Atmos. Chem. Phys. Discuss.,
doi:10.5194/acp-2015-970-RC1, 2016

Atmospheric Chemistry and Physics Discussions
Open Access

Interactive comment on "Controls on phase composition and ice water content in a convection permitting model simulation of a tropical mesoscale convective system" *by* C. N. Franklin et al.

Anonymous Referee #1

Received and published: 18 February 2016

Overview

This manuscript explores a very interesting topic with many outstanding questions, namely the controls on mixed phase hydrometeor properties in deep convection. A large number of sensitivity simulations are performed of an MCS case near Darwin, Australia, for which there are ground radar and aircraft in situ measurements for comparison. Given the unique observational dataset and important topic, the manuscript is certainly worthy of eventually being published in ACP, but there are many major issues that need to be addressed before this can happen, and it will likely take the authors a

Printer-friendly version

Discussion paper

long time to address all of these issues in a satisfactory manner. The most important of these issues are the flawed methodology for comparing very limited observations with very ample model output and the large amount of unsubstantiated assertions that are passed off as conclusions without evidence. These and other issues are discussed in much more detail below, and a number of suggestions are offered that provide paths forward to overcoming the major shortcomings of the manuscript.

Major Comments

1. The comparison of a single sounding with the model sounding is nowhere close to representative of environmental differences between the model and observations. In fact, the observed sounding is a classic “onion” sounding in a stratiform region where the low level air is completely stable and mid level air is dried out because of the mesoscale downdraft. This is not the air that is feeding the system (it thermodynamically cannot be since it is stable), and convective cloud base from lifted boundary layer parcels would be below 1 km, as it nearly always is in Darwin active monsoon conditions. In the stratiform region, where these soundings are taken, the cloud base is typically around the melting level, which is where the soundings approximately show it. The humidity profile will vary depending on where you take the sounding in the stratiform region, so you also cannot draw conclusions about upper tropospheric humidity. The likelihood that the model sounding is in a stratiform region location that is exactly like the one observed is practically zero, so no conclusions regarding model environmental biases can be drawn from this comparison. The winds are also not representative and examination of CPOL radial velocity shows that low-mid level winds are quite variables because of the MCS forming in a trough convergence region and the mesoscale circulations induced by the stratiform precipitation. Therefore, you should remove all conclusions based on comparison of these soundings.

The prior Darwin sounding at 12Z (attached as Fig. 1) before the system initiates shows a classically active monsoon environment and one that is probably similar to the one that the convection develops in 6 hours later, so you can compare that to the model, but it is still not okay to draw conclusions about model environmental biases from one sounding because humidity, winds, and instability are highly variable across mesoscale domains (you can prove this to yourself by plotting them using the model output), so if you choose to include a comparison of 12Z soundings, you should plot a spread of model soundings outside of clouds and precipitation and place the observed sounding in this spread. If the spread covers the one observed sounding, you cannot conclude that there are biases in model environmental representation. Otherwise, simply remove the comparison of observed and modeled soundings. There could be environmental representation biases, but it is nearly impossible to show that given the available observations, and this is not the purpose of the manuscript anyway.

2. The timing and location of this observed MCS are very important for how it needs to be compared to model output. After the initial deep convective stage when convection is most intense, a large stratiform region forms and the most intense convective cells push westward outside of CPOL coverage before the aircraft even begins sampling the system. The aircraft then samples the remnant stratiform region and weak convection that is not representative of the convection that forms the MCS. It is unlikely that the simulations reproduced this lifecycle (in fact Figure 3 shows that they did not), but this lifecycle strongly impacts interpretation of comparisons between model output and observed reflectivity and aircraft observations in some of the figures (i.e., is some of the model error because of a different system evolution in terms of timing and location?). Therefore, the figures showing statistical comparisons would be greatly aided by showing observed and simulated (just pick a representative simulation – the time series show that they have similar evolutions) low level reflectivity during a couple times between the initial intense convection and the decaying stages when the aircraft was making observations.

