Response to Reviews: Vegetation Influence and Environmental Controls on Greenhouse Gas Fluxes from a Drained Thermokarst Lake in the Western Canadian Arctic

The authors would like to thank the reviewers for their valuable comments and criticisms. Point by point responses to both reviewers are included in this document. Reviewer comments are in bold, responses are noted below each comment and shown in plain text. Changes to the manuscript are in red. Following that, is a marked-up version of the revised manuscript showing all changes made.

Reviewer #1

This paper by Skeeter et al. looked at vegetation and environmental conditions influencing greenhouse gas exchange in a drained lake basin in the Western Canadian Arctic. I enjoyed the opportunity to review this paper and I thank the authors for what is a well written paper overall (but with some tweaks needed). Given the lack of studies outside of the Barrow Peninsula, is a worthy addition to the literature.

I will add however, I apologise, I am not an expert on the eddy covariance data cleaning, gapfilling and analyses and I am therefore unable to comment fully on those sections.

Introduction: I find this section well written if rather short. I think a little more context could be given for the reader. You could include more information from non-DLB work but still relevant arctic tundra literature.
[Responses]
We added more context on why thermokarst landscapes are important (spatial extent, significant soil carbon storage) to the first paragraph [Lines 32-34].

Lake thermokarst landscapes are widespread in poorly drained, sedimentary permafrost lowlands with excess ground ice volume and constitute about a third of all thermokarst area (French, 2013; Olefeldt et al., 2016).

Further we added another paragraph on Arctic carbon budgets [Lines 40-49]. Including a review of pan-arctic NEE chamber studies by Virkkala et al. 2017.

Net ecosystem exchange (NEE), ecosystem respiration (ER) and gross primary productivity (GPP), where NEE=ER-GPP are lower in the Arctic than warmer regions but have significant seasonal cycles and variability between vegetation types (Virkkala et al., 2018). Future trajectories in NEE will in large part be governed by ER (Biasi et al., 2008; Cahoon et al.,2012). Dominant vegetation types in the Western Canadian Arctic are erect-shrub tundra and

wetlands (Walker et al., 2005). Growing season NEE is typically negative across these units throughout the Arctic indicating a net CO2 sink as GPP exceeds ER in part due to cold and/or anoxic soil conditions (Virkkala et al., 2018; Lafleur et al., 2012). Annual NEE can be positive or negative with large variation in GPP linked to annual weather variability (Virkkala et al., 2018, McGuire et al., 2009). Arctic net methane exchange (NME) is positive because wetland areas are strong methane (CH4) sources while upland areas with better drainage can be net sinks (Whalen and Reeburgh, 1990; McGuire et al., 2009; Sturtevant and Oechel, 2013).

Line 54: NEE should be ER – GPP.

[Responses]

We corrected the equation to NEE = ER – GPP

Line 85: Can you include somewhere the dominant species found in each vegetation class. It would be interesting and useful to know what sort of sedge dominated the sedge class – is it Carex aquatilis or Eriophorum angustifolium for example?

[Responses]

We added a sentence to mention the specific species [Lines 90-92].

Current vegetation at Illisarvik is diverse relative to the dwarf-shrub tundra of the surrounding uplands (Table 1); the basin hosts a mix of woody shrubs (Salix spp., Betula spp., & Alnus spp), wetland vegetation (Carex aquatilis, Arctophila fulva, etc.), and various grasses (Pocacea spp.) (Wilson et al. 2019).

This information was partly available in Table 1. We made the table more detailed to include the dominant species present within each class/subclass where known/applicable (see comment below).

Line 100: Completely an aside but COOL!

- Yes, it was an amazing thing to see.

Section 2.3: I would remove any mention of N2O – you don't present the data, so it is unnecessary.

[Responses]

We removed mentions of N_2O throughout the manuscript.

Line 150: Can you be more explicit with how many collars were used? 2 per site – 10 sites total? I know the main focus of this paper is not the collars, but I'm not sure if a replicate of 2 per vegetation type over an 11-day period is very representative.

[Responses]

We updated the text to be more specific about numbers of replications per vegetation type [Lines 176-177].

There were three replicates (six collars) for the Shrub class, two for the Sedge, Grass, and Upland tundra, and no replicates for the Sparse class.

There were 19 collars and 10 sites. The bare ground site only had 1 collar, thus making 2 per sites for the rest. The number of collars that could be shipped in via helicopter were limited and there was a high amount of heterogeneity in soil and vegetation characteristics within the basin. The chamber study was designed to better understand the relationship between soil properties and carbon loss in a situation where permafrost had aggraded within the lake bed to potentially protect 'old' carbon from mineralization and accumulate 'new' carbon since the lake drained. We expected saturated soils (where there were wet sedges) to have higher organic carbon accumulation and be dominated by anaerobic respiration processes, which was interesting to us, therefore we chose two different wet sites populated by sedges. We also focused on different statures of will (low, tall and dense) as we expected different amounts of snow accumulation and different impacts on permafrost at these sites.

Line 152: How soon after installation were the collars fluxed for the first time?

[Responses]

Collars were installed on July 11th and first set of measurements was taken on July 12th, so about 24h.

Line 154: Why not use a clear chamber so you could get GPP then cover with a dark sleeve in order to get ER?

[Responses]

We appreciate the comment. The chosen chambers were not designed for NEE measurements. Although measurements of GPP would have been informative, logistics limited us to use the existing collars, and the number of measurements we could make.

Section 2.4 (and subsequently Appendix A): Unfortunately, I do not have the expertise in these methods so I do not feel comfortable commenting on it in a reviewer context.

Section 3: I think it would be better to separate this section out into Results and Discussion rather than combine them. As it stands, it's quite hard to follow.

We have added a discussion (section 4) and rewrote the results to solely contain the objectively retrieved data, so we hope the manuscript to be more straightforward and easier to follow with separate "Results" and "Discussion" sections.

3 Results

During the 29-day study, half-hourly T_a and T_s ranged between 0.4 and 26.2°C and 4.4 and 11.0°C, respectively (Fig 2a). Day length and maximum solar altitude decreased from 24 hours to 19.25 hours and 41.6˚ to 35.4˚, but daily *PPFD* was more influenced by variations in cloud cover. Precipitation (19 mm) fell on 14 of the 28 days with trace snowfall on three of those days, but *VWC* of the soils decreased throughout the period (Fig 2b). At the onset of the study period, *VWC* was high and soils were saturated with ponding in the sedge areas. By the end of the study most of this surface water had dried up. On July 11th average thaw depth (cm) was 37, 45, 51, 64, 81 at Upland, Sedge, Grass, Shrub, and Sparse classes, respectively. By August 6th, average thaw depth had increased to 45, 62 and 66 cm at Upland, Sedge and Grass surface classes and over 100 cm at both the Shrub and Sparse classes.

A strong low-pressure system stalled off the coast between day of year (DOY) 199 and 204. This caused westerly winds to occur much more frequently than is typical for July and August. The 50%, 80% and 90% flux F_{Clim} contours are shown in Fig 1a. Mean source area fractions indicate the EC observations were skewed towards the Grass surface class and under-sampled the Shrub class, but the range of surface classes sampled was diverse enough to allow for testing of the impact of source area fraction on the fluxes (Table 2).

3.1 EC Observations

Half hourly observations of F_{CO2} and F_{CH4} along with the NEE_{NN} and NME_{NN} used to gap-fill the time series are shown in (Fig 2c & d). Gap-filled daily NEE ranged from -3.7 to -0.2 g C-CO_2 m^{-2} d^{-1} with a mean -1.5 [$CI_{95\%}$ ± 0.2] g C-CO_2 m^{-2} d^{-1}. Day to day variability was considerable but there was no notable trend in NEE over the peak growing season. The half hourly NEE during the study period reached a minimum of -10.4 µmol CO_2 m^{-2} h^{-1} just before solar noon and peaked at 4.7 µmol CO_2 m^{-2} h^{-1} around midnight (Fig 2c). NEE_{NN} was used to gap-fill the flux data because it was in good agreement with F_{CO2} observation (r^2 = 0.91). Daily ER_{NN} was estimated to be 2.2 [$CI_{95\%}$ ± 0.9] g C-CO_2 m^{-2} d^{-1} with corresponding GPP of 3.7 g C-CO_2 m^{-2} d^{-1}. ER_{NN} was in poor agreement (R^2 = 0.35, n= 95) with night time F_{CO2} observations. For comparison, Eq. 3 provided a better fit (R^2 = 0.47) with night-time EC data, and ER_{Q10} was estimated to be 3.0 g C-CO_2 m^{-2} d^{-1}. However, NEE_{Q10} did not fit F_{CO2} as well (r^2 = 0.80) as NEE_{NN}.

Gap-filled daily NME was modest and decreased over the study period. It ranged from 2.0 to 25.1 mg C-CH_4 m^{-2} d^{-1} with a mean of 8.7 [$CI_{95\%}$ ± 0.4] mg C-CH_4 m^{-2} d^{-1} (Fig 2d). NME_{NN} was used to gap-fill the flux data because it provided a reasonable fit (r^2 = 0.62) to F_{CH4} observations. NME did not constitute a significant component of the carbon balance and thus the flux footprint area was a carbon sink during the peak growing season with negative GWP after accounting for the greater GWP of CH_4.

3.2 Chamber Observations

ER was highest in the Sedge, Upland, and Grass classes where fluxes were very similar at 5.5 [$CI_{95\%}$ ± 1.2], 5.4 [$CI_{95\%}$ ± 1.2] and 4.9 [$CI_{95\%}$ ± 0.7] g C-CO_2 m^{-2} d^{-1}. Shrub ER was significantly less (3.5 [$CI_{95\%}$ ± 0.6] g C-CO_2 m^{-2} d^{-1}) than the other vegetated classes and Sparse ER was the lowest among the classes (2.0 [$CI_{95\%}$ ± 0.3] g C-CO_2 m^{-2} d^{-1}) (fig 3a). The Q_{10} and R_{10} values also differed between vegetation classes: ER in the Sedge was the most sensitive to changes in air temperature and modelled values provided the best fit (R^2 = 0.82) to observations. Upland and Grass had the highest base respiration and fit observations moderately well (Table 3).

NME was more variable between vegetation classes than ER (Fig 3b & c). Sedge was a very strong CH_4 source at 114.7 [$CI_{95\%}$ ± 15.3] mg C-CH_4 m^{-2} d^{-1}. Shrub and Grass were very weak sources, 0.7 [$CI_{95\%}$ ± 0.3] and 0.4 [$CI_{95\%}$ ± 0.3] mg C-CH_4 m^{-2} d^{-1}, respectively. Sparse was neutral. Upland was a net CH_4 sink -1.1 [$CI_{95\%}$ ± 0.4] mg C-CH_4 m^{-2} d^{-1}. Sedge and Shrub were NME were positively correlated with T_s (r=0.61, p < 0.01; r=0.35, p = 0.04) respectively and *VWC* (r=0.58, p < 0.01; r=0.5, < 0.01) respectively. They also had a positive correlation with T_a, while Upland NME was negatively correlated with T_a. Grass and Sparse didn't have any significant correlations.

Footprint scaled chamber fluxes were 59% and 47% higher than ER_{NN} or gap-filled NME, respectively. Mean ER_{FS} was 3.5 g C-CO_2 m^{-2} d^{-1} [$CI_{95\%}$ ± 0.1], it fit ER_{Q10} very well (R^2 = 0.95) as would be expected and ER_{NN} moderately

well (R^2 = 0.46). Mean NME_{FS} was 12.8[$CI_{95\%}$ ± 1.3] mg C-CH_4 m^{-2} d^{-1}, it did not fit NN_{NME} well (R^2 = 0.30). At the basin scale, ER_{BS} (3.4 [$CI_{95\%}$ ± 0.1] g C-CO_2 m^{-2} d^{-1}) was slightly lower than ER_{FS} because of the exclusion of upland areas. NME_{BS} was higher (15.2 [$CI_{95\%}$ ± 0.1] g C-CO_2 m^{-2} d^{-1}) because of the greater sedge fraction in the basin than the footprint because the (Table 2).

3.3 NEE Response to Environmental Factors and Vegetation Type

NEE_{NN} (r^2 = 0.91) was estimated using four factors: *PPFD*, *VPD*, *VWC*, and F_{Shrub}. *PPFD* is the primary control over NEE: a NN trained on *PPFD* alone provided a reasonable fit (r^2 = 0.83). The three additional factors, *VPD*, *VWC*, and F_{Shrub}, helped NN_{NEE} fit a wider variety of conditions. Examining the partial first derivative of NN_{NEE} under different conditions provides interpretation of the modelled light response curves (Fig 4). The minimum values represent the peak light use efficiency and are analogous to α in eq. 5 (Fig 4b). With increasing *PPFD*, light use becomes less efficient and approaches zeros as the light response nears light saturation (Fig 4b).

VPD was a secondary control over NEE. Increasing *VPD* increased peak light use efficiency and net CO_2 uptake until a threshold, above which it had a strong limiting effect (Fig 4a & b). For example, under dry atmospheric conditions (e.g. VPD = 1.5 kPa), peak light use is less efficient (-12 nmol CO_2 μmol^{-1} photon) than under more humid conditions (-18 nmol CO_2 μmol^{-1} photon). The value of this VPD threshold was dependent upon soil moisture: from 1 kPa when *VWC* was highest to 0.2 5Pa when *VWC* was low. Mapping NN_{NEE} and NN_{ER} at F_{Shrub} = 100%, F_{Shrub} = 0%, and F_{Shrub} = 36% (F_{Clim}), shows that *VWC* and F_{Shrub} were the primary controls over ER and thus influenced NEE (Fig 4c & d). We can see from the partial first derivates of NN_{ER} that increasing *VWC* increases ER from Shrub areas. In the absence of shrubs, increasing *VWC* inhibits ER although it is important to note that variations in VWC were subtle ranging from 51.7% to 59.0%. The partial first derivative of NN_{NEE} shows that *VWC* slightly limits NEE from non-Shrub areas and significantly reduces it in Shrub areas.

3.4 NME Response to Environmental Factors and Vegetation Type

NME was estimated using NME_{NN} (r^2 = 0.62) which had five factors: F_{Sedge}, F_{Shrub} , *VWC*, T_S, and *U*. NME was more variable and less dependent on any one factor than NEE which is why the NN_{NME} needed an extra factor and had a lower r^2 score. Source area had a significant effect on NME, and it was encouraging that the model contained F_{Sedge} and F_{Shrub} since Sedge and Shrub were the strongest CH_4 source and largest footprint component, respectively. These two factors can combine to map NME under three general situations: we can extrapolate to F_{Sedge} = 100 % and F_{Shrub} = 0 % or F_{Sedge} = 0 % and F_{Shrub} 100 %, or represent actual F_{Clim} where F_{Sedge} = 11% and F_{Shrub} = 37% (Table 2). Some upland tundra was included in the F_{Clim} estimate, which reduced NME.

VWC was the primary climatic driver identified by NN_{NME}. Wetter soils had a consistent positive effect on NME which was strongest when F_{Sedge} was high (Fig 5a & b). Between driest and wettest conditions, estimated NME increased: by an order of magnitude at F_{Sedge} = 100 %, 4-fold at F_{Shrub} = 100%, and from neutral to a source at F_{Clim} (Fig 5a). Higher T_s generally had a negative effect on NME (Fig 5c & d). The negative correlation between T_s and VWC (r = 0.54, < 0.01) may have contributed to this result. NN_{NME} performance improved less with the addition of

U indicating the NN_{NME} was near saturation and its effects are less relevant. Higher U had a weak limiting effect on NME when *VWC* was high and increased NME when *VWC* was low (not shown).

4 Discussion

4.1 Carbon Balance and Controlling Factors

Compared to other DTLB, Illisarvik has drier soils and greater shrub and grass cover (Table 4). Peak growing season CO_2 uptake at Illisarvik was greater than at most wet sedge-dominated DTLB (Table 4; Zona et al. 2010, Sturtevant and Oechel, 2013; Lara et al. 2015). These differences may be due to differences in the periods of observation and year to year variability but may also be due to the presence of more productive shrubs and slightly warmer climate at Illisarvik. Mean 1980-2010 T_a at Utqiagvik (formerly Barrow, AK) is -11.2 °C (US National Climate Data Centre, 2020). Tuktoyaktuk, the closest station to Illisarvik is 1.1° warmer. Shrub cover is expected to have a number of impacts on the microclimate and carbon cycle of Arctic tundra (eg. Myers-Smith et all, 2011). Typically, greater deciduous shrub cover is expected to increase GPP as a result of greater leaf area and photosynthetic potential compared to graminoid-dominated tundra (Sweet et al. 2015; Street et al., 2018). GPP was greater at Ilisarvik compared to the young wet-sedge dominated DTLB in Alaska (Zona et al., 2010). It was more similar to Katyk which has significant dwarf shrub cover, predominately *Betula nana* and *Salix pulchra* (van der Molen et al. 2007).

Differences in ER among tundra environments can be related to substrate availability, soil moisture and temperature and thaw depth, among other factors (Sturtevant and Oechel, 2013). The 'snow-shrub hypothesis' (Sturm et al. 2001) describes the potential for greater snow trapping in shrub communities which insulates soils in winter, leads to increased decomposition and nutrient availability and promotes further shrub growth. At Illisarvik, snow blowing in off the Arctic Ocean results in large snow drifts within the basin where snow depth correlates with vegetation height (Wilson et al., 2019). Wilson et al. (2019) concluded that the soils within the Illisarvik basin were warmer than those of the surrounding dwarf-shrub tundra in part through these snow-shrub interactions. Although our chamber observations suggested Shrub ER is lower than ER from other vegetation classes, this may have been an artifact as the taller shrubs (>40 cm) could not fit inside the chambers. In another study, chamber ER increased with greater shrub cover in upland tundra (Ge et al., 2017). ER at Illisarvik was greater than the ER observed at both the young wet-sedge DTLB in Barrow (Zona et al., 2010) and at the shrub/wet sedge DTLB at Katyk where thaw depth was much shallower (45 to >100 cm at Illisarvik vs. 25 to 40 cm at Katyk; van der Molen et al. 2007). The importance of F_{Shrub} in describing temporal variations in half hourly NEE within the flux footprint at Ilisarvik is further evidence of the importance of shrub cover on tundra carbon cycle processes in this environment.

PPFD and *VPD* were the most important factors for predicting half hourly NEE. This was to be expected as they are typically the primary controls over GPP (Aubinet et al., 2012). The limiting effects of VPD are consistent another study using NN to analyse NEE at a deciduous forest site (Moffat et al., 2012) and has been found at other tundra sites (Euskirchen et al. 2012; López-Blanco et al. 2017). *VWC* was also important at Illisarvik. Zona et al. (2010) found *VWC* could explain 70% of the variability in daily peak season ER in young DTLB. Similarly, Kittler et al. (2016)

found drier soils increased ER and decreased NEE after a wet tundra drainage experiment in Siberia, consistent with our results at Illisarvik when F_{Shrub} was low.

As expected, NME at Illisarvik was about half that observed at the Alaskan DTLB sites where soils were wetter with greater sedge cover (Table 4, Zona et al., 2009; Lara et al. 2015). NME at Katyk was even higher than the Barrow DTLB and had a significant impact on the greenhouse gas (GHG) balance for this site (van der Molen et al. 2007; Parmentier et al., 2011). In our NN modelling of NME at Illisarvik, F_{Sedge} was the most important factor for predicting half hourly F_{CH4}. Sedges are aquatic plant species with arenchymatous tissues that act as conduits for CH_4 from below the water table to the atmosphere and limits CH_4 oxidation by methanotrophs in aerobic surface soils (Lai et al. 2009). The inclusion of F_{Shrub} further refined the model, allowing it to better fit the site-specific distribution of vegetation types. Budishchev et al. (2014) found shrub and sedge fraction had a significant influence on F_{CH4} at Katyk. Vegetation type is the dominant control over NME across multiple tundra landscapes and our results further support that (Davidson et al., 2016).

VWC was the second most important factor, which was expected as CH_4 production occurs in anaerobic environments and has been linked to variability in CH_4 emission in many other studies (e.g. Zona et al., 2009; Nadeau et al., 2013; Olefeldt et al., 2013). Soil temperature (T_s) was the third most important factor. Higher T_s increase the oxidation potential of methanotrophs (Liu et al., 2016; King and Adamsen, 1992), so this result was expected for the drier portions of the basin and upland tundra. However, this was not expected for the sedge areas because most studies find NME in sedges is positively correlated to T_s (Olefeldt et al., 2013). The negative correlation between T_s and *VWC* may partly explain this.

4.2 Upscaling

ER_{FS} and NME_{FS} were about 59% and 47% greater than the EC estimates. Discrepancies between EC and chamber observations are common and have been attributed to differences in measurement techniques, the small sample size of chamber observations, and sampling bias since all chamber measurements were taken during the day with fair weather (Katayanagi et al., 2005; Chaichana et al., 2018). Meijide et al. (2011) found that chamber NEE could be up to twice as large as EC observations and Riederer et al. (2014) also found chamber NME estimates were about 30% higher than EC estimates. Others have been more successful, yielding upscaled chamber NME fluxes within 10% of EC observations (Zhang et al., 2012; Budishchev et al., 2014; Davidson et al., 2017). A potential reason for the disagreement with ER_{FS} may be the lack of direct observations by the EC system under low-light conditions. Another potential source of error for the upscaling is inaccuracies in the vegetation map.

4.3 Future Trajectories

Presently, peak growing season carbon uptake at Illisarvik is greater than similarly aged landscape features on the Barrow Peninsula, Alaska and more similar to levels observed at Katyk, Siberia. NME is well below levels observed at any other DTLB studied, making this site a stronger GHG sink than other DTLB. However, the basin at Illisarvik will continue to evolve and the trajectory it takes could significantly alter its carbon balance. Historically, DTLB on Richards Island and the Tuktoyaktuk Peninsula evolve into sedge wetlands, as do DTLB on the Barrow Peninsula

(Ovendend, 1986; Lara et al., 2015). Active maintenance of the outlet channel at Illisarvik has artificially lowered soil moisture and flooding and potentially limited this transition thus far (C. Burn, personal communication 2016). If Illisarvik follows the same trajectory as older DTLB in the area and becomes dominated by sedge wetlands, NME will increase significantly. With extrapolations to full Sedge cover (F_{Sedge} = 100%), NME would be similar to values on the Barrow Peninsula (Zona et al., 2009). If the basin instead transitions into a shrub dominated DTLB similar to those of Old Crow Flats, Yukon (Lantz et al., 2015), NME_{NN} would remain similar to current levels meaning the basin would remain a weak source of CH_4. These are projections well beyond F_{Clim} fractions observed so confidence in the specific values predicted is low.

The effects of changing shrub/sedge cover on Illisarvik's growing season NEE are less straightforward than on NME. Partly because Shrub cover had less overall influence on NEE_{NN}. Our model suggests ER decreases and GPP increases with increasing shrub coverage when soils are slightly drier, but has the opposite effect under wetter conditions. To our knowledge, only few winter season (e.g. Zona et al. 2016) and no year-round studies of NEE and NME have yet to be published to help evaluate the factors influencing DTLB carbon losses through the non-growing season months. Further observation year-round is needed to better understand the implications of continued vegetation change on the carbon balance of DTLB such as Illisarvik.

Line 241: You only mentioned thaw depth twice? Why not measure it on each day chambers were used?

[Responses]

Thaw depth tends to increase over time but at different rates at different locations within the basin as a result of varying soil and surface properties. We measured thaw depth at the start and end of the measurement period to highlight these differences rather than develop a variable that could be related to the fluxes. In past studies we found that day to day variations in respiration correlate best with near surface soil temperature and moisture rather than thaw depth while spatial variations in average fluxes can sometimes correlate to max thaw depth. In the revised manuscript we refer to thaw depth more and use it to compare Illisarvik to Katyk Line 384

ER at Illisarvik was greater than the ER observed at both the young wet-sedge DTLB in Barrow (Zona et al., 2010) and at the shrub/wet sedge DTLB at Katyk where thaw depth was much shallower (45 to >100 cm at Illisarvik vs. 25 to 40 cm at Katyk; van der Molen et al. 2007).

Section 3.1: I think this section needs an overhaul unfortunately. Many of the sentences do not make sense in their current format. For example: Lines 254-257: 'NEE was greater than (ie. Less carbon uptake) 255 EC observations of from four wetter, sedge dominated DTLB, where peak season NEE was -2.5 g C-CO2 m-2 d-1, ER (1.5 g C-CO2 m-2 d-1) was lower than at Illisarvik while GPP (4.0 g C-CO2 m-2 d-1) was slightly higher (Zona et al., 2010).'

[Responses]

The entire section has been rewritten and the references to other work were moved out from the results to the discussion section. See previous comment response.

I also think it might be useful to separate the EC results and the chamber results into subsections. By referring your measured values to other studies in the results, it makes it hard to follow for the reader.

[Responses]

We agree and have separated the EC and chamber results into Sections 3.1 and 3.2. These changes are shown in the above comment about section 3.

Further, by combining the results, there is a lack of discussing the results (for example, it feels like only section 3.4 is really doing this). It sadly reads as a lot of results statements and then suddenly we are at the conclusion.

[Responses]

We added a separate "discussion" section (Section 4). These changes are shown in the above comment about section 3.

Line 273: Why compare methane to ER here?

[Responses]

We thought it is relevant to contextualize the differences. Methane emissions are far more spatially variable than ecosystem respiration. We revised the sentence to get rid of the comparison of the magnitude because that is less relevant, but left the point about spatial variability being enhanced [Line 316]. We think it is an important finding to show that NME is more influenced by spatial heterogeneity than ER.

NME was more variable between vegetation classes than ER (Fig 3b & c).

Line 286: Although discrepancies do occur between upscaling chamber measurements and EC measurements – some studies have done it successfully and I think it would be good to include here as a caveat;

- **Budischev et al. 2014: Evaluation of a plot-scale methane emission model using eddy covariance observations and footprint modelling. Biogeosciences 11. 4651-4664**
- **Zhang et al. 2012: Upscaling methane fluxes from closed chambers to eddy covariance based on a permafrost integrated model. Global Change Biology, 18, 1428-1440.**
- **Davidson et al. 2017. Upscaling CH4 fluxes using high-resolution imagery in Arctic Tundra Ecosystems. Remote Sensing, 9, 1227; doi:10.3390/rs9121227**
- **Sachs et al. 2010 Environmental controls on CH4 emissions from polygonal tundra on the microsite scale in the Lena river delta, Siberia. Global Change Biology, 16, 3096-3110**

[Responses]

We reviewed this literature and added it to the newly separated "discussion" section [Lines 417-418].

Others have been more successful, yielding upscaled chamber NME fluxes within 10% of EC observations (Zhang et al., 2012; Budishchev et al., 2014; Davidson et al., 2017).

We also decided to use the footprint weighted upscaling method discussed in Budishchev et al. (2014) for the chamber upscaling [Line 207-208], but it did not make an appreciable difference in the upscaled chamber ER or NME.

Chamber fluxes of ER were upscaled from the plot scale (individual chamber) to the footprint scale using the footprint weighted average method and to the basin scale using the area weighted average method (Budishchev et al., 2014).

I think more discussion of the results in the context of other GHG literature from other tundra ecosites would be useful. Although this study is focused on drained lake basins, the results are comparable to wet-sedge dominated tundra landscapes. I feel this would be a good addition and strengthen what is already a useful paper.

