~~Measuring~~Benefits of a robotic chamber system for determining evapotranspiration ~~on~~in an ~~eroded~~erosion affected, heterogeneous cropland ~~by an automated and mobile chamber system: gap filling strategies and impact of soil type and topsoil removal~~

Adrian Dahlmann[1], Mathias Hoffmann[1], Gernot Verch[2], Marten Schmidt[1], Michael Sommer[3,4], Jürgen Augustin[1], Maren Dubbert[1]

[1]Isotope Biogeochemistry and Gas Fluxes, Leibniz Centre for Agricultural Landscape Research, Müncheberg, 15374, Germany
~~[2]Experimental Station Dedelow,~~[2] Experimental Infrastructure Platform (EIP), Leibniz Centre for Agricultural and Landscape Research, Prenzlau, 17291, Germany
[3]Landscape Pedology, Leibniz Centre for Agricultural Landscape Research, Müncheberg, 15374, Germany
[4] Institute of Geography and Environmental Science, University of Potsdam, 14476, Potsdam, Germany

Correspondence to: Adrian Dahlmann (adrian.dahlmann@zalf.de)

Abstract. In light of the ongoing global climate crisis and related increases in extreme hydrological events, it is ~~increasingly~~ crucial to assess ecosystem resilience and - in agricultural systems - to ensure sustainable management and food security. For that purpose, comprehensive understanding of ecosystem water cycle budgets and spatio-temporal dynamics are indispensable. Evapotranspiration (ET) plays a pivotal role returning up to 90 % of incoming ~~(and throughout)~~ precipitation back to the atmosphere. Here, we studied impacts of soil types and management on an ~~agroecosystems main water budgets~~agroecosystem's seasonal cumulative ET (ET_{sum}) and agronomic water~~-~~-use efficiencies (WUE_{agro}, dry matter per unit of water used by the crop). To do so, a plot experiment with winter rye (September 17, 2020 to June 30, 2021) was conducted at an eroded cropland which is located in the hilly and dry ground moraine landscape of the Uckermark region in NE Germany. Along the experimental plot (110 m x 16 m), two closed chambers were mounted on a robotic gantry crane ~~mounted mobile and automated two chamber~~ system (FluxCrane as part of the AgroFlux platform ~~within the CarboZALF-D research site) was~~) and used to continuously determine ~~evapotranspiration for the first time~~ET. Three soil types representing the full soil erosion gradient related to the hummocky ground moraine landscape (extremely eroded: Calcaric Regosol, strongly eroded: Nudiargic Luvisol, non-eroded: Calcic Luvisol) and additional top-soil ~~manipulation~~dilution (topsoil removal and subsoil admixture) were investigated (randomized block design, 3 replicates per treatment). Five different gap-filling approaches were used and compared in light of their potential for reliable ~~water budgets~~ET_{sum} over the entire crop ~~growth~~cultivation period as well as to reproduce short-term (day-to-day, diurnal) water~~-~~-flux dynamics. ~~The best calibration performance was achieved with approaches based on~~While machine learning~~,~~ approaches such as support vector ~~machine~~machines (SVM) and artificial neural networks (with Bayesian regularization; ANN_BR~~), while especially~~) generally performed well during calibration, SVM ~~yielded in best predictions~~also provided a satisfactory prediction of measured ET during validation~~.~~ (k-fold cross validation, k = 5).

We found significant, major differences in dry biomass (DM) and ~~minor~~small trends in ~~evapotranspiration~~ET_{sum} between soil types, resulting in different ~~water use efficiencies (~~WUE_{agro}~~).~~. The extremely eroded Calcaric Regosol ~~(extremely eroded) shows a maximum of around 37~~showed an up to 46 % lower ~~evapotranspiration~~ET_{sum} and ~~a maximum of around 52~~up to 54 % lower ~~water use efficiency (~~WUE_{agro}~~)~~ compared to the non-eroded Calcic Luvisol. The key period contributing to ~~~~70 % of ~~overall ET of the entire growth period~~ET_{sum} was from the beginning of stem elongation in April ~~until June (~~to harvest~~),~~ ~~however~~ in June. However, differences in the ~~overall ET budget (~~ET_{sum}~~)~~ between soil types and ~~manipulation~~topsoil dilution resulted predominantly from small differences between the treatments ~~over~~throughout the ~~entire growth period~~cultivation, rather than only during this short period of time.

1 Introduction

Only 12 % of the world's land area is suitable for food and fiber production due to its highly productive soils (Blum, 2013). Much of this land is already in use to ensure food security, mandated by a still growing human population paired with the ongoing climate crisis (Searchinger et al. 2018). Worldwide, land area is largely affected by soil degradation (Jie et al., 2002) and agriculture is closely related, since at least six degradation processes (e.g. erosion or compaction) are associated with it (Louwagie et al., 2011). In hummocky landscapes, erosion and associated topsoil dilution caused by, e.g. wind, water or tillage, affects the crop yields (Bakker et al. 2007; Biggelaar et al. 2003). In addition, weaker rootability on eroded soils suggests a higher susceptibility towards droughts (Schneider and Don 2019). However, methodologically studying the influence of small scale soil heterogeneity (e.g. soil erosion) and land use (e.g. soil management) on the dynamics of the water balance (especially evapotranspiration (ET)) separately has been challenging. The effect of both factors can be significantly different with complex interactions, e.g. soil erosion can lead to differences in soil water storage capacity and management affects soil organic matter and water retention (Bakker et al. 2007; Biggelaar et al. 2003) Thus, a separate response analysis is an indispensable prerequisite for the development of site-specific land use procedures adapted to the changing climate conditions. Moreover, the climate crisis is affecting the amount and spatio-temporal distribution of precipitation worldwide, leading to more frequent and stronger precipitation events in high-precipitation regions (e.g. increase of 10 – 40 % in northern Europe; DWD, 2019) and fewer and weaker events in low-precipitation regions (e.g. up to 20 % decrease in the Mediterranean region and southeastern Europe; Trenberth, 2011). In Germany, annual precipitation is more than 800 mm in most regions of west and south Germany but only 400 - 500 mm y^{-1} in the northeast (e.g. areas in Brandenburg and Mecklenburg-Western Pomerania; Schappert, 2018). Here, dry hydrological conditions and erosion shaped landscapes meet. As crop yields and related crop productivity depend on various factors such as soil properties or water availability, such agriculturally used precipitation limited regions could face increasing problems.

ET describes the total amount of water that evaporates from a given area and is thus defined as the sum of soil evaporation (E), transpiration (T) and interception evaporation (Fohrer et al., 2016; Rothfuss et al., 2021). Generally, ET is one of the most important components of the hydrological cycle in terrestrial ecosystems, accounting for up to 100 % of ingoing

precipitation (Hanson 1991). With a share of up to 90 %, it is largely dominated by T in most terrestrial ecosystems, indicating that terrestrial vegetation is a dominant force in the global water cycle (Jasechko et al. 2013). Due to the expected increasing dependency of a systems productivity on sufficient water supply with an accelerating climate crisis, quantifying the ET plays an important role to achieve a process-based understanding of the mitigation potential of different crops to drought in the future and to, e.g., establish a more efficient supplemental irrigation. Moreover, there is a tight link of carbon and water cycling in precipitation limited systems because water loss by ET and the water use efficiency of a system can largely define its productivity (Tallec et al., 2013).

A particular challenge in current ET research is combining high frequency with multi treatment approaches. At the field scale for example, eddy covariance systems provide high frequency estimates of ET of a homogeneous system while currently dominant manual chamber approaches are able to precisely capture multi treatment effects (<1m²) on ET at the plot scale, but lack the high frequency. In this regard, modern automated chamber systems allow a combination of high frequency measurements and thus high temporal resolution with multi-treatment observation. They provide the unique opportunity to test advanced gap-filling strategies, able to reproduce not only seasonal cumulative ET (ET_{sum}) but day to day and diurnal variability in ET. Modern gap-filling methods (e.g. artificial intelligence and neural network approaches) have previously been limited to eddy covariance measurements. Coupling such advanced gap-filling strategies with modern automatic chamber systems might be an ideal fusion of measurement frequency and the ability to capture treatment effects like small scale soil differences (Falge et al. 2001a; KIŞI and ÇIMEN 2009). AgroFlux – a newly developed sensor platform centered around closed chambers mounted on an robotic gantry crane (FluxCrane) – was initially built to capture the effect of soil type and management on GHG emissions and in particular CO_2 fluxes with high spatial and temporal resolution (Vaidya et al. 2021). The adaption of the system to measure ET provided us with the opportunity to analyze stand scale ET fluxes including the development of a data analysis tool for measured ET-fluxes, and test different gap-filling strategies. We tested five different gap-filling strategies including basic statistic and advanced approaches including machine learning approaches. During the cultivation period of winter rye from mid-September 2020 to the end of June 2021, ET and relevant environmental and plant growth parameters were measured to identify the corresponding drivers of crop ET and productivity. The FluxCrane system covers a field where three different soil types are present, which reflect the erosion gradient typical for the hillside of the hummocky ground moraine landscape of northeast Germany. This made it possible to evaluate the impact of soil type as well as soil management on ET_{sum}, seasonal development and agronomic water use efficiency (WUE_{agro}; dry matter per unit of water used by the crop).

In the following we will examine i) soil type and top soil dilution effects on crop yield, ET_{sum} and WUE_{agro}, ii) the spatio-temporal variability of ET fluxes over the growing season, and iii) the suitability of various gap-filling strategies. The paper's aim is to establish an approach that would provide reliable predictions of ET fluxes both in terms of ET_{sum} as well as diurnal trends of ET fluxes. We hypothesize that: i) eroded soils and top-soil dilution lead to decreased ET controlled by weaker plant growth, ii) WUE_{agro} declines from least to most eroded soil type and with top soil dilution; iii) the automated, continuous FluxCrane measurements result in unique insights into small scale dynamics such as night time ET fluxes and ET fluxes during

the non-growing season. Here, we hypothesize, that iv) the uncommonly (compared to manual chamber systems) large data set allows for a robust use of gap-filling strategies based on machine learning. We envisage that this will greatly improve ET_{sum} and subsequently WUE_{agro} based on automated closed chamber systems.

2 Material and Methods

2.1 Study Site and experimental design

More than 70% of the global ice-free land area is directly influenced by humans in distinct ways (Shukla et al., 2019). In fact, only 12% of the world's land area is suitable for food and fiber production due to its highly productive soils, 65% of which are located in temperate and Mediterranean climatic zones and only 35% in tropical ones (Blum, 2013). Due to a still growing human population paired with the ongoing climate crisis, much of this land is already in use worldwide to ensure food security. Additionally, worldwide land area is largely affected by soil degradation (Jie et al., 2002). According to the European Union, agriculture is closely related to soil degradation, since at least six degradation processes (e.g. erosion or compaction) are associated with it (Louwagie et al., 2011). In hummocky landscapes, erosion and associated topsoil removal caused by different elements (e.g. wind, water or tillage) affects the crop yields (Bakker et al. 2007; Biggelaar et al. 2003). Moreover, the global climate crisis, for example, is affecting the amount and spatio-temporal distribution of precipitation worldwide, leading to more frequent and stronger events in high precipitation regions and fewer and weaker events in low precipitation regions (Trenberth, 2011). This clear dichotomy is shown, with increases of 10 – 40 % of precipitation mainly in northern Europe and decreases of up to 20 % in the Mediterranean region and southeastern Europe (DWD, 2019). In Germany, annual precipitation budgets of more than 800 mm occur in most regions in west and south Germany but only 400 – 500 mm y^{-1} in the northeast (e.g. areas in Brandenburg and Mecklenburg-Western Pomerania; Schappert, 2018). As crop yields and related crop productivity depend on various factors such as soil properties or water availability, such agricultural used arid regions could face increasing problems.

The AgroFlux experimental platform is located in Brandenburg, a federal state in northeast Germany including , near Dedelow within the Uckermark region (53° 23´ N, 13° 47´ E; ~50-60 m a.s.l). Brandenburg, which includes some of the driest regions in entire Germany, uses 48.6% or about 1.44 million hectares of its area for agriculture (Amt für Statistik Berlin-Brandenburg, 2020). Brandenburg is located in the continental characterized climate zone and has a water deficit of about 150 mm over the growing season (Wessolek and Asseng 2006). The focus of agriculture in Brandenburg is on grain production, which faces a variety of challenges due to increasingly dry conditionsIt is located in the continental climate zone and has a water deficit of about 150 mm during the growing season (Wessolek and Asseng 2006). The long-term (1991 to 2020; ZALF) mean annual air temperature in this region is 8.8°C with a mean annual precipitation and potential evapotranspiration of 467 mm and 637 mm, respectively (ZALF research station, Dedelow). The focus of agriculture in Brandenburg is on grain production, which

faces a variety of challenges due to increasingly dry conditions during the main growing season (Amt für Statistik Berlin-Brandenburg, 2020). The Uckermark region is the most ~~fertile~~productive region for agriculture ~~in~~within Brandenburg. It is shaped by glaciation with a hilly to flat-wavy ground moraine landscape whose soils are strongly influenced by soil erosion (Nudiargic Luvisol, Calcaric Regosols, Colluvic Regosols) as well as redoximorphic soils (Stagno-, Gleysols) (MLUK, 2020). The strong soil heterogeneity and ongoing soil erosion, mainly by tillage, has a great influence on the ~~fertility~~productivity of the cultivated areas ~~(Sommer et al. 2016)~~(Sommer et al. 2016)~~. In combination with climate crisis, it is proving difficult to develop land use methods that allow reliable arable farming under these changing conditions.~~

~~There is a tight link of carbon and water cycling in such precipitation limited systems. In particular, on eroded soils with limited rootability it becomes highly important because water loss by evapotranspiration (ET) and the water use efficiency of a system (WUE) can largely define its productivity (Tallec et al., 2013). To date, methodologically studying the influence of small scale soil heterogeneity (e.g. soil erosion) and land use (e.g. soil management) on the dynamics of the water balance (especially ET) separately is challenging. However, the effect of both factors can be significantly different with complex interaction (Bakker et al. 2007; Biggelaar et al. 2003), making a separate response analysis an indispensable prerequisite for the development of site-specific land use procedures adapted to the changing climate conditions. ET describes the total amount of water that evaporates from a given area and is thus defined as the sum of evaporation (E), transpiration (Tr), and interception evaporation (Fohrer et al., 2016; Rothfuss et al., 2021). Generally, ET is one of the most important components of the hydrological cycle in terrestrial ecosystems accounting for up to 100 % of ingoing precipitation (Hanson 1991) and is dominated by transpiration in most terrestrial ecosystems with a share of up to 90 %, indicating that terrestrial vegetation is a dominant force in the global water cycle (Jasechko et al. 2013). Due to the expected increased dependency of a systems productivity on sufficient water supply with ongoing climate change, quantifying the ET plays an important role to achieve a process based understanding~~. Today,~~the mitigation potential of different crops to drought in the future and to, e.g., establish a more efficient supplemental irrigation.~~

~~Methods to quantify evapotranspiration have been developed from disciplines ranging from plant physiology to hydrology and meteorology and range from plot scale to global estimations. To date, each community and method refers to a specific measurement scale, contains its own uncertainties and comparability between methods is rather low (Drexler et al., 2004; Hamel et al., 2015). A particular challenge in current evapotranspiration research apart from standardized approaches is combining high frequency with multi treatment approaches. At the plot scale for example, eddy covariance systems deliver high frequency estimates of evapotranspiration of a homogeneous system (few hundred to millions of square meters) while currently prevailing manual chamber approaches are able to precisely capture multi treatment effects (<1m²) on evapotranspiration but lack temporal resolution. In this regard, modern automated chamber systems allowing a combination of high temporal with multi treatment observation provide the unique opportunity to test advanced statistical gap filling approaches able to reproduce not only seasonal budgets but day to day and diurnal variability in ET. Previously often limited to eddy covariance approaches, coupling such advanced statistical approaches (e.g. artificial intelligence and neural network~~

approaches) might be an ideal fusion of measurement frequency and the ability to capture treatment effects like small scale soil differences (Falge et al. 2001a; KIŞI and ÇIMEN 2009).

AgroFLUX – a newly developed research platform centered around an automated gantry crane (FluxCrane) mounted chamber system at the "CarboZALF-D" research site – was initially built to capture the effect of soil type and management on CO_2 fluxes with high spatial and temporal resolution (Vaidya et al., 2021). The adaption of the system to measure ET provided us with the opportunity to analyze stand scale ET fluxes including the development of a data analysis tool for measured ET fluxes, and test different gap-filling strategies. The paper's aim is to establish an approach that would provide reliable predictions of ET fluxes both in terms of gross seasonal budgets as well as representing diurnal trends for automated canopy chamber systems. We tested five different gap-filling approaches including basic statistic and advanced approaches including artificial intelligence and machine learning approaches. During the growth period of winter rye from mid-September 2020 to the end of June 2021 ET and relevant environmental and plant growth parameters were measured to identify the corresponding drivers of crop evapotranspiration and productivity. The gantry crane system covers a field where three different soil types are present, which reflect the erosion gradient typical for the hillside of the hummocky ground moraine landscape of NE Germany. This enabled to assess the impact of soil type as well as soil management on crop season ET budgets, seasonal development and agronomic WUE (WUE_{agro}). In the following we will examine i) soil type and management effects on budgets as well as spatio-temporal variability of ET fluxes over the growing season as well as on WUE_{agro} and ii) the suitability gap filling strategies used in this study as well as potential ways forward to improve these approaches.

2 Material and Methods

2.1 Study Site and experimental design

The experimental site is located near Dedelow in the NE of Germany within the Uckermark region (53° 23´ N, 13° 47´ E; ~50-60 m a.s.l). The typical surface is characterized by the last ice ages, when glaciers shaped the land surface to a hummocky moraine landscape (MLUK 2020). Today, this landscape is characterized by an intensive agriculture, where only 20% of the land is not affected by past and present soil erosion due to tillage and water (Sommer et al. 2008; Wilken et al. 2020)(Sommer et al. 2008; Wilken et al. 2020) resulting in a very high spatial variability of soils (Wehrhan and Sommer 2021)(Wehrhan and Sommer 2021) and associated growing conditions for crops (Wehrhan et al. 2016)(Wehrhan et al. 2016). The study were carried out on the 100 x 16 m FluxCrane experimental field within the AgroFlux site as a part of the interdisciplinary research area CarboZALF-D (Fig. 1a). CarboZALF was initially built to capture the effect of soil type and management on CO_2 fluxes and face the challenge to establish a robust approach to determine carbon budgets of erosion deposition processes within the soil-plant-atmosphere system with high spatial and temporal resolution (Sommer et al. 2016). There is an elevation difference of one meter and all relevant erosion stages are covered (non-eroded Calcic Luvisol (LV-cc), strongly eroded Nudiargic Luvisol (LV-ng) and extremely eroded Calcaric Regosol (RG-ca); see Fig. 1b, c; (Sommer et al. 2008; Wehrhan et al. 2016; Vaidya et al. 2021)) , according to the soil classification of the IUSS working group (WRB 2014). For our study. In combination with the ongoing climate crisis, it is proving difficult to develop land-use methods that allow reliable and sustainable arable

farming under these challenging conditions ~~we used 18 plots in total, 6 per soil type (Fig. 1c). For the 6 plots per soil type, a randomized, full-factorial design, each repeated three fold, was adopted for topsoil dilution vs. non topsoil dilution (topsoil modification - first 8 to 9 cm). During the study period from September 2020 to June 2021 (287 days), winter rye of the hybrid variety "SU Piano" was grown with a density of 200 plants per m² on an area of 0.176 ha. The novel gantry crane automatic chamber system (Fig. 1d) was installed over this study site in 2019 (see Vaidya et al. 2021); the gas exchange chambers are lowered on each plot on round iron frames with a diameter of 1.59 m and a basal area of 1.99 m²~~.

The study was carried out on the 100 x 16 m FluxCrane experimental field, an integral part of the AgroFlux sensor platform located at the interdisciplinary research area CarboZALF-D (Fig. 1a). There is an elevation difference of one meter and all relevant local erosion stages are covered (WRB 2014): non-eroded Calcic Luvisol (LV-cc), strongly eroded Nudiargic Luvisol (LV-ng) and extremely eroded Calcaric Regosol (RG-ca); see Fig. 1b, e; (Sommer et al. 2008; Wehrhan et al. 2016; Vaidya et al. 2021). Here we used 18 plots in total, 6 per soil type (Fig. 1c). For the 6 plots per soil type, a randomized, full-factorial design, each repeated three fold, was adopted for topsoil dilution vs. non-topsoil dilution (first 8 to 9 cm). During the study period from September 2020 to June 2021 (286 days), winter rye of the hybrid variety SU Piano was grown with a density of 200 plants per m² on an area of 0.176 ha. The novel gantry crane automatic chamber system (Fig. 1d) was installed on this study site in 2019 (see Vaidya et al. 2021). The attached gas exchange chambers were lowered on each plot on round structural steel frames with a diameter of 1.59 m and a basal area of 1.99 m^2.

