Author reply to Referee comments from **Anonymous Referee # 1** from 19 November 2018 (https://doi.org/10.5194/essd-2018-98-RC1, 2018) on:

A long-term (2002 to 2017) record of closed-path and open-path eddy covariance CO₂ net ecosystem exchange fluxes from the Siberian Arctic
by David Holl et al.

Reviewer comments (RC)
Author comments (AC)
Manuscript changes (MC)
Please note that figure numbers in this document refer to figures in this document and not to the numbering in the original draft unless denoted otherwise.

RC 1
1.) Inclusion of a 'scientific overview' In the 'Site description' Section, the first 5 paragraphs give a comprehensive overview on the site conditions, while the last paragraph is clearly detached from this material, and in its present form does not belong there. Still, I believe it will be of use to the reader to demonstrate what has been found so far based on the flux time series presented in this manuscript. My recommendation is to move this paragraph to a new chapter 4, i.e. between methods and data availability, and extend it to a length of 3-4 paragraphs in total. This would give ample room to summarize the main findings based on Samoylov eddy-covariance (and other) data so far, therefore highlighting the value of the dataset presented herein, and the role of the site in general for Arctic climate change research.

I moved and extended the paragraphs about scientific findings from the "Site Description" section to a new chapter 5 "Scientific overview" before the conclusions as suggested. Please also note my reply to RC 4 below.

While results on methane exchange fluxes and the soils' methane production and oxidation potential are more prominent in the publication record (e.g. Wagner et al., 2003; Kutzbach et al., 2004; Liebner and Wagner, 2007; Knoblauch et al., 2008; Sachs et al., 2008; Wille et al., 2008; Schneider et al., 2009; Sachs et al., 2010; Liebner et al., 2011; Knoblauch et al., 2015), literature on CO₂ flux time series recorded with the same measurement system presented in this publication is available for distinct years. Flux processing has, however, been streamlined only now. The length of the time series, the addition of detailed footprint information, the site-specific correction of OP fluxes and the coherent processing and quality filtering distinguishes the data set at hand from past publications like the contribution made to the FLUXNET2015 data set (Kutzbach et al., 2015).

Ongoing analysis of the long-term data set (Kutzbach, unpublished) *inter alia* confirms what has been found in the past (Kutzbach, 2006; Kutzbach et al., 2007; Runkle et al., 2013). The polygonal tundra of Samoylov Island appears to be a robust growing season CO₂-C sink whereas this sink strength can vary that much interannually that prolonged low-level respiratory CO₂-C loss during the cold season can offset CO₂-C uptake during the vegetation period. Reduced summer uptake has been observed for both the coldest and warmest summers. Runkle et al. (2013) found that with frequent early season heat spells, the temperature-induced increase in respiratory release can exceed the rise in photosynthetic uptake. Recently, all data from this publication has been contributed to the Arctic Data Center's chamber and EC synthesis project *Reconciling historical and contemporary trends in terrestrial carbon exchange of the northern permafrost-zone* that aims at identifying seasonal and interannual C flux dynamics and its drivers based on a newly established pan-arctic data base.

In context with the improvement of earth system models (ESMs), carbon dioxide fluxes from Samylov Island can be especially of use due to the site's comparably high moss cover. Using data from Samoylov, Chadburn et al. (2017) found that current ESMs miss an observed early season CO₂ uptake peak suspected to be connected to the earlier onset of moss photosynthesis in comparison with vascular plants. Although there have been advances and e. g. Porada et al. (2013) developed a dynamic moss model for JSBACH (Raddatz et al., 2007), Chadburn et al. (2017) noted that the simulated CO₂ uptake and release terms combining vascular vegetation and moss carbon fluxes did not agree with observational data. The fact that the Samoylov Island NEE data set has now been extended and its

quality has been greatly improved holds the opportunity to estimate the performance of updated ESM versions that are set up to represent carbon fluxes in the moss layer better.

RC 2

2.) Ensure that tower locations do not disrupt continuous time series The combination of text in Section 3.1, Figure 1 and Table 1 provides a good overview on the different site setups used to form this 16-year data record. However, the material also raises the question how the shifts in tower position and sensor configuration, including sensor height, may have influenced the signal captured by the EC system, and therefore maybe biased the long-term time series. I therefore recommend moving Section 3.6 upward as a new Section 3.2, and extending the discussion of the footprint issue. You can use parts of the conclusions section for this, but more details need to be provided how the shifts in landscape element fraction in the footprints may have compromised the continuity of the flux observations. See also my comment on Section 3.6 in the 'line comments' below.

I added a new "Discussion" section to the manuscript addressing the effects of tower location shifts and other possible disruptions of the time series' coherency.

Although we did our best to ensure the consistency and appropriateness of the data processing workflow for the presented NEE time series, due to technical and logistical constraints during 16 years of field work, disparities in the experimental setup exist which may challenge its integrity. The EC tower was relocated twice, the measurement height was changed three times (see Figure 1 and Table 1 *(in original draft)*). These changes of tower location and measurement height affected the source area and hence the surface types sampled during flux measurements. Most notably, between July 2007 and June 2009, the EC tower was placed about 650 m south-west of its original position at the center of Samoylov Island, in an area with an increased coverage of the surface class wet tundra. This is revealed by the footprint analysis (Figure 1). While the EC footprint is dominated by the surface class *dry tundra* throughout the time series, during subperiods 2007, 2008 and 2009 I the contributions of *wet tundra* to the measured flux are significantly higher.

To check the effect of the shifts in tower location and measurement height on cumulative CO_2-C fluxes, we calculated flux sums for a period when flux time series without gaps were available in most years. The overlapping period covers days of year 200 to 234, i.e. part of the growing season in all years except for 2004 (see Figure 2). Interannual variability of cumulative C fluxes in years with constant tower location (and measurement height) appears to be large and driven by a more complex set of variables than shifts in surface class contributions only. Flux sums from the periods when EC tower relocation led to a significant shift in EC footprint composition are well within the range of the distribution of cumulated fluxes from years with a more homogeneous EC fetch area. We therefore assume that, at least with respect to budget calculations, the presented long-term time series is not disrupted and can be regarded as representative for a polygonal tundra site dominated by dry tundra. For a more in depth analysis of flux dynamics, footprint information should and can be considered by users of the data set. Recently, a comparison between surface class level NEE models based on chamber measurements with EC fluxes, using the half-hourly footprint information provided in this data set for scaling, yielded good agreement between the results obtained with both methods Eckhardt et al. (2018). We regard the availability of half-hourly footprint information in the presented NEE data set an attribute that sets it apart from other studies and holds chances for comprehensive analyses.

Apart from the changes in anemometer height, other deviations of the general instrument setup occurred due to limitations in data storage during two winter periods when the acquisition frequency was reduced to 5 Hz and 10 Hz respectively. Rinne et al. (2008) demonstrated in a field experiment that fluxes calculated from raw data recorded at frequencies below 20 Hz compare well with fluxes derived from high frequency raw data. Differences arise as an increase of random noise and not as a systematic bias. High frequency noise removal before ensemble spectra estimation in EddyPro is effective in limiting the effect of increased noise on the quality of transfer function estimation in the process of spectral correction. Overall spectral correction in EddyPro is expressed as a spectral correction factor SCF which comprises the effect of all applied compensations for high and low frequency loss. Raw fluxes are multiplied with the respective SCFs during processing. We compared the SCF distributions of the two above mentioned winter periods with statistics of the remaining parts of the time series when data was recorded at 20 Hz. SCF deviations between the different acquisition frequencies are minor (see Figure 03) implying that systematic differences between fluxes calculated form raw data of different temporal resolutions are in fact small, random uncertainties increase, however.
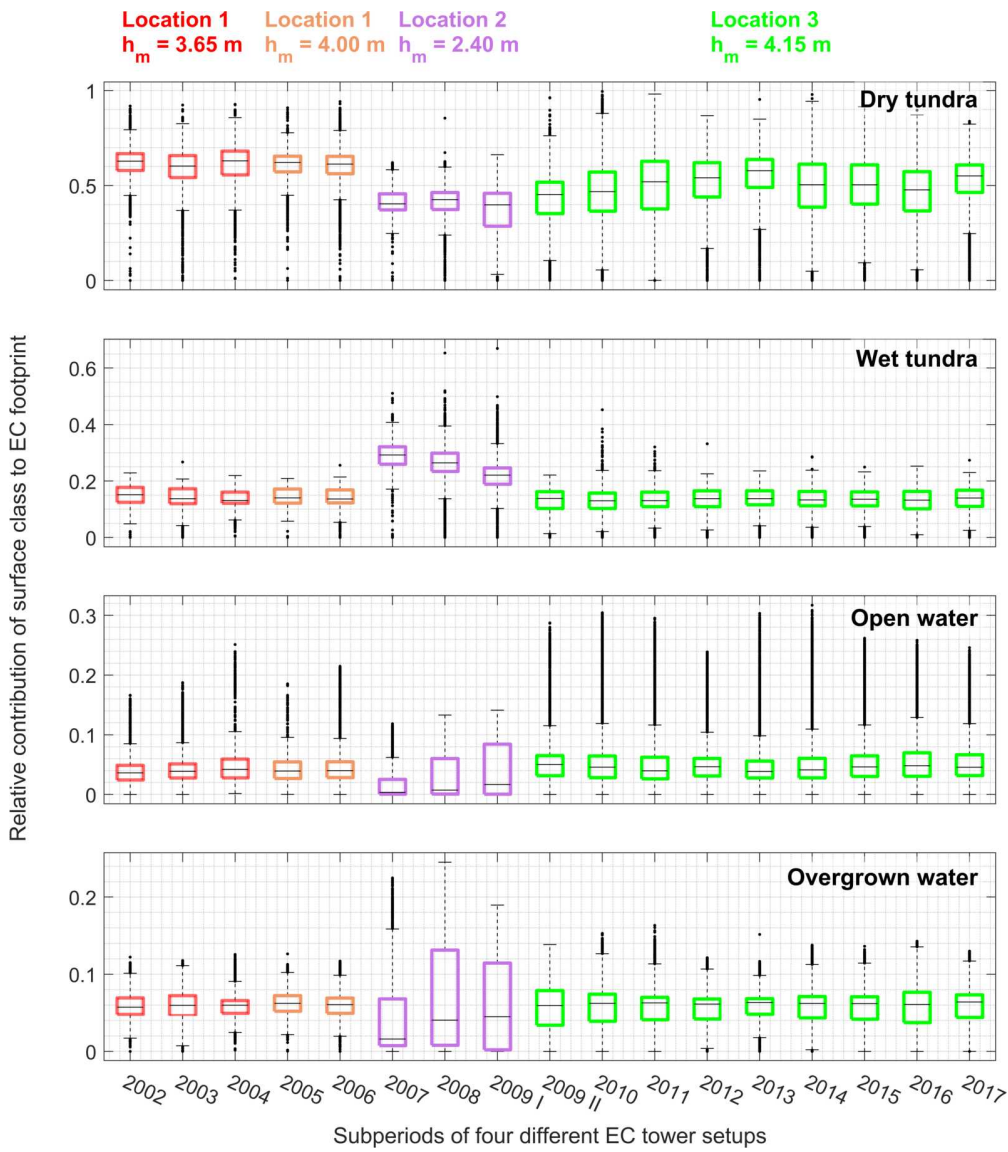
Fig. 1   Mean surface class composition of the eddy covariance footprint during 17 subperiods of four different tower setups at three locations on Samoylov Island.
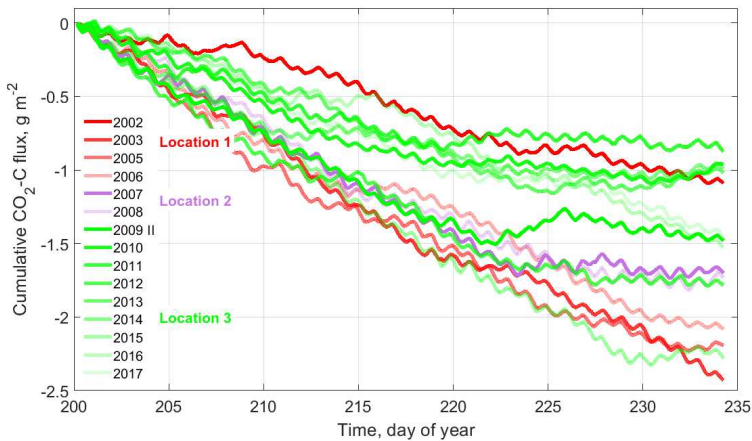
Fig. 2   Comparison of cumulative CO_2 flux sums of different years during the same day of year range.
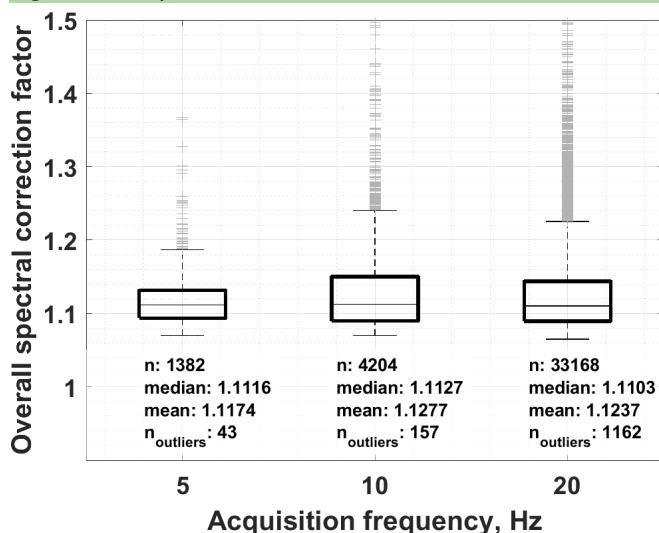

Fig. 3   Spectral correction factor statistics for periods with different acquisition frequencies.

RC 3
3.) Flux uncertainty description, and discussion A clear definition of data uncertainty is mandatory for publications in this journal. In Section 3.2, you briefly mention that you used the standard EddyPro feature to estimate random flux uncertainties – which is a good start, but certainly deserves more attention. So please work out in a separate paragraph what these random uncertainties consist of, and how exactly those were addressed in EddyPro. Moreover, there are also potential sources of systematic uncertainties in eddy covariance flux measurements, e.g. data-processing errors, or instrument calibration issues. These should ideally be covered directly in your uncertainty assessment of the flux data. Since you obviously decided to ignore them here, you should at least provide a convincing rationale why this simplification is justified.

I added a new part "Flux uncertainty estimation" to the "Methods" section.

Flux uncertainty can be regarded as a combination of a systematic and a random part. While the attempt should be made to remove systematic biases, random errors cannot be corrected for Richardson et al. (2012). However, statistical methods exist to estimate the uncertainty of a flux measurement due to random errors. We used three different approaches from literature to quantify random uncertainty and addressed fluxes with a suspected large bias by correcting for it during processing or by filtering in the course of quality assessment.

Most importantly, systematic errors are introduced when underlying EC assumptions are not met. Using the method of Mauder and Foken (2004) that combines an assessment of well developed turbulence and steady state conditions, we identified biased fluxes and flagged them. Other sources of systematic errors that we addressed include for example the angle of attack correction of faulty sonic anemometer readings, filtering for low instrument signal strength, the OP self-heating correction and compensations for high frequency loss and air density fluctuations (see sections 3.2.2, 3.3 and 3.4). Although we are confident that we applied corrections for systematic errors both rigorously and carefully enough, biases were certainly not always removed entirely. The quality flags included in the data set, reflect a level of confidence based on the assessment of general EC assumptions and our six additional quality filtering steps (see section 3.3).

To be able to include a random uncertainty estimate for each individual OP and CP flux in the provided data set, we set EddyPro to calculate random uncertainty estimates following Finkelstein and Sims (2001). The authors developed a method that aims at quantifying flux uncertainty associated with turbulence sampling errors. These errors can contribute largely to the total random error as they refer to the insufficient sampling of large eddies with high spectral energy. Due to the stochastic nature of turbulence, this type of error is random. To estimate its magnitude, the so-called integral turbulence timescale (ITS) is first determined by expressing the covariance of vertical wind velocity and gas concentration as a function of a lag time between these two time series. The ITS is then given by integrating the cross-correlation function theoretically from 0 to infinity, in practice, however, until an

upper lag time limit is reached. The upper limit can be defined in three different ways in EddyPro. We used the definition of the normalized cross-correlation function reaching a value of $1/e = 0.369$ to determine an upper lag time limit used for integration. While the normalized cross-correlation should reach zero with increasing lag time in theory, in practice it sometimes does not. The setting we used on the one hand provides the least conservative estimate of the ITS but on the other hand offers computational efficiency and makes sure that an upper limit for integration can reliably be found. With the ITS, a flux uncertainty can be determined by calculating the variance of an EC flux or, as Finkelstein and Sims (2001) put it, by calculating the variance of the covariance. This ensemble variance would approach zero with the averaging time approaching infinity. In the data set available for download, a random uncertainty estimate calculated with the method of Finkelstein and Sims (2001) is given for each OP and CP flux (see Table 6 in original draft). Random uncertainties based on ITS estimation observations increase with absolute fluxes with mean values of 0.16 and 0.05 μmol m^{-2} s^{-1} for OP and CP fluxes (see Figure 4). OP random uncertainty estimates are generally larger and more scattered with respect to the corresponding flux values.

As the above described random uncertainty estimate specifically addresses the turbulence sampling error, other sources of random flux errors such as the noise introduced by the different components of the measurement system are neglected. With simultaneous measurements from two sensors, we could additionally estimate random errors for the measurement system as a whole during times when the data sets from both sensors overlapped. We followed the paired observations approach as presented by Dragoni et al. (2007) and calculated a random error estimate ε as

$$\epsilon = \frac{1}{\sqrt{2}} \cdot (F_{CP} - F_{OP})$$

with the closed-path and open-path CO_2 fluxes F_{CP} and F_{OP} of quality classes 0 and 1 in μmol m^{-2} s^{-1}. The distribution of ε estimates is shown in Figure 5. The ε values calculated with OP fluxes corrected for the self-heating error have a mean close to zero and are distributed more symmetrically than the ε values calculated with uncorrected OP fluxes. The mean of this distribution is shifted from its mode as well as from zero, indicating a much stronger systematic component whithin the measurement error. This result increases our confidence that the OP self-heating correction we applied was successful in removing a systematic bias from the data.

Further following Dragoni et al. (2007), we used the ε system error data set from the overlap period to generate flux uncertainty estimates for bins of increasing OP flux ranges. We sorted the ε values in 20 corresponding flux bins between -2 and 2 μmol m^{-2} s^{-1} and calculated an uncertainty estimate for each bin $\sigma(\varepsilon)_i$ as

$$\sigma(\epsilon)_i = \sqrt{2} \frac{1}{N_j} \sum_{j=0}^{N_j} |\epsilon_{i,j} - \bar{\epsilon}_i|$$

Results show (see Figure 4) a similar data range and pattern of uncertainty estimates in relation to associated fluxes like the half-hourly values calculated after Finkelstein and Sims (2001).

As a third method of random uncertainty estimation we simplified the successive observations approach from Richardson et al. (2006) by using results of the quality run performed during MDS gap filling (see section 3.5). We selected the time steps when an flux observation and a MDS value that was estimated using a one day window and the MDV technique were available. We used the standard deviation of the fluxes measured at the same hour of day within a one day window, as an uncertainty estimate of the observed flux. Results are shown in Figure 4 and also increase with rising absolute fluxes in the same ranges as random uncertainties due to turbulence sampling error or measurement system error do.

We included the results obtained with ITS estimation into the uploaded data set considering the similarity between the uncertainty-flux relations calculated with independent methods as well as due to the advantage of a distinct uncertainty estimate for each sensor and time step.
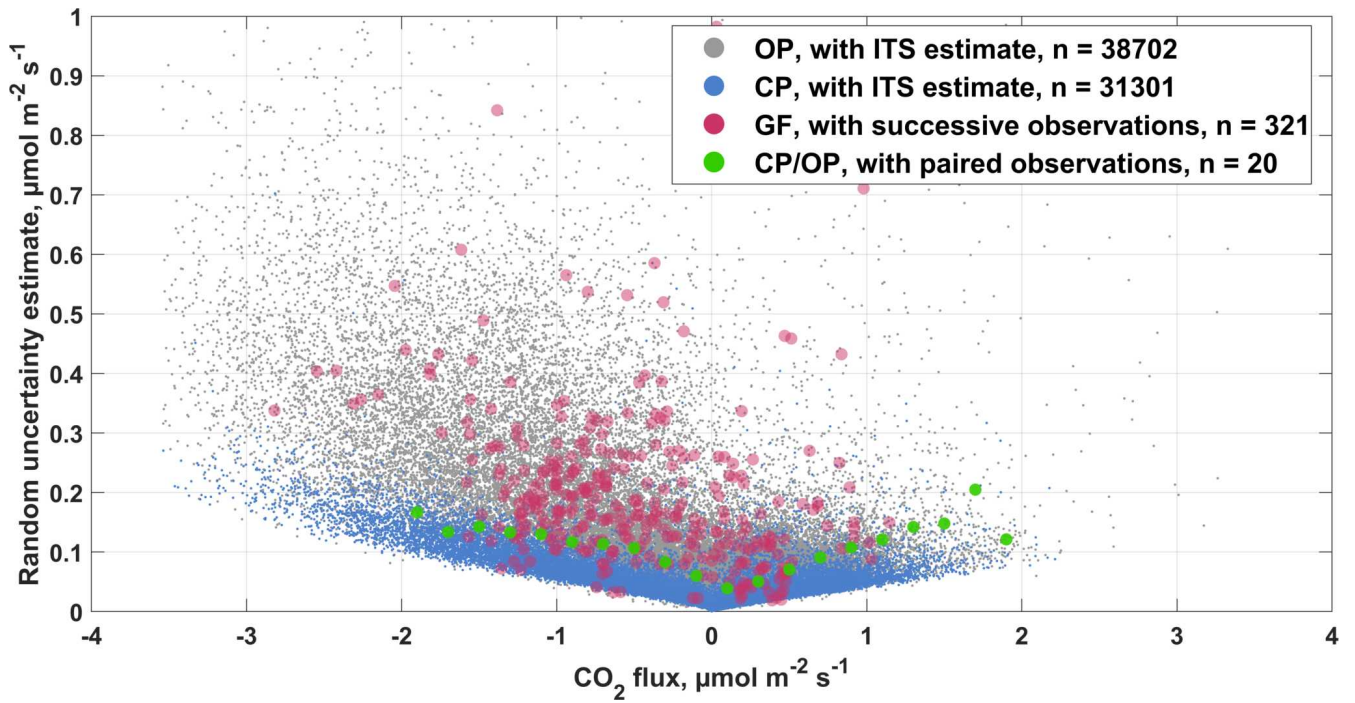
Fig. 4   Random uncertainty estimates for all closed path (CP) and open-path (OP) CO2 fluxes calculated using (1) estimates of the integral turbulence time scale (ITS), (2) the successive observations approach and results from gap filling (GF) and (3) the paired observations approach during periods with simultaneous OP and CP records.
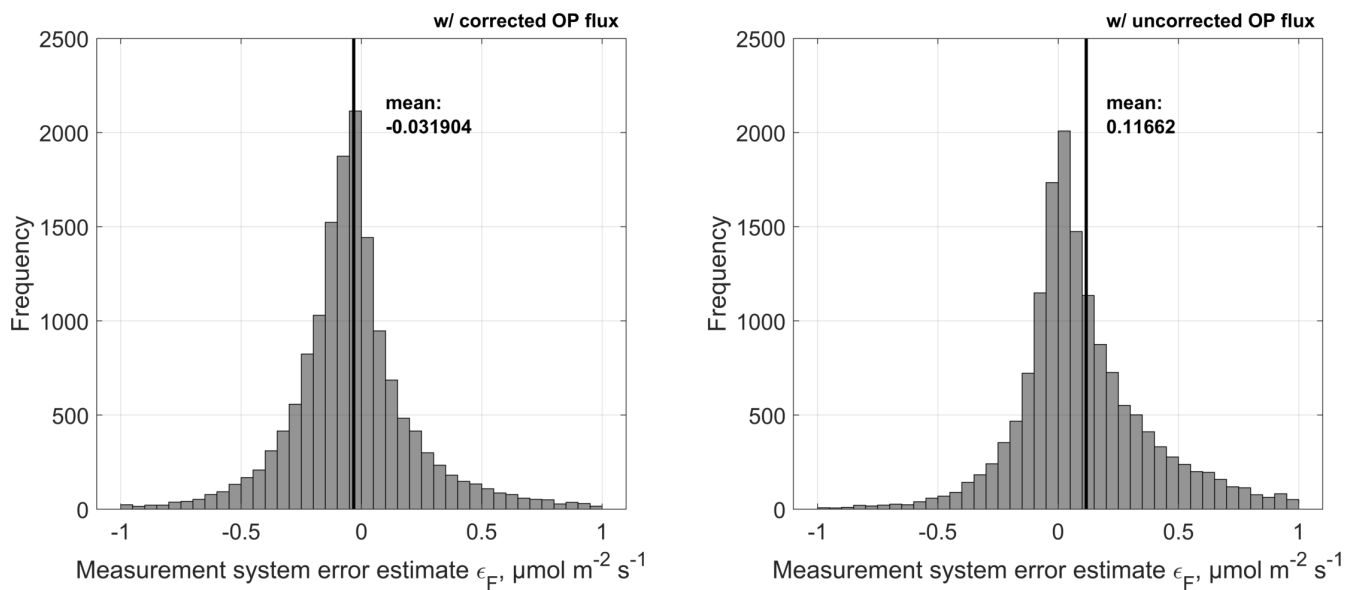

Fig. 5   Distributions of the measurement system errors ε estimated using the paired observations approch for differences between closed path and corrected (left panel) as well as uncorrected (right panel) open-path (OP) fluxes.

Line comments:

RC 4

p.1, abstract & introduction: Within these sections, I'm missing data-driven insights. Having a 16-year data record at hand, I would first think about analyzing the data directly to determine long-term trends in surface-atmosphere exchange processes. Next, I would aim at generating process insights, e.g. what causes interannual and inter-seasonal variability in flux rates, Only then I would start thinking about the time series being a useful resource for calibrating and validating process models. I think these data-driven topics deserve additional attention in both sections.

We regard this dataset publication as a starting point for analysis of flux dynamics done by us and other members of the scientific community. We are aiming at publishing those types of results in the future (Kutzbach, unpublished). In this paper, however, we wanted to focus on the methods we used to process the data rather than its interpretation. To our understanding, this proceeding is in line with the "Aims and scope" of ESSD, which is one reason why we selected this journal.

"Articles in the data section may pertain to the planning, instrumentation, and execution of experiments or collection of data. Any interpretation of data is outside the scope of regular articles. Articles on methods describe nontrivial statistical and other methods employed (e.g. to filter, normalize, or convert raw data to primary published data) as well as nontrivial instrumentation or operational methods. Any comparison to other methods is beyond the scope of regular articles." (https://www.earth-system-science-data.net/about/aims_and_scope.html)

RC 5

p.1, l.6: FLUXNET is not restricted to CO2 fluxes

True, this is a lapse. I changed "The site is part of the international network of carbon dioxide flux observation stations (FLUXNET, Site ID: Ru-Sam)." to

The site is part of the international network of eddy covariance flux observation stations (FLUXNET, Site ID: Ru-Sam).

RC 6

p.2, l.7: excessive use of references for a single statement

I do not agree. The reference list is meant to express that many authors agree on the importance of permafrost carbon pools in the context of climate change. The references are thought to proof the statement of "wide recognition" of the topic.

RC 7

p.2, l.16: not sure what inversion model have to do with the scope of this paper. They are trained on mixing ratio observations, not fluxes.

Thank you for pointing out this fact. I changed the sentence to:

McGuire et al. (2012) conclude that reducing uncertainties of regional estimates based on observational data relies on high quality ground-based measurements that should be placed strategically, e. g. along hydrological or vegetation gradients.

RC 8

p.2, l.30f: this section could use a map to show location of the delta, and the island itself

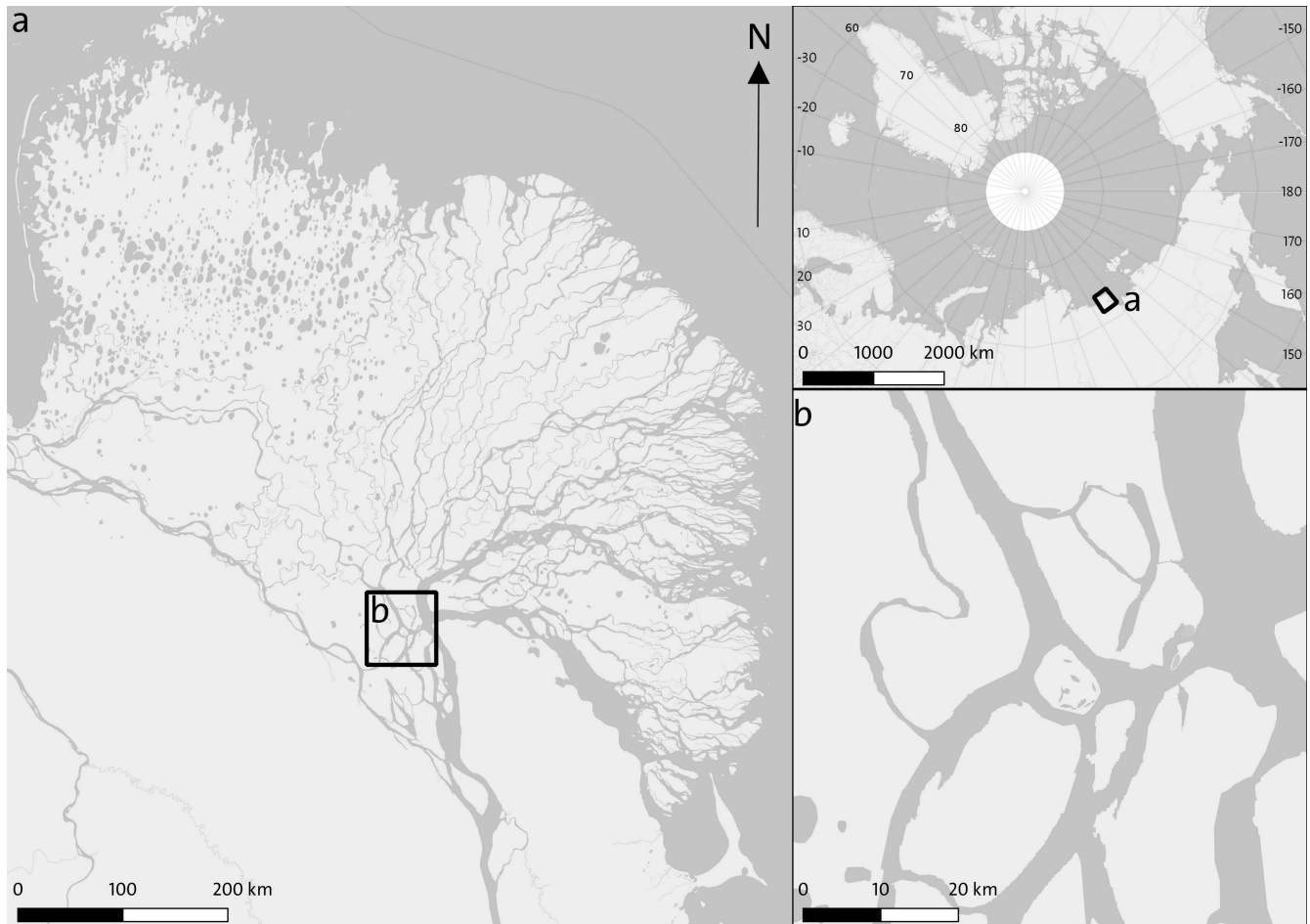I added an overview map (Figure 6 in this document).

Fig. 6   Location of Samoylov Island (center of panel b) in the Lena River Delta (panel a). Map data from: OpenStreetMap contributors, under Open Database License

RC 9

p.3, l.17: there is no high-centered polygon on the entire island ..??

Yes, there are some high-centered polygons on Samoylov. I changed the sentence.

In contrast to the modern floodplain, the river terrace's surface is patterned due to frost-action that formed a wet polygonal tundra landscape consisting of mostly low-centered and some high-centered ice-wedge polygons as well as thermokarst lakes and channels.

RC 10

p.3, l.29ff: climatology information given here is certainly useful, but only based on a #20 year record from the site itself. It may be helpful to compare to longer-term climate records from the region (e.g., for Tiksi there is data starting in the 1930s).

I agree, I added longer-term meteorological information from Tiksi.

The closest WMO (World Meteorological Organisation) weather station is located on the continent, around 110 km southeast from Samoylov Island in the city of Tiksi. Between 1936 and 2017 the mean air temperature reported from Tiksi is − 12.74 °C, mean annual precipitation amounts to 304.5 mm (AARI, 2018). While the mean air temperature in Tiksi is very similar to the 20-year mean from Samoylov Island, average annual precipitation appears to be much higher in Tiksi than in the delta region. Boike et al. (2013) explain this divergence with the fact that Tiksi is located at the coast of the Laptev sea and surrounded by mountains.

RC 11

p.4, l.1f: is there any record of snow depth, and its variability?

I added information on snow depth from Boike et al. (2018).

..., the snow-free period 138 ± 18 days. Snow depth was reported by Boike et al. (2018) averaging 0.3 m between 2002 and 2017 with a maximum of 0.8 m in 2017. Beginning in early to mid-June,...

RC 12

p.5, l.6ff: you may add the power consumption as another important difference between CP and OP systems.

I added a remark on power consumption to the sentence starting in line 11 of page 5.

OP sensors are commonly installed in close proximity to the anemometer and do not require a pump that greatly reduces the power consumption of OP instruments compared to CP setups.

RC 13

p.6, l.4ff: even though you spend a few sentences to describe the WPL-approach, you fail to mention that this is about accounting for the influence of density fluctuations

I agree, the WPL-approach needs a more thorough and clear introduction. I therefore rewrote the section from page 5, line 14 (starting with "CP analyzers have the...") until page 6, line 8 (before "Major drawbacks...") and moved it to a new paragraph.

Infrared gas analyzers typically measure gas densities and report the number of molecules per volume of air. To be able to refer the mass of a gas to the mass of air, gas densities are transformed to mixing ratios using air density. However, as the optical path of an OP gas analyzer is exposed to the varying temperature, pressure and humidity conditions of the atmosphere, air density in the measurement cell fluctuates mainly due to thermal expansion/contraction and water dilution/concentration. This effect, that leads to faulty concentration readings of OP instruments and thereby to incorrect flux estimates, has first been described by Webb et al. (1980). The authors proposed two flux correction terms to compensate for these density fluctuation effects that are referred to as Webb-Pearman-Leuning (WPL) terms and have since been verified experimentally and theoretically and are routinely applied in OP EC studies. Especially at times of low gas fluxes, WPL terms can become orders of magnitude larger than raw gas fluxes (Munger et al., 2012). CP analyzers have the advantage of controlled temperature and pressure conditions in the measurement cell, allowing for the sample-wise calculation of mixing ratios rather than molar densities (Ibrom et al., 2007b) and thereby avoiding the need to apply air density fluctuation correction terms after raw flux calculation.

RC 14

p.7, Section 3.3: It's a bit odd that you start describing some elements of quality flagging already in Section 3.2, and continue with this material here, in the main quality section. This should be cleaned up. Also, you fail to reference Table 3 in the text. Moreover, you should improve the structure of this Section. You begin with a too short general overview on additional quality filters, and how they are used in the overall QC flagging scheme. You then close the section with very similar statements. This should be merged to a single introductory paragraph that clearly states that you applied 6 more quality checks, and if any of them indicated problems, the quality flag was set to 2.

Thank you for the suggestion, I agree, the "Quality filtering" section (3.3) should be more clear. I moved the end of section 3.2 (from page 7 line 21 to the end of the paragraph) to a newly formulated introduction of section 3.3. As suggested, I also moved the end of section 3.3 (page 8, line 23, starting from "In the dataset available...") into this new introductory paragraph. Section 3.3 now begins with:

We set EddyPro to calculate quality flags according to Mauder and Foken (2004) that represent flux quality in three classes (0, 1 and 2) with 0 denoting the highest and 2 denoting the lowest quality class. This quality evaluation is based on tests for stationarity and developed turbulence and thereby indicates whether general EC assumptions about atmospheric conditions were met during a flux calculation period. Flux quality assessment was largely based on the scheme of Mauder and Foken (2004). In the data set available for download, we included one column for each analyzer type containing this quality flag. Additionally, we applied six further screening steps and flagged fluxes of low quality. If a flagged flux was not already assigned to class 2 according to Mauder and Foken (2004), we set the quality flag to 2. Fluxes of quality class 2 should be omitted from further analysis. They are included in the reported dataset for the sake of completeness. We performed the six additional flagging steps in the following

sequence. An overview of these filtering steps including the number of flagged values is given in Table 3 (*in original draft*).

RC 15
p.8, l.14: The choice of 450ppm as the upper concentration limit seems rather narrow. Can you please justify?
I want to stress that this limit refers to half-hourly average concentrations, the absoute concentration filter applied to the high frequency data during raw data screening in EddyPro (following Vickers & Mahrt, 1997) allowed a much wider range (200 ppm to 900 ppm) of concentrations. The limit of half-hourly average concentrations was decided for after calculating the 95th percentile of closed-path (440 ppm) and open-path (410 ppm) averages for timesteps with flux qualities 0 and 1.

RC 16
p.9, Fig.2: Figure 2 isn't really informative, since it's hard to distinguish between corrected and uncorrected time series in such a cloud of values. Please think about a different format (box plots?), or just leave out the plots, and show the regression statistics instead in a table.
I replaced the figure and added a table with the regression statistics.
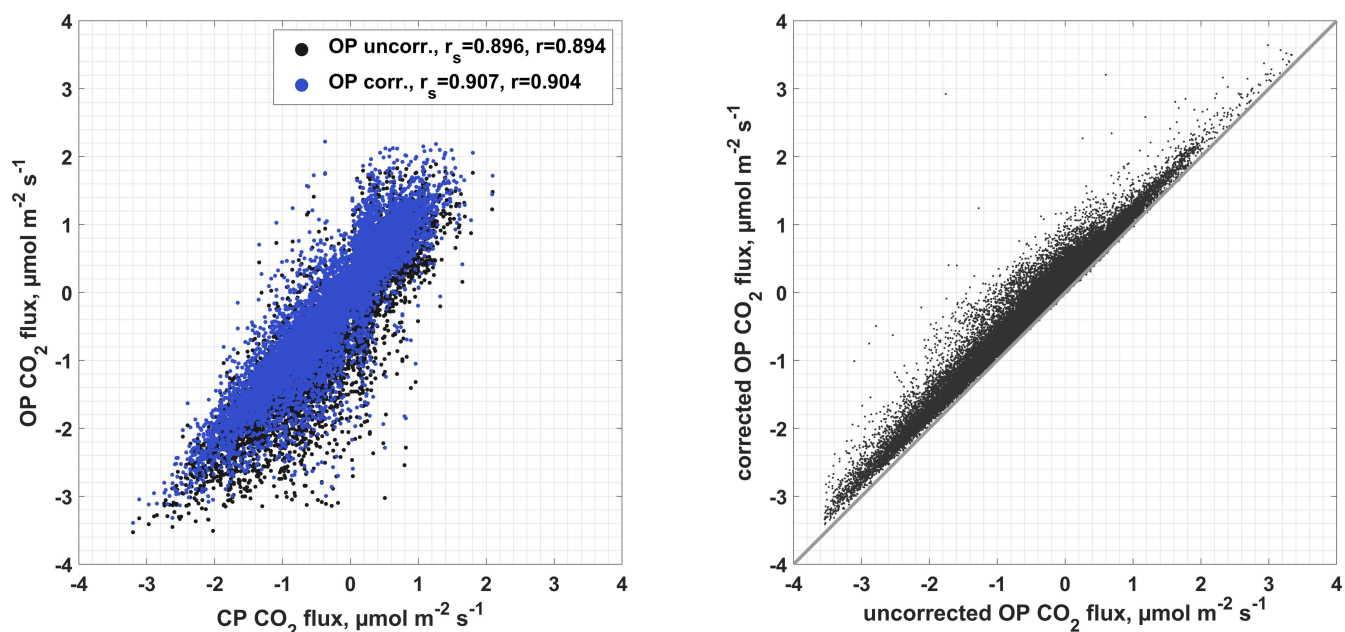

Fig. 7   Effect of the self-heating correction on the correlation between open-path (OP) and closed-path (CP) fluxes (left panel). Only quality class 0 is shown. Negative fluxes are affected more strongly by the correction than positive fluxes (right panel).

Table 1   Spearman's rank correlation coefficient rs and Pearson's correlation coefficient r between closed-path (CP) and open-path (OP) fluxes with and without the applied self-heating correction. The agreement between CP and OP fluxes increases throughout all quality classes after OP correction.

		Quality class 0	Quality classes 0,1	Quality classes 0, 1, 2
r_s	OP uncorrected	0.896	0.866	0.508
	OP corrected	0.907	0.871	0.512
r	OP uncorrected	0.894	0.871	0.042
	OP corrected	0.904	0.877	0.055

p.10, Section 3.5: I suppose Figs. 3 & 4 should belong to this section. They are not referred to in the text. Moreover, it's not necessary to show Fig.4, since given the minor absolute shifts in fluxes after Burba correction in this case, the differences between figures are not discernible. As an alternative for Fig.4, it may be interesting to show the gap-filled time series, maybe even in cumulative form?

I agree, the gain of information from Figures 3 and 4 in the original draft is limited. I replaced both with one new Figure (Fig. 2 in this document) showing the measured time series that we compiled from open-path and closed-path records as well as the gap-filled time series.

gap-filled **measured**

F_{CO_2} , µmol m^{-2} s^{-1}

2002	n = 2564	n = 1794
2003	n = 4557	n = 3211
2004	n = 2569	n = 1731
2005	n = 2196	n = 1834
2006	n = 5087	n = 3917
2007	n = 2011	n = 1431
2008	n = 7571	n = 3634
2009	n = 9455	n = 3995

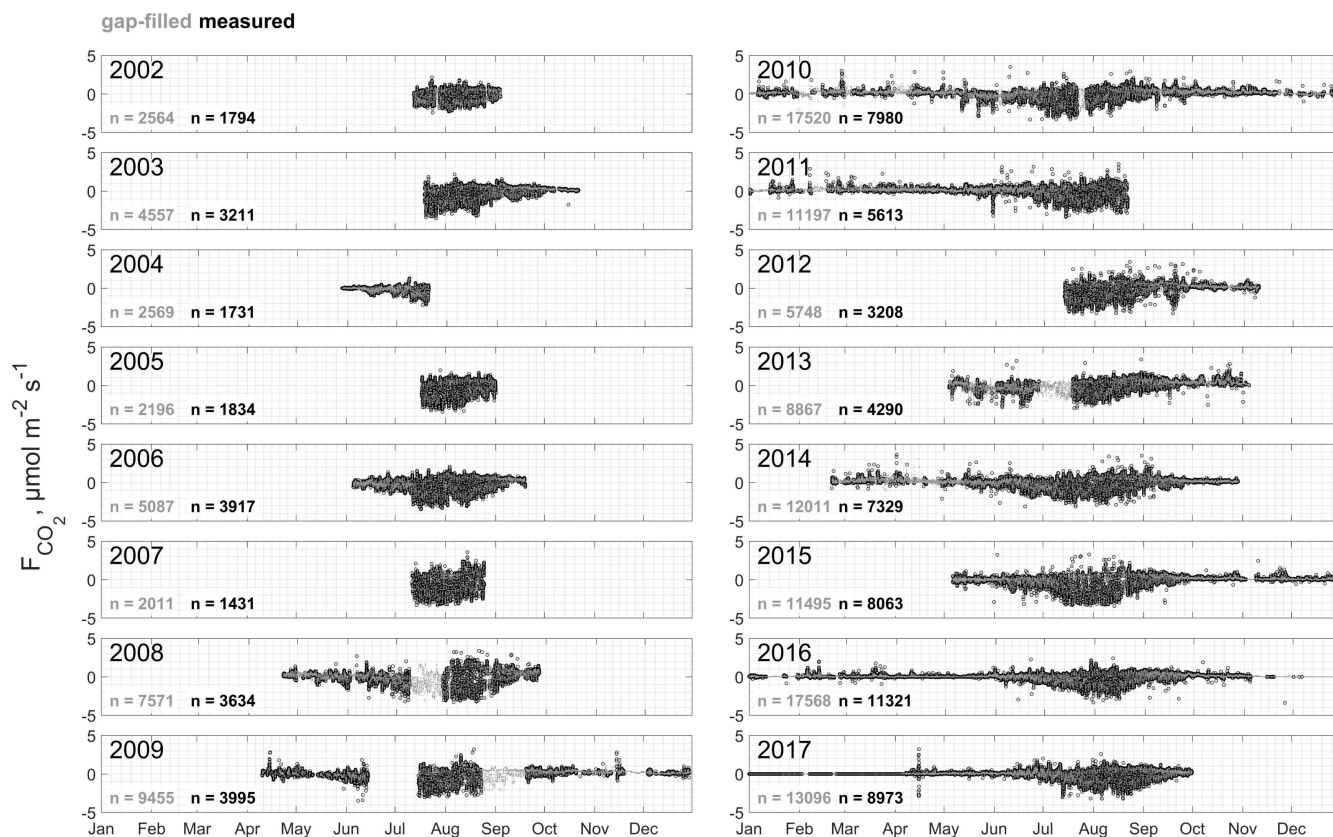| 2010 | n = 17520 | n = 7980 |
2011	n = 11197	n = 5613
2012	n = 5748	n = 3208
2013	n = 8867	n = 4290
2014	n = 12011	n = 7329
2015	n = 11495	n = 8063
2016	n = 17568	n = 11321
2017	n = 13096	n = 8973

Fig. 8   Multiannual carbon dioxide flux time series compiled from fluxes measured with closed-path and open-path sensors on Samoylov Island's river terrace. Fluxes of quality class 2 are not shown. Self-heating errors in the OP data set have been corrected for. Additionally, the result from gap filling this time series with the MDS method is shown. The given numbers of values for the gap-filled time series include measured fluxes.

p.11, Section 3.6: while the method applied to calculate footprints is sufficiently detailed, it is not fully clear how footprint results were combined with the land cover map. What's completely missing here is a reference to the findings, a.k.a. a bottom line. As already mentioned in the 'medium comments' above, this is an important piece of information, since (as shown in Table 1) multiple positions with multiple sensor heights were used over the 16 year data record. The authors clearly need to point out that this mixture of setups is still suitable to form a coherent, long-term time series of flux exchange for this site. It's not sufficient to just briefly mention these results in the conclusions. In particular, the results in Table 5 emphasize that the southernmost tower position, used within the years 2007-2009, featured a quite different composition of landscape elements than the northern site position. The authors need to make an effort to convince the readers that these differences did not result in a significant deviation of flux patterns, and therefore would bias the long-term trends.

I added a new "Discussion" section detailing the effects of tower relocations. See my response to RC 2 above.

I added more information on how the footprint results were combined with the land cover map to the end of section "Footprint modeling"

...We evaluated the footprint model at the same resolution that was used by Muster et al. (2012) to classify the surface (i. e. 0.14 m x 0.14 m). We could thereafter assign a probability of being the EC source area to each classified pixel and sum up the probabilities of all pixels belonging to the same surface class to estimate the contribution of each class. This proceeding to combine an EC source area estimation with a land cover classification is similar to what has been applied and described in more detail by Forbrich et al. (2011).

RC 19

p.11, Section 4: It's good to list the parameters given in the PANGAEA dataset in a separate table. However, since this dataset is obviously restricted to CO2 fluxes and their QC parameters, it would be good to also list the source for ancillary meteorological information, if available, since those will be necessary to put the flux time series into contex

I added a reference to ancillary measurements to the "Data availability section.

Ancillary long-term time series of meteorological and soil variables from Samoylov Island are available from Boike et al. (2018) and can be accessed through https://doi.pangaea.de/10.1594/PANGAEA.891142

I also added a new paragraph to the conclusions pointing out the importance of these ancillary data.

Furthermore, analysis of this NEE time series is not limited to the gas flux data only. An extensive data stream of meteorological and soil variables between 2002 and 2017 has recently been published by Boike et al. (2018). The authors made their records publicly accessible on the two long-term repositories Pangaea (https://doi.pangaea.de/10.1594/PANGAEA.891142) and Zenodo (https://zenodo.org/record/2223709). The fact of parallelly available ancillary ecosystem variables enables a potential user to put the gas flux dynamics reported in this publication into context with the variability of other ecosystem properties and potential flux drivers. We regard this type of analysis as vital to understand inter-annual variability of CO2 fluxes on Samoylov Island and are working on it ourselves (Kutzbach, unpublished).

New References

Chadburn, S. E., Krinner, G., Porada, P., Bartsch, A., Beer, C., Belelli Marchesini, L., Boike, J., Ekici, A., Elberling, B., Friborg, T., Hugelius, G., Johansson, M., Kuhry, P., Kutzbach, L., Langer, M., Lund, M., Parmentier, F.-J. W., Peng, S., Van Huissteden, K., Wang, T., Westermann, S., Zhu, D., and Burke, E. J.: Carbon stocks and fluxes in the high latitudes: using site-level data to evaluate Earth system models, Biogeosciences, 14, 5143–5169, 2017.

Eckhardt, T., Knoblauch, C., Kutzbach, L., Simpson, G., Abakumov, E., and Pfeiffer, E.-M.: Partitioning CO2 net ecosystem exchange fluxes on the microsite scale in the Lena River Delta, Siberia, Biogeosciences Discussions, 2018, 1–27, 2018.

Porada, P., Weber, B., Elbert, W., Pöschl, U., and Kleidon, A.: Estimating global carbon uptake by lichens and bryophytes with a processbased model, Biogeosciences, 10, 6989–7033, 2013.

Raddatz, T., Reick, C., Knorr, W., Kattge, J., Roeckner, E., Schnur, R., Schnitzler, K.-G., Wetzel, P., and Jungclaus, J.: Will the tropical land biosphere dominate the climate–carbon cycle feedback during the twenty-first century?, Climate Dynamics, 29, 565–574, 2007.

Richardson, A. D., Aubinet, M., Barr, A. G., Hollinger, D. Y., Ibrom, A., Lasslop, G., and Reichstein, M.: Uncertainty quantification, in: Eddy Covariance: A Practical Guide to Measurement and Data Analysis, edited by Aubinet, M., Vesala, T., and Papale, D., pp. 173–209, Springer, 2012.

Richardson, A. D., Hollinger, D. Y., Burba, G. G., Davis, K. J., Flanagan, L. B., Katul, G. G., Munger, J. W., Ricciuto, D. M., Stoy, P. C., Suyker, A. E., et al.: A multi-site analysis of random error in tower-based measurements of carbon and energy fluxes, Agricultural and Forest Meteorology, 136, 1–18, 2006.

Rinne, J., Douffet, T., Prigent, Y., and Durand, P.: Field comparison of disjunct and conventional eddy covariance techniques for trace gas flux measurements, Environmental pollution, 152, 630–635, 2008.

Dragoni, D., Schmid, H. P., Grimmond, C. S. B., and Loescher, H. W.: Uncertainty of annual net ecosystem productivity estimated using eddy covariance flux measurements, Journal of Geophysical Research: Atmospheres, 112, 2007.