We would like to thank two reviewers for their valuable comments and contributions to improve this manuscript.

*Table 1 (GOES-r channel description) will be removed based on reviewer 2's comment
*Table 4 that was not included in the manuscript by mistake will be added as Table 1 in the revised manuscript.

	MRMS-C	MRMS-NC
GOES-C	2.73%	0.46%
GOES-NC	3.30%	93.51%

*Figure 3 is edited to have close-up subfigure.

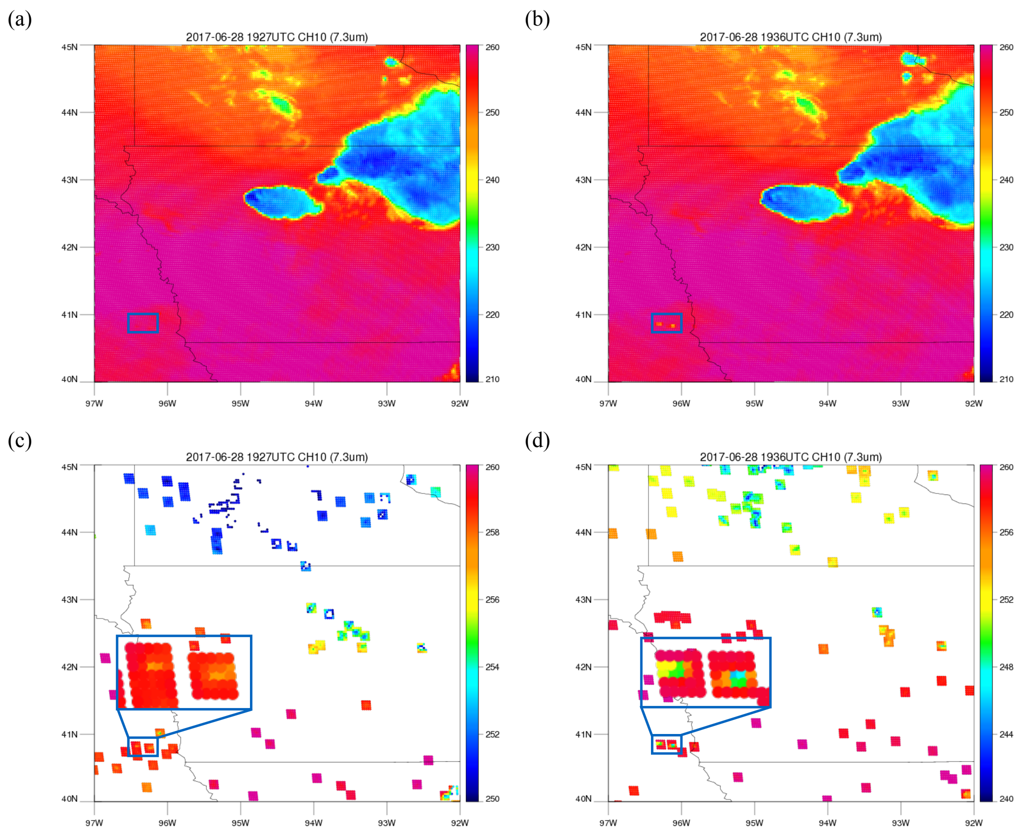


*Figure 5 (laying GOES detection on top of MRMS detection) will be added based on both reviewers' comments

