For clarity, the mathematics is based on three assumptions:

(1) alpha depends only on scale and the two heights a and b.
(2) The mean and variance in cloud cover fraction is the same at both heights a and b.
(3) The mean and variance in cloud cover fraction is the same in both grid boxes j and j+1.

Response to Referee #1

Major Comments:

The note up to Eq. (12) seemed clear to me. But Eqs. (12-15) and (19-21) could use some extra explanation or the presentation of intermediate mathematical steps. For Eq. (12), I'd note that c_a = c_b = mu and that <c_a'2> = sigma.

(1) *It is noted in the final version of the paper that for Eq. 12 $c_a(j) = c_b(j)$. Hence, the mean cloud cover fraction, μ, is the same at both heights (as is the variance, σ^2). [NB. The assumption that $c_a(j) = c_b(j)$ is dropped later on, but the mean and variance are always assumed fixed at μ and σ^2 respectively]. This is explained in the new version of the paper. Also, between Eqs. (12-15) and (19-21) we have put in some more explanation and presented intermediate mathematical steps.*

For Eq. (13), I'd provide an extra step in the derivation and note that sigma and mu retain their definitions from Eq. (12) (if that's true).

(2) *(The old) Eq. (13) follows from an additional assumption that μ and σ are also the same for both grid boxes j and j+1. This is explained in the text. Eq. 13 it is a somewhat complicated to derive, but we have now put in the intermediate mathematical steps we have used.*

For Eq. (14), I'd clarify how this equation depends on c_max in Eq. (11) and what the value of c_max is.

(3) *This now explained in the text near Eq. 11.*

In Eq. (15), what are 'a' and 'b'? Are they related to the two altitude levels 'a' and 'b' (see Eq. 1)? If not, can you change the variable names?

(4) *No, a and b in this case are the parameters of the Beta distribution. We have changed the variable names to avoid this confusion.*

In addition, can you write mu and sigma in terms of 'a' and 'b' for the convenience of readers?

(5) *This is explained in the new text and (1) above.*

line 21, p. 9807: What is *<c_a>* and how is it different than c_a and C_a?

(6) *This is given in (1) above in that <c_a> = mu is the long-term average of c_a*

I don't understand the derivation of Eqs. (19) and (20).

(7) *More explanation and mathematical steps are included to show how (the old) Eqs. 19 and 20 were derived. These comes from the references mentioned in our 'technical note' or, in a more general form, from the following paper, which is referenced in the new text:*

S. Nadarajah and S. Kotz, "Exact distribution of the max/min of two Gaussian random variables," IEEE Trans. Very Large Scale Integr. Syst.,vol. 16, no. 2, pp. 210–212, Feb. 2008.
http://ieeexplore.ieee.org/xpl/articleDetails.jsp?arnumber=4403040

Minor Comments:

line 17, p. 9804: "data is discarded" should be "data are discarded".
line 19, p. 9804: This would be somewhat redundant, but I would replace "two adjacent grid boxes" with "two horizontally adjacent grid boxes".
line 10, p. 9805: Replace "Where . . ." with "where", and do not indent.
line 25, p. 9807: Replace "aproach" with "approach",

(8) These are corrected in the final version (as are some references that were missing in the reference list, but given in the text).

Response to reviewer#2

The reviewer became lost in the algebra and assumptions discussed used around Eqns. 12 and 13, and onwards, which made following the rest of the paper very challenging.

(1) In the revised manuscript more mathematical steps are given to show how the equations after Eq.12 are derived.

Some further discussion on the wealth of available data and how it can be used to address these issues should be discussed.

(2) In the revised manuscript, we reference work using 'CloudSat and CALIPSO for the vertical profiling of multiple cloud layers and deriving vertical correlations, and from imagers such as MODIS'.

The abstract starts off clearly enough, but after line 13 it gets detailed and it is unclear as to how these details should be considered take-home messages. Keep the abstract clear and to the point because this is as far as most readers will get.

(3) The lines beyond line 13 in the abstract are dropped.

Line 13: clouds are deeper

(4) This is typographical error is fixed in the new text.

Section 2. It is difficult to tell apart the uppercase and lowercase 'c' for cloud fraction. Furthermore, the 'rand' and 'max' subscripts are lowercase and uppercase depending on the case of 'c', but this is not true for the subscripts 'a' and 'b' (e.g., eqn. 5). Would it help if 'a' and 'b' changed to uppercase if 'C' was uppercase? 'C_T' and 'c_t' follow this convention.

(5) This would make reading easier. However, we have kept it 'as is' because a and b are the two fixed altitudes, which don't change with scale. Changing them may bring its own confusions.

