Author reply to Referee comments from **Anonymous Referee # 1** from 21 January 2020 (https://doi.org/10.5194/bg-2019-432-RC1) on:

"Comparison of eddy covariance CO_2 and CH_4 fluxes from mined and recently rewetted sections in a NW German cutover bog" by David Holl et al.

Reviewer comments (RC)
Author comments (AC)
Mentioned line numbers refer to the originally submitted manuscript
Manuscript changes (MC)

This manuscript reports carbon dioxide and methane fluxes for the period June 2012 to May 2014. Using a combination of a single eddy covariance tower, footprint modeling, and manual spatial cover classification using remotely sensed images, the authors distinguish, separately gap-fill, and quantify annual sums for, both actively mined and recently rewetted peat sections. The authors find that rewetting increases methane and decreases carbon dioxide emissions, but those effects manifest themselves much more strongly in the second year after rewetting, indicating lags. Overall the paper is clearly written but could be much shorter. The strongest aspects of the study are the comprehensive scholarship and the clarity of the methods. For example, there is a clear description of eddy covariance data processing for methane, which seems to have been considered with great care, and is an active area of research in the flux community (e.g., European RINGO initiative, perhaps should be linked more specifically). The exploration of gap-filling approaches is also a nice addition, though I think it takes up too much of the paper overall, given that is not the primary focus of the study (not even in the title). There are however some issues with the paper that I think need to be addressed which I outline below.

Major Comments

Soil conditions
In year 2 the authors report a substantial amount of soil data being recorded, including temperature, redox, and water table height. These in turn are included via their variable selection procedure in the predictive models of methane flux. Unfortunately, these data are not presented to the reader at all. This is disappointing as the focus of the paper implied by the title is the difference in fluxes between the two cover types, and soil conditions are likely the mechanism underlying those differences by year 2. I would encourage the authors to explore visualizations of those soil data in the paper, perhaps by substituting it for some of the discussion of either the machine learning or the CO2 discussion.
As mentioned by the referee, our modeling approach does include an identification procedure for likely flux drivers. We also present a short (section 3.1) and extended (Appendix B) discussion on how these drivers can explain flux variability in a mechanistic way. Most likely because our data set was measured over heterogeneous terrain, we did not find a comparably simple flux—driver relation (e. g. with soil temperature or water table) which explained the observed flux variability to a sufficient degree so it could be used to gap-fill our high-frequency data in order to calculate annual flux balances. Due to the complexity of the flux data set, we decided to use a more complex modeling approach. Nevertheless, we agree that it is necessary to depict the site conditions more clearly so a reader can more easily compare to conditions at similar sites and grasp our data set quicker and more comprehensively. We therefore implemented a new modeling approach representing methane flux as a function of soil temperature and water table and explored the results in a new figure and an additional paragraph in section 3.1.

To further investigate the relation between CH_4 flux and the identified likely drivers, we fitted an exponential model of water table and soil temperature (in 40 cm depth) to the CH_4 fluxes from the rewetted section (see Figure XX1). With the exponential dependence of CH_4 flux on soil temperature, a fair amount ($R^2 = 0.55$) of the flux variability can be explained while the added water table term allows for the optimized temperature-F_{CH4} curve to take two distinct paths above and below an approximate water table threshold of 20 cm below the surface (see Figure XX1, panel A). Half-hourly flux variability is, however, substantial due to the heterogeneity of the site's surface and other confounding factors like for example the above-mentioned air pressure variations and is comparably better explained by our neural network models (see Figure XX1, panel B).

