

Revision of « The Microphysics of Clouds over the Antarctic Peninsula – Part 2: modelling aspects within Polar WRF »

We thank the reviewers for their comments.

This document is organised as follows :

- p. 1 (of the pdf) Introduction to the answers to the reviewers (some general comments)
- p. 2 Answers to reviewer #1
- p. 11 Answers to reviewer #2
- p. 18 The highlighted version of the new manuscript

Introduction to our answers :

The motivation for this paper is the hypothesis by King et al. (2015) that radiative biases over the Larsen C ice shelf (Eastern Peninsula) were due to the lack of cloud liquid water. They worked at 5 km resolution and use three models – including AMPS (the Antarctic Mesoscale Prediction System, based on Polar WRF). Bromwich et al. (2013a) also reached this conclusion with Polar WRF at coarser resolution (15 km, and 60 km) continent-wide. However Bromwich et al. (2013a) not investigate the phase simulated by the clouds (only the cloud cover as a whole).

The main goal of the present paper is to provide with a first evaluation of the cloud schemes as they are using the (averaged) flight measurements, to show and discuss their impact on the surface radiative bias, to highlight some important aspects like the need to use double-moment ice microphysics parameterization (unlike WSM5/WDM6) to better capture supercooled cloud water, to show that the current scheme used in AMPS (WSM5) has a lesser ability to model the clouds, to discuss the biases in water vapour and temperature, and to discuss the INP parameterisation. The conclusion has been simplified to clearly emphasize those points.

Note that section 5.2 (« The ice phase parameterisation », now « The INP parameterisations ») and section 5.3 (« Simulating the clouds thermodynamic phase ») have been swaped to putt less emphasis on the effect of the INP parameterisation because it is indeed the case that we do not investigate its effect in details (although we do discuss it).

Note that in sectin 5.3 (former section 5.2), we have decided to delete the paragraph and Figure about the Hallett Mossop process as this discussion was making the section lengthy and would be more adapted to more in-depth investigation of the processes producing the ice using case studies (specific flights). The deleted text and Figure are shown at the very end of this document, as a reminder.

Additionally a section 5.4 has been added to emphasize the role of water vapour, and temperature biases (and other model sensitivity consideration on the initialization of the model, or the number of vertical levels). In doing so we insist on the fact that changing the INP parameterisation is not enough for improving the simulation of the clouds. The previous version of the paper suggested that the INP parameterisations were the main reason of missing liquid cloud in the simulation, and we wanted to correct this.

We have simplified the conclusion to highlight more clearly the main aspects of this work, while we now insist on the need for further work on improving/investigating the temperature/water vapour biases as well. (The abstract has been amended, accordingly.)

Please note that : - Edited/new text is highlighted in the paper. A highlighted section title means that the section is new (5.4) or fully changed/reformulated (conclusion).

- The reviewers comments below are in black. Our answers are in blue color. Below each answer, the page, and line numbers of the related edited/new text of the paper are indicated.

Anonymous Referee #1

This paper presents comparisons of simulations of clouds over the Antarctic Peninsula using different cloud microphysical schemes in WRF, and includes comparisons against observations described in Part 1 of the manuscript. The main finding of Part 2 is that, regardless of which scheme is used for the simulations, a large misrepresentation of the cloud thermodynamic phase explains a lot of the large radiative biases derived at the poles continent wide (i.e., all schemes fail at predicting as much supercooled liquid mass as seen in observations). The authors conclude that a parameterization scheme for ice nucleation that depends on both temperature and aerosol content should be implemented in WRF to get a more realistic representation of primary ice production, and hence a better representation of supercooled liquid phase in Antarctic clouds.

We have emphasized that the main finding of the paper is that changing from the AMPS scheme WSM5 to other schemes does reduce the radiative bias and some schemes substantially reduce the bias on the Larsen C ice shelf. We feel that this makes the paper clearer. In particular we find that double moment schemes (Morrison, Thompson, Milbrandt) perform better than the single moment scheme that is used as standard in AMPS, to simulate liquid cloud. We feel that this is an important finding as it is easy to implement this simple change within operation model while more detailed changes such as implementing new INP parameterisation would require more investigation first. Moreover we also stress the fact that working on the water vapour/temperature biases is also a mandatory step.

The study should be published because unique modeling simulations are compared with a unique set of data over the Antarctic Peninsula. However, the paper can be improved in a number of ways. It is lacking in that it only focus on sensitivities to cloud microphysics, ignoring impacts from other parameterization schemes that could also be affecting the simulations and comparison with observed quantities. The uncertainties associated with the representation of the microphysics needs to be better placed in context of uncertainties associated with the representation of other processes such as boundary layer parameterization schemes. Otherwise, it is possible that the sought after agreement between models and observations may occur due to reasons not associated with the representation of microphysics (i.e., are the right answers being obtained for the wrong reasons). I'm not convinced that simulations that do not consider uncertainties due to other processes can truly assess the ability of the microphysics schemes to model realistic clouds across the Antarctic Peninsula (page 5, line 10) in the absence of these other sensitivities. This is especially true given the statement that none of the microphysics schemes adequately predict the supercooled water: maybe some other process other than microphysics is causing this problem.

We have concentrated on the cloud physics schemes as we felt that this scheme was most likely to be the one that was causing the errors seen by King et al (2015), and because no study focused on Antarctic clouds have investigated this before. Apart from the cloud scheme we have not altered any other physics schemes as although they can be responsible for large errors we feel that many have already been investigated eg : the boundary layer parameterization, although it is one of the key parameterization, due to its implication in the heat/energy transfer between the surface and the troposphere. We are using the one used in the operational forecast model AMPS ; furthermore Deb et al. (2016) showed that the model performance at the surface is most sensitive to the choice of PBL scheme and show that the Mellor Yamada-Janjic (MYJ) scheme is the best performer in terms of the temperature diurnal cycle (in west Antarctica) at 5 km resolution. Hence we are confident that we are using a relevant framework to investigate the cloud schemes' behavior.

See page 5, L.7-12

 For example, I think that the amount of fetch off the ocean and its representation would be very important.

It is possible that larger scale dynamical effects can be affecting the clouds but on the whole the wind field is well forecast by WRF (for instance, wind direction measured at the flights/station is well captured by the model as recalled at the beginning of section 4.2) and so it is unlikely that errors in the fetch over the open ocean are the problem.

See p. 13 , Lines 16-19

 The sensitivity studies in the paper emphasize how different schemes affect the modeled cloud parameters. But, in addition to sensitivity in choice of scheme, there can also be sensitivity to some of the parameters (i.e., constants) that are assumed within an individual scheme. Was any effort made to look at the sensitivity to constants within a scheme?

There are many constants in the cloud schemes but we were concerned in this study with difference between the schemes – if we had altered constants in every scheme we would have had to do many more runs and this was not feasible with the resources we had available. In this paper we focus on more general issues like the performance of the schemes as they are implemented, and the fact that the schemes are single/double moment for the ice, the water vapour/temperature biases, and the INP parameterisation (in the discussion) and this allows a manageable number of runs. However we do show some new results discussing the change of the INP parameterisation :

See p. 24 Line 18-26, and Figure 11

Also in the new section 5.4, page 30.

 Another weakness of the paper is that the paper does a reasonable job in describing the differences between the simulations and the simulations that are most consistent with the observations (note, I think most consistent with the observations is a better way of wording it rather than saying the best simulation). However, a paper in ACP should do a better job in showing why these differences between parameterization schemes are occurring. This can be done by looking at various prognostic terms in the model (i.e., production rates due to various microphysical processes). This could explain why different schemes produce different modeled fields rather than just stating that different schemes produce different cloud fields; the result that different microphysical parameterization schemes gives different modeled fields is not an incredibly new or exciting result in that it has been previously demonstrated in a wide number of other meteorological conditions.

Although we appreciate the reviewer's point this study is intended to give a general picture of the cloud simulation performances of different schemes over the Antarctic Peninsula, and as this has never carried out before we did not feel it necessary to start investigating the production terms of all relevant processes of the different schemes. A major difference as we show it is the fact of having single-moment ice microphysics (and the related dependency between the number of crystals, and the ice water content) (WSM5/WDM6) versus double-moment ice microphysics (Morrison, Thompson, and Milbrandt). Some additional simulation results have been carried out and are presented (section 5.2, and new section 5.4) or just reported (section 5.1) to improve the discussion.

See p. 23 Line 17-21, in section 5.1

See p. 24 Line 18-26, and the related new Figure 11, in section 5.2

See p. 30 section 5.4, and the related new Figure 15

These different additions are further mentioned in the answers below.

 The paper attributes a lot of the disagreement between the modeled and observed fields to the performance of the ice nucleation parameterization: namely, it is stated that the ice nucleation parameterization is a major reason why supercooled water is underestimated. For example, it is stated that forcing a dependence on aerosol amount and temperature, rather than just temperature, could give better agreement. Can more be done with adjusting constants in the ice nucleation schemes within the different cloud parameterizations to better demonstrate this? This way it would be easier to define how much of the difference is due to the different microphysical schemes, and how much is due to the microphysical parameterization schemes (rather than just stating that different parameterization schemes use different representations of INPs that may be causing some of the differences).

We feel that we may have been misunderstood and so have changed the way the results are presenting in the paper. We have now put less emphasis on the importance of INP parameterisation as the main cause of the differences between WSM5/WDM6, and Thompson/Milbrandt/Morrison, is the fact that the former rely on single-moment ice microphysics with a parameterisation linking N_i (number of crystals) and q_i (mass of the ice phase), that lead to enhanced production of ice.

See p. 24 Line 18-26, and the related new Figure 11, in section 5.2

To this respect, the choice of the INP appears to be secondary in comparison whether the scheme is single or double moment. Also we stress the need of reducing the water vapour/temperature biases to help improving the cloud simulation but this is probably outside the scope of the present study.

See p. 30 section 5.4, and the related new Figure 15

As such these are interesting results to Polar WRF/AMPS users.

 The authors conclude that the Morrison scheme is the best performing scheme. It would have been beneficial to show some production and depletion rates of processes producing various hydrometeor categories (especially supercooled water) so that there could be a focus on specific terms and processes that are well represented in that scheme, and hence to better determine where the supercooled water is coming from. That might also help explain why the Morrison scheme is the better performing scheme.

In this study we have concentrated on the average performance of the schemes over many different cases – those that were reported in part one of this study – rather than look in detail at a particular case. We think that the detailed study of production/depletion rates is best left to future work focusing on higher resolution simulation and case-studies (specific flights) and rather than the mean values we use here. These average results show that the double moment microphysics schemes (Morrison, Thompson, Milbrandt) perform better than (WSM5, WDM6) in modelling the liquid phase. We find that the Morrison scheme is performing better but no longer claim that it is the « best performing scheme » as we agree more detailed study of the processes will be needed.

The authors use the Brown and Francis (1995) mass-dimension parameterization to derive the ice water content. However, this is a very crude parameterization whose use may be inappropriate depending on the mixtures of the shapes and sizes of crystals present. What types of habits were noted in the CIP images? Are they consistent with those assumed in the Brown and Francis parameterization? How sensitive are the calculated IWCs to the assumed parameterization? Does this assumption affect any of the qualitative conclusions?

Although there are more recent mass diameter laws that we could have considered using Brown and Francis was the only scheme available within our image processing software (OASIS) and considering the errors we would expect anyway from our low resolution CIP probe. It also meant that our results were directly comparable with other results from polar studies who had used the same method (for example Lloyd et al (2011)).

 The smallest domain used in the model simulations is 5 km. This seems quite coarse for a study looking at the impacts of microphysical processes. What vertical resolution is used? Does the choice of horizontal and vertical resolution have any effect on the most important findings of this study? In addition, how sensitive were the simulated fields to the choice of initial conditions (e.g., either the product used or the time at which the simulations were initialized)?

We answer the various points below :

* It is a similar horizontal resolution to previous work (ie King et al. 2015) that we wish compare this study too. It is also a higher resolution to Bromwich et al. 2013a (15-60 km resolution), who investigated radiative biases over the continent using Polar WRF. We have completed many runs for this study – we have been looking at more than 1 month for each scheme, each year (2010, 2011) during which many observational flights were completed. This has meant that resources would not have allowed a significantly higher horizontal resolution. (However, we can say that other unpublished/preliminary simulations comparing WSM5 and the Morrison scheme at slightly higher resolution (2 km), and doubling the number of vertical eta levels (60, instead of 30), over the Weddell sea, and related to our 2015 campaign do lead to the same conclusions presented in the manuscript : that the WSM5 scheme underpredicts significantly more the supercooled liquid water, compared to the Morrison scheme, that the radiative biases are significantly reduced using the Morrison scheme, that the water vapour/temperature biases are also responsible for disagreement between the model and the cloud measurements). (For the vertical resolution, see further below).

* Bromwich et al. (2013a) showed that using ERA-Interim reanalysis products for initial, and boundary conditions produces the best skills within Polar WRF, and thus we decided to use ERA-Interim data.

See p.6 , L. 1-2

* 30 (eta) levels are used, with the levels being closer to each other closer to the surface. See Table 1, and explanation when deriving the averaged vertical LWC profiles (p17. L4-7) ; the maximum distance between model vertical levels is approximately 500m below 4500 meters, with an average distance of approximately 350m, and down to 10-50 m above the surface).

See p. 17 L4-7

* A doubling of the number of levels does allow to reduce the biases of water vapor and temperature in some places, but it should not affect our general result, as a significant improvement in the cloud simulation is not observed. The initialisation closer to the date of observation reduces the water vapour bias (mainly), but it does not necessarily need to better cloud simulation, and the general picture we

provide here when comparing the schemes will not be altered. A discussion with small sensitivity test on specific flights was added to raise those concerns, as we agree they must be discussed to some extent. Hence, the additional section 5.4.

See new section 5.4 p. 30.

 I am also worried about some of the inconsistencies between the microphysics and radiative parameterization schemes that seem to exist in the paper. For example, radiative schemes may make specific assumptions about the shape distributions of various hydrometeor categories. Are these assumptions consistent with what is assumed in the microphysical schemes themselves? Otherwise, an inconsistency between the radiative and microphysical parameterizations could be responsible for some of the discrepancies in the radiative parameterization schemes. Further, the authors themselves seem to state that the assumption of a constant effective radius of 10 micrometers in the radiative scheme may not be consistent with what the microphysical parameterization scheme is assuming.

We understand completely that radiation scheme may not be consistent with the cloud physics scheme and we did consider changing both at the same time. However we felt that it would be better to only change one scheme at a time. More generally, those inconsistencies mentioned by the reviewer are part of the drawbacks in using a flexible model like WRF where different parameterizations can be used in different combinations.

However, and importantly, the approach used in the SW radiative scheme (where r_{eff} is held constant) is to parameterize the optical depth for water/ice cloud as a function of r_{eff} based on the fact that (A solar radiation parameterization for atmospheric studies. NASA Tech. Rep. NASA/TM-1999-10460, vol.15, 38 pp) : « Theoretical considerations and radiative transfer calculations have shown that cloud single-scattering properties are not significantly affected by details of the particle size distribution and can be adequately parameterized as functions of the effective particle size » (Fu, 1996; Hu and Stamnes, 1993; Tsay et al., 1989). We appreciate that this may be a strong assumption, especially that the width of the size distribution should impact the radiative properties of the cloud (but then the width also impact the effective radius value). Nevertheless, as the parameterization used does not assume any type of size distribution, we think that it can be used with the different cloud schemes despite their own different assumptions on the hydrometeors size distributions.

See p. 23 L. 11-14

Also, we derived an averaged r_{eff} for the droplets in the measurements of about 7.6 μm in 2010, and 7.2 μm in 2011, close to the 10 μm assumed (constant) in the SW radiative scheme. We have ran the model over the period 11th-20th 2011 setting $r_{\text{eff}}=7\mu\text{m}$ and observed no consistently better radiative bias. Having this r_{eff} constant through the simulation is certainly problematic (more than the fact of having a fixed value either to 7 or to 10 μm) and could be one of the reason of larger shortwave biases. The LW bias is more sensitive to the ice/liquid content, hence is systematically reduced (to the east of the Peninsula) with cloud schemes simulating more liquid clouds (Morrison, Thompson, Milbrandt).

See p. 23 L. 17-21

SPECIFIC COMMENTS

Page 4, line 13: What diameter is the distinct peak located at? Also, recommend using size distribution instead of spectra. Spectra typically refers to radiative quantities.

It is for a diameter between 8 and 12 μm .

See p. 4 L 19-20

and we have replaced spectrum by size distribution.

Page 5, line 14, recommend “graupel is” rather than “graupel are”

We have corrected that.

Page 8, line 9-10: Degree of agreement is always a very relative term. I think when talking about whether simulations agree or disagree, the writing should be more quantitative stating what the degree of agreement is.

We made the writing more quantitative.

See p 8 L. 24-29

Page 8, line 19: How much graupel was in the observations? Can you give some indication of how much agreement there is between modeled/observed graupel?

We don't think we can resolve graupel properly in our observations. It is difficult to tell the difference between an irregular ice crystal and a heavily rimed one (graupel) by looking at the shadow images.

Page 9, line 16: Is the radiative scheme consistent with each of the microphysical parameterization schemes?

Please refer to our answer to the last major comment above, and related to this specific point.

Page 10, line 4: Shouldn't the radiative schemes be consistent with each of the microphysical parameterization schemes rather than just the K15 scheme?

Please refer to our answer to the related major comments above.

Page 12, line 8: It would be nice to know the specific terms that are producing the supercooled liquid mass content in the model.

We did not perform simulations allowing to target specifically the production terms, as it is not the goal of this paper. However, the production of supercooled liquid results from the supersaturation remaining after the ice has condensed. The process is condensational growth and is depending on how efficient the ice formation is. In particular, in this case, we show that the double-moment schemes for ice are performing better than the single-moment schemes for ice WSM5/WDM6 (which are single- and double- moment for the liquid phase, respectively). The latter have Nice and IWC not evolving independently (single-moment), and the relationship used to link both leads to high values of IWC. We now explain this in section 5.2, particularly :

See p. 24 L19-27

Line 12, line 22: If two schemes use the INP parameterizations, is it fair to say that the microphysical parameterizations are causing the differences, or should it just be attributed to the INP parameterization. This sentence is clear, but I am wondering how this affects some of the preceding discussions in the paper.

To avoid confusion, we removed this sentence, as it anticipates too much on what corresponds to the discussion part, and since – as stated in our introduction to this document – we decided to put let emphasis on the INP parameterisation, when revising our work.

Page 13, line 19: If the effects of the mountains are variable/valuable, what are the prognosed values that are affecting the generation of the LWC? This might be especially good to look at given that quantities east and west of the mountain range are being compared.

Note that we compare quantities to the east and to the west of the Peninsula because the statistics of the observations were much better on each side as the aircraft tended to avoid clouds when crossing the mountains. Also the data have been grouped in East/West « bins » to improve the statistics when comparing averaged quantities (as in Part 1).