Section 3, lines 19-23: Regarding the question of averaging two adjacent grid boxes, the idealized nature of this study is appreciated and well taken. But, if that grid box is averaged in the zonal or meridional direction, could there be anisotropies in certain cloud regimes that would lead to a breakdown of this approach in a practical setting, or may blur out the signal shown in this paper in real data?

(6) True, anisotropies would make R and alpha_2 directionally dependent. This wouldn't affect the mathematics in our technical note, but could blur the signal when applied to real data if all pairs of adjacent grid boxes are used when finding alpha_2. This is mentioned in the new text. This could be addressed by giving a direction to j with, say, grid box j+1 being zonally or meridionally adjacent to grid box j.

Furthermore, can there be 'scale breaks' in particular cloud regimes that could cause different values of R depending on whether the grid box was averaged over a scale in which a scale break in power density or variance is observed? For instance, see Wood and Hartmann, 2006, J. Climate for low cloud examples (there are non overlapping examples). I could not find an obvious reference for this issue relating to overlapping clouds.

(7) In this note, we assume that R and the variance do not change between grid boxes. If they did then alpha_1 would be different in the two grid boxes and alpha_2 would be a weighted average of the two values for alpha_1. This would potentially blur the linear relationship that we present if the two values for alpha_1 were very different and would make the mathematics more complicated. This is mentioned in the revised manuscript.

p. 9805, line 15: With regard to the time averaging, over what time scales are we talking about here? A day? Week? Month? Season? Since this is an idealized study, at what time scale would the averaging need to occur at for this study's results to hold?

(8) For the idealized case this is not that important. However, we do need the mean and variance in the cloud cover to be stable and similar at both heights. We mention that most published work on overlap is based on monthly or seasonal averages.

Line 16: in the parentheses, should it say 'and the altitude between a and b'?

(9) This is fixed.

p. 9806, line 6: not sure if this is an error or the mixed notation wasn't defined. A lowercase 'c' is mixed with an uppercase subscript 'MAX'.

(10) This is also fixed.

Before line 12, I was able to follow the algebra and assumptions after multiple readings. After line 12, it was impossible to figure out all of the details and steps. How does eqn.12 follow from eqn. 2? I don't see it. Same for eqn.13. Through eqns 17, it appears the authors are deriving forms of the algebraic relationships that will be functions of R so that relationships between alpha_1, alpha_2, and R (i.e., Fig. 1) can be calculated. Some discussion and clear description of what the authors are doing in simple words will be very helpful here.

(11) More mathematical steps are included (see also response to reviewer#1).

p. 9807, lines 9 to 11: Can't this depend on the cloud regime of interest?

(12) Yes, this would likely be associated with vertically deep convective clouds.

From lines 20 and onwards, now the authors are rewriting the algebraic relationships in terms of rho to gather additional insight on the vertical correlation issue. Again, a few additional and simple words on what is being done will benefit the reader.

(13) This is done in the new manuscript.

p. 9808, eqns 19 and 20, where did they come from? How do you get these from two triangularly distributed random variables?

(14) This is explained in the revised manuscript and is explained in the pdf attachment in response to reviewer#1.

Conclusions, lines 7-10: R and rho can be obtained from real data. How does this study shine light on the use of remote sensing data for the cloud overlap problem and its relation to horizontal scale dependence?

(15) If R, rho, mu and sigma are found from real data then this publication allows the value of alpha_2 to be calculated from alpha_1 directly. Mu, sigma and R are all horizontal cloud properties and can be found from the passive or active remote sensing of clouds. However, rho would require knowledge of cloud vertical structure which could come from active remote sensing (e.g. from CloudSat, Calipso, AEOLUS etc.). This mentioned in the new text.

Line 13, which published results? Please describe.

(16) This refers to figures, such as Figure 1 in Oreopoulos and Norris (2011), where alpha_1 and alpha_2 are plotted together on the same graph against height separation, rather than against one another. This is mentioned in the revised text.

Lines 17-20: again, can test with real data. Also, same comment for lines 21-25. How do the authors conclude R must be small? Can't they say something more quantitative and definitive based on real data?

(17) The generally observed increase in alpha with scale (in the publications referenced) implies that R must be positive and less than 1. As the scale dependence increase as R tends to zero, we say R is small as the observed increase with scale is usually strong. Based on published data on alpha, or cloud data it is possible to determine R if there is enough data to determine rho, mu and sigma.

As an illustration, in Oreopoulos and Norris (2011) for June, July and August (their Figure 1) for an altitude separation of 10km alpha_1(75km)≈0 and alpha_2(150km)≈0.04. Based on our note, this would indicate that if rho=1 then R would have a maximum value of 0.8 (our figure 1). However, R could equal zero, provided that rho is at least 0.2 (our figure 2). As rho is likely close in value to alpha_1 this would seem to imply that R is closer to 0 than 0.8. Hence, our conclusion that R would likely be small. This is discussed in the revised text.