Grassland yield estimations – potentials and limitations of remote sensing in comparison to process-based modelling and field measurements

Sophie Reinermann[1,2], Carolin Boos[3], Andrea Kaim[4,5], Anne Schucknecht[6], Sarah Asam[2], Ursula Gessner[2], Sylvia H. Annuth[4], Thomas M. Schmitt[3,4], Thomas Koellner[4], Ralf Kiese[3]

[1]University of Würzburg, Institute of Geography and Geology, Department of Remote Sensing, Am Hubland, 97074 Würzburg, Germany
[2]German Aerospace Center (DLR), German Remote Sensing Data Center (DFD), Oberpfaffenhofen, Wessling 82234, Germany
[3]Karlsruhe Institute of Technology (KIT), Institute of Meteorology and Climate Research, Atmospheric Environmental Research (IMKIFU), Garmisch-Partenkirchen, Germany
[4]University of Bayreuth, Professorship of Ecological Services, Bayreuth Center of Ecology and Environmental Research (BayCEER), Universitaetsstr. 30, 95447 Bayreuth, Germany
[5]Helmholtz Centre for Environmental Research - UFZ, Department of Computational Landscape Ecology, Leipzig, Germany
[6]OHB System AG, EO Applications Team, Manfred-Fuchs-Str. 1, 82234 Wessling, Germany

Correspondence to: Sophie Reinermann (sophie.reinermann@dlr.de)

Abstract. Grasslands make up the majority of agricultural land and provide fodder for livestock. Information on grassland yield is very limited as the fodder is directly used at the farms. Data on grassland yields would be needed, however, to inform politics and stakeholders on grassland ecosystem services and inter-annual variations. Grassland yield patterns are often varying on small scales in Germany and estimations are further complicated by missing information on grassland management. Here, we compare three different approaches to estimate annual grassland yield for a study region in southern Germany. We apply (i) a novel approach based on a model derived from field samples, satellite data and mowing information (RS), (ii) the biogeochemical process-based model LandscapeDNDC (LDNDC) and (iii) a rule-set approach based on field measurements and spatial information on grassland productivity (RVA) to derive grassland yields per parcel for the Ammer catchment area in 2019. All three approaches reach plausible results of annual yields of around 4-9 t/ha and show overlapping as well as diverging spatial patterns. For example, direct comparisons show that higher yields were derived with LDNDC compared to RS and RVA, in particular related to the first cut and for grasslands mown only one or two times per year. The mowing frequency was found to be the most important influencing factor for grassland yields of all three approaches. There were no significant differences found in the effect of abiotic influencing factors, such as climate or elevation, on grassland yields derived from the different approaches. The potentials and limitations of the three approaches are analysed and discussed in depth, such as the level of detail of required input data, or the capability of regional and inter-annual yield estimations. For the first time, three different approaches to estimate grassland yields were compared in depth resulting in new insights in their

35 potentials and limitations. Grassland productivity maps provide the basis for long-term analyses of climate and management impacts and comprehensive studies of the functions of grassland ecosystems.

1 Introduction

Grassland ecosystems provide fodder for livestock, apart from many other ecosystem services, such as carbon storage,
40 provision of habitats, water purification, recreation and erosion control (Bengtsson et al., 2019; Le Clec'h et al., 2019; Gibon, 2005; Gibson, 2009; Richter et al., 2021; White et al., 2000). In Germany, grasslands cover almost one third of the agriculturally used area (Statistisches Bundesamt, 2023) and are of central importance for the meat and dairy industry (Schoof et al., 2020b; Soussana and Lüscher, 2007). In large parts of Europe, grassland ecosystems are managed, hence strongly shaped by human activities. In Germany, for example, almost all of the grassland is under some form of agricultural use; i.e. grazed
45 and/or mown in different frequency (Dengler et al., 2014; Schoof et al., 2020a, c). Grassland management and use intensity, i.e. the number and timing of grazing and/or mowing as well as fertilization events, have a strong impact on grassland functions and ecology (Gossner et al., 2016; Neyret et al., 2021; Socher et al., 2012). Apart from climate and soil conditions, grassland management determines the productivity, thus yields, and species diversity of these ecosystems (Gilhaus et al., 2017). In Germany, grasslands are managed on small units (parcels) individually, resulting in a wide variety of combinations of the
50 number and timing of mowing events on small spatial scales. As a consequence, grassland landscapes can show high spatial and temporal variability in their biomass availability and species composition (Gerowitt et al., 2013).

Grassland biomass is usually directly used on farms as fodder for livestock and not traded, which is why there is usually no data on grassland yields resulting from sales statistics. The lack of information on yields exacerbates extensive spatio-temporal
55 analyses of drivers of grassland productivity, as well as modelling of grassland ecosystem services, e.g. nitrogen and carbon fluxes. Long-term effects of climate change as well as short-term weather extremes influence grassland productivity and yields (Beniston, 2003; Berauer et al., 2019). In the Alpine and pre-Alpine regions of southern Germany this is of particular importance, since temperature increases twice as fast as the global average (Auer et al., 2007; Kiese et al., 2018). In addition, drought and heat episodes are expected to increase in the region. Therefore, information on grassland yields and the
60 dependency on climate conditions is needed to support the planning of fodder production and imports for farmers and to inform administration and politics. Furthermore, information on grassland yields is required for a comprehensive assessment of grassland ecosystem services and sustainable management also under changing climate conditions. Despite these information needs, continuous and large-scale monitoring is lacking.

65 There are different approaches to retrieve grassland yield information. Ground-measurements alone such as cutting and removing herbage from the grassland for direct analysis or estimating yields by the use of a rising plate meter are usually time

intensive, can hardly provide regular information and might not represent the conditions on broader spatial scales (Murphy et al., 2021). This holds in particular in grassland ecosystems characterized by a high small-scale variability, like in southern Germany. To retrieve spatially continuous and multi-temporal information, grassland yields can be (i) modelled empirically

70   in different degree of complexity, e.g. taking in-situ and remote sensing data into account, (ii) modelled bio-geochemically, e.g. with process-based models, or (iii) derived from simple rule sets used by authorities based on yield surveys and further spatially extensive data, e.g. elevation and soil fertility index.

Remotely sensed reflectance, and in particular vegetation indices derived from them, depict vegetation greenness, structure

75   and photosynthetic activity and, thus, relate to vegetation biomass (Holtgrave et al., 2020; Huete et al., 2002). Grassland traits, such as above-ground biomass, can be estimated using an empirical relationship employing remote sensing and in-situ data to train and validate models, as shown in many studies summarized in Reinermann et. al. (2020). Space-borne remote sensing-based biomass models have been applied in many different grassland ecosystems using various sensors and regression models. Using satellite remote sensing data to quantify vegetation properties enables large-scale, continuous, reproducible and

80   comparatively cost-sensitive monitoring. Compared with the relatively frequent application of empirical remote sensing data-based biomass models for mostly grazed grassland ecosystems (Wu et al., 2024; Yao and Ren, 2024), the number of studies using this approach for grasslands dominated by mowing is more limited (Reinermann et al., 2020). Previous studies from regions characterized by mown grasslands investigated the potential of various vegetation indices derived from medium resolution sensors (moderate resolution imaging spectroradiometer (MODIS), moderate resolution imaging spectrometer

85   (MERIS)) to estimate grassland biomass for single sites in Ireland and the Netherlands (Ali et al., 2017a; Ullah et al., 2012). Based on Landsat and Sentinel-2, grassland biomass and height were estimated for study regions in Germany, France, Spain and Austria using various regressors, such as multi-linear regression, random forest or deep learning models (Barrachina et al., 2015; Dusseux et al., 2022; Eder et al., 2023; Muro et al., 2022; Schwieder et al., 2020). However, despite the strong influence of grassland management on the productivity, to our knowledge none of the previous remote sensing-based studies have

90   directly included mowing information in the biomass estimation approach. Further, grassland biomass estimates are only a snapshot in time. In particular for grasslands dominated by frequent mowing activities, the amount of standing biomass varies a lot in the course of a year. A single biomass estimation is therefore not sufficient to inform on annual grassland productivity and yields. One way to approach this is the combination of multi-temporal biomass estimations informed by timing of mowing events to retrieve annual grassland yields. To our knowledge, there is no remote sensing-based study that estimated annual

95   grassland yields using this approach so far.

Another method to obtain grassland yield estimates are deterministic process-based models, such as LandscapeDNDC (Haas et al., 2013; Kraus et al., 2015; Petersen et al., 2021), Daycent (Del Grosso and Parton, 2019; Parton et al., 1998), PaSIM (Riedo et al., 1998, 2000), LPJmL (Bondeau et al., 2007; Schaphoff et al., 2018), APSIM (Holzworth et al., 2014), or

100   ORCHIDEE-GM (Chang et al., 2013). An advantage of process-based models is the possibility to assess all spatial levels

ranging from the site (Chang et al., 2013; Liebermann et al., 2020; Petersen et al., 2021) to continental (Vuichard et al., 2007) and global (Rolinski et al., 2018) scale. Additionally, the application of process-based models opens the possibility to evaluate ecosystem productivity under various scenarios including climate change (Petersen et al., 2021), or management changes like adaptions in fertilization regimes (Hong et al., 2023; Reis Martins et al., 2024), or shifts in cutting frequencies (Rolinski et al.,

105   2018). However, as process-based models rely on the availability of data for model development, testing, and, in particular, up-scaling large-scale applications are limited.

A third approach to estimate grassland yield is by making use of measurements from field experiments or regional census statistics (Smit et al., 2008). In Germany, some federal states provide reference values for grassland yields at county level that

110   can be used by farmers to derive their grassland's fertilizer requirements. For instance, the reference values provided in the guideline for fertilization of crop- and grassland by the Bavarian State Institute for Agriculture (LfL) (LfL, 2018) are aggregated values for Bavaria based on LfL internal research and field experiments (Diepolder et al., 2016).

Here, a novel remote sensing-based approach to derive grassland yield in southern Germany is presented and applied. Annual

115   grassland yields are compared to yield estimates of the same year and region derived from two other approaches – either established in the scientific community or used by authorities, however optimized for the study region. Annual yields are compared and advantages and disadvantages are highlighted. We estimate annual grassland yields for a study area in southern Germany in 2019 using (a) a novel empirical satellite remote sensing model (RS), which is compared with results of (b) a process-based biogeochemical model (LandscapeDNDC) and with (c) a simple rule-set reference value approach used by

120   authorities (RVA). To represent grassland management intensity, all three approaches use information on satellite-derived mowing dates retrieved from Reinermann et al. (2023) and (2022). To examine under which conditions the results of the three different approaches differ, we examined the influence of various factors – management and climate – on the grassland yields resulting from the three methods. By examining the spatial and temporal patterns of yields derived from the three approaches and analyzing the influence of various factors, we assess the differences and similarities in the methods. We aim to determine

125   which method best represents specific conditions and under which circumstances it is most reliable.

2 Study area

The study area is located in the pre-Alpine and Alpine region of southern Germany (Figure 1) and consists of the broader Ammer catchment area including the Pre-Alpine TERENO Observatory (Kiese et al., 2018). The area belongs to the temperate oceanic climate according to Köppen and Geiger (Kottek et al., 2006). The mean annual temperature was 8.9 °C in 2019, the

130   long-term average (2012-2021) is 8.1 °C for the region. The mean precipitation sum was 1175 mm in 2019 and the long-term average 1141 mm (Boos et al., 2024; Petersen et al., 2021). The elevation of the study area with grassland land use ranges between 500 and 1100 m a.s.l. with grasslands dominating agricultural land use which totals to about 38% of the region area

(Kiese et al., 2018). In the region, grasslands are of economic importance, in particular for meat and dairy production, but also for tourism (Schmitt et al., 2024; Soussana and Lüscher, 2007). The grasslands of the Ammer region are grazed and/or mown
135   at intensities ranging from extensive (with one to two mowing events) to highly intensive use with up to six mowing events per year (Reinermann et al., 2022, 2023). The timing of the management activities varies from grassland parcel to parcel. Here, we focus on meadows and mowing pastures which make up around 657 km² (27138 parcels).
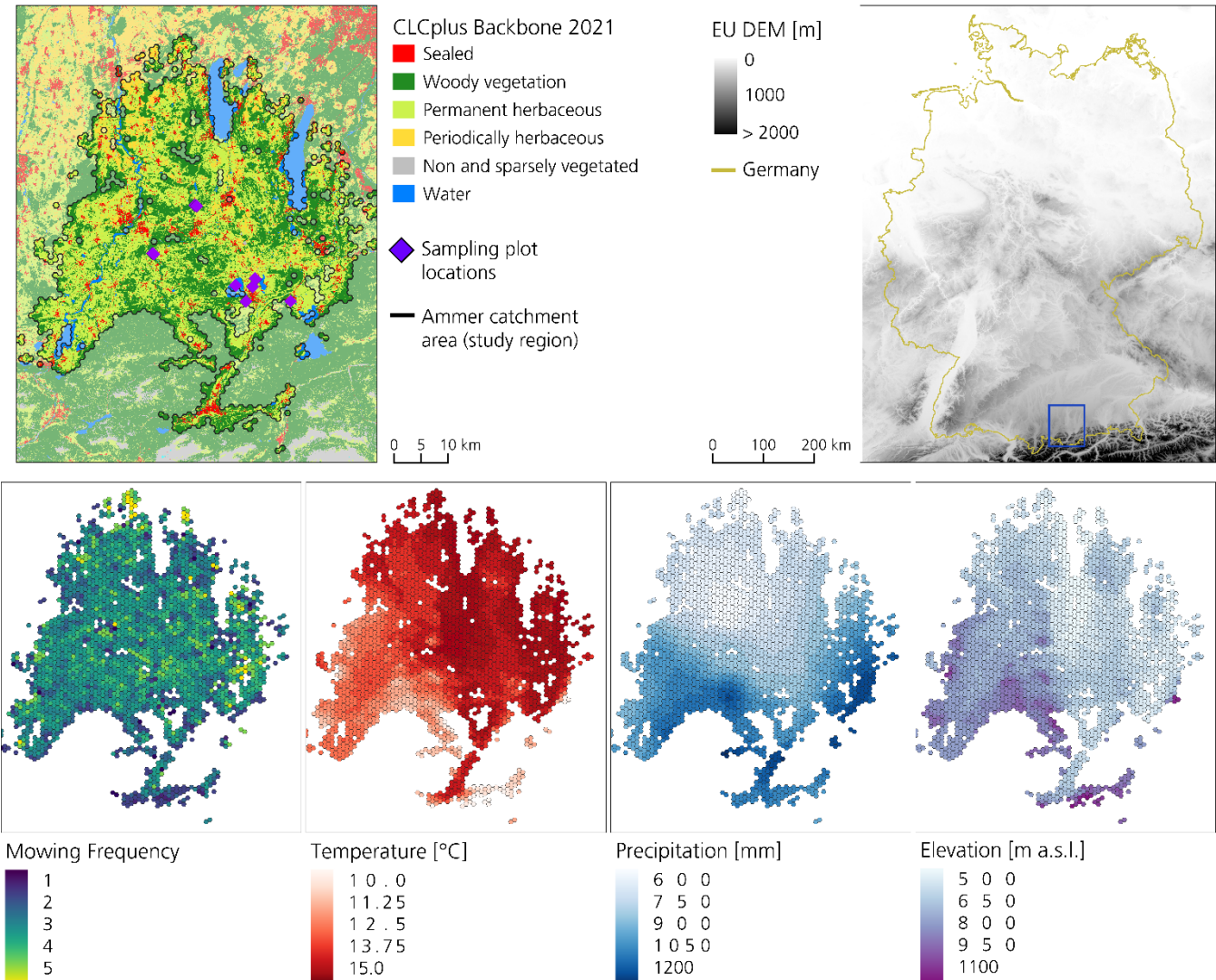
Figure 1: Study area in southern Germany showing land use, elevation and hexagon averages of 2019 of the mowing frequency,
140   mean annual air temperature, mean annual precipitation, and elevation (see Section 3.2.2). The hexagon diagonal size is 1 km.

