The authors would like to thank referee 2 for taking the time to review and offer their comments/suggestions to improve this paper. We think the review was very good and it has improved the paper substantially.

The authors believe that the methodology used in this study is totally valid and can be reproduced in any other study area, although it is true that it would be necessary to previously validate and readjust the coefficients to the features of each particular location. However, the methodology presented is explained in a manner that is completely reproducible.

The combination of channels for constructing the CM and HM algorithms was not done in any unusual way, as we applied logistic regression models that are widely used in the literature. The resulting combination of channels quite accurately discriminates cumulonimbus (in the case of CM) and hailfall (in the case of HM), and has a physical interpretation in terms of the sign and weight in line with the physical arguments discussed in the literature. We therefore consider that this methodology is perfectly valid in order to achieve the aims of the paper.

The originality of the work lies in the direct use of albedos and BTs in producing the logistic regression models. Using physical properties to discriminate cumulonimbus has already been carried out in previous studies (Henken et al., 2011). We also believe that this methodology is completely valid, and that it has the major advantage of the products being obtained more quickly and not having to depend on processed products. In this case, the microphysical and optical properties of the clouds – OP, LWP, R_e, were obtained for a single example, thereby helping in the physical interpretation of the channels that were introduced.

The results of the HDT tool are shown in Figure 6 for the study case, and a new figure has been included for this same example showing the result of the HDT together with the NMDH of the radar. The rest of the results of the verification are shown in Table 9.

Our replies to the specific comments are shown below:

Line 2, page 5455, how is this statement relevant to the paper? You do not have any satellite simulation work in this paper.

It is true that a satellite simulation was not carried out in this study, but the study of Cattanni et al, (2007) is a reference for this study, as they examined the influence of the physical properties of the cloud top with the reflectances and BT of the Meteosat channels. The sentence was change introducing a new example of SAF-NWC application.

Line 17, I think you should be precise here and say that SEVIRI is an advance over the Meteosat First Generation imager, since Meteosat may have more sensors than just the imager.

The sentence has been changed to specifically refer to SEVIRI.

Line 12, page 5456, one can compute updraft speeds only during the developing phase with cloud top cooling, not mature phase. Once the storm becomes mature, there are very small temporal oscillations in IR temp, but this definitely does not mean that the storm has a weak updraft.

The sentence “during the developing phase” has been added to avoid confusion.

Line 1, 5457, I would insert the word “can” prior to “have long-lived cold rings”. Your current statements suggests that all storms have rings.

Done.

Line 3-5, 5457, you state that none have provided an unsupervised hail detection algorithm. While this may be true, several of the studies you cite in the beginning of this paper (and one that you’ve neglected, i.e. Dworak et al. 2012 (Wea. Forecasting)) have shown strong relationships between their products and hail falls. The way you’ve phrased your argument would suggest that your method is the only possible way to objectively recognize hailstorms which isn’t very fair to the hard work of the authors above.

The sentence has been modified, as the aim is not to be disparaging to the work of other authors, but instead on the contrary their work has been decisive in carrying out this study. Also, the reference Dworak et al. 2012 (Wea. Forecasting) has been included.

Line 26, 5462, I suggest you insert 15 min ahead of temporal resolution for this and all other occurrences

Done.

Table 2, some of the parameters you’re using in combination with each other do not make sense. What exactly does “interaction” mean when you’re examining the combination of two variables? When I see (Channel 8.7*Channel 1.6) does this mean that the two parameters are multiplied by each other? Subtracted? Regardless, for the 8.7 um channel discussion is invalid. You say that the 8.7 channel permits distinction between different cloud top phase. This is true, but only when used in combination with the 10.8 micron channel (i..e 8-7 – 10.8 BTD). With this BTD, one is taking advantage of differences in the refractive index of ice in the two channels. The 8.7 alone does not offer the information you state, nor has some sort of “interaction” between 1.6 and 8.5 been published as far as I know

In this study, to construct the algorithms based on the MSG channels, we used logistic regression. This method is a statistical method which selects the metric explanatory variables to construct the probability of a weather phenomenon occurring, based on the information content of the explanatory variables themselves. It is widely used in the literature for combinations of variables in order to discriminate between different classes. This method includes variables or combinations of variables (multiplying their coefficients) using purely statistical criteria, as explained in the paper:

“The predictive equations were constructed with purely statistical criteria. However, the combination of variables must have a meteorological justification for the analyzed weather phenomenon (Doswell and Schultz, 2006)”.

Therefore, in this study, in order to construct the CM and HM, the logistic regression included the variables shown in tables 2 and 5. Apart from seeing that the method really worked, a physical interpretation was made of the channels that were included, finding a clear relationship between the variables introduced by the models and physical explanation of these variables.