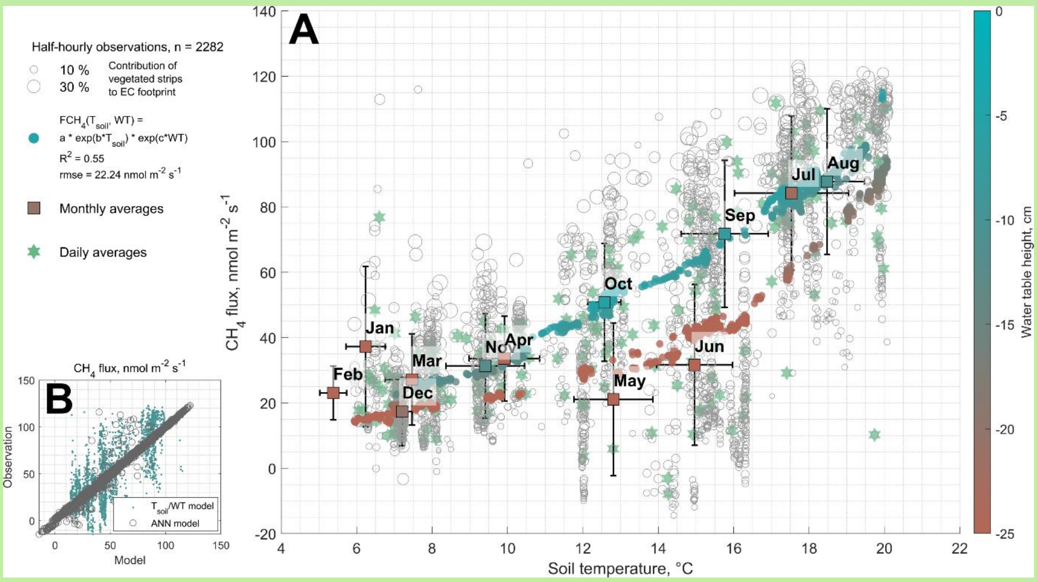


Figure XX1 *Panel A: Observed Half-hourly methane (CH4) fluxes from the rewetted section of Himmelmoor modeled as an exponential function of soil temperature in 40 cm depth and water table (FCH4(Tsoil/WT)). Monthly and daily flux and temperature averages are also given. Panel B: Comparison of a more complex artificial neural network (ANN) model with the exponential model from Panel A. Although methane flux variability can be explained by the exponential model to a reasonable degree, the level of complexity in flux—driver relations appears to be represented considerably better by the ANN.*

Flux Partitioning

Why was the net ecosystem exchange flux partitioning done at the monthly timestep? Can this not be performed at half-hour timesteps in EddyPro? I assume this was done intentionally but the justification is not clear.

No, flux partitioning is not a capability EddyPro provides. We applied our own model (Eq. 1, page 13), and used half-hourly fluxes to optimize the parameters of Eq. 1 (as stated in line 2, page 15). Due to the surface heterogeneity of our site, the independent parameters we chose differ from those used in other common approaches (e.g. Reichstein et al., 2005). We included half-hourly footprint characteristics and radiation but not temperature as drivers of net ecosystem exchange (NEE). Therefore, total ecosystem respiration (TER) in our model is a parameter. To depict the seasonal course of the parameter time series, we optimized one set of parameters for each month of the two-year data set (Fig. 4, page 15). The resolution of the ecosystem respiration time series is therefore limited to monthly steps. In an effort to yield NEE models, which are less likely to be overfitted, we reduced the number of model inputs and parameters and chose monthly, rather than for example daily, flux ensembles for optimization. We were not able to achieve a higher temporal resolution of TER confidently. We could, however, have calculated gross primary production (GPP) for half-hourly intervals using thirty minute radiation measurements and the two parameters Pmax and α (panels C and D in Fig. 4, page 15). Due to the focus of this paper (annual NEE balances and gap-filling), we decided to omit this step at this point. Instead, we compared the determined photosynthesis parameters, which directly relate to plant characteristics, to literature values of plants that also occur at our site in order to examine the credibility of our land use-specific gap-filling models (as stated on page 15, line 1).

Synthesis Literature summaries in the introduction and the discussion need to avoid listing. I am referring to the carbon dioxide flux sections, whereas the methane section is better synthesized (I especially like the comparison to IPCC values). The comparisons made in the results to other studies might be better tabulated. If they are noted in the main text, they should be synthesized better.
We agree, literature synthesis was a bit wordy in the running text. We created three new tables. One for the introduction that summarizes methane gap-filling methods and two in the results section (3.2) where we compare our CO_2 model parameters to literature values. We propose to move the tables to the appendix.