3 Material and methods

3.1 Spatial and field data

3.1.1 Spatial management data

All three yield modelling approaches used the same parcel boundary and mowing information data. Parcel boundaries were taken from the EU's Integrated Administration and Control System (IACS) provided by the Bavarian State Institute for Agriculture (LfL). The same data source was used to exclude all parcels which were not used as meadows or mowing pastures in 2019.

The dates of mowing events originate from Reinermann et al. (2023) and (2022) where mowing events were detected based on Sentinel-2 time series. However, differing from the original approach, for this study no grassland mask was used to ensure that all meadows and mowing pastures identified by the IACS data are covered. To transfer the 10x10 m² pixel-based mowing dates to parcel level, detected dates within a time frame of three weeks were agglomerated per parcel using the majority vote. Only when at least for 20 % of the parcel the date was detected, the mowing event remained in the dataset. Regional validation using information from farmers and webcam images showed an accuracy (F1-Score) of 0.65 for the mowing dates on parcel level in the Ammer region in 2019. The validation was conducted only with data from the study region but in the same manner as in Reinermann et al. (2022).

3.1.2 Biomass field data

To train and validate the empirical remote sensing model as well as for regional quality evaluation of LandscapeDNDC, in-situ biomass measurements were used. A total of 14 grassland plots of 20 m x 20 m in the Ammer region (Figure 1) were sampled in the period 2019-2021 to obtain in-situ above-ground biomass (AGB) information (Schucknecht et al. (2023), Schucknecht et al. (2020)). The sampling plots are constituted by homogeneous vegetation coverage and placed to be representative for the entire grassland parcel. The sampled grassland parcels are characterized by different land management intensities ranging from one to six mowing events per year. The sampling campaigns took place at multiple times during the growing season to ensure that biomass samples from a variety of growth stages before and after mowing events were included. For each plot, above-ground biomass was collected on four randomly placed subplots of 50 cm x 50 cm. The plot was divided into four equal quadrants, with each subplot being randomly positioned in one of the quadrants. The position of each subplot varied with every measurement. To account for mowing height, AGB was sampled from >7cm on all subplots and complemented by one biomass sampling from 2-7 cm vegetation height at one of the four subplots. The samples were dried for at least 48 hours until constant weight at 60 °C and weighed. The weight of the dried biomass of the four samples from above 7 cm were averaged and added to the measurements of 2-7 cm to obtain total AGB per site and date as well as scaled to 1 m². In total, 111 biomass samples were collected. Apart from three grassland plots which were used as alternatives when the

actual plots were freshly mown at the time of the campaign, all plots were sampled five times in 2019, three times in 2020 and one time in 2021 between March and October. Among the 14 grassland plots, one was mown six times, two five times, four four times, three three times, one two times and three only once per year.

3.2 Yield estimation approaches

The three grassland yield estimation approaches applied in this study, i.e., the remote sensing (RS) approach, the process-based LandscapeDNDC model approach (LDNDC) and an estimation based on reference values (RVA), are described in detail in Sect. 3.2.1-3.2.3. (Figure 2).
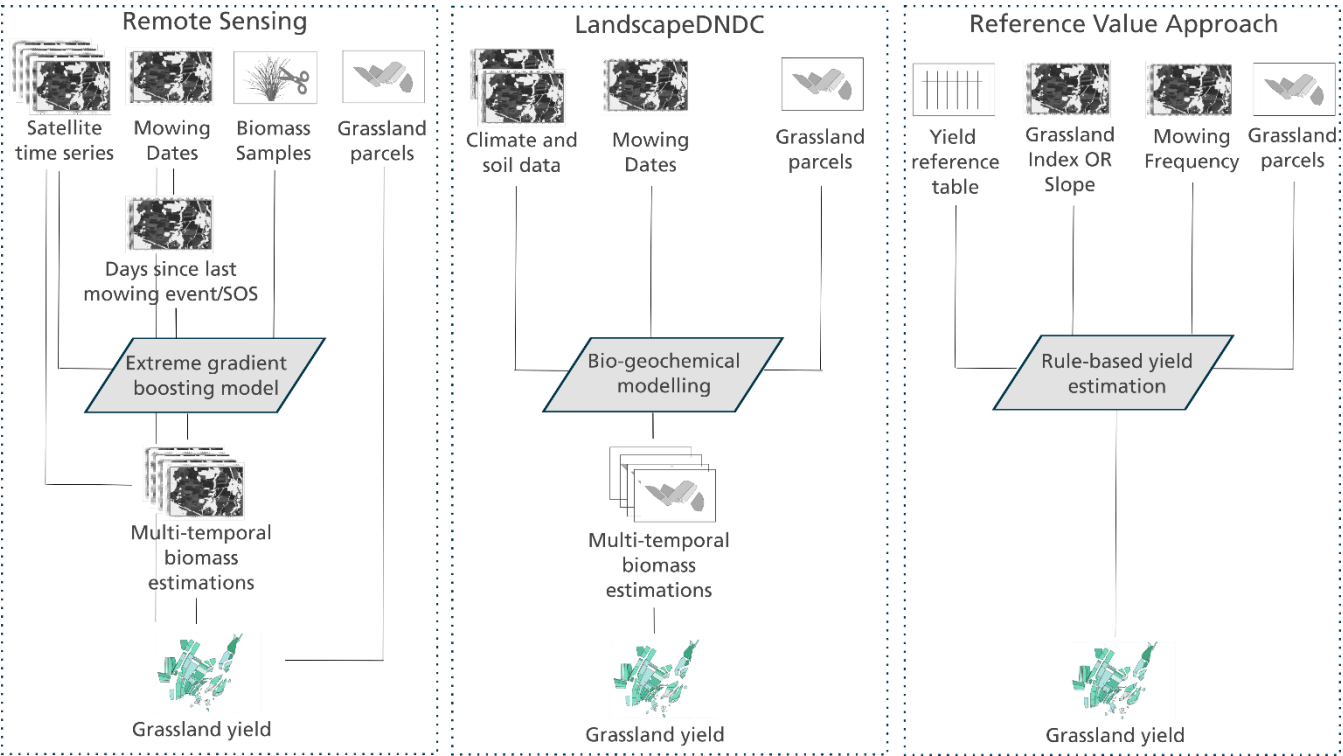
Figure 2: Conceptual scheme of the three yield estimation approaches.

3.2.1 Remote sensing-based approach

For the remote sensing model, MAJA (version 3.3) Sentinel-2 (S2) level 2a (Hagolle et al., 2017) time series from the years 2019-2021 and tiles 32TPT and 32UPU were used to match the biomass sampling campaigns. The optical satellite reflectance data consists of acquisitions from two identical satellites (S2A and S2B) acquiring information in 12 spectral bands (Drusch et al., 2012). The bands used here are the 10 bands relevant for vegetation monitoring, namely bands 2, 3, 4, 5, 6, 7, 8, 8A, 11

and 12, covering the red, green, vegetation red edge, near-infrared and shortwave-infrared wavelengths. The bands which have a 20 m spatial resolution were resampled to 10 m by the nearest-neighbour method to achieve a consistent spatial resolution of 10 m for all bands.

Based on S2 satellite data and in-situ total AGB samples (Section 3.1.2), an empirical model was trained and optimized to estimate grassland biomass. Satellite data influenced by clouds, cloud shadows or unfavourable terrain conditions were excluded according to the MAJA algorithm. The empirical model was built based on the S2 reflectances from the 10 selected bands and additional spectral indices as predictor variables. Specifically, the Enhanced Vegetation Index (Huete et al., 2002) (EVI, Equation 1) and the Tasseled Cap Wetness Index (Indexdatabase, 2024; Krauth and Thomas, 1976) (wetness, Equation 2) were calculated and included as they relate to vegetation biomass:

$$EVI = 2.5 \frac{B8 - B4}{B8 + 6*B4 - 7.5*B2 + 1}, \tag{1}$$

$$Wetness = 0.1509 * B2 + 0.1973 * B3 + 0.3279 * B4 + 0.3406 * B8 - 0.7112 * B11 - 0.4572 * B12, \tag{2}$$

where B2, B3, B4, B8 and B11 are reflectance bands in the blue, green, red, near infrared and short-wave infrared area, respectively.

In addition, information on timing of mowing events (Section 3.1.1) was directly included in the modelling process by adding an additional predictor variable representing the days since the last mowing event. For each S2 acquisition, a layer was calculated giving the days since the last mowing event on pixel basis. When no mowing event took place before the S2 acquisition, the number of days since the start of the growing season was calculated. To retrieve the start of the growing season, the Copernicus Land Monitoring Service (CLMS) High Resolution Vegetation Phenology and Productivity (HR VPP) Start of Season product was used (CLMS, 2019). Further, the S2 acquisition date was included into the mode as predictor variable. This resulted in 14 input features for the empirical modelling, i.e. 10 spectral bands, 2 spectral indices, the days since last mowing/start of growing season, and the date of the satellite acquisition.

To prepare the input data for model training, pairs of cloud-free S2 acquisitions and corresponding in-situ biomass samples were built by allowing a maximum of five days between satellite acquisition and field sampling in both directions. If there were multiple satellite acquisitions in the allowed range, closer ones were preferred as well as satellite acquisitions after field sampling dates. It was also checked that there was no mowing event in between a satellite acquisition and a field sampling to maintain representative data pairs. Due to cloud conditions in 2021, only data from 2019 and 2020 remained in the data table after this procedure. Data pairs from sampling campaigns from every month between April and October, apart from July and August, were available.

An extreme gradient boosting model was trained on the input features and the corresponding AGB values (Friedman, 2001). Initial tests showed that the extreme gradient boosting model outperformed others, such as Random Forest, Support Vector Machines or multi-linear models. The xgboost package (version 1.5.2) was used in Python. In total, 74 data pairs were available

225 from which 82% (n=61) were used for training and testing and 18% (n=13) as independent test of the trained model. A stratified sampling of the test data was conducted to ensure that the value range of the test data was representative. With the data used for training the hyperparameters for an optimized model were searched for, using grid search, 5-fold cross-validation (CV) and ten iterations each. To find the best model the coefficient of determination (R^2) was used. The best model according to the training was then tested against the independent test data set.

230

The best trained biomass model was applied to estimate the AGB of all available S2 scenes to generate a biomass time series. This biomass time series was used in combination with the mowing dates and IACS parcel information to estimate annual yields per parcel. This was approached by going through the parcel-based mowing dates. For each mowing date, the pixel-based biomass estimates from all observations of up to three weeks before and one week after the mowing date were extracted.

235 The 95 % percentile was calculated from this biomass data to estimate the yield per mowing event and parcel, minimizing the influence of parcel boundaries. This time frame was used to ensure that the biomass was captured shortly before a mowing event as there is an uncertainty in the timing of the mowing dates. These single mowing event yields were afterwards summed up to annual yields per parcel.

3.2.2 LandscapeDNDC modelling approach

240 The process-based biogeochemical model LDNDC was run for the whole study region with individual high-quality input data combinations of soil, climate and management for every field (Boos et al., 2024). This became possible by the availability of accurate small-scale grassland soil profile data, interpolated reference climate data based on continuous measurements from weather stations, and cutting dates from remote sensing on a field level.

245 The model calibration and validation for yields was performed on extensive measurements on lysimeters from the TERENO pre-Alpine observatory covering three sites in different elevations within the study area with intensive and extensive management (Kiese et al., 2018; Petersen et al., 2021). Running the model with regional input data for the sites, which are used for RS-training and validation, compare Sect. 3.1.2., and comparing to standing biomass before cutting events, lead to a coefficient of determination of 0.67 and a RMSE of 1.46 t ha^{-1} a^{-1} (Boos et al., 2024).

250

Climate inputs were generated from the reference data of the Climex project (Poschlod et al., 2020; Willkofer et al., 2020) - interpolated from station measurements of the German Weather Service -, which we aggregated to daily climate inputs (minimum, maximum, and mean air temperature, precipitation sum, mean relative humidity, mean global radiation, and mean wind speed) and assigned the nearest virtual climate station to every field to run LDNDC. Regional soil data was derived from

the soil database of the Bavarian State Office of the Environment (LfU) (LfU, 2020). Only mineral grassland soils were considered and a unique, i.e. with a single profile per polygon, soil map was compiled. More details can be found in Boos et al. (2024).

The model simulates plant growth depending on factors like photosynthesis, nitrogen and water availability, phenology, and temperature (Petersen et al., 2021). At a prescribed cutting date, the above ground biomass is reduced to a pre-set value for the remaining biomass, which equals the standing biomass after the cutting event and is according to farmers' practice calibrated to a cutting height of about 7 cm. The harvested biomass from all events in a year is then summed up to calculate annual yields per field. Therefore, the management is another key model driver and was set for every parcel individually. The cutting dates in the study year 2019 were taken from the dataset generated by Reinermann et al. (2022), as described in Sect. 3.1.1. Fertilizer in the form of slurry was applied according to the mowing information following farmers' practice in the study region. For parcels with three or more mowing events per year, the number of manuring events equalled the number of cuts. For parcels with less than three mowing events per year, the number of fertilizer applications was one less than the number of cuts which corresponds to local farmers practices. The amount of manure varied between 40 and 55 kgNha^{-1} per event and decreased per application. For further details on the applied regional model drivers for LDNDC, see Boos et al. (2024).

For every grassland field simulation (N=27138), climate, soil, and management input were derived from superimposing field boundaries with the respective spatial products. The model (LDNDC revision: 10786, Crabmeat revision: 8136) was run with an hourly time step and the submodels CanopyECM (Grote et al., 2009) as the microclimate module, WatercycleDNDC (Kiese et al., 2011) as the watercycle module, MeTrx (Kraus et al., 2015) as the soil-chemistry module, and PlaMox (Kraus et al., 2016; Liebermann et al., 2020) employing the PhotoFarquhar model (Ball et al., 1987; Farquhar et al., 1980) for photosynthesis as the physiology module. For a general description of LDNDC and the functioning and interaction of the different sub-modules see Petersen et al. (2021).

