 2018 and a dense time series of Landsat 7/8 and Sentinel-2 A/B data can be transferred with little decreases in accuracy to map 12 consecutive years from 2010 to 2021. The Transformer Network was slightly more accurate, while the 1D-CNN was much three times faster. Furthermore, the proposed models could achieve similar performances in the same years with and without fully available satellite information. The TE with ROS and RDS appears well suited for improving temporal transferability to support long term historic crop mapping.

1. Introduction

The rapid growth of the global population has led to an increasing demand for agricultural products. Thus, monitoring agricultural land use has become a crucial task in aiding decision-makers to enforce and assess policies that ensure food security and biodiversity, for example, through annual crop type mapping (Karthikeyan et al., 2020). In this context, Earth observation (EO) data has been widely used (Weiss et al., 2020). The intra-annual information from EO time series reveals the distinctive growing patterns or phenologies of different crops and, hence, shows great potential for distinguishing crop types (Zeng et al., 2020). In recent years, openly available EO data at high (10 m) and medium (30 m) spatial resolutions have considerably increased in terms of temporal frequency because of the existence of Sentinel-2A/B images (revisit interval of 5 days) coupled with the already existing Landsat images (revisit interval of 16 days) (Li and Roy, 2017). As a response, recent studies used this dense temporal EO information, particularly harmonized time-series information derived from Landsat and Sentinel-2 observations, to map crop types from regional to national scales (Griffiths et al., 2019; Piedelobo et al., 2019; Liu et al., 2020; Pan et al., 2021).

One of the major challenges when using multi-temporal EO data for

* Corresponding author.
E-mail address: vudong.pham@uni-greifswald.de (V.-D. Pham).

https://doi.org/10.1016/j.jag.2024.103867
Received 23 December 2023; Received in revised form 5 April 2024; Accepted 22 April 2024
Available online 26 April 2024

mapping over multiple years and larger areas is the need for regular temporal features that align better with the requirement for data completeness of machine learning models. Due to different orbit constellations and the highly irregular occurrence of clouds in optical EO imagery, temporal aggregation methods are frequently employed to capture consistent time-series information from multi-temporal data (Gómez et al., 2016). In the context of crop mapping, numerous methods adopted in the literature aim to sample crop phenological information and transform it into equidistant features for classification workflows. These methods include, but are not limited to, best-pixel compositing (Griffiths et al., 2019), spectral temporal metrics (Asam et al., 2022), temporal smoothing (Blickensdörfer et al., 2022), phenology compositing (He et al., 2021), and temporal mosaicking (Vaudour et al., 2021). While the strength of each method has been demonstrated in previous studies, they all have specific data requirements such as a temporal resolution that fits the specific application (Qiu et al., 2023). Another data aggregation method that is commonly used to capture crops' seasonal information is linear interpolation (Qiu et al., 2014; Inglada et al., 2017; Sadeh et al., 2021; Yan et al., 2021). Linear interpolation does not require any parameterizations and has been shown to perform principally well within a consistent temporal density range (Valero et al., 2016). However, the performance of linear interpolation over multiple years with varying temporal data densities has not been tested. Hence, previous temporal aggregation methods often limit the application to similar annual data situations and hinder a model transfer, e.g., to periods with different sensor constellations.

In recent years, newly developed deep learning (DL) architectures have been proposed within the context of tackling the irregularity of temporal data. For example, Garnot et al. (2020) used positional encoding as a means of representing the date of acquisition for Sentinel-2 data from a single year, to be used with the self-attention mechanism of the Transformer Network (Vaswani et al., 2017) for crop classification. The positional encoding allowed the classification model to predict crop types using stacked Sentinel-2 data from varying numbers of acquisitions. However, the variation is considered at the image level, not at the pixel level. Therefore, cloud cover differences or orbit constellations required the authors to utilize predetermined input vector lengths, which had to be generated through linear interpolation (Garnot et al., 2020). Providing consistent feature lengths at the pixel level is an important requirement for machine learning models to make use of batch processing for pixel-based mapping, especially when mapping over large areas. Rußwurm and Körner (2020), on the other hand, showed that DL models can map crop types using "raw" temporal information by including the cloud observations in the stacked time-series data. This way, the stacked data leads to equal feature length in every pixel without the need for interpolation. However, the study solely examined data in one year (2018) with dense Sentinel-2 observations, where the clear observations exceeded the cloudy observations by far for most pixels. Additionally, the approach is limited when facing no-data problems in regions at the edge of the observation swaths when combining multiple satellite sensors. Still, this approach has not been tested in unknown years with fewer temporal images, more clouds, or when using multiple satellite sensors. Overall, while previous studies have examined some potentials of different DL architectures to handle irregular temporal data, it is still important to acquire input features of equidistant length for every pixel to enable parallelism of batch processing for large area mapping. Given the complexity of the choice of aggregation methods regarding temporal density, there is a need for more generalized and flexible feature vectors that can capture the highly irregular time-series to map crop types based on their distinct phenologies.

Despite the high number of studies investigating the efficacy of different data sources or classification algorithms, another challenge that has so far gained rather little attention is the temporal transfer of classification models with multi-temporal remote sensing data, which are characterized by inter-annual differences in data density and distribution. Such a model transfer is needed to derive a consistent time series of crop type maps to enable the analyses of crop rotations over time. Commonly, crop classifications based on supervised machine-learning methods require crop reference information to train the classifier (Kamilaris and Prenafeta-Boldú, 2018). However, crop reference information is scarce in many regions, especially for multiple years. To monitor annual crop rotations, a common approach is to use crop reference data from all available years as training data to obtain a generalized crop classification model to then limit the analysis to the years where reference information existed (Blickensdörfer et al., 2022; Liu and Zhang, 2023; Xuan et al., 2023). However, transferring supervised models to different domains with limited or no training data through domain adaptation is often challenging (Tuia et al., 2016; Kellenberger et al., 2021; Capliez et al., 2023; Wang et al., 2023b). In recent studies, different data augmentation methods have been used to simulate the variations in temporal data for generalizing classification models. For hyperspectral data, some augmentation methods were proposed such as self-supervised (Liu et al., 2023), deep learning generative (Wang et al., 2019) or integrating vector transformation (Wang et al., 2023a). For multi-spectral data, Sainte Fare Garnot et al. (2022) introduced a temporal dropout method that randomly removed observations in time-series data during training. This method was applied to the fusion of Sentinel-2 and Sentinel-1 data in one year (2019) and showed some improvements in overall crop mapping accuracy in the same year. In another study, Nyborg et al. (2022) introduced a temporal shifting method, i.e., by modifying the acquisition date by certain days of the Sentinel-2 imagery during training. The authors showed that the method improved the spatial transferability of the crop model between two countries (Denmark and France) where similar crop types have different phenological growths. Overall, these augmentation methods were only tested in a single period, where both temporal density and crop phenology are consistent across space with fixed numbers of satellite data. Hence, the potential of such methods in terms of temporal transferability, particularly mapping over multiple years, has not been fully investigated.

Against the background of these challenges, the overarching aim of our study was to develop a set of methods for improving the temporal transferability of crop classification models trained with single-year reference data. We addressed two major challenges: 1) creating a generalized method for aggregating optical data to capture irregular time-series information and 2) temporally transferring a classification model trained in a single year. To achieve this, we first introduced a data structure concept that encodes annual time-series information into equidistant features, which is applicable regardless of the temporal resolution and sensor constellation. Secondly, inspired by previous studies on utilizing augmentation techniques for training, we applied two data augmentation methods to the proposed data structure to simulate the temporal density and distribution information of crop data. Subsequently, we adopted a 1-dimensional Convolutional Neural Network (1D-CNN) and a Transformer Network to construct the crop classification models, where we trained the models based on the augmented dataset. We experimented with the proposed methods by mapping crop types for the entire state of Brandenburg (Germany) using all available Landsat (5 − TM, 7 − ETM+, 8 − OLI) and Sentinel −2 (A, B) imagery. We selected the year 2018 with the highest temporal information to train the model and applied the model to map crop types in 12 consecutive years (2010 – 2021) where there were large variations between years in the density of the time series of satellite images. We demonstrated the effectiveness of the proposed method in terms of:

- Temporal transferability when mapping annual crop types over 12 years
- Consistent mappings despite artificially omission of satellite data

2. Materials and methods

2.1. Study area

Our study area covers the German federal state of Brandenburg with a total area of 29,640 km^2 (Fig. 1a). Nearly 45 % of its land cover is used for agricultural purposes (Schindler, 2004). In the area, grassland accounts for approximately 23 % of the total agricultural area, 77 % of the area for cropland (Troegel and Schulz, 2018). Due to low rainfall, there is only a small percentage of irrigated crop types (around 2 % of cropland area), where major crop types such as maize, rye, wheat, barley, and fodder crop are dominant products.

2.2. Earth observation data

For the study area, we acquired all available Landsat images (5 – TM, 7 – ETM+, 8 – OLI) with cloud cover < 75 % from 2010 to 2021 provided by the United States Geological Survey (USGS) and all Sentinel-2 (A, B) images with cloud cover < 85 % from 2017 to 2021 provided by the European Space Agency (ESA).

Due to orbit constellations and cloud cover, the number of acquisitions differs between sensors and years (Fig. 1b). Landsat and Sentinel-2 data were harmonized and pre-processed to Level-2 surface reflectance using the Framework for Operational Radiometric Correction for Environmental monitoring (FORCE, Frantz, 2019). To map crop types, we used the six reflectance bands available in both Landsat and Sentinel-2 data (Red, Green, Blue, Near-Infrared, Shortwave-Infrared 1, Shortwave-Infrared 2). Three additional indices were included: Normalized Difference Vegetation Index (NDVI, Tucker, 1979); Normalized Difference Water Index (NDWI, McFeeters, 1996); and Soil-Adjusted Vegetation Index (SAVI, Huete, 1988). This combination of bands and indices has been shown to give the best result for crop classification in recent studies (Oliphant et al., 2019; Blickensdörfer et al., 2022). For Sentinel-2 data, the reflectance bands and indices were resampled to 30 m resolution using an approximated point spread function according to Frantz (2019) in order to match the Landsat data and allow consistent mapping from 2010 to 2021. We extracted all annual clear sky observations (CSO) from the study area using the cloud masking approach of Frantz et al. (2018). In our study area, there are large differences in annual CSO in terms of density (Fig. 1b) and distribution (see Supplement 1), especially for periods of different sensor constellations, but also dependent on annual weather conditions and related cloud cover.