2.2 Cultivation and ~~manipulation~~top-soil dilution

~~Because the AgroFLUX sensor platform site is located on a conventionally farmed agricultural area intended to represent a variety of soils in the region, soil manipulation, tillage, cultivation and fertilization measures were implemented before and during the experiment. For soil manipulation, the first 5 cm of topsoil (1.2 t per plot) was removed from 6 selected plots of the 12 plots per soil type using an excavator on July 14-15, 2020. Then, a soil pit was excavated (20-21 July, 2020) at the adjacent field to expose the soil horizons (A, B and C) and for application of an equivalent mass to the former prepared plots. Thus, Al horizon was applied to the prepared plots of the non eroded Calcic Luvisol (LV-cc), Bt horizon on the strongly eroded Nudiargic Luvisol (LV-ng) and Ck horizon to the extremely eroded Calcaric Regosol (RG-ca). In order to exchange equal amounts, both the removed and the applied material were weighed. The next day (22 July, 2020) a heavy tine cultivator was used for slow, deep chiseling (25 cm) of all plots in east-west direction with subsequent mixing of all manipulated plots by hand. Finally, the chamber frames were reinstalled. In the following, the resulting treatments of the same soil types are labelled as "manipulated" (m) and "non-manipulated" (n-m).~~

~~The actual tillage prior to sowing took place just before seeding on September 17, 2020. For this, the frames were removed and the soil was loosened to a depth of 25 cm in west-east direction. Sowing was done with a power harrow-drill combination. Organic fertilization was applied to all plots per soil type, with each 3 plots manipulated and 3 plots non-manipulated, using digestate (10 $m^3 ha^{-1}$) on September 16, 2020. Due to uneven field emergence, replanting had to be done in all non-manipulated~~

~~plots within the frames (LV-cc: 13 plants per plot; LV-ng: 40 plants per plot; RG-ca: 82 plants plot). For general plant protection and soil treatment, herbicides were applied to the field during the growing season (e.g. glyphosate).~~

The AgroFLUX sensor platform site is located on a conventionally farmed agricultural area that is intended to represent a variety of soils in the region. Hence, top-soil dilution, tillage, cultivation and fertilizer application measures were implemented before and during the experiment. The manipulative field experiment was originally established to study the feedbacks of a dynamic disequilibrium in the carbon cycle of arable lands. Deep tillage or soil erosion lead to an admixture of subsoil material into the plough layer (Doetterl et al. 2016) which alters topsoil properties (SOC, clay content etc.). The resulting changes in the main rooting zone might reduce crop growth (Öttl et al. 2021). We mimic these common landscape processes in our top-soil dilution experiment under controlled conditions (Vaidya et al. 2021). After topsoil removal (1.2 t per plot; first 8-9 cm; 3 of the 6 plots per soil; July 14-15, 2020) we added the equivalent mass (1.2 t) of the respective subsoil horizons (E, Bt, Ck) taken from a large soil pit nearby. Thus, E horizon was applied to the prepared plots of the non-eroded Calcic Luvisol (LV-cc), Bt horizon on the strongly eroded Nudiargic Luvisol (LV-ng) and Ck horizon to the extremely eroded Calcaric Regosol (RG-ca). Subsequently we mixed the added subsoil material with the remaining local Ap horizon. Finally, the chamber frames were reinstalled. In the following, the resulting treatments of the same soil types are labelled as non-diluted (n-d) and diluted (d). The actual tillage prior to sowing took place just before seeding on September 17, 2020. For this, the frames were removed and the soil was loosened to a depth of 25 cm in west-east-direction. Sowing was done with a power harrow-drill combination. Fertilization was applied to all plots per soil type before and during the growing season using digestate from Pflanzenbauhof GbR (Uckermark, Germany), Triple Super Phosphate (TSP) and grain potash (Table B1). Due to initial changes in the topsoil structure (after the addition of subsoil material), germination differed between manipulated and non-manipulated plots. In order to achieve similar plant densities in all plots, replanting had to be done in all non-diluted plots within the frames (LV-cc: 13 plants per plot; LV-ng: 40 plants per plot; RG-ca: 82 plants per plot). For general plant protection and soil treatment, herbicides were applied to the field prior to the growing season (e.g. glyphosate; September 3, 2020).

2.3 Gantry crane system description and gas exchange measurements

The ET flux measurements were carried out by a novel automated chamber system (FluxCrane) using a 5-meter-high gantry crane traveling on two ~~110m~~110 m tracks which has been described in detail ~~(Vaidya et al. 2021)~~(Vaidya et al. 2021)~~.~~. Briefly, the ~~Gantry Crane carries two transparent chambers made of polymethyl methacrylate (PMMA; A: 1.986 m^2; V: 4.621 m^3). The system,~~system designed by Pfannenstiel ProProject GmbH (Bad Tölz, Germany), is capable of moving in three dimensions: the x-axis for movement along the track, the y-axis for movement perpendicular to the track, and the z-axis for vertical chamber movement. The FluxCrane carries two transparent chambers made of polymethyl methacrylate (PMMA; A: 1.986 m^2; V: 3.756 m^3). Since the two chambers do not move independent from each other along the track, frames where arranged in rows, from which each half was measured by one chamber ~~is parallel~~. To ensure airtight sealing during chamber deployment, steel frames ~~were used~~ with a diameter of 1.59 m and ~~additionally~~ a depth of 5 cm into the soil equipped with a foam ring were used to further increase the chambers bearing surface, while deployed.

2.4 Input parameters for gap filling

2.4.1 Environmental Parameters

Relative humidity (RH) [%] was measured simultaneously during the ET flux measurements outside the chambers, while temperature (T) [°C] and photosynthetically active radiation (PAR) [µmol $m^{-2} s^{-1}$] were measured both outside as well as inside the chambers. Like precipitation (Pr) [mm], soil moisture (SM; 13 to 18 cm depth) [%] was measured at an adjacent (< 50m) experimental field.

2.4.2 Plant specific parameters

Spectral plant indices, such as the ratio vegetation index (also simple ratio (SR); RVI) were manually recorded weekly for all 18 subplots using a near infrared (NIR)/visible light (VIS) double, 2 canal sensor device (SKR 1850, Skye Instruments Ltd., UK) mounted on a 1.8 m handheld pole (Görres et al. 2014; Kandel et al. 2013), connected to a CR1000 data logger (Campbell Scientific Ltd., USA). The double, 2 channel sensor device consist of an up- and downward sensor, measuring the incoming (VISi) and reflected (VISr) VIS at a wavelength of 656 ±10 nm and incoming (NIRi) and reflected (NIRr) NIR at 780 ±10 nm. The upward sensor was fitted with a cosine-correction diffusor for measurements of the incident radiation, while the downward sensor had a 25° cone field of view, thus covering an area of 0.5 m^2 during measurements. For each plot, three successive 10 seconds (=10 records) measurements were performed. The RVI was used as an indicator for standing crop biomass and is close to zero for a fallow surface and increases as plant cover increases. The RVI was calculated following Equ. 1:

$$RVI = \frac{\frac{NIR_r}{NIR_i}}{\frac{VIS_r}{VIS_i}} \quad (1)$$

Since only weekly measurement data were available from plot wise RVI, daily RVI data was obtained by fitting a sigmoidal function for initial plant growth in autumn up to a stagnation in winter and polynomial function for shoot elongation and later on senescence during spring growth and summer ripening, respectively (Fig. A1). No plant growth was assumed during the that we term the "non-growing season" from November 24, 2020, to March 22, 2021, because average daily temperatures were mostly below 5°C (<3 consecutive days).

2.5 ET flux calculation and gap filling

The workflow included various steps to pre-process data obtained by the gantry crane automatic chamber system, calculate ET fluxes and finally applying and validating the different gap-filling procedures (Fig. A2).

2.5.1 ET Flux calculation

ET fluxes were determined by measuring the development of chamber headspace H_2O concentrations (4 sec frequency) over chamber deployment time of 7 minutes in a flow-through non-steady-state (FT-NSS) mode (Livingston and Hutchinson 1995),

using two infrared gas analyzers (one per chamber; LI-COR 850, Licor Biosciences, UK). The chambers have an average light transmittance of about 76 % (74% for chamber 1 and 78% for chamber 2), but a reduction in ET due to reduced light availability is not expected (Pape et al. 2009). Temperature differences during chamber closure were minimized by the short measurement time and ventilation (<1.5°C) with two small axial flow fans (5.61 m^3 min^{-1}) used to homogenize the chamber headspace air. To compensate for the difference in tubing length between the chambers and the analyzer (chamber 1: 15 m vs. chamber 2: 22 m), a flow rate of 2.3 l min^{-1} and 3.6 l min^{-1} was set to obtain a similar sensor death time of 13 seconds. A CR6 data logger and CDM-A116 analog multiplexer (Campbell Scientific Inc., USA) were used for data recording and storage. The plots were measured hourly up to 24 times a day in order to be able to detect daily variations. Due to the randomized measurement design, each plot was measured approximately twice per week, which would theoretically result in approximately 2200 measurements per plot throughout the entire season. However, the system was designed to shut down under high winds and cold temperatures, resulting in a true average of only 724 measurements per plot per season. Diurnal ET day- and nighttime fluxes considered in this study were calculated for the cultivation period from September 17, 2020 (sowing of winter rye), until harvest of winter rye on June 30, 2021.

2.4 Input parameters for gap filling

2.4.1 Environmental Parameters

Relative humidity (RH) [%] (WXT520, Vaisala, FI) was measured during the ET flux measurements outside the chambers while temperature (T) [°C] (109, Campbell Scientific Ltd., USA) and incoming photosynthetically active radiation (PAR) [µmol m^{-2} s^{-1}] (SKP 215, Skye Instruments Ltd., UK) were measured both outside as well as inside the chambers. Precipitation (Pr) [mm] (Tipping Bucket Rain Gauge 52203, R. M. Young Company, USA) and relative soil moisture (SM; 13 to 18 cm depth) [%] (ML2x, Delta-T Devices Ltd., UK) were measured at an adjacent field (< 25m; Fig. 1).

2.4.2 Plant specific parameters

Spectral plant indices, such as the ratio vegetation index (RVI; also simple ratio SR) were manually recorded weekly for all 18 plots using a near-infrared (NIR)/visible light (VIS) double, 2 channel sensor device (SKR 1850, Skye Instruments Ltd., UK) mounted on a 1.8 m handheld pole (Görres et al. 2014; Kandel et al. 2013), connected to a CR1000 data logger (Campbell Scientific Ltd., USA). The double, 2 channel sensor device consisted of an upward- and downward–facing sensor, measuring the incoming (VISi) and reflected (VISr) VIS at a wavelength of 656 ± 10 nm and incoming (NIRi) and reflected (NIRr) NIR at 780 ±10 nm. The upward sensor was fitted with a cosine-correction diffusor for measurements of the incident radiation, while the downward sensor, installed 1.8m above the ground, had a 25° cone field of view, thus covering an area of 0.5 m^2 during measurements (Görres et al. 2014). Each plot was measured once a week for 30 seconds, resulting in one mean value including 30 measurement points. The RVI was used as an indicator for standing crop biomass and is close to zero for a fallow surface and increases as plant cover increases. The RVI was calculated following Equ. 1:

$$RVI = \frac{\frac{NIRr}{NIRi}}{\frac{VISr}{VISi}} \quad (1)$$

Since only weekly plot-wise RVI data were available, daily RVI data were obtained by fitting a sigmoidal function for initial plant growth in the fall up to stagnation due to plant inactivity in the winter and a polynomial function for shoot elongation and later senescence during spring growth and summer maturation, respectively (Fig. A1). During the period from November 24, 2020 to March 22, 2021, which we refer to as the non-growing season, no plant growth was assumed due to average daily temperatures below 5°C (<3 consecutive days).

2.5 ET flux calculation and gap filling

2.5.1 ET Flux calculation

~~The plots were measured up to 24 times a day in order to be able to detect daily variations. Diurnal ET fluxes considered in this study were calculated for the period from September 17, 2020 (sowing of winter rye), until harvest of winter rye on June 30, 2021. ET flux calculation was performed based on the ideal gas equation (Eq. 2) modified by (Hamel et al. 2015) using an adapted R-script, based on those presented by (Hoffmann et al. 2015).~~

$$ET_{flux} = \frac{c_{H2O} \times P \times M_{H2O}}{R \times T} \quad (2)$$

The workflow included various steps to pre-process data obtained by the FluxCrane, calculate ET fluxes and finally applying and validating the different gap-filling procedures (Fig. A2). ET flux calculation was performed based on the ideal gas equation (Eq. 2) modified by (Hamel et al. 2015) using an adapted R-script, based on those presented by (Hoffmann et al. 2015).

$$ET_{flux} = \frac{c_{H2O} \times P \times M_{H2O}}{R \times T} \left[\frac{mm}{d}\right] \quad (2)$$

$$ET_{flux}[mmol\ m^{-2}s^{-1}] = \frac{ET_{flux}\left[\frac{mm}{d}\right]}{(t \times 1000)} * \left(\frac{1}{M_{H2O}}\right) \quad (3)$$

With ET_{flux} [~~mmol m^{-2} s~~mm d^{-1}] being the evapotranspiration rate, c_{H2O} the moles of water per total moles present, P the gas pressure [Pa], M_{H2O} the molar mass of water [18 g mol^{-1}], R the gas constant [8.314 m^3 Pa K^{-1} mol^{-1}] and T the temperature [K] inside the chamber. ~~To prevent a disturbance due to initial chamber headspace air homogenization, the~~The ET flux in mmol m^{-2} s^{-1} (Equ. 3) was also calculated to ensure comparability with other studies. The first 15% of each measurement were discarded prior to flux calculation, to prevent a disturbance due to initial homogenization of the chamber headspace air. The temporal change was determined by linear regressions on several subsets of the data generated based on a variable moving window with a starting window size of 1:20 minutes (20 consecutive data points) and a maximum length of 2 minutes (30 consecutive data points). This procedure resulted in several ET fluxes for each measurement, from which the best flux was subsequently selected using a set of soft and hard criteria. Soft criteria included: (i) checking whether the conditions for the

application of a linear regression were fulfilled (normality, variance homogeneity, linearity); (ii) no outliers were present (±~~6xIQR~~6x interquartile range); (iii) temperature variation during the measurement was < 1.5 °C. Calculated fluxes per measurement that did not meet the quality criteria were discarded. Afterwards applied hard-criteria reduced potentially remaining multiple fluxes per measurement further to the ideal ET flux. Since the air in the chamber headspace reached higher water saturation with increasing time, hard criteria were based on the selection of the flux which showed the shortest temporal distance to the start of measurement and had the maximum length.

During the measurements, various events could lead to erroneous ET fluxes such as e.g. ~~fog,~~fog (saturation of the chamber interior), sensor failures, or chamber leakage due to failure in chamber deployment. Erroneous fluxes were hence discarded~~, such as e.g., in May 2021 (sensor failure) or after 23 June 2021.~~. In addition, potential differences of the measurements between the sensors of both chambers were evaluated by the measurements of ambient air during periods of no chamber deployment.

~~2.5.2 ET flux Gap-filling~~

A complete data set of hourly data points for the 286 days of the cultivation period would consist of 6888 measurements per treatment. After the flux calculation and filtering using the soft and hard criteria, a total of 13,011 ET flux measurements were performed, resulting in approximately 2,169 measurements per treatment. For individual plots, an average of 723 (624 to 1210; 10.5 %) measurements were measured and the remaining were predicted by the gap approaches (Table B2; 89.5% on average).

2.5.2 Gap-filling ET flux

To gap fill ET fluxes, five different gap filling approaches were used and compared with each other. Gap-filling procedures included: 1.) ~~-~~a simple statistical approach: ~~"~~Mean diurnal variation~~"~~ (MDV); two empirical approaches: namely 2.) ~~"~~non-linear regression~~"~~ (NLR) and 3.) ~~"~~Look-Up-Tables~~"~~ (LUT) as well as two machine learning approaches: with 4.) ~~"~~Support Vector machine~~"~~ (SVM) and 5.) ~~"~~artificial neural network with Bayesian regularization~~"((~~ANN_BR).

~~MDV (Falge et al. 2001b; Moffat et al. 2007) is used to calculate missing hourly values through interpolation of values measured at the same hour during adjacent days. Thus, for the season with 286 days, the missing values were calculated for every hour, generating 24 values per day.~~

~~NLR is based on parameterized non-linear equations using the mean least square method to express the relation between ET and T, RH, SM, PAR and RVI.~~MDV (Falge et al. 2001b; Moffat et al. 2007) is used to calculate missing hourly values through interpolation of values measured at the same hour during adjacent days. Thus, for the season with 286 days, the missing values were calculated for every hour, generating 24 values per day.

NLR is based on parameterized non-linear equations using the mean least square method to express the relation between the total seasonal data of ET and T, RH, SM, PAR and RVI. Half-hourly values were predicted using the predictor variables and obtained function parameters.

Gap-filling missing ET fluxes using the LUT approach is based on the assumption of similar ET fluxes during similar environmental conditions, whereby similarity is defined through a number of thresholds for the different environmental

variables. Thus, missing ET-fluxes can be predicted based on the environmental conditions as well as the RVI associated with the missing data. To do so, measured ET-fluxes per subplot were split into different classes ($c_{sturges}$) based on T, RH, SM, PAR and RVI, with their class limits determined by Sturges rule (Equ. ~~34~~, ~~(~~Harkins 2022~~)~~). Within this study, on average, 12 classes of equal size were formed covering the range of all parameters.

$$c_{sturges} = \frac{1+3.32 * log\,(n)}{log\,(10)} \qquad (~~34~~)$$

Gaps in ET-fluxes were subsequently filled with the average ET flux of the class corresponding to obtained environmental parameters within the gap. In case, no class could attributed to measured environmental conditions within a gap, the average ET flux was used.

~~SVM is a black-box-model, where a computer algorithm learns by teaching data to assign values to objects or classes (Noble 2006). As mentioned by (Kim et al. 2020), the SVM uses a slack variable to circumvent unseparated parameters due to noise or extreme values in the data, as well as the radial basis kernel function based on previous SVM studies for upscaling fluxes (e.g. Ichii et al. 2017; Xu et al. 2018).~~

~~In comparison, ANN_BR is a combination of a purely empirical nonlinear regression model with a stochastic Bayesian algorithm for regularizing the network training (Schmidt et al. 2018). An artificial neural network (ANN) consists of nodes connected by weights representing the regression parameters (Bishop and others 1995; Hagan et al. 1996; Moffat et al. 2007; Kubat 1999; Rojas 1996). The network is trained by providing it with sets of input data such as the environmental and plant-specific parameters mentioned earlier and the associated output data in the form of e.g. ET flux values. Similar to (Moffat et al. 2007), all techniques evaluated use the classical back-propagation algorithm, where the training of the ANN is performed by propagating the input data through the nodes via the weighted connections and then back-propagating the error and adjusting the weights so that the network output optimally approximates the ET fluxes. Subsequent to this training, the underlying dependencies of the ET fluxes on the environmental and plant-specific input variables are mapped to the weights and the ANN is used to predict half-hourly ET fluxes, where the performance of the ANN depends on several criteria.~~

~~2.6 Seasonal budgets of~~SVM is a black-box-model, where a computer algorithm learns by teaching data to assign values to objects or classes (Noble 2006). As mentioned by Kim et al. (2020), the SVM uses a slack variable to circumvent unseparated parameters due to noise or extreme values in the data, as well as the radial basis kernel function based on previous SVM studies for upscaling fluxes (e.g. Ichii et al. 2017; Xu et al. 2018).

In comparison, ANN_BR is a combination of a purely empirical nonlinear regression model with a stochastic Bayesian algorithm for regularizing the network training (Schmidt et al. 2018). An artificial neural network (ANN) consists of nodes connected by weights representing the regression parameters (Bishop and others 1995; Hagan et al. 1996; Moffat et al. 2007; Kubat 1999; Rojas 1996). The network is trained by providing it with sets of input data such as the environmental and plant-specific parameters mentioned earlier and the associated output data in the form of e.g. ET flux values. Similar to (Moffat et

al. 2007), all techniques evaluated use the classical back-propagation algorithm, where the training of the ANN is performed by propagating the input data through the nodes via the weighted connections and then back-propagating the error and adjusting the weights so that the network output optimally approximates the ET-fluxes. Subsequent to this training, the underlying dependencies of the ET fluxes on the environmental and plant-specific input variables are mapped to the weights and the ANN is used to predict half-hourly ET fluxes, where the performance of the ANN depends on several criteria.

2.6 Seasonal cumulative ET ~~and,~~ Water Use Efficiency

~~**Seasonal budgets of** and crop **ET were determined using half-hourly ET values predicted by SVM. Daily averages and seasonal budgets of ET were formed in order to view the development over the entire growth period. The amount of plant biomass in dry mass (DM) [kg] was recorded during harvest treatment wise, which, in combination with ET_{sum} yields the agricultural water use efficiency WUE_{agro} ((Hatfield and Dold 2019), Equ. 4). This is the WUE_{ABG} variant of WUE_{agro}, as the dry mass is total aboveground biomass (Katerji et al. 2008).~~**

ET_{sum} were determined using half-hourly or hourly ET values predicted by all five gap filling approaches (Figure 6, Figure A3 - A6). Daily averages [mm d^{-1}] and ET_{sum} (mm cultivation $period^{-1}$) were formed in order to view the development over the entire cultivation period. The amount of plant biomass in dry mass (DM) [kg] was recorded during harvest for each treatment, which, in combination with ET_{sum} yields the agricultural water use efficiency WUE_{agro} (Hatfield and Dold 2019, Equ. 5). This is the WUE_{ABG} variant of WUE_{agro}, as the dry mass is total aboveground biomass (Katerji et al. 2008).

$$WUE_{agro} = \frac{DM}{ET_{sum}} \quad (45)$$

To obtain a comparable value for the ET_{sum} calculated by the FluxCrane, crop evapotranspiration (ET_c) was calculated (Allen 1998). ET_c (Equ.6) was calculated from the crop factor K_c and the potential evapotranspiration ET_0 using monthly averages (DWD 2022).

$$ET_c = K_c \times ET_0 \quad (6)$$

2.7 Statistical analysis

All calculations were performed using the ~~statistic~~ statistical software R (R Core Team, 2021) version ~~"~~4.0.4~~"~~. Therefore, several packages (Table ~~B1~~B3) were used to calculate the ET fluxes and to perform subsequent gap-filling as well to visualization of results. To check the general precision and accuracy of all gap-filling approaches, all measured values were compared with associated predicted values for each treatment ~~wise~~. Additionally, k-fold cross validation (k = 5) was performed on the resulting ET data to test the predictive outcome of the approaches and ensure robust statistics. Then, all the coefficient of determination (~~r²~~R^2), mean absolute error (MAE), normalized root mean square error (NRMSE) and Nash-Sutcliffe efficiency (NSE) were calculated and used to define ~~the most accurate approach~~performance classes (Table 1) to evaluate the

accuracy of the approaches for the given setup (Moriasi et al. 2015). To determine parameter impact on ET, linear and non-linear models were used. Lastly, differences of ET_{sum}, DM and WUE_{agro} between treatments were tested with the Kruskal-Wallis-test.