[Responses]

We added section 4.1 (see above comment) where we discuss NEE and NME observations at Illisarvik relative to natural shrub vs. sedge-dominated DTLB to highlight the differences among these environments rather than attempt to fully contrast Illisarvik to a myriad of arctic tundra types/sites. These comparisons are always challenging given different years, time periods within a year, instrumentation, and data presentation. However, we believe we make a strong argument that shrub vs. sedge-dominated DTLB have the potential to differ and Illisarvik differs in particular from all other DTLB in its low methane emissions. (Table 4). We now further highlight the important implications of vegetation succession on CO2 and CH4 fluxes at our site in the discussion section 4.3 "Future Trajectories". This is one of the key messages associated with DTLBs – they undergo relatively rapid vegetation change over a number of decades that will influence their C budgets.

Table 4: Growing season (gs) daily range in eddy covariance-derived NEE and NME from drained thermokarst lake basins (DTLB) and other select wetland/coastal tundra sites across the Arctic. The period of studies measurements for the studies observations are: a) mid-June – end of July b) June 12 – August 28, 2007, Fig 4 c) June 11 – August 25, 2011 d) upscaled chamber estimates, exact dates not specified, e) mean June 15 –August 31 2003-2006, f) July 5 – Aug 4, 2009.

Site	Site Characteristics	NEE g C-CO_2 m^{-2} d^{-1}	NME mg C-CH_4 m^{-2} d^{-1}	Studies
Illisarvik	Young DTLB, Low & Tall Shrub/Grass/Wet Sedge	-1.5	8.7	(this study)
Various DTLB, Barrow Peninsula, Alaska	Young DTLB, Wet Sedge Tundra	-1.1[b], -0.9[d], -0.8[c]	18.4[a], 26.1[d], 44.0[c]	Zona et al. 2009[a] & 2010[b], Sturtevant and Oechel, 2013[c]; Lara et al. 2015[d]
	Medium DTLB, Wet Sedge Tundra	-0.7[b], -0.6[d], -0.4[c]	27.0[d], 41.3[c]	
	Old DTLB, Wet Sedge Tundra	-1.0[b], -0.4[d], 0.1[c]	24.2[d], 38.7[c]	
	Ancient DTLB, Wet Sedge Tundra	0.4[d]	21.7[d]	

Katyky, Indigirka lowlands, Siberia	Ancient DTLB, Dwarf-Shrub and Wet Sedge Tundra	-1.3[e]	36.0[f]	Van der Molen et al. 2007[e], Budishchev et al. 2014[f]

Figure 2a: Could you change the colour of the T_a line? Red on orange is difficult to read.

[Responses]

Agreed. We have changed the bar colour, setting the orange to grey to make it easier to distinguish.

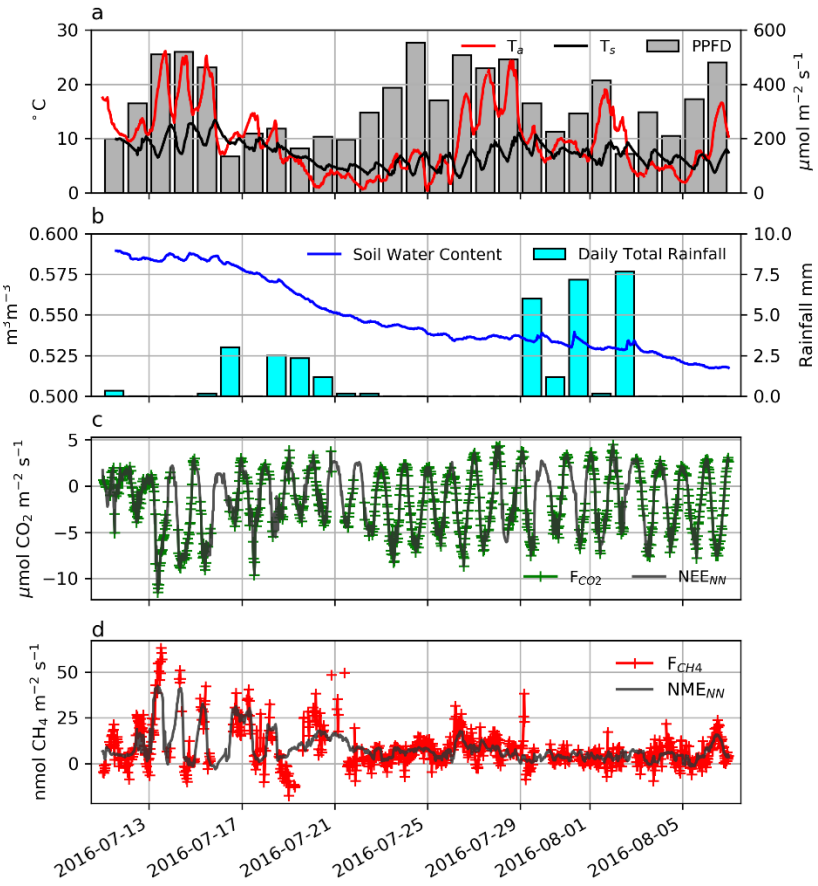


Figure 3: I will leave this up to the author's discretion, but I wonder if this figure (and in fact, all figures) would benefit from having a plain white background. I find all the lines distracting. Especially when other lines are being used to annotate.

[Responses]

We agree the grid is distracting for Figure 3 and removed it, but left grids in all the other figures.

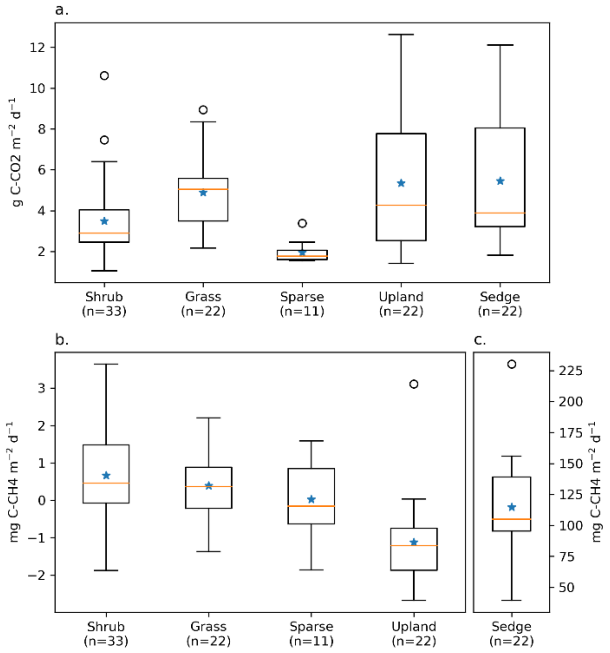


Figure 4c: Please use another three colours for Shrub. It is confusing that they are the same colour as VPD on the left-hand panel.

[Responses]

We changed the colour scheme to address the concern.

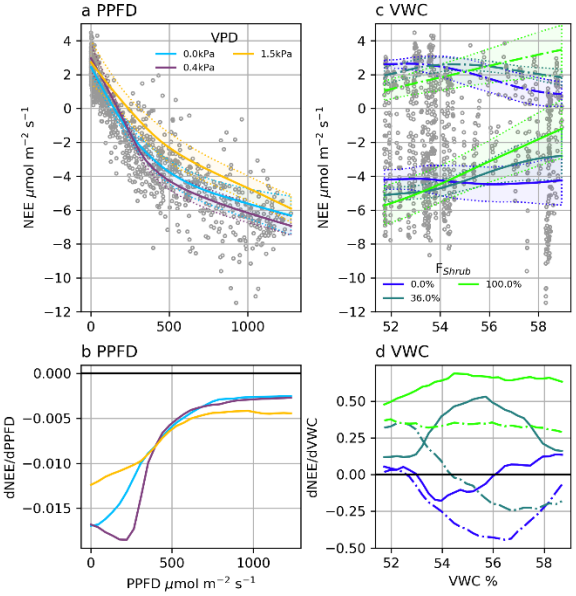


Figure 4 and 5: I think these figures would benefit from having a title for each panel – it was not clear to me initially the difference between Figure 5a and b. I think just putting VWC above left hand panel and T_s above right hand panel, this would make it much clearer.

[Responses]

We added subtitles to all panels in Figures 4 and 5.

Tables: Caption should go above the Tables.

[Responses]

Captions were all moved above Tables.

I think a table including the dominant vegetation species for each class would be super useful for the reader.

[Responses]

We updated Table 1 to be more specific and included additional information where known/applicable.

Table 1: Dominant species or landscape feature within the vegetation/cover classes. Unit codes correspond to the map Figure 1a.

Unit Code	Vegetation Class	Dominant Species/Landscape feature
1a	Shrub	*Salix alaxnesis* (Tall Willow)
1b	Shrub	*Salix glauca* (Low Willow)
1c	Shrub	*Alnus viridis* subsp. *crispa* (Alder)
2a	Sedge Marsh	*Carex aquatilis* (Sedge)
2b	Sedge Marsh	*Arctophila fulva* (Pendant Grass)
3	Grass Meadow	*Pocacea* spp. (Grasses), *Eriophorum angustifolium* (Cotton Grass)
4a	Sparse Cover	Sparse Vegetation
4b	Sparse Cover	Bare Ground
5	Ponds	*Hippuris vulgaris* (Mare's Tail), Open Water
6a	Outside of Basin	Dwarf shrub tundra: *Salix* spp. & *Betula nana* (Birch)
6b	Outside of Basin	Fen
6c	Outside of Basin	Ocean

Reviewer # 2:

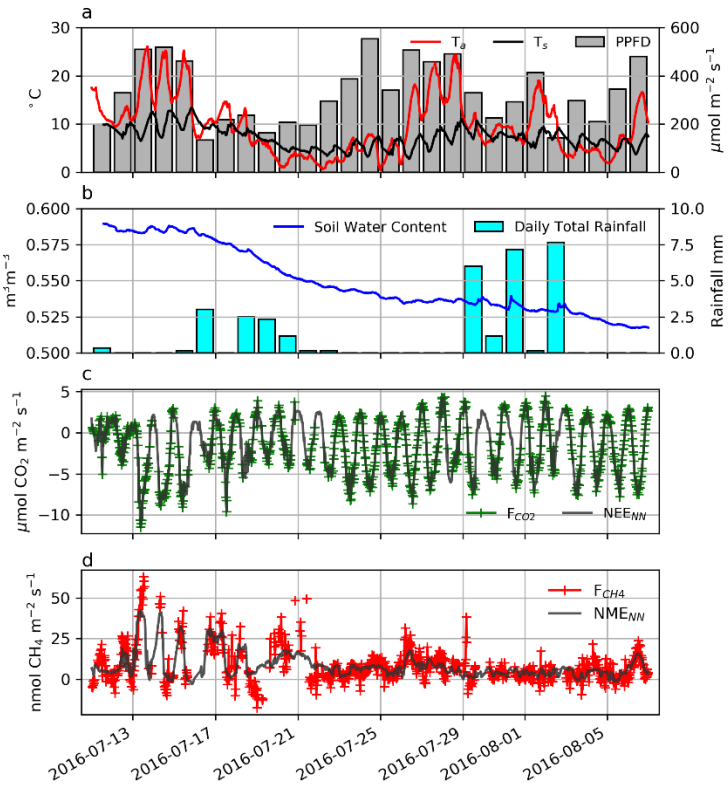
General comments

The paper "Vegetation Influence and Environmental Controls on Greenhouse Gas Fluxes from a Drained Thermokarst Lake in the Western Canadian Arctic" by June Skeeter et al. reports CO2 and CH4 flux measurements from a permafrost tundra site in Western Canada. Eddy covariance and chamber flux measurements were taken during the growing season 2016, and analysed accounting for the spatial variability of vegetation cover. Statistical gap-filling and an analysis of the environmental controls of the fluxes is performed using artificial neural networks. I think the chosen methods are properly applied and explained. The results are presented clearly and the conclusions are supported by the results. Also, the paper is very well written.

Given the remote and rather special site location, this study should be very valuable for the arctic carbon flux community. As the flux time series collected in your study may be used and referred to in future studies, it would be nice if you could present the time series in a more raw format than you do in Figure 2. For example, a plot of the 30-minute flux time series would help to understand the character of the data. This is also relevant, because I guess the performance and output of your NNs could be susceptible to noise or outliers in the EC time series.

[Responses]

We changed Fig 2 to show the half hourly F_{CO2} and F_{CH4} observations, along with the NEE_{NN} and NME_{NN}.



Also, several of your results (cf. Line 221 and Line 342) are based on extrapolations into parts of the parameter space where the flux response could be governed by processed not captured in your NNs. Perhaps these statistical uncertainties could be discussed.

[Responses]

We added a sentence at the end of section 2.5.1 discussing the impact of calculating ER by extrapolation and its impact on the confidence of ER estimates relative to NEE [Lines 265-267]. We refer the reader to Appendix A for details on the calculation of confidence intervals around NN outputs [Lines 501 -521].

This is a projection outside of the observed parameter space resulting in greater uncertainty and a wider confidence interval around ER_{NN} than NEE_{NN}. Calculation of confidence intervals for NN outputs is discussed in Appendix A

We also added a sentence to section 4.3 noting that projecting to Sedge = 100% is well outside of parameter space [Lines 433-434].

These are projections well beyond F_{Clim} fractions observed so confidence in the specific values predicted is low.

I understand there is little research from DTLB sites, but it would be good in your discussion to relate your findings to those from other tundra sites with (and without) thermokarst. In this discussion, it would be good to elaborate further on the peculiarities of the artificial draining performed at your site. Given the title of this paper, readers will probably expect more of these aspects discussed.

[Responses]

We split the results into two separate sections "results" (Section 3) and "discussion" (Section 4). In section 4.1 we discuss NEE and NME observations at Illisarvik relative to natural shrub vs. sedge-dominated DTLB to highlight the differences among these environments rather than attempt to fully contrast Illisarvik to a myriad of arctic tundra types/sites. These comparisons are always challenging given different years, time periods within a year, instrumentation, and data presentation. However, we believe we make a strong argument that shrub vs. sedge-dominated DTLB have the potential to differ and Illisarvik differs in particular from all other DTLB in its low methane emissions. (Table 4). We now further highlight the important implications of vegetation succession on CO2 and CH4 fluxes at our site in the discussion section 4.3 "Future Trajectories". This is one of the key messages associated with DTLBs – they undergo relatively rapid vegetation change over a number of decades that will influence their C budgets.

3 Results

During the 29-day study, half-hourly T_a and T_s ranged between 0.4 and 26.2°C and 4.4 and 11.0°C, respectively (Fig 2a). Day length and maximum solar altitude decreased from 24 hours to 19.25 hours and 41.6˚ to 35.4˚, but daily *PPFD* was more influenced by variations in cloud cover. Precipitation (19 mm) fell on 14 of the 28 days with trace snowfall on three of those days, but *VWC* of the soils decreased throughout the period (Fig 2b). At the onset of the study period, *VWC* was high and soils were saturated with ponding in the sedge areas. By the end of the study most of this surface water had dried up. On July 11th average thaw depth (cm) was 37, 45, 51, 64, 81 at Upland, Sedge,

Grass, Shrub, and Sparse classes, respectively. By August 6th, average thaw depth had increased to 45, 62 and 66 cm at Upland, Sedge and Grass surface classes and over 100 cm at both the Shrub and Sparse classes.

A strong low-pressure system stalled off the coast between day of year (DOY) 199 and 204. This caused westerly winds to occur much more frequently than is typical for July and August. The 50%, 80% and 90% flux F_{Clim} contours are shown in Fig 1a. Mean source area fractions indicate the EC observations were skewed towards the Grass surface class and under-sampled the Shrub class, but the range of surface classes sampled was diverse enough to allow for testing of the impact of source area fraction on the fluxes (Table 2).

3.1 EC Observations

Half hourly observations of F_{CO2} and F_{CH4} along with the NEE_{NN} and NME_{NN} used to gap-fill the time series are shown in (Fig 2c & d). Gap-filled daily NEE ranged from -3.7 to -0.2 g C-CO_2 m^{-2} d^{-1} with a mean -1.5 [$CI_{95\%}$ ± 0.2] g C-CO_2 m^{-2} d^{-1}. Day to day variability was considerable but there was no notable trend in NEE over the peak growing season. The half hourly NEE during the study period reached a minimum of -10.4 µmol CO_2 m^{-2} h^{-1} just before solar noon and peaked at 4.7 µmol CO_2 m^{-2} h^{-1} around midnight (Fig 2c). NEE_{NN} was used to gap-fill the flux data because it was in good agreement with F_{CO2} observation (r^2 = 0.91). Daily ER_{NN} was estimated to be 2.2 [$CI_{95\%}$ ± 0.9] g C-CO_2 m^{-2} d^{-1} with corresponding GPP of 3.7 g C-CO_2 m^{-2} d^{-1}. ER_{NN} was in poor agreement (R^2 = 0.35, n= 95) with night time F_{CO2} observations. For comparison, Eq. 3 provided a better fit (R^2 = 0.47) with night-time EC data, and ER_{Q10} was estimated to be 3.0 g C-CO_2 m^{-2} d^{-1}. However, NEE_{Q10} did not fit F_{CO2} as well (r^2 = 0.80) as NEE_{NN}.

Gap-filled daily NME was modest and decreased over the study period. It ranged from 2.0 to 25.1 mg C-CH_4 m^{-2} d^{-1} with a mean of 8.7 [$CI_{95\%}$ ± 0.4] mg C-CH_4 m^{-2} d^{-1} (Fig 2d). NME_{NN} was used to gap-fill the flux data because it provided a reasonable fit (r^2 = 0.62) to F_{CH4} observations. NME did not constitute a significant component of the carbon balance and thus the flux footprint area was a carbon sink during the peak growing season with negative GWP after accounting for the greater GWP of CH_4.

3.2 Chamber Observations

ER was highest in the Sedge, Upland, and Grass classes where fluxes were very similar at 5.5 [$CI_{95\%}$ ± 1.2], 5.4 [$CI_{95\%}$ ± 1.2] and 4.9 [$CI_{95\%}$ ± 0.7] g C-CO_2 m^{-2} d^{-1}. Shrub ER was significantly less (3.5 [$CI_{95\%}$ ± 0.6] g C-CO_2 m^{-2} d^{-1}) than the other vegetated classes and Sparse ER was the lowest among the classes (2.0 [$CI_{95\%}$ ± 0.3] g C-CO_2 m^{-2} d^{-1}) (fig 3a). The Q_{10} and R_{10} values also differed between vegetation classes: ER in the Sedge was the most sensitive to changes in air temperature and modelled values provided the best fit (R^2 = 0.82) to observations. Upland and Grass had the highest base respiration and fit observations moderately well (Table 3).

NME was more variable between vegetation classes than ER (Fig 3b & c). Sedge was a very strong CH_4 source at 114.7 [$CI_{95\%}$ ± 15.3] mg C-CH_4 m^{-2} d^{-1}. Shrub and Grass were very weak sources, 0.7 [$CI_{95\%}$ ± 0.3] and 0.4 [$CI_{95\%}$ ± 0.3] mg C-CH_4 m^{-2} d^{-1}, respectively. Sparse was neutral. Upland was a net CH_4 sink -1.1 [$CI_{95\%}$ ± 0.4] mg C-CH_4 m^{-2} d^{-1}. Sedge and Shrub were NME were positively correlated with T_s (r=0.61, p < 0.01; r=0.35, p = 0.04) respectively and *VWC* (r=0.58, p < 0.01; r=0.5, < 0.01) respectively. They also had a positive correlation with T_a, while Upland NME was negatively correlated with T_a. Grass and Sparse didn't have any significant correlations.

Footprint scaled chamber fluxes were 59% and 47% higher than ER_{NN} or gap-filled NME, respectively. Mean ER_{FS} was 3.5 g C-CO_2 m^{-2} d^{-1} [$CI_{95\%}$ ± 0.1], it fit ER_{Q10} very well (R^2 = 0.95) as would be expected and ER_{NN} moderately well (R^2 = 0.46). Mean NME_{FS} was 12.8[$CI_{95\%}$ ± 1.3] mg C-CH_4 m^{-2} d^{-1}, it did not fit NN_{NME} well (R^2 = 0.30). At the basin scale, ER_{BS} (3.4 [$CI_{95\%}$ ± 0.1] g C-CO_2 m^{-2} d^{-1}) was slightly lower than ER_{FS} because of the exclusion of upland areas. NME_{BS} was higher (15.2 [$CI_{95\%}$ ± 0.1] g C-CO_2 m^{-2} d^{-1}) because of the greater sedge fraction in the basin than the footprint because the (Table 2).

3.3 NEE Response to Environmental Factors and Vegetation Type

NEE_{NN} (r^2 = 0.91) was estimated using four factors: *PPFD*, *VPD*, *VWC*, and F_{Shrub}. *PPFD* is the primary control over NEE: a NN trained on *PPFD* alone provided a reasonable fit (r^2 = 0.83). The three additional factors, *VPD*, *VWC*, and F_{Shrub}, helped NN_{NEE} fit a wider variety of conditions. Examining the partial first derivative of NN_{NEE} under different conditions provides interpretation of the modelled light response curves (Fig 4). The minimum values represent the peak light use efficiency and are analogous to α in eq. 5 (Fig 4b). With increasing *PPFD*, light use becomes less efficient and approaches zeros as the light response nears light saturation (Fig 4b).

VPD was a secondary control over NEE. Increasing *VPD* increased peak light use efficiency and net CO_2 uptake until a threshold, above which it had a strong limiting effect (Fig 4a & b). For example, under dry atmospheric conditions (e.g. VPD = 1.5 kPa), peak light use is less efficient (-12 nmol CO_2 μmol^{-1} photon) than under more humid conditions (-18 nmol CO_2 μmol^{-1} photon). The value of this VPD threshold was dependent upon soil moisture: from 1 kPa when *VWC* was highest to 0.2 5Pa when *VWC* was low. Mapping NN_{NEE} and NN_{ER} at F_{Shrub} = 100%, F_{Shrub} = 0%, and F_{Shrub} = 36% (F_{Clim}), shows that *VWC* and F_{Shrub} were the primary controls over ER and thus influenced NEE (Fig 4c & d). We can see from the partial first derivates of NN_{ER} that increasing *VWC* increases ER from Shrub areas. In the absence of shrubs, increasing *VWC* inhibits ER although it is important to note that variations in VWC were subtle ranging from 51.7% to 59.0%. The partial first derivative of NN_{NEE} shows that *VWC* slightly limits NEE from non-Shrub areas and significantly reduces it in Shrub areas.

3.4 NME Response to Environmental Factors and Vegetation Type

NME was estimated using NME_{NN} (r^2 = 0.62) which had five factors: F_{Sedge}, F_{Shrub} , *VWC*, T_S, and *U*. NME was more variable and less dependent on any one factor than NEE which is why the NN_{NME} needed an extra factor and had a lower r^2 score. Source area had a significant effect on NME, and it was encouraging that the model contained F_{Sedge} and F_{Shrub} since Sedge and Shrub were the strongest CH_4 source and largest footprint component, respectively. These two factors can combine to map NME under three general situations: we can extrapolate to F_{Sedge} = 100 % and F_{Shrub} = 0 % or F_{Sedge} = 0 % and F_{Shrub} 100 %, or represent actual F_{Clim} where F_{Sedge} = 11% and F_{Shrub} = 37% (Table 2). Some upland tundra was included in the F_{Clim} estimate, which reduced NME.

VWC was the primary climatic driver identified by NN_{NME}. Wetter soils had a consistent positive effect on NME which was strongest when F_{Sedge} was high (Fig 5a & b). Between driest and wettest conditions, estimated NME increased: by an order of magnitude at F_{Sedge} = 100 %, 4-fold at F_{Shrub} = 100%, and from neutral to a source at F_{Clim} (Fig 5a). Higher T_s generally had a negative effect on NME (Fig 5c & d). The negative correlation between T_s and

VWC (r = 0.54, < 0.01) may have contributed to this result. NN_{NME} performance improved less with the addition of *U* indicating the NN_{NME} was near saturation and its effects are less relevant. Higher U had a weak limiting effect on NME when *VWC* was high and increased NME when *VWC* was low (not shown).

4 Discussion

4.1 Carbon Balance and Controlling Factors

Compared to other DTLB, Illisarvik has drier soils and greater shrub and grass cover (Table 4). Peak growing season CO_2 uptake at Illisarvik was greater than at most wet sedge-dominated DTLB (Table 4; Zona et al. 2010, Sturtevant and Oechel, 2013; Lara et al. 2015). These differences may be due to differences in the periods of observation and year to year variability but may also be due to the presence of more productive shrubs and slightly warmer climate at Illisarvik. Mean 1980-2010 T_a at Utqiagvik (formerly Barrow, AK) is -11.2 °C (US National Climate Data Centre, 2020). Tuktoyaktuk, the closest station to Illisarvik is 1.1° warmer. Shrub cover is expected to have a number of impacts on the microclimate and carbon cycle of Arctic tundra (eg. Myers-Smith et all, 2011). Typically, greater deciduous shrub cover is expected to increase GPP as a result of greater leaf area and photosynthetic potential compared to graminoid-dominated tundra (Sweet et al. 2015; Street et al., 2018). GPP was greater at Ilisarvik compared to the young wet-sedge dominated DTLB in Alaska (Zona et al., 2010). It was more similar to Katyk which has significant dwarf shrub cover, predominately *Betula nana* and *Salix pulchra* (van der Molen et al. 2007).

Differences in ER among tundra environments can be related to substrate availability, soil moisture and temperature and thaw depth, among other factors (Sturtevant and Oechel, 2013). The 'snow-shrub hypothesis' (Sturm et al. 2001) describes the potential for greater snow trapping in shrub communities which insulates soils in winter, leads to increased decomposition and nutrient availability and promotes further shrub growth. At Illisarvik, snow blowing in off the Arctic Ocean results in large snow drifts within the basin where snow depth correlates with vegetation height (Wilson et al., 2019). Wilson et al. (2019) concluded that the soils within the Illisarvik basin were warmer than those of the surrounding dwarf-shrub tundra in part through these snow-shrub interactions. Although our chamber observations suggested Shrub ER is lower than ER from other vegetation classes, this may have been an artifact as the taller shrubs (>40 cm) could not fit inside the chambers. In another study, chamber ER increased with greater shrub cover in upland tundra (Ge et al., 2017). ER at Illisarvik was greater than the ER observed at both the young wet-sedge DTLB in Barrow (Zona et al., 2010) and at the shrub/wet sedge DTLB at Katyk where thaw depth was much shallower (45 to >100 cm at Illisarvik vs. 25 to 40 cm at Katyk; van der Molen et al. 2007). The importance of F_{Shrub} in describing temporal variations in half hourly NEE within the flux footprint at Ilisarvik is further evidence of the importance of shrub cover on tundra carbon cycle processes in this environment.

PPFD and *VPD* were the most important factors for predicting half hourly NEE. This was to be expected as they are typically the primary controls over GPP (Aubinet et al., 2012). The limiting effects of VPD are consistent another study using NN to analyse NEE at a deciduous forest site (Moffat et al., 2012) and has been found at other tundra sites (Euskirchen et al. 2012; López-Blanco et al. 2017). *VWC* was also important at Illisarvik. Zona et al. (2010) found *VWC* could explain 70% of the variability in daily peak season ER in young DTLB. Similarly, Kittler et al. (2016)

found drier soils increased ER and decreased NEE after a wet tundra drainage experiment in Siberia, consistent with our results at Illisarvik when F_{Shrub} was low.