The phenological profiles, as depicted by satellite time series, for the different agricultural land use types for a full calendar year are shown in Fig. 2. Agricultural land use in Germany is dominated by single cropping systems with one main crop that is harvested between June and November. However, a cover crop is often grown after harvest or during the winter period. Distinct phenological stages, such as growing, ripeness, and harvest, can be clearly recognized but also ambiguities especially for crops from the same seasons, i.e., winter crops or spring crops. In our work, we focus only on the main crop.

2.3. Crop reference data

We acquired annual crop reference information for the entire study area from the Integrated Administration and Control System – IACS. The crop information was provided in the Land Parcel Information System (LPIS), with each agricultural parcel digitized into georeferenced polygons. LPIS data for Brandenburg was provided annually over 12 consecutive years (2010 – 2021) based on reports by farmers and landowners (MLUK – Version 2.0; data can be accessed through https

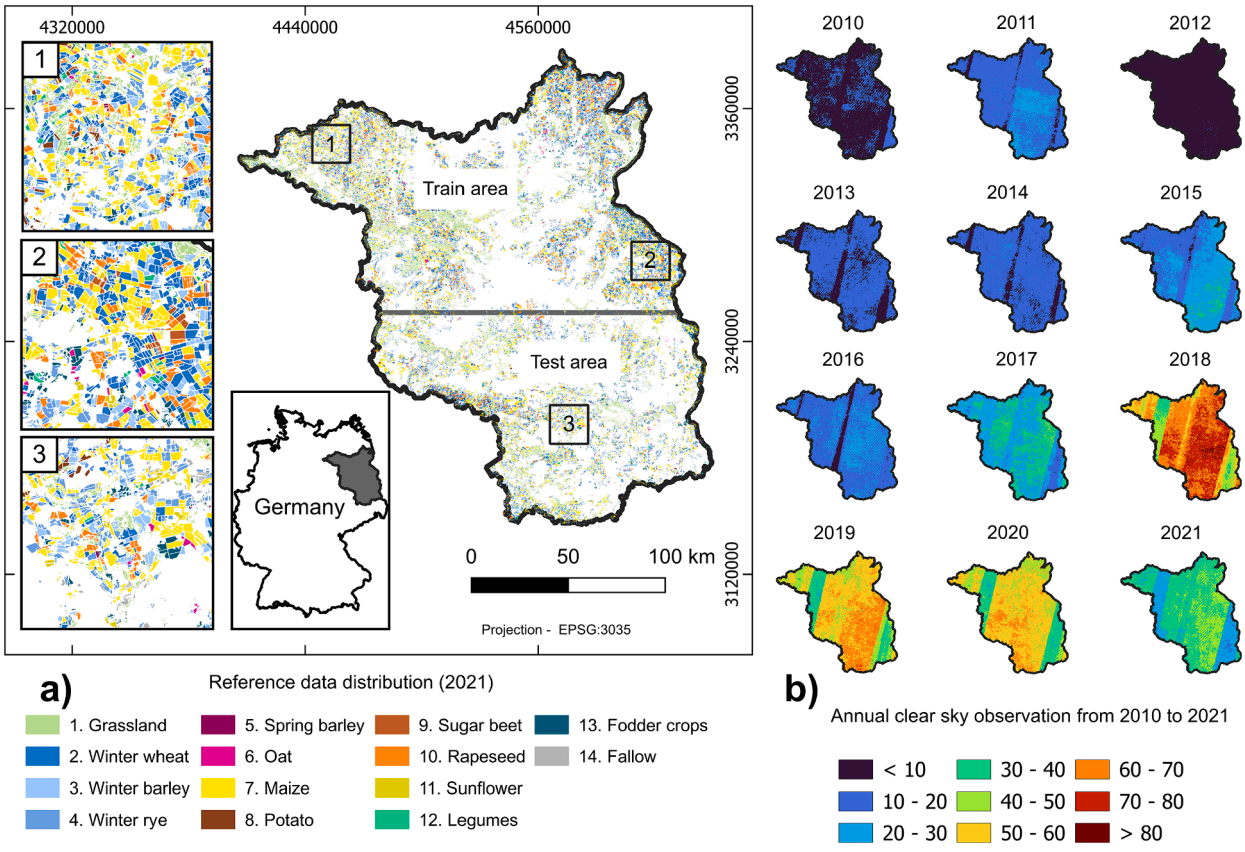


Fig. 1. A) the study area (brandenburg, germany) and examples of crop reference data in one year (2021); b) annual clear sky observations from all available observations from landsat (tm, etm+, oli) and sentinel-2 (a, b) from 2010 to 2021.

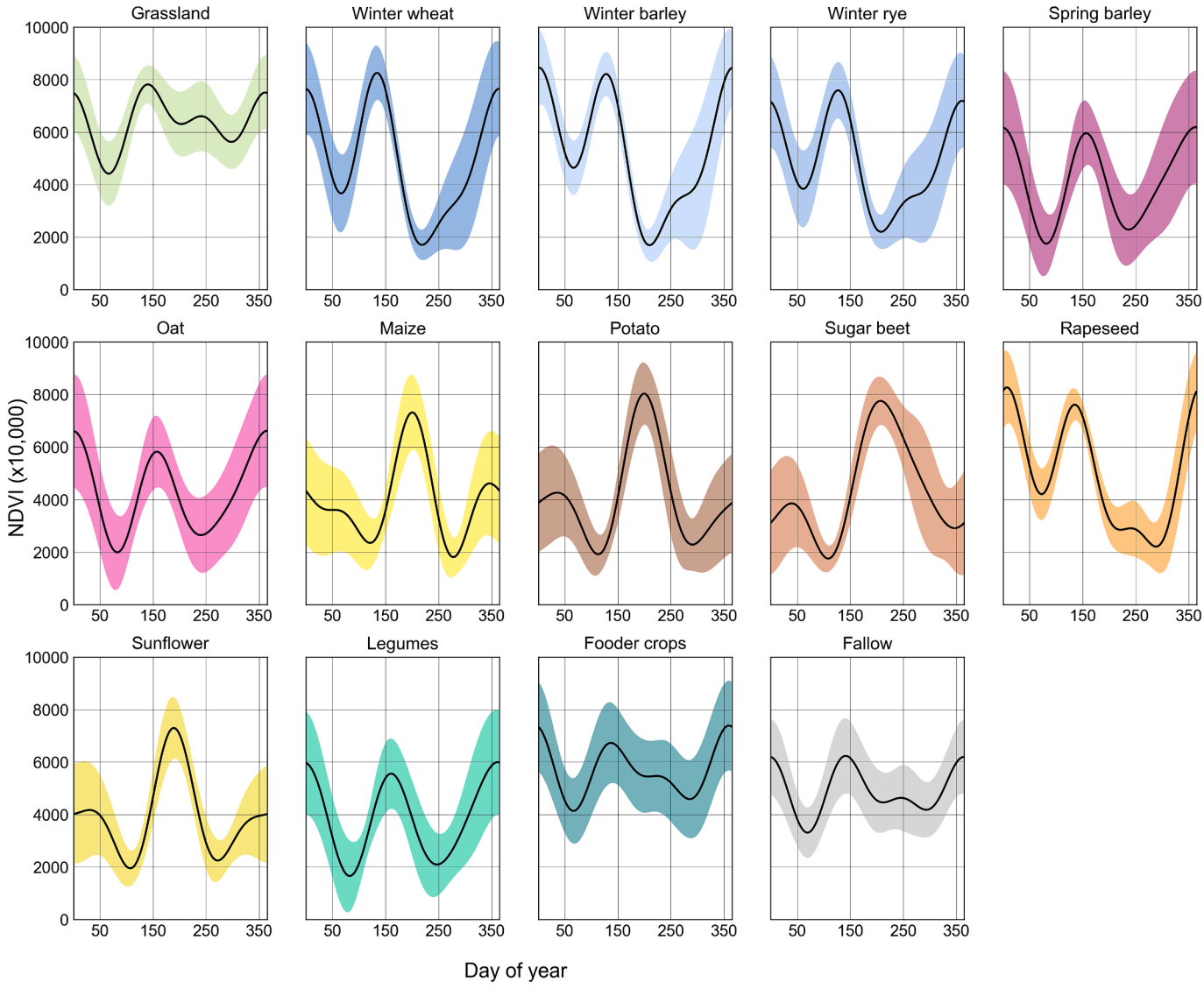


Fig. 2. Annual NDVI profiles (2018) of 14 crop types. Daily NDVI data (1000 reference data samples per crop type) was interpolated using a harmonic function (Zhu et al., 2015). Mean values (black line) and +/- standard deviation (filled color) are shown.

