Interactive comment on "On the relationship between ecosystem-scale hyperspectral reflectance and CO_2 exchange in European mountain grasslands" by M. Balzarolo et al.

Anonymous Referee #1
Received and published: 17 July 2014

General comments:
This manuscript uses hyperspectral data to identify spectral regions that can be used to estimate biophysical characteristics of three grassland sites in Europe using three simple types of vegetation indices. The models are very simplistic considering the complexity of the BPCs examined, i.e. Gross Primary Production (GPP). Most approaches to estimate GPP use multiple inputs, thus, the approach in this study to estimate GPP using VIs developed by all possible band combinations performed poorly across all three sites. While non-linear relationships with BPCs and VIs may produce low error estimates in calibration, they perform poorly when validated, especially when applied to sites not included in the validation. The study should focus on the linear models and provide readers a sense of stability of the bands selected by using a calibration/validation or cross-validation approach. The authors have a rich data set that can be very beneficial to the scientific community; however, the approach to analyze this data needs improvement. Other smaller issues include (1) a very weak methods section that did not provide enough detail regarding the data collection, (2) poor presentation of the results in complicated tables and figures, and (3) needless duplication of figures in the supplemental that could be presented in the manuscript.

The authors thank the anonymous reviewer #1 for his/her constructive comments and the helpful suggestions. We believe that the manuscript will be improved by addressing these comments. We will revise the manuscript accordingly.

Specific comments

2.2 Hyperspectral reflectance measurements

- **P10328L13: Reflectance should be collected near solar noon. In many locations the midday times may be offset from this ideal period of data collection due to local/national rules and regulations such as the implementation of daylight savings. Indicate when the reflectance measurements were collected in reference to solar noon at the summer solstice or the rough time for much of the growing season.**

RESPONSE
We agree with this comment but in general the uncertainty due to a different sun position should not play an important role collecting data close to solar noon. The hyperspectral measurements were collected close to solar noon that occurs around 13:00 Central European time (i.e. not taking daylight savings into account) during the growing season at the location of the study sites. In particular, the hyperspectral measurements were taken in between 11:00 to 14:00 Central European time.

- **P10328L14: Indicate the model number here. While all details are probably not warranted, do not expect readers to read the previous publication. Even the cited publication is lacking some details**

and refers to another publication. Why not refer to the original here? It is already cited in the manuscript?

RESPONSE
We referred to Vescovo et al. (2012) since hyperspectral and biophysical data used in this manuscript were part of the previous publication and therefore the sampling strategy was exactly the same as reported in this paper. However, we agree with this comment and we will give more details about the hyperspectral sampling methodology in the revised manuscript. The reference number of the spectrometers will also be provided.
In Vescovo et al. 2012, the authors referred to Gianelle et al. 2009 for describing the use of the cosine diffuser foreoptic. We already cited in the manuscript this publication for discussing other issues. Nevertheless, we believe that the reference to Gianelle et al. 2009 can be helpful in this part. It can also help to answer to the following question about the foreoptic diffuser.

REFERENCE
Gianelle, D., Vescovo, L., Marcolla, B., Manca, G., and Cescatti, A.: Ecosystem carbon fluxes and canopy spectral reflectance of a mountain meadow, Int. J. Remote Sens., 30, 435–449, 2009.

- **P10328L18-22: Hemispherical reflectance is very unusual as it is easy to have contamination of the nadir view by the sky as the field of view (FOV) is very wide. What is the model of the cosine diffuser used? What is the FOV of the diffuser? What steps were taken to reduce/eliminate the user/tripod from contaminating the FOV?**

RESPONSE
We agree on the fact that hemispherical reflectance is not usual (e.g. of the use of the same experimental set-up given by Fava et al. 2009; Gianelle et al., 2009; Vescovo et al. 2012). On the other hand, when measurements are carried out close to the canopy (e.g. on a small EC tower with a reduced height), a cosine diffuser is able to provide a more scale-appropriate observation. The setup and spatial dimension of spectral measurements at EC sites (e.g., the connection between the sensor support and the EC footprint) is an area of great debate. This issue is related to the spatial representativeness of spectral data (Balzarolo et al, 2011). When the objective of optical sampling is to provide measurements to be coupled with EC data, the footprints of the two systems should be as much as possible comparable. As shown in Meroni (2011), the cosine foreoptic was selected as an optimal compromise for measuring standard sky irradiance values, canopy irradiance from near-surface optical measurement, and comparing the aforementioned observations with carbon fluxes. For our observations, we used the ASD's Remote Cosine Receptor which was provided and calibrated by the manufacturer. The form of ASD's cosine receptor is referred to as a diffusion-disc collector (DDC). The DDC is constructed of a tube with one end covered by a diffusion-disc, designed with a geometry and material that provides a hemispherical FOV (180˚) and optimizes the cosine response. The FOV contamination is very difficult for hemispherical view sensors, both for sky irradiance and for canopy irradiance. To reduce the nadiral FOV contamination, the instrument was placed on a 1.5 m long horizontal arm. To avoid the zenithal FOV contamination, measurements were taken at least at a 15 meters distance from the EC tower (maximum height of the tower was 6 meters).

<u>In detail, to answer to the comments related to the Section 2.2 "Hyperspectral reflectance measurements" this section (P10328L12- P10329L3) will be rewritten as:</u>

The canopy hyperspectral reflectance measurements were collected at each site under clear sky conditions close to solar noon (between 11:00 to 14:00 Central European time) using the same model of a portable spectroradiometer (ASD FieldSpec HandHeld, Inc., Boulder, CO, USA; serial numbers: 1275 for Amplero, 6354 for Monte Bondone in 2006 and 1191 for Neustift and Monte Bondone in 2006) at all sites. The spectroradiometer acquires reflectance values between 350 and 1075 nm with a Full Width Half Maximum (FWHM) of 3.5 nm and a spectral resolution of 1 nm. In order to achieve a better match between the eddy covariance flux footprint and optical measurements, a cosine diffuser foreoptic (ASD Remote Cosine Receptor, Inc., Boulder, CO, USA), calibrated by the manufacture, was used for nadir/zenith measurements (Gianelle et al., 2009; Fava et al., 2009; Meroni et al. 2011). The ASD's cosine receptor is designed with a geometry and material that provides a hemispherical field of view (FOV) of 180˚ and optimizes the cosine response. To reduce the nadir FOV contamination (i.e. sky irradiance and for canopy irradiance) due to the hemispherical view of the sensor the instrument was placed on a 1.5 m long horizontal arm at a height of 1.5 m above the ground. To avoid the zenithal FOV contamination, the measurements were taken at least at a 15 m distance from the eddy covariance tower (maximum height of the tower was 6 m). The vegetation irradiance (sensor pointing nadir) and sky irradiance (sensor pointing zenith) were measured by rotating the spectoradiometer alternately to acquire spectra from the vegetation and from the sky. Hemispherical reflectance was derived as the ratio of reflected to incident radiance. Each reflectance spectrum was automatically calculated and stored by the spectroradiometer as an average of 20 readings. Before starting each spectral sampling, a dark current measurement was done. For more details on experimental set-up see Vescovo et al. (2012). Spectral measurements were collected from spring until the cutting date at Amplero and Monte Bondone, while at the site in Neustift, which is cut three times during the season, spectral measurements were taken about once per week throughout the growing season of 2006.

REFERENCES

Fava, F., Colombo, R., Bocchi, S., Meroni, M., Sitzia, M., Fois, N., and Zucca, C.: Identification of hyperspectral vegetation indices for Mediterranean pasture characterization, Int. J. Appl. Earth Obs. Geoinf., 11, 233–243, 2009.

