**Comments from the editors and reviewers:**  
-**Reviewer 1**  
  
  - The paper presents a method for reflectance retrieval from uncalibrated photogrammetric imagery by normalisation with MODIS data. Overall, this method is sound and straight forward and it is nice to see that it could be applied successfully in a quite convincing test case. There are only minor things which could be improved in the paper but some sort of error budget for the overall process would be nice to have.

Detailed comments:

Abstract: 'the mean absolute difference was 4.18%, which compares well..." I doubt that this figure is accurate to 3 digits; 4.2 or even 4% would be more appropriate as a resulting figure. As DMC is uncalibrated, there's hardly another method which could be applied, so one should omit the second part of the sentence.

Page 2: mentioning the correct terminology, one should name the output according to the MODIS reflectance product output, as it has been normalized to this product. Please mention and use the corresponding terminology (they call it a reflectance corrected for BRDF effects) - one should not name it a bidirectional reflectance, as this implies a BRDF value (with a maximum at 1/pi). One could name it a bidirectional reflectance factor, as it has its maximum at 1 and it has been normalized to a relative stanard (not a spectralon, in this case, however).

Page 5, Eq 3: rho\_t in the denominator is only the same as rho\_t in the nominator for the most simple case assuming an uniform ground; normally it should be the adjacency reflectance, differing from the pixel reflectance.

'typically around 0.07': The typical atmospheric albedo varies with wavelength - this should be mentioned.

Eq 4: by this assumption, adjacency effects are neglected in the correction; please mention. This will have a strong impact on dark object signatures.

page 7: step3: the resampling step should be more detailed described as you're going back from 500m to 2m: is it a linear interpolation, some triangulation, how much smoothing has been applied?

page 8: the linear relation between various sensors and wavelengths does not imply that using one spectral band to correct another is a correct approach. One will automatically neglect the inherent differences in absolute reflectance between the two; absolute values may differ even if the relation is linear and correlation is high. This needs to be reformulated.

page 11: "C was assumed to be small enough": what does 'small enough' mean? did you check the blue spectral band, as C may be considerable larger there?

page 17: the validation is mainly done on the relative R-square value. Indeed, the validation should be against a 1:1 linear relation, as absolute reflectances should agree to each other. reflectance offsets are neglected by this analysis.

page 18: it is not clear enough if the spot imagery (and MODIS) has been corrected for terrain influences using ATCOR or whatever other method, and how much impact this could have on the results

Figure 13: same again: validation shoud be against 1/1 line and all the reflectance graphs should be at the same scale (up to reflectance=1) on all axis to make them easier to read.

Validation/Conclusions: It would be interesting to see how the method performs over critical targets such as lakes, shadows or very bright objects. Some sample spectra after correction in comparison to SPOT would be interesting to see and analyze. As all validation is done statistically only so far, the performance on average spectra is well analyzed but more critical cases are not outlined.

-**Reviewer 2**  
  
  -

The paper makes a good contribution by proposing a method for relative radiometric normalization and homogenization of radiometric changes between aerial images to succeed in mosaicing processes by freely available satellite data. The methodology serve as an improved alternative to other smoothing techniques such as dodging or LUTs. However, as the methodology is raised (different data acquisition days, image resampling for calibrating DMC with MODIS, significant disparities between RSRs of sensors, etc.) rigorous quantitative remote sensing evaluations by analysing the reflectance of land covers can not be performed.

The paper has to highlight and discuss some limitations, including:

-          The influence of performing the aerial campaign during 4 different days and at different times of the day

-          The consequences of performing the radiometric calibration with a satellite sensor with different spectral range and bandwidths regarding the DMC camera.

-          How does the fact that calibrated reference images and aerial DCM images were not taken on the same days affect the accuracy of results. On the other hand, how does this fact affect in the calibration validation with SPOT5 images? (SPOT5 images were taken on 21 January, in contrast to DMC images whose data acquisition started on 22 January)

-          Discuss the partial validation of the experiment due to both facts: SPOT5 has not blue band and SPOT5 images do not completely cover the DMC mosaiced area.

For all that:

I propose to make some changes in the title so that it is more consistent with the experiment and results, for example by adding "an approach to the ..." at the beginning. Furthermore, some improvements must be made in the paper especially regarding the specification of important data that should be included and detailed in the accuracy assessment section.

On the basis of the above considerations I recommend a **major revision**of the paper. Below you can find some specific comments (I attached the paper with numbered lines to help in the review process).

Comments:

\*\*\*\*\*\*\*\*\*\*

**Abstract**

**Line 12**: a simple yet effective technique (without comma after simple).

**Line 20**: The quantity of images does not have much significance. On the contrary, provide info about the percentage of overlap as well as the problems encountered (arising from the limitations of the flight itself: angles, verticality, etc.) is necessary. Maybe it is not required to specify them here at the abstract but it is necessary in 2.2 or 2.3 section.