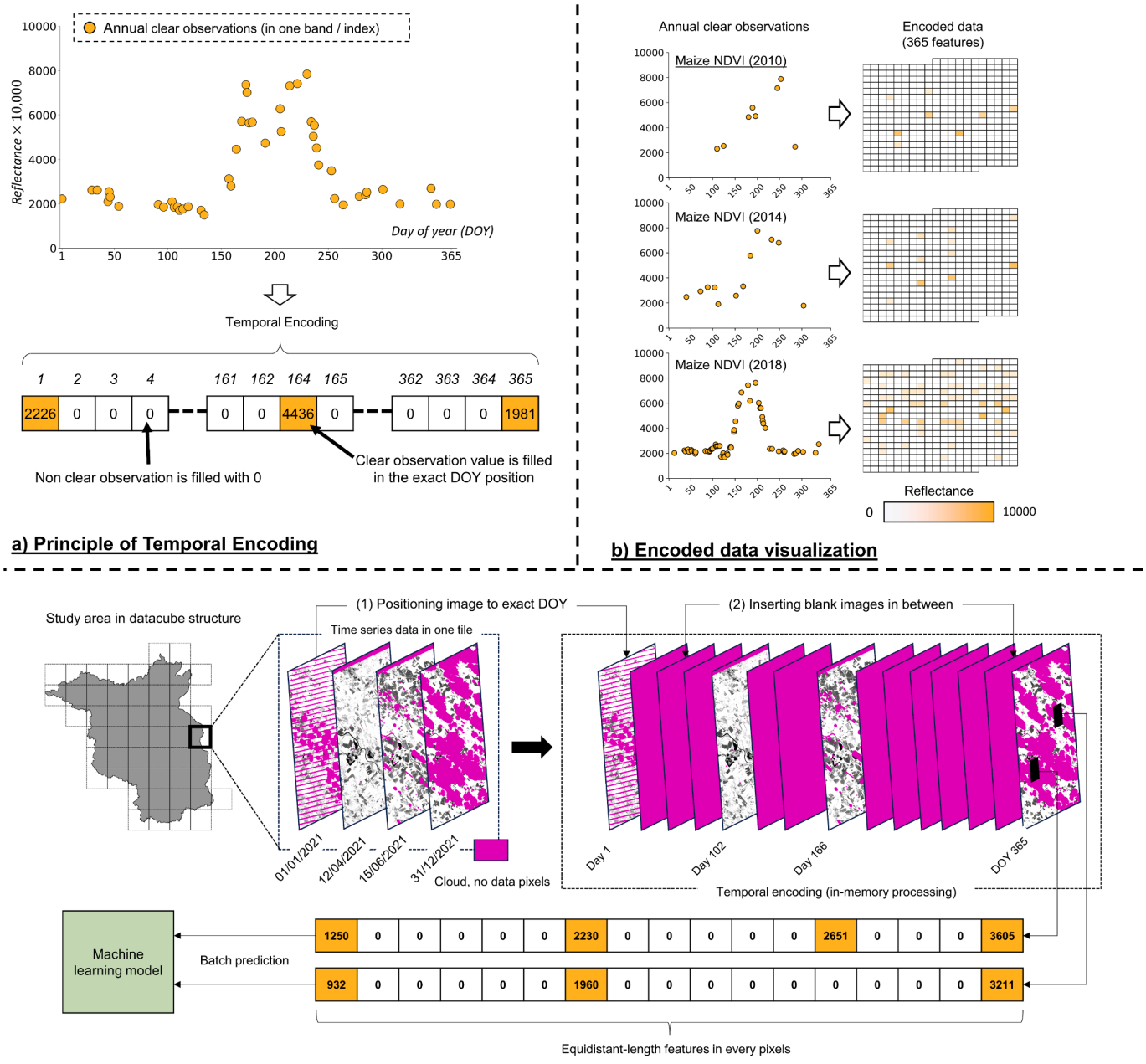


Fig. 3. A) principle ofTemporal Encoding (TE); b) Example of annual NDVI (2010, 2014, 2018) values encoded into TE features; c) Demonstration of applying TE in batch processing over large area with datacube structure.

://geobroker.geobasis-bb.de/ last access: 28/02/2024). Originally, the LPIS dataset consisted of more than 300 crop classes for the entire Germany, of which 30 major classes existed in Brandenburg state. Hence, to achieve sufficient class areas for the experiments we aggregated the original classes into 14 crop types (including grassland) for the study area (see Supplement 2 for crop aggregation details). We split the study area into two parts, one for training (northern part) and one for testing (southern part, Fig. 1a). This spatial separation was done to allow for an unbiased evaluation of our proposed method in terms of generalization. To demonstrate the temporal transferability of the proposed method when mapping years with sparser temporal information, we selected crop reference data in 2018 for the training process, i.e., the year that had the highest number of clear sky observations or highest temporal data density. As a baseline, we selected 4000 training samples per crop type in 2018 (56,000 samples in total for the 14 crop types). This selection was based on a pre-analysis between the amount of training data and mapping accuracy (see Supplement 4). The training pixels were randomly selected across the training area (northern Brandenburg). For testing, since we have the reference for all crop parcels in the area for 12 years, we evaluated the classification model by comparing wall-to-wall annual mapping results to all available crop reference information from 2010 to 2021. Details of numbers of crop reference data are provided in Supplement 3.

2.4. Temporal encoding for annual observations

To capture the crops' phenological information from highly irregular annual temporal information in the EO data, we developed a universally applicable method to stack images, which we refer to as Temporal Encoding (TE). TE represents all observations (clear and non-clear) relative to their acquisition dates in a standardized and equidistant structure (Fig. 3). The process follows:

- First, for each location (pixel) and reflectance band (or index), we extracted reflectance values and the corresponding day-of-year (DOY) for all CSO on an annual basis.
- Next, we encoded this time-series of observations into a fixed-length data feature consisting of 365 features corresponding to 365 days of the year.
- Subsequently, we filled the cells in the data feature with their corresponding reflectance values.
- For the DOYs without observation (no-data observations) and non-clear observations (cloud, cloud shadow), 0 is assigned.

The average reflectance value per band was used on days with more than one observation (observations from multiple sensors). In a year with 366 days, if there is a clear observation on day 366, the reflectance value will be encoded on day 365 (if there is no clear observation on day 365), otherwise day 366 is ignored. With the annual TE, each band (or index) contains 365 features and when n bands (or indices) are used the encoded data is a 2D-matrix of the dimension $365 \times n$ (in this study, 365×9, since we used 6 bands and 3 indices).

The EO data in our study area was tiled in a datacube structure using FORCE, which divided the area into 51 equally sized tiles of 30×30 km (1000×1000 pixels) (Fig. 3c). Thus, for TE in each tile, we created an image cube with a fixed size of $1000 \times 1000 \times 365 \times 9$. All pixels in the image cube were initialized with values of 0. Next, time-series images were inserted into the image cube where the position of each image corresponded to its acquisition date. Note that, within each image, pixels flagged with cloud and no-data were assigned with 0 values. Thus, all pixels are guaranteed with fixed feature lengths regardless of the total number of time-series images or the presence of cloud cover as well as images at the edge of the observation swaths. This way, pixel-based mapping can be performed in batches i.e., $m \times 365 \times 9$ where m is the number of pixels that can be mapped in one feed-forward.

2.5. Applying data augmentation to TE data

By using the TE data structure as input feature, we can flexibly modify the time series information to simulate, e.g., different densities and distributions.

a) Simulating data sparsity with Random Observation Selection

Adapted from the dropout observations techniques introduced in Sainte Fare Garnot et al. (2022), we used what we call Random Observation Selection (ROS) with our proposed TE data structure. By using ROS, we simulated temporal data sparsity and enabled the machine learning model to learn how the same crop phenology is captured with varying data availability. We implemented ROS within TE by randomly selecting a fraction of all annual CSO and then removing unselected observations by changing their encoded value to 0. In our implementation, the proportion of selected observations is fully random within the range of 5 % – 100 % of the original data (Fig. 4a).

b) Random Day Shifting – Simulating meteorological effects on phenology.

Inspired by the temporal shifting method introduced in Nyborg et al. (2022), we applied this approach to TE data, but shifted each observation individually, which we refer to as Random Day Shifting (RDS). Based on the assumption of crop phenology changing between years, we applied RDS to modulate inter-annual phenological variations. For all annual CSO, we randomly shifted their DOY, i.e., its position in the first dimension of the 365x9 feature array, forward or backward (Fig. 6b). Here, we randomly specified the shifting range from −16 to +16 days (including 0), i.e., the maximum date shifting was related to the 16-day revisit interval of the Landsat satellite (Fig. 4b).

c) Implementation of ROS and RDS

During training, we first extracted the annual CSO for each training data location and converted it into the proposed TE data structure. Then, we applied ROS and RDS sequentially to create augmented data (Fig. 4c). This way, we continuously generated new training data to train the model until it reached sufficient accuracy (Fig. 5). Additionally, ROS and RDS were also implemented individually to demonstrate the impact of each method in Section 3.2.

2.6. Crop classification with deep learning models

In this study, we tested two different deep learning classifiers. First, we used 1-dimensional Convolutional Neural Network (1D-CNN) architecture, which has been widely used for crop types and land cover classifications (Zhong et al., 2019; Debella-Gilo and Gjertsen, 2021; Zhang et al., 2023). 1D-CNN has been shown to be a lightweight architecture that is well capable of exploiting the time-series sequence data at high computational efficiency. Particularly, the TE data structure is highly beneficial from the 1D-CNN, by simply applying the 1D convolutional filters along the temporal dimension to reduce the number of model's parameters. Specifically, the pixel-based classification 1D-CNN model takes an input of $i = 365 \times 9$ (Time step × Bands) and applies a 1D kernel convolution with the kernel shape of 16×64 (filter size × filter numbers). A batch normalization was applied to the resulting feature maps (Ioffe and Szegedy, 2015), followed by a rectified linear unit activation (ReLU)

$$ReLU(x) = \max(0, x) = \begin{cases} x \ if \ x > 0 \\ 0 \ otherwise \end{cases}$$

Subsequently, a 1D max pooling operator with window size of 2 was applied to reduce the temporal dimension of the data. The process of 1D convolution −> batch normalization −> ReLU −> max pooling was repeated two more times. At the end of the network, the feature maps were flattened and a fully connected layer with 14 units was applied to predict the crop type probabilities.

To evaluate the effect of data augmentation in a more sophisticated network architecture, we ran an additional experiment with a state-of-