3.2.3 Reference values approach

The reference values approach (RVA) is mainly based on a look-up table from the Bavarian State Institute for Agriculture (LfL) that includes yield reference values of the farmer's yield for Bavaria for (i) different types of grassland uses and intensities (number of cutting events, low/medium/high grazing intensity) as well as yield levels (low, medium, high) (see Table A1). Apart from the mowing information (Section 3.1.1), cattle numbers (Section 3.1.1) were used to identify the management intensity (i). We further used the land appraisal dataset (LDBV, 2018) to obtain grassland indices (German "Grünlandzahl") for each field to get the yield levels (ii). This index represents the quality of a location for grassland production, considering factors such as soil type, soil properties, climate, and water availability. The index ranges from 1 (poor) to 100 (best). As the grassland index can vary within a parcel, we assigned the value that covered the largest portion of the parcel area. In cases, where the grassland index was unavailable, we substituted it with the field's maximum slope (ASTER

GDEM, 2018) instead. The assumption was that with increasing slope the management intensity decreases, consequently also the grassland yields. The management intensity dataset described in Sect. 3.1.1 was used to allocate the number of cutting events to each grassland field. For mowing pastures, firstly, the share between mowing and grazing had to be identified in order to determine their use intensities. Mowing pastures with one cutting event were defined as being used up to 60 % for grazing and all mowing pastures with at least two cutting events as being used up to 20 % for grazing. Secondly, the management intensity of mowing pastures was approximated by the farm's stocking rate (SR) for all parcels with zero to three cuts, i.e. the higher the SR, the higher the use intensity. SR is defined as LSU/P, with LSU being the number of cattle per farm in livestock units and P the farm's total grazing area. Mowing pastures with more than three cuts were assumed to have a high use intensity (Table A2). All assumptions regarding slope, share of grazing in mowing pastures, and SR were discussed with and approved by grassland experts from science and the Office of Food, Agriculture and Forestry (AELF) Weilheim, Germany – a local stakeholder from the study region. More information can be found in Kaim et al. (2025).

3.3 Spatial aggregation of yield data and comparisons

The annual yields estimated by all three approaches are compared for any individual parcel (N = 27138 meadows and mowing pastures) as well as for mean values of hexagons (1 km diagonal) with a total of 2571 hexagons covering the full study region. The yields were averaged per hexagon by weighting with the parcel areas. The aggregation on hexagons has several reasons: Firstly, it is not allowed to publish the parcel shapes and locations, due to data privacy regulations, secondly, the visibility of spatial patterns is improved and, lastly, outlier effects are minimized.

To compare the yields per parcel and hexagon resulting from the three different approaches, the Pearson correlation coefficient was calculated. To analyse the effect of influencing factors, the relationships between yields and mowing frequency, temperature, precipitation and elevation were plotted and Pearson correlation coefficients calculated.

4 Results

4.1 Grassland biomass estimation

4.1.1 Biomass estimation based on the RS approach

315 Sub-parcel biomass estimations are intermediate products of the RS approach. Fig. 3 shows the estimated biomass for single satellite observations on pixel resolution (10 m x 10 m) highlighting the potential to capture small scale variability in patterns of standing biomass and grassland productivity of the RS approach.
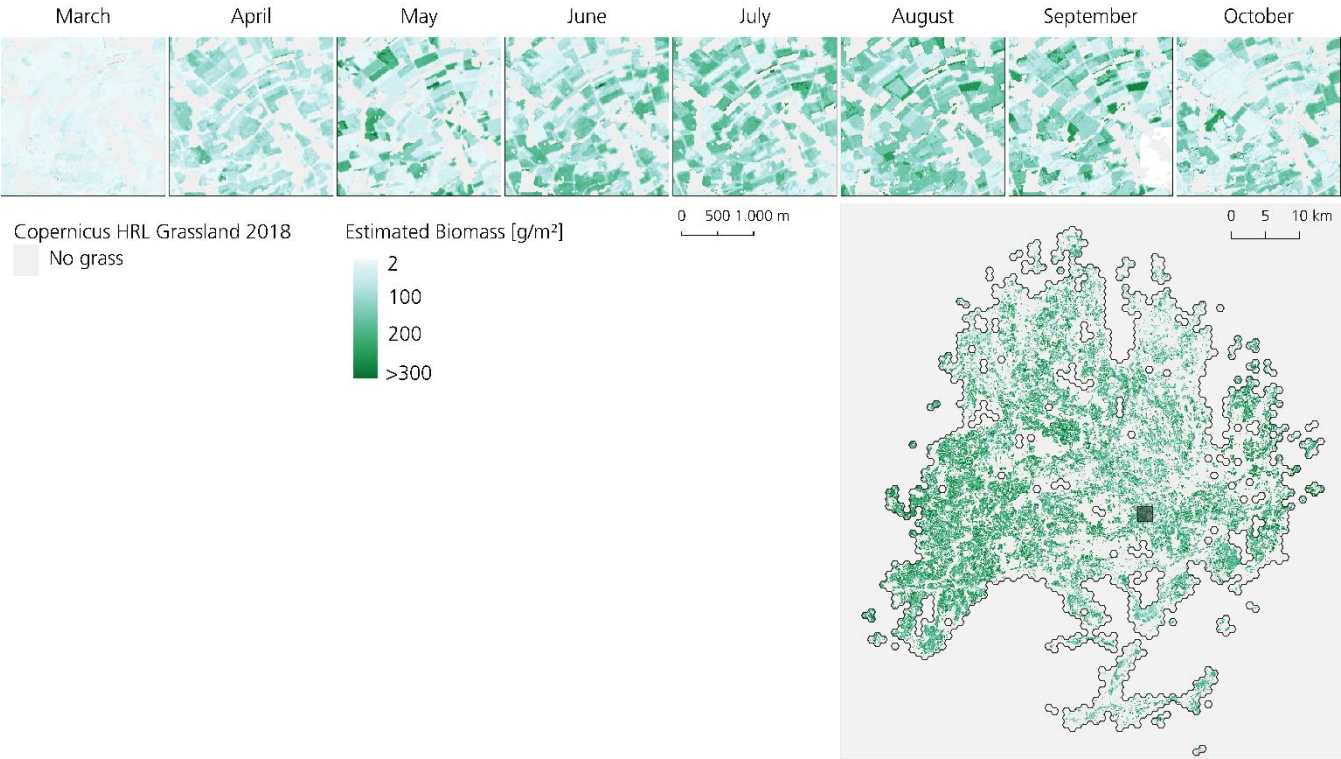
Figure 3: Estimated grassland biomass based on the RS approach for single time steps (30th March, 19th April, 24th May, 13th
320 June, 23rd July, 27th August, 16th September and 16th October) and annual mean biomass for the entire study region in 10x10 m² pixel resolution. The top panel location is indicated as a square in the map of the whole study region.

The estimated biomasses from RS were validated with a part (18%) of the in-situ measurements that were not used for model training (i.e., the test data). The best RS model – extreme gradient boosting regressor, parameterized with a learning rate of
325 0.05, a maximum depth of each tree of three and a number of features used in each tree of 40 % – reached an average R² (CV) of 0.97 and a root mean square error (RMSE) (CV) of 0.18 t ha^{-1} during the internal validation. Band 12 (short-wave infrared), wetness index and days since the last mowing were the most important features according to the relative influence measure. The validation of the model with the test dataset (n=13) lead to a R² of 0.68 and a RMSE of 0.43 t ha^{-1} (compare Figure 4).
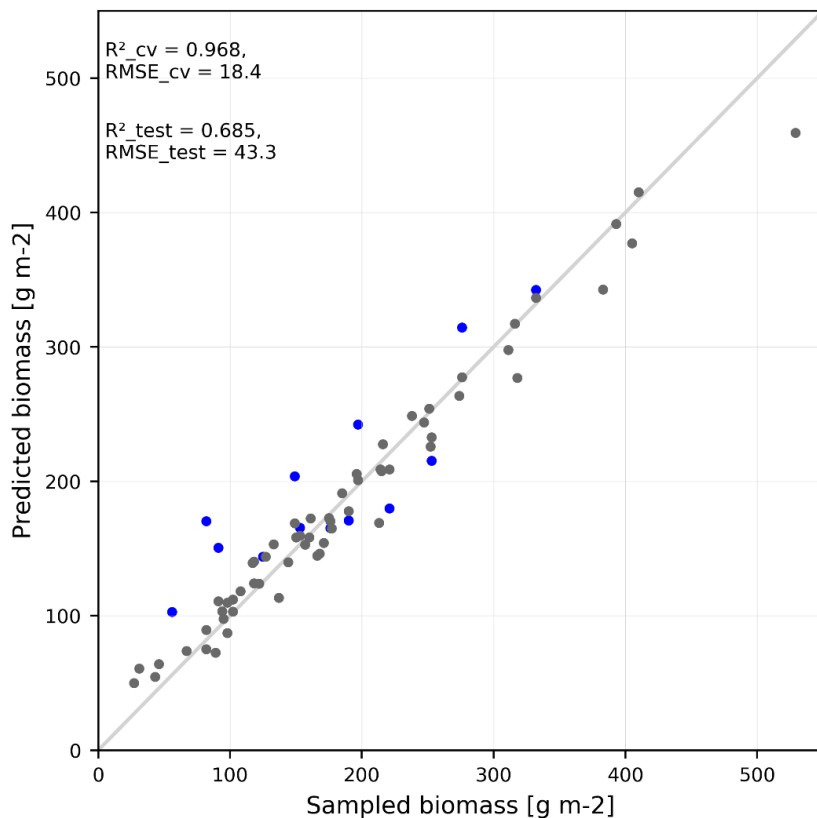
Figure 4: Predicted against sampled biomass (AGB) [g m⁻²] for cross-validation and test data of the RS approach.

4.1.2 Comparison of RS and LDNDC biomass estimations

Time series of AGB for the year 2019 are illustrated in Fig. 5 for three example parcels from the measurement campaign which were used for the RS biomass model training and validation (Section 3.1.2.). The figures include the AGB of in-situ measurements, estimated AGB derived by the RS and LDNDC approach (not available for RVA) and the annual yield based on all three estimation methods. The temporal pattern of LDNDC biomass estimations shows an increase in spring, drops after mowing and increase thereafter. The LDNDC biomass at the first cut is the highest. The temporal profile of the RS-based biomass follows the mowing dates less strictly as compared to LDNDC; however, for most mowing events there is a peak before the mowing event, followed by a drop and a regrowth pattern (Figure 5). When comparing the two biomass time series it becomes clear that the yield of the first mowing event from LDNDC is notably higher than from RS. In contrast to that, RS shows only at times a gradual decline in the local maxima during the growing season (e.g. Figure 5 top).
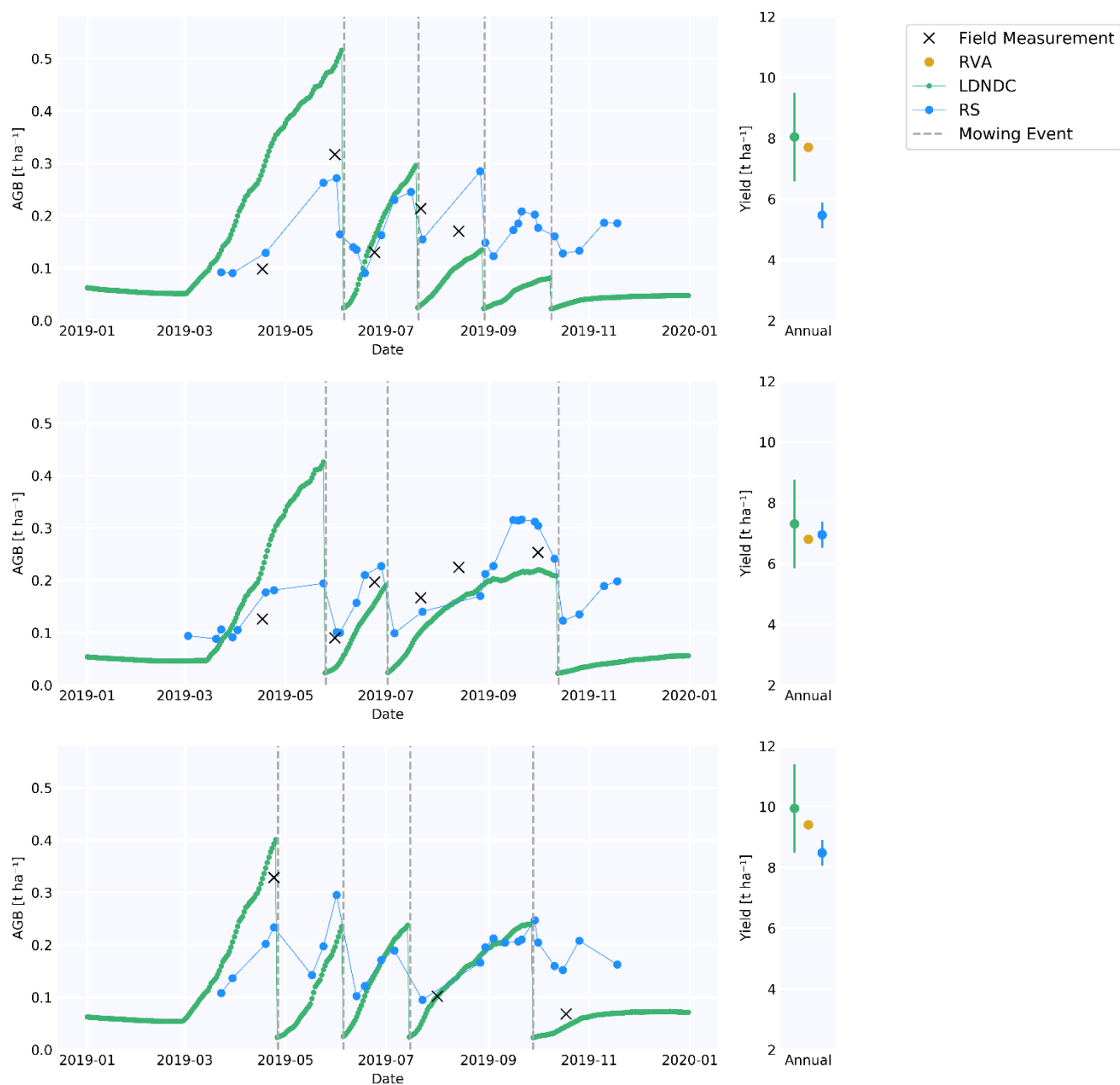
Figure 5: Temporal patterns of grassland AGB estimated by the RS and LDNDC models, in-situ measurements of AGB, annual yields of all three models (on the right of each sub-figure), and mowing dates (horizontal dashed lines) of three grassland parcels in the study area.

The LDNDC model together with the regional input data was also evaluated against the AGB measurements/in-situ data of the above-mentioned field campaign (Section 2.2.2) in a previous study (Boos et al., 2024). To this end, only biomass

measurements from within one week before a mowing event were considered and summed to yearly values per parcel. With this procedure an R² of 0.67 and a RMSE of 1.46 t ha^{-1} a^{-1} were found.

4.2 Estimated annual grassland yields

The spatial patterns of annual grassland yield averaged per hexagon estimated with each method are depicted in Fig. 6. All
355 three models achieve plausible results of annual grassland yields ranging mostly between 3 to 9 t ha^{-1} for the Ammer region in 2019. The annual grassland yields of the entire study region are on average 6.5 t ha^{-1} (372.9 kt) estimated by RS, 7.4 t ha^{-1} (445.5 kt) estimated by LDNDC and 6.9 t ha^{-1} (419 kt) estimated by RVA. The yield maps highlight the spatial variability of grassland yields in the study region, which is consistent in many cases for the three modelling approaches. Noticeable patterns are grasslands with relatively high annual yields in the north of the study region according to the RS and LDNDC models and
360 to a lesser degree also according to the RVA. Grasslands with lower annual yields are present in the northeast and the east of the study area, which is mostly visible in the RS and RVA maps. The centre of the study area shows grasslands with above-average yields mostly based on the RS and even more the LDNDC model results. Grassland in the south of the study area located within Alpine valleys show lower yields in particular for the RS and RVA maps. The entire south-western part of the region shows overall lower grassland yields from the RS and LDNDC, but not the RVA models (Figure 6).