The supercooled LWC will be generated mainly depending on the supersaturation removal by the ice microphysics processes. This is shown for instance by the additional simulation using WSM5 scheme with a modified relationship between q_{ice} (mass) and n_{ice} (number), where the factor linking q_{ice} to n_{ice} is divided by 100 in the empirical relationship used for WSM5's ice crystals. This edit makes WSM5 generating supercooled LWC (say comparable to Morrison). In comparison, changing the INP parameterisation does not increase the supercooled liquid simulated with WSM5 across the Peninsula. As presented above, this is now emphasize in the discussion part section 5.2.

Page 14, line 7: the statement that the schemes perform worst on the West compared to the East is interesting. But, the paper would be much more insightful if it could identify the processes that are at work so that the schemes are performing better on the West compared to on the East.

It is actually hard to say.

The average temperature bias is larger to the west compared to the east, and now it is clearly recalled in the discussion, as well as the topography which is completely different to the west and would require higher resolution simulation in future work.

See p. 26 L. 1-2

and

See p. 27 L. 7-12

Page 15, line 4: There is an overemphasis on what is the best scheme. You might be getting the right answer for the wrong reasons. You should also look at what are some of the prognosed terms allowing this scheme to perform better, and how the sensitivity to microphysical parameterizations is affected by sensitivity to other parameterization schemes.

We agree that we emphasized too much on what the best scheme could be and we changed that in the paper (as already stated above). As for the second part of the comment, we rather state that the double-moment ice schemes is more important than a double-moment liquid scheme to improve the supercooled liquid water simulation, and has to be preferred to a single-moment ice scheme. In terms of other parameterization schemes we explain above that choosing the AMPS (Antarctic operational model) configuration is a relevant choice, and mention Deb et al. (2016)'s paper as for validating the most relevant PBL scheme in an Antarctic environment.

See p. 5 L. 7-12

Page 17, line 9: How much of the missing mass might not be detected? I think a simple estimate of this can be made if you extend the mass distributions to larger sizes (simple exponential or lognormal fit) to estimate how much of the mass that you are missing.

We have approximated the crystal size distribution by exponential distributions to get some estimate of the error related to the particles that are missed above $D=1.5$ mm. Overall, the average relative error on the water ice content is about 5% in 2010, and 8% in 2011. We also introduced a new Figure 8 to show this.

See P18 L.12-32, and the related new Figure 8.

Page 19, line 1: How much of a difference is important? What difference is acceptable in terms of being a good match with observations?

In this sentence we compare the averaged value of the mixed-phase ratio, and its variability to the observed quantities, and state which scheme produces the closest agreement to the observations. Here, the best match is considered as the one which gives the closest value of the so-called mixed-phase ratio to the observed one. (but we might be missing your point, here.)

This paragraph is now p. 19 L17-22.

Page 19, line 7: This paragraph has a very good description of errors associated with these measurements. It would be nice if a similar comprehensive discussion of the uncertainties and errors associated with the microphysical measurements were included, especially given that microphysics is the focus of this study.

The errors associated with the microphysical measurements and the instruments are considered in part one of this study (Lachlan-Cope et al. 2016, ACP). For instance we explained that the « the CAS data for most flights agrees within 15% with the Hotwire LWC Sensor – part of this discrepancy is attributed to the Hotwire LWC Sensor's tendency to underread at high values of LWC. The CIP images particles between a diameter of $25 \mu\text{m}$ and 1.5 mm, with $25 \mu\text{m}$ pixel resolution, and had not at the time of this campaign been fitted with anti-shatter tips. However, a study of the particle inter-arrival times indicated very few shattered particle, and these were removed by eliminating particles that arrive within $1 \mu\text{s}$. » (Part 1, Lachlan-Cope et al. 2016, section 2.1)

In the present paper :

- the errors (or bias) we can speak of, related to the LWC, are explained in the section 4.3.1 (first paragraph) already and are related to the inability to identify correctly the entire population of droplet in the large size-range (small size range of the CIP/crystals). Errors on the CAS measurements itself as such are in comparison negligible. Also, the hotwire probe (bulk measurements) backs up the LWC derived from the Cloud Aerosol Spectrometer (it is said in part I, although we do not recall this here).

See p14 L. 1-2 and p15 L1-9

- As explained above, we now present some errors for the IWC in section 4.3.2 (Ice phase) related to the icy particles not detected above $D=1.5$ mm (using exponential distributions to extrapolate the number of crystals), that lead to an underestimation of the real ice water content.

See p18 L 7-26

Page 21, line 12: Should you state that WSM5 and WDM6 should not be used for these studies when the other boundary layer/other parameterization schemes are being used? This conclusion might not apply if some other boundary layer parameterization schemes are being used.

We think we can state that given our settings WSM5/WDM6 which rely on the same ice microphysics parameterization (the fact that WDM6 has also graupel does not change the results since graupel are

very few and almost always absent in the simulation output) should not be used because they perform less well in terms of supercooled water simulation. Since we explain that the ice microphysics part of the cloud parameterization is responsible for this, we don't think using another PBL would lead to different result. Also, the PBL scheme we are using was shown by Deb et al. (2016) to perform the best and is used in AMPS, and as such is a relevant framework to work with.

Page 21, line 17: Is the assumption of a constant 10 micrometer effective radius consistent with the assumptions that are made in the microphysical parameterization schemes? See earlier comment about consistency between microphysical and radiative schemes.

Please refer to our answer in the major comments.

Page 21, lines 29-30: If the INP parameterization is so important, why diagnose it only from temperature if an alternate parameterization is available? It would seem that this could be a greater focus of the study if the INP parameterization is so important.

Although we have decided to keep some discussion about the INP parameterization in the paper - because we believe such report can motivate further work - we toned down the statement that INP are mostly responsible for the differences between schemes. Indeed, for our work, the most important factor is the ice microphysics single-moment param. of WSM5/WDM6 (See page 24 L14-24) and local vapour/temperature biases that can lead to missing clouds in the model (see new section 5.4)

Page 22, paragraph beginning line 22: Can you show how much specific terms are contributing to the ice production processes rather than having generic descriptions of these processes?

We have reorganised the discussion and made it insist less on the possible importance of the INP parameterisation (e.g. compared to the single-moment param. for the ice in the WSM5/WDM6 which has a more obvious effect. See additional simulation in section 5.2)

The goal of the present paper is to use some of the WRF schemes as they are, comparing averaged quantities at 5 km resolution. We keep the paper general as to provide with a first report on the importance of choosing a relevant cloud scheme when using Polar WRF, and showing that the double-moment ice microphysics schemes simulate a LWC in better agreement with observations (and reduced radiative biases over the Larsen C ice shelf).

We plan to use the aircraft campaigns for further modelling case-studies (at higher horizontal/vertical resolution) in the future (along with the recent 2015 campaign that took place at Halley) where specific production terms will be more easily investigated by focusing on specific flights, rather than deriving/comparing averages.

Anonymous Referee #2

The authors present a modeling study of various microphysics schemes and compare model results to observations to aircraft measurements in the Antarctic. They find significant differences in predictions of the cloud thermodynamic phase and conclude that the Morrison scheme is the best scheme as it leads to the best model/observation comparison in terms of clouds and radiation. Fore The number of model studies on cloud in Antarctica are sparse, even though the region is very sensitive to changes in climate. Therefore, the current study might represent an important contribution to the literature if my comments below will be addressed.

General comments

1) I am not convinced that the authors can indeed conclude on ‘the best performing’ microphysics scheme based on their model studies. They show that the Morrison scheme predicts best the super-cooled liquid in clouds, followed by the Milbrandt and Thompson schemes since all three have a sophisticated description of the various ice categories. Given the large uncertainties that are associated with the representation of clouds in general and of mixed-phase clouds in particular, the comparison to only a few parameters might not be sufficient to identify ‘the best’ scheme. Uncertainties in the radiation and boundary layer schemes might lead to a prediction bias such as the microphysics scheme may give the right answer for the wrong reason. I suggest a more careful wording throughout the manuscript. At several places in the manuscript, the language is rather colloquial; I listed several instances below but encourage the authors to carefully read and improve language where possible.

We rather state now that the Morrison, the Milbrand, and the Thompson, lead to simulation results that are in closer agreement to the observation than the WSM5/WDM6 schemes, which are single-moment schemes for the ice crystals. We avoid the « best performing » expression, as we agree it is not adapted.

In terms of the other physics schemes (radiation, PBL, etc.) :

Please note that the physics schemes (other than the cloud scheme) we are using are the ones used in the operational forecast model AMPS (used by King et al. 2015, whose conclusion motivated the present study at 5 km resolution). We decided to discuss the microphysics schemes in the framework of the AMPS physics settings. Many choices arise when working with WRF in terms of physics parameterization – indeed - but it is very difficult to test all the parameterizations and possibilities and so using the AMPS ones seemed to be a reasonable choice. In this framework it is relevant to discuss the ability of the schemes to simulate the clouds. As for the boundary layer parameterization, we agree that it is one of the key parameterization, due to its implication in the heat/energy transfer between the surface and the troposphere. Again, we are using the one used in AMPS ; furthermore Deb et al. (2016) showed that the model performance at the surface is most sensitive to the choice of PBL scheme and show that the Mellor Yamada-Janjic

MYJ scheme is the best performer in terms of the temperature diurnal cycle (in west Antarctica) at 5 km resolution. Hence we are confident that we are using a relevant framework to investigate the cloud schemes’ behavior.

See page 5, L.7-12

2) The authors state that none of the applied microphysics schemes was specifically developed for Antarctic clouds and, therefore, their study is a first step in exploring the skills of the different schemes for such scenarios. However, I am missing a more conclusive

statement on what improvements should be done in these schemes to optimize them for Antarctic clouds.

The paper concludes on what is required for a Polar WRF user in order to have a better simulation of the supercooled liquid water (a double moment parameterisation for the ice microphysics, and not a single-moment like WSM5/WDM6). We show that the ice double-moment schemes are also the one allowing for reduction of the LW bias (in addition to simulating clouds in better agreement with observations). It is of importance as WSM5 is the scheme implemented in the AMPS operational model, widely used in the antarctic community. Also, we take advantage of having ice crystals measurements to compare to the INP parameterisations to suggest a best choice for future simulations/work. However, we do now emphasize that water vapour, and temperature biases are also responsible for the failure of cloud simulation (new section 5.4), and improving the INP parameterisation is not the only thing to do.

Section 5.4 p. 30 with additional results related to sensitivity of water vapour/temperature bias:
See new conclusion

3) It is mentioned that the liquid phase of clouds (drop activation) is either described based on a fixed number of droplets or it is a function of the CCN (p. 7, l. 5). Drop activation is an essential process that determines the microphysical properties of a cloud. How much difference is caused between the different microphysics schemes due to differences in the prescription of the cloud droplet number?

We have tested a different CCN concentration for the Morrison scheme (for instance, 100/cc instead of 250/cc) in four flights to see the difference, and it lead to a reduction of the liquid formation, without being conclusive for the discussion and the conclusions we want to present. We do mention another additional simulation with CCN=100 cm⁻³ for the Morrison scheme in section 5.2., but again the differences induced by these changes is much less important than the difference between WSM5/WDM6 and Morrison/Thompson/Milbrandt that we want to stress.

See p. 24 L. 28-33

However, a complete picture addressing this would require many more simulations, in prescribing different CCN parameterizations to all the different schemes used, but this is not what the present paper is intended to present. The main aspect of the present paper is to provide with a first evaluation of the schemes as they are comparing with the (average) flight measurements, to show and discuss their impact on the surface radiative bias, and to highlight some important aspects like the need to use a double-moment ice microphysics parameterization (unlike WSM5/WDM6). It is also important to show that the current scheme used in AMPS (WSM5) has a lesser ability to model clouds in agreement with observations. Finally we discuss the biases in water vapour and temperature, and the INP parameterisation.

4) The authors describe in length the differences in the results predicted by the microphysics schemes in Section 4.1. However, in order to understand the differences and to assess the skill of the various schemes, a more detailed discussion of the underlying processes is needed. Such analysis will help to identify the ‘best scheme’ for the right reasons and to improve existing microphysics schemes.

The present paper aims at comparing several schemes within Polar WRF (among which the one used in AMPS) in an averaged way and we decided to group the schemes in categories : the one having a single-moment ice microphysics parameterization (WSM5/WDM6), and the one having a double-moment ice microphysics parameterization (Morrison/Thompson/Milbrandt). One of the result is that the latter is best in simulation supercooled liquid. The good/poor ice microphysics simulation is directly responsible for the good/poor supercooled liquid simulation as it is explained in section 5.2 :

See p. 24 L.18-27

5) Several previous studies have highlighted the importance of ice particle shapes for ice growth and sedimentation and therefore for the partitioning between ice and liquid phases (e.g., Sulia and Harrington, JGR, 2001). How are ice particle shapes treated here? Are the predicted shapes in general agreement with observations?

The ice particle shapes are not handled in any of those schemes. Each scheme assumes a fixed shape for the ice particles, and uses a mass-diameter ($m=cD^d$) law, which is needed in the single- or double-moment parameterisations used. For instance the Morrison scheme has $c=\pi\rho/6$ (where ρ is the particle density), and $d=3$: thus the scheme assumes that all particles are spheres. In the Thompson scheme, all hydrometeors are modelled as spheres, except from the snow with $c=0.069$ and $d=2$ (fractal like particles). In the Milbrandt scheme, all hydrometeors are assumed to be spherical, except for the ice crystals, which are assumed to be bullet rosettes ($c=440 \text{ kg m}^{-3}$ and $d=3$). In WSM5/WDM6 the mass-diameter equation is derived based on empirical formulas giving the terminal velocity of falling crystals. The fixed shape of the ice particles is a commonly used assumption in the derivations of moments of the size distributions and allow for explicit equations to be used for cloud microphysics’ bulk parameterisation in atmospheric models. The sphericity often assumed is a crude assumption but is largely used, for simplicity. We agree that this aspect will need to be addressed in case-studies simulation.

6) Only in Section 5, it is mentioned that the radiation scheme assumes a drop radius of cloud droplets. How does this value compare to observations and how does it affect in general the radiation prediction? Are there any studies (not necessarily in Antarctica) that have discussed biases in predicted radiation due to a constant effective radius?

Our derived value from the observation is an averaged effective radius of 7.5 μm for both campaign (7.2 μm for 2010, and 7.6 for 2011), so that it is close to 10 μm . Also, we have rerun a simulation over the period 11th -20th Jan. 2011 to compare the new radiative bias to the former one and found no improvement in terms of bias. We believe a time-dependent effective radius should be implemented as an input to the radiative scheme, but this is out of scope of the present paper. We could not find studies discussing biases due to a constant reff. We have added some additional discussion about this in the paper, also related to the other reviewer’s comment on this aspect.

See p. 23 L 11-21

7) Each microphysics scheme includes a different parameterization of ice formation. Different INP parameterizations can lead to significant differences, as it has been shown in several literature studies (e.g. Eidhammer et al., JGR, 2009). How much of the predicted differences in comparison to observations and model/model comparisons can be ascribed to the differences in INP parameterizations? Why not using the DeMott parameterization as the base case as it has been shown to perform better than the temperature-dependent-only parameterization in previous studies?

We have changed the way we present the results. We put less weight on the INP parameterisation, as we rather emphasize the main difference between WSM5/WDM6 (single-moment for the ice microphysics) and Morrison/Thompson/Milbrandt (double-moment for the ice microphysics). We do comment on the INP parameterisation as we believe it is a matter of future improvements of the cloud scheme, however we also emphasize the need of better capturing the water vapour, and temperature.

In this paper, we do not intend to investigate the detailed reason of the differences, rather to show the differences between the two groups of schemes, and explain the main reason of their different behavior (which is the single-moment parameterization for ice). Also, in shorter runs we have not observed the DeMott parameterization being enough to lead alone to a better prediction of supercooled liquid as we explain it in the new section 5.4. We also emphasize the importance of water vapour, and temperature biases.

Obviously, more simulations would be needed, certainly at higher resolution to make an even better use of the aircraft measurement, and this is what we are planning in future work.

8) The authors should better justify their choice of a 5 km resolution. At the end of Section 5, they state that convective processes are badly resolved on this resolution; however, they fail to discuss the consequences of this caveat for simulating clouds.

* We chose the 5 km resolution because it was the resolution used in King et al. (2015) where the authors examine the surface energy budget on the Larsen C Ice Shelf, using three models including AMPS (the Antarctic Mesoscale Prediction System, based on Polar WRF). The authors noted the very little amount of liquid formed in the AMPS model, and they suggested that the radiative biases measured were pointing towards the misrepresentation of the clouds. Also Bromwich et al. (2013) reached similar conclusion using Polar WRF at coarser resolution (15 km, and 60 km), however continent-wide. We decided to work with the higher resolution of 5 km, used in King et al. (2015).

p. 4, Line 30-31 and p.5, Line 1-2

* We mention that « warmer oceanic and sea-ice free influence of the western Peninsula and the Southern Ocean implies more convective processes (compared to the east) that are badly resolved at 5 km resolution and prevent better matching with the aircraft observations ». Hence we expect convective processes to play a bigger role to the west of the Peninsula.

See p.26 L.30-32

* Using the convective parameterisation in the 5km resolution domain did not allow to produce better agreement with the observations in terms of clouds and clouds phase, for instance on the specific flights

mentioned in the new section 5.4 (we do not mention it, though). We believe working at 1 km horizontal resolution on case studies in future work will help better address the possible issues related to convection processes by capturing them, without having to rely on the parameterisation.

9) Throughout the paper, expression such as ‘the xxx scheme forms more liquid’ etc should be avoided (e.g. p. 13, l. 15). The schemes themselves do not form or produce any ice or water. Wording such as ‘Using the xx scheme, it is predicted : : :’ (or similar) should be used instead.

We corrected the wording, when appropriate.

Minor comments

p. 2, l. 34: Not clear what ‘the latter case’ refers to. – Do you mean all three, i.e. immersion, contact or condensation freezing? If so, it might not be fully correct, since droplets may scavenge INP before ice nucleation and therefore INP and CCN are not the same.

We meant the « condensation freezing » case. We corrected this by saying « in the condensation case ».
See p3 L 4

p. 7, l. 18: The beginning of the sentence is not clear. Does the AMPS model predict a liquid phase or not?

AMPS simulate clouds predominantly composed of ice with very little or even zero liquid water.
We corrected the sentence.

See p8 L3-5

p. 8, l. 10: What do you mean exactly with ‘twice to four times more liquid clouds’? – Does it refer to the number of clouds or to the liquid water content of the clouds?
Figures 2 and 3: In particular, in the upper panels, the dotted lines are pretty hard to distinguish.

We have corrected the wording :

* Using the Morrison scheme, twice to four times more liquid cloud mass is simulated than when using the WSM5 scheme.

* We made the dotted lines appear more clearly in Figures 2, and 3.

See p.8 L23-28

p. 11, l. 6: I think (ice) should be removed – or SWC0 added at the beginning of the sentence (?)

We edited the couple of sentences.

