Dear Referee two. We thank you for the constructive and detailed comments and suggestions. We agree with most of the comments and have made some further analysis to improve the study. We have made most of the required updates to the manuscript.

Specific comments:

In several places the approach seems to take too simplistic a view, without properly discussing the assumptions or their impact. For example, a key result presented is the difference between GPP derived from EC measurements in an outbreak year (2012) compared to five other years without insect utbreaks. Unfortunately, it seems EC data were only available for one outbreak year, but there is no analysis of differences between years due to factors other than insect damage. Inter-annual variability in meteorological conditions (rainfall/soil moisture, solar radiation and temperature in particular) can result in different annual total GPP. The values given in Table 1 should therefore be analysed with respect to meteorological conditions. This should also allow the 2012 value to be given some context – if the insect outbreak had not occurred, would the 2012 total be lower/higher/similar to the average based on meteorological conditions alone? Furthermore, is it possible that the outbreaks in 2012 and 2013 contributed to the lower GPP obtained in 2014, or is this attributable to meteorological conditions?

Response: We agree that these important considerations need to be discussed, and we have consequently studied relationships between meteorological conditions and GPP as well as $NDVI_{DL}$. One important note here is that the annual GPP values given in Table 1 (p. 15 in manuscript) were not displayed in the correct order for the years 2009, 2010, 2011 and 2014. The correct figures in Table 1 are:

	Years without insect outbreak					**Outbreak**
Year	**2007**	**2009**	**2010**	**2011**	**2014**	**2012**
GPP (g C m^{-2} yr^{-1})	451	531	373	401	448	180

where annual GPP was largest in 2007 and 2009, and the lower annual GPP in 2010 and 2011 indicates that there were minor defoliation already in these years.

Due to the limited number of years with EC derived GPP we did not consider it reliable to study correlations between annual GPP measured at the EC tower and meteorological variables. Instead we modeled GPP for the birch forest around the tower with PAR, and compared EC derived and PAR-modeled GPP. The comparison suggests that in the two years (2010 and 2011) prior to the outbreak, measured GPP was lower than PAR-modeled GPP, indicating that there were signs of defoliation by growing larval densities. Also in 2014 when the birch forest most likely was recovering from the outbreak, measured GPP was lower than PAR-modeled GPP. For the earlier years (2007 and 2009) when the birch forest was likely closer to undisturbed conditions, EC derived GPP and GPP modeled with PAR data agreed well; measured GPP was slightly lower compared to PAR-modeled in 2007 and slightly larger in 2009. For the outbreak year 2012 the difference between EC derived GPP and PAR-modeled GPP was 286 g C m^{-2} yr^{-1}, which is similar to the decrease of 261 g C m^{-2} yr^{-1} we found in our study. In addition, we ran the LUE model with meteorological data from the scientific research station in Abisko (ANS) for the year 2008 to fill the gap in the time-series with measured GPP and to study how well it agreed with the years 2007 and 2009. According to the LUE model the annual GPP at the EC tower was 440 g C m^{-2} yr^{-1} in 2008. This indicates that the GPP for undisturbed years of 441 g C m^{-2} yr^{-1} that we used is reasonable.

However, with data from the EC tower available for more years it would be a potentially important improvement to include meteorological data when estimating the decrease in annual GPP. We have added a section about these results in the discussion and the figure showing EC derived and PAR-modeled GPP was added to the supplementary material.

We also studied correlations between NDVI and meteorological data available from ANS, where we used mean of the highest seasonal $NDVI_{DL}$ value derived from 200 MODIS pixels with birch forest. To minimize the influence of insect induced defoliation we excluded the outbreak years and years immediately prior to and after outbreaks. No linear relationships between PAR and GPP were found. There were, however, negative correlations between temperature and $NDVI_{DL}$, with the strongest correlation between $NDVI_{DL}$ and the mean temperature in May-June. The influence of the temperature on NDVI was however, weak and due the large estimated uncertainties of the LUE model (30%) we did not include these correlations in the analysis. We do, however, mention these results in the discussion but due to the limited amount of data we do not further elaborate on the results as that would be speculation.