**Lines 21-24:** A discussion on the fact that the SPOT 5 has no blue band (limitations on the validation process) should be made. Furthermore, the mean value of the coefficient of determination (R2) achieved should be specified at the abstract.

**Lines 24-27**: Redrafting this portion. (One option: The technique allow producing seamless mosaics corrected for atmospheric and coarse-scale bidirectional reflectance distribution….).

**Lines 27-28**: Values that support and justify this affirmation are required.

**Introduction**

**In general,**there are repetitive references along this section. In addition, some of them are not current references. Try to update them and add some new current references as for example the following one (it can be mentioned at line 133-134):

*\*Del Pozo, S., Rodríguez-Gonzálvez, P., Hernández-López, D., & Felipe-García, B. (2014). Vicarious radiometric calibration of a multispectral camera on board an unmanned aerial system. Remote Sensing, 6(3), 1918-1937.*

**Lines 58-63**: Some of the limitations of the paper mentioned in the first part of the review became evident with this sentence.

**Lines 69-72**: To long sentence, rewrite and/or split it.

**Lines 88-89:** Put references in chronological order.

**Lines 88-89:** Put references in chronological order.

**Lines 89-92**: Add a reference that justify this sentence (“techniques that…source of error”).

**Line 101:** Put references in chronological order.

**Line 127:** Put references in chronological order.

**Line 133**: I do not entirely agree with that. The vicarious calibration can be an inexpensive experiment (as it is proposed in\*) by using low-cost calibrated surfaces or natural covers with homogeneous radiometric behaviour (grass, ground, etc.).

**Lines 133-134:** Put references in chronological order.

**Lines 142-143:** “…avoids the need to perform time consuming…” Ok, but there is no reference along the paper about the percentage of time saving or about in which sense there is a time saving by following the proposed methodology.

**Line 146:** More info is required (here or in 2.2 section). Percentage of overlapping, flying height, ground sample distance, etc.

**Methods**

**In general,**there are many assumptions and simplifications without scientific support in the research study.

**Line 158:** Assumption nº 1: “…can be modelled as a spatially varying linear relationship…”. It is required a literature reference that check it or evidences this affirmation.

**Line 168:**Define DN (Digital Numbers).

**Line 192:**Put references in chronological order.

**Line 200**: The accuracy of the calibrated satellite image must be specified (it is only specify at conclusions section).

**Lines 201-203**: Some of the limitations of the paper mentioned in the first part of the review became evident with this sentence.

**Lines 207-209:** Assumption nº 2: It is required a reference to other paper that check it or evidences this affirmation.

**Lines 209-212:** Assumption nº 3: In this case, it is referenced.

**Lines 224-226:** Assumption nº 4: the camera offset is zero. Why? There are no evidences of that. Justified this assumption.

**Lines 245-246:** Assumption nº 5: It is assumed that the spectral responses of the MODIS bands and the DMC camera are identical. How great are the errors arising from accepting this assumption?

**Line 276:** Assumption nº 6: The effect of differing PSFs is negligible. A literature reference where make the same assumption or a justification of that is required.

**Lines 295-299:** If a previous DMC image correction has been made (with Intergraph Z/I Post-Processing Software) there should not be problems with the radiometry of boundary pixels. In this way they should not been discarded.

**Lines 300-304**: What you really mean with “90 percent coverage”? Is it the footprint of the MODIS images about the mosaic aerial image?

**Line 309:** How much overlap you consider sufficient? Specify and justify it.

**Line 318:** There is an error when Figure 1 is mentioned.

**Figure 2:** Specify the MODIS bands inside the graph by, for example, writing “b1”, “b2”, “b3” and “b4” in the appropriate spectral response.

**Line 382:** How does the fact that calibrated reference images and aerial DCM images were not taken on the same days affect the accuracy of results.

**Line 382:** The site was selected as…

**Line 384:** Specify what kind of corrections were performed, their magnitude, its importance and influence in the next steps of the study.

**Line 386:**How does the fact that the spatial resolution of the DEM (5 m) and NGI imagery (0.5 m) are different affect?

**Line 392:**Bands 4, 1, 3 and 2 from MODIS sensor have not a 500m resolution. Correct it because bands 1 and 2 have 250m resolution. By the way, did you make a downsampling process with these two 250m resolution images to have the four bands with the same spatial resolution (500 m)? Explain it.

**Line 405:**Assumption nº 7: ***C*** was ignored.

**Line 406:**The derived calibration coefficients (gain in this case) of the DMC camera should be specified for each DMC band in a table or written throughout the paper body, here or in the results section.

**Lines 433-435:**Discuss the partial validation of the experiment because SPOT 5 images do not completely cover the DMC mosaiced area.