365
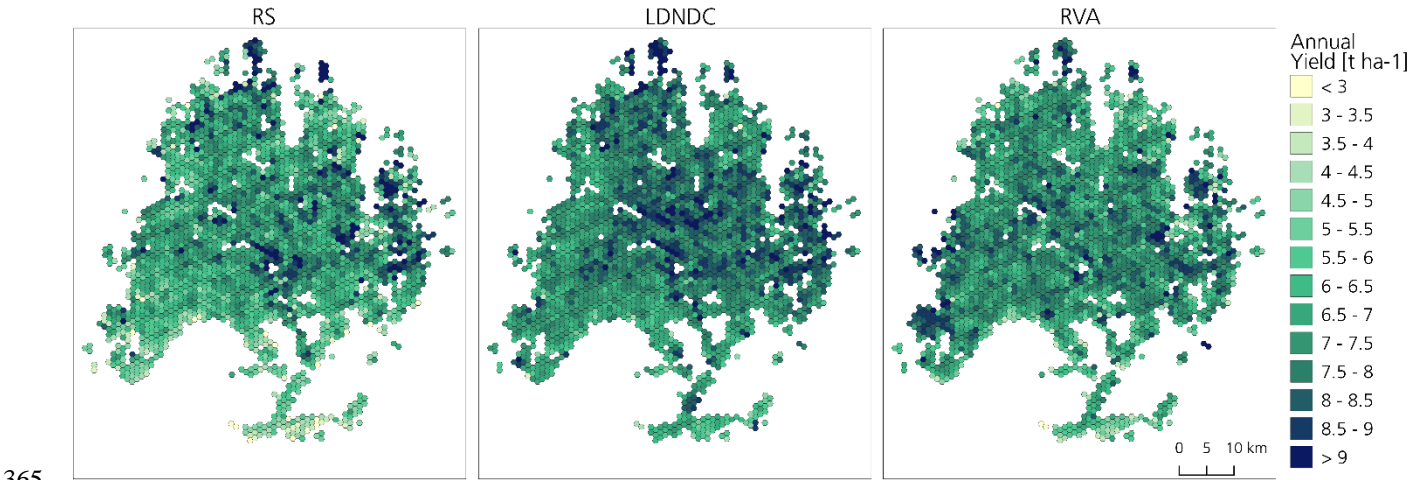
Figure 6: Spatially aggregated (hexagon diagonal length of 1km) annual yield estimates for meadows and mowing pastures in the study area in 2019 based on Remote Sensing, LandscapeDNDC simulations and the reference values approach. Hexagons for which the grassland area is smaller than 1 ha are not shown.

The differences in yields based on RS and LDNDC are not concentrated to specific locations but rather distributed throughout
370 the Ammer region (Figure 7). The yield differences between RS and LDNDC to RVA shows in both cases a north-east, south-west pattern as yields derived from RVA are higher in the south-western part of the study region compared to the yields based on RS and LDNDC. Direct comparisons of hexagon yields reveal that 36 % of the Ammer region show differences smaller than 1 t ha^{-1} among all three approaches. The standard deviation of yield averages of all three methods shows no distinctive spatial patterns (compare A2).
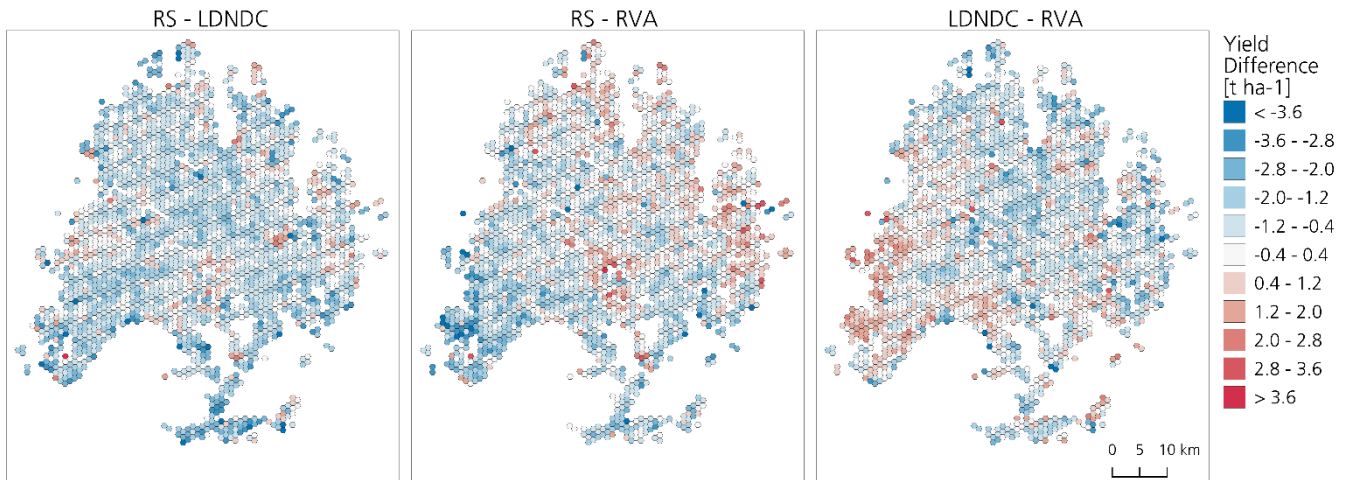
Figure 7: Differences between spatially aggregated (hexagon diagonal length of 1km) annual yield estimates for meadows and mowing pastures in the study area in 2019 based on Remote Sensing, LandscapeDNDC simulations and the reference values approach. Hexagons for which the grassland area is smaller than 1 ha are not shown.

In order to compare the estimated annual grassland yields averaged per hexagon of the three approaches, the frequency distributions were plotted together with the correlation of the annual yield estimates from the different approaches (Figure 8). The frequency distribution shows the largest range for the RS method (variance of 2.0 compared to 0.9 for LDNDC and 1.2 for RVA). The peaks showing the highest number of estimated yield ranges around 7 t ha^{-1} for RS and 7.8 t ha^{-1} for LDNDC and RVA. The comparison of the hexagon yields of the RS and LDNDC approaches shows they largely overlap, in particular for the most common yield values of 6-8 t ha^{-1}. The Pearson's correlation coefficient is 0.67 for the hexagon yields based on RS and LDNDC. It shows significant relationships in all combinations of methods. In general, but particularly for yields approximately below 5 t ha^{-1}, LDNDC shows higher values compared to the RS (and RVA) results. The comparison of annual yields averaged per hexagon from the RS and RVA models also show a generally good agreement (Pearsons's r = 0.64), but an overall larger scattering. The hexagon yields of RVA compare well with the RS yields and also show a tendency for over- or underestimation, respectively, for smaller value ranges below approximately 6 t ha^{-1} when compared to LDNDC yields (Figure 8). The relationship between RVA and LDNDC yields is the weakest with a Pearson's r of 0.47.
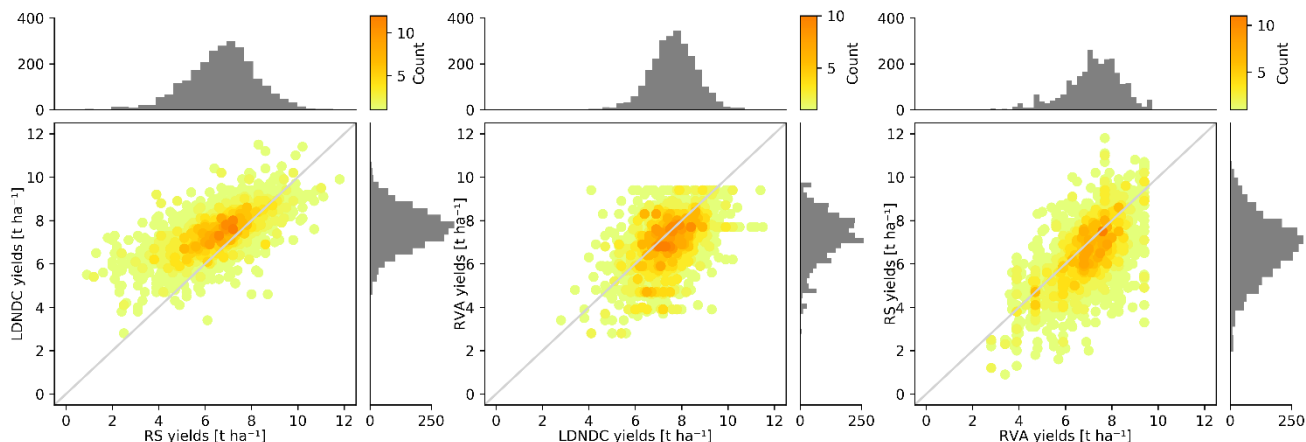
Figure 8: Pairwise relationships of spatially aggregated (hexagon diagonal length of 1km) annual yield estimates for meadows and mowing pastures in the study area in 2019 based on Remote Sensing, LandscapeDNDC simulations and the reference values approach with histograms of the three modelling approaches.

395   The frequency distribution and pairwise comparison of the estimated annual grassland yields per single parcel is shown in Fig. A1. The higher spatial resolution leads to increased scattering for all three approaches compared to the hexagon averages. The yields derived from RS show a variance of 4.6, 95%-percentile of 9.7 and 5%-percentile of 2.5 t ha^{-1}. LDNDC and RVA yields have a variance of 2.2 and 2.7, a 95%-percentile of 9.8 and 9.4 as well as a 5%-percentile of 4.9 and 3.9 t ha^{-1}. The RVA method results in discrete values in contrast to the other two approaches and does not predict values higher than 10 t ha^{-1}. Due

400   to the discrete values of RVA, there is a much higher overlap of yield values resulting in higher counts for the relationships including RVA and varying count scales in Fig. A1. The relationships between estimated yield based on the three approaches are relatively similar as for the hexagon means with a Pearson's r of 0.64 between RS and LDNDC, 0.54 between RS and RVA and 0.44 between LDNDC and RVA. While the LDNDC yields shows maximum values around 12 t ha^{-1}, there are a few values reaching up to 15 t ha^{-1} based on the RS model. LDNDC overestimates lower yields compared to RVA and RS.

405

4.3 Yield estimates in relation to influencing factors

4.3.1 Impact of mowing frequency

410   It can be assumed that the number of mowing and associated fertilization events per year largely influences temporal grassland vegetation growth dynamics and annual yields. We investigated the estimated annual yields per mowing frequency to compare how the relationship found for the remote sensing-based approach differs from the results of the other two approaches. The RS and LDNDC models consider the same mowing dates and all approaches the same frequencies of mowing events. Boxplots

of parcel-based annual yields per mowing frequency show that the estimated yield rises with the number of mowing events per year for all models (Figure 9). The mowing frequency has the strongest impact on the yield derived by the RS method, as the estimated yields show a continuous and clear increase with each additional number of annual mowing events. The Pearson's correlation coefficient, which is significant for all three approaches, is 0.81 for the number of mowing events and the RS yields, 0.74 for LDNDC and 0.66 for RVA. While the average yields for parcels mown 3-5 times correspond relatively well for all three methods, the yields for parcels mown only 1-2 times per year are lower for the RS model compared to the other two models. For a single (two-) cut field, the RS approach shows an average annual yield of 2.1 t ha^{-1} (4.3 t ha^{-1}), whereas LDNDC predicts 4.4 t ha^{-1} (6.5 t ha^{-1}) and RVA 4.4 t ha^{-1} (5.7 t ha^{-1}).
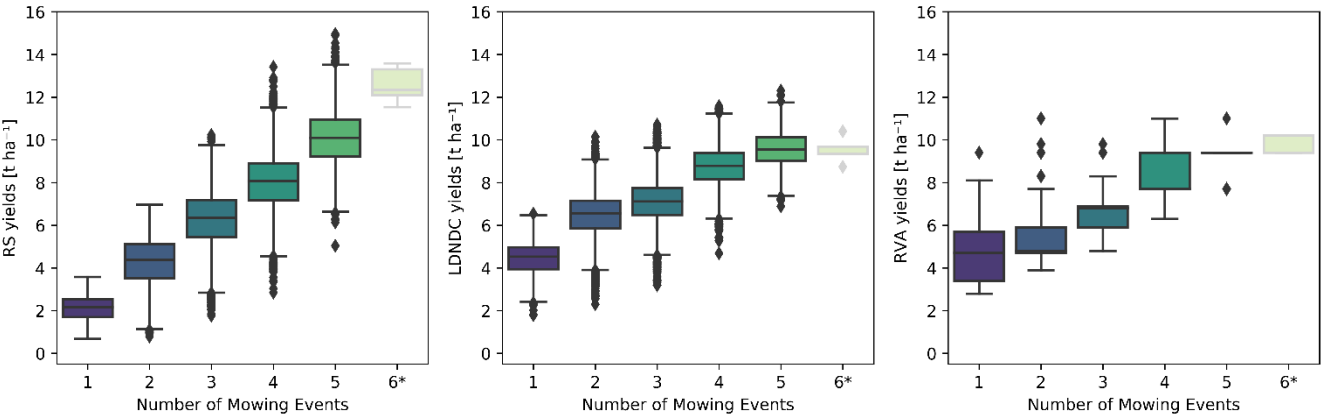
Figure 9: Estimated annual grassland yields per mowing frequency, based on the three models RS, LDNDC and RVA. *Only six parcels were mown six times per year which might not be representative.