The comparison with the 2004 results of Heliasz et al. (2011) on page 20 is useful and indicates that closer analysis of the temporal evolution of the EC data may be beneficial. Currently data are separated into years with and without insect outbreaks. During those years with outbreaks, does the reduction in GPP over the course of the growing season agree with the timing of insect population growth/insect damage? Is this also supported by the NDVI data? In the years classified as being without insect outbreaks, are there any effects of (albeit smaller) insect populations on the GPP or NDVI values? Perhaps such analyses could offer further insight into the refoliation effect; it is currently hard to draw meaningful conclusions on this subject from the information given in the article. Some evidence to support the assumption that NDVIDL captures refoliation would also be useful; Fig 3 is not very convincing in this respect.

Response: For the year 2004, when we had EC data for the later part of the insect defoliation and the following refoliation at the EC tower, we can see that GPP is low during the defoliation events and increasing later in the growing season with the new leaves appearing. The raw NDVI values have a similar pattern with lower values (around 0.6) until early August when refoliation results in a late season peak in NDVI. This is illustrated in Figure 3 where raw NDVI has a similar seasonal development in 2013, when there was substantial refoliation around the EC tower: raw NDVI stayed around 0.6 during June, but increased to pre-outbreak levels in early July when refoliation occurred. In 2012 when there was no refoliation around the EC tower, raw NDVI stayed around 0.6 during the entire growing season. These different seasonal trajectories are utilized in the defoliation detection method (Olsson et al. 2016). $NDVI_{DL}$ does, however, not capture the typical trajectory for refoliation years with sharply increasing NDVI values when the growing season starts, that level off and start increasing again later in the season. The higher NDVI values in the later part of the growing season result in $NDVI_{DL}$ values that are higher than in 2012 but still lower than for years without defoliation (even though the actual timing of the defoliation is not captured in the LUE modeled GPP during years with refoliation). A new version of TIMESAT, currently being developed and tested, will capture also more detailed seasonal trajectories with smooth fitting of curves. These new curve fitting methods have a potential to improve the performance of the LUE model. We have added a section about our response to this comment to the discussion.

As mentioned in our previous comment we found influence on GPP at the EC tower the two years prior to the outbreak that are likely due to increasing insect population before reaching outbreak levels. This pattern is suggested in time-series of NDVI where there seems to be weak signs of defoliation 1-2 years prior to the outbreak.

P15, L6-10: Section 2.2.3 states that EC data were available from 1 May to 30 Sep, covering most of the growing season. Do the EC observations and values given in Table 1 agree with this timeframe, or is it possible GPP in Table 1 is underestimated if the growing season extended beyond these dates? This study focuses on GPP, but could the authors comment on other potential impacts of the insect outbreak on the carbon balance? For example, how might respiration rates be affected, and might this impact the partitioning into Reco and GPP?

Response: Budburst usually occurs early in June or very late in May so including data from first of May means that we capture the start of the growing season. For the years included in this study GPP was approaching zero by the last week of September implying that there were no underestimates of annual GPP.

Respiration is affected by insect outbreak in two ways: (1) Autotrophic respiration is reduced as defoliated trees cannot photosynthesize and (2) heterotrophic respiration increase when dead larvae decompose. The amount of carbon respired by larvae should be the same as the amount of carbon in eaten leaves so we should only observe a shift of respiration in time. In addition, larvae move nutrients from trees to fungi and bacteria living in soil which further increase respiration. The increase in heterotrophic respiration did not offset decrease in autotrophic respiration and R_{eco} for outbreak year was decreased in comparison to non-disturbed years.

On a similar note, many of the decisions taken in the presentation of results and development of the model rely on data collected during non-outbreak years. Is the gap-filling approach also suitable in defoliated years?

Response: The gap-filling approach is suitable also for defoliated years since the gap-filling function is created based on data from short time windows, usually seven days, even though the window can be longer if there are insufficient data for correct fitting. This short time window adjusts the fitting parameters for changing ecosystem conditions. We have clarified this in the manuscript.