**Lines 436-437:**The accuracy of the SPOT 5 calibration is required.

**Line 438:**Discuss the partial validation of the experiment because SPOT 5 has not blue band.

**Results and discussion**

**Figure 4:**The title of the X/Y axes should be extended. What kind of magnitude is compared between DMC and MODIS?

**Table 1:**Why did you specify both the RMS and SDD?

**Table 2**: Why RMS and SDD are so different? It is required to discuss the values of the R2, they are quite low.

**Conclusions**

Somehow, reflect the observed limitations that are evidenced through this review.

-**Reviewer 3**  
  
  -

The paper describes a process of mosaic homogenization on the basis of concurrent satellite imagery. The title leads to think that a vicarious calibration was carried out to convert the aerial digital numbers to reflectance values. However, the methodological approach is the proper for image mosaicking. Moreover, the paper has two mayor issues: the first regarding the high number of hypothesis considered in the different steps, which were not quantitatively supported:

\* Assumption of the same BRDF and spectral signature of a surface cover in a 500 x 500 m region (25 hectares).

\* Claiming that the viewing geometry, terrain effects and their consequent BRDF effects are negligibles.

\* Generation of reflectance values for 0.5 m/pixel on the basis of a single coefficient fitting (arithmetic mean) of 500 m/pixel reference values. This assumption of land cover homogeneity implies an error propagation in the final reflectance values that invalidate their use for radiometric analysis at this very high spatial resolution.

\* Assumption that the atmospheric effects are negligible for the different sensors, the flight /orbital heights, and the different days.

\* The removal of radiometric coefficients under the assumption of null camera offset (without any empirical study).

\* The claim that all radiometric effects could be corrected by one resampled pixel in the aerial image.

\* Evaluation of the least square fitting by the analysis of the fitted points, instead of check points not involved in the computation.

\* As well as, confusing conclusions, such as the wavelength discrepancy of VHR and MODIS sensors are negligible in the calibration computation, but a similar one between the VHR and SPOT in the final comparison is the reason of higher final differences.

And secondly, the lack of quantitative results and technical information about the experiment:

\* The whole experiment is summed up with only one number, while no other internal analysis were carried out in the different steps (least square a-posteriori deviation, residuals, …)

\* The authors employed the MODIS data as a ground truth (zero error), but, as they stated in the last paragraph of the manuscript, its accuracy is about 5%. This a-priori error must be taken into account for a proper analysis.

\* The radiometric calibration is summed up to the least square fitting of only one parameter, namely, an arithmetic mean.

\* The aerial campaign for images data acquisition is confusing. The authors only mention that more than 2000 images were acquired, but no information about how many images were used, how many redundant points were involved in the calculations, or why the 20-days gap among acquisitions.

\* The geo-referencing error was only taken into account to support the discrepancy with the SPOT imagery, but not in the calibration phase (assumed as perfect).

\* The efficacy tests of the interpolation algorithms were not shown.

This work appears to be the continuation of a initial work of 2014 (“Very high resolution mapping of Spekboom canopy cover”) where the same sensors (0.5 m VHR, 500 m MODIS and SPOT image) were used to the radiometric correction of aerial photographs, but obtaining different results: “The mean absolute reflectance difference of the calibrated imagery was found to be 3.92% compared to an atmospherically corrected SPOT reference”. It would have been interesting to add this information to the discussion of the work, to highlight the differences.

For all these reasons, and the following comments, I have to propose the rejection of the manuscript in its actual form.

**Comments**:

 Abstract

L1: Expand the acronym (as well as in section 1)

L4: “A near-concurrent MODIS MCD43A4 image” -> “A near-concurrent MODIS nadir BRDF-Adjusted Reflectance image (MCD43A4)…”

L13: “near-concurrent SPOT 5 reflectance image”: specify that there is neither a spatial nor neither temporal concurrency. This last one (temporal) is a relevant issue that affects the results interpretation.

 “BRDF”: The authors claim that the BRDF effects were corrected; however, the constraint of nadiral image invalids this claim. So, indicate this limitation to only nadiral BRDF correction.

 “known reflectance targets”: Since the authors did not test the use of artificial targets or pseudoinvariant features, they should not claim that their method improve the mosaic results, especially when so many assumptions and hypothesis were required.

 “relative simplicity and computational efficiency”: These features were not tested and evaluated in the manuscript. Since no figures about processing time were shown, this claim should be removed.

Keywords: Remove “atmospheric correction” since it was not explicitly evaluated. Replace “calibration” for “radiometric calibration” (no geometry calibration was carried out), and “BRDF correction” for “BRDF” (since a complete BRDF characterization was not carried out).

Section 1:

“but extraction of surface reflectance from aerial imagery remains a challenge” -> please, elaborate.

 “spatial […] variations in aerial imagery limit the extent over which quantitative remote sensing techniques can be successfully applied” -> Authors said in the abstract that “VHR aerial imagery holds great potential for quantitative remote sensing”. These both phrases are contradictory. Clarify them.

 “viewed at nadir” -> This is a limitation of the proposed methodology. The issue about the drones oblique imagery to obtain reflectance values are never addressed. This is an interesting topic for the readers, due to the rapid growth of these platforms in relation to the VHR platforms.