4.3.2 Precipitation, temperature and elevation

The annual grassland yields increase with increasing mean annual temperature (MAT) mainly for RS and LDNDC but less for RVA (Figure 10), however the increase of yields is associated with a higher variability of yields at higher temperature classes. RVA yields stagnate at MAT above 12.25 °C. The Pearson's r is significant for all relationships of annual yield and MAT but shows in general a low positive correlation (RS: r = 0.2; LDNDC: r = 0.26; RVA: r = 0.1). The yields estimated by all three methods stay relatively constant for all precipitation levels present in the study region. Pearson's r values are -0.2 (RS), -0.18 (LDNDC) and -0.05 (RVA). The relationship between yield and elevation show a negative relationship as annual yields decrease with elevation for all methods. Pearson's r for the relationships between the estimated yields and elevation are -0.19 (RS), -0.23 (LDNDC) and -0.08 (RVA). However, for the RVA yield estimates stay on average constant for an elevation of 500-900 m a.s.l. and only afterwards decline. Overall, the relationships between estimated yields and site conditions, such as temperature, precipitation and elevation, are relatively low and similar for the three methods (Figure 10). These patterns stayed the same when the relationships were tested for individual mowing frequencies (compare Fig. A3).
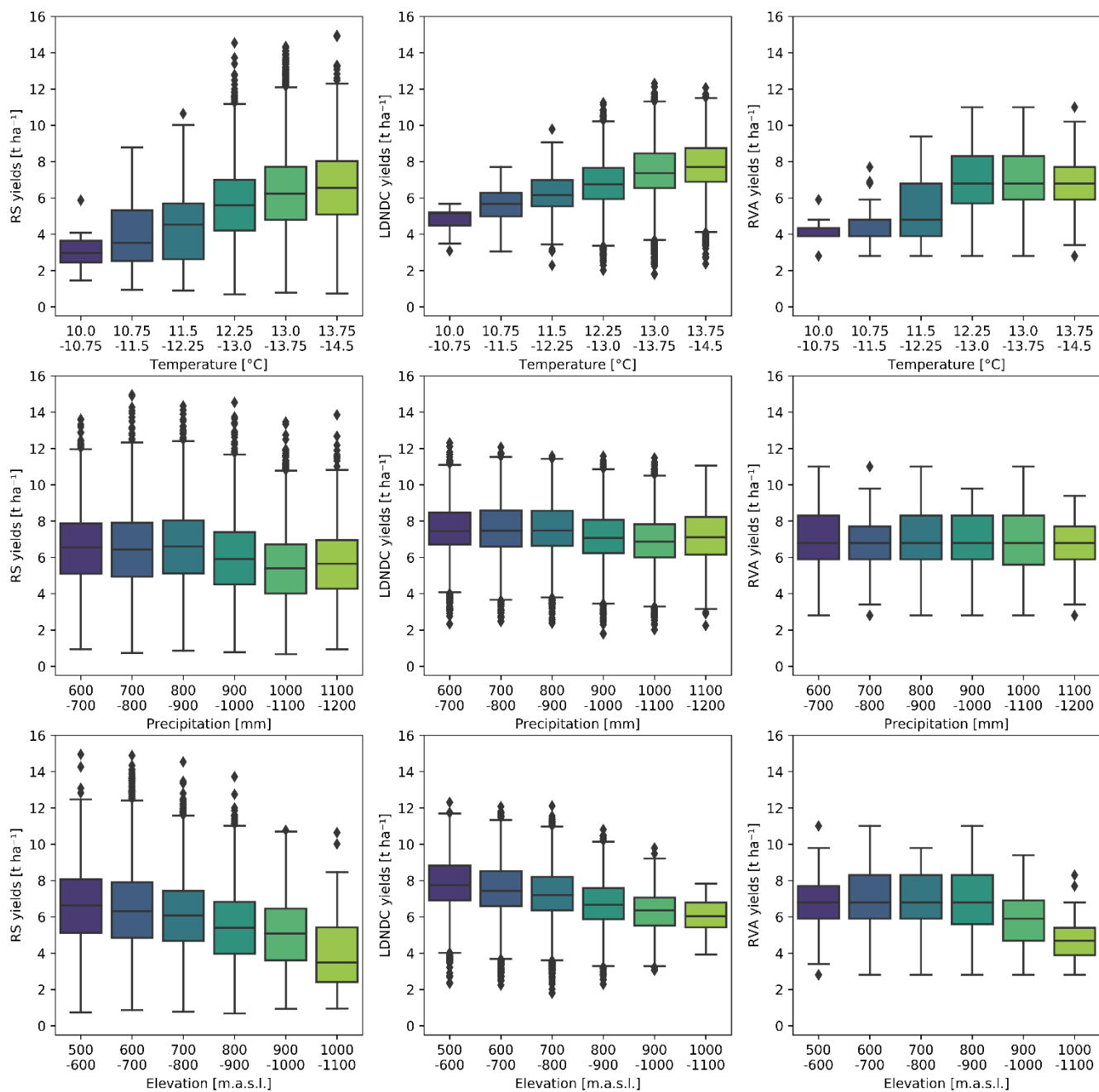
Figure 10: Aggregation statistics of estimated annual yields based on the three models per temperature, precipitation and elevation classes.

5 Discussion

5.1 Performance of biomass modelling results

Despite the numerous studies of empirical grassland biomass modelling based on satellite and field data, only few studies were carried out in areas characterized by heterogeneous and small grassland parcels which are mowed multiple times at different dates during the year. Of these previous studies the potential of various vegetation indices derived from medium resolution sensors (MODIS, MERIS) to estimate grassland biomass for single sites in Ireland and the Netherlands were investigated (Ali et al., 2017b; Ullah et al., 2012). Based on Landsat and Sentinel-2 data, grassland biomass and height were estimated for study regions in Germany, France, Spain and Austria using various regressors, such as multi-linear regression, random forest or deep learning models, resulting in accuracies (R^2) of 0.45-0.79 (Barrachina et al., 2015; Dusseux et al., 2022; Eder et al., 2023; Muro et al., 2022; Schwieder et al., 2020). The performance of empirical models seems to depend more on the number of the training data and the variety of grassland types and use intensities included than the tested regressors and model parameters. Regarding the tested indices and bands, wetness indices (Barrachina et al., 2015) and red-edge, near-infrared and short-wave infrared bands (Dusseux et al., 2022) were found to be valuable for grassland biomass modelling. With an R^2 of 0.68 for the test data set, we reach comparable results with the RS approach. In our case, the R^2 of the cross-validation was relatively high (0.97) which might indicate overfitting which should be avoided. As far as we know, information on the time since the last mowing event was not included in previous studies so far. This parameter, however, was among the most important input features for the extreme gradient boosting model applied. Including information on mowing dates seems to be advantageous for grassland biomass estimation in study regions characterized by intensive grassland management.

The performance of LDNDC to reproduce grassland yields of individual mowing events was found to be comparable or better than other process-based models. For the annual yields, where regional input data was employed, the performance measures were even better than for single cuts. For further details, see Boos et al. (2024). We assume that the results of LDNDC can be transferred to other process-based models (e.g. Daycent, APSIM).

The comparison of the temporal patterns on estimated above-ground biomass shows that both results (RS and LDNDC) follow the mowing dynamics closely. This behaviour is expected as the mowing dates are directly included in both modelling approaches. However, the RS biomass estimates fluctuate more than the ones stemming from LDNDC and clearly depend on the availability of cloud-free satellite observations. LDNDC biomasses show a very high peak in the first growth cycle which is often double the amount of the RS-based estimates. It is known from previous works that LandscapeDNDC overestimates yields from the first cut of the year and underestimates the ones of later cuts (Boos et al., 2024). However, it is also possible that the RS method underestimates the first cut yield as the AGB estimation might be prone to a saturation effect. In addition, AGBs at the higher end of the training data distribution are less likely predicted and strongly depend on a well-balanced training data set.

Another important aspect is the AGB that needs to be estimated. Here, the goal was to estimate the biomass per cut, which corresponds to the yield. Farmers are generally advised to cut at 7 cm, though this can vary in practice. In the RS approach,

475 total AGB was used because the sensor not only provides data from a certain height upwards. The LDNDC was calibrated using AGB samples taken from 7 cm. The cutting height for the reference values in the RVA approach likely varies considerably and cannot be reconstructed. Therefore, potential uncertainties due to AGB information of different cutting heights must be assumed.

5.2 Spatial patterns of annual grassland yield and influencing factors

480 The spatial yield patterns from all three methods match in some regions and differ in others. The spatial patterns mostly resemble the mowing frequency which is a major influencing factor determining grassland yields (Bernhardt-Römermann et al., 2011). In many cases, areas with high yield match with areas showing a higher number of mowing events. For example, in the East or the North of the study area, for many grasslands, high annual yield estimates, in particular from the LDNDC and RS models, fit well to a large number of mowing events. This relationship is also underlined by the significant Pearson

485 correlation coefficients of 0.66-0.81 between annual yield and number of mowing events for the three models. Other regions, for example, East of lake Starnberg and in the centre of the study area show high yield estimates but rather low to intermediate mowing frequencies (compare Figures 1 and 6). This discrepancy must be explained by other factors influencing grassland yields, which were not looked into in detail here, such as soil conditions or an optimal interplay of influencing factors. Intensified mowing, usually accompanied by more fertilization, enhances biomass production and changes species composition

490 towards more productive vegetation with less species in systems not strongly limited by other factors (Isbell et al., 2013; Mayel et al., 2021; Savage et al., 2021). The influence of site conditions, in particular climatic conditions, on the spatial annual yield patterns is twofold. Firstly, the conditions in the year of interest influence vegetation growth in that particular year. Secondly, climatic site conditions determine species composition and management options of grasslands in the long-term as well as soil properties like soil organic carbon. There are overall smaller yields visible in the south and southwest of the study region which

495 matches the temperature patterns. Apart from that, the resemblance between annual yield maps and temperature, precipitation and elevation is relatively low (compare Figures 1 and 6). Pearson correlation coefficients were significant but low for all combinations (-0.23-0.26). For correlation tests with as many data as in our case, correlations tend to be significant (Rouder et al., 2009). It can be assumed that climatic effects have a more significant influence on grassland yields on a larger spatial scale, e.g. continental scale (Emadodin et al., 2021; Goliski et al., 2018; Zhang et al., 2018). Either the climatic gradients are

500 not large enough in our study area to explain grassland yields or climatic conditions play only a minor role as yields are mostly determined by management. A beforehand anticipated difference in the relationship between site conditions and yields for the three models, as LDNDC includes these data directly and RS captures the effects indirectly, was not found. Furthermore, in this study only one year was examined which showed relatively normal climatic conditions. The differences between the models are most probably higher in extreme years, e.g. 2018, as extreme climatic effects can be depicted by LDNDC and RS,

505 but not the RVA (compare Boos et al. (2024)).

5.3 Differences between the modelling results

When comparing the annual yield values, the RS method leads to overall lower estimates compared to LDNDC and RVA, which can be explained in several ways. On the one hand, the RS method is prone to underestimation of annual yields as the empirical model of above-ground biomass is rather underestimating. This is related to the available training data which stem from field measurements throughout the year with only a few samples from shortly before a mowing event. In this regard, the distribution of training samples might not be optimal and rather missing data of high biomass, which results in an underestimation of the empirical model. In addition, the biomass estimation and the mowing detection are both dependent on the availability of cloud-free satellite observations. Biomass estimates from periods shortly before mowing events might be missed and, consequently, the yield related to the mowing event is potentially lower than in reality. Further, mowing events are potentially missed entirely due to cloud coverage (Reinermann et al., 2023) which additionally leads to these yields missing in the annual estimate. As the mowing information is included in all three approaches, missed mowing events due to clouds also affects the LDNDC and RVA yields.

On the other hand, while the RS method is more likely underestimating yields, LDNDC and RVA are more prone to overestimation. This is related to the fact that some factors which negatively influence vegetation growth and, therefore, yields are not included in the LDNDC and RVA models. For LDNDC, these are neglected local factors, like north-facing slopes and lateral run-off as well as the calibration on lysimeter data taken under favourable conditions (Haas et al., 2013). The input into the RVA model is very limited in the sense that local factors influencing the current growing conditions as well as yearly climate data are not included. Both models therefore tend to represent optimal growing conditions and hence overestimate yields. Even though the RS model is not specifically addressing these factors, they are captured by the spatially varying reflectance signal.

When examining the lower yields resulting from the RS model compared to the other model estimates, it becomes clear that this effect is most prominent for grasslands mown one or two times per year (compare Figure 9). This could be related to a grazing effect. Such extensively used grasslands are very often also grazed (Schoof et al., 2020a, c). In the LandscapeDNDC simulations, the amount of grazed material remains on the field and leads to an overestimation of the yield from the next cut, either directly within the same year or via plant storage over the winter in the first cut of the following year. The RS method does not specifically account for grazing, and as the estimated biomass before mowing events is used, there is no such accumulation effect. Assuming that the difference between the RS and LDNDC yields of extensively used grasslands stems mostly from the grazing effect, it might be used to calculate grazed yields (e.g. Chang et al. (2015)). The RVA considers grazing on mowing pastures. However, it follows a rather simple approach estimating the grazing intensities with the farm's stocking rate. The stocking density, i.e. the number of LSU per ha field area, would be a more meaningful measure but the IACS data does not include any information on the type of animal husbandry or grazing intensities. Therefore, yields of

mowing pastures might be over- or underestimated in the RVA resulting from lack of detail in the information on grazing intensity.

It is also important to mention that all three approaches depend on mowing information as input data. RS and LDNDC use the mowing dates and RVA the mowing frequency of grassland parcels in their algorithms. As the same mowing data was used for all three approaches, the results are comparable to each other. The importance of the mowing data for model performance is difficult to assess, as the methods do not work without it, however one could investigate the impact of using various sources for mowing dates. In addition, there is also an uncertainty in the mowing data as it was derived from Sentinel-2 time series which is prone to gaps due to cloudy weather conditions. The R^2 for the mowing detection in the study area is 0.65.

Other studies investigating grassland yield in Europe mostly focus on a continental scale and are not conducted on parcel level. However, the estimated yields are in similar value ranges compared to our results, e.g. Chang et al., 2015; Smit et al., 2008.

5.4 Advantages and limitations of the approaches and implications drawn from them

All of the three approaches investigated in this study hold individual advantages and limitations to estimate annual grassland yields in southern Germany (compare Table 1). The RS approach depends on well distributed training data and cloud-free satellite observations. The value distribution of the training data determines the range of potential predictions of the empirical model and is, therefore, crucial. This also plays an important role considering the transferability (in space and time) of the approach. Without additional training data, the approach can hardly be applied in regions with different conditions. The necessity of satellite data can be particularly problematic for regions or time periods with relatively high cloud coverage. For instance, in Germany, the year 2021 was characterized by an overall relatively low number of cloud-free satellite observations making satellite data-based products as the mowing detection or biomass estimation more challenging (Reinermann et al., 2023). Despite these limitations, estimating grassland yields based on RS has several advantages. Even small-scale spatial effects are depicted enabling the detection of parcels with reduced yields. Current spatio-temporal variation is mirrored, such as yield reduction through drought periods. In addition, despite the need of training data, no large input data set or parametrisation is needed for the RS approach, which might even enhance the transferability.

Major advantages of LandscapeDNDC – as a bio-geochemical model – are that firstly, spatial and temporal variations are accounted for, secondly, the direct relation to input data like climate, soil, and air chemistry is possible, thirdly, carbon, nitrogen and water budgets are modelled as well, and at last, scenarios (climate or management) can be studied. This also means, that high resolution input data (climate, soil, management and air chemistry) needs to be available for the modelled domains. Generally, the model performs best, if individual detailed simulations are performed, which are than aggregated to a larger spatial-temporal scale (e.g. field scale to hexagons). However, slope and orientation as well as species composition not accounted for. To transfer the model, it is ideally recalibrated based on local measurements even though for crops it has

performed very well even on a global scale without this step (Jägermeyr et al., 2021). For grasslands, so far, LDNDC has been used successfully in Switzerland, the UK, and Germany (Houska et al., 2017; Molina-Herrera et al., 2016; Petersen et al., 2021).

575

The RVA is based on field measurements and mostly static input data, apart from the mowing frequency and stocking rate. It is therefore not able to depict spatial and temporal variations, such as drought episodes (Diepolder et al., 2016). It also needs input data, first of all the yield reference values which are rarely provided on a larger scale. Further, cattle numbers, land-use type (i.e. meadow or mowing pasture), and mowing frequencies are needed, which is – especially on scales larger than farm

580   scale – at times difficult to obtain as it is not openly accessible or does not exist. An advantage of the RVA is that the approach is relatively straight forward and does not need large amounts of computational power. Additionally, reference values are usually used by farmers to calculate their field's fertilizer requirements. The approach is thus more useful at farm scale where data can easily be obtained and becomes more difficult to use at larger scales due to restricted data availability.

