Response to review by D. Raymond:

The authors of this manuscript make parcel trajectory calculations for the failed tropical cyclone Gaston (2010), based on ECMWF analyses that included dropsonde data from the PREDICT project. There is general agreement that Gaston decayed after the second PREDICT mission on 3 September 2010 as a result of ingestion of dry air. However, the current manuscript as well as a couple of other papers cited by the authors further assert that the decay of Gaston between 2 and 3 September was also due to the incorporation of dry air.

Gjorgjievska and Raymond (2014; GR2014, cited in the manuscript) do not dispute that dry air was instrumental in the decay of Gaston after 3 September. However we also demonstrate that a more subtle process was likely occurring in the 2-3 September interval that led to the subsequent flood of dry air invading Gaston.

The first hint that dry air did not affect the convection in Gaston prior to 3 September comes from in figure 3d of GR2014. (Note that due to an unfortunate transposition error, the images for figures 3 and 4 are switched, so the image for figure 3 is shown with the figure 4 caption.) This figure demonstrates that the relative humidity averaged over a 4 by 4 degree box centered on roughly on the 5 km vortex center changed very little between 2 and 3 September. The main difference is an increase in the relative humidity near the 5 km vortex center in the 7-9 km range between these two dates. (See also figure 5 of GR2014).

With due respect, just because the area-averaged RH changes little on the 4 by 4 degree box moving with the system does not unequivocally imply that the inner pouch was isolated and protected from its environment. As an example, the convection that was observed during Gaston 1 would be expected to moisten the middle levels inside the pouch; if dry air was intruding into the system (which we show to be the case in Figure 6), this convective moistening could work to offset the dry air entrainment. Given the demonstrated intrusion of dry air into the pouch, we would expect the convection to be negatively impacted by the dry air according to the findings of Kilroy and Smith (2012).

Figure 6 of the current manuscript shows the analyzed equivalent potential temperature at 500 hPa (approximately 5 km) on 1-3 September. There is indeed a dry tendril of air sweeping around the south and east side of Gaston 1 and 2 (on 2 and 3 September respectively), but the actual 5 km circulation centers in the co-moving frame were at (39W, 15N) and (42W, 15N) on these two days, i.e., 2-4 degrees to the north and west of the dry air intrusions (see GR2014 figure 5).

We believe we can explain the apparent discrepancy of the circulation center as a misunderstanding of the timing of the data and the appropriate figures. The dropsondes during the first flight into Gaston (GR2014's Gaston 1) were deployed during 1532-1906 UTC on 2 September, which is closer to 0000 UTC 3 September rather than 2 September. The Reviewer uses our 2 September plots (Fig. 6, middle)

to compare with his Gaston 1 data, but our 3 September plots (Fig. 6, bottom) are actually more appropriate. The GR2014 Gaston 1 position of (39W, 15N) actually agrees well with our 3 September Fig. 6 (bottom) plots. It should be noted that the Gaston 2 flight that occurred late on 3 September (closer to 4 September) is not represented by any of our figures. However, upon inspection of our other analyses not included in this manuscript, we believe that our pouch positions are actually in agreement with the Reviewer on that later flight as well. Considering the circulation's WNW motion, GR2014's Gaston 2 position of (42W, 15N) based upon dropsondes deployed during 1444-1849 UTC 3 September corresponds fairly well with the subsequent ECMWF 0000 UTC 4 September pouch position of about (43.5W, 15.5N).

Bottom line: There is no 2 – 4 degree positioning error of the pouch center as suggested by the reviewer.

Figure 6 of GR2014 shows the vertical mass flux pattern at 700 hPa in Gaston 1 on 2 September. The strongest upward motion (representing deep convection) is centered at (39.5W, 14.5N), or slightly to the SW of the 500 hPa circulation center. There is evidence of downward motion roughly 1.5-2 degrees to the east of this ascent, possibly representing the effects of the intruding dry air. Nevertheless, convection responds to the thermodynamics of 1 air in its immediate vicinity, not to air 150-200 km away, indicating that the convective core of Gaston on this date was still narrowly protected from dry air by the pouch.

