

Reviewer = Black
Guimond et al. = Blue

Responses to Reviewer A

1. As mentioned above, the authors do a very thorough job of identifying uncertainties and possible sources of error. One area that did not seem to be specifically addressed, however, was the fact that much of the retrieval algorithm is based on output from a numerical model run at 2-km grid length. What impact would higher-resolution models have on the relationships derived? For example, the model output is used to confirm the assumption (based on flight-level data) that updrafts > 5 m/s are saturated, while weaker updrafts may or may not be saturated, thus requiring the estimate of Q_{net} . Is it reasonable to assume that a simulation with 1 km, or 100 m, grid length may have a lower updraft threshold for assuming saturation? Would that impact your resulting algorithm? What about if different microphysical parameterizations (e.g., double-moment or bin schemes) were used? I'm not saying you need to rederive the retrieval algorithm using these alternate model configurations, but I think it would be illuminating to at least discuss these possible uncertainties.

We believe that changes in model resolution and microphysics would probably change the vertical velocity saturation threshold. At a resolution finer than 2 km, entrainment and turbulence start becoming resolved which will affect updrafts and saturation. We have modified the manuscript to mention this problem. However, in the manuscript, we focus on observations of vertical velocities and relative humidity, which do not suffer from the modeling issues discussed above (although they are dependent somewhat on scale). In addition, the aircraft flight level data is processed at high resolution (0.5 km). The model data is used only in a supportive role to test the latent heating algorithm. One drawback of the observations is that there is not a very large sample size. This is mentioned in the paper. We have added a sentence describing the modeling shortcomings in the manuscript.

2. Another question I have deals with the fact that only the latent heat of condensation/evaporation is considered in this algorithm, i.e., it only extends up to about 8-10 km altitude. However, other observational and modeling studies (e.g., Zipser 2003, Fierro et al. 2008, 2009, Kelley 2010) have shown a high-altitude updraft peak that is thought to be driven by latent heat of fusion. What impact could neglecting that term have on resulting latent heating calculations, in particular in the upper troposphere? The magnitude of the latent heating may be low, compared with in the lower troposphere, but it should alter the vertical profiles at least.

We only consider warm rain microphysics for several reasons. (1) warm rain heating is easier to understand and compute than mixed phase heating, and (2) the total latent heat budget is dominated by warm rain processes (this was discussed in the manuscript and the work of Tong et al. 1998 was referenced. Their study focused on Florida deep convection, which is probably not too different than deep convection in a hurricane. The

hurricane modeling study of Zhang et al. (2002) that was referenced also came up with a similar conclusion). Including mixed-phase heating would, of course, add some heating in the upper-levels but this would be of secondary importance. As far as the dynamic response of the different profiles in hurricanes, warm-rain heating will have a greater impact because it is coupled to the high vorticity in the mid to lower levels. Shifting the peak heating in altitude does have an impact in balance models as described in Hack and Schubert (1986). We added a sentence in the introduction clarifying the results of the Zhang et al. (2002) study and mentioning that mixed phase processes will alter the vertical heating profile somewhat.

3. Figure 13 could be improved considerably. What about, instead of showing isosurfaces, you show vertical cross sections of latent heating alongside the same sections showing reflectivity from the tail Doppler radar? At the very least it would be good to provide more labeling on the figures, including cardinal directions. Plotting the shear vector would be helpful too. Also, what about including additional contour levels other than just ± 100 K/h? Of course you couldn't do isosurfaces then, but it'd be helpful to see additional contour levels, which again raised the prospect of showing vertical cross sections rather than isosurfaces.

Figure changed due to similar suggestions from the other reviewers. Figure now shows the following: vertically averaged horizontal field of latent heat along with the azimuthal mean vertical profile of heating at the RMW for each pass. Text is updated to reflect new figure and a discussion was added.

4. Would there be any utility in showing examples of latent heating distributions from satellites? You discuss it a little in the introduction – what is gained by the Doppler retrieval compared with satellite (e.g., TRMM)-based retrievals? Is it only resolution? I know that there was no TRMM satellite during the Guillermo flights (or it was just launched), but you would at least show an example plot from TRMM or discuss in more depth what is gained by using airborne Doppler.

It would be interesting to show examples of latent heating retrievals from satellites but that is for another paper. We mention the vast array of problems with satellite retrievals in the introduction and this should convince the reader that these current retrievals are less than optimal. References are given for the reader to find satellite examples. TRMM does not have Doppler capability and thus no winds (a major disadvantage for latent heat retrievals).

5. It would be interesting to discuss a bit more what the limitations are in terms of the retrievals using a 2-km grid of Doppler data. You're essentially only capturing the largest up- and downdrafts, and are presumably missing a significant portion of the spectrum where weaker vertical velocities may reside. What if you were to apply your algorithm to ELDORA data, with resolutions on the order of 0.5 km? And can you apply it to EDOP profiles, other than the mean EDOP hot tower profile shown in Fig. 14? Again, this

suggestion is intended to stimulate discussion in the manuscript, rather than generating a whole new set of results and plots.

The wind retrieval is performed on a 2 km horizontal grid and then a Laplacian filter is applied to the data. Thus, the effective resolution of the data is on the order of 5 – 10 km, which represents only the largest vertical velocities as the reviewer mentions. This can be seen clearly in Fig. 13. The large updraft pulse is captured well, but the smaller-scale oscillations are smoothed out. This is discussed in the manuscript. Applying the algorithm to ELDORA data would be interesting (application to other radars is currently being done). We have added a sentence that highlights the importance of applying the retrieval algorithm to other radar systems.

6. Finally, while the article is generally well-written, I do have one stylistic quibble – there too many parenthetical inserts. For some paragraphs nearly every single sentence has such an insert. This can serve to disrupt the flow of the reading. If you could go through the manuscript and try to remove these inserts whenever possible it would help the flow of the paper.

We agree, and have reduced the parenthetical inserts.

Responses to Reviewer B

Major Point and Recommendations:

1. The current manuscript exceeds the AMS length threshold by roughly five pages. The most obvious way to reduce the length is to remove all sections related to the EDOP radar. The authors introduce and discuss the EDOP radar, but the latent heating retrieval (as best as I can tell) was developed primarily from the Bonnie numerical simulation and the Guillermo radar data (supplemented with Katrina liquid water contents and high-altitude dropsonde data), with no contribution by the EDOP data. Then, at the very end of Section 4a, latent heating profiles are presented for the EDOP data. At this point, showing the EDOP data is a distraction that raises more questions than it answers (or clarifies). For example, if the EDOP has trouble estimating the horizontal winds, how reliable are the heating profiles obtained using Equation (2) – which clearly contains a horizontal divergence term?

In order to reduce the number of pages, we have removed the EDOP material. The paper was not dependent on this information. The paper will still be a few pages greater than the AMS length threshold, but it takes a bit more space to present a new algorithm and convince the reader of its merit including error characteristics.

Minor Points and Recommendations:

1. Pages 8-9: If the EDOP material is retained (see above), please clarify your definition of a hot tower. For example, do they represent the strongest updrafts observed at any altitude, or for a deep-layer average? Also, you provide a hot tower definition for this

study, but never use the definition anywhere else in Part I. Rather, you only use the 5 m/s threshold. If hot towers are exclusively discussed in Part II, then move the definition and background to Part II.

EDOP material removed.

2. Page 12, Figure 3: I recommend adding a sentence or two to the caption of Figure 3 stating that these convective-scale updraft events were defined after a scale separation method was applied. As a result, the total vertical velocity can be negative when a convective updraft is superimposed upon a stronger mesoscale downdraft.

Added the sentences the reviewer suggested.

3. Pages 13-18 and Figure 6: Each time I read through this section and view Figure 6, the same question arise: how are the precipitation water contents obtained from the radar reflectivity field? Granted, your method is described in detail later, but at this stage the reader may be confused. Thus, I recommend inserting a sentence or two at the end of Section 3a (and possibly a box Figure 6 – see below) stating that these methods will be discussed in Section 4a.

Added a sentence clarifying the application to radar observations at the end of section 3a and put a sentence in the caption of Figure 6.

4. Page 16 and 23: Please clarify what you mean by “...magnitude of saturation”. I think you are implying a spectrum ranging from sub-saturation through super-saturation, but one could easily be confused since the term “saturation” often implies a single state described as 100% relative humidity.

We added this clarifying sentence to page 16...

“Put another way, the algorithm is only dependent on knowing if precipitation is being produced not on the precise value of precipitation production.”

Also added some clarification to a similar sentence on page 23.

5. Page 22: For clarity, define total precipitation water content (q_p) as the sum of the liquid water content (LWC) and ice water content (IWC).

Done.

6. Page 23: Replace “...a composite high-altitude () dropsonde...” with “...a composite sounding derived from ten high-altitude () dropsondes...” for greater clarity.

Done.

7. Page 23 and Figure 13: I find the 3D imagery difficult to interpret, much less

compare to Figure 2. In particular, the altitudes of the heating/cooling maxima are nearly impossible to determine, and the azimuthal distributions are not very clear. I recommend converting these to top-down images that show the horizontal structure of the vertically averaged heating/cooling. You could then add a second field to each panel denoting the altitude of the heating/cooling maxima at each grid point.

Figure changed due to similar suggestions from the other reviewers. Figure now shows the following: vertically averaged horizontal field of latent heat along with the azimuthal mean vertical profile of heating at the RMW for each pass. Text is updated to reflect new figure and a discussion was added.

8. Page 25: Your retrieval method assumes a horizontally uniform density profile, yet the eyewall contains strong thermodynamic (and thus density) gradients. Did you test the retrieval's sensitivity to such density gradients or just different horizontal uniform density profiles? Please clarify in the text.

Density only appears in the computation of saturation and perturbations from the composite sounding have a small effect. We added a parenthetical insert to acknowledge this, "(perturbations to the density profile in all directions had a small effect on calculations)".

Temperature and pressure appear in the calculation of the magnitude of latent heat release but do not require horizontal gradients. Sensitivity to the thermodynamic information was small for the magnitude calculation. This is discussed extensively in the manuscript.

9. Page 29-32: The summary and conclusions section could be streamlined to further reduce manuscript length.

We streamlined the summary and conclusions section a bit, but this section is only 2 – 2.5 pages anyway. Removal of the EDOP information will help with length.