585   All three approaches are in a way calibrated with regional data and, therefore, their applicability limited to regions with relatively similar grassland ecosystems. The RS approach relies on biomass samples, LDNDC on regional model calibration (input data) and RVA on the grassland index. As the management is highly relevant for the productivity and yield of grasslands in intensively used areas, such as in our study area, not only the geographic conditions and species composition but also the management strategies and use intensities of grassland ecosystems need to be comparable to transfer the calibrated models.

590   Including additional ground-truth data from the target region for calibration would largely improve the transferability. However, the models also rely on data, such as IACS or mowing information, which need to be available in potential target regions.

Table 1: Overview of methodological properties, both advantages and drawbacks, comparing the three yield estimation approaches.
595   **The strongest relative fit between method and property is indicated by +++, the weakest fit by +.**

Methodological Properties	Yield Estimation Approaches		
	RS	LDNDC	RVA
Applicability without expert knowledge	+	+	++
Low in necessary hardware and processing time	+	+	+++
Independent of closed access data	++	+	+
Low in required data	++	+	+++
Sub-parcel variability	+++	+	+
Interannual variability	+++	+++	+
Internal test accuracy	++	++	*
Plausibility of results	+++	+++	++

Distribution range and data continuity of results	+++	+++	+
Transferability	+	++	++

*unknown

6 Conclusions and outlook

Within this study, annual grassland yield was estimated for a study region in southern Germany in 2019 based on three different
600   approaches: i) an empirical remote sensing model, ii) a bio-geochemical model (LandscapeDNDC) and iii) a rule-based reference value approach. It was shown that grassland yields can be estimated based on three completely different approaches, as plausible and comparable results were reached for the study area. The three models contain varying input data sets, however, all use mowing information as a driver. The mowing frequency was found to be the most important influencing factor for grassland yields in the study region for all three approaches.

605

Yield patterns and comparisons between the approaches showed that all three methods can be legitimately used for yield estimation (value ranges of approximately 4-9 t ha^{-1}) in pre-Alpine grassland ecosystems considering individual limitations. All approaches need the mowing information on parcel level as input data. When a training data set of well-balanced AGB samples and cloud-free satellite observations are available, it is advisable to use the RS approach. Depending on the training
610   data distribution the RS approach is capable to estimate grassland yields on a regional level and also capture small-scale patterns on field level (and beyond). LandscapeDNDC is also recommended to be used on a regional or even continental level. Detailed data on climate, soil and management is needed and strongly determines the performance. For the field scale, also the RVA can be used as the needed data can be obtained more easily for this rather than the regional level. To investigate single cut yields only RS and LDNDC can be used as RVA provides solely an annual yield estimate.

615

The expected biases include a likely overestimation of LDNDC of the first cut yield and an underestimation of the first cut yield – and to a lesser degree annual yields in general – of the RS approach. To investigate yield patterns over time, only RS and LDNDC are useful as the RVA does not include actual conditions. Improved grassland yield estimations could be obtained with more AGB sample data, in particular when analysing years with climatic extremes as the AGB data might not be
620   representative. In addition, validation data on annual grassland yields would be needed to evaluate the approaches in more detail.

The study presents synergies of grassland yield estimation approaches which is particularly important as spatial information on grassland yield is limited. Based on robust grassland yield estimations multi-annual analyses can be conducted and effects

625 of climate change, for instance, investigated. In addition, a comprehensive understanding of grassland ecosystems is facilitated supporting authorities and science.

Author contribution

SR, CB, AK and RK designed the study. SR and AS collected the field data. SR carried out the remote sensing model, CB the process-based model and AK the reference value approach with the help from all co-authors. SR prepared the manuscript and
630 all co-authors supported with writing or suggestions for adjustment.

Acknowledgements

We would like to thank all state authorities which provided data and the multiple helpers during field campaigns. In addition, we would like to thank the two anonymous reviewers for very clear and helpful comments during the reviewing process.

References

635 ASTER GDEM Global Digital Elevation Model V003, 2018.

Ali, I., Barrett, B., Cawkwell, F., Green, S., Dwyer, E., and Neumann, M.: Application of Repeat-Pass TerraSAR-X Staring Spotlight Interferometric Coherence to Monitor Pasture Biophysical Parameters: Limitations and Sensitivity Analysis, IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, 10, 3225–3231, https://doi.org/10.1109/JSTARS.2017.2679761, 2017a.

640 Ali, I., Cawkwell, F., Dwyer, E., and Green, S.: Modeling Managed Grassland Biomass Estimation by Using Multitemporal Remote Sensing Data - A Machine Learning Approach, IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, 10, 3254–3264, https://doi.org/10.1109/JSTARS.2016.2561618, 2017b.

Auer, I., Böhm, R., Jurkovic, A., Lipa, W., Orlik, A., Potzmann, R., Schöner, W., Ungersböck, M., Matulla, C., Briffa, K., Jones, P., Efthymiadis, D., Brunetti, M., Nanni, T., Maugeri, M., Mercalli, L., Mestre, O., Moisselin, J., Begert, M., Müller–
645 Westermeier, G., Kveton, V., Bochnicek, O., Stastny, P., Lapin, M., Szalai, S., Szentimrey, T., Cegnar, T., Dolinar, M., Gajic–Capka, M., Zaninovic, K., Majstorovic, Z., and Nieplova, E.: HISTALP—historical instrumental climatological surface time series of the Greater Alpine Region, International Journal of Climatology, 27, 17–46, https://doi.org/10.1002/joc.1377, 2007.

Ball, J. T., Woodrow, I. E., and Berry, J. A.: A Model Predicting Stomatal Conductance and its Contribution to the Control of Photosynthesis under Different Environmental Conditions, in: Progress in Photosynthesis Research, edited by: Biggins, J.,
650 Springer Netherlands, Dordrecht, 221–224, https://doi.org/10.1007/978-94-017-0519-6textunderscore 48, 1987.

Barrachina, M., Cristóbal, J., and Tulla, A. F.: Estimating above-ground biomass on mountain meadows and pastures through remote sensing, International Journal of Applied Earth Observation and Geoinformation, 38, 184–192, https://doi.org/10.1016/j.jag.2014.12.002, 2015.

Bengtsson, J., Bullock, J. M., Egoh, B., Everson, C., Everson, T., O'Connor, T., O'Farrell, P. J., Smith, H. G., and Lindborg, R.: Grasslands-more important for ecosystem services than you might think, Ecosphere, 10, e02582, https://doi.org/10.1002/ecs2.2582, 2019.

Beniston, M.: Climatic Change in Mountain Regions: A Review of Possible Impacts, Climatic Change, 59, 5–31, https://doi.org/10.1023/A:1024458411589, 2003.

Berauer, B. J., Wilfahrt, P. A., Arfin-Khan, M. A. S., Eibes, P., Heßberg, A. von, Ingrisch, J., Schloter, M., Schuchardt, M. A., and Jentsch, A.: Low resistance of montane and alpine grasslands to abrupt changes in temperature and precipitation regimes, Arctic, Antarctic, and Alpine Research, 51, 215–231, https://doi.org/10.1080/15230430.2019.1618116, 2019.

Bernhardt-Römermann, M., Römermann, C., Sperlich, S., and Schmidt, W.: Explaining grassland biomass–the contribution of climate, species and functional diversity depends on fertilization and mowing frequency, Journal of Applied Ecology, 48, 1088–1097, https://doi.org/10.1111/j.1365-2664.2011.01968.x, 2011.

Bondeau, A., SMITH, P. C., ZAEHLE, S., Schaphoff, S., Lucht, W., Cramer, W., Gerten, D., LOTZE–CAMPEN, H., Müller, C., Reichstein, M., and SMITH, B.: Modelling the role of agriculture for the 20th century global terrestrial carbon balance, Global Change Biology, 13, 679–706, https://doi.org/10.1111/j.1365-2486.2006.01305.x, 2007.

Boos, C., Reinermann, S., Wood, R., Ludwig, R., Schucknecht, A., Kraus, D., and Kiese, R.: Drought impact on productivity: Data informed process-based field-scale modeling of a pre-Alpine grassland region, https://doi.org/10.5194/egusphere-2024-2864, 2024.

Chang, J., Viovy, N., Vuichard, N., Ciais, P., Campioli, M., Klumpp, K., Martin, R., Leip, A., and Soussana, J.-F.: Modeled Changes in Potential Grassland Productivity and in Grass-Fed Ruminant Livestock Density in Europe over 1961-2010, PLOS ONE, 10, e0127554, https://doi.org/10.1371/journal.pone.0127554, 2015.

Chang, J. F., Viovy, N., Vuichard, N., Ciais, P., Wang, T., Cozic, A., Lardy, R., Graux, A.-I., Klumpp, K., Martin, R., and Soussana, J.-F.: Incorporating grassland management in ORCHIDEE: model description and evaluation at 11 eddy-covariance sites in Europe, Geoscientific Model Development, 6, 2165–2181, https://doi.org/10.5194/gmd-6-2165-2013, 2013.

CLMS, High Resolution Vegetation Phenology and Productivity (HR-VPP) Start-of-Season Date (SOSD), https://doi.org/10.2909/c1c46cb2-b02b-4013-aae5-a54a8c018b1e 2019.

Le Clec'h, S., Finger, R., Buchmann, N., Gosal, A. S., Hörtnagl, L., Huguenin-Elie, O., Jeanneret, P., Lüscher, A., Schneider, M. K., and Huber, R.: Assessment of spatial variability of multiple ecosystem services in grasslands of different intensities, Journal of environmental management, 251, 109372, https://doi.org/10.1016/j.jenvman.2019.109372, 2019.

Dengler, J., Janišová, M., Török, P., and Wellstein, C.: Biodiversity of Palaearctic grasslands: a synthesis, Agriculture, Ecosystems & Environment, 182, 1–14, https://doi.org/10.1016/j.agee.2013.12.015, 2014.

Diepolder, M., Heinz, S., Kuhn, G., and Raschbacher, S.: Ertrags- und Nährstoffmonitoring Grünland Bayern: Zusammengefasste Ergebnisse des Forschungsprojektes, https://www.lfl.bayern.de/mam/cms07/iab/dateien/ertrags_naehrstoffmonitoring_gruenland_bayern_diepolder_sub.pdf. 2016.

Drusch, M., Del Bello, U., Carlier, S., Colin, O., Fernandez, V., Gascon, F., Hoersch, B., Isola, C., Laberinti, P., Martimort, P., Meygret, A., Spoto, F., Sy, O., Marchese, F., and Bargellini, P.: Sentinel-2: ESA's Optical High-Resolution Mission for GMES Operational Services, Remote Sensing of Environment, 120, 25–36, https://doi.org/10.1016/j.rse.2011.026, 2012.

Dusseux, P., Guyet, T., Pattier, P., Barbier, V., and Nicolas, H.: Monitoring of grassland productivity using Sentinel-2 remote sensing data, International Journal of Applied Earth Observation and Geoinformation, 111, 102843, https://doi.org/10.1016/j.jag.2022.102843, 2022.

Eder, E., Riegler-Nurscher, P., Prankl, J., and Prankl, H.: Grassland Yield Estimation Using Transfer Learning from Remote Sensing Data, KI - Künstliche Intelligenz, 37, 187–194, https://doi.org/10.1007/s13218-023-00814-9, 2023.

Emadodin, I., Corral, D. E. F., Reinsch, T., Kluß, C., and Taube, F.: Climate Change Effects on Temperate Grassland and Its Implication for Forage Production: A Case Study from Northern Germany, Agriculture, 11, 232, https://doi.org/10.3390/agriculture11030232, 2021.

Farquhar, G. D., Caemmerer, S. von, and Berry, J. A.: A biochemical model of photosynthetic CO_2 assimilation in leaves of C_3 species, Planta, 149, 78–90, https://doi.org/10.1007/BF00386231, 1980.

Friedman, J. H.: Greedy Function Approximation: A Gradient Boosting Machine, The Annals of Statistics, 29, 1189–1232, 2001.

Gerowitt, B., Schroeder, S., Dempfle, L., Engels, E.-M., Engels, J., Feindt, H. Peter, Graner, A., Hamm, U., Heißenhuber, A., Schulte-Coerne, H., and Wolters, V.: Biodiversität im Grünland – unverzichtbar für Landwirtschaft und Gesellschaft: Stellungnahme des Wissenschaftlichen Beirats für Biodiversität und Genetische Ressourcen beim Bundesministerium für Ernährung, Landwirtschaft und Verbraucherschutz. 2013.

Gibon, A.: Managing grassland for production, the environment and the landscape. Challenges at the farm and the landscape level, Livestock Production Science, 96, 11–31, https://doi.org/10.1016/j.livprodsci.2005.05.009, 2005.

Gibson, D. J.: Grasses and grassland ecology, Oxford University Press, Oxford, UK, 2009.

Gilhaus, K., Boch, S., Fischer, M., Hölzel, N., Kleinebecker, T., Prati, D., Rupprecht, D., Schmitt, B., and Klaus, V. H.: Grassland management in Germany: effects on plant diversity and vegetation composition, https://doi.org/10.14471/2017.37.010, 2017.

Goliski, P., Czerwiski, M., Jørgensen, M., Mølmann, J. A. B., Goliska, B., and Taff, G.: Relationship Between Climate Trends and Grassland Yield Across Contrasting European Locations, Open life sciences, 13, 589–598, https://doi.org/10.1515/biol-2018-0070, 2018.

Gossner, M. M., Lewinsohn, T. M., Kahl, T., Grassein, F., Boch, S., Prati, D., Birkhofer, K., Renner, S. C., Sikorski, J., Wubet, T., Arndt, H., Baumgartner, V., Blaser, S., Blüthgen, N., Börschig, C., Buscot, F., Diekötter, T., Jorge, L. R., Jung, K., Keyel, A. C., Klein, A.-M., Klemmer, S., Krauss, J., Lange, M., Müller, J., Overmann, J., Pašali, E., Penone, C., Perovi, D. J., Purschke, O., Schall, P., Socher, S. A., Sonnemann, I., Tschapka, M., Tscharntke, T., Türke, M., Venter, P. C., Weiner, C. N., Werner, M., Wolters, V., Wurst, S., Westphal, C., Fischer, M., Weisser, W. W., and Allan, E.: Land-use intensification causes multitrophic homogenization of grassland communities, Nature, 540, 266–269, https://doi.org/10.1038/nature20575, 2016.

Del Grosso, S. J. and Parton, W. J.: History of Ecosystem Model Development at Colorado State University and Current Efforts to Address Contemporary Ecological Issues, in: Bridging Among Disciplines by Synthesizing Soil and Plant Processes, edited by: Wendroth, O., Lascano, R. J., and Ma, L., American Society of Agronomy and Soil Science Society of America, Madison, WI, USA, 53–69, https://doi.org/10.2134/advagricsystmodel8.2017.0012, 2019.

Grote, R., Lavoir, A.-V., Rambal, S., Staudt, M., Zimmer, I., and Schnitzler, J.-P.: Modelling the drought impact on monoterpene fluxes from an evergreen Mediterranean forest canopy, Oecologia, 160, 213–223, https://doi.org/10.1007/s00442-009-1298-9, 2009.

Haas, E., Klatt, S., Fröhlich, A., Kraft, P., Werner, C., Kiese, R., Grote, R., Breuer, L., and Butterbach-Bahl, K.:
730 LandscapeDNDC: a process model for simulation of biosphere–atmosphere–hydrosphere exchange processes at site and regional scale, Landscape Ecology, 28, 615–636, https://doi.org/10.1007/s10980-012-9772-x, 2013.