3. Results

3.1 Environmental parameters

The ~~long-term (1991 to 2020; ZALF) mean annual air temperature in this region is 8.8°C with a mean annual precipitation and potential evapotranspiration of 467 mm and 637 mm, respectively (ZALF research station, Dedelow). Compared to that, the~~ study ~~period~~year was significantly warmer (mean temperature 9.6 °C) and wetter ~~with mean temperature of 9.6°C and an annual precipitation of (~~508 mm annual precipitation) between July 1, 2020 and June 30, 2021~~. Temperatures~~, compared to mean annual air temperature (8.8 °C) and precipitation (467 mm). In particular, temperatures (Fig. 2a) were above average in the fall and winter period in 2020 as well as June, 2021. On the other hand, April and May, which are crucial for crop growth, were colder and also drier. Daily mean ~~relative humidity (rH)~~RH ranged between 50 % and 92.4 % with increasing diurnal variation in warm periods. ~~Photosynthetically active radiation (~~PAR~~)~~ (Fig. 2b) largely reflected the seasonal variation of the day length with a maximum of 1860 µmol m^{-2} s^{-1} (half hourly measurements), and reduced values during longer storm events and high cloud cover~~. The relative soil moisture~~ (e.g. through changes in photosynthesis). The SM at 13 to 18 cm depth largely reflects the intensity of precipitation events (Fig. 2c), ranging from 12 % to 29 %. One exception is a prominent increase in mid-February that can be attributed to low temperatures and subsequent snowmelt. The largest precipitation events (> 10 mm d^{-1}) occurred ~~in the observed period~~ on September 26, 2020 with 12 mm, on December 24, 2020 with 15 mm and on February 3, 2021 with 16 mm. ~~With respect to soil moisture, a~~A sharp ~~downward~~declining trend in SM and no response to precipitation events is evident from ~~late~~ April ~~until~~ (about 25 %) to harvest in June~~.~~ (about 12 %). However, this can be explained by a high water consumption ~~by~~of the fully developed crop stand and canopy interception. Shallower ~~soil moisture~~SM sensors at 3 to 8 cm (not shown) indeed responded to these precipitation events albeit weakly, indicating the infiltration to deeper soil layers was impaired.

3.2 Plant development

RVI estimates are based on weekly measurements. Two temporal periods in particular were relevant for plant growth: i) the period from germination to the non-growing season in winter; and ii) the growing period after winter until harvest (Fig. A1). The maximum RVI values were all reached at a similar time (May ~~09~~15, 2021 to May ~~13~~18, 2021). In this regard, the ~~manipulated~~non-diluted non-eroded soil ~~of "~~LV-cc~~"~~ n-d had the highest RVI ~~(~22.~~(16.46 on average), while the diluted non-~~manipulated~~eroded soil LV-cc d showed ~~minor~~lower values ~~(~17.19). In comparison, the~~ (13.88 on average). The strongly eroded soil of ~~"~~LV-ng~~"~~ revealed ~~a reversed~~the same pattern with a ~~high~~higher RVI for non-~~manipulated (~18.53~~diluted (12 on average) and ~~low~~lower RVI ~~manipulated (~13.81~~on diluted (10.35 on average) treatments. The extremely eroded soil of ~~"~~RG-

ca~~",~~, on the other hand, showed ~~negligible~~huge differences ~~for manipulated~~between the non-diluted and ~~non-manipulated~~diluted treatments ~~(~12.72, ~12.49).~~(10.95 vs. 5.87 on average). Apart from that, ~~manipulated treatments showed a greater~~ the maximum standard deviation ~~than~~differed between non-~~manipulated~~diluted and diluted treatments for the three soil types (LV-cc: ~~5.2 < 6.6~~1.65 < 3.29; LV-ng: ~~4.7 < 6.~~1.09 < 1.94; RG-ca: 1.~~6 < 6.2~~17 < 0.82). Higher RVI values were already reached in non-eroded and strongly eroded soils compared to extremely eroded soil ~~only~~ during the ~~first~~initial growing season in fall of 2020 until the non-growing season. Thus, mean RVI values of ~~~~4.~~9647 to ~~7.81~~6.63 were obtained for non-eroded and strongly eroded soils, while the extremely eroded soils had mean RVI values of only ~~~3.22 (manipulated) and ~~3.61 (~~non-manipulated~~n-d) and 2.31 (d).

3.3 ET Fluxes

~~A total of 13,011 ET flux measurements were performed, resulting in approximately 2,169 measurements per treatment. ET fluxes differ mainly between soil types (Fig. 3) with minor differences between non-manipulated and manipulated treatments of the same soil type.~~ The seasonal ~~trend~~development (Fig. 3) of the quality-screened measured ET-fluxes is similar for all treatments: a short growth phase after germination (1 - 2 mmol m^{-2} s^{-1}) is followed by a ~~drop in~~decrease of fluxes until and during the non-growing season in winter~~,~~ (< 0.1 mmol m^{-2} s^{-1}), when hardly any plant activity is found due to low temperatures. ~~At the end of~~After the non-growing season, fluxes ~~in the magnitude of those observed in fall occur~~ quickly ~~and subsequent~~return to their maximum fall level (1 - 2 mmol m^{-2} s^{-1}) and then increase rapidly ~~increase.~~(> 5 mmol m^{-2} s^{-1}). On the non-eroded soil (LV-cc~~)~~), this rapid increase continued into June, while ET fluxes on the eroded soils (LV-ng and RZ-ca) ~~showed their peak~~ already ~~in May.~~peaked in May. In addition, there is a clear difference in the maximum fluxes measured between soil types with 6.7 mmol m^{-2} s^{-1} for both treatments of non-eroded LV-cc, 5.6 / 6.5 mmol m^{-2} s^{-1} (n-d / d) for LV-ng, and 5.8 / 5.1 mmol m^{-2} s^{-1} (n-d / d) for RG-ca. Notably, there is a data gap from late April to late May due to sensor failure.

3.4 Gap filling and validation

~~Figure 4 shows the calibration model performance of the different tested gap-filling approaches as 1:1 agreement plots with associated coefficients of determination (r^2).~~ Pronounced differences of tested gap-filling approaches in terms of respective calibration statistics could be found. Calibration model-performance differ in their scatter around the 1:1-agreement plots (Fig. 4) and associated coefficients of determination (R2). NLR shows a clear overestimation at low and underestimation of higher ET fluxes. Compared to that, MDV more accurately predicts low/high ET fluxes, but is characterized by a much lower precision due to a higher variance. Among all gap filling approaches, displayed ~~r^2 (> 0.89) and NSE (> 0.9; see table 1)~~ calibration statistics (Table 2) indicate a very good or good (Table 1) prediction for SVM, ANN_~~Br~~BR, MDV and LUT over the entire range of observed ET fluxes. Considering the k-fold~~-~~ cross-validation (Fig. 5, ~~table 2), one can quickly see that~~ Table 3), ANN_BR and SVM perform good, while MDV shows partially satisfactory statistics, and LUT ~~is not suitable~~shows unsatisfactory statistics due to allocation problems ~~while ANN_Br and SVM also perform well.~~that arise when no class is found for the given climate conditions and the mean is used. Statistically, ANN_BR and SVM were similarly good in predicting

observed ET fluxes (~~table 1 and~~Table 2~~)~~ and 3), even if they show a small tendency to overestimate low ET fluxes. However, gap-filled ET fluxes using ANN~~ BR showed a large number of negative ET fluxes over the entire non-growing season, which is generally less realistic. Hence, we chose SVM to gap-fill ET flux data and predict crop season budgets.~~ BR showed a large number of predicted negative ET fluxes (1547 on average per plot; Figure A6) throughout the cultivation period. These fluxes occurred to an unrealistic degree during times when RH was significantly below saturation and plants were active (e.g., during the daytime period), resulting in a reduction in seasonal cumulative ET between 1 and 51 mm, depending on the treatment. This is most likely a method specific extrapolation problem (see discussion) and the reason we use SVM for final budget estimations.

3.5 ~~Seasonal ET budgets, WUE temporal dynamics, treatment~~Treatment differences and crop ~~productivity~~ET

~~Statistical analysis of the correlations revealed that no significant difference was found between ET_{sum} of the different treatments.~~In general, eroded soils tend to have a negative effect on ET_{sum}. However, this trend was not statistically significant (Kruskal-Wallis-Test, ET_{sum}: $\chi2 = 3.04$, df = 5, p = 0.69). DM and WUE, on the other hand, differed significantly between treatments (Kruskal-Wallis-Test, DM: $\chi^2 = 14.58$, df = 5, p = 0.01; WUE: $\chi^2 = 11.$~~15~~12, df = 5, p = 0.05). ~~–~~The amount of plant biomass in dry mass (DM) [kg] is decreasing from non-~~manipulated~~diluted to ~~manipulated~~diluted and less eroded soil types to more eroded soil types. DM ranges from 1.5 ± 0.13 kg m^{-2} for ~~"~~LV-cc n-~~m"~~d to 0.85 ± 0.2 kg m^{-2} for ~~"~~RG-ca ~~m". d.~~ WUE_{agro} is decreasing from less eroded to more eroded soil types ranging from 7.~~73~~25 ± 1.~~07~~23 g DM kg^{-1} H_2O to 4.~~98~~69 ± 0.~~76~~71 DM kg^{-1} H_2O (Fig. 8).

In order to compare the individual treatments, daily ET and ~~cumulative ET budgets~~ET_{sum} were calculated (Fig. 6). ET was affected by T, RH, PAR, and RVI, whereas only a small correlation was found with SM (Fig. 7~~; table B2~~). Higher ET fluxes were induced by increases in T, PAR and RVI, whereas increasing RH resulted in lower ET fluxes. ~~The seasonal ET budgets (~~ET_{sum}~~,~~ (Fig. 8a) ~~range~~ranges between ~~197.74 ± 36.36~~212 ± 45 mm (LV-cc n-~~m~~d) and ~~169.66 ± 24.28~~180 ± 29 mm (RG-ca ~~m). In general, increasing erosion tends to a decrease in ET_{sum}.~~d).

ET_0 for the observed study period (September 2020 – June 2021) and region (Uckermark) was 370 mm (DWD 2022). We used the monthly values to calculate the ET_c using ET_0 and the crop coefficient (Allen 1998), resulting in an ET_c of 263 mm for the cultivation period.

4 Discussion

~~ET is the main ecosystem response variable to drought and mostly measured or modelled on a larger scale by e.g. measurement systems such as eddy covariance, Bowen ratio or even satellite remote sensing. These methods are unsuitable for assessing small scale differences and do require great experimental care or errors can lead to serious economic losses due to incorrect water management (Allen et al. 2011; Burba and Verma 2005; Hargreaves 1989). Here, the used, new automated chamber system, which is based on a gantry crane combined with complex gap-filling approaches, offered an unprecedented opportunity~~

~~to combine benefits of all methods (high temporal resolution of e.g. eddy covariance with high spatial resolution of chamber measurements) with fewer disadvantages (e.g. atmospheric disturbances due to permanently installed chambers). It measure ET reliably in high resolution, while allowing for the assessment of treatment effects on ET (in this case soil type and management). We can see that differences in ET over the growing season as well as between soil types and management options result from a variety of factors. In the following we will discuss i) soil type and management effects on budgets along with spatio-temporal variability of ET fluxes over the growing season as well as WUE_{agro} and ii) the suitability of gap-filling strategies used in this study as well as potential ways forward to improve these approaches.~~

~~4.1 Seasonal variability, soil and treatment differences in ET fluxes and WUE~~

~~Over the entire growing period, ET fluxes responded on the long term particularly to crop growth, first during the establishment period in fall (mid of October to mid of November) and then again during the main growth period in spring (end of March to mid of May). The close relation of measured ET flux dynamics to obtained RVI (Fig. 7) (e.g. Hanks et al. 1969), can be associated with increasing transpiration rates that superimpose the suppressive effects of increasing canopy biomass (e.g. increases in LAI or RVI) on soil evaporation (Dubbert et al. 2014; Groh et al. 2020). More dynamically, ET reacted to changes in environmental conditions, particularly temperature and relative humidity which together define the vapour pressure deficit (VPD) as well as PAR. In particular, crops that have been bred to prioritize carbon gain over water conservation will tend to respond to rising VPD strongly (Dubbert et al. 2014; Massmann et al. 2019). In summary, air temperature, humidity and PAR together with increasing biomass (expressed as higher RVI) control ET variability during the peak growth period in spring until harvest. Soil moisture, on the other hand, did not have a limiting effect on ET during the period considered. This is noteworthy since the Uckermark region has an ET_{pot}/Pr ratio of nearly 1 which is regarded as a threshold value between systems that are energy (<1) vs. water limited (>1; DWD 2022).~~

~~The soil types in the region~~In the following, we will discuss i) the effects of soil-type and top-soil dilution on ET_{sum}, yield (DM) and WUE_{agro}, along with ii) the spatio-temporal variability of ET fluxes over the cultivation period, and iii) the suitability of the gap-filling strategies used in this study as well as potential ways forward to improve our approaches.

4.1 Effects of soil-type and top-soil dilution on ET

In the studied region, soil types vary in their suitability for agricultural cultivation (MLUK 2020). Luvisols ~~show a good~~support large water ~~budget~~fluxes due to their clay-depleted deep ~~topsoils~~top-soils in combination with the clay-enriched and mostly decalcified ~~subsoils and thus belong to~~sub-soils. They are among the most ~~fertile~~productive soils in Brandenburg (MLUK 2020; Stahr 2022). ~~Regosols are generally only moderately suitable for arable farming but cultivated in the studied region. They are usually found on hilltops and are characterized by C horizons near to the surface, lack of depth development, and limited rootability due to the dense, high carbonate parent material, resulting in a low water availability and plant growth (Herbrich et al. 2018). It is a typical soil type, formed by erosion of agricultural Luvisols. As a result of these conditions and the associated increase in drought and heat events in Central Europe (Spinoni et al., 2018), regional farmers might become~~

~~limited in their choice of crops grown due to water availability. Here, our results indicate significant differences between measured soil types in DM and WUE_{agro} but no significant differences between ET_{sum}, although we see a clear decreasing trend along the erosion gradient and topsoil manipulation of the three studied soil types (198 ± 36 mm to 170 ± 25 mm). This was expected due to the supposed lower rootability and soil fertility, but the difference was likely reduced by the comparatively wet year of observation. A dry spell during the main growth period could lead to larger differences here. Since there are no significant differences in ET_{sum}, it can be concluded that the observed agroecosystems were mostly energy limited, however it is very likely that the ET_{pot}~~Regosols are generally only moderately suitable for arable farming. They are usually found on hilltops and are characterized by parent material near to the surface, lack of depth development, and limited rootability due to the dense, carbonate-rich parent material. They typically have low water availability and plant growth (Herbrich et al. 2018). They are formed by erosion of agricultural Luvisols as relatively organic matter rich top-soil is removed and deeper, nutrient-depleted lower soil layers are mixed into the cultivated layer (Arriaga and Lowery 2003; Pimentel and Kounang 1998). In addition, the carried out top-soil dilution aimed at testing one of the processes of an approach to enhance soil C storage through top-soil deepening. Topsoil deepening through deeper ploughing might store originally top-soil bound SOC in the deeper subsoil and generate SOC recharge in the diluted C poor top soil (Sommer et al. 2016). The latter being tested during this study by the carried out top-soil dilution. However, side effects include, similar to erosion, nutrient deficiency and weaker rootability leading to decreased crop growth and yield (Al-Kaisi and Grote 2007; Schneider et al. 2017; Feng et al. 2020). The boundary soil conditions established by erosion and top-soil dilution may not only impact crop growth and yield but also disrupt the crop water balance, especially with the expected increase in drought and heat events in Central Europe (Spinoni et al. 2018). Consequently, farmers might become limited in their choice of crops due to water availability.

As predicted, we observed a significant decline in yield with erosion and top-soil dilution during the study period. However, the impact of soil-type specific erosion intensity and top-soil dilution on ET_{sum} was not as pronounced and the trend of declining ET_{sum} with soil-type and top-soil dilution was not statistically significant among all treatments (212 ± 45 mm on non-eroded Calcic Luvisol to 180 ± 29 mm on extremely eroded top-soil diluted Calcaric Regosol). Notably, the studied year 2020/21 was comparatively wet (231.1 mm precipitation during the observed period), and potential effects of lower rootability and enhanced drought stress were not observed during the main growth period. This is of great importance because the Uckermark region generally has an overall water balance of about 1 (precipitation input equals ET_{pot} output) and is therefore water or energy limited depending on the annual precipitation and ET_{pot} of each year. For example, the extremely dry year of 2018 was very likely water limited with an annual precipitation of < 450 mm and a predicted ET_{pot} of > 650 mm and thus by far exceeding annual precipitation. However, the year 2021 had an annual precipitation of about 600 mm and a predicted ET_{pot} of < 575 mm (DWD 2022). Hence, in rather wet years, like the observed 2021, plant growth in the region is rather energy limited (of course dependent on precipitation during the growth period). This fits with our results, as during the studied period most plots had a lower ET_{sum} than cumulative precipitation. However, it is very likely that the ET_{pot}/Pr ratio, and in fact also the observed actual ET_{sum}/Pr ratio will vary considerably between wetter and drier years and between different crops (particularly winter vs. summer crops). ~~Winter crops and especially winter rye, are more resilient to drought (Ehlers 1997) and an investigation with~~

~~other crops and management strategies could lead to a higher decrease in ET with soil erosion. The difference in DM for the observed year was much larger compared to the ET differences in turn driving soil type or erosion effects on the water use efficiency (WUE_{agro}). In a period of consecutive dry years e.g., a lower WUE~~

Additionally, the observed imbalance of response in yield vs. ET_{sum} led to significantly reduced WUE_{agro}. In a period of consecutive dry years, a lower WUE_{agro} could additionally have a negative effect on the water and carbon balance of the whole system, since the water consumption becomes ~~largely unproductive~~less efficient, especially for the Calcaric Regosol (up to 36% less yield per used amount of water; Meena et al. 2020). This could further exacerbate the drought and potentially lead to legacy effects. Finally, winter crops and especially winter rye, are more resilient to drought (Ehlers 1997) due to their head start in growth. Hence, a long-term investigation spanning with other crops (e.g. summer cereal crops) and management strategies would be particularly interesting, as a greater decrease in ET may be observed with soil-specific erosion intensity.

~~One of our expectations was that differences in ET crop season budgets would result mainly from differences during the main vegetation period from April to harvest due to variations in biomass and thus transpiration. In contrast, differences in ET budgets between soil types, rather occur continuously throughout the period with very little variation. It is noticeable that ET fluxes of eroded treatments prior the growing season are lower than those of non-eroded plots for long periods. Although the growing season between April and June is responsible for a large portion of total evapotranspiration for the entire period, ranging from ~ 64% to ~ 70%, it is only responsible for a small portion of differences, with a maximum of ~ 12.3 mm from the non-eroded plot. The prior period, on the other hand, is responsible for a difference of up to ~ 15.8 mm from non-eroded plots, although it accounts for only ~ 30 % to 36 % of total evapotranspiration. However, it should be noted that mean differences per day of about 0.14 mm in the growing season was about 75% higher than in the previous period with about 0.08 mm per day. This is interesting, because it suggests that the reason behind the soil type differences in ET are caused by static differences (e.g. lower biomass) rather than dynamic differences (e.g. the vegetation responses to environmental drivers or drought). Here, partitioning ET into evaporation and transpiration (e.g. Rothfuss et al. 2021) could help explain the reasons for small scale ET variability.~~

~~Beside this static differences, high ET fluxes could be detected at the beginning of the growing season which could be explained by a lower soil cover. This may be related to the fact that vegetation cover due to uneven field emergence, which is visible in the RVI, can largely reduce soil evaporation (Dubbert et al. 2014; Hu et al. 2009; Raz-Yaseef et al. 2012; Wang et al. 2012). However, the vegetation period studied can be described as rather wet compared to long-term environmental conditions and it will be interesting to monitor this for a number of years differing in environmental and hydrological conditions as well as for different crops. For example, the more eroded soil types are reported to lead to decreases in rooting depths (Ötttl et al. 2021; Sommer et al. 2016; Herbrich et al. 2018). In dry years, this may very well lead to a more pronounced response in ET flux dynamics between the treatments particularly in the spring-summer transition compared to results from the present study, when the crop benefits stronger from deep water sources during pronounced dry spells (Thorup-Kristensen et al. 2020). Finally, results for evapotranspiration from nearby lysimeters suggested annual gross ET budgets of around 300 mm to 600 mm in years between 2014 and 2018 with varying crops (Groh et al. 2020). The budgets estimated in (Groh et al. 2020), however,~~

~~spanned the entire year, whereas in this study budgets were calculated to the crop period (only ~9 months and the summer months with high ET_{pot} excluded), explaining the overall higher budgets. In addition, temperatures and precipitation were below average in April and May during the growing season under consideration which result in lower ET thus matching our results. Thus, our ET budgets seem sensible overall, however, a direct comparison between the new gantry crane chamber system, lysimeters and potentially drone based observations of ET would be advisable particularly in light of ongoing discussions surrounding method constraints of estimating ET across scales (Ding et al. 2021; Ghiat et al. 2021; Hamel et al. 2015).~~

~~4.2~~ 4.2 Seasonal variability of ET fluxes and WUE

Over the entire cultivation period, ET fluxes responded particularly to crop growth, first during the establishment period in fall (mid of October to mid of November) and then again during the main growth period in spring (end of March to mid of May). The close relation of measured ET flux dynamics to RVI (Fig. 7; e.g. Hanks et al. 1969) can be associated with increasing T rates that strongly compensate for the suppression of E, as canopy biomass increases (Dubbert et al. 2014; Groh et al. 2020). Over the diurnal cycle, ET reacted to changes in environmental conditions, particularly temperature and RH, which together define the vapor-pressure deficit (VPD), as well as PAR. In particular, crops that have been bred to prioritize carbon gain over water conservation will tend to respond to rising VPD strongly (Dubbert et al. 2014; Massmann et al. 2019). Air temperature, humidity and PAR together with increasing biomass (expressed as higher RVI) controlled ET variability during the peak growth period in spring until harvest. SM, on the other hand, did not have a limiting effect on ET, which we attribute to the wet conditions during the observation period (see above), confirming that the observed crop cycle was not limited by water availability.