 “Techniques that account for land cover specific BRDF’s […] introduces another potential source of error” -> How is possible that an approximate BRDF correction will introduce more error in reflectance value than not applying any?

 “these overlapping portions of different images should be identical” -> it is not clear why this constraint is so relevant. Please, clarify.

 “reflectance is beneficial as this is an invariant property” ->  It is not invariant, since the cover could change in time.

 “course-resolution” -> coarse-resolution

 “Unlike existing methods [….] of the same area” -> These phrases about the experimental setup are redundant with the abstract (and they should not be in the introduction) and should be removed.

Section 2:

“linear relationship between surface reflectance and sensor measurement” -> This is an assumption that should be verified accordingly to the sensor employed. Please specify it, since other sensors could have a different behavior.

 “BRDF effects and so also varies with the viewing geometry” -> But, since the proposed method is only capable to corrected the BRDF effect for the nadir, how its error was propagated to the final computed reflectances.

 “obtained from a well calibrated satellite image” -> In spite of being a well calibrated image it has an error that has to be taken into account in the methodology. This information was obviated until the penultimate paragraph (error about 5%). The 5% error is as significant error in relation to the claimed results of 4.18%.

 “should have been captured at a similar time” -> Define how much time. Hours, days? How it will affect the results?

 MIRS -> define.

 “500m” -> add an space between the number and the unit.

 “it is assumed that atmospheric and BRDF effects vary little at a small spatial scale […]Real BRDF can vary significantly over short distances” -> First, this is contradictory since the BRDF effect is significant, but for the methodology implementation is assumed as not. Secondly, assuming that every terrain region of 25 hectares has the same BRDF is not admissible to correct centimetric pixels.

 Equation 8 -> Specify the equation to be minimized by least squares (AX+t=V).

 “The sliding window should consist of at least two pixels to solve for the two parameters” -> It is highly dubious that all the atmospheric, BRDF, spatial and time variations could be corrected by only two DN. Indicate, according to the error propagation, what is the minimum number of observations to achieve an acceptable result.

 “In circumstances where the camera offset, &#x1d450;1, is zero and atmospheric reflectance, &#x1d70c;&#x1d44e;, is small, C may be regarded as being sufficiently small to be omitted from the model […]one pixel may be used” -> This reduces the method to a simple arithmetic mean. These assumptions have to be supported by a previous check.

 “Resample uncalibrated aerial image” -> Authors obviated that the resampling/interpolation modify the acquired DN (and in the step 3 the computed reflectance values), adding an additional uncertainty factors to the whole process.

 “In practice, this does not hold true” -> This could be solved by selecting a more suitable sensor to the calibration and/or verification.

 “In our approach it is assumed that the effect of differing PSFs is negligible” -> Too much hypotheses to obtain reflectance values from a DN mean correction.

 “Cubic spline interpolation […] was found to best satisfy this requirement” -> Add the results and comparison to support this claim.

 “single wide swath width reference satellite image” -> Is it mandatory?

 “When upsampling to the aerial resolution” -> downsampling?

 “(Figure 1Error! Reference source not found” -> misprint.

 “The study site is covered by 2228 images” -> this single number is insufficient to the reader. First, why so many images where required? Secondly, what was the mean number of overlapped images?

 “C was assumed to be small enough to be ignored” -> On the basis of what results? This claim requires results to support it.

 The first paragraph of section 2.3 is not relevant for an Accuracy Assessment section. Remove or relocate it.

Section 3.

 Figure 4: What are the units? What is their range?

 “Each DMC image was converted to surface reflectance” ->  Additional information is required, such as the individual image residuals, relation among the number of overlapped images and the fitted error,…

 “Figure 9 is effectively comparing calibrated values to the values that were used for calibration” -> this comparison has to be made it with the pixels that were not used for the computation, not with the ones that enter in the equation.

 “Mean absolute difference” -> Why there were used absolute values for the mean, and signed values for the standard deviation?  They should be similar.

 Figure 5 -> What are the units of Y axis?

 Figure 6 -> How many points were plotted?

 Table 1 -> RMS and standard deviation are redundant. Remove one.

 “extracted DMC surface reflectance” -> But, how was the calibration coefficient upsampled from the MODIS grid (500 m) to the DMC grid (0.5 m), to be, at next, downsample to match with the SPOT grid (10 m)?

 “Disparities in the mountainous areas are [..] caused by variations in the time of day when the images were captured” -> Authors said that in section 2 that “Their method model spatially varying functions of the viewing geometry and atmospheric conditions”. So, how it is explained that error, if it was already taken into account in the methodology?

 “Abrupt changes in BRDF may occur between adjacent fields in cultivated areas along the major rivers. These changes may not be captured at the MODIS resolution and could also be contributing to the NIR differences in these regions” -> Again, these effects were, in theory, assumed not relevant by the authors: “it is assumed that atmospheric and BRDF effects vary little at a small spatial scale and that the reference resolution is sufficient to capture gradual variations in atmospheric conditions and BRDF”. So, if it is relevant, why it was not considered in the methodology?