In the case of the interpretation of the variable “Channel 8.7 µm*Channel 1.6 µm” included to discriminate cumulonimbus is of great physical significance. We know that in order to

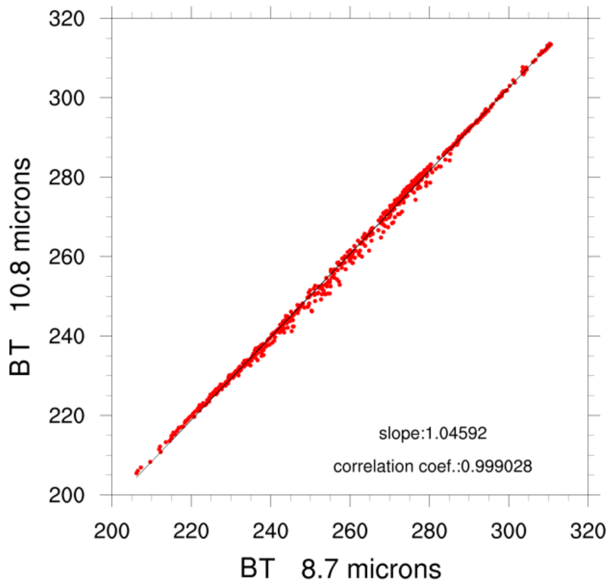
precisely determine the thermodynamic phase of the cloud top, the combination of channels 8.7 – 10.8 BTD is used.

However, in this case the reflectances of the 1.6 µm channel are strongly dependent on the thermodynamic phase of the particles, as cloud tops formed by liquid water hydrometeors have lower absorption than those formed by ice crystals, so the reflectances are higher for water than for ice particles. This means that the sign of this channel has to be modulated by an infrared channel, as included in the model:

"Thus, channel 1.6 µm has a positive effect for channel 8.7 µm BTs less than 226.91 K, and a negative effect for BTs greater than that value. This result has a physical interpretation, since clouds with 8.7 µm BTs less than 226.91 K (lower than the level of homogeneous nucleation, according to Rosenfeld et al. (2008)) are formed by ice crystals. Greater albedo values at 1.6 µm mean that particles at cloud top are small, and updrafts are more vigorous (Rosenfeld et al., 2008); that is, the probability of cumulonimbus increases. On the contrary, clouds with 8.7 µm BTs warmer than this temperature may be formed by liquid water, with large albedos at 1.6 µm, which diminishes the probability of cumulonimbus."

The interpretation of 8.7 µm in the model has been carried out considering their category (Infrared windows) as this channel is selected according the statistical performance in the model. This means that introducing new channels of the same category does not lead to a statistically significant improvement in the model. This fact was checked by substituting the 8.7 µm channel with another Infrared Window channel (10.8 µm; 12 µm). The results of this new model were very similar to the original, after adjusting the coefficient.

Also, the BT of the 8.7 µm and 10.8 µm channels of the CM events have been correlated, obtaining a correlation coefficient very close to 1 (figure below). This is indicative of the proximity of both measurements in the CM events, and so in terms of constructing the logistical model, the inclusion of a new Infrared Windows channel does not provide any significant statistical improvement to the model.



An interpretation which in terms of the sign of the variables, accurately corresponds with the expected physical interpretation.

Line 4, 5464, you never state how you acquired the VISST product, what its characteristics (spatial/temporal resolution) are and how you are "extracting" these properties? I see some acknowledgement at the end of the paper but I think this should be moreprominent in the paper body since this VISST data is used quite often in the paper. You also discuss the LWP, OT, and Re parameters if they were all independently retrieved datasets, when in fact, LWP is a function of OT and Re so LWP would offer no unique information here. You should consult with the VISST algorithm developers to better understand their products.

The physical properties of the channels were only used on one example day, as clearly stated in the text:

"**The cloud physical properties were examined for only one case study as example, on 12 Aug 2011 at 1400 UTC."**

And only to improve the physical interpretation of the results:

"To assist in physical interpretation of the logistic model results, the cloud physical properties optical thickness (OT), effective radius (R_e) and liquid water path (LWP) were extracted via the Visible Infrared Solar-Infrared Split-Window Technique (VISST; Minnis et al., 1995)".

In addition, more information about the VISST algorithm was added.

"Data were obtenined from VISST website (http://angler.larc.nasa.gov/) with the same temporal and spatial resolution of MSG data".

It is true that LWP is a function of OT and Re, but in this study this fact did not have any effect on the results, as these properties were not used in constructing the algorithms.

Line 10, 5467, you should explicitly state that you're using the reflected component of the 3.9 micron channel and indicate the methodology used to compute this reflected component.