GR2014 recognize that asymmetries may alter averages, saying that any way of averaging may give inconsistent results. In at least two places in their mss., they recognize the role that time-dependent non-linear dynamics could have played. The first is when they note that the increase in the mid-level vorticity went against a negative vorticity tendency. The second, is when they note that transport of dry air could have been a factor in the decay of Gaston, however, the pouch (according to GR2014) was apparently closed. The difference in sign between vorticity tendency and actual vorticity evolution is a very strong indication that time-dependent dynamics plays a role. Our study here uses the Lagrangian manifolds as a way to measure the role of time-dependent dynamics objectively, that is, without any sensitivity to the choice of spatial location in averaging.

Comparison of Gaston 1 and 2 with the developing cyclone Karl shows that the relative humidity profiles in the early stages of Karl were very similar to that of Gaston 2. Yet Karl developed into a major hurricane. The most obvious difference between the two cases is that Gaston 1 and 2 experienced SSTs of 28.2 C and 28.4 C respectively, whereas the first 3 Karl missions showed SSTs of 30-30.2 C (see table 1 of GR2014). Thus the SSTs in the early Karl stages exceeded those in Gaston by almost 2 C. In addition, as figure 9 of GR2014 shows, the tropical cyclone heat potential for Gaston was quite small in its initial stages and quite large for Karl.

The GR2014 argument appears to be that Karl had 30 C SSTs and Gaston had 28.5 C SSTs, therefore Gaston did not develop because it had greater *convective inhibition* than Karl. This is a case where the models predicted the SSTs fairly well, unlike Nate (2011) where the upwelling effect of the storm led to errors in the SSTs, yet the models had no trouble predicting development.

It seems highly plausible that there is a feedback between less convection and greater permeability of the pouch boundary, since the vorticity gradients caused by convection and its associated convergence determine the strength of the boundary. In other words, a weaker boundary leads to dry air intrusions, which further weakens the convection.

GR2014 do not explicitly consider the aforementioned feedback between dry air intrusion and weakening of convection. Their focus is rather on the supposition that the SSTs played an essential role, and only then because Karl happened to have higher SSTs and did develop.

Our study has focused on the proposed potential feedback effect and how a weaker boundary further hinders development. This line of study is different from the GR2014 line (and the reviewer's line above) that lower SSTs underneath ex-Gaston's pouch were the most obvious difference between it and the developing Karl. It is noteworthy to point out that in numerical cyclogenesis studies, an SST of 28.5C is not subcritical provided, inter alia, the initial vorticity is favorable. From a larger-scale perspective, however, the mid-level vorticity for Gaston was unfavorable; there was significant vertical shear. We can't absolutely say that the vertical shear would or would not have broken the pouch boundary if convection had been stronger. What we do show is that dry air intruded, and the dry air was not moistened until it entered the pouch. That suggests that the dry air played at least some role in limiting the convection.

Summary: GR2014 do show ocean tropical cyclone heat potential, in addition to noting the minor SST difference between ex-Gaston and Karl. We acknowledge that there is a potential feedback between the ocean heat content and dynamics of the pouch boundary, but it seems unreasonable to dismiss lateral intrusions when: i) many of GR2014's results are possibly artifacts of spatial averaging and ii) the actual dynamics suggest that Lagrangian versus Eulerian temporal sampling are so different.

Figure 10a of GR2014 shows that the environment of convection in Gaston 1 had strong convective inhibition near 2 km, and that considerable energy had to have been expended by the convection in breaking this inhibiting layer in the convective region. Such an inhibiting layer did not exist in the vicinity of convection in Karl 3 (11 September 2010), as figure 10c shows. The inhibiting layer in Gaston relative to Karl is almost certainly related to the lower SST experienced by Gaston.

Convective inhibition (and the corresponding acronym `CIN') is mentioned only

once in the GR2014 mss. on page 3068 (section 4, right column, middle paragraph) in a Background section reviewing Raymond and Sessions 2007 and its argued application to the real world.

On that page of the GR2014 mss., CIN is mentioned in the context of 'parcel buoyancy' below 2 km altitude. The term `parcel buoyancy' is used four times in this paragraph and this is the only place in their mss. where the term parcel buoyancy is mentioned. (Parcel buoyancy is not defined mathematically in the GR2014 mss.)