See p13 L7-11

p. 17, l. 5: Can an estimate be given how much the observed mass is biased due to the small and large cut-off diameters? – An extrapolation of the measured size spectrum

might be better than the current complete omission.

We have approximated the crystal size distribution by exponential distributions to get some estimate of the error (bias) related to the particles that are missed above $D=1.6$ mm. Overall, the average relative error (bias) due to this aspect on the ice water content is about 5% in 2010, and 8% in 2011. See explanations :

See p18 L12-32 and Figure 11

Technical comments

- p. 1, l. 9: 'struggle' is rather colloquial
- p.2, l. 10; p. 3, l. 7, and other places: Antarctic
- p. 3, l. 6: Southern
- p. 4, l. 14: discriminated
- p. 4, l. 26: one domain
- p. 5, l. 11: developed
- p. 6, l. 10: cloud formation
- p. 6, l. 25: flight measurements
- p. 7, l. 21: observed
- p. 8, l. 1: domain
- p. 8, l. 27: demonstrate
- p. 9, l. 2: tests
- p. 11, l. 8: Add unit to LWC (0.02)
- p. 12, l. 28: wind regimes
- p. 12, l. 33: either 'shows a : : transect' or 'shows : : transects
- p. 16, l. 19: flights
- p. 17, l. 14: equal
- p. 19, l. 27: increases
- p. 21, l. 9: simulation
- p. 21, l. 24: are consistent
- p. 23, l. 7: agrees
- p. 25, l. 3: crystal
- p. 26, l. 15: 'The transects in Figure 5a and b clearly show..'

These are corrected.

Deleted paragraph and figure from the previous version of the paper, about the secondary ice process Hallett-Mossop (former section 5.2, now section5.3)

As discussed in Part 1, the distribution of ice crystals as a function of the temperature shows that below -10 C the production of ice crystals is dominated by primary ice production processes, which are represented in the cloud schemes by the INP parameterisations. However, a secondary ice production process peaking around -5 C, and identified as the Hallett-Mossop (HM) process (Hallett and Mossop, 1974) (see Part 1, section 3.2) has to be accounted for with a dedicated parameterisation in the temperature range -10 C to -3 C. Except from WSM5 and WDM6, which do not account for that process, the Milbrandt, Thompson and Morrison schemes use the same parameterisation for that mechanism (Reisner et al., 1998). It is derived using the equations of collection of water droplets by snow and graupels, multiplied by a temperature-dependent parameter based on Figure 2 of (Hallett and Mossop, 1974). Figure 11 shows the median number density of ice crystals and snow particles hydrometeors

per temperature bin along all the flight tracks. The primary ice production peak which relies on the INP parameterisations clearly appears for the three schemes although with different amplitudes, that can be related to the above discussion on the various efficiencies of the INP parameterisations. In the HM temperature regime the Thompson scheme does show a clear increase of the number density around -5 C, while the Morrison schemes does not. The Milbrandt scheme shows more or less steady and larger concentrations than the Thompson, and the Morrison schemes, but with no clear signature of the HM process. The one to two orders of magnitude difference in the predicted total number concentrations of icy condensates between the Milbrandt scheme, and the other schemes (Figure 10b) below -10 C suggests that the INP parameterisations used in the Milbrandt scheme would blur any possible signal from the HM parameterisation (the -5 C peak). The fact that the Morrison scheme does not show the HM process while the Thompson scheme does can be explained by a threshold effect. The mechanism is triggered only if the snow content is $> 0.1 \text{ g kg}^{-1}$ and LWC $> 0.5 \text{ g kg}^{-1}$, values 5 which are above simulated averages for the snow and for LWC in the Morrison scheme (Figure 4c, and Figure 6). Interestingly the median number of ice and snow hydrometeors for the Thompson scheme agree within a factor of 2 with the observed median around -5 C. This shows that the Hallett-Mossop process can be to some extent correctly accounted for at this spatial resolution with the current parameterisation.

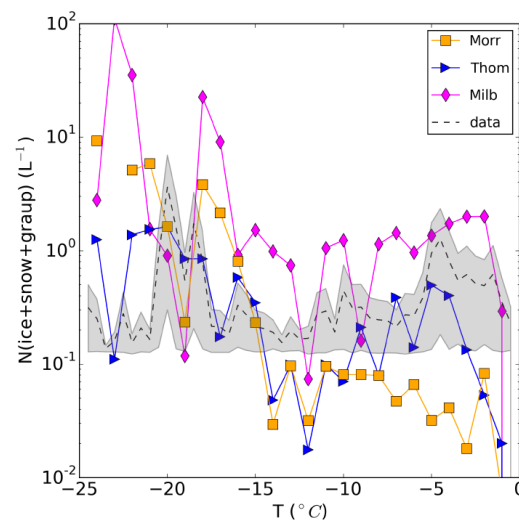


Figure 11. Median of the observations versus the temperature, as long as the simulated total number concentrations of hydrometeors along the flight tracks for Morrison, Thompson, and Milbrandt.

The Microphysics of Clouds over the Antarctic Peninsula – Part 2: modelling aspects within Polar WRF

Constantino Listowski¹ and Tom Lachlan-Cope²

¹British Antarctic Survey, NERC, High Cross, Madingley Rd, Cambridge, CB3 0ET, UK

¹now at: LATMOS/IPSL, UVSQ Université Paris-Saclay, UPMC Univ. Paris 06, CNRS, Guyancourt, France

²British Antarctic Survey, NERC, High Cross, Madingley Rd, Cambridge, CB3 0ET, UK

Correspondence to: C. Listowski (constantino.listowski@latmos.ipsl.fr)

Abstract. The first intercomparisons of cloud microphysics schemes implemented in the Weather Research and Forecasting (WRF) mesoscale atmospheric model (version 3.5.1) are performed in the Antarctic Peninsula using the polar version of WRF (Polar WRF) at 5 km resolution, along with comparisons to the British Antarctic Survey's aircraft measurements (presented in Part 1 of this work, Lachlan-Cope et al., 2016). This study follows previous works suggesting the misrepresentation of the cloud thermodynamic phase in order to explain large radiative biases derived at the surface in Polar WRF continent-wide (at 15 km or coarser horizontal resolution), and in the Polar WRF-based operational forecast model Antarctic Mesoscale Prediction System (AMPS) over the Larsen C Ice shelf at 5 km horizontal resolution. Five cloud microphysics schemes are investigated: the WRF Single-Moment 5-class scheme (WSM5), the WRF Double-Moment 6-class scheme (WDM6), the Morrison double-moment scheme, the Thompson scheme, and the Milbrandt-Yau Double-Moment 7-class scheme. WSM5 (used in AMPS), and WDM6 (an upgrade version of WSM5) lead to the largest biases in observed supercooled liquid phase, and surface radiative biases. The schemes simulating clouds in closest agreement to the observations are the Morrison, Thompson, and Milbrandt schemes for their better average prediction of occurrences of clouds, and cloud phase. Interestingly, those three schemes are also the ones allowing for significant reduction of the longwave surface radiative bias over the Larsen C Ice Shelf (Eastern Peninsula). This is important for surface energy budget consideration with Polar WRF since the cloud radiative effect is more pronounced in the infrared over icy surfaces. Overall, the Morrison scheme compares better to the cloud observation and radiation measurements. The fact that WSM5 and WDM6 are single-moment parameterisations for the ice crystals is responsible for their lesser ability to model the supercooled liquid clouds compared to the other schemes. However, our investigation shows that all the schemes fail at simulating the supercooled liquid mass at some temperatures (altitudes) where observations show evidence of its persistence. An ice nuclei parameterisation relying on both temperature and aerosol content like DeMott et al. (2010) (not currently used in WRF cloud schemes) is in best agreement with the observations, at temperatures and aerosol concentration characteristic of the Antarctic Peninsula where the primary ice production occurs (Part 1), compared to parameterisation only relying on the atmospheric temperature (used by the WRF cloud schemes). Overall, a realistic double-moment ice microphysics implementation is needed for the correct representation of the supercooled liquid phase in Antarctic clouds. Moreover, a more realistic INP alone is not enough to improve the cloud modelling, and water vapour and temperature biases also need to be further investigated, and reduced.

1 Introduction

Tropospheric Clouds in Antarctica are amongst the least well observed on Earth due to the remote environment and harsh conditions that make field observation difficult. As a result of this, no modelling study has ever focused on comparing the performances of WRF cloud microphysics schemes to in-situ cloud measurements. Yet, this is a necessary step to improve our ability to model the Antarctic atmosphere. Better understanding the meteorology is also crucial for providing reliable forecast to aircraft or ground operations in the Antarctic.

Much attention has focused on Antarctica's energy budget in recent years notably due to the West Antarctic Ice Sheet warming (O'Donnell et al., 2011; Bromwich et al., 2013b), and on large ice mass loss (gain) recorded in West (East) Antarctica (Harig and Simons, 2015). Assessing how atmospheric driven processes affect the evolution of Antarctica's ice mass and surface energy budget requires to improve our understanding, and modelling of the clouds in that region. Importantly, changes in microphysical properties of Antarctic clouds impact the atmosphere dynamics at lower southern latitudes and even at northern latitudes, since their altered radiative properties modify the North-South temperature gradient (Lubin et al., 1998).

The Antarctic Peninsula is characterised by high mountains forming a barrier to the dominant westerlies, and which roughly extends across the longitudes 67°W to 65°W at the latitude of Rothera station (67.586°S), with altitudes up to around 2500 meters in some places. This major topography feature causes significant differences between each side in terms of temperatures (Morris and Vaughan, 2003), and precipitations (King and Turner, 1997), and also aerosols and cloud microphysics (as concluded in Part 1 of this work, Lachlan-Cope et al., 2016). Significant climate changes have been recently observed across the Peninsula during the last few decades (O'Donnell et al., 2011; Turner et al., 2016). Interestingly, oceanic driven mechanisms are the main contributor to glaciers melting in the Peninsula (Wouters et al., 2015). In this context, improving the modelling of the different components of the energy budget of the Antarctic Peninsula is required to better understand its climatological evolution, and how atmospheric-driven processes act along with ocean-driven processes to impact Antarctica's ice mass balance, and temperatures. Clouds are one of the least well understood of the atmospheric components (Boucher et al., 2013; Flato et al., 2013).

Recent studies have pointed towards Antarctic clouds being responsible for large shortwave (SW) and longwave (LW) surface radiative biases (several tens of W m^{-2}) in high resolution models over the whole continent (Bromwich et al., 2013a), and more specifically over the Eastern Peninsula's Larsen C Ice Shelf (King et al., 2015). Improved cloud physics allowing for realistic ice supersaturations led to lower the surface energy budget biases in the RACMO2 high resolution climate model (van Wessem et al., 2014a). King et al. (2015) compared three mesoscale models simulations over the Larsen C Ice Shelf during a summer month, and showed how they differed in the amount of cloud liquid and cloud ice that were simulated. The authors suggested that this explained the comparatively different surface biases, and they pointed towards issues in modelling the thermodynamic phase of clouds, and more specifically the supercooled liquid component (liquid maintained at $T \leq 0^{\circ}\text{C}$). The modelling of the mixed-phase clouds needs to be improved in models, and the misrepresentation (underestimation) of supercooled liquid over Antarctica can be related to its poor representation over the surrounding Southern Ocean as a whole (Lawson and Gettelman, 2014).

A related issue deals with the initiation of the ice phase in clouds, which is driven by the Ice nucleating particles (INPs). They are the substrates needed to activate ice crystals growth either directly from the vapour condensing on the INP (deposition freezing), or from the freezing of supercooled droplets following immersion of, contact with, or condensation on, an INP (Hoose and Möhler, 2012). In the condensation case the INP act as a Cloud Condensation Nuclei (CCN) first to form a droplet. Homogeneous freezing of droplets (ie without the intervention of an INP) can occur at temperatures usually believed to be colder than -38° (Hoose and Möhler, 2012) although there are possible significant effects already below -30° (Herbert et al., 2015). In a remote place like Antarctica little is known about the exact nature of the INPs although studies have been identifying various plausible sources: biological sources from the snowy surface, blowing snow, sulfate particles resulting from sea-surface emissions, mineral dust lifted from ice-free regions or brought by winds from continental landmasses at lower latitudes e.g South-America. Many candidates are found in the literature to explain the presence of INPs in Antarctica (see Bromwich et al., 2012, for a review). Similar questions arise for INPs in marine air in remote places like in the middle of the Southern Ocean (Burrows et al., 2013), which surrounds the Antarctic continent. Regarding CCNs, which are needed to activate cloud droplet growth, sea salt is known to be an efficient substrate. Interestingly, its emission in polar region's boundary layer is believed to be enhanced in places where brine-rich snow covering sea ice can be lifted by the winds (Yang et al., 2008).

In the last decades, very localized ground-measurements using in-situ or remote sensing techniques have allowed to characterise microphysical properties of clouds (particle phase, particle size, crystals shape), however these observations are sparse (Lachlan-Cope, 2010; Grosvenor et al., 2012). Ground based remote-sensing measurements provide local continuous measurements making it possible to link clouds properties to precipitations or accumulation events (Gorodetskaya et al., 2015).

Two aircraft campaigns led by the British Antarctic Survey (BAS) took place during the summer 2010, and 2011, measuring cloud properties on both sides of the Antarctic Peninsula (Lachlan-Cope et al., 2016, hereafter referred to as Part 1). Analysis of some of the 2010 flights were already presented in Grosvenor et al. (2012) with a focus on cloud ice, and secondary ice multiplication processes. These two campaigns, and the surface radiative biases pointing towards a misrepresentation of antarctic clouds within high resolution models at 5 km resolution (King et al., 2015) or at coarser resolution (Bromwich et al., 2013a) motivate this first attempt to compare some of the existing cloud microphysics schemes implemented in the Weather Research and Forecasting atmospheric model (Skamarock et al., 2009) (WRF) v3.5.1, with simulations performed at 5 km resolution. We use the polar version of WRF (Polar WRF, Hines and Bromwich, 2008), which has optimised representation for polar regions in terms of surface properties (ice, snow, sea ice, and sea water) and processes (heat transfer between the surface and the atmosphere). Polar WRF is widely used by the Antarctic community and it is used by the Antarctic Mesoscale Prediction System (AMPS, Powers et al., 2012), which is an operational forecast model that provides support for international Antarctic efforts. Both Bromwich et al. (2013a) and King et al. (2015) relied on Polar WRF, and AMPS, respectively, in their study.

In section 2 we present the model settings along with the microphysics schemes used in this work, and explain their main characteristics. In section 3 we discuss simple results of radiation biases to illustrate the importance of cloud schemes on the Peninsula energy budget. In section 4 we compare modelling results to in-situ measurements already presented in Part 1, and evaluate the performance of the cloud microphysics schemes. In section 5 we comment on the performances of the cloud

schemes investigated, discuss sensitivity issues of the present study, and comment on the aspects to consider in future work for improving cloud microphysics parameterisations in Antarctica. In section 6 the main aspects of this work are summarized.

2 Observations, atmospheric model, and the cloud microphysics schemes

2.1 Overview of the airborne observations

5 Two campaigns of in situ cloud measurements took place on both sides of the Antarctic Peninsula (61–73°W) in February 2010, and January 2011, respectively (Part 1). The observations were made with the British Antarctic Survey’s instrumented Twin Otter aircraft (King et al., 2008) based at Rothera research station (67.586°S, 68.133°W). ERA-Interim reanalysis show an intensified northerly flow in 2011 to the west of the Peninsula, expected to bring warmer air. However colder temperatures were observed in the reanalysis, as well as in the radiosonde ascents made at Rothera (not shown), and from the aircraft
10 measurements (a tendency correctly reproduced in the simulations, see section 4.4). This can be explained by colder air being pulled from the Weddell Sea (to the east of the Peninsula) during the 2011’s campaign, following intensification and eastward movement of the Amundsen Sea low to the west of the Peninsula (Part 1, their Figure 3). Results on average cloud properties (predominantly stratus, or altostratus) comparing both campaigns and both sides of the Peninsula are presented in Part 1, and detailed results on some 2010 flights are presented in Grosvenor et al. (2012). The aircraft was fitted with various instruments
15 measuring notably temperature, pressure, humidity, turbulence, radiation as well as with a Droplet Measurement Technology Cloud, Aerosol, and Precipitating Spectrometer (CAPS) (Baumgardner et al., 2001). The CAPS has a Cloud and Aerosol Spectrometer (CAS), a Cloud Imaging Probe (CIP), and a Hotwire Liquid Water Content (LWC) Sensor. The CAS measures particle size (diameter) between 0.5 and 50 μm , and the Hotwire was used to validate the supercooled LWC as derived from the CAS, which cannot discriminate between liquid and ice. Also, the CAS showed a distinct peak in the size distribution in
20 the range 8 and 12 μm (in diameter) indicative of drop formations. The CIP images particles with sizes between 25 μm and 1.5 mm, with 25 μm pixel resolution. Particles smaller than 200 μm in size cannot be discriminated between crystals and droplets. Their number concentration is very small compared to the CAS, and therefore they were ignored (see also section 4.3.1 of this paper for the impact on the LWC). The ice water content was calculated using the Brown and Francis mass-dimension parameterisation (Brown and Francis, 1995). More details on data processing and the derivation of the ice crystal number
25 concentration are given in Part 1. Finally, the CIP samples at a rate of a little less than 10 L s^{-1} , hence the lower limit for the measured crystal number concentration of a little more than 0.1 L^{-1} .

2.2 Model settings

Polar WRF v3.5.1 was used with a downscaling method (Figure 1a) where a 45 km resolution domain contains a smaller 15 km resolution nest, which itself contains a smaller nest at 5 km resolution centered over the regions where the 2010 and 2011
30 flights took place (Figure 1b). The topography is from Fretwell et al. (2013). The simulation outputs of the highest resolution domain were used for the present analysis. We work at a similar horizontal resolution to K15 (5 km), and at a higher resolution

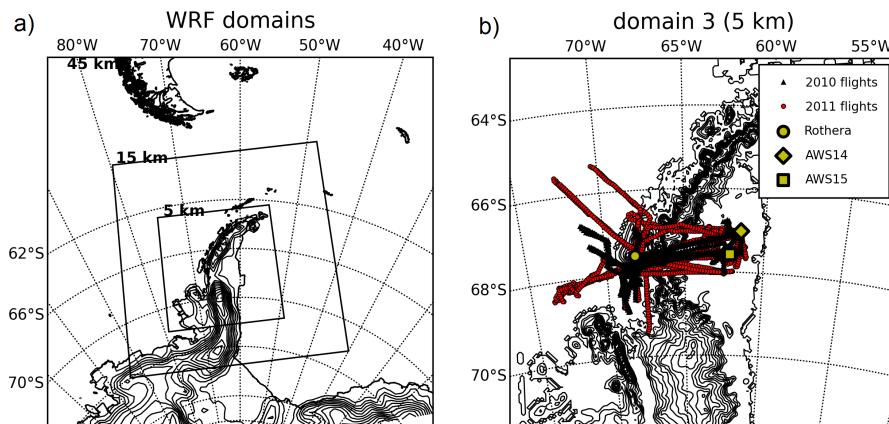


Figure 1. (a) WRF configurations for the three domains used for all simulations and (b) close-up on the highest resolution domain with detailed topography from Fretwell et al. (2013). The black triangles indicate the flight tracks of the 2010 campaign, while the red circles indicate the flight tracks of the 2011 campaign. The other markers indicate station (circle), and the Automatic Weather Stations (AWS) 14 (diamond) and 15 (square) located on the Larsen C Ice Shelf.

than Bromwich et al. (2013a) (60 km, and 15 km). Both studies pointed towards the clouds being responsible for the surface radiative biases measured, but they did not investigate the actual effect of using a different cloud microphysics schemes.

The simulation is one-way, in the sense that no information is passed in return from one domain to its parent domain. Table 1 summarizes the WRF settings used for the main physical processes, except for the cloud microphysics schemes, which are addressed in section 2.3. King et al. (2015) (hereafter referred to as K15) were interested in the surface radiative biases on the Larsen C Ice Shelf on the eastern side of the Peninsula. They used outputs from AMPS (built on Polar WRF v3.0.1 for the 2011 period). For consistency we use the same shortwave and longwave radiation schemes as in K15 (Table 1). More generally we are using the set of WRF physics parameterization used by the operational model AMPS, which should be a relevant framework to testing the cloud microphysics schemes in Antarctica. Regarding the boundary layer parameterization, Deb et al. (2016) showed that Polar WRF performances at the surface are most sensitive to the choice of the PBL scheme and show that the Mellor-Yamada-Janjic (MYJ) scheme is the best performer in terms of the temperature diurnal cycle (in west Antarctica) at 5 km resolution, and it is the one used in AMPS, and in our study.

One of the main differences with K15 is that our simulation is constrained horizontally and vertically, at the boundaries of the 45 km resolution domain, with ERA-Interim reanalysis data instead of Global Forecast System data (GFS, run by the U.S. National Centers for Environmental Prediction). ERA-Interim is the latest global atmospheric reanalysis (Dee et al., 2011) provided by the European Centre for Medium-Range Weather Forecasts (ECMWF). This reanalysis is based on archived observations from 1989 onward. It is obtained through data assimilation into an atmospheric model running at a resolution of 0.75×0.75 degrees, which roughly corresponds to 30 km in longitude by 80 km in latitude, at the latitude of Rothera station

(67.586°S). Bromwich et al. (2013a) showed that using ERA-Interim reanalysis for initial, and boundary conditions produces the best skills within Polar WRF.

We ran two sets of simulations. The first set spans the period 1st February 2010 to 5th of March 2010, and the second set goes from 1st of January 2011 to 11th of February 2011 (the first two days were not included in the analysis as part of the model spin-up). These two periods cover the time of the two aircraft campaigns (see Part 1), including the period during 2011 when a camp was set up on the Larsen C Ice Shelf, close to the Automatic Weather Station fourteen (AWS14, see Figure 1b), as described in K15.

2.3 Cloud microphysics schemes

We used five different microphysics scheme to assess their ability to model realistic clouds across the Antarctic Peninsula. None of the WRF microphysics schemes has been specifically developed for Antarctic clouds modelling. As no work has been done so far on comparing microphysics schemes implemented in Polar WRF with respect to their performances for Antarctic clouds, we used them as such with no modification. This appeared to be the most reasonable first step that can then help guide further development of Antarctic clouds microphysics modelling.

Generally speaking, each scheme is a bulk microphysics parameterisation (BMP) where either mass (Single-Moment – SM – scheme), or both mass and number density (Double-Moment – DM – scheme) of the various types of hydrometeors are independently predicted. In the DM case, the scheme allows for a more realistic behaviour of clouds. Indeed, predicting both the mass and the number density of hydrometeors allows the average particle size to be predicted, which in turn allows the modelling of all size-dependent processes like sedimentation, accretion, growth to be improved (Igel et al., 2015). All schemes have non-precipitable, and precipitable hydrometeors. The former (cloud droplets, ice crystals) are considered as having zero sedimentation velocity in the collection or accretion processes, in contrast to the latter (rain drops, snow crystals, graupel, or hail) that act as collector particles. Finally, we did not use any microphysics radius bin model (as opposed to the BMPs). They predict the evolution of cloud particles within given size bins, and allow for the prediction of the actual particle size distributions. Bin models are missing from WRF v3.5.1. However, a bin model is more demanding in terms of computer time, and BMPs are used in current global or regional atmospheric models, and in operational forecast models like AMPS. Table 2 highlights some aspects of the cloud microphysics schemes investigated in this study.

The actual default microphysics scheme of WRF is WRF Single-Moment 3 (WSM3), which has been discarded here because it does not allow for the existence of supercooled liquid droplets. Thus, our default reference scheme is the WRF Single-Moment 5 (WSM5) which allows for mixed-phase cloud formation (Hong et al., 2004). WSM5 is a SM scheme for all the hydrometeors. It is used in the operational model AMPS.

The WRF double-moment 6 (WDM6, Lim and Hong, 2010) is an improvement of WSM5, in which droplets and rain are both treated with DM schemes, graupel is included, and all the ice phase particles are treated with a SM scheme. It is used here in order to test the improvement of the prediction of the supercooled liquid phase that one could expect from the use of a more sophisticated parameterisation for the liquid phase (DM instead of SM as in WSM5).

The Morrison scheme (Morrison et al., 2005; Morrison et al., 2009) is a full DM scheme for all icy hydrometeors, and rain, and SM for water droplets. The Morrison scheme requires the coupling to the WRF Chemistry module (Peckham et al., 2011) in order to act as a DM scheme for the cloud droplets, and since such coupling was not available we used the Morrison scheme as a SM scheme for the water droplets. This scheme is used in the Arctic System Reanalysis (ASR), which is based on
5 Polar WRF as well. It slightly improved the modelling of the clouds in the northern polar summer compared to WSM5 at 30 km resolution (Wesslén et al., 2014), and this paper investigates its ability to better represent the clouds in the southern polar region.

The Thompson scheme (Thompson et al., 2008) has a state-of-the art parameterisation of snow, which relies on extensive flight measurements, and it uses a more realistic size dependent density for snow particles. The latter are treated as non-
10 spherical, and their density decreases with increasing size. This was identified as having a major influence on the production of supercooled drops mainly because of a decreased efficiency in the riming process resulting in longer lasting supercooled drops (Thompson et al., 2008).

Finally, The Milbrandt-Yau scheme (Milbrandt and Yau, 2005a, b) (hereafter designated as Milbrandt) is a full double moment scheme (with the shape parameters of the particle distribution being fixed). It is used here in order to test the ability
15 of a full double moment scheme to predict supercooled drops better than the Morrison or the Thompson schemes.

Table 3 details the way the cloud schemes treat the initiation of the cloud ice phase, and the cloud liquid. The initiation of the ice phase is the most complex part, and it relies on INP parameterisations. They diagnose the number of INPs, hence the number of activated crystals, accounting for the various freezing modes described in the introduction. The INP parameterisations rely on the atmospheric temperature only. They are used in various ways by the different cloud microphysics schemes as illustrated
20 in Table 3. They deal with primary ice production (droplets or vapour converted to ice through interaction with INPs), as opposed to secondary ice processes, which result from the interaction of already formed crystals with other crystals or with supercooled droplets. Finally, the liquid phase relies on a fixed number of droplets or a predicted number of activated CCN (hence number of drops), depending on the cloud scheme. The liquid phase is formed after the ice microphysics is computed provided there is still an excess of vapour compared to equilibrium (ie if $S_w > 1$, where S_w is the saturation ratio with respect
25 to liquid water).

3 Preliminaries : results in radiation biases

Large biases in both surface downward shortwave (SW, solar flux) and longwave (LW) radiation were reported east of the Peninsula over the Larsen C Ice Shelf by K15. The authors compared the summertime surface energy budget as simulated for January 2011 by three mesoscale models: AMPS, The UK Met Office Unified Model (UM) (see Wilson and Ballard, 1999, for
30 the cloud scheme), and the Regional Atmospheric Climate MOdel (RACMO2) version 2.3 (see van Wessem et al., 2014a, for the cloud scheme). A field camp was established close to AWS14 (see Figure 1b) where radiosonde ascents allowed the water vapour column density to be calculated. AWS14 is fitted with SW and LW radiometers. K15 showed that all models mostly overestimated SW radiation by several tens of $W m^{-2}$ (positive bias) while they underestimated LW radiation (negative bias).

They pointed towards the lack of simulated clouds that blocked the incoming shortwave solar radiation, and emitted thermal radiation back to the surface. The only exception was noted for the UM model which had several tens of W m^{-2} of negative bias in SW suggesting an overestimation of the cloud cover. AMPS simulated clouds predominantly composed of ice with very little or even zero liquid water, during this period over AWS14, providing an explanation to the very large surface radiative biases, especially in SW to which small droplets are the most responsive. Ice clouds, however, were simulated, and K15 pointed towards a misrepresentation of the actual phase of the clouds to explain the biases observed.

Following K15, Table 4 shows average biases of daily averaged SW and LW downward surface fluxes. They were derived by subtracting the observed value to the modelled value. Three sites were selected: British Antarctic Survey's Rothera station, on the western side of the Antarctic Peninsula, and two Automatic Weather stations - AWS14 and AWS15 - on the Eastern side of the Peninsula on the Larsen C ice shelf (see Figure 1b). Both AWS are about 70 km apart on the ice shelf. Table 4 also indicates whether the difference between WSM5 (used in AMPS) and the other schemes is statistically significant (with a Student t-test).

3.1 The particular case of AWS14 in January 2011

We first compare results obtained by K15 with the AMPS model over the period 8 January 2011 to 8 February 2011 at AWS14 (see their Table 3), to our results obtained with the WSM5 scheme over the same period (Table 4, fourth column of results). Their computed biases are 56 W m^{-2} and -10 W m^{-2} in SW and LW, respectively). Ours are 53 W m^{-2} and -20 W m^{-2} , respectively. Discrepancies in biases can result from the combination of different settings in the AMPS (forcing, number of vertical levels, domain boundaries). However, we do obtain the same orders of magnitude and same signs of biases as K15, consistent with a lack of clouds. A striking result is that the Morrison scheme reduces the biases in both SW and LW in a statistically significant way at the 99% level, while the Milbrandt, and the Thompson schemes reduce it significantly in LW only.

Figure 2 (bottom) shows the cloud liquid mass integrated over the entire atmospheric column (kg m^{-2}) for the different simulations as a function of time, in the model gridbox corresponding to the AWS14 location in 2011. Figure 2 (top) shows the simulated column density of water vapour compared to the radiosonde ascent measurements from the field camp at AWS14 (presented in K15, and plotted in their Figure 7). The modelled water vapour is consistent with observations in terms of trend and value (within $\pm 1 \text{ kg m}^{-2}$), between day 8 and day 32. The simulations give similar values within $\pm 1 \text{ kg m}^{-2}$, except between day 15 and day 22 (where no observation is available), and all the simulations capture the sharp increase by 6 kg m^{-2} measured around day 28. Using the Morrison scheme, twice to four times more liquid cloud mass is simulated than when using the WSM5 scheme (Figure 2 bottom). The Milbrandt, and Thompson schemes lead to intermediate amounts between WSM5 and Morrison, and WDM6 is similar to WSM5. The larger amount of liquid clouds simulated with the Morrison scheme compared to WSM5 is consistent with its smaller SW and LW biases at AWS14 in 2011 (Table 4). This is also in line with K15's conclusion that the thermodynamic phase of the clouds was responsible for the SW and LW biases they found in AMPS over the Larsen C. The Thompson scheme, and the Milbrandt schemes do have lower SW biases than WSM5 as well, however the

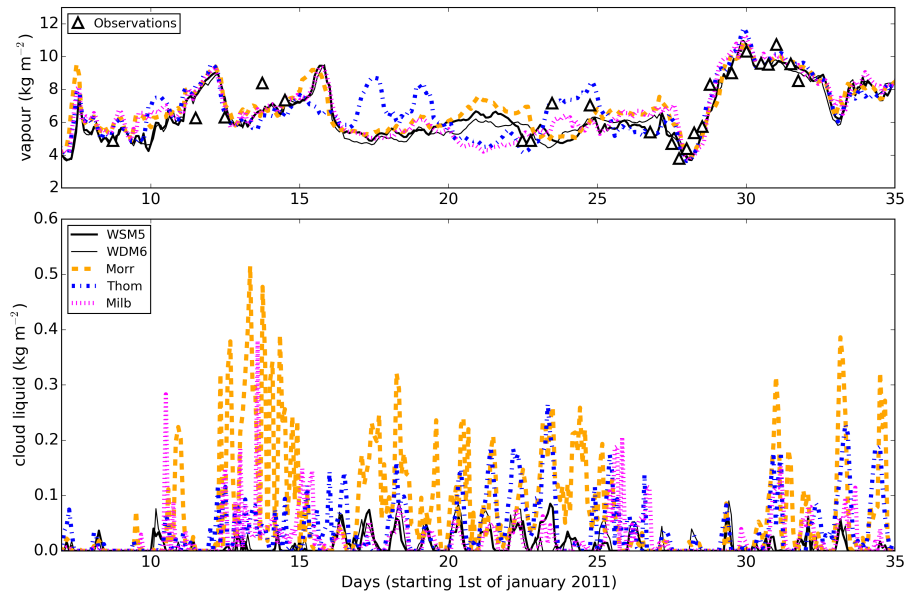


Figure 2. (Top) Time series of column density of water vapour (kg m^{-2}) for the different simulations computed in the model gridbox corresponding to the AWS14 location, in 2011, along with the radiosonde measurements from the field camp (described in K15, see their Figure 7). (Bottom) Time series of the column density of the cloud liquid (kg m^{-2}) for the different simulations.

improvement is smaller than with the Morrison scheme, and less (or not) statistically significant. However, it is still significant for LW radiation.

The total ice mass is similar from one scheme to another, as shown by Figure 3 (top). However an important difference arises when considering the cloud ice crystals mass only (ie the pristine ice – ignoring the main precipitable particles like snow and graupel particles); WSM5 and WDM6 simulate three to four times more ice crystals mass than the other schemes (Figure 3, bottom). The Milbrandt scheme leads only occasionally to as much ice mass as WSM5, and WDM6, around the 19th, the 25th and the 30th of January. Graupel is mainly absent except when using the Milbrandt schemes which leads to low amounts around 0.05 kg m^{-2} on rare occasions (not shown). Overall, the main difference in the cloud microphysics between the various simulations at AWS14 is the ability of the cloud schemes to sustain supercooled liquid drops, which in turn can explain differences in the SW and LW surface biases. The other difference lies in the distribution of the mass within the total ice phase between cloud ice crystals and snow particles.

3.2 General results in radiation biases

For the eastern side of the Peninsula (AWS14 and AWS15), the biases shown in Table 4 (right part) demonstrate the importance of the choice of the microphysics scheme for the surface energy budget of the Larsen C in Polar WRF. Similar biases (sign, and order of magnitude) are observed at a given year and for a given scheme, between AWS14 and AWS15. This is consistent with the stations being 70 km apart from each other on the ice shelf, which consists of a relatively flat surface covered with

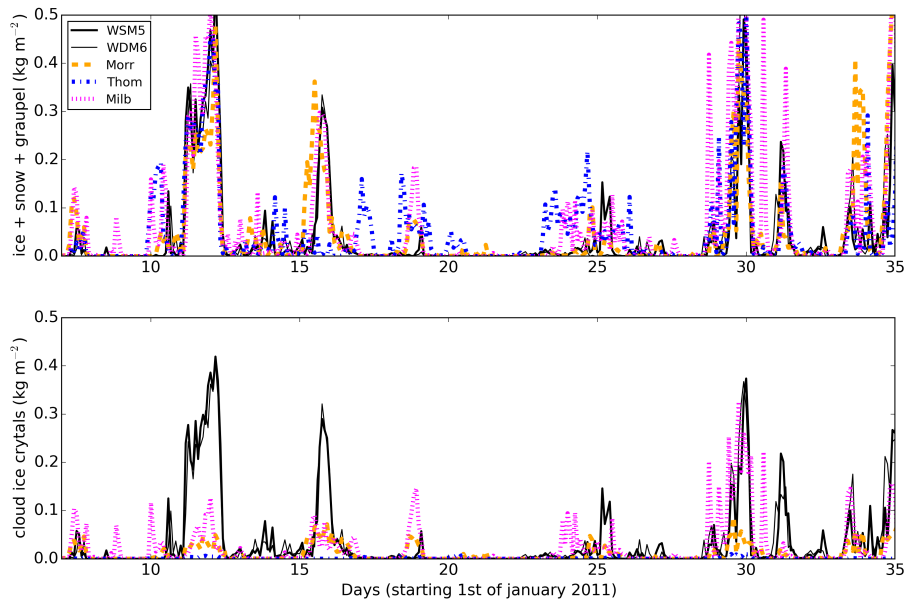


Figure 3. (Top) Same as Figure 2 for the total ice phase (ice, snow, and graupel particles). (Bottom) Same as Figure 2 for cloud ice particles only.

snow, and where the large scale influences are likely to be similar in absence of significant local variations in the topography or the nature of the surface. A remarkable result is that the LW bias is significantly reduced during both periods of interest on the Larsen C ice shelf using the Morrison, Thompson, or Milbrandt schemes, compared to WSM5 (or WDM6 – not shown) as can be seen from the lower right part of Table 4. However, the Thompson, and the Milbrandt schemes still have a negative LW bias, while Morrison’s is slightly positive and gives on average the smallest LW bias at both AWS stations for both years. The standard deviation of daily-averaged measurements remains high but statistical tests show that the differences between WSM5 and the three other schemes are significant, mainly at the 99% level in 2011, and mainly at the 95% level in 2010. The SW bias is significantly reduced only with the Morrison scheme in 2011. However no improvement occurs in 2010 for the SW bias. The Milbrandt, and the Thompson scheme’s SW biases are slightly lower in 2011 with differences to WSM5 that are significant at the 90% level. In 2010 all the schemes have a large negative SW bias with the largest amplitude attributed to the Morrison scheme.

For the western side of the Peninsula (Rothera station), SW biases are always positive, and LW biases always negative, whatever the cloud scheme or the year considered (left part of Table 4). All simulations consistently show this imbalance suggesting no improvement in cloud simulation. Besides, no statistically significant difference is observed between WSM5, and the other schemes. Note that almost no cloud liquid water (not shown) is simulated above Rothera (as opposed to AWS14, Figure 2), whatever the cloud scheme used, in line with the persistent large SW and LW biases. Ice and snow (graupel),

however, are formed in similar amounts to the ones shown in Figure 3 (not shown). Overall, Table 4 demonstrates the high sensitivity of the simulated downward radiation fluxes to the microphysics scheme used in Polar WRF.

A major issue in assessing the performances of the cloud microphysics schemes by investigating radiation biases is that it does suppose that the appropriate information is passed on from the cloud scheme to the radiative scheme. This aspect can explain the apparent paradox of the significant improvement of the LW bias to the east of the Peninsula with three schemes, while no concomitant SW bias improvement is being observed. The Discrepancies in SW and LW biases improvements will be further discussed in section 5.1. Radiative schemes themselves also require careful examination as they also rely on various assumptions, and simplified geometry to retrieve SW and LW fluxes. The radiative schemes that we used was chosen for consistency with K15, in order to compare their conclusions (using AMPS) to ours (using Polar WRF). We do not intend here to investigate the radiative schemes implemented in WRF. For further assessment of cloud microphysics schemes performances and behaviours at a much wider scale, we now compare the simulations outputs to each other and to the cloud microphysics properties as measured during the BAS aircraft campaigns that took place over the Antarctic Peninsula (presented in Part 1).

4 Results: simulated clouds as compared to observations

4.1 General trends for simulated clouds across the Peninsula

The topography of the Antarctic Peninsula (Figure 1) makes it interesting to focus on zonal distribution of latitudinal averages for the Liquid Water Content (LWC, in g kg^{-1}), and the Solid Water Content (SWC, g kg^{-1}). SWC comprises ice, snow and graupel mass. It is different from the Ice Water Content (IWC), which consists only in the mass of the cloud ice crystals. LWC and SWC were respectively averaged between latitudes 65.5°S and 68.5°S and altitudes below 4500 m. This geographical area includes the region where both flight campaigns took place in summer 2010, and 2011 (Figure 1b). For simplicity we designate each simulation by using the name of its cloud microphysics scheme.

Both periods of interest display the same relative trends, and we present an average over both periods to give an overview. Averages are computed considering either all values, including null instances (LWC_0 and SWC_0 in Figure 4a and Figure 4b, respectively) or only strictly positive values (LWC and SWC in Figure 4c and Figure 4d, respectively). Thus, we always have $\text{LWC}_0 \leq \text{LWC}$, and $\text{SWC}_0 \leq \text{SWC}$. LWC (resp. SWC) gives the liquid (resp. ice) content that is simulated disregarding how often the clouds form. LWC_0 (resp. SWC_0) describes a more realistic average behaviour since it also accounts for the ability of the scheme to lead to liquid (resp. ice) cloud formation, more or less often.

For all the simulations LWC, and LWC_0 are in the interval $0.05\text{-}0.14 \text{ g kg}^{-1}$, and $0.002\text{-}0.03 \text{ g kg}^{-1}$, respectively. SWC, and SWC_0 are in the interval $0.02\text{-}0.08 \text{ g kg}^{-1}$, and $0.01\text{-}0.035 \text{ g kg}^{-1}$, respectively. The lower limit of LWC_0 (0.002 g kg^{-1}) is due to WSM5 decreasing over the mountains down to 0.002 g kg^{-1} around 65°W . For the other cloud schemes, $\text{LWC}_0 \geq 0.01 \text{ g kg}^{-1}$. There is roughly a factor of 5 to 10 between LWC_0 and LWC while there is a factor of 1.2 to 2 between SWC_0 and SWC. The liquid phase features more important changes (from null to non-null values) than the total ice phase, which is simulated more frequently.

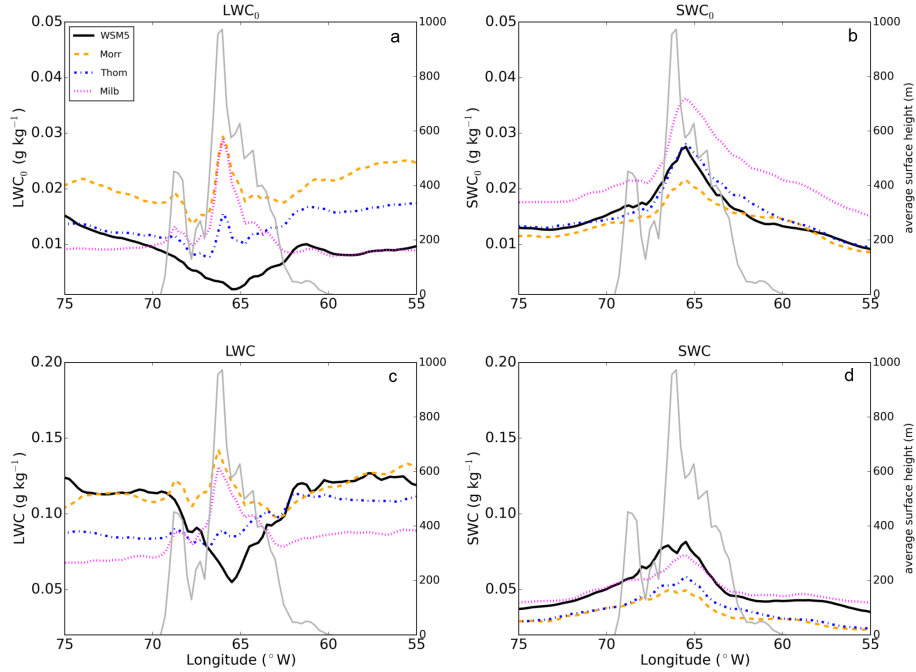


Figure 4. Longitudinal distribution of latitudinally ($65.5\text{--}68.5^\circ\text{S}$) averaged LWC and SWC (g kg^{-1}) over both periods of interest for schemes WSM5, Morrison, Thompson, and Milbrandt. The average is computed over all gridboxes and times leading to (a) LWC_0 and (b) SWC_0 , or only over gridboxes and times where values are non-null leading to LWC (c) and SWC (d). The grey thick line shows the surface height averaged over the same region and labelled on the right vertical axis of each plot. Note the identical scales used for the vertical axes for LWC_0 and SWC_0 , and LWC and SWC, respectively.

WSM5 strikingly differs from the Morrison, Thompson, and Milbrandt schemes in that its LWC and LWC_0 decrease above the Peninsula mountains. LWC drops from $\sim 0.12 \text{ g kg}^{-1}$ by more than 50% from 70°W to 65°W , before increasing back from 65°W to 60°W to $\sim 0.12 \text{ g kg}^{-1}$ (Figure 4c). Except East of 62°W where WDM6's LWC is larger than WSM5 by less than 0.03 g kg^{-1} (not shown), both schemes display very similar averages for LWC and SWC, and we only show WSM5. LWC is much steadier for the three other schemes, and a sharp increase for the Milbrandt and Morrison scheme is observed above the highest terrains, caused by the orographic forcing induced by the westerlies or the easterlies (see section 4.2).

The first obvious assessment with respect to the ability of the cloud schemes in forming (supercooled) liquid clouds is that WSM5 (WDM6) leads to less supercooled liquid mass compared to the Morrison, Thompson, and Milbrandt schemes across the Antarctic Peninsula. Eastward of 62°W , WSM5's LWC_0 is 100% (50%) smaller than Morrison's (Thompson's), while it is similar to Milbrandt's (Figure 4a). However, in the central region over the mountains, WSM5 (and WDM6) leads to less liquid mass by up to an order of magnitude than the three other schemes. Westward of 62°W LWC_0 is similar for WSM5, Milbrandt, and Thompson; they are all twice as small as Morrison's LWC_0 . WSM5 does not lead as often as the other schemes to supercooled liquid formation, what is illustrated by its lowest LWC_0 values, yet it does simulate as large average liquid water

contents as the other schemes when and where liquid forms (similar LWC), except in the central region where orographically induced clouds have systematically less liquid water. The ice phase instead shows a similar behaviour across the different cloud schemes with an increasing SWC closer to the high altitude topography, due to orographic forcing. SWC_0 is similar for WSM5, Morrison and Thompson while with Milbrandt it reaches 50% larger value largely due to the graupel mass (not shown).

5 Comparing LWC_0 and SWC_0 , we see that the simulation with the Morrison scheme is the only one sustaining supercooled liquid mass more frequently than ice mass ($LWC_0 > SWC_0$ by a factor of >1 to 2). For the Thompson scheme, $LWC_0 \sim SWC_0$ on average (but $LWC_0 < SWC_0$ over the mountains, and $LWC_0 > SWC_0$ East of the Larsen C). The Milbrandt scheme leads more often to ice mass formation than liquid mass ($LWC_0 < SWC_0$) by a factor of less than 2 at all longitudes, and WSM5 by a factor comprised between 1 and 5. Finally, the simulation with WSM5 (WDM6) is the only one resulting in an anticorrelation
10 between LWC (LWC_0) and SWC (SWC_0) with an increase (resp. decrease) for the cloud ice (resp. cloud liquid) over the mountains.

4.2 Dynamics, and microphysics structure of the simulated clouds

The Antarctic Peninsula mountains acts as a barrier to the westerly, or easterly winds that drive the formation of orographic clouds. As a complement to the general picture given above, we identified two periods of sustained westerly, and easterly
15 wind regime, respectively. We isolated the period 7 January to 10 January 2011 when westerlies prevailed almost exclusively, and similarly the period 11 January 2011 to 17 January 2011 when the easterly regime prevailed. Note that the average wind directions and speed, and their relative variations are in agreement with upper-air measurements performed daily from Rothera station (not shown), if not always quantitatively at least qualitatively, as well as with measurements from the aircraft (not shown).

20 Figure 5 shows time- and space-averaged transects of the hydrometeors' mass (including null instances) across the Peninsula on a ~ 100 km wide (67 - 68° S) latitudinal band approximately centered on Rothera station, for WSM5 (Figure 5a, and Figure 5b), and the Morrison scheme (Figure 5c, and Figure 5d). The westerly cases (b + d), and easterly cases (a + c) show the orographic clouds microphysical structure. It also illustrates the very different contexts of Rothera station on the western side, and of AWS14 and AWS 15 on the eastern side. Rothera itself is in the lee of a mountaneous feature (Adelaide island) and the
25 topography adds to the complexity in simulating the clouds compared to the flat Larsen C ice shelf.

WSM5 predicts completely glaciated clouds on the Peninsula, and liquid clouds only away from the mountains with a very limited vertical extent (up to 500 meters above surface). The Morrison scheme maintains mixed-phase clouds across the region and at much higher altitudes. This fact alone is in better agreement with observations from the aircraft, which measured almost exclusively mixed-phase clouds (Part 1). Over the Ice shelf the snow increases from 0.01 to 0.05 g kg^{-1} as we get closer to the
30 mountain barrier for all the schemes (similar amounts are simulated on the western side during the westerly regime). However, WSM5 simulates an IWC (orange lines) as large as the snow particles mass (red lines) down to the surface, contrasting very much with the Morrison scheme.

Note that WDM6 (not shown) gives similar results as WSM5. Also, the microphysical structure of the clouds predicted by the Thompson scheme, and the Milbrandt scheme (not shown) are similar to Morrison's. On either side of the Peninsula,

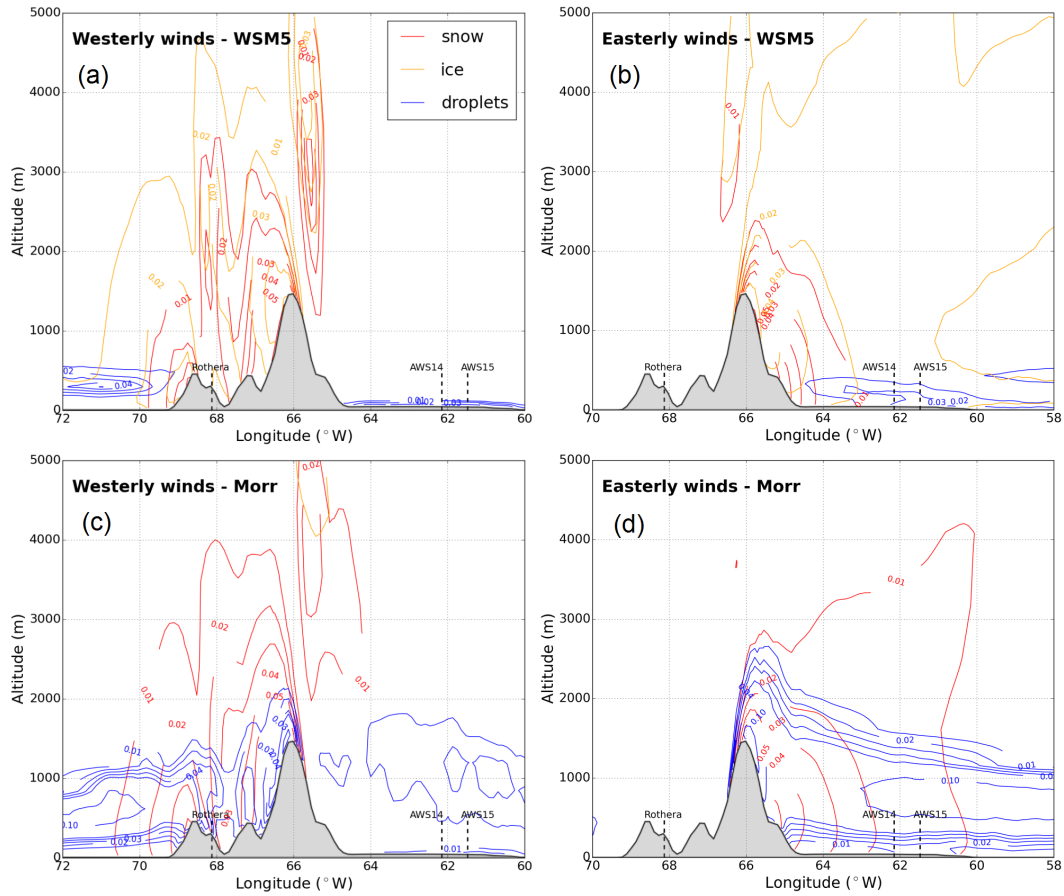


Figure 5. Transect of the cloud microphysics for WSM5 (a and b), and Morrison (c and d) averaged over a period (7-10 Jan 2011) dominated by westerly winds (a and c), and over another period (11-17 Jan 2011) dominated by easterly winds (b and d). The transect is approximately centered on Rothera Station (67.586°S), and it is an average over a 100 km wide latitudinal band. The longitudes of Rothera, AWS14, and AWS15 are indicated.

downwind, the Morrison schemes forms the most abundant mixed-phase cloud layer with $\text{LWC} \sim 0.1 \text{ g/kg}$, and the clouds extends almost down to the surface ($\text{LWC} \sim 0.01 \text{ g kg}^{-1}$), whereas the Milbrandt, and the Thompson schemes form less than half of that maximum amount, in line with the general picture given in section 4.1. Also, The Milbrandt scheme forms a significant amount of cloud ice crystals (IWC) above 3000 meters, as well as graupel in the mixed-phase orographic clouds above the windward slopes (not shown), which are absent from the average transects of the Morrison, and the Thompson simulations.

4.3 Microphysics schemes performances West and East of the Antarctic Peninsula

4.3.1 Liquid phase

To assess the performances of the different cloud schemes we compare the LWC measured from the aircraft to the simulated LWC by restraining the latter to the model gridboxes corresponding to the flight tracks. We only consider non-null LWC values (LWC > 0.001 g kg⁻¹). For each data point, the closest (both in time and space) gridbox value is extracted from the model. Latitudinal averages are derived for each flight per 0.5° longitude bins, for simulations and observations. At the latitude of Rothera station (67.586°S) this corresponds to ~10 kilometres (ie two gridboxes). Then, global west and east averages are derived, corresponding to longitudes westward of 67°W, and to longitudes eastward of 65°W, respectively (as in Part 1). LWC is derived as presented in Part 1 using the droplet size distribution obtained from the Cloud and Aerosol Spectrometer (CAS). The unknown thermodynamic phase of the smallest particles seen by the Cloud Imaging Probe (CIP, see Part 1), but not resolved, and that could either be drops or small crystals, may induce a bias in the derivation of LWC. However, if all of them were counted as droplets, they would increase LWC by ≤ 8% for all flights, except two flights in 2010 (13%, and 30%) and one flight in 2011 (12%). This bias does not alter the results, and conclusions below. **More information on the instruments and the measurement can be found in Part 1 (their section 2.1).**

Figure 6 shows the scatter plots of simulated LWC versus observed LWC for 2010 (Figure 6a, and Figure 6b) and 2011 (Figure 6c, and Figure 6d), and for either side of the Peninsula, West (a, and c), and East (b, and d). Regional (East or West) averages are represented by the largest bold markers while smaller markers relate to individual flight averages. Note that the width of the large markers is larger than the length of the errorbar associated with the aforementioned error (bias) related to the LWC derivation. The numbers shown next to each scheme's markers in the legend (in the form n₅;n₅₀/N) indicate the number of flight tracks for which the simulation forms at least an average of 5% (n₅) or 50% (n₅₀) of the observed average LWC, over the total number of flight tracks (N) having measured cloud liquid. We refer to those as the n₅ criterion, and the n₅₀ criterion, respectively.

Three results stand out. First, all the schemes perform worst on the West compared to the East in terms of number of tracks with simulated clouds (n₅, and n₅₀ criteria), except for the WSM5 scheme which performs equally bad on both sides. Second, the WSM5 scheme has the lowest numbers of flights with some liquid clouds simulated (n₅ criterion). For the Morrison, Thompson, and Milbrand schemes, about (Figure 6a) or much less (Figure 6c) than 30% of flights are predicted with some substantial supercooled liquid (n₅₀ criterion) in the West, and more than 60% of them in the East (Figure 6b and d). Third, the Morrison scheme performs on average the best in reproducing observed LWC on the western and the eastern portions of the flight tracks, with larger values of LWC simulated compared to the other schemes. When considering the n₅ criterion, the Thompson, and Milbrandt schemes show equally good scores compared to Morrison suggesting the same ability to initiate a non-negligible supercooled liquid phase, as opposed to WSM5 (and especially on the eastern side). However, overall the Morrison scheme performs better because it has an averaged simulated LWC closer to the observed one within a factor of less than 2 (except in East 2010 – Figure 6b – where it simulates an average LWC three times larger than the observations).

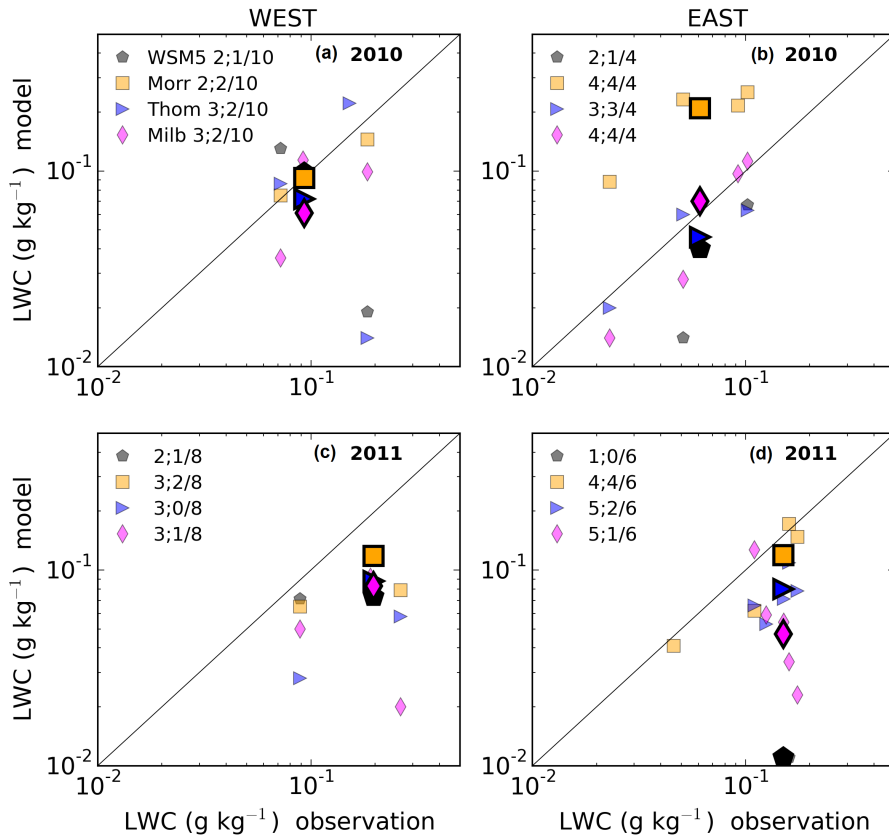


Figure 6. Scatter plot of simulated LWC versus observed LWC in 2010 (a and b) and in 2011 (c and d) on the western side of the Peninsula (a and c) and on the Eastern side of the Peninsula (b and d). Light markers show averages per flight track while bold markers give the average of all the tracks on each side of the Peninsula, respectively. The numbers next to each scheme's marker in the legend ($n_5;n_{50}/N$) gives the number n_5 (resp. n_{50}) of simulated flights with a simulated LWC at least 5% (resp. 50%) of the observed LWC, to the total number N of flights measuring an average LWC. Note that in (c) the bold markers (total average) overlay some light markers (flight averages), what explains the actual higher position of the total average on the graph compared to the other discernible lower flight averages.

Those averages do not take into account the duration over which such values are observed. Thus, we use an additional metric that is the average time spent in cloud (or instances of cloud occurrences) on either side of the Peninsula. The average ratio of the time spent in clouds in the model (with $LWC > 0.01 \text{ g kg}^{-1}$) over the one in the measurements, is given in Table 5 for each side, and year. The average is derived as an average of the flight-averages. Over both periods the best scheme appears as the Morrison since the Thompson, and the Milbrandt schemes have very low occurrences of clouds compared to the observation on the western side in 2011 with 4% in 2011 and 5%, respectively. On the eastern side, WSM5 has the poorest performance (<1%), and the Morrison scheme has twice as much occurrences of clouds (although still quite low) in 2011 compared to the two other schemes, and it overpredicts the formation of clouds in 2011 (215%), although not the average LWC (Figure 6d).

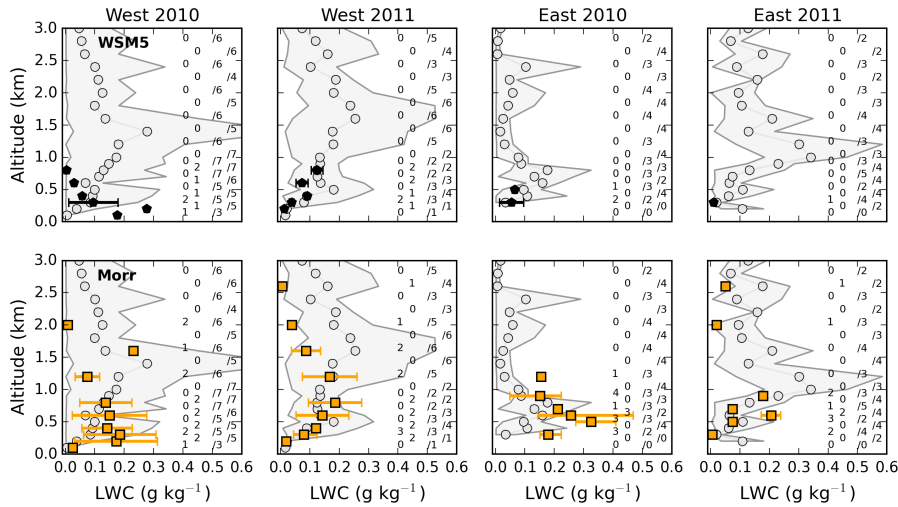


Figure 7. Averaged vertical profiles of non-null LWC for WSM5 (top), and Morrison (bottom) and for the observations, west (left) and east (right) of the Peninsula. Grey markers indicate the measured average at each altitude, while the shaded area gives the range of the observed flight averages at each altitude. Similarly, coloured markers and error bars relate to the cloud schemes. The numbers indicate how many simulated flight averages were used to derive the global average at each altitude for each scheme, as compared to (“/”) for the observations.

Average vertical profiles of cloud liquid (and ice) were also derived for flights measurements as well as for the model outputs. The altitude grid on which flights observations, and model outputs, were averaged on is finer in its lower layers with one level every 100 meters below 1100 meters, and every 200 meters above 1100 meters. At each altitude levels the average of the flight averages is computed so that every flight has the same weight. Model altitude levels are separated by less than 1000 meters at the highest altitude levels of the atmospheric column. However, below 4500 meters, where the flights took place, the maximum model level separation is approximately 500 meters. Thus, any data point level is less than 250 meters away from the closest model level (less than 100 meters below 1100 m).

Figure 7 compares vertical distribution of observed (grey circles) and simulated (coloured markers) non-null average LWC ($>0.001 \text{ g kg}^{-1}$), for WSM5 (top), and the Morrison scheme (bottom). The grey shaded area shows the spread of all flight averages. The errorbars show the spread of the simulated flight averages. The numbers at each level indicate how many simulated flights with non-null averages are used to derive the total average of each level, for the simulations as compared to (“/”) the observations.

WSM5 scheme does not form liquid clouds above 800 meters on the western side of the Peninsula, as well as no liquid clouds above 500 meters on the eastern side, during both periods of interest. Liquid clouds were observed as high as 4400 meters. The numbers at each level show that WSM5 simulates fewer occurrences of liquid compared to the Morrison scheme, which still underpredicts the occurrences of liquid clouds. The Morrison scheme shows liquid cloud formation up to 2500 meters although only, very few instances above 1500 m.

WDM6 shows no improvement compared to WSM5 (not shown). The Milbrandt, and the Thompson schemes simulate liquid clouds more often than WSM5 in the lowest layers, but no clear systematic difference emerges between those two and the Morrison scheme. The Morrison scheme simulates best the increasing trend of LWC with altitude in the West in 2011. It has the largest LWC below 1000 meters (by 0.1-0.2 g kg⁻¹) on either side of the Peninsula in 2010 compared to other schemes, while LWC is comparable for all the three schemes in 2011 (not shown).

4.3.2 Ice phase, and mixed phase

For completeness we compare the simulated Solid Water Content (SWC, g kg⁻¹) to the observed ice mass. Figure 9 is the same as Figure 6 but for SWC, with the addition of the corresponding Ice Water Content (IWC) regional averages shown as grey bold markers. (The latter are slightly shifted rightwards by 50% of the observed value on the x-axis, in order to be visible). The smaller and lighter markers are individual flight averages. The same n₅ criterion, and n₅₀ criterion as in Figure 6 are used and referenced on the Figure 9 next to each cloud scheme's name.

As mentioned in Part 1, there is an uncertainty in the smallest particles detected by the Cloud Imaging Probe (CIP), however they contribute to a negligible amount of the total measured ice mass. To the other end of the size distribution, the maximum cut-off for detected ice particles is about 1.5 mm in size (diameter). Thus possible larger particles that could significantly add to the mass are not detected. However, in order to have an estimate of the error caused by the missed larger particles we approximated and extrapolated the average size distribution of the crystals for each flight (examples are shown in Figure 8a, and 8b), using an exponential distribution of the form $N(D) = N_0 \exp(-\lambda D)$ (known as Marshall-Palmer distributions) commonly used for the rain and the ice hydrometeors in the cloud microphysics schemes (e.g Morrison et al., 2009). Using the exponential distribution and the mass-diameter law, we derive an ice water content below 1.5 mm, and above 1.5 mm, respectively. In order to derive the mass for $D \leq 1.5$ mm we integrated over the crystal sizes starting from the peak diameter of the distribution, and up to 1.5 mm. The peak diameter of the observed ice crystal distribution is located in the range 250-425 μm (with an average of 315 μm), value from which the exponential distribution can approximate the distribution of the largest crystals. Then, the ice water content for particles with $D > 1.5$ mm, and up to an arbitrary limit of 3.2 mm, was derived (setting the upper limit to even larger sizes does not change the resulting additional ice mass given the even lower amounts of crystal number concentration predicted by the exponential distribution) The ratio of both ice water contents allow to estimate the relative error caused by the undetected particles on the measured SWC when assuming an exponential distribution. For the 2010 flights, this average error is about 5%, including an outlier flight with a 33% error (ignoring this flight brings the average error to 2%). For the 2011 flights, the average relative error is about 8%, including an outlier with a 65% error (shown in Figure 8c) (ignoring this flight brings the average error to 2%). The large error derived for the two flights is related to a shoulder of the crystal distribution for the larger particles, leading to an exponential distribution predicting a number concentration of the largest particles likely to be in large excess compared to the actual one (Figure 8c). Overall, these estimates of the relative errors on SWC do not alter the main conclusions presented here.

Table 3 gives the cut-off radii between ice particles and snow particles in the different cloud schemes. The different definitions of the icy hydrometeors across the cloud schemes add to the difficulty of performing comparisons between the schemes

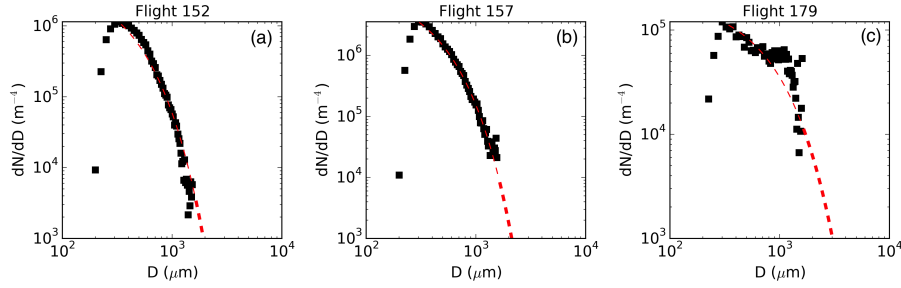


Figure 8. Average Size distribution of the crystals identified with the CIP for the flights (a) 152, (b) 157, and (c) 179 (black squares), along with the exponential distribution approximating them (red dashed line). The relative increase in ice mass when further integrating from 1.5 mm to larger diameters (equal to the relative error on the actual ice mass used in this study) is about 3% (a), 2.5% (b), and 65% (c).

as well as to the observations. The observed particles identified unambiguously as crystals in Part 1 span the diameter range 200 μm to 1.5 mm. Hence, because the cloud microphysics scheme have a lower limit size smaller than 200 μm for the ice crystal and an upper limit size larger than 1.5 mm for the precipitating ice particles (snow, graupel)(see Table 3), we expect the simulated IWC, and SWC to bracket the observations. However, the measured ice mass should be closer to SWC than to IWC
 5 given the relatively low additional mass expected from particles with $D > 1.5$ mm using the estimates presented above.

In 2010, the instances where SWC and IWC do bracket the observations happen on both sides of the Peninsula (Figure 9a, and Figure 9b) for the Morrison, and the Thompson schemes (note that the Thompson's IWC is comprised between 10^{-4} and 10^{-3} g kg^{-1}). Both WSM5's SWC and IWC equal the observation showing that a significant part of the simulated SWC is on average in the form of cloud ice crystals (IWC) (i.e. radii < 250 μm , see Table 3). In 2010, west of the Peninsula, Milbrandt's
 10 SWC and IWC are lower than the observations suggesting not enough ice formation.

In 2011 (Figure 9c, and Figure 9d), all the scheme have both averaged SWC and IWC lower than the observations, except for WSM5 to the East of the Peninsula, where the averaged IWC exceeds the observed value. East of the Peninsula, all the schemes predict equally well some non-negligible ice phase (n_5 criterion) with Morrison, Thompson, and Milbrandt performing better than WSM5 when considering the n_{50} criterion. However, the schemes perform worst west of the Peninsula with less than 40%
 15 of ice occurrences actually simulated (n_5 criterion). Overall, As for the liquid phase, the occurrences of the ice phase are less well simulated on the western side of the Peninsula, as compared to its eastern side.

Finally, we focus on the partition of water between the condensed phases, LWC, and SWC, by looking at the total average mixed-phase ratio $f_m = \text{LWC}/(\text{LWC} + \text{SWC})$ as a function of temperature along the flight tracks. Table 6 summarizes the statistics on f_m derived from measurements, and from simulations. First, none of the schemes sustain liquid clouds at temperatures
 20 below -15°C , or even below -9°C for the WSM5 (WDM6) scheme (leading to $f_m = 0$). This will be further commented in section 5.2. Second, between -15°C and 0°C , the Morrison scheme (0.91 ± 0.1), and the Milbrandt (0.78 ± 0.1) schemes have an average f_m in closest agreement with observations (0.83 ± 0.08). WSM5 performs the least well with f_m around 0.6 on average and down to 0.07 at its minimum. WSM5 ($\sigma = 0.24$), and the Thompson scheme ($\sigma = 0.2$) have a variability of f_m more than twice larger than the observations ($\sigma = 0.08$). Practically, for WSM5 and the Thompson scheme, it results in a highly

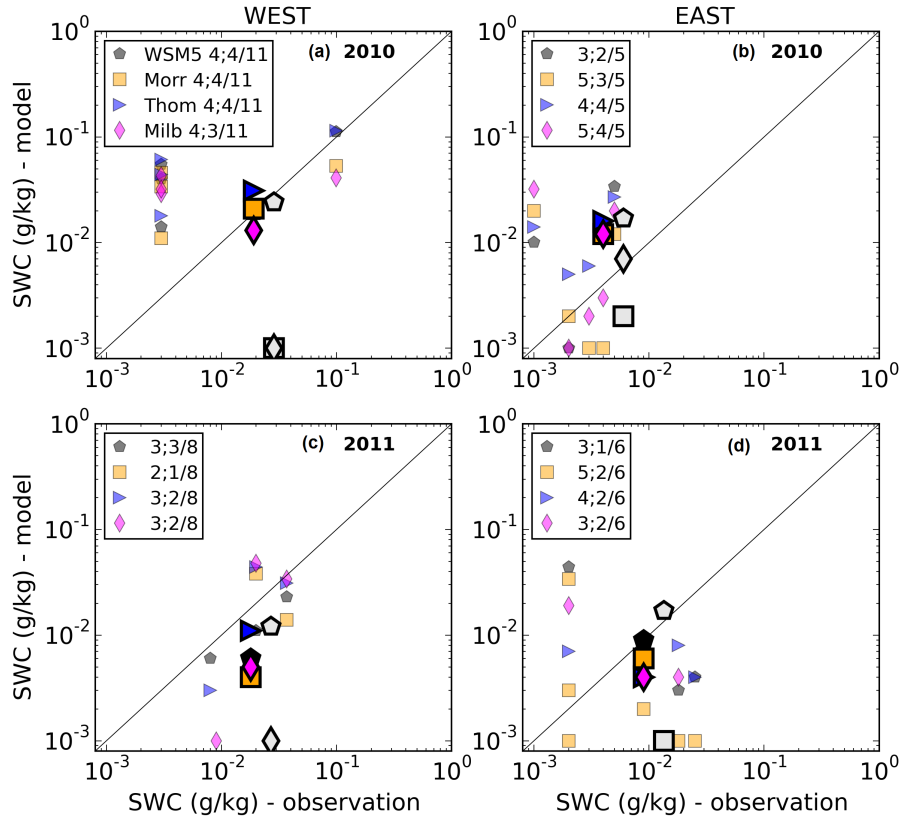


Figure 9. Scatter plot of simulated Solid Water Content, SWC (ice, snow, graupel) versus observed SWC in 2010 (a and b) and in 2011 (c and d) on the Western side of the Peninsula (a and c) and on the Eastern side of the Peninsula (b and d). Light markers show average per flight while bold markers gives the average over the whole tracks on each side of the Peninsula, respectively. Each bold grey marker indicates the Ice Water Content (IWC) average corresponding to the SWC average and each IWC marker is slightly shifted on the x-axis by 50% of the observed value for readability. The numbers next to each scheme's name ($n_5; n_{50}/N$) gives the number n_5 (resp. n_{50}) of simulated flights with SWC at least 5% (resp. 50%) of the observed SWC, to the total number N of flights measuring an average SWC of at least 0.0001 g kg^{-1} .

variable mixed-phase ratio from one 0.5° temperature bin to the next, what is not observed in the measurements (not shown).

The Morrison scheme ($\sigma=0.1$), and the Milbrandt schemes ($\sigma=0.1$) have a steadier f_m across the investigated temperature range where mixed-phase clouds are simulated, in closer agreement to the observations.

4.4 Temperatures and Water vapour in Polar WRF over the flight campaigns

- 5 We take advantage of temperature and water vapour measurements performed along with the cloud in-situ measurements to compare with the averaged simulation outputs. Latitudinal averages (in 0.5° longitude bins) for both observations and simulations, are shown for temperatures ($^\circ\text{C}$) and water vapour mass mixing ratios (g kg^{-1}) in Figure 10a, and Figure 10b

respectively. The variability of the water vapour, and of the temperature (shown as the standard deviation of the flight averages in each longitude bin) is indicated with shaded area for the observations, and with errorbars for the different cloud schemes. The measurements uncertainty for the temperature measured with a Rosemount probe is about 0.3°C (Stickney et al., 1994), corresponding to less than the width of the solid blue and red lines, respectively. The darkest narrow shaded areas bracketing
5 solid lines on both years correspond to a conservative estimate of uncertainty on water vapour ($\pm 0.15 \text{ g kg}^{-1}$) as derived using the relative humidity measured with a Vaisala Humicap HMP45 ($\pm 3\%$ estimated relative error), and the atmospheric temperature measurements from the Rosemount probe. A Buck 1011C cooled mirror hygrometer also present on board was used to correct for an offset in the Humicap measurements. At low temperatures and humidity the cooled mirror hygrometer occasionally has difficulty in identifying the frost point correctly and tends to hunt over a wide range. Therefore the Humicap
10 measurements were used once corrected using the cooled mirror hygrometer during periods when we are confident that it has correctly identified the frost point.

For the temperature, in 2010 all the simulations show best agreement with the measurements to the east of the Peninsula where the overestimation of the temperature ranges between 0 and 1°C (Figure 10a, top). Westward of 65°W the positive biases are larger and range between 1 and 2°C . In 2011 and East of the Peninsula, the temperature bias lies between 1 and 2°C ,
15 whereas West of 69°W it ranges between 2 and 3°C with the exception of the Thompson scheme leading to overestimations as large as 4°C (Figure 10a, bottom).

For the water vapour, the 2011 observed average is underestimated at almost all longitudes except between 68.5°W and 64°W where it is overestimated by 0.15 g kg^{-1} on average (Figure 10b, top). Eastward of 63°W , the underestimation increases up to values closer to 1 g kg^{-1} , while westward of 71°W it remains around 0.5 g kg^{-1} . In 2010 the average water vapour
20 is underestimated by $0.2\text{-}0.5 \text{ g kg}^{-1}$, except west of 68°W where it reaches 1 g kg^{-1} (Figure 10b, bottom). The bias then decreases to around 0.25 g kg^{-1} in the area $67.5^{\circ}\text{W}\text{-}63.5^{\circ}\text{W}$, except for WSM5, which remains closer to 0.5 g kg^{-1} . Eastward of 62°W the underestimation increases up to 1 g kg^{-1} but only one flight measured water vapour, hence the poor statistics (as shown by the absence of shaded area). WSM5 has the largest biases in averaged water vapour during both years, 0.6 g kg^{-1} and 0.45 g kg^{-1} in 2010 and 2011 respectively, mostly consisting in an underestimation of the observed water vapour. Other
25 schemes also mostly underestimate the water vapour, however less than WSM5 by $0.05\text{-}0.1 \text{ g kg}^{-1}$. No cloud scheme clearly stands out in terms of reducing the negative bias.

Overall across the Peninsula the simulations underestimate the measured water vapour by an average value of 0.5 g kg^{-1} ($\pm 0.2 \text{ g kg}^{-1}$, depending on schemes or regions across the Peninsula), and the temperatures are overestimated by 1°C ($\pm 0.2^{\circ}\text{C}$ depending on the scheme) in 2010 and 2 ($\pm 0.5^{\circ}\text{C}$ depending on the scheme) in 2011. Interestingly, the variabilities of the
30 observations (shaded area), and of the simulations (error bars), are consistent with each other. This suggests a good performance of the model average, and variability. The broad agreement in temperature and water vapour between the simulations suggest that their differences in average simulated clouds cannot be mainly related to the differences in water vapour and temperature, but rather to their microphysics. The biases compared to the observation will be further commented in the section 5.4.

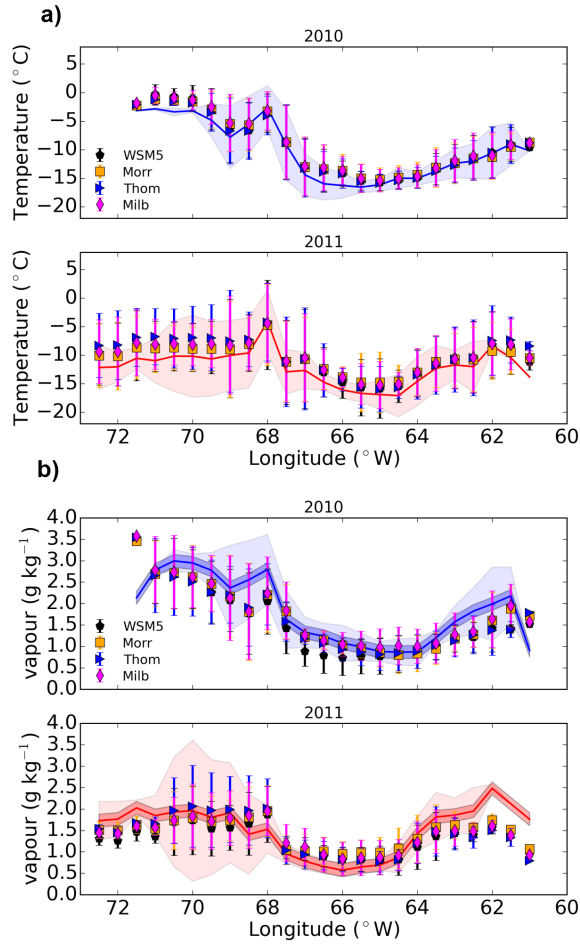


Figure 10. (a) Zonal distribution for 2010 and 2011 flight campaigns of (a) averaged temperatures and (b) water vapour (g kg^{-1}). Measurements are shown as a solid line, and simulations as markers. Both shaded areas, and error bars, give the standard deviation in each 0.5° longitude bin for the observation, and the simulations, respectively. The dark shaded area correspond to a conservative estimate of the uncertainty on water vapour (see text for details).

5 Discussion

5.1 On the radiative biases

A deficiency of downward LW radiation responsible for a cold summer surface (temperature) bias in Polar WRF simulations was spotted at a continental scale by Bromwich et al. (2013a) at 60 km, and 15 km resolution, and the authors related this bias to a deficiency in the cloud cover. Bromwich et al. (2013a) showed that using ERA-interim analysis forcing at the domain boundaries (instead of GFS analysis) helped to significantly reduce the average cold summer bias (see their Table 5), although the improvement for surface pressure or dew point is not clear. Bromwich et al. (2013a)'s simulation outputs as well as K15's

AMPS results relied on the WSM5 scheme. We did find a similar negative bias for the LW radiation to K15 over the Larsen C Ice Shelf, as well as a similar positive bias for the SW radiation (section 3.1).

When schemes different from WSM5 or WDM6 are used with Polar WRF in our simulations, a strong decrease in the LW bias for both periods of interest over the Larsen C ice shelf are measured. It suggests that schemes like the Morrison, Thompson or Milbrandt schemes should be preferred to the WSM5, and WDM6 schemes in studies dealing with the evolution of the energy budget of the Larsen C Ice Shelf within Polar WRF. The strong decrease in LW surface biases (by as much as 20 W/m², see table 4) when using the three cloud schemes, which have a more sophisticated ice microphysics parameterisation (double-moment) is systematically statistically significant on both years and at both AWS (AWS14 and AWS15). Note that the smallest biases are obtained using the Morrison scheme.

The explanation as to why the LW bias is significantly improved while the SW is not always (especially in 2010) is most probably because the variations of the cloud droplets effective radius is not accounted for in the model. The radiative scheme (Goddard scheme, see Table 1) parameterises the optical depth for water and ice as a function of the particle effective radius (Chou and Suarez, 1999). This parameterization does not assume any type of droplet (crystal) size distribution, so it can be used with the different cloud schemes despite their own different assumptions on the hydrometeors size distributions. Also, the SW radiative scheme assumes a constant value of 10 μm for the cloud droplets effective radius. However, for a given LWC, the SW radiation is scattered in different ways depending on the effective radius of the droplets, with smaller radii reflecting more efficiently SW radiation. The droplet effective radius derived for both campaigns is about 7 μm , close to the 10 μm assumed in the radiative scheme. Running a simulation over a shorter period (11-20th January 2011) replacing the constant effective radius of 10 μm by a constant 7 μm in the radiative scheme did not lower the SW radiative bias over the Larsen C Ice shelf (not shown). This was expected as it is rather the variations of the effective radius with time that could be expected to improve the SW bias. As noted by Bromwich et al. (2013a), the SW bias is of secondary importance for the surface energy budget because SW radiations not reflected by missing clouds in the model will be reflected by the icy or snowy surface underneath. The cloud radiative effect dominates in the LW radiation over icy surfaces (as opposed to over the ocean).

The poorer performances of the various simulation in terms of surface radiation biases at Rothera station (Table 4, left part) and especially the similarly large LW surface biases for all the schemes are consistent with a poorer representation of the supercooled liquid clouds in the western part of the Peninsula (Figure 6). Indeed, only a few flight tracks were simulated with supercooled liquid phase (3 out of 11 and 3 out of 10 at best, in 2010 and 2011 respectively). This is further commented in section 5.2, where we discuss the simulation of the cloud phase.

5.2 Simulating the clouds thermodynamic phase

Cloud schemes form supercooled liquid provided the growth of the activated ice phase does not consume the entire excess of water vapour (compared to RH=100%). Except for the Milbrandt scheme which completely removes the supersaturation by conversion of the excess of water vapour into liquid, the other schemes explicitly derive a condensation growth rate. Thus, the cloud microphysics schemes mainly differ in their ice microphysics, and mixed-phase interactions, which will determine their ability to form and maintain supercooled liquid in the atmosphere.

WSM5 (WDM6) is the only microphysics scheme showing an anticorrelation of the liquid water content and solid water content on the Peninsula, suggesting a systematic depletion of water vapour in favor of the ice phase (Figure 4). Close to, and above the topography WSM5 has a deficit in liquid clouds due to orographic forcing which favours ice clouds, whereas the Morrison, Milbrandt, and Thompson schemes have a steady if not increasing LWC. One of the main differences between WSM5 and WDM6 (hereafter called the WRF schemes), and the other three schemes is that the former are single-moment schemes for the icy hydrometeors, whereas the latter are double-moment schemes for the ice crystals (only the Thompson scheme is not a double-moment scheme for snow/graupel particles). A consequence of The WRF schemes being single-moment schemes for the ice crystals is the use of a relationship for linking the number concentration of ice crystals (N_i) to the ice water content (q_i), since they cannot evolve independently. N_i is diagnosed from q_i based on empirical relationships (see equations 5a-d Hong et al., 2004, where N_i is proportional to $q_i^{3/4}$). In addition to that, the INP parameterisation used in the WRF schemes (WSM5 and WDM6) is predicting significantly more INPs than any other parameterisation above -15°C (as it will be shown in the next section 5.3). Since it is used to predict the initial q_i when firt ice appears, the latter is biased towards larger than expected values, and so N_i will also be biased towards larger values (because of the empirical relationship linking both). This, in turn, impacts the growth rate of the ice crystal mass, which depends on N_i (see equations 9 Hong et al., 2004), favoring an increasing ice water content.

The transects in Figure 5a and b clearly show the ubiquitous ice simulated with WSM5 during westerly and easterly events due to the orographic forcing, whereas the Morrison scheme leads to snow and supercooled liquid formation in both cases (Figure 5c and d) as the Milbrandt, and the Thompson schemes (not shown). As an additional experiment, we made a simulation over the period 6-10th January 2011 where westerlies are simulated (Figure 5a and c). We alter the WSM5 scheme in changing the empirical relationship linking q_i to N_i described above. We divide by 100 the resulting N_i (variable `xni` in the code where the empirical equation 5c in Hong et al. (2004) is implemented), which is diagnosed from q_i in the cloud scheme. We plot a similar transect to Figure 5a. This results in more supercooled liquid being simulated (Figure 11a), closer although not yet similar to what - for instance - the Morrison scheme is leading to (Figure 5c). Then, a second simulation with WSM5 was performed (Figure 11b) by changing the INP parameterisation to the one used in the Morrison scheme (more realistic, as shown in section 5.3). It shows no major difference with the original simulation (Figure 5a). The hypotheses (empirical relationships) used for the single-moment ice crystals parameterisation of WSM5 have a more determining impact on the ability to sustain supercooled water than the nature of the INP parameterisation itself in the WRF schemes.

Note that a third simulation was performed with the Morrison scheme using a lower concentration of CCN set to 100 cm^{-3} (instead of 250 cm^{-3}). Using a lower CCN concentration does affect the amount of supercooled liquid formed in reducing it (not shown), but the overall distribution of liquid (the order of magnitude) remains similar, as well as the ice, thus not altering the main conclusions of this work. Note that, according to Part 1, the observed average number of drops is $100\text{-}120\text{ cm}^{-3}$ in 2010, and $150\text{-}200\text{ cm}^{-3}$ in 2011. Hence the Morrison, and Thompson schemes use a drop number similar to the observed upper limit and lower limit, respectively (see Table 3). More simulations focused on particular flights (case studies) would be required to assess in greater details the impact of the CCN concentration compared to the one of the INP parameterisation.

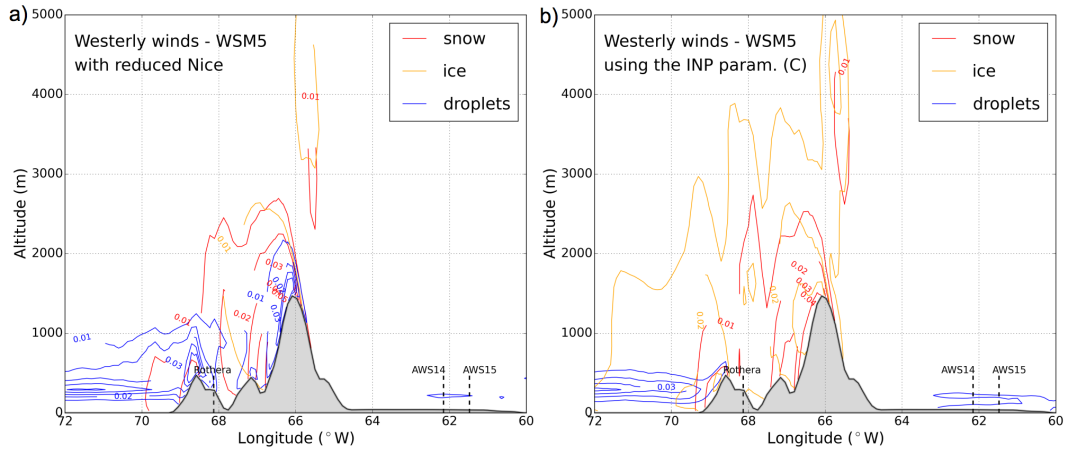


Figure 11. Same as Figure 5a with (a) a modified version of WSM5 where the empirical relationship linking the number of ice crystals to IWC is altered (see text for details), and (b) a modified version of WSM5 where the INP parameterisation used (Fmod) is changed to (C) (Cooper, 1986), which predicts less INPs (see section 5.3).

Figure 12 shows the distribution of the cloud mass as a function of the temperature for the transects shown in Figure 5 during westerlies (solid line), and easterlies (dashed line), respectively. The top row is the median simulated mass per 1°C bin for (a) the liquid droplets, (b) the ice crystals, and (c) the snow particles. The bottom row shows the corresponding number of non-null occurrences ($> 0.001 \text{ g kg}^{-1}$) over which the median values are derived in each bin. Figure 12a shows that LWC simulated down to -10°C by WSM5 is similar to Milbrandt's and Thompson's, and lower than Morrison's for both easterlies and westerlies scenarios. However, the frequency of liquid cloud formation for WSM5 is lower by a factor of two to four compared to other schemes (Figure 12d). At colder temperatures, WSM5 ability to simulate cloud liquid is drastically reduced for both scenarios (Figure 12d). This can be related to the much shallower vertical extent of the WSM5 simulated cloud liquid (Figure 7), which is limited to the warmest subfreezing temperatures. The observations show liquid clouds up to higher altitudes (Figure 7), and the Morrison, Milbrandt, and Thompson schemes account better for the liquid at these higher altitudes (lower temperatures) than the WRF schemes. Yet, these three schemes do show a decreasing trend in the supercooled liquid mass at temperatures lower than -10°C , despite their steady (slightly growing) ability to simulate ice and snow crystals (Figure 12e and f). Measured vertical profiles do show indeed the presence of ice clouds at temperatures lower than -10°C , and the schemes simulate their occurrences better than the supercooled liquid ones (not shown). Interestingly, the frequency of simulated ice crystals are the most different above -10°C (Figure 12e), where there is an order of magnitude difference between WSM5, and the other schemes. WSM5's IWC is a factor of five to ten larger than for the other schemes (Figure 12b) and this can be mainly explained by its single-moment parameterisation for ice (as described above). This feature also appears in the model outputs at stations AWS14 (Figure 3) and AWS15 (not shown), where the WRF schemes simulate three to four times more cloud ice crystal mass than the other cloud schemes.

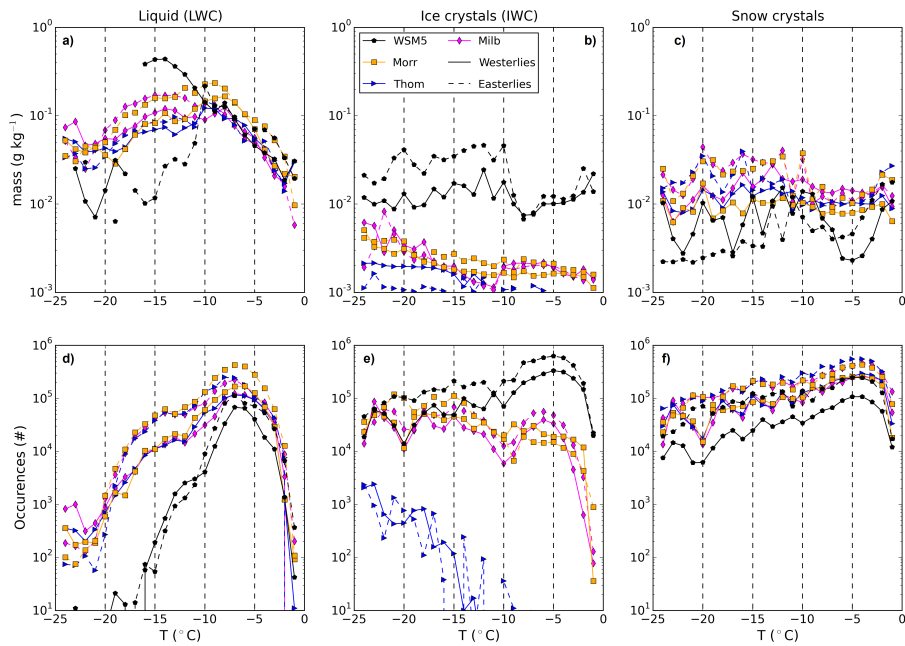


Figure 12. Distribution of (a) LWC, (b) IWC, and (c) snow crystal mass, as a function of temperature (per 1°C bin), and the corresponding distribution of non-null occurrences of LWC, ice, and snow (d, e, f, respectively), used to derive those mass distributions, for the transects shown in Figure 5 for westerly (solid line), and easterly wind periods.

Overall the Morrison, Thompson, and Milbrandt schemes are better able than the WRF schemes to form supercooled liquid in both years on both sides of the Peninsula. They perform better in simulating liquid cloud occurrences, and the Morrison scheme gives an average LWC closer to the measured value, except on the east in 2011, when it overestimates the liquid content (section 4.3.1, Figure 6b an d). However, the observed interannual variations of the LWC from 2010 to 2011 East of the Peninsula (Figure 6) and described as statistically significant in Part 1, is not captured by any of the cloud schemes. Part 1 reported on the role of the nature and number of aerosols, of which a subset act as CCN or INP. Based on Part 1, it is likely that the observed regional and interannual cloud microphysics variabilities need an adequate aerosol model in order to be properly simulated. The clouds measured by the aircraft campaigns were exclusively mixed-phase clouds and Table 6 shows that at least down to -13°C the Morrison scheme, and the Milbrandt schemes are the more capable of simulating the observed relative proportions of liquid and ice across this temperature range. However, as shown by Figure 7 little supercooled liquid is simulated above 1500 meters along the flight tracks.

The poorer performances of the schemes to the west of the Peninsula can be seen in the poorer ability to predict the occurrences of both cloud liquid (Figure 6a and c) and cloud ice (Figure 9a and c) in that region. The number of tracks predicted with some liquid phase represent about 20% (2010) or 40% (2011) of the total observed. Figure 9a and c show a slightly better ability to predict the ice phase but still less than 50% of the tracks are predicted. The associated failure for any scheme to lower the LW surface radiation biases at Rothera station (Table 4) suggest an overall inability to correctly simulate liquid

clouds where they are observed. As noted in section 4.4, the average temperature biases are larger to the West of the Peninsula than to the East of the Peninsula, by 1 to 2°C (Figure 10a). On both sides the average temperature is overestimated. Since the supersaturation depends exponentially on the temperature, the lowest ability of the schemes to predict liquid cloud formation to the West is consistent with the larger temperature biases measured in that region. The warmer oceanic and sea-ice free influence of the western Peninsula implies more convective processes (compared to the east) that are badly resolved at 5 km resolution and prevent better matching with the aircraft observations. The 5 km horizontal resolution lies in the so-called "grey zone" (resolution 1–10 km) where convective processes are badly simulated and parameterised. An explanation for the bad performances of the schemes above Rothera station may be the complex topography as shown in Figure 5. The station sits in the eastern part of Adelaide island, and is surrounded by mountaineous features. More generally these features will also affect the air flow reaching the regions where the flights took place, to the west of the Peninsula. By contrast, the eastern part of the Peninsula with the Larsen C Ice shelf has much less complex topographical features, and this should be helping the modelling of the clouds on that side.

Finally, it is worth recalling that working with ice categories, but also different definitions for these ice categories from one scheme to another, makes overall comparisons to flight measurements difficult. The Thompson scheme shows very little formation of ice crystals, which is readily converted to snow crystals (Figure 12f). Every cloud scheme defines a radius cut-off between ice crystals and snow crystals ranging between 100 and 250 μm (Table 3), with Thompson having the second largest value at 200 μm . It is not clear why the Thompson scheme cloud ice crystals numbers are so low (Figure 6, and Figure 12b). Thompson scheme's gives much less frequent and much less abundant crystals at radii below 200 μm (Figure 12e) compared to the ones above that radius (Figure 12f), and this is at odds with the other schemes, and with the observations. Finally, given that the observations show an average crystal radius of 150-250 μm in 2010, and 200-250 μm in 2011 (not shown), this is probably not ideal to work with cloud schemes having an ice-snow radius cut-off artificially set around those sizes.

5.3 The INP parameterisations

All the cloud microphysics schemes investigated in this work rely on INP parameterisations to initiate the ice phase, and here we comment on those. The number concentration of INPs is diagnosed from the modelled atmospheric temperature only. These empirical parameterisations address the different ice nucleation mechanisms (see introduction). They are triggered at different temperatures or supersaturation thresholds, depending on the cloud scheme (Table 3). They increase exponentially with decreasing temperatures, and can lead to very different INPs concentrations (Figure 13a, coloured lines). The direct consequence of this are clear differences between icy hydrometeors number concentrations, as a function of temperature. To illustrate this Figure 13b shows the median non-null number concentration of total icy condensates (ice crystals, snow and graupel particles) over the transects shown in Figure 5 for both the westerly (solid lines), and easterly (dashed lines) cases. For deposition/condensation freezing (which does not require the presence of supercooled droplets) the Milbrandt scheme uses the INP parameterisation from Meyers et al. (1992) (Their equation 2.4), while Morrison, and Thompson use the one from Cooper (1986). Both parameterisations are now referred to as (M) and (C), respectively (Table 3). This translates into the drastically different number concentrations at temperatures above -15°C (Figure 13b), because INPs concentrations predicted by (M)

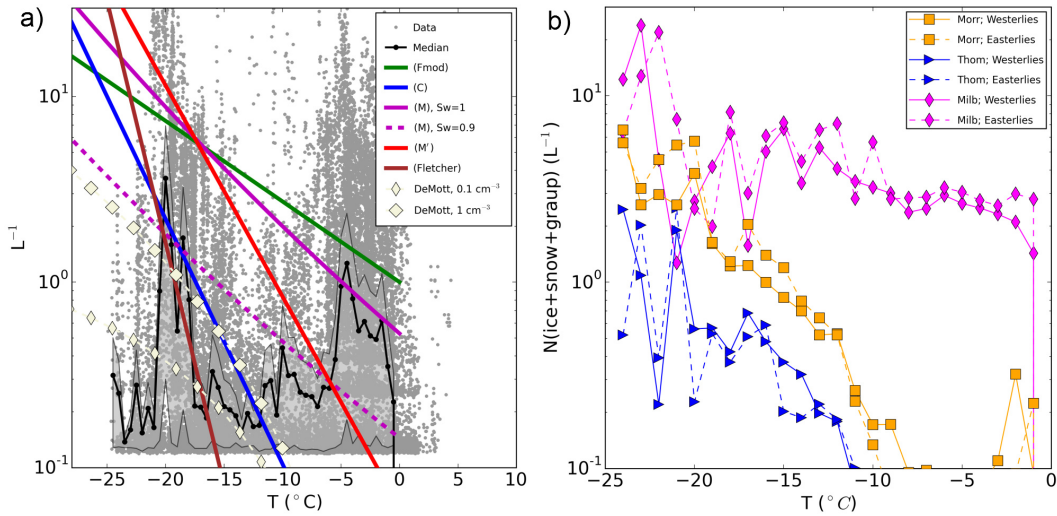


Figure 13. (a) Ice crystals measurements data points as a function of the temperature (grey) with their median per 0.5°C bins (black solid line), along with the absolute median deviation in shaded grey. The labelled INP parameterisations used by the different cloud schemes (Table 3) are overlaid. DeMott refers to DeMott et al. (2010)'s INP parameterisation (see text for details). (b) Same as (a) for the Median of the total number concentrations of icy hydrometeors in the same transects used in Figure 5 for both cases, the westerly case (solid lines), and the easterly case (dashed lines).

(Figure 13a, purple lines) are much larger than the ones predicted by (C) (blue line). For contact freezing the Milbrandt, and the Morrison schemes use the INP parameterisation from Meyers et al. (1992) (their equation 2.6), which is referred to as (M') in Table 3. The Thompson scheme does not explicitly parameterise contact freezing. The consequence is that the Morrison scheme predicts larger amount of icy condensates than the Thompson scheme since (M') predicts much larger INPs concentrations than (C). The latter effect is enhanced by the more constraining thresholds on temperature and ice supersaturation for the Thompson scheme to allow for ice formation (Table 3). The Milbrandt scheme relies on the INP parameterisation (M), which predicts much larger amounts of INPs in the deposition mode, which does not depend on the scheme ability to simulate supercooled liquid water in a first place. Interestingly, the Milbrandt scheme has an average solid water content (SWC₀, and SWC) almost twice as big as the Morrison, and the Thompson schemes (Figure 4b, and Figure 4b, respectively).

DeMott et al. (2010) developed an INP parameterisation (hereafter called DeMott) using both the temperature and the observed aerosol number concentration, for aerosols larger than $0.5 \mu\text{m}$ in diameter (as presented in Part 1) believed to be the main contributors to the worldwide INP population (DeMott et al., 2010). Aircraft measurements used in Part 1 made it possible to derive out-of-clouds aerosol concentration for diameters larger than $0.5 \mu\text{m}$. Using this information to describe the aerosol background, we compared the measured total ice crystals number concentrations to INPs predictions by DeMott, and to the other INP parameterisations implemented in the WRF cloud schemes. The comparison to the observations was done at temperatures below -9°C , as this is the temperature range over which DeMott parameterisation was derived. This also allows

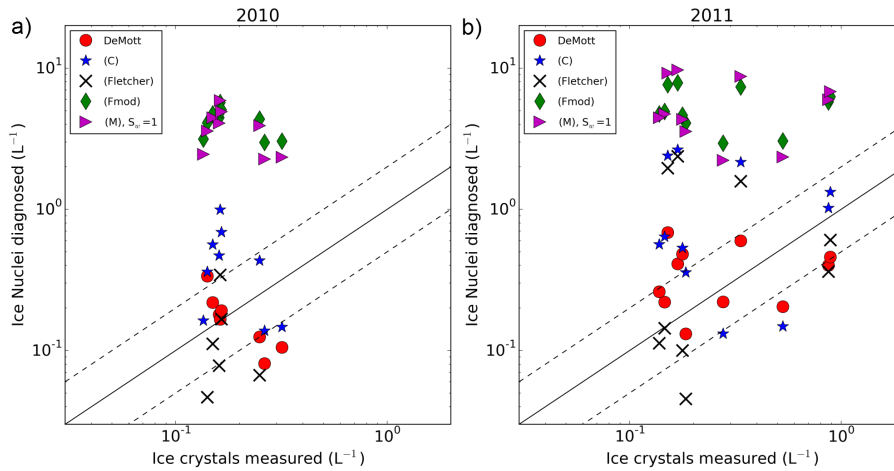


Figure 14. Predicted INP number densities versus observed number densities of crystals (flights averages) for various INP parameterisations for (a) 2010, and (b) 2011. The solid line corresponds to the one-to-one line, and the dashed lines correspond to a factor of two difference between y and x axes. See Table 3 for the references of the various INP parameterisations.

to discard the warmer temperatures where a secondary ice process was identified as responsible for the ice crystal production around -5°C (see Part 1). The INP parameterisations are meant to account only for the primary ice production process.

For each measurement of crystals below -9°C (one data point every second), a corresponding number of INPs is derived for each parameterisation using the measured temperature. For the background aerosol input to the DeMott parameterisation, we derived a 1 min averaged out-of-cloud aerosol number concentration (n_{aer}) within ± 30 s of any crystal measurement. As shown in Part 1 (see their Figure 13) the average n_{aer} ranges between 0.1 cm^{-3} and 1 cm^{-3} . Figure 13 shows the DeMott parameterisation for those two values (white markers). We computed flight averages for the observations, and for each INP parameterisation. Figure 14 shows the observed and predicted average values for both years. Table 7 gives the median relative difference (ϵ) between flight-averaged observations (Obs), and the INP parameterisations predictions (INPparam), along with the associated median absolute deviation ($\Delta\epsilon$).

The DeMott parameterisation performs better than any other INP parameterisation as suggested by Figure 14a and Figure 14b. Table 7 shows that DeMott ($\epsilon=0.5\text{-}0.6$) performs better than Cooper's (C) ($\epsilon=1.6\text{-}2$) (used by the Morrison, and the Thompson's schemes). This is because of its ability to take into account the number concentration of aerosols. For instance if we force a constant value of $n_{\text{aer}}=1\text{ cm}^{-3}$, the DeMott parameterisation performs as poorly as (C) and worse than the original Fletcher's (as opposed to Fmod used by Hong et al., 2004, see their equation 8). However, if we force $n_{\text{aer}}=0.1\text{ cm}^{-3}$ (the average n_{aer} across the Antarctic Peninsula above 2000 m where most primary ice production occurs – see Part 1, and their Figures 13, and 14) then DeMott is performing better than (C), and than any other parameterisation. It performs as good as DeMott with a varying n_{aer} (compare the two first lines of Table 7). The modified version of Fletcher's (Fmod) used in

WSM5 is the worst performer ($\epsilon > 20$), followed by the Meyer's (M) parameterisation, which is used by the Milbrandt scheme. ($\epsilon = 4.5-24$, for relative humidities characteristics of mixed-phase clouds, 90–100%, respectively).

It should be recalled that the DeMott parameterisation is based on analysis of aerosols, which exclude strong marine influence, and so sea salts were not included (DeMott et al., 2010). Also, aerosol concentrations below 0.3 cm^{-3} have less weight in the DeMott parameterisation's analytical derivation compared to the larger values ($0.5-5 \text{ cm}^{-3}$), as shown by (DeMott et al., 2010)'s Figure S1 of their supplementary materials. Despite these caveats, the strength of the DeMott parameterisation is to be able to account for the low aerosol number densities at altitudes higher than 2000 m, where primary ice production occurs in the Antarctic Peninsula (Part 1). This makes it, on average, a better candidate than any other IN parameterisations for future work meant to improve the cloud microphysics scheme for Antarctic clouds.

Finally, note that the comparisons are made in times and places where ice crystals were indeed measured, ignoring instances where cloud ice was not measured, but where any INP parameterisation would still predict some crystal production. This challenging issue could probably be dealt with only in managing the coupling of the cloud scheme to an aerosol model able to predict the absence of INPs. Moreover, given existing biases in water vapour and temperature along each flight tracks separately (as opposed to the averages discussed in section 4.4), better calibrating the INP parameterisation consists in only one of the needed improvements for Antarctic clouds modelling, as discussed below.

5.4 Additional results on water vapor and temperature biases, and cloud nuclei parameterisations

In section 4.4 it was shown that the model was able to capture the average temperature within $0.5-2.5^\circ\text{C}$, and the average water vapour within $0.3-0.7 \text{ g kg}^{-1}$. The average simulated temperature and water vapour are within the variability (standard deviation within each longitude bin) of the observations. Although the average behavior of the model matches the average observations, it should be noted that water vapour and temperature biases do hamper the good prediction of the clouds by the model. As an example, Figure 15a shows the time series of the water vapour, and of the temperature, as measured (black line) and as simulated when using the Morrison scheme (red line) for the flight 150. The model fails in simulating the liquid cloud before 18.5h, and after 18.7h (not shown), where the water vapour and temperature biases are the greatest (see the red solid line in Figure 15a), while it does simulate liquid cloud (not shown) where the bias is much reduced (at warmer temperatures, lower altitudes) between those two times (red solid line, in Figure 15a).

An additional simulation was performed over the period 11-20th January 2011 using the Morrison scheme, initializing it ten days later than the simulations used so far (on the 11th of January instead of starting on the 1st). During this period, four flights took place (150-153) and we show flights 150, and 152 (Figure 15). The result is shown as the blue solid line in the Figure 15a and b. The initialization of the model closer to the dates of the airborne measurements does lead to a lower bias in water vapour for both flights. However, the bias in terms of temperature is relatively less reduced across the flights, suggesting that the initialisation of the model has a greater impact on the quality of the water vapour prediction. However the improvement in terms of water vapour does not lead to an improvement of the liquid cloud prediction along the flight tracks, and it even leads to the suppression of the liquid cloud initially simulated along flight 152's flight tracks (not shown). Note that further doubling the number of vertical levels for the above shorter simulation (using 60 eta levels, instead of 30 eta levels) leads to

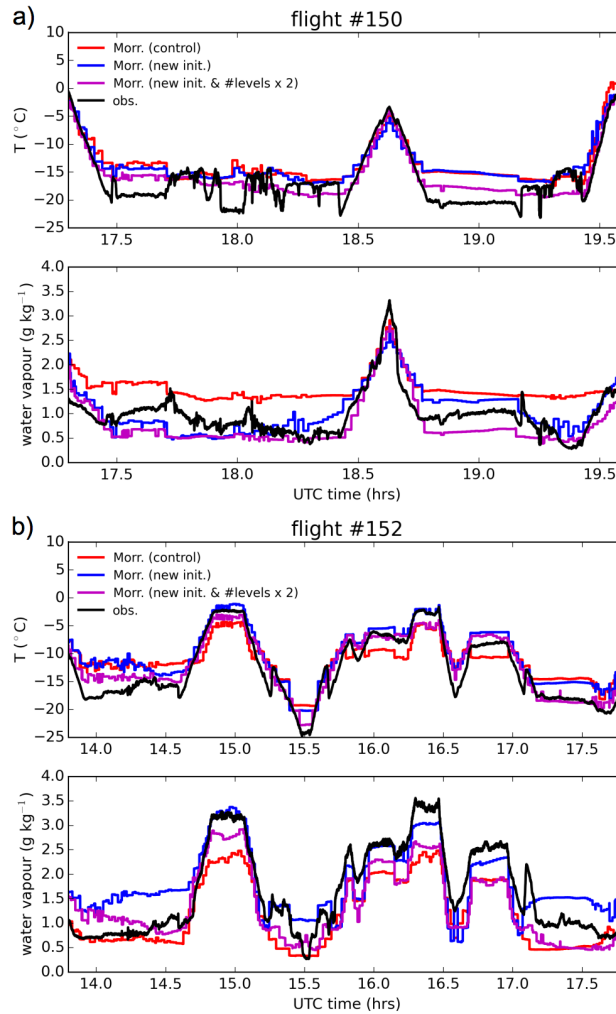


Figure 15. Time series of water vapour (g kg^{-1}), and temperature ($^{\circ}\text{C}$) measured during (a) flight 150, and (b) flight 152. For (a), and (b), the observation is the black solid line. Simulation outputs are overplotted: the simulations used to derive the averages presented in this work (red line), a different simulation over the period 11-20th January 2011 (thus with an initialisation closer to the date of the measurements)(blue solid line), and the latter simulation with – additionally – a doubling of the number of vertical levels (60 instead of 30) (magenta solid line).

a very limited further reduction of the water vapour bias over the four flights, while it does reduce the temperature bias over the flights (magenta solid line in Figure 15a and b). Overall, those results suggest that initializing the model at a closer date to the observations reduces on average the water vapour bias, while doubling the number of levels helps reducing the temperature bias. However, this improvement is not systematic along the flights, and not significant enough, so that the cloud prediction is

5 not really impacted, at least in the investigated cases.

Another run was performed over the same period of the four flights, this time only replacing the INP parameterisation used in the Morrison scheme (see section 5.3), by the Demott parameterization. The result was a much reduced ice crystal water content during the flights (actually lowering the quality of the ice phase prediction; not shown), but no improvement was obtained for the supercooled liquid, what may be explained by the remaining water vapour and temperature biases (not shown) preventing from supersaturation with respect to liquid water to be simulated.

This discussion stresses the need for further work investigating the water vapour and temperature biases in addition to using appropriate cloud scheme (double-moment scheme for the ice crystals), and INP parameterisation, for improving antarctic clouds simulation.

6 Summary and Perspective

In this work we provide the first intercomparison of WRF microphysics schemes performances in Antarctica over the Antarctic Peninsula within Polar WRF at 5 km resolution, as well as the first comparisons with in-situ cloud measurements on both sides of the Peninsula. The specificities and properties of the schemes are summarized in Tables 2 and 3. We compared the simulations to averaged aircraft measurements of cloud microphysics properties (Part 1) as well as other atmospheric properties on both sides of the Peninsula and over the two periods of interest (February 2010 and January 2011). This paper was motivated by King et al. (2015) which pointed towards possible problems in the thermodynamic phase simulation in three high resolution model at 5 km resolution over the Eastern Peninsula's Larsen C Ice Shelf, as well as Bromwich et al. (2013a), which demonstrated the presence of Antarctic-wide surface radiative biases within Polar WRF at coarser (60–15 km) resolution. This study is a first step towards the improvement of cloud modelