Authors' reply to Anonymous Referee comments (RC C864) on "Monitoring of carbon dioxide fluxes in a subalpine grassland ecosystem of the Italian Alps using a multispectral sensor" (Sakowska et al.):

Dear Anonymous Reviewer,

thank you for the evaluation of the manuscript and constructive comments. We addressed all the issues raised in the review. The reviewer will find below the responses to general and specific comments (typed in bold characters). We hope that thanks to the comments and suggestions addressed during the reviewing process the scientific value of the article will increase.

General comment:

The work in this paper is very solid and the analysis is good, but the authors do little to expand the science. This work repeats studies done by others without showing us anything really new. To me, there are a number of questions that can be addressed by this analysis that would make the paper much more interesting and useful to the community.

Response to the general comment:

In this study we used and tested a 16-band multispectral system which is: 1) not very commonly used in the European proximal sampling and especially in the flux community, 2) relatively low cost (total cost of approximately 4.5 KEUR, including datalogger which is not strictly necessary), 3) commercially available, and 4) easy to configure and use.

We believe that the use of such sensors should be encouraged within the EC networks (Fluxnet, ICOS) e.g. to simulate SENTINEL bands and to investigate the ability of the upcoming sensors to provide reliable estimates of biophysical parameters and CO_2 fluxes across different ecosystems.

Also, if it is true that the link between spectral observations and carbon dioxide fluxes on grasslands is a well investigated topic, we think that the database of this study (5 years) is very solid and allows us to answer a critical question, regarding the applicability of the simple optical sampling models across different years.

Specific comments:

C1: First, one omission in the methods; there is no description of instrument calibration. Over the long study period, what was done to prevent instrument drift? How stable was the instrument? Is this an issue for anyone else using this type of instrument?

A1: According to the reviewer's question, the following information was added to the manuscript section "Multispectral reflectance and narrow-band vegetation indices" (P4735L20):

"Before the beginning of each growing season, the system was calibrated using the method recommended by the manufacturer, based on the use of a white reference panel with known reflectance (http://www.cropscan.com/wsupdn.html). Additionally, CROPSCAN, Inc. provided cosine response calibration data with each upward facing MSR16 module and temperature sensitivity calibration data. Both cosine and temperature corrections were included in the postprocessing software (POSTPROC program) provided with the MSR system."

C2: In the introduction the light use efficiency equation (LUE) was introduced (Eq 1). However, it is not mentioned again in the paper. Of the four different statistical models, only Model 2 directly relates to the LUE, and Model 2 is stated to do poorly. As the LUE is widely used, what do the results of this study say about its applicability? Is PAR unnecessary in the LUE model? If you do need PAR, why did the statistical models that used PAR in them do poorly? Should there be a direct/diffuse ratio added to the model? These are important questions that fall out of your analysis and should be addressed.

A2: From a remote sensing perspective, a strong argument for the use of the concept of LUE model is that all LUE model input parameters can in principle be derived from remote sensing measurements. Spectral vegetation indices presented in the paper are non-direct measures of f_{APAR}, which is one of the components of LUE model. In our view, even not using all the components of the LUE model, but only its simplified version, allows for definition of the general idea behind using VIs as a LUE model concept. This will be especially valid in "dynamic" canopies where f_{APAR} shows high seasonal variations and appears to be the main driver of GEP.

According to the reviewer's comment, the section "Models for GEP_m estimation", describing models formulations presented in the paper, was supplemented with clarification that we refer to the LUE model concept (P4736L8). Later in the article we refer to terms defined in this section:

"In order to estimate GEP_m we used two approaches, one based on linear regression (using the concept of the LUE model) and the other on multiple regression."

Also, we agree with both reviewers that the complex relationship between GEP and PAR should be further discussed in the paper. For this reason, in the "Discussion" section we reworded the sentence in P4742L19-P4742L22 into:

"One reason for this is that sunlight is used by plants more efficiently under cloudy than clear sky conditions due to a more uniform illumination of the canopy, and thus a smaller fraction of the canopy likely to be light saturated (Baldocchi and Amthor, 2001; Chen et al., 2009; Mercado et al., 2009)."

In the same section we added the following paragraph (P4743L1):

"A recent study of Peng et al. (2013) confirmed that the use of PAR in the model can introduce noise and unpredictable uncertainties in GEP estimations. As suggested by these authors, the response of productivity to changes in PAR is quite complex and is influenced by many variables such as vegetation physiological status, canopy structure and light distribution in the canopy. Some other authors also brought to light some important aspects related to the use of PAR. Sims et al. (2008) showed that the variation in PAR is a more relevant determinant of GEP over very short timescales, and appears to be important for diurnal trends. Gitelson et al. (2012) demonstrated that seasonal variation of PAR potential (defined as the maximal value of incident PAR that may occur when the concentrations of atmospheric gasses and aerosols are minimal) can be used to improve the performance of the models."

Also, we reworded the sentence in P4743L1-P4743L4 into:

"Therefore, further analyses of the response of different vegetation types to various levels of diffuse radiation are required, and the hypothesis that the DI and PAR potential can improve the performance of the models including radiation as an input parameter needs to be verified."

And the sentence: "Also, the assessment of the influence of radiation quality on canopy reflectance should be further investigated." (P4743L4-P4743L5) was removed.

C3: The authors suggest that these types of reflectance measurements could be used to determine carbon fluxes and productivity and it would be much cheaper and easier to deploy these optical sensors than flux towers. I wish the authors explored this idea a little farther. How robust are their best models? If the model were parameterized using data from one year, how well would it have performed in the other years?

A3: Considering the long data series presented in the study (characterized by a high variability in both precipitation and air temperature - covering approximately 88% and 54% of the variability observed in a 20 year period for precipitation and temperature, respectively) and the obtained results (robust relationship between GEP_m derived from EC measurements and GEP_m derived from general model 1, 3 and 4), we see the use of ground spectral measurements for monitoring GEP_m in a long-term framework as very promising. However, taking into account the limitations of both methods (EC and optical sampling of vegetation), they cannot be used interchangeably, but only complement each other.

Following the suggestion of both reviewers we performed the validation of the best performing general models (model 1 and 4). Sections "Statistical analysis" (P4737L21), "Results" (P4741L5) and "Discussion" (P4743L21) have been enhanced with the information about the validation procedure and results:

2.5 Statistical analysis:

"Additionally, a validation of the best performing general models using training/validation splitting approach, in which one year at a time was excluded from the dataset, was conducted. The remaining 4 years subset was used as a training set and the excluded year as a validation set. The model was fitted (calibrated) against each training set and the resulting parameterization was used to predict the GEP_m of the excluded year. Validation accuracy was evaluated in terms of RMSE."

3 Results:

"Validation of model 1 based on $NDVI_{red-edge}$ showed that there was no relevant difference in prediction accuracy among validation years (RMSE was varying between 3.12 and 3.85 μmol m^{-2} s^{-1}, Figure 6). Validation results of general model 4 showed that considering all the 5 validated years RMSE was on average 3.26 μmol m^{-2} s^{-1}."

4 Discussion:

"Validation results of general model 1 fed with $NDVI_{red-edge}$ showed that RMSE increased on average from 3.41 to 3.48 μmol m^{-2} s^{-1}, compared to non-validated general model 1 (averaging the values obtained from the 5 different validation years). Validation results of general model 4 showed that RMSE increased on average from 3.06 to 3.26 μmol m^{-2} s^{-1}, compared to non-validated general model 4. The highest decrease of the GEP_m estimation accuracy was noted in the growing season of 2012 (Table 4, Figure 6), which was presumably caused by the unusual drought which occurred just after the cut event. The precipitation to temperature ratio for a 15 day period after the cut in the growing season of 2012 was more than 10 times lower than in the other years and this fact could have affected GEP_m to a higher extent than VIs related to canopy "greenness". As a consequence, models calibrated with the first four years of the dataset overestimated the GEP_m measured in the second part of the growing season of 2012."

C4: Are there particular times or conditions (e.g. rain or very cloudy conditions) where errors in flux estimation are particularly bad?

A4: There are only a few times or conditions under which errors in flux estimation are particularly bad. We are aware of three cases when the data should be discarded from the analysis: 1) when the site was covered by snow, 2) when precipitation was recorded 2 hours prior or during the midday averaging period, and 3) when the weather conditions did not allow for the removal of the cut biomass from the footprint of Cropscan system (and EC tower) straight after the cut event. In either case the data were omitted in the deliberation. Also, in order to check the performance of the models in cloudy conditions we established and compared the relationships between EC derived GEP_m and $NDVI_{red-edge}$ in the growing season 2012 for: 1) sunny conditions (diffusion index − DI<0.3), 2) cloudy conditions (DI>0.7) and 3) regardless the quality of

incoming radiation (DI<0.3 and DI>0.7). The obtained results showed that cloudy conditions did not affect the model performance significantly.

The paragraph of "Multispectral reflectance and narrow-band vegetation indices" section (P4735L26-P4736L2) was slightly extended as follows:

"In addition, data were excluded when: 1) the site was covered by snow, 2) precipitation was recorded 2 hours prior or during the midday averaging period, and 3) the weather conditions did not allow for the removal of the cut biomass from the footprint of Cropscan system (and EC tower) straight after the cut event. According to these quality criteria, 24% of the data were discarded, mainly due to the meteorological conditions."

C5: Are the relationships developed during the spring green-up the same as those for the summer green-up after cutting? Can a brief (say, month-long) training dataset provide a good solution for the rest of the season (or other years)?

A5: In order to check the above mentioned possible seasonal effect, we established and compared the relationships between EC derived GEP_m and $NDVI_{red-edge}$ measured during the 5 years of observations for: 1) the periods before the cut event, and 2) the periods after the cut event. Slopes and y-intercepts of both linear regressions were statistically indistinguishable (p>0.72) (Figure A).

C6: If the optical data provide a reliable estimate of GEP, could that then be used to estimate daytime respiration?

A6: In order to answer the reviewer's question, we tested whether our optical data are able to provide reliable estimates of mean midday ecosystem respiration (R_{eco}). The obtained results showed that all our models were performing poorly in R_{eco} predictions. We think that the reason for this is the lack of a direct relationship between reflectance and both, autotrophic and heterotrophic components of the respiration (Wohlfahrt et al., 2010). Moreover, at the Monte Bondone grassland site, temperature plays a major role in affecting both, diurnal and seasonal patterns of ecosystem respiration (Marcolla et al., 2011).

C7: It was also suggested that the optical data could be used to fill in gaps in the flux data. It would be nice to see a test of that idea, by creating gaps of varying sizes at different times of the year and filling them using the optical data. Does a single parameterization work well, or is it better to tune the equations from data surrounding the gap? Perhaps the authors intend to address these kinds of questions in future papers, but not adding something to the discussion in this paper leaves it with little lasting to say.