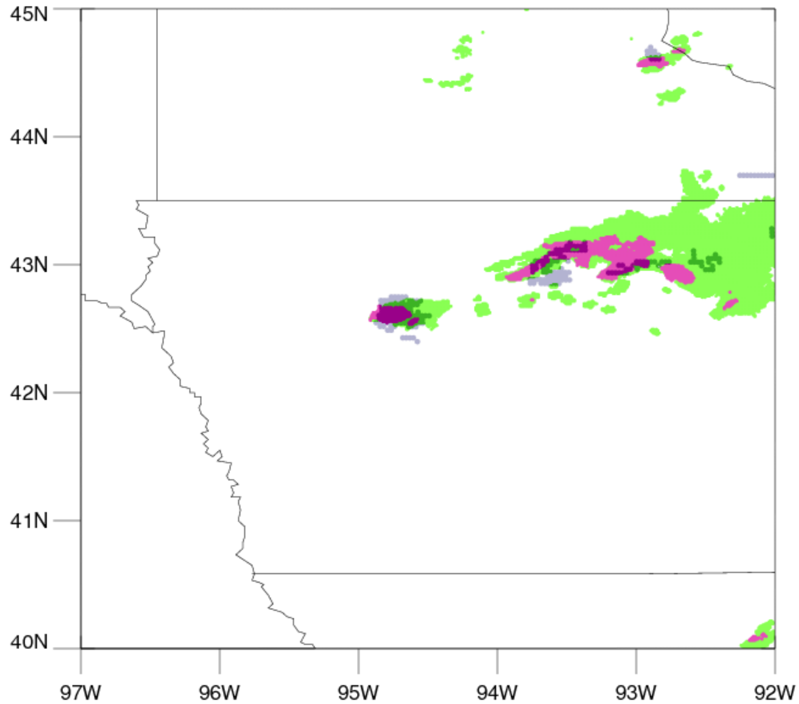


Figure 5: Convective regions detected by GOES-16 (white regions in Fig. 4d) are colored in navy on top of MRMS PrecipFlag at 2230UTC 18 June 2018 (Same figure on Fig. 2d. Pink represents convective while green represents stratiform)

*Line number in the parenthesis is the number in the revised manuscript.

Response to Reviewer #1

Major 1

The introduction will be revised to elaborate more on literature overview, including suggested references from reviewer1.

Major 2

More discussion will be added regarding WV channels, and limitations of visible channel method will be more discussed in the text.

Major 3

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

Major 4

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 T_b method against 30 minute (including future 20minute) MRMS data.

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.

l.28: Maybe you want to add something general like Gustafsson et al. 2018 or something very close to your motivation point like Scheck et al. 2020. (major 1)

Both papers will be added (line 29).

l.84: The use of geostationary VIS and IR texture signals was introduced in automatic detection already by Zinner et al. 2008 (WV texture, Zinner et al. 2013). Another important tool forming an early reference for the use of IR and WV imagery and time trends in it is the EUMETSAT RDT algorithm (Morel and Senesi, 2002, Autones et al. 2009, Guillou et al 2011, see below). (major 1)

These references will be added in the introduction (line 65 and 87).

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.

l.114f: Are you aware of Mueller et al 2019 "A Novel Approach for the Detection of Developing Thunderstorm Cells". That should be discussed somewhere. (major 1)
It will be added (line 88).

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.

l.145: Shouldn't "grids" be "grid cells". This sounds like lab slang to my non-native English ear.
The term will be changed to "grid points".

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, T_bs that are sensitive to water vapor will decrease due to condensed cloud water droplets aloft generated by a strong updraft." (line 163-165).

l.176: "the difference between two matrices will be small.". Which two? Please clarify.
It will be clarified to "the absolute value of the difference between the T_b matrix and the inverse Gaussian matrix" (line 182-183).

l.185: "smaller than -1K/min for channel 10 or -0.5K/min for channel 8". Why is there a difference? A growing cloud top is cooling at the same rate in both channels. Unless there still is considerable (colder) WV above it. Thus, it first shows up in the channel 10, later in the channel 8. You will increase the sensitivity of channel 8 to match channel 10 detections by lowering the slope threshold. You will earn a lot of uncertainty without adding any additional insight. Once the cloud top reaches the upper mid-troposphere above WV background, they will show exactly the same temperatures and trends. Please discuss, perhaps revise. (major 2)
As you pointed out, the cooling rates are similar once clouds are mature enough because their T_bs themselves are similar. However, when clouds are in the very beginning stage, the T_b difference between channel 8 and 10 is high, and the cooling rate is observed to be different.

This makes sense because in order to exhibit the same T_b at both channels at their mature stage, T_b at channel 10 which is usually lower than T_b at channel 8 has to increase faster than T_b at channel 8. Sentences "Growth rate observed at channel 8 is smaller than channel 10 due to higher absorption at channel 8. Channel 8 senses moisture at higher altitude and thus, when water vapor starts to condensate at lower levels, it is less affected, and its T_b does not decrease as much as in channel 10. As clouds grow thicker, signals in water vapor absorption bands are dominated by the clouds, less from water vapor, and their T_bs becomes similar. Therefore, it makes sense again that the growth rate at channel 10 has to be bigger to catch up lower T_b in channel 8." will be added (line 372-376).

l.221ff: Once more : : : What about low sun lumpiness? Shadows cast onto the cloud itself might dampen VIS reflectivity below 0.8. Please discuss. (major 2)
It will be discussed using Figs. 4b and 4d. Sentences "Using reflectance threshold sometimes limits detecting shaded convective regions that exhibits lower reflectance than the threshold of 0.8, and white regions surrounded by colored regions in Fig. 4b are such regions. However, these regions are relatively small, and once they are upsampled into 2km map with nearest neighbour interpolation, some of these regions are included in the detection as shown in Fig. 4d." will be added (line 271-274).

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.

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 T_b method. This sentence will be changed to "Growing clouds shown in purple, blue, yellow, and green boxes are detected by the T_b 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)

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).

l.298ff: Again, it is still unclear for the reader, why you use both WV channels? Are there any channel 8 detection windows not contained in the channel 10 detected windows already? Please clarify or simplify.
Yes, there were windows detected by one channel but not by the other channel. It will be discussed in the text using convective cloud in the blue box in the second case study. (line 293)

l.303ff: "Future MRMS convective flags up to 30 minutes were included : : :" I do not fully understand. It was your goal to detect convection before the radar, wasn't it? That means, it is just logical to check the next 15 minutes/30 minutes. You should check the lliterature on MRMS and give us some details here. Using it, you have to discuss the choice of the future time span : : : the longer it is, the better your scores. (major 4)
As you pointed out, clouds detected by the T_b method are often detected earlier than radar while clouds detected by the reflectance method are usually detected at the detection time by radar, although figure 8 suggests that it still has an ability to detect earlier. This was why two sets of POD and FAR were given in this section, and it was confusing for readers. Therefore, as mentioned earlier when major 4 was discussed, it will be changed to show one set of POD and FAR from combined results by validating results from the reflectance method using only 10 minute data and results from the T_b method using 30 minute (including additional future 20 minute) data. This result will be presented in the beginning of section 4.3.

l.306: Where do you get the "constant speed" from? Please add information.
This sentence will be changed to "assuming convection moves at the same speed that clouds moved during the initial ten minutes" (line 364).

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).

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).

l.315f: The reasoning here is unclear. What about virga? I would just say, it is the typical turbulent, highly statistical nature of the chances of convective cells. Some just do not do it the moment latter.

It is modified to "This would be due to mixing between convective cells and their dry environment or highly non-linear nature of chances of precipitation." (line 377-378).

l.335: For the reader, in order to be able to understand the impact on data assimilation, you have to give proper references or explain a lot more.
It is modified to "To make this method effective and reduce FAR as much as possible for its potential use in the short-term forecast" line 335-336, and suggested reference (Gustafsson et al. 2018) is added in the introduction.

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.

l.345ff: Checking of just one of the two examples you show, it is obvious how to improve it. In addition, the missed regions there are neither cirrus covered nor in decaying mode. You should accept and talk about shortcomings of your very simple method. There are good reasons out there that full detection and warning schemes are far more complex than your approach. Please discuss that. (major 3)
Its limitations are elaborated in section 4.3 using different thresholds. As mentioned in section 4.3, most of the missed regions had a flat cloud top surface, which is a key feature of stratiform clouds in this study. And since the results are compared with radar products observed from the ground, convectively raining pixels might not perfectly align with bubbling pixels. However, as you can see from the two case studies, the locations of convective cloud clusters detected by GOES and MRMS are very close to each other, and most convective regions in each scene are detected by GOES. Nevertheless, limitation regarding the reflectance threshold will be added based on your comment in "l.221ff" and it will be mentioned again in the last paragraph of section 4.3 (line 380-384).

