# Referee #1

Overall I think this is a well described and comprehensive study that provides a valuable contribution to the phenology- and flux-related literature. The main point I’d raise is that the manuscript could be improved with a more detailed description of the radiative transfer modelling, as well as a fuller explanation of the function linking the reflectance that is output from the PROSAIL model and the DN values derived from the camera images (i.e. equation 2). This will aid the interpretation of results in Section 3.2.

Response: We thank referee #1 for their positive and constructive comments. Both yourself and referee #2 presented a strong case for providing more details on the model and the function that relates the reflectance outputs of PROSAIL to the DN values extracted from digital images. In order to make the modelling component of this paper more transparent and to encourage the testing of the model at other sites we created a git repository for the documented code and included the data contained in Figure 12 on a Bitbucket account (https://bitbucket.org/jerome\_ogee/webcam\_network\_paper). Should the paper be accepted for publication in Biogeosciences, this code and dataset will be open to public access. Our hope is that by placing it in a git repository the research community will actively contribute to the improvement of the code and tools to assess its sensitivity. We will place the link information for this git repository in the methods section 2.3 of the paper and again in the legend of Fig. 12.

In addition, the assumptions behind this type of RT model should be discussed.

Response: The RT model used here, SAIL, assumes that diffusers are randomly distributed in space (turbid medium assumption). We completely agree that such assumption will not be applicable for very clumped canopies such as sparse forests or crops (e.g. vineyards, orchards…). Also the RT model assumes only one type of foliage, and therefore cannot deal with species mixtures. For mixed forests, PROSAIL can still be used to interpret RGB signals if the ROI used on the images is dominated by one single species (e.g. 90% Oak cover for the forest simulated in Fig. 12). These caveats have now been raised in the revised version of the manuscript.

Finally, the use or possible applications of these data and the modeling framework could be expanded upon in the discussion or conclusions. Perhaps it is worth mentioning here what complementary information these data will bring to phenology metrics derived from satellite data. For example the data are available smaller scale at the location of flux towers which opens up the opportunity to understand differences between growing season length and carbon uptake period, in addition to linking the trends in phenology with net C fluxes in order to determine the impacts of greening on the balance of C uptake and respiration. Or this could be added in the discussion.

Response: We thank the referee for pointing out the advantages of seasonally resolved camera data for interpreting fluxes and differences between GSL and the CUP. We have added a paragraph covering these topics in the conclusion.

P8 Lines 14-16: Please can you describe why this is necessary in a bit more detail – for readers who are not familiar with more technical aspects of cameras?

Response: We are not sure we fully understand this comment or more exactly what parts of the manuscript it refers to. Our statement on page 8 lines 14-16 is that it is *necessary* to have the same ROI between images and we clearly explain why by saying it is otherwise *problematic* (for seasonal characterization) and *impractical* (would require to check the ROI on every single image). Maybe the comment refers to another page as the next comment referenced to a page below (page 9 line 18) clearly refers to page 8 line 18…

Also are the same lenses used on each camera? Are the aperture size, shutter speed, ISO and sensor gain kept constant? What is the difference in sensor size between the two types of cameras?

Response: Unfortunately the network currently uses different types of cameras, and manufacturers do not always give detailed information on the optics, the sensors or the image processing of the cameras. As part of the ICOS infrastructure, a protocol has been proposed to homogenise as much as possible the settings of these different cameras and is briefly described in section 2.1 of the manuscript. Also, based on the results shown in Fig. 12 and S7 (now S8) we believe that, as long as we are interested in colour fractions over large ROIs, details regarding the physics and signal processing of the camera (sensor size, aperture, shutter speed…) can all be accounted through the spectral response of the camera (*G*RBG()), the color balance factor (*B*RGB) and the UV and IR cut-off wavelengths defined in equation 2.

P 9 Line 18: For the calculation of the mean color fraction is the mean value of ncolor, nred etc is calculated for all pixels in the ROI, and then the Color Fraction is calculated? Or is the color fraction calculated per pixel and then the mean taken?

Response: In our analysis the mean values for nred, ngreen & nblue were calculated over all pixels in the ROI and the colour fraction was computed from these mean DN values.

Did you investigate what the breakpoint-derived dates might correspond to in the images? For example, the first breakpoint may be 30% leaf out, instead of 50%? Does it correspond to a particular change in photosynthetic parameters, or C fluxes?

Response: In some cases where information was available yes. From Figure 3 we see that the first breakpoint is usually detected before 50% of the ROI contains green leaves (our criteria for leaf out detection). The hardest breakpoints to investigate were those of the evergreen needle-leaf canopies for example in Hyytiala (Fig. 5). We described in the manuscript how these were often linked directly to the appearance and disappearance of snow and indirectly to changes in temperature. In the case of grasslands and evergreen broadleaves, many of these breakpoints were linked to management practices such as mowing (eg. Neustift and Fruebuel), appearance and disappearance of snow and importantly flowering events. In the deciduous broadleaf forests the breakpoints nearly always picked out leaf on and leaf off, but in addition the often detected the maximum and minimum dates of the ‘spring hump’. As discussed in the paper and after simulations with the model we see that these breakpoints are linked strongly to the relationships between pigment concentrations shown in Fig 11. Thus these extra breakpoints in deciduous ecosystems are providing indicative dates of how quickly the photosynthetic apparatus is being assembled and when it reaches its photosynthetic maximum.

It may also be interesting to investigate if there was as consistent a (or similar) bias if: a) A different breakpoint was analysed for each phenophase (for example using the 2nd breakpoint for leaf out). b) A different automatic detection method was used, such as the DOY at half maximum which could be site specific and therefore might account for the issue of different camera set-ups?