This was clarified at the end of the section 4.1. Additionally, the text was modified accordingly.

"The 3.9 µm radiances were converted to BT taking both contributions into account."

Line 13, 5468, you should be more precise when stating that radiances from two satellite channels with broads weighting functions informs one on the "water vapor concentration". The term concentration implies a quantitative retrieval which cannot be done with the satellite information here.

The word "concentration" has been deleted to avoid confusion. The physical interpretation of this ingredient can be found in the first paragraph of section 4.3.2.

Line 25, 5468, you say that cloud thickness is related to Optical Thickness (OT), are you referring to the geometric thickness, i.e. height difference between cloud top and base? If so, your statement is not necessarily true in that one can have very high OT for low stratus clouds given that they are very reflective and spatially uniform, but have fairly low geometric thickness.

We agree. To avoid confusion, "cloud thickness" has been changed to "cloud optical thickness"

Line 21, 5472, you say that high 0.8 um albedo is required for hail storms. Could you provide a physical explanation as to why this parameter has more value than the 0.6 µm channel?

Please see, previous answers. In this case the model has introduced the 0.8 µm channel and not the 0.6 µm channel, because in statistical terms its inclusion did not provide any new information. In physical terms the measurements in these two channels are very similar, which means that it is enough to only include one of them. It would have been a cause for concern if no visible channel had been included.

This fact was checked by substituting the 0.8 µm channel for the 0.6 µm channel. The results of this new model were very similar to the original, after adjusting coefficient. Thus, the physical interpretation of these channels in the model has been carried out considering their category, as they can be replaced by channels in the same category.

Line 25, 5473, a storm need not have a overshooting cloud or V-shape to produce hail. This should be clarified. In fact hail tends to fall upon collapse of an overshooting region.

The next line states: "Bedka (2011) found that 53% of cumulonimbus with overshooting clouds produce hail on the ground" which means it is clear that it not always happens, but that these structures are associated with hailfall on the ground.

General question, you devote a lot of effort to relating various SEVIRI channel information and some strange channel combinations to retrieved cloud microphysical parameters from NASA LaRC VISST. The NWC SAF produces cloud property retrievals in real time, why not just use their retrieval fields (or the LaRC VISST) for your work instead of the individual channel information and combinations?

It is perfectly feasible to develop an algorithm using cloud microphysical parameters, something which in fact has already been developed in other studies (Henken et al., 2011). However, the objective and originality of this study is precisely using albedos and BTs from the MSG channels. The combination of channels is obtained using logistic regression models, commonly used for discrimination purposes in meteorology (Bastarrika et al., 2011; López et al., 2007). The resulting combination of channels quite precisely discriminates cumulonimbus (in the case of CM) and hailfall (in the case of HM), and they have a physical interpretation in terms of the sign and weight according to the physical arguments presented in the literature. We therefore consider that this methodology is perfectly valid for achieving the goals of the paper.

Another general question, page 5474, how could one ever have high optical thickness and strong updrafts (inferred from NIR) without a lot of upper tropospheric water vapor? It would seem that your analysis of microphysical cloud properties items 1-3 are not necessarily independent of one another.

Precisely all of the component elements needed to identify hailfalls are related. Logically, the presence of strong updrafts causes the formation of clouds with a high optical thickness and upper tropospheric water vapor. Here the most important aspect is that the channels selected by the logistic model faithfully reflect the elements needed for hailfall to occur.

Section 6.1, it seems that the sample size of events becomes smaller and smaller as the paper progresses, you start with 700 events for the CM, then 200 events for the hail training, then 78 events for the verification. Why such a small verification database?

Normally, studies that use databases to construct algorithms and for validation purposes use the largest database for training, so that the result is more robust, and a smaller database for

validation. In this case, the limitation lies in obtaining more hailfall events, as we used all of those that were available in the weather radar database from 2006-2011 that met the characteristics of point 5.1, which are essential to compare them with satellite data.

In the case of CM we did not have any problems in including a more extensive database, because the events it includes are much more frequent than in the case of the HM database.

