Response to reviewers:

We would like to thank both reviewers for their comments. We feel that these comments (shown in bold), and the changes made in response to these suggestions, have improved the manuscript. Our point by point response to each comment/suggestion is embedded below.

LM Merbold (Referee # 1)

The manuscript by Watts et al. provides a consistent overview of carbon flux (CO2 and CH4) measurements in the arctic and evaluated the potential of the Terrestrial Carbon Flux model to represent site specific flux measurements at different time scales. Up to date few studies integrating both flux measurements and modeling approaches to determine the net ecosystem carbon balance of arctic ecosystems exist. Therefore the presented results contribute to the currently available knowledge on CO2 and CH4 exchange in the Arctic and also provides an outlook for ongoing and future research on this important topic. The authors provide a very well structured and analyzed manuscript with an excessive material and methods paragraph (which seems appropriate for this integrative study) and a sometimes lengthy to read results section. I have only minor technical/structural comments and recommend this paper to be published in Biogeosciences.

As stated before I recommend to restructure the Results paragraph since the amount of information (numbers and abbreviations) make it rather difficult to follow. Therefore one step could either be moving some results in an additional table and state only the most important results.

Response: As suggested by the reviewer, we have reduced the amount of information presented in the results.

As a second recommendation I would like to encourage the authors to be consistent with naming: e.g. TCF model instead of TCF only- please adjust this throughout the manuscript.

Response: We apologize for this, and have modified the manuscript accordingly.

P16493, l9: TCF model simulations

Response: Corrected.

P16495, l1: I would argue that this is not even a network yet, unfortunately more likely a dozen of sites that are still or have been active.

Response: We agree with the reviewer, and have changed this to read, “the scarcity of in-situ observations.”

P16495, l15: remove NECB

Response: We have removed NECB.

P16496, l6-9: This is correct, but why does the labile carbon increase CH4 production? An additional sentence on the process could be useful.

Response: We have added that the labile carbon fractions require shorter decomposition pathways to produce the substrates most commonly used for methanogenesis (i.e. acetate and $H_2 + CO_2$).

P16496, l35: This is nice to read but is this actually needed in this manuscript, I suggest to delete this.

Response: We have deleted this sentence.

P16498, l6: is -> are

Response: We have made the suggested change.

P16498, l11-13: name an example

Response: Following the MOD17 GPP approach described in Running *et al* (2004), minimum daily air temperature and vapor pressure deficit are the primary environmental conditions used to constrain photoassimilation in the light use efficiency (LUE) model. We have clarified this in the manuscript.

Running, S. W., Nemani, R. R., Heinsch, F. A., Zhao, M., Reeves, M., and Hashimoto, H.: A continuous satellite-derived measure of global terrestrial primary production, BioScience, 54, 547–560, 2004.

P16498, l14-16: try to combine the information with the previous sentence.

Response: We have combined these sentences (i.e. l11-16) as suggested.

P16498, l17: Why? You should guide the reader towards an easy understanding, e.g. if the reader is a non ecologist or non-firm in arctic ecosystems one would not understand.

Response: We have included additional text to explain that the sensitivity of bryophytes to surface drying is primarily due to the lack of vascular tissues necessary for water transport. We have avoided a detailed explanation of the physiological mechanisms by which water stress can limit plant metabolism and CO_2 assimilation, as this is outside of the scope of this paper.

P16499, l19: recalcitrant? But you have a separate pool dealing with this.

Response: The use of “recalcitrant” here was unintentional, and we have removed it. We thank the reviewer for observing this.

P16500, l6-11: this is unclear to me, I see the need for this but cannot follow the explanation of your multipliers- please improve.

Response: We have modified this sentence for greater clarity.

P16502, l20-21: F = Flux of what since F has been used before for CO2, this might be confusing.

Response: We have changed F_{total} to F_{CH4} to avoid confusion.

P16509, l26: the word "where" seems not to fit in here.

Response: We have adjusted this sentence accordingly.

P16513, l17-19: I disagree. The uncertainty can not be directly attributed to the partitioning of EC measured NEE into GPP and Reco. Certainly there are severe differences in flux partitioning, but one could try different approaches which commonly result in a range of GPP and Reco. This range could be compared to the TCF model estimates.