3. In the paragraph that starts at the bottom of page 9 and continues onto page 10, I disagree with your reasoning that the underestimate in precipitation at later times is a result of drier low-mid levels. First, you can't determine whether they are drier or

not given the available observations, and second, Figure 3 shows that the simulated MCS develops about 2 hours earlier than the observed one. If you shift the simulated precipitation time series to 2 hours later, then the evolution of the precipitation is very similar in the simulations and observations.

In the second part of this paragraph, you state that lack of stratiform rainfall is not caused by excessive evaporation (even though earlier in the paragraph you partly blame drier low-mid level air) and instead blame overly strong convection that detrains too high in the troposphere. This could be going on, but you show no evidence of low biased stratiform rainfall or overly strong convection, so this is purely speculation and should be removed unless you add evidence to support it.

4. In Figure 3d, the satellite retrieved OLR looks incorrect. I checked the satellite observations between 18 and 21Z and they show OLR less than 125 W m-2 covering the entire domain (see attached Fig. 2 for 21Z OLR), whereas your figure shows 160 W m-2. Therefore, your conclusions on page 11, lines 16-24 are incorrect. Perhaps you are averaging over too large of a region for the comparison?

5. Because so many model sensitivities are examined, I don't think that any single sensitivity is given the detail that it deserves to understand the mechanisms behind changes in model output. This leads to a lot of speculation throughout the manuscript without much evidence shown. Some speculation is fine, but the speculation is passed off as facts in a number of spots including in the conclusions. Here is a list of examples: a. On page 10, lines 26-31, your explanation regarding differences in RH profiles between simulations with different ice PSDs may be reasonable, but you do not provide evidence showing riming rates connected to latent heating connected to convective updraft strength. b. On Pages 10-11, you discuss convective entrainment but you are showing domain mean horizontal mass divergence in Figure 5, which incorporates all regions (convective, stratiform, neither) so you can't assume that differences in mass divergence profiles are related to convection alone. Furthermore, more than entrainment impacts convection. The location of the convection (differing surrounding envi-

ronment) and the low level convective forcing influence the strength of the convective updrafts, and mass divergence incorporates all updrafts and more, so one simulation can simply have more updrafts reaching a certain height level than another simulation, but the entrainment and strength characteristics of the updrafts may be the same. To claim what you claim, you'd have to isolate convective updrafts (perhaps by using a vertical velocity threshold) and compute their buoyancy and detrainment. I also don't understand your argument on lines 3-6. Why would simulations that have the least mass divergence at upper levels be consistent with updrafts that penetrate higher and higher mean cloud tops? c. On page 18, lines 12-14, differences in entrainment and water loading may impact the convective updraft strength and max reflectivity profile, but this is speculation and the correlation between lines in Figure 11c and Figure 10b is far from perfect. To show this, you could plot these variables vs. one another to provide evidence. Another cause of simulation differences includes possible differences in the positioning and/or timing of convection. For example, for 17-18 UTC, the max reflectivity profile comparison looks quite different than for 23-24 UTC. If entrainment and water loading buoyancy differences caused by the turbulence or microphysics parameterizations are primarily controlling updraft strength and max reflectivity, then why is this the case? d. On page 18, lines 27-29, how do you know extra latent heating is occurring without compensation by entrainment or water loading in the ENDGame simulation? Latent heating is one component of buoyancy, but the environment could also be different. e. On page 21, lines 4-6, why can't increases in IWC with vertical velocity be the result of higher vertical velocities lofting more condensate upward? f. On page 21, lines 17-22, how can you draw any conclusion regarding change in IWC with height in observations with so few samples? If you look data from all of the flights and RASTA, they would disprove this result. Furthermore, where do the simulations support the drop in IWC between -20 to -10°C and -30 to -20°C? The distributions for both temperature regimes look very similar. g. On page 22, lines 23-25, why do you bring up the aerosol invigoration effect if your figures do not support it? For example, Figure 11c shows weaker max vertical velocities when cloud droplet number concentration is increased while Figure 16 shows that total ice mass is not changed. h. On page 23, lines 13-16, I don't see a change in 90th percentile cloud vertical velocity in Figure 11, but they aren't as relevant as convective vertical velocity anyway, since it is in convective updrafts (not reflected in 90th percentile cloud upward motion of 0.2 m/s) where Hallett-Mossop is operating. If you examine the max vertical velocity in Figure 11, which is convective, it shows a decrease in vertical velocity by including Hallet-Mossop. Also, how do you know that including Hallett-Mossop increases latent heating? Can you show this? i. On page 24, lines 19-21, you claim that a bimodal PSD representation or a larger observational dataset to generate a more applicable PSD parameterization that correctly represents snow sizes. This is not necessarily true, and I don't see any evidence presented that the two modes of the ice size distribution are important to represent. In fact, the simulated ice size distribution is already bimodal or trimodal because of 2-3 separate ice categories. The fact is that a single-moment scheme will always struggle if it has to represent convective regions dominated by small ice particles and stratiform regions dominated by aggregating large ice particles. This instead suggests that a two-moment scheme that predicts number concentration in addition to mass is needed, and even then, as you show in the manuscript, microphysical and turbulent processes need to be properly parameterized as well, since they impact the predicted PSD moments that define the PSD. j. On page 26, lines 4-12, you say that you show convective updraft buoyancy, but you don't show this or latent heating in the manuscript. Everything related to convective buoyancy and entrainment/detrainment is speculation. k. On page 26, lines 15-23, you don't have a figure where it is possible to discern the slope of reflectivity above the melting level. This is not shown by Figure 6, which shows that the coverage of different reflectivity thresholds is different in simulations and observations, but doesn't show profiles of reflectivity. Furthermore, the slope of snow mean size in Figure 4c looks similar in observations and simulations using the generic PSD and the difference in diameters for 0.5 g m-3 in Figure 17 is not robust and strongly affected by very few observation samples between 0 and -5°C. So overall, I don't see a lot of evidence that implicit aggregation based on the shifting temperaturedependent PSD is too weak. I. Of your 4 listed model shortcomings on page 28, “too much rain above the freezing level”, “too little entrainment”, “increases the stratiform cloud and rain area”, and “too efficient depositional growth” are all statements that are not supported by any evidence shown. They are speculation for explaining the figures that you show, but they are not the only possible explanations for the figures that you show.