Gianelle, D., Vescovo, L., Marcolla, B., Manca, G., and Cescatti, A.: Ecosystem carbon fluxes and canopy spectral reflectance of a mountain meadow, Int. J. Remote Sens., 30, 435–449, 2009.

Vescovo L., Wohlfahrt, G., Balzarolo, M., Pilloni, S., Sottocornola, M., Rodeghiero, M., and Gianelle, D.: New spectral vegetation indices based on the near-infrared shoulder wavelenghts fro remote sensign detection of grassland phytomass, Int. J. Remote Sens., 33, 7, 2012.

Balzarolo, M., Anderson, K., Nichol, C., Rossini, M., Vescovo, L., Arriga, N., Wohlfahrt, G., Calvet, J.-C., Carrara, A., Cerasoli, S., Cogliati, S., Daumard, F., Eklundh, L., Elbers, J. A., Evrendilek, F., Handcock, R. N., Kaduk, J., Klumpp, K., Longdoz, B., Matteucci, G., Meroni, M., Montagnani, L., Ourcival, J.-M., Sanchez-Canete, E. P., Pontailler, J.-Y., Juszczak, R., Scholes, B., and Pilar Martin, M.: Ground-based optical measurements at European flux sites: a review of methods, instruments and current controversies, Sensors, 11, 7954–7981, 2011.

Meroni, M., Barducci, A., Cogliati, S., Castagnoli, F., Rossini, M., Busetto, L., Migliavacca, M., Cremonese, E., Galvagno, M., Colombo, R., and Morra di Cella, U.: The hyperspectral irradiometer, a new instrument

for long-term and unattended field spectroscopy measurements, Rev. Sci. Instrum. 82, 043106; 2011; http://dx.doi.org/10.1063/1.3574360.

- **P10328L20: It is assumed that the 1.5 m was above the ground, but it was not explicitly stated. Why not above the canopy? This would result in the same area seen by the sensor at the top of canopy. Thus, when the grasses are taller, the less of an area the sensor will be seeing.**

RESPONSE
The height of the hyperspectral measurements was 1.5 m above the ground. The vegetation of the three grasses was very homogenous and dense and the maximum eighth of vegetation was less than 0.7 m for that we assumed that the reduction of the footprint area of optical measurements during the growing season was negligible.

2.3 CO2 flux measurements

- **P10329L5-16: More details are needed to describe the CO_2 flux measurements. What brand/models were used? Describe the methods as they are critical to the interpretation. The book by Aubinet et al 2012 describe multiple methods and it is CRITICAL that readers know exactly which methods were used and why. How was Reco modeled? Daytime estimates are confounded by both plant photosynthesis and different meteorological conditions (e.g. temperature, wind speed). It is quite possible that different models and methods are driving the differences between the sites.**

RESPONSE
As for more details on flux measurements we will add more details to Table 1. As will be possible to see from the data of the new Table 1, instruments and post-processing steps were very similar at all sites.
As written in L8-20 on p. 10330 of the BGD paper, Reco was determined by fitting Eq. (4) to both day and nighttime data. The same model and estimation procedure was used at all sites.

Were key supporting meteorological variables also measured (soil heat flux, humidity, incident solar radiation, etc.)? If so, at least list these variables so users understand what kind of gap-filling strategies could be used without needing to directly look up the cited publication and to see if the suggest gap-filling methods make sense for the site.

RESPONSE
Yes, the relevant meteorological variables were measured at each flux tower and the most important ones will be added to the text (see details below on the section 2.3)

- **P10329L5-8: For an empirical study with mixed results, this seems like a very small sample size (1 year of data for 2 sites and 2 years of data for 1 site).**

RESPONSE
We agree with this comment but joint hyperspectral reflectance and eddy covariance flux measurements have been done at very few sites and even fewer of those have long-term data. In addition, there exists no common protocol for hyperspectral measurements in the eddy covariance networks (Balzarolo et al., 2011). Therefore, available hyperspectral measurements are not standardized and comparable between different sites: how to standardize measurements and make results comparable are still open questions. The dataset used in this paper was built based on a coordinated

field experiment by three groups in order to standardize in-situ hyperspectral measurements and make these measurements comparable.

In detail, to answer to the comments related to the Section 2.3 “CO_2 fluxes measurements” this section (P10329L5- 16) will be rewritten as:

Continuous measurements of the net ecosystem CO_2 exchange (NEE) were made by the eddy covariance (EC) technique (Baldocchi et al., 1996; Aubinet et at., 2012) at the three sites. The three wind components and the wind speed were measured using ultra-sonic anemometers, and CO_2 concentrations using open-path infrared gas analysers (IRGA), as detailed in Table 1. Raw data were acquired at 20 Hz and averaged over 30 min time windows in post-processing. Turbulent fluxes were obtained from raw data by applying block averaging (Monte Bondone, Neustift) or linear de-trending (Amplero) methods with a time window of 30 minutes. A 3D coordinate correction was performed according to Wilczak et al. (2001). The CO_2 flux densities were corrected for the effect of air density fluctuations as proposed in Webb et al. (1980). Low- and high-pass filtering was corrected for following Aubinet et al. (2000) (Amplero, Monte Bondone) or Moore (1986) (Neustift). Data gaps due to sensors malfunctioning or violation of the assumptions underlying the EC method were removed and filled using the gap-filling and flux-partitioning techniques as proposed in Wohlfahrt et al. (2008). Ecosystem respiration (Reco) was calculated from the y-intercept of the light response model (see eq. 4). Gross primary productivity (GPP) was calculated as the difference between NEE and Reco. Half-hourly NEE and GPP values were averaged between 11:00 to 14:00 solar local time (at the time window of optical measurements) to allow for direct comparison with the hyperspectral data, and daily sums were also computed. At each site the following supporting environmental measurements were acquired: photosynthetically active radiation (PAR; quantum sensors), air temperature (Ta; PT100, thermistor and thermoelement sensors), and humidity (RH; capacitance sensors) at some reference height above the canopy, and soil temperature (Ts; PT100, thermistor and thermoelement sensors) and water content (SWC; dielectric and time-domain reflectometry sensors) in the main rooting zone. In this study we used CO_2 flux and meteorological data of the years 2005 and 2006 for Monte Bondone and of 2006 for the other sites.

In addition, the following new Table 1 will be added to the revised version of the manuscript:

Table 1. Description of the study sites and period.

Site characteristics	Amplero (IT-Amp)	Neustift (AT-Neu)	Monte Bondone (IT-MBo)
Latitude	41.9041	47.1162	46.0296
Longitude	13.6052	11.3204	11.0829
Elevation (m)	884	970	1550
Mean annual temperature (˚C)	10.0	6.5	5.5
Mean annual precipitation (mm)	1365	852	1189
Vegetation type	Seslerietum apenninae	Pastinaco–Arrhenatheretum	Nardetum Alpigenum
Study period[1]	111-170, 2006 (9)	122-303, 2006 (16)	129-201, 2005 (13) 124-192, 2006 (12)
Sonic anemometer model	R3, Gill Instruments Ltd., Lymington, UK	R3, Gill, Instruments Ltd., Lymington, UK	R3, Gill Instruments Ltd., Lymington, UK
Infrared gas analyser model	Li-7500, Li-Cor Inc., Lincoln, Nebraska, USA	Li-7500, Li-Cor Inc., Lincoln, Nebraska, USA	Li-7500, Li-Cor Inc., Lincoln, Nebraska, USA
Data acquisition frequency (Hz)	20	20	20
Post-processing software	Developed by University of Viterbo (IT)	EdiRE (Version 1.4.3.1021, R. Clement, University of Edinburgh)	EdiRE (Version 1.4.3.1021, R. Clement, University of Edinburgh)
Outlier removal (method)	Wickers and Mahrt (1997)	-	-
CO_2/H_2O signal lag removal	Covariance maximization	Covariance maximization	Covariance maximization
Coordinate rotation (method)[2]	3D	3D	3D
Detrending of time series (method)	Linear detrending	-	-
Density corrections applied[3]	x	x	x
Sonic buoyancy to sensible heat flux conversion and cross-wind correction[4]	x	x	x
Low- and high-pass filtering corrected for (method)	Aubinet et al. (2000)	Moore (1986)	Aubinet et al. (2000)

[1] from-to DOY, year (number of hyperspectral measurement dates); [2] according to Wilczak et al. (2001); [3] according to Webb et al. (1980); [4] according to Schotanus et al. (1983); [5] according to Mauder et al. (2008).