Response: We fully agree that further analysis of these additional breakpoints may vary consistently with the leaf on dates and coincide with key parameters derived from other automatic detection methods. However, for the time being consistent information about key phonological phases is not available at all the sites, hampering a comprehensive analysis as the one suggested by the referee. This is certainly a study that should be conducted in the future.

4) Also for sites where more than these 3 breakpoints were detected, did you investigate what the other breakpoints might correspond to?

Response: Yes when possible (see our response two comments above).

Do you think the same procedure should be used for all vegetation/canopy types?

Response: I think it really depends on the question to be tackled within a particular study and the ecosystem under investigation. For systems with management and ecological studies interested in all the phenophases displayed by canopies including flowering and fruiting, I believe this type of breakpoint approach could be very useful. For studies that want to study snow dynamics and vegetation responses this again could also be an attractive choice for example in conifer sites or in semi-arid grasslands where phenology is regulated by rain pulses. In most canopy phenology studies or remote sensing studies, the start, 50% max and end dates are often the most important dates that need to be resolved and I believe this approach may be as good as others on deciduous ecosystems and non-managed grasslands to obtain such dates. In our case, using this approach identified significant breakpoint changes in the green signal during spring that were consistently found at all the deciduous forests studied. This was part of the motivation to understand the emergent properties of the green fraction time-series during spring, explain it mechanistically and understand better how it related to the peak in GPP.

One final point here: perhaps this validation could be added as a section to the be- ginning of the results, given it is an analysis, rather than a description, of the method used?

Response: We agree that figure 3 shows results from an analysis but it is needed here to show the success of our method to automatically detect break points in the RGB signals. In that sense we felt it should rather appear in the methods section, allowing the results section to focus on more ecological aspects.

P 12 Line 10: How is the spectral efficiency defined? Is it just the spectral response curve that is used? A fuller explanation of this and the derivation of the BRGB parameter would help to explain the function (here I will call it an observation operator or OO for the sake of brevity) define in Equation 2. It would be valuable to detail the issues with the camera specifications such as the response functions and how they impact the derivation of GRGB. Also, an initial sentence explaining that you need this type of function to match the DN values measured by the camera sensor to the reflectance simulated by the PROSAIL (taking into account the camera/set-up specific characteristics) model may help readers that are not familiar with this topic. Are any of the camera sensor characteristics not taken into account? And if so are these mostly lumped into the BRGB parameter?

Response: We have added a sentence to the manuscript as suggested by the referee to clarify the link between the camera and PROSAIL outputs. In addition we have provided the derivation in the documented model code and placed it in a git repository (https://bitbucket.org/jerome\_ogee/webcam\_network\_paper).

5) P12 Line 21: Perhaps for the general reader it may help to detail that the SAIL part of the model essentially scales from leaf to canopy. A bit more information on what type of RT model it is (i.e. turbid medium) and the assumptions that are made about the canopy structure for this model.

Response: See our response to the second comment of referee 1.

6)It would be good to explain why this type of RT model was used and not another (e.g. a Geometric-Optic model, or a multi-layer or 3D RT model).

Response: Our choice of the RT model was driven by simplicity and efficiency. More complex RT models such as ray tracing and 3D models require a very heavy parameterisation that would be un-applicable within a network covering so many different ecosystem types. It would also result in an over-pararameterisation of the system with too many degrees of freedom to address ecological questions.

Section 3.1.1

How might the automatic detection of phenophases be improved for coniferous sites do you think? Might it be improved if all the color fraction time series were included in the analysis?

Response: Detecting the phenophases of conifers is a real challenge as they do not produce very large shifts in any of the colour fraction signals that can be easily related to phenology. However, we have not done any breakpoint analysis on these other colour signals but we intend to systematically implement such analysis in the phenopix package (http://r-forge.r-project.org/projects/phenopix/) in the future.

By visual comparison do the breakpoints detected for evergreen sites correspond to any obvious phenophases? Is the leaf flush and browing that you describe on P 15 Line 19 onward evident when you examine the images, or are the dates of this leaf flush etc (i.e. arrows in Figure S4) for each particular site/year known from field measurements?

Response: In the case of the Hyytiala conifer site shown in Figure S4, the shoot elongation and needle flushing episodes are visually identified in the photos but are not detected by the breakpoint approach (i.e. 5 breakpoints maximum).

P15 Line 24 – it would be easier to examine Figure 4 if you gave rough dates for the phenological events you describe in the text (as you have done for leaf flushing) given the multiple points of increase and decrease for the evergreen sites.

Response: In the case of Hyytiala it was possible to identify the dates of shoot elongation and leaf flushing manually from the digital images (see Figure S4). However for most of the other sites the distance between the camera and the trees is large making it extremely difficult to resolve such detail and identify the phenophases. Locating cameras closer to the canopy at these sites would certainly improve phenology measurements. However, we are aware that there are some logistical constraints often imposed at flux sites given also that the camera should not be placed too close to the canopy to still get images that are representative of the flux footprint.

Figure 4 (and 6): given these are latitudinal comparisons it may be nice to have the approximate latitude given on each plot.

Response: We have added the latitude coordinates on each panel.

Figure 5: it might be beneficial to extent the dashed lines up through the temperature and PPFD?

Response: We have extended the lines through all the panels on Fig 5.

7)P16 Lines 16-18: Interesting that although the green fraction increased around the time of a short spell of increased temperature, the GPP did not change (bp4). How might this be explained?

