

Reply to anonymous Referee#1

We thank referee#1 for reviewing the paper and we address in detail the reviewer's concerns. As suggested, the manuscript was revised in order to improve clarity and readability, and to correct English grammar

1. "By focusing on VIs make the novelty of the work questionable. There are loads of similar paper."

5 Answer:

We agree with the referee that numerous studies have already been published on the use of spectral vegetation indices (VIs) for GPP estimation and frequently they have shown a limited applicability. Some of these studies are reported in the introduction of our manuscript. However, our study intends to explore the use of spectral information beyond the use of known VIs. We also assessed the potential of reflectance values in the 350-2500nm range of the spectra to estimate GPP, grouped in intervals of contiguous highly correlated reflectance (named "bands"). Our results (see Table 5) compare the best models, considering VIs, bands and the combination of both. Moreover, we apply statistical tests to determine if there are significant differences between models. As far as we know, this provides a new understanding of the potential use and limitations of VIs to model GPP, and explains at which extent VIs can be replaced or complemented by bands.

- 15 2. "The methodology used to choose relevant variables for model construction seems not optimal. There is a dimensionality reduction done at 3 different levels while that could be done in one step by choosing a more appropriated method"

Answer:

In fact, the reduction of dimensionality was achieved in progressive steps. Firstly, we determine groups of wavelengths (we call them "bands") that satisfy a clear criteria: they are the smallest set such that correlation between observations is always larger than 90% within each band. This allows us to preserve the spectral continuity of the signal, since we aggregate contiguous individual 1 nm intervals of wavelength into broader wavelength bands. Secondly, we select the bands and/or VIs that best explain GPP. In both cases, we apply methods which are optimal for the criteria we consider. For selecting the best subset of variables in Multiple Linear Regression (MLR) models and further reduce the dimensionality of the data, we rely on the leaps and bounds technique (LEAPS), which is suitable for large number of predictor variables (up to approximately 50). We haven't used Principal Component Analysis (PCA), which has been suggested by the referee for dimensionality reduction, since PCA replaces the original variables by linear combinations. Given that the goal of the analysis is to determine the set of VIs and /or bands that best model GPP, we believe that we need to apply a method that identifies the best subset of single variables for a straightforward interpretation of results. Unlike alternative heuristic approaches (Cadima et al., 2004), LEAPS returns the optimal subset of predictors according to a given criteria (we used the adjusted R^2 as the criteria in our study). The referee also suggests using Machine Learning (ML) algorithms. Although the expressiveness of those models (ML) is stronger than MLR, we believe that linear models are better at making explicit the relation between predictors and GPP, while offering enough flexibility by including variables in a high dimension representation space. Moreover, linear models allow us to derive confidence intervals for our results, apply statistical tests to

compare models at a given significance level, and prevent overfitting. Finally, the analysis of residuals for MLR doesn't contradict the assumptions of the linear model, which confirms that MLR is an adequate approach for the problem at hand.

After the best MLR model (with the highest adjusted R^2) is determined, we do simplify it by exhaustive backwards stepwise selection of predictors at a 5% significance level. We agree with the reviewer that it would be more elegant to incorporate this directly into the variable selection procedure. This could have been done if we had used a heuristic stepwise selection method, which would not guarantee however the optimality of the solution. Here, we chose instead to use LEAPS to find the optimal model (the one that maximizes the adjusted R^2), and then we identify all its sub models that do not differ in quality at the 5% significance level and choose the best and most parsimonious one.

The following text was inserted in section 2.5 (Data analysis):

10 Given that the goal of the analysis is to determine the set of VIs and /or bands that best model GPP, we identify the best subset of explanatory variables to model the response variable (GPP) by multiple linear regression (MLR). We preferred MLR to other methods, (e.g. Principal Component Analysis, PCA) where dimensionality reduction is achieved through replacing variables by their linear combinations. MLR was also preferred to non-linear models (e.g. in the field of machine learning) because it provides a clearer interpretation of the relation between predictors and GPP, while offering enough flexibility by including variables in a high dimension representation space. Moreover, MLR allows us to derive confidence intervals for our results, apply statistical tests to compare models at a given significance level, and it is less prone to overfitting than complex non-linear models.

3- "There is nothing in the discussion about the influence of the fertilization treatment on PAI and GPP" – "Fertilization has an effect on PAI(gr) but not on GPP - PAIgr continues to increase with P treatment until May20 (and not with NPK treatment) Indication that P fertilization enables to keep photosynthetic active leaves for a longer period?"

Answer:

We appreciate the comments of the referee on the subject. As suggested, we added a discussion point on the impact of fertilization treatment on green biomass, duration of the growth cycle and NEE. The text added to the discussion is :

25 4.1 The impact of fertilization treatment

The fertilization treatment influenced the growth rate and the composition of the herbaceous plots more than carbon sequestration. In line with a five year and ongoing study at the same grassland site, the fertilization treatment resulted in differences in aboveground biomass and functional groups proportion for NPK and P treatments, while the single addition of N had no effect. An earlier growth response was also observed in the NPK treatment (larger PAIgr at the first measurement day). Higher percentage of graminoid species with on average higher growth rates as compared to most forb species (Ansquer et al., 2009; Craine et al., 2001; Westoby et al., 2002) could explain early differences in PAI and PAIgr. As leaf area is usually positively related to GPP (e.g. Aires et al., 2008; Xu and Baldocchi, 2004), it would be expected that higher PAIgr in NPK treatments would induce in increased GPP. However, confounding factors such as increased water demand

associated with higher growth rates in NPK might have downscaled differences between treatments (e.g. Weisser et al., 2017).

While the fertilization treatment had no impact on NEP or GPP a higher R rate was observed at the 4th experimental day (June, 3) in the NPK and P treatments. Differences can be ascribed to the higher PAI, and precipitation at the end of May, which must have stimulated soil respiration (Jarvis et al., 2007; Reichstein et al., 2003).

4-“I do not really get why it is interesting to calculate different VI values for different treatments? (fig. 5)”

Answer: Indeed, no significant differences were observed in GPP among treatments. However, since the discussion was extended focusing on the impact of the treatments on vegetation growth, greenness and carbon sequestration, we consider more interesting and easier for the reader to keep differences among treatments and dates in figure 5. The legend was corrected into:

Fig 5. Average values of several vegetation indices retrieved from reflectance measurements of herbaceous plots undergoing different fertilization treatments. Vertical bars represent standard errors. Different letters indicate significant differences among treatments within the same date (p<0.05).

Detailed comments:

Page 1 line 20. Define LEAPS. Ok.

Page 1 line 21-23. Rephrased.

Page 1 line 24. Corrected.

Page 1 line 28. It is not clear what the referee meant with the comment: “what about the simulated hyperspectral sensors?”. In this study, we did not simulate hyperspectral sensors. The hyperspectral measurements were collected in the field by using a Field Spec 3 portable equipment.

Page 2 line 9. Corrected.

Page 2 line 22. Corrected

Page 2 line 31. We thank the referee for the observation. Indeed, one of the references about the vegetation index PRI mentioned in the introduction relates to a leaf level study (Peñuelas et al., 1995). The sentence was rephrased and references were changed accordingly. Many studies applied PRI at the canopy scale as shown by the useful review from Galbursky et al (2011) on the use of PRI in different biomes and scale (leaf/canopy/ecosystem) which is now cited in the text. In that review, authors shows that the relationship between PRI and Light Use-Efficiency holds across spatial and temporal scales, in spite of a variable strength of relationship.

Page 3 line 26. Thanks for pointing out this incorrect statement. We have replaced the previous sentence by “The recently launched S2 covers the regions of the visible and near-infrared and the SWIR in 13 bands with at least five days revisiting time for the combination of S-2A and S-2B platforms (Drusch et al., 2012)”

Page 3 line 29. Changed

Page 3 line 31. Done

Page 4 line 5. Done.

Page 5 line 8. Rephrased

Page 5 line 27. The theoretical link between NEE and GPP is explained at the bottom of the paragraph, page 6 line 6.

5 Page 6 line 4 and 5. The necessity for a standardized time window is explained in Perez-Priego (2015), cited in the manuscript. In addition the linear trend during the measurement period was also visually verified at each measurement.

Page 6 line 6. Field spectroradiometric measurements need to be collected close to the sun zenith and under clear sky conditions. Gas exchange measurements were performed as close as possible to Fieldspec measurements to increase the representativeness of spectral data.

10 Page 6 line 10. The paragraph title was rephrased and PAI defined. The equipment used for indirect measurement of leaf area, the ceptometer, estimates LAI through the indirect measurement of the interception of light by all vegetation structures above ground (leaves, stalks, and eventually flowers, fruits and seeds). Therefore, the term Plant Area Index (PAI) was preferred to LAI (Bréda, 2003).

Page 6 line 15. The fPAR is defined as the fraction of PAR intercepted by the canopy at line 12.

15 Page 6 line 21-22. As explained in the text (2.4.2), the light extinction coefficient, K depends on both the canopy structure and the zenith angle, which is the angle of the sun with a vertical line at the local where measurements are being taken. In this study, while the zenith angle is calculated with basis on the geographical coordinates of the local and date and the time of measurements, the canopy structure was considered constant and equal for all plots, assuming a “spherical leaf distribution”. This concept is defined by Jones (1992) as the leaf canopy distribution in which leaves have equal probability of any orientation in such a way that could be envisioned as rearranged on the surface of a sphere. A reference was added to

20 the text.

Page 7 line 6 and 7. In this paragraph (2.4.3) the equipment used for the collection of hyperspectral measurements (FieldSpec3, ASD) is described, as well as technical specifications, data collection and processing. (see also Cerasoli et al., 2016; Jongen et al., 2013). The paragraph was rephrased for clarity. The new text in this section is:

2.4.3 Hyperspectral measurements of vegetation reflectance

25 At each field campaign, hyperspectral observations of all plots were also acquired with a FieldSpec3 spectroradiometer (ASD Inc., Boulder, USA), which provides reflectance of vegetation in the range of 350-2300 nm. The spectral resolution (Full-Width-Half-Maximum) is 3 nm at 700 nm and 10 nm at 1400 nm and 2100 nm. The sampling interval is 1.4 nm for the spectral region of 350-1000 nm (visible and near infrared) and 2 nm for the spectral region of 1000-2500 nm (short-wave infrared). A white reference of known reflectance (Spectralon panel, Labsphere, Inc., North Sutton, USA) was used to

30 normalize for variations in atmospheric conditions and to convert the measurements into absolute reflectance. Spectra were collected using a bare fibre optical cable (with an instantaneous field of view of 25°) inserted into a pistol grip at approximately 90 cm above the canopy and a nadir view.

Five spectra were recorded for each plot, each one representing the average of 25 observed spectra, All measurements were conducted immediately after grassland gas exchange measurements, within two hours around solar noon, to minimize the effects of shadowing and solar zenith changes.

5 The term absolute reflectance refers to the reflectance measurements after correction by multiplying by the spectrum of the calibration panel (Spectralon) measured under similar conditions. (details are available in the Field Spec3 user manual).

Page 7, line 15. As explained in the referred literature a linear mixed effects model is a linear model that incorporates both fixed and random effects (Bates et al., 2014). We believe that the term is self-explanatory for most readers. Additional details can be found in the cited literature. Random and fixed effects considered are also reported in the text (Page7, line 16).

10 Page 7, line17. Corrected

Page7, line 24. We have replaced “dimensionality” by “the number of predictors”.

Page7. Line 25. The output of the spectroradiometer are reflectance values at 1nm interval. No further post-processing was adopted before the statistical analysis reported in section 2.5. The sentence referring to a cubic spline interpolation was deleted.

15 Page 7, line 27-Page8, line 3. We have simplified the text and clarified the mathematical procedures described in this section.

The new text is:

The full spectra of vegetation reflectance retrieved from the Fieldspec was used to model GPP, after excluding noisy values in the range 1350-1400 nm and 1800-1950 nm. Our P=1748 original explanatory variables are x_{350}, \dots, x_{2299} where x_i

20 represents the reflectance in the narrow band $[\lambda, \lambda + 1]$ (nm) and our response variable is the GPP ($\mu\text{mol m}^{-2} \text{s}^{-1}$). A total number of 96 observations were available (4 treatments X 2 replicates X 3 blocks X 4 dates). Since we have 1748 explanatory variables and just 96 observations, hence a high level of redundancy in our data, the number of predictors was

first reduced by grouping variables that belong to intervals of wavelengths where all variables are highly correlated. A hierarchical cluster analysis was performed to reduce the number of predictors from P=1748 to P=25 groups of contiguous

25 variables named “bands”. The basic idea is to aggregate contiguous and highly correlated individual 1 nm intervals of wavelength into broader wavelength bands. Two original predictors $x_{\lambda a}, x_{\lambda b}$ are clustered together if their correlation

coefficient $r(x_{\lambda a}, x_{\lambda b})$ is larger than 0.90. Bands correspond to the largest contiguous intervals where all pairs of original predictors satisfy that condition. If a band groups all original predictors between λ_1 and λ_2 , then it is represented by a new

30 variable $x_{[\lambda_1, \lambda_2]}$, which is the arithmetic mean of the original variables $x_{\lambda_1}, \dots, x_{\lambda_2}$. The procedure is repeated to obtain bands that partition the full x_{350}, \dots, x_{2299} spectrum.

Page 8, line 4-10. The referee suggests moving the paragraph to the introduction. The purpose of this paragraph was only to provide the reader with a list of the vegetation indices selected for this study. To avoid any overlap with the introduction, the

paragraph was summarized and the reader is invited to read table 2 where the name, equation, reference and biophysical property represented are reported for each vegetation index.

Page 8, Line 15. The referee questioned the utility of a multiple linear regression adopted to establish mathematical relationships between the explanatory variables (bands and VIs) and the response variable (GPP) suggesting the adoption of machine learning algorithms.

Please see the comments at the beginning of this reply, which addresses the questions raised by the referee.

Page 8 line 15. The referee asked for more information about the LEAPS algorithm. The text was edited to clarify that LEAPS performs an exhaustive search for the best subset of predictor variables x of y in linear regression, hence it is useful to reduce the dimensionality of predictors while keeping the original variables. Further details can be found in the cited literature.

Page 8 line 22. The referee criticizes the procedure adopted for determining the most parsimonious sub model. Please see the comments at the beginning of this reply, which address the criticism.

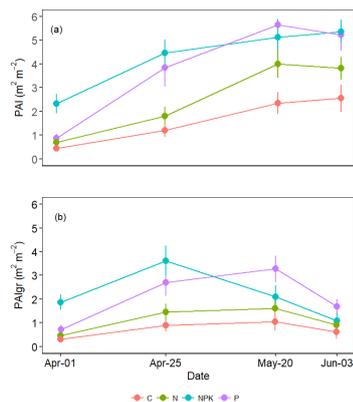
Page 8 line 32. Corrected.

Page 9 line 8. The year was added.

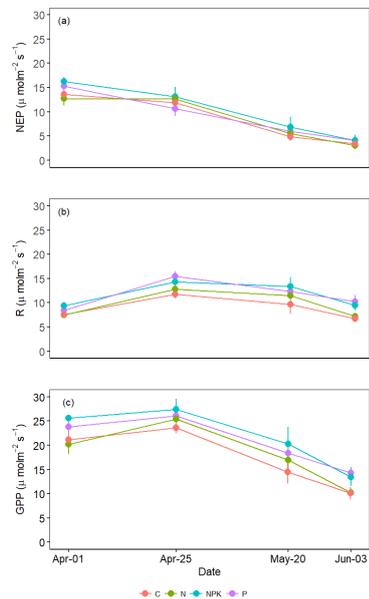
Page 11 line 3. The sentence was deleted.

Page 13 line 3. Corrected.

Page 28 Fig.2. The figure was changed as suggested and equal y-axis was adopted for the two plots. The new figure is shown below:



Page 29 Fig.2. In this study we adopted the atmospheric sign convention where the Net flux is considered negative when detracting CO₂ from the atmosphere (Baldocchi, 2008). In order to improve the readability of the figure, we changed to the NEP (Net Ecosystem Productivity) which is considered positive. The figure and text were changed accordingly. The new figure is shown below:



Page 31 Fig. 5. The legend was corrected. The VI for each picture is indicated on y-axis and the fertilization treatment is indicated below the figure. The designation of the VI, the biophysical property represented and reference are listed in table 2.

5 Please see comments above.

The English grammar was accurately reviewed.

References:

- Aires, L. M. I., Pio, C. A. and Pereira, J. S.: Carbon dioxide exchange above a Mediterranean C3/C4 grassland during two climatologically contrasting years, *Glob. Chang. Biol.*, 14(3), 539–555, doi:10.1111/j.1365-2486.2007.01507.x, 2008.
- 10 Ansquer, P., Duru, M., Theau, J. P. and Cruz, P.: Convergence in plant traits between species within grassland communities simplifies their monitoring, *Ecol. Indic.*, 9(5), 1020–1029, doi:10.1016/J.ECOLIND.2008.12.002, 2009.
- Baldocchi, D.: TURNER REVIEW No. 15. ‘Breathing’ of the terrestrial biosphere: Lessons learned from a global network of carbon dioxide flux measurement systems, *Aust. J. Bot.*, 56(1), 1–26, doi:10.1071/BT07151, 2008.
- 15 Bates, D., Mächler, M., Bolker, B. and Walker, S.: Fitting Linear Mixed-Effects Models using lme4, *J. Stat. Softw.*, 67(1), 1–48, doi:10.18637/jss.v067.i01, 2014.
- Bréda, N. J. J.: Ground-based measurements of leaf area index: A review of methods, instruments and current controversies, *J. Exp. Bot.*, 54(392), 2403–2417, doi:10.1093/jxb/erg263, 2003.
- Cadima, J., Cerdeira, J. O. and Minhoto, M.: Computational aspects of algorithms for variable selection in the context of principal components, *Comput. Stat. Data Anal.*, 47(2), 225–236, doi:10.1016/J.CSDA.2003.11.001, 2004.
- 20 Cerasoli, S., Costa e Silva, F. and Silva, J. N.: Temporal dynamics of spectral bioindicators evidence biological and ecological differences among functional types in a cork oak open woodland, *Int. J. Biometeorol.*, 1–13, doi:10.1007/s00484-015-1075-x, 2016.