REFERENCES

Aubinet, M., Grelle, A., Ibrom, A., Rannik, U., Moncrieff, J., Foken, T., Kowalski, P., Martin, P., Berbigier, P., Bernhofer, C., Clement, R., Elbers, J., Granier, A., Grunwald, T., Morgenster, K., Pilegaard, K., Rebmann, C., Snijders, W., Valentini, R., and Vesala, T.: Estimates of the annual net carbon and water exchange of Europeran forests: the EUROFLUX methodology, Adv.Ecol.Res., 30, 113–75, 2000.

Balzarolo, M., Anderson, K., Nichol, C., Rossini, M., Vescovo, L., Arriga, N., Wohlfahrt, G., Calvet, J.-C., Carrara, A., Cerasoli, S., Cogliati, S., Daumard, F., Eklundh, L., Elbers, J. A., Evrendilek, F., Handcock, R. N., Kaduk, J., Klumpp, K., Longdoz, B., Matteucci, G., Meroni, M., Montagnani, L., Ourcival, J.-M., Sanchez-Canete, E. P., Pontailler, J.-Y., Juszczak, R., Scholes, B., and Pilar Martin, M.: Ground-based optical measurements at European flux sites: a review of methods, instruments and current controversies, Sensors, 11, 7954–7981, 2011.

Mauder, M., Foken, T., Bernhofer, C., Clement, R., Elbers, J., Eugster, W., Gr¨unwald, T., Heusinkveld, B., and Kolle, O.: Quality control of CarboEurope flux data – Part 2: Inter-comparison of eddy-covariance software, Biogeosciences, 5, 451–462, 2008, http://www.biogeosciences.net/5/451/2008/.

Moore, C.J.: Frequency response corrections for eddy correlation systems, Bound.-Lay. Meteorol., 37, 17–35, 1986.

Schotanus, P., Nieuwstadt, F. T. M., and De Bruin H. A. R.: Temperature measurement with a sonic anemometer and its application to heat and moisture fluxes, Bound. Lay. Meteorol., 26 81–93, 1983.

Webb, E. K., Pearman, G. I., and Leuning, R.: Correction of flux measurements for density effects due to heat and water vapour transfer, Q. J. Roy. Meteorol. Soc., 106, 85–100, 1980.

Vickers, D. and Mahrt, L.: Quality control and flux sampling problems for tower and aircraft data, J. Atmos. Oceanic. Tech., 14, 512–526, 1997.

Wilczak, J. M., Oncley, S. P., and Stage, S. A.: Sonic anemometer tilt correction algorithms, Bound.-Lay. Meteorol., 99, 127–150, 2001.

2.4 Estimation of grassland ecophysiological parameters

- **P10330L1: Be specific on how the extinction coefficient was calculated. It would be assumed for grasslands, but it should be explicitly stated not just referenced. Also identify that this k was determined for grasslands different from the site.**

RESPONSE
We will add in the revised manuscript that we used a value of k = 0.4 defined for southern mixed-grass prairie in Texas.

- **P10330L15: Citation needed for the Levenberg-Marquardt algorithm.**

RESPONSE
The following citation will be added: Marquardt, 1963

REFERENCE
Marquardt, D. W.: An Algorithm for Least-Squares Estimation of Nonlinear Parameters, SIAM J. Appl. Math., 11, 431–441, doi:10.1137/0111030, 1963.

- **P10330L10-19: It is not clear how respiration was measured and/or fitted. Were nighttime measurements used to estimate daytime measurements?**

RESPONSE
P10330L16 will be modified to „… by fitting Eq. (4) to both day and nighttime data …"

2.5 Hyperspectral data analysis

- **P10331L13-17: Most individuals know how to calculate R2 and RMSE. These equations and associated descriptive text can be deleted.**

RESPONSE
We agree and we will delete this part from the text.

- **P10332L12-18: While AIC is a valid approach to determine if added complexity improves the model accuracy, the purpose of this study was to "develop a common framework for predicting grassland GPP based on optical remote sensing data." Thus, a model that is high accuracy in calibration may not be very useful when validated. This is a critical concern when using non-linear models as the VI becomes insensitive to the biophysical characteristic (BPC; e.g. GPP, NEE). This will reduce scatter (thus increase R^2 and reduce RMSE), but be unusable for practical purposes as similar VI values can represent a wide range of BPC (this problem is especially prominent when using NDVI to estimate LAI). A better metric to use for both linear and non-linear relationships would be noise equivalent (NE). Unfortunately for non-linear models, the NE will change based on the value of the BPC. Thus, in some ranges of BPC they will work better than others. This information could not be easily presented in correlograms. This reviewer suggest eliminating non-linear relationships and focus on linear ones as they are (a) easier to use and (b) more reliable throughout the entire dynamic range of each BPC if the relationship is truly linear. This could be easily tested by plotting the best bands for each VI against the BPC.**

RESPONSE
We agree with this comment and will remove the non-linear statistics from the paper and refer to R^2 and RMSE instead of AIC. In addition, the a new figure C1 representing the linear models for the selected bands for all BCPs for each site and all sites pooled will be added to the revised version of the paper. In these plots the results for the leave-one-out cross validation will be presented. In particular, the cross validated R-squared (R^2cv) and Root Mean Square Error (RMSEcv) will shown in the figure C1.

- **P10330L20-P10332L18: Why not divide the data into calibration/validation data sets or use a leave-one-out procedure to test the sensitivity of these selected bands? If the goal is to estimate GPP using remote sensing data, then determining a robust set of wavebands that works for each site should be the initial goal with a secondary goal of finding a set of wavebands that works for all three sites.**

RESPONSE
Thanks for the comment, we agree. In the revised version of the manuscript we will test the sensitivity of the selected bands for all BCPs for each site and all sites pooled by using leave-one-out cross validation procedure and validating the models against new sites.

In the revised version of the manuscript the metrics obtained by leave-one-out procedure applied to BGD dataset will be reported in the new figure C1 (see previous comment). In particular, the cross validated R-squared (R^2cv) and Root Mean Square Error (RMSEcv) will shown in the figure C1.

In order to evaluate the performance of the found relationships and the new selected bands the authors included three new sites in their database (validation sites – see table S1 below). These three additional sites were already part of the preceding study by Vescovo et al. (2012) and used exactly the same methodology as applied at the main three study sites and thus fully comply with our own standards of intercomparability. Validation will be performed applying to the three new sites all the three site specific models (Amplero, Neustift and Monte Bondone) and a model parameterized grouping Neustift and Monte Bondone since the two sites show similar structural characteristics. Figure C2 here below shows the results of the validation of the models against validation sites. As shown in this figure, correspondence between simulated and measured VIs was reasonable when using the models developed for Monte Bondone and Neustift or both sites pooled, but less so with the models of Amplero. This is understandable as Monte Bondone and in particular Neustift are structurally and functionally much more similar to the validation sites compared to Amplero. Overall, the validation shows that the models developed are transferable.

The following new Table S1 will be added to the revised version of the manuscript:

Table S1. Description of the validation study sites and period.

Site characteristics	Längenfeld (AT-Lan)	Leutasch (AT-Leu)	Scharnitz (AT-Sch)
Latitude	47.0612	47.3780	47.3873
Longitude	10.9634	11.1627	11.2479
Elevation (m)	1180	1115	964
Mean annual temperature (˚C)	5.8	4.8	6.4
Mean annual precipitation (mm)	733	1309	1418
Vegetation type	*Phyteumo-Trisetion*	*Astrantio-Trisetetum*	*Arrenatherum montanum*
Study period[1]	163, 2006 (1)	227, 2006 (1)	184-284, 2006 (5)
Sonic anemometer model	R3, Gill Instruments Ltd., Lymington, UK	R3, Gill Instruments Ltd., Lymington, UK	R3, Gill Instruments Ltd., Lymington, UK
Infrared gas analyser model	Li-7500, Li-Cor Inc., Lincoln, Nebraska, USA	Li-7500, Li-Cor Inc., Lincoln, Nebraska, USA	Li-7500, Li-Cor Inc., Lincoln, Nebraska, USA
Data acquisition frequency (Hz)	20	20	20
Post-processing software	EdiRE (Version 1.4.3.1021, R. Clement, University of Edinburgh)	EdiRE (Version 1.4.3.1021, R. Clement, University of Edinburgh)	EdiRE (Version 1.4.3.1021, R. Clement, University of Edinburgh)
Outlier removal (method)	-	-	-
CO_2/H_2O signal lag removal	Covariance maximization	Covariance maximization	Covariance maximization
Coordinate rotation (method)[2]	3D	3D	3D
Detrending of time series (method)	-	-	-
Density corrections applied[3]	X	x	x
Sonic buoyancy to sensible heat flux conversion and cross-wind correction[4]	X	x	x
Low- and high-pass filtering corrected for (method)	Moore (1986)	Moore (1986)	Moore (1986)

[1] from-to DOY, year (number of hyperspectral measurement dates); [2] according to Wilczak et al. (2001); [3] according to Webb et al. (1980); [4] according to Schotanus et al. (1983); [5] according to Mauder et al. (2008).

REFERENCES

Mauder, M., Foken, T., Bernhofer, C., Clement, R., Elbers, J., Eugster, W., Gr¨unwald, T., Heusinkveld, B., and Kolle, O.: Quality control of CarboEurope flux data – Part 2: Inter-comparison of eddy-covariance software, Biogeosciences, 5, 451–462, 2008, http://www.biogeosciences.net/5/451/2008/.

Moore, C.J.: Frequency response corrections for eddy correlation systems, Bound.-Lay. Meteorol., 37, 17–35, 1986.

Schotanus, P., Nieuwstadt, F. T. M., and De Bruin H. A. R.: Temperature measurement with a sonic anemometer and its application to heat and moisture fluxes, Bound. Lay. Meteorol., 26 81–93, 1983.

Webb, E. K., Pearman, G. I., and Leuning, R.: Correction of flux measurements for density effects due to heat and water vapour transfer, Q. J. Roy. Meteorol. Soc., 106, 85–100, 1980.

Wilczak, J. M., Oncley, S. P., and Stage, S. A.: Sonic anemometer tilt correction algorithms, Bound.-Lay. Meteorol., 99, 127–150, 2001.

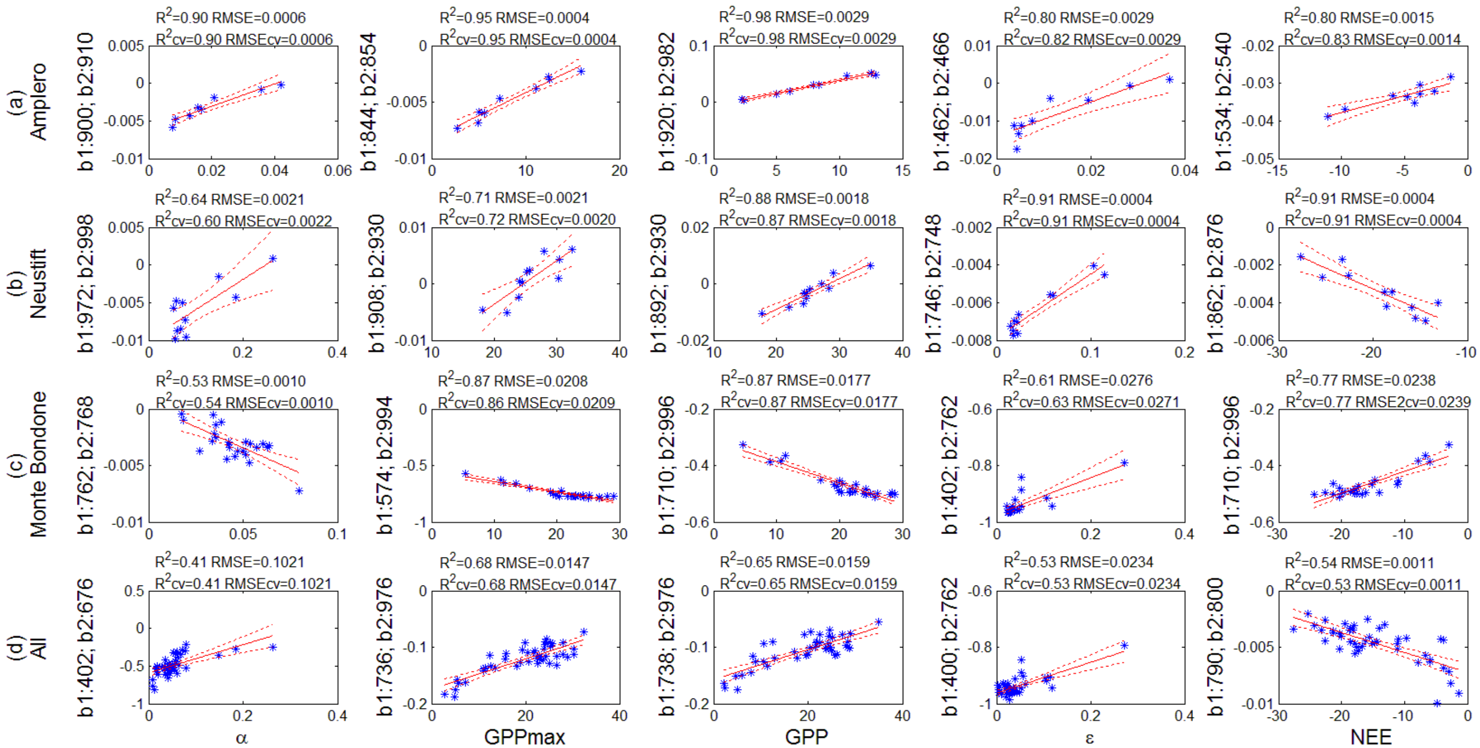

Fig. C1: Results of linear correlation analysis for α, GPPmax and midday averaged GPP, ε and NEE and selected NSD-type indices for (a) Amplero, (b) Neustift,(c) Monte Bondone (both study years pooled) and (d) all sites pooled. R^2—Coefficient of determination; RMSE—Root Mean Square Error; R^2cv—Cross-validated coefficient of determination; $RMSE_{cv}$— Cross-validated root Mean Square Error. The red lines indicate the fitted models and the red dotted lines represent the 95% upper and lower confidence bounds.

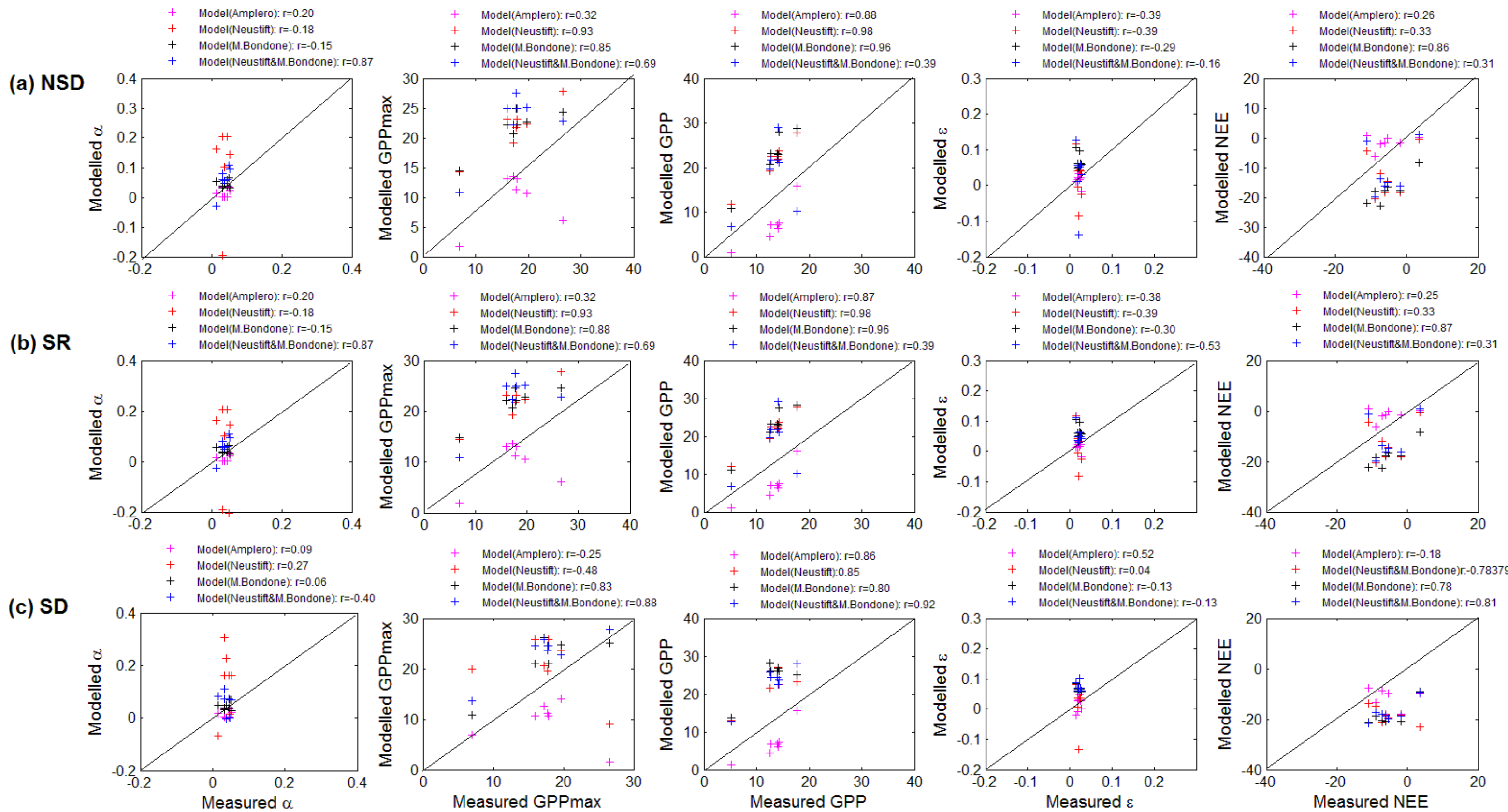

Fig. C2 – Results of validation of linear regression models between VIs ((a) NSD type; (b) SR-type; (c) SD-type) and ecophysiological parameters: α, ε (midday average), GPP_{max} and midday average CO_2 fluxes (NEE and GPP). r—coefficient of correlation. Different colours represent results of the validation performed applying to the three new sites the model for Amplero (in magenta), Neustift (in red) and Monte Bondone (in blue) and a model parameterized grouping Neustift and Monte Bondone (in black).

3.2 Hyperspectral data and their relation to CO2 fluxes and ecophysiological parameters

- **P10333L23-25: Figure 3c does not follow this pattern. The spectra for the highest LAI have lower NIR than the spectra for the next two highest LAI.**

RESPONSE
Thanks for the correction. We agree and we will correct it in the revised version of the manuscript.

- **P10333L27-P10334L1: It is impractical to compare exponential relationships using R^2 values (even RMSE values should not be used) as different slopes/intercepts make it very difficult to conceptualize their real differences. As these relationships were never presented, it is impossible to compare these relationships in this manuscript.**

RESPONSE
Exponential relationships will be removed from the paper.

- **P10334L15-18: One of the problems with correlograms is the end result does not explain causation, only that some correlation exists. There has been quite a bit of research in understanding why specific spectral regions can explain various BPCs. There is no discussion of this research and how it supports the results from the correlograms.**

RESPONSE
We agree, although we would like to stress that no statistical analysis implies causation, and we will add some explanations of the existence of the correlation in some spectral regions. We agree with the reviewer and in the revised version of the manuscript we will add some explanations of the existence of the correlation in some spectral regions. To investigate more the basis of the correlation between the selected band combinations and ecophysiological variables (e.g. alpha, GPPmax, GPP, epsilon) in the revised version of the paper we will analyse the relationship between the selected bands and biophysical parameters such as dry phytomass, nitrogen and water content collected during the field campaign in the same footprint of hyperspectral measurements. The new tables S2 and S3 (here below) will be added to the revised version of the manuscript. This analysis confirmed that the that the spectral response in the selected band combinations for NDS, SR and SD-type indices is strongly related to structural characteristics of the vegetation of the three grasslands (e.g. nitrogen and photomaps) that impact on their spectral response in NIR and VIS regions. For the Mediterranean site (Amplero site) and for all eco-physiological parameters (i.e. a, GPPmax, GPP, epsilon) the dry phytomass is the main driving factor of the spectral response in the selected bands while nitrogen content drive the spectral response in NIR region for Neustift site. For Monte Bondone both dry phytomass and nitrogen content effect spectral response of the grassland. Similar results were obtained for SR and SD-type indices.

Therefore, according to the obtained results, more studies are needed to understand the physical basis of this correlation. The results are somehow confirming the findings of Vescovo et al, (2012) which highlighted a strong relationship, for several grassland types, between an NSD-type index and phytomass. In addition, these new analysis will substantial contribute to the analysis of the structural effect on the ability to estimate canopy nitrogen content that is still a controversial issue (Knyazikhin et al., 2012).

These results will be discussed in detail in the revised version of the manuscript and the corresponding table for the estimation of daily parameters will be added to the supplemental material.

Table S2. Results of the correlation (r – correlation coefficient) between the best NDS, SR and SD-type indices and dry phytomass and nitrogen content for Amplero, Neustift, Monte Bondone for the α, GPPmax, midday GPP, midday ε and midday NEE.

			α		GPPmax		GPP		ε		NEE	
Index	**Site**	**Parameter**	**Band center [i,j]** (nm)	**r** (-)	**Band center [i,j]** (nm)	**r** (-)	**Band center [i,j]** (nm)	**r** (-)	**Band center [i,j]** (nm)	**r** (-)	**Band center [i,j]** (nm)	**r** (-)
NSD-type	Amplero	Dry phytomass (g m-2)	[900, 910]	-0.81**	[844, 854]	-0.85**	[920, 982]	-0.76*	[462, 466]	-0.87**	[534, 540]	0.60
	Amplero	Nitrogen content (%)		0.54		0.57		0.44		0.70*		-0.39
	Amplero	Water content (%)		0.53		0.73*		0.75*		0.66		-0.74*
	Neustift	Dry phytomass (g m-2)	[972, 998]	-0.04	[908, 930]	0.51	[892, 930]	0.59	[746, 748]	-0.66*	[862, 876]	0.15
	Neustift	Nitrogen content (%)		0.40		-0.39		-0.46		0.88**		0.18
	Neustift	Water content (%)		-0.07		0.03		-0.18		0.77*		0.31
	Monte Bondone	Dry phytomass (g m-2)	[762, 768]	-0.13	[574, 994]	-0.77***	[710, 996]	-0.70***	[402, 762]	-0.74***	[710, 996]	-0.70***
	Monte Bondone	Nitrogen content (%)		0.29		0.72***		0.62**		0.69***		0.62**
	Monte Bondone	Water content (%)		0.31		0.69***		0.59**		0.65***		0.59**
	All	Dry phytomass (g m-2)	[402, 676]	0.23	[736, 976]	0.12	[738, 976]	0.14	[400, 762]	-0.22	[790, 800]	0.06
	All	Nitrogen content (%)		0.51***		0.19		0.13		0.64***		0.30
	All	Water content (%)		0.03		-0.09		-0.09		0.32*		0.05
SR-type	Amplero	Dry phytomass (g m-2)	[900, 910]	-0.81***	[844, 854]	-0.85**	[920, 982]	-0.76*	[462, 466]	-0.87*	[534, 540]	0.60
	Amplero	Nitrogen content (%)		0.54		0.57		0.43		0.70*		-0.39
	Amplero	Water content (%)		0.53		0.73**		0.74*		0.66		-0.74*
	Neustift	Dry phytomass (g m-2)	[972, 998]	-0.04	[908, 930]	0.51	[892, 930]	0.59	[746, 478]	-0.66*	[862, 876]	0.15
	Neustift	Nitrogen content (%)		0.40		-0.39		-0.46		0.88**		0.18
	Neustift	Water content (%)		-0.07		0.03		-0.18		0.77*		0.31
	Monte Bondone	Dry phytomass (g m-2)	[762, 768]	-0.13	[570, 994]	-0.77***	[714, 996]	-0.73***	[402, 762]	-0.74***	[570, 574]	0.61**
	Monte Bondone	Nitrogen content (%)		0.29		0.73***		0.64***		0.69***		-0.53**
	Monte Bondone	Water content (%)		0.31		0.69***		0.61**		0.64***		-0.50*
	All	Dry phytomass (g m-2)	[402, 676]	0.28	[736, 976]	0.14	[738, 976]	0.15	[400, 762]	-0.22	[790, 800]	0.06
	All	Nitrogen content (%)		0.51***		0.19		0.13		0.63***		0.30
	All	Water content (%)		-0.01		-0.10		-0.10		0.33*		0.05
SD-type	Amplero	Dry phytomass (g m-2)	[900, 910]	-0.80***	[844, 866]	-0.90**	[920, 982]	-0.77*	[492, 496]	-0.76*	[422, 432]	-0.50
	Amplero	Nitrogen content (%)		0.47		0.55		0.46		0.55		0.19
	Amplero	Water content (%)		0.41		0.67*		0.77*		0.43		0.70*
	Neustift	Dry phytomass (g m-2)	[474, 494]	-0.45	[736, 968]	0.20	[878, 922]	0.61	[732, 942]	-0.45	[402, 456]	-0.04
	Neustift	Nitrogen content (%)		0.33		0.09		-0.34		0.90**		-0.28
	Neustift	Water content (%)		0.15		0.51		-0.04		0.80*		-0.72*
	Monte Bondone	Dry phytomass (g m-2)	[762, 768]	-0.38	[444, 482]	0.65***	[436, 488]	0.60**	[658, 682]	0.67***	[450, 486]	0.60**
	Monte Bondone	Nitrogen content (%)		0.53***		-0.58**		-0.58**		-0.62**		-0.59**
	Monte Bondone	Water content (%)		0.52***		-0.58**		-0.58		-0.56**		-0.55**
	All	Dry phytomass (g m-2)	[822, 824]	0.45***	[550, 560]	0.12	[414, 470]	0.00	[732, 928]	0.55***	[468, 660]	-0.11
	All	Nitrogen content (%)		-0.09		-0.09		-0.15		0.16		-0.33*
	All	Water content (%)		-0.08		0.18		0.19		-0.53***		0.24

Statistical significance is indicated as * ($p < 0.05$), ** ($p < 0.01$), and *** ($p < 0.001$).

Table S3. Results of the correlation (r – correlation coefficient) between the best NDS, SR and SD-type indices and dry phytomass and nitrogen content for Amplero, Neustift, Monte Bondone for daily GPP, ε and NEE.

Index	Site	Parameter	GPP Band center [i,j] (nm)	r (-)	ε Band center [i,j] (nm)	r (-)	NEE Band center [i,j] (nm)	r (-)
NSD-type	Amplero	Dry phytomass (g m-2)	[868, 878]	-0.82**	[896, 904]	-0.89**	[902, 922]	-0.83**
	Amplero	Nitrogen content (%)		0.61		0.54		0.56
	Amplero	Water content (%)		0.81**		0.53		0.85**
	Neustift	Dry phytomass (g m-2)	[972, 988]	-0.14	[722, 942]	-0.54	[422, 516]	-0.25
	Neustift	Nitrogen content (%)		0.27		0.91**		0.15
	Neustift	Water content (%)		0.19		0.80*		0.06
	Monte Bondone	Dry phytomass (g m-2)	[580, 986]	-0.75***	[658, 682]	0.69***	[712, 714]	-0.52*
	Monte Bondone	Nitrogen content (%)		0.71***		-0.66***		0.50*
	Monte Bondone	Water content (%)		0.67***		-0.61**		0.43*
	All	Dry phytomass (g m-2)	[736, 976]	0.12	[404, 944]	-0.21	[790, 798]	0.02
	All	Nitrogen content (%)		0.19		0.68***		0.32*
	All	Water content (%)		-0.09		0.36*		0.08
SR-type	Amplero	Dry phytomass (g m-2)	[868, 878]	-0.82**	[896, 904]	-0.89**	[902, 922]	-0.83**
	Amplero	Nitrogen content (%)		0.61		0.54		0.561
	Amplero	Water content (%)		0.81*		0.53		0.85**
	Neustift	Dry phytomass (g m-2)	[868, 878]	-0.14	[722, 942]	-0.54	[422, 516]	-0.254
	Neustift	Nitrogen content (%)		0.27		0.92**		0.142
	Neustift	Water content (%)		0.19		0.80*		0.056
	Monte Bondone	Dry phytomass (g m-2)	[600, 608]	0.56**	[658, 682]	0.69***	[712, 714]	-0.52*
	Monte Bondone	Nitrogen content (%)		-0.52**		-0.66***		0.50*
	Monte Bondone	Water content (%)		-0.54**		-0.61**		0.43*
	All	Dry phytomass (g m-2)	[736, 976]	0.14	[404, 944]	-0.21	[790, 798]	0.021
	All	Nitrogen content (%)		0.19		0.67***		0.32*
	All	Water content (%)		-0.10		0.37*		0.083
SD-type	Amplero	Dry phytomass (g m-2)	[894, 998]	-0.81**	[844, 856]	-0.89**	[816, 834]	-0.84**
	Amplero	Nitrogen content (%)		0.49		0.51		0.56*
	Amplero	Water content (%)		0.76*		0.59		0.84**
	Neustift	Dry phytomass (g m-2)	[972, 988]	-0.16	[732, 942]	-0.45	[400, 410]	0.092
	Neustift	Nitrogen content (%)		0.33		0.90**		-0.672
	Neustift	Water content (%)		0.25		0.80*		-0.293
	Monte Bondone	Dry phytomass (g m-2)	[444, 502]	0.57**	[658, 680]	0.72***	[468, 496]	0.47*
	Monte Bondone	Nitrogen content (%)		-0.56**		-0.67***		-0.55**
	Monte Bondone	Water content (%)		-0.57**		-0.63***		-0.48*
	All	Dry phytomass (g m-2)	[424, 446]	0.28	[734, 928]	-0.44**	[444, 464]	0.167
	All	Nitrogen content (%)		0.14		0.56***		0.070
	All	Water content (%)		-0.16		0.16		-0.032

Statistical significance is indicated as * ($p < 0.05$), ** ($p < 0.01$), and *** ($p < 0.001$).