10. Figure 1: Should be removed if all EDOP discussion is removed (see above).

Done.

11. Figure 4: What about the quasi-linear "spikes" in the scatter associated with high precipitation production for both warm and cold processes?

This is more obvious in the ice phase precipitation and is probably associated with some feature of the microphysics scheme (there may be quite a few conditional statements in the code that act to organize the output into quasi-linear clusters). Either way, we focus on warm rain processes so it's not a big issue.

12. Figure 6: You may wish to add a box between the Doppler grid volume and Equation (2) showing the conversion of radar reflectivity to liquid/ice water content.

We added a sentence to the caption and in the manuscript. See comment 3.

13. Figure 14: Should be removed if all EDOP discussion is removed (see above).

Done.

14. Figure 15: My version is not very clear (it looks scanned), please contact the author and obtain an original.

I actually made that plot. It should look clearer in encapsulated postscript or pdf format for the final production.

Responses to reviewer C

Major Points

(1) Title and direction of the paper: The title needs to reflect what this paper is really about. The work depends upon EDOP measurements of updrafts in eyewalls from other TCs, deep soundings from several TCs, model runs from Bonnie, vertical velocity data from many TCs, radar and LWC observations from Katrina besides observations from Guillermo. Wait till part II when you apply the new scheme to Guillermo to put the name into the title. Try something akin to: “A new latent heat retrieval scheme for hurricanes. Part I: Methodology”.

We think the title is fine as it stands. We have Guillermo in the part I title because the end result of the retrieval algorithm is to apply it to radar observations of Hurricane Guillermo, as will be further demonstrated in part II. The other data sources are used to assist the main algorithm. The reader will understand this. Also, the title the reviewer suggests isn't very different than what we already have minus mentioning Guillermo.

(2) The abstract does not describe the paper fairly. All the rather unanticipated steps are glossed over. No one would know that they are going to see data from so many different sources. It mentions Doppler estimates of w and a model when there are several other ingredients that are crucial to the story. In fact it stresses the Doppler retrieval suggesting that we will see the w field which drives the latent heat pattern. Actually the w retrieval was done in Reasor et al. (2009) and the fields are simply applied here.

We don't agree. We can't explain all the details in the abstract, that is what the paper is for and readers understand this point. The abstract provides all the main bullet points and does indicate that other data sources besides radar will be used, “... (2) identifying algorithm sensitivities through the use of ancillary data sources”.

(3) How is the scheme supported by simply placing it in a model? One would think that

the scheme might be used in lieu of whatever the model did (it needs latent heat release too) to see if the TC forms and behaves in a realistic fashion, but this was not done.

Ideally, we would like to have measurements of temperature, pressure, precipitation processes and winds on a high-resolution grid over the entire storm every 30 seconds or so for about 24 hours. Of course, this does not exist and so as a useful alternative, model data is used that does have these features. There are no errors in the model budgets and the contributions from each term can be determined exactly. It is difficult to do budgets from observational datasets due to the various sources of error and unknowns that do not allow the budget to be closed. The model will not replicate the storm as there are many sources of uncertainty inherent in the system, and presumably errors in the model physics. However, the Bonnie simulation was shown by Braun et al. (2006) and Braun (2006) to reproduce several observations of the storm. In addition, several previous authors have found that using a model to test radar retrieval algorithms is useful. This discussion is already in the manuscript: “Although the simulated TC does not replicate the observed storm, the dynamically consistent nature of the model budgets allows the assessment of the qualitative and, to some degree, quantitative accuracy of the method. Gao et al. (1999) used numerical model output to test the accuracy of a Doppler radar wind retrieval algorithm and found errors (see their table 1) that are consistent with those computed from *in situ* data using a similar retrieval algorithm (e.g. see table 2 of Reasor et al. 2009). More real cases are needed to determine if the quantitative aspects of the Gao et al. (1999) results are valid, but the qualitative accuracy appears robust.”

The reviewer mentions in “Minor points” that “the Gao material has no link to the determination of saturation does it? If not then you can trim this, it is an aside”. However, based on the major comment above, I believe the Gao material *is* useful for the reader to gain some confidence in the practice of using model output to test retrieval algorithms. It is by no means perfect, but it does provide a useful proving ground. We think the quoted paragraph (see above) from the manuscript does a reasonably good job of communicating this message. We also added this sentence to the manuscript at the end of the quoted paragraph, “As a result, we believe that testing the latent heat retrieval algorithm in the context of a numerical model provides a useful first step towards a reliable product.”

The second part of the reviewer’s comment is interesting. I encourage him/her to read the results of part 2: it presents modeling results on this exact topic.

(4) The manuscript suffers from casually using data that are not well supported. Some of these steps require a leap of faith on the part of the reader. Examples are:

(a) Comparison of the LWC and dBZ. No discussion of LWC errors which are large. Volume that radar sees and what the PMS probes sample are different.

(b) Use of GPS sondes to discern eyewall thermodynamic structure. Here what the sonde fell through matters but is not discussed. Later satellite observations are mentioned but how they are used is not presented.

(a) Errors in LWC values derived from reflectivity can be large. This was stated in the

introduction. We clarified this in section 4a: “Note that relationships between radar reflectivity factor and water content parameters are not unique and therefore, uncertainty in Q_{net} will exist. This uncertainty is similar to rainfall rate (discussed in section 1) with random errors as large as a factor of four (Doviak and Zrnich 1984).” We also already have details on the probe vs. radar sampling issues: “The cloud particle data are averaged over a period of 6 s in an attempt to match the sampling volumes of the particle probe and Doppler radar pulses (Robert Black, personal communication).” Note that we only care about the condition of saturation in our algorithm and not the magnitude, so some of the error is reduced.

(b) Yes, we do discuss what the sondes fell through and also discuss the satellite observations. Here is the discussion in the manuscript: “To approximate the thermodynamic structure, a composite sounding derived from ten high-altitude (using NASA aircraft that fly at altitudes of 10 and 20 km) dropsondes representative of eyewall convection in TCs is utilized. The storms sampled were: Hurricane Bonnie (1998), Tropical Storm Chantal (2001), Hurricane Gabrielle (2001), Hurricane Erin (2001) and Hurricane Humberto (2001) yielding ten independent thermodynamic profiles of eyewall convection. The sampling of eyewall convection is verified using winds and relative humidity from the dropsondes as well as satellite (infrared and passive microwave) observations.” We didn’t specifically say we were looking for high winds and high relative humidity with clouds appearing on the satellite images, but we think the reader will understand.

(5) What updrafts are saturated? I think horizontal scales matter a lot here, not just the magnitude of the updraft. If the updraft exists for 1 km length scale or more I’ll wager that you are at saturation. For smaller length scales or corresponding periods the cooled mirrors, depending on how they were tuned by the flight director, may not have time to reach saturation. The overall result seems that the authors have selected an extremely high value of w (5 m/s) where they assume saturation. (Note that the entire updraft in Rotunno and Emanuel’s model would be unsaturated if one made saturated versus unsaturated decisions based on w alone!) I’ll bet that updrafts greater than 0.5 m/s for more than 5 seconds with at least one second reaching 2 m/s are probably saturated. What would be the impact if you assumed a lower threshold than 5 m/s? Did Eastin et al. (2005) use the FSSP or J-W sensors?

The raw flight-level data in Eastin et al. (2005) is recorded at a 1 Hz rate and processed into 0.5 km radial bins. The convective scale vertical velocities defined in Eastin et al. (2005) and shown in Fig. 3 (along with relative humidity) clearly show that many of the updrafts are not saturated. Our threshold of 5 m/s for saturation is motivated by this figure and results from the Bonnie numerical simulation. We state in the manuscript: “This threshold should only be used as a guide as updrafts likely do not obey strict rules, but rather evolve through a continuum. Furthermore, the saturation threshold has uncertainty: the observational data shown in Fig. 2 has a small sample size and the model statistics are likely dependent on grid spacing and parameterized physics.” Note sure what sensors Eastin et al. (2005) used but they went through extensive error characteristics in their study and the data should be reliable.

(6) The scheme neglects the latent heat of fusion (end of page 17). Ice processes, however, have been shown to have an impact on TC structure (Lord et al .1984). I hope that this scheme is compared against one with ice to demonstrate the efficacy of the proposed scheme.

We only consider warm rain microphysics for several reasons: (1) warm rain heating is easier to understand and compute than mixed phase heating, and (2) the total latent heat budget is dominated by warm rain processes (this was discussed in the manuscript and the work of Tong et al. 1998 was referenced. Their study focused on Florida deep convection, which is probably not too different than deep convection in a hurricane. The hurricane modeling study of Zhang et al. (2002) was referenced that also came up with a similar conclusion). Including mixed-phase heating would, of course, add some heating in the upper-levels but this would be of secondary importance. As far as the dynamic response of the different profiles in hurricanes, warm-rain heating will have a greater impact because it is coupled to the high vorticity in the mid to lower levels. Shifting the peak heating in altitude does have an impact in balance models as described in Hack and Schubert (1986). We added a sentence in the introduction clarifying the results of the Zhang et al. (2002) study and mentioning that mixed phase processes will alter the vertical heating profile somewhat.

(7) The tail radar from the WP-3D is used to estimate the precipitation field. What range from the radar was accepted, what choices are made about what part of the field is attenuated? This discussion might be better in part II. Should at least mention what Reasor et al. (2009) did because it is so important.

The Doppler analysis is from Reasor et al. (2009). They used the same domain and did not correct for attenuation. Note, the center of action is the eyewall at ~ 30 km radius from the radar on average. There will be attenuation here, but it will not be as bad as near the domain boundaries. Note that for some passes, two aircraft were used to construct the radar analysis which will help reduce attenuation effects. We added some discussion in the manuscript about this issue: “The TA radar reflectivity field used in this study has not been corrected for attenuation. We focus our attention mostly on the inner portion of the domain (the eyewall, which is ~ 30 km from the radar on average) to minimize these effects. Note, for many passes, two aircraft were used to construct the radar analysis which will help reduce attenuation effects (see table 1 in Reasor et al.2009).” We hope to apply the latent heat algorithm to attenuation corrected fields in a future study.

(8) Large uncertainty in the heating due to weak updrafts (156%) coupled with the observation that the majority of the upward mass flux is in the weak updrafts suggests that getting the weak updrafts right is very important. This uncertainty gets glossed over in the conclusions. Recommend that you discuss this more near the end of page 31. It seems that you are suggesting that if one gets the 5 m/s updrafts right then the latent heat estimate will be ok.

Depends on what the reviewer means by “weak updrafts”. The uncertainty estimate of 156% was for a 1 m/s updraft (this applies to a single updraft, not the uncertainty in the retrieval for the system). We stated in the introduction that “...full-physics modeling studies (Braun 2002) and observational composites (Black et al. 1996) show that small-scale, intense convection contributes the largest percentage of the total upward mass flux (~ 65 % from updrafts stronger than 2 m s^{-1}). “ For updrafts of 1 m/s or less, Braun (2002) shows that these vertical velocities only contribute 15 – 20 % of the total mass flux in the eyewall. This is consistent with the observational study of Black et al. (1996). There are a lot of small (1 m/s or less) updrafts, but they don’t carry the majority of the mass flux. However, 20 % could still be important to get correct (some of this will be discussed in part 2). We have added a sentence discussing the above on the page in question... “Even though errors in the vertical velocity can lead to large uncertainties in the latent heating field for small updrafts/downdrafts ($|w| \leq 1 \text{ m s}^{-1}$), in an integrated sense the errors are not as drastic. Furthermore, the majority (65 – 85 %) of the upward mass flux in TCs come from updrafts greater than $1 - 2 \text{ m s}^{-1}$ (Braun 2002; Black et al. 2006), which have smaller errors.”

(9) All the EDOP material – really not relevant to what is going to happen in Guillermo is it? Fig. 1 and 14 seem like extras that are not crucial to your story.

EDOP info removed.

Minor points: page on pdf file and line from top or bottom (-) given for reference

3, 12: of the structure ...to... of its’ structure

Done. Should be “ its ”, not “ its’ ”.

3, 15: to the structural characteristics ...to... to its’ structural characteristics

Done. Should be “ its ”, not “ its’ ”.

4, 9: a higher resolution than what? (other older satellite studies, not the radar derived work)

Higher relative to the satellite estimates just previously stated at 25 km. Should be fine as it stands.

4, -1: 4.3 km on a side? We knew that that is too coarse to resolve convection based on the in-situ measurements by Jorgensen et al. (1985), Black et al. (1996), and Lucas to name a few.

Added Black et al. (1996) reference.

6, 10: I would say that a comprehensive retrieval would show the vertical velocity fields and how they were estimated but this has been done by Reasor et al. (2009). I would vote

for a different goal....how about the application of a new latent heat scheme...later to be applied to the dBZ and w fields of Guillermo.

The retrieval is actually much more comprehensive than the earlier studies of Roux (1985) and Roux and Ju (1990). We go into much more detail including error characteristics. From the manuscript, "The goal of the first part of this work is to perform a comprehensive, high-resolution, 4D, airborne Doppler radar retrieval of the latent heat of condensation in a rapidly intensifying TC. New additions to existing retrieval methods will be highlighted including detailed error characteristics." This is the goal of the work and we follow through on this promise in the manuscript. The reviewer's "new" goal for the paper is essentially the same as what is already written. A far more detailed application will be in Part II of this paper. Note that including this application with the methods would grossly exceed the AMS page limits.

7, 1: the first surprise, what is the ER-2 doing here? Did the ER-2 fly in Guillermo, no.....

EDOP info removed.

7, 10: The long-track...(this sentence needs a rewrite) 100 m is at 20 km altitude, .300 m at 10 km altitude and 550 m at the surface?

EDOP info removed.

7, -4: seems odd that you need a cardinal heading...why?

EDOP info removed.

7, -2: use tail (TA) for the uninitiated the first time.

TA is already defined above this point in the manuscript.

8, 12: the authors fail to provide some info here: How far away from the aircraft will they accept data? Attenuation through an eyewall has what impact? Aircraft level in Guillermo is?

This particular page was intended to provide some general info on the P-3 TA radar; specifics on inclusion of data and aircraft heights for the Guillermo case are presented on page 9 (just below the general info). We did, however, insert a sentence on attenuation on this page as requested by the reviewer since this is more general..." Additionally, the attenuation of the beam at 10 GHz through strong convective cores can be significant."

9, 1: why the top 38%? (I could guess but I would rather you tell the readers)

EDOP info removed.

10, 2: oddly here are some of the details about the WP-3D radar (about two pages later

from where one expected to see it). Are you going to accept data 60 km from the radar...this strikes me as extreme, prior studies used 20 to 30 km. The F/AST technique does not resolve updrafts that are 4-5 km very well. If the updrafts are this scale what % of the w signal can you expect to resolve? Wouldn't this error be your biggest issue?

On this page, we are presenting some details of the radar analysis specific to Guillermo. The previous info on the TA radar is more general.

The Doppler analysis is from Reasor et al. (2009). They used the same domain. Is the reviewer concerned about attenuation? Note, the center of action is the eyewall at ~ 30 km radius from the composite center of the radar. There will be attenuation here, but it will not be as bad as near the domain boundaries. We added this discussion on the page in question: "The TA radar reflectivity field used in this study has not been corrected for attenuation. We focus our attention mostly on the inner portion of the domain (the eyewall, which is ~ 30 km from the radar on average) to minimize these effects. Note, for many passes, two aircraft were used to construct the radar analysis which will help reduce attenuation effects (see table 1 in Reasor et al.2009)."