6. By heterogeneous rain freezing, do you mean heterogeneous nucleation by ice nuclei or all freezing mechanisms other than homogeneous freezing? This is unclear in the text. You state that because including heterogeneous rain freezing produces better agreement between observations and simulations, it must be important in tropical convective cloud systems (e.g., page 15, lines 11-12), but the simulation including heterogeneous rain freezing only slightly improves on the simulation without it, getting nowhere near observations. With such a difference between the simulation and observations, can you confidently trust that a change in the model is reflective of a change in the real world? For example, what if real tropical convective updrafts loft fewer raindrops than the model does for a given updraft strength. Then the effect of heterogeneous rain freezing in the model will have a larger impact than in real life.

7. The discussion about cloud base on page 19 is incorrect since the inferred cloud base from the stratiform sounding (as discussed in point #1) is incorrect, so I suggest removing this discussion. Cloud base for rising low level air is certainly not 3 km. The argument in lines 15-17 does not make sense to me either. Latent heating by condensation can make air buoyant, but only if this heating makes the air warmer than the environment, which is never guaranteed. Buoyancy accelerates air, so vertical velocity is a function of vertically integrated buoyancy. Therefore, any peak in updraft strength will occur at higher altitudes than peak buoyancy and peak buoyancy is often offset from peak latent heating. In this paragraph and later discussions in the manuscript referencing Figure 11, there is also confusing wording equating in-cloud upward vertical velocity with convective updraft vertical velocity. These are not the same. The

90th percentile upward vertical velocity in Figure 11e is $\sim$ 0.2 m/s, which can easily be achieved in many non-convective cloud types. To confine your analysis to convective updrafts would require some minimum threshold vertical velocity of 1-2 m/s.

8. Be careful interpreting aircraft humidity measurements in convective updrafts. Such measurements can and often do have large errors. Because of this and the small number of updraft samples biasing any statistical comparison, I would not trust any of your conclusions in the second paragraph on page 23.

9. Your reasoning on page 24, lines 10-15, doesn't make sense to me. For the generic ice PSD, if mean sizes are overestimated for IWC > 0.5 g m-3, that means that this PSD has larger concentrations of large particles than observed, not smaller as is stated. This is the only way that mean sizes can be larger for a given IWC.