 Table 2: There is so much difference between RMS and Std Dev Diff. It should not be. Please, check it. Moreover, the R2 values, which are very relevant for the discussion, were not commented. They are pretty low and represent, among other error sources, the effect of surface heterogeneity.

Section 4.

“course-resolution” -> coarse-resolution.

 “The proposed technique is simpler and computationally more efficient” -> the image requirements were not shown in text, as well as the computation times.

 “NBAR data used in the case study is accurate to “well less than 5% albedo at the majority of the validation sites” -> this is a very relevant info that has to appear prior in the manuscript, and taken into account in the least squares computation.

“The relatively higher (5.88%) NIR reflectance difference between the DMC mosaic and the SPOT 5 values […] are likely due to the more exaggerated differences in NIR RSR’s between the MODIS, DMC and SPOT sensors” -> Confusing phrase, since the authors assumed that the differences in the sensors’ RSR are not significant for the method: MODIS NBAR “was also selected as it has similar spectral bands to the Intergraph DMC”, and it was supported by the results of figure 4: “The correlation between the DMC and MODIS band-averaged values (Figure 4) is surprisingly strong”.

-**Reviewer 4**  
  
  -

It is a well written paper with an interesting topic.

The methodic section lacks the definition of Mean Abs. Diff., "Root Mean Square" and "Std. Dev. Diff." from Table 1 and 2.  
Probably Mean Abs. Diff. and Std. Dev. Diff. are the mean and standard deviation of the E(x,y) distribution in formula 10.

Sec 3.4 reports larger differences when comparing MODIS calibrated DMC images with the SPOT image. It would be interesting how much of the difference originates from the difference of the MODIS and the SPOT image itself. This should be analysed by comparing the MODIS and the SPOT image.

The discussion about the methodology is hidden in the results section. A summary of the shortcomings and restrictions should be added and how they could be treated.

Since there is a large discrepancy of the GFOV between DMC and MODIS, the method acts as a reflectance calibration grid.  
The additive term will subtract the path radiance and the multiplicative term will perform the irradiance normalization of the data.

It will not correct for radiometric nonlinearity, so a stable aerial sensor and a good relative calibration is still a prerequisite for high reflectance accuracy.  
Also the adjacency effect (atmospheric contrast reduction) is not addressed. A weak sharpening filter might do the work. A rigorous treatment requires averaging a large neighbourhood around each pixel and is prohibitively slow.

Any radiometric fault in the satellite master images like local haze or clouds will spoil the result.

In case the aerial image project is spread over several satellite images a mosaicking of the latter has to be performed beforehand.

Please consider also the comments in the revised pdf.

**My Summary of Important Comments etc**

**Reviewer 1:** (note some are reviewer 1’s comments in the text are actually reviewer 4).

There are only minor things which could be improved in the paper but some sort of error budget for the overall process would be nice to have.

* Something weird about linear rel bet sesnsors not meaning the lin model can be used i.e. he misunderstands somehting
* Justify that C was assumed small enough to ignore. Esp in blue band
* MODIS etc validation should compare r2 against 1:1 (see above about lin rel) again I think this is bogus as I do include MAD etc
* Suggests comparison against typical spectra such as lakes, shadows and very bright objects i.e. more specific stuff

**Reviewer 2:** is good for mosaicking but more rigorous reflectance analyses can not be performed. Limitations must be highlighted (fair) eg time differences, differernt spectral sensitivities, Furthermore, some improvements must be made in the paper especially regarding the specification of important data that should be included and detailed in the accuracy assessment section. Suggests to give different spin eg “The technique allow producing seamless mosaics corrected for atmospheric and coarse-scale bidirectional reflectance distribution”. Major revision.

* Justify my claims about computational speed & simplicity
* Info on aerial survey – overlap GSD etc
* Some more details needed in theoretical formulation i.e. adjacency effects are ommitted
* Clarify limitations around assumptions of coarse scale BRDF. It is not assumed that it varies little at small spatial scale – we know it does vary i.e. it is a necessary approximation
* Justify offset C is zero
* Can we quantify the error from assuming spectral bands are identical
* Justify differing PSF negligible assumption (or state it as a limitation)
* Some misunderstandings around boundary pixel treatment
* Discuss how the differing times the images were taken affect the results i.e. you are correcting from one time to another time
* Specify PPS corrections
* Provide accuracy of SPOT5 correction
* Why do I specify RMS and SDD – when they are both the same? (stupid – let me check this out)
* Discuss low SPOT-DMC R2
* Discuss the limitations of the method

**Reviewer 3:**

However, the methodological approach is the proper for image mosaicking. Moreover, the paper has two mayor issues: the first regarding the high number of hypothesis considered in the different steps, which were not quantitatively supported. And secondly, the lack of quantitative results and technical information about the experiment:

\* Assumption of the same BRDF and spectral signature of a surface cover in a 500 x 500 m region (25 hectares). More or less true (interpolation means <500m and the assumption is of the same average spectral sig not the same spectral signature) – so state this as a limitation

\* Claiming that the viewing geometry, terrain effects and their consequent BRDF effects are negligibles. Not true in this wording but the limitation of coarse scale BRDF must be stated – he has a point here. I’m not claiming they’re negligible but can be lumped into the linear model i.e. the effect of dmc viewing geom variation inside a sliding window has negligible impact on reflectance or the the dmc BRDF stays fairly constant inside a sliding window for the same surface even though the viewing geometry is changing.