Response: We think this increase in the green and red fractions and decrease in the blue signal is linked to the very fresh re-growth of grass shoots after the mowing event and the gradual build up of LAI and pigment content (see Fig 11). However, at the very beginning of this particular period when the green increased the temperature in fact seems to be decreasing and rises again after a few days. This may be the reason that the GPP increase slows down a little during this period but then increases again as temperatures become higher.

Overall what’s your advice for using these data for evergreen trees? Are the difficulties regarding detection of phenophases using piecewise regression a limitation? Do you think a slightly different detection protocol is needed for evergreen stands?

Response: As mentioned above our first advice would be to install cameras close enough to the evergreen canopy so that phenological events can be verified visually. However from our experience it seems very difficult to derive phenological events automatically from digital images. But if phenological events in the conventional sense (i.e. needle elongation) do not cause large variations in the green signal perhaps we should understand better the breakpoints or changes that are picked out, such as the snow-free periods, the dates when the ‘re-organisation’ of pigments occur with acclimation to leaf temperature and when the maximum green signal is observed in the season. Perhaps these seasonal landmarks are more important for understanding the inter-annual variability of evergreen CO2 sequestration?

Section 3.1.2

P18 Lines 4-6: The fact you can detect the impact of flowering and cutting is indeed interesting. What type of grass and flowers are at this site? Do you have any photos that you could add into the supplementary information (as for Figure S5)?

Response: The vegetation at this site (Neustift) has been classified as Pastinaco-Arrhenatheretum and is characterized of a few dominant grass (*Dactylis glomerata*, *Festuca pratensis*, *Phleum pratensis*, *Trisetum flavescens*) and forb (*Trifolium pratense* and *repens*, *Ranunculus acris*, *Taraxacum officinale*, *Carum carvi*) species. On day 131 one sees the yellow flowers of *R. acris* and *T. officinale* and on day 181 the white flowers are from *C. carvi*. A figure (S6) has now been added to the supplementary information showing the mowing and flowering events.

P18 Lines 9 and 10: “Even more challenging” is repeated. I am unclear as to why having 8 breakpoints made it more difficult to detect the start and end of the growing season. Please could you explain this in more detail?

Response: The main challenge with using 8 breakpoints when you are perhaps only interested in 2 of them (leaf growth and leaf death) is that you will need to check manually all 8 breakpoints to be sure you have the right date for the right event and this will take time. If there is only one site that has this issue then it is probably not a challenge. However, if there are many sites like this, then this can start to take up some time. Surely ecologists working on flowering events would like the choice to have more breakpoints and the software we developed can accommodate this type of scenario. We have added a clarification on this in the manuscript.

P19 Line 2: Migliavacca repeated.

Response: The text has been duly corrected.

P19 Lines 2-4: This is a very modestly described caveat of the method used, but surely it would be the same with any method used? If the grass is buried under litter than no sensor (or method) will detect the start of new leaf growth.

Response: We completely agree with the reviewer on this issue.

P19 Lines 6-25: This analysis is of course true but I suspect that it is not a lack of knowledge of different crop sowing and harvest dates that results in the crops being treated as grasses in models but rather the difficulties of prescribing these dates and different crop management strategies. I do not know but are there not country/EU-wide datasets that give the broad dates of sowing and harvest dates for different crop types? Nonetheless these data do indeed present a good demonstration of this issue.

Response: We fully agree with this comment and re-formulated the sentence to clarify this point.

Section 3.1.3

P20 Lines 4 and 5: For a second I wondered why the red and blue signals decreased, and then I remembered that these are relative signals. Perhaps others will not need this clarification but it may be even more informative to remind readers of that here?

Response: We thank the referee for this reflection and we have added a clarification on the relativity of the signals in the text.

P20 Line 12: It may be useful to give examples of the Mediterranean sites in the text as you have done for the continental sites.

Response: We are unsure what is required here. There is already a fair bit of discussion about the Mediterranean sites.

P20 Lines 15-17: What do you think is the cause of this variability if not a climatic driver?

Response: It is not entirely clear what variability is being referred to in this comment. If it is the length of the growing season (?) there could be multiple environmental factors at play including the light quality, temperature and moisture status of the site.

Another possibility is that the species composition of the forest may introduce additional variability as different species often have different phenological attributes and environmental thresholds. For example, just from looking at the data presented one might also conclude that ecosystems comprised of deciduous Oak species have longer growing seasons than those dominated by Beech, however these sites are also oceanic or Mediterranean, thus we would need some Oak continental sites to test this statement.

Section 3.2.1 This section presents very thorough and informative sensitivity analyses. Figure S6: The axes font is very small and may not be readable on paper

Response: We agree with the referee that this Figure has small font. We have re-orientated the figure to landscape for the final version and have improved the ticks for the Hotspot parameter that were too difficult to read.

P 23 Line 15 on: This is a nice extension of the first sensitivity analysis. Was the sensitivity analysis presented in Figure 11 conducted over the same time period as in Figure S6, even though the constraints are only defined for the spring green up? If so, it may be helpful to the reader to put the LAI, Chl and Car columns of the original sensitivity analysis in Figure S6 next to the 3 columns in Figure 11 to see the impact on the sensitivity brought by the extra constraints, but this is not needed. If not, then this should be detailed in the text and caption of Figure 11.

Response: We thank the referee for this comment and we have improved the text to prompt the reader to compare the LAI, Chl and Car panels in S6 to Figure 11. In addition we confirm that the sensitivity analyses for Figs S6 and 11 were performed for the same timeframe.