CIN is never used in the GR2014 mss. to compare Gaston and Karl. Rather, the non-standard 'instability index' (their Eq. (6), same page, left column) is used as a measure of a *system instability (our words and emphasis)*. Logic suggests that one should not use a non-standard definition of 'CIN' to compare against the CIN from another study using a standard definition.

Traditionally, CIN is defined as the work done required to lift a moist air parcel to its level of free convection (e.g., Emanuel 1994). Since this work depends on the parcel, one needs to take a further step and make the definition unambiguous. In Smith and Montgomery (2012, QJRMS; hereafter SM12), CIN was defined and calculated using the *minimum work required to lift a test parcel to its level of free convection*. Although the minimum work usually corresponds to parcels lifted from the surface, this is not always the case. SM12 calculated CIN defined in this (standard) way for Gaston, Karl and Matthew pre-storms. In particular, the CIN calculated for ex-Gaston on 2 September is shown in their Figure 6 and the CIN calculated for both flights into pre-Karl on 10 September are shown in their Figure 10. After carefully examining the data plotted in these figures, one does not find that the CIN for ex-Gaston is larger than that for Karl. In fact, in the vicinity of the sweet spot, one finds that the opposite is true! As an example, we calculate below the arithmetic average of CIN values within a horizontal radius of approximately 2 degrees from the center of the ECMWF-analyzed sweet spot (indicated by the green curve) in SM12's Figure 6 and the top right plot of SM12's Figure 10 (first flight of Karl). (As a point of clarification, the 2 degree radius circle corresponds roughly with the 4 by 4 degree box averages used by GR2014.)

We find the following results:

Ex-Gaston Average CIN within 2 degrees radius from analyzed sweet spot: 9 + 29 + 0 + 11 + 0 + 0 + 5 + 58 J kg^{-1} / 8 = 112 J kg^{-1}/8 = 8 J Kg^{-1}

Karl 1 Average CIN within 2 degrees radius from analyzed sweet spot: 4 + 18 + 40 + 16 + 14 + 7 + 47 + 36 + 21 J kg^{-1} / 9 = 203 J kg^{-1}/9 = 22.2 J Kg^{-1}

Karl 1 CIN is more than a factor of two larger than that for ex-Gaston.

A similar result is found for Karl 2 (the second GV flight on 10 September).

Karl 2 Average CIN within 2 degrees radius from analyzed sweet spot: 37 + 11 + 33 + 20 + 29 + 1 + 24 + 9 + 33 + 16 J kg^{-1} / 10 = 213 J kg^{-1}/10 = 21.3 J Kg^{-1}

Admittedly, we have chosen a 2 degree radius around the sweet spot and we could have chosen a larger radius or a smaller radius. But the results will not change significantly so as to render ex-Gaston CIN > Karl 1,2 CIN.

Thus, based on the evidence presented in Smith and Montgomery (2012), the hypothesis that ex-Gaston did not develop because its CIN was larger than that of Karl is rejected. Our finding here is in accord with one of the primary conclusions of Smith and Montgomery (p1738) who stated that:

"Even so, the evolution and distribution of CAPE and CIN by themselves did not reveal an obvious distinction between developing and non-developing systems."

As Figure 7 of GR2014 shows, the convective mass flux profile for Karl 3 was vastly different from that of Gaston 1, with extreme top-heavy convection in Karl 3 and extreme bottom-heavy convection in Gaston 1. This resulted in much stronger convergence below 2 km and a corresponding increase in the strength of the low-level circulation between Gaston 1 and Gaston 2 (see figure 3a – shown under the figure 4 caption as noted above). However, strong divergence above 3 km in Gaston 1 resulted in the destruction of an initially strong mid-level vortex, as figure 3a shows. GR2014 argue that the elimination of the mid-level vortex weakened the pouch sufficiently to allow the ingestion of dry air, resulting in the subsequent decay of Gaston. Given that the relative humidity profiles for Gaston 1 and 2 and Karl 3 were nearly identical, as are the parcel buoyancy profiles above 3 km, the existence of strong convective inhibition in the environment of Gaston 1, undoubtedly related to the lower SST, is the most plausible explanation for the dramatic differences between the convection in the two cases. (As noted by the authors of this manuscript, the most extreme convective inhibition, as represented by a trade wind inversion, occurred well to the west of Gaston 1. However, relatively strong convective inhibition, as noted above, existed on all sides of the convective core in this case.)