Hagolle, O., Huc, M., Desjardins, C., Auer, S., and Richter, R.: Maja Algorithm Theoretical Basis Document, https://doi.org/10.5281/zenodo.1209633, 2017.

Holtgrave, A.-K., Röder, N., Ackermann, A., Erasmi, S., and Kleinschmit, B.: Comparing Sentinel-1 and -2 Data and Indices
735 for Agricultural Land Use Monitoring, Remote Sensing, 12, 2919, https://doi.org/10.3390/rs12182919, 2020.

Holzworth, D. P., Huth, N. I., deVoil, P. G., Zurcher, E. J., Herrmann, N. I., McLean, G., Chenu, K., Oosterom, E. J. van, Snow, V., Murphy, C., Moore, A. D., Brown, H., Whish, J. P. M., Verrall, S., Fainges, J., Bell, L. W., Peake, A. S., Poulton, P. L., Hochman, Z., Thorburn, P. J., Gaydon, D. S., Dalgliesh, N. P., Rodriguez, D., Cox, H., Chapman, S., Doherty, A., Teixeira, E., Sharp, J., Cichota, R., Vogeler, I., Li, F. Y., Wang, E., Hammer, G. L., Robertson, M. J., Dimes, J. P., Whitbread,
740 A. M., Hunt, J., Rees, H. van, McClelland, T., Carberry, P. S., Hargreaves, J. N. G., MacLeod, N., McDonald, C., Harsdorf, J., Wedgwood, S., and Keating, B. A.: APSIM – Evolution towards a new generation of agricultural systems simulation, Environmental Modelling & Software, 62, 327–350, https://doi.org/10.1016/j.envsoft.2014.07.009, 2014.

Hong, M., Zhang, Y., Braun, R. C., and Bremer, D. J.: Simulations of nitrous oxide emissions and global warming potential in a C4 turfgrass system using process-based models, European journal of agronomy : the journal of the European Society for
745 Agronomy, 142, 126668, https://doi.org/10.1016/j.eja.2022.126668, 2023.

Houska, T., Kraus, D., Kiese, R., and Breuer, L.: Constraining a complex biogeochemical model for CO_2 and N_2O emission simulations from various land uses by model–data fusion, Biogeosciences, 14, 3487–3508, https://doi.org/10.5194/bg-14-3487-2017, 2017.

Huete, A., Didan, K., Miura, T., Rodriguez, E. ., Gao, X., and Ferreira, L. .: Overview of the radiometric and biophysical
750 performance of the MODIS vegetation indices, Remote Sensing of Environment, 83, 195–213, https://doi.org/10.1016/S0034-4257(02)00096-2, 2002.

Indexdatabase: https://www.indexdatabase.de/, last accessed 19.12.2024.

Isbell, F., Tilman, D., Polasky, S., Binder, S., and Hawthorne, P.: Low biodiversity state persists two decades after cessation of nutrient enrichment, Ecology letters, 16, 454–460, https://doi.org/10.1111/ele.12066, 2013.

755 Jägermeyr, J., Müller, C., Ruane, A. C., Elliott, J., Balkovic, J., Castillo, O., Faye, B., Foster, I., Folberth, C., Franke, J. A., Fuchs, K., Guarin, J. R., Heinke, J., Hoogenboom, G., Iizumi, T., Jain, A. K., Kelly, D., Khabarov, N., Lange, S., Lin, T.-S., Liu, W., Mialyk, O., Minoli, S., Moyer, E. J., Okada, M., Phillips, M., Porter, C., Rabin, S. S., Scheer, C., Schneider, J. M., Schyns, J. F., Skalsky, R., Smerald, A., Stella, T., Stephens, H., Webber, H., Zabel, F., and Rosenzweig, C.: Climate impacts on global agriculture emerge earlier in new generation of climate and crop models, Nature food, 2, 873–885,
760 https://doi.org/10.1038/s43016-021-00400-y, 2021.

Kaim, A., Schmitt, T.M., Annuth, S.H., Haensel, M., and Koellner, T. An agent-based model to simulate field-specific nitrogen fertilizer applications in grasslands, European Journal of Agronomy, 165, 127539, https://doi.org/10.1016/j.eja.2025.127539, 2025.

Kiese, R., Heinzeller, C., Werner, C., Wochele, S., Grote, R., and Butterbach-Bahl, K.: Quantification of nitrate leaching from
765 German forest ecosystems by use of a process oriented biogeochemical model, Environmental pollution (Barking, Essex : 1987), 159, 3204–3214, https://doi.org/10.1016/j.envpol.2011.05.004, 2011.

Kiese, R., Fersch, B., Baessler, C., Brosy, C., Butterbach-Bahl, K., Chwala, C., Dannenmann, M., Fu, J., Gasche, R., Grote, R., Jahn, C., Klatt, J., Kunstmann, H., Mauder, M., Rödiger, T., Smiatek, G., Soltani, M., Steinbrecher, R., Völksch, I., Werhahn, J., Wolf, B., Zeeman, M., and Schmid, H. P.: The TERENO Pre-Alpine Observatory: Integrating Meteorological, Hydrological, and Biogeochemical Measurements and Modeling, Vadose Zone Journal, 17, 180060, https://doi.org/10.2136/vzj2018.03.0060, 2018.

Kottek, M., Grieser, J., Beck, C., Rudolf, B., and Rubel, F.: World Map of the Köppen-Geiger climate classification updated, Meteorologische Zeitschrift, 15, 259–263, https://doi.org/10.1127/0941-2948/2006/0130, 2006.

Kraus, D., Weller, S., Klatt, S., Haas, E., Wassmann, R., Kiese, R., and Butterbach-Bahl, K.: A new LandscapeDNDC biogeochemical module to predict CH4 and N2O emissions from lowland rice and upland cropping systems, Plant and Soil, 386, 125–149, https://doi.org/10.1007/s11104-014-2255-x, 2015.

Kraus, D., Weller, S., Klatt, S., Santabárbara, I., Haas, E., Wassmann, R., Werner, C., Kiese, R., and Butterbach-Bahl, K.: How well can we assess impacts of agricultural land management changes on the total greenhouse gas balance (CO2, CH4 and N2O) of tropical rice-cropping systems with a biogeochemical model?, Agriculture, Ecosystems & Environment, 224, 104–115, https://doi.org/10.1016/j.agee.2016.03.037, 2016.

Krauth, J. R and Thomas, S. G: The tasseled Cap - A Graphic Description of the Spectral-Temporal Development of Agricultural Crops as Seen by LANDSAT, in: Proceedings of the Symposium on Machine Processing of Remotely Sensed Data, 1976.

LDBV (Landesamt für Digitalisierung, Breitband und Vermessung): Bodenschätzung (BOSCH), 2018.

LfL (Landesamt für Landwirtschaft): Leitfaden für die Düngung von Acker- und Grünland: Gelbes Heft, 2018.

Liebermann, R., Breuer, L., Houska, T., Kraus, D., Moser, G., and Kraft, P.: Simulating Long-Term Development of Greenhouse Gas Emissions, Plant Biomass, and Soil Moisture of a Temperate Grassland Ecosystem under Elevated Atmospheric CO2, Agronomy, 10, 50, https://doi.org/10.3390/agronomy10010050, 2020.

Mayel, S., Jarrah, M., and Kuka, K.: How does grassland management affect physical and biochemical properties of temperate grassland soils? A review study, GRASS AND FORAGE SCIENCE, 76, 215–244, https://doi.org/10.1111/gfs.12512, 2021.

Molina-Herrera, S., Haas, E., Klatt, S., Kraus, D., Augustin, J., Magliulo, V., Tallec, T., Ceschia, E., Ammann, C., Loubet, B., Skiba, U., Jones, S., Brümmer, C., Butterbach-Bahl, K., and Kiese, R.: A modeling study on mitigation of N2O emissions and NO3 leaching at different agricultural sites across Europe using LandscapeDNDC, The Science of the total environment, 553, 128–140, https://doi.org/10.1016/j.scitotenv.2015.12.099, 2016.

Muro, J., Linstädter, A., Magdon, P., Wöllauer, S., Männer, F. A., Schwarz, L.-M., Ghazaryan, G., Schultz, J., Malenovský, Z., and Dubovyk, O.: Predicting plant biomass and species richness in temperate grasslands across regions, time, and land management with remote sensing and deep learning, Remote Sensing of Environment, 282, 113262, https://doi.org/10.1016/j.rse.2022.113262, 2022.

Murphy, D. J., Murphy, M. D., O'Brien, B., and O'Donovan, M.: A Review of Precision Technologies for Optimising Pasture Measurement on Irish Grassland, Agriculture, 11, 600, https://doi.org/10.3390/agriculture11070600, 2021.

Neyret, M., Fischer, M., Allan, E., Hölzel, N., Klaus, V. H., Kleinebecker, T., Krauss, J., Le Provost, G., Peter, S., Schenk, N., Simons, N. K., Plas, F. van der, Binkenstein, J., Börschig, C., Jung, K., Prati, D., Schäfer, D., Schäfer, M., Schöning, I., Schrumpf, M., Tschapka, M., Westphal, C., and Manning, P.: Assessing the impact of grassland management on landscape multifunctionality, Ecosystem Services, 52, 101366, https://doi.org/10.1016/j.ecoser.2021.101366, 2021.

805   Parton, W. J., Hartman, M., Ojima, D., and Schimel, D.: DAYCENT and its land surface submodel: description and testing, Global and Planetary Change, 19, 35–48, https://doi.org/10.1016/S0921-8181(98)00040-X, 1998.

Petersen, K., Kraus, D., Calanca, P., Semenov, M. A., Butterbach-Bahl, K., and Kiese, R.: Dynamic simulation of management events for assessing impacts of climate change on pre-alpine grassland productivity, European journal of agronomy : the journal of the European Society for Agronomy, 128, None, https://doi.org/10.1016/j.eja.2021.126306, 2021.

810   Poschlod, B., Willkofer, F., and Ludwig, R.: Impact of Climate Change on the Hydrological Regimes in Bavaria, Water, 12, 1599, https://doi.org/10.3390/w12061599, 2020.

Reinermann, S., Asam, S., and Kuenzer, C.: Remote Sensing of Grassland Production and Management—A Review, Remote Sensing, 12, 1949, https://doi.org/10.3390/rs12121949, 2020.

Reinermann, S., Gessner, U., Asam, S., Ullmann, T., Schucknecht, A., and Kuenzer, C.: Detection of Grassland Mowing
815   Events for Germany by Combining Sentinel-1 and Sentinel-2 Time Series, Remote Sensing, 14, 1647, https://doi.org/10.3390/rs14071647, 2022.

Reinermann, S., Asam, S., Gessner, U., Ullmann, T., and Kuenzer, C.: Multi-annual grassland mowing dynamics in Germany: Spatio-temporal patterns and the influence of climate, topographic and socio-political conditions, Frontiers in Environmental Science, 11, [1040551], https://doi.org/10.3389/fenvs.2023.1040551, 2023.

820   Reis Martins, M. dos, Ammann, C., Boos, C., Calanca, P., Kiese, R., Wolf, B., and Keel, S. G.: Reducing N fertilization in the framework of the European Farm to Fork strategy under global change: Impacts on yields, N2O emissions and N leaching of temperate grasslands in the Alpine region, Agricultural Systems, 219, 104036, https://doi.org/10.1016/j.agsy.2024.104036, 2024.

Richter, F., Jan, P., El Benni, N., Lüscher, A., Buchmann, N., and Klaus, V. H.: A guide to assess and value ecosystem services
825   of grasslands, Ecosystem Services, 52, 101376, https://doi.org/10.1016/j.ecoser.2021.101376, 2021.

Riedo, M., Grub, A., Rosset, M., and Fuhrer, J.: A pasture simulation model for dry matter production, and fluxes of carbon, nitrogen, water and energy, Ecological Modelling, 105, 141–183, https://doi.org/10.1016/S0304-3800(97)00110-5, 1998.

Riedo, M., Gyalistras, D., and Fuhrer, J.: Net primary production and carbon stocks in differently managed grasslands: simulation of site-specific sensitivity to an increase in atmospheric CO2 and to climate change, Ecological Modelling, 134,
830   207–227, https://doi.org/10.1016/S0304-3800(00)00356-2, 2000.

Rolinski, S., Müller, C., Heinke, J., Weindl, I., Biewald, A., Bodirsky, B. L., Bondeau, A., Boons-Prins, E. R., Bouwman, A. F., Leffelaar, P. A., Roller, J. A. te, Schaphoff, S., and Thonicke, K.: Modeling vegetation and carbon dynamics of managed grasslands at the global scale with LPJmL 3.6, Geoscientific Model Development, 11, 429–451, https://doi.org/10.5194/gmd-11-429-2018, 2018.

835   Rouder, J. N., Speckman, P. L., Sun, D., Morey, R. D., and Iverson, G.: Bayesian t tests for accepting and rejecting the null hypothesis, Psychonomic bulletin & review, 16, 225–237, https://doi.org/10.3758/PBR.16.2.225, 2009.

Savage, J., Woodcock, B. A., Bullock, J. M., Nowakowski, M., Tallowin, J. R. B., and Pywell, R. F.: Management to Support Multiple Ecosystem Services from Productive Grasslands, Sustainability, 13, 6263, https://doi.org/10.3390/su13116263, 2021.

Schaphoff, S., Bloh, W. von, Rammig, A., Thonicke, K., Biemans, H., Forkel, M., Gerten, D., Heinke, J., Jägermeyr, J.,
840   Knauer, J., Langerwisch, F., Lucht, W., Müller, C., Rolinski, S., and Waha, K.: LPJmL4 – a dynamic global vegetation model

with managed land – Part 1: Model description, Geoscientific Model Development, 11, 1343–1375, https://doi.org/10.5194/gmd-11-1343-2018, 2018.

Schmitt, T., Haensel, M., Kaim, A., Lee, H., Reinermann, S., and Koellner, T.: Recreation and its synergies and trade-offs with other ecosystem services of Alpine and pre-Alpine grasslands, Regional Environmental Change, 24, https://doi.org/10.1007/s10113-024-02213-8, 2024.

Schoof, N., Luick, R., Ackermann, A., Baum, S., Böhner, H., Röder, N., Rudolph, S., Schmidt, T. G., Hötker, H., and Jeromin, H. (Eds.): Auswirkungen der neuen Rahmenbedingungen der Gemeinsamen Agrarpolitik auf die Grünland-bezogene Biodiversität, 2nd ed., Bundesamt für Naturschutz, Bonn - Bad Godesberg, Germany, 2020a.

Schoof, N., Luick, R., Jürgens, K., and Jones, G.: Dairies in Germany: Key Factors for Grassland Conservation?, Sustainability, 12, 4139, https://doi.org/10.3390/su12104139, 2020b.

Schoof, N., Luick, R., Beaufoy, G., Jones, G., Einarsson, P., Ruiz, J., Stefanova, V., Fuchs, D., Windmaißer, T., Hötker, H., Jeromin, H., Nickel, H., Schumacher, J., and Ukhanova, M. (Eds.): Grünlandschutz in Deutschland: Treiber der Biodiversität, Einfluss von Agrarumwelt- und Klimamaßnahmen, Ordnungsrecht, Molkereiwirtschaft und Auswirkungen der Klima- und Energiepolitik, 2nd ed., Bundesamt für Naturschutz, Bonn - Bad Godesberg, Germany, 2020c.