- Craine, J. M., Froehle, J., Tilman, D. G., Wedin, D. A. and Chapin, III, F. S.: The relationships among root and leaf traits of 76 grassland species and relative abundance along fertility and disturbance gradients, *Oikos*, 93(2), 274–285, doi:10.1034/j.1600-0706.2001.930210.x, 2001.
- 5 Drusch, M., Del Bello, U., Carlier, S., Colin, O., Fernandez, V., Gascon, F., Hoersch, B., Isola, C., Laberinti, P., Martimort, P., Meygret, A., Spoto, F., Sy, O., Marchese, F. and Bargellini, P.: Sentinel-2: ESA's Optical High-Resolution Mission for GMES Operational Services, *Remote Sens. Environ.*, 120, 25–36, doi:10.1016/j.rse.2011.11.026, 2012.
- Jarvis, P., Rey, A., Petsikos, C., Wingate, L., Rayment, M., Pereira, J., Banza, J., David, J., Miglietta, F., Borghetti, M., Manca, G. and Valentini, R.: Drying and wetting of Mediterranean soils stimulates decomposition and carbon dioxide emission: the "Birch effect", *Tree Physiol.*, 27(7), 929–940, doi:10.1093/treephys/27.7.929, 2007.
- 10 Jones, H. G.: *Plants and microclimate: A quantitative Approach to Environmental Plant Physiology*, 2nd ed., Cambridge University Press, Cambridge, 1992.
- Jongen, M., Unger, S., Fanguero, D., Cerasoli, S., Silva, J. M. N. and Pereira, J. S.: Resilience of montado understorey to experimental precipitation variability fails under severe natural drought, *Agric. Ecosyst. Environ.*, 178, 18–30, 2013.
- 15 Peñuelas, J., Filella, I. and Gamon, J. A.: Assessment of photosynthetic radiation-use efficiency with spectral reflectance, *New Phytol.*, 131(3), 291–296, doi:10.1111/j.1469-8137.1995.tb03064.x, 1995.
- Pérez-Priego, O., López-Ballesteros, A., Sánchez-Cañete, E. P., Serrano-Ortiz, P., Kutzbach, L., Domingo, F., Eugster, W. and Kowalski, A. S.: Analysing uncertainties in the calculation of fluxes using whole-plant chambers: random and systematic errors, *Plant Soil*, 393(1–2), 229–244, doi:10.1007/s11104-015-2481-x, 2015.
- 20 Reichstein, M., Rey, A., Freibauer, A., Tenhunen, J., Valentini, R., Banza, J., Casals, P., Cheng, Y., Grünzweig, J. M., Irvine, J., Joffre, R., Law, B. E., Loustau, D., Miglietta, F., Oechel, W., Ourcival, J.-M., Pereira, J. S., Peressotti, A., Ponti, F., Qi, Y., Rambal, S., Rayment, M., Romanya, J., Rossi, F., Tedeschi, V., Tirone, G., Xu, M. and Yakir, D.: Modeling temporal and large-scale spatial variability of soil respiration from soil water availability, temperature and vegetation productivity indices, *Global Biogeochem. Cycles*, 17(4), n/a-n/a, doi:10.1029/2003GB002035, 2003.
- 25 Weisser, W. W., Roscher, C., Meyer, S. T., Ebeling, A., Luo, G., Allan, E., Beßler, H., Barnard, R. L., Buchmann, N., Buscot, F., Engels, C., Fischer, C., Fischer, M., Gessler, A., Gleixner, G., Halle, S., Hildebrandt, A., Hillebrand, H., de Kroon, H., Lange, M., Leimer, S., Le Roux, X., Milcu, A., Mommer, L., Niklaus, P. A., Oelmann, Y., Proulx, R., Roy, J., Scherber, C., Scherer-Lorenzen, M., Scheu, S., Tschardtke, T., Wachendorf, M., Wagg, C., Weigelt, A., Wilcke, W., Wirth, C., Schulze, E.-D., Schmid, B. and Eisenhauer, N.: Biodiversity effects on ecosystem functioning in a 15-year grassland experiment: Patterns, mechanisms, and open questions, *Basic Appl. Ecol.*, 23, 1–73, doi:10.1016/j.baec.2017.06.002, 2017.
- 30 Westoby, M., Falster, D. S., Moles, A. T., Vesk, P. A. and Wright, I. J.: Plant Ecological Strategies: Some Leading Dimensions of Variation Between Species, *Annu. Rev. Ecol. Syst.*, 33(1), 125–159, doi:10.1146/annurev.ecolsys.33.010802.150452, 2002.
- Xu, L. and Baldocchi, D. D.: Seasonal variation in carbon dioxide exchange over a Mediterranean annual grassland in California, *Agric. For. Meteorol.*, 123(1–2), 79–96, doi:10.1016/j.agrformet.2003.10.004, 2004.

35

Reply to anonymous Referee#2

We thank the reviewer for the comments/suggestions that much improved the manuscript. We reply to referee#2 comments point by point.

Comment1:

- 5 “The study design introduces nutrient treatments which influence species composition GPP to some degree but the successive analyses discard this structure when testing the skill of reflectance metrics as predictors of GPP. GPP varies over time, across replicates and across treatments but all of this variability is pooled when testing the reflectance metrics. The paper might provide more insights by structuring the analysis to test treatment and temporal variability separately. Perhaps a repeated measures ANOVA could help, for example, by testing for significant effects of date, treatment, and one reflectance metric on variability in GPP. This would be akin to the linear mixed effect analysis of GPP and PAIgr”.

10 Answer: The ability of reflectance metrics in predicting GPP were analyzed altogether without considering different treatments and dates because the objective of the study was to test the skill of spectral information to represent GPP across treatments and time. However no statistical differences were found in GPP among treatments (see section 3.3).

15 Comment2:

- “GPP and PAIgr both vary over time and PAIgr varies across treatments (GPP appears to as well but apparently the statistical testing does not support this). These patterns are displayed well with Figs 2 and 3 but what is missing is display of a scatter of GPP versus PAIgr, and also of GPP versus (selected) reflectance metrics. These relationships should be shown, with symbols that differentiate the treatments and display individual replicates. The relationship that emerges (slope) would offer insights about the effective light use efficiency per unit green leaf area. The term ‘effective’ here refers to the combination of a maximal LUE with any limitations by water, light, or nutrients. This is the sort of parameterization that would be needed in a functional model. In fact, it would be interesting to test if any of the reflectance metrics have skill in predicting variability in GPP / PAIgr, thus capturing patterns in LUE rather than just green plant area.”

- 25 Answer: We thank the referee for this valuable comment that improved our insight on the relationship between green plant area and GPP. The scatter plot produced is shown below.

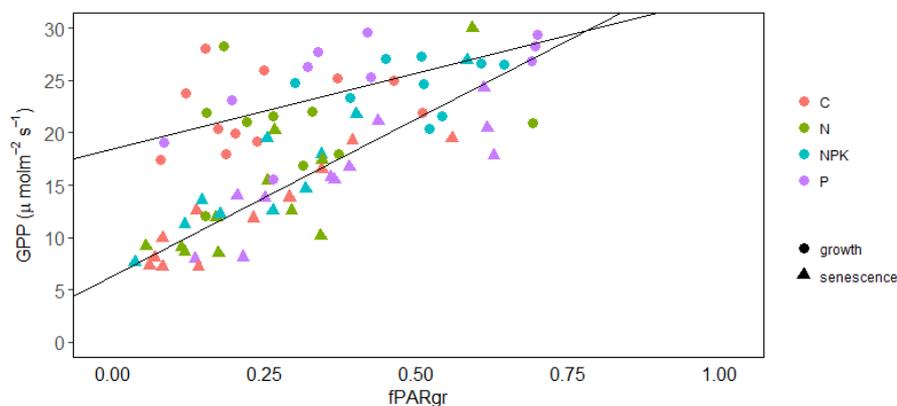


Figure caption: Relationship between GPP and fPAR_{gr} observed in grassland subjected to different fertilization treatment (C, N, NPK and P). Measurements performed during the vegetation growth are indicated as circles (corresponding to measurement days 1 and 2) while those realized during the senescence period are indicated as triangles (measurement days 3 and 4). Linear regression lines were fitted separately to the two periods ($GPP = 14.48 fPAR_{gr} + 18.44$, $R^2=0.39$, $P<0.001$ and $GPP = 30.08 fPAR_{gr} + 6.39$, $R^2=0.65$, $P<0.001$, respectively for the growth and the senescence period).

The GPP-fPAR_{gr} relationship showed to be different during the vegetation growth (circles in the figure) and the senescence phase (triangles in the figure).

The differences observed in the slope of the regression lines reveal the occurrence of marked changes in the “effective LUE” along the growth life cycle and confirms previous results (Nestola et al., 2016; Perez-Priego et al., 2015).

The figure will be integrated in the manuscript and the text modified accordingly.

In addition, we also verified the correlation between LUE and selected vegetation indices calculated (GNDVI, MTCI, NDWI, PSRI) (See table below). For all vegetation indices correlations are weak.

Table: Correlation coefficient (*R*) between LUE and selected vegetation indices observed in herbaceous plots undergoing different fertilization treatments.

VI	<i>R</i>
GNDVI	-0.145
MTCI	-0.076
NDWI	-0.214
PSRI	-0.022

Altogether these results further confirm the need for an empirical approach, instead of a classic LUE model, to assess the ability of spectral retrievals for GPP estimation.

We agree that the approach has limited applicability as compared to a full functional model. On the other hand it allows testing different spectral metrics, independently of their ability to represent green plant area or LUE.

Comment3:

“Table 5 shows the skill of various multivariate linear models that include a suite of reflectance metrics selected to represent those available from different observing system types. This is a highly empirical approach to analysis and does not seem particularly useful in my opinion. The results are likely to be very heavily tuned to the specific dataset on hand and is not likely to be generalizable beyond the current study. For example, the Hyp-B step one selection includes a simple, linear model involving 13 unique bands. Biophysical or ecological functional models tend to use one or two metrics to represent structural (PAI_{gr}) and functional (LUE) attributes of an ecosystem’s capacity for primary production. This paper’s approach throws every possible indicator and combination at the variability in the data and thus lends little practical insight into the theory with very limited capacity for transferability. A more thoughtful approach grounded in theory and practice would be more useful.”

Answer: We agree with the referee that our results are tuned to our data set and results cannot be generalized. However, our dataset concerns herbaceous communities with a wide range of species composition and growth rates, resulting from the fertilization treatments. Hence, our empirical results can be considered valid for Mediterranean grasslands in a wide variety of situations. Also, the theoretical approach based on the LUE model already showed, in several studies (examples are reported also in the manuscript introduction), a limited capacity in expressing dynamic GPP changes in grasslands. This is also confirmed by the GPP-fPAR_{gr} relationships and the low correlation between LUE and VIs observed (see answer to comment 2).

We agree with the reviewer that considering reflectance bands, obtained by grouping highly correlated reflectance can be impractical but, in our opinion, it presents the advantage of indicating areas of the spectra of potential interest for GPP estimates in grasslands. For example, our results show a large influence of bands in the SWIR region in both step one and two (see Table 5). These results provide basis for further studies in the area.

5 **Comment4:**

“The study’s test of linear models includes VIs and bands, but not band ratios. Given that the approach is highly empirical in nature, there does not seem to be a good reason to omit band ratios or other simple mathematical combinations of bands (e.g. unique normalized difference ratios). Testing a wider range of combinations could be warranted to see if any other indicators happen to rise to the top in terms of predictive skill.”

10 Answer: As explained in the manuscript, we considered VIs that are widely used. We didn’t want to expand even more the set of explanatory variables, which was already large, to be able to interpret our results in terms of simple bands and well-known indices, while considering a large pool of predictors. Moreover, there is in general a 1 to 1 relation between VIs and band ratios. For instance, $NIR/RED=k$ implies that $NDVI=(k-1)/(k+1)$. Therefore, band ratios are implicitly incorporated into the model (although this is not equivalent to incorporating them as additional explanatory variables).

15 **Comment5:**

20 “The paper’s interpretations and conclusions suggest that bands are better than VIs as predictors of GPP but this is not reasonably supported by the quantitative results. Table 5 shows a small, marginal, and questionable increase in adjusted R2 for Hyp-B compared to Hyp-VIs, and a decrease in adjusted R2 for S2-B compared to S2-VIs. In any case, the differences in explanatory power over all of these cases is less than 0.0247 R2, or 2.5% of the variability in GPP, indicating that all are equally good at predicting GPP. For L8, a case might be made, however the band metric has many more variables thrown at the problem (6 bands compared to just NDVI) and when these other bands are included in a step two selection, the NDVI model with bands had high skill. Surely bands and VIs are equally skillful for the other observing system types. Corresponding edits need to be made to section 4.2.”

25 Answer: We thank the referee for this thoughtful comment on the discussion of results shown in table 5. The text was changed accordingly. The modified text is:

4.2 Are spectral bands better GPP estimators than VIs?

30 Our results suggest a marginal improvement in GPP estimates obtained adopting bands (Hyp-B, S2-B, Table 5) instead of vegetation indices (Hyp-VIs, S2-VIs, Table 5). A larger impact was observed in the case of L8-VIs+B models when compared with the L8-VIs model which included only NDVI (L8-VIs, Table 5).

35 Although normalized VIs are important in establishing strong relationships between biophysical and optical properties of vegetation, our results showed that the selection of the proper band is equally important to the mathematical formulation of the indices for the explanatory power of spectral retrievals as predictor variables. Previous studies comparing the explanatory power of VIs and bands in grasslands evidenced also the importance of the selection of the proper spectral range (Balzarolo et al., 2015; Matthes et al., 2015). The approach adopted in our analysis assured a high correlation among responses within a band, which determined that spectral bands were representative.

Comment6:

“One of the advantages of VIs is that they normalize for a wide range of background reflectance, sun-sensor geometry, and atmospheric effects in ways that direct bands do not. This point seems to be lost on the authors and is important for developing indicators that can be transferred to remote sensing (space or airplane) over large areas and across large gradients in surface and atmospheric conditions. Discussion about this should be included in the paper.”

5

Answer: We thank the referee for this valuable comment. The following sentence was included in the text.

However, VIs offer, in comparison with spectral bands, the advantage of providing spatial and temporal comparable representation of vegetation features, since differences resulting from background reflectance, sun-sensor geometry and atmospheric effects are minimized by normalization of spectral values.

10

Comment7:

“Akaike or Bayesian Information Criteria need to be adopted to evaluate the relative skill of the selected linear models, penalizing models that select more variables.”

15

Answer: We agree that models should be compared with indicators that penalize complex models and prevent overfitting. This is why we used the adjusted R^2 instead of the standard R^2 . The optimization techniques we considered (LEAPS in particular) can be applied using adjusted R^2 (as we did), AIC, BIC or other similar criteria. It is easy to show that stepwise selection always chooses, at each step, the same variable to be excluded or included in the model regardless of the criteria (AIC, BIC or adjusted R^2). In that restricted sense, they are equivalent among them but distinct from R^2 . Our preference for the adjusted R^2 over BIC or AIC was determined by the fact that we applied bootstrap techniques (available for the adjusted R^2) to derive confidence intervals for the goodness-of-fit of the models we fitted to the data.

20

Comment8:

“It is worth noting that soil moisture is essentially equal on all four reflectance observation dates, while temperature increased steadily from the first to the last observation date. Correspondingly, the statement on P12, L19 that suggests that the Hyp model represents changes in canopy water content might need to be revised. Canopy water content was not observed and soil water content did not differ over the four sampling dates. It is possible that canopy water content differed substantially from soil water content over this time series but that has not been established with quantitative, direct observations.”

25

30

Answer: The aim of the mentioned test is not to establish a quantitative relationship between canopy water content and the vegetation indices NDWI and WBI but to underline the known ability of these vegetation indices as indicators of canopy water content. On the other side, differences observed between fPAR and fPAR_{gr} reveal a progressive increase of dry biomass, which is related with the senescence process that depends more on increased temperatures than soil water content. The sentence was changed into:

35

The Hyp model also put in evidence the importance of changes in canopy water content, as both the NDWI (Gao, 1996) and the WBI (Penuelas et al., 1997) were included in the model. Considering changes observed in NDWI along the experiment and the good correlation observed between NDWI or WBI and GPP it is reasonable to assume that GPP is largely affected by the progressive senescence of vegetation, (Balzarolo et al., 2015; Vescovo et al., 2012). In a previous study, Vicca (et al., 2016) found that NDWI was able to estimate GPP in semiarid grasslands better than other indices, allowing to distinguish the effect of drought.

40

Comment9:

“It is unfortunate that the study did not include an additional observation period in the mid to late June as PAIgr continued to decline.”