One of our expectations was that expected differences in ET_{sum} would result mainly from differences during the main vegetation period from April to harvest due to variations in biomass and thus T. However, while the growing season between April and June is responsible for a large portion of ET_{sum}, ranging from 66 % to 73 %, it is only responsible for a small portion of differences between treatments, with a maximum of 14.3 mm from the non-eroded soil-types. The combined fall and winter period, on the other hand, is responsible for a difference of up to 17.5 mm in ET_{sum} between non-eroded and extremely eroded soil-types, although it accounts for 27 % to 34 % of ET_{sum} only. This is interesting, because it suggests that the reason behind the soil type differences in ET for winter rye are caused by static differences (e.g. lower biomass) and suppressed E (e.g. a shift in the T/ET ratio) rather than dynamic differences (e.g. the vegetation responses to environmental drivers or drought). This should be further evaluated by partitioning ET into T and E. The described FluxCrane is particularly suited for such an approach by combining flux and in-situ stable isotope approaches (Dubbert et al. 2014; Rothfuss et al. 2021). Beside the overall slight reduction of ET_{sum} on eroded soil-types and top-soil diluted treatments, measured ET fluxes were larger on extremely eroded plots at the beginning of the growing season before canopy closure which could be explained by a lower soil cover. This may be related to the fact that lower vegetation cover, which is visible in the RVI, can lead to higher E prior to canopy closure (Dubbert et al. 2014; Hu et al. 2009; Raz-Yaseef et al. 2012; Wang et al. 2012).

4.3 Gap-filling approaches

The experimental design of the used gantry crane system with a high amount of different treatments and measurements of diurnal cycles coupled with unsuitable climate conditions causes data gaps that need to be filled. The challenge was composed of two main elements: i) the correct quantification of ET_{sum} and ii) a realistic prediction of diurnal variations both during winter (low ET) and summer (high ET). With respect to the gap-filling approaches compared, it was determined that all approaches provided fairly plausible results for accounting of ET_{sum}. While LUT showed good calibration results, allocation problems caused the least predictive ability during validation over the entire range of measured ET fluxes. While some studies received also fairly plausible results for LUT and MDV (Boudhina et al. 2018; Falge et al. 2001a; Moffat et al. 2007), adjusting the classes of the LUT could further improve the results of this approach. However, due to the significantly good results of the other approaches, this was not done yet. Additionally, other studies found out that MDV (as well as LUT) especially performs worse with increasing large gaps (Moffat et al. 2007; Kim et al. 2020). Especially for conditions, where no measurements could take place due to environmental conditions, the fact that MDV takes averaged values of close measurements with good measurement conditions could explain the bad predictions. Same for LUT, since no classes could be created for conditions no measurements took place. The machine learning approaches SVM and ANN_BR, on the other hand, also performs well here. For seasonal variability and budgets we also achieved the best performance with approaches ANN_BR and SVM. However, the best approach for gap filling can vary depending on the application and investigated parameter. For example, in gap-filling methane fluxes using eddy covariance (Kim et al. 2020), ANN_BR was superior to SVM. Additionally, results of other studies like (Shrestha and Shukla 2015), who tried to predict actual evapotranspiration (ET_a) measured with lysimeters with different approaches (e.g. ANN_BR and SVM) and crops (pepper, watermelon) in a sub-tropical environment resulting in best predictions with SVM (pepper: 204.7 mm lysimeter vs. 181.8 mm SVM; watermelon: 231.71 mm lysimeter vs. 189.83 mm SVM) found a overestimation of low- and underestimation of high-fluxes of ANN_BR and SVM. We were able to detect this tendency only with a slight overestimation of small fluxes. In comparison to empirical calculations, another study by (KIŞI and ÇIMEN 2009) worked out, that SVM performed even better for ET_0 predictions (FAO-56 PM) with solar radiation, temperature, relative humidity and wind speed compared to commonly used empirical approaches like CIMIS-Penman, Hargreaves, Ritchie or Turc and similar with reduced parameters (solar radiation and temperature). Thus, using other environmental parameters such as wind speed could also improve SVM results for our application. Additionally, it must be noted that the quality of SVM (and ANN_BR) predictions is highly dependent on the amount of data (Chia et al. 2020; Abudu et al. 2010).

4.3 Conclusion and outlook

Filling data gaps using statistical and empirical approaches is used in many fields like calculations of reference ET (ET_0) with limited meteorological parameters (Chia et al. 2020) or actual ET (ET_a) from eddy-covariance measurements as well as plant chamber-measurements (Hui et al. 2004; Moffat et al. 2007; Falge et al. 2001a; Falge et al. 2001b; Hamel et al. 2015; Kübert

et al. 2019). The connection between gap-filling approaches in combination with the described continuous high-resolution long-term ET measurements of numerous small-scale treatments gives additional opportunities to observe the progression of ET over an entire growing season. In comparison to eddy covariance measurements (e. g. Boudhina et al. 2018; Simpson et al. 2019), this approach is able to point out small-scale treatment differences like soil type differences and associated erosion stages at a heterogeneous field with a relatively high number of different treatments simultaneously. Additionally, there is not such a big impact through the system compared to e.g. permanently installed plant chambers or manually conducted approaches which deal with condensation problems due to fix installed tubing and inappropriate air mixing inside the chamber (e. g. (Hamel et al. 2015). This problem does not exist with the automated gantry crane system, since there are permanent measurements and a consistent air flow through the system.

The study could demonstrate that the gantry crane system used provided plausible ET fluxes. While it has many benefits over other measurement systems, there are some particular challenges. For example, the crane and the installed measuring equipment have to be checked regularly and in case of strong winds a manual shutdown and fixation have to be performed. With respect to the ET measurements, there might also be an underestimation of the automated chambers, which has already been found in comparison to the measurement with the eddy covariance in other studies (e.g. Simpson et al., 2019). To shed light on this problem, a more detailed comparison could be made in the future with nearby lysimeters with the same soil types. In order to the gap-filling approaches, only the support vector machine (SVM), artificial neural network with Bayesian regularization (ANN_BR) and non-linear regression (NLR) approaches are suitable for a consistently consideration of diurnal variations at this stage, while all approaches provided plausible data for the seasonal period. Thereby, gap filling is highly dependent on data quality, with larger gaps leading to greater problems. Since the gantry crane was unable to provide data for extended periods (up to approx. 1. month), this led to uncertainties. However, the studied measurement period was the first full crop period observed by the novel infrastructure and these initial problems will be very likely mitigated in the future through reliable and regular maintenance. One gap in particular should be highlighted in the data considered: hardly any data could be collected with chamber 1 for an extended period of time due to a sensor problem in May (May 1 to May 29.). Although the gap could be filled reasonably well with the SVM, with validation statistics indicating a reasonable predicting performance (Table 2), an underestimation is possible due to the large gaps importance of this month for plant growth. Additionally, an interruption of the non-growing season could cause better results, since it can be assumed that already in short warmer period's plant activity is to be found (<3 days less than 5°C). Here, e.g. several non-growing seasons could represent the plant growth much better and lead to a further increase of the predicted accuracy.

Despite these small flaws, the system has a very large potential to bring new insights into water flux dynamics and budgets and, in combination with measurements of NEE into growths season dynamics of water use efficiencies in the future. On the one hand, the different effects of soil types and related erosion levels on other crops could be investigated, and on the other hand, related knowledge for cultivation could be gained. Furthermore, in contrast to other established approaches to quantify ET at the plots scale, the novel system is unique in its potential to combine it with innovative measurements such as stable water isotopes (Dubbert et al. 2014; Kübert et al. 2020). Stable water isotopes could be used to separate the evapotranspiration

~~into transpiration and evaporation, [which is of crucial importance for the terrestrial water balance and for the prediction of future ecosystem feedbacks] (Groh et al. 2020). This could lead to further plant specific knowledge like e.g. root water uptake and uptake depths of different crops. Especially the possibility of automated high-resolution long-term measurements of these parameters is unique and could provide further knowledge about the ratio of evaporation and transpiration in the course of the season, since we expect huge changes through drought here.~~

Methodologically, the study faced two main challenges: accurately quantifying ET_{sum} and realistically predicting diurnal variations during both, the low ET winter and high ET summer periods. Among the gap-filling approaches compared in this study, only NLR showed calibration statistics less than good (Table 2). While the LUT showed very good calibration results, the allocation problems that occur when no class is found (Fig. 5) and the mean is used resulted in the lowest predictive ability during validation over the full range of measured ET fluxes. Some studies also obtained quite plausible results for LUT and MDV (Boudhina et al. 2018; Falge et al. 2001a; Moffat et al. 2007), and adjusting the classes of the LUT could further improve the results of this approach. However, with the available dataset, the only way to avoid allocation problems was to use fewer classes. This resulted in a loss of variability, making diurnal differences invisible and ET estimates less accurate. MDV, on the other hand, partially showed only satisfactory values during validation (Table 3), while SVM and ANN performed good or very good according to the defined classes (Table 1). Additionally, previous studies found that MDV (as well as LUT) performs particularly weakly with increasingly large gaps (Moffat et al. 2007; Kim et al. 2020). Especially for conditions where no measurements could take place due inter-alia environmental conditions (large gaps in winter), the fact that MDV takes averaged values of adjacent measurements could explain the rather bad predictions. This is similar for LUT, since no classes could be created for conditions where no measurements took place. The machine learning approaches SVM and ANN_BR, on the other hand, are not as sensitive to larger observational gaps because their training includes all measurements. For seasonal variability and budgets, we achieved the best performance with the SVM approach, while ANN showed reduced daily and seasonal cumulative ET due to an unrealistic amount of predicted negative fluxes (up to 51 mm; Fig. A6). However, the best approach for gap filling can vary depending on the application and investigated parameters. For example, in gap-filling methane fluxes using eddy covariance (Kim et al. 2020), ANN_BR was superior to SVM.

Another important aspect of gap-filling is potential over- or underestimation of fluxes. Shrestha and Shukla (2015), for example, attempted to predict actual lysimeter ET using different approaches (e.g. ANN_BR and SVM) and crops (pepper, watermelon) in a subtropical environment. Best predictions were obtained with SVM (pepper: 204.7 mm lysimeter vs. 181.8 mm SVM; watermelon: 231.71 mm lysimeter vs. 189.83 mm SVM). However, they reported overestimation of low fluxes and underestimation of high fluxes by ANN_BR and SVM. In our study, we observed a tendency to slightly overestimate small fluxes using SVM based gap-filling. In this regard, using plot specific multi-depth SM data could also improve the predicted ET_{sum} based on SVM in the future. Similarly, we expected considering wind speed to improve ET prediction, but could not find an effect on observed ET for the study period.

Furthermore, it must be noted that the quality of SVM (and ANN_BR) predictions is highly dependent on the amount of data available (Chia et al. 2020; Abudu et al. 2010). Consequently, we tested the minimum amount of data necessary to provide predicted ET fluxes of good quality (see criteria in M&M). For the particular dataset already 50 % of the total data available (minimum 300 measurements) provided good results. Thus, we emphasize that capturing a large variability of fluxes under different environmental conditions seems to be more important than a merely large data set.

4.4 Accuracy of the new system

ET_c was 263 mm from the cultivation period. This is comparable to our observed results (ET_{sum}) of 212 mm for non-eroded Calcic Luvisol, given that ET_0 calculations using the Penman-Monteith equation (FAO56-PM) are reported to overestimate ET_0 and thus ET_c (Allen 1998). Thus, our ET_{sum} seem sensible overall, however, a direct comparison between the FluxCrane, lysimeters and potentially drone based observations of ET would be advisable. This is particularly true in light of ongoing discussions surrounding method constraints of estimating ET across scales (Ding et al. 2021; Ghiat et al. 2021; Hamel et al. 2015).

For example, a multi treatment lysimeter experiment is located nearby the FluxCrane and Groh et al. (2020) report a wide range of ET_{sum} for the period between 2014 and 2018 (300 to 600 mm), with the lower range boundary being comparable to our results (considering that we only calculated budgets for the 9 months growths period and exclude the fallow period during the summer months with high ET). It has to be noted however that not only environmental conditions but also crops studied in Groh et al. (2020) varied from year to year and more importantly from our study, hampering comparability between studies. However, the direct vicinity of two large scale set-ups able to estimate ET_{sum} should be utilized in the future. Another lysimeter based study conducted in Brunswick (Lower Saxony, Germany) for a cultivation season of winter rye report a range of observed ET fluxes very comparable to our study, with less than 1 mm day^{-1} in winter to a maximum of 6 - 7 mm day^{-1} in summer (bfg 2023).

Finally, filling data gaps using statistical and empirical approaches is used in many fields e.g. to calculate of reference ET (ET_0) with limited meteorological parameters (Chia et al. 2020) or ET from eddy-covariance measurements as well as canopy chamber measurements (Hui et al. 2004; Moffat et al. 2007; Falge et al. 2001a; Falge et al. 2001b; Hamel et al. 2015; Kübert et al. 2019). The connection between gap-filling approaches, in combination with the described continuous high-resolution long-term ET measurements of numerous small-scale treatments, gives additional opportunities to observe the progression of ET over an entire cultivation season and, for example, to identify key periods that drive overall treatment differences.

Compared to other methods for estimating ET, such as eddy covariance measurements (e. g., Boudhina et al. 2018; Simpson et al. 2019), our approach is able to highlight small-scale treatment differences, such as soil type differences and associated erosion stages, in a heterogeneous field with a relatively high number of different treatments simultaneously. In addition, the system is not as disruptive to plant growth. For example, permanently installed canopy chambers or manually conducted approaches, tend to physically harm the canopy and have condensation issues due to permanently installed tubing and inappropriate air mixing within the chamber (e. g. Hamel et al. 2015). The FluxCrane eliminates these problems by providing

continuous measurement and a constant flow of air through the attached canopy chambers. Moreover, the ability to observe nighttime fluxes has great potential to study previously overlooked short-term dynamics in ET and to improve the representation of underlying processes in process-based hydrological modeling, compared to other measurement systems and especially to manually operated chambers.

4.4 Conclusion and outlook

Here, we present a possibility to obtain not only plausible ET_{sum} but also diurnal cycles of ET by using the novel FluxCrane system in combination with SVM based gap filling. We expected strong negative effects of eroded soils and top-soil dilution on ET_{sum} as well as yield. However, crop yield responded much more strongly to eroded soils and top-soil dilution than ET_{sum} in the observed rather wet year, leading to strong negative shifts in WUE_{agro}. The novel FluxCrane with its potential to observe temporal dynamics and seasonal budgets for distinct landscape elements simultaneously, combines the contrasting benefits of eddy covariance and manual chamber techniques. Thus the new system has a large potential to bring new insights into water-flux dynamics and budgets and, in combination with measurements of NEE into growth season dynamics of WUE in the future. This is particularly relevant for the studied region in the Uckermark with its strong spatial heterogeneity in soils and its generally low precipitation. Finally, the novel FluxCrane is unique in its potential to combine it with innovative measurements such as in-situ stable water isotopes (Dubbert et al. 2014; Kübert et al. 2020). Stable water isotopes could be used to separate the ET into T and E. This separation is of crucial importance for the terrestrial water balance and for the prediction of future ecosystem feedbacks (Groh et al. 2020). Water isotopes might also be used to study root water-uptake dynamics (Deseano Diaz et al. 2023; Kühnhammer et al. 2020).

5 Code availability

The codes produced for this study are available from the corresponding author upon request.

6 Data availability

The data sets produced for this study are available from the corresponding author upon request.

7 Author contributions

AD, MD conceived and planned the study. MS, GV, MH supervised automated measurements and conducted complimentary field measurements. AD analyzed the data and wrote a first version of the manuscript. All authors contributed to manuscript writing.

8 Competing interests

The authors declare that they have no conflict of interest.

9 Acknowledgements

The authors acknowledge funding by the German Federal Ministry of Food and Agriculture (FNR Grant: 22404117) and the German Science Foundation (DFG Grant DU1688/6-1). The authors are very grateful to the Pfannenstiel ProProject GmbH for the excellent collaboration in designing as well as constructing the gantry crane system and ATTEC Automation GmbH for programming the system control. The study was a part of the CarboZALF project of the Leibniz Centre for Agricultural Landscape Research (ZALF). The CarboZALF project was financially supported by the German Federal Ministry of Food, Agriculture and Consumer Protection (BMELV) and the Ministry of Environment, Health and Consumer (MLUV) of the State of Brandenburg. The authors thank the ~~Research Station Dedelow of the~~ Experimental Infrastructure Platform (EIP) of ZALF for assistance with field measurements and John Marshal for proofreading the manuscript.

Publication bibliography

Abudu, Shalamu; Bawazir, A. Salim; King, J. Phillip (2010): Infilling Missing Daily Evapotranspiration Data Using Neural Networks. In *J. Irrig. Drain Eng.* 136 (5), pp. 317–325. DOI: 10.1061/(ASCE)IR.1943-4774.0000197.

~~Allen, Richard G.; Pereira, Luis S.; Howell, Terry A.; Jensen, Marvin E. (2011): Evapotranspiration information reporting: I. Factors governing measurement accuracy. In *Agricultural Water Management* 98 (6), pp. 899–920. DOI: 10.1016/J.AGWAT.2010.12.015.~~

Al-Kaisi, Mahdi M.; Grote, Jesse B. (2007): Cropping Systems Effects on Improving Soil Carbon Stocks of Exposed Subsoil. In *Soil Sci. Soc. Am. J.* 71 (4), pp. 1381–1388. DOI: 10.2136/sssaj2006.0200.

Allen, Richard G. (1998): Crop evapotranspiration. Guidelines for computing crop water requirements. Reprinted 2002. Rome: FAO (FAO irrigation and drainage paper, 56).

Amt für Statistik Berlin-Brandenburg (2020): Ernteberichterstattung über Feldfrüchte und Grünland im Land Brandenburg 2019. Available online at https://www.statistik-berlin-brandenburg.de/, checked on 8/18/2022.

Arriaga, F.J; Lowery, B. (2003): Corn production on an eroded soil: effects of total rainfall and soil water storage. In *Soil & Tillage Research* 71 (1), pp. 87–93. DOI: 10.1016/S0167-1987(03)00040-0.

Bakker, Martha M.; Govers, Gerard; Jones, Robert A.; Rounsevell, Mark D. A. (2007): The Effect of Soil Erosion on Europe's Crop Yields. In *Ecosystems* 10 (7), pp. 1209–1219. DOI: 10.1007/s10021-007-9090-3.

bfg (2023): 2.12 Mean Annual Potential Evaporation Depth as Grass Reference Evapotranspiration. Edited by Bundesanstalt für Gewässerkunde. Available online at https://geoportal.bafg.de/dokumente/had/212GrassReferenceEvapotraspiration.pdf, checked on 3/10/2023.

Biggelaar, Christoffel den; Lal, Rattan; Wiebe, Keith; Breneman, Vince (2003): The Global Impact Of Soil Erosion On Productivity: I: Absolute and Relative Erosion-induced Yield Losses. In *Advances in Agronomy* 81, pp. 1–48.

Bishop, Christopher M.; others (1995): Neural networks for pattern recognition: Oxford university press.

Blum, Winfried E.H. (2013): Soil and Land Resources for Agricultural Production: General Trends and Future Scenarios-A Worldwide Perspective. In *International Soil and Water Conservation Research* 1 (3), pp. 1–14. DOI: 10.1016/S2095-6339(15)30026-5.

Boudhina, Nissaf; Zitouna-Chebbi, Rim; Mekki, Insaf; Jacob, Frédéric; Ben Mechlia, Nétij; Masmoudi, Moncef; Prévot, Laurent (2018): Evaluating four gap-filling methods for eddy covariance measurements of evapotranspiration over hilly crop fields. In *Geosci. Instrum. Method. Data Syst.* 7 (2), pp. 151–167. DOI: 10.5194/gi-7-151-2018.

~~Burba, George G.; Verma, Shashi B. (2005): Seasonal and interannual variability in evapotranspiration of native tallgrass prairie and cultivated wheat ecosystems. In *Agricultural and Forest Meteorology* 135 (1-4), pp. 190–201. DOI: 10.1016/J.AGRFORMET.2005.11.017.~~

Chia, Min Yan; Huang, Yuk Feng; Koo, Chai Hoon (2020): Support vector machine enhanced empirical reference evapotranspiration estimation with limited meteorological parameters. In *Computers and Electronics in Agriculture* 175, p. 105577. DOI: 10.1016/j.compag.2020.105577.