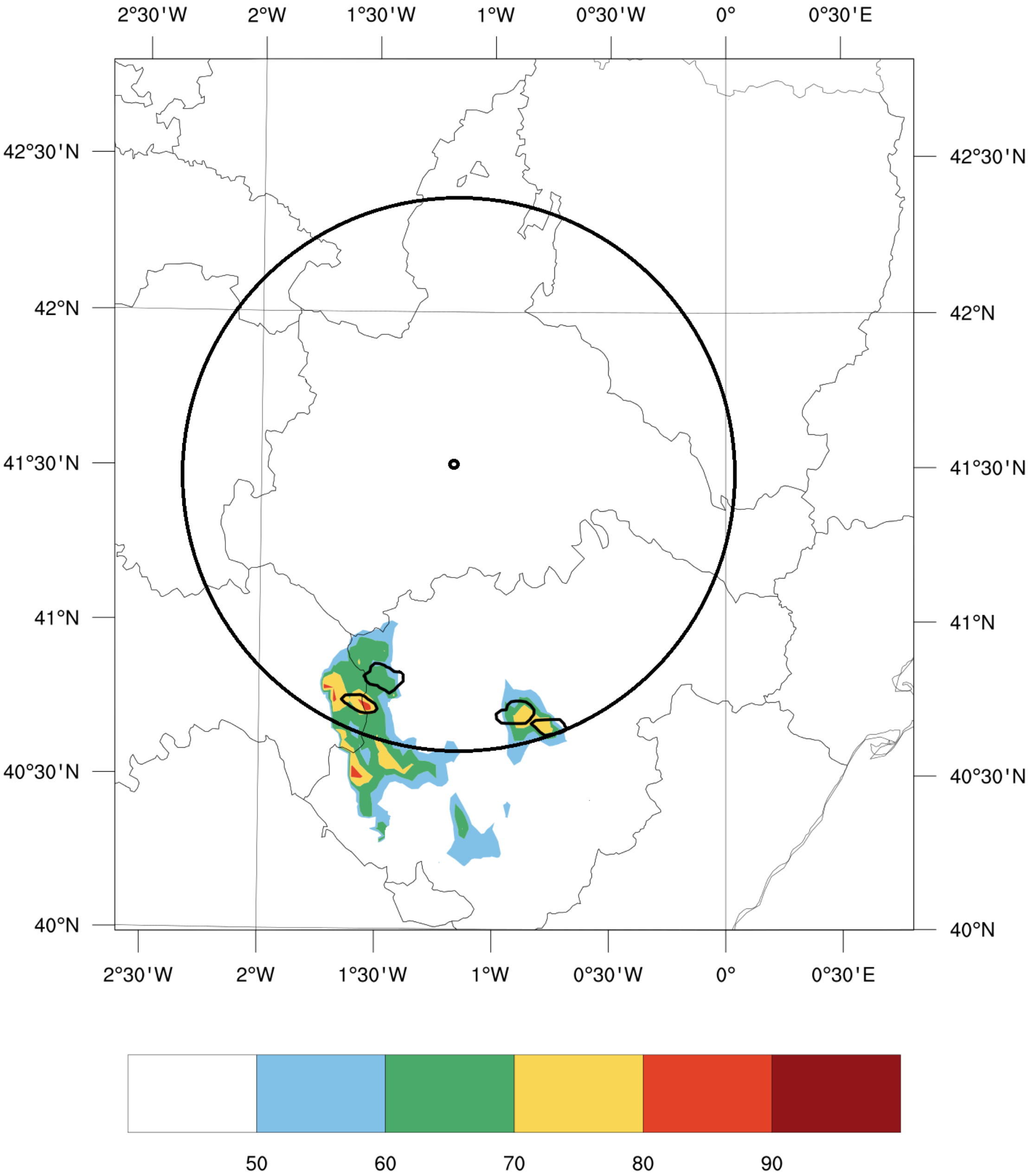
Line 5, 5477, it is unclear how the HM is transitioned into a probabilistic product in the HDT, this should be clarified.

From the end of the introduction onwards there are several mentions in the text to how the detection tool was constructed: "First, the deep convection is identified using a Convective Mask algorithm (CM); second, the Hail Mask algorithm (HM) is used to identify hail falls within the clouds."

Nevertheless, in case this is not clear enough, we have included the following sentence: "The HDT is obtained applying the HM to pixels identified as cumulonimbus by the CM".

Figure 6, the locations of hail fall at this time should be included on the figure.

A new figure has been included, showing the result of the HDT together in the NMHD of the radar.



The HDT result is shown in colours, with the black outlines showing the zones with a high likelihood of hailfall obtained from the NMHD by radar. It can be seen that the results of the HDT closely approximate the results provided by the radar. However, small deviations can be expected, as indicated in the text:

"Deviations may occur, one source of which being error in cloud-top estimation with the radar, and another computation of the parallax effect. Other deviations may be from strong wind shear, which can tilt the storm. In all these cases, the area on the ground where hail is recorded does not coincide exactly with the cloud top."

General question, how do you account for variations in visible and NIR reflectance as a function of solar zenith angle in your HDT product?

At the end of section 4.1 we state: "The reflectances were transformed into albedos following the Lambert method to avoid their dependence on the solar zenith angle".