We replaced page 3, line 34 to page 4, line 12 with:
…the relations between environmental drivers and CH4 flux often appear to be more complex than for CO_2. An overview of methods applied in EC literature is given in Table A 1. Basic gap-filling methods include for example interpolation between measured values or the use of an average to replace all gaps. Simple linear models have also proven to be applicable in certain settings. A common approach is to fit Arrhenius-type non-linear functions to the flux as a function of various environmental drivers. However, as stated by Brown et al. (2014), there is evidence that these functional relationships do not necessarily behave monotonically. Artificial neural networks (ANNs) form a category of non-parametric models that have frequently been used to fill gaps in EC CO2 flux time series. Mostly, multilayer perceptrons (MLP) were chosen (Papale and Valentini, 2003; Moffat et al., 2007; Moffat, 2012; Järvi et al., 2012; Pypker et al., 2013; Menzer et al., 2015). Most recent literature on CH_4 flux gap-filling assess MLP models to be the most robust. MLPs are recommended within the processing for the pan-European Integrated Carbon Observation System (ICOS) by Nemitz et al. (2018) and for the new methane component of FLUXNET and the Global Carbon Project’s efforts better constrain the global methane budget respectively (Knox et al., 2019).

Appendix A: Gap-filling methods from literature

Table A1. Overview of methods applied in literature to gap-fill eddy covariance methane flux time series.

Method		References
Interpolation		Hanis et al. (2013); Dengel et al. (2011)
Averaging		Hatala et al. (2012); Mikhaylov et al. (2015)
Arrhenius-type non-linear functions	half-hourly	Kroon et al. (2010); Forbrich et al. (2011); Hommeltenberg et al. (2014); Goodrich et al. (2015)
	downsampled	Suyker et al. (1996); Friborg and Christensen (2000); Rinne et al. (2007); Long et al. (2010); Wille et al. (2008); Jackowicz-Korczyński et al. (2010); Parmentier et al. (2011); Brown et al. (2014); Shoemaker et al. (2015); Mikhaylov et al. (2015)
Look-up tables		Pypker et al. (2013); Hommeltenberg et al. (2014); Bhattacharyya et al. (2014)
Mean diurnal variation		Dengel et al. (2011); Jha et al. (2014)
Marginal distribution sampling		Alberto et al. (2014); Shoemaker et al. (2015)
Machine learning	Artificial neural networks	Dengel et al. (2013); Deshmukh et al. (2014); Knox et al. (2015); Goodrich et al. (2015); Nemitz et al. (2018); Knox et al. (2019); Kim et al. (2019)
	Support vector machines	Kim et al. (2019)
	Random forest	Kim et al. (2019)

I replaced page 15, line 1 to page 17, line 9 with:

As a fourth method to evaluate the applicability of our land use-specific flux decomposition, we fitted a combined respiration-photosynthesis model (see Eq. 1) to monthly ensembles of the half-hourly CO_2 SCTS in order to check if the resultant parameters are reasonable in relation to each other and to literature data. In general, the vegetation period, with its productivity maximum between June and July and its cessation between mid-October and November is well depicted in the seasonal course of the model parameters throughout both years. The parameter courses relating to the vegetated strips of the drained and rewetted areas (Figure 4, panels B – D) develop fairly similar. Distinctions between the drained and rewetted areas are more pronounced with respect to CO_2 release from bare peat surfaces (Figure 4, panel A). Ditch-blocking of a rewetted sector close to the EC tower (which therefore made up a large part of the EC footprint) was only performed one year before our measurements started. In summer of 2012 this area therefore was not yet permanently flooded leading to TER_{bare} fluxes exceeding those from the active mining site. From winter 2012/2013 on, inundation of the rewetted bare peat area progressively increased, resulting in lower TER_{bare} fluxes from the rewetted compared to the drained section. Our TER_{bare} fluxes are in concordance with findings from two studies that were also conducted on the active peat extraction area in Himmelmoor with manual chambers; TER data reported from similar peat extraction sites also agree with our results (see Table 3). As model includes the relative contributions of the vegetated strips to the EC footprint we compare the extracted model parameter time series (see Figure 4, panels B – D) with estimates of these plant species-specific values from other studies investigating similar plants and plant communities as found in the vegetated strips in Himmelmoor. Reported averages and ranges agree well with our findings (see Table 4). Additionally, we could distinguish between CO_2 release from decomposing bare peat (TER_{bare}, see Figure 4, panel A and Table 3) and from the vegetated strips (TER_{veg}, see Figure 4, panel B and Table 4) where respiratory CO_2 release also includes autotrophic respiration of plants. In our data set, TER is between twofold and fourfold larger in areas with than without vegetation. TER_{veg} from the rewetted area is mostly larger than from the drained area. Progressive inundation led to a hydrological connection of SC_{veg} and the flooded bare peat areas. An increased input of dead plant material as a result of higher water tables might have promoted heterotrophic respiration. Hampered plant productivity due to flooding is also expressed in lower peak values of P_{max} at the vegetated strips of the rewetted site.