\* Generation of reflectance values for 0.5 m/pixel on the basis of a single coefficient fitting (arithmetic mean) of 500 m/pixel reference values. This assumption of land cover homogeneity implies an error propagation in the final reflectance values that invalidate their use for radiometric analysis at this very high spatial resolution. Invalidates their use for radiometric analysis yes (there is no VHR product that can claim this and I’m pretty sure I never claimed it could be used for radiometric analysis) but not for remote sensing in general. Also, it is not a single coefficient fitting – “upsampling” means likely coefficients are interpolated. The example uses a single coefficient, the method does not.

\* Assumption that the atmospheric effects are negligible for the different sensors, the flight /orbital heights, and the different days. No – it assumes they can be modelled by a spatially varying linear relation. It does not take into account different days though – how could it? Most vicarious calibrations also are not same day measurements. UM, OK maybe based on Vermote – we should double check the eq for aerial sensor that includes flying height. Again, perhaps this is a misunderstanding.

\* The removal of radiometric coefficients under the assumption of null camera offset (without any empirical study). OK – PPS corrected imagery is null offset – it is not an assumption. The assumption of clear sky C=0 is suspect though and needs some justification or clarification.

\* The claim that all radiometric effects could be corrected by one resampled pixel in the aerial image. Clarify “the claim” then that it is an approximation. This is same point as above.

\* Evaluation of the least square fitting by the analysis of the fitted points, instead of check points not involved in the computation. This is not true – we use spot which is not involved in validation and we make exactly this point in the paper - but perhaps we should omit the modis validation. But then he complains about us not having enough error analysis.

\* As well as, confusing conclusions, such as the wavelength discrepancy of VHR and MODIS sensors are negligible in the calibration computation, but a similar one between the VHR and SPOT in the final comparison is the reason of higher final differences. Seriously? They have an effect but a small one that needs to be ignored to formulate the method. The bigger the difference in RSR will introduce bigger error of course. Hence my conclusion

And secondly, the lack of quantitative results and technical information about the experiment:

\* The whole experiment is summed up with only one number, while no other internal analysis were carried out in the different steps (least square a-posteriori deviation, residuals, …) seriously? There are scatter plots of the band avg relations, scatter plots of dmc-modis-spot relations, images of mosaics, error images etc. the only step we can carry out an error analysis is the final one where we have “reflectance”. Let’s see what he means by “least square a-posteriori deviation”. Since he keeps making the point that we are using a single pixel arithmetic mean for M – obviously the least square error is 0

\* The authors employed the MODIS data as a ground truth (zero error), but, as they stated in the last paragraph of the manuscript, its accuracy is about 5%. This a-priori error must be taken into account for a proper analysis. Yes- fine

\* The radiometric calibration is summed up to the least square fitting of only one parameter, namely, an arithmetic mean. No it is not. What about the up and downsampling steps, the linear 2 parameter fit inside a window. He is talking about a skewed version of the application, not the method. And even if it is – that is advantage of the method – it is simple and it works as well as anything else out there…

\* The aerial campaign for images data acquisition is confusing. The authors only mention that more than 2000 images were acquired, but no information about how many images were used, how many redundant points were involved in the calculations, or why the 20-days gap among acquisitions. It states clearly that all images were used. We should state that the 20 day? Gap was the closest we could get in available imagery. I don’t know what he means about redundant points or why that matters.

\* The geo-referencing error was only taken into account to support the discrepancy with the SPOT imagery, but not in the calibration phase (assumed as perfect). Well the geo-referencing errors in the calibration phase will also contribute to the errors. As it is at 500m pixels, it is a lot less of an issue than at 10m.

\* The efficacy tests of the interpolation algorithms were not shown. Perhaps we need to add something here but it is a can of worms.

This work appears to be the continuation of a initial work of 2014 (“Very high resolution mapping of Spekboom canopy cover”) where the same sensors (0.5 m VHR, 500 m MODIS and SPOT image) were used to the radiometric correction of aerial photographs, but obtaining different results: “The mean absolute reflectance difference of the calibrated imagery was found to be 3.92% compared to an atmospherically corrected SPOT reference”. It would have been interesting to add this information to the discussion of the work, to highlight the differences.

