We thank Referee 1 for their help comments (in italics below) and provide responses inline.

This manuscript presents methane flux measurements at a temperate forest in Canada by using the eddy covariance (EC) method. Since measurements of methane flux at forest ecosystems by the EC is currently a few, the manuscript potentially contributes the FLUXNET community who will measure methane fluxes in future. Totally, the study does a good job to provide example for the eddy covariance measurement and to analyze environmental regulations over the methane fluxes. However, some points, mainly for uncertainty issues, are needed to be addressed before the publication.

General comments: Presented detection limit in Fig. 1 and Fig. S5 is precision of the system rather than the detection limit. Since null fluxes were calculated from actual vertical wind speed and methane concentration at the calibration, examined errors were distributed around zero value as shown in Fig. S5. In the actual case, however, those errors must be distributed around true fluxes; consequently, the precision must be drawn as error bars centered on each flux rather than on zero value in Fig. 1; if error bar crosses the zero value, the data is considered to be statistically zero. Please redraw the figure and show what percent of observed data did statistically differ to zero for daytime and nighttime.

This is correct, and we will specify in text and in the figures that the determined precision of the system or "contribution of noise to the flux signal" is distributed around the true fluxes rather than zero.

Precision of the gas analyzer must be presented as the Allan variance (Allan, 1966) as well as time-series data as shown in Fig. S3, because time-series data cannot provide direct measure of the precision. Since the low flow setting of the analyzer cannot be available for the eddy covariance measurement, data obtained at the high flow setting must be shown also. Such information is necessary in measurements in forests, because range of the actual fluxes was near the detection limit.

An Allan variance of the system operated at high flow using a calibration gas for 30 minutes will be included in the revised manuscript.

Conpectra presented in Fig. S4 was results from ensemble mean. Although the ensemble provide a general characteristic, standard error in each data point is also necessarily, which can provide information how the general characteristics were hold in each run. In addition, please explain how you determine the cutoff frequency of 0.2Hz.

The cutoff frequency of 0.2 Hz was determined by visually inspecting the averaged ensemble of select cospectra. The normalized cospectra using CH_4 began to significantly deviate from the normalized cospectra using sonic temperature at frequencies above 0.2 Hz. Because the deviations at higher frequencies were generally positive, this was interpreted as a contribution of noise to the CH_4 flux signal. The purpose of this analysis was to assess a potential cutoff where noise begins to dominate the higher frequency portion of the cospectrum. Therefore, the time periods chosen for the ensemble averaging were those with 'well-behaved' individual cospectra that could be compared to the cospectra for other scalars. This will be clarified in the revised

manuscript. Adding standard deviation as error bars to the averaged cospectrum would be difficult given the normalizations and processing that occurs in producing a cospectrum, and would not contribute to easier interpretation of the figure.

When you compare mean values, please add the standard error in order to understand that differences are statistically significant. For example, in page 17753 lines from 5 – 14, you compared the values from the EC and chambers, but standard error was not shown. In page 17756 line 5, you compared fluxes from the EC and BML methods, but errors in each method were not shown.

A standard error will be included in fluxes where applicable. When comparing between EC fluxes and fluxes calculated from chamber measurements, the standard error can be misleading. For the EC fluxes, the standard error represents the temporal variability in the tower-level fluxes. For chamber fluxes, the standard error represents the variation across the sites spatially. This was described in the Fig. 7 caption, and will be clarified in the text in the revised manuscript.

Specific issues:
Page 17750 line 20-22: Please show the diurnal variations obtained at chamber measurements in the results section.

Information about the diurnal variations in the chamber data is from an intensive sampling campaign on a single day in August (measurements at 06:00, 12:00, 18:00 and 24:00). There was no clear pattern in the diurnal cycle on this day, but a figure can be included in the supplementary information, if the reviewers find it useful.

Page 17755 Line 4: "longer timescale" is ambiguous. Please be concrete.

It will be clarified in the revised manuscript as a trend between fluxes and the soil moisture that showed a stronger relation over the measurement period (e.g. seasonally) rather than with diurnal variations.

Page 17760: If substrate limitation was significant, relationship between methane concentration and flux may be obvious. Please examine the relationship.

The variability in the methane concentration at the EC sensor height was very small, and has no discernible influence on the magnitude or direction of the fluxes. The relationship between methane concentrations and fluxes in the chamber data will be explored in a forthcoming manuscript by Winsborough *et al.*

Fig. 2 – 6: The pried for the analysis used in the figure must be shown in the captions.

Data for the entire measurement period was used in these figures, and that information will be clarified in the revised manuscript.

Fig. 6: Please show p-value for the slope.

Because of the large number of data points, the p-value for each panel in Figure 6 is less than $1x10^{-6}$, therefore we feel the uncertainty in the slope is a better measure of the significance of the relationship. In the Fig 6 caption, we will specify that the p-value is less than 0.001 for each panel.