How general are these constraints? Could they be applied for all ecosystems (albeit with different ratios for different ecosystems)? Or do these constaints break down elsewhere? Have these constraints ever been applied to the PROSAIL model before?

Response: We believe these constraints are pretty general over most of the growing season. Feret et al., 2008 show in a number of different ecosystems that there are strong correlations between the concentrations of chlorophylls a/b and carotenoids (see Fig. 3 of Feret et al., 2008). Thus the general opinion is that these pigment characteristics co-vary in nature and they are not statistically independent.

However, it does seem that there are some events, particularly senescence, when a decoupling of the chlorophyll from some of the xanthophyll carotenoids may occur. For example Lu et al. (2001) observed in field-grown wheat plants a drop in [Chl] around 20 days after flowering that was also associated with a strong increase in the xanthophyll/chlorophyll concentration ratios. Their results suggest that the degradation of the chlorophyll pigment appears to be extremely rapid whilst the degradation of other leaf pigments can be played out over a few more weeks.

Another possible moment in the growing season could be during budburst or early in the season for very cold ecosystems such as in the boreal region. Some studies have indicated higher amounts of carotenoids in comparison to chlorophylls in aspen at the beginning of the vegetation season (Hillker et al., 2011, Fig. 5a) when foliage requires protection from photo-oxidative damage, as the photosynthetic apparatus is not completely mature (Lewandowska and Jarvis 1977). Further studies may now be warranted to verify these constraints during budburst and senescence for different plant types.

Lastly, we are not aware of these constraints being applied to the PROSAIL model before. Most previous studies have used PROSAIL in the inverse mode to solve for LAI or Chl content. Typically in these inversions each of the parameters is solved independently without any such constraints.

Section 3.2.2

P 24 Line 22: By adapted PROSAIL model I assume you mean with the fucntional constraints applied in the previous section?

Response: When we refer to the adapted PROSAIL model we are simply indicating that it has been adapted to make the link between RGB reflectance outputs and camera digital numbers. We have removed the word adapted to avoid any confusion in the text.

In Table 2 do you mean to refer to Figure 12 and not 13?

Response: We thank the referee for spotting this typo, it has now been changed to Fig 12.

Figure S7: Do you mean to refer to Figure 12 and not 13?

Response: We have double-checked the Figure S7 legend and it seems to be correct.

|  |  |
| --- | --- |
| P 26 Line 2: Indeed it would be good to re-do this analysis for years where pigment concentration data (or other parameters) are available for extra validation.Response: We completely agree with the referee and hope this study will encourage the community to measure some of these useful parameters. Section 3.3This seems to be a very rigorous assessment of the technical difficulties. P28 Line 20: Mizunuma et al. repeated.Response: This citation has been removed.Conclusions P 31 Line 14: spelling mistake for “archives” Response: Corrected.Further possible points for discussion:* I have already mentioned this and do not wish to repeat myself, but what are your overall suggestions of improvements to or usage of the piecewise regression method for detecting phenological events. It seems that that visual inspection may be needed for evergreen and managed ecosystems as you’ve described. Or would another metric like a threshold useful?

Response: The breakpoint approach was used as a consistent methodology allowing the extraction of phenophases from a variety of natural and managed ecosystems belonging to the network. We agree that some breakpoints raise questions and require visual validation, but at the same time this approach is successful in identifying the main phonological transitions in natural ecosystems and, more importantly, management practices in grasslands and croplands, where other methods such as curve fitting and threshold extraction would have failed.The objective of this paper was to apply a consistent methodology across an entire network of very different ecosystems. In other cases, e.g. for an ecosystem-specific analysis, we would not necessarily advocate a breakpoint approach as the best solution, but rather we suggest the use of fitting and extraction methods that are best suited to the data under analysis.  | **BGD**12, C3114–C3121, 2015Interactive CommentFull Screen / Esc Printer-friendly Version Interactive Discussion Discussion Paper |

- These data and the RT modeling could be used to see if you can see the same events (and validate) satellite data. Coming back to Section 3.1.3 (P21 Lines 17-21) following the modeling discussion. Perhaps the model could be used to derive NDVI which could be more readily compared with ground-based measurements as well as satllite data products?

Response: We believe this is a good point to raise in the conclusion and have added a brief paragraph on this.

- Of course this is a preliminary exploratory study; however, the thorough analysis could be complemented by discussing the impact of this type of RT (turbid medium) model and the assumptions used. For more complex canopies (vertical heterogeneity and/or a mix of under- and over-story) the PROSAIL model assumptions may break down, particularly as these are near canopy measurements and therefore local-scale effects such as clumping are probably not properly accounted for.

Response: We believe this suggestion could be a relevant and interesting study to make across the network in the future. However given the length and density of the manuscript it is currently beyond the scope of this study to go into too much further detail on these matters.

As well as using the modeling to quantify and monitor plant physiological status, how else might you use this framework? You mentioned improving the phenology models of dynamic vegetation models in the introduction what other applications can you see from these data? Are any of the time series long enough for trend analysis for example?

Response: We believe that this camera network along with the other networks in the USA and Asia when sustained over decades will provide a valuable archive of ecosystem responses to climatic change and variability. Presently many of the time-series in the European network are too short to conduct robust and meaningful trend analyses but of course this is the long-term goal of such networks. However, we also believe these networks and these types of analyses can provide us with information on the impact of stress events such as frost or insect outbreaks on canopy colour signals and their impact on the energy and carbon balance of ecosystems.