Not sure what "% of the w signal" the analysis can resolve, but detailed error characteristics are described in section 4b. Yes, errors in vertical velocity are a big issue. This topic is covered extensively in the "Observations and errors" section (section 4b).

11, -5: ...the dynamically consistent nature of the model budgets (aren't they all consistent?) allows for the assessment...what would really demonstrate if the new scheme was an improvement was if the run with this scheme simulated the TC better, all other subroutines in the model being kept the same.

Yes, model budgets should be all dynamically consistent, but observational budgets are usually not. That is why we specify in the sentence in question.

Second part of the reviewer's comment: this is the topic of part II of the paper.

11, -4: the Gao material has no link to the determination of saturation does it? If not then you can trim this, it is an aside

See response to major point 3.

12, 3: stating a Met 101 class point here – try a reword if you feel you must remind us about phase changes

We believe the reviewer is talking about this sentence: "Therefore, an important question is: does a threshold of vertical velocity exist where saturation and the release of latent heat can be assumed?" We don't think there is anything wrong with this sentence. It mentions both saturation and the release of latent heat just to be clear to readers.

12, 7: no references are required for such a general statement. People recognized latent

heat releases' impact long before Scott and Matt were born. It reads as if these two gentlemen discovered this.

“Above 5 m s^{-1} , vertical accelerations are dominated by local buoyancy forcing while below 5 m s^{-1} various physical processes may play a role in the evolution such as perturbation pressure gradient forces (that are not generated by heating) and turbulence (Braun 2002; Eastin 2005).” References are required for these statements. Most earlier studies believed that very little if any vertical acceleration occurred in the TC eyewall.

12, -8: the WP-3D has trouble identifying small clouds with the mirror dew point sensors; I would expect it not to be very discerning for times less than 5 seconds depending on who set the response time of the sensor. If a weak updraft exists for more than a few hundred m in the horizontal I'll bet it is shortly to become saturated, smaller turbulent eddies may not be. Did Matt count updrafts of a certain minimum time-scale? If he simply counted drafts of even a few seconds then many of these features could be saturated.

The raw flight-level data in Eastin et al. (2005) is recorded at a 1 Hz rate and processed into 0.5 km radial bins (4 seconds with $\sim 125 \text{ m/s}$ ground speed). Thus, the convective scale vertical velocities (and relative humidity) defined in Eastin et al. (2005) have a resolution of 0.5 km. Despite this fact, Fig. 3 clearly shows that many of the updrafts are not saturated. Not sure what sensors Eastin et al. (2005) used but they went through extensive error characteristics in their study and the data should be reliable.

12,-3: are these unsaturated updrafts continuous in z for more than 1 km or so? Hope they aren't, otherwise the model has some strange structures.

We only counted grid points here, not coherent updraft structures. I don't think the model has any “strange structures”.

13, -4: should at least briefly explain what negative mixing ratios are

The negative mixing ratios are due to numerics (advection discretization errors). It would take too long to explain this in the manuscript so the appropriate reference is given (Braun 2006).

13, -2: awfully big convective-scale...more like meso-gamma and bigger than all but the top 1-2% of updrafts

We removed the definition of the scale since we got the same result using smaller definitions as well.

14, 6: nice that the terms can be combined, but in the end you can't do it...so why mention as it becomes an aside

Part of this paper is to present a new retrieval algorithm that can be applied by users in

the future. We are illustrating how someone might reduce the errors in the budgets, which we think could be useful for future applications.

15, -4: when could the model be saturated without a positive Q_{net} ?

Sentence in question: “In summary, Figs. 3 and 4 demonstrate that by acquiring information on Q_{net} and determining where $Q_{net} > 0$ (net production of precipitation), we are able to distinguish where the air is saturated, which is required before the release of latent heat can take place.” We don’t see anything wrong with this sentence. Note that we try to remind readers of the association of Q_{net} and saturation throughout the paper to be clear.

16, 8: what about the latent heat release for regions where there are no precipitation-sized particles? Are you arguing that where precipitation-sized particles are the only place where there is saturation? What about where there is rain but it is subsaturated (below cloud base)?

We are using a precipitation radar (10 GHz), which will not be able to detect cloud-sized particles. However, the vast majority of latent heat release will be associated with precipitating particles for which warm rain dominates (see major point 6 response). The algorithm solves for saturation so it should (in principal) be able to discern unsaturated air below cloud base where it exists (although there will be errors from various sources-- as mentioned in the manuscript).

16, 11: what is the maximum range that you have accepted data from the aircraft? What have you done about obvious attenuation situations?

See response to major point (7).

Fig. 6: the tail radar detects precipitation and the wind field - then you solve for the equation (here don’t you have an issue at cloud edge where q_p is zero so all entrainment would make the second term on the right negative?) also how does the radar provide the storage term given that you sample a volume only once? Then Q_{net} is (+) or (-)volume is considered saturated and latent heat release determined chiefly by w and rate of change of q_s with height. Well, the top of page 18 really doesn’t describe the flow chart. Later on p. 18 you discuss the fact that the storage term period (34 min) is so long as to not affect the precipitation budget. You aren’t applying a parameterization scheme for a model time to the observation period are you?

Cloud edge can be a problem. This is mentioned in the manuscript. The radar samples the same volume ~ 10 times during the Guillermo obs (~ 5.5 h). This time step is much too large to estimate the storage term so we used the numerical simulation with high temporal output to examine this issue. We find a way to parameterize the storage term in observations based on the physics present in the model simulation (advection and the divergence theorem).

The flow chart is a summary of the information that has been discussed already in that section. All the details about the equations, variables and explanations are given in the section and mentioned in the flowchart caption.

18, -5: now you celebrate the importance of the storage term for shorter periods for a model – what has that got to do with the Guillermo obs?

See response given above.

19, 6: you seemed surprised that Q_{net} would be a large term...why?

We want to convince the reader that the signal-to-noise ratio is significant here, which is important for the algorithm.

19, 9: u is total wind is it not?

No, it is storm-relative.

19, 11: an explanation of morphing would help here (a sentence or so)

We changed the sentence to, “This relationship indicates that morphing (advecting precipitation features forward in time; Wimmers and Velden 2007) the radar reflectivity and derived precipitation fields using the Doppler wind analyses to generate a storage term tendency shows promise.”

19, -5: got references for the radar studies?

These were completed by the first author. Sentence reads: “The storage term values produced through the model-based parameterization are very similar to those calculated by the authors using ground-based radar (refresh time of ~5 minutes) and P-3 LF radar (refresh time of 30 s) observations of mature TCs (not shown).”

20, 2: what approximations?

The approximations for calculating Q_{net} described above this sentence. This should be clear for the reader (its also explained in the figure captions).

20, 6: does the inclusion of the storage term reduce Q_{net} ?

It can, but we are talking about the error: “...reduces the error in Q_{net} by ~ 16 %..”

20, 10: a nice reduction for the short time scale of the model...but for Guillermo won't it be far less?

Should be reasonably accurate for Guillermo. There is no “model offset” in the real atmosphere. The impact of turbulent diffusion in the real atmosphere is uncertain, but the

model showed this term was smaller relative to the other terms. Again, we are using the model as a proving ground because we don't have the observational data necessary to do these budgets.

21, 6: when would something be quantitatively significant and not physically significant?

Plenty of times. For example, suppose one can improve the temperature structure in the hurricane far from the center (maybe 500 km radius) by a quantitatively significant amount (maybe 2 K or 50 % improvement). Well that is great, but that quantitatively significant improvement will likely not be physically significant for structure and intensity changes occurring near the eyewall of the storm. These changes are governed, for the most part, by inner-core ($\sim < 50$ km) dynamics and latent heat release.

22, 6: ...data are averaged....

Changed. Thank you.

22, 8: are you using the tail radar? How do you compare the in-situ volume with the tail given that you are not sampling the same volumes at any given time? Aren't side lobes very close to the aircraft an issue?

Yes, tail radar. From the manuscript: "The cloud particle data are averaged over a period of 6 s in an attempt to match the sampling volumes of the particle probe and Doppler radar pulses (Robert Black, personal communication)." This attempts to align the two samples in space/time, but is not perfect. Yes, side lobes can be an issue although HRD folks (Robert Black) didn't indicate too much concern over this issue.

23, 6: for a 2 km model grid just how much of the volume has 5 m/s or greater?

Small relative to the weaker updrafts. Don't know the actual percentage, however.

23, 9: wow – this argument about a dropsonde being representative of the eyewall is a reach. Did it fall in the updraft, downdraft, or some of both? Where did the sonde go with respect to a reflectivity maximum?

Not just one dropsonde, a composite dropsonde representative of the eyewall. Here is the discussion in the manuscript: "To approximate the thermodynamic structure, a composite sounding derived from ten high-altitude (using NASA aircraft that fly at altitudes of 10 and 20 km) dropsondes representative of eyewall convection in TCs is utilized. The storms sampled were: Hurricane Bonnie (1998), Tropical Storm Chantal (2001), Hurricane Gabrielle (2001), Hurricane Erin (2001) and Hurricane Humberto (2001) yielding ten independent thermodynamic profiles of eyewall convection. The sampling of eyewall convection is verified using winds and relative humidity from the dropsondes as well as satellite (infrared and passive microwave) observations." We didn't specifically say we were looking for high winds and high relative humidity with clouds appearing on the satellite images, but we think the reader will understand. Note that the

temperature structure did not change much in the eyewall and furthermore, our latent heat retrieval algorithm is fairly insensitive to the details of the thermodynamics (discussed in the uncertainty section).

24, 6: calling any of passes 2 through 5 as symmetric is a stretch. Latent heating remains insignificant to the NW and W through out almost all the frames of Fig. 2.

The sentence in question has been altered a bit to: "...revealing a slightly more axisymmetric distribution of convection." We think this is an accurate statement. Note they are more axisymmetric *relative* to the other passes.

24, 10: during the intensification period the eyewall also decreased.

This section has been changed due to a new figure. Not really sure what the reviewer means by this statement. However, we do mention a decrease in the azimuthal mean heating towards the end of the period.

