Response to anonymous Referee #2

We welcome this reviewer's thoughtful comments and suggestions and have provided responses to individual comments below. Our responses describe changes that will be submitted in the revised manuscript.

General comments

The authors present soil respiration quantity and 14CO2 signature as well as soil pore 14CO2 data of high-latitude soils during the thawing seasons of 2012 -2014. Acquiring this type of data in these regions is not trivial and the data provides some interesting insights into the seasonally and spatially variable carbon respiration in this region. The data is provided under the given link and the R codes used for the graphs and modeling are provided on github. Overall, the data and presented insights are valuable and merit publication.

Some aspects regarding the sampling documentation, blank analyses, statistics as well as data use in models could be improved upon in order for this dataset in order to reach its full potential. Specifically, as micro topographic variability was identified as a key driver for variability in 14CO2 trends, a visualization and quantification of that topography would be helpful.

In response to this comment and a comment from Reviewer #1, we will add a schematic diagram of polygon microtopography to the revised manuscript. In the study site section, we have also added values for the percent coverage of each polygon type across the Utqiaġvik region land surface, as estimated by a classification algorithm in Lara et al. (2014). Quantitative estimates of the percent cover of each polygon type are unavailable for the specific study region, but we have provided a relative ranking of the coverage of each polygon type, based on the classification in Wainwright et al. (2015).

Furthermore, respiration data of soil carbon is directly linked to the solid bulk soil carbon – however, that there is no data on this pool.

The explicit focus of this analysis was the carbon contributing to respiration. Because the most fast-cycling carbon pool dominates the decomposition flux but generally makes up only a small portion of the total soil carbon pool, $\Delta^{14}C$ of ecosystem or soil respiration is not directly linked to that of bulk soil carbon (Trumbore, 2000). For this reason, given the high cost of radiocarbon analyses, we chose to focus this study's analytical effort on CO_2 data. We do agree, however, that bulk soil radiocarbon values provide a metric describing carbon stabilization within soil, and that a comparison between spatial patterns in the CO_2 flux and bulk soil values could potentially be informative. In a separate study using laboratory incubations of soil cores, we measured $\Delta^{14}C$ of shallow and deep bulk soil increments collected in September 2014 from the same locations as this study's measurements. Based on this reviewer's suggesiton, we have added an additional panel to figure 2 showing these bulk soil $\Delta^{14}C$ values from each polygon type. These bulk soil $\Delta^{14}C$ measurements will be added to the data archive as well.

Although analytical precision data was provided, no data on procedural blanks were provided. This type of sampling and measurements take many intermediate steps, and if available it would be good to have procedural blank data in addition to just the OXAs (e.g., Hanke et al., 2017). Especially because the samples were processed not in one batch but rather over a number of years (2014-2017).

The paper referenced by the reviewer addresses compound-specific radiocarbon analysis, in which the small-size samples are particularly prone to contamination by non-sample carbon during the various processing steps. In such cases, procedural blanks are important for quantifying contamination during the sample preparation steps. With environmental CO_2 samples as in this study, our understanding is that procedural blanks are not typically used. In place of blanks (which would not yield a carbon quantity large enough for the radiocarbon analysis system used), we evaluated our air sampling canisters with repeated leak testing and an ethanol-derived certified CO_2 reference gas standard with empirically determined ^{14}C (Airgas, USA) to confirm that no contamination was introduced during sample storage. Additionally, the vacuum line used for CO_2 purification was tested with USGS coal and oxalic acid II standards for contamination or leakage. Standards used were well characterized with respect to ^{14}C. Our understanding is that these quality control measures are standard and sufficient for this sample and analysis type. Details on the standards used have been added to the methods section of the revised manuscript.

Regarding statistics, there are some points were statistical terms such as standard deviation is used when n=2, or box plots with quartiles are implemented when n = 3 or 4. This is could be seen as misleading and result in misinterpretation. See specific comments for suggestions how this can be improved.

We have responded to these comments in the line-by-line responses below.

Considering that this is a journal aimed at geo-data and use of this data in other studies, the authors could add details regarding which specific types of models could benefit from this data. This could improve how this data is used in the future stages.

