Review of version 2 of "A simplified method for the detection of convection using high resolution imagery from GOES-16", by Yoonjin Lee et al.

The manuscript is still not fit for publication. The presentation has been improved slightly, but still major issues have not been solved. Still I think, this could become an interesting paper after revision. This is why I commented in some detail again. However, I'm not willing to review it again. It is not the task of a reviewer to improve argumentation, the order of presentation, the completeness of information, and the language details at the level the manuscript is at in this version. This has to be worked out among authors.

List of issues still open from the review of version 1 (reply by authors, new comment):

Major issues:

2. In some parts, the understanding of the underlying physics has to be discussed in more detail. For the first core concept, the information content of the WV channels 8 and 10, the discussion has to be improved at several places throughout the manuscript. Perhaps this method could even be further simplified by skipping the use of the less sensitive channel 8 data. For another central method, the visible channel texture, its limitations (at least for this manuscript's purpose) have to be evaluated and introduced in more detail.

More discussion will be added regarding WV channels, and limitations of visible channel method will be more discussed in the text.
Not really improved. Discussable argumentation was replaced by imprecise argumentation. See below.

3. Although the argument of convective precipitation information for data assimilation is touched upon in the beginning and shortly mentioned in the end again, the goal of the manuscript stays unclear. For quite some pages, the text seems to present a new, complete, very simple solution for a thunderstorm detection and early warning task. A task on which the satellite community has been working on for quite some 20 years. Major improvement seems to be reachable, because of the new GOES capabilities. Only when it comes to test cases and systematic verification, the limitations become obvious. This is where the simple solution presented would have to become more complex – as many working detection codes are. These limitations have to be discussed in the light of existing detection and nowcasting methods in literature as well as possible integration of the investigated aspects in such tools.

First paragraph in the introduction will be revised to clarify the purpose of the manuscript.

The clarification only clarifies the use of latent heating nudging in models. It still does not explain the aim of this manuscript. The title suggest it is storm detection and not provision of heating fields for data assimilation. This issue comes down to a simple question: What is your paper aiming at? Please give this answer in the introduction.

You do not give a clear aim in the introduction and you do not state whether you reached any in the final conclusions. Your method is too simple to be a detection or warning tool. It is not a full algorithm, but only proposing and testing concepts. In the current state, the paper still does provide neither a closed concept nor an independent validation (you use the month data set to "calibrate" your method). You are missing more than 50% of all convective cell, according to MRMS-C radar data. The quality of your Tb algorithm for early convection is still unclear to me, because of the limitations of your presentation. Does it add any lead time compared to radar MRMS-C detection? Or is this all not important for you, because you aim for latent heat nudging fields?

4. The "statistical results" section is not well presented up to now. Definitions and basis are not given clearly. Some of the statements and numbers seem questionable, in part, because of the limitations of the presentation.

The whole section of 'statistical results' will be modified to present one-month results first and then discuss results choosing different thresholds for the methods. Definitions of POD and FAR will be provided in the beginning of this section. The authors agree that providing results using different period of MRMS data in the validation separately can be confusing to readers.
Therefore, in the revised manuscript, only one set of POD and FAR will be presented from the contingency table that is recreated validating results from the reflectance method against 10 minute MRMS data and results from the Tb method against 30 minute (including future 20minute) MRMS data.
Especially this issue is not solved yet. The presentation of section 4 improved, but is still full of vague wording, fuzzy description, and missing definitions. Even a new questionable issue evolved there with the mixing of the two methods into one contingency table.

Specific and minor issues (line numbers are still copied from the version 1 review):

l.20ff: In the light of the many questions left by the presentation of the statistical evaluation/tuning chapter, the numbers here are not useful. Either present some details of the used definitions and scoring basics or leave them out. I do not understand why there is one set of values for two independent methods (major 4).
As mentioned in Major 4, now only one set of POD and FAR is provided for better readability. Since two methods are used to detect convection in different stages and validation dataset includes convection in all stages, results from the two methods are combined.
How can you combine POD and FAR from two independent methods with two independent goals? This is not useful. See comments below.

l.112ff: You state that Channel 2 data is "normalized by solar zenith angle". Please tell us how you do that. This is not a simple or straightforward task. You could normalize reflectivity, but for the texture signal cos(SZA) will not do the trick. The apparent lumpiness increases following a complex dependence on SZA and is strongly dependant on the cloud top structure. (major 2)
Channel 2 data was divided by cos(SZA). A measure of lumpiness used in this study is horizontal gradients of the cloud top surface and thus, shadows (low reflectance) or even brighter surface enhanced by the effect of complexity on SZA actually helps distinguish convective regions from flat cloud top surfaces that are less affected by SZA.
I understand the background idea. But this is not a sufficient normalization, because brightness decreases with decreasing cos(SZA). You are correct, you could compensate for that by dividing by cos(SZA). But for lumpiness it works the other way round: lumpiness/gradient values increase for decreasing cos(SZA)! Lumpiness is at minimum at noontime and strong for low sun!