In addition, to investigate more the basis of the correlation between the NIR band combinations and GPP, we analyzed a similar dataset collected in summer 2013 on Monte Bondone. Measurements were acquired using the same ASD FieldSpec spectrometer used for Monte Bondone in 2006 (serial number: 6354). The measurements were taken on the tower at a height of 6 m, with a field of view of 25°. To obtain reflectance values, white panel radiance spectra and canopy radiance spectra were acquired at approximately weekly intervals. At the same time of the hyperspectral measurements, measurements of the canopy chlorophyll and canopy water content were performed within the spectrometer footprint (5 m2). In the Figure C3, it is possible to see that, NSD- and SR-type indices for the selected bands for estimating GPP (i.e. 710 nm and 996 nm) are strongly correlated with canopy total chlorophyll content (R^2 > 0.90). For the band combinations < 750 nm, the correlation is related to chlorophyll content while for band combinations > 750nm (which is the most common situation; e.g. 761 and 770, 761 and 850, 800 and 850, etc.) there is a structural effect which needs to be further investigated (confirmed by Gitelson by a personal communication). In fact, the literature indicates that the wavelenghts in the NIR (>750nm) are not sensitive to chlorophyll content. They are sensitive to leaf and canopy structure (and around the 970nm area to water). These new analysis will substantial help to the analysis of the structural effect on the ability to estimate canopy nitrogen content that is still a controversial issue (Knyazikhin et al., 2012).

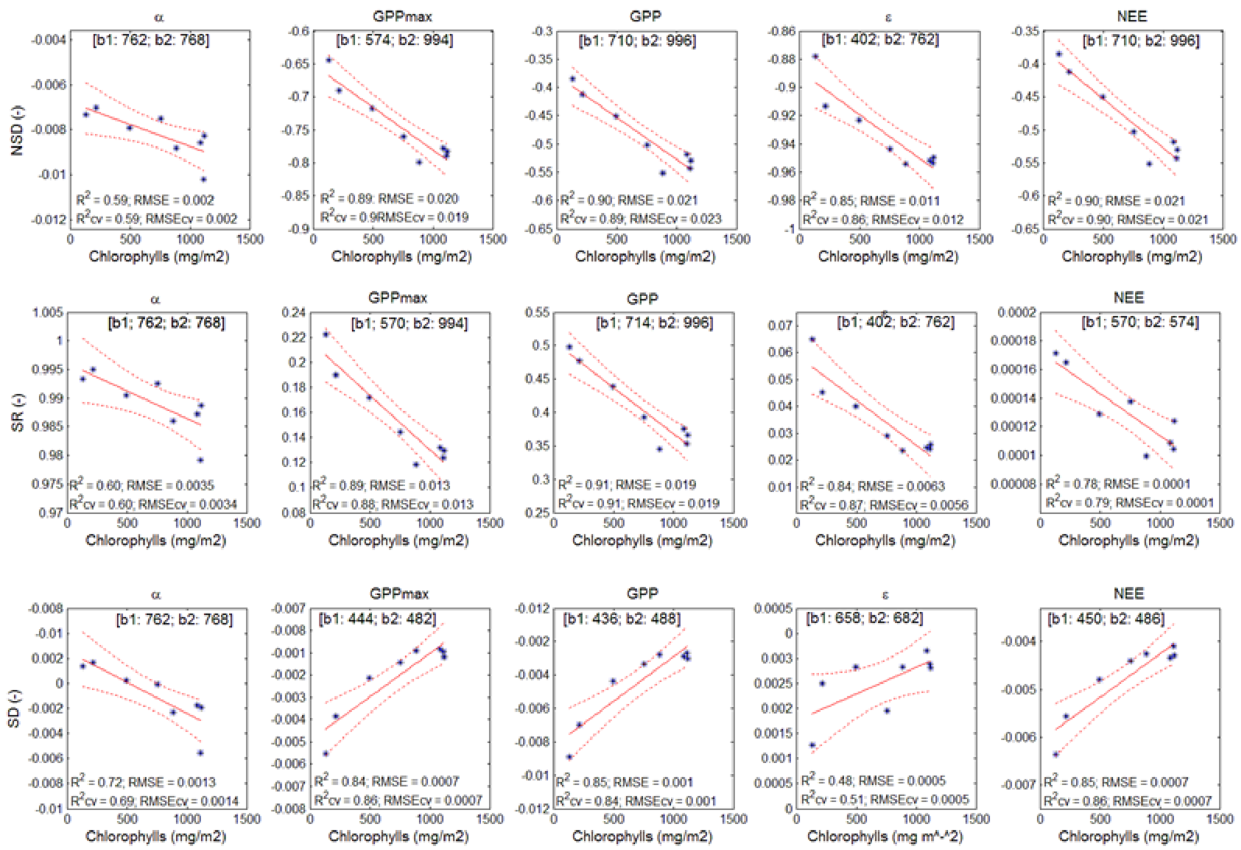

Fig. C3 – Correlation between selected NSD-, SR- and SD-type indices and the total chlorophyll content content for Monte Bondone in 2013. R^2—coefficient of correlation; RMSE—root mean square error; R^2cv— cross-validated coefficient of correlation; RMSEcv— cross-validated root mean square error. The red lines indicate the fitted models and the red dotted lines represent the 95% upper and lower confidence bounds. In the brackets are reported the selected bands to compute NSD-, SR- and SD-type indices.

In the revised version of the manuscript CI index reported in the Table 2 will be re-defined using the following formula: CI = (R750/R720) – 1 as reported in Gitelson et al. (2005). In addition, Red-edge (Red-edge NDVI = (R750–R720)/ (R750+R720); Gitelson and Merzlyak (1994)) will be added to the elaboration. Consequently, Figure 1 of BGD paper will be modified by showing the positions of these new indices and removing the old CI. In addition, in the revised version of the manuscript the Table 3 and Table 4 of the BGD paper and the Table S1 in the BGD supplemental will be modified showing the results of the correlation analysis between biophysical parameters and the new indices.

- **P10334L24-28: Significance is not a good predictor of accuracy as significance can be improved by sample size.**

RESPONSE
We agree and we will correct the text in the revised version of the manuscript.

- **P10334L18-28: It seems that this model is extremely simple and this is why it fails. It is already well known that GPP is controlled by many different factors (temperature, water stress, etc.). One reason VIs are widely used is that they remove some variation (i.e. two different sets of reflectance can yield the same VI value). Thus, VIs may not capture all of the necessary variation to explain GPP. There are GPP models that use multiple VIs to help address each of these components.**

RESPONSE
We agree with this comment. However, our general aim was to understand if there were some spectral regions where VIs and BPCs showed the same performance for each site or for all sites pooled together. Our intention was also to suggest simple measurements to do at all the sites where sensors with few main bands are prioritized.