For all these reasons, and the following comments, I have to propose the rejection of the manuscript in its actual form.

* State limitations – we can’t assume negligible BRDF variation in 500m
* Make it clear that we are modelling atmospheric and BRDF effcts with spatially varying linear model
* Cannot be used for radiometric analysis
* What is the effect of different days
* State PPS corrections and null offset
* Justify C=0 assumption
* Discuss varying atmospheric and BRDF effects and that atmospheric effects are more likely to vary slowly over 500m pixels. Also make it clear about the spatially varying linear model.
* Omit MODIS validation?
* (least square a-posteriori deviation) ?
* Include MODIS accuracy more formally in the error analysis
* Include more info on the images, overlap etc
* Talk about geo-referencing with ref to both MODIS and SPOT
* Interpolation study? The paper is already long.
* Either remove claim about computational efficiency or justify it
* Some clarification is needed around BRDF variation with viewing angle and correction to nadir
* Clarify that interpolation adds uncertainty
* We need to re-align the method with mosaicking rather than radiometric correction perhaps as we can’t expect reflectance accuracy on a 0.5m scale but only on average
* Individual DMC image residuals? All 2000 of them? Scatter plot? Why though? What does this tell us?
* Specify how the DMC mosaic was resampled to SPOT resolution
* Sort out MAD, SDD, RMS necessity

**Reviewer 4** generally happy. Has some good points

* It would be interesting how much of the difference originates from the difference of the MODIS and the SPOT image itself. This should be analysed by comparing the MODIS and the SPOT image
* The discussion about the methodology is hidden in the results section. A summary of the shortcomings and restrictions should be added and how they could be treated.
* Also the adjacency effect (atmospheric contrast reduction) is not addressed. A weak sharpening filter might do the work.
* It will not correct for radiometric nonlinearity, so a stable aerial sensor and a good relative calibration is still a prerequisite for high reflectance accuracy.

**Summary of common valid points**

* ~~Make limitations and assumptions clear. Don’t pretend that it does brdf at fine scale. Point out that because of this reflectance will not be accurate at source resolution. But that this is a property shared by other correction methods.~~
* ~~Perhaps align the method more as a mosaicking technique with coarse brdf and atmospheric correction. Cannot be used for radiometric analysis but remote sensing.~~
* ~~State what the PPS corrections were and that the DMC imagery is zero offset.~~
* ~~Better justify assuming C=0.~~
* ~~Sort out RMS and SDD and MAD etc~~
* ~~Justify claims of computational speed and simplicity or leave them out.~~
* ~~Talk about the differing times the images were taken and what effect that has~~
* Further error analyses like per image, a priori least squares, contribution of differing RSR and the up front MODIS error. Talk about the limitations of the error analyses i.e. we have no absolute reference.
* ~~Discuss low SPOT-DMC R2`~~
* ~~Provide accuracy of SPOT correction if possible.~~
* ~~Mention that the adjacency effect is not addressed.~~
* ~~It will not correct for radiometric nonlinearity, so a stable aerial sensor and a good relative calibration is still a prerequisite for high reflectance accuracy.~~
* ~~Include MODIS accuracy more formally in the error analysis~~
* ~~More info on the images, overlap etc~~

**My notes on things to add**

~~Section on sliding window assumptions. Mention similar assumptions are made by other methods (eg an entire scene is modelled by one BRDF ??). Crude but necessary and produce good results. Possibly MODIS BRDF params could be used to explicitly correct for DMC view geom. Eg you change MODIS NBAR view geom to match DMC, then correct for only atmospheric effects using lin model. Then you re-apply the inverse BRDF adjustment to atmospherically corrected DMC images. This would likely have its own challenges though. BRDF parameters are available is MCD43A1 and the actual kernels are documented in the “MODIS BRDF/albedo product: algorithm theoretical basis document version 5.0”~~

~~Section saying that other sensor combinations require validation: RSR and viewing geom variation in a ref pixel.~~

~~Section re-iterating uniform and lambertian assumptions.~~

~~Possible eval on likely BRDF variation due to viewing geom.~~

~~Possible section comparing performance of method to other actor RTM or at least referencing the problem of time-consuming actor algorithms.~~

~~DMC focal len and fov rather than footprint~~

Use the words empirical and qualitative, “first approach” as applicable rather than implying we have theoretical basis.

Limitations and assumptions

* ~~The theory is to support the model. We acknowledge it is not rigorous or complete. Similar empirical models are used in other methods.~~
* ~~BRDF and atmospheric coupling is ignored or assumed approximate by the model~~
* ~~Clear sky assumption~~
* ~~MODIS error~~
* ~~Assume BRDF and atmospheric effects are constant in sliding window. This is both land cover and viewing geometry. It is a necessary limitation. Other methods are also limited by eg per frame BRDF models~~
* ~~C=0 is a first approach justified partially by clear sky, PPS corrections and Collings~~
* Evaluation is partly qualitative

Organisation of sections

* The linear model
* The algorithm with resampling
* Viewing geometry and RSR considerations.