A7: Following the reviewer's suggestion we tested the ability of the optical data for GEP_m gap-filling.

Accordingly, a new section was added to the manuscript: "2.6 The gap scenarios" (P4737L24), and sections "Results" (P4741L9) and "Discussion" (P4744L21) were extended:

2.6 The gap scenarios:

"In order to evaluate the ability of spectral models to gap-fill CO_2 flux data, secondary datasets were generated by flagging ~16 % of the 5 growing seasons data as unavailable (artificial gaps constituting 90 observation days out of 573 available observation days). The percentage of artificial gaps was chosen due to the fact that during the observation period of the study (~ May to November, 2008-2012) the EC dataset had an average of 16 % of missing or rejected values of NEE data collected during midday hours. Following Moffat et al. (2007) these artificial gaps were superimposed on the already incomplete data, without regard for the distribution of real gaps in the time series. Three gap length scenarios were considered: gaps of 1, 3 and 5 observation days. The artificial gaps were distributed randomly and each of the three scenarios was permuted 10 times and results were averaged (Moffat et al., 2007). Secondary datasets with artificial gaps were used to calibrate the models that were applied for filling GEP_m data. The gap-filling statistical metrics (adjR^2, RMSE, PRMSE) were calculated using the EC derived GEP_m in these artificial gaps to validate the predictions of filling technique."

3 Results:

"The differences in the adjR^2 performance of the gap-filling scenarios showed that the accuracy of gap filling decreased slightly with gap length, while the range of the goodness of fit statistics (adjR^2, RMSE, PRMSE) generally increased with gap size (Table 6). However, on average, GEP_m gaps were filled with an accuracy of 73% with model 1 fed with $NDVI_{red-edge}$ (RMSE=3.40 μmol m^{-2} s^{-1}, PRMSE= 16.48 %), and with an accuracy of 76% (RMSE=3.14 μmol m^{-2} s^{-1}, PRMSE= 15.25 %) with model 4 using reflectance at 681, 720 and 781 nm and PAR_m data."

4 Discussion:

"The results of a simple gap filling approach presented in this study (based on creating and superimposing randomly distributed artificial gaps of three different lengths on the real dataset and comparing GEP_m values derived from EC with GEP_m values filled with the best performing spectral models) encourage the use of spectral data in the gap filling procedures of EC flux time series. The spectral based models were able to predict GEP_m values with a performance comparable with others methods (Moffat et al., 2007) with adjR^2 ranging from 0.70 (5 days long gap, general model 1 parameterized with $NDVI_{red-edge}$) to 0.78 (1 day long gap, general model 4 based on reflectance at 681, 720 and 781 nm and PAR_m data) (Table 6)."

References

Baldocchi, D. D. and Amthor, J. S.: Canopy Photosynthesis: History , Measurements, and Models, in Terrestrial Global Productivity:Past, Present and Future, edited by J. Roy, B. Saugier, and H. Mooney, pp. 9–31, Academic Press, San Diego., 2001.

Chen, J., Shen, M. and Kato, T.: Diurnal and seasonal variations in light-use efficiency in an alpine meadow ecosystem: causes and implications for remote sensing, J. Plant Ecol., 2(4), 173–185, doi:10.1093/jpe/rtp020, 2009.

Gitelson, A. A., Peng, Y., Masek, J. G., Rundquist, D. C., Verma, S., Suyker, A., Baker, J. M., Hatfield, J. L. and Meyers, T.: Remote estimation of crop gross primary production with Landsat data, Remote Sens. Environ., 121, 404–414, doi:10.1016/j.rse.2012.02.017, 2012.

Marcolla, B., Cescatti, A., Manca, G., Zorer, R., Cavagna, M., Fiora, A., Gianelle, D., Rodeghiero, M., Sottocornola, M. and Zampedri, R.: Climatic controls and ecosystem responses drive the inter-annual variability of the net ecosystem exchange of an alpine meadow, Agric. For. Meteorol., 151, 1233–1243, doi:10.1016/j.agrformet.2011.04.015, 2011.

Mercado, L. M., Bellouin, N., Sitch, S., Boucher, O., Huntingford, C., Wild, M. and Cox, P. M.: Impact of changes in diffuse radiation on the global land carbon sink., Nature, 458, 1014–1017, doi:10.1038/nature07949, 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, Agric. For. Meteorol., 147, 209–232, doi:10.1016/j.agrformet.2007.08.011, 2007.

Peng, Y., Gitelson, A. A. and Sakamoto, T.: Remote estimation of gross primary productivity in crops using MODIS 250 m data, Remote Sens. Environ., 128, 186–196, 2013.

Sims, D., Rahman, a, Cordova, V., Elmasri, B., Baldocchi, D., Bolstad, P., Flanagan, L., Goldstein, a, Hollinger, D. and Misson, L.: A new model of gross primary productivity for North American ecosystems based solely on the enhanced vegetation index and land surface temperature from MODIS, Remote Sens. Environ., 112(4), 1633–1646, doi:10.1016/j.rse.2007.08.004, 2008.

Wohlfahrt, G., Pilloni, S., Hörtnagl, L. and Hammerle, A.: Estimating carbon dioxide fluxes from temperate mountain grasslands using broad-band vegetation indices, Biogeosciences, 7, 683–694 [online] Available from: http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=3856878&tool=pmcentrez&rendertype=abstract, 2010.