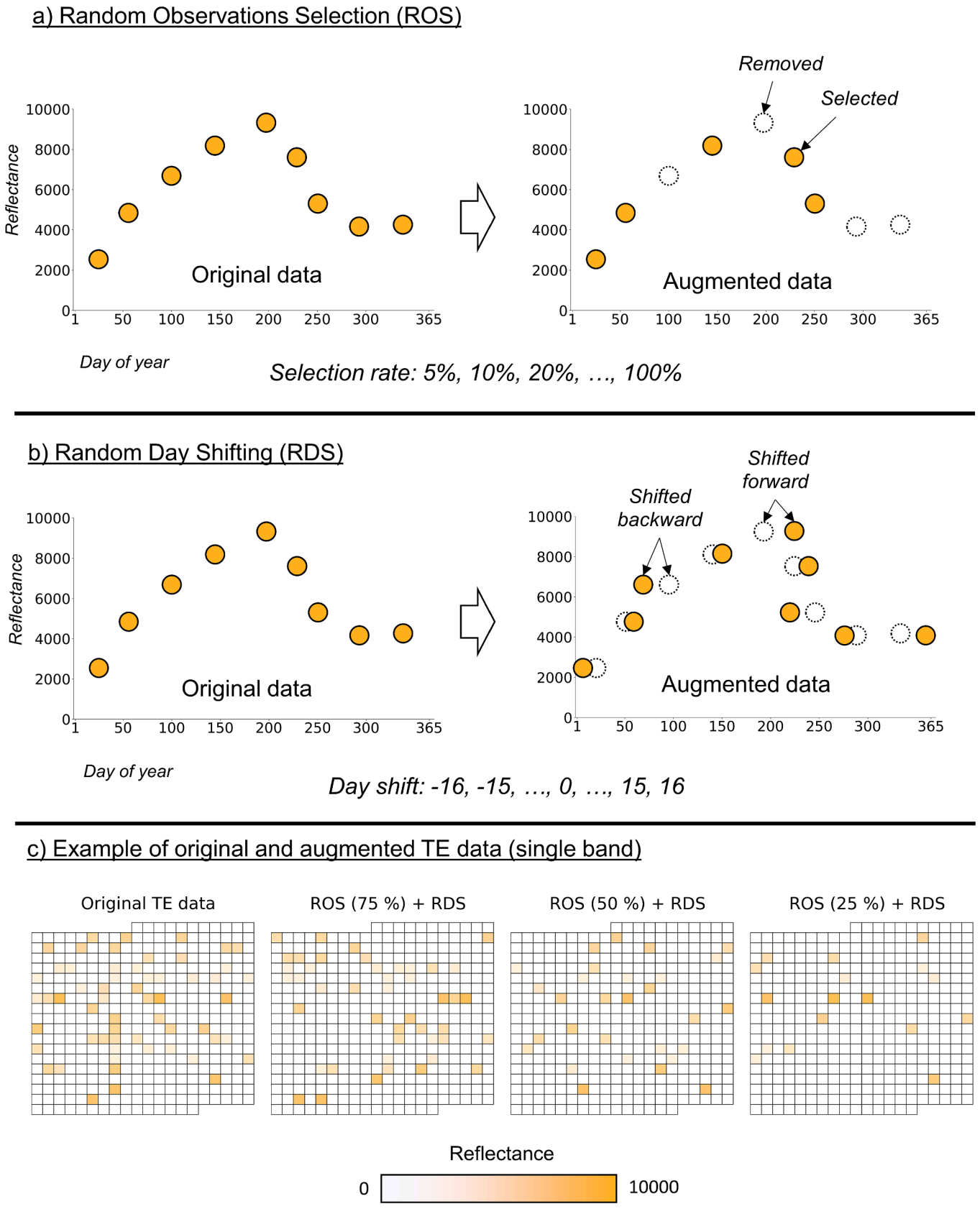


Fig. 4. a) Random Observation Selection (ROS) – randomly selecting a proportion of observations from the original time-series data; b) Random Day Shifting (RDS) – randomly shifting the date of every observation by +/- 16 days; c) An example of original time-series data encoded with Temporal Encoding (365 features) and its three corresponding augmented versions are applied with ROS (75 %, 50 % and 25 %) and RDS.

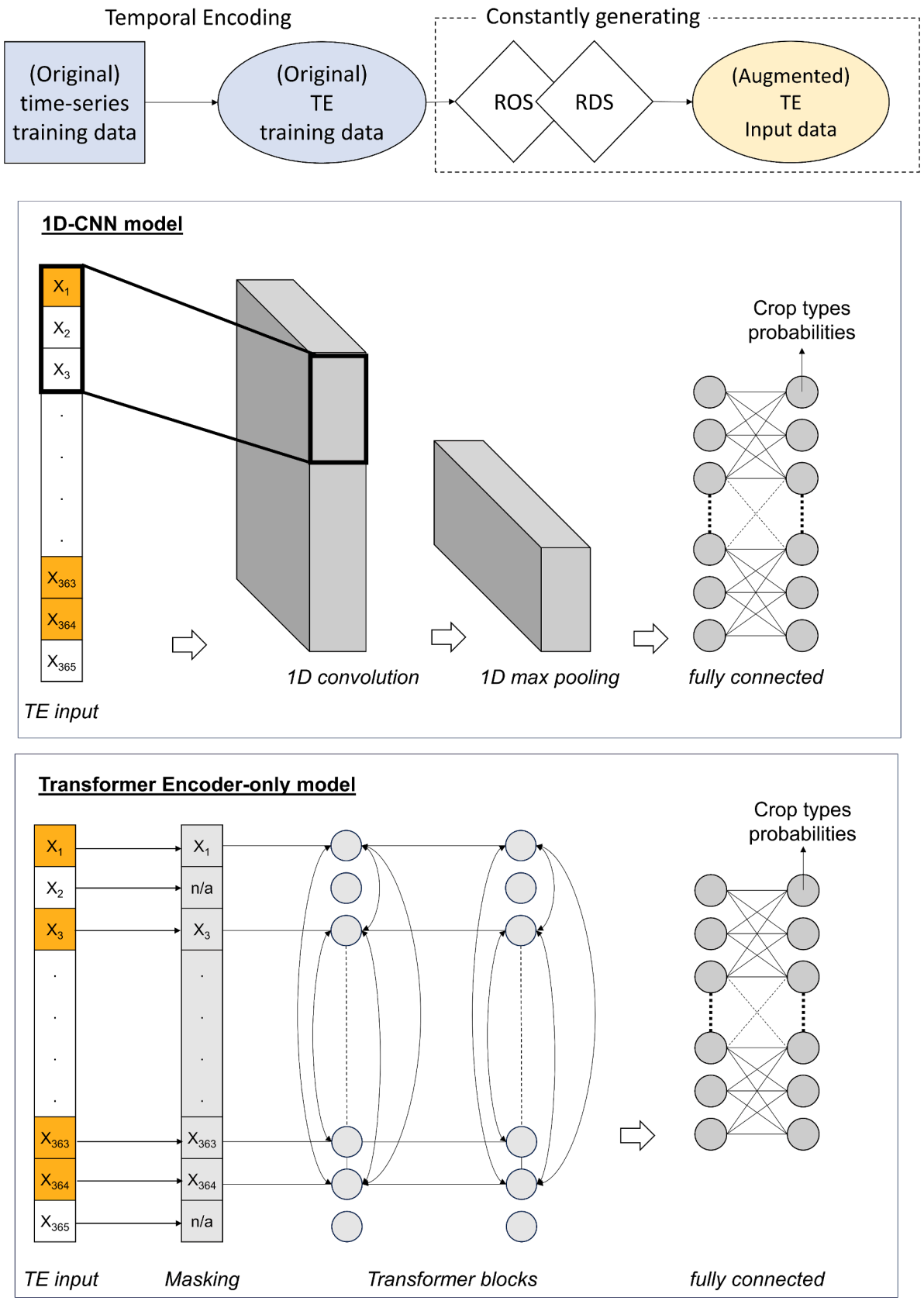


Fig. 5. Workflow of generating training data with ROS and RDS to train the deep learning models.

the-art Transformer network, which has been shown to outperform traditional CNN-based models (Xu et al., 2020; Rußwurm et al., 2023; Zhang et al., 2024). We used the Transformer model proposed by Zhang et al. (2024), which has an encoder only architecture and has been used for land cover classification based on satellite time-series data. We changed the dimension of input and output of the Transformer network to fit with our dataset, while the rest of the network architecture remained as proposed in the original study.

2.7. Experiment setup and accuracy assessments

We first tested for differences in performance between the two deep learning models (1D-CNN and Transformer). Due to the heavy processing of the Transformer network (see Section 3.1), in the second part, we only used the results from the 1D-CNN model for a detailed evaluation of the TE and data augmentation methods.

To establish the baseline performance of our approach, we trained a model with the proposed TE combined with the two augmentation methods (ROS and RDS). We evaluated the temporal transferability by comparison with a linearly interpolated data structure, which is one of the most common temporal aggregation methods for crop mapping that has been widely used (Qiu et al., 2014; Inglada et al., 2017; Sadeh et al., 2021; Yan et al., 2021). Hence, to compare with the proposed approach, we used linear interpolation to generate daily CSO, resulting in the same feature length as TE data (365 features for each band). Additionally, to analyze the impact of the two augmentation methods, we also trained three additional models with TE data structures but using only either ROS, RDS, or neither one.

In summary, we evaluated the following models:

- *Baseline model* (model trained with TE features, ROS and RDS are applied, 1D-CNN and Transformer classifiers)
- *TE-only model* (model trained with TE features, no augmentation applied, 1D-CNN and Transformer classifiers)
- *TE-ROS model* (model trained with TE features, only ROS is applied, 1D-CNN classifier)
- *TE-RDS model* (model trained with TE features, only RDS is applied, 1D-CNN classifier)
- *LI model* (model trained with linearly interpolated features, 1D-CNN classifier)

All models were trained with crop reference data from 2018, with the same number of training pixels specified at 4,000 per crop type (a total of 56,000 training samples for 14 crop types). Performance analysis is based on comparing annual crop maps from 2010 to 2021 with reference information from the test area in the respective years. We measured the Overall Accuracy (OA) which refers to wall-to-wall comparisons between annual crop maps produced by each model to the test reference.

We additionally tested the transferability of the *Baseline model* regarding the intra-variations of satellite data. For every individual year, we intentionally reduced the annual data density using only one single satellite sensor (e.g., using only Landsat 7 data to map crop types in 2019). We then evaluated the crop maps' OA for different single or multi-modal sensor collections.

3. Results

3.1. Comparison of different deep learning models

We compared the performances of the 1D-CNN and Transformer models for mapping transferability with and without augmentation methods (ROS and RDS) (Fig. 6). Without using the augmentations, the Transformer model outperformed the 1D-CNN when transferred to most years by roughly 10 % – 25 % in absolute OAs. After applying ROS and RDS, the two models showed similar performances with OAs higher than 60 % in all years. Except for 2010 where the Transformer model had around 7 % higher OA, it only marginally outperformed the 1D-CNN by 1–3 %. In 2012, the 1D-CNN model yielded about 4 % higher OA than the Transformer.

To test mapping speed, we deployed the two models on higher-performance clusters. We assigned each model to 20 compute nodes; each node has 18 CPUs (2.9 GHz) and 60 Gb RAM. For each node, we used the pre-trained model to map crop type in one random tile (1000 × 1000 pixels) from a random year. The results showed that the Transformer model took 3 times longer (60 min on average) than the 1D-CNN model (18 min on average). Therefore, we focus on the results from the 1D-CNN in the following.