**Collings**

Collings does a per-pixel linear model with a per-frame kernel BRDF i.e. they assume uniform cover within a frame. By assuming uniform cover within a frame they can say that all pixels should be normally stributed around a mean refl. This is (somehow) how they get around needing to know actual reflectances. Another important point about Collings is that it is an empirical model, not a theoretical model. They assume a model that describes the rel betw dn and refl. We have perhaps a semi-theoretical or semi-empirical model in that our model is partially supported by theory.

Ref Collings that they find that multiplicative term is more important than additive and covers most of the correction p2580 this is for a situation where cover is assumed uniform but actually varies.

Collings also do an acquisition over a long time i.e. 39 days. Although they don’t use a single ref.

**Chandelier**

We are at same order of speed for Chandelier – rather than claim that we are faster, we could simply state the processing time of x images.

Processing time is nb for aerial due to 1000s of ims.

Physical models req knowledge of atmospheric conditions which are generally not available. “They also require atmospheric characteristic information to be provided, such as the atmospheric composition or the aerosol types, or specific spectral channels to be present in order to estimate the required information from the data themselves”

They call their method “semi-empirical”

They assume const atmospheric conditions during the day

Their brdf model is for hot spots only and is a multiplicative factor applied to “real” refl

They also use a L = M\*refl + C type model where M varies spatially and is parameterised non-lineary. **NB when we talk of others using spatially varying linear models it is the rel betw radiance/dn and reflectrance that is linear. The actual model may be non-lin in the parameters**

“Parame- ters of the radiometric model are determined as the ones minimizing the radiometric differences of matching areas in every image.”

Their evaluation is purely qualitative. Their method is purely relative

**Lopez**

“Nevertheless, it is an approximated way to correct the atmospheric effect and it has to be understood as a first approach, opening the proposed methodology for a well characterized atmosphere experiment”

1 BRDF model per image.

but these coeffi- cients include the errors due to the non perfect atmosphere char- acterization and the temporal variation of reflectance between the image acquisition and the field measurements. As

they solve physical brdf and atmospheric equations using ground truth for a few images. Their method relies on field acquired ground truth.

There is a time gap between flight and ground truth. They ref haest t al 2009 for this. Their time difference is August 20th 2006 and ?? – they never seem to say

Essentially, they seem to adopt a std RTM and kernel-BRDF approach using vicarious calibration on a few images. It is not really a mosaic technique-

**Gehrke 2016 (2010)**

Uses a spatially varying linear model

Parameters are interpolated between radiometric fix locations – similar to me. They have fewer fixes than MODIS points per image.

Think about the possibility of different window sizes for different channels or different window sizes for different parameters? Eg – first fit a per image haze parameter and then a per siding window gain parameter i.e. assume haze constant across image. But how to find that offset – it could be the min or low percentile val for that channel in an image, but depending on the cover, that is no guarantee that it is pure haze – so that won’t work. We could look at other methods of actor such as the straight line method and see if that works. Thinking about it, constraining C to 0 probably helps regularise the problem otherwise there could be a of noise in the M and C values. Working out a per siding window C could make things worse not better and not just because of the sliding window size. We could also assume only the blue channel has a C worth bothering with. Another possibility is to pre-process with a basic atmospheric correction procedure like dark pixel subtraction. Then the assumption of C=0 should be true(r).

In talking about dealing with temporal variations, mentions that we do deal with changing atmosphere. Well not really since C=0…

**To DO**

~~Plots with 1:1 regression line~~

~~MODIS DN plot units~~

X Number of points

~~Discussion of terrain and shadow effects~~

Summary of limitations

~~Additional assumptions of model~~

Eg of procedure with images

~~Comparison of specific spectra between SPOT and DMC~~

~~Make a note about drones having > viewing geom variations~~

~~Note that RSR’s should be checked~~

~~we don’t know if it is the spot image that is 4% out or the dmc image that is 4% out. Only that they are 4% different. We should make this explicit.~~

Justify the resampling algorithm with results

Hillman: 10pg

Berry: 40pg

Sells: 60pg

A3: 7hrs reading – 8 hrs writing = 15 hrs

Moving house: 40 boxes \* 30mins each = 20 hrs @ 2hrs/day = 2 weeks

Unpacking…

Sort out Cedar Lodge…

GEF: 16 hrs

4 wks until A3

Phone Bronnie

Order books

A3: 3 days

Packing for A3+Otter: 1 days

Packing Cedar Lodge: 20 hrs ~ 4 days

Move: 1 days

Move Logistics + rent cedar lodge: 3 days

GEF Report: 3 days

Unpacking: 2 days

Other admin: 2 days

19 days ttl – we have 30 days

Order of activities:

Call Bronnie

Arrange with Dad

Make a decision

Arrange with mover and Judy

Packing house

Move

Unpacking

A3

GEF

Otter prep