25, 2: but you already did the saturation state error analysis. Might want to rethink how you organize this paper.

From the manuscript: "There are two main calculations in the retrieval that require error analysis: the computation of the saturation state and the magnitude of the latent heat. The approximate errors associated with determining saturation are analyzed in section 3b and thus, the focus here is on the magnitude of the latent heat fields."

25, -3: people have long known, simply by comparing the horizontal resolution for the tail Doppler to in-situ estimates of drafts, that the smaller scale drafts are not well represented by the radar. See the in-situ measurements by Jorgensen, Zipser and LeMone and compare to the eyewall or rainband structures reported in the literature with the Doppler for eyewalls, rainbands, and convective cells.

Added the Marks et al. (1992) reference.

28, 1: calling the Guillermo dataset sampling uncertainty is a little misleading. We know it is one storm but sampling uncertainty it is not.

This constitutes sampling uncertainty in the context of Guillermo. Obviously sampling errors must be calculated for each application, and cannot be expected to apply to other studies of different systems. The aircraft gathers one snapshot of the storm every 30 minutes or so for about 5.5 h. Quite a bit of evolution occurs during those snapshots on the convective scale and out side of the 5.5 h window of observations. It is very likely there are important convective events the aircraft have missed.

30, 4: again, show me a situation where there are strong w's well above the boundary layer (away from mountains) and I'll bet there is buoyancy. You state the obvious.

We believe it is better to be explicit so points are as clear as possible.

30, 12: this is Eastin et al. result, not this paper. You have not conducted the analysis so rephrase to give Matt and company the finding....you are going to apply it.

Added, "... (courtesy of Eastin et al. 2005)..." to the manuscript.

30, 15: I think everyone knows you need saturation to have latent heat release

They should. However, we believe it is better to be explicit so points are as clear as possible.

30, -7: what else would explain the precipitation production?

Is this a rhetorical question? We believe it is better to be explicit so points are as clear as possible.

31, 4: this is for the model or Guillermo?

From the model. We clarified this in the manuscript..." A parameterization for the storage term based largely on the tangential advective flux of precipitation (a consequence of the divergence theorem) was developed using output from the Bonnie numerical simulation that shows promise for reducing the steady state uncertainties in TCs."

31, 5: not sure that it is a consequence of the DIV theorem.....

The divergence theorem is a statement of conservation (for say a grid cell) where the sum of the sources and sinks of a quantity (the divergence of that quantity) is equal to the net flow across the boundary of that cell. For a grid cell in a hurricane, the flow across the cell boundary is dominated by the tangential flux. Referring to our precipitation continuity equation...

$$\frac{\partial \rho q_p}{\partial t} \cong -\nabla \cdot (\rho q_p \mathbf{v}) - \frac{\partial (\rho q_p w)}{\partial z} + \frac{\partial (\rho q_p V_t)}{\partial z} + \rho Q_{net}$$

The Q_{net} term is the sum of the sources and sinks of precipitation. This term plus the fallspeed term result in a small value relative to other terms. This leaves the flux terms on the RHS for which the tangential flux dominates and thus, a relationship between the storage term and the tangential flux exists. This relationship is not perfect, but a large percentage of the variability in the storage term can be explained by the tangential flux (at least in the model). This is essentially a consequence of the divergence theorem.

31, 11: prior studies that estimate LE release through precipitation have their errors dominated by the estimate of the precipitation field itself. In your case your main source

of error is controlled by the w field and when to assume saturation. You may want to tell the reader the explicit gains you have made – showing heat release as a function of z .

Already mentioned in the summary/conclusions section.

31, -4: how does the EDOP analysis help with Guillermo estimates of w ?

EDOP info removed.

32, 2: It seems that you will apply this scheme to a model to see if the new scheme does a better job with Guillermo's RI. You'll initialize with Guillermo's w field. Now will you use the model's regular scheme as well as the new one to see if you get a different and better result?

See part II.

Fig.1: I suspect that the EDOP estimates have some errors in it. A 5 m/s updraft at about 400 m altitude would demand a convergence approaching 50 m/s over about 4 km width.... $-1.25 \times 10^{-2} \text{ s}^{-1}$; seems quite unrealistic given that there is virtually no buoyancy at this level.

EDOP info removed.

Fig. 2: Cells to the west never achieve the same heights or rain rates as cells to the east and north of the circulation center.

Is this a comment or observation? Not sure.

Fig. 5. is this for a cell that is mature? (Surprised by the production of rain below 2 km)

Yes. Mentioned in manuscript.

Fig. 13. A pretty fig. that conveys little quantitative information. The reader won't be able to discern the height where the LE release is, but we already know it is largely controlled by dq_s/dz and w , therefore it will be a maximum in the lower troposphere. We already know from prior work (Eastin et al, Reasor et al., Sitkowski and Barnes) that the west side of Guillermo is inactive.

Figure changed due to similar suggestions from the other reviewers. Figure now shows the following: vertically averaged horizontal field of latent heat along with the azimuthal mean vertical profile of heating at the RMW for each pass. Text is updated to reflect new figure and a discussion was added.

Fig. 14. an aside that is not relevant to Guillermo. You don't need to advertise other projects.

EDOP info removed.