__Authors' reply to Referee comments (Anatoly Gitelson, RC C933) on "Monitoring of carbon dioxide fluxes in a subalpine grassland ecosystem of the Italian Alps using a multispectral sensor" (Sakowska et al.):__

Dear Reviewer,

thank you for the evaluation of the manuscript and constructive comments. We addressed all the issues raised in the review. The reviewer will find below the responses to general and specific comments. We hope that, thanks to the comments and the suggestions, the scientific value of the article will be enhanced.

General comment:

The paper aimed to estimate gross primary production of subalpine grasslands remotely. Using simple radiometer measuring reflectance in 16 spectral bands synchronously with CO_2 fluxes, very valuable data set consists 5 years of observation has been collected. Authors tested two models based on vegetation indices, one of them including incident PAR, and two regression models. The results showed that in vegetation studied, a main factor affecting productivity is total chlorophyll content and, thus, primary production could be accurately estimated via remote detection of chlorophyll content. The results of this study are very interesting and convincing. I believe, more explicit presentation of the results greatly improve value of this paper.

Specific comments:

C1: Firstly, more explanation is required the fact that performance of model 1 (that does use any meteorological data, e.g., incident PAR) is better than model 2 where PAR was used. Recently many studies brought empirical evidence that using incident PAR in gross primary production (GPP) models, requiring meteorological data, does not increase accuracy of GPP estimation. Moreover, the models, which do not use any meteorological information and based only on remotely sensed data, perform better (e.g., Sims et al, 2008; RSE; Yang et al., 2013; GRL). Sakamoto et al., (2011; RSE) showed that the use of vegetation index alone allowed for accurate estimation of crop GPP up to the point where seasonal decrease of PAR became significant. Thus, seasonal change of PAR was found one of the factors affecting GPP. GPP is affected by incident PAR and the response of productivity to change in PAR relates to many factors such as vegetation physiological status and light climate inside the canopy, which affects absorbed PAR and LUE, among others. Therefore, the use of incident PAR in the model may introduce noise and unpredictable uncertainties (see figure below from Peng et al., 2013; RSE, showing it explicitly). As a result, it was suggested using calculated seasonal variation of PAR in the

model (Gitelson et al., 2012; RSE). Thus, authors' conclusion that "the photosynthesis process is more efficient under diffuse compared to direct radiation, …the accuracy of GEPm estimation decreased after including incident PARm into the model" is only one factor in very complicated interaction GEP/PAR. I suggest to refer Sims et al., (2008; RSE) paper discussing this issue.

A1: We agree with both reviewers that the complex relationship between GEP and PAR should be further discussed in the paper. For this reason, in the "Discussion" section we reworded the sentence in P4742L19-P4742L22 into:

"One reason for this is that sunlight is used by plants more efficiently under cloudy than clear sky conditions due to a more uniform illumination of the canopy, and thus a smaller fraction of the canopy likely to be light saturated (Baldocchi and Amthor, 2001; Chen et al., 2009; Mercado et al., 2009)."

In the same section we added the following paragraph (P4743L1):

"A recent study of Peng et al. (2013) confirmed that the use of PAR in the model can introduce noise and unpredictable uncertainties in GEP estimations. As suggested by these authors, the response of productivity to changes in PAR is quite complex and is influenced by many variables such as vegetation physiological status, canopy structure and light distribution in the canopy. Some other authors also brought to light some important aspects related to the use of PAR. Sims et al. (2008) showed that the variation in PAR is a more relevant determinant of GEP over very short timescales, and appears to be important for diurnal trends. Gitelson et al. (2012) demonstrated that seasonal variation of PAR potential (defined as the maximal value of incident PAR that may occur when the concentrations of atmospheric gasses and aerosols are minimal) can be used to improve the performance of the models."

Also, we reworded the sentence in P4743L1-P4743L4 into:

"Therefore, further analyses of the response of different vegetation types to various levels of diffuse radiation are required, and the hypothesis that the DI and PAR potential can improve the performance of the models including radiation as an input parameter needs to be verified."

And the sentence: "Also, the assessment of the influence of radiation quality on canopy reflectance should be further investigated." (P4743L4-P4743L5) was removed.

C2: Secondly, the performance of the model 1 was very consistent among 4 years of observation (2008-2011); however, it was not a case for 2012. I do not see a problem that "the slopes of these linear relationships in 2011 and 2012 were significantly different from the general model". Slope is not the only factor affecting relationship, there is also intercept. Relationship for 2011 was very close to five year line (Fig. 4). What has to be addressed and explained is very different performance of the model in 2012. I suggest

establishing GEP vs. VI relationship for four years (2008-2011) and explain discrepancy between this close relationship and that for 2012. The reason for this discrepancy is very important to understand; it brings crucial information about validity of the model.