5 Answer: After our last measurement PAIgr declined very fast, in mid June all plots were already dried. There was no time for a further round of measurements.

Comment10:

“The introduction is very well written and cited. One paper that might be useful to add to the framing and discussion is that of Asner et al. 2004 in PNAS (“Drought stress and carbon uptake in an Amazon forest :”)”.

10

Answer: The following sentence was inserted in the introduction.
A promising new technology is the use of spaceborne imaging spectroscopy. The hyperspectral resolution of these images allows to identify canopy properties with higher sensitivity than traditional vegetation indices. For example, the use of spaceborne imaging spectroscopy was able to detect changes in canopy leaf area and water stress in the humid tropical forest, while NDVI and other vegetation indices failed (Asner et al., 2004).

15

Comment11:

“Page 5, L12: “All nutrients were added at” seems to suggest N for nitrogen, or is N for nutrients here?”

Answer: The referee is right; it should be “All nutrients were added at a rate of 10 g. m⁻² yr⁻¹...”. In the text N refers to nitrogen. The sentence was corrected.

20

Comment12:

“Measurement of soil moisture at only 10 cm depth may not be adequate to represent the soil water content being experienced by the grassland plants. It would be best to also measure a deeper profile of moisture.”

25 Answer: We agree with the referee that the soil water measurements along a soil profile would have provided more detailed information. However our grassland is dominated by annual species with winter seasonality. In these systems mean rooting depth is approximately 20 cm (Schenk and Jackson, 2002) with most of the rooting zone at 10 cm depth as observed in a similar grassland in Portugal (Jongen et al., 2013).

Comment13:

30 “It is interesting that the nutrient treatments allegedly altered the functional composition of the grassland plots, however pre-treatment data are not presented and this would be essential to demonstrate that the compositional shifts were due to the treatments themselves. Unless it can be established with data, the corresponding statement (P9, L23) should be corrected to omit suggestion that the treatments caused the compositional differences.”

35 Answer: As mentioned in the manuscript, the grassland is part of the Nutrient Network experiment (www.nutnet.umn.edu) which is still on going. Since 2012 (pre-treatment), species composition is evaluated every year. In 2012 species composition showed a large dominance of forbs (73%), while graminoids (24%) and legumes (3%) were less represented. These results are only partially similar to the observed in the present study (Table 3) in which forbs were the dominant functional type

(56.85%), and grams and legumes have similar proportion (21.22% and 21.93% respectively). However, it must be also considered that species composition can change from year to year as a result of precipitation amount and distribution as observed also in previous studies (DeMalach et al., 2017). The sentence was rephrased into:

5 Plant species composition has been measured every year since 2012 (pre-treatment) under an ongoing long-term nutrient addition experiment on this grassland site. In line with results from that study (Nogueira et al., personal communication), the fertilization treatments influenced the functional composition of grasslands (Table 3). In the NPK treatment the percentage of graminoids was higher than in any of the other treatments. P treatment showed a higher percentage of legumes and in the C and N treatments forbs were the dominant functional group.

Comment14:

10 “It is surprising that treatment effect was significant for respiration but not for GPP considering that both have similar spreads and error bars. Double check results here.”

Answer: Results were double checked and they are correct.

On estimating Gross Primary Productivity of Mediterranean grasslands under different fertilization regimes using vegetation indices and hyperspectral reflectance

Sofia Cerasoli¹, Manuel Campagnolo¹, Joana Faria¹, Carla Nogueira¹, Maria da Conceição Caldeira¹

¹CEF, Centro de Estudos Florestais, Instituto Superior de Agronomia, Universidade de Lisboa, PT.

Correspondence to: Sofia Cerasoli (sofiac@isa.ulisboa.pt)

Key-words: hyperspectral reflectance, remote sensing, carbon sequestration, pasture

Running-title: GPP estimates in grasslands by optical sensors

Abstract: We applied an empirical modelling approach for Gross Primary Productivity (GPP) estimation from hyperspectral reflectance of Mediterranean grasslands undergoing different fertilization treatments. The objective of the study was to identify combinations of vegetation indices and bands that best represent GPP changes between the annual peak of growth and senescence dry out in Mediterranean grasslands.

In-situ hyperspectral reflectance measurements of vegetation and were collected at the same time as CO₂ gas exchange measurements were measured concurrently performed in control (C) and fertilized plots with added nitrogen (N), phosphorus (P) or the combination of N, P and potassium (NPK). Reflectance values were aggregated, according to their similarity ($r > 90\%$), in 26 continuous wavelength intervals (Hyp). In addition, also, the same reflectance values were resampled reproducing the spectral bands of both Sentinel-2A Multispectral Instrument (S2) and Landsat 8 Operation Land Imager (L8) simulating the signal that would be captured in ideal conditions by either Sentinel-2A or Landsat 8.

An optimal procedure for selection of the best subset of predictor variables. The (LEAPS) procedure was applied to select-identify the best most effective set of the-vegetation indices or spectral bands for GPP estimation using Hyp, S2 or L8. The-LEAPS selected some-vegetation indices putting in evidence their-according to their explanatory power, -showing their importance as indicators of the dynamic changes occurring in community vegetation properties such as canopy water content (NDWI) or chlorophyll and carotenoids/chlorophyll ratio (MTCI, PSRI, GNDVI) and underlining their importance usefulness for grasslands GPP estimates.

For Hyp and S2, bands showed similar explanatory power were as performant as than vegetation indices to estimate GPP. A two-step LEAPS procedure allowed us also To to identify spectral bands with potential for improving vegetation indices based GPP estimates we applied a two-step LEAPS procedure which- This procedure clearly indicates pointed to indicated the shortwave infrared region of the spectra as the most relevant promising for this purpose. A The comparison between of S2 and L8 based models showed similar explanatory power for of the two simulated satellite sensors when both vegetation indices and bands were included in the model when spectral bands were adopted.

Altogether, our results show-describe the potential of sensors on board of Sentinel 2 and Landsat 8 satellites for monitoring grasslands phenology and improving GPP estimates in support of a sustainable agriculture management.

1. Introduction:

Mediterranean grasslands are high biodiverse ecosystems, covering around 22% of the European Union land area, and providing important ecosystem services such as forage production (Bugalho and Abreu, 2008; Dfáz-Villa et al., 2003). These ecosystems are subjected to large pressures under global change (Sala, 2000), namely by the increasing availability of nutrients (e.g., phosphorus (P) and nitrogen (N)) due to human use of fertilizers, and N deposition (Ceulemans et al., 2014; Galloway et al., 2004; Peñuelas et al., 2013) and by a decrease and shift in seasonal patterns of precipitation (Costa et al., 2012; Kovats et al., 2014). The contemporary changes in water and nutrients supply can affect species composition, biomass and phenology along the life cycle of annual grasslands (Harpole et al., 2007), compromising their productivity. In particular, the onset and duration extension of the senescence period, largely dependent on soil water availability, can be affected in Mediterranean grasslands with great impacts in their functioning and large consequences on Gross Primary Productivity (GPP) functioning (Aires et al., 2008a, 2008b; Jongen et al., 2013b; Xu and Baldocchi, 2004). This uncertainty scenario increases the need for frequent monitoring of GPP along the growing season.

Using remote sensing based information to evaluate GPP brings important advantages both from a scientific and management point of view. Spectral retrievals collected from optical sensors on board of remote platforms may provide information on many biophysical properties of vegetation and can be usefully employed for monitoring and modelling ecosystems GPP in a cost and time-effective way (Schimel et al., 2015). Also, for land managers, the capability of making timely grassland management decisions may improve the use and sustainability of these ecosystems.

GPP estimation models integrating remote sensed observations increased considerably in the last decades (Beer et al., 2010; Grimm et al., 2008). Such models are generally based on the Light Use-Efficiency (LUE) concept (Monteith, 1972, 1977), which defines GPP as a function of the fraction of radiation absorbed by vegetation ($fPAR$), which in turn depends on green leaf area and the efficiency by which light energy is used to fix carbon during photosynthesis (i.e. LUE) (Cheng et al., 2014; Yuan et al., 2014).

Based on this approach large efforts have been put to derive vegetation indices able to represent the green leaf area and LUE.

The Normalized Difference Vegetation Index (NDVI) is widely used for its known linear relationship with the absorbed radiation $fPAR$ (Fensholt et al., 2004; Joel et al., 1997; Myneni and Williams, 1994). However, some exceptions are reported in the literature. For example in highly productive environments, such as grasslands, NDVI becomes easily saturated, not responding to increased leaf area and LUE, and the regression observed is no more linear (Vescovo et al., 2012; Viña and Gitelson, 2005).

In annual grasslands, such as the Mediterranean, control on ecosystem carbon balance is generally considered related mainly to the amount of green leaf area, while little LUE changes are expected (Gamon, 2015). Nonetheless, several studies reported a hysteresis in LUE in grasslands when the duration of the study encompasses the whole life cycle (Nestola et al., 2016; Pérez-Priego et al., 2015b).

The Photochemical Reflectance Index (PRI) is frequently adopted as a proxy of LUE (Gamon et al., 1997) at leaf and canopy scale (Peñuelas et al., 1995) (Garbulsky et al., 2011). PRI in the short term represents mirrors the dynamic of the xanthophylls cycle (Peñuelas et al., 1995) which is related to thylakoid energization and hence to light harvesting by photosynthesis. In the long term, PRI was found to be correlated with the ratio of carotenoids to chlorophyll (Filella et al., 2004; Porcar-Castell et al., 2012) and hence to plant senescence, since chlorophyll degradation and N export is a distinctive process of leaf ageing (Thomas, 2013). However, also PRI shows some drawbacks, since it is largely affected by species identity, leaf age or environmental conditions (Peñuelas et al., 1995) and by sensors geometry and atmospheric factors (Moreno et al., 2012). Hence the performance of models integrating PRI is frequently below the expected expectation (Pérez-Priego et al., 2015b). As a result, other vegetation indices have been tested as alternatives to NDVI and PRI for GPP estimation. Rossini et al. (Rossini et al., 2012), in a subalpine grassland obtained the best model to estimate GPP adopting together the MERIS Terrestrial Chlorophyll Index (MTCI) (Dash and Curran, 2004), a proxy of chlorophyll, and PRI. In another study, in a subalpine grassland, Sakowska et al. (2014), found that the red-edge NDVI, a modified NDVI, where the infrared band is substituted with a red-edge band (Gitelson and Merzlyak, 1994) improved GPP estimates. In Mediterranean grasslands with different N and P fertilization level, PRI together with solar induced fluorescence improved GPP estimates (Pérez-Priego et al., 2015b). In a semi-arid grassland, Vicca et al. (2016) observed that several vegetation indices including NDVI and the Normalized Different Water Index (NDWI) (Gao, 1996), a proxy of vegetation water content, were able to capture the drought effect on GPP.

Altogether these results clearly indicate that, inspite in spite of the usefulness of VIs to represent dynamic changes in biophysical properties of vegetation, the need for further studies are needed, aiming to identify the vegetation indices and the regions of the spectral regions that can be of potentially interesting for to estimate grassland GPP estimates of grasslands under different environmental constraints, such as nutrients availability.

The adoption of a specific model and vegetation index depends also frequently on the availability of remote sensed products at a suitable spatial and temporal scale. In the case of local scale monitoring of managed grasslands, sensors with high spatial resolution will produce better results than sensors with coarse spatial resolution. In this study we opted for simulating using data from Sentinel-2A MSI (Multi-Spectral Instrument), (hereafter named S2) and Landsat8 OLI (Operational Land Imager) (hereafter named L8), for their spatial resolution, (10-20m for S2 and 30m for L8) more suitable for representing grasslands spatial heterogeneity and hence better adapted to implement management options from a precision agriculture perspective.

The L8 provides reflectance in 7 bands ranging from the visible to the short wave infrared region (SWIR) (Loveland and Irons, 2016), but its main drawback is the long revisiting time of 16 days. The recently launched S2 covers the regions of the visible and near-infrared and the SWIR in 13 bands with at least five days revisiting time when both S-2A and S-2B platforms will become operational. The recently launched S2 covers the visible and near-infrared regions and also the SWIR in 13 bands with at worst five days revisiting time for the combination of S-2A and S-2B platforms (Drusch et al., 2012).

Field collection of vegetation reflectance by hyperspectral sensors is less cost-effective and more time consuming than satellite remote sensed data but presents the advantage of providing reflectance in numerous, high resolution, wavelengths

(Porcar-Castell et al., 2015). Therefore, it can be usefully employed for identifying which wavelengths best ~~reflect mirror~~ biophysical properties and physiological status of vegetation (Balzarolo et al., 2015; Matthes et al., 2015) and ~~put in evidence point to~~ regions of the spectra ~~of potential interest~~ potentially interesting for GPP modelling which until now have not been actually not exploited by remote sensors. The high detail of spectral resolution (1 nm nominal) is a further advantage of hyperspectral measurements. In particular, it allows comparing the performance of similar vegetation indices available from different satellite platforms resampling hyperspectral information to match spectral bands of different remote sensors.

A promising new technology is the use of ~~spaceborne~~ space borne imaging spectroscopy. The hyperspectral resolution of these images allows identifying canopy properties with higher sensitivity than traditional vegetation indices. For example, the use of ~~space borne~~ imaging spectroscopy was able to detect changes in canopy leaf area and water stress in ~~the~~ humid tropical forest, whereas NDVI and other vegetation indices failed (Asner et al., 2004).

The aim of this study was to identify combinations of vegetation indices and bands that better ~~represent-reflect~~ GPP changes in the period comprised between the annual-peak of growth and senescence dry out in Mediterranean grasslands subjected to different fertilization treatments.

To achieve this goal, in situ hyperspectral measurements of vegetation reflectance were ~~employed-used~~ to estimate GPP in a Mediterranean grasslands north-east of Lisbon, in central Portugal, before and after the annual peak of growth ~~was achieved~~, which generally occurs in May, with large inter-annual differences (Jongen et al., 2011). A set of vegetation indices proposed in the literature were calculated and the performance of models to estimate GPP based on linear combinations of vegetation indices and bands were compared.

~~Whenever possible,~~ Whenever comparable spectral range is available for the S2 and L8, vegetation indices were also calculated simulating the respective S2 and L8 bands and the performance of GPP estimates based on remote platforms and in situ hyperspectral measurements compared.

The specific objectives of the study were: (i) Test the impact of different nutrient availability on GPP in Mediterranean grassland. (ii) Identify ~~the a~~ set of vegetation indices ~~useful~~ to optimize ~~a~~ GPP model in our experimental conditions for Mediterranean grasslands; (iii) Compare the performance of GPP models employing only vegetation indices, spectral bands or a combination of both ~~only and in combination with spectral bands~~; (iv) ~~Finally~~ Finally, compare GPP models using spectral information obtained from hyperspectral sensors with similar models obtained from S2 and L8 platforms.

2. Material and Methods

2.1 The study site

Our study was conducted in a semi-natural Mediterranean grassland at Companhia das Lezírias, an estate of approximately 15 000 ha, located north-east of Lisbon, Portugal (38°49'45.13'N, 8°47'28.61'W). The grassland plant community is composed ~~mainly of~~ by annual C3 species. The climate is Mediterranean, with mild, wet winters and hot, dry summers. Long-term (1961–1990) mean annual rainfall is 709 mm. Mean annual temperature is 15.9 °C (INMG, 1991). Site topography is flat and the soil is a well-drained deep Haplic Arenosol (WRB, 2006).

2.2 Experimental design

The grassland studied is part of the Nutrient Network experiment (<http://www.nutnet.umn.edu>; Borer et al., 2017; Seabloom et al., 2013). Plots (5m x 5m) were established in 2012, in a randomized block design. Factorial combinations of nitrogen (N), phosphorus (P), and potassium plus micronutrients (K), a total of eight treatments per block, including ~~the~~ controls (C) with no added nutrients, were ~~considered~~ reestablished. All nutrients were added at a rate of 10 g. N·m⁻² yr⁻¹. N was added as slow-release urea (60-90 days), P was added as triple-super phosphate and K as potassium sulphate. Micronutrients (6% Ca, 3% Mg, 12% S, 0.1% B, 1% Cu, 17% Fe, 2.5% Mn, 0.05% Mo, and 1% Zn) were added with K only once, at the start of the study to avoid possible micronutrient toxicity. In this study, only four fertilization treatments were considered: C, N, P and NPK. Each one of these treatments was repeated twice per block, a total of 24 plots were ~~considered~~ measured (2 replicates X 4 treatments X 3 blocks).

2.3 Environmental measurements

Temperature, PAR and relative humidity were measured in situ using a VP-3 humidity temperature and vapour pressure sensor and QSO-S PAR Photon Flux sensor (Decagon Devices, Pullman, USA) logged every 30 min (EM50 data logger, Decagon Devices, Pullman, USA). Precipitation was recorded using a tipping bucket rain gauge (RG2, Delta-T Devices, Cambridge, UK). Soil water content (SWC) was continuously measured, at a depth of 10 cm, which corresponds to the main rooting zone (Jongen et al., 2013a; Schenk and Jackson, 2002), using EC-5 soil moisture sensors (Decagon Devices, Pullman, USA). The rain gauge and soil sensors were connected to a CR1000 and AM16/32B multiplexer data logger (Campbell Scientific, Logan, USA).