Deseano Diaz, Paulina Alejandra; van Dusschoten, Dagmar; Kübert, Angelika; Brüggemann, Nicolas; Javaux, Mathieu; Merz, Steffen et al. (2023): Response of a grassland species to dry environmental conditions from water stable isotopic monitoring: no evident shift in root water uptake to wetter soil layers. In *Plant Soil* 482 (1-2), pp. 491–512. DOI: 10.1007/s11104-022-05703-y.

Ding, Jie; Li, Sien; Wang, Hongshuo; Wang, Chunyu; Zhang, Yunxuan; Yang, Danni (2021): Estimation of Evapotranspiration and Crop Coefficient of Chinese Cabbage Using Eddy Covariance in Northwest China. In *Water* 13 (19), p. 2781. DOI: 10.3390/w13192781.

~~Drexler, Judy Z.; Snyder, Richard L.; Spano, Donatella; Paw U, Kyaw Tha (2004): A review of models and micrometeorological methods used to estimate wetland evapotranspiration. In *Hydrol. Process.* 18 (11), pp. 2071–2101. DOI: 10.1002/hyp.1462.~~

Doetterl, Sebastian; Berhe, Asmeret Asefaw; Nadeu, Elisabet; Wang, Zhengang; Sommer, Michael; Fiener, Peter (2016): Erosion, deposition and soil carbon: A review of process-level controls, experimental tools and models to address C cycling in dynamic landscapes. In *Earth-Science Reviews* 154, pp. 102–122. DOI: 10.1016/j.earscirev.2015.12.005.

Dubbert, Maren; Piayda, Arndt; Cuntz, Matthias; Correia, Alexandra C.; Costa e Silva, Filipe; Pereira, Joao S.; Werner, Christiane (2014): Stable oxygen isotope and flux partitioning demonstrates understory of an oak savanna contributes up to half of ecosystem carbon and water exchange. In *Frontiers in plant science* 5, p. 530. DOI: 10.3389/fpls.2014.00530.

DWD (2019): Klimareport Brandenburg. 1. Auflage. Edited by Deutscher Wetterdienst. Offenbach am Main, Germany.

DWD (2022): ~~Klimakarten Deutschland.~~Climatological maps of Germany. Edited by Deutscher Wetterdienst. Available online at https://www.dwd.de/~~DE/leistungen~~EN/ourservices/klimakartendeutschland/klimakartendeutschland.html, checked on ~~9/8/2022~~3/10/2023.

Ehlers, W. (1997): Zum Transpirationskoeffizienten von Kulturpflanzen unter Feldbedingungen. In *Pflanzenbauwissenschaften* 1 (3), pp. 97–108.

Falge, Eva; Baldocchi, Dennis; Olson, Richard; Anthoni, Peter; Aubinet, Marc; Bernhofer, Christian et al. (2001a): Gap filling strategies for defensible annual sums of net ecosystem exchange. In *Agricultural and Forest Meteorology* 107 (1), pp. 43–69. DOI: 10.1016/S0168-1923(00)00225-2.

Falge, Eva; Baldocchi, Dennis; Olson, Richard; Anthoni, Peter; Aubinet, Marc; Bernhofer, Christian et al. (2001b): Gap filling strategies for long term energy flux data sets. In *Agricultural and Forest Meteorology* 107 (1), pp. 71–77. DOI: 10.1016/S0168-1923(00)00235-5.

Feng, Qi; An, Chunjiang; Chen, Zhi; Wang, Zheng (2020): Can deep tillage enhance carbon sequestration in soils? A meta-analysis towards GHG mitigation and sustainable agricultural management. In *Renewable and Sustainable Energy Reviews* 133, p. 110293. DOI: 10.1016/j.rser.2020.110293.

Fohrer, Nicola; Bormann, Helge; Miegel, Konrad; Casper, Markus; Schumann, Andreas; Bronstert, Axel; Weiler, Markus (2016): Hydrologie. 1. Auflage. Stuttgart: UTB GmbH; Haupt (Utb Basics, 4513).

Ghiat, Ikhlas; Mackey, Hamish R.; Al-Ansari, Tareq (2021): A Review of Evapotranspiration Measurement Models, Techniques and Methods for Open and Closed Agricultural Field Applications. In *Water* 13 (18), p. 2523. DOI: 10.3390/w13182523.

Görres, C.-M.; Kutzbach, L.; Elsgaard, L. (2014): Comparative modeling of annual CO2 flux of temperate peat soils under permanent grassland management. In *Agriculture, Ecosystems & Environment* 186, pp. 64–76. DOI: 10.1016/j.agee.2014.01.014.

Groh, Jannis; Diamantopoulos, Efstathios; Duan, Xiaohong; Ewert, Frank; Herbst, Michael; Holbak, Maja et al. (2020): Crop growth and soil water fluxes at erosion-affected arable sites: Using weighing lysimeter data for model intercomparison. In *Vadose zone j.* 19 (1). DOI: 10.1002/vzj2.20058.

Hagan, M. T.; Demuth, H. B.; Beale, M.; Jesus, O. de (1996): Neural Network Design, Boston. In *PWS Pub. Co. USA*.

Hamel, Perrine; Mchugh, Ian; Coutts, Andrew; Daly, Edoardo; Beringer, Jason; Fletcher, Tim D. (2015): Automated Chamber System to Measure Field Evapotranspiration Rates. In *J. Hydrol. Eng.* 20 (2), Article 04014037, p. 4014037. DOI: 10.1061/(ASCE)HE.1943-5584.0001006.

Hanks, R. J.; Gardner, H. R.; Florian, R. L. (1969): Plant Growth-Evapotranspiration Relations for Several Crops in the Central Great Plains 1. In *Agronomy Journal* 61 (1), pp. 30–34. DOI: 10.2134/agronj1969.00021962006100010010x.

Hanson, Ronald L. (1991): Evapotranspiration and droughts. In *National Water Summary 1988-89: Hydrologic Events and Floods and Droughts (US Geological Survey Water-Supply Paper 2375)*, pp. 99–104.

~~Hargreaves, George H. (1989): Accuracy of Estimated Reference Crop Evapotranspiration. In *J. Irrig. Drain Eng.* 115 (6), pp. 1000–1007. DOI: 10.1061/(ASCE)0733-9437(1989)115:6(1000).~~

Harkins, Ray (2022): Sturge's Rule: A Method for Selecting the Number of Bins in a Histogram, checked on 8/19/2022.

Hatfield, Jerry L.; Dold, Christian (2019): Water-Use Efficiency: Advances and Challenges in a Changing Climate. In *Frontiers in plant science* 10, p. 103. DOI: 10.3389/fpls.2019.00103.

Herbrich, Marcus; Gerke, Horst H.; Sommer, Michael (2018): Root development of winter wheat in erosion-affected soils depending on the position in a hummocky ground moraine soil landscape. In *J. Plant Nutr. Soil Sci.* 181 (2), pp. 147–157. DOI: 10.1002/jpln.201600536.

Hoffmann, Mathias; Jurisch, Nicole; Albiac Borraz, Elisa; Hagemann, Ulrike; Drösler, Matthias; Sommer, Michael; Augustin, Jürgen (2015): Automated modeling of ecosystem CO 2 fluxes based on periodic closed chamber measurements: A standardized conceptual and practical approach. In *Agricultural and Forest Meteorology* 200, pp. 30–45. DOI: 10.1016/J.AGRFORMET.2014.09.005.

Hu, Zhongmin; Yu, Guirui; Zhou, Yanlian; Sun, Xiaomin; Li, Yingnian; Shi, Peili et al. (2009): Partitioning of evapotranspiration and its controls in four grassland ecosystems: Application of a two-source model. In *Agricultural and Forest Meteorology* 149 (9), pp. 1410–1420. DOI: 10.1016/J.AGRFORMET.2009.03.014.

Hui, Dafeng; Wan, Shiqiang; Su, Bo; Katul, Gabriel; Monson, Russell; Luo, Yiqi (2004): Gap-filling missing data in eddy covariance measurements using multiple imputation (MI) for annual estimations. In *Agricultural and Forest Meteorology* 121 (1-2), pp. 93–111. DOI: 10.1016/S0168-1923(03)00158-8.

Ichii, Kazuhito; Ueyama, Masahito; Kondo, Masayuki; Saigusa, Nobuko; Kim, Joon; Alberto, Ma. Carmelita et al. (2017): New data-driven estimation of terrestrial CO 2 fluxes in Asia using a standardized database of eddy covariance measurements, remote sensing data, and support vector regression. In *JGR Biogeosciences* 122 (4), pp. 767–795. DOI: 10.1002/2016JG003640.

Jasechko, Scott; Sharp, Zachary D.; Gibson, John J.; Birks, S. Jean; Yi, Yi; Fawcett, Peter J. (2013): Terrestrial water fluxes dominated by transpiration. In *Nature* 496 (7445), pp. 347–350. DOI: 10.1038/nature11983.

Jie, Chen; Jing-zhang, Chen; Man-zhi, Tan; Zi-tong, Gong (2002): Soil degradation: a global problem endangering sustainable development. In *J. Geogr. Sci* 12 (2), pp. 243–252. DOI: 10.1007/BF02837480.

Kandel, Tanka P.; Elsgaard, Lars; Laerke, Poul E. (2013): Measurement and modelling of CO 2 flux from a drained fen peatland cultivated with reed canary grass and spring barley. In *GCB Bioenergy* 5 (5), pp. 548–561. DOI: 10.1111/gcbb.12020.

Katerji, Nader; Mastrorilli, Marcello; Rana, Gianfranco (2008): Water use efficiency of crops cultivated in the Mediterranean region: Review and analysis. In *European Journal of Agronomy* 28 (4), pp. 493–507. DOI: 10.1016/j.eja.2007.12.003.

Kim, Yeonuk; Johnson, Mark S.; Knox, Sara H.; Black, T. Andrew; Dalmagro, Higo J.; Kang, Minseok et al. (2020): Gap-filling approaches for eddy covariance methane fluxes: A comparison of three machine learning algorithms and a traditional method with principal component analysis. In *Global change biology* 26 (3), pp. 1499–1518. DOI: 10.1111/GCB.14845.

KIŞI, OZGUR; ÇIMEN, MESUT (2009): Evapotranspiration modelling using support vector machines / Modélisation de l'évapotranspiration à l'aide de 'support vector machines'. In *Hydrological Sciences Journal* 54 (5), pp. 918–928. DOI: 10.1623/hysj.54.5.918.

Kubat, Miroslav (1999): Neural networks: a comprehensive foundation by Simon Haykin, Macmillan, 1994, ISBN 0-02-352781-7. In *The Knowledge Engineering Review* 13 (4), pp. 409–412. DOI: 10.1017/S0269888998214044.

Kübert, Angelika; Götz, Miriam; Kuester, Emma; Piayda, Arndt; Werner, Christiane; Rothfuss, Youri; Dubbert, Maren (2019): Nitrogen Loading Enhances Stress Impact of Drought on a Semi-natural Temperate Grassland. In *Frontiers in plant science* 10, p. 1051. DOI: 10.3389/fpls.2019.01051.

Kübert, Angelika; Paulus, Sinikka; Dahlmann, Adrian; Werner, Christiane; Rothfuss, Youri; Orlowski, Natalie; Dubbert, Maren (2020): Water Stable Isotopes in Ecohydrological Field Research: Comparison Between In Situ and Destructive Monitoring Methods to Determine Soil Water Isotopic Signatures. In *Frontiers in plant science* 11, p. 387. DOI: 10.3389/fpls.2020.00387.

Kühnhammer, Kathrin; Kübert, Angelika; Brüggemann, Nicolas; Deseano Diaz, Paulina; van Dusschoten, Dagmar; Javaux, Mathieu et al. (2020): Investigating the root plasticity response of Centaurea jacea to soil water availability changes from isotopic analysis. In *New Phytologist* 226 (1), pp. 98–110. DOI: 10.1111/nph.16352.

Livingston, G. P.; Hutchinson, G. L. (1995): Enclosure-based measurement of trace gas exchange: applications and sources of error. In *Biogenic trace gases: measuring emissions from soil and water* 51, pp. 14–51.

Louwagie, G.; Gay, S. H.; Sammeth, F.; Ratinger, T. (2011): The potential of European Union policies to address soil degradation in agriculture. In *Land Degrad. Dev.* 22 (1), pp. 5–17. DOI: 10.1002/ldr.1028.

Massmann, Adam; Gentine, Pierre; Lin, Changjie (2019): When Does Vapor Pressure Deficit Drive or Reduce Evapotranspiration? In *Journal of advances in modeling earth systems* 11 (10), pp. 3305–3320. DOI: 10.1029/2019MS001790.

Meena, Ram Swaroop; Kumar, Sandeep; Yadav, Gulab Singh (2020): Soil Carbon Sequestration in Crop Production. In Ram Swaroop Meena (Ed.): Nutrient Dynamics for Sustainable Crop Production. Singapore: Springer Singapore, pp. 1–39.

MLUK (2020): Steckbriefe Brandenburger Böden. Sammelmappe. Edited by Ministerium für Ländliche Entwicklung.

Moffat, Antje M.; Papale, Dario; Reichstein, Markus; Hollinger, David Y.; Richardson, Andrew D.; Barr, Alan G. et al. (2007): Comprehensive comparison of gap-filling techniques for eddy covariance net carbon fluxes. In *Agricultural and Forest Meteorology* 147 (3-4), pp. 209–232. DOI: 10.1016/j.agrformet.2007.08.011.

Moriasi, Daniel N.; Gitau, Margaret W.; Prasad Daggupati, Naresh Pai (2015): Hydrologic and Water Quality Models: Performance Measures and Evaluation Criteria. In *Trans. ASABE* 58 (6), pp. 1763–1785. DOI: 10.13031/TRANS.58.10715.

Noble, William S. (2006): What is a support vector machine? In *Nature biotechnology* 24 (12), pp. 1565–1567. DOI: 10.1038/nbt1206-1565.

Öttl, Lena Katharina; Wilken, Florian; Auerswald, Karl; Sommer, Michael; Wehrhan, Marc; Fiener, Peter (2021): Tillage erosion as an important driver of in-field biomass patterns in an intensively used hummocky landscape. In *Land Degradation & Development* 32 (10), pp. 3077–3091. DOI: 10.1002/LDR.3968.

Pape, L.; Ammann, C.; Nyfeler-Brunner, A.; Spirig, C.; Hens, K.; Meixner, F. X. (2009): An automated dynamic chamber system for surface exchange measurement of non-reactive and reactive trace gases of grassland ecosystems. In *Biogeosciences* 6 (3), pp. 405–429. DOI: 10.5194/bg-6-405-2009.

Pimentel, David; Kounang, Nadia (1998): Ecology of Soil Erosion in Ecosystems. In *Ecosystems* 1 (5), pp. 416–426. DOI: 10.1007/s100219900035.

Raz-Yaseef, Naama; Yakir, Dan; Schiller, Gabriel; Cohen, Shabtai (2012): Dynamics of evapotranspiration partitioning in a semi-arid forest as affected by temporal rainfall patterns. In *Agricultural and Forest Meteorology* 157, pp. 77–85. DOI: 10.1016/J.AGRFORMET.2012.01.015.

Rojas, Raúl (Ed.) (1996): Neural Networks. Berlin, Heidelberg: Springer Berlin Heidelberg.

Rothfuss, Youri; Quade, Maria; Brüggemann, Nicolas; Graf, Alexander; Vereecken, Harry; Dubbert, Maren (2021): Reviews and syntheses: Gaining insights into evapotranspiration partitioning with novel isotopic monitoring methods. In *Biogeosciences* 18 (12), pp. 3701–3732. DOI: 10.5194/bg-18-3701-2021.

Schappert, Sebastian (2018): Wie wird Niederschlag beim DWD gemessen und wo fällt am meisten? Deutscher Wetterdienst. Offenbach am Main, Germany. Available online at https://www.dwd.de/DE/wetter/thema_des_tages/2018/11/28.html, checked on 8/18/2022.

Schmidt, Andres; Creason, Whitney; Law, Beverly E. (2018): Estimating regional effects of climate change and altered land use on biosphere carbon fluxes using distributed time delay neural networks with Bayesian regularized learning. In *Neural networks : the official journal of the International Neural Network Society* 108, pp. 97–113. DOI: 10.1016/j.neunet.2018.08.004.

Schneider, Florian; Don, Axel (2019): Root-restricting layers in German agricultural soils. Part I: extent and cause. In *Plant Soil* 442 (1-2), pp. 433–451. DOI: 10.1007/s11104-019-04185-9.

Schneider, Florian; Don, Axel; Hennings, Inga; Schmittmann, Oliver; Seidel, Sabine J. (2017): The effect of deep tillage on crop yield – What do we really know? In *Soil & Tillage Research* 174, pp. 193–204. DOI: 10.1016/j.still.2017.07.005.

Searchinger, Timothy D.; Wirsenius, Stefan; Beringer, Tim; Dumas, Patrice (2018): Assessing the efficiency of changes in land use for mitigating climate change. In *Nature* 564 (7735), pp. 249–253. DOI: 10.1038/s41586-018-0757-z.

Shrestha, N. K.; Shukla, S. (2015): Support vector machine based modeling of evapotranspiration using hydro-climatic variables in a sub-tropical environment. In *Agricultural and Forest Meteorology* 200, pp. 172–184. DOI: 10.1016/j.agrformet.2014.09.025.

Simpson, Gillian; Runkle, Benjamin R.K.; Eckhardt, Tim; Kutzbach, Lars (2019): Evaluating closed chamber evapotranspiration estimates against eddy covariance measurements in an arctic wetland. In *Journal of Hydrology* 578, p. 124030. DOI: 10.1016/j.jhydrol.2019.124030.

Sommer, M.; Augustin, J.; Kleber, M. (2016): Feedbacks of soil erosion on SOC patterns and carbon dynamics in agricultural landscapes—The CarboZALF experiment. In *Soil & Tillage Research* 156 (156), pp. 182–184. DOI: 10.1016/j.still.2015.09.015.

Sommer, M.; Gerke, H. H.; Deumlich, D. (2008): Modelling soil landscape genesis — A “time split” approach for hummocky agricultural landscapes. In *Geoderma* 145 (3-4), pp. 480–493. DOI: 10.1016/j.geoderma.2008.01.012.

Spinoni, Jonathan; Vogt, Jürgen V.; Naumann, Gustavo; Barbosa, Paulo; Dosio, Alessandro (2018): Will drought events become more frequent and severe in Europe? In *Int. J. Climatol* 38 (4), pp. 1718–1736. DOI: 10.1002/joc.5291.

Stahr, Alexander (2022): Bodentypen. Available online at http://www.ahabc.de/bodentypen/, checked on 8/19/2022.

Tallec, Tiphaine; Béziat, Pierre; Jarosz, Nathalie; Rivalland, Vincent; Ceschia, Eric (2013): Crops’ water use efficiencies in temperate climate: Comparison of stand, ecosystem and agronomical approaches. In *Agricultural and Forest Meteorology* 168, pp. 69–81. DOI: 10.1016/j.agrformet.2012.07.008.

~~Thorup-Kristensen, Kristian; Halberg, Niels; Nicolaisen, Mette; Olesen, Jørgen Eivind; Crews, Timothy E.; Hinsinger, Philippe et al. (2020): Digging Deeper for Agricultural Resources, the Value of Deep Rooting. In *Trends in plant science* 25 (4), pp. 406–417. DOI: 10.1016/J.TPLANTS.2019.12.007.~~

Trenberth, K. E. (2011): Changes in precipitation with climate change. In *Clim. Res.* 47 (1), pp. 123–138. DOI: 10.3354/cr00953.

Vaidya, Shrijana; Schmidt, Marten; Rakowski, Peter; Bonk, Norbert; Verch, Gernot; Augustin, Jürgen et al. (2021): A novel robotic chamber system allowing to accurately and precisely determining spatio-temporal CO2 flux dynamics of heterogeneous croplands. In *Agricultural and Forest Meteorology* 296, p. 108206. DOI: 10.1016/j.agrformet.2020.108206.

Wang, L.; D'Odorico, P.; Evans, J. P.; Eldridge, D. J.; McCabe, M. F.; Caylor, K. K.; King, E. G. (2012): Dryland ecohydrology and climate change: critical issues and technical advances. In *Hydrol. Earth Syst. Sci.* 16 (8), pp. 2585–2603. DOI: 10.5194/HESS-16-2585-2012.

Wehrhan, Marc; Rauneker, Philipp; Sommer, Michael (2016): UAV-Based Estimation of Carbon Exports from Heterogeneous Soil Landscapes--A Case Study from the CarboZALF Experimental Area. In *Sensors (Basel, Switzerland)* 16 (2), p. 255. DOI: 10.3390/s16020255.

Wehrhan, Marc; Sommer, Michael (2021): A Parsimonious Approach to Estimate Soil Organic Carbon Applying Unmanned Aerial System (UAS) Multispectral Imagery and the Topographic Position Index in a Heterogeneous Soil Landscape. In *Remote Sensing* 13 (18), p. 3557. DOI: 10.3390/rs13183557.

Wessolek, Gerd; Asseng, Senthold (2006): Trade-off between wheat yield and drainage under current and climate change conditions in northeast Germany. In *European Journal of Agronomy* 24 (4), pp. 333–342. DOI: 10.1016/j.eja.2005.11.001.