10. The overall text could be shortened and streamlined. It reads like a "stream of consciousness" at times, which makes finding the key points difficult. This is particularly true because of the large number of sensitivity simulations that you want to describe. I recommend cutting out minor points so that the readers do not get so easily distracted away from the key points. One way to do this is to simply focus on the couple of model component changes that create the biggest effects for whatever variable you are examining. This would also free up space to show evidence supporting your theories (as listed in point #5) for why these specific changes cause the observed effects. You could also cut out some of the simulations if they don't make much of a difference and just say that they don't make a difference. This would unclutter the plots.

11. For comparisons between model output using 1-km grid spacing and 1-Hz aircraft observations ($\sim$150-m sampling), do you average the aircraft observations to a 1-km grid before making comparisons? If not, please do this and include this information in the manuscript. Also include information for how the vertical velocity is retrieved from aircraft measurements, how water vapor is subtracted out of IKP evaporator probe measurements, and why IKP retrievals are assumed to be IWC rather that

TWC (a combination of liquid and ice). If they are rather used as TWC, then making comparisons to simulated TWC (IWC + LWC) would potentially change some of the conclusions in the manuscript.

12. The comparisons of model output with aircraft observations are not robust because of the low observational sample size in updrafts and downdrafts (e.g., Figures 11c, 12, 15, 17). In fact, the aircraft only penetrated 4 updraft cores at -12°C, 1 at -18°C, and then flew through the edges of a few others around -25°C. You admit as much in a few places in the manuscript, but then attempt to draw conclusions from the comparison about which simulations are most realistic, which isn't possible in convective updrafts or downdrafts for this case alone. Therefore, the plots with these comparisons are not appropriate since the model output is a mean relationship with many samples (essentially a population mean) while the observations are but one, likely unrepresentative sample. There are two ways that this issue can be corrected: a. Include aircraft data from the other field campaign flights to make the sample size more robust. These are different cases, but the sampling for this one case is already biased anyway as mentioned in point #2. Furthermore, the aircraft avoided cells with lightning (the most intense cells) in all cases and the most intense cells in this 18 February case had plenty of lightning, so no matter what, the aircraft is always sampling convection in all flights that is weaker than the most intense convection in this case. Furthermore, as your coauthors know, there are RASTA W-band radar retrievals of vertical velocity and IWC that can be used at temperatures colder than -20°C and would increase the observational sample size to make comparisons with model output more robust. b. Sample the model output with pseudo-flight tracks (E-W or N-S is fine) and limit the total sample size to the same as that observed. Do this a number of times to get a population of samples that are each directly comparable to the observed sample. Then the observed sample can be compared to the distribution of samples drawn from the model to see if it fits into the model spread or not. If it does, you cannot say that the model is wrong. If it doesn't, then you can say that the simulation and observations are different. Without this method, any conclusions drawn on the difference between the model output and

aircraft observations are unfounded.

13. You restate many of the results in the conclusions section making it rather long (4 pages). I suggest cutting much of this repetitive text out and focusing on key general points like you attempt to do at the very end of the conclusions section.

Minor Comments

1. On Page 6, lines 17-18, you say that graupel formation does not including freezing rain. Do you mean heterogeneous freezing of rain by ice nuclei? Surely, if a raindrop homogeneously freezes or freezes through contact with an ice particle, it should go in the graupel category, no?

2. On the bottom of page 7, you should also note whether the particle probes have anti-shattering tips or not.

3. On page 8, line 11, you should note the resolution of the peak ice water content since ice water contents strongly depend on resolution.

4. On page 8, lines 24-26: The problems with moisture related to domain size are related to periodic lateral boundaries, but you use a nested simulation where moisture can leave the innermost domains, so I'm unsure as to why this discussion is relevant. As I note in major comment #1 though, your conclusion that the model has a moisture bias is not robust because the soundings are not representative, so I would remove all discussion of it or replace it with the comparison I suggest.

5. For your comparisons in Section 3.1, please state whether you are using the full model domain or the CPOL domain defined by the range ring in Figure 1 to calculate model domain mean quantities.

6. On page 9, lines 21-23: I'm not sure why you cite Fridlind et al. (2012) here to say that the simulated domain mean precipitation rate is outside of the radar-derived precipitation rate range of uncertainty. You also don't show the uncertainty range. If you examined that, why not show it using vertical bars in Figure 3a?

7. On page 10, line 23, in Figure 4, and throughout the manuscript, when you say "mean ice particle sizes", how are mean sizes calculated? Are these mass-weighted mean diameters or something else? Please clarify this throughout the manuscript.

8. On page 11 and for Figure 3, how do you define cloud top in simulations?

9. On page 12, line 21, 23, and 29: A C-band radar cannot observe cloud top or the fraction of the domain covered by hydrometeors since it is only sensitive to precipitation sized hydrometeors, so clarify this by referring to the reflectivity echo coverage.

10. How can you tell that the control simulation evolves from scattered to more organized convection with stratiform regions from Figure 6? I suggest showing this as I state in major comment #2.

11. On page 13, lines 27-28, the excess large particles above the freezing level can also be related to insufficient representation of the rain DSD, warm rain processes, and/or rain sedimentation (representation of fall speeds and size of updrafts being too large).

12. On page 13, line 31: This is true of raindrops and cloud drops, but the lower temperature limit should be 0°C as many raindrops freeze quickly at relatively warm temperatures from contacting entrained ice particles starting at 0°C.

13. On page 14, lines 16-19, I doubt this is the reason for the non-prominent bright band in observations. It is much more plausible that the radar beam smears the bright band out because this data is taken from volumetric scans and more data is far away from the radar than close to it (because of radar coverage increasing as range ring radius squared). Despite this, you still see a bump at 4 km height corresponding to the bright band.

14. On page 15, lines 14-16, single moment schemes typically do increase the number concentration as IWC increases. Aggregation is a decrease in number concentration for no change or an increase in IWC. This can also be diagnostically represented in

single moment schemes by altering the PSD as a function of temperature though. For example, the Thompson microphysics scheme (Thompson et al. 2008) commonly produces the best agreement with observed stratiform reflectivity profile above the melting level. Two-moment schemes can explicitly represent aggregation through predicting the number concentration, but also typically overestimate reflectivity aloft because other factors include excessive size sorting, mass-size relationships, and the assumed PSD shape.

15. On page 17, line 4, the aircraft observations are mostly in stratiform precipitation (plot the flight track on top of the CPOL reflectivity and you'll see this clearly) even though the aircraft penetrates a few weak deep convective cores. The highest concentrations are found in convective cores, not in stratiform regions, so having convective observations does not make them lesser than the ones in Field et al. (2007), which also include convective observations. The observations in Field et al. (2007), however, may suffer from ice shattering artifacts, so they may not be directly comparable to these new aircraft observations that mitigate and control for shattering.

16. From Fig. 10, it looks like there is an issue in limiting hydrometeor sizes to realistic values in the microphysics scheme you are using. A rain reflectivity of 75 dBZ is physically impossible because raindrops begin breaking apart at large sizes. In the real world, rain reflectivities are limited to less than $\sim$55-60 dBZ. Some schemes implement limits on the slope of the rain DSD, and that may need to be done for this scheme.

17. On page 17, line 18, the observed decrease in max reflectivity above 2 km may also be from raindrops falling through weak updrafts and collecting cloud droplets in the classic warm rain process.

18. On page 17, lines 22-24: This is true that different subgrid turbulent mixing decreases max reflectivity, but only for 23-24 UTC and not for 17-18 UTC. Why?

19. On page 17, lines 24-27, I can't clearly see the reduction in max reflectivity caused by implementing the heterogeneous rain freezing parameterization. Perhaps increase

the symbol sizes so that the different lines can be seen more clearly.

20. On page 19, lines 19-21, the upper level vertical velocity peak is also a result of vertical velocity being related to vertically integrated buoyancy. CAPE is usually distributed over a significant depth and the updraft will accelerate as CAPE is used up, primarily being limited by entrainment and opposing pressure gradients. Freezing of condensate and unloading of condensate simply help to push the peak higher.

21. On page 20, lines 23-24, you state that the reduction in rain by heterogeneous freezing reduces accretion of cloud water and thus increases the cloud water mass. Why don't the graupel particles formed by the freezing raindrops accrete the cloud water through riming? Is this related to lower cloud droplet collection efficiency by graupel than rain?

22. On page 20, lines 25-28, how do fast fall speeds of particles help to generate downdrafts? I think of the loading and evaporation, mostly relating to rain in the tropics, as primary drivers. Do fast fall speeds impact loading and evaporation? Also, on lines 28-30, why does more accumulated graupel mass being correlated with the largest IWC in downdrafts support the argument that fast graupel fall speeds generate downdrafts? Do the strongest downdrafts have the most graupel? If so, that would be a supportive argument.

23. On page 22, lines 29-30, I don't see a reduction in total accumulated ice mass in Figure 16. Am I missing something?

24. On page 25, line 5, you claim that the simulations capture the timing of the deepest convection well, but Figure 3 suggests that the simulations initiate and organization deep convection earlier than observed, as you suggest on lines 9-10.

25. On page 25, lines 16-19, what is your definition of "large" particles? Reflectivity is more sensitive to large particles than small particles but a large number of small particles can give the same reflectivity as a small number of large particles, so it seems

Printer-friendly version

Discussion paper

that you are using an arbitrary reflectivity value here to define large vs. small particles.

26. On page 25, line 32, and page 26, line 2, you mention the percentiles of updraft speed, but your figure shows 90th percentile cloud upward motion, which isn't necessarily correlated with max reflectivity since most of the cloud volume is not convective updrafts where the max reflectivities are occurring.

27. On page 26, lines 24-25, do you mean that the heterogeneous rain freezing parameterization reduces raindrops above the freezing level rather than reducing the lofting of raindrops? A freezing mechanism shouldn't impact raindrops lofting above 0°C, right?

28. On page 26, lines 26-28, raindrops not being lofted above the freezing level cannot be detected by radar reflectivity and the aircraft was clearly observing the MCS during its decaying stage, not its mature stage, based on the time series shown in Figure 3. Updrafts, even weak ones, commonly loft raindrops above the 0°C level, but it is true that most of them freeze rather quickly. That is different though than what you state here, that raindrops are not lofted above the 0°C level, which is not supportable from available observations.

Interactive comment on Atmos. Chem. Phys. Discuss., doi:10.5194/acp-2015-970, 2016.

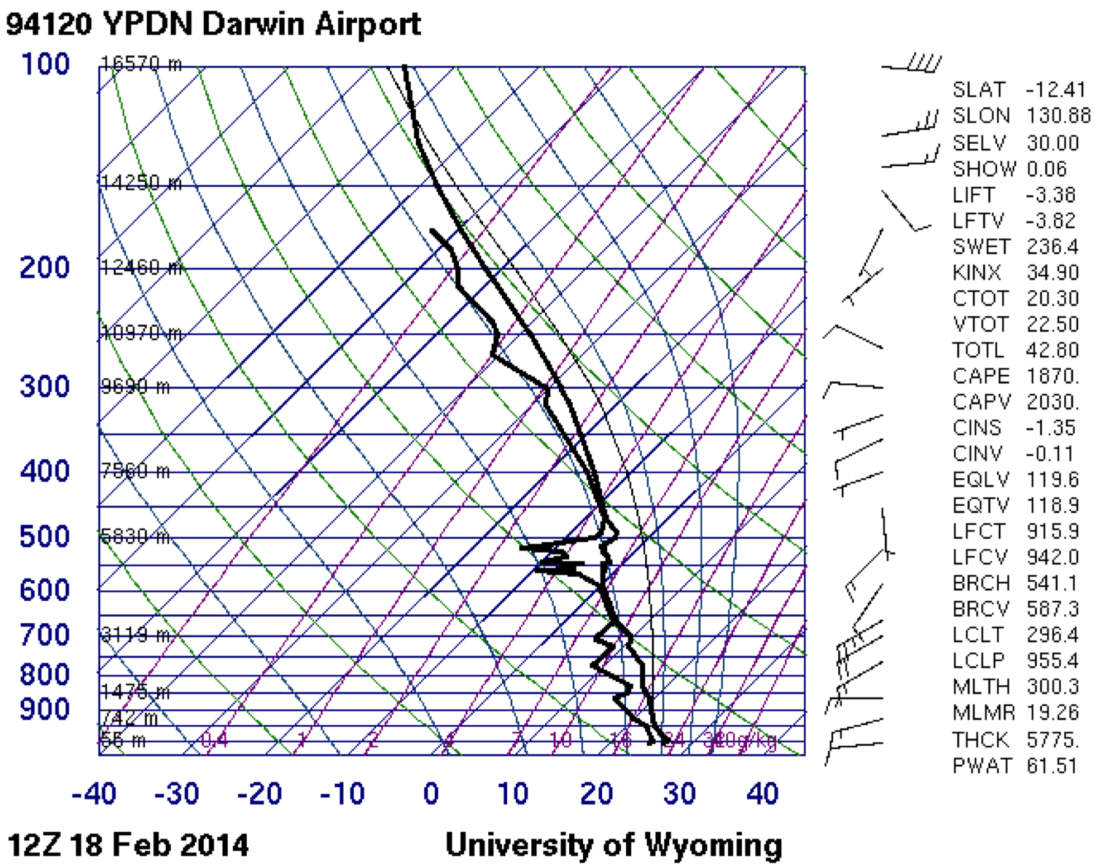

Fig. 1. 12Z 18 February 2014 Darwin sounding

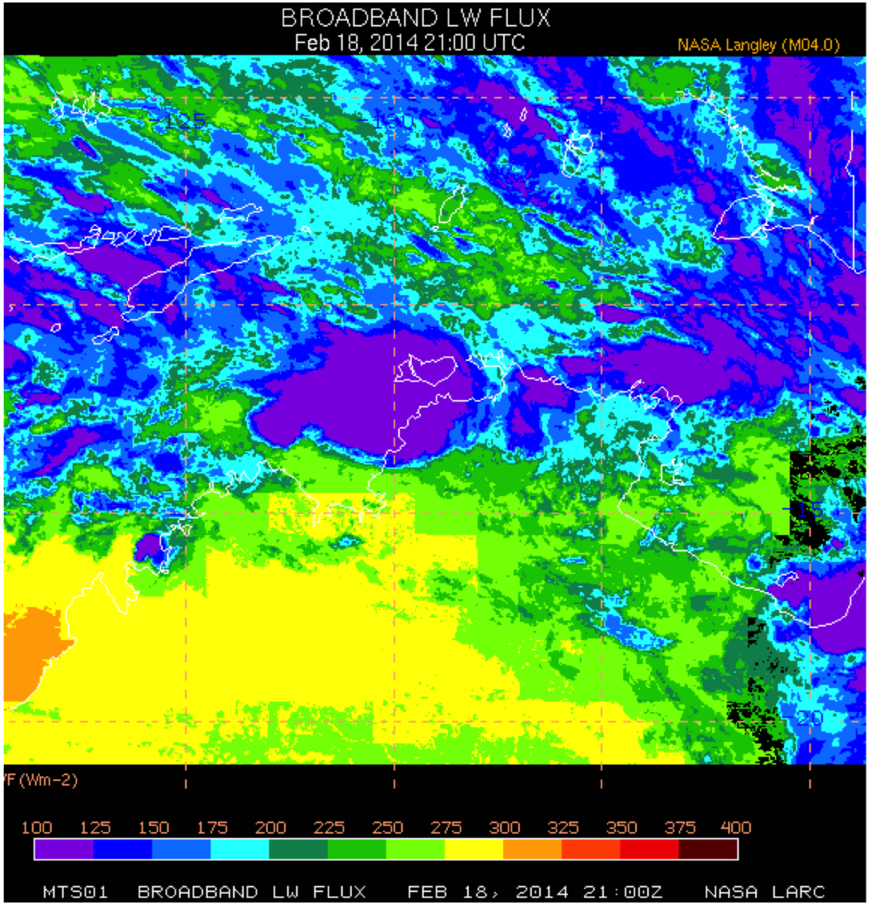

Fig. 2. MTSAT observed OLR at 21Z 18 February 2014