Radiocarbon measurements have potential applications for both bottom-up and top-down models of the carbon cycle. These uses include constraining the decomposition rate parameters used in bottom-up models with empirical distributions of pool-specific carbon cycling rates, evaluating the abilities of such models to produce realistic decomposition rate distributions, and determining temporal variations in ^{14}C of ecosystem respiration, important for top-down (inverse) models that use atmospheric radiocarbon time series to evaluate net CO_2 sources and sinks. Such models are quite broad in scope, and could include any models with pool-specific decomposition rate parameters (which are otherwise challenging to determine in situ), models with a built-in ^{14}C output (which can be used as a direct benchmarking method), or top-down models that use radiocarbon in atmospheric CO_2.

These applications were briefly mentioned in the manuscript's introduction, but we have expanded this section at the reviewer's request. In the revised manuscript, this passage now reads: "Radiocarbon measurements of soil respiration may be particularly useful for models of the carbon cycle. First, any bottom-up carbon cycle models that employ pool-specific respiration rate parameters may benefit from radiocarbon-derived soil carbon pool structures and decomposition rates. Without the radiocarbon tracer, such metrics are challenging to measure in situ, particularly where carbon cycling rates are slow. Second, by including a ^{14}C calculation for carbon pools and fluxes, such models can use radiocarbon measurements in surface CO_2 emissions and soil pore gas as a benchmarking tool. Additionally, the radiocarbon signature of ecosystem respiration can be used to constrain the terrestrial signal in top-down carbon cycle analyses (He et al., 2016; Randerson et al., 2002)."

Specific Comments

Regarding the novelty of the data and methods, the data is novel and focuses on a region which is undergoing rapid environmental change. Data is based on established sampling and analytical methods.

Regarding the data being used in the future: The authors refer to the data being used in 'models', but could perhaps provide more detail. At some point, they refer to a publication (He et al., 2016) that uses Earth System models, but I can imagine that there could be a larger scope of the implementation of this data in models.

We have addressed this comment in our previous response. We would also welcome any suggestions from the community for additional model applications.

One major piece of information which is missing - which is likely important for these types of models – is the size and 14C signature of the solid soil carbon. It is unknown how the amount of respired carbon relates to this –presumably much larger – pool.

As discussed above, in response to this comment we have provided Δ^{14}C values for bulk soil carbon collected in September 2014 from the same locations as the ecosystem respiration samples. We stress, however, that the radiocarbon values of bulk soil carbon are not necessarily well correlated with those of the respired fraction.

One additional aspect which can be improved upon in order to increase the value of this data is an improved description of the morphology of the micro topographic features that were sampled. The authors underline that micro topography has a strong impact on 14C signature of respired carbon, but there are no pictures or sketches. I would recommend adding photographs or a sketch showing the features. Furthermore, in order to be able to extrapolate this data, it would be helpful to know how often which type of polygon types occurs. That way, the signal could be averaged for this region.

As discussed in our response above, we have provided additional information on polygon microtopography in the manuscript's introduction. This includes background information

on how controls on carbon cycling relate to polygon morphology, as well as a distribution of polygon types across the land surface at our site.

The materials and methods are described in sufficient detail and the R code and data themselves are provided. The article itself is appropriate for the publication of this dataset because it clearly describes how the samples was acquired and data was measured.

As mentioned previously, if there is data on the bulk carbon quantity and signature from other papers, it could be valuable to cite and integrate that. Presently it is not present in the paper.

As discussed above, we will add bulk soil carbon data to the manuscript and data archive.

Data quality and accessibility

The data is accessible & codes are available on git. There are even some 2-pool model codes on git which I believe are not used in the paper. Error estimates could be improved upon. Mostly only the analytical error is provided. I am missing the procedural blank, which is especially important as the samples were measured over three years. The sample sizes for 14C measurements are also not provided. Smaller samples are more susceptible to contamination (Hanke et al., 2017). Ideally, there would be more sample replicates. However, considering the difficulty of the sampling and the significant cost of 14C measurements it makes sense that no extensive replication done in this case.

We agree with the reviewer that susceptibility to contamination increases with smaller sample sizes. To address this concern, we will add sample sizes to the data archive. These sample sizes are all greater than 0.1 mg C, with the exception of one small sample (0.044 mg C; soil pore-space CO_2 collected from HC1-center in August 2012). As discussed at length in our response above, our understanding is that procedural blanks are not typically used for this type of ^{14}C sample. Instead, standard analyses and leak testing were used as alternative quality control measures.

There is also a potential bias in the soil respiration data that has not yet been addressed. The authors state that when the soil was water saturated, it was not possible to measure pore CO2. It is very likely that there is a difference in carbon decomposition speed between water saturated and non-water saturated soils. Therefore, if only the non-saturated soils soil pore CO2 is measured, there could be a bias in the interpretation. Authors should address this in the discussion.

We agree with this astute point raised by the reviewer. We were cautious in our interpretation not to make statistical inferences from our data because of sampling issues such as this.   This issue was briefly addressed in the final paragraph of section 3.2, but we agree that this consideration deserves greater attention. Accordingly, we have added several sentences to this paragraph, which now reads: "Due to sampling limitations, ^{14}C data from soil pore-space CO_2 are available from only a subset of polygon types, positions, sapling dates, and depths. No samples were collected from polygon rims, and

only few samples were obtained where soils were saturated (Table 2). For this reason, our dataset does not capture the full range of microtopographic variations in deep soil decomposition rates and controls. Soil temperature profiles, soil pore-space oxygen availability, soil organic matter concentration and chemistry, and vegetation composition, rooting distribution, and productivity all influence microbial activity and vary among profiles and with time in the thawed season (Lipson et al., 2012; Olivas et al., 2011). In particular, these controls on decomposition dynamics likely differ between saturated and unsaturated soils, due both to the direct effects of saturation and to covarying effects of microtopography. For this reason, radiocarbon values presented in this manuscript cannot be scaled across the full range of environmental variation present at the site. To better characterize the relationships between these variables and soil carbon decomposition rates, further measurements are needed of ^{14}C in soil pore-space DIC and CO_2, across spatial, seasonal, and hydrological gradients.”

Dataset quality. Data seems in the range of what can be expected from these types of cold soils.

Dataset presentation

The graphs can be improved upon on a number of points, especially concerning the usage of some statistical terms, as detailed in the section technical comments. Note that box plots are not designed for data where n = 3.

Nor is standard deviation applicable if n = 2. Standard deviation is a measure to see what the spread of data is, but with n= 2 you don’t really have a spread.

See below for responses to each of these points.

Publication length

The length of the paper is good, the figures and table are appropriate. If the authors provide more information of the sampling and micro topography, the dataset could be used in the future after reading the paper & downloading the data

Technical comments

Page 1 line 11: Specify model

We have changed this to read, “models of the carbon cycle.” We chose not to go into greater detail in the abstract, but have added further discussion of model applications in the conclusion section.

Line 12: Specify use of D14C

We have added the units $\Delta^{14}C_{Reco}$ and $\Delta^{14}C_{CO2p}$.

Line 25: You state that the flux is critical – is there any quantification?

This introductory sentence is meant to emphasize that the variability in ecosystem-atmosphere flux over space and time makes it challenging to quantify, both empirically and through models. Accordingly, we feel that in this context, it is more appropriate to leave this as a qualitative statement rather than provide quantification either at the global scale or for a specific location.

Page 2 lines 12-15: Sentence is very long and not clear

In response to the reviewer's comment, we have broken this sentence into two shorter sentences and have changed the wording. This passage now reads, "There, an estimated 1300 Pg of soil carbon has been protected from decomposition by cold temperatures and often frozen or anoxic conditions. With climate change, these controls on decomposition rates are expected to change as soils warm, hydrological changes occur, and permafrost degradation intensifies."

Study site section: If possible, add a sketch of polygon and how dominant each polygon type is.

This comment has been addressed at two points above.

Page 5 lines 2-3: 'measurements lacking a linear range were not included in the dataset' Do you mean samples that fall outside of the linear range, or complete measurement series?

For clarity, this has been changed to "curves lacking a clear linear range."

Line 14: Low-concentration CO_2 measurements were omitted except when value was clearly not-atmospheric. Does that mean that the excluded samples all had atmospheric signature?

Yes, the $\Delta^{14}C$ values of the excluded samples ranged from +13.6 ‰ to +59.4 ‰, all indicating a large percentage of carbon with a near-atmospheric radiocarbon signature. Because the CO_2 yields for these samples were at or below the expected yields at atmospheric concentration and pressure, we believe there was a high likelihood of leakage around the soil probe during sampling. As discussed in the paper, subsurface $\delta^{13}C$ of CO_2 can vary considerably at this site due to methane production and oxidation, so we were unable to calculate the degree of atmospheric carbon in soil pore-space samples. For this reason, we believed the conservative approach of excluding low-concentration samples was appropriate. While we believe this choice was based on sound reasoning, the reviewer's comment has led us to use more precise language in this passage. In the revised manuscript, this now reads: "With soil pore-space radiocarbon data, we omitted measurements with CO_2 yields below the expected yield for atmospheric measurements due to possible leakage and atmospheric contamination during sampling"

Line 16: Pore CO₂ was not measurement for water saturated soils. This may bias final interpretations. Please address

As discussed above, we took care in our interpretation not to draw statistical inferences because of this and other potential sampling biases. In response to the reviewer's comment and to emphasize this consideration, we have changed this sentence to read, "For this reason, the final sample set represents only a subset of depths and sampling locations and is not a random sample of the entire landscape." Additionally, we have added a passage discussing this point to the discussion.

Page 6 line 11: Tr or mean the of carbon in plants is either assumed to be 0 or 5 year. Whilst cited sources give a which wider range. Even decomposing pine needles in a temperate zone are generally 8 years old. In such cold environments with slow growth rates one could expect even older ages of plant-derived carbon into the soil. Is this assumption realistic? Please add some further justification.

We agree with the reviewer that a portion of plant-derived carbon resides for considerably longer than 5 years within plant tissues before being incorporated into soil organic matter. As the reviewer points out, the sources cited in our manuscript indicate a range of as long as 15 years for certain plant species and organs. We recognize as well that in these cold environments, adaptations of select species may lead to even longer carbon residence times within certain living plant tissues. However, we emphasize that the transit time described by the 0-5 year T_R range approximates the flux-weighted mean of *all* carbon passing through living plant tissues. In Barrow and other Arctic sites, the majority of aboveground vegetation tissues turn over annually. Annual leaf growth, for example, accounts for the majority of total annual production in Arctic tundra sites, even though it amounts to a much smaller fraction of the standing biomass (Shaver and Kummerow, 1991). Turnover of this aboveground leaf litter (with a < 1-year transit time) thus represents a large proportion of the plant-derived carbon flux. In addition, some non-structural photosynthates enter the soil within one or several days of fixation, contributing an additional source of carbon with very low T_R values. Given the importance to total annual carbon assimilation of leaf production and rapid photosynthate transport to the soil, we believe the range of 0-5 years is an appropriate—and conservative—T_R estimate.

We recognize that more justification for this range could have been provided in the manuscript, so we have included the following explanation in the revised manuscript: "The mean residence time of carbon in plants reflects a mixture of materials with varying transfer rates. In Arctic vegetation dominant at our site, some photosynthates enter the soil within 1 day of fixation (Loya et al., 2002); leaf tissues with transit times of < 1 year represent the majority of annual production (Shaver and Kummerow, 1991); roots and shoots are estimated to live for ranges of 3-7 and 1-8 years respectively (Chapin III et al., 1980; Shaver and Billings, 1975); and rhizomes and stem bases, the longest-lived of belowground tissues, have estimated turnover times of 2.7-15.6 years (Dennis, 1977). Overall, the mean residence time of carbon in plants (T_R) represents the flux-weighted average of various pools. Given the large percentage of carbon flux dedicated to annual

leaf production and rapid transport to soils, we estimate that across plant organs, plant species, and seasons, the mean value of T_R lies between 0 and 5 years."

Lines 21-25: When two turnover times are equally likely (Graven, 2015), it could also be helpful to have measured the bulk soil signature. If this is indeed not available, this is the best fix possible, but it is not ideal.

We respectfully disagree with the reviewer's comment. Even when bulk soil radiocarbon values are available, the present-day bulk soil turnover time offers little to no information about the turnover time of the rapidly decomposing fraction. Please see Trumbore *et al*., (2000) and Torn *et al*., (2009) for further background and discussion of this issue.

Page 7 line 14: When describing the effects of the polygons and microtopography, a sketch or image would be helpful for the understanding of the reader.

Based on the reviewer's suggestion, we will include a sketch of polygon microtopography to the revised manuscript's introduction.

Page 8 line 10: Authors state that pre $^{14}CO_2$ becomes increasingly negative with depth, but this is not the case for three profiles (as mentioned later in the text on line 18)

This section has been restructured in a way that clarifies the fact that three profiles do not display the depth trend seen in the other profiles. This paragraph now begins, "In most soil profiles, $\Delta^{14}C_{CO2p}$ became increasingly negative with depth in the soil. This depth trend is similar to that seen in the bulk soil carbon profiles, indicating that the cycling rates of both bulk soil organic matter and the decomposing carbon fraction tend decrease with depth in the soil. Because frozen or near-freezing temperatures slow decomposition from the deep active layer throughout the majority of summer, this pattern is what would be expected in the absence of vertical mixing or rapid contributions of fast-cycling carbon at depth. In contrast with this general trend, soil pore-space CO_2 from three soil profiles (HC3-trough, HC3-center, and FC2-center became enriched in radiocarbon near the permafrost table (Fig. 4)…"

Page 8 line 13: Author state that there are minimal contributions from plant-respired carbon. Did they do a mass-balance? Or is this a qualitative statement? Please specify.

This is a qualitative statement based on the fact that $\Delta^{14}C_{CO2p}$ values were negative—in most cases highly negative—with the modeled mean age of respired carbon ranging from 410 to 3350 years. Without known isotopic end-members for the plant-respired and soil-respired fractions, it is not possible to do a mixing model (mass-balance) to quantify these contributions. In theory, it is possible to perform incubations with vegetation and soils from a range of depths and use the ^{14}C and ^{13}C abundances of the evolved CO_2 to constrain this type of mixing model. In practice, however, this approach cannot be used reliably at our site where spatially- and temporally-variable methane processes influence ^{13}C abundance in soil pore-space CO_2 (Vaughn et al., 2016). This issue is discussed in the methods section (page 5, lines 16-18 of the original manuscript).

At the reviewer's request, we added a sentence to the end of this paragraph that clarifies this point. This paragraph now ends, "Detecting and characterizing the decomposition of older, deeper soil organic carbon requires direct measurements of soil pore-space CO_2. Such measurements provide a qualitative indicator of old carbon decomposition; as with surface CO_2 effluxes, proportional or absolute contributions from distinct carbon pools cannot be calculated without well-resolved vertical distributions of ^{14}C and ^{13}C source pools, as well as gas transport within the profile."

Lines 13-17: These statements support Figure 4. For better flow, maybe move towards end of section

We agree with the reviewer that these statements support Figure 4, and have added a reference to this figure in the text. We have chosen not to move this statement to the end of the section, but instead have changed the wording to improve flow: "These consistently negative values indicate that pore-space CO_2 was derived primarily from older soil organic matter, with minimal contributions from plant-respired carbon or fast-cycling soil organic carbon. In contrast, ecosystem respiration was generally enriched in radiocarbon relative to soil pore-space CO_2, even at only 10 cm depth (Fig. 4)."

Line 19: Figure 2 should be 4

This typo has been corrected.

Lines 25-26: When trying to understand soil carbon stability and decomposability, it would be good to not only determine the age of respired CO_2 but also the age of the bulk soil carbon. If this is not available for these profiles, but similar ones, it would be good to benchmark the data.

As discussed above, we have included bulk soil radiocarbon profiles for a subset of the locations in this dataset.

Page 9 line 9: Authors state: "due to sampling limitations, soil profile 14C data are available from only a subset of polygon types[...] In tables 1 & 2 I see only soil CO2 measurements, no bulk. Did they mean to refer to pore CO2?

Yes, this statement refers to soil pore-space CO_2. To clarify, we have changed this to read, "^{14}C data from soil pore-space CO_2"

Page 9 line 35: 'later in season' please be more specific

This line has been changed to "In September and October, over half the $\Delta^{14}CO_2$ measurements differed greatly from the local atmosphere…"

Table 1: Standard deviation with n = 2 does is not appropriate and can be seen as misleading (go back to calculation). It is supposed to give the range of sample variability, but for n = 2 that is not applicable. Consider an alternative, e.g., giving the average

Based on the comments from Reviewer #1, we have chosen no longer to include Table 1 in the main manuscript. Instead, we have added two figure panels that summarize the soil temperature and thaw depth data. All data originally listed in Table 1 are included in the data archive, where both the mean and standard deviation of each set of measurements are listed. For consistency, this includes values for which n = 2. While the standard deviation when n = 2 is not particularly informative, we do not believe that it is statistically misleading, as it is not being used for statistical inference.

Table 2: HC1-center: Measurements were done at different depths during different years. Or is it a typo? If not, considering the inter annual variability, is it reasonable to compare the data? Please clarify.

This is not a typo. As stated in the text, we were unable to obtain measurements from only a subset of polygon types and positions. For clarity, we have expanded this explanation to state that samples were unavailable not only from all polygon types and positions but also from all sampling dates and depths. When presenting these data in Figure 4, we grouped measurements by soil profile *and* sampling date, in order to emphasize that measurements from different depths and years should be considered separately. In Table 2, we believe it is appropriate to include the full dataset. For clarity, we have added a line break between samples collected on different dates.

Table 2: For HC3 center for the same depth there are large differences in 14C

This is true. As stated above, we have added line breaks between sampling dates to clarify that these measurements were made on different dates. The observed differences in 14C are likely due not only to inter-annual variability but also to seasonal differences in soil temperature profiles, root respiration, and soil carbon decomposition rates.

Figure 1: It appears that boxplot R are used for n = 2-5. It may not be the most appropriate way of presenting the range of data. Also add in the explanation that when R boxplots are visualized, data that falls outside the range is statistically speaking an outlier. Also indicate what the solid line in the box means. Also considering adding a striped line for the mean.

Following the reviewer's suggestion, we have modified this figure so that it is no longer a box plot. The revised figure is a dot plot, with day of year on the x-axis instead of month. We agree that this formatting more accurately reflects the distribution of the data and the exact sampling times.

Figure 1: Considering using this plot to highlight the important finding that during high-summer, all respiration is from year to decadal old C, whilst later in the season, whilst the deeper soil continues to warm, older (stabilized) C is lost.

We agree with the reviewer's interpretation of this figure. We believe this finding is thoroughly addressed in the original manuscript in the second paragraph of section 3.2: "As shallow soils warm and plant activity increases in this early summer period, ecosystem respiration includes high proportions of ^{14}C-enriched CO_2 from autotrophic respiration and heterotrophic decomposition of shallow, rapid-cycling soil carbon. Later

in summer and into the autumn, the balance of respiration shifts toward increased importance of deeper soil decomposition. Autotrophic respiration peaks in July or August, and decreases substantially into the fall after plants senesce (Hicks Pries et al., 2013). During the autumn season, surface soils refreeze while deep soils continue to warm (Zona et al., 2016), limiting heterotrophic respiration from shallow soils while enhancing decomposition from deeper, more ^{14}C-depleted soil carbon pools. The effect of these changes is a seasonal shift in respiration from primarily shallow, fast-cycling source carbon pools to more deep, ^{14}C-depleted soil organic matter." This passage follows a reference to Figure 1, but we have added an additional reference to Figure 1 to highlight the connection between the figure and this interpretation.

Figure 2: Consider changing axis (starting at 0.3) in a to highlight differences in respiration

We thank the reviewer for the suggestion, but prefer not to change this figure's axis. Given that the standard error of each measurement is far greater than any differences in mean respiration rates, we do not believe it is appropriate to highlight these differences.

Figure 3: Legend should be month-year.

This has been changed to month-year in the revised manuscript.

Figure 3: This graph shows an interesting trend which has not been explicitly discussed in the text, but which could be valuable for the data interpretation. There could be two end member type of behaviours: High summer (2013-07), high respiration of topsoil C releases bomb-enriched carbon. Late summer (2014-09) releases a low amount of old, stabilized carbon. 2013-09 could be an intermediate type where there is mixing from both pools, providing a spread of ages and concentrations.

To an extent, we agree with the reviewer's interpretation. This figure suggests that during high summer (July 2013), all respiration has a bomb-enriched radiocarbon signature. In September of 2014, in contrast, we observed a wide range in radiocarbon values, with the samples divided between bomb-enriched and ^{14}C-depleted signatures. However, because respiration rates were highly variable in July 2013 and Δ^{14}C values were highly variable in September 2014, we hesitate to classify these sampling periods into distinct "high respiration of bomb-enriched C" vs. "low respiration of old, stabilized C" regimes. Additionally, because one of three dates (September 2013) does not fit into either regime, we do not think the data are sufficient to support this interpretation. Instead, as discussed in the final paragraph of section 3.1, we feel that they key takeaway from this figure is that old, stabilized carbon constitutes a large percentage of the total CO$_2$ flux *only* when respiration rates are low.

Figure 4: Colour spread is not optimal, some colors are near-identical. Considering changing. Also consider highlighting the 3 soils which some younger respired C at depth, or grouping samples by polygon type.

Based on these comments and those from Reviewer 1, we have changed this figure substantially in the revised manuscript. Figure 4 now groups $\Delta^{14}C_p$ profiles by position within polygon and polygon type and no longer lists the specific profile ID and measurement month/year. We have also added a horizontal line at 0 representing the soil surface and moved the chamber flux data to +2 cm depth. As a consequence of these changes, the color scheme is simplified and improved.