- **P10339L29-P103340L1-3: The MOD17 algorithm is a very low bar. Most researchers active in the field know it is too simplistic, thus for most site-specific applications, they do incorporate at least several of these aspects.**

RESPONSE
Agreed – sentence will be removed.

Tables

- **Table 2: Chlorophyll index is not a normalized difference VI. It is more similar to simple ratio with the exception of the ratio being subtracted by 1. The CI presented in the table would be more accurately called the Red Edge NDVI. A better citation for CI would be Gitelson et al. 2005, doi:10.1029/2005GL022688.**

RESPONSE
Thanks for the correction. We agree and we will modify the table 2 accordingly. CI will be moved in the block of simple ratio indices and will be defined as: CI= (R750/R720) – 1. The reference will be changed in in Gitelson et al. (2005). Red-edge NDVI will be added to the block of normalized different vegetation indices and will be defined as: Red-edge NDVI = (R750–R720)/ (R750+R720) (Gitelson and Merzlyak (1994).
Figure 1 of BGD paper will be modified by showing the positions of these new indices and removing the old CI. In addition, in the revised version of the manuscript the Table 3 and Table 4 of the BGD paper and the Table S1 in the BGD supplemental will be modified showing the results of the correlation analysis between BPCs and the new indices.

<u>The following new Table 2 will be added to the revised version of the manuscript:</u>

Index name and acronym	Formula	Use	Reference
Simple Spectral Ratio Indices			
Simple Ratio (SR or RVI)	$SR = R_{830}/R_{660}$	Greenness	Jordan (1969)
Green Ratio Index (GRI)	$GRI = R_{830}/R_{550}$	Greenness	Peñuelas and Filella (1998)
Water Index (WI)	$WI = R_{900}/R_{970}$	Water content, leaf water potential, canopy water content	Peñuelas et al. (1993)
Simple Ratio Pigment Index (SRPI)	$SRPI = (R_{430})/(R_{680})$		Peñuelas et al. (1995)
Chlorophyll Index (CI)	$CI= (R_{750}/R_{720}) - 1$	Chlorophyll content	Gitelson et al. (2005)
Normalized Spectral Difference Vegetation Indices			
Normalized Difference Vegetation Index (NDVI)	$NDVI= (R_{830}–R_{660})/ (R_{830}+R_{660})$	Greenness	Rouse et al. (1973)
Normalized Phaeophytinization Index (NPQI)	$NPQI = (R_{415}–R_{435})/ (R_{415}+R_{435})$	Carotenoid /Chlorophyll ratio	Barnes et al. (1992)
Normalized Pigment Chlorophyll Index (NPCI)	$NPCI = (R_{680} - R_{430})/ (R_{680} + R_{430})$	Chlorophyll ratio	Peñuelas et al. (1994)
Red-edge NDVI (Red-edge NDVI)	$\text{Red-edge NDVI} =(R_{750}–R_{720})/ (R_{750}+R_{720})$	Chlorophyll content	Gitelson and Merzlyak (1994)
Structural Independent Pigment Index (SIPI)	$SIPI=(R_{800}–R_{445})/ (R_{800}+R_{445})$	Chlorophyll content	Peñuelas et al. (1995)

REFERENCES

Gitelson, A. A., Viña, A., Ciganda, V., Rundquist, D. C., and Arkebauer, T. J.: Remote estimation of canopy chlorophyll content in crops, Geophys. Res. Lett., 32, L0840, doi:10.1029/2005GL022688, 2005.

Gitelson, A. and Merzlyak, M. N.: Quantitative experiments estimation of chlorophyll-a using reflectance spectra: experiments with autumn chestnut and maple leaves, J. Photochem. Photobio., 22, 247–252, 1994.

- **Tables 3 and 4: These are too complex to be able for readers to digest easily. Eliminate the exponential relationships. The number of significant digits is not appropriate for all metrics. For example the RMSE for _ and " is not 0.0. Readers cannot make any valid comparisons with insufficiently presented tables.**

Eliminate the AIC from the tables as the AIC values should only be compared between models with increasing complexity using the same data set (i.e. the same VI estimating the same BPC). The table makes it appear that these AIC values can be compared across VIs, when this is not the case due to the different values/dynamic ranges of VIs.

RESPONSE
We agree and we will correct these tables in the revised manuscript.

Figures

- **Figure 1: It is difficult to read the VI text on the figure. Define all abbreviations in figure captions so readers do not need to find them in the text.**

RESPONSE
We agree and we will add the VIs definition in the figure caption.

- **Figures 4-9: Use the figures in the supplemental. The information in the poor relationships are just as valuable.**

RESPONSE
Thanks for the suggestion. We agree and we will substitute figures 4-6 with equivalent figures of the supplemental (figures from S1 to S3). Figures 7-9 will be removed from the text and equivalent figures of the supplement (figures from S7 to S9) will be removed as well.

- **Figure 10: It does not matter if the model is more "accurate" if a significant portion of the dynamic range is insensitive to changes in GPP. It would be helpful to readers if a figure using the proposed VIs vs. BPCs were presented.**

RESPONSE
Thank for the suggestion, the Figure 10 of GBD paper will be removed. A new figure C1 (see above) using the proposed VIs vs. BPCs for each site and all sites pooled together will be added to the revised version of the manuscript. The fitting models and the cross validation metric are presented in the figure.

Other Notes:

- **These correlation matrices are not ideal for identifying the best bands except for very simple cases. An approach that would have yielded a more informative conclusion would be a GA-PLS analysis which can provide insight into more complex interactions between different wavebands.**

RESPONSE
We believe that genetic algorithms (i.e. GA-PLS) are stronger techniques than correlation matrices to identify best bands. In the revision version of the paper we will explore also the use of hybrid feature selection strategy based on genetic algorithm and random forests (GA–rF). The first method was used for the feature selection and the second as regression for predicting the target variables. First of all we aggregated the original dataset to 10 nm in order to lessen the effects of spatial autocorrelation. Li et al. 2010 suggested for the same purpose the use of the information theory. This approach is limited to a specific measurements and can't be applied for this our study since we aimed to compare multiple sites and find generalizable hot spots in the spectral domain. The genetic algorithm is based on an evolutionary principle: "the survival of the fittest". It generates a number of possible model solutions

(chromosomes) and uses these to evolve towards an approximation of the best solution of the model. In our case the genes of each chromosome are the wavelengths. We made use of 5 genes for each chromosome. We opted for such a length to overcome overfitting. Each population of 1000 chromosomes evolved for 200 generations. The mutation chance was set to the inverse of population size increased by one. The fitness of each chromosome was measured by means of the applying the random forest algorithm (Breiman, 2001). This was used as ensemble method for regression that is based on the uncontrolled development of decision trees (n=100). We opted for this method because of it is demonstrated the efficiency with large datasets. In combining the two methods we choose the mean squared error as target variable.

The new section 2.4 “Band selection based on the combination of random forests and genetic algorithm (GA–rF)” containing the above information will be added to the revised version of the manuscript.

REFERENCES

Breiman, L., Random Forests, Mach. Learn., 45, 5-32, 2001.

Li, S., Wu, H., Wan, D., Zhu, J.: An effective feature selection method for hyperspectral image classification based on genetic algorithm and support vector machine, Knowl.-Based Syst., 24, 40–48, 2011.

Technical Corrections:

- **P10325L4: The word “lately” implies very recent papers, 2007-2010 are recent, but not very recent. Delete the word or find more recent publications.**

RESPONSE
Thanks for the correction. We will delete this word in the revised manuscript.

- **P10326L12: Misplaced comma after “and”**

RESPONSE
Thanks, we will correct it in the revised manuscript.