A2: According to our observations, 2012 was a very particular year, with both high average precipitation rates and high average temperatures during the growing season (Figure 2 of the manuscript: http://www.biogeosciences-discuss.net/11/4729/2014/bgd-11-4729-2014.pdf), which in this type of ecosystem led to optimal growing conditions and particularly green grassland with high green herbage ratio (GR) values (ratio between the green biomass and the total biomass) throughout the season. Fig. B referring to the growing season of 2012 shows that the vegetation index values after the cut are the lowest (pattern dots). The same pattern is visible in the other years of observation (Fig. A). In 2012, the values of $NDVI_{red-edge}$ right after the cut were higher compared to the $NDVI_{red-edge}$ values after the cut of the other years (Figure 4 in the manuscript:http://www.biogeosciences-discuss.net/11/4729/2014/bgd-11-4729-2014.pdf); which means that also the f_{APAR} green values in 2012, right after the cut, were presumably higher than in the other years due to the favorable climatic conditions before the cut. On the contrary, the GEP_m values of 2012 after the cut, were lower than in the other years for the same index values (Figure 4 in the manuscript: http://www.biogeosciences-discuss.net/11/4729/2014/bgd-11-4729-2014.pdf), and this is expected to affect the slope and the intercept of the 2012 relationship. We presume that the reason for these low values of GEP_m is that the grassland was undergoing stress right after the cut event. To check this hypothesis, we tested the precipitation pattern after the cut and we found that in 2012 no significant daily precipitation (>3mm) was recorded during the 15 days after the cut. For other years, no dry periods (>5 days without a daily precipitation > 3 mm) were recorded. Also, we calculated the Precipitation/Temperature ratio (P/T) for a 15 day period after the cut. The P/T ratio in 2012 during this period was more than 10 times lower than in the other years. According to these results, we can confirm that the difference in the 2012 relationship is likely due to the drought event occurring right after the cut. Although the grass was greener compared to the other years, the GEP_m values were lower than expected.

Following the suggestion of both reviewers we performed the validation of the best performing general models (model 1 and 4). Sections "Statistical analysis" (P4737L21), "Results" (P4741L5) and "Discussion" (P4743L21) have been enhanced with the information about the validation procedure and results:

2.5 Statistical analysis:

"Additionally, a validation of the best performing general models using training/validation splitting approach, in which one year at a time was excluded from the dataset, was conducted. The remaining 4 years subset was used as a training set and the excluded year as a validation set. The model was fitted (calibrated) against each training set and the resulting parameterization was

used to predict the GEP_m of the excluded year. Validation accuracy was evaluated in terms of RMSE."

3 Results:

"Validation of model 1 based on $NDVI_{red-edge}$ showed that there was no relevant difference in prediction accuracy among validation years (RMSE was varying between 3.12 and 3.85 µmol $m^{-2} s^{-1}$, Figure 6). Validation results of general model 4 showed that considering all the 5 validated years RMSE was on average 3.26 µmol $m^{-2} s^{-1}$."

4 Discussion:

"Validation results of general model 1 fed with $NDVI_{red-edge}$ showed that RMSE increased on average from 3.41 to 3.48 µmol $m^{-2} s^{-1}$, compared to non-validated general model 1 (averaging the values obtained from the 5 different validation years). Validation results of general model 4 showed that RMSE increased on average from 3.06 to 3.26 µmol $m^{-2} s^{-1}$, compared to non-validated general model 4. The highest decrease of the GEP_m estimation accuracy was noted in the growing season of 2012 (Table 4, Figure 6), which was presumably caused by the unusual drought which occurred just after the cut event. The precipitation to temperature ratio for a 15 day period after the cut in the growing season of 2012 was more than 10 times lower than in the other years and this fact could have affected GEP_m to a higher extent than VIs related to canopy "greenness". As a consequence, models calibrated with the first four years of the dataset overestimated the GEP_m measured in the second part of the growing season of 2012."

C3: Thirdly, in discussion authors should address limitations of the applied models. I am not sure that authors used right expression ("simultaneous estimates of "can be redundant") about necessity to assess light use efficiency in non-stressed ecosystems characterized by strong seasonal dynamics such as grasslands and croplands. But why "non-stressed" vegetation mentioned? Authors study natural vegetation that does stressed. LUE relates to electron transport that in turn relates to chlorophyll content. Thus, detecting chlorophyll content does help to take into account some aspects of plant physiological status but there are many other factors affecting plant status and, thus, assessing LUE is extremely important especially for natural stressed vegetation. Obvious lag between stress and decrease in chlorophyll content does affect accuracy of the model and it should be explicitly mentioned.

A3: We agree on the reviewer's observation highlighting the existing lag between stress and chlorophyll content, especially in shorter time observations, thus in the "Discussion" the limitation of our approach was mentioned (P4743L28). Also, "Introduction" part (P4732L8-P4732L13) was slightly reworded as follows:

1 Introduction:

"In non-stressed ecosystems characterized by strong seasonal dynamics such as some managed croplands, independent estimates of ε may be unnecessary due to its relation with the chlorophyll content (Gitelson et al., 2012; Peng and Gitelson, 2012; Peng et al., 2011; Rossini et al., 2012; Wu et al., 2009), and this is particularly true when integrating GEP over longer time scales, e.g. days (Gitelson et al., 2008). Therefore most of the variations in plant productivity in such ecosystems should be reflected by changes in APAR (Lobell et al., 2002)."

4 Discussion:

"We must however emphasize that the possible limitation of the approach based on VIs related to "canopy greenness" is that variations of GEP due to the short term environmental stresses cannot be monitored by these VIs, unless these stresses affect chlorophyll content (Gitelson et al., 2008)."

C4: Forth, I suggest authors to select only figures those are really necessarily for clear an understandable presentation obtained results. These results are valuable and would be much better presented by selecting few self-explained figures.

A4: According to the reviewer comment, we think that the Figure 6 of the manuscript (http://www.biogeosciences-discuss.net/11/C2641/2014/bgd-11-C2641-2014.pdf), although it provides an overview of the time series of mean midday gross ecosystem production (GEP$_m$) estimated from EC measurements and GEP$_m$ obtained with the various models, is not strictly necessary in the paper. Also, the paragraph referring to Fig. 6 (P4741L5-P4741L8) was removed from the manuscript.

C5: Finally, abstract does not seems to me very informative and conclusions requite thoughtful revision.