Response to Reviewer #2

General comments

- No comparisons with past studies. It would also be feasible to compare with the results from ABI data of 15 min. on the same dates.

The code that is used in previous studies is not publically available. Although previous studies also present their POD and FAR, they are validated against different dataset (lightning data or reflectivity threshold), and it makes it hard to compare with the previous studies.

- It is hard to see the details in some figures.

Resolutions of the figures will be updated.

- A table for accuracy is not found in the manuscript.

We apologize for that mistake. It will be added in the modified version.

Line 7: What is the meaning of the proper heating?

'Proper heating' means reasonable heating to drive convection in the model. This phrase will be modified as 'The ability to detect convective regions and adding heating in these regions is the most important skill in forecasting severe weather systems.' (line 6).

Line 9: Why is the latent heating especially mentioned here?

It was mentioned because in the operational model, after convection is detected, latent heating is added to drive convection, and methods developed in this paper are also intended to be used in the short-term forecast model. The purpose of this study is more clarified (line 29-36).

Line 11-12: Shouldn't it be more sensitive to the drop size?

This sentence is changed to "Visible and Infrared sensors on a geostationary satellite can provide data that are more sensitive to small droplets" (line 11-12).

Line 14: I don't understand how better spatial and temporal resolutions could be a solution to the intrinsic problem that optical sensors can only get information from the top layers of clouds.

You're correct that it is not a perfect solution to the intrinsic problem. However, by having high spatial and temporal resolution, we can better observe clouds bubbling, which is an indicator of convective clouds. Even though we can't still see through inside of clouds, bubbling cloud top allows us to guess where convection is occurring, and it is the main feature used in this study to detect convective clouds.

Line 15: What are the life stages to be analyzed?

Actively growing clouds in the vertical and mature convective clouds. This sentence will be modified to "This study develops two algorithms to detect vertically growing clouds and mature convective clouds using 1-minute GOES-16 ABI data." (line 15).

Line 17-18: Does this mean that the detection accuracy of the method for the clouds at early stages was 71%?

Yes, but the accuracy increases when MRMS data up to 30min are included because MRMS tends to miss early convection with less or no reflectivity. However, these sentences in the abstract seem confusing and thus, will be modified based on changes made in 'statistical results' section.

Line 19: How the rapid temporal evolution is identified? It needs to be clear, : : : rapid temporal evolution of what?
"the lumpy texture, and rapid temporal evolution" will be changed to "lumpy texture from rapid development" (line 19).

Line 21: Do the convective clouds here are clouds at all different life stages?
'These convective clouds' meant convective clouds that are missed by ground-based radar. This sentence will be also changed based on changes made in 'statistical results' section.

Line 22: It seems that the statement does not match with what is mentioned above in Line 14.
The authors are aware of the intrinsic problem of VIS and IR sensors as mentioned above. Thus, we wanted to address here that this intrinsic problem can't be solved.

Line 26: What is 'this issue'?
This issue means 'initiating convection in the right location and intensity'.

Line 55: What does 'to initiate convection' mean?
In the short-term forecast model, if radar echo exceeds certain values, latent heating is added to increase buoyancy in the atmosphere and trigger or initiate convection.

Line 71: Is cooling really not seen in mature clouds? The sentence needs to be corrected.
The whole sentence in line 70 will be changed to "Convective clouds in their mature stage sometimes do not grow much in the vertical, and T_b decrease is not a main feature that is applicable to such clouds." (line 73).

Line 88: Where are the mesoscale sectors? What's the size?
'Mesoscale sectors' are moved around manually by a person whenever there is an interesting weather event. Its size is 2000x2000 (1000kmx1000km) for channel 2 and 500x500 for IR channels.

Line 89-94: It seems that the last part of the introduction is a bit detailed. They would be rather explained more concisely, and the details would be addressed in the method section.
It was briefly mentioned here to describe how the methods presented in this study differ from previous studies. Literature overview is revised (line 61-70 and line 84-89) and more references are added with help from reviewer1.

Line 91: What are the errors from cloud movements?
Previous studies use atmospheric motion vector to track clouds, and there can be errors in its algorithm.

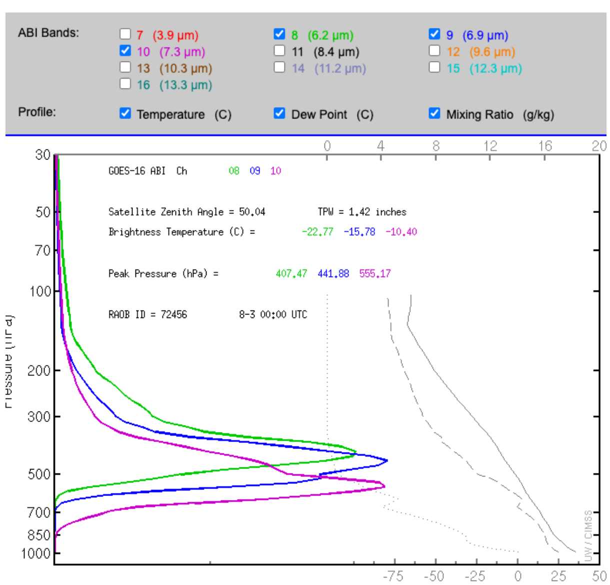
Line 107: It seems that Table 1 is not really necessary. It can be removed and explained in the text.
As you suggested, this table will be removed because they're already explained in section 2.1.

Line 118-120: Need to add references. What about using channel differences? Past studies on detecting convective initiation have widely used channel differencing between water vapor channels and IR channels (c.f. Mecikalski2006, Lee2017)
Mecikalski paper was already mentioned in the introduction, but since Lee 2017 paper wasn't mentioned, this paper will be added in the introduction. As mentioned above, literature review will be revised.

Line 140-141: Clouds do not necessarily reach the tropopause. Clouds form when air parcels reach the equilibrium level.
You're correct that not all clouds reach the tropopause. However, when clouds are mature enough, they sometimes reach tropopause and move horizontally, rather than vertically. (Please refer to Zinner et al., 2013; Validation of the Meteosat storm detection and nowcasting system Cb-TRAM with lightning network data – Europe and South Africa)

Line 142-143: How can the availability of higher temporal resolution data simplify the method to use two channels?
It's worded in a wrong way. This sentence will be removed.

Line 159: What height does each channel usually reflect?
There is no certain height that can represent each channel since height can vary depending on the situation. However, as mentioned in line 125~127, the height that channel 8 represents is higher than channel 10 because channel 8 is more sensitive to water vapor. Figure below is a "realtime" weighting function obtained from https://cimss.ssec.wisc.edu/goes/wf/. You can see from this figure that weighting function of channel 8 peaks at higher altitude.



Line 189-190: "Clouds that develop into deep convective clouds are eventually captured by these thresholds in later times even if they had small decrease in the beginning." This sentence is a bit unclear.
This sentence will be changed to "Clouds that develop into deep convective clouds are eventually captured by these thresholds in later times as they show rapid intensification sooner or later." (line 196-197).

Line 227-229: "It is intentionally chosen so that the method considers warmer convective clouds without those features in the next step when evaluating lumpiness of cloud top." This sentence is a bit unclear.
This sentence is changed to "Warmer threshold is intentionally chosen so that the method considers warmer convective clouds without those features in the next step when evaluating lumpiness of cloud top." (line 233-234).

Results -> Results and Discussion
It is changed to Results and Discussion.

Line 253: It is almost impossible to see the Gaussian shape in Figure 3c. Maybe a close-up subfigure could be used here.
Subfigure is added.

Line 258-259: "Since the same method is used in each time step, the same window can be captured throughout an overlapping time period despite the starting time. Therefore, this method can be used continuously in time." This sentence is a bit unclear.
These sentences are removed.

Line 261-275: It would be better to move this paragraph to the beginning of this section.
Paragraphs before these sentences were descriptions of the scene used to derive the results, and we think that they should come before the results.