Wilken, Florian; Ketterer, Michael; Koszinski, Sylvia; Sommer, Michael; Fiener, Peter (2020): Understanding the role of water and tillage erosion from $^{239+240}$Pu tracer measurements using inverse modelling. In *SOIL* 6 (2), pp. 549–564. DOI: 10.5194/soil-6-549-2020.

WRB (2014): International soil classification system for naming soils and creating legends for soil maps. World Soil Resources Reports No. 106: Food and Agriculture Organization of the United Nations Rome, Italy.

Xu, Tongren; Guo, Zhixia; Liu, Shaomin; He, Xinlei; Meng, Yangfanyu; Xu, Ziwei et al. (2018): Evaluating Different Machine Learning Methods for Upscaling Evapotranspiration from Flux Towers to the Regional Scale. In *JGR Atmospheres* 123 (16), pp. 8674–8690. DOI: 10.1029/2018JD028447.

Tables:

Table 1: Performance classes to evaluate gap-filling approaches.

Class	MAE	NRMSE	NSE	R2
Very Good	< 0.35	< 30	> 0.85	> 0.85
Good	0.35 <= 0.67	30 <= 40	0.85 => 0.75	0.85 => 0.75
Satisfactory	0.67 <= 1	40 <= 50	0.75 => 0.5	0.75 => 0.6
Not Satisfactory	> 1	> 50	< 0.5	< 0.6

Table ~~1: Statistical values for the calibration~~2: Calibration statists of all gap-filling approaches and treatments.

Approach	MAE	NRMSE	NSE	R2
		~~RG-ca nm~~LL-cv n-d		
MDV	0.27	~~44.5~~39.6	0.~~88~~4	0.~~82~~85
LUT	0.~~69~~09	~~72~~14.3	0.~~48~~98	0.~~48~~98
NLR	0.~~51~~56	~~50.9~~49.0	0.~~74~~76	0.~~75~~77
SVM	0.~~3~~26	~~33.9~~25.7	0.~~89~~93	0.~~89~~93
ANN_BR	0.31	~~32.1~~28.0	0.~~9~~92	0.~~9~~92
		LL-ng ~~nmn~~-d		
MDV	0.25	~~31~~29.3	0.~~9~~91	0.~~91~~92
LUT	0.~~74~~09	~~65.9~~13.6	0.~~57~~98	0.~~57~~98
NLR	0.~~48~~54	41.~~7~~1	0.83	0.~~83~~84
SVM	0.28	~~25.4~~22.9	0.~~94~~95	0.~~94~~95
ANN_BR	0.~~3~~31	~~26.2~~24.6	0.~~93~~94	0.~~93~~94
		~~LL-cv nm~~RG-ca n-d		
MDV	0.~~24~~26	~~33.7~~30.9	0.~~89~~90	0.~~89~~91
LUT	0.~~72~~09	~~72.3~~15.7	0.~~48~~98	0.~~48~~98
NLR	0.~~44~~50	42.~~2~~4	0.82	0.83
SVM	0.26	~~28.2~~25.6	0.~~92~~93	0.~~92~~93
ANN_BR	0.~~28~~30	~~29.8~~27.3	0.~~91~~93	0.~~91~~93

Approach	MAE	NRMSE	NSE	R2
		~~RG-ca m~~LL-cv d		
MDV	0.~~23~~25	35.~~4~~2	0.~~87~~88	0.88
LUT	0.~~65~~08	~~71.1~~13.4	0.~~49~~98	0.~~5~~98
NLR	0.~~49~~53	~~48~~46.8	0.~~76~~78	0.~~77~~79
SVM	0.~~28~~23	~~32~~23.0	0.~~9~~95	0.~~9~~95
ANN_BR	0.27	~~29.6~~25.7	0.~~91~~93	0.~~91~~93
		LL-ng ~~m~~d		
MDV	0.25	~~31~~30.6	0.~~9~~91	0.~~9~~91
LUT	0.~~78~~10	~~67.7~~14.6	0.~~54~~98	0.~~54~~98
NLR	0.~~49~~55	~~41.4~~40.6	0.~~83~~84	0.84
SVM	0.29	~~26~~23.6	0.~~93~~94	0.~~93~~94
ANN_BR	0.~~29~~32	~~26.3~~25.0	0.~~93~~94	0.~~93~~94
		~~LL-cv m~~RG-ca d		
MDV	0.~~21~~22	~~30.3~~29.8	0.91	0.91
LUT	0.~~65~~09	~~68.8~~14.3	0.~~53~~98	0.~~53~~98
NLR	0.~~43~~48	~~42~~41.8	0.82	0.83
SVM	0.23	~~26.3~~23.4	0.~~93~~95	0.~~93~~95
ANN_BR	0.~~28~~29	~~29.5~~26.0	0.~~91~~93	0.~~91~~93

Table ~~2:~~ ~~Statistical values for the validation~~3: Validation statists of all gap-filling approaches and treatments.

Approach	MAE	NRMSE	NSE	R2	Approach	MAE	NRMSE	NSE	R2
	LL-cv n-d					LL-cv d			
MDV	0.33	46.0	0.79	0.81	MDV	0.26	35.8	0.87	0.88
LUT	0.74	69.8	0.51	0.51	LUT	0.72	70.0	0.51	0.51
NLR	0.57	50.9	0.74	0.75	NLR	0.55	48.8	0.76	0.77
SVM	0.34	33.7	0.89	0.89	SVM	0.31	32.1	0.90	0.90
ANN_BR	0.35	32.2	0.90	0.90	ANN_BR	0.32	29.6	0.91	0.91
	LL-ng n-d					LL-ng d			
MDV	0.31	32.1	0.90	0.90	MDV	0.31	33.3	0.89	0.89
LUT	0.78	62.7	0.61	0.61	LUT	0.83	64.7	0.58	0.58
NLR	0.54	41.7	0.83	0.83	NLR	0.55	41.5	0.83	0.84
SVM	0.32	25.4	0.94	0.94	SVM	0.33	26.5	0.93	0.93
ANN_BR	0.33	25.9	0.93	0.93	ANN_BR	0.34	26.6	0.93	0.93
	RG-ca n-d					RG-ca d			
MDV	0.28	34.4	0.88	0.89	MDV	0.27	31.8	0.90	0.90
LUT	0.79	71.8	0.48	0.49	LUT	0.69	65.9	0.57	0.57
NLR	0.49	42.2	0.82	0.83	NLR	0.48	42.0	0.82	0.83
SVM	0.29	28.1	0.92	0.92	SVM	0.26	26.2	0.93	0.93
ANN_BR	0.33	29.9	0.91	0.91	ANN_BR	0.31	28.8	0.92	0.92

Figures:

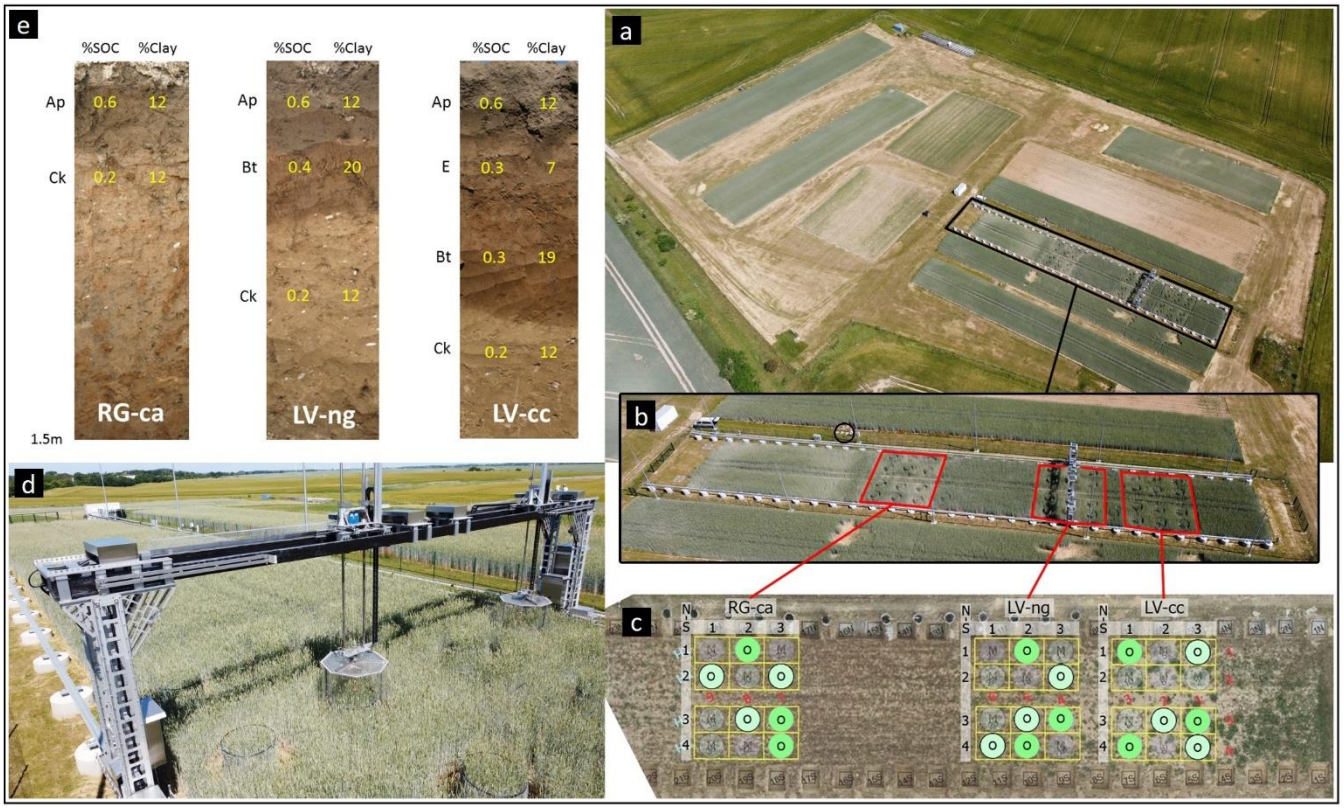


Figure 1: (a) AgroFLUX research site in the CarboZALF-D experimental area with (b) the 110 x 16 m field where (d) the FluxCrane operates on (c) 18 measurement plots ~~from~~of (e) three different soil types (LV-cc: non-eroded ~~calcareous~~calcic Luvisol, LV-ng: highly eroded nudiargic Luvisol, and RG-ca: extremely eroded ~~calcareous~~calcaric Regosol). Soil moisture and precipitation measurements were taken in the marked area (black circle, b). The separation of non-diluted (unframed green) and diluted (framed light green) plots can be seen in (c).

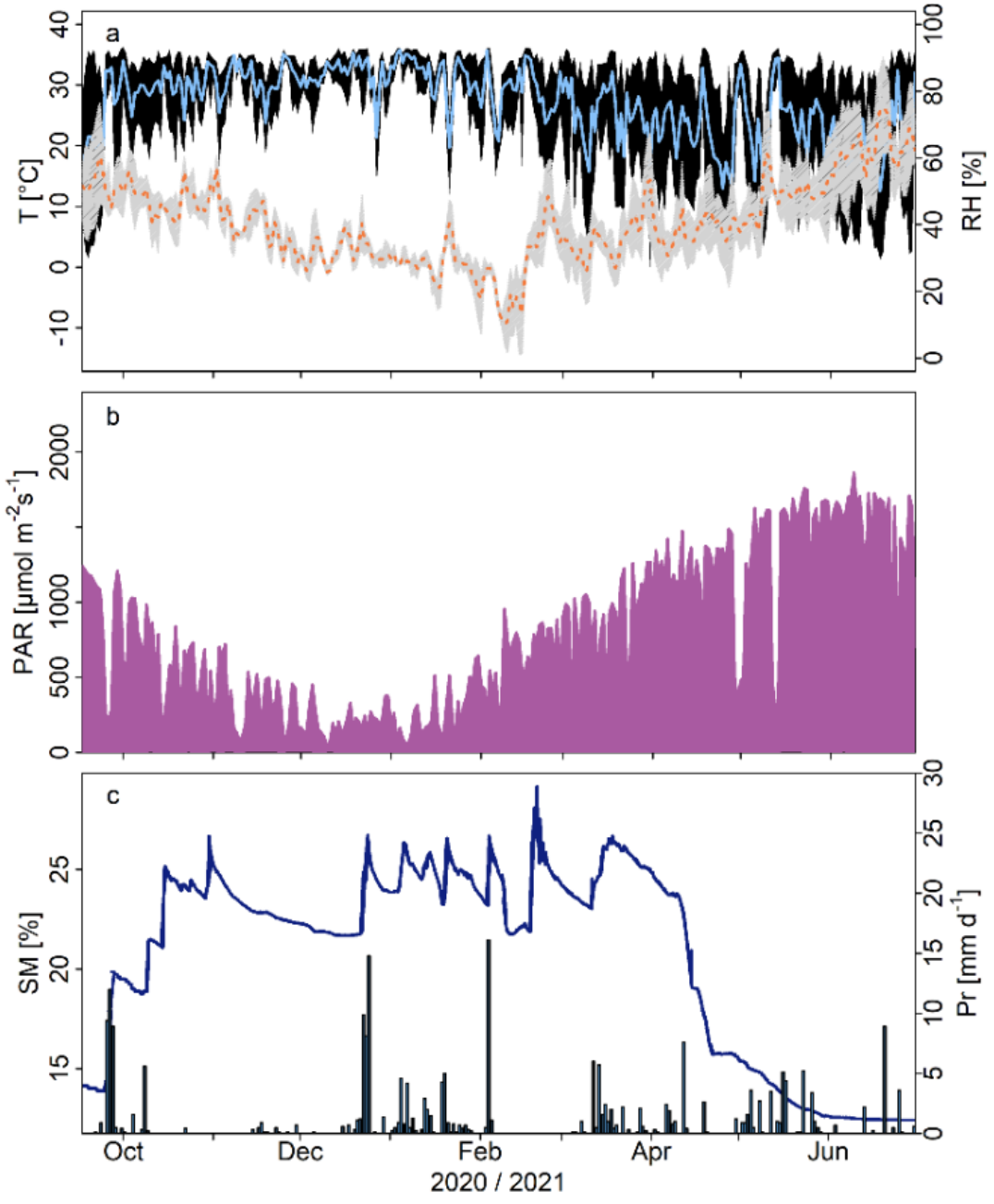


Figure 2: Environmental parameters during the measurement period with (a) daily mean temperatures (T; orange line; light ~~grey = associated~~gray = corresponding variation) and daily mean relative humidity (RH; black line; dotted lines = ~~associated~~corresponding variation), (b) incoming hotosynthetically active radiation (PAR; purple) and (c) soil moisture (SM; blue line) and precipitation (PR; blue bars).

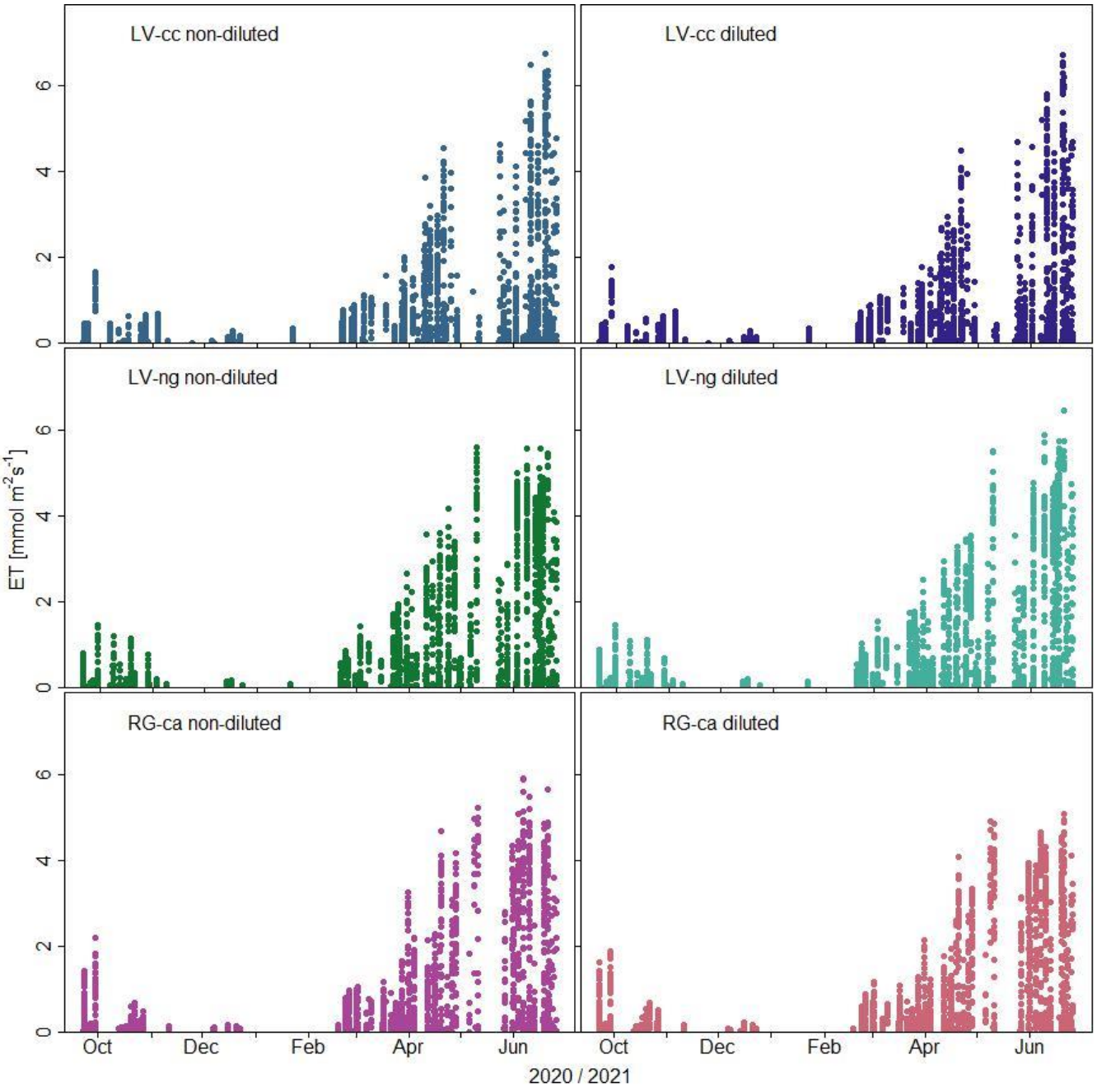


Figure 3: Measured and quality-screened (by soft and hard criteria) ET fluxes of the three soil types ~~(replicates summarized)~~ over the entire observation period (non-~~manipulated~~diluted treatments on the left, ~~manipulated~~diluted treatments on the right).

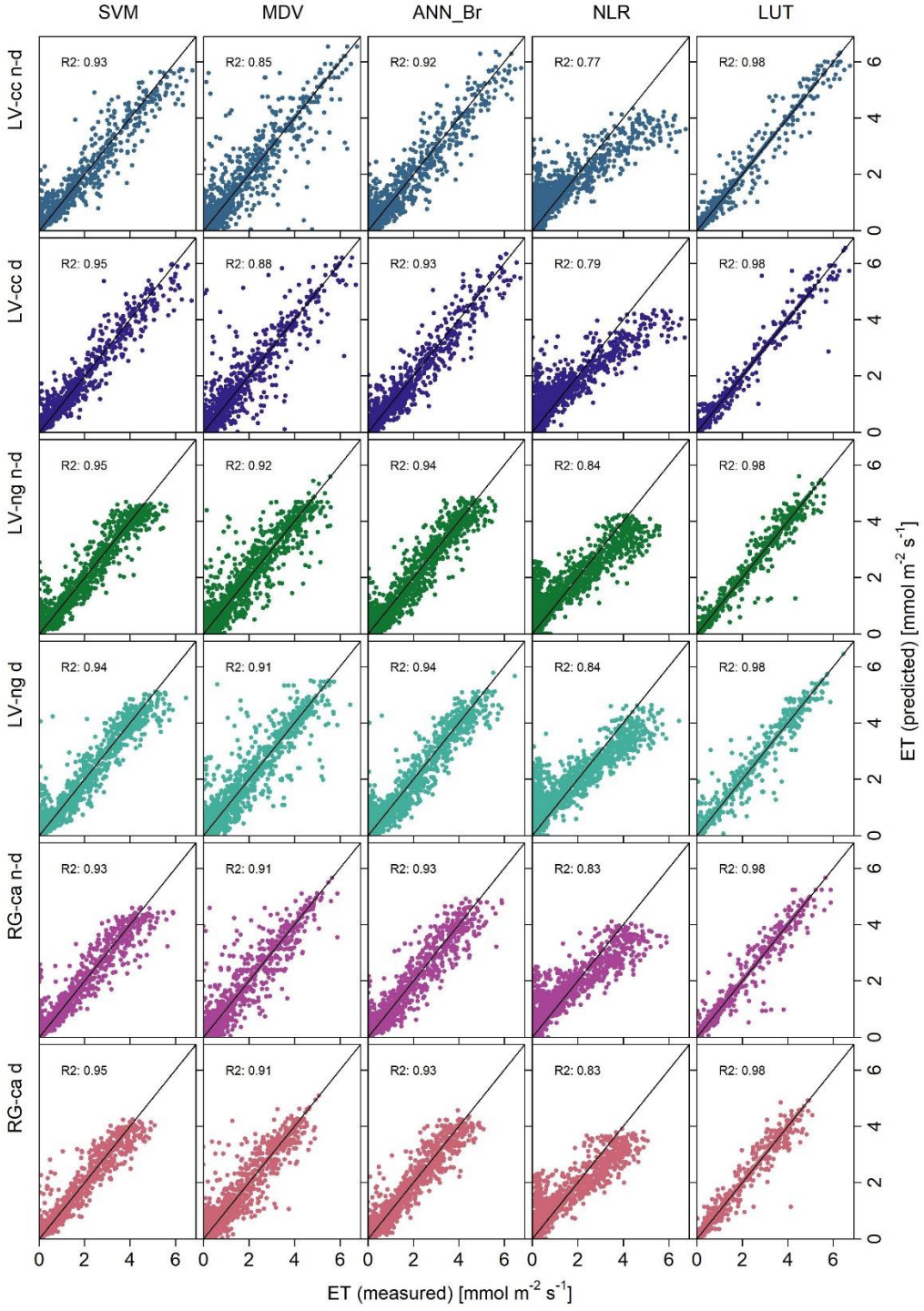


Figure 4: Comparison of the measured (bottom) with the predicted (left) ET fluxes and associated r-~~square~~squared values (R2) of the calibration results of all approaches. The black line ~~forms~~represents the 1/1 line. The different treatments are shown ~~above~~on top, the approaches on the right.