2.4 Field Measurements

2.4.1 NEPE and R from a closed system IRGA

Grassland net ecosystem ~~productivity~~ exchange (NEPE) was measured with a closed chamber (40 cm X 40 cm X 54 cm) of polymethylmethacrylate (3 mm thick) inserted into a permanent frame buried 5 cm into the soil. Radiation transmittance was

higher than 95%. The same chamber was covered with a reflective cloth for dark respiration (-R-) measurements. Air temperature inside the chamber was continuously monitored and PAR was measured at beginning and end of measurements with a ceptometer (AccuPAR-LP80, Decagon Devices, Inc. Pullman, WA, USA). Fans in the chamber ensured air circulation. The chamber was connected to an infrared gas analyser (LI-840, Li-Cor, Lincoln, NE, USA) measuring CO₂ and water vapour. Each measurement was no longer than 3 min. Fluxes were calculated based on the rate of change of CO₂ inside the chamber, after an initial period of at least 10 seconds. Flux calculations and corrections for CO₂ water vapour dilution followed Pérez-Priego [et al. \(2015a\)](#). GPP was obtained by detracting R from NEP at each measurement. All plots were measured between 11:00 and 13:00 on clear sky sunny days, [as close as possible to field spectroradiometric measurements](#). Measurements were performed during the 2016 growing season. Two field campaigns were carried out during vegetation growth, day 1 (31st March ~~to~~ 1st April) and day 2 (24th ~~to~~ 25th April) and two during the senescence phase, day 3 (19th – 20th May) and day 4 (1st-3rd June).

2.4.2 Leaf area and biomass Green Plant Area Index and biomass

The Plant Area Index (PAI), [a measure of all aboveground plant structure](#), —was indirectly measured with a linear PAR ceptometer (AccuPAR LP-80 Decagon Devices Inc., Pullman, WA, USA). The ceptometer measures the fraction of PAR intercepted by the canopy (*f*PAR) according to equation (1):

$$fPAR = \frac{1 - PAR_t}{PAR_i} \quad (1)$$

[Where \$PAR_i\$ is the incoming PAR measured above the canopy and \$PAR_t\$ is the PAR transmitted through the canopy, measured below it.](#)

The *f*PAR was considered approximately equal to absorbed radiation, as the amount of reflected radiation in the PAR range is usually low (Gower et al., 1999). For each plot, 6-8 measurements above (PAR_i) and below (PAR_t) the canopy were taken and averaged.

The PAI is calculated by inversion of the Beer-Lambert law (=equation 2):

$$fPAR = 1 - e^{-K \cdot PAI} \quad (2)$$

where K is the light extinction coefficient, which depends on the leaf angle distribution of the canopy, in this study considered spherical distributed (Jones, 1992), and on the zenith angle of the probe, calculated by the ceptometer with basis on the geographic coordinates of the local and date and time of measurements. To avoid low solar zenith angles all measurements were performed around solar noon.

As the growing season progressed some species started to senesce. In order to estimate the fraction of PAR absorbed only by photosynthesizing components of the canopy ("green" [f](#)PAR and PAI, [f](#)PAR and PAI_{gr} respectively), [f](#)PAR and PAI were multiplied by a normalized (by scaling between 0 and 1) greenness index (GI, calculated as a ratio between the digital number values of green and the sum of red, green, and blue digital number values) derived from the analysis of digital

pictures of the plots taken at each measurements day around solar noon (Cyber-shot DSC-W530, SONY), using the Phenopix R package (Filippa et al., 2016).

To determine aboveground productivity, A strip of vegetation (0.1 m x 1 m) within each plot was also collected close to the peak growth and biomass divided into functional types (legumes, forbs, graminoids) and dried in an oven until constant weight at 60°C.

2.4.3 Hyperspectral measurements of vegetation reflectance

At each field campaign, hyperspectral observations of all plots were ~~also acquired~~ performed with a FieldSpec3 spectroradiometer (ASD Inc., Boulder, USA), which provides reflectance of vegetation in the range of 350-2300 nm. The spectral resolution (Full-Width-Half-Maximum) is 3 nm at 700 nm and 10 nm at 1400 nm and 2100 nm. The sampling interval is 1.4 nm for the spectral region of 350-1000 nm (visible and near infrared) and 2 nm for the spectral region of 1000-2500 nm (short-wave infrared). A white reference of known reflectance (Spectralon panel, Labsphere, Inc., North Sutton, USA) was used to normalize for variations in atmospheric conditions and to convert the measurements into absolute reflectance (Ref.). Spectra were collected using a bare fibre optical cable (with an instantaneous field of view of 25°) inserted into a pistol grip at approximately 90 cm above the canopy and a nadir view.

~~The spectral resolution (Full Width Half Maximum) is 3 nm at 700 nm and 10 nm at 1400 nm and 2100 nm. The sampling interval is 1.4 nm for the spectral region of 350-1000 nm (visible and near infrared) and 2 nm for the spectral region of 1000-2500 nm (short wave infrared). Spectra at 1nm intervals are obtained from a cubic spline interpolation function. Five spectra were recorded/collected for each plot, each one representing the average of 25 observations/spectra, employing a bare fibre optic cable (with an instantaneous field of view of 25°) inserted into a pistol grip at approximately 90 cm above the canopy. A white reference of known reflectance (Spectralon panel, Labsphere, Inc., North Sutton, USA) was used to normalize for variations in atmospheric conditions and to convert the measurements into absolute reflectance (Ref.).~~ All measurements were conducted immediately after grassland gas exchange measurements, within two hours around solar noon, to minimize the effects of shadowing and solar zenith changes.

2.5 Data analysis

All statistical analyses were performed using open-source R (R Core Team, 2016). We used the lme4 package (Bates et al., 2014) to perform linear mixed effect analyses of the effect of the fertilization and control treatments on NEPE, R, GPP and PAIgr. Treatment and date were the fixed effects and the block was the random effect. Conditions of homoscedasticity and normality were always verified by visual inspection ~~offer~~ residuals. *P*-values were obtained by likelihood ratio tests of the full model with the effect in question against the model without the effect in question. A Tukey test was used for post-hoc comparison using the multcomp package (Hothorn et al., 2008).

The full spectra of vegetation reflectance retrieved from the Fieldspec was used to model GPP, after excluding noisy values in the range 1350-1400 nm and 1800-1950 nm. Our ~~PP~~=1748 original explanatory variables are x_{350}, \dots, x_{2299} where x_i

represents the reflectance in the narrow band $[\lambda, \lambda + 1]$ (nm) and our response variable is the GPP ($\mu\text{mol m}^{-2} \text{s}^{-1}$). A total number of 96 observations were available (4 treatments \times 2 replicates \times 3 blocks \times 4 dates). Since we have 1748 explanatory variables and just 96 observations, hence a high level of redundancy in our data, the [number of predictors dimensionality](#) was [first](#) reduced by grouping variables that belong to intervals of wavelengths where all variables are highly correlated. A hierarchical cluster analysis was performed to reduce the number of predictors from $P=1748$ to $pP=25$ groups of contiguous variables named “[bBands](#)”. [The basic idea is to aggregate contiguous and highly correlated individual 1 nm intervals of wavelength into broader wavelength bands. Two original predictors \$x_{\lambda_a}, x_{\lambda_b}\$ are clustered together. The distance between two variables is the correlation coefficient and the distance admitted \$s\$ within a band is given by the complete link criterion to guarantee that if their correlation coefficient \$r\(x_{\lambda_a}, x_{\lambda_b}\)\$ is larger than \$\geq 0.90\$. Bands correspond to the largest contiguous intervals where all pairs of original predictors satisfy that condition. If a band groups ~~for any pair of variables~~ \$\(x_{\lambda_a}, x_{\lambda_b}\)\$ within each group all original predictors between \$\lambda_1\$ and \$\lambda_2\$, then it is represented by a new variable \$x_{\[\lambda_1, \lambda_2\]}\$. Formally, wherever all pairs of variables \$x_{\lambda_a}\$ and \$x_{\lambda_b}\$ such that \$\lambda_1 \leq \lambda_a < \lambda_b \leq \lambda_2\$ are highly correlated, i.e. \$r\(x_{\lambda_a}, x_{\lambda_b}\) > 0.9\$, then the original variables \$x_{\lambda_1}, \dots, x_{\lambda_2}\$ in the interval \$\[\lambda_1, \lambda_2\]\$ are replaced by a new variable \$x_{\[\lambda_1, \lambda_2\]}\$ which is the arithmetic mean of the original variables \$x_{\lambda_1}, \dots, x_{\lambda_2}\$. The procedure is repeated to obtain \[groups of highly correlated reflectances of contiguous wavelengths\]\(#\). all bands that partition the full \$x_{350}, \dots, x_{2299}\$ spectrum.](#)

Reflectance values were also resampled to simulate bands of Sentinel-2A MSI (S2) and Landsat8 OLI (L8). [Since each band of S2 or L8 has a spectral response which is not perfectly uniform, we use the spectral response function of each sensor \(Barsi et al., 2014; ESA, 2018\) to weigh the contribution of each original predictor. ~~adopting the~~ As a result, for each sensor and band \$\[\lambda_1, \lambda_2\]\$, we calculated the \[reflectance response\]\(#\) as a weighted mean of \$x_{\lambda_1}, \dots, x_{\lambda_2}\$, ~~in the interval of the corresponding band~~ where the weights \[are given by the are the coefficients of the spectral response function of the sensor \\(Barsi et al., 2014; ESA, 2018\\): spectral response.\]\(#\) The list of S2 and L8 bands used in this study is shown in table 1.](#)

Vegetation indices (VIs) (Table 2) were calculated from hyperspectral (Hyp), or simulated S2 and L8 sensors (Table 1). The VIs [were selected were retrieved](#) from the literature with basis [based](#) on their relation to biophysical properties of vegetation affecting GPP. [The NDVI and the NDVIre, are considered a proxy of fPAR; the Green Normal Difference Vegetation Index \(GNDVI\), the MERIS Terrestrial Chlorophyll Index \(MTCI\) and the chlorophyll index \(CI\) are representative of chlorophyll a and N content, while the Photochemical Reflectance Index \(PRI\) and the Plant Senescence Reflectance Index \(PSRI\) are expected to mirror changes in the ratio of carotenoids to chlorophyll. Finally, the Normalized Difference Water Index \(WBI\) and the Normalized Difference Water Index \(NDWI\) are considered proxy of tissue water content.](#)

[Given that the goal of the analysis is to determine the set of VIs and /or bands that best model GPP, we apply a data analysis method that identifies the best subset of single variables. This is distinct from Principal Component Analysis \(PCA\) where dimensionality reduction is achieved through replacing variables by their linear combinations, which still involve all the variables. We adopted linear regression \(MLR\) to model the relation between our explanatory variables \(bands and VIs\) and the response variable \(GPP\). Although the expressiveness of non-linear models \(e.g. in the field of machine learning\) is](#)

stronger than MLR, we believe that linear models provide a clearer interpretation of the relation between predictors and GPP, while offering enough flexibility by including variables in a high dimension representation space. Moreover, and as discussed below, linear models allow us to derive confidence intervals for our results, apply statistical tests to compare models at a given significance level, and they are less prone to overfitting than complex non-linear models.

~~A multiple linear regression (MLR) was adopted to model the relation between our explanatory variables (bands and VIs) and the response variable (GPP).~~ Since the number of observations is only roughly twice as large as the number of new explanatory variables we performed a variable selection and excluded those variables that do not contribute significantly to the goodness-of-fit of our model. Although the dimensionality of the problem is very large, it can be solved efficiently by the LEAPS algorithm (Furnival and Wilson, 1974) available through the R package L4eaps (Lumley, 2009). Unlike alternative heuristic approaches (Cadima et al, 2004), LEAPS returns the optimal subset of predictors according to a given criteria. In our analysis, the criteria was the adjusted R², so LEAPS returns the sub-model with the highest adjusted R² among all possible 2^p sub-models, where p is the number of predictors.

A nested approach was adopted to formally test which model better explained GPP. A preliminary test showed that better results were obtained with exponential regressions and therefore lnGPP was adopted as the response variable in all analyses. The general model was $\ln GPP \sim \sum_{j=1}^n v_j$, where v are vegetation indices (VIs) or optical bands (B) from Hyp grouping procedure or from simulated S2 or L8 data. The subset of v_j was selected by maximizing the adjusted R² among all possible combination of predictors.

The LEAPS procedure returns an optimal model ~~named that we called~~ L. However, L may include variables which contribute only marginally for the overall adjusted R². To further reduce the dimensionality of the predictors, we test sub-models of L (obtained by backwards stepwise selection of predictors) against the LEAPS optimal model L. When sub-models of L were found not to be significantly worse than L, at a significance level alpha=0.05, then we considered the most parsimonious of those sub-models as the optimal solution. A F-test was used to perform those comparisons. The analysis was repeated separately for all vegetation indices (VIs) and bands (B) from Hyp, S2 or L8 data, obtaining an optimal model for each sensor. We performed an analysis of residuals for each selected model, which showed no evidence of violation of the linear model assumptions.

Besides determining the adjusted R² for the optimal model from the full sample, we applied a bootstrap procedure (N=10000 iterations) to estimate the distribution of the adjusted R² in the whole population (Ohtani, 2000). This allowed us to estimate quantiles (25%-75%) for adjusted R² and also compare the adjusted R² distributions among models. In particular, it permits to estimate the probability that the some model A has a higher adjusted R² than the an alternative model B.

Two-step models were also used to investigate if optical bands had the potential to improve models based only on vegetation indices (VIs). Toward that end, bands (B) were added to the optimal models obtained by the procedure above described denoted by Hyp-VIs, S2-VIs and L8-VIs (step 1). Using step 1 as the base model, we applied LEAPS to determine the subset of bands that maximized the overall adjusted R². As before, we applied a F-test (alpha=0.05) to possibly reduce

the number of bands in the optimal model. As a result, we defined the optimal two-step models: Hyp-VIs+B, S2-VIs+B and L8-VIs+B. Finally, for Hyp, S2 and L8, we performed a F-test to compare the one step optimal model with the correspondent two-steps optimal model. A low p-value for this F-test indicates that the two-step model is significantly better and means that bands, in addition to vegetation indices, contribute for an improved modelling of GPP.

5 3. Results

3.1 Conditions during the experimental period

During the period of measurements, from March 31 to June 3, 2016, the average daily PAR and temperature increased progressively, ranging from 630 $\mu\text{molm}^{-2}\text{s}^{-1}$ to 1000 $\mu\text{molm}^{-2}\text{s}^{-1}$ and from 9.6 °C to 17°C, respectively (Fig.1a). Soil water content (SWC) (Fig. 1b) showed fluctuations according to rainfall events, ranging from 0.05 to 0.2 m^3m^{-3} . During the experimental period, rainfall was concentrated in the first half of April and at the beginning of May. Along the experimental period rainfall recorded was 195 mm, corresponding to the 33% of the whole year.

3.2 The effect of fertilization on leaf Plant Area Index -area and species composition functional groups proportion

From the beginning to the end of the study period, PAI increased on average 4 fold from 1 to 4 (Fig.2a). In all treatments, the increase in PAI was completed by May 20 and no further increase was observed in the last measurement (June 3). On the contrary, at the beginning of the experiment, PAIgr (Fig2b) showed an increasing tendency similar to PAI but from May 20 onwards, the ~~tendency~~ trend changed and a decreasing trend was then observed corresponding to the onset of grassland senescence.

The fertilization treatments influenced both the PAI ($P<0.000$) and the PAIgr ($P<0.000$) being both significantly higher for treatments NPK and P than for treatment C ($P<0.001$). No differences were observed between C and N treatments ($P>0.05$). In both PAI and PAIgr the treatment P showed similar values to NPK, with the exception of the first measurements day (April 1). The grassland communities fertilized with NPK had a higher and earlier leaf area growth when compared to the other treatments.

Plant species composition has been measured every year since 2012 (pre-treatment) under an ongoing long-term nutrient addition experiment on this grassland site. In line with results from that study (Nogueira et al., personal communication submitted). (The fertilization treatments also influenced the functional composition of grasslands (Table 3). In the NPK treatment the percentage of graminoids was higher than in any of the other treatments. P treatment showed a higher percentage of legumes and in the C and N treatments forbs were the dominant functional group.

3.3 The effect of fertilization on GPP

The ability of grasslands to sequester atmospheric carbon dioxide was not affected by fertilization treatments. The NEPE (Fig.3a) and the GPP (Fig.3c) did not reveal any statistical significant difference among treatments ($P>0.05$). On the

contrary, the rate of respiration (Fig.3b, R) was affected by the fertilization treatment ($P<0.05$), being on average higher for treatments NPK and P than C. CO₂ gas exchanges were influenced by the grassland life cycle and marked trends were observed along the measurement period. NEPE showed an average drop of 74%, shifting from $-14.47 \mu\text{molm}^{-2}\text{s}^{-1}$ to $-3.67 \mu\text{molm}^{-2}\text{s}^{-1}$, from April 01 (day 1) to June 03 (day 4) ($P<0.000$). This decrease in NEPE rate was particularly evident from the second to the third measurement day, after the annual peak of grassland growth was achieved (Fig. 3a). R also showed differences along the experimental period ($P<0.000$) but the trend observed was different. R increased from the first to the second measurement day, from $8.22 \mu\text{molm}^{-2}\text{s}^{-1}$ to $13.65 \mu\text{molm}^{-2}\text{s}^{-1}$ and then decreased toward the end of the experiment (Fig. 3b). GPP also changed significantly along the studied period ($P<0.000$), decreasing from $25.72 \mu\text{molm}^{-2}\text{s}^{-1}$ on April 25 (day 2) to $12.12 \mu\text{molm}^{-2}\text{s}^{-1}$ on June 3 (day 4). A linear relationship was observed between GPP and fPARgr (Fig.4)-, however the slope of the regression line changed along the experimental period and marked differences were observed between the vegetation growth (day 1 and 2-) and the senescence phase (day 3 and 4) (Fig.4).