Schucknecht, A., Krämer, A., Asam, S., Mejia-Aguilar, A., Garcia-Franco, N., Schuchardt, M. A., Jentsch, A., and Kiese, R.: Vegetation traits of pre-Alpine grasslands in southern Germany, Scientific data, 7, 316, https://doi.org/10.1038/s41597-020-00651-7, 2020.

Schucknecht, A., Reinermann, S., Kiese, R., and Bayer, A.: Towards The Retrieval of Plant Traits in Grasslands with Hyperspectral Enmap Data, in: 2023 13th Workshop on Hyperspectral Imaging and Signal Processing: Evolution in Remote Sensing (WHISPERS), 1–5, https://doi.org/10.1109/WHISPERS61460.2023.10431190, 2023.

Schwieder, M., Buddeberg, M., Kowalski, K., Pfoch, K., Bartsch, J., Bach, H., Pickert, J., and Hostert, P.: Estimating Grassland Parameters from Sentinel-2: A Model Comparison Study, PFG – Journal of Photogrammetry, Remote Sensing and Geoinformation Science, 88, 379–390, https://doi.org/10.1007/s41064-020-00120-1, 2020.

Smit, H. J., Metzger, M. J., and Ewert, F.: Spatial distribution of grassland productivity and land use in Europe, Agricultural Systems, 98, 208–219, https://doi.org/10.1016/j.agsy.2008.07.004, 2008.

Socher, S. A., Prati, D., Boch, S., Müller, J., Klaus, V. H., Hölzel, N., and Fischer, M.: Direct and productivity-mediated indirect effects of fertilization, mowing and grazing on grassland species richness, Journal of Ecology, 100, 1391–1399, https://doi.org/10.1111/j.1365-2745.2012.02020.x, 2012.

Soussana, J. and Lüscher, A.: Temperate grasslands and global atmospheric change: a review, GRASS AND FORAGE SCIENCE, 62, 127–134, https://doi.org/10.1111/j.1365-2494.2007.00577.x, 2007.

Statistisches Bundesamt, S.: Land- und Forstwirtschaft, Fischerei - Bodennutzung der Betriebe: Fachserie 3, 2023.

Ullah, S., Si, Y., Schlerf, M., Skidmore, A. K., Shafique, M., and Iqbal, I. A.: Estimation of grassland biomass and nitrogen using MERIS data, International Journal of Applied Earth Observation and Geoinformation, 19, 196–204, https://doi.org/10.1016/j.jag.2012.05.008, 2012.

LfU (Landesamt für Umwelt): private communication, 2020.

Vuichard, N., Soussana, J., Ciais, P., Viovy, N., Ammann, C., Calanca, P., Clifton–Brown, J., Fuhrer, J., Jones, M., and Martin, C.: Estimating the greenhouse gas fluxes of European grasslands with a process–based model: 1. Model evaluation from in situ measurements, Global Biogeochemical Cycles, 21, https://doi.org/10.1029/2005GB002611, 2007.

White, R. P., Murray, S., and Rohweder, M.: Pilot analysis of global ecosystems - Grassland ecosystems, World Resources Institute, Washington DC, USA, 2000.

Willkofer, F., Wood, R. R., Trentini, F. von, Weismüller, J., Poschlod, B., and Ludwig, R.: A Holistic Modelling Approach for the Estimation of Return Levels of Peak Flows in Bavaria, Water, 12, 2349, https://doi.org/10.3390/w12092349, 2020.

Wu, H., An, S., Meng, B., Chen, X., Li, F., and Ren, S.: Retrieval of grassland aboveground biomass across three ecoregions in China during the past two decades using satellite remote sensing technology and machine learning algorithms, International Journal of Applied Earth Observation and Geoinformation, 130, 103925, https://doi.org/10.1016/j.jag.2024.103925, 2024.

Yao, Y. and Ren, H.: Estimation of grassland aboveground biomass in northern China based on topography-climate-remote sensing data, Ecological Indicators, 165, 112230, https://doi.org/10.1016/j.ecolind.2024.112230, 2024.

Zhang, H., Fan, J., Wang, J., Cao, W., and Harris, W.: Spatial and temporal variability of grassland yield and its response to climate change and anthropogenic activities on the Tibetan Plateau from 1988 to 2013, Ecological Indicators, 95, 141–151, https://doi.org/10.1016/j.ecolind.2018.05.088, 2018.

Financial Support

Funding was provided by the Federal Ministry of Education and Research (BMBF) via the BonaRes project SUSALPS (031B0027A, 031B0027F).

Data availability

The mowing frequency and date of first mowing event are available at the EOC Geoservice (https://geoservice.dlr.de/web/datasets/agriculture), the mowing dates upon request.

Competing interests

The authors declare that they have no conflict of interest.

Appendix

Table A1: Bavarian reference values for farmer's grassland yield from field measurements for meadows and mowing pastures of different use types and yield levels.

Use type and intensity	Farmer 's yield [t ha-1 a-1]		
	Low	Medium	High

Meadow 1 cut	2.8	3.4	4.0
Meadow 2 cuts	3.9	4.7	5.5
Meadow 3 cuts	5.6	6.8	8.0
Meadow 4 cuts	6.3	7.7	9.0
Meadow 5 cuts	7.7	9.4	11.0
Meadow 6 cuts	8.4	10.2	12.0
Mowing pasture extensive, 20 % pasture	4.8	5.9	6.9
Mowing pasture medium intensive, 20 % pasture	6.9	8.3	9.8
Mowing pasture intensive, 20 % pasture	7.7	9.4	11.0
Mowing pasture extensive, 60 % pasture	4.7	5.7	6.7
Mowing pasture medium intensive, 60 % pasture	5.7	6.9	8.1
Mowing pasture intensive, 60 % pasture	6.6	8.0	9.4

Table A2: Grassland yield level (/grazing intensity) according to grassland index or slope (/stocking rate) which is used as input for table A1.

Yield level / grazing intensity	Grassland index (GI)	Slope	Stocking rate (SR)
Low	$0 \leq GI \leq 33$	$50\ \% \leq slope$	$0 = SR \leq 1.5$
Medium	$33 < GI \leq 67$	$25\ \% \leq slope < 50\ \%$	$1.5 < SR \leq 3$
High	$67 < GI \leq 100$	$0\ \% \leq slope < 25\ \%$	$3 < SR$

905
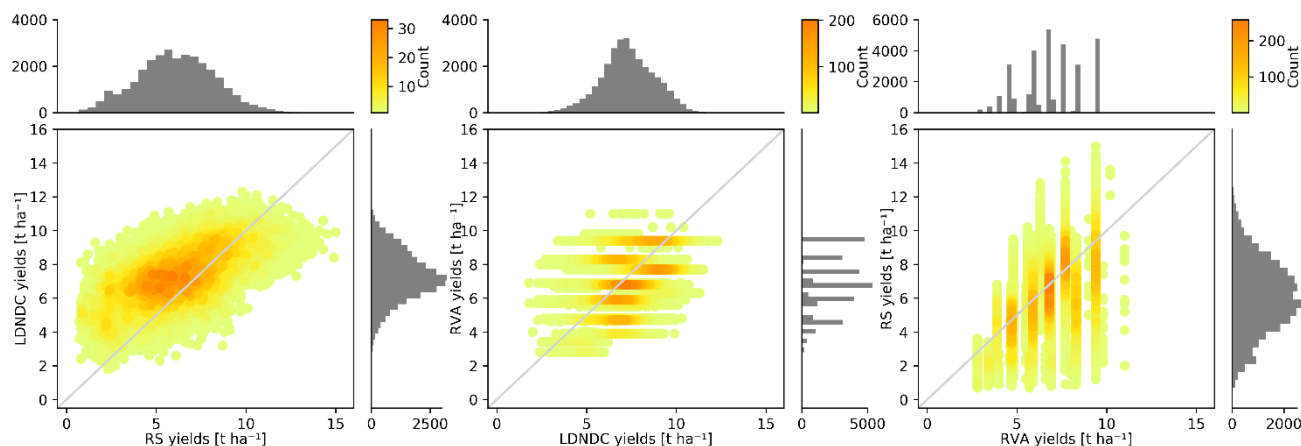
Figure A1: Pairwise relationships of parcel-level annual yield estimates for meadows and mowing pastures in the study area in 2019 based on Remote Sensing, LandscapeDNDC simulations and the reference values approach with histograms of the three modelling approaches. Note, that the colour ranges differ for the subplots.

910
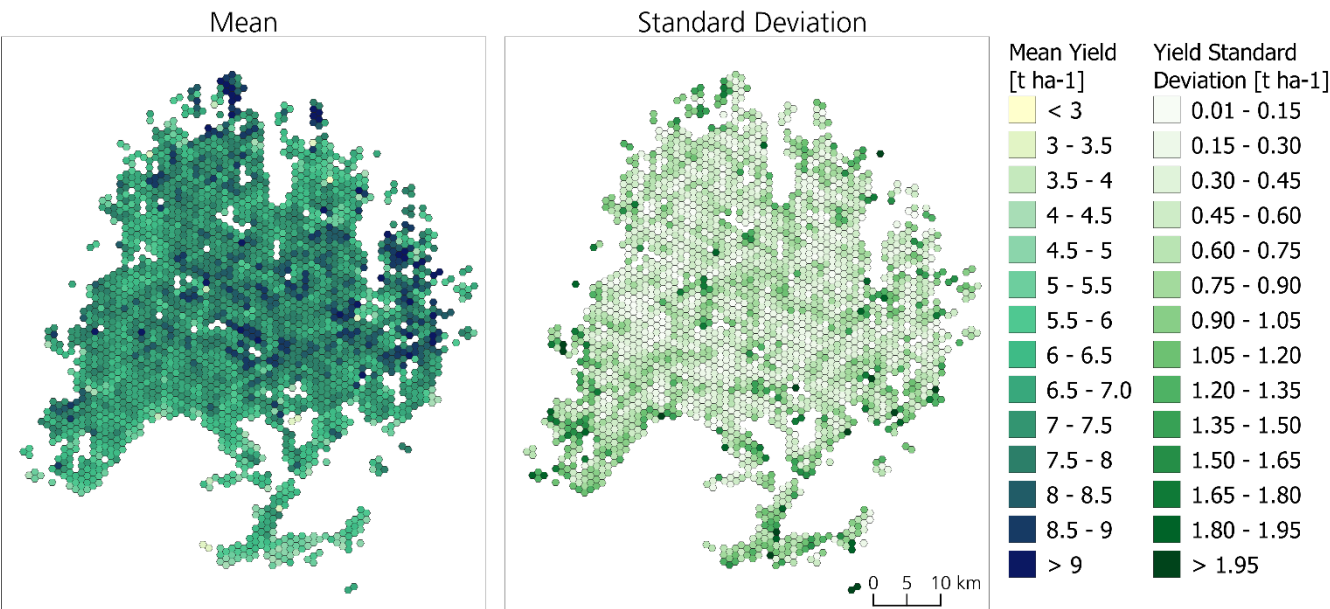
Figure A2: Mean and standard deviation of spatially aggregated (hexagon diagonal length of 1km) annual yield estimates for meadows and mowing pastures in the study area in 2019 of all three approaches. Hexagons for which the grassland area is smaller than 1 ha are not shown.
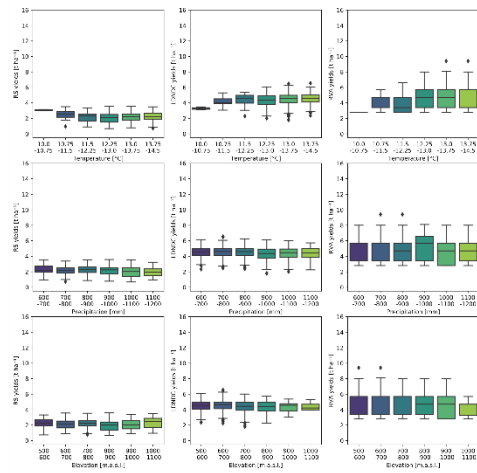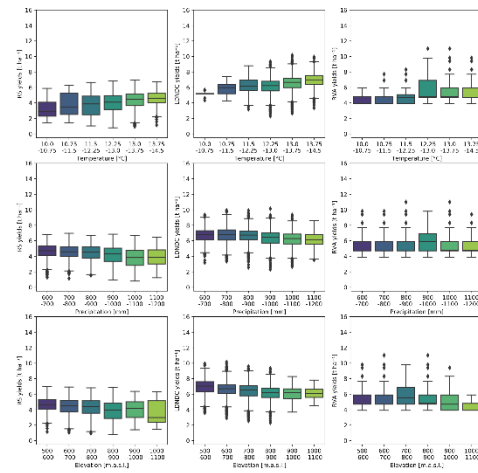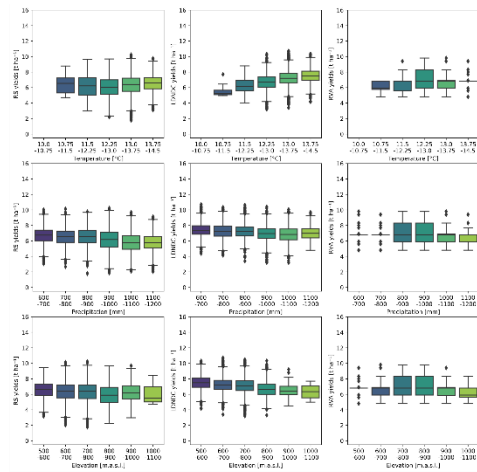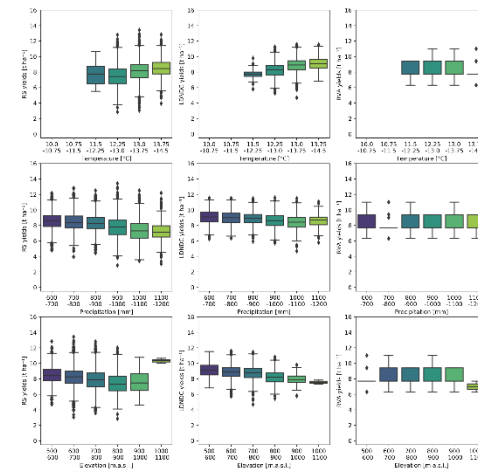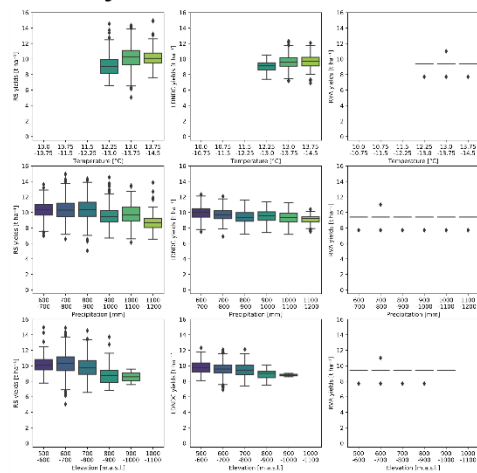
Figure A3: Aggregation statistics of estimated annual yields based on the three models per temperature, precipitation and elevation classes for grasslands mown one to five times per year.