# Referee #2

The breakpoint analysis for transition dates presents an interesting alternative for ex- tracting transition dates, but is not strongly justified in the paper. As the authors admit, the maximum number of breakpoints must be specified, as well as the minimum segment size. This places a two parameter constraint on the fit, where as thresholding and other techniques only place one. In addition, though the first and final breakpoint locations correspond well to leaf out and senescence, the middle breakpoints don’t seem to correspond to phenological transitions. Figures 4-8,10 all note that the breakpoint changes identify “important transitions”, though it is unclear from the data presented that these transitions are actually important for the canopy or ecosystem.

Response: We agree that the breakpoint analysis requires some arbitrary (but sensible) constraints on the time-series, regarding the maximum number of breakpoints or the minimum segment size. We are convinced that the method could be improved and adapted for each site. However, the application of a general parameterization of the method as in this study provided good results in terms of detection of transition days. Actually the method presented here is already an improvement of a method we had tried before using the bcp package that required even more arbitrary (i.e. not physically identifiable) information and seemed to require site-by-site adjustments.

We do not completely understand why the referee considers that using simple thresholds would be superior or less parametric, because we believe that we would need to specify a minimum segment size for a threshold to be detected or require some sort of rescaling of the timeseries using extrema (or percentiles) and then arbitrary thresholds. Fitting the data with parametric functions and then thresholds on the derivatives of the fit would also require several a priori assumptions. Furthermore, fitting methods or threshold approaches would not have worked on managed grasslands, thus for this study breakpoint analysis that would provide a common extraction method with fairly good results across all ecosystems. As said above we do not claim that our method is the best for detecting transition dates in the time-series of color indices, but it is for sure the best we could find so far for an application on a large network and we are very open about using another method if it can be proved to be more appropriate.

We agree that transitions in the green fraction identified by our method are not always obviously linked to changes in canopy structure or physiology. The identified breakpoints are thus not all “important” in this sense, but nonetheless real. We thus removed the qualification “important” in the legends of figures 4-8 and 10 (i.e. “Vertical dashed lines indicate breakpoints corresponding to transitions in the green fraction”).

The RGB signal modeling of section 3.2 is overshadowed by network-wide analysis of section 3.1. It would be nice to understand more of how the work in section 3.2 was performed, including a full description of the algorithm, parameter values and un-certainties, parameter starting ranges that link PROSAIL results and camera sensor properties to output color fraction curves (Fig 12, panel 3). The results shown in Figure 12 are impressive, and this section of the paper is likely to be of greatest interest to readers, but readers are left without the tools necessary to reproduce or extend the results.

Response: We thank the referee for highlighting that figure 12 and the modelling section that accompanies it is likely to be of great interest to readers. We also understand that a better description of the tools used in the section could be given. In order to address this and also encourage the testing of the model at other sites we created a repository containing the documented code and data needed to generate Fig. 12 (and also Figs. 11, 13, 14, S1, S2, S3, S7 and S8) on a Bitbucket account (https://bitbucket.org/jerome\_ogee/webcam\_network\_paper). Should the paper be accepted for publication in Biogeosciences, this code will be open to public access. Our hope is that by placing it in a git repository the research community will actively contribute to the improvement of the code and tools to assess its sensitivity. We will place the link information for this git repository in the methods section 2.3 of the paper and again in the legend of Fig. 12.

Fig 12, panels 1-2 need to include standard deviation envelopes around the curves for Chl, Car, C\_brown, and N.

Response: As pointed out by the referee in his/her appraisal of the present manuscript, one of our aims was to provide a potential mechanistic framework that tries to link the seasonal RGB fraction datasets collected by the network of cameras to plant function. For the seasonal time-series presented in Fig. 12 we ran the model in the forward mode and prescribed how each of the parameters Chl, Car, C\_brown, and N were hypothesised to vary seasonally. As these are virtual estimates from the literature we did not feel at this point a fully blown uncertainty analysis was required. However, we do provide a repository for the code and the seasonally prescribed parameter set. Another study using real pigment measurements and their uncertainty with PROSAIL is also underway.

When discussing the sensitivity analysis of the RGB signal modeling, the note about the impact of diffuse light and leaf inclination angle could use further detail and discussion. Why is it that these two parameters, along with at least 4 others have an impact on the green signal, but not blue or red?

Response: This comment seems to be a little inconsistent with the text in the manuscript. We wrote that the green fraction was affected by at least 5 model parameters (Chl, Car, Cbrown, LAI and N) and that diffuse light and leaf inclination seemed to affect the blue and red fractions but *not* the green fraction (see last sentence at the end of first paragraph in section 3.2.1). We referred to Fig. S6 (now S7) to justify this statement. In order to best address the referee’s comment we did however elaborate more on the effect of diffuse light and leaf inclination angle on the RGB signals.

In the case of diffuse light the difference in the RGB response is a direct consequence of the spectra used in our analysis for direct and diffuse light (Fig. S3). From the spectra shown in this figure we can deduce that more diffuse light will bring more blue and less red, thereby influencing directly the red and blue fractions and less so the green fraction. We felt it important to highlight this influence of diffuse light on colour fractions given the large changes in sky conditions that are often experienced in the field between one day and the next. The lack of sensitivity of the green fraction to diffuse light is probably one of the best reasons for using the green fraction for phenology studies. We have added a little bit of text to clarify this point in section 3.3.

In the case of leaf inclination angle our statement was based on results shown in Fig. S6 (now S7) and in response to previous studies that had suggested leaf inclination may have a role to play on the green signal. From Fig. S6 it looked like this parameter was clearly not the most sensitive one probably because of the view angle of the camera (nearly horizontal). The idea that more sensitivity could be found at other view angles was added to the text although more tests would be needed to verify this possibility.