3.4 Vegetation reflectance

The reflectance of vegetation (Ref) varied on average between 0 and 0.4 (Figure-Fig. 54a). The cluster analysis created 25 bands (Fig. 54b) based on Ref similarity of contiguous wavelengths ($r>90\%$). Bands were narrower in the visible region (350 -nm to 750 nm) than in the NIR (750 nm to 1350 nm) and in the SWIR (1350 nm to 2300 nm) region. In particular, in the red-edge region, between 698 and 732nm 6 different bands were identified, corresponding to a steep increase in reflectance observed in this region of the spectra.

3.5 Vegetation indices

Adopting wave bands obtained by cluster analysis (Fig.4Fig.5b) several vegetation indices were calculated (Table 2). The average values of the indices GNDVI, NDWI, PSRI and MTCI are shown in figure 5. Other indices are omitted from the figure for showing very similar trends to the ones represented (NDVI, NDVIre, WBI, CI) or not being significantly correlated with the response variable (PRI). All of them showed larger changes during the study period, particularly after April 25 (day 2), when the annual peak of growth was achieved.

The GNDVI (Fig. 65a) showed small changes among treatments and dates, with a significant drop of 20% observed from April 1 to June 3 ($P<0.000$), and significant differences in the NPK and P treatments ($P<0.001$) as compared to C but differences were not evident anymore on June 3 (day 4). The MTCI (Fig.65b) showed a large drop particularly evident after April 25 (day 2). At April 1 (day 1), the effect of fertilization was evident in treatments NPK and P as compared to C ($P<0.001$), however along the experimental period differences among treatments diminished and by June 3 (day 4) no differences among treatments were observed. The NDWI (Fig.65c) showed a similar temporal trend with a marked decrease from April 1 to June 3 ($P<0.001$). Also for MTCI, the NPK and P treatments showed always higher values than C ($P<0.001$), suggesting a positive impact of the higher nutrient availability on tissue water content. The PSRI had an opposite trend,

showing on average a threefold increase from April 1 to June 3 ($P<0.000$) and a tendency to lower values in fertilized treatments as compared to C ($P<0.001$) for NPK and P and $P<0.01$ for N).

Significant regressions were established between GPP and all the vegetation indices considered (Table 4) with the exception of PRI. The NDWI was the index that explained the higher proportion of variability of GPP, ~~which is the result of the progressive drying out of vegetation toward the end of the growing season.~~

3.6 GPP estimates by multiple linear regression models

The LEAPS procedure selected VIs or Bands as predictor variables retrieved from Hyperspectral data (Hyp) or simulating Sentinel-2 MSI (S2) and Landsat8 OLI (L8) sensors. ~~We performed an analysis of residuals for each selected model, which showed no evidence of violation of the linear model assumptions.~~

Models adopting only VIs as predictor variables (Hyp-VIs and S2-VIs) performed similarly with a considerable overlap of adjusted R^2 (Table 5). On the contrary, ~~the~~ L8-VIs model showed a lower performance (lower adjusted R^2) than Hyp-VIs and S2-VIs models (Table 5). Bootstrap results allowed us to conclude, at a confidence level of 90%, that Hyp-VIs has higher adjusted R^2 than L8-VIs.

The selection of VIs in the Hyp-VIs and S2-VIs models exhibited a similar spectral pattern. Both models included PSRI and GNDVI. On the contrary, NDVI, the most frequently adopted index as green leaf area proxy was not included in the Hyp-VIs model but only in the S2-VIs. Two of the indices included in the Hyp model are related with water balance (WBI) and water tissue content (NDWI). The S2 model includes also MTCI, which represents chlorophyll-a and N.

Models including only bands (-B) showed similar performance to respective models employing vegetation indices (-VIs).

Only in the case of L8, where just one vegetation index (NDVI) was available, bands (L8-B) led to better modelling of GPP than vegetation indices (L8-VIs). Similar spectral patterns were also observed in the selection of bands for GPP estimate for all sensors (Hyp, S2, L8). A common pattern is the inclusion of bands in the SWIR region strongly represented in the Hyp-B ($R_{1951-2299}$, $R_{1209-1327}$, $R_{1328-1349}$), S2-B (B11) and L8-B (B6 and B7) models. The red edge region of the spectra was also largely represented in the Hyp-B ($R_{724-732}$, $R_{706-710}$, $R_{702-705}$, $R_{698-701}$, $R_{716-723}$) and S2-B (B5, B6 and B7) underlining the importance of this region for vegetation reflectance.

The LEAPS two-step procedure allowed us to identify bands with potential to improve the VIs based models, identifying regions of the spectra generally not adopted in vegetation indices. For both S2Hyp and L8 the two-step model (VIs + B) increased significantly ($P<0.010$) the performance of the model, while for HYPs2 the difference between HypS2-VIs and HypS2-VIs+B model, in spite of still being significant, is less marked ($P<0.05$). The bootstrap procedure indicated that the probability of the Hyp-VIs+B being significantly better ($\alpha=0.05$) is 83% when compared to S2-VIs+B and 81% when compared to L8-VIs+B. On the contrary, the S2-VIs+B and ~~the~~ L8-VIs+B models ~~do not differ significantly exhibit roughly the same explanatory power.~~ In all ~~the~~ VIs+B models, bands in the SWIR region were included. The second region of the spectra more represented in the Hyp-B and Hyp-VIs+B model was the red-edge.

4. Discussion

4.1 The impact of fertilization treatment

The fertilization treatment influenced the growth rate and the composition of the herbaceous plots more than carbon sequestration. In line with a five year and ongoing study at the same grassland site, the fertilization treatment resulted in differences in aboveground biomass and species composition/functional groups proportion for NPK and P treatments as compared to C, while the sole addition of N had no effect.

An earlier growth response was also observed in the NPK treatment measurement. The higher percentage of graminoid species with on average higher growth rates as compared to most forb species (Ansquer et al., 2009; Craine et al., 2001; Westoby et al., 2002) could explain early differences in PAI and PAIgr in this treatment. As leaf area is usually positively related to GPP (e.g. Aires et al., 2008b; Xu and Baldocchi, 2004), it would be expected that higher PAIgr in NPK treatments would induce increased GPP. However, confounding factors such as increased water demand associated with higher growth rates in NPK might have downscaled differences between treatments (e.g. Weisser et al., 2017).

The relationship GPP-/PARgr showed some differences along the experiment. The slope of the regression line was considerably lower during the vegetative growth than during the senescence period (Fig.4) revealing the occurrence of marked changes in the effective LUE along the growth life cycle and confirming previous results (Nestola et al., 2016; Pérez-Priego et al., 2015b).

While the fertilization treatment had no impact on NEP or GPP a higher R rate was observed at the measurement day 4 (June, 3) in the NPK and P treatments. Differences can be ascribed to the higher PAI- and precipitation at the end of May, which must have stimulated soil respiration (Jarvis et al., 2007; Reichstein et al., 2003).

4.2 Best VIs for GPP estimation

The LEAPS procedure selected several indices as significant predictor variables for GPP in the Hyp-VIs and S2-VIs model (Tab. 5). The vegetation indices selected in the Hyp-VIs and S2-VIs models are known to represent different properties of vegetation, specifically: the green fraction of the leaf area (GNDVI and NDVI), the chlorophyll-a and N concentration (MTCI), the ratio carotenoids/chlorophyll (PSRI) and the tissue water content (NDWI_WBI). Each of these traits has a major role in GPP.

Among the vegetation indices selected in the multiple linear models, both the Hyp-VIs and the S2-VIs included the PSRI (Merzlyak et al., 1999) which is generally applied to detect the occurrence of vegetation senescence. PSRI is able to capture changes in the carotenoids/chlorophyll ratio which occur during vegetation senescence since chlorophyll declines more rapidly than carotenoids (Merzlyak et al., 1999). In this study, PSRI increased in all treatments after April 25, when the maximum peak of growth (Fig.2,b) was achieved and close to the onset of canopy drying out. Another index known to be related with the carotenoids/chlorophyll ratio, the PRI (Filella et al., 2009), showed no correlation with GPP in our study. These results are in contrast with previous studies (Pérez-Priego et al., 2015b). However, a low performance of PRI in

representing the carotenoids/chlorophyll ratio has been already observed in semiarid grasslands (Vicca et al., 2016). In crops, a good agreement between PRI and pigment pools was observed at leaf (Gitelson et al., 2017a) but not at stand level. (Gitelson et al., 2017b). Differences in the last two studies were ascribed to changes in canopy structure (e.g., changes in leaf inclination angle) over the growing season.

5 The Hyp model also put in evidence the importance of changes in canopy water content, as both ~~the~~ NDWI (Gao, 1996) and ~~the~~ WBI (Penuelas et al., 1997) were included in the model. Considering ~~C~~changes observed in NDWI along the experiment and the good correlation observed between NDWI or WBI and GPP it is reasonable to assume that GPP is largely affected by the progressive senescence of vegetation, is an evidence of the importance of the onset of drought for grassland ~~vegetation as senescence marks the end of the growing season in Mediterranean grasslands~~ (Balzarolo et al., 2015; Vescovo
10 et al., 2012). In a previous study, Vicca et al. (2016) found that NDWI was able to estimate GPP in semiarid grasslands better than other indices, allowing to distinguish the effect of drought.

Other indices, sensitive to changes in chlorophyll-a concentration, MTCI (Dash and Curran, 2004) and GNDVI (Gitelson and Merzlyak, 1998) were also included in the model. The fertilization treatment resulted in an increase in MTCI during the first stage of the experiment in the NPK and P treatments, followed by a decrease observed in all treatments as the season
15 progressed toward the end of the annual growth cycle. A similar trend was observed in a study by (Pérez-Priego et al., 2015b) in which Mediterranean grasslands were subjected to fertilization with N or NP. The primary role of chlorophyll in photosynthesis is well known and justifies the positive relationship observed between GPP and MTCI. However, in the present study, no differences were observed in GPP among fertilized and non-fertilized treatments suggesting that the expected increase in photosynthesis due to the increase in chlorophyll and nitrogen was constrained by other environmental
20 and physiological factors.

Notably NDVI, the most frequently applied index in GPP estimates by LUE models (Yuan et al., 2014) was not selected in the Hyp-VIs model and showed a poorer coefficient of determination than other indices, (e.g. NDWI). ~~NDVI is expected to~~ reflect ~~mirror~~ changes in green leaf area, being generally linearly related with the fraction of photosynthetically absorbed radiation /PAR (Myneni and Williams, 1994). However, previous studies reported a saturation of NDVI and consequent lack
25 of linearity in the ~~NDVI/PAR~~ regression in high productive vegetation communities (Gianelle et al., 2009; Vescovo et al., 2012) such as grasslands and sometimes other indices showed a better performance. For example, in grasslands subjected to water and nutrient stress, the NDVI green index (GNDVI), which adopts a green band instead of the red band of NDVI and hence is more sensitive to chlorophyll-a concentration (Gitelson et al., 1996), showed a better performance than NDVI as
30 ~~PAR~~ proxy of leaf area (Cristiano et al., 2010; Gianelle et al., 2009). Also in this study, the GNDVI explained a larger proportion of GPP variance than ~~NDVI~~ in the Hyp-VIs and S2-VIs models being selected in both and before NDVI in the S2-VIs model.

The indices selected by the LEAPS (i.e. NDVI, GNDVI, NDWI, MTCI, PSRI and WBI) also showed high significant relationship with GPP (tab. 4) in simple regressions explaining 63% to 72% of the variability observed. The functional convergence (Ollinger, 2011) of different traits participating in the photosynthetic process may have hampered results

observed in the regression for each single vegetation index, showing a high degree of correlation for most of them (Table 4). However, the selection of several VIs, representatives of different structural and functional traits in the multiple linear models and the lower performance observed in the L8 model, including solely the NDVI index, clearly indicate the importance of considering the contribution of different traits with different temporal dynamics to capture GPP temporal changes in models integrating vegetation indices.

4.3.2 Are spectral bands better GPP estimators than VIs?

Our results suggest a marginal improvement in that better GPP estimates can be obtained by adopting bands (Hyp-B, S2-B, L8-B; Table 5) instead of vegetation indices (Hyp-VIs, S2-VIs, L8-VIs; Table 5). However, a larger impact was observed in the case of L8-VIs+B models when compared with the L8-VIs model which included only NDVI (Table 5).

Although normalized VIs are important in establishing strong relationships between biophysical and optical properties of vegetation, our results showed that the selection of the proper band is equally important to more important than the mathematical formulation of the indices for the explanatory power of spectral retrievals as predictor variables. Previous studies comparing the explanatory power of VIs and bands in grasslands showed similar results (Balzarolo et al., 2015; Matthes et al., 2015).

However, it cannot be disregarded that VIs offer, in comparison with spectral bands, the advantage to provide of being more robust in representing spatial and temporal comparable representation of vegetation features, since differences resulting from background reflectance, sun-sensor geometry and atmospheric effects are minimized by normalization of spectral values (Glenn et al., 2008).

Our results also evidenced the importance of the SWIR region of the spectra, as bands in this region were selected in all one- and two-step models, which is rarely adopted in vegetation indices with few exceptions. The SWIR region is known to correlate with canopy water content (Casas et al., 2014).

Studies investigating the potential of spectral bands to estimate canopy chlorophyll content and green fAPAR, found that the SWIR region was strongly positively correlated with them in grasslands (Sakowska et al., 2016) and also GPP in a semi-arid savanna (Tagesson et al., 2015).

Bands in the red-edge region were also largely represented in the Hyp-B, S2-B and in the Hyp-VIs+B models. The red-edge corresponds to the steep increase in reflectance at the boundary between the red region where chlorophyll is absorbed and the leaf scattering at the NIR region. Red-edge bands were successfully employed for estimating chlorophyll content in maize (Zhang and Zhou, 2017) and LAI in crops (Kira et al., 2017). For these reasons they were integrated into numerous VIs, such as MTCI and PSRI, also applied in this study, which explain the reason for the lack of red-edge bands in the second step of the S2 model (S2-VIs+B) while strongly represented in the S2-B.

4.34 Satellite sensors as estimators of GPP

Differences in the selection of the vegetation indices among sensors had apparently no effect on the performance of the S2-VIs and Hyp-VIs models (Table 5), while the limited number of available vegetation indices for the L8 resulted in a lower performance of the model.

Our results show ~~that an equal potential of~~ S2 and L8 ~~sensor spectral signatures are equally suitable~~ for assessing GPP.

GPP estimates obtained simulating S2 and L8 sensors showed a similar performance in the -B and -VIs+B models, while when only VIs were adopted, the S2 model had clearly a better performance than L8. These results suggest a need for testing new vegetation indices adopting L8 bands.

In agreement with our results, other studies comparing linear additive models showed similar ability for estimating canopy cover and LAI adopting the S2 or L8 sensors (Korhonen et al., 2017).

An important difference between the S2 and L8 availability of wavebands is the lack of reflectance values in the red-edge region ~~for in the~~ L8, which limited the possibility of computing VIs, such as MTCI and PSRI (Korhonen et al., 2017). However, the limitation imposed by the lack of bands in the red-edge region, had apparently more importance for the -VIs model, while differences in the performance of the model between S2 and L8 decreased for -B and VIs+B models.

In this study, S2 and L8 data comparison was based only on the simulation of the respective bands not taking into consideration other factors possibly affecting sensors spectral response such as sun-sensor viewing geometry (Tagesson et al., 2015). Nonetheless, in a recent study (Korhonen et al., 2017) the comparisons of satellite data from the two platforms showed no differences between S2 and L8 reflectance values in the NIR, SWIR1 and SWIR 2 bands. In other regions of the spectra, such as the green and blue bands reflectance values were considerably smaller in the S2 than in the L8 but still proportional, suggesting that comparisons between S2 and L8 simulated bands can largely be representative of the actual differences obtained by the two remote platforms.

~~At the same time, our~~ results confirm the importance of performing hyperspectral measurements. Indeed, ~~inferential in this study; bootstrap results show that for the whole population, and with probability 80%, the the~~ Hyp-VIs+B model ~~is showed to be superior to to~~ the corresponding S2 and L8 models ~~with over than 80% of probabilities~~. The high detailed resolution and the wide range of wave bands makes hyperspectral sensors ~~as~~ unique in identifying regions of the spectra of high interest for representing different ~~vegetation~~ properties ~~of vegetation~~ (Porcar-Castell et al., 2015)

5. Conclusions

In agreement with previous studies (Pérez-Priego et al., 2015b; Rossini et al., 2012; Vicca et al., 2016), our results clearly indicate the need to integrate into GPP models spectral information representing both structural and functional traits of vegetation along the whole grasslands life cycle. Specifically, water content (NDWI), chlorophyll (MTCI, GNDVI) and the ratio of chlorophyll to carotenoids (PSRI) were indicated as best predictor variables for GPP estimates. Altogether these

vegetation indices describe the loss of photosynthetic pigments and efficiency and dry out of vegetation occurring and when considered together improved considerably GPP estimates in comparison with models adopting only NDVI.