A5: Both, "Abstract" and "Conclusions" sections were reworded as follows:

Abstract:

"The study investigates the potential of a commercially available proximal sensing system - based on a 16-band multispectral sensor - for monitoring mean midday gross ecosystem production (GEP$_m$) in a dynamic subalpine grassland ecosystem of the Italian Alps equipped with an eddy covariance flux tower. Reflectance observations were collected for five consecutive years, characterized by different climatic conditions, together with turbulent carbon dioxide fluxes and their meteorological drivers. Different models based on linear regression (vegetation indices approach) and on multiple regression (reflectance approach) were tested to estimate GEP$_m$ from optical data. The overall performance of this relatively low-cost system was positive.

Chlorophyll-related indices including the red-edge part of the spectrum in their formulation (Red-Edge Normalized Difference Vegetation Index, $NDVI_{red-edge}$; Chlorophyll Index, $CI_{red-edge}$) were the best predictors of GEP_m, explaining most of its variability during the observation period. The use of the reflectance approach did not lead to considerably improved results in estimating GEP_m: the adjusted R^2 (adjR^2) of the model based on linear regression - including all the 5 years - was 0.74, while the adjR^2 for the multiple regression model was 0.79. Incorporating mean midday photosynthetically active radiation (PAR_m) into the model resulted in a general decrease in the accuracy of estimates, highlighting the complexity of the GEP_m response to incident radiation. In fact, significantly higher photosynthesis rates were observed under diffuse as regards to direct radiation conditions. The models which were observed to perform best were then used to test the potential of optical data for GEP_m gap-filling. Artificial gaps of three different lengths (1, 3 and 5 observation days) were introduced in the GEP_m time series. The values of adjR^2 for the three gap-filling scenarios showed that the accuracy of the gap filling slightly decreased with gap length. However, on average, the GEP_m gaps were filled with an accuracy of 73% with the model fed with $NDVI_{red-edge}$, and of 76% with the model using reflectance at 681, 720 and 781 nm and PAR_m data."

Conclusions:

"This study investigated the potential of a commercially available system - based on a 16 band multispectral sensor - for monitoring mean midday gross ecosystem production (GEP_m) in a dynamic subalpine grassland ecosystem of the Italian Alps. Chlorophyll-related indices including the red-edge part of the spectrum in their formulation (such as $NDVI_{red-edge}$ and $CI_{red-edge}$) were the best predictors of GEP_m, and were able to explain most of its variability (adjR^2 = 0.74 for $NDVI_{red-edge}$, adjR^2 = 0.73 for $CI_{red-edge}$) during the five consecutive years of observations, characterized by different climatic conditions. Our results confirm the findings of the literature regarding the complexity of the response of ecosystem productivity to change in PAR (Peng et al., 2013). This response is influenced by many variables and in fact, in our study, the accuracy of GEP_m estimation decreased after including incident PAR_m into the linear regression model and the photosynthesis process was shown to be more efficient under diffuse compared to direct radiation. Further investigations are planned in order to explore the utility of including DI and PAR potential in the models to improve their performances. Also, the use of the reflectance approach instead of the VIs approach did not lead to considerably improved results in estimating GEP_m. Although a more detailed analysis of the full vegetation spectrum is desirable (for providing best performing algorithms and a solid basis for in-situ validation and up-scaling of optical models to the airborne and satellite platforms), the results indicate that such relatively low-cost multispectral sensors can be adopted to provide a significant contribution in monitoring carbon dioxide fluxes and biophysical parameters in dynamic ecosystems, for improving gap-filling techniques and for further integration into more complex biogeochemical models."

References

Baldocchi, D. D. and Amthor, J. S.: Canopy Photosynthesis: History , Measurements, and Models, in Terrestrial Global Productivity:Past, Present and Future, edited by J. Roy, B. Saugier, and H. Mooney, pp. 9–31, Academic Press, San Diego., 2001.

Chen, J., Shen, M. and Kato, T.: Diurnal and seasonal variations in light-use efficiency in an alpine meadow ecosystem: causes and implications for remote sensing, J. Plant Ecol., 2(4), 173–185, doi:10.1093/jpe/rtp020, 2009.

Gitelson, A. A., Viña, A., Masek, J. G., Verma, S. B. and Suyker, A. E.: Synoptic Monitoring of Gross Primary Productivity of Maize Using Landsat Data, IEEE Geosci. Remote Sens. Lett., 5(2), 133–137, 2008.

Gitelson, A. A., Peng, Y., Masek, J. G., Rundquist, D. C., Verma, S., Suyker, A., Baker, J. M., Hatfield, J. L. and Meyers, T.: Remote estimation of crop gross primary production with Landsat data, Remote Sens. Environ., 121, 404–414, doi:10.1016/j.rse.2012.02.017, 2012.

Lobell, D. B., Asner, G. P., Ortiz-Monasterio, J. I. and Benning, T. L.: Remote sensing of regional crop production in the Yaqui Valley, Mexico: estimates and uncertainties, Agric. Ecosyst. Environ., 1944, 1–16, 2002.

Mercado, L. M., Bellouin, N., Sitch, S., Boucher, O., Huntingford, C., Wild, M. and Cox, P. M.: Impact of changes in diffuse radiation on the global land carbon sink., Nature, 458, 1014–1017, doi:10.1038/nature07949, 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, Agric. For. Meteorol., 147, 209–232, doi:10.1016/j.agrformet.2007.08.011, 2007.