# Referee #3

My main concern is that this study includes two themes which are not joined together well and in fact each of these two could be the basis for a separate manuscript…. Moreover, the wealth of data existing within this network could allow a broad synthesis that could result in novel and more general knowledge of phenology patterns emerging across the different biomes.

Response: The referee points out that this paper presents two important and related themes. Originally back in 2011 we wrote a previous manuscript presenting the network and focusing on the phenology patterns and their automatic detection, with the idea of having a separate paper for the PROSAIL modelling. However, reviewers requested that the interpretation of the datasets with our model was necessary.

We took this onboard especially because the network is immature and has only been running for at most several years and not decades. Thus we believe it would be premature to conclude any broad or large-scale results on phenological patterns or inter-annual variability across the network. For example, Arguez & Vose, 2011 (Bull. Am. Met. Soc.) argued that at least 30 years of met data would be necessary to construct a WMO climate normal. We would imagine this would also be applicable to phenology given the strong link to climate and currently this data does not exist.

Thus as the referee points out besides presenting the network the main focus of our paper is about developing the tools that will be necessary to analyse these growing datasets. We believe there is a natural progression and link in the manuscript between the results of breakpoint analysis and the need to understand better where possible what mechanism(s) underlie certain breakpoints and seasonal trends. We believe the addition of the PROSAIL modelling explains these breakpoints and patterns in deciduous broadleaf forests and grasslands well.

We believe the comments incorporated from the three referees comments have really improved the manuscript and we have made an effort to make the link stronger in the objectives, the transitions text and in the discussion.

It is not clear why the maximum amount of breakpoints was set to 5 for both managed and unmanaged ecosystems (Page 7988, line 2-4). In my opinion it would be more logic to allow more breakpoints in managed systems than in natural ecosystems. The choice of the number 5 is also not well justified. Moreover, while breakpoints 1 and 5 are relevant to quantify the start and end of the vegetation period, it is often not clear to which event the additional breakpoints in between relate to (see Fig 5 and 6), specifically in the natural systems.

This is crucial, see comment above ref 1

Response: This was also raised by Referee 1 and we completely agree that in managed ecosystems or in other applications such as studies of flowering phenology running the algorithm to obtain more breakpoints would be important. For example some sites like Fruebuel in good years can have up to 7 mowing events and thus at least 8 possible changes would need to be identified.

Page 7983, line 16 and at other places: avoid the term ‘dramatic’

Response: We have substituted the use of ‘dramatic’ to ‘large’

Page 7985, line 7-8. It might be merely a matter of wording, however, the current for- mulation of the first objective ’i) how well can digital images be automatically processed to reveal the key phenological events. . .’ is weak. It is known that images can be easily processed automatically with standard software routines. I believe the actual relevant question in this objective could be ‘how well do color indices derived from digital image analysis describe phenological patterns. . .’

Response: We completely agree with the referee that objective 1 could be better formulated. Following the referees suggestion we have amended objective 1 to read.

‘how well do colour indices derived from digital image analysis describe key phenological events such as….’

Page 7985, line 26: Having only 1 image per day is not enough to derive robust color indices since the effects of illumination for this specific image might introduce consid- erable noise. Moreover, the method by Sonnentag et al 2012 used in in this study (method section Page 7987, line 14-15) was developed for image archives with more than 1 image per day.

Response: We agree to some extent with the referee about this comment. It is definitely true that several images per day are preferred for the reconstruction of seasonal variability over only one image. This is mainly to overcome noise that can be introduced in colour signals through changes in sky conditions. However, of the digital image archives available to date from this network, not all sites have collected more than one image per day in the past years. Thus we are currently addressing this problem at the moment and a protocol has been written recently (as part of the ICOS infrastructure) to ensure each site in the network will collect several images during daylight hours in the future. You will also see from our analysis that at sites where multiple images per day were available we incorporated them.

Nonetheless, the sites where only one image per day are available for the seasonal analysis still appear to provide strong indications of changes in the different colour signals over the growing season. In addition the seasonal patterns in the deciduous sites where only one photo per day is available still present seasonal features consistent with those sites where many images are available. One example is reported in Fig 2a, where the differences between the original (raw) and reconstructed green fraction (applying the Sonnentag filter) show a good agreement, also during the spring and autumn transition phases, where the application of the method to archives with 1 image per day could be more problematic. To demonstrate this we performed an analysis using an example dataset with 8+ images per day. We then randomly removed daily images to have only 1, 2, 3 or 4 images per day and then applied our filtering procedure. The random removal was boot-strapped 100 times (to sample different images for a single day) and then a min-max envelope was computed as shown in the plot below for the Italian site Torgnon. As expected the more images per day reduce the thickness of the envelope. However, the overall shape of the seasonal curve is exactly the same, in terms of phenological events, and also in terms of the minimum and maximum green fraction values. Thus although the filtering algorithm of Sonnentag et al., 2012 was developed for datasets with several images per day, because it is based on a 3-day moving window it is robust enough to be used on datasets with less than one image per day. Thus we believe that even a dataset of one image per day can be analysed robustly with Phenopix.



In addition, it is also worth noting that sites with only one photo per day seem to provide reasonable agreement with the start and end of the CO2 uptake period (Figs. 5, 7, 8 and 12) and visual observations (Fig. 3). So we believe that it is still possible to identify the principal changes in canopy features over the season in deciduous and grassland/cropland sites even when very few photos are taken each day. We would consider data gaps of several days or changes in the camera ROI far more problematic.