Table 3. Comparison of total ecosystem respiration fluxes from bare peat areas without vegetation (TER_{bare}) between our study (see Figure 4, panel A for full time series) and literature values (closed chamber methods) from the same and similar peat extraction sites.

Reference	Site: Land use	Site: Name, Location	TER_{bare}, µmol m^{-2} s^{-1}
This study	Active mining	Himmelmoor, NW-Germany, 53°N	1.1 ± 0.5 (annual average and standard deviation)
	Ceased mining, rewetted		0.8 ± 0.7 (annual average and standard deviation)
Vanselow-Algan et al. (2015)	Active mining		0.5 ± 0.1 (annual average and uncertainty)
Vybornova et al. (2019)	Active mining		0 to 1 (annual range), 3 (maximum)
	Ceased mining, rewetted		0 to 0.5 (annual range), 1.4 (maximum)
Waddington et al. (2002)	Ceased mining, wet year	Sainte-Marguerite-Marie, SE Canada, 48°N	0.8 (May to August average)
	Ceased mining, dry year		3 (May to August average)
Shurpali et al. (2008)	Active mining	Linnansuo, SE Finland, 62 °N	1.3 (end of August maximum)
			0.2 (mid-November minimum)

Table 4. Total ecosystem respiration (TER_{veg}, µmol m^{-2} s^{-1}), maximum photosynthesis (P_{max}, µmol m^{-2} s^{-1}) and initial quantum yield (α, dimensionless) from the vegetated strips of this study compared to literature values from similar plant species. As the literature record of combined plant and soil respiration measurements of the species that occur at the site of this study is limited, autotrophic respiration respiration (R_a, µmol m^{-2} s^{-1}) estimates of plants from the same genera are also given. Note that R_a values were determined on leaf scale and therefore refer to leaf area. Since shrubs and trees can have a leaf area index larger than 1, fluxes referring to ground surface area could be higher. Model parameter values from this study are given as averages and standard deviation. The latter statistic expresses the value range throughout two annual courses as shown in Figure 4 rather than parameter uncertainty.

Reference	Plant species	TER_{veg}	R_a	P_{max}	α
This study, drained section	*Betula pubescens, Salix spp., Eriophorum vaginatum, E. angustifolium, Molinia caerulea, Calla palustris, Typha latifolia, Carex spp., Juncus effusus, Calamagrostis canescens*	2.5 ± 2.9		14.4 ± 13.8	0.04 ± 0.04
This study, rewetted section	—"—	3.0 ± 2.1		10.5 ± 10.6	0.05 ± 0.03
Vanselow-Algan et al. (2015)	*Molinia caerulea, Betula pubescens, Eriophorum angustifolium*	> 10 (summer)			
Beyer and Höper (2015)	*Molinia caerulea*	≤ 7		15 to 30	
	Eriophorum angustifolium	≤ 5		20 to 70	
Patankar et al. (2013)	*Salix pulchra*		≤ 2		
	Eriophorum vaginatum		≤ 3		
	Carex bigelowi		≤ 1		
Körner (1982)	*Carex curvula*		1		
Murchie and Horton (1997)	*Carex flacca*		1.5		
Kaipiainen (2009)	*Salix dasyclados*		0.8 to 1.2		0.04 to 0.08
Patankar et al. (2013); Gu et al. (2008)	*Betula spp.*		1 to 5	10 to 15	
Lienau (2014)	*Betula pubescens*			32 to 41	
Nygren and Kellomäki (1983)	*Betula pubescens*			4 to 17	
Hoogesteger and Karlsson (1992)	*Betula pubescens*			8	
Chen et al. (2010)	*Typha latifolia*			25	0.02 to 0.07
Ögren (1993)	*Salix spp.*			16 to 29	
Vernay et al. (2016)	*Molinia caerulea*			7 to 15	0.03

Machine Learning In Appendices A and B, the authors outline the machine learning approach used (artificial neural networks). Can the authors justify why they used a single data split as opposed to a k-fold cross validation approach, which tends to give a more stable performance evaluation? Using the alternative year as a "test" set for generalizability is interesting. Can the authors also comment on whether gaps were artificially created during validation. or whether the data splits were performed randomly on all observations?

I assume the referee refers to the division of target data (fluxes) into training and validation sets in the course of network optimization. To my understanding, we actually did use a simple 2-fold cross validation by dividing the data set into two groups. Due to the large number of fluxes (especially in case of methane) that we discarded during quality filtering, a division into more groups would have resulted in a lower number of fluxes per group, impairing network training. As we used ensemble averages of 1000 networks and therefore performed network optimization 1000 times, we also (randomly) divided the target data differently each time and in my opinion sufficiently counteracted effects of overfitting by this proceeding.

Style
I personally enjoyed the descriptive style of the writing, but it is unfortunately much too verbose for a modern readership. I would encourage the authors to mercilessly edit to reduce text. They might be surprised how much shorter the paper is if written in a more declarative style.
Thank you for the feedback. I agree and cut down on verbosity by replacing large parts of running text with tables as suggested in previous comments from Referee #1.

An example:
"We used a factor of 34 to convert FCH4 into CO2e release. This value is given in the Fifth Assessment Report of the Intergovernmental Panel on Climate Change (IPPC AR5, Myhre et al., 2013), refers to a 100-year time horizon and includes climate–carbon feedbacks"
...could be shortened to:
"We used a CH4 global warming potential of 34 kg CO2-eq (IPCC AR5, Myhre et al., 2013), which assumes a 100-year time horizon and includes climate-carbon feedbacks."
Although I understand the referee's general notion, I do not think this is a good example. The edited sentence is shorter mostly because the abbreviation "IPCC AR5" is not explained. In my opinion, also commonly used and widely known abbreviations should be explained when they first occur for consistency.

Or:
"Nevertheless, on an annual basis the periods when the sink character of SCrew prevails do not compensate for CO2 release during periods of reduced plant activity."

"Nevertheless, annually integrated ecosystem respiration outweighs photosynthesis in SCrew."
I agree, sentence replaced.

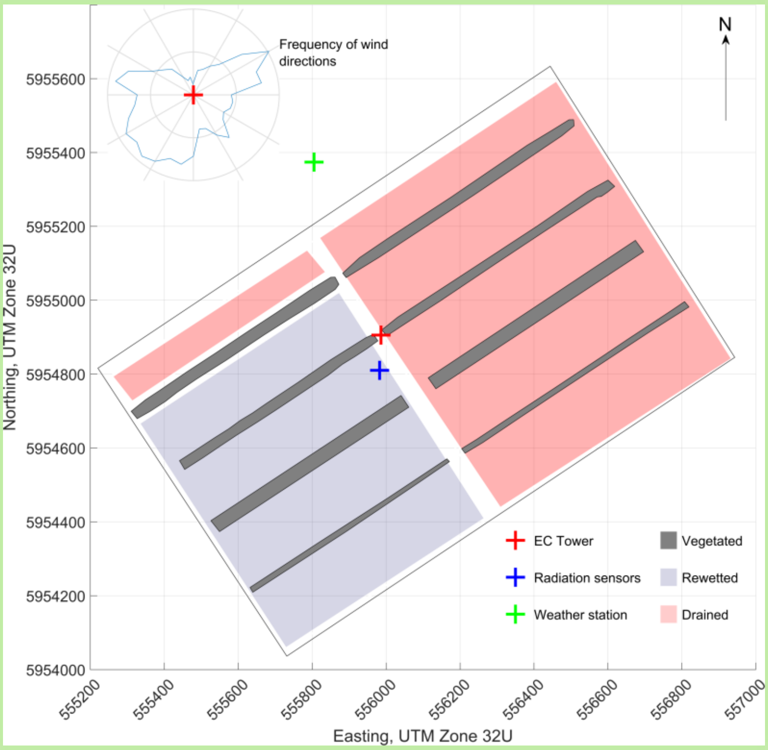
Minor Comments
Page 2
Line 9: Perhaps "land-use or climate change" rather than "men"
Changed
Line 14: "of carbon dioxide"
Changed
Line 20: "inheres the potentials" is ambiguous phrasing
"Inheres" replaced with "has".
Line 29: perhaps "oxidized" rather than "decomposed"
"decomposed" replaced with "converted to CO_2".
Page 3
Line 3: perhaps "strongly" rather than "gravely"
Changed
Lines 4-20: This is a nice minireview, but could be stronger if structured more systematically, or if the points could be linked more, to sound less like a list.
No change made. The questions are: What is known from literature about the development of methane emissions after peatland rewetting? What is to be expected for a largely vegetation-free site like Himmelmoor? To me the structure is systematic and the points are linked. I am not sure what to change.
Page 7 Line 14: "brown"
Changed
Lines 20-: Can you briefly justify the variable positions of these sensors? How representative is the water-level sensor of the general footprint?
I extended the description of sensor positions:

“A second HMP45 was installed together with a NR01 4-component net radiometer (Hukseflux, Netherlands) 70 m southwest of the EC tower on a tripod at 2~m height. The radiation sensors were not mounted on the EC tower because the field of view of the downward-facing sensors would have covered the peat dam and therefore not be representative for a dominant surface type at the site. These additional HMP45 and NR01 data were logged on a CR-3000 (Campbell Scientific, UK). Another logger of this type was used at the weather station which was taken over from a previous project and for data consistency was left at a position approximately 500~m north of the EC tower.

The water level within the footprint is highly variable as the surface consists of drained and rewetted sections. Our single sensor is representative for the rewetted bare peat strip to the southwest of the EC tower making up a large part of the EC footprint when wind comes from southwesterly directions.

Figure 1: Can you please add some more points for the other sensor installations.

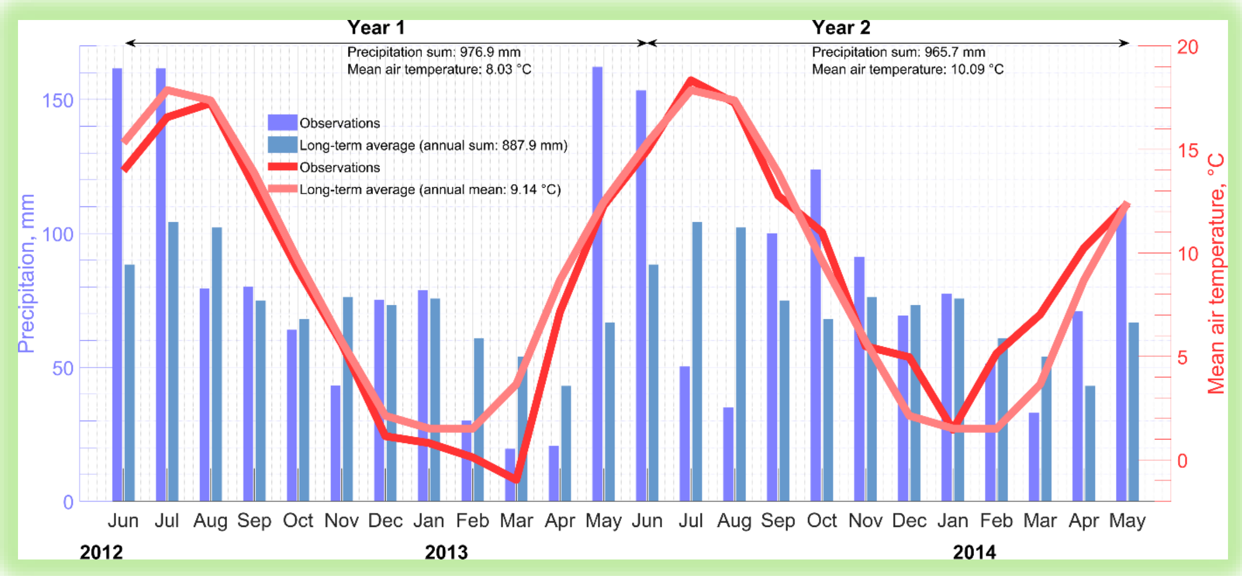
I updated the map with locations of all measurement systems



Page 8

Figure 2: This figure can be more useful to visualize how each true calendar year deviates from the long-term average if it just showed the full timeseries in one series (June 2012-May 2014). The problem currently is that it is difficult to visually integrate the deviations from the mean.

Ok. Figure restyled.



Line 8: Is WPL strictly a correction?

It is true that there is a discussion about this topic as compensation for air density fluctuations can also be seen as part of the eddy covariance method itself and not as a post-processing step and therefore does not qualify as a correction. On the other hand, the term WPL-correction is still widely used in the community, likely for historical reasons. I did not change this terminology.

Page 18 Table: Acres are not SI units. Please report in m2, hectares (ha), or km2

"a^{-1}" stands for "per annum/year". We do report area in m^2. No change made.

Line 33: I think the values in parentheses should be reversed given the order of the sentence.

True, thank you for the hint, order was reversed.