3.2. Temporal transferability of different input features

We assessed the temporal transferability of the approach with the *Baseline model (1D-CNN)* in comparison to the linear interpolation

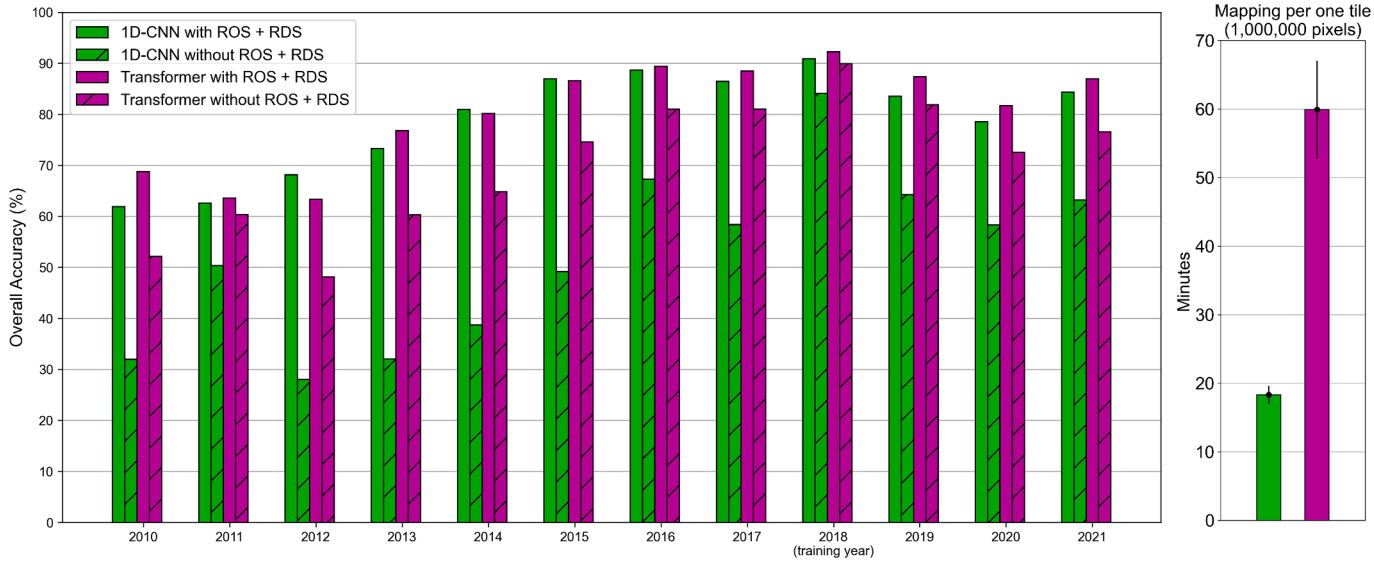


Fig. 6. (Left) Overall Accuracy (%) for annual crop maps for 12 years (2010 – 2021) mapped with the 1D-CNN and Transformer models (with and without ROS + RDS). (Right) Mapping time for one tile (1,000,000 pixels) for each model.

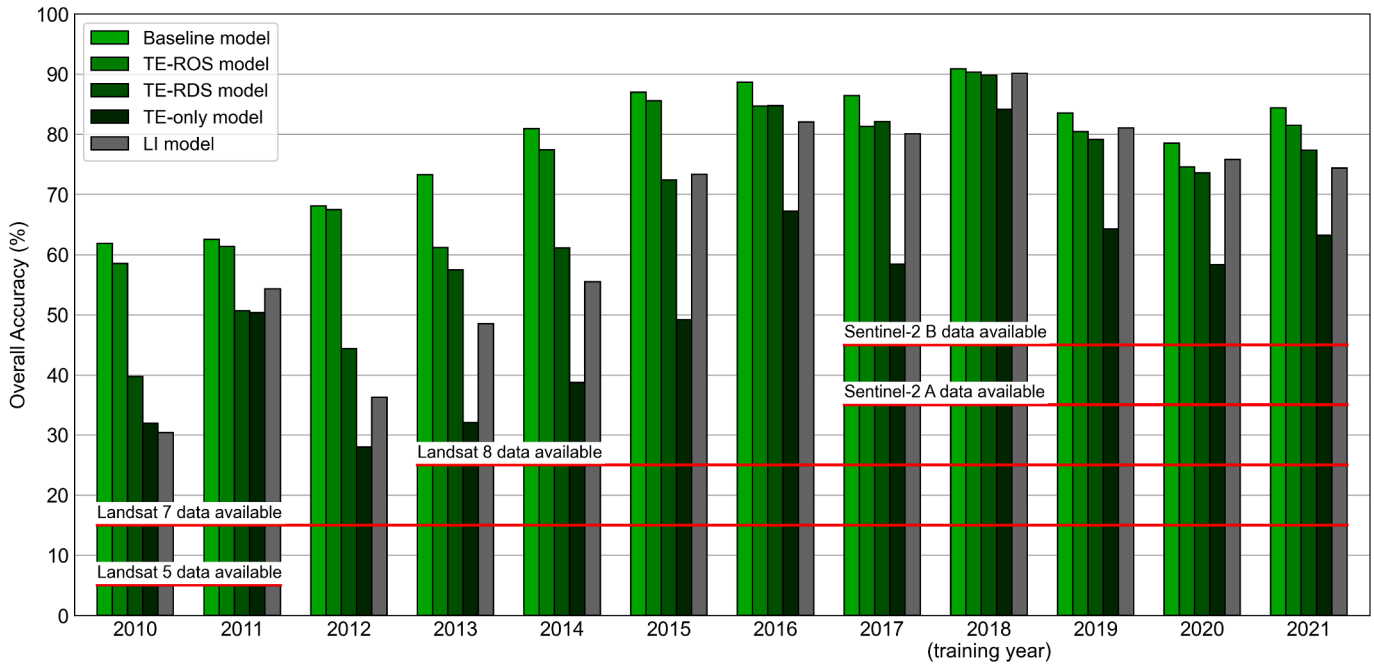


Fig. 7. OA (%) for annual crop maps from 12 years (2010 – 2021) for five models: 1) Baseline model 2) TE-ROS model, 3) TE-RDS model, 4) TE-only model and 5) LI model.

approach (*LI model*) using the annual OA of each model (Fig. 7) and class-wise performances (Supplement 6).

In the training year of 2018, both models yielded similar results with OAs of around 90 %. However, when transferred to other years, the *Baseline model* outperformed the *LI model* by 2–10 % in 2019–2021 and 4–12 % in 2015–2017. In 2010–2014, where temporal data was sparse, the gaps in accuracy were wider. The *LI model* achieved 30 % to 54 %, while the *Baseline model* produced results exceeding 60 % for the first four years.

Visual assessment (Fig. 8) revealed comparable performances in 2018 and 2020 of both models. For those years, all predicted maps showed accurate and clear distinctions of crop parcels when compared with the reference maps. In the examples for 2010 and 2012, the *LI model* showed high errors, with most crop types being misclassified as

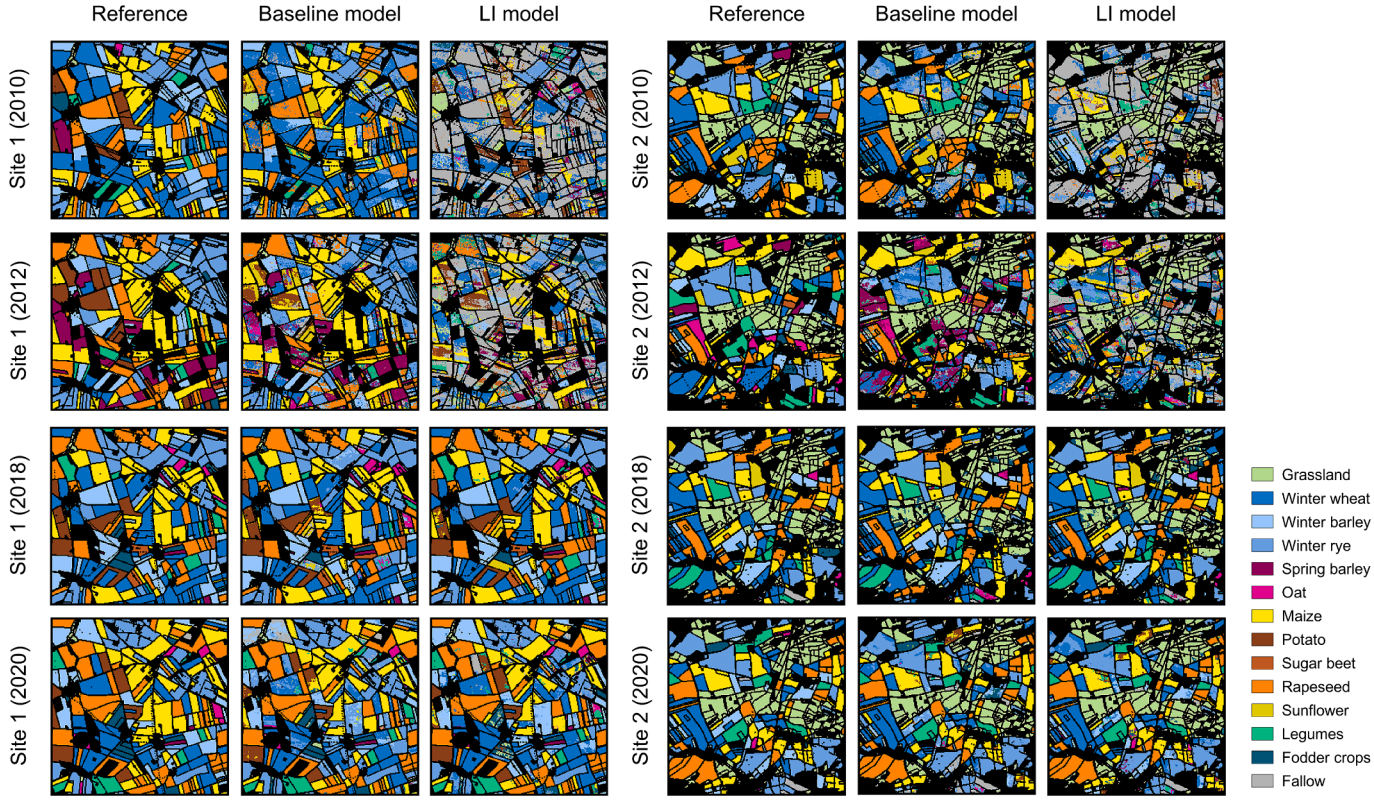


Fig. 8. Visual comparison of crop maps from *Baseline model* and *LI model* for four different years (2010, 2012, 2018, 2020) and for two sites within the test area.

Fig. 9. Annual overall accuracies (OAs) produced by *Baseline model* when using single and all available satellite data. (*) Sentinel-2B data in 2017 was reduced, because the nominal phase started late in the year.