We also discuss in the manuscript that the green fraction variability is relatively small within and between consequent measurement days in comparison to those of the red and blue signals over the growing season, as it is the green signal is least affected by diffuse sky conditions. We would also add that assimilating data on the fraction of diffuse radiation within and between days with PROSAIL can allow one to simulate the typical daily and seasonal variability in RGB signals. Thus in the future even these limited datasets may still provide useful information for time-series trend analyses.

Page 7986, line 2. Define ‘LT’. Using images between 11am and 1pm would provide 3 images, assuming hourly resolution. This is a limited number of images and their daily RGB means are likely sensitive to illumination changes. This uncertainty should be addressed.

Response: Local time has been defined in the manuscript. We believe this comment has been addressed in the response to the previous comment.

Page 7986, line 4-5: Describe in more detail what the ‘fixed’ and ‘manual’ white balance settings are. Usually the fixed ‘daylight’ setting is recommended since it results in a color temperature of around 5200K. At this setting, the RGB digital numbers are the most neutral across all wavelengths. For lower and higher color temperature settings, especially the red and blue digital numbers deviate substantially for shorter and longer wavelengths. Consequently, this would hamper the comparison of the red and blue fraction among cameras with different white balance settings. This is also an important consideration with regards to the discussion on Page 8004, line 5-26.

Response: The most important point about our statement in the Material and Methods is that to track seasonal changes in colour fractions the ‘automatic’ setting must be off, otherwise the camera will default to a ‘grey world’ algorithm, meaning the entire picture will be averaged and thus = grey. Red total = Green total = Blue total. The result of this problem is demonstrated nicely in the manuscript of Mizunuma et al., (2013).

We also believe that in order to track seasonal changes in phenology, inter-camera differences do not seem to be a critical issue as demonstrated by Sonnentag et al., 2012. However, we agree that a network containing the exact same camera model would make future site inter-comparisons easier.

We also agree with the referee that differences in how colour balance settings vary between cameras must be known or characterised if the link between colour signals and canopy pigment content are to be made. This is why so far we have only attempted the PROSAIL modelling for two camera models where we have the required information available to complete this step (Figs S1 and S2).

Page 7986, line 11. Why should soil not be included in the ROI? This could provide valuable information on the fraction on ground covered by plants (e.g. in croplands).

Response: We agree with the referee that there is no reason why soil should not be included in the ROI especially if the objective of the study is to look at the fraction of ground covered by a canopy. However, this would require that the camera is mounted looking downward in a manner similar to those installed in the Japanese PEN network and used to generate Figs S8 (former S7) and S2.

However, the majority of cameras in the European network are not suited to measure soil and are specifically mounted on towers to contain an ROI that can be easily related to the net ecosystem exchange also measured at the site and plant phenology.

Page 7986, line 12-17. The term ‘Automated segmentation methods’ is not fully clear in this context. Does segmentation refer to defining a region of interest? Moreover, since this method was not used after all, I think there is no need to include a paragraph on it.

Response: On reflection we agree with this comment and we will remove this statement from the text.

Page 7987, line 14. What is the amount of images (in %) that has been removed as outliers?

Response: For the Migliavacca filtering there is a very small removal (<1%) for most sites. However, for some sites this can increase demonstrating the effectiveness of the filtering procedure when necessary.

Page 7989, line 10-17. Since this study is about the phenology of European ecosystems, it is not clear why this analysis was done also for the Nikon camera since this model is used only at two sites within the European network.

Response: We agree this manuscript is focused on European ecosystems. However, although the wider implications of the present manuscript can be linked to phenology it more specifically aims to develop tools that can help us interpret RGB signals from digital cameras and make links to canopy development and physiology.

In particular we explored the novel use of PROSAIL to simulate the signals obtained from digital cameras. In order to test this approach we felt that it was necessary to test the model at a site where more than one model of camera was installed. Over the years we have been very fortunate to have a successful collaboration with researchers from the PEN network and have both a Stardot and a PEN system (with a Nikon camera) installed and running at the same site. These systems are quite different in the way they are set-up, one looking across the canopy and another looking down (Mizunuma et al., 2013) as well as having slightly different RGB sensor characteristics (Figs S1 & S2). Thus they present a challenging test of the PROSAIL parameterisation and approach. The test shows that with information on the camera angle and sensor characteristics it is possible to model the RGB signals of different camera models with our seasonal inputs of pigment, LAIproxy and radiation. We hope it is now clearer the reasons behind our local choice but we also feel that by doing this we have demonstrated the modelling approach can work on the two camera systems that dominate the ‘global’ network of cameras at flux sites, making the result internationally important.

Page 7993, line 4-5. Bp 2 and bp 4 occur on day 110 and 310, respectively, and cannot be assigned to a range of 10-20 days. Furthermore, there is actually no clear change visible in the gcc pattern shown in Fig 5 around the bp2 (day 110) and no clear change in GPP at bp4 (day 310). In my opinion, the timing of these breakpoints and their importance for linking GPP and gcc patterns has been over-interpreted in this specific analysis.

Response: We thank the referee for pointing out this oversight and have corrected the day numbers accordingly.

We also agree that the link between bp4 and the decrease in GPP is not as clear as in spring and have tamed down our interpretation in the text.

Page 7993, line 19-21: The sites selected from the network are mostly located within central Europe, thus it is not surprising that the differences are limited. The example of the alpine site Torgnon however indicates that including sites with more contrasting climate (i.e. maritime and Nordic sites) in the analysis would likely result in much greater differences among the patterns. Moreover, the breakpoint analysis does not capture well the onset of the greening up at the Klingenberg, Grillenburg and Neustift sites.

Response: We agree that the recent addition of Nordic sites to the network will likely provide some very nice contrasts in the RGB signal when compared with the continental and Mediterranean sites in the future.

We agree that the analysis for Klingenberg and Grillenburg is not great and is likely affected by the lack of temporal resolution (weekly) in the digital images.

Page 7994, line 17-22. I don’t see an issue with allowing breakpoint 1 to represent snowmelt and only breakpoint 2 describing leaf out, as long as this pattern is realistic for the specific ecosystem.

Response: We completely agree with the referee on this point and depending on the study having more breakpoints could be an advantage.

Page 7995, line 22. This statement is not well supported since the current study does not show any field data on how well the greenness color and leaf area correlate.

Response: We have modified this sentence.

Page 7999, line 20-26 and Page 8000, line 20. Based on Fig 11, the slope of the Chl concentration rise is greater than that of Car concentrations at any time during the green-up phase, especially right around the time of the ‘greenness hump’. It is therefore not clear why at some point the synchronous increase in Chl and Car should switch from an increase to a decline of the greenness fraction. Moreover, the model outputs are currently not validated with measured concentrations of Chl and Car. This is a limitation to take into account when interpreting the model outputs.

Response: The increasing pigment content of the leaves causes the decrease in the green fraction. Chlorophyll concentrations above 30 ug cm-2 and Car concentrations above 7 ug cm-2 cause a decline in the green fraction, this is demonstrated with the sensitivity analysis in Fig. 11. We also agree that measurements at the site would be great to confirm this model response, this is something we are currently working on. Preliminary results however indicate that the seasonal trends and values in pigment content we report for Oak are consistent with the present parameterisation and those found in a number of other published studies providing confidence in our interpretation.

Page 8001, line 6 and Page 7996, line 6. What mechanism is changing the blue fraction? Is it possible that the blue fraction merely changes passively due to changes in the green and red signals? In that case the importance of humps and other patterns in the blue fraction would be limited.

Response: The blue fraction is responding directly to both changes in chlorophyll and carotenoid content thus we would not classify this as a passive response. For example in the Fig. 8 of Feret et al., the absorption for chlorophyll and carotenoids is not the same in all wavelengths and can account for most of the colour signal trends. However, in addition the spectral efficiency of the RGB colour sensors of the camera (Figs S1 & S2) is also sensitive to wavelength and thus also feeds into the RGB camera signal trends. Thus, to summarise the pigment content, the spectral efficiency of the camera sensors, the conversion to colour fractions and changes in sky conditions (particularly for Blue and Red) all contribute partially to the RGB signals observed by the camera. To understand these different components we have created a repository containing the documented code and data needed to generate Fig. 12 (and also Figs. 11, 13, 14, S1, S2, S3, S7 and S8) on a Bitbucket account (https://bitbucket.org/jerome\_ogee/webcam\_network\_paper). Should the paper be accepted for publication in Biogeosciences, this code will be open to public access. We will place the link information for this git repository in the methods section 2.3 of the paper and again in the legend of Fig. 12.

Page 8002, line 25ff: I appreciate the discussion on current challenges; however, I suggest presenting only those solutions that offer a realistic option. For instance, sug- gesting the use of a color checker to quantify drifts is meaningless if within the same section it is acknowledged that the color checker itself might drift over time.

Response: A colour checker could be used on a very limited (hrs) but regular basis (a few times a year), thus we do not rule out this method as a possibility, on the other hand we think it is important to point out that it would not be our recommendation to leave a colour checker in the field for long periods of time for calibration purposes.

Language: Page 7988, line 16; Page 7992, line 7-9; Page 7995, line 10-14; and at several other places: Avoid subjective terms like ‘few’, steep’ , ‘slow’, ‘gentle’ , ‘rapid’ or ‘fairly similar’, ‘slightly shorter’ etc and instead provide some quantitative information such as numbers and dates, e.g. ‘within 5 days’ , ‘From April 1 to 5’, etc.

Response: We have searched the manuscript for ‘subjective’ terms and where appropriate and when it does not break the flow of the text we have made changes.

Page 7995, line 1, Page 8004, line 27 to pg 8005, line 57 and at other places: the discussion is based on initial and ‘preliminary’ results at too many places. Preliminary result may be shown but in a limited quantity within a scientific publication. However, I suspect that this might be a language issue and that the authors actually refer to robust findings and analyses in these cases which however require further research. I suggest to avoid the term ‘preliminary’ where possible and/or to exclude results that provide no solid evidence.

Response: We agree with the referee on this issue and have corrected the language in 4 places to convey that the results are robust and further research is now required.

Avoid weak phrases such as ‘we tried to’, it seems that’, ‘it appears that’ or ‘tended to’ etc which occur frequently throughout the manuscript. It leaves the reader wondering about the robustness of the results and implications of such weak statements.

Response: We agree with the referee on this issue also and have corrected the language in 4 places to make stronger statements.

Table 1 & Figure 1: It seems odd that peatland sites are presented here while no dedicated section was included describing phenology patterns for this ecosystem type in the first section of the manuscript. I suggest including also a section on peatlands if the goal is to present a network synthesis (theme 1) or to otherwise remove these (and other) sites not used in the current analysis from the Table and Figure.

Response: Peatland sites are an extremely recent and exciting addition to the network. However, incorporating further analysis on these new sites is currently beyond the scope of the present manuscript. Nonetheless, these sites will undoubtedly be analysed in future studies and thus we believe it is important to state in this paper that these sites are now contributing to the network or will be very shortly in 2015.