Our study also confirms the importance of hyperspectral in-situ measurements for exploratory analysis of the relationship between biophysical and optical properties of vegetation providing a wide spectral range and high resolution of spectral retrievals.

The hyperspectral reflectance values, together with the two-step procedure adopted for the selection of predictor variables allowed also to identify critical region of the spectra, not included in the initial selection of vegetation indices but that revealed their usefulness in estimating GPP. For example, the LEAPS two-step procedure evidenced which bands could improve significantly a model including only vegetation indices, identifying the red edge and SWIR regions of the spectra as of major importance for improving GPP estimates. This information can be critical in the development of new spectral indices and sensors.

Our results also evidenced the potential of S2 and L8 sensor in assessing GPP, since models obtained by simulating bands from the two sensors showed similar performance. The possibility of using remote sensing information for monitoring and modelling vegetation at a suitable spatial resolution, such as in S2 and L8 sensor, allows [for](#) attempted vegetation monitoring and modelling in a cost-effective way, in support of sustainable agriculture management practice.

Author Contribution

S Cerasoli designed the experiment and together with Joana Faria performed field measurements. M.L. Campagnolo developed the [variable selection](#) code. C. Nogueira and M.Caldeira set up the experimental plots. S.Cerasoli prepared the manuscript with contributions from all the authors.

Acknowledgments

This work was funded by Fundação para a Ciência e a Tecnologia, ministry of Science and education, through the project MEDSPEC, Monitoring Gross Primary productivity in Mediterranean oak woodlands through remote sensing and biophysical modelling (PTDC/AAG-MAA/3699/2014) and the research activities of Forest Research Centre (PEsst-OE/AGR/UI0239/2011). Sofia Cerasoli is post-doc fellow (BPD/SFRH/78998/2011), Carla Nogueira is the recipient of a PhD Grant (SFRH/BD/88650/2012). Authors wish to thank Georg Wohlfarth for a critical reading of the manuscript and Joaquim Mendes, Madalena Silva, Lurdes Marçal, Alicia Horta and Paula Pais for technical help in the field and laboratory. Authors are also grateful to FLAD/NSF for funding (A1/Proj 124/12) and to the Companhia das Lezírias for providing access to the grassland site and permission to undertake research.

References

Aires, L. M., Pio, C. A. and Pereira, J. S.: The effect of drought on energy and water vapour exchange above a mediterranean C3/C4 grassland in Southern Portugal, Agric. For. Meteorol., 148(4), 565–579, doi:10.1016/j.agrformet.2007.11.001, 2008a.

- Aires, L. M. I., Pio, C. A. and Pereira, J. S.: Carbon dioxide exchange above a Mediterranean C3/C4 grassland during two climatologically contrasting years, *Glob. Chang. Biol.*, 14(3), 539–555, doi:10.1111/j.1365-2486.2007.01507.x, 2008b.
- Ansquer, P., Duru, M., Theau, J. P. and Cruz, P.: Convergence in plant traits between species within grassland communities simplifies their monitoring, *Ecol. Indic.*, 9(5), 1020–1029, doi:10.1016/J.ECOLIND.2008.12.002, 2009.
- 5 Asner, G. P., Nepstad, D., Cardinot, G. and Ray, D.: Drought stress and carbon uptake in an Amazon forest measured with spaceborne imaging spectroscopy., *Proc. Natl. Acad. Sci. U. S. A.*, 101(16), 6039–44, doi:10.1073/pnas.0400168101, 2004.
- Balzarolo, M., Vescovo, L., Hammerle, A., Gianelle, D., Papale, D., Tomelleri, E. and Wohlfahrt, G.: On the relationship between ecosystem-scale hyperspectral reflectance and CO₂ exchange in European mountain grasslands, *Biogeosciences*, 12(10), 3089–3108, doi:10.5194/bg-12-3089-2015, 2015.
- 10 Barsi, J., Lee, K., Kvaran, G., Markham, B. and Pedelty, J.: The Spectral Response of the Landsat-8 Operational Land Imager, *Remote Sens.*, 6(10), 10232–10251, doi:10.3390/rs61010232, 2014.
- Bates, D., Mächler, M., Bolker, B. and Walker, S.: Fitting Linear Mixed-Effects Models using lme4, *J. Stat. Softw.*, 67(1), 1–48, doi:10.18637/jss.v067.i01, 2014.
- Beer, C., Reichstein, M., Tomelleri, E., Ciais, P., Jung, M., Carvalhais, N., Rödenbeck, C., Arain, M. A., Baldocchi, D.,
 15 Bonan, G. B., Bondeau, A., Cescatti, A., Lasslop, G., Lindroth, A., Lomas, M., Luysaert, S., Margolis, H., Oleson, K. W., Rouspard, O., Veenendaal, E., Viovy, N., Williams, C., Woodward, F. I. and Papale, D.: Terrestrial gross carbon dioxide uptake: global distribution and covariation with climate., *Science*, 329(5993), 834–838, doi:10.1126/science.1184984, 2010.
- Borer, E. T., Grace, J. B., Harpole, W. S., MacDougall, A. S. and Seabloom, E. W.: A decade of insights into grassland ecosystem responses to global environmental change, *Nat. Ecol. Evol.*, 1(5), 0118, doi:10.1038/s41559-017-0118, 2017.
- 20 Bugalho, M. N. and Abreu, J. M.: The multifunctional role of grasslands, in *Sustainable Mediterranean Grasslands and their Multifunctions*. vol. 79. *Option Méditerranéennes*, edited by C. Porqueddu and M. M. Tavares de Sousa, pp. 25–30., 2008.
- Casas, A., Riaño, D., Ustin, S. L., Dennison, P. and Salas, J.: Estimation of water-related biochemical and biophysical vegetation properties using multitemporal airborne hyperspectral data and its comparison to MODIS spectral response, *Remote Sens. Environ.*, 148, 28–41, doi:10.1016/j.rse.2014.03.011, 2014.
- 25 Ceulemans, T., Stevens, C. J., Duchateau, L., Jacquemyn, H., Gowing, D. J. G., Merckx, R., Wallace, H., van Rooijen, N., Goethem, T., Bobbink, R., Dorland, E., Gaudnik, C., Alard, D., Corcket, E., Muller, S., Dise, N. B., Dupré, C., Diekmann, M. and Honnay, O.: Soil phosphorus constrains biodiversity across European grasslands, *Glob. Chang. Biol.*, 20(12), 3814–3822, doi:10.1111/gcb.12650, 2014.
- Cheng, Y.-B., Zhang, Q., Lyapustin, A. I., Wang, Y. and Middleton, E. M.: Impacts of light use efficiency and fPAR
 30 parameterization on gross primary production modeling, *Agric. For. Meteorol.*, 189–190, 187–197, doi:10.1016/j.agrformet.2014.01.006, 2014.
- Costa, A. C., Santos, J. A. and Pinto, J. G.: Climate change scenarios for precipitation extremes in Portugal, *Theor. Appl. Climatol.*, 108(1–2), 217–234, doi:10.1007/s00704-011-0528-3, 2012.
- Craine, J. M., Froehle, J., Tilman, D. G., Wedin, D. A. and Chapin, III, F. S.: The relationships among root and leaf traits of

- 76 grassland species and relative abundance along fertility and disturbance gradients, *Oikos*, 93(2), 274–285, doi:10.1034/j.1600-0706.2001.930210.x, 2001.
- Cristiano, P. M., Posse, G., Di Bella, C. M. and Jaimes, F. R.: Uncertainties in fPAR estimation of grass canopies under different stress situations and differences in architecture, *Int. J. Remote Sens.*, 31, 4095–4109, doi:10.1080/01431160903229192, 2010.
- 5 Dash, J. and Curran, P. J.: The MERIS terrestrial chlorophyll index, *Int. J. Remote Sens.*, 25(23), 5403–5413, doi:10.1080/0143116042000274015, 2004.
- Díaz-Villa, M. D., Marañón, T., Arroyo, J. and Garrido, B.: Soil seed bank and floristic diversity in a forest-grassland mosaic in southern Spain, *J. Veg. Sci.*, 14(5), 701–709, doi:10.1111/j.1654-1103.2003.tb02202.x, 2003.
- 10 Drusch, M., Del Bello, U., Carlier, S., Colin, O., Fernandez, V., Gascon, F., Hoersch, B., Isola, C., Laberinti, P., Martimort, P., Meygret, A., Spoto, F., Sy, O., Marchese, F. and Bargellini, P.: Sentinel-2: ESA's Optical High-Resolution Mission for GMES Operational Services, *Remote Sens. Environ.*, 120, 25–36, doi:10.1016/j.rse.2011.11.026, 2012.
- ESA: Sentinel-2 Data Quality Report Issue 24 (February 2018) - Sentinel-2 MSI Document Library - User Guides - Sentinel Online, [online] Available from: [https://earth.esa.int/web/sentinel/user-guides/sentinel-2-msi/document-library/-](https://earth.esa.int/web/sentinel/user-guides/sentinel-2-msi/document-library/-/asset_publisher/Wk0TKajlSaR/content/)
- 15 /asset_publisher/Wk0TKajlSaR/content/ (Accessed 25 January 2018), 2018.
- Fensholt, R., Sandholt, I. and Rasmussen, M. S.: Evaluation of MODIS LAI, fAPAR and the relation between fAPAR and NDVI in a semi-arid environment using in situ measurements, *Remote Sens. Environ.*, 91, 490–507, doi:10.1016/j.rse.2004.04.009, 2004.
- Filella, I., Peñuelas, J., Llorens, L. and Estiarte, M.: Reflectance assessment of seasonal and annual changes in biomass and
- 20 CO₂ uptake of a Mediterranean shrubland submitted to experimental warming and drought, *Remote Sens. Environ.*, 90(3), 308–318, doi:10.1016/J.RSE.2004.01.010, 2004.
- Filella, I., Porcar-Castell, A., Munné-Bosch, S., Bäck, J., Garbulsky, M. F. and Peñuelas, J.: PRI assessment of long-term changes in carotenoids/chlorophyll ratio and short-term changes in de-epoxidation state of the xanthophyll cycle, *Int. J. Remote Sens.*, 30(17), 4443–4455, doi:10.1080/01431160802575661, 2009.
- 25 Filippa, G., Cremonese, E., Migliavacca, M., Galvagno, M., Forkel, M., Wingate, L., Tomelleri, E., Morra di Cella, U. and Richardson, A. D.: Phenopix: A R package for image-based vegetation phenology, *Agric. For. Meteorol.*, 220, 141–150, doi:10.1016/j.agrformet.2016.01.006, 2016.
- Furnival, G. M. and Wilson, R. W.: *Regressions by Leaps and Bounds*, *Technometrics*, 16(4), 499–511, doi:10.1080/00401706.1974.10489231, 1974.
- 30 Galloway, J. N., Dentener, F. J., Capone, D. G., Boyer, E. W., Howarth, R. W., Seitzinger, S. P., Asner, G. P., Cleveland, C. C., Green, P. A., Holland, E. A., Karl, D. M., Michaels, A. F., Porter, J. H., Townsend, A. R. and Vosmarty, C. J.: Nitrogen Cycles: Past, Present, and Future, *Biogeochemistry*, 70(2), 153–226, doi:10.1007/s10533-004-0370-0, 2004.
- Gamon, J. A.: Reviews and Syntheses: optical sampling of the flux tower footprint, *Biogeosciences*, 12(14), 4509–4523, doi:10.5194/bg-12-4509-2015, 2015.

- Gamon, J. A., Peñuelas, J. and Field, C. B.: A narrow-wave band spectral index that tracks diurnal changes in photosynthetic efficiency, *Remote Sens. Environ.*, 41(1), 35–44, 1992.
- Gamon, J. A., Serrano, L. and Surfus, J. S.: The photochemical reflectance index: an optical indicator of photosynthetic radiation use efficiency across species, functional types, and nutrient levels, *Oecologia*, 112, 492–501, doi:10.1007/s004420050337, 1997.
- Gao, B.: NDWI—A normalized difference water index for remote sensing of vegetation liquid water from space, *Remote Sens. Environ.*, 58(3), 257–266, doi:10.1016/S0034-4257(96)00067-3, 1996.
- Garbulsky, M. F., Peñuelas, J., Gamon, J., Inoue, Y. and Filella, I.: The photochemical reflectance index (PRI) and the remote sensing of leaf, canopy and ecosystem radiation use efficiencies A review and meta-analysis, *Remote Sens. Environ.*, 115(2), 281–297, doi:10.1016/j.rse.2010.08.023, 2011.
- 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(2), 435–449, doi:10.1080/01431160802314855, 2009.
- Gitelson, A. and Merzlyak, M. N.: Spectral reflectance changes associated with autumn senescence of *Aesculus-Hippocastanum* L. and *Acer-Platanoides* L leaves - spectral features and relation to chlorophyll estimation, *J. Plant Physiol.*, 143, 286–292, 1994.
- Gitelson, A. A. and Merzlyak, M. N.: Remote sensing of chlorophyll concentration in higher plant leaves, *Adv. Sp. Res.*, 22, 689–692, doi:http://dx.doi.org/10.1016/S0273-1177(97)01133-2, 1998.
- Gitelson, A. A., Kaufman, Y. J. and Merzlyak, M. N.: Use of a green channel in remote sensing of global vegetation from EOS-MODIS, *Remote Sens. Environ.*, 58(3), 289–298, doi:10.1016/S0034-4257(96)00072-7, 1996.
- Gitelson, A. A., Gamon, J. A. and Solovchenko, A.: Multiple drivers of seasonal change in PRI: Implications for photosynthesis 1. Leaf level, *Remote Sens. Environ.*, 191, 110–116, doi:10.1016/j.rse.2016.12.014, 2017a.
- Gitelson, A. A., Gamon, J. A. and Solovchenko, A.: Multiple drivers of seasonal change in PRI: Implications for photosynthesis 2. Stand level, *Remote Sens. Environ.*, 190, 198–206, doi:10.1016/J.RSE.2016.12.015, 2017b.
- Glenn, E., Huete, A., Nagler, P. and Nelson, S.: Relationship Between Remotely-sensed Vegetation Indices, Canopy Attributes and Plant Physiological Processes: What Vegetation Indices Can and Cannot Tell Us About the Landscape, *Sensors*, 8(4), 2136–2160, doi:10.3390/s8042136, 2008.
- Gower, S. T., Kucharik, C. J. and Norman, J. M.: Direct and Indirect Estimation of Leaf Area Index, fAPAR, and Net Primary Production of Terrestrial Ecosystems, *Remote Sens. Environ.*, 70, 29–51, doi:10.1016/S0034-4257(99)00056-5, 1999.
- Grimm, N. B., Grove, J. M., Pickett, S. T. a and Redman, C. L.: Integrated approaches to long-term studies of urban ecological systems, *Bioscience*, 50(7), 571–584, doi:10.1641/0006-3568, 2008.
- Harpole, W. S., Potts, D. L. and Suding, K. N.: Ecosystem responses to water and nitrogen amendment in a California grassland, *Glob. Chang. Biol.*, 13(11), 2341–2348, doi:10.1111/j.1365-2486.2007.01447.x, 2007.
- Hothorn, T., Bretz, F. and Westfall, P.: Simultaneous Inference in General Parametric Models, *Biometrical J.*, 50(3), 346–

- 363, doi:10.1002/bimj.200810425, 2008.
- INMG: O clima de Portugal. Normais climatológicas da região Ribatejo e Oeste correspondentes a 1951-1980., 1991.
- Jarvis, P., Rey, A., Petsikos, C., Wingate, L., Rayment, M., Pereira, J., Banza, J., David, J., Miglietta, F., Borghetti, M., Manca, G. and Valentini, R.: Drying and wetting of Mediterranean soils stimulates decomposition and carbon dioxide emission: the "Birch effect", *Tree Physiol.*, 27(7), 929–940, doi:10.1093/treephys/27.7.929, 2007.
- Joel, G., Gamon, J. A. and Field, C. B.: Production efficiency in sunflower: The role of water and nitrogen stress, *Remote Sens. Environ.*, 62, 176–188, doi:10.1016/s0034-4257(97)00093-x, 1997.
- Jones, H. G.: *Plants and microclimate: A quantitative Approach to Environmental Plant Physiology.*, 2nd ed., Cambridge University Press, Cambridge., 1992.
- 10 Jongen, M., Pereira, J. S., Aires, L. M. I. and Pio, C. A.: The effects of drought and timing of precipitation on the inter-annual variation in ecosystem-atmosphere exchange in a Mediterranean grassland, *Agric. For. Meteorol.*, 151(5), 595–606, doi:10.1016/j.agrformet.2011.01.008, 2011.
- Jongen, M., Lecomte, X., Unger, S., Fangeiro, D. and Pereira, J. S.: Precipitation variability does not affect soil respiration and nitrogen dynamics in the understorey of a Mediterranean oak woodland, *Plant Soil*, 372(1–2), 235–251, 15 doi:10.1007/s11104-013-1728-7, 2013a.
- Jongen, M., Unger, S., Fangeiro, D., Cerasoli, S., Silva, J. M. N. and Pereira, J. S.: Resilience of montado understorey to experimental precipitation variability fails under severe natural drought, *Agric. Ecosyst. Environ.*, 178, 18–30, 2013b.
- Kira, O., Nguy-Robertson, A. L., Arkebauer, T. J., Linker, R. and Gitelson, A. A.: Toward generic models for green LAI estimation in maize and soybean: Satellite observations, *Remote Sens.*, 9(4), 318, doi:10.3390/rs9040318, 2017.
- 20 Korhonen, L., Hadi, Packalen, P. and Rautiainen, M.: Comparison of Sentinel-2 and Landsat 8 in the estimation of boreal forest canopy cover and leaf area index, *Remote Sens. Environ.*, 195, 259–274, doi:10.1016/j.rse.2017.03.021, 2017.
- Kovats, R. S., Valentini, R., Bouwer, L. M., Georgopoulou, E., Jacob, D., Martin, E. and Rounsevell, M. Soussana, J. F.: Europe, in *Climate Change 2014: Impacts, Adaptation, and Vulnerability. Part B: Regional Aspects. Contribution of Working Group II to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change*, pp. 1267–1326, 25 Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA., 2014.
- Loveland, T. R. and Irons, J. R.: Landsat 8: The plans, the reality, and the legacy, *Remote Sens. Environ.*, 185, 1–6, doi:10.1016/j.rse.2016.07.033, 2016.
- Lumley, T.: Thomas Lumley using Fortran code by Alan Miller (2009). Leaps: regression subset selection. R package version 2.9., 2009.
- 30 Matthes, J. H., Knox, S. H., Sturtevant, C., Sonnentag, O., Verfaillie, J. and Baldocchi, D.: Predicting landscape-scale CO₂ flux at a pasture and rice paddy with long-term hyperspectral canopy reflectance measurements, *Biogeosciences*, 12(15), 4577–4594, doi:10.5194/bg-12-4577-2015, 2015.
- Merzlyak, M. N., Gitelson, A. A., Chivkunova, O. B. and Rakitin, V. Y.: Non-destructive optical detection of pigment changes during leaf senescence and fruit ripening, *Physiol. Plant.*, 106(1), 135–141, doi:10.1034/j.1399-