Fallow. In contrast, the maps based on the *Baseline model* produced estimates closer to the reference data. Major crop types such as *Maize*, *Winter wheat* and *Rapeseed* were mapped correctly, while other minor classes observed some misclassifications.

3.3. The impact of ROS and RDS

Looking at the impact of augmentation methods (Fig. 7), specifically comparing the *Baseline* and the *TE-only* models, the positive impact of ROS and RDS on temporal transferability becomes apparent. The augmentation methods improve OA by around 20 % in four years (2016, 2019, 2020, 2021) and nearly 30 % in 2017. The improvements were further emphasized in earlier years. Except for 2011, the augmentation methods resulted in nearly double OAs compared with the standard model from 2010 to 2015. In the training year (2018), however, ROS and RDS only improved the OA marginally, around 5 %.

Comparing ROS and RDS individually, the results suggested that *TE-ROS model* outperformed *TE-RDS model* in most years. Most significantly, the OAs of *TE-ROS model* was 20 % higher than those of *TE-RDS model* in 2010, 2012, 2014 and 2015. In the remaining years, training with both ROS and RDS always improved the transferability compared with the *TE-only model*.

3.4. Single and multiple satellite information mapping

We evaluated the temporal transferability of the *Baseline model* for all years after changing the intra-annual temporal information by selecting only single satellite sensors (Fig. 9).

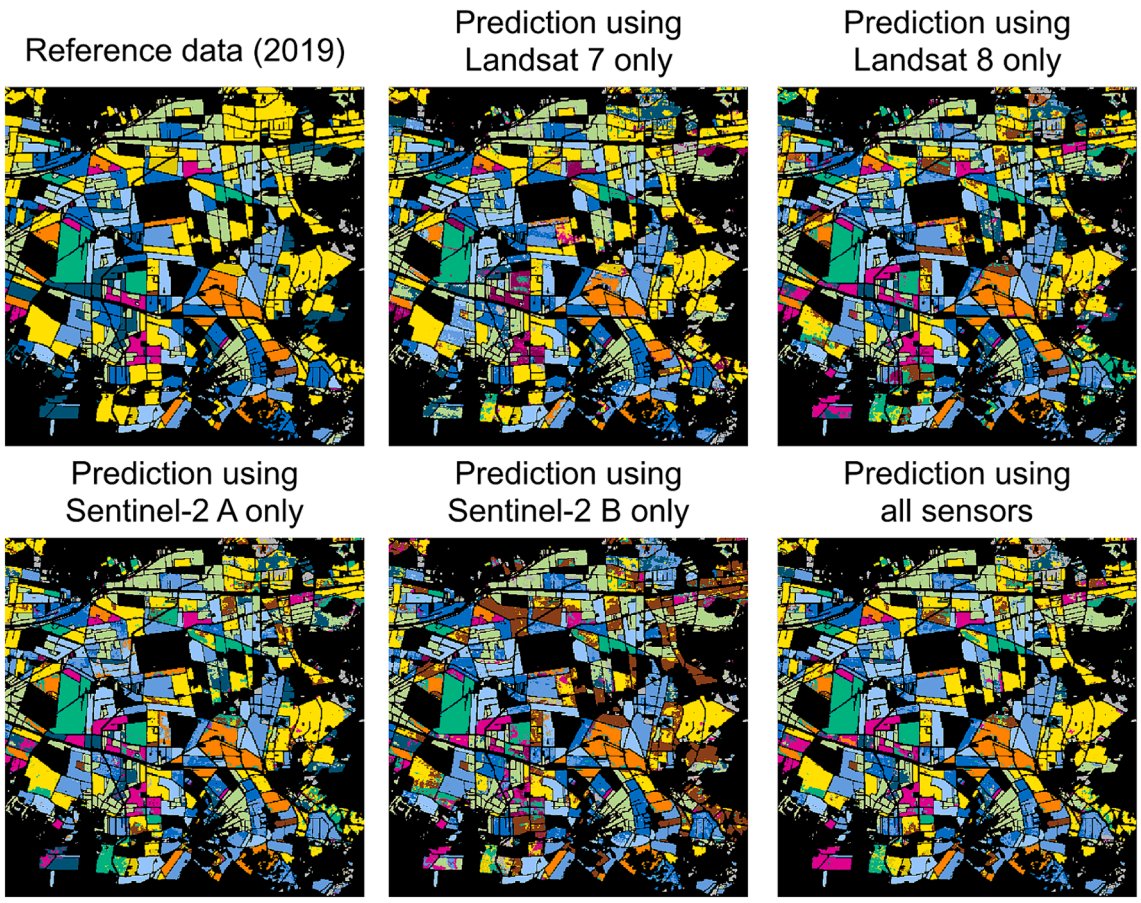


Fig. 10. Visual assessment of the *Baseline model* when mapping an example area in 2019 using data from single satellites (Landsat 7/8 and Sentinel-2 A/B) and all available satellite data.

In years with multiple sensors available, utilizing all satellite data for predictions resulted in higher OA than using data from a single satellite. Although Sentinel-2 A/B data offer a denser annual time-series (with a nominal 5-day revisit) compared to Landsat data (with a nominal 16-day revisit), the model's performance was not significantly impacted by excluding observations from Sentinel-2 A/B. For example, in 2018, the model achieved 81.5 % and 89 % OAs, respectively for Landsat 7 and Landsat 8. Those two OAs were similar to results achieved from Sentinel-2 A/B and only 9 % and 2 % lower than using all available satellite data (91 %). When comparing Landsat 8 and Landsat 7 data, the model tended to produce higher OA with Landsat 8. Except for 2014 and 2020, where there were few CSO from Landsat 8 in the test area, which resulted in a lower performance of (60.9 % and 71.5 %) compared to the slightly higher OAs (76.3 % and 73.3 %) produced with Landsat 7. In the first two years of the study period (2010 and 2011), maps based on Landsat 7 data were comparable to the results using combined Landsat 5 and 7 data. Mapping with Landsat 5 data only produced 42.2 % OA in 2010 and 53.9 % OA in 2011.

Visual assessment (Fig. 10) showed the consistency of mapping results produced by the *Baseline model* regarding the different input data sources. Overall, the results were comparable to the reference data, with accurate mapping of major crop fields. Mapping with only Landsat 7 data resulted in noisier maps. Compared with the reference data, the model based on only Landsat 7 data misclassified *Potato* as *Maize* and *Fodder crops* as *Grassland*. Mapping respectively with Landsat 8, Sentinel-2 A and B showed relatively similar results. Using all available satellite data resulted in the best result, yet the visual impression of classified maps was similar.

4. Discussion

In this study, we addressed two major challenges of mapping annual crop types with multi-temporal EO data over extended periods based on multiple sensor constellations: 1) creating a generalized method for aggregating optical data to capture irregular time-series information and 2) temporally transferring a classification model trained in a single year. We introduced the Temporal Encoding (TE) method, which converts annual time-series data into equidistant features. This is achieved by positioning clear observations on their respective acquisition dates while encoding the no-data and non-clear observations as zeroes. The TE data structure is easily integrated with two augmentation techniques: Random Observation Selection (ROS) and Random Day Shifting (RDS). The combination of data encoding and data augmentations showed great potential for improving the model transfer.

4.1. Temporal transferability improvements

In our experiment, all models were trained with single-year reference data (2018). We showed that when training data and target (mapping) data come from the same year, the *Baseline model* produced accuracies that were in line with recent studies that mapped similar crop types in the same regions (Griffiths et al., 2019; Blickensdörfer et al., 2022; Orynbaikyzy et al., 2022) and similar to the *LI- model.* This suggests that when the training and mapping data share similar characteristics, e.g., temporal density and distribution as well as crop phenology, the choice of temporal aggregation methods becomes less relevant.

When comparing the temporal transferability of the *Baseline model* and *LI model*, as expected, both models showed decreases in overall accuracies (OAs) when compared to the 2018 results. The level of

decrease was closely related to the number of annual CSOs. In the four transfer years where Sentinel-2 data were available (2017, 2019, 2020, 2021), the *Baseline model* outperformed the *LI model* by around 3 % – 10 % in absolute OAs. The earlier years' improvements are highlighted when mapping with coarser annual CSO data (prior to 2017, when only Landsat data was available). In some years (e.g., 2010, 2012, 2013, 2014) the *Baseline model* showed more than 20 % absolute improvements compared to the *LI model*. This could be because interpolation methods tend to suffer from long gaps between clear-sky observations (Roy and Yan, 2020). Thus, interpolating sparse temporal data with LI often results in a high number of unrealistic features. On the other hand, the data structure of TE captured the annual time-series information without compressing or extrapolating the original information. This was achieved by simply positioning the clear observation in the TE features and applying the values of 0 for non-existent or cloudy observations.

A potential concern of TE may be raised due to its "bulky" data structure. In years with high numbers of CSO (e.g., on average, pixels have more than 80 CSO in 2018), the daily encoded pixels still contain more than 75 % features with no-data values (zero values). This might impact the CNN-based models as all input features are participating in the feed-forward processing. On the other hand, the Transformer network with its masking mechanisms can ignore the no-data values from the input features (Zhang et al., 2024). Hence, with the Transformer model, applying dropout and masking to the input data could be equivalent to using ROS with CNN-based models. We also showed that the Transformer model outperformed the 1D-CNN model (Fig. 6a), but only when no augmentations were applied. Using ROS and RDS benefited both deep learning models and 1D-CNN nearly matched the Transformer's performance while being 3 time faster in mapping (Fig. 6a). Nevertheless, the potential of the Transformer network to handle irregular time-series data has been shown in recent studies (Rußwurm and Körner, 2020; Sainte Fare Garnot et al., 2022; Zhang et al., 2024). Hence, optimizing the network architecture with our proposed methods could greatly improve both mapping accuracies and mapping speed. In addition, the 1D-CNN and Transformer networks in our study only focused on pixel-based mapping, while other 2D and 3D CNN networks have been shown to benefit from the spatial context for improving the land cover classification (Cui et al., 2022; Hong et al., 2022; Ji et al., 2023). Nevertheless, such networks require more sophisticated training data (patches instead of pixels), which the impacts of TE data structure with augmentation methods ROS and RDS should be further explored.