As expected, NME at Illisarvik was about half that observed at the Alaskan DTLB sites where soils were wetter with greater sedge cover (Table 4, Zona et al., 2009; Lara et al. 2015). NME at Katyk was even higher than the Barrow DTLB and had a significant impact on the greenhouse gas (GHG) balance for this site (van der Molen et al. 2007; Parmentier et al., 2011). In our NN modelling of NME at Illisarvik, F_{Sedge} was the most important factor for predicting half hourly F_{CH4}. Sedges are aquatic plant species with arenchymatous tissues that act as conduits for CH_4 from below the water table to the atmosphere and limits CH_4 oxidation by methanotrophs in aerobic surface soils (Lai et al. 2009). The inclusion of F_{Shrub} further refined the model, allowing it to better fit the site-specific distribution of vegetation types. Budishchev et al. (2014) found shrub and sedge fraction had a significant influence on F_{CH4} at Katyk. Vegetation type is the dominant control over NME across multiple tundra landscapes and our results further support that (Davidson et al., 2016).

VWC was the second most important factor, which was expected as CH_4 production occurs in anaerobic environments and has been linked to variability in CH_4 emission in many other studies (e.g. Zona et al., 2009; Nadeau et al., 2013; Olefeldt et al., 2013). Soil temperature (T_s) was the third most important factor. Higher T_s increase the oxidation potential of methanotrophs (Liu et al., 2016; King and Adamsen, 1992), so this result was expected for the drier portions of the basin and upland tundra. However, this was not expected for the sedge areas because most studies find NME in sedges is positively correlated to T_s (Olefeldt et al., 2013). The negative correlation between T_s and *VWC* may partly explain this.

4.2 Upscaling

ER_{FS} and NME_{FS} were about 59% and 47% greater than the EC estimates. Discrepancies between EC and chamber observations are common and have been attributed to differences in measurement techniques, the small sample size of chamber observations, and sampling bias since all chamber measurements were taken during the day with fair weather (Katayanagi et al., 2005; Chaichana et al., 2018). Meijide et al. (2011) found that chamber NEE could be up to twice as large as EC observations and Riederer et al. (2014) also found chamber NME estimates were about 30% higher than EC estimates. Others have been more successful, yielding upscaled chamber NME fluxes within 10% of EC observations (Zhang et al., 2012; Budishchev et al., 2014; Davidson et al., 2017). A potential reason for the disagreement with ER_{FS} may be the lack of direct observations by the EC system under low-light conditions. Another potential source of error for the upscaling is inaccuracies in the vegetation map.

4.3 Future Trajectories

Presently, peak growing season carbon uptake at Illisarvik is greater than similarly aged landscape features on the Barrow Peninsula, Alaska and more similar to levels observed at Katyk, Siberia. NME is well below levels observed at any other DTLB studied, making this site a stronger GHG sink than other DTLB. However, the basin at Illisarvik will continue to evolve and the trajectory it takes could significantly alter its carbon balance. Historically, DTLB on Richards Island and the Tuktoyaktuk Peninsula evolve into sedge wetlands, as do DTLB on the Barrow Peninsula

(Ovendend, 1986; Lara et al., 2015). Active maintenance of the outlet channel at Illisarvik has artificially lowered soil moisture and flooding and potentially limited this transition thus far (C. Burn, personal communication 2016). If Illisarvik follows the same trajectory as older DTLB in the area and becomes dominated by sedge wetlands, NME will increase significantly. With extrapolations to full Sedge cover (F_{Sedge} = 100%), NME would be similar to values on the Barrow Peninsula (Zona et al., 2009). If the basin instead transitions into a shrub dominated DTLB similar to those of Old Crow Flats, Yukon (Lantz et al., 2015), NME_{NN} would remain similar to current levels meaning the basin would remain a weak source of CH_4. These are projections well beyond F_{Clim} fractions observed so confidence in the specific values predicted is low.

The effects of changing shrub/sedge cover on Illisarvik's growing season NEE are less straightforward than on NME. Partly because Shrub cover had less overall influence on NEE_{NN}. Our model suggests ER decreases and GPP increases with increasing shrub coverage when soils are slightly drier, but has the opposite effect under wetter conditions. To our knowledge, only few winter season (e.g. Zona et al. 2016) and no year-round studies of NEE and NME have yet to be published to help evaluate the factors influencing DTLB carbon losses through the non-growing season months. Further observation year-round is needed to better understand the implications of continued vegetation change on the carbon balance of DTLB such as Illisarvik.

Specific comments

Line 16: "During the study period". Please be more specific here, because the upscaled average fluxes you mention in lines 18 and 20/21 don't tell much if you don't know the study period.

[Responses]

Changed wording to "peak growing season" Line 16

Line 24: Your abstract lacks a broader conclusion

[Responses]

We added another sentence to make a broader conclusion about plant succession and Illisarvik's carbon balance, Lines 25-26

Presently, Illisarvik is a carbon sink during the peak growing season. However, these results suggest that rates of growing season CO2 and CH4 exchange rates may change as the basin's vegetation community continues to evolve.

Line 100: Could the grazing have a measurable effect on e.g. NEE? It could be a point to add to your discussion.

[Responses]

Good point, we added a few words to mention that grazing may have affected GHG Fluxes [Lines 121-122].

which may have affected greenhouse gas fluxes.

It is possible that grazing had some impact, but we cannot answer this based on the data collected. According to images from a fish eye camera mounted on the tripod (taken at 5-minute intervals), the animals spent about an hour gazing within the footprint of the eddy-covariance tower. In other areas of the basin where they stayed for longer, there was definitely a more significant impact. They were only spotted within the footprint the morning of July 12th. In addition

to the fish eye camera images, we were present at the field site during the full campaign and observed the reindeer's movements.

Line 116: You discarded a sector because its flow could be disturbed by the tower.

But did you see this effect in any of your quality checks? Maybe it's not necessary to discard this data.

[Responses]

It is standard practice to discard winds affected in the wake of the tower and sensor head. We have added a reference to Aubinet et al., 2012 to support this choice. During light winds, windspeeds can be reduced as much as 50% in the wake of a tower/instrument mount and turbulent eddies are artificially created, significantly violating the assumptions that go into eddy-covariance flux calculations. We oriented the tower such that this wind sector was the least frequent (according to climatology from Tuktoyaktuk). It only resulted in 6.7% (86 of 1279) half hourly observations being discarded.

Line 146: Maybe be more specific about the Python modules you used, otherwise this sentence adds very little to the understanding of your analysis.

[Responses]

We removed this portion of the manuscript. Most of the code was written specifically for the project by the first author, the footprint model of Kljun et al. 2015 is available in multiple programming languages, and we mention the python module for the neural networks on Line 243 and discuss the procedures in more in the appendix.

Line 182: Shouldn't there be five times more vials than flux estimates, if you used 5 gas samples per flux measurement?

[Responses]

Yes, that is correct. The sentence has been corrected [Lines 202-203].

After removal of spurious point measurements (72 vial samples were rejected out of 1135 vials), dc/dt was determined using three or more gas sample concentrations resulting in coefficients of determination that ranged from 0.71 to 0.99

The '681 flux measurements' referred to the three different gas fluxes each measurement produced (CO_2, CH_4, and N_2O). For clarity, we have removed this number (681) and we now only refer to CO_2 and CH_4 flux measurements (see also reply to reviewer #1)

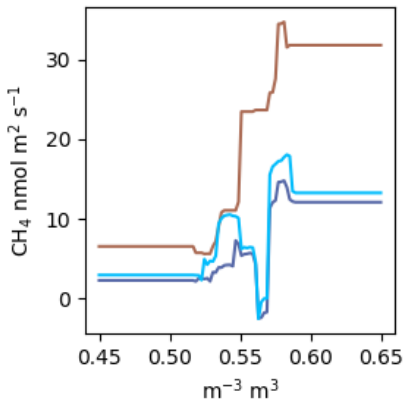
Line 292: "Random forest regression tree". Did you use only one decision tree, or the ensemble mean of several?

[Responses]

It was the ensemble mean of 100 trees. We removed this from the text however. In retrospect, it was beyond the scope/point of the paper. Discussing the choices made for the random forest (RF) analysis would have required a new section in the methods. But since we didn't use it in the results (beyond this one comparison), this didn't seem necessary. Instead, we added a paragraph to the appendix [Lines 522-532] discussing why we RF weren't the best choice for this analysis and we added Figure A2 to support this.

Random forests (RF) are said to be among the best performing gap filling methods for NME (Kim et al., 2020). and it has been claimed that aggregating many regression trees in a RF prevents overfitting (Breiman, 2001;). We did not find this to be the case. Following the methods outlined in Kim et al. (2020): a RF with 400 trees and no restrictions on tree size fit FCH4 nearly perfectly (R2 = 0.98). Without considerable limitations on tree size, the RF will just learn

the dataset rather than the relationships present. It is our view that this tree is extremely overfit, as highlighted by the example in Figure A2. Further, RF do not allow for straightforward visualization functional relationships in a dataset. Plotting FCH4 against VWC, which is the dominant environmental control identified does not reveal a meaningful relationship like Figure 5 a & c. You can look at an individual decision tree within the RF, but those are difficult to interpret beyond the first few splits, and each tree will be different. Lastly, RF are incapable of projecting beyond the parameter space observed which limited their applicability for this study (Fig A2). This presents an issue because may gaps in EC data arise from data filtering (e.g. clear calm nights, precipitation events) and are by definition outside the parameter space observed.



Line 296: Maybe refer to an equation defining alpha.

[Responses]

Alpha in this context is analogous to the minimum of the first derivative of the neural network output; which was calculated numerically. We added a new equation (Eq. 5) in section 2.5.1 [Line 258] to show a light response curve

$$NEE = \frac{1}{2c}\left(\alpha \mathrm{PPFD} + \beta - \sqrt{(\alpha \mathrm{PPFD} + \beta)^2 - 4\alpha\beta c \mathrm{PPFD}}\right) + \mathrm{ER}$$

and clarified section 3.3 to better describe this [Line 332-334].

The minimum values represent the peak light use efficiency and are analogous to α in eq. 5 (Fig 4b).

Technical corrections

Line 75: Did you really mean 100 m, or maybe km?

[Responses]

Yes, the antient basin, is just 100m to the south, it can be seen in Figure 1a (labeled 6b) and 1c in the top left of the drone image.

Lines 302/303: Pa, with a capital P Line 310: "both"?

[Responses]

Corrected, we also decided to use kPa instead

Please check and correct the names of your references in the text, as several have spelling mistakes ("Whalen and Reedburgh", "Merbould", "Meyer-Smith")

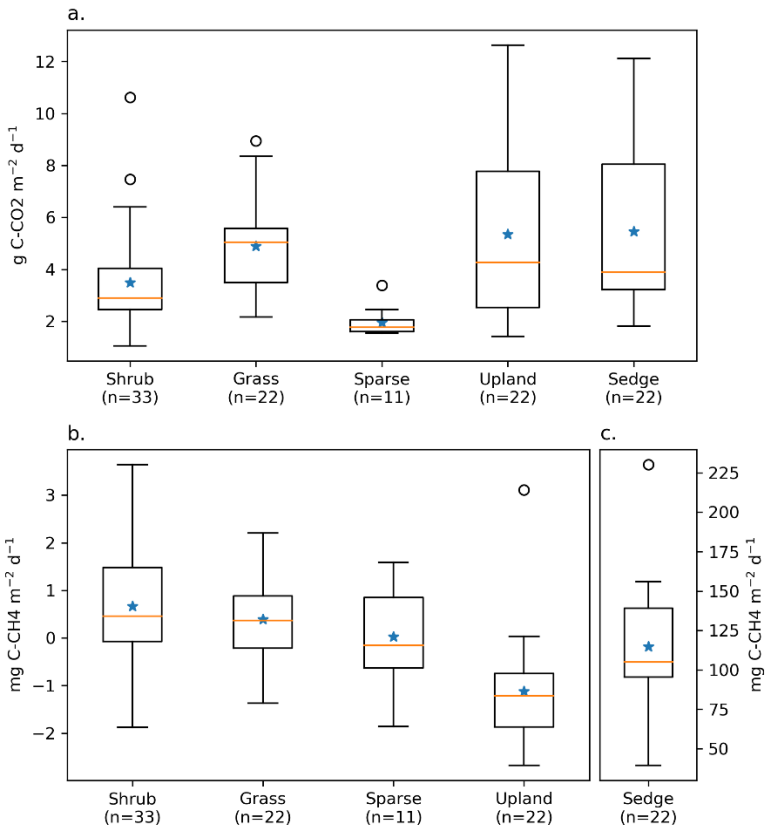
[Responses]

Thank you. We corrected these spelling mistakes.

Figure 3b: Can you add a little bit of horizontal white space between the the Sedge plot and the rest? I think this could prevent confusion and make it clear that the y-axis for this box has a different scale

[Responses]

We added the requested horizontal space and put “Sedge” into a separate subplot of the same figure.



Vegetation Influence and Environmental Controls on Greenhouse Gas Fluxes from a Drained Thermokarst Lake in the Western Canadian Arctic

June Skeeter[1], Andreas Christen[2], Andrée-Anne Laforce[3], Elyn Humphreys[3], Greg Henry[1]

[1]Department of Geography, The University of British Columbia, Vancouver, V6T1Z2, Canada
[2]Environmental Meteorology, Faculty of Environment and Natural Resources, Albert-Ludwigs Universität Freiburg, Freiburg, Germany
[3]Department of Geography and Environmental Studies, Carleton University, Ottawa, K1S5B6, Canada

Correspondence to: June B. Skeeter (skeeter1@mail.ubc.ca)

Abstract. Thermokarst features are widespread in ice-rich regions of the circumpolar Arctic. The rate of thermokarst lake formation and drainage is anticipated to accelerate as the climate warms. However, it is uncertain how these dynamic features impact the terrestrial Arctic carbon cycle. Methane (CH_4) and carbon dioxide (CO_2) fluxes were measured during peak growing season using eddy covariance and chambers at Illisarvik, a 0.16 km^2 thermokarst lake basin that was experimentally drained in 1978 on Richards Island, Northwest Territories, Canada. Vegetation in the basin ~~differes~~differs markedly from the surrounding dwarf-shrub tundra and included patches of tall shrubs, grasses and sedges with some bare ground and a small pond in the centre. During the ~~study period~~peak growing season, temperature and wind conditions were highly variable and soil water content decreased steadily. Basin ~~-~~scaled net ecosystem CO_2 exchange (NEE) measured by eddy covariance was -1.5 [~~-~~$CI_{95\%}$ ± 0.2] g C-CO_2 m^{-2} d^{-1}; NEE followed a marked diurnal pattern with no day-to-day trend during the study period. Variations in half-hourly NEE ~~was primary~~were primarily controlled by photosynthetic photon flux density and influenced by vapor pressure deficit, volumetric water content and the presence of shrubs~~. By contrast, net~~ within the flux tower footprint, which varied with wind direction. Net methane exchange (NME) was low (8.7 [$CI_{95\%}$ ± 0.4] mg CH_4 m^{-2} d^{-1}~~and~~[1] and had little impact on the growing season carbon balance of the basin~~ during the study period.~~. NME displayed high spatial variability~~,~~ and sedge areas in the basin were the strongest source of CH_4 while upland areas outside the basin were a net sink. Soil moisture and temperature were the main environmental factors influencing NME~~, having a positive and negative effect respectively.~~. Presently, Illisarvik is a carbon sink during the peak growing season. However, these results suggest that rates of growing season CO2 and CH4 exchange rates may change as the basin's vegetation community continues to evolve.

Keywords: Climate Change, Arctic, Permafrost, Thermokarst, Carbon Dioxide, Methane

1 Introduction

The northern permafrost region stores approximately 50% of global organic soil carbon in 16% of the terrestrial land area (Tarnocai et al. 2009). Thermokarst ~~lakes~~landscapes account for approximately 20% of the land area in this

region and hold about half of its organic soil carbon (Olefeldt et al., 2016). Lake thermokarst landscapes are ~~a~~ widespread ~~feature~~ in poorly drained, sedimentary permafrost lowlands with excess ground ice volume and constitute about a third of all thermokarst area (French, 2013~~). These~~; Olefeldt et al., 2016). Thermokarst lakes drain, sometimes catastrophically, via bank overflow, ice wedge erosion, coastal erosion, and stream migration (Billings and ~~Petersons~~Peterson, 1980; Mackay, 1999). Thermokarst lakes and drained thermokarst lake basins (DTLB) are prominent landscape features of the ~~western~~Western Canadian Arctic (Mackay, 1999; Marsh et al, 2009; Lantz & Turner, 2015). Lake formation and drainage is a natural part of the thaw lake cycle, but it is anticipated that climate change will accelerate or disturb this cycle ~~resulting in more lake formation and drainage~~, potentially altering the regional carbon balance (Jones et al., 2018).

Net ecosystem exchange (NEE), ecosystem respiration (ER) and gross primary productivity (GPP), where $NEE = ER - GPP$ are lower in the Arctic than warmer regions but have significant seasonal cycles and variability between vegetation types (Virkkala et al., 2018). Future trajectories in NEE will in large part be governed by ER (Biasi et al., 2008; Cahoon et al.,2012). Dominant vegetation types in the Western Canadian Arctic are erect-shrub tundra and wetlands (Walker et al., 2005). Growing season NEE is typically negative across these units throughout the Arctic indicating a net CO2 sink as GPP exceeds ER in part due to cold and/or anoxic soil conditions (Virkkala et al., 2018; Lafleur et al., 2012). Annual NEE can be positive or negative with large variation in GPP linked to annual weather variability (Virkkala et al., 2018, McGuire et al., 2009). Arctic net methane exchange (NME) is positive because wetland areas are strong methane (CH_4) sources while upland areas with better drainage can be net sinks (Whalen and Reeburgh, 1990; McGuire et al., 2009; Sturtevant and Oechel, 2013).

Thermokarst lakes are well recognized sources of ~~methane (~~CH_4 (Walter et al., 2007) which is 28 times as potent as carbon dioxide (CO_2) on a 100-year time scale (~~Walter et al., 2007~~IPCC, 2014). Thermokarst lake formation and expansion is expected to exert a positive feedback on climate change and accelerate Arctic warming in the near term, but ~~one~~ modelling ~~study~~ suggests that drainage may limit expansion and result in decreased lake area by the end of the century (van Huissteden et al., 2011). Post drainage, DTLB undergo rapid ecological succession. In colder tundra environments, wet meadows or polygonal landscapes dominated by sedges, grasses and rushes will form (Lara et al., 2015). ~~-~~In slightly warmer, boreal ~~or~~and transitional regions, DTLB often become dominated by willows and other shrubs (Lantz and Turner, 2015).

Carbon exchange in DTLB of various ages has been examined ~~by~~in a few studies, ~~all of which~~almost exclusively focused on the Barrow Peninsula in Northern Alaska. ~~In general,~~ DTLB ~~are net sinks for CO_2~~NEE during the growing season is negative with greatest CO_2 uptake in younger basins and decreasing net uptake ~~of CO_2 as the basin ages~~as basins age in this region (Zona et al., 2010; Zulueta et al., 2011; Sturtevant and Oechel, 2013; Lara et al., 2015). DTLB source/sink strength of CH_4 was found to be highly variable depending on vegetation and ground conditions (Lara et al., 2015). ~~Net methane exchange (~~NME~~)~~ is highest in wet ~~or flooded~~ meadows and remnant ponds ~~while upland tundra surrounding DTLB can be a methane sink~~but considerably reduced in areas with better drainage (Zona et al., 2009; Zona et al., 2012; Lara et al., 2015). There may be regional variations in the carbon balance of DTLB. For example, a shrub dominated ancient DTLB known as Katyk in the Indigirka lowlands of Siberia shows considerably higher growing season carbon uptake than young Alaskan DTLB with comparable NME (van der Molen

et al., 2007; Parmentier et al. 2011).~~; Sturtevant~~ Similarly, DTLB in the Western Canadian Arctic may have different carbon fluxes due to differences in climate and ~~Oechel, 2013).~~vegetation composition.

In this study, fluxes of CO_2 and CH_4 were measured at Illisarvik, an experimentally drained thermokarst lake ~~bed~~basin on ~~Richard's~~Richards Island in the ~~western~~Western Canadian Arctic, Northwest Territories, Canada. Fluxes of CO_2 and CH_4 were measured during the peak growing season using a combination of closed chamber and eddy covariance (EC) measurements. ~~Net ecosystem exchange (~~NEE~~) of CO_2~~ was calculated from fluxes and storage change and ~~NEE~~ was separated into ~~ecosystem respiration (~~ER~~)~~ and ~~gross primary productivity (~~GPP~~), $NEE = ER + GPP$~~. Here we report on: 1) the spatial and temporal variability of the NEE and NME, 2) the vegetation and environmental factors influencing NEE and NME, 3) how the growing season carbon balance at Illisarvik compares to other DTLB, and 4) potential future carbon balance trajectories as ~~Illisarvik continues~~Illisarvik's vegetation communities continue to evolve.

2 Methods

2.1 Study Site and Data Collection

The study took place at Illisarvik, a DTLB on Richards Island (69˚28'47.5" N, 134˚35'18.7" W), that was drained experimentally in 1978 (Mackay, 1981). Illisarvik has since served as the focus of studies on permafrost growth, active layer development and vegetation succession (Ovenden, 1986; Mackay and Burn, 2002; O'Neil et al., 2012; Wilson et al., 2019). At the nearby Tuktoyaktuk climate station mean annual air temperature (T_a) is -10.1 °C, July is the warmest month with a mean of 11°C and January is the coldest at -27°C. Mean annual precipitation is 160.7 mm yr^{-1}, the majority falling as rain in the summer and autumn. Snow cover typically lasts from mid-September or early October to late May (Environment Canada. 2016). Tuktoyaktuk is 60 km east of Illisarvik and in similar proximity to the coast so the climatology is expected to be similar at Illisarvik.

In the 39 years since drainage, Illisarvik has undergone rapid vegetation succession. After drainage, there were two remnant ponds. In the first five years after drainage, vegetation colonized the basin margins and wetter areas (Ovenden, 1986). By 1999, low vegetation had proliferated across most of the basin and taller willows had become established along the basin margins (Mackay and Burn; 2002). By 2010, some of the willows had grown to be 3 m in height (O'Neil and Burn; 2012). Current vegetation at Illisarvik is diverse relative to the dwarf-shrub tundra of the surrounding uplands~~;~~ (Table 1); the basin hosts a mix of woody shrubs (*Salix* spp., *Betula* spp., & *Alnus* spp), wetland vegetation (*Carex* ~~spp.,~~*aquatilis*, *Arctophila fulva*, etc.), and various grasses (*Pocacea* spp~~.).~~) (Wilson et al. 2019). The basin is partly ringed by a terrace of peat that formed after a partial drainage event ~ ~~2000~~5000 years BP and supports vegetation similar to the uplands (~~C Burn, personal communication 2016~~Michel et al., 1989). An ancient DTLB is located 100 m to the south of the Illisarvik basin and the Arctic Ocean is to the west of the basin, separated by a ridge of upland tundra about 50 m wide at its narrowest (Fig 1).

A vegetation survey of species ~~presence~~composition and ~~percent cover~~abundance was done on a 50 m grid in and around the basin during the 2016 study period (Wilson~~,~~ et al., 2019). A vegetation map was created with ten units based on plant functional type and vegetation structure, with sub-units denoting sub-canopy vegetation. The unit

boundaries between grid points were estimated visually by traversing the grid lines. Additional survey data on vegetation units ~~&~~and canopy height were collected manually with a GPS in the proximity of the EC station because greater resolution was needed for footprint modelling. ~~Drone~~Aerial imagery was collected on July 23rd over two flights using a Phantom 2 drone (DJI, Shenzhen, China). The GPS points and drone imagery were used to cross reference and modify the map of Wilson et al. (2019). The ten units were then aggregated into ~~6~~six broader surface cover classes (listed from largest to smallest areal fraction within the footprint climatology (F_{Clim}) see Section 2.3 for definition): shrub, grass, sedge, upland, sparse, and water classes (Fig 1 & Table 1).

2.2 Weather and Soil Measurements

Weather data were logged on a CR1000 datalogger (Campbell Scientific Inc, Logan, UT, USA; CSI) at 5-minute intervals. ~~A~~ Net all-wave radiation (R_n) and photosynthetic photon flux density ($PPFD$) were measured with a NRLite net radiometer (Kipp & Zonen, Delft, Netherlands) ~~measured net all-wave radiation (R_n),~~and a SQ-110 quantum sensor (Apogee Instruments, Logan, UT, USA~~) measured photosynthetic photon flux density ($PPFD$), and a~~), respectively 3.2 m above the grass surface on the main EC system tripod (Fig 1). A shielded HMP35 (CSI) ~~measured~~recorded T_a and relative humidity (RH~~),~~) 2 m above the surface. A tipping bucket rain gauge (R.M Young Company, Travers City, MI, USA) was placed 3 m to the west of the main tripod. Soil temperature and moisture ~~data~~ were ~~recorded at 30-minute intervals on CR10x dataloggers (CSI)~~measured within ~~two~~ soil pits ~~with~~in two different vegetation types near the tripod: Grass (30 m to the east) and Shrub (40 m to the north). ~~Each system measured~~Measurements were made of ground heat flux (G) with custom ~~made~~ heat flux plates, soil temperatures (T_s) with custom ~~made~~ type-T thermocouples at depths of 0.08 m, and 0-20 cm integrated? volumetric water content (~~θ_w~~)VWC) with CS616 water content reflectometers (CSI). ~~There was one repetition of each observation per pit.~~The soil measurements were recorded at 30-minute intervals on CR10x dataloggers (CSI). The climate and soil stations operated uninterrupted from July 10th (day 192) and July 11th (day 193~~)~~), respectively, until August 7th, 2016 (day 220). On July 11th and August 6th thaw depth was measured at each of the 10 chamber sites~~.~~ (see below). Thaw depth was measured by inserting a graduated steel probe into the ground to point of refusal. Each site was probed five times: the median value has been used as the thaw depth at each location. ~~Between day 193~~On July 12th and ~~197~~15th, a large herd of reindeer (500 + animals) visited Illisarvik. They mostly avoided the tripod but did graze near it ~~on a few occasions.~~for about an hour on July 12th which may have affected greenhouse gas fluxes.

2.3 EC Fluxes

An EC system was placed in the southwestern portion of the basin (69° 28' 47.82", -134° 35' 18.6") and measured fluxes of CO_2 (F_{CO2}) and CH_4 (F_{CH4}) for the full study period~~.~~ between July 10th and Aug 7th, 2016. The EC system consisted of an open-path infrared CO_2/H_2O gas analyser (IRGA) (model LI-7500, LI-COR Inc., Lincoln, NK, USA; LI-COR), an open-path CH_4 analyser (model LI-7700, LI-COR) and a CSAT3 sonic anemometer (CSI) mounted on a tripod at a measurement height (z_m) of 3 m (Fig 2~~.~~).). The EC data and air pressure (P_a) were logged at 10 Hz on the LI-7550 Analyzer Interface Unit (LI-COR). The CSAT3 was oriented to the northeast (40°) because climatology for Tuktoyaktuk indicated northerly and easterly winds are typical for July and August (Environment Canada, 2016).