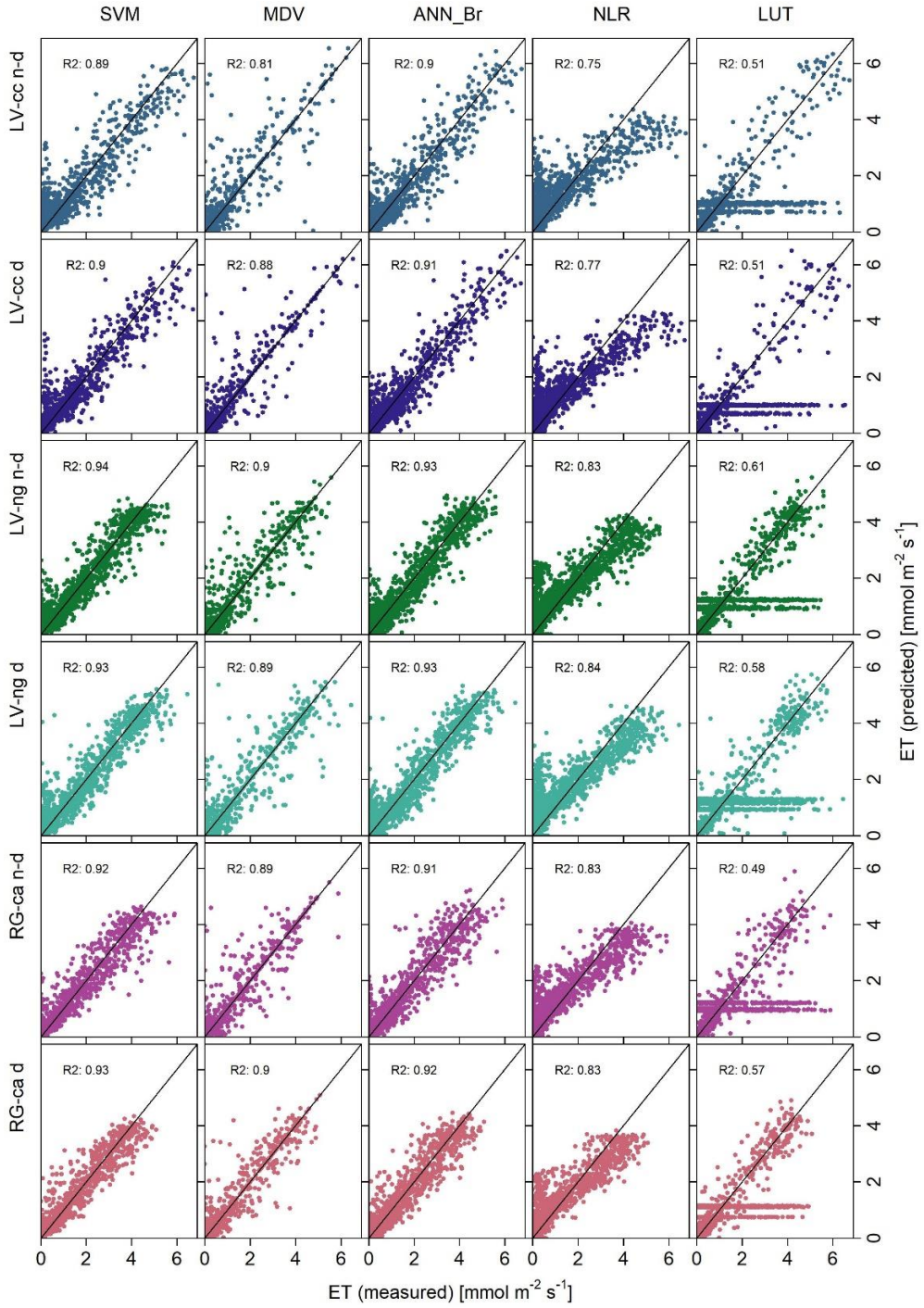


Figure 5: Comparison of the measured (bottom) with the predicted (left) ET fluxes and the associated r-~~square~~squared values (R2) of the validation results of all approaches. The black line ~~forms~~represents the 1/1 line. The different treatments are shown ~~above~~on top, the approaches on the right.

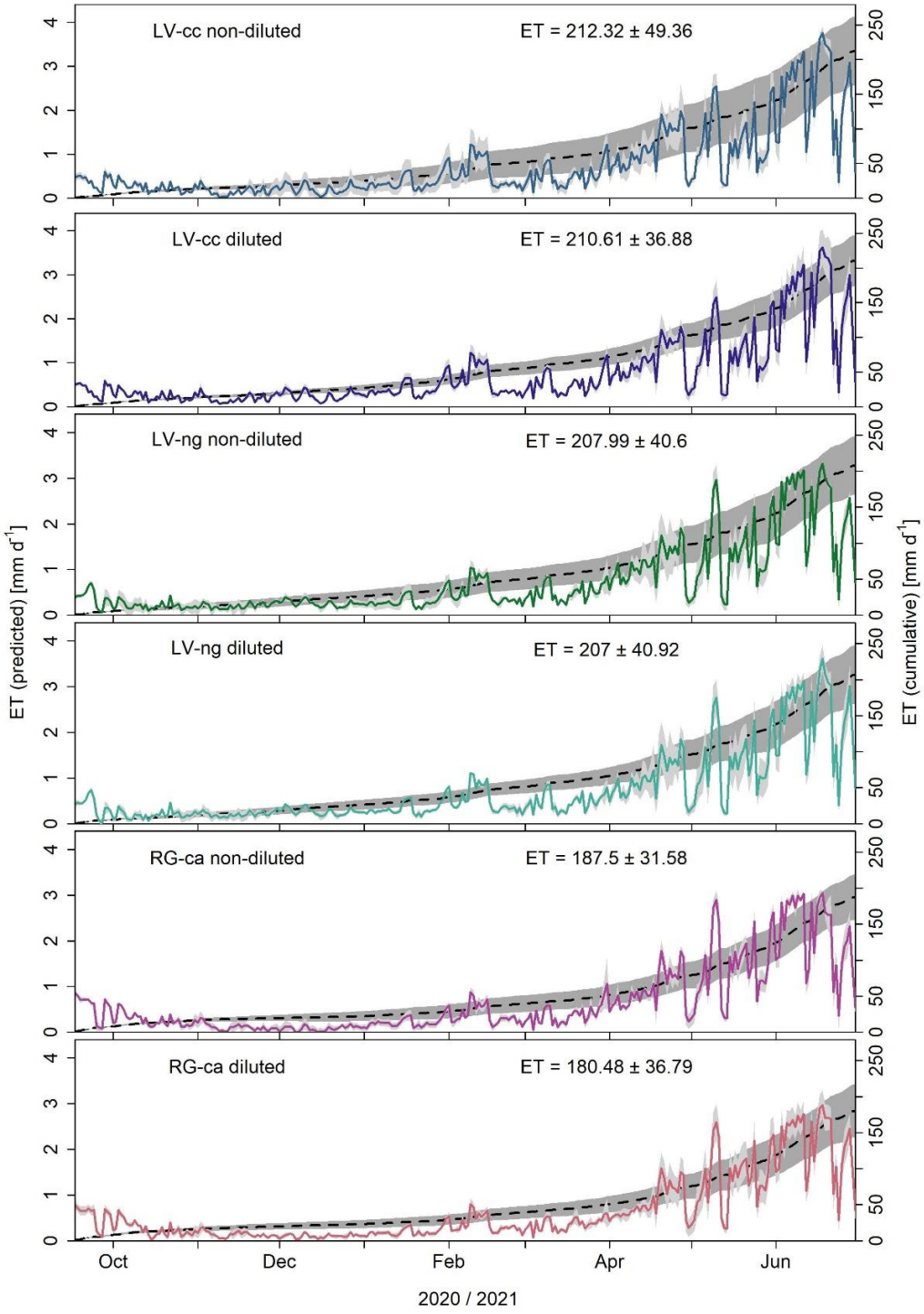
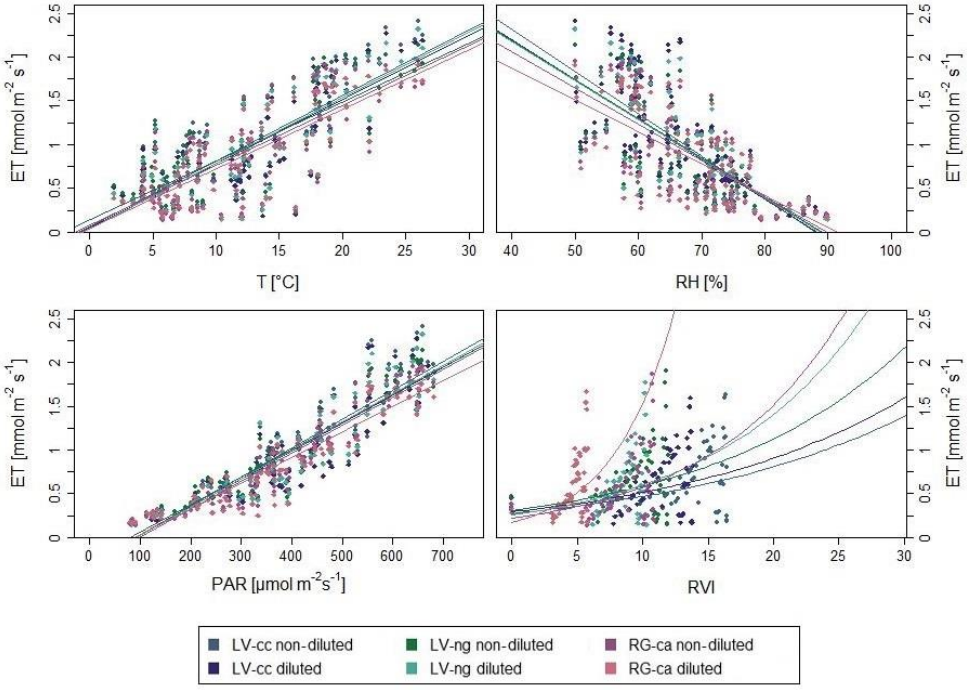


Figure 6: Daily mean ET sums (colored lines) of the different treatments and ~~mean~~ seasonal cumulative ET ~~budgets (~~(ET_{sum}; dashed lines) with ~~variations~~standard deviation between replicates (light and dark ~~grey~~gray).



Treatment	Parameter	r2	p
LV-cc n-d	T	0.59	1.19e-20
LV-cc d	T	0.61	1.59e-21
LV-ng n-d	T	0.56	5.63e-19
LV-ng d	T	0.63	4.94e-23
RG-ca n-d	T	0.58	2.16e-20
RG-ca d	T	0.65	3.26e-24
LV-cc n-d	RH	0.55	1.84e-18
LV-cc d	RH	0.53	1.30e-17
LV-ng n-d	RH	0.54	2.32e-18
LV-ng d	RH	0.50	1.70e-16
RG-ca n-d	RH	0.48	1.70e-15
RG-ca d	RH	0.41	4.94e-13
LV-cc n-d	PAR	0.82	9.61e-38
LV-cc d	PAR	0.82	3.38e-38
LV-ng n-d	PAR	0.87	6.50e-45
LV-ng d	PAR	0.85	3.88e-42
RG-ca n-d	PAR	0.90	1.74e-51
RG-ca d	PAR	0.86	3.72e-44
LV-cc n-d	RVI	0.09	2.67e-02
LV-cc d	RVI	0.07	4.59e-02
LV-ng n-d	RVI	0.07	4.73e-02
LV-ng d	RVI	0.08	3.20e-02
RG-ca n-d	RVI	0.13	7.56e-03
RG-ca d	RVI	0.21	5.17e-04

Figure 7: Relationship between ET and temperature (T) [°C], relative humidity (~~H~~RH) [%], photosynthetically active radiation (PAR) [µmol m^{-2} s^{-1}~~]~~], and ratio vegetation index (RVI) [mmol m^{-2} s^{-1}~~] as well as~~], and associated regression lines. Statistical values (r^2 and p) for the relationship between ET and response variables (environmental parameters) are presented in the table.

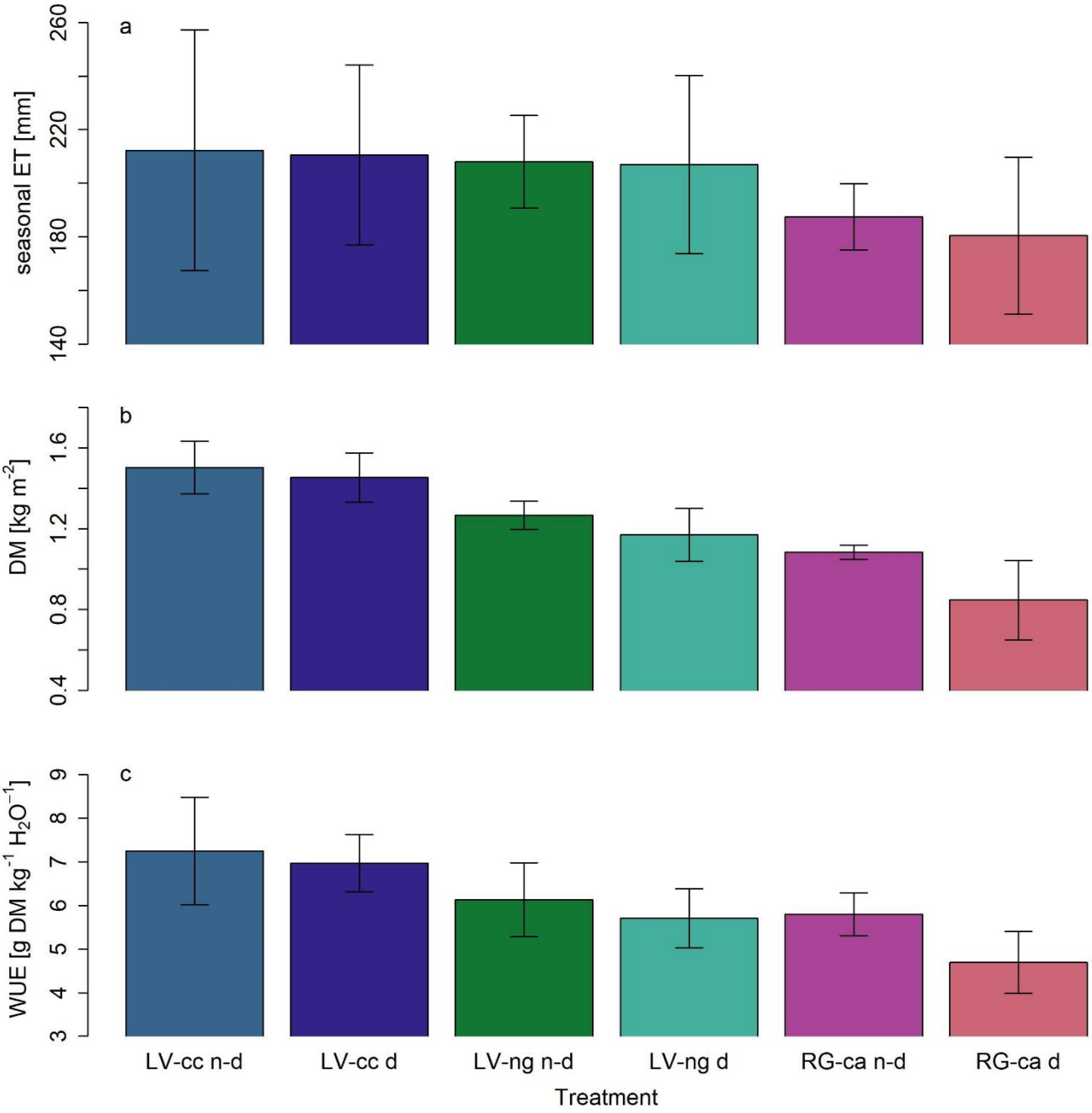


Figure 8: Averaged seasonal ~~ET totals~~cumulative ET (ET_{sum}) [mm] (a), harvest in form of dry mass (DM) [kg] (b), and WUE_{agro} of the different treatments and the associated standard deviation.

Appendix A:

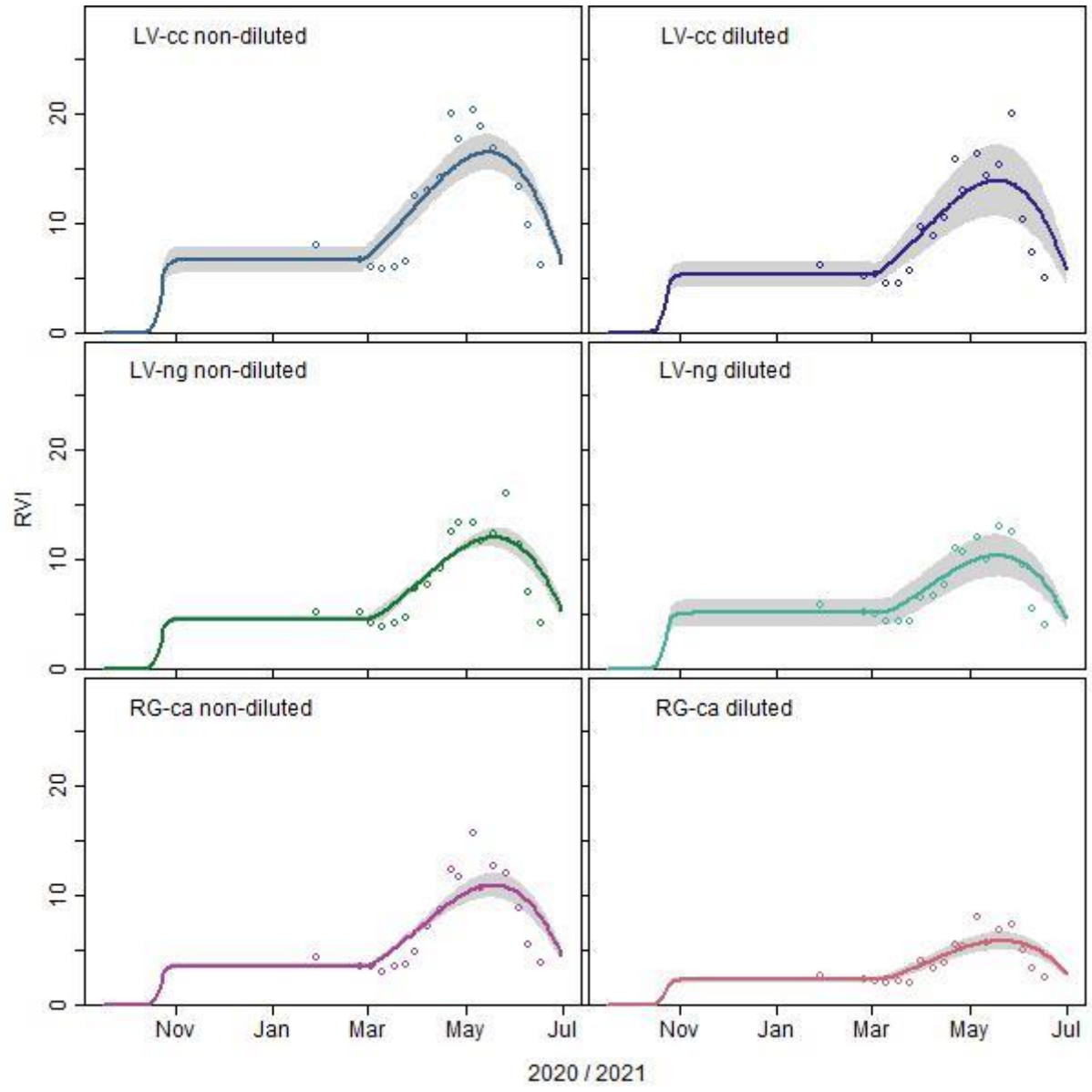


Figure A1: RVI fit (colored lines) of the different treatments with ~~variations~~the standard deviation between replicates (light ~~grey~~gray) and the corresponding averages of the daily measurements (points).