- 3054.1999.106119.x, 1999.
- Monteith, J. L.: Solar Radiation and Productivity in Tropical Ecosystems, *J. Appl. Ecol.*, 9(3), 747–766, doi:10.2307/2401901, 1972.
- Monteith, J. L.: Climate and efficiency of crop production in Britain, *Philos. Trans. R. Soc. London Ser. B-Biological Sci.*,
5 281, 277–294, 1977.
- Moreno, A., Maselli, F., Gilabert, M. A., Chiesi, M., Martinez, B. and Seufert, G.: Assessment of MODIS imagery to track light-use efficiency in a water-limited Mediterranean pine forest, *Remote Sens. Environ.*, 123, 359–367, doi:10.1016/j.rse.2012.04.003, 2012.
- Myneni, R. B. and Williams, D. L.: On the relationship between FAPAR and NDVI, *Remote Sens. Environ.*, 49(3), 200–
10 211, 1994.
- Naudts, K., Chen, Y., McGrath, M. J., Ryder, J., Valade, A., Otto, J. and Luyssaert, S.: Europes forest management did not mitigate climate warming, *Science* (80-.), 351(6273), 597–600, doi:10.1126/science.aad7270, 2016.
- Nestola, E., Calfapietra, C., Emmerton, C. A., Wong, C. Y. S., Thayer, D. R. and Gamon, J. A.: Monitoring grassland seasonal carbon dynamics, by integrating MODIS NDVI, proximal optical sampling, and eddy covariance measurements,
15 *Remote Sens.*, 8(3), 260, doi:10.3390/rs8030260, 2016.
- Ohtani, K.: Bootstrapping R2 and adjusted R2 in regression analysis, *Econ. Model.*, 17(4), 473–483, doi:10.1016/S0264-9993(99)00034-6, 2000.
- Ollinger, S. V.: Sources of variability in canopy reflectance and the convergent properties of plants, *New Phytol.*, 189(2), 375–394, doi:10.1111/j.1469-8137.2010.03536.x, 2011.
- 20 Penuelas et al., J.: The reflectance at the 950-970nm region as an indicator of plant water status, *Int. J. Remote Sens.*, 14(10), 1887–1905, doi:0143-1161/93, 1993.
- Penuelas, J., Pinol, J., Ogaya, R. and Filella, I.: Estimation of plant water concentration by the reflectance Water Index WI (R900/R970), *Int. J. Remote Sens.*, 18(13), 2869–2875, doi:10.1080/014311697217396, 1997.
- Peñuelas, J., Filella, I. and Gamon, J. A.: Assessment of photosynthetic radiation-use efficiency with spectral reflectance,
25 *New Phytol.*, 131(3), 291–296, doi:10.1111/j.1469-8137.1995.tb03064.x, 1995.
- Peñuelas, J., Poulter, B., Sardans, J., Ciais, P., van der Velde, M., Bopp, L., Boucher, O., Godderis, Y., Hinsinger, P., Llusia, J., Nardin, E., Vicca, S., Obersteiner, M. and Janssens, I. A.: Human-induced nitrogen–phosphorus imbalances alter natural and managed ecosystems across the globe, *Nat. Commun.*, 4, 2934, doi:10.1038/ncomms3934, 2013.
- Pérez-Priego, O., López-Ballesteros, A., Sánchez-Cañete, E., Serrano-Ortiz, P., Kutzbach, L., Domingo, F., Eugster, W. and
30 Kowalski, A.: Analysing uncertainties in the calculation of fluxes using whole-plant chambers: random and systematic errors, *Plant Soil*, 393(1–2), 229–244, doi:10.1007/s11104-015-2481-x, 2015a.
- Pérez-Priego, O., Guan, J., Rossini, M., Fava, F., Wutzler, T., Moreno, G., Carvalhais, N., Carrara, A., Kolle, O., Julitta, T., Schrupf, M., Reichstein, M. and Migliavacca, M.: Sun-induced chlorophyll fluorescence and photochemical reflectance index improve remote-sensing gross primary production estimates under varying nutrient availability in a typical

- Mediterranean savanna ecosystem, *Biogeosciences*, 12(21), 6351–6367, doi:10.5194/bg-12-6351-2015, 2015b.
- Porcar-Castell, A., Garcia-Plazaola, J. I., Nichol, C. J., Kolari, P., Olascoaga, B., Kuusinen, N., Fernandez-Marin, B., Pulkkinen, M., Juurola, E. and Nikinmaa, E.: Physiology of the seasonal relationship between the photochemical reflectance index and photosynthetic light use efficiency, *Oecologia*, 170(2), 313–323, doi:10.1007/s00442-012-2317-9, 2012.
- 5 Porcar-Castell, A., Mac Arthur, A., Rossini, M., Eklundh, L., Pacheco-Labrador, J., Anderson, K., Balzarolo, M., Martín, M. P., Jin, H., Tomelleri, E., Cerasoli, S., Sakowska, K., Hueni, A., Julitta, T., Nichol, C. J. and Vescovo, L.: EUROSPEC: at the interface between remote-sensing and ecosystem CO₂ flux measurements in Europe, *Biogeosciences*, 12(20), 6103–6124, doi:10.5194/bg-12-6103-2015, 2015.
- Reichstein, M., Rey, A., Freibauer, A., Tenhunen, J., Valentini, R., Banza, J., Casals, P., Cheng, Y., Grünzweig, J. M., Irvine, J., Joffre, R., Law, B. E., Loustau, D., Miglietta, F., Oechel, W., Ourcival, J.-M., Pereira, J. S., Peressotti, A., Ponti, F., Qi, Y., Rambal, S., Rayment, M., Romanya, J., Rossi, F., Tedeschi, V., Tirone, G., Xu, M. and Yakir, D.: Modeling temporal and large-scale spatial variability of soil respiration from soil water availability, temperature and vegetation productivity indices, *Global Biogeochem. Cycles*, 17(4), n/a-n/a, doi:10.1029/2003GB002035, 2003.
- 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(7), 2565–2584, doi:10.5194/bg-9-2565-2012, 2012.
- 15 Rouse, J. W., Haas, R. H., Schell, J. A. and Deering, D. W.: Monitoring vegetation systems in the great plains with ERTS, in *Third ERTS Symposium*, vol. 1, edited by N. SP-351, pp. 309–317, NASA, Washington, DC, USA., 1974.
- Sakowska, K., Vescovo, L., Marcolla, B., Juszczak, R., Olejnik, J. and Gianelle, D.: Monitoring of carbon dioxide fluxes in a subalpine grassland ecosystem of the Italian Alps using a multispectral sensor, *Biogeosciences*, 11, 4695–4712, doi:10.5194/bg-11-4695-2014, 2014.
- 20 Sakowska, K., Juszczak, R. and Gianelle, D.: Remote Sensing of Grassland Biophysical Parameters in the Context of the Sentinel-2 Satellite Mission, *J. Sensors*, 2016, 1–16, doi:10.1155/2016/4612809, 2016.
- Sala, O. E.: Global Biodiversity Scenarios for the Year 2100, *Science* (80-.), 287(5459), 1770 LP-1774, 2000.
- 25 Schenk, H. J. and Jackson, R. B.: Rooting depths, lateral root spreads and below-ground/above-ground allometries of plants in water-limited ecosystems, *J. Ecol.*, 90(3), 480–494, doi:10.1046/j.1365-2745.2002.00682.x, 2002.
- Schimel, D., Pavlick, R., Fisher, J. B., Asner, G. P., Saatchi, S., Townsend, P., Miller, C., Frankenberg, C., Hibbard, K. and Cox, P.: Observing terrestrial ecosystems and the carbon cycle from space, *Glob. Chang. Biol.*, 21(5), 1762–1776, doi:10.1111/gcb.12822, 2015.
- 30 Seabloom, E. W., Borer, E. T., Buckley, Y., Cleland, E. E., Davies, K., Firn, J., Harpole, W. S., Hautier, Y., Lind, E., MacDougall, A., Orrock, J. L., Prober, S. M., Adler, P., Alberti, J., Michael Anderson, T., Bakker, J. D., Biederman, L. A., Blumenthal, D., Brown, C. S., Brudvig, L. A., Caldeira, M., Chu, C., Crawley, M. J., Daleo, P., Damschen, E. I., D’Antonio, C. M., DeCrappeo, N. M., Dickman, C. R., Du, G., Fay, P. A., Frater, P., Gruner, D. S., Hagenah, N., Hector, A., Helm, A., Hillebrand, H., Hofmockel, K. S., Humphries, H. C., Iribarne, O., Jin, V. L., Kay, A., Kirkman, K. P., Klein, J. A., Knops, J.

- M. H., La Pierre, K. J., Ladwig, L. M., Lambrinos, J. G., Leakey, A. D. B., Li, Q., Li, W., McCulley, R., Melbourne, B., Mitchell, C. E., Moore, J. L., Morgan, J., Mortensen, B., O'Halloran, L. R., Pärtel, M., Pascual, J., Pyke, D. A., Risch, A. C., Salguero-Gómez, R., Sankaran, M., Schuetz, M., Simonsen, A., Smith, M., Stevens, C., Sullivan, L., Wardle, G. M., Wolkovich, E. M., Wragg, P. D., Wright, J. and Yang, L.: Predicting invasion in grassland ecosystems: is exotic dominance the real embarrassment of richness?, *Glob. Chang. Biol.*, 19(12), 3677–3687, doi:10.1111/gcb.12370, 2013.
- 5 Tagesson, T., Fensholt, R., Huber, S., Horion, S., Guiro, I., Ehammer, A. and Ardö, J.: Deriving seasonal dynamics in ecosystem properties of semi-arid savanna grasslands from in situ-based hyperspectral reflectance, *Biogeosciences*, 12(15), 4621–4635, doi:10.5194/bg-12-4621-2015, 2015.
- Thomas, H.: Senescence, ageing and death of the whole plant, *New Phytol.*, 197(3), 696–711, doi:10.1111/nph.12047, 2013.
- 10 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 wavelengths for remote detection of grassland phytomass, *Int. J. Remote Sens.*, 33(7), 2178–2195, doi:10.1080/01431161.2011.607195, 2012.
- Vicca, S., Balzarolo, M., Filella, I., Granier, A., Herbst, M., Knohl, A., Longdoz, B., Mund, M., Nagy, Z., Pintér, K., Rambal, S., Verbesselt, J., Verger, A., Zeileis, A., Zhang, C. and Peñuelas, J.: Remotely-sensed detection of effects of
- 15 extreme droughts on gross primary production, *Sci. Rep.*, 6(1), 28269, doi:10.1038/srep28269, 2016.
- Viña, A. and Gitelson, A. A.: New developments in the remote estimation of the fraction of absorbed photosynthetically active radiation in crops, *Geophys. Res. Lett.*, 32(17), doi:10.1029/2005GL023647, 2005.
- Weisser, W. W., Roscher, C., Meyer, S. T., Ebeling, A., Luo, G., Allan, E., Beßler, H., Barnard, R. L., Buchmann, N., Buscot, F., Engels, C., Fischer, C., Fischer, M., Gessler, A., Gleixner, G., Halle, S., Hildebrandt, A., Hillebrand, H., de
- 20 Kroon, H., Lange, M., Leimer, S., Le Roux, X., Milcu, A., Mommer, L., Niklaus, P. A., Oelmann, Y., Proulx, R., Roy, J., Scherber, C., Scherer-Lorenzen, M., Scheu, S., Tschardtke, T., Wachendorf, M., Wagg, C., Weigelt, A., Wilcke, W., Wirth, C., Schulze, E.-D., Schmid, B. and Eisenhauer, N.: Biodiversity effects on ecosystem functioning in a 15-year grassland experiment: Patterns, mechanisms, and open questions, *Basic Appl. Ecol.*, 23, 1–73, doi:10.1016/j.baae.2017.06.002, 2017.
- Westoby, M., Falster, D. S., Moles, A. T., Vesk, P. A. and Wright, I. J.: Plant Ecological Strategies: Some Leading
- 25 Dimensions of Variation Between Species, *Annu. Rev. Ecol. Syst.*, 33(1), 125–159, doi:10.1146/annurev.ecolsys.33.010802.150452, 2002.
- WRB, I. W. G.: World reference base for soil resources 2006. 2nd edition. World Soil Resources Reports No.103., 2006.
- Xu, L. and Baldocchi, D. D.: Seasonal variation in carbon dioxide exchange over a Mediterranean annual grassland in California, *Agric. For. Meteorol.*, 123(1–2), 79–96, doi:10.1016/j.agrformet.2003.10.004, 2004.
- 30 Yuan, W., Cai, W., Xia, J., Chen, J., Liu, S., Dong, W., Merbold, L., Law, B., Arain, A., Beringer, J., Bernhofer, C., Black, A., Blanken, P. D., Cescatti, A., Chen, Y., Francois, L., Gianelle, D., Janssens, I. A., Jung, M., Kato, T., Kiely, G., Liu, D., Marcolla, B., Montagnani, L., Raschi, A., Rouspard, O., Varlagin, A. and Wohlfahrt, G.: Global comparison of light use efficiency models for simulating terrestrial vegetation gross primary production based on the LaThuile database, *Agric. For. Meteorol.*, 192–193, 108–120, doi:10.1016/j.agrformet.2014.03.007, 2014.

Zhang, F. and Zhou, G.: Deriving a light use efficiency estimation algorithm using *in situ* hyperspectral and eddy covariance measurements for a maize canopy in Northeast China, *Ecol. Evol.*, 7(13), 4735–4744, doi:10.1002/ece3.3051, 2017.

|

Table 1. Spectral bands range and spatial resolution of Sentinel-2A MSI and Landsat 8 OLI sensors simulated in this study.

Band	Sentinel-2A MSI			Landsat 8 OLI		
	Spectral region	Wavelength range (nm)	Resolution (m)	Spectral region	Wavelength range (nm)	Resolution (m)
B1				Blue	435-451	30
B2	Blue	458-523	10	Blue	452-512	30
B3	Green peak	543-578	10	Green	533-590	30
B4	red	650-680	10	Red	636-673	30
B5	Red-edge	698-713	20	NIR	851-879	30
B6	Red-edge	733-748	20	SWIR1	1566-1651	30
B7	Red-edge	773-793	20	SWIR2	2107-2294	30
B8	NIR	785-899	10			
B8A	NIR narrow	855-875	20			
B11	SWIR	1565-1655	20			
B12	SWIR	2100-2280	20			

Table 2. Selection of vegetation indices adopted in this study with their formulation using hyperspectral (Hyp) grouped bands, S2 or L8 simulated sensors, biophysical properties represented according to the literature and original bibliographic reference.

	Vegetation Index	Hyp	S2	L8	Biophysical property	Reference
NDVI	Normalized Difference	$\frac{R_{800} - R_{670}}{R_{800} + R_{670}}$	$\frac{B8A - B4}{B8A + B4}$	$\frac{B5 - B4}{B5 + B4}$	Green biomass and area	(Rouse et al., 1974)
	Vegetation Index					
GNDVI	Green Normalized Diff.	$\frac{R_{750} - R_{550}}{R_{750} + R_{550}}$	$\frac{B7 - B3}{B7 + B3}$	$\frac{B5 - B3}{B5 + B3}$	Green biomass and area	(Gitelson and Merzlyak, 1998)
	Veg. Ind.					
NDVIre	Red-edge Normalized	$\frac{R_{750} - R_{720}}{R_{720} + R_{750}}$			Green biomass	(Gitelson and Merzlyak, 1994)
	Diff. Veg. Ind.					
CI	Chlorophyll index	$\frac{R_{750} - R_{705}}{R_{750} + R_{705}}$			Chlorophyll	(Gitelson and Merzlyak, 1994)
MTCI	MERIS Terrestrial	$\frac{R_{754} - R_{709}}{R_{709} + R_{681}}$	$\frac{B6 - B5}{B5 + B4}$		Chlorophyll, nitrogen	(Dash and Curran, 2004)
	chlorophyll Index					
PRI	Photochemical	$\frac{R_{570} - R_{531}}{R_{570} + R_{531}}$			Radiation Use-Efficiency, Carotenoid/chlorophyll	(Gamon et al., 1992)
	Reflectance Index					
PSRI	Plant Senescence	$\frac{R_{680} - R_{500}}{R_{750}}$	$\frac{B4 - B3}{B6}$		Carotenoid/chlorophyll	(Merzlyak et al., 1999)
	Reflectance Index					
NDWI	Normalized Difference	$\frac{R_{860} - R_{1240}}{R_{860} + R_{1240}}$			Tissue water content	(Gao, 1996)
	Water Index					
WBI	Water Band Index	$\frac{R_{970} - R_{900}}{R_{970} + R_{900}}$			Tissue water content	(Peñuelas Peñuelas et al.,

1993)

Tab.3 - Percentage of each plant functional type (Forbs, Graminoids and Legumes) in above ground biomass of grasslands under different fertilization treatments (C, N, NPK and P). Values represent means of 6 replicates; standard errors are shown in parenthesis.