4.2. The impact of ROS and RDS

We showed that TE is only better than LI in terms of transferability under the condition of utilizing augmentation methods (ROS and RDS). Hence, our proposed TE method is not as effective when training a model without the augmentations. In fact, the *TE-only model* is highly overfitted without the use of augmentation methods, i.e., only performed well in the training year and poorly in unseen years (Fig. 7). Here, the transferability of *TE-only model's* performance without ROS and RDS was severely weakened in relation to the sparsity of annual CSO. However, the advantage of TE is that its data structure allows a very flexible modification of the original data, which enables the possible simulation of inter-annual temporal variations or differences in crop phenology. Hence, augmentation methods like ROS and RDS added useful variations during the classification model training. This way, overfitting was avoided, and the model generalized better. As a result, the Baseline model was highly tolerant to the changes in annual temporal density.

When training and testing data are from the same year, implementing ROS, RDS, or both only yielded marginal OA improvements of around 5 % compared to *TE-only model's* performances (Fig. 7). This degree of improvement was expected, as such differences also exist spatially, due to differences in orbits and clouds, and it was already demonstrated in the previous studies (Metzger et al., 2021; Nyborg et al., 2022; Sainte Fare Garnot et al., 2022). However, our study revealed that the augmentation methods were highly effective in terms of improving temporal transferability, when the time-series observations of training and mapping data differed. We also showed that each individual augmentation method constantly improved accuracies during transfer, yet to varying extents (Fig. 7). Using only ROS for training is better than using only RDS in most years. The most striking difference can be observed in 2012, where the ROS model outperformed the RDS model greatly with nearly 30 % higher OA. Here, the simulation of information sparsity in the training data allowed the ROS model to perform exceptionally well in 2012, probably by better simulating the uneven distribution between different months that was discussed earlier. In principle, the impact of low annual CSO counts being simulated by the ROS augmentation is higher than that of RDS simulating shifts in phenologies between years.

This also raises another aspect of ROS, which is that it helps the model to learn from the dense temporal data to map in coarser temporal situations. However, in the reverse situation, i.e., when the model is trained with data from coarse years, we expect the impact of ROS on improving transferability to be less effective or maybe even negative. Thus, the dense CSO of the training year also plays an important role in utilizing the ROS augmentation. Nevertheless, for most studies covering long time series, both data density and reliability of reference information are better in very recent years and going back in time implies going back to lower densities.

The performance of the RDS model in this study could also be partly limited due to the current parameters. We used RDS to simulate the crop phenology in training data by randomly shifting individual observation dates back and forth (within 16 days). Despite the consistently positive impact, however, shifting individual observation might potentially lead to illogical sequences and increase confusion or ambiguity. An alternative is to randomly shift the entire set of observations within TE-encoded pixel values by a fixed number of days. This appears more meaningful regarding crop phenological growth. Hence, future sensitivity studies on the maximum number of days during RDS and on the shifting rule could be useful to further improve the impact of this augmentation method.

4.3. Consistent mappings with missing satellite information

The *Baseline model* mapped crop types at relatively consistent accuracies despite the artificial degradations of intra-annual temporal densities, i.e., mapping with data from single satellites (Fig. 9, Fig. 10). For example, we showed that crop predictions based only on Landsat 8 data can be nearly as good as the combination of Landsat 7/8 and Sentinel-2 A/B data (3 % – 10 % OA decrease compared to all available satellite data in 2017 – 2021). This finding demonstrates the possibility of using only one pre-trained classification model to map crop types in different periods or regions where the coverage of all available satellite information could not be met.

For Landsat 7 data, crop mapping with solely this sensor often results in visual artifacts in the classification due to the failure of the Scan Line Corrector (SLC) (Tatsumi et al., 2015). Our approach of using TE with ROS and RDS reduced the line-wise artifacts of Landsat 7 maps almost entirely (Fig. 10, more examples in Supplement 7) and achieved OAs higher than 60 % for all years (Fig. 9). Moreover, the exclusion of Landsat 7 data might be considered in the future due to the drifting orbit of this satellite that has been observed in recent study (Qiu et al., 2021). In this regard, based on the ability of the readily trained model to effectively ingest Landsat-5 data, which was not existent for the training year, we expect the model to work with data from the newly launched Landsat 9 (Masek et al., 2020).

Mapping the past accurately remains challenging, however. Maps based on Landsat 5 data- frequently revealed lower OA. Notably, in 2010 and 2011, Landsat 5 was near the end of its mission and witnessing some failing components (Loveland and Dwyer, 2012) and orbit drifts (Roy et al., 2020). The combination of Landsat 5 observations with SLC-

affected Landsat will have caused poor performance of satellite data mapping in 2010 and 2011. Nevertheless, we expect reasonable results also for earlier years (2000 – 2009), which we excluded here due to the lack of reference data and hence quantitative assessments. The performances prior to the 2000s, however, are expected to be negatively affected by the absence of Landsat 7 data and low density of Landsat 4/5 caused by the lack of a Long-Term Acquisition Plan (LTAP) (Kovalskyy and Roy, 2013). Still, our suggested TE-encoded data with ROS and RDS augmented training data appears to be the proper approach to explore the potential of satellite-based crop mapping until the mid-1980ies.

5. Conclusion

We introduced TE as a simple and universal method for obtaining time-series information that is independent of temporal densities, when annually mapping crop types with optical remote sensing data over longer periods. We demonstrated that incorporating TE with augmentation methods enhanced the temporal transferability especially of a standard DL model (1D-CNN) over time and led to consistently better results. The benefit of data augmentation is mostly compensated when using more complex network architectures. However, for large area mapping (e.g. nation-wide), more simple and robust workflows are preferred. Another aspect to be studied in the future is the utility of the TE with ROS and RDS when classifying more general land cover classes. Separating crop types is mostly related to temporal patterns and not so much about the differences between spectral bands at a single acquisition date. When land cover, including temporally invariant, non-vegetated classes, is analyzed, the differences between spectral wavelength regions might have more relevance. Here the ROS and RDS cannot be expected to directly improve transferability between years. Nevertheless, as phenology is described, the separation of classes with and without phenology, as well as the separation between the different vegetation classes should improve. Our study supports future research aimed at transferring DL models trained with recent data further back in time, as these applications rely heavily on training data, which often is only available in recent years.

CRediT authorship contribution statement

Vu-Dong Pham: Writing – review & editing, Writing – original draft, Visualization, Validation, Methodology, Investigation, Formal analysis, Data curation, Conceptualization. **Gideon Tetteh:** Writing – review & editing. **Fabian Thiel:** Writing – review & editing. **Stefan Erasmi:** Writing – review & editing. **Marcel Schwieder:** Writing – review & editing. **David Frantz:** Writing – review & editing. **Sebastian van der Linden:** Writing – review & editing, Methodology, Funding acquisition, Conceptualization.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

Data will be made available on request.

Acknowledgments

This research was conducted in the frame of the Interdisciplinary Research Center for the Baltic Sea Region Research (IFZO) of University of Greifswald, Germany, and the research project Fragmented Transformations, which is funded by the German Federal Ministry of Education and Research (FKZ 01UC2102). We thank the three anonymous reviewers for their constructive comments and suggestions.

Appendix A. Supplementary material

Supplementary data to this article can be found online at https://doi.org/10.1016/j.jag.2024.103867.

References

Asam, S., Gessner, U., Almengor González, R., Wenzl, M., Kriese, J., Kuenzer, C., 2022. Mapping Crop Types of Germany by Combining Temporal Statistical Metrics of Sentinel-1 and Sentinel-2 Time Series with LPIS Data, Remote Sensing.

Blickensdörfer, L., Schwieder, M., Pflugmacher, D., Nendel, C., Erasmi, S., Hostert, P., 2022. Mapping of crop types and crop sequences with combined time series of Sentinel-1, Sentinel-2 and Landsat 8 data for Germany. Remote Sens. Environ. 269, 112831.

Capliez, E., Ienco, D., Gaetano, R., Baghdadi, N., Salah, A.H., 2023. Temporal-domain adaptation for satellite image time-series land-cover mapping with adversarial learning and spatially aware self-training. IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens. 16, 3645–3675.

Cui, Y., Xia, J., Wang, Z., Gao, S., Wang, L., 2022. Lightweight spectral-spatial attention network for hyperspectral image classification. IEEE Trans. Geosci. Remote Sens. 60, 1–14.

Debella-Gilo, M., Gjertsen, A.K., 2021. Mapping Seasonal Agricultural Land Use Types Using Deep Learning on Sentinel-2 Image Time Series, Remote Sensing.

Frantz, D., 2019. FORCE—Landsat + Sentinel-2 Analysis Ready Data and Beyond. Remote Sens. (Basel) 11.

Frantz, D., Haß, E., Uhl, A., Stoffels, J., Hill, J., 2018. Improvement of the Fmask algorithm for Sentinel-2 images: Separating clouds from bright surfaces based on parallax effects. Remote Sens. Environ. 215, 471–481.

Sainte Fare Garnot, V., Landrieu, L., Chehata, N., 2022. Multi-modal temporal attention models for crop mapping from satellite time series. ISPRS Journal of Photogrammetry and Remote Sensing 187, 294-305.

Garnot, V.S.F., Landrieu, L., Giordano, S., Chehata, N., 2020. Satellite image time series classification with pixel-set encoders and temporal self-attention. In: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, pp. 12325–12334.

Gómez, C., White, J.C., Wulder, M.A., 2016. Optical remotely sensed time series data for land cover classification: a review. ISPRS J. Photogramm. Remote Sens. 116, 55–72.

Griffiths, P., Nendel, C., Hostert, P., 2019. Intra-annual reflectance composites from Sentinel-2 and Landsat for national-scale crop and land cover mapping. Remote Sens. Environ. 220, 135–151.

He, Y., Dong, J., Liao, X., Sun, L., Wang, Z., You, N., Li, Z., Fu, P., 2021. Examining rice distribution and cropping intensity in a mixed single- and double-cropping region in South China using all available Sentinel 1/2 images. Int. J. Appl. Earth Obs. Geoinf. 101, 102351.

Hong, D., Han, Z., Yao, J., Gao, L., Zhang, B., Plaza, A., Chanussot, J., 2022. SpectralFormer: rethinking hyperspectral image classification with transformers. IEEE Trans. Geosci. Remote Sens. 60, 1–15.

Huete, A.R., 1988. A soil-adjusted vegetation index (SAVI). Remote Sens. Environ. 25, 295–309.

Inglada, J., Vincent, A., Arias, M., Tardy, B., Morin, D., Rodes, I., 2017. Operational High Resolution Land Cover Map Production at the Country Scale Using Satellite Image Time Series, Remote Sensing.