Peng, Y. and Gitelson, A. A.: Remote estimation of gross primary productivity in soybean and maize based on total crop chlorophyll content, Remote Sens. Environ., 117, 440–448, doi:10.1016/j.rse.2011.10.021, 2012.

Peng, Y., Gitelson, A. A., Keydan, G., Rundquist, D. C. and Moses, W.: Remote estimation of gross primary production in maize and support for a new paradigm based on total crop chlorophyll content, Remote Sens. Environ., 115, 978–989, doi:10.1016/j.rse.2010.12.001, 2011.

Peng, Y., Gitelson, A. A. and Sakamoto, T.: Remote estimation of gross primary productivity in crops using MODIS 250 m data, Remote Sens. Environ., 128, 186–196, 2013.

Rossini, M., Cogliati, S., Meroni, M., Migliavacca, M., Galvagno, M., Busetto, L., Cremonese, E., Julitta, T., Siniscalco, C., Morra di Cella, U. and Colombo, R.: Remote sensing-based

estimation of gross primary production in a subalpine grassland, Biogeosciences, 9, 2565–2584, doi:10.5194/bg-9-2565-2012, 2012.

Sims, D., Rahman, a, Cordova, V., Elmasri, B., Baldocchi, D., Bolstad, P., Flanagan, L., Goldstein, a, Hollinger, D. and Misson, L.: A new model of gross primary productivity for North American ecosystems based solely on the enhanced vegetation index and land surface temperature from MODIS, Remote Sens. Environ., 112(4), 1633–1646, doi:10.1016/j.rse.2007.08.004, 2008.

Wu, C., Niu, Z., Tang, Q., Huang, W., Rivard, B. and Feng, J.: Remote estimation of gross primary production in wheat using chlorophyll-related vegetation indices, Agric. For. Meteorol., 149, 1015–1021, doi:10.1016/j.agrformet.2008.12.007, 2009.

Table 6. Summary of the statistical metrics of gap filling procedure: adjusted R^2 (adjR^2), root mean square error (RMSE) and percentage root mean square error (PRMSE).

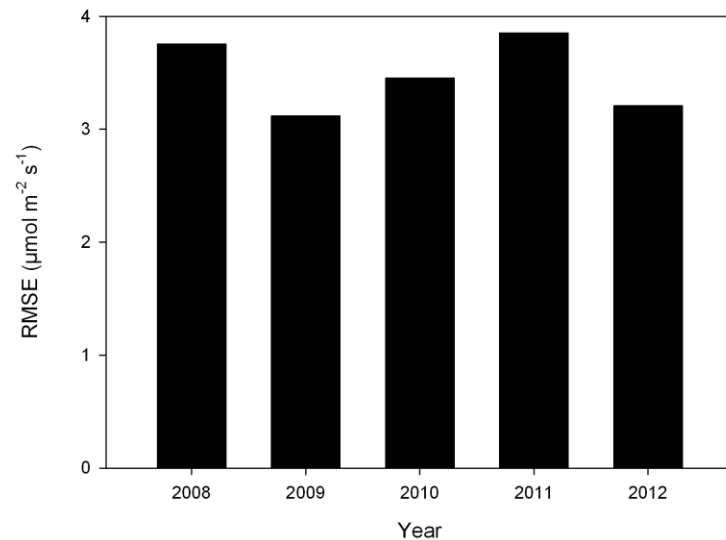| Model		Gap length								
		1 observation day			3 observation days			5 observation days		
		adjR^2	RMSE	PRMSE	adjR^2	RMSE	PRMSE	adjR^2	RMSE	PRMSE
		-	$\mu molm^{-2}s^{-1}$	%	-	$\mu molm^{-2}s^{-1}$	%	-	$\mu molm^{-2}s^{-1}$	%
1	mean	0.76	3.41	16.45	0.72	3.43	16.71	0.70	3.34	16.28
	range	0.16	0.73	3.80	0.28	1.19	5.45	0.46	0.95	6.50
4	mean	0.78	3.16	15.25	0.77	3.10	15.08	0.73	3.17	15.42
	range	0.14	0.46	2.72	0.18	0.81	4.23	0.33	0.75	5.13