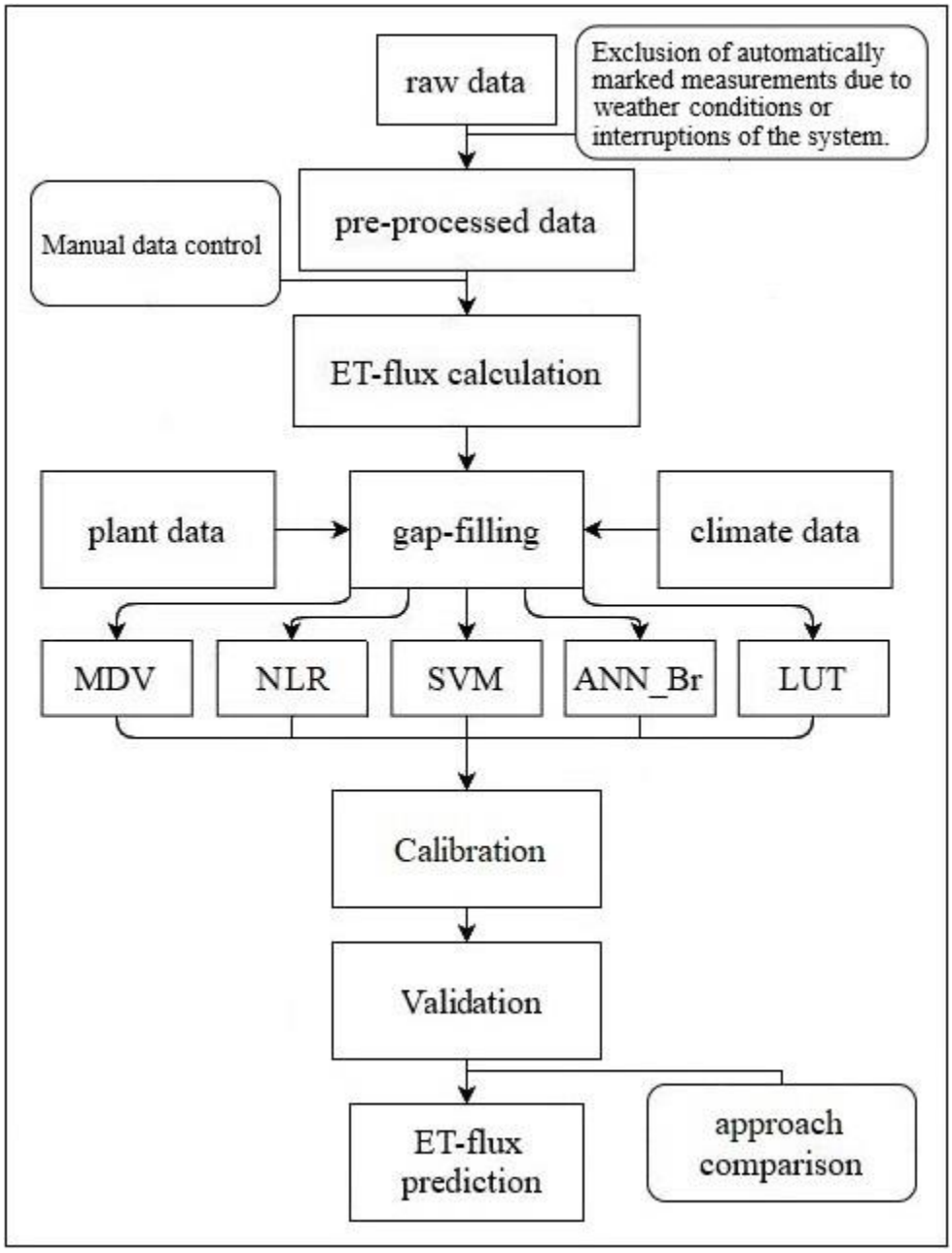


Figure A2: Schematic representation of the main steps of the presented data processing: raw data preparation was followed by a campaign-specific ET-flux calculation. Then, environmental parameters were used for gap filling using five different approaches. After calibration and validation, the most accurate approach was used for gap filling.

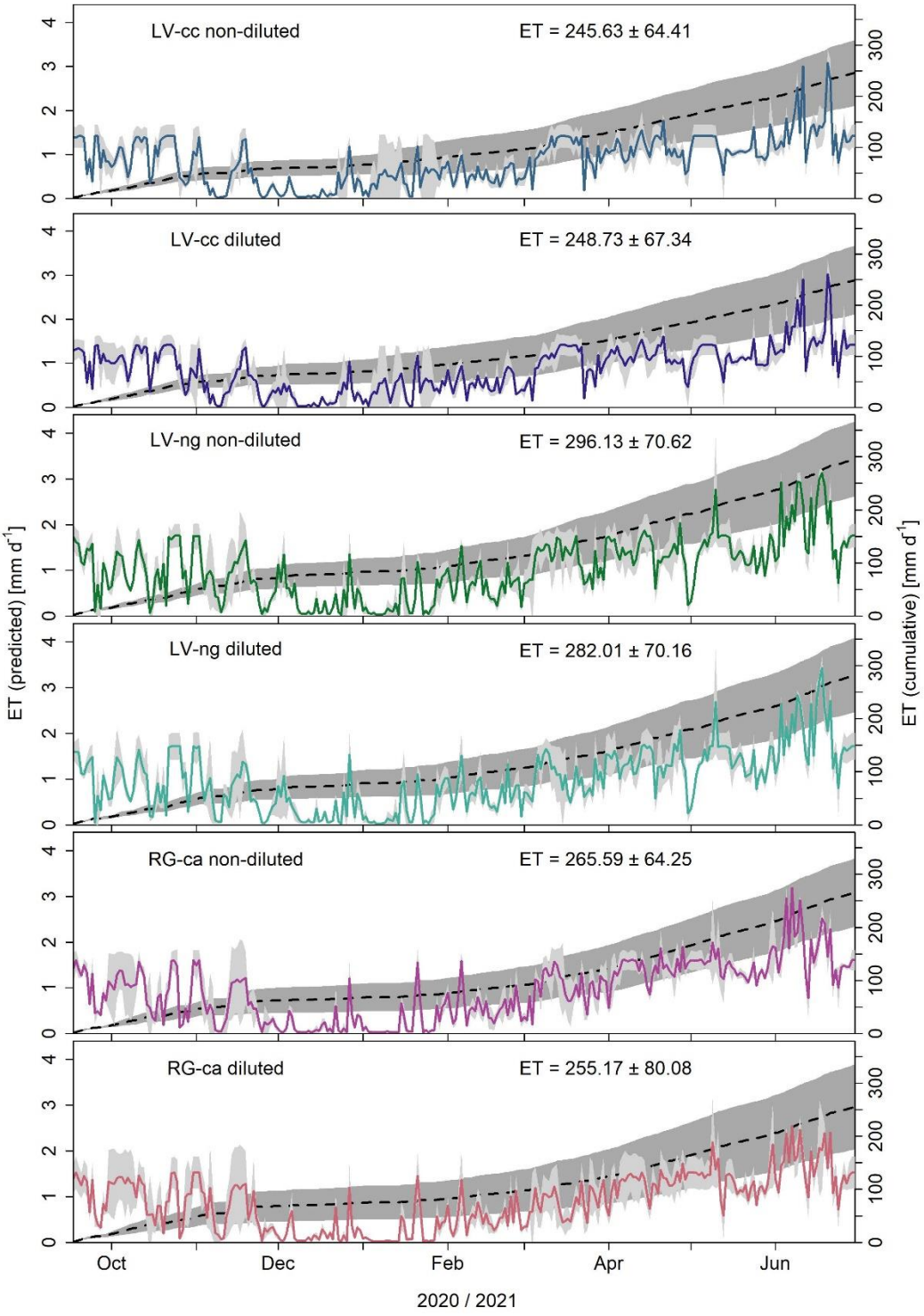


Figure A3: LUT predicted daily mean ET sums (colored lines) of the different treatments and seasonal cumulative ET (ET_{sum}; dashed lines) with standard deviation between replicates (light and dark gray).

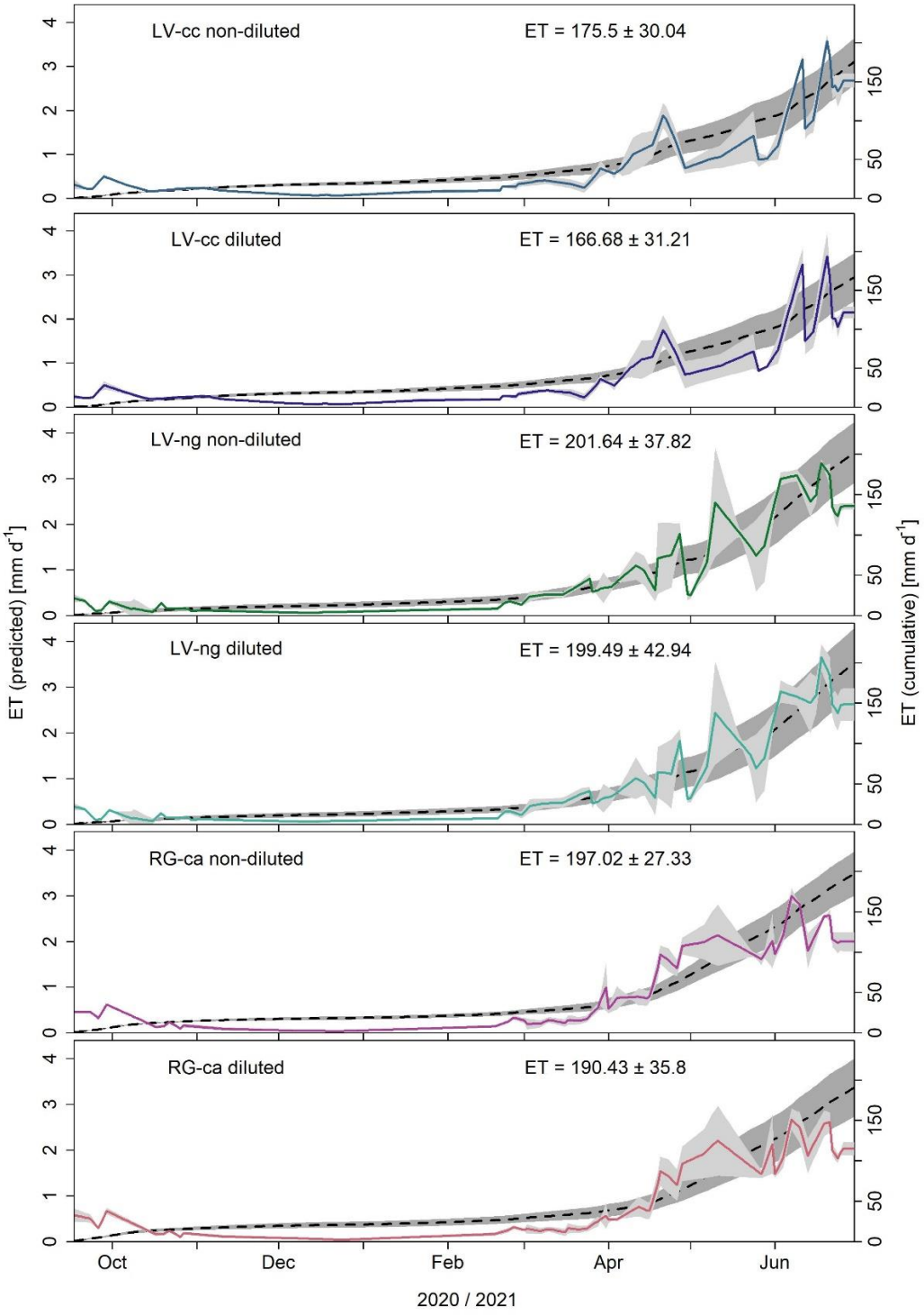


Figure A4: MDV predicted daily mean ET sums (colored lines) of the different treatments and seasonal cumulative ET (ET_{sum}; dashed lines) with standard deviation between replicates (light and dark gray).

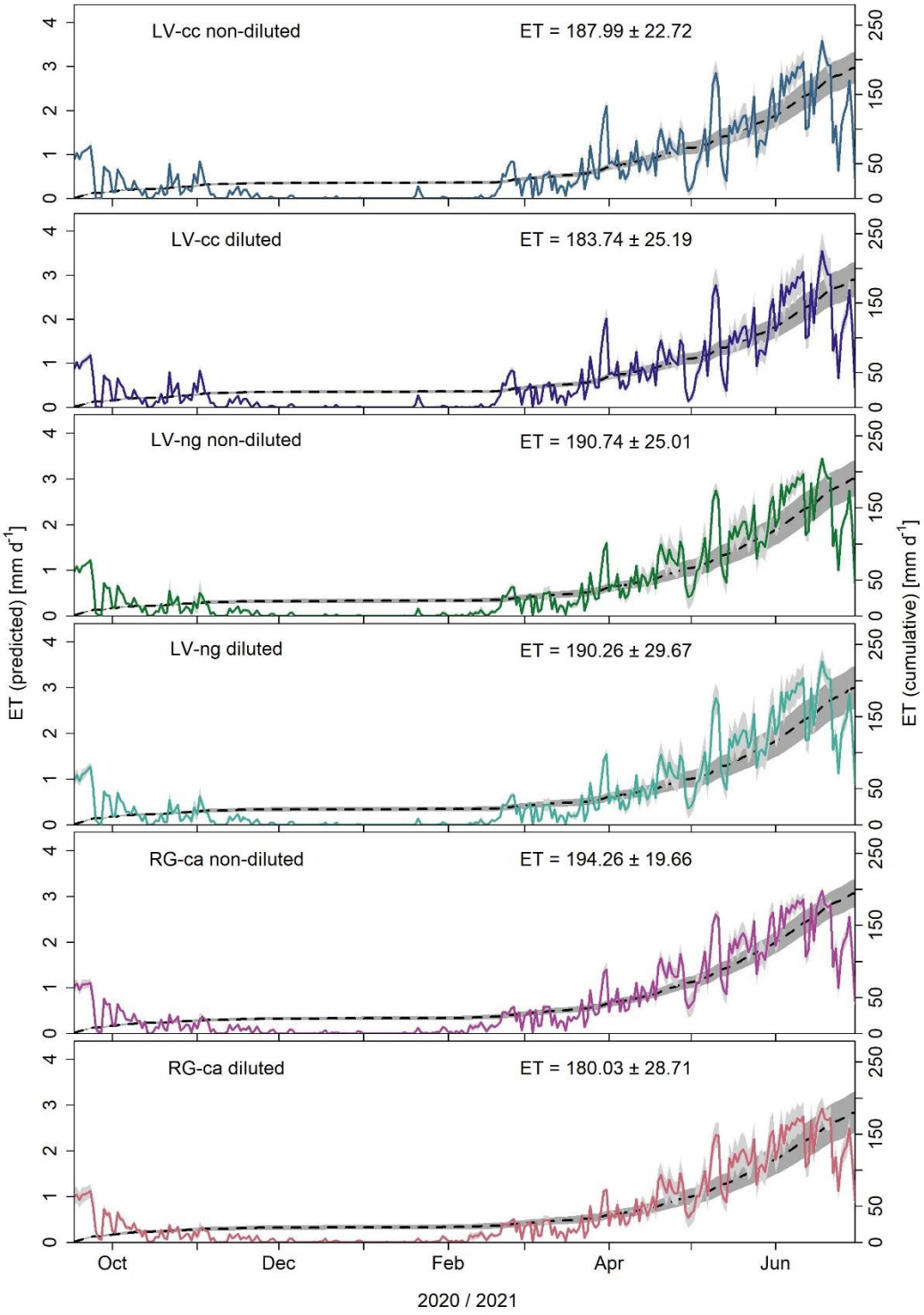


Figure A5: NLR predicted daily mean ET sums (colored lines) of the different treatments and seasonal cumulative ET (ET_{sum}; dashed lines) with standard deviation between replicates (light and dark gray).

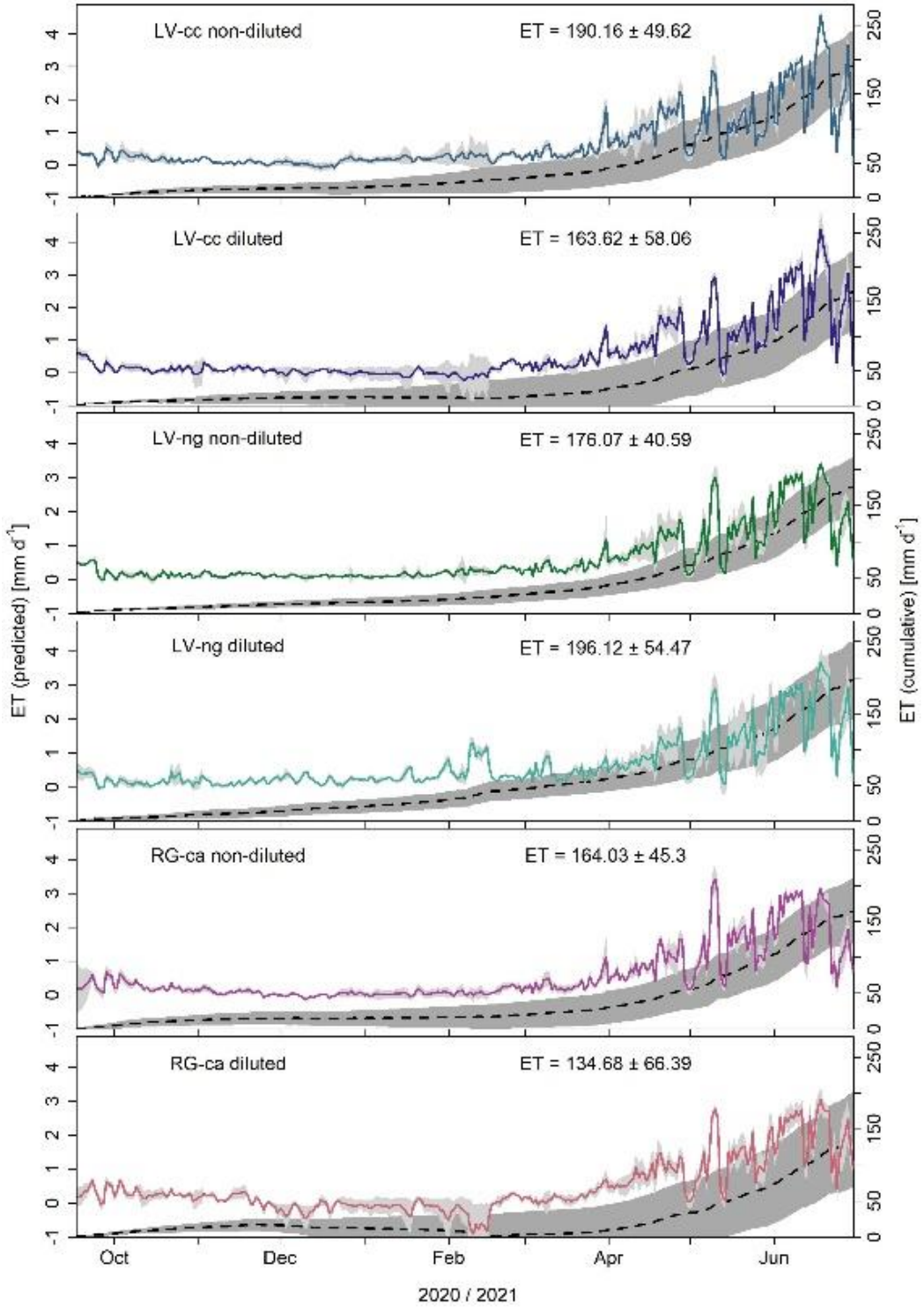


Figure A6: ANN_BR predicted daily mean ET sums (colored lines) of the different treatments and seasonal cumulative ET (ET_{sum}; dashed lines) with standard deviation between replicates (light and dark gray).

Appendix B:

Table B1: Fertilization information for the field.

Date	Amount	Details
15.10.2020	161 kg P_2O_5 ha^{-1}	applied on 6 plots of LL as TSP
22.03.2020	77 kg P_2O_5 ha^{-1}	as Triple Super Phosphate (TSP)
22.03.2020	259 kg K_2O ha^{-1}	as 40% grain potash
16.09.2020	30 kg N ha^{-1}	10 m^3 ha^{-1} digestate
10.03.2021	91 kg N ha^{-1}	30 m^3 ha^{-1} digestate
08.04.2021	45 kg N ha^{-1}	12 m^3 ha^{-1} digestate

Table B2: The number of measurements per treatment and the percentage of gap-filling data.

Plot		Measurements [n]	gap-filled [%]
LV-cc n-d	1	990	85.63
LV-cc n-d	2	624	90.94
LV-cc n-d	3	996	85.54
LV-cc d	1	624	90.94
LV-cc d	2	735	89.33
LV-cc d	3	989	85.64
LV-ng n-d	1	1210	82.43
LV-ng n-d	2	1210	82.43
LV-ng n-d	3	705	89.76
LV-ng d	1	718	89.58
LV-ng d	2	1215	82.36
LV-ng d	3	1205	82.51
RG-ca n-d	1	657	90.46
RG-ca n-d	2	772	88.79
RG-ca n-d	3	669	90.29
RG-ca d	1	669	90.29
RG-ca d	2	1130	83.59
RG-ca d	3	1129	83.61

Table B3: Used R packages and associated sources.

package	source
Akima	Akima & Gebhardt (2021)
Andrews	Myslivec (2012)
Base	R Core Team (2021)
Boot	Davison & Hinkley (1997)
Caret	Kuhn (2021)
data.table	Dowle & Srinivasan (2021)
e1071	Meyer et al. (2021)
FSA	Ogle et al. (2022)
ggplot2	Wickham (2016)
gridExtra	Auguie (2017)
gt	Iannone et al. (2022)
hydroGOF	Mauricio Zambrano-Bigiarini (2020)
Kernlab	Karatzoglou et al. (2004)
Lattrice	Sarkar (2008)
Lmtest	Zeileis & Hothorn (2002)
lookupTable	Jia & Maier (2015)
Lubridate	Grolemund & Wickham (2011)
Neuralnet	Fritsch et al. (2019)
Nortest	Gross & Ligges (2015)
Plotrix	J (2006)
Plyr	Wickham (2011)
Reshape	Wickham (2007)
Shape	Soetaert (2021)
Tibble	Müller & Wickham (2021)
tidyr	Wickham & Girlich (2022)
Vioplot	Adler & Kelly (2020)
webshot	Chang (2022)
Zoo	Zeileis & Grothendieck (2005)