Ioffe, S., Szegedy, C., 2015. Batch normalization: Accelerating deep network training by reducing internal covariate shift. Int. Conf. Mach. Learn. Pmlr 448–456.

Ji, R., Tan, K., Wang, X., Pan, C., Xin, L., 2023. PASSNet: A spatial-spectral feature extraction network with patch attention module for hyperspectral image classification. IEEE Geosci. Remote Sens. Lett. 20, 1–5.

Kamilaris, A., Prenafeta-Boldú, F.X., 2018. Deep learning in agriculture: A survey. Comput. Electron. Agric. 147, 70–90.

Karthikeyan, L., Chawla, I., Mishra, A.K., 2020. A review of remote sensing applications in agriculture for food security: Crop growth and yield, irrigation, and crop losses. J. Hydrol. 586, 124905.

Kellenberger, B., Tasar, O., Bhushan Damodaran, B., Courty, N., Tuia, D., 2021. Deep domain adaptation in earth observation. Deep Learning for the Earth Sciences 90–104.

Kovalskyy, V., Roy, D.P., 2013. The global availability of Landsat 5 TM and Landsat 7 ETM+ land surface observations and implications for global 30m Landsat data product generation. Remote Sens. Environ. 130, 280–293.

Li, J., Roy, D.P., 2017. A Global Analysis of Sentinel-2A, Sentinel-2B and Landsat-8 Data Revisit Intervals and Implications for Terrestrial Monitoring, Remote Sensing.

Liu, W., Liu, K., Sun, W., Yang, G., Ren, K., Meng, X., Peng, J., 2023. Self-supervised feature learning based on spectral masking for hyperspectral image classification. IEEE Trans. Geosci. Remote Sens. 61, 1–15.

Liu, L., Xiao, X., Qin, Y., Wang, J., Xu, X., Hu, Y., Qiao, Z., 2020. Mapping cropping intensity in China using time series Landsat and Sentinel-2 images and Google Earth Engine. Remote Sens. Environ. 239, 111624.

Liu, W., Zhang, H., 2023. Mapping annual 10 m rapeseed extent using multisource data in the Yangtze River Economic Belt of China (2017–2021) on Google Earth Engine. Int. J. Appl. Earth Obs. Geoinf. 117, 103198.

Loveland, T.R., Dwyer, J.L., 2012. Landsat: building a strong future. Remote Sens. Environ. 122, 22–29.

Masek, J.G., Wulder, M.A., Markham, B., McCorkel, J., Crawford, C.J., Storey, J., Jenstrom, D.T., 2020. Landsat 9: Empowering open science and applications through continuity. Remote Sens. Environ. 248, 111968.

McFeeters, S.K., 1996. The use of the normalized difference water index (NDWI) in the delineation of open water features. Int. J. Remote Sens. 17, 1425–1432.
Metzger, N., Turkoglu, M.O., D'Aronco, S., Wegner, J.D., Schindler, K., 2021. Crop classification under varying cloud cover with neural ordinary differential equations. IEEE Trans. Geosci. Remote Sens. 60, 1–12.
Nyborg, J., Pelletier, C., Lefèvre, S., Assent, I., 2022. TimeMatch: Unsupervised cross-region adaptation by temporal shift estimation. ISPRS J. Photogramm. Remote Sens. 188, 301–313.
Oliphant, A.J., Thenkabail, P.S., Teluguntla, P., Xiong, J., Gumma, M.K., Congalton, R.G., Yadav, K., 2019. Mapping cropland extent of Southeast and Northeast Asia using multi-year time-series Landsat 30-m data using a random forest classifier on the Google Earth Engine Cloud. Int. J. Appl. Earth Obs. Geoinf. 81, 110–124.
Orynbaikyzy, A., Gessner, U., Conrad, C., 2022. Spatial Transferability of Random Forest Models for Crop Type Classification Using Sentinel-1 and Sentinel-2, Remote Sensing.
Pan, L., Xia, H., Yang, J., Niu, W., Wang, R., Song, H., Guo, Y., Qin, Y., 2021. Mapping cropping intensity in Huaihe basin using phenology algorithm, all Sentinel-2 and Landsat images in Google Earth Engine. Int. J. Appl. Earth Obs. Geoinf. 102, 102376.
Piedelobo, L., Hernández-López, D., Ballesteros, R., Chakhar, A., Del Pozo, S., González-Aguilera, D., Moreno, M.A., 2019. Scalable pixel-based crop classification combining Sentinel-2 and Landsat-8 data time series: Case study of the Duero river basin. Agr. Syst. 171, 36–50.
Qiu, B., Zhong, M., Tang, Z., Wang, C., 2014. A new methodology to map double-cropping croplands based on continuous wavelet transform. Int. J. Appl. Earth Obs. Geoinf. 26, 97–104.
Qiu, S., Zhu, Z., Shang, R., Crawford, C.J., 2021. Can Landsat 7 preserve its science capability with a drifting orbit? Science of Remote Sensing 4, 100026.
Qiu, S., Zhu, Z., Olofsson, P., Woodcock, C.E., Jin, S., 2023. Evaluation of Landsat image compositing algorithms. Remote Sens. Environ. 285, 113375.
Roy, D.P., Li, Z., Zhang, H.K., Huang, H., 2020. A conterminous United States analysis of the impact of Landsat 5 orbit drift on the temporal consistency of Landsat 5 Thematic Mapper data. Remote Sens. Environ. 240, 111701.
Roy, D.P., Yan, L., 2020. Robust Landsat-based crop time series modelling. Remote Sens. Environ. 238, 110810.
Rußwurm, M., Courty, N., Emonet, R., Lefèvre, S., Tuia, D., Tavenard, R., 2023. End-to-end learned early classification of time series for in-season crop type mapping. ISPRS J. Photogramm. Remote Sens. 196, 445–456.
Rußwurm, M., Körner, M., 2020. Self-attention for raw optical satellite time series classification. ISPRS J. Photogramm. Remote Sens. 169, 421–435.
Sadeh, Y., Zhu, X., Dunkerley, D., Walker, J.P., Zhang, Y., Rozenstein, O., Manivasagam, V.S., Chenu, K., 2021. Fusion of Sentinel-2 and PlanetScope time-series data into daily 3 m surface reflectance and wheat LAI monitoring. Int. J. Appl. Earth Obs. Geoinf. 96, 102260.
Tatsumi, K., Yamashiki, Y., Canales Torres, M.A., Taipe, C.L.R., 2015. Crop classification of upland fields using Random forest of time-series Landsat 7 ETM+ data. Comput. Electron. Agric. 115, 171–179.
Troegel, T., Schulz, C., 2018. Ergebnisse der Agrarstrukturerhebung 2016 für das Land Brandenburg. Zeitschrift Für Amtliche Statistik Berlin Brandenburg 1, 44–60.
Tucker, C.J., 1979. Red and photographic infrared linear combinations for monitoring vegetation. Remote Sens. Environ. 8, 127–150.
Tuia, D., Persello, C., Bruzzone, L., 2016. Domain adaptation for the classification of remote sensing data: an overview of recent advances. IEEE Geosci. Remote Sens. Mag. 4, 41–57.
Valero, S., Pelletier, C., Bertolino, M., 2016. Patch-based reconstruction of high resolution satellite image time series with missing values using spatial, spectral and temporal similarities. In: 2016 IEEE International Geoscience and Remote Sensing Symposium (IGARSS), pp. 2308–2311.
Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A.N., Kaiser, Ł., Polosukhin, I., 2017. Attention is all you need. Advances in neural information processing systems 30.
Vaudour, E., Gomez, C., Lagacherie, P., Loiseau, T., Baghdadi, N., Urbina-Salazar, D., Loubet, B., Arrouays, D., 2021. Temporal mosaicking approaches of Sentinel-2 images for extending topsoil organic carbon content mapping in croplands. Int. J. Appl. Earth Obs. Geoinf. 96, 102277.
Wang, Y., Feng, L., Zhang, Z., Tian, F., 2023b. An unsupervised domain adaptation deep learning method for spatial and temporal transferable crop type mapping using Sentinel-2 imagery. ISPRS J. Photogramm. Remote Sens. 199, 102–117.
Wang, X., Tan, K., Du, Q., Chen, Y., Du, P., 2019. Caps-TripleGAN: GAN-assisted CapsNet for hyperspectral image classification. IEEE Trans. Geosci. Remote Sens. 57, 7232–7245.
Wang, X., Tan, K., Du, P., Han, B., Ding, J., 2023a. A capsule-vectored neural network for hyperspectral image classification. Knowl.-Based Syst. 268, 110482.
Weiss, M., Jacob, F., Duveiller, G., 2020. Remote sensing for agricultural applications: a meta-review. Remote Sens. Environ. 236, 111402.
Xu, J., Zhu, Y., Zhong, R., Lin, Z., Xu, J., Jiang, H., Huang, J., Li, H., Lin, T., 2020. DeepCropMapping: A multi-temporal deep learning approach with improved spatial generalizability for dynamic corn and soybean mapping. Remote Sens. Environ. 247, 111946.
Xuan, F., Dong, Y., Li, J., Li, X., Su, W., Huang, X., Huang, J., Xie, Z., Li, Z., Liu, H., Tao, W., Wen, Y., Zhang, Y., 2023. Mapping crop type in Northeast China during 2013–2021 using automatic sampling and tile-based image classification. Int. J. Appl. Earth Obs. Geoinf. 117, 103178.
Yan, S., Yao, X., Zhu, D., Liu, D., Zhang, L., Yu, G., Gao, B., Yang, J., Yun, W., 2021. Large-scale crop mapping from multi-source optical satellite imageries using machine learning with discrete grids. Int. J. Appl. Earth Obs. Geoinf. 103, 102485.
Zeng, L., Wardlow, B.D., Xiang, D., Hu, S., Li, D., 2020. A review of vegetation phenological metrics extraction using time-series, multispectral satellite data. Remote Sens. Environ. 237, 111511.
Zhang, H.K., Roy, D.P., Luo, D., 2023. Demonstration of large area land cover classification with a one dimensional convolutional neural network applied to single pixel temporal metric percentiles. Remote Sens. Environ. 295, 113653.
Zhang, H.K., Luo, D., Li, Z., 2024. Classifying raw irregular time series (CRIT) for large area land cover mapping by adapting transformer model. Science of Remote Sensing 9, 100123.
Zhong, L., Hu, L., Zhou, H., 2019. Deep learning based multi-temporal crop classification. Remote Sens. Environ. 221, 430–443.
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.