l.141: "GOES-R CI algorithm". Can you please give a reference?
It is no longer operational products but, it is based on Mecikalski and Bedka (2006) or Mecikalski et al. (2010) cited in this paper.
Give the explanation and reference in the paper not here!

l.155ff: "…updrafts of water vapour…", "…GOES-ABI… can." – You seem to formulate a misconception here. You cannot really see the rising water vapor. The signal is not strong enough. In a WV channel, you do see the water vapor background in midtroposphere. You cannot see low-level dry-convection below condensation level. If you start to see convection cells in this data, it is the cloud body itself you see. Only once the cloud has formed, the emissivity is large enough to dominate the thermal signal in the WV channel. The cloud top "punches through" the background water vapor. Unless the mid-troposhere is very dry, you cannot see what's going on at lower. Please clarify and adjust the discussion here. (major 2)

We agree that these sentences were misleading. "It will be modified to Operational weather radars cannot observe small cloud water, but water vapor absorption bands in GOES-ABI, are more sensitive to these small droplets. During the early convective stages, Tbs that are sensitive to water vapor will decrease due to condensed cloud water droplets aloft generated by a strong updraft." (line 163-165).

What is "small cloud water"? Probably droplets. It's not always liquid water! "WV bands are more sensitive to small droplets". Discussable. "During early convective ... Tb WV will decrease ... by strong updraft ..."? You replace the statement that WV updrafts are visible in WV channels, by the statement that strong updrafts create water droplets and are visible in WV channel due to that. Weak updrafts create the same signal in WV, even cirrus anvil clouds create the same signal. Your argumentation is not precise.

L.270: "… most of convective regions align well with high reflectivity regions in Fig. 2c…", You should not only talk about false alarms, but also about the POD. You are missing large regions with coldest temperatures and, thus, a quite obvious signal just next to the region you detected along 43 N and 93 W to 94 W! These regions shows up clearly in a cold absolute 11.2 mu data and in 11.2 lumpiness! This is opposed to your above statement on IR lumpiness and is a large area completely missed by your mature storm detection. Please discuss. (major 3)

It seems like it wasn't clear from two figures being separate and thus, overlaying figure will be added as Figure 5 (with the description of the figure "For a better comparison between detection from GOES and MRMS, convective regions detected by GOES (Fig. 4d) are parallax corrected with a constant cloud top height of 10km and plotted on top of the MRMS map (Fig. 2d), and it is shown in Fig. 5. Most of convective regions align well with high reflectivity regions in Fig. 2c and convective regions in Fig. 2d." in line 276-278). You will see that over 43N and 93W to 94W, convective regions are also detected by the method.

. By showing this overlay, at least, one can see better that for the most western parts of the GOES detection there are convective radar matches. However "Most of convective regions align well …" is not true. Most (= the majority) of the convective radar area does not show overlap with your detections. You can argue (in the manuscript), but you cannot simply deny it!

l.281: "Growing clouds…" Are these boxes result of your method or did you place them by hand as marker to talk about certain areas. You are talking about the purple and blue boxes next. What about yellow and green? Did you miss them? Please make clear. (major 4)

Boxes are clouds that were detected by the Tb method. This sentence will be changed to "Growing clouds shown in purple, blue, yellow, and green boxes are detected by the Tb method, but all starting from different time." to make it clear for readers (line 288-289). Clouds in yellow and green boxes are discussed few sentences after this sentence (line 294 and 295)

Sorry, I missed the description of yellow and green obviously.

Still the problem of this figure and section is that it is purely descriptive and confusing. It still sounds as if you labelled four arbitrary clouds in these two images at two arbitrary points of time. Two clouds were detected early enough, i.e. before MRMS convective flag (purple, yellow), one late (blue), for one the reader can't tell whether detected in time (green) and at least one can be seen for which no Tb detection was issued at all (98W, 38.3 N). If you want to demonstrate the functioning of your method, you have to think about a different way to show it. Maybe time-step-by-time-step image series with automatic display of Tb detection at exactly the moment they happen.

l.295, Section 4.3, Statistical results: Please start this chapter with a clear definition of the "truth" you compare to, of a hit, miss, false alarm and false positive, all derived skill scores. What is the basic element of your scoring? Is it a grid point, a storm, or a 5x5 window? Please state that for all scores you derive for the early convection as well as the mature convection steps. Right now, this important information is (in part) hidden in the following chapter, but the reader has to guess most of the time. (major 4)
The definition of POD and FAR and its application in this study will be added in the beginning of section 4.3 (line 307-315)
Improved but still not complete. You give this strange mixed contingency table now. It is not useful. You have to do that separately for both your methods. You have to describe all parameters. See next points. What is the basis of your comparison? My question still is: Is it a grid point, a storm, or a 5x5 window.?

l.310: What is the "accuracy" you are talking about? You have to introduce it. It seems to be the correct positives. Please clarify in the beginning of the chapter. (major 4)
It will be clarified "100% accuracy of detecting convection as in MRMS" (line 369).
What is 100% accuracy? This is still not defined. Does that mean 100% of Tb method warnings are MRMS-C cases within 20 min? Or is it 100% of all MRMS-C cases have been detected before. Or is it something else?

l.311: "because most of early convection does not have such a strong updraft". No. It's because it is detected late. See my comment on the WV channels misconception above. In some situations, convection has to reach a considerable height before it can be detected. This is the reason why Mecikalski, Zinner or Guillou did not just use a WV channel to detect early stages. (major 2)
This sentence was misleading as well. It will be changed to "it misses much of the convection and loses an ability to detect convection earlier than radar because not all convective clouds have such a strong updraft." (line 369).
As before, wording is still imprecise. The strength of the updraft (the vertical wind) can not be seen in the WV channels. It can be seen only (in cloud top cooling), if the cloud top is high enough. If you agree, you have to write this. If not, you have to convince me.

l.333f: "improvements in both FAR and POD (lower FAR and higher POD) when later data are included." This is not surprising and it is just tuning values. It would improve further, if you would include another 10 minutes, or even -10 minutes. Unless you can tell us a very good reason resulting from the function of the MRMS algorithm, I would suggest not showing the alternative numbers. They are not much different anyway. (major 4)
The whole section of 'statistical results' is modified to reflect this comment.
This problem still exists unchanged.

Excuse me … some new issues (line numbers from new version 2 now):

l.220: Please state which VIS channel!

Section 4.3: You need to split the analysis into section "4.3.1 Mature convection detection method" and "4.3.2 Early convection method". It makes no sense to mix up the results as you do now.

l. 306: What is both methods? You should give unique names to the methods and use them throughout the manuscript. GOES-C seems to be the VIS maturity detection. You do not state this. Do these numbers refer to pixel? To area? To objects? Not clear. And I asked this in the review 1.

l.320: Before this line it should be section headline "4.3.1 Mature…". And here you state yourself: it makes no sense to mix up results of both methods.

l.322: “(reflectance at channel 2 … cloud top surfaces)” - This all has to be defined and fixed in section 3.2. Otherwise you have to make it clear there that you intend to tune all these parameters here.

l. 337 ff: You should not show this test. Your maturity detection is aimed at detecting mature convection at the time t of detection. Obviously you include MRMS-C cases of the time "t+10min" IN ADDITION to MRMS-C cases at time "t", right? Otherwise not all the numbers would improve.
If this is the case, this is just number tuning. If you would include time "t+5min" and "t+15" and so on, it would always improve numbers without any improvement of quality of the mature convection detection. And I requested not to do this in review 1.

l.355: Here is the point for headline “4.3.2 Early …”.

l.360: “27971 and 73204” This means the basis of your analysis is everything your method detects as convection? With which thresholds? How do you get these “windows”? If this is right you can not derive a POD, because you do not analyse an independent truth. Please state this in the manuscript.

l.366: What about the +30 min you just mentioned?

l. 368: “Accuracy” is still not defined.

l.372: “Growth rate observed”: You rather mean “Cooling rate observed”. Similar in line 375.

L385: “14.4% is achieved, and 96.4% of false alarm pixels”. Can the reader check this statement anywhere in the presented results? Please tell him in the manuscript.

Fig. 2: Do (c) and (d) really show the same point in time as stated? The two precipitation fields (d) and the derived product (d) do not look like! The precipitation areas north of 44 N have hardly any matching feature.

Fig. 5: Suddenly it’s 2230 UTC and not 1930 UTC as before? Something wrong with the time and date?

Fig. 7: Please give all parameters kept constant in the caption or image. And please extend caption with some information on method discussed. Reflectance method for mature convection detection.

Fig. 8: Skip figure b. And again. Extend caption as before.

Fig. 9: “when the pixel was” ...Should better read “…if a pixel was assigned to be convective by MRMS, but not detected by method XXX …“ Correct?

Tab. 1: Please extend caption again. This is too short to be understood.