Treatment	Forbs	Grams	Legumes
C	56.85 (5.10)	21.22 (3.62)	21.93 (4.20)
N	65.00 (1.89)	25.04 (2.53)	9.95 (1.45)
NPK	34.07 (3.43)	52.55 (3.29)	13.37 (1.03)
P	25.60 (3.06)	31.43 (4.47)	42.96 (3.82)

Formatted: Font: Not Italic

Table 4. Linear regressions established between lnGPP and vegetation indices (VI) selected for this study (see Table 2). Best regression is shown in bold.

Vegetation Index	R^2	RMSE	P
NDVI	0.6853	0.2364	0.0000
GNDVI	0.6360	0.2543	0.0000
NDVire	0.6872	0.2357	0.0000
CI	0.7161	0.2246	0.0000
MTCI	0.6303	0.2563	0.0000
PRI	0.0209	0.4171	0.1715
PSRI	0.6745	0.2405	0.0000
NDWI	0.7205	0.2228	0.0000
WBI	0.6491	0.2497	0.0000

Table 5. Best selection of linear models for GPP estimate according to the general equation: $\ln GPPP \sim \sum_{j=1}^n v_j$, where v are vegetation indices (VIs) or optical bands (B). Bands and vegetation indices are obtained from Hyperspectral measurements grouped in clusters with 90% similarity (Hyp), or resampled for simulating Sentinel 2/MSI (S2) and Landsat8/OLI (L8) sensors. Vegetation indices formulation is shown in Table 2. The order of the variables (most important first) reflects their importance in the model. The Quantiles 25% and 75% of the $adj R^2$ are obtained from a bootstrap with 10000 iterations. The [two-step](#) models add a selection of bands to the variables (VIs) selected at step one. A low p-value indicates that the model including VIs and bands (step 2) is significantly better than the model just with VIs (step 1).

Model	Step one	$Adj R^2$	$Adj R^2$ Q-25%	$Adj R^2$ Q-75%	Step two	$Adj R^2$	$Adj R^2$ Q 25%	$Adj R^2$ Q75%	p
Hyp-VIs	NDWI; PSRI; WBI; GNDVI	0.7659	0.7431	0.8047					
S2-VIs	MTCI; PSRI; GNDVI; NDVI	0.7426	0.7225	0.7822					
L8-VIs	NDVI	0.6792	0.6405	0.7194					
Hyp-B	$R_{1951-2299}$; $R_{724-732}$; $R_{1328-1349}$; $R_{706-710}$; $R_{449-466}$; $R_{566-582}$; $R_{519-532}$; $R_{350-397}$; $R_{398-411}$; $R_{1209-1327}$; $R_{702-705}$; $R_{698-701}$; $R_{716-723}$;	0.7884	0.7906	0.8392					
S2-B	B7; B11; B5; B2; B8; B6	0.7412	0.7222	0.7848					
L8-B	B7; B5; B6; B4; B3; B1	0.7557	0.7367	0.7974					
Hyp-VIs+B	NDWI; PSRI; WBI; GNDVI				$R_{698-701}$; $R_{412-448}$; $R_{716-723}$; $R_{467-518}$; $R_{706-710}$; $R_{449-466}$; $R_{350-397}$; $R_{1209-1327}$; $R_{1412-1505}$; $R_{1951-2299}$; $R_{702-705}$; $R_{724-732}$; $-R_{1328-1349}$	0.7986	0.8083	0.8550	0.0260
S2-VIs+B	MTCI; PSRI; GNDVI; NDVI				B11; B3; B12	0.7684	0.7542	0.8104	0.0081
L8-VIs+B	NDVI				B6; B3; B7	0.7686	0.7472	0.8047	0.0000

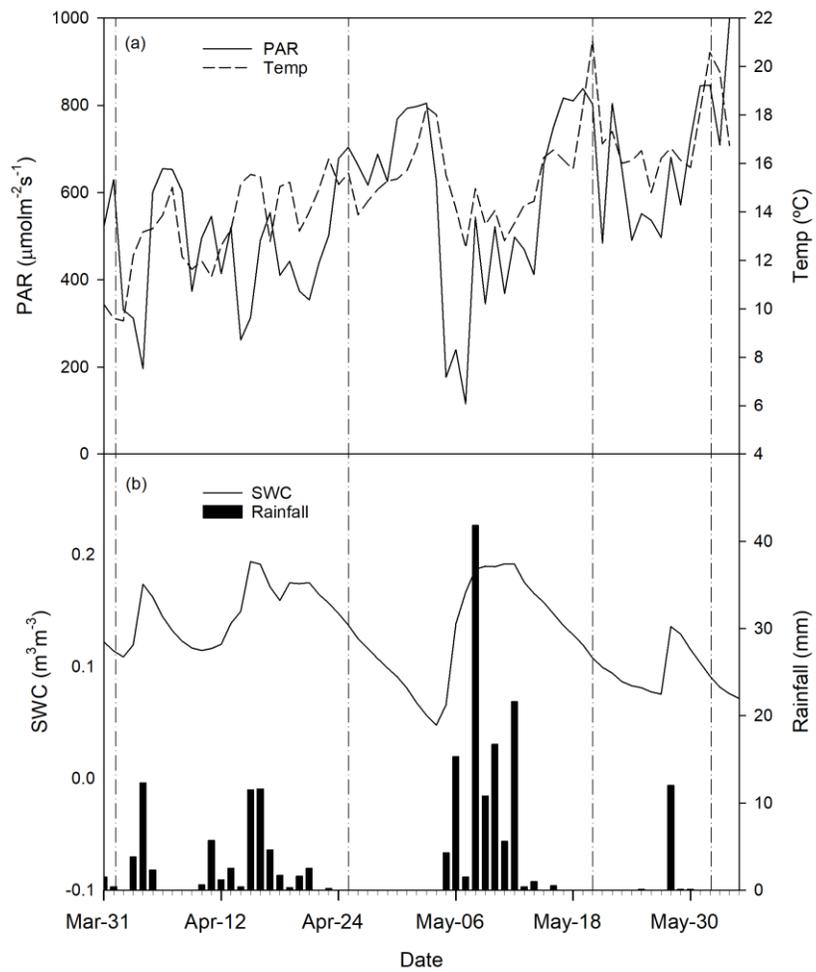


Figure 1: Daily average PAR, temperature (a), soil water content for different treatments and total rainfall (b) recorded on the site during the experimental period. Dates of field measurements are indicated by vertical dash-dotted lines.

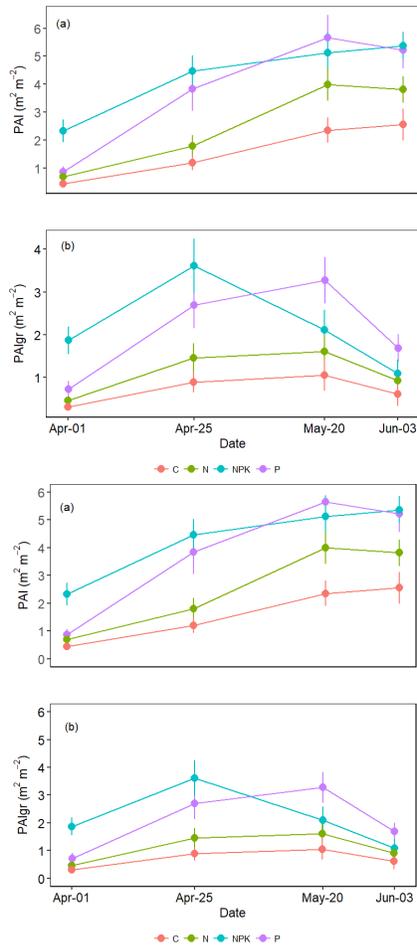
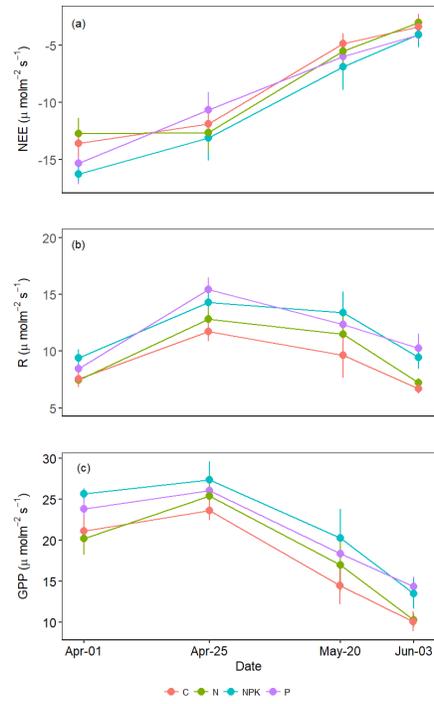


Figure 2 Average green Plant Area index (PAI) and the green fraction of PAI (PAIgr) observed in grasslands subjected to different fertilization treatments (C, N, NPK or P). Each point is the average of 6 replicates. Vertical bars represent error bars.



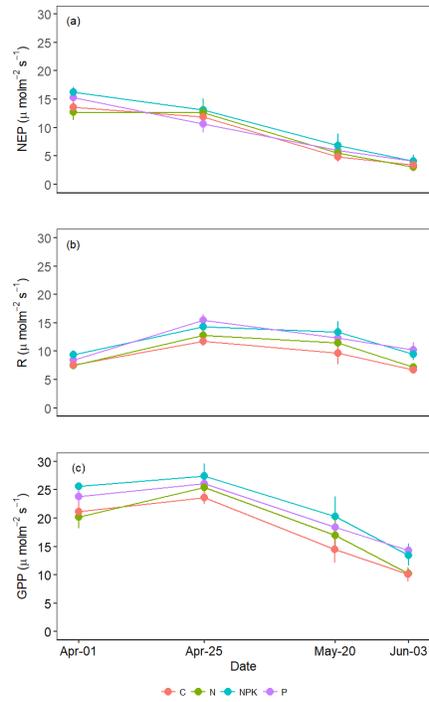


Fig.3 Average Net Ecosystem **ProductivityExchange** (NEPE)(a), dark Respiration (R) (b) and Gross Primary Productivity (GPP) (c) measured in grasslands under different fertilization regimes (C, N, NPK and P). Each point is the average of 6 replicates. Vertical bars represent standard errors.

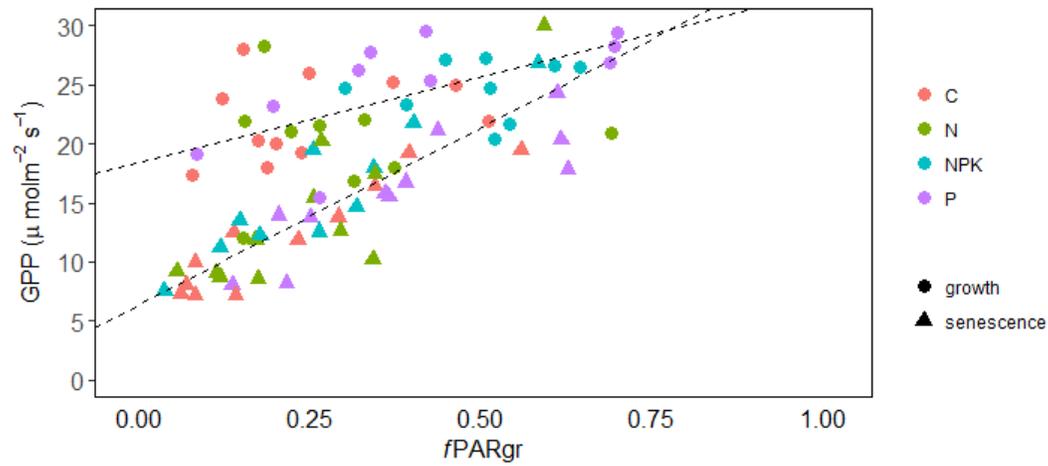


Figure 4: Relationship between GPP and $fPAR_{gr}$ observed in grassland subjected to different fertilization treatment (C, N, NPK and P). Measurements performed during the vegetation growth are indicated as circles (corresponding to measurement days 1 and 2) while those realized during the senescence period are indicated as triangles (measurement days 3 and 4). Linear regression lines were fitted separately to the two periods ($GPP = 14.48 fPAR_{gr} + 18.44$, $R^2 = 0.39$, $P < 0.001$ and $GPP = 30.08 fPAR_{gr} + 6.39$, $R^2 = 0.65$, $P < 0.001$, respectively for the growth and the senescence period).

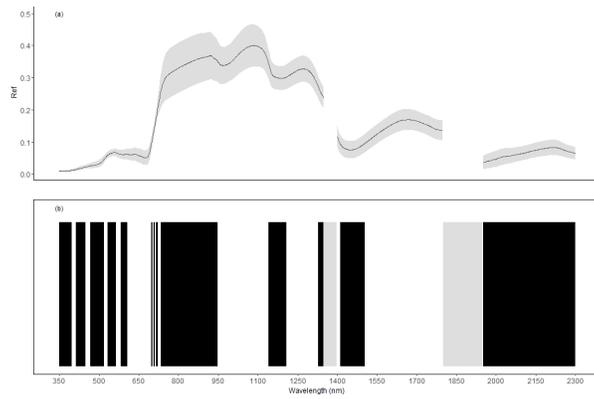


Fig.54 Average reflectance values and standard deviation (grey ribbon) observed in herbaceous plots undergoing different fertilization treatments (A). The bottom picture (B) shows the bands obtained by grouping Ref for similarity ($\geq 90\%$) of contiguous hyperspectral measurements with 1nm resolution in the range 350-2300nm, bands are alternated black and white. Grey bars represent areas of the spectrum not considered for being noise.

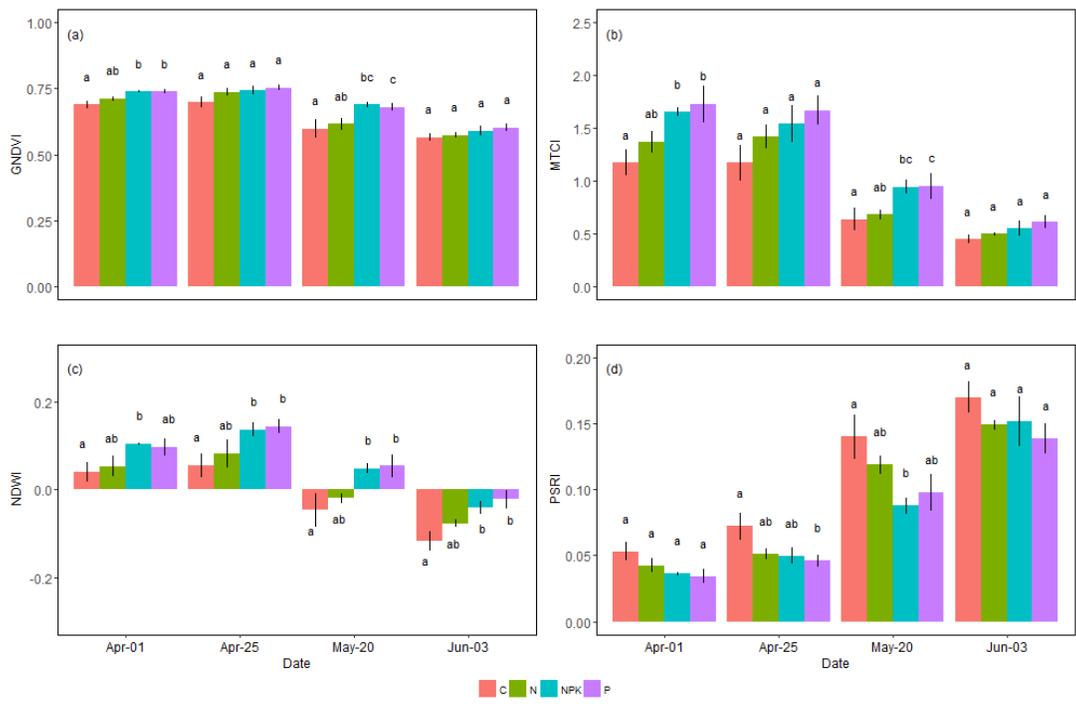


Fig 65. Average values of several vegetation indices retrieved from reflectance measurements of herbaceous plots undergoing different fertilization treatments. Vertical bars represent standard errors. Different letters indicate significant differences among treatments within the same date (p<0.05).