Figure 6. Root mean square error (RMSE) of the validated models based on the Red-Edge Normalized Difference Vegetation Index (NDVI_{red-edge}).
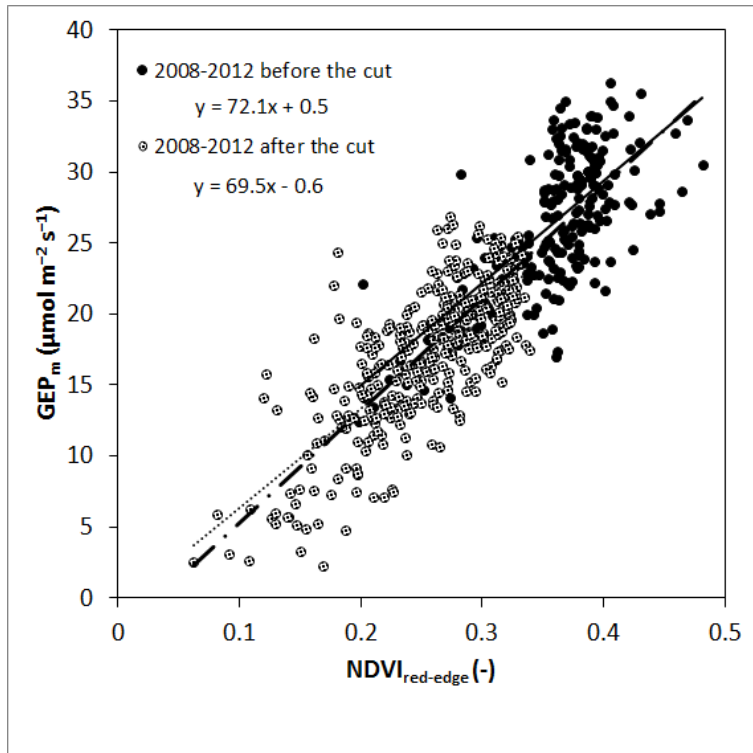
Figure A. Relationship between the Red-Edge Normalized Difference Vegetation Index (NDVI$_{red-edge}$) and mean midday gross ecosystem production (GEP$_m$) considering all the 5 years of observations. Solid dots and solid trend line refer to the periods before the cut event (I sub period of the growing season), pattern dots and dotted trend line refer to the periods after the cut event (II sub-period of the growing season), respectively. Dashed trend line refers to the 5 years of observations without dividing the time series into two sub-periods related to the cut.
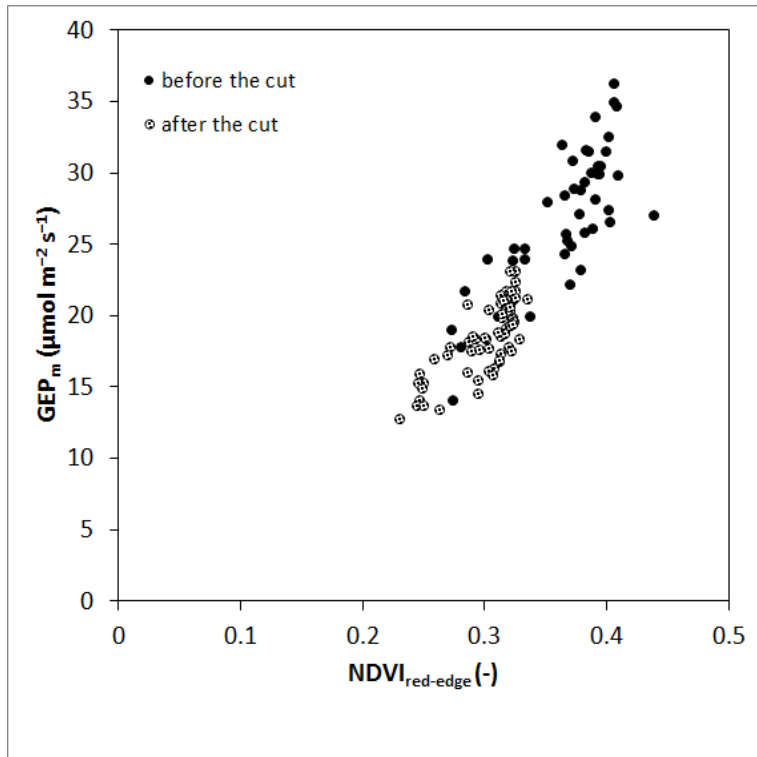
Figure B. Relationship between the Red-Edge Normalized Difference Vegetation Index ($NDVI_{red-edge}$) and mean midday gross ecosystem production (GEP_m) in the growing season of 2012.

__List of other relevant changes made in the manuscript (http://www.biogeosciencesdiscuss. net/11/4729/2014/bgd-11-4729-2014.pdf), besides those presented in the Authors' reply to Referees comments:__

1. P4729: the following affiliation was added: "{Department of Matter and Energy Fluxes, Global Change Research Center, AS CR, v.v.i. Belidla 986/4a, 603 00 Brno, Czech Republic}"
2. P4733L8, another aim of the study was added: "iiii) to evaluate the potential of spectral models to gap-fill GEP_m data"
3. P4739L27- P4739L28, the sentence: "The PRMSE was on average 14.64 % lower in model 2 than in model 1 considering all of the 5 years of observations." was corrected into: "The PRMSE was on average 14.64 % higher in model 2 than in model 1 considering all of the 5 years of observations."
4. P4745L23, the following acknowledgements were added: "The authors would like to acknowledge Maurizo Bagnara, PhD student of Fondazione Edmund Mach, for help in R programming and John Gamon, Profesor from University of Alberta, for fruitful discussions. The authors would like to thank the editor and the two reviewers (Anatoly

Gitelson, and the anonymous reviewer) of this manuscript for their valuable comments which have helped to improve the overall quality of the paper."

5. P4749L17, the following reference was added: "Gitelson, A. A., Viña, A., Masek, J. G., Verma, S. B. and Suyker, A. E.: Synoptic Monitoring of Gross Primary Productivity of Maize Using Landsat Data, IEEE Geosci. Remote Sens. Lett., 5(2), 133–137, 2008."

6. P4752L5, the following reference was added: "Peng, Y., Gitelson, A. A. and Sakamoto, T.: Remote estimation of gross primary productivity in crops using MODIS 250 m data, Remote Sens. Environ., 128, 186–196, 2013."

7. P4753L5, the following reference was added: "Sims, D., Rahman, a, Cordova, V., Elmasri, B., Baldocchi, D., Bolstad, P., Flanagan, L., Goldstein, a, Hollinger, D. and Misson, L.: A new model of gross primary productivity for North American ecosystems based solely on the enhanced vegetation index and land surface temperature from MODIS, Remote Sens. Environ., 112(4), 1633–1646, doi:10.1016/j.rse.2007.08.004, 2008."