Line 269-271: It would be much better to illustrate this as one Figure in the manuscript by merging both figures together.
Figure 5 that shows convective regions by GOES on top of MRMS convective regions is added in the manuscript with the description of the figure (line 276-278).

Line 288-290: "These results show that even though the thresholds for the Tb method can be strict for some growing clouds, the thresholds were adequate for detecting convective storms in their earliest stages." This sentence is a bit unclear.
It was changed to "These results show that even though the thresholds for the T_b method can miss some convective clouds that grow slowly in the beginning, the thresholds were adequate for detecting rapidly growing convective storms which are of more interest during the forecast." (line 296-298).

*Note that the whole section 4.3 will be modified based on two reviewers' comments. It 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 T_b method against 30 minute (including future 20minute) MRMS data.

Line 299: "Since clouds do not grow at the same speed, : : :", which is a bit unclear.
It is changed to "Since growth rate can vary depending on the surrounding environment and different evolution stages" (line 357).

Line 302: Why is the number of samples different?
It is because channel 8 and 10 have different sensitivity to water vapor and represent different height of moisture.

Line 311: Why is it important to have the ability to detect convection earlier than radar? You mentioned earlier that the method of this study is to complement ground-based networks for either off-shore or other regions lacking coverage of radar data.
Radar reflectivity is observed from bigger drops (eg. rain), and thus it takes time for radar to observe signals as cloud water becomes rain. On the other hand, T_bs are sensitive to water vapor, and they are expected to observe condensation by updrafts of water vapor before cloud water becomes rain. It is important to detect convection as early as possible because the ultimate goal of convection detection is to help forecast models to produce precipitation in the right place at the right time, and there is also a spin-up time after the addition of latent heating.

Line 318-322: It seems to be redundant.
It will be put in the parenthesis (line 322-323).

Line 326: "The upper threshold does not change results much, : : :" The result for upper threshold is not shown here.
We thought that it is not necessary to put figure for this because it did not have much change. But '(not shown)' will be added in the text (line 325).

Line 328: However, the choice of 0.4 seems to lose a lot of convective regions.
We preferred to have less POD and less FAR for its potential use in the short-term forecast model.

Line 342-343: ": : : in preventing the method from detecting convective regions.", which needs to be corrected
This sentence sounds misleading so it will be changed to "Therefore, regions that were missed are evaluated further to investigate which threshold contributed most to missing those regions." (line 344-345).

Figure 8 caption: ": : : due to only one of the thresholds.", which is a bit unclear.

It will be changed to "Histograms of T_b, reflectance, and texture values when the pixel was not detected by the GOES detecting method due to each of the thresholds." Note that Figure 8 will be Figure 9 in the revised manuscript.

Line 345: ": : : have flat cloud top surfaces." What percentage was this case? It would be good to provide quantitative values for one-month data.
It was visually shown in figure 8 (now figure 9 in the modified version of the manuscript), but as you suggested, the percentage will be presented (line 346).

Line 348: Q: Why are convective clouds in a decaying mode not considered?
The main purpose of application of this method is to add latent heating in active convective regions and produce precipitation in the forecast model. If convective clouds are in a decaying mode, precipitation will slow down, and therefore no need to add heating.

Line 349-350: "It is also possible that it is due to a misclassification of trailing stratiform regions using radars. Previous studies (Qi et al. 2013; Shusse et al. 2011) have indeed tried to improve the radar classification schemes." The sentences are a bit unclear.
It will be changed to "It is also possible that it is due to a misclassification of trailing stratiform regions using radars. It is indeed an ongoing research in the radar community since better convective/stratiform classification scheme improves QPE retrieval (Qi et al., 2013; Veljko et al., 2019)." (line 351-353).

Line 355-360: Reporting accuracy would be placed at the beginning of this section.
As the whole section is revised, it will be placed in the beginning.

Line 354: to avoid FAR -> to avoid high FAR
It will be changed.

Line 356: There is no Table 4 in the manuscript.
It will be added as table 1 since "POD and FAR" are now placed in the beginning.

All "technical corrections" has been made.