For the reasons given above, with due respect we do not accept the loose association of lower SSTs and convective inhibition and we think one should not use non-standard definitions of 'CIN' to compare against the CIN from another study using a standard definition. Based on the calculations presented by SM12, and the additional calculations summarized above, we do not find that the dropsonde data supports the reviewer's assertion that the "convective inhibition" was relatively strong on all sides of Gaston 1.

In summary, the evidence for our view of the decay of Gaston before 3 September consists of 2 parts: (1) The relative humidity did not decrease and in fact increased at upper levels between Gaston 1 and Gaston 2 in a region centered on the 5 km circulation center. The convective core was very close to the circulation center in these two cases. (2) The low SSTs and increased static stability near the convective core of Gaston likely had a negative effect on convection in Gaston 1 even if there was technically no trade wind inversion in the convectively active area. This stands out particularly in comparison to Karl 3, in which convective inhibition was weak over the entire region, and for which the SSTs were much higher.

Part of the discrepancy between the results of GR2014 and the current manuscript may be due to the location of the pouch. For both Gaston 1 and Gaston 2, the 5 km circulation 2 centers are on the NW edge of the pouch positions as defined in the manuscript (see figures 5a and 5b in GR2014 in comparison with figure 6 in the manuscript). Furthermore, the convective cores in these cases are much closer to the 5 km circulation centers than to the center of the pouches defined in the manuscript under review (see figure 6a in GR2014 for Gaston 1; Gaston 2 not shown). One can of course define the pouch in accordance with the circulation center at any level one desires, and Montgomery and colleagues tend to define this at 850 hPa (or perhaps 700 hPa in this paper – this is not clear). For reasons set forth in Raymond et al. (2014; Tropical cyclogenesis and mid-level vorticity. Australian Meteorological and Oceanographic Journal, 64, 11-25.) we prefer a higher level, i.e., near 5 km in many cases. Given that the convection tends to occur near the 5 km circulation center on both days, the higher level would seem to be more appropriate in this case.

The nominal center depicted in this mss. is defined as the intersection between the wave trough and the local critical curve at the tracking level (700 mb). We have clarified this point in the revised mss.

The apparent discrepancy between the GR2014 pouch positions and our sweet spot positions has been largely resolved by our response to the reviewer above.

The dependence on a global analysis for very delicate Lagrangian trajectory calculations also raises at least a yellow flag. Analyses incorporate sounding data in competition with model prejudices with opaque weighting factors. Our analyses depend on PREDICT dropsonde data only.

While individual trajectory computations are sensitive to the trajectory integration scheme and the quality of the wind data, Lagrangian coherent structure identification is surprisingly robust (Haller 2002, DOI:10.1063/1.1477449). This study provides a detailed analysis of the wind fields, which over any finite time interval must satisfy momentum conservation within the model. Therefore, for the purposes at hand we feel that the model wind fields provide a better depiction of time-dependent velocities than instantaneous wind fields derived from dropsonde data. As an additional affirmation of the consistency of our methodology employed here, the Lagrangian manifolds that we have computed are in agreement with both the model moisture fields and vorticity fields.

I feel that I am perhaps too close to this whole argument to give an objective recommendation on this paper, so I shall leave that to the other reviewers and the editor. However, though I do appreciate the authors' attempt to represent our position in their manuscript, I would like to see the whole story told, which explains the length of this commentary. Technically, the manuscript is well written, though some of the figures, such as figure 5, are very hard to decipher.

We thank the reviewer for his careful reading of the manuscript, and for his perceptive and thorough review. For the reasons given in our manuscript and in our responses above, we have a very different interpretation of the failed development of Gaston (before 03 September). In our study, we show (1) the pouch is open and vulnerable as early as 1 September (Figure 6), (2) dry environmental air was entrained into the pouch (Figures 7 and 8) as early as 1 September, and (3) vorticity and vertical mass flux decrease with decreasing relative humidity (Figure 4).

We think the data supports the foregoing dynamical interpretation that builds on the new insights by Kilroy and Smith (2012) concerning the negative impact of dry air on vorticity amplification within the pouch of pre-storm disturbances.