Half-hourly fluxes were calculated with EddyPro V.6.2.0 (LI-COR). The software performed statistical assessments (Vickers and Mart, 1997), low and high frequency spectral corrections (Moncrieff et al., 1997 and 2004), a double rotation (Wilczak et al., 2001), applied the WPL correction to account for density fluctuations (Webb et al., 1980), and computed quality control (qc) flags (Mauder and Foken, 2004). Post processing treatments included: storage correction (calculating the net flux as the sum of the observed scalar flux and the rate of change in scalar concentration at z_m), filtering fluxes by friction velocities (u_*) below 0.1 m s^{-1}, ~~and~~ removing ~~spurious half hourly measurements~~qc flags = 2 (Mauder and Foken, 2004), and the mean absolute deviation spike removal algorithm (Papale et al., 2006). Additionally, observations with mean winds from 220˚ ± 30˚ were removed ~~as these~~ to avoid uncertainties associated with the wake of the sonic anemometer, and observations were removed during precipitation events and when the open-path ~~analyzers~~analysers indicated there were any other obstructions within the path~~.~~ (Aubinet et al., 2012). The data were gap-filled using neural networks (NN) which have been applied to F_{CO2} and F_{CH4} in other studies (Moffat et al., 2010; Dengel et al., 2013). Details of the NN methodology are ~~discussed~~described in Appendix A.

The flux footprint represents the influence of upwind areas on a measured scalar flux and the footprint climatology is the average of individual footprints over a time period. Evaluation of the flux footprints and climatology help evaluate the reliability of the dataset and estimate the source area of each individual ~~datapoint of the~~half-hourly EC flux measurement. A scalar flux F_c sampled at $(0,0,z_m)$, where z_m is the height of the EC instrumentation, can be represented as the integral of the flux footprint function *f(x,y)* and the distribution of sources/sinks (Q_c) over a domain *D* (Kljun et al., 2015):

$$F_c(0,0,z_m) = \int_D Q_c(x,y) f(x,y) \quad (1)$$

The flux contribution of upwind source areas increases sharply upwind from the measurement location to a peak then decrease gradually with increasing distance (Schmid, 2002). The empirically derived flux footprint function of ~~Klujn~~Kljun et al~~.,~~. (2015) was used to estimate the source area of each half hourly flux measurement.

The model requires boundary layer heights which were not measured onsite. Half hourly boundary layer heights were interpolated from three-hour estimates obtained from the Global Data Assimilation System of the U.S. National Oceanic and Atmospheric Administration. The model also requires the aerodynamic roughness length (z_0) which is influenced by the canopy height and spacing. Canopy height (C_h) varied considerably within the basin (from >1 m in the north to ~0 m in the bare ground areas). Canopy height variability was lower in the vicinity of the EC tripod but ranged from 0.35- 0.55 m with a few taller shrubs approaching 1 m. Median z_0 was calculated for 30˚ wind sectors following Paul-Limoges et al. (2013). This calculation was performed for near neutral conditions $-0.05 \leq \frac{Zm}{L} \leq 0.05$, where *L* is the Obukhov length. The z_0 for each wind sector was found to be insensitive to ~~changes in the~~ zero-plane displacement height, *d*, as $z_m >> d$, so the mean value of *d* around the tripod was used, where $d = 2/3\ C_h$. Zero-plane displacement did not change significantly over the course of the study so z_0 remained fixed over the study period for each wind sector.

For each ~~30 minute~~half-hourly flux observation, $f(x,y)_i$ was solved at ~~one meter~~1 m^2 resolution over a 1 km^2 domain centred on the EC tripod. Then, $f(x,y)_i$ were intersected with the surface classes to determine ~~their~~the relative contribution of each surface type to each flux observation (referred to as ~~Shrub%, Sedge%,~~F_{Shrub}, F_{Sedge}, etc.). The footprint function is technically infinite so a fraction of each $f(x,y)_i$ was not contained within the model domain. The out-of-domain source fraction ranged from 1.8% ~~-~~to 4.9% with a mean of 3.2% and was assumed to have minimal impact on the analysis. The flux footprint climatology (F_{Clim}) was calculated by averaging the half hourly flux footprints ~~were then averaged~~ over the study ~~to calculate the flux footprint climatology. All post processing~~period and ~~footprint modelling were carried out using Python~~is shown in Figure 1. Table 2 shows the flux contribution of each vegetation class.

2.~~3~~4 Closed Chamber Measurements

In addition to EC measurements, fluxes of CO_2~~, CH_4,~~ and ~~nitrous oxide (N_2O)~~CH_4 were sampled using a static non-steady state chamber flux technique on 11 dates between July 12 and August 5, 2016 (Laforce, 2018). ~~Only the bare ground was a significant source of N_2O, so those results are not shown here. Chamber~~Nineteen chamber collars were located at ten sites, eight sites within and two outside the basin (~~Figure~~Fig 1). Each surface cover class was represented by at least one chamber site, except for open water. At each vegetated site a pair of collars were installed 20 cm apart ~~and the~~, except at the 'sparse' site where only one collar was installed. The above ground biomass was removed from one of the collars~~, with~~ at each vegetated site. There were three replicates (six collars) for the ~~exception of the 'sparse' cover~~Shrub class ~~where only one collar was installed.~~ , two for the Sedge, Grass, and Upland tundra, and no replicates for the Sparse class. PVC collars 30 cm long and 24.3 cm in diameter were inserted to a depth of approximately 15 cm. The chambers were 34 cm tall and made out of polycarbonate covered in black opaque tape to maintain dark conditions inside the chamber (for more details, see Martin et al., 2018). The chambers contained a small vent (10 cm coiled 1/8" diameter copper pipe) to ensure a constant pressure during measurements. The ~~use of~~ opaque chambers ~~means that F_{CO2} represents only ER. The~~fluxes of CO_2 provided an independent estimation of ER. This helped characterize ER ~~is important as it is difficult to estimate ER using~~given the challenges with standard ~~EC~~NEE partitioning techniques at high latitude sites during the Arctic summer ~~at this latitude. Standard approaches require night time data, a condition which was only found during a small fraction of the study~~ ~~period~~as noted below. ~~Measurements~~Chamber flux measurements were made between 9:00 and 17:00 starting at a different ~~site~~collar set each day to randomize the sampling order to avoid a bias due to diurnal ~~ranges on the 11 days when performing chamber measurements~~patterns. During gas flux measurements, the chambers were sealed to the top of the collars within a groove filled with water and five 24 mL air samples were collected into evacuated 12 mL vials sealed with doubled septa. Each vial contained a small amount of magnesium perchlorate to dry the air sample. Samples were collected at 0, 5, 10, 15 and 20 minutes after the chambers were set on the collars. Air within the chamber was mixed with a 60 mL syringe attached to a three-way stopcock before each air sample was taken. Samples were stored until analysis ~~the following fall~~one month later at Carleton University. To monitor the integrity of the vials through shipping, storage and analysis, a number of the evacuated vials were filled with helium in the lab before the field season began.

Concentrations of CO_2, CH_4 and N_2O were determined ~~at Carleton University,~~ using a CP 3800 gas chromatograph (Varian Inc., Pao Alto, CA, USA) as described by Wilson and Humphreys (2010). Three replicates of five CO_2/ CH_4 standards varying from 383.1 to 15212.6 ppm CO_2 and from 1.08 to 22.11 ppm CH_4 were included in every set of measurements to create a linear relationship between gas concentration and chromatogram area. The chamber fluxes of CO_2 and CH_4 (F_C) were calculated as follows:

$$F_C = \frac{VP}{ART}\frac{dc}{dt} \quad (2)$$

where (dc/dt) is the linear rate of change in the mixing ratio of the gas, A is the chamber area (0.0464 m^3), V is the chamber volume (between 0.0182 and 0.0242 m^3 adjusted for ~~m^2~~ collar depth at each collar location), R is the ideal gas constant, P is pressure in Pa and T is the air temperature in Kelvin. P and T values corresponding to the time of each measurement were obtained from the EC station. After removal of spurious point measurements (72 vial samples were rejected out of 1135 vials), dc/dt was determined using three or more gas sample concentrations resulting in coefficients of determination that ranged from 0.71 to 0.99. No flux measurements were removed from the analysis. Positive fluxes indicate emissions of gases to the atmosphere and negative fluxes indicate uptake by the surface. ~~After removal of spurious point measurements (if more than 2 points were rejected, the flux measurement was rejected), dc/dt coefficients of determination ranged from 0.71 to 0.999. Overall, 72 vial samples were rejected out of 1135 vials and no flux measurements were rejected out of 681.~~

2.4.1 Upscaling

Chamber ~~measurements~~fluxes of ER were upscaled from the plot scale (individual chamber) to the footprint scale using the footprint weighted average method and to the basin scale using the area weighted average method (Budishchev et al., 2014). The chamber ER and air temperature from the EC tripod (T_a) were used to determine R_{10}, the base respiration at 10 C°, and Q_{10}, the temperature sensitivity coefficient, using eq 3 for five of the six surface classes (Fig 1) (Laforce, 2018) (Table 3).

$$ER = R_{10} Q_{10}^{\frac{(T_a - 10)}{10}} \quad (3)$$

~~(comparison to EC observations. ER was~~

Half hourly footprint scale estimates (ER_{FS}) were calculated ~~using eq. 4 and the Q_{10} and R_{10} coefficients~~by multiplying ER derived ~~by Laforce (2018) explained in section 2.4.1. ER was calculated~~from eq. 3 for each surface class ~~separately and then weighted~~ by ~~footprint climatology (see~~ the footprint source area fraction and summing over classes. Basin scale estimates (ER_{BS}) were estimated the same way but using the mean source area fractions of the basin (Table 2). ~~For NME~~As there were no open water class ER estimates, ER from open water was assumed to be zero.

In contrast to ER, there are no standard empirical functions~~, so mean NME~~ to estimate temporal variations in NME. Instead, we used ordinary least squares regression (OLS) to estimate NME. The most important environmental controls over F_{CH4} were VWC and T_s (discussed below). Continuous observations of these factors at the flux chambers were not available, instead chamber NME were grouped by vegetation class and fit to VWC and T_s measured in the soil pits near the EC station. Half hourly footprint scale (NME_{FS}) and basin scale (NME_{BS}) estimates were then estimated using the OLS parameters for each ~~class was weighted by the footprint climatology instead. The weighted~~

~~estimates were compared to EC observations and to help verify our findings.~~surface class using the same procedures for ER_{FS} and ER_{BS}.

2.~~4~~5 Factor Selection and Gap Filling

We used an exploratory approach to identify the smallest set of factors that best predicted half hourly EC-derived NEE and NME without overfitting the dataset~~.~~ using a series of neural networks (NN). We started with 10 factors: four meteorological variables [(*PPFD*), T_a, vapor pressure deficit (*VPD*~~),~~) computed using the T_a and RH data, three-dimensional wind speed (*U*~~)],~~) measured using the CSAT3 sonic anemometer], two soil variables ~~[volumetric water content ([~~(*VWC*) and T_s averaged between the two soil ~~temperature (T_s)],~~pits near the EC tripod], and four source area fractions [~~shrub~~Shrub (F_{Shrub}), ~~grass~~Grass (F_{Shrub}), ~~sedge~~Sedge (F_{Sedge}), and ~~upland~~Upland, (F_{Upland})]. The four source area variables correspond to surface classes sampled by the ~~chamber samples.~~chambers. We excluded ~~sparse~~Water (F_{Water}) and Sparse (F_{Sparse}) fractions because its average contribution to the EC observations was only 0.2~~.1% and~~% and 2.2%, respectively, and there were no chamber measurements for the Water class while chamber ~~samples~~measurements indicated ER was low and NME was not significantly different from zero~~.~~ for the Sparse class. A number of these ~~variables are~~prediction factors were highly correlated but it was necessary to include them so the model could account for source area heterogeneity.

~~A series of neural networks (NN)~~The NNs were trained iteratively on bootstrapped datasets. First NN were trained on each factor individually and the one with the lowest MSE was selected. Next, NN were trained on that factor in combination with one of the remaining nine. The best performing additional factor was again selected and this process was repeated until MSE failed to improve. The most parsimonious model was identified using the one standard error (*SE*) rule. Dybowski and Roberts (2001) give the standard error of a bootstrap estimate of a given error metric (~~eg.~~e.g., $\theta = MSE$) to be

$$SE_{boot}(\theta) = \sqrt{\frac{1}{B-1}\sum_{b=1}^{B}(\theta_b - \theta_{boot})^2} \quad (~~34~~)$$

where θ_{boot} is the mean of the bootstrapped samples. The smallest set of factors where θ_{boot} was within one SE_{boot} of the minimum θ_{boot} for both NEE and NME were selected for further analysis. The outputs from the selected models ~~trained on those factor sets~~ are referred to as ~~NN_{NEE}~~NEE_{NN} and ~~NN_{NME}~~NME_{NN}, respectively. NN modelling was done using the Keras Python library (Chollet et al., 2015), see the Appendix A for a more detailed explanation of the NN analysis.

Multiple Imputation (MI) was then used to gap fill the NEE and NME with the ~~outputs from NN_{NEE}~~NEE_{NN} and ~~NN_{NME}~~NME_{NN}, respectively (Vitale et al., 2018). Of the 1296 half hourly flux observations 28.9% of F_{CO2} and 31.3% of F_{CH4} were missing or filtered out. There were a few gaps in the source area fractions needed to gap-fill the flux time series because the footprint function is not valid when $u_* < 0.1$ m s^{-1}. When source area fractions were missing, they were gap-filled by using the mean source are fraction observed for winds within ± 5˚of the observed wind direction. The meteorological and soil data were continuous and did not need to be gap-filled.

2.45.1 Flux Partitioning

NEE is negative when there is net uptake of CO_2 by the ecosystem and positive when there is net emission. ER and GPP are always positive, ER represents the sum of heterotrophic and autotrophic respiration and GPP represents photosynthetic uptake of CO_2. Night-time NEE observations (~~eg.~~e.g., *PPFD* <= 10 µmol m^{-2} s^{-1}) are typically used to quantify ER because GPP ~ 0 (Aubinet et al., ~~2012). Some NN analyses of NEE have trained separate models for night-time and daytime conditions (Papale & Valenini, 2003). However, these methods are not practical during the Arcitc summer, the sun did not set at Illisarvik until July 28th, over half way through the study period. There were not enough night-time samples (n=100) to be worth training a separate NN. Instead, we estimated ER by calculating the intercept of NN_{NEE} at PPFD = 0 µmol m^{-2} s^{-1} for all observations. This estimate is referred to as NN_{ER}. Laforce (2018) fit a Q_{10} equation to the chamber ER observations:~~ 2012). We fit the limited night-time EC observations available (n=95) to equation 3 for comparison with the ER measured using the chambers. We used the fitted values to model daytime ER and approximate NEE by fitting the daytime data to a light response curve Aubinet et al. (2012).

$$\textit{ER}NEE = \frac{1}{2c}\left(\alpha\text{PPFD} + \beta - \sqrt{(\alpha\text{PPFD} + \beta)^2 - 4\alpha\beta c\text{PPFD}}\right) + \text{ER} \quad (5)$$

~~$= R_{10} Q_{10}^{\frac{(T_s - 10)}{10}}$ (4)~~

where ~~R_{10}~~α is the ~~base respiration~~initial slope of the light response curve, β is GPP at ~~10 C°~~saturation, and ~~Q_{10} is the temperature sensitivity coefficient. We fit the~~ c is a curvature parameter. These estimates are referred to as ER_{Q10} and NEE_{Q10}.

Some NN analyses of NEE have trained separate models for night-time ~~EC~~and daytime conditions for partitioning purposes (Papale & Valentini, 2003). However, these methods are not practical during the Arctic summer as the sun did not set at Illisarvik until July 28th, over halfway through the study period. There were not enough night-time samples to train a separate NN. Instead, we estimated ER by calculating NEE_{NN} at PPFD = 0 µmol m^{-2} s^{-1} for all observations ~~available (n=100) to this Q_{10} equation for comparison (Aubinet et al., 2012).~~, henceforth referred to as ER_{NN}. This is a projection outside of the observed parameter space resulting in greater uncertainty and a wider confidence interval around ER_{NN} than NEE_{NN}. Calculation of confidence intervals for NN outputs is discussed in Appendix A.

2.~~4.1~~5.2 Factor Analysis ~~and Upscaling~~

The trained NNs were used to investigate how individual factors influenced NEE and NME. The partial first derivative of the model response to one controlling factor was calculated while keeping all other inputs fixed. For example, the partial first derivative, $\frac{\partial NEE}{\partial \text{PPFD}}\frac{\partial \text{NEE}}{\partial \text{PPFD}}$, is an approximation of the NEE light response curve under a specific set of conditions. Similarly, NN_{NME} can be used to approximate NME response to controls like *VWC* or T_s. For both fluxes, the selected ~~model~~models contained at least one source area fraction variable, indicating the vegetation type(s) which had significant influence over ~~both~~ NEE and NME. ~~We~~Additionally, we mapped NN_{NEE} and NN_{NME} to ~~full~~100% coverage for ~~the~~individual surface classes to ~~approximate their~~see how fluxes at Illisarvik may change as vegetation succession continues. For example, to project to 100% Sedge coverage, we set the other surface classes to 0% and

~~compare with~~left the ~~chamber observations.~~other environmental factors unchanged. This allows for an estimation of how carbon fluxes may change if vegetation succession leads Illisarvik to look more like the DTLB studied in Alaska.

3 Results ~~& Discussion~~

During the 29-day study, half-hourly T_a and T_s ranged between 0.4 and 26.2 ~~C°~~°C and 4.4 and 11.0 ~~C°~~°C, respectively (Fig 2a). Day length and maximum solar altitude decreased from 24 hours to 19.25 hours and 41.6˚ to 35.4˚, but daily *PPFD* was more influenced by variations in cloud cover. Precipitation (19 mm) fell on 14 of the 28 days with trace snowfall on three of those days, but *VWC* of the soils decreased throughout the period (Fig 2b). At the onset of the study period, *VWC* was high and soils were saturated with ponding in the sedge areas. By the end of the study most of this surface water had dried up. On July 11th average thaw depth (cm) was 37, 45, 51, 64, 81 at ~~upland, sedge, grass, shrub~~Upland, Sedge, Grass, Shrub, and ~~sparse~~Sparse classes, respectively. By August 6th, average thaw depth ~~was~~had increased to 45, 62 and 66 cm at ~~upland, sedge~~Upland, Sedge and ~~grass~~Grass surface classes and over 100 cm at both the ~~shrub~~Shrub and ~~sparse~~Sparse classes.

A strong low-pressure system stalled off the coast between day of year (DOY) 199 and 204. This caused westerly winds to occur much more frequently than is typical for July and August. The 50%, 80% and 90% flux ~~footprint climatology~~F_{Clim} contours are shown in Fig 1a. Mean source area fractions indicate the EC observations were skewed towards the ~~grass~~Grass surface class and under-sampled the ~~shrub~~Shrub class, but the range of surface classes sampled was diverse enough to allow for testing of the impact of source area fraction on the fluxes (Table 2).

3.1 EC ~~& Chamber~~ Observations

~~EC~~Half hourly observations ~~indicate that~~of F_{CO2} and F_{CH4} along with the ~~flux footprint area was a carbon sink during~~NEE_{NN} and NME_{NN} used to gap-fill the ~~peak growing season, but fluxes varied considerably from day to day.~~time series are shown in (Fig 2c & d). Gap-filled daily NEE ranged from -3.7 to -0.2 g C-CO_2 m^{-2} d^{-1} with a mean -1.5 [$CI_{95\%}$ ± 0.2] g C-CO_2 m^{-2} d^{-1} ~~(Figure 2c). Estimated ER was~~. Day to day variability was considerable but there was no notable trend in NEE over the peak growing season. The half hourly NEE during the study period reached a minimum of -10.4 µmol CO_2 m^{-2} h^{-1} just before solar noon and peaked at 4.7 µmol CO_2 m^{-2} h^{-1} around midnight (Fig 2c). NEE_{NN} was used to gap-fill the flux data because it was in good agreement with F_{CO2} observation (r^2 = 0.91). Daily ER_{NN} was estimated to be 2.2 [$CI_{95\%}$ ± 0.9] g C-CO_2 m^{-2} d^{-1} with corresponding GPP of 3.7 g C-CO_2 m^{-2} d^{-1}. ~~There were no notable trends in NEE or ER over the study period. Our~~ER_{NN} was in poor agreement (R^2 = 0.35, n= 95) with night time F_{CO2} observations ~~are within ranges observed from young DTLB on the Barrow Peninsula. NEE was greater than (ie. less carbon uptake) EC observations of from four wetter, sedge dominated DTLB, where peak season NEE was -2.5 g C-CO_2 m^{-2} d^{-1}, ER (1.5 g C-CO_2 m^{-2} d^{-1}) was lower than at Illisarvik while GPP (4~~. For comparison, Eq. 3 provided a better fit (R^2 = 0.47) with night-time EC data, and ER_{Q10} was estimated to be 3.0 g C-CO_2 m^{-2} d^{-1}~~) was slightly higher (Zona et al., 2010). But NEE was less than (ie. more carbon uptake) upscaled chamber estimates for wet meadow DTLB (-0.9 g C-CO_2 m^{-2} d^{-1}), while ER (2.7 g C-CO_2 m^{-2} d^{-1}) and GPP (3.5 g C-CO_2 m^{-2} d^{-1}) were higher and lower than Illisarvik (Lara et al., 2015). Mid-day NEE at was comparable to aircraft observations from young DTLB (Zulueta et al., 2011).~~ However, NEE_{Q10} did not fit F_{CO2} as well (r^2 = 0.80) as NEE_{NN}.

Gap-filled daily NME was modest~~, ranging from~~ and decreased over the study period. It ranged from 2.0 to 25.1 mg C-CH_4 m^{-2} d^{-1} with a mean of 8.7 [$CI_{95\%}$ ± 0.4] mg C-CH_4 m^{-2} d^{-1} (Fig ~~3d). Even accounting for the greater GWP of CH_4,~~2d). NME_{NN} was used to gap-fill the flux data because it provided a reasonable fit (r^2 = 0.62) to F_{CH4} observations. NME did not constitute a significant component of the carbon balance~~. Daily NME decreased over the study period as soils in the basin dried. NME at Illisarvik was much less than EC observations from in a wet sedge dominated DTLB (18.4 mg C-CH_4 m^{-2} d^{-1}) and chamber observations from DTLB (26.1 mg C-CH_4 m^{-2} d^{-1}) on the Barrow Peninsula (Zona et al., 2009; Lara et al., 2015).~~ and thus the flux footprint area was a carbon sink during the peak growing season with negative GWP after accounting for the greater GWP of CH_4.

~~**Chamber ER was higher than EC estimates and showed modest variability among vegetation classes (Fig 3a). Greatest ER was observed**~~**3.2 Chamber Observations**

ER was highest in the Sedge, Upland, and Grass classes where fluxes were very similar at 5.5 [$CI_{95\%}$ ± 1.2], 5.4 [$CI_{95\%}$ ± 1.2] and 4.9 [$CI_{95\%}$ ± 0.7] g C-CO_2 m^{-2} d^{-1}~~,~~. Shrub ER was significantly less (3.5 [$CI_{95\%}$ ± 0.6] g C-CO_2 m^{-2} d^{-1}) than the other vegetated classes and Sparse ~~had~~ER was the lowest ~~ER~~among the classes (2.0 [$CI_{95\%}$ ± 0.3] g C-CO_2 m^{-2} d^{-1}~~). These values were generally higher than chamber observations from DTLB or tundra sampled by Lara et al. () (fig 3a). The2015).~~ Q_{10} and R_{10} values also differed between vegetation classes~~,~~: ER in the Sedge ~~were~~was the most sensitive to changes in air temperature~~, while~~ and modelled values provided the best fit (R^2 = 0.82) to observations. Upland and Grass had the highest base respiration and fit observations moderately well (Table 3)~~. EC estimated Q_{10} and R_{10} (r^2 = 0.47) were generally lower than the vegetated chambers, but most similar to Shrub, which was the largest component of the footprint.~~

~~Chamber~~ NME was ~~much smaller than ER but there was significantly~~ more ~~variability among~~ variable between vegetation classes than ER (Fig 3b ~~&~~ c). Sedge was a very strong CH_4 source at 114.7 [$CI_{95\%}$ ± 15.3] mg C-CH_4 m^{-2} d^{-1}~~, which is comparable to chamber observations from vegetated ponds on the Barrow Peninsula (Lara et al., 2015). Sites across the arctic with sedges have higher NME than sites without sedges (Olefeldt et al., 2013).~~. Shrub and Grass were very weak sources, 0.7 [$CI_{95\%}$ ± 0.3] and 0.4 [$CI_{95\%}$ ± 0.3] mg C-CH_4 m^{-2} d^{-1}, respectively~~ and~~. Sparse was neutral. Upland was a net CH_4 sink ~~for CH_4~~ -1.1 [$CI_{95\%}$ ± 0.4] mg C-CH_4 m^{-2} d^{-1}. ~~Lara et al. (2015) did not find upland areas on the Barrow Peninsula to be CH_4 sinks, but upland tundra is known to be a globally significant methane sink (Whalen and Reedburgh, 1990; Whalen et al., 1996) Upland fluxes were~~ Sedge and Shrub were NME were positively correlated with T_s (r=0.61, p < 0.01; r=0.35, p = 0.04) respectively and *VWC* (r=0.58, p < 0.01; r=0.5, < 0.01) respectively. They also had a positive correlation with T_a, while Upland NME was negatively correlated with ~~T_s,~~ T_a. Grass and Sparse ~~and Grass had no clear relationships, while Shrub and Sedge had~~didn't have any significant ~~positive~~ correlations~~ with *VWC* and T_s, but there were no straightforward empirical functions to model NME like there is for ER.~~.

~~Footprint scaled chamber estimates of ER and NME were about 32% and 31% greater than the EC estimates. Mean estimated ER estimated was 3.2 g C-CO_2 m^{-2} d^{-1} and NME was 12.9 [$CI_{95\%}$ ± 8.1] mg C-CH_4 m^{-2} d^{-1}. The discrepancies between the EC and chamber observations is often observed and has~~Footprint scaled chamber fluxes were 59% and 47% higher than ER_{NN} or gap-filled NME, respectively. Mean ER_{FS} was 3.5 g C-CO_2 m^{-2} d^{-1} [$CI_{95\%}$ ± 0.1], it fit ER_{Q10} very well (R^2 = 0.95) as would be expected and ER_{NN} moderately well (R^2 = 0.46). Mean NME_{FS} was 12.8[$CI_{95\%}$ ±

1.3] mg C-CH_4 m^{-2} d^{-1}, it did not fit NN_{NME} well (R^2 = 0.30). At the basin scale, ER_{BS} (3.4 [$CI_{95\%}$ ± 0.1] g C-CO_2 m^{-2} d^{-1}) was slightly lower than ER_{FS} because of the exclusion of upland areas. NME_{BS} was higher (15.2 [$CI_{95\%}$ ± 0.1] g C-CO_2 m^{-2} d^{-1}) because of the greater sedge fraction in the basin than the footprint because the (Table 2).

3.3 been attributed to differences in measurement techniques, the small sample size of chamber observations, and sampling bias since all chamber measurements were taken during the day with fair weather (Katayanagi et al., 2005; Chaichana et al., 2018). Meijide et al. (2011) found that chamber NEE could be up to twice as large as EC observations and Riederer et al. (2014) also found chamber NME estimates were about 30% higher than EC estimates.

3.2 NEE Response to Environmental Factors and Vegetation Type

$NN_{NEE}$$NEE_{NN}$ (r^2 = 0.91) had was estimated using four factors: *PPFD*, *VPD*, *VWC*, and *Shrub*. For comparison, NEE estimated using the Q_{10} paired with a light response curve (r^2 = 0.8) or a random forest regression tree (r^2 = 0.86) trained with the factors selected by NN_{NEE} were not as accurate as NN_{NEE}. F_{Shrub}. *PPFD* is the primary control over NEE;: a NN trained on *PPFD* alone provided a reasonable fit (r^2 = 0.83). The three additional factors:, *VPD*, *VWC*, and *Shrub*F_{Shrub}, helped NN_{NEE} fit a wider variety of conditions. Looking atExamining the partial first derivative of NN_{NEE} under different conditions, we can inspect provides interpretation of the modelled light response curves (αFig 4). The minimum of α isvalues represent the peak light use efficiency. and are analogous to α in eq. 5 (Fig 4b). With increasing *PPFD*, light use becomes less efficient and α approaches zeros as the light response nears light saturation. (Fig 4b).

VPD was a secondary control over NEE, which is consistent with the findings of another study using NN to analyse NEE (Moffat et al., 2012). Increasing *VPD* increased peak light use efficiency and net CCO_2 uptake until a threshold, above which it had a strong limiting effect (figFig 4a & b). For example, under dry atmospheric conditions (ege.g. VPD = 1500 Pa1.5 kPa), peak light use is less efficient (-12 nmol CO_2 μmol^{-1} photon) than under optimalmore humid conditions (-18 nmol CO_2 μmol^{-1} photon). The value of this VPD threshold was dependent upon soil moisture: from 1000 pa1 kPa when *VWC* was highest to 200 pa0.2 5Pa when *VWC* was low.

Mapping NN_{NEE} and NN_{ER} at *Shrub*%F_{Shrub} = 100%, *Shrub*%F_{Shrub} = 0%, and *Shrub*%F_{Shrub} = 36% (footprint climatologyF_{Clim}), shows that *VWC* and Shrub%F_{Shrub} were the primary controls over ER and thus influenced NEE (figFig 4c & d). We can see from the partial first derivates of NN_{ER} that increasing *VWC* increases ER from shrubShrub areas. In the absence of shrubs, increasing *VWC* inhibits ER. although it is important to note that variations in VWC were subtle ranging from 51.7% to 59.0%. The partial first derivative of NN_{NEE} shows that *VWC* slightly limits NEE from non-Shrub areas and significantly reduces it in Shrub areas. The *VWC* relationships support Zona et al. (2010) who found *VWC* explained 70% of the variability in daily peak season NEE and ER and Kittler et al. (2016) who found drier soils increased both NEE after a wet tundra drainage experiment in Siberia. The chamber data supports the inclusion of Shrub% because Shrub ER was significantly lower than the other vegetation classes.

3.34 NME Response to Environmental Factors and Vegetation Type

NN_{NME}NME was estimated using NME_{NN} (r^2 = 0.62) which had five factors: *Sedge*%, *Shrub*%F_{Sedge}, F_{Shrub} , *VWC*, T_S, *VPD*, and *U*. A random forest regression tree (r^2 = 0.51) and a GLM (r^2 = 0.25) trained on the factors selected by

~~NN_{NME} were not as accurate as NN_{NME}.~~ NME was more ~~chaotic~~variable and less dependent on any one factor than NEE which is why the NN_{NME} needed an extra factor and had a lower r^2 score. Source area had a significant effect on NME, and it ~~is~~was encouraging that the model ~~contains *Sedge*%~~contained F_{Sedge} and ~~*Shrub*%~~F_{Shrub} since Sedge and Shrub were the strongest CH_4 source and largest footprint component, respectively. These two factors can combine to map NME under three general situations: we can extrapolate to ~~Sedge%~~F_{Sedge} = 100 % ~~& Shrub%~~and F_{Shrub} = 0 % or ~~Sedge%~~F_{Sedge} = 0 % ~~& Shrub%~~and F_{Shrub} 100 %, or ~~to footprint climatology~~represent actual F_{Clim} where F_{Sedge} = 11% and F_{Shrub} = 37% (Table 2). ~~The footprint climatology means that some~~Some upland tundra ~~is also~~was included in the F_{Clim} estimate, which ~~brings~~reduced NME ~~down~~.

VWC was the primary climatic driver identified by NN_{NME}. Wetter soils had a consistent positive effect on NME which was strongest when ~~*Sedge*%~~F_{Sedge} was high (Fig 5a & b). Between driest and wettest conditions, estimated NME increased: by an order of magnitude at ~~*Sedge*%~~F_{Sedge} = 100 %, 4-fold at ~~*Shrub*%~~F_{Shrub} = 100%, and from neutral to a source at ~~footprint climatology. The~~ F_{Clim} (Fig 5a). Higher T_s generally had a negative effect ~~of soil moisture on CH_4 production~~NME (Fig 5c & d). The negative correlation between T_s and VWC (r = 0.54, < 0.01) may have contributed to this result. NN_{NME} performance improved less with the addition of *U* indicating the NN_{NME} was near saturation and its effects are less relevant. ~~emission has been noted in many other studies (e.g.~~ Higher U had a weak limiting effect on NME when *VWC* was high and increased NME when *VWC* was low (not shown).~~Zona et al., 2009; Nadeau et al., 2013; Olefeldt et al., 2013).~~

4 Discussion

4.1 Carbon Balance and Controlling Factors

Compared to other DTLB, Illisarvik has drier soils and greater shrub and grass cover (Table 4). Peak growing season CO_2 uptake at Illisarvik was greater than at most wet sedge-dominated DTLB (Table 4; Zona et al. 2010, Sturtevant and Oechel, 2013; Lara et al. 2015). These differences may be due to differences in the periods of observation and year to year variability but may also be due to the presence of more productive shrubs and slightly warmer climate at Illisarvik. Mean 1980-2010 T_a at Utqiagvik (formerly Barrow, AK) is -11.2 °C (US National Climate Data Centre, 2020). Tuktoyaktuk, the closest station to Illisarvik is 1.1° warmer. Shrub cover is expected to have a number of impacts on the microclimate and carbon cycle of Arctic tundra (eg. Myers-Smith et all, 2011). Typically, greater deciduous shrub cover is expected to increase GPP as a result of greater leaf area and photosynthetic potential compared to graminoid-dominated tundra (Sweet et al. 2015; Street et al., 2018). GPP was greater at Ilisarvik compared to the young wet-sedge dominated DTLB in Alaska (Zona et al., 2010). It was more similar to Katyk which has significant dwarf shrub cover, predominately *Betula nana* and *Salix pulchra* (van der Molen et al. 2007). Differences in ER among tundra environments can be related to substrate availability, soil moisture and temperature and thaw depth, among other factors (Sturtevant and Oechel, 2013). The 'snow-shrub hypothesis' (Sturm et al. 2001) describes the potential for greater snow trapping in shrub communities which insulates soils in winter, leads to increased decomposition and nutrient availability and promotes further shrub growth. At Illisarvik, snow blowing in off the Arctic Ocean results in large snow drifts within the basin where snow depth correlates with vegetation height

(Wilson et al., 2019). Wilson et al. (2019) concluded that the soils within the Illisarvik basin were warmer than those of the surrounding dwarf-shrub tundra in part through these snow-shrub interactions. Although our chamber observations suggested Shrub ER is lower than ER from other vegetation classes, this may have been an artifact as the taller shrubs (>40 cm) could not fit inside the chambers. In another study, chamber ER increased with greater shrub cover in upland tundra (Ge et al., 2017). ER at Illisarvik was greater than the ER observed at both the young wet-sedge DTLB in Barrow (Zona et al., 2010) and at the shrub/wet sedge DTLB at Katyk where thaw depth was much shallower (45 to >100 cm at Illisarvik vs. 25 to 40 cm at Katyk; van der Molen et al. 2007). The importance of F_{Shrub} in describing temporal variations in half hourly NEE within the flux footprint at Ilisarvik is further evidence of the importance of shrub cover on tundra carbon cycle processes in this environment.

PPFD and *VPD* were the most important factors for predicting half hourly NEE. This was to be expected as they are typically the primary controls over GPP (Aubinet et al., 2012). The limiting effects of VPD are consistent another study using NN to analyse NEE at a deciduous forest site (Moffat et al., 2012) and has been found at other tundra sites (Euskirchen et al. 2012; López-Blanco et al. 2017). *VWC* was also important at Illisarvik. Zona et al. (2010) found *VWC* could explain 70% of the variability in daily peak season ER in young DTLB. Similarly, Kittler et al. (2016) found drier soils increased ER and decreased NEE after a wet tundra drainage experiment in Siberia, consistent with our results at Illisarvik when F_{Shrub} was low.

As expected, NME at Illisarvik was about half that observed at the Alaskan DTLB sites where soils were wetter with greater sedge cover (Table 4, Zona et al., 2009; Lara et al. 2015). NME at Katyk was even higher than the Barrow DTLB and had a significant impact on the greenhouse gas (GHG) balance for this site (van der Molen et al. 2007; Parmentier et al., 2011). In our NN modelling of NME at Illisarvik, F_{Sedge} was the most important factor for predicting half hourly F_{CH4}. Sedges are aquatic plant species with arenchymatous tissues that act as conduits for CH_4 from below the water table to the atmosphere and limits CH_4 oxidation by methanotrophs in aerobic surface soils (Lai et al. 2009). The inclusion of F_{Shrub} further refined the model, allowing it to better fit the site-specific distribution of vegetation types. Budishchev et al. (2014) found shrub and sedge fraction had a significant influence on F_{CH4} at Katyk. Vegetation type is the dominant control over NME across multiple tundra landscapes and our results further support that (Davidson et al., 2016).

VWC was the second most important factor, which was expected as CH_4 production occurs in anaerobic environments and has been linked to variability in CH_4 emission in many other studies (e.g. Zona et al., 2009; Nadeau et al., 2013; Olefeldt et al., 2013). ~~Higher T_s generally had a negative effect on NME (Fig 5c & d). Higher soil temperatures~~Soil temperature (T_s) was the third most important factor. Higher T_s increase the oxidation potential of methanotrophs (Liu et al., 2016; King and Adamsen, 1992), so this result was expected for the drier portions of the basin and upland tundra. However, this ~~wasn't~~was not expected for the sedge areas because most studies find NME in sedges is positively correlated to T_s (Olefeldt et al., ~~2013). There was diurnal cycle in NME with NME peaking in the morning when T_s was at its daily minimum, which supports this finding~~2013). The negative correlation between T_s and *VWC* may partly explain this~~. NN_{NME} performance improved less with the addition of *U* indicating the NN_{NME} was near saturation and its effects are less relevant. Higher wind speed had a weak limiting effect on NME when VWC was~~

high and increased NME when VWC was low (not shown). High winds were mainly associated with two strong storm events. In order to better resolve this relationship a longer dataset would be needed.

3.4.2 Upscaling

ER_{FS} and NME_{FS} were about 59% and 47% greater than the EC estimates. Discrepancies between EC and chamber observations are common and have been attributed to differences in measurement techniques, the small sample size of chamber observations, and sampling bias since all chamber measurements were taken during the day with fair weather (Katayanagi et al., 2005; Chaichana et al., 2018). Meijide et al. (2011) found that chamber NEE could be up to twice as large as EC observations and Riederer et al. (2014) also found chamber NME estimates were about 30% higher than EC estimates. Others have been more successful, yielding upscaled chamber NME fluxes within 10% of EC observations (Zhang et al., 2012; Budishchev et al., 2014; Davidson et al., 2017). A potential reason for the disagreement with ER_{FS} may be the lack of direct observations by the EC system under low-light conditions. Another potential source of error for the upscaling is inaccuracies in the vegetation map.

4.3 Future Trajectories

Presently, NEE and ER peak growing season carbon uptake at Illisarvik are within ranges observed within similar is greater than similarly aged landscape features on the Barrow Peninsula but NME was considerably lower., Alaska and more similar to levels observed at Katyk, Siberia. NME is well below levels observed at any other DTLB studied, making this site a stronger GHG sink than other DTLB. However, the basin at Illisarvik will continue to evolve and the trajectory it takes could significantly alter theits carbon balance. MostHistorically, DTLB on Richards Island and the Tuktoyaktuk Peninsula evolve into wet sedge-moss peatlands (Ovendend, 1986). Assedge wetlands, as do DTLB on the Barrow Peninsula (Ovendend, 1986; Lara et al., 2015). This would cause NME to Active maintenance of the outlet channel at Illisarvik has artificially lowered soil moisture and flooding and potentially limited this transition thus far (C. Burn, personal communication 2016).

If Illisarvik follows the same trajectory as older DTLB in the area and becomes dominated by sedge wetlands, NME will increase significantly. NN_{NME} estimated that at With extrapolations to full Sedge$_{\%}$ cover (F_{Sedge} = 100% mean%), NME would be 17.9 [$CI_{95\%}$ ± 10.6] mg C CH_4 m^{-2} d^{-1}, which is comparablesimilar to EC observations from a wet sedge DTLBvalues on the Barrow Peninsula (Zona et al., 2009). If the basin becomes wetter and the shrubs are displaced by sedges and grasses net carbon uptake may increase. NN_{NEE} and NN_{ER} estimated that at $Shrub_{\%}$ = 0% NEE and ER over the study period would be -1.9 [$CI_{95\%}$ ± 0.5] g C CO_2 m^{-2} d^{-1} and 1.8 [$CI_{95\%}$ ± 1.1] g C CO_2 m^{-2} d^{-1} respectivelyIf the basin instead transitions into a shrub dominated DTLB similar to those of Old Crow Flats, Yukon (Lantz et al., 2015), NME_{NN} would remain similar to current levels meaning the basin would remain a weak source of CH_4. These are projections well beyond F_{Clim} fractions observed so confidence in the specific values predicted is low. However, active maintenance of the basin's outlet channel (C. Burn, personal communication 2016) has artificially lowered soil moisture and potentially limited this transition. This coupled with climate change will promote shrub expansion and Illisarvik could end up more like the shrub dominated DTLB of Old Crow Flats, Yukon (Lantz et al., 2015). NN_{NEE} and NN_{ER} estimated that at $Shrub_{\%}$ = 100% NEE and ER over the study period would be -0.9 [$CI_{95\%}$ ±

~~1.5] g C CO_2 m^{-2} d^{-1} and 2.4 [$CI_{95\%}$ ± 1.6] g C CO_2 m^{-2} d^{-1} respectively. Other studies have also suggested that shrub expansion could influence NEE by increasing ER (Merbould et al., 2009; Meyer Smith et~~ The effects of changing shrub/sedge cover on Illisarvik's growing season NEE are less straightforward than on NME. Partly because Shrub cover had less overall influence on NEE_{NN}. Our model suggests ER decreases and GPP increases with increasing shrub coverage when soils are slightly drier, but has the opposite effect under wetter conditions. To our knowledge, only few winter season (e.g. Zona et al. 2016) and no year-round studies of NEE and NME have yet to be published to help evaluate the factors influencing DTLB carbon losses through the non-growing season months. Further observation year-round is needed to better understand the implications of continued vegetation change on the carbon balance of DTLB such as Illisarvik.

~~5al., 2011). NN_{NME} estimated NME would be 13.1 [$CI_{95\%}$ ± 9.4] mg C CH_4 m^{-2} d^{-1} at $Shrub_{\%}$ = 100%, meaning the basin will likely remain a net source of CH_4.~~

~~4~~ Conclusions

This study investigated NEE, GPP, ER and NME in the Illisarvik experimental DTLB using EC and chamber data. To our knowledge this is the first such study conducted in a DTLB outside of the Barrow Peninsula~~. Our observations were generally in agreement with other studies but show how Illisarvik differs from the colder, wetter DTLB on the Barrow Peninsula.~~ or Siberia. Illisarvik is a carbon sink during the growing season with NME only having a small ~~positive~~ effect on the net carbon balance. ~~Illisarvik NEE was similar to young~~ Our flux observations were generally in agreement with other studies but show how shrub-dominated DTLB such as Illisarvik and Katyk in Siberia differ from sedge-dominated DTLB on the Barrow Peninsula. Illisarvik's growing season net carbon uptake was greater than young and ancient DTLB on the Barrow Peninsula~~, while~~ and more similar to the shrub dominated ancient DTLB in Siberia. NME at Illisarvik was lower than all published DTLB studies likely due to better drainage~~, drier conditions,~~ and more diverse vegetation. ~~However, higher NME in early July indicated we likely missed a key period of CH_4 emissions earlier in the season.~~ A longer, more comprehensive study would be needed to resolve the annual carbon budget for Illisarvik.

Chamber measurements of ER and NME from different land cover classes within and outside the Illisarvik basin added context to the EC observations. Vegetation class (and associated difference in terrain and soil properties) had only a ~~slight~~small but significant impact on NEE and ER but was one of the dominant controls over NME. Sedge areas were a strong source of CH_4, other vegetation types in the basin were weak sources, and upland areas were a net sink. These results suggest that NME in particular ~~is expected to shift~~will change as ~~both the terrain and the vegetation of~~ the Illisarvik DTLB ~~continues~~vegetation communities continue to evolve.

Appendix A: Neural Networks analysis and uncertainty calculations

Typically, NEE is gap-filled using flux-partitioning algorithms that model ER and GPP separately using T_S and *PPFD*, respectively (e.g. Lee et al., 2017; Aubinet, 2012). However, this method requires night-time observations and thus does not perform well for Arctic summertime measurements due to the limited number of samples available during

low light conditions. There are no widely agreed upon functional relationships for gap-filling NME since CH_4 production and consumption vary considerably both between different landcover types and environmental conditions. Some methods that have been used include~~: classification and regression trees (CART) (Nadeau et al., 2013; Sachs et al., 2008),~~ general linear models (GLM) (Zona et al., 2009), ~~and~~ mean diurnal variation (Nadeau et al., 2014), and classification and regression trees (CART) (Nadeau et al., 2013; Sachs et al., ~~).~~2008). We attempted to use a GLM~~, CART,~~ and ~~random forest regression trees~~CART but they were not flexible enough to account for source area variability.

Neural networks (NN) are flexible machine learning methods that are ideally suited to perform non-linear, multivariate regression. They make no a priori assumptions about the functional relationships between the factors and responses. (Melesse and Hanley, 2005; Desai et al., 2008). NN are universal approximators; given enough hidden nodes a NN is capable of mapping any continuous function to an arbitrary degree of accuracy (Hornik et al., 1991). If all relevant climate and ecosystem information are available to a network, the remaining variability can be attributed to noise in the measurement (Moffat et al., 2010).

NN have been shown to be among the best performing methods for gap-filling NEE data for temperate forest and wetland sites ~~in Europe (~~(Papale et al., 2003; Moffat et al., 2007; Knox et al., 2016). They have also been used to gap-fill NME time series in sub-arctic wetlands, tundra sites, and wet sedge tundra (Dengel et al., 2013). NN have been used to identify and model factors influencing NEE and to partition NEE into ER and GPP (Moffat et al., 2010). NNs have even been used to upscale fluxes from the ecosystem level to the continental scale (Dou and Yang, 2018; Papale et al, 2003).

A NN approximates a true regression function $F(X)$:

$$F(X) = t(X) - \varepsilon(X) \quad \text{(A1)}$$

where $t(X)$ is the target function and $\varepsilon(X)$ the noise (Khosravi and Nahavandi, 2010). $X = [x_0, x_1, \dots, x_M]$ where $x_0 = 1$ is a bias term and $[x_1, \dots, x_M]$ are the independent variables. M denotes the number of independent variables. The network approximates $F(X)$ as $f(X, w)$ by mapping the relationship between X and the target. Here we used feed-forward dense NN with a single hidden layer:

$$f(X, w) = \sum_{h=1}^{H} \beta_h g\left(\sum_{m=0}^{M} \gamma_{hm} x_m\right) \quad \text{(A2)}$$

$g(\cdot)$ is a non-linear transfer function, here we used the rectified linear activation unit (ReLu) (Anders and Korn, 1999). H denotes the number of hidden nodes in the network and must be assigned before training. Too many hidden nodes and the NN will overfit the training data, too few and it will underfit. Early stopping will prevent NN from overfitting training sets (Weigend, 1993; Sarle, 1995; Tetko et al., 1995). Therefore, it is more important to ensure a NN has enough hidden nodes to adequately map the target function (Smith, 1994). We set H to a function M, the number of training samples (N), and the number of targets (1):

$$H = \frac{N}{a*(M+1)} \quad \text{(A3)}$$

This rule of thumb ensures a NN has sufficient flexibility to approximate the target response. The weights $w = [\beta_1 \dots \beta_N, \gamma_{10} \dots \gamma_{NM}]$ are randomly initialized and after each model iteration is updated by backpropagating the error through the network. N denotes the number of observations or targets. The error metric most commonly used is the mean squared error, MSE:

$$MSE = \sum_{i=1}^{N}(f(X_i) - t_i)^2 \quad \text{(A4)}$$

The weights are adjusted in the direction that will decrease the error and training continues until a stopping criterion is reached. We chose to set aside 20% of the training data as a test set to be used for early stopping, and terminated training when the MSE of the test set failed to improve for 10 consecutive iterations.

Bootstrapping is used to account for model variability and estimate confidence and prediction intervals by training NN on *B* different realizations of the dataset, where *B* is the number of bootstrapped samples, we used *B* = 30 (Heskes, 1997; Khosravi & Nahavandi, 2010). An individual NN generates point outputs approximating a target function with no information on the confidence in those estimates (Khosravi & Nahavandi, 2010). However, there are usually multiple $f(X, w)$ that approximate $F(X)$ because of the random weight initializations (Weigend & LeBaron, 1994). As such, there are two sources of error we are concerned with, the accuracy of our estimation of $F(X)$ and the accuracy of our estimates with respect to the target. A confidence interval describes the first (e.g. $F(X) - f(X, w)$) while a prediction interval describes the latter (eg. $t(X) - f(X, w)$) (Heskes, 1997). By definition, a prediction interval contains the confidence interval because:

$$t(X) - f(X, w) = [F(X) - f(X, w)] + \varepsilon(X) \quad \text{(A5)}$$

For b = 1 … *B*, a random sample with replacement of size *p* is drawn from the original dataset. Setting *p* equal to the size of the original dataset yields a set of *B* training sets each containing approximately 67% of the original dataset. The 33% leftover from each bootstrap sampled can be used for model validation (Heskes, 1997). The average of our ensemble of networks can then serve as our approximation of F(X):

$$F(X) = \frac{1}{B}\sum_{b=1}^{B} f_b(X, W) \quad \text{(A6)}$$

The variance of the model outputs is:

$$\sigma^2(X) = \frac{1}{B-1}\sum_{b=1}^{B}\left(f_b(X, W) - F(X)\right)^2 \quad \text{(A7)}$$

A confidence interval (CI) for $F(X)$ can be calculated as $F(X) \pm t_{(1-\propto, df)}\sigma(X)$, where *tscore* is the students t-score, 1-α is the desired confidence level, and df are the degrees of freedom which are set to the number of bootstrapped samples B. NN performance can be seen to improve with the inclusion of more factors, until the model saturates and becomes over-parametrized (~~Figure~~Fig A1).

Random forests (RF) are said to be among the best performing gap filling methods for NME (Kim et al., 2020). and it has been claimed that aggregating many regression trees in a RF prevents overfitting (Breiman, 2001;). We did not find this to be the case. Following the methods outlined in Kim et al. (2020): a RF with 400 trees and no restrictions on tree size fit F_{CH4} nearly perfectly (R^2 = 0.98). Without considerable limitations on tree size, the RF will just learn the dataset rather than the relationships present. It is our view that this tree is extremely overfit, as highlighted by the example in Figure A2. Further, RF do not allow for straightforward visualization functional relationships in a dataset. Plotting F_{CH4} against VWC, which is the dominant environmental control identified does not reveal a meaningful relationship like Figure 5 a & c. You can look at an individual decision tree within the RF, but those are difficult to interpret beyond the first few splits, and each tree will be different. Lastly, RF are incapable of projecting beyond the parameter space observed which limited their applicability for this study (Fig A2). This presents an issue because

may gaps in EC data arise from data filtering (e.g. clear calm nights, precipitation events) and are by definition outside the parameter space observed.

Data & Code availability

Our data and code are available on github: https://github.com/June-Spaceboots/Illisarvik_CFluxes

Author Contribution

JS, AC, and GH designed the EC study. AL and EH designed the chamber study. JS collected, processed, and analysed the EC data. AL and EH collected the chamber data with help from JS. AL and EH processed the chamber data. JS designed and conducted the NN analysis. JS prepared the manuscript with input from all co-authors.

Competing interests

The authors declare they have no competing interests.

Acknowledgements

We would like to thank: the staff at the Aurora Research Institute in Inuvik for providing logistical support, Chris Burn for allowing us to work at Illisarvik and the knowledge he shared, Alice Wilson for sharing vegetation survey data, and Tony ~~Lekowicz~~Lewkowicz for collecting and sharing the drone images~~.~~, and Rick Ketler for providing logistical support. Funding for this study was provided by the Canada Foundation for Innovation ~~and~~, the ~~National Science~~Natural Sciences and Engineering Research Council ~~(NSERC).~~of Canada, and the Polar Continental Shelf Program, Natural Resources Canada.

References

Anders, U. and Korn, O.: Model selection in neural networks, Neural Networks, 12(2), 309–323, doi:10.1016/S0893-6080(98)00117-8, 1999.

Aubinet, M., Vesala, T. and Papale, D., Eds.: Eddy Covariance: A Practical Guide to Measurement and Data Analysis, 2012 edition., Springer, Dordrecht ; New York., 2012.

Biasi, C., Meyer, H., Rusalimova, O., Hämmerle, R., Kaiser, C., Baranyi, C., Daims, H., Lashchinsky, N., Barsukov, P. and Richter, A.: Initial effects of experimental warming on carbon exchange rates, plant growth and microbial dynamics of a lichen-rich dwarf shrub tundra in Siberia, Plant Soil, 307(1), 191–205, doi:10.1007/s11104-008-9596-2, 2008.

Billings, W. and Peterson, K.: Vegetational Change and Ice-Wedge Polygons Through the Thaw-Lake Cycle, Arctic and Alpine Research, 12(4), 413–432, doi:10.2307/1550492, 1980.

Budishchev, A., Mi, Y., van Huissteden, J., Belelli-Marchesini, L., Schaepman-Strub, G., Parmentier, F. J. W., Fratini, G., Gallagher, A., Maximov, T. C. and Dolman, A. J.: Evaluation of a plot-scale methane emission model using eddy covariance observations and footprint modelling, Biogeosciences, 11(17), 4651–4664, doi:https://doi.org/10.5194/bg-11-4651-2014, 2014.

Canada, E. and C. C.: Canadian Climate Normals - Climate - Environment and Climate Change Canada, [online] Available from: https://climate.weather.gc.ca/climate_normals/index_e.html (Accessed 3 December 2019), 2011.

Cahoon, S. M. P., Sullivan, P. F., Shaver, G. R., Welker, J. M. and Post, E.: Interactions among shrub cover and the soil microclimate may determine future Arctic carbon budgets, Ecology Letters, 15(12), 1415–1422, doi:10.1111/j.1461-0248.2012.01865.x, 2012.

Chaichana, N., Bellingrath-Kimura, S. D., Komiya, S., Fujii, Y., Noborio, K., Dietrich, O. and Pakoktom, T.: Comparison of Closed Chamber and Eddy Covariance Methods to Improve the Understanding of Methane Fluxes from Rice Paddy Fields in Japan, Atmosphere, 9(9), 356, doi:10.3390/atmos9090356, 2018.

Chollet, F. and others: Keras, [online] Available from: https://keras.io, 2015.

Davidson, S. J., Sloan, V. L., Phoenix, G. K., Wagner, R., Fisher, J. P., Oechel, W. C. and Zona, D.: Vegetation Type Dominates the Spatial Variability in CH4 Emissions Across Multiple Arctic Tundra Landscapes, Ecosystems, 19(6), 1116–1132, doi:10.1007/s10021-016-9991-0, 2016.

Davidson, S. J., Santos, M. J., Sloan, V. L., Reuss-Schmidt, K., Phoenix, G. K., Oechel, W. C. and Zona, D.: Upscaling CH4 Fluxes Using High-Resolution Imagery in Arctic Tundra Ecosystems, Remote Sensing, 9(12), 1227, doi:10.3390/rs9121227, 2017.

Dengel, S., Zona, D., Sachs, T., Aurela, M., Jammet, M., Parmentier, F. J. W., Oechel, W. and Vesala, T.: Testing the applicability of neural networks as a gap-filling method using ~~CH₄~~CH_4; flux data from high latitude wetlands, Biogeosciences, 10(12), 8185–8200, doi:10.5194/bg-10-8185-2013, 2013.

Desai, A. R., Richardson, A. D., Moffat, A. M., Kattge, J., Hollinger, D. Y., Barr, A., Falge, E., Noormets, A., Papale, D., Reichstein, M. and Stauch, V. J.: Cross-site evaluation of eddy covariance GPP and RE decomposition techniques, Agricultural and Forest Meteorology, 148(6), 821–838, doi:10.1016/j.agrformet.2007.11.012, 2008.

Dou, X. and Yang, Y.: Estimating forest carbon fluxes using four different data-driven techniques based on long-term eddy covariance measurements: Model comparison and evaluation, Science of The Total Environment, 627, 78–94, doi:10.1016/j.scitotenv.2018.01.202, 2018.

Dybowski, R. and Roberts, S. J.: Confidence intervals and prediction intervals for feed-forward neural networks, Clinical applications of artificial neural networks, 298–326, 2001.

Euskirchen, E. S., Bret-Harte, M. S., Scott, G. J., Edgar, C. and Shaver, G. R.: Seasonal patterns of carbon dioxide and water fluxes in three representative tundra ecosystems in northern Alaska, Ecosphere, 3(1), art4, doi:10.1890/ES11-00202.1, 2012.

Foken, T., Göockede, M., Mauder, M., Mahrt, L., Amiro, B. and Munger, W.: Post-Field Data Quality Control, in Handbook of Micrometeorology, edited by X. Lee, W. Massman, and B. Law, pp. 181–208, Springer Netherlands., 2004.

French, H. M.: The Periglacial Environment, John Wiley & Sons., 2013.

Heskes, T.: Practical Confidence and Prediction Intervals, in Advances in Neural Information Processing Systems 9, pp. 176–182, MIT press., 1997.

Ge, L., Lafleur, P. M. and Humphreys, E. R.: Respiration from Soil and Ground Cover Vegetation Under Tundra Shrubs, [online] Available from: https://pubag.nal.usda.gov/catalog/5943221 (Accessed 28 April 2020), 2017.

Hornik, K.: Approximation capabilities of multilayer feedforward networks, Neural networks, 4(2), 251–257, 1991.

van Huissteden, J., Berrittella, C., Parmentier, F. J. W., Mi, Y., Maximov, T. C. and Dolman, A. J.: Methane emissions from permafrost thaw lakes limited by lake drainage, Nature ~~Clim.~~Climate Change, 1(2), 119–123, doi:10.1038/nclimate1101, 2011.

Intergovernmental Panel on Climate Change (IPCC). 2014. Climate Change 2014: Synthesis Report. Fifth Assessment Report of the Intergovernmental Panel on Climate Change. IPCC, Geneva, Switzerland.

Jones, B. M., Grosse, G., Arp, C. D., Jones, M. C., Anthony, K. M. W. and Romanovsky, V. E.: Modern thermokarst lake dynamics in the continuous permafrost zone, northern Seward Peninsula, Alaska, Journal of Geophysical Research: Biogeosciences, doi:10.1029/2011JG001666@10.1002/(ISSN)2169-8961.TKLCARBON1, 2018.

Katayanagi, N.: Spatial variability of greenhouse gas fluxes from soils of various land uses on a livestock farm in southern Hokkaido, Japan, Phyton (Austria) Special Issue:" APGC 2004", 45, 309–318, 2005.

Khosravi, A. and Nahavandi, S.: A Comprehensive Review of Neural Network-based Prediction Intervals and New Advances, ~~, 17,~~ 2010.

Kim, Y., Johnson, M. S., Knox, S. H., Black, T. A., Dalmagro, H. J., Kang, M., Kim, J. and Baldocchi, D.: Gap-filling approaches for eddy covariance methane fluxes: A comparison of three machine learning algorithms and a traditional method with principal component analysis, Global Change Biology, 26(3), 1499–1518, doi:10.1111/gcb.14845, 2020.

King, G. M. and Adamsen, A. P. S.: Effects of Temperature on Methane Consumption in a Forest Soil and in Pure Cultures of the Methanotroph Methylomonas rubra, Appl. Environ. Microbiol., 58(9), 2758–2763, 1992.

Kittler, F., Burjack, I., Corradi, C. A. R., Heimann, M., Kolle, O., Merbold, L., Zimov, N., Zimov, S. and Gockede, M.: Impacts of a decadal drainage disturbance on surface-atmosphere fluxes of carbon dioxide in a permafrost ecosystem, Biogeosciences, 13(18), 5315–5332, doi:10.5194/bg-13-5315-2016, 2016.

Kljun, N., Calanca, P., Rotach, M. W. and Schmid, H. P.: A simple two-dimensional parameterisation for Flux Footprint Prediction (FFP), Geoscientific Model Development, 8(11), 3695–3713, 2015.

Knox, S. H., Matthes, J. H., Sturtevant, C., Oikawa, P. Y., Verfaillie, J. and Baldocchi, D.: Biophysical controls on interannual variability in ecosystem-scale CO2 and CH4 exchange in a California rice paddy, Journal of Geophysical Research: Biogeosciences, 121(3), 978–1001, doi:10.1002/2015JG003247, 2016.

Lantz, T. C. and Turner, K. W.: Changes in lake area in response to thermokarst processes and climate in Old Crow Flats, Yukon, J. Geophys. Res.-Biogeosci., 120(3), 513–524, doi:10.1002/2014JG002744, 2015.

Lafleur, P. M., Humphreys, E. R., St. Louis, V. L., Myklebust, M. C., Papakyriakou, T., Poissant, L., Barker, J. D., Pilote, M. and Swystun, K. A.: Variation in Peak Growing Season Net Ecosystem Production Across the Canadian Arctic, Environ. Sci. Technol., 46(15), 7971–7977, doi:10.1021/es300500m, 2012.

Lara, M. J., McGuire, A. D., Euskirchen, E. S., Tweedie, C. E., Hinkel, K. M., Skurikhin, A. N., Romanovsky, V. E., Grosse, G., Bolton, W. R. and Genet, H.: Polygonal tundra geomorphological change in response to warming alters

future CO2 and CH4 flux on the Barrow Peninsula, Global Change Biology, 21(4), 1634–1651, doi:10.1111/gcb.12757, 2015.

Lee, S.-C., Christen, A., Black, A. T., Johnson, M. S., Jassal, R. S., Ketler, R., Nesic, Z. and Merkens, M.: Annual greenhouse gas budget for a bog ecosystem undergoing restoration by rewetting, Biogeosciences, 14(11), 2799–2814, doi:10.5194/bg-14-2799-2017, 2017.

Liu, Y., Liu, X., Cheng, K., Li, L., Zhang, X., Zheng, J., Zheng, J. and Pan, G.: Responses of Methanogenic and Methanotrophic Communities to Elevated Atmospheric CO2 and Temperature in a Paddy Field, Front Microbiol, 7, doi:10.3389/fmicb.2016.01895, 2016.

López-Blanco, E., Lund, M., Williams, M., Tamstorf, M. P., Westergaard-Nielsen, A., Exbrayat, J.-F., Hansen, B. U. and Christensen, T. R.: Exchange of CO 2 in Arctic tundra: impacts of meteorological variations and biological disturbance, Biogeosciences, 14(19), 4467–4483, doi:10.5194/bg-14-4467-2017, 2017.

~~Mackay, J., R.: An experiment in lake drainage, Richards Island, Northwest Territories: a progress report., Geological Survey of Canada, Paper, (81–1A:), 63–68, 1981.~~

Mackay, J. R.: Periglacial features developed on the exposed lake bottoms of seven lakes that drained rapidly after 1950, Tuktoyaktuk Peninsula area, western Arctic coast, Canada, Permafrost Periglacial Process., 10(1), 39–63, doi:10.1002/(SICI)1099-1530(199901/03)10:1<39::AID-PPP305>3.3.CO;2-I, 1999.

Mackay, J. R. and Burn, C. R.: The first 20 years (1978-1979 to 1998-1999) of active-layer development, Illisarvik experimental drained lake site, western Arctic coast, Canada, Can. J. Earth Sci., 39(11), 1657–1674, doi:10.1139/E02-068, 2002.

Marsh, P., Russell, M., Pohl, S., Haywood, H. and Onclin, C.: Changes in thaw lake drainage in the Western Canadian Arctic from 1950 to 2000, Hydrol. Process., 23(1), 145–158, doi:10.1002/hyp.7179, 2009.

Martin, A. F., Lantz, T. C. and Humphreys, E. R.: Ice wedge degradation and CO2 and CH4 emissions in the Tuktoyaktuk Coastlands, Northwest Territories, Arctic Science, 4(1), 130–145, 2017.

Mauder, M. and Foken, T.: Documentation and Instruction Manual of the Eddy Covariance Software Package TK2, Univ., Abt. Mikrometeorologie., 2004.

McGuire, A. D., Anderson, L. G., Christensen, T. R., Dallimore, S., Guo, L., Hayes, D. J., Heimann, M., Lorenson, T. D., Macdonald, R. W. and Roulet, N.: Sensitivity of the carbon cycle in the Arctic to climate change, Ecological Monographs, 79(4), 523–555, doi:10.1890/08-2025.1, 2009.

Meijide, A., Manca, G., Goded, I., Magliulo, V., Tommasi, P. di, Seufert, G. and Cescatti, A.: Seasonal trends and environmental controls of methane emissions in a rice paddy field in Northern Italy, Biogeosciences, 8(12), 3809–3821, doi:https://doi.org/10.5194/bg-8-3809-2011, 2011.

Melesse, A. M. and Hanley, R. S.: Artificial neural network application for multi-ecosystem carbon flux simulation, Ecological Modelling, 189(3), 305–314, doi:10.1016/j.ecolmodel.2005.03.014, 2005.

~~Merbold, L., Kutsch, W~~Michel, F. A., Fritz, P. and Drimmie, R. J.: Evidence of climatic change from oxygen and carbon isotope variations in sediments of a small arctic lake, Canada, Journal of Quaternary Science, 4(3), 201–209, doi:10.1002/jqs.3390040302, 1989.

. L., Corradi, C., Kolle, O., Rebmann, C., Stoy, P. C., Zimov, S. A. and Schulze, E. D.: Artificial drainage and associated carbon fluxes (CO2/CH4) in a tundra ecosystem, Global Change Biology, 15(11), 2599–2614, doi:10.1111/j.1365-2486.2009.01962.x, 2009.

Moffat, A. M., Papale, D., Reichstein, M., Hollinger, D. Y., Richardson, A. D., Barr, A. G., Beckstein, C., Braswell, B. H., Churkina, G., Desai, A. R., Falge, E., Gove, J. H., Heimann, M., Hui, D., Jarvis, A. J., Kattge, J., Noormets, A. and Stauch, V. J.: Comprehensive comparison of gap-filling techniques for eddy covariance net carbon fluxes, Agricultural and Forest Meteorology, 147(3–4), 209–232, doi:10.1016/j.agrformet.2007.08.011, 2007.

Moffat, A. M., Beckstein, C., Churkina, G., Mund, M. and Heimann, M.: Characterization of ecosystem responses to climatic controls using artificial neural networks, Global Change Biology, 16(10), 2737–2749, doi:10.1111/j.1365-2486.2010.02171.x, 2010.

Moncrieff, J., Clement, R., Finnigan, J. and Meyers, T.: Averaging, Detrending, and Filtering of Eddy Covariance Time Series, in Handbook of Micrometeorology, edited by X. Lee, W. Massman, and B. Law, pp. 7–31, Springer Netherlands., 2004.

Moncrieff, J. B., Massheder, J. M., de Bruin, H., Elbers, J., Friborg, T., Heusinkveld, B., Kabat, P., Scott, S., Soegaard, H. and Verhoef, A.: A system to measure surface fluxes of momentum, sensible heat, water vapour and carbon dioxide, Journal of Hydrology, 188, 589–611, doi:10.1016/S0022-1694(96)03194-0, 1997.

Myers-Smith, I. H., Forbes, B. C., Wilmking, M., Hallinger, M., Lantz, T., Blok, D., Tape, K. D., Macias-Fauria, M., Sass-Klaassen, U., Lévesque, E., Boudreau, S., Ropars, P., Hermanutz, L., Trant, A., Collier, L. S., Weijers, S., Rozema, J., Rayback, S. A., Schmidt, N. M., Schaepman-Strub, G., Wipf, S., Rixen, C., Ménard, C. B., Venn, S., Goetz, S., Andreu-Hayles, L., Elmendorf, S., Ravolainen, V., Welker, J., Grogan, P., Epstein, H. E. and Hik, D. S.: Shrub expansion in tundra ecosystems: dynamics, impacts and research priorities, Environ. Res. Lett., 6(4), 045509, doi:10.1088/1748-9326/6/4/045509, 2011.

Nadeau, D. F., Rousseau, A. N., Coursolle, C., Margolis, H. A. and Parlange, M. B.: Summer methane fluxes from a boreal bog in northern Quebec, Canada, using eddy covariance measurements, Atmospheric Environment, 81, 464–474, doi:10.1016/j.atmosenv.2013.09.044, 2013.

Olefeldt, D., Turetsky, M. R., Crill, P. M. and McGuire, A. D.: Environmental and physical controls on northern terrestrial methane emissions across permafrost zones, Glob Change Biol, 19(2), 589–603, doi:10.1111/gcb.12071, 2013.

O'Neill, H. B.,. and Burn, C. R. and Gajewski, K.: Physical and temporal factors controlling the development of near-surface ground ice at Illisarvik, western Arctic coast, Canada, Canadian Journal of Earth Sciences, 49(9), 1096–1110, 2012.

Ovenden, L.: Vegetation Colonizing the Bed of a Recently Drained Thermokarst Lake (illisarvik), Northwest-Territories, Can. J. Bot.-Rev. Can. Bot., 64(11), 2688–2692, 1986.

Papale, D. and Valentini, R.: A new assessment of European forests carbon exchanges by eddy fluxes and artificial neural network spatialization, Global Change Biology, 9(4), 525–535, doi:10.1046/j.1365-2486.2003.00609.x, 2003.

Papale, D., Reichstein, M., Aubinet, M., Canfora, E., Bernhofer, C., Kutsch, W., Longdoz, B., Rambal, S., Valentini, R., Vesala, T. and others: Towards a standardized processing of Net Ecosystem Exchange measured with eddy covariance technique: algorithms and uncertainty estimation, Biogeosciences, 3(4), 571–583, 2006.

Parmentier, F. J. W., Huissteden, J. van, Molen, M. K. van der, Schaepman-Strub, G., Karsanaev, S. A., Maximov, T. C. and Dolman, A. J.: Spatial and temporal dynamics in eddy covariance observations of methane fluxes at a tundra site in northeastern Siberia, Journal of Geophysical Research: Biogeosciences, 116(G3), doi:10.1029/2010JG001637, 2011.

Paul-Limoges, E., Christen, A., Coops, N., Black, T. and Trofymow, J.: Estimation of aerodynamic roughness of a harvested Douglas-fir forest using airborne LiDAR, Remote Sensing of Environment, 136, 225–233, doi:10.1016/j.rse.2013.05.007, 2013.

Riederer, M., Serafimovich, A. and Foken, T.: Net ecosystem CO2 exchange measurements by the closed chamber method and the eddy covariance technique and their dependence on atmospheric conditions, Atmospheric Measurement Techniques, 7(4), 1057–1064, doi:https://doi.org/10.5194/amt-7-1057-2014, 2014.

Rößger, N., Wille, C., Holl, D., Göckede, M. and Kutzbach, L.: Scaling and balancing carbon dioxide fluxes in a heterogeneous tundra ecosystem of the Lena River Delta, Biogeosciences, 16(13), 2591–2615, doi:https://doi.org/10.5194/bg-16-2591-2019, 2019a.

Rößger, N., Wille, C., Veh, G., Boike, J. and Kutzbach, L.: Scaling and balancing methane fluxes in a heterogeneous tundra ecosystem of the Lena River Delta, Agricultural and Forest Meteorology, 266–267, 243–255, doi:10.1016/j.agrformet.2018.06.026, 2019b.

Sachs, T., Wille, C., Boike, J. and Kutzbach, L.: Environmental controls on ecosystem-scale CH4 emission from polygonal tundra in the Lena River Delta, Siberia, Journal of Geophysical Research-Biogeosciences, 113, G00A03, doi:10.1029/2007JG000505, 2008.

Sarle, W. S.: Stopped Training and Other Remedies for Overfitting, in Proceedings of the 27th Symposium on the Interface of Computing Science and Statistics, pp. 352–360., 1995.

Schmid, H. P.: Footprint modeling for vegetation atmosphere exchange studies: a review and perspective, Agricultural and Forest Meteorology, 113(1–4), 159–183, doi:10.1016/S0168-1923(02)00107-7, 2002.

Skeeter, J.: June-Spaceboots/Illisarvik_CFluxes. [online] Available from: https://github.com/June-Spaceboots/Illisarvik_CFluxes (Accessed 3 December 2019), 2019.

Smith, M.: Neural Networks for Statistical Modeling, John Wiley & Sons, Inc., New York, NY, USA., 1993.

Sturtevant, C. S. and Oechel, W. C.: Spatial variation in landscape-level CO2 and CH4 fluxes from arctic coastal tundra: influence from vegetation, wetness, and the thaw lake cycle, Glob. Change Biol., 19(9), 2853–2866, doi:10.1111/gcb.12247, 2013.

Street, L. E., Subke, J.-A., Baxter, R., Dinsmore, K. J., Knoblauch, C. and Wookey, P. A.: Ecosystem carbon dynamics differ between tundra shrub types in the western Canadian Arctic, Environ. Res. Lett., 13(8), 084014, doi:10.1088/1748-9326/aad363, 2018.

Sturm, M., Racine, C. and Tape, K.: Increasing shrub abundance in the Arctic, Nature, 411(6837), 546–547, doi:10.1038/35079180, 2001.

Sweet, S. K., Griffin, K. L., Steltzer, H., Gough, L. and Boelman, N. T.: Greater deciduous shrub abundance extends tundra peak season and increases modeled net CO2 uptake, Glob Chang Biol, 21(6), 2394–2409, doi:10.1111/gcb.12852, 2015.

Tarnocai, C., Canadell, J. G., Schuur, E. A. G., Kuhry, P., Mazhitova, G. and Zimov, S.: Soil organic carbon pools in the northern circumpolar permafrost region: Soil organic carbon pools, Global Biogeochemical Cycles, 23(2), n/a-n/a, doi:10.1029/2008GB003327, 2009.

Tetko, I. V., Livingstone, D. J. and Luik, A. I.: Neural network studies. 1. Comparison of overfitting and overtraining, Journal of Chemical Information and Modeling, 35(5), 826–833, doi:10.1021/ci00027a006, 1995.

Vickers, D. and Mahrt, L.: Quality Control and Flux Sampling Problems for Tower and Aircraft Data, http://dx.doi.org/10.1175/1520-0426(1997)014<0512:QCAFSP>2.0.CO;2 [online] Available from: http://journals.ametsoc.org/doi/abs/10.1175/1520-0426%281997%29014%3C0512%3AQCAFSP%3E2.0.CO%3B2 (Accessed 6 March 2017), 199721997.

Virkkala, A.-M., Virtanen, T., Lehtonen, A., Rinne, J. and Luoto, M.: The current state of CO 2 flux chamber studies in the Arctic tundra: A review, Progress in Physical Geography, 42, 030913331774578, doi:10.1177/0309133317745784, 2017.

Mackay, J., R.: An experiment in lake drainage, Richards Island, Northwest Territories: a progress report., Geological Survey of Canada, Paper, (81–1A:), 63–68, 1981.

Vitale, D., Bilancia, M. and Papale, D.: A Multiple Imputation Strategy for Eddy Covariance Data, Journal of Environmental Informatics, 34(2), 68-87–87, 2018.

Walker, D., Raynolds, M., Daniëls, F., Einarsson, E., Elvebakk, A., Gould, W., Katenin, A., Kholod, S., Markon, C., Melnikov, E., Moskalenko, N., Talbot, S., Yurtsev, B. and Team, T.: The Circumpolar Arctic Vegetation Map, Journal of Vegetation Science, 16, 267–282, doi:10.1111/j.1654-1103.2005.tb02365.x, 2005.

Walter, K. M., Smith, L. C. and Chapin, F. S.: Methane Bubbling from Northern Lakes: Present and Future Contributions to the Global Methane Budget, Philosophical Transactions: Mathematical, Physical and Engineering Sciences, 365(1856), 1657–1676, 2007.

Whalen, S. C. and Reeburgh, W. S.: Consumption of atmospheric methane by tundra soils, Nature, 346(6280), 160–162, doi:10.1038/346160a0, 1990.

Webb, E., Pearman, G. and Leuning, R.: Correction of Flux Measurements for Density Effects Due to Heat and Water-Vapor Transfer, Q. J. R. Meteorol. Soc., 106(447), 85–100, doi:10.1002/qj.49710644707, 1980.

Weigend, A. S. and Lebaron, B.: Evaluating Neural Network Predictors by Bootstrapping, Humboldt University of Berlin, Interdisciplinary Research Project 373: Quantification and Simulation of Economic Processes. [online] Available from: https://ideas.repec.org/p/zbw/sfb373/199435.html (Accessed 21 January 2019), 1994.

Whalen, S. C. and Reeburgh, W. S.: Consumption of atmospheric methane by tundra soils, Nature, 346(6280), 160–162, doi:10.1038/346160a0, 1990.

Wilczak, J. M., Oncley, S. P. and Stage, S. A.: Sonic Anemometer Tilt Correction Algorithms, Boundary-Layer Meteorology, 99(1), 127–150, doi:10.1023/A:1018966204465, 2001.

Wilson, M. A., Burn C. R. and Humphreys E. R.: Vegetation Development and Variation in Near-Surface Ground Temperatures at Illisarvik, Western Arctic Coast, Cold Regions Engineering 2019, 687–695, doi:10.1061/9780784482599.079, 2019.

Zhang, Y., Sachs, T., Li, C. and Boike, J.: Upscaling methane fluxes from closed chambers to eddy covariance based on a permafrost biogeochemistry integrated model, Global Change Biology, 18(4), 1428–1440, doi:10.1111/j.1365-2486.2011.02587.x, 2012.

Zona, D., Oechel, W. C., Kochendorfer, J., Paw U, K. T., Salyuk, A. N., Olivas, P. C., Oberbauer, S. F. and Lipson, D. A.: Methane fluxes during the initiation of a large-scale water table manipulation experiment in the Alaskan Arctic tundra, Glob. Biogeochem. Cycle, 23, GB2013, doi:10.1029/2009GB003487, 2009.

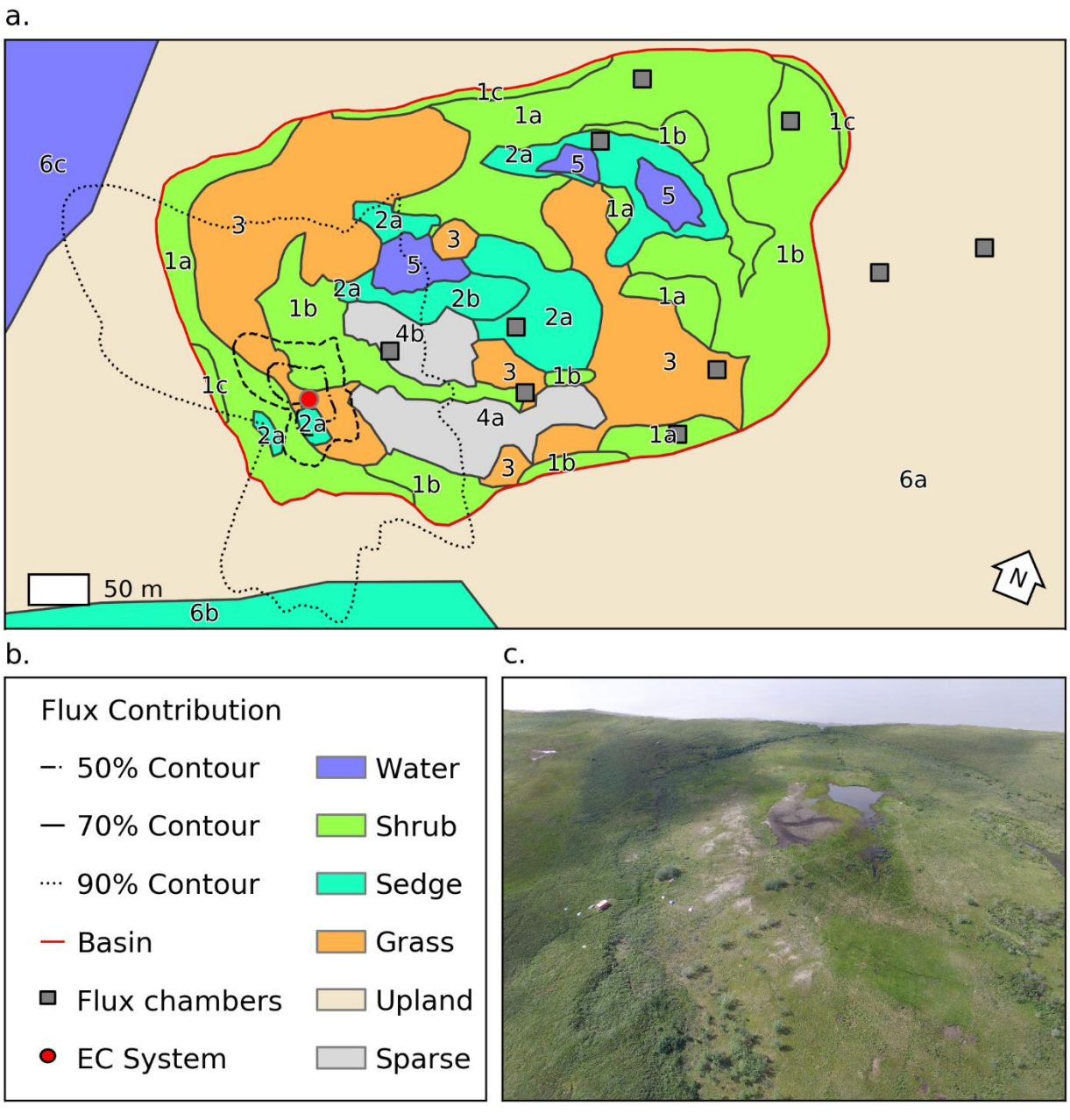
Zona, D., Oechel, W. C., Peterson, K. M., Clements, R. J., U, K. T. P. and Ustin, S. L.: Characterization of the carbon fluxes of a vegetated drained lake basin chronosequence on the Alaskan Arctic Coastal Plain, Glob. Change Biol., 16(6), 1870–1882, doi:10.1111/j.1365-2486.2009.02107.x, 2010.

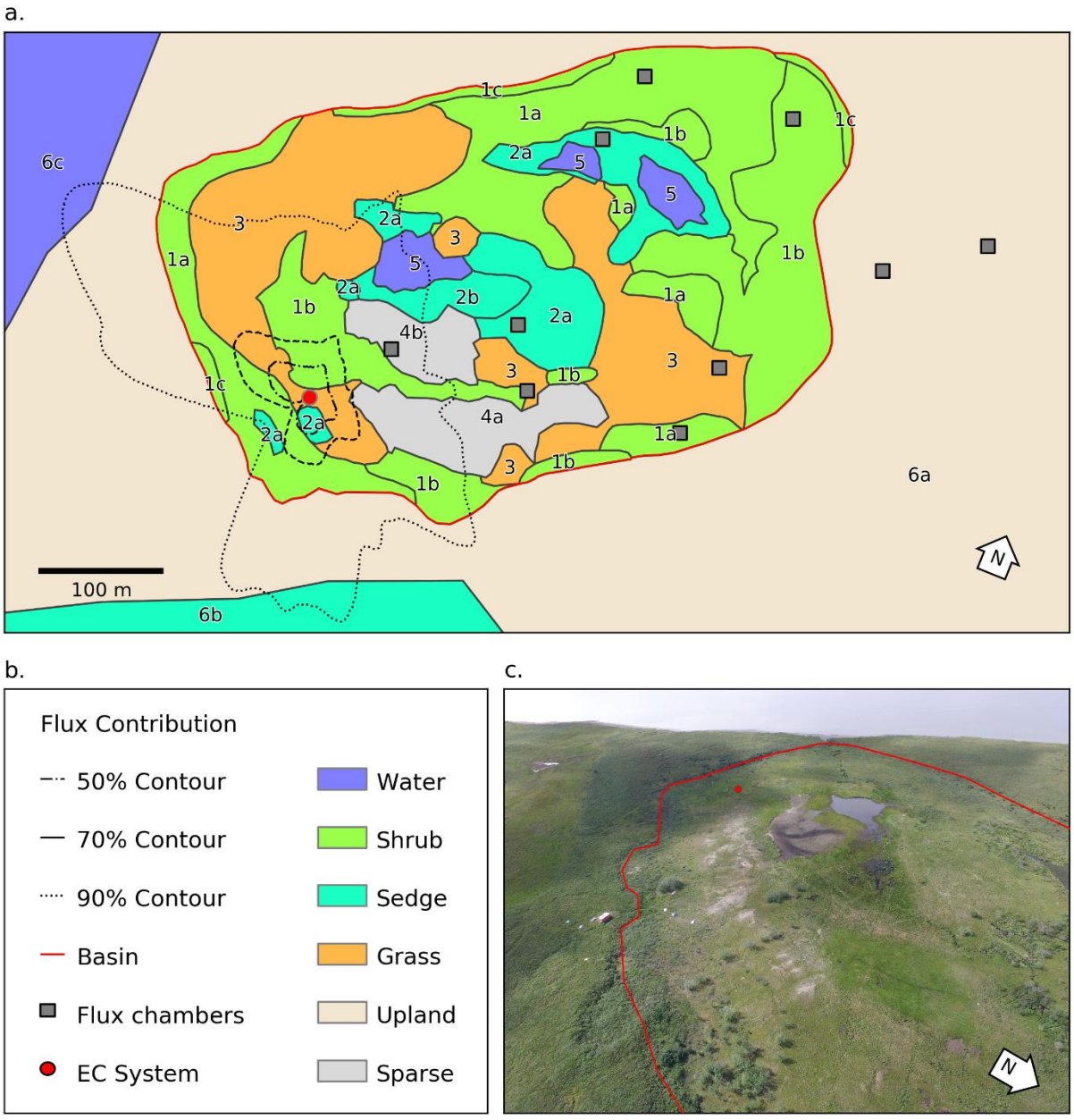
Zona, D., Lipson, D. A., Paw U, K. T., Oberbauer, S. F., Olivas, P., Gioli, B. and Oechel, W. C.: Increased CO2 loss from vegetated drained lake tundra ecosystems due to flooding, Glob. Biogeochem. Cycle, 26, GB2004, doi:10.1029/2011GB004037, 2012.

Zona, D., Gioli, B., Commane, R., Lindaas, J., Wofsy, S. C., Miller, C. E., Dinardo, S. J., Dengel, S., Sweeney, C., Karion, A., Chang, R. Y.-W., Henderson, J. M., Murphy, P. C., Goodrich, J. P., Moreaux, V., Liljedahl, A., Watts, J. D., Kimball, J. S., Lipson, D. A. and Oechel, W. C.: Cold season emissions dominate the Arctic tundra methane budget, Proc. Natl. Acad. Sci. U.S.A., 113(1), 40–45, doi:10.1073/pnas.1516017113, 2016.

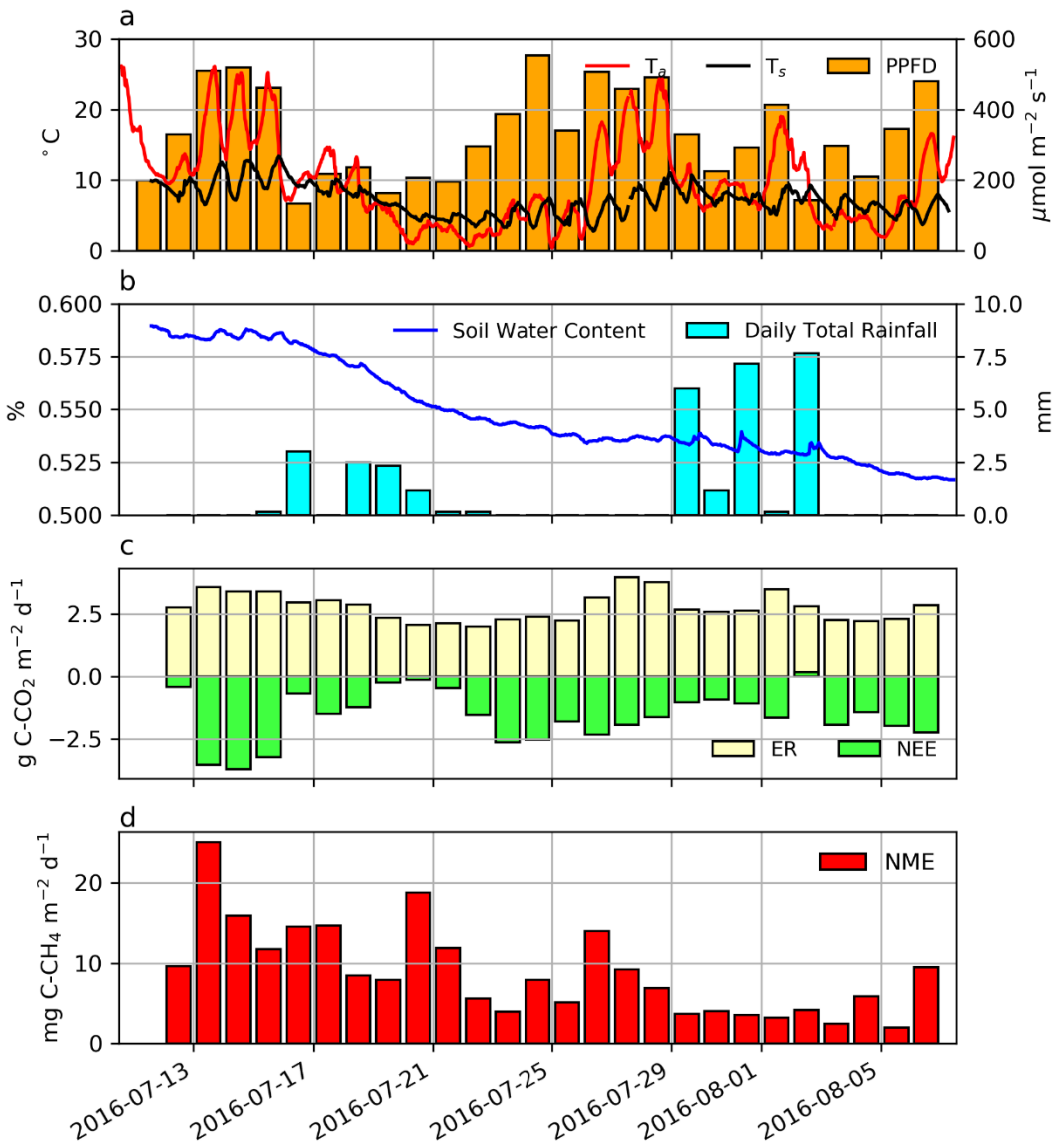
Zulueta, R. C., Oechel, W. C., Loescher, H. W., Lawrence, W. T. and Paw U, K. T.: Aircraft-derived regional scale CO2 fluxes from vegetated drained thaw-lake basins and interstitial tundra on the Arctic Coastal Plain of Alaska, Glob. Change Biol., 17(9), 2781–2802, doi:10.1111/j.1365-2486.2011.02433.x, 2011.

Figures & Tables





905 **Figure 1: a) Map of the distribution of vegetation classes at Illisarvik, with the footprint climatology (F_{Clim}) over the study period, the locations of the chambers and the eddy covariance (EC) system. The alphanumeric labels correspond to the unit codes in Table 1. b) Legend for the map in a. c) Oblique drone image of Illisarvik, take at 16:40 July 23rd 2016 view from E of DTLB towards W. The Basin and EC system are shown on the image using the same symbology as a).**


a
b
c
d
°C
T_a
T_s
PPFD
μmol m^{-2} s^{-1}
%
Soil Water Content
Daily Total Rainfall
mm
g C-CO$_2$ m^{-2} d^{-1}
ER
NEE
mg C-CH$_4$ m^{-2} d^{-1}
NME
2016-07-13
2016-07-17
2016-07-21
2016-07-25
2016-07-29
2016-08-01
2016-08-05

910

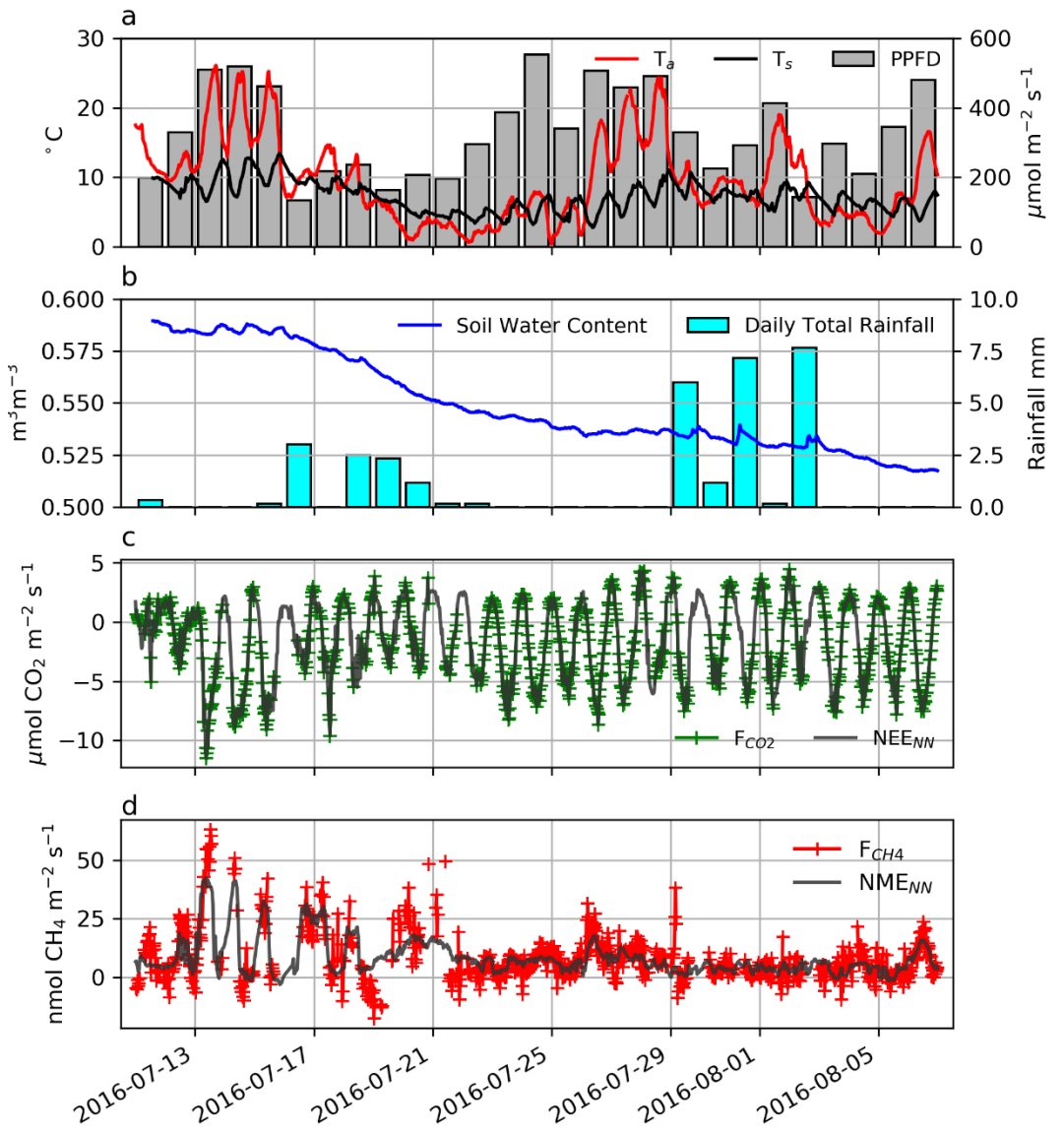
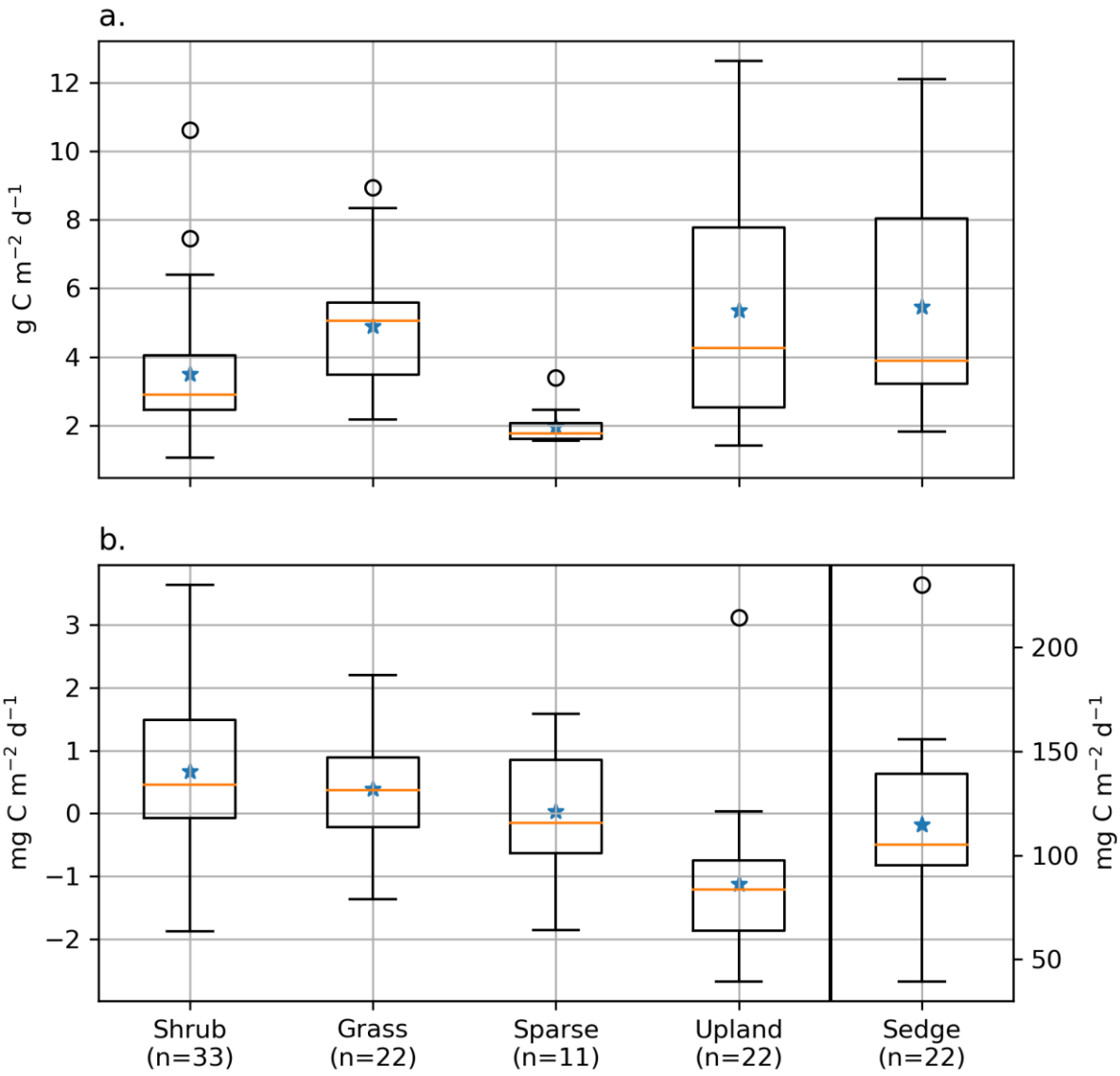


Figure 2: a) Half hourly air and soil temperatures, displayed along with photosynthetic photon flux density (PPFD). b) Hourly soil volumetric water content and daily total precipitation. ~~c)Gap-filled daily total NEE (dark~~ c) Half hourly F_{CO2} (green~~), ER~~) and NEE_{NN} (grey), and d) half hourly F_{CH4} (red) and ~~d) daily total NME.~~ NME_{NN} (grey).


a.
g C m^{-2} d^{-1}
12
10
8
6
4
2
b.
mg C m^{-2} d^{-1}
3
2
1
0
−1
−2
200
150
100
50
mg C m^{-2} d^{-1}
Shrub
(n=33)
Grass
(n=22)
Sparse
(n=11)
Upland
(n=22)
Sedge
(n=22)

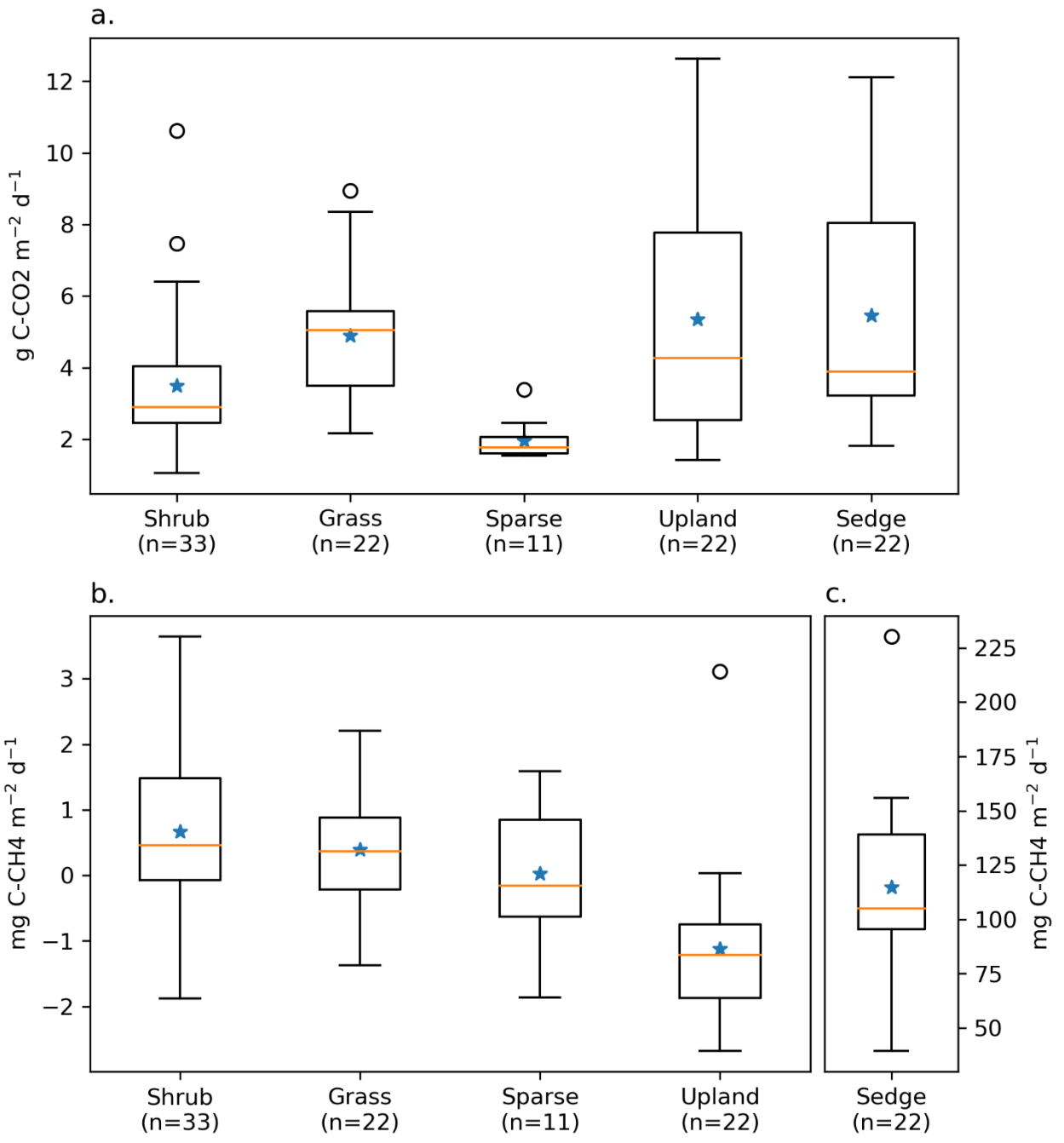
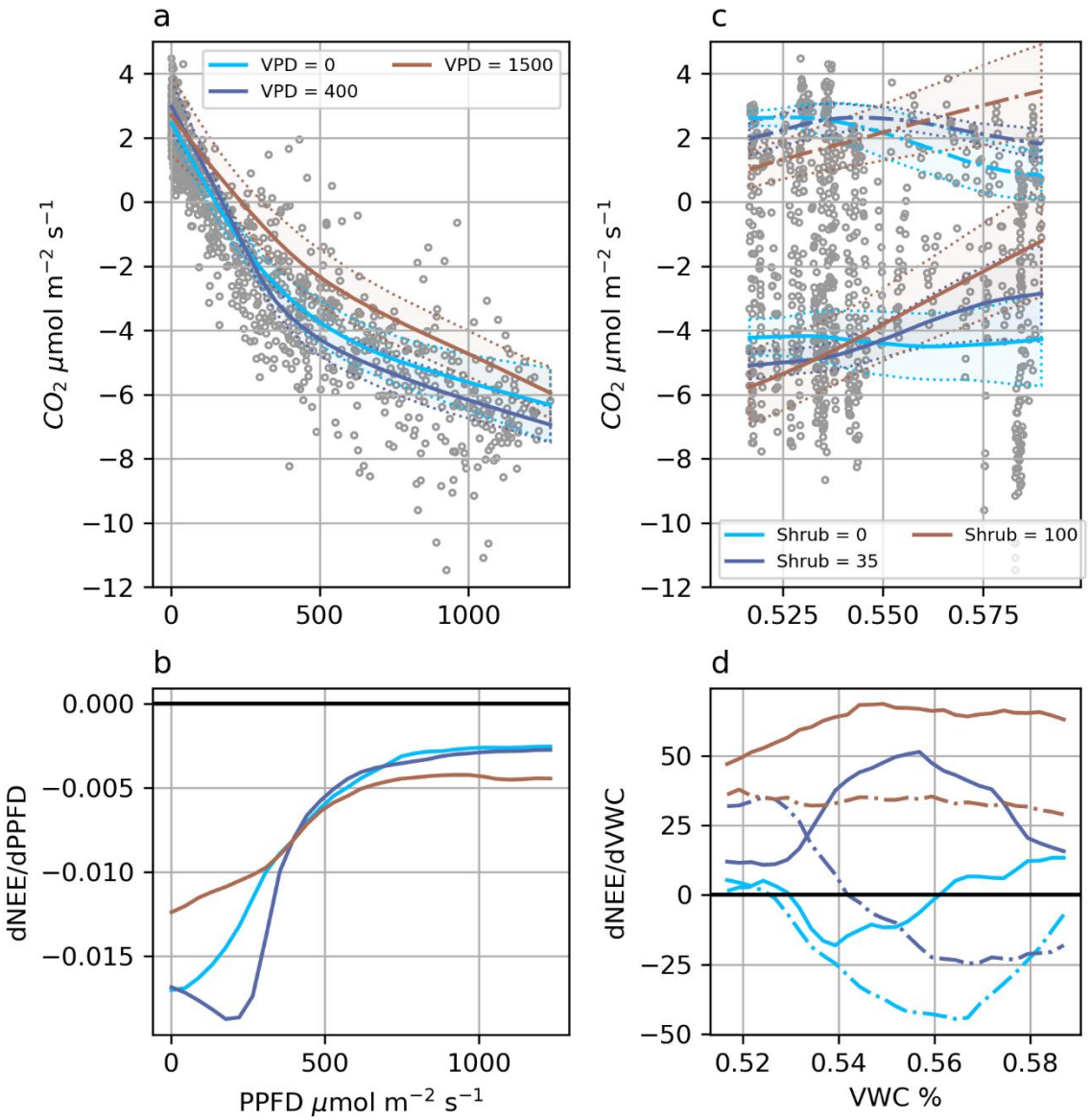


Figure 3: Boxplot of a) ER, b) NME and c) NME fluxes measured using closed chambers, grouped by vegetation class. The orange lines represent the median, blue stars represent means, the boxes indicate the interquartile range (Q3-Q3), the whiskers indicate Q1 –(1.5*IQR) and Q3+(1.5*IQR), and the circles represent outliers extending beyond the whiskers. Note the scale for c) Sedge ~~in b~~ is different.


a
c
b
d
VPD = 0
VPD = 400
VPD = 1500
Shrub = 0
Shrub = 35
Shrub = 100
CO_2 μmol m^{-2} s^{-1}
dNEE/dPPFD
dNEE/dVWC
PPFD μmol m^{-2} s^{-1}
VWC %

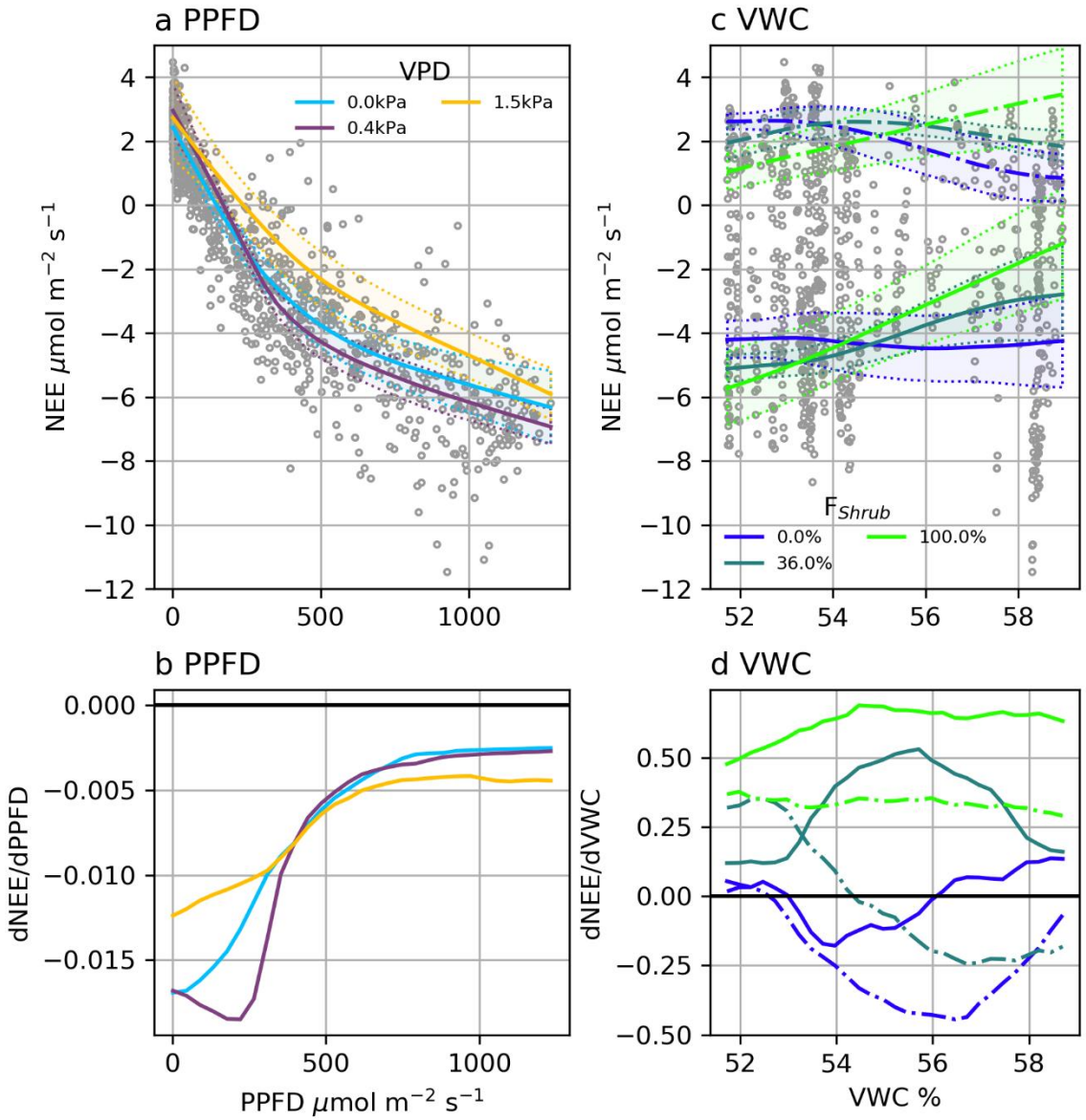
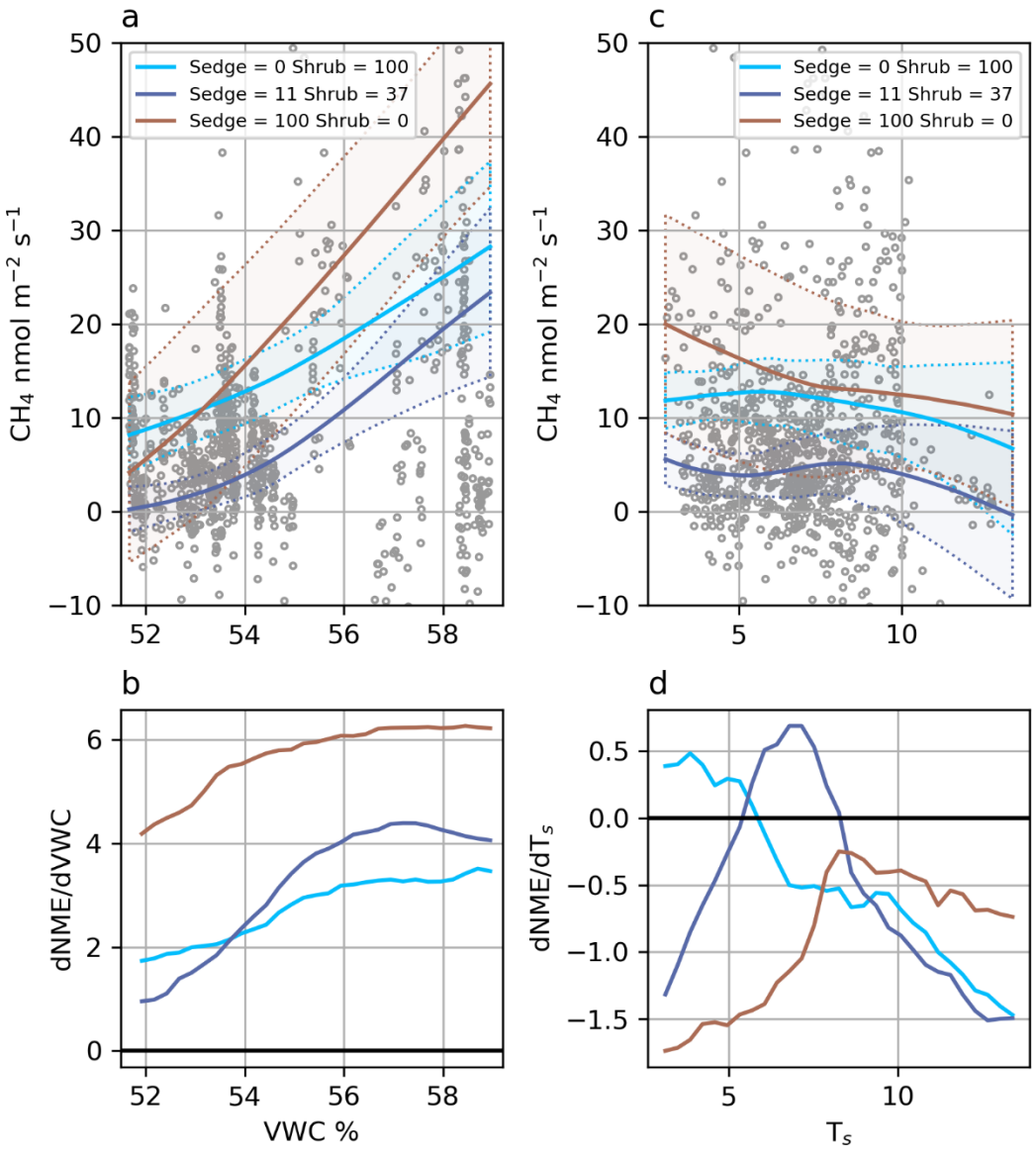
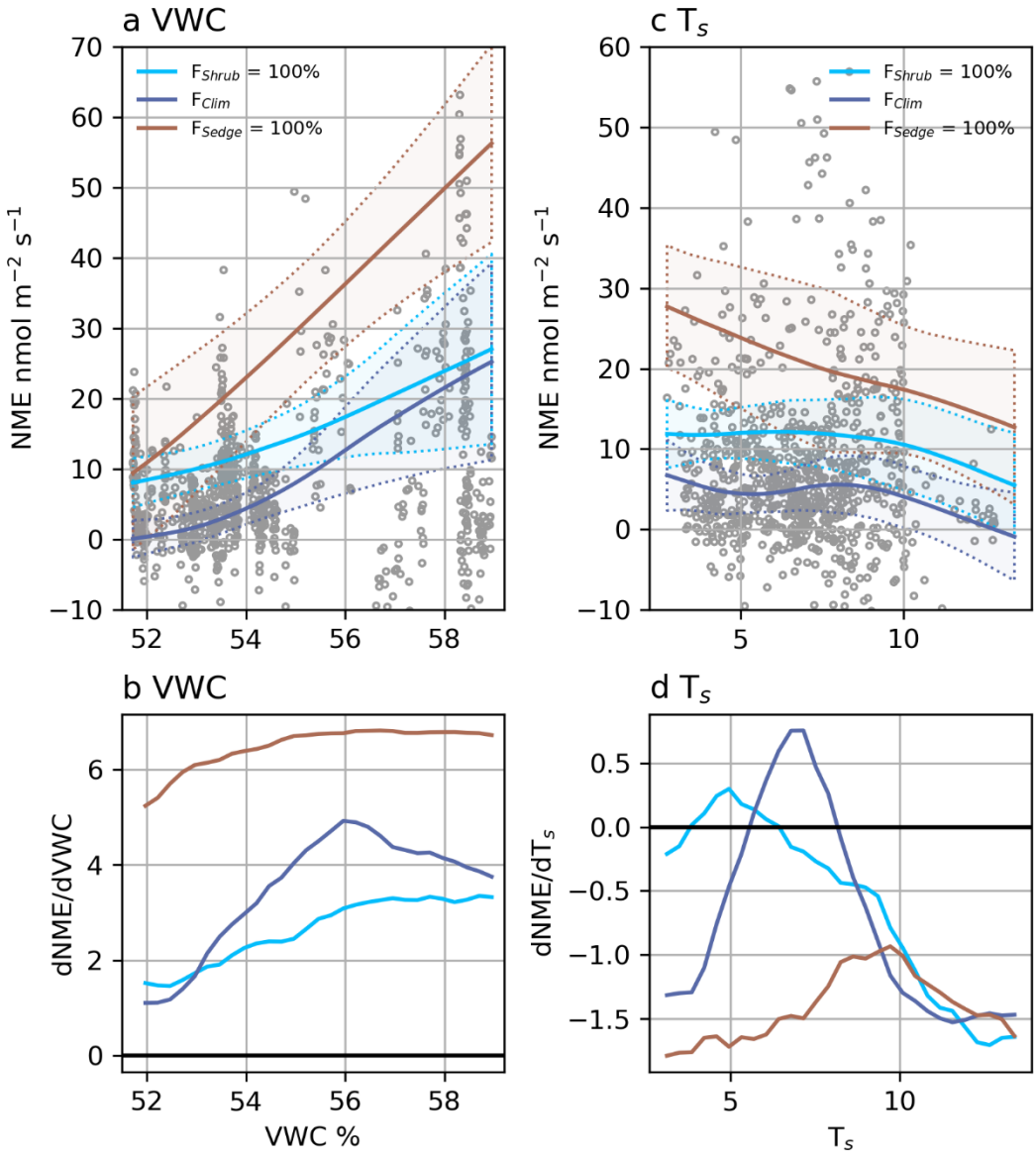
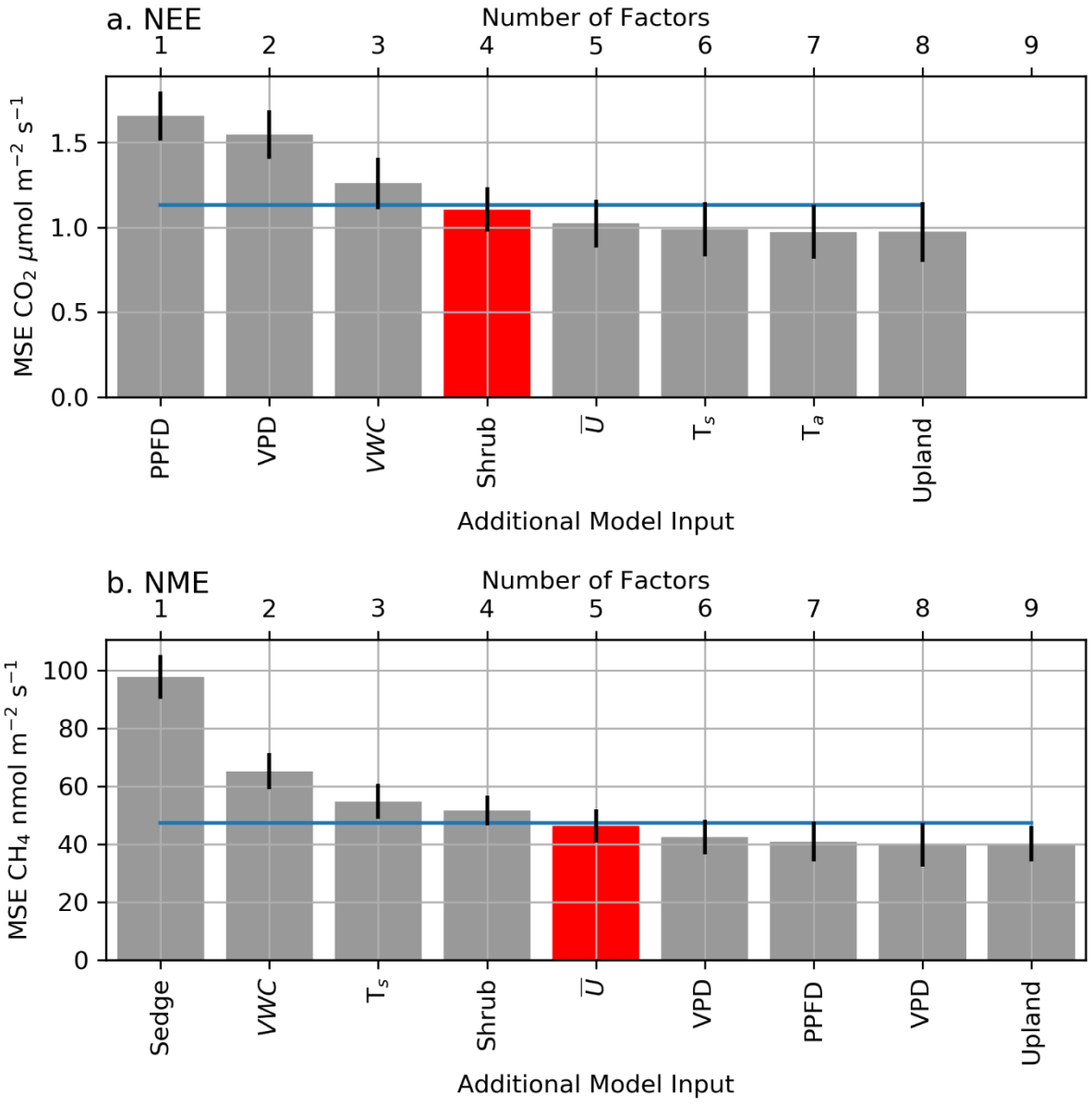


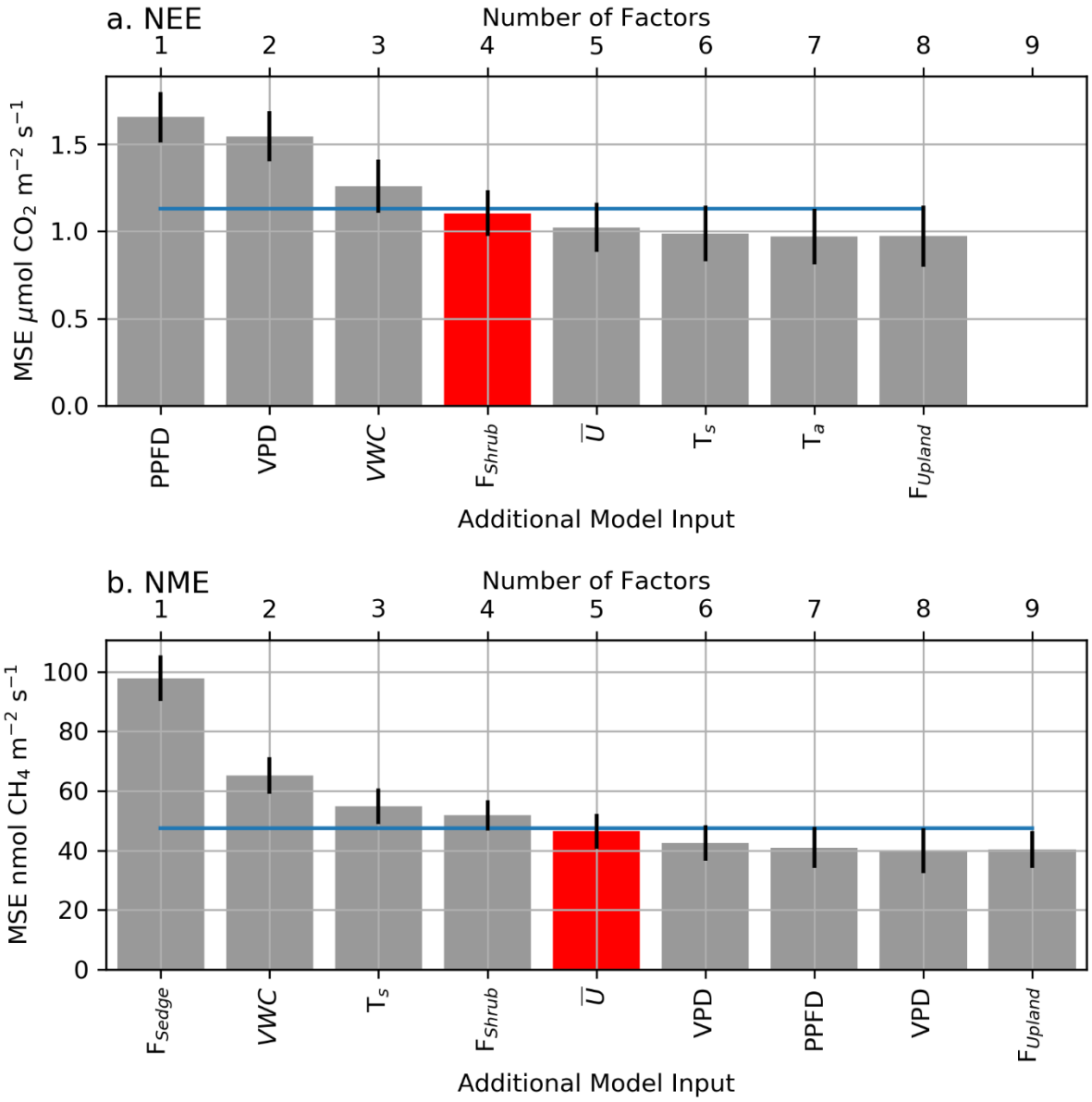
Figure 4: a. Modelled NEE response to PPFD under different VPD conditions and b. the partial first derivatives of NEE with respect to PPFD. c. Modelled ER (dashed line) and NEE (solid line) response to VWC at different Shrub$_{\%}$ and d. the partial first derivatives of ER (dashed lined) and NEE (solid line) with respect to VWC. NEE in c was calculated at PPFD = 600 ~~umol~~μmol m^2 s^{-1}. The shaded areas in a & c are 95% confidence intervals and grey circles are the EC observations.


a
c
b
d
Sedge = 0 Shrub = 100
Sedge = 11 Shrub = 37
Sedge = 100 Shrub = 0
CH4 nmol m−2 s−1
dNME/dVWC
dNME/dTs
VWC %
Ts



930 **Figure 5: a. Modelled NME response to VWC at different source area fractions and b. the partial first derivatives of NME with respect to VWC. c. Modelled NME response to T_s at different source area fractions and d. the partial first derivatives of NME with respect to T_s. The shaded areas in a & c are 95% confidence intervals and grey circles are the EC observations.**


a. NEE
Number of Factors
1
2
3
4
5
6
7
8
9
MSE CO_2 μmol m^{-2} s^{-1}
0.0
0.5
1.0
1.5
PPFD
VPD
VWC
Shrub
$\overline{U}$
T_s
T_a
Upland
Additional Model Input
b. NME
Number of Factors
1
2
3
4
5
6
7
8
9
MSE CH_4 nmol m^{-2} s^{-1}
0
20
40
60
80
100
Sedge
VWC
T_s
Shrub
$\overline{U}$
VPD
PPFD
VPD
Upland
Additional Model Input



935

Figure ~~A1~~A2: The averaged mean squared error (θ) of the bootstrapped Neural Network model validation datasets, with error bars showing one standard error (SE). The x axis shows models of increasing size from left to right (1-9 factors), and the label indicates the factor added to the model at each step. The blue line indicates the 1-SE rule threshold and the red bar indicates the model selected by the 1-SE rule.

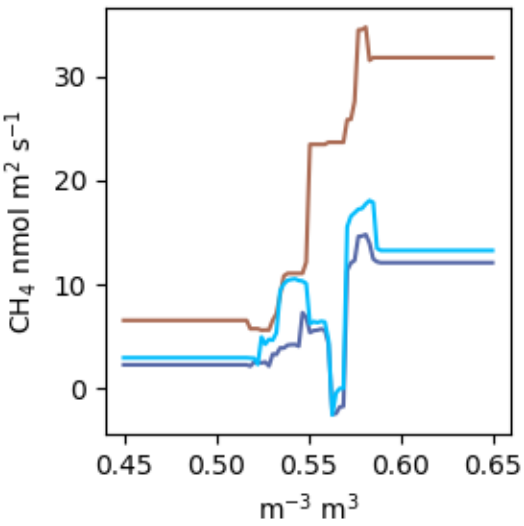


Figure A2: F_{CH4} estimated by a RF using the same factors as the NN model. The colours correspond to the scenarios in Fig 5a. *VWC* was estimated over the range 0.45 to 0.65.

Table 1: Dominant species or landscape feature within the vegetation/cover classes. Unit codes correspond to the map Figure 1a.

Unit Code	Vegetation Class	~~Vegetation~~Dominant Species/Landscape feature
1a	Shrub	*Salix alaxnesis* (Tall Willow)
1b	Shrub	*Salix* ~~spp.~~*glauca* (Low ~~Willows~~Willow)
1c	Shrub	*Alnus* ~~spp.~~*viridis* subsp. *crispa* (Alder)
2a	Sedge Marsh	*Carex* ~~spp.~~*aquatilis* (Sedge)
2b	Sedge Marsh	*Arctophila fulva* (Pendant Grass)
3	Grass Meadow	*Pocacea* spp. (~~Grass~~Grasses), *Eriophorum angustifolium* (Cotton Grass)
4a	Sparse Cover	Sparse Vegetation
4b	Sparse Cover	Bare Ground
5	Ponds	*Hippuris vulgaris* (Mare's Tail), Open Water
6a	Outside of Basin	~~Upland Tundra~~Dwarf shrub tundra: *Salix* spp. & *Betula nana* (Birch)
6b	Outside of Basin	Fen
6c	Outside of Basin	Ocean

~~**Table 1: Primary species and notable landscape feature present within the vegetation/cover classes.**~~

Table 2: The surface cover class fractions of the basin, along with the mean source area fractions of the footprint climatology (F_{Clim}) and the range of source area fractions for individual half hourly observations shown in brackets.

~~**- Unit codes correspond to the map Figure 1a.**~~

Surface Class	Basin	~~Footprint~~F_{Clim}
Shrub	48.3 %	36.0 % [0.0 – 79.0%]
Grass	27.9 %	39.0 % [1.1-78.1%]
Sedge	12.3 %	10.9 % [0.0 – 55.6%]
Sparse	8.4 %	2.2% [0.0 – 33.6%]
Water	3.1 %	0.2% [0.0 – 4.4%]
Upland	0%	6.2% [0.6 – 15.0%]
Outside Basin	0%	12.3% [0.2 - 28.0%]

~~**Table 2: The surface cover class fractions of the basin, along with the mean source area fractions of the footprint climatology and the range of source area fractions for individual observations shown in brackets.**~~

	~~Q_{10}~~	~~R_{10} μmol m^{-2} s^{-1}~~
~~Sedge~~	~~2.1~~	~~3.7~~
~~Upland~~	~~1.9~~	~~4.1~~

~~Grass~~	~~1.7~~	~~3.8~~
~~Shrub~~	~~1.7~~	~~2.7~~
~~Sparse~~	~~1.0~~	~~1.9~~
~~Night-time EC observations (n=100)~~	~~1.6~~	~~2.8~~

Table 3: The ER temperature sensitivity (Q_{10}) and base respiration (R_{10}) estimated by Laforce (2018) and estimated from nighttime EC footprint observations.

	Q_{10}	R_{10} µmol m^{-2} s^{-1}	R^2
Sedge	2.1	3.8	0.82
Upland	1.9	4.1	0.55
Grass	1.6	4.0	0.55
Shrub	1.8	2.7	0.46
Sparse	1.0	1.9	0.01
Night-time EC observations (n=100)	1.6	2.9	0.47

Table 4: Growing season (gs) daily range in eddy covariance-derived NEE and NME from drained thermokarst lake basins (DTLB) and other select wetland/coastal tundra sites across the Arctic. The period of studies measurements for the studies observations are: a) mid-June – end of July b) June 12 – August 28, 2007, Fig 4 c) June 11 – August 25, 2011 d) upscaled chamber estimates, exact dates not specified, e) mean June 15 –August 31 2003-2006, f) July 5 – Aug 4, 2009.

Site	Site Characteristics	NEE g C-CO_2 m^{-2} d^{-1}	NME mg C-CH_4 m^{-2} d^{-1}	Studies
Illisarvik	Young DTLB, Low & Tall Shrub/Grass/Wet Sedge	-1.5	8.7	(this study)
Various DTLB, Barrow Peninsula, Alaska	Young DTLB, Wet Sedge Tundra	-1.1[b], -0.9[d], -0.8[c]	18.4[a], 26.1[d], 44.0[c]	Zona et al. 2009[a] & 2010[b], Sturtevant and Oechel, 2013[c]; Lara et al. 2015[d]
	Medium DTLB, Wet Sedge Tundra	-0.7[b], -0.6[d], -0.4[c]	27.0[d], 41.3[c]	
	Old DTLB, Wet Sedge Tundra	-1.0[b], -0.4[d], 0.1[c]	24.2[d], 38.7[c]	
	Ancient DTLB, Wet Sedge Tundra	0.4[d]	21.7[d]	
Katyky, Indigirka lowlands, Siberia	Ancient DTLB, Dwarf-Shrub and Wet Sedge Tundra	-1.3[e]	36.0[f]	Van der Molen et al. 2007[e], Budishchev et al. 2014[f]