P8, L10-11: More detail is needed about the quality control. Under which 'bad' weather/measuring conditions were data removed? How much data remained after quality control and what proportion was gap-filled?

Response: Data were removed mainly due to precipitation since an open path gas analyzer was used, or if the atmosphere was not fulfilling the turbulent conditions required for eddy covariance measurements. Available data after cleaning: 2007 61%; 2009 71%; 2010 66%; 2011 65%; 2012 58%; 2014 65%. We have clarified this in the manuscript.

P10, L11-2: Here, it is not clear which years have been used or why. The years for which the EC data are available should be stated in Section 2.2.3. Why was 2012 the only year used to calculate "max with insect defoliation? Why were data from 2008 and 2013 not included/not available?

Response: We have clarified that EC data from undisturbed years are from 2007, 2009, 2011, 2012 and 2014. In 2008 data were missing due to instrument failure. In 2013 the measurements were disturbed by larvae which unfortunately leave us with EC data from one year only.

P12, L13: It is not clear how these statistics were calculated and they don't seem to follow from Fig 4. Please provide more details/clarification

Response: We do agree that these statistics were not well described. Previously the statistics were based on the entire study area. We have changed the statistics to include only the pixels around the EC tower to correspond to the figure and clarified this. "The influence of observations with $NDVI_{DL}$ values < 0.4 and with f_{8day} > 0 was small. For the years with data available from the EC tower 8% of the eight day periods had $NDVI_{DL}$ < 0.4 and f_{8day} > 0 in the MODIS pixels surrounding the tower. For these time periods average f_{8day} was 0.068."

P14, Fig 6: The two green lines for NDVIDL show higher NDVI values early in the growing season for the defoliated year. Possible reasons for this should be investigated and the impact on the results commented on.

Response: Thanks for pointing this out. This is due to a weak fitting of the double logistic functions in TIMESAT. In a currently developed version of TIMESAT the fitting of the functions will be more robust. We have added a clarification in Section 3.3.1:"In Figure 6, $NDVI_{DL}$ has higher values in the year with defoliation compared to undisturbed years in May (period 16-18). These high $NDVI_{DL}$ values are due to poor fitting of the double logistic function during winter and early spring in 2012 (see Figure 3, where $NDVI_{DL}$ increases earlier in 2012 compared to the other years). The impact on the result is however small since these eight periods are in the early part of the growing season and the reduction factor (f_{8day}) is zero."

Minor comments:

We agree with most of the minor comments. To keep this response short we only include the minor comments that we want to give any specific response to. For all minor comments that are not listed below the manuscript has been updated accordingly.

P2, L26: Would be useful to give the land cover type for the southern France site
Response: We have added that the defoliation occurred in holm oak (Quercus ilex L.).

P3, L17: Change to read 'of the form'
P3, L18: Number this equation and update the others accordingly
P3, L20-1: Change to read 'and with variability in meteorology'
Response: The section is updated and these comments are no longer relevant.

P4, L24-6: It is not clear that these recent outbreaks are for the study site – please state the area they apply to. Also give some information about the EC tower in that study (i.e. location, and mention the flux measurements were also for the birch forest)
Response: We have clarified that the outbreaks mentioned were in the study area, and that the EC tower mentioned was located in birch forest: "The latest outbreaks in the study are occurred in 2004, with a documented reduction in carbon sink strength of 89% at an EC tower located in birch forest…"

P11, L22-23 Suggest choosing alternative notation for GPPreduction, as it is a ratio of GPPs, rather than GPP itself

Response: We have changed the notation of the reduction factor to *GPPredfact*

P14 L5-6 and Fig 7: The text mentions 'low GPP observations' but in Fig 7 it looks as though the modelled GPP values are lower than the EC observed values, with several zero values. Please clarify

Response: The low GPP observations with several zero values for LUE modelled GPP are from May, before budburst for the birch forest, when there is photosynthetic activity in understory captured in the EC data. These low GPP values have little influence on annual GPP. We have added a note about this in the discussion but do not further elaborate in potential causes since we have not studied this.