Response: We agree with the reviewer. A separate study is needed to thoroughly investigate the range of GPP and Reco that might result from the use of different flux partitioning approaches. We have added an additional sentence in the manuscript to address this.

P16515: I guess the larger discrepancies of the Barrow results are basically caused by the experimental manipulation.

Response: Yes, it is possible that the modification of site conditions during the water table manipulation experiment at Barrow was not adequately reflected in the model simulations. We have added a sentence in the manuscript to address this.

P16515, l28: define similar, 20%, 80%? Unclear.

Response: We have added additional information here to clarify this.

Figure 4: very difficult to see, this is either due to the typesetting at BGD. If not, please enlarge to the full page.

Response: Unfortunately, this resulted from the typesetting. We will work with the publisher to correct this is.

Figure 8: Please include the years in the bars, since this is otherwise difficult to see within the figure.

Response: We have included the years in the figure, as suggested.

Anonymous Referee #2

Due to huge carbon stock in the northern high latitude region, it is very important to get clear understanding of net ecosystem carbon balance (NECB). In this manuscript, the authors use a terrestrial carbon flux model to analyze NECB at six sites by using two sets of input data, and compare the simulation results with field observation through eddy covariance. Also, one merit of this study is to quantify the global warming potential by considering carbon dioxide and methane together. Overall, this manuscript is well organized and clearly stated.

Comments are given below for the authors and editors to consider.

First, to address the relative uncertainty introduced by using both remote sensing data and in-situ data, it is good to analyze the difference of input data itself, which is more straightforward to illustrate the point. Meanwhile, the uncertainty caused by input data is a restriction not only for site-level study, but also regional estimation. If the authors could provide deeper exploration of how the current results have an implication for the uncertainty of upscaling estimation, that could be more interesting.

Response: In this analysis, we used NASA GMAO (GEOS-5) MERRA reanalysis to drive the TCF model simulations. The MERRA input meteorology are currently being incorporated into global TCF model CO_2 simulations used to inform algorithm development for the upcoming NASA Soil Moisture Active Passive mission Level 4 carbon products, and have been evaluated against satellite, in-situ, and other reanalysis (GEOS-4) records (see Yi *et al*, 2011 in the manuscript).

We fully agree that a deeper exploration of uncertainty in the reanalysis data is needed, especially for the northern regions, and we highlight observed differences between the reanalysis data and in-situ measurements in our manuscript. Although outside the scope of this study, we intend to further investigate these uncertainties in a subsequent analysis that will address the potential implications for regional modeling and up-scaling of CO_2 and CH_4 fluxes.

Yi, Y., Kimball, J. S., Jones, L. A, Reichle, R. H., Nemani, R., and Margolis, H. A.: Recent climate and fire disturbance impacts on boreal and arctic ecosystem productivity estimated using a satellite-based terrestrial carbon flux model, J. Geophys. Res.-Biogeo., 118, 1–17, 2013.

Second, it seems the model simulation could catch the temporal patterns of different variables of the field observation in most sites. But still, there are some mismatch between model simulation and site observation in some sites (for example, KY 2009 in Fig. 2 and Fig. 4). It is better to discuss the potential reason for why there are great differences between them.

Response: For Kytalyk (KY 2009), the differences between the TCF model and EC derived GPP estimates (Figure 2) early in the growing season are attributed to the limited ability of the light use efficiency approach to account for vegetation community specific (e.g. *Betula nana* and *Salix pulchra*) phenology. To investigate this, we ran an additional simulation (described on P16512 in the manuscript) using a temperature driven phenology model (see Parmentier *et al* 2011) to account for the limited early season bud break observed at this site due to a 2-4 week delay in active layer deepening following spring warming and snowmelt. This step reduced the model RMSE by 56%, as we report in the manuscript.

At Kytalyk, the TCF model simulations overestimated ecosystem respiration (R_{eco}) during the growing season relative to the EC CO_2 records (Figure 4). Although some of the positive bias results from the higher TCF model GPP estimates, it may also reflect higher vegetation carbon use efficiency (CUE; i.e. NPP/GPP) at this site and seasonally variations in sub-surface soil thermal conditions that were not reflected in the TCF model simulations. We found that using in-situ soil temperature from an 8-cm sampling depth (instead of the 5 cm depth) reduced the TCF model overestimation of EC based R_{eco} by 8%. We also found that increasing the CUE parameter from 0.55 to 0.65 further reduced the overestimation of cumulative R_{eco} by 36% at this site; this may reflect lower autotrophic losses from shrub communities dominant in drier portions of the landscape (Parmentier *et al* 2011). Although this higher CUE is within the range reported for shrub vegetation, we did not use this parameter value in the model simulations as it falls further from the reported mean (0.55; Choudhury, 2000). We have added a sentence in the manuscript to address the resulting CUE based differences in R_{eco}. We intend to further investigate the range of expected CUE values for Arctic-specific plant functional groups in a subsequent study; this step will allow us to develop posterior probability distributions for the TCF model CO_2 parameters, using additional EC CO_2 data available in the FLUXNET database.

Choudhury, B. J.: Carbon use efficiency, and net primary productivity of terrestrial vegetation. Adv. Space Res., 26, 1105-1108, 2000.

Parmentier, F. J. W., van der Molen, M. K., van Huissteden, J., Karsanaev, S. A., Kononov, A. V., Suzdalov, D. A., Maximov, T. C., and Dolman, A. J.: Longer growing seasons do not increase net carbon uptake in the northeastern Siberian tundra, J. Geophys. Res.-Biogeo., 116, G4, doi:10.1029/2011JG001653, 2011.

Third, I notice that your parameter values (in the supplementary material) for different sites are different even they are same biomes, e.g. Ro, Q10p etc. How did you determine the value of those parameters? Did you calibrate the model? If so, which set of data you are using for calibration and which set of data are using for model validation.

Response: Our model parameter values were based on published literature and a prior biome-based calibration of the TCF model for GPP and R_{eco} using a Markov Chain Monte Carlo (MCMC) approach (described in Yi *et al* 2013). We then performed a site based optimization using the EC data to provide a balance between RMSE and the temporal correspondence between flux observations and model estimates. We plan to conduct a regional adaptive MCMC based TCF model CO_2 and CH_4 calibration/validation in a subsequent study, which will include additional EC CO_2 and CH_4 information. This process with allow us to assign posterior probability distributions to the TCF model parameter values using a Bayesian framework, and should help to better constrain estimate uncertainties for regional simulations.

Yi, Y., Kimball, J. S., Jones, L. A, Reichle, R. H., Nemani, R., and Margolis, H. A.: Recent climate and fire disturbance impacts on boreal and arctic ecosystem productivity estimated using a satellite-based terrestrial carbon flux model, JGR-Biogeosciences, 118, 1-17, 2013.

Fourth, this study is to focus on northern peatland and tundra carbon dioxide and methane flux. Only two sites have the observation beyond the growing season. It would be more interesting to have more sites which consider the greenhouse gas fluxes during the spring-thaw period as well. Some recent studies indicates that the methane flux during spring thaw period could be even larger than the growing season.

Response: Unfortunately, as was noted by the first reviewer, the availability of EC CH_4 data for the high latitude regions is rather limited. For this study, we were fortunate to pool together datasets from six tower sites. We recognize the need for observations that span beyond the growing season, and are optimistic that these will be made increasingly available with the publication of newly collected data and the installation of additional EC sites in Arctic regions. We advocate for the development of a pan-Arctic EC network to simultaneously measure CO_2 and CH_4 fluxes throughout the summer and spring/fall shoulder seasons over a multi-year period. The availability of these data, provided in a comprehensive format similar to that used by the FLUXNET CO_2 community, will greatly benefit future regional and global modeling studies by providing up-to-date EC information in a standardized manner that may help to better characterize spatial and temporal variability in carbon fluxes across the high latitudes, and better quantify the range of variability in the flux data resulting from differences in data processing/partitioning methods.

We acknowledge the necessity of an EC network capable of measuring CO_2 and CH_4 in our manuscript (P16523), and we have added additional text to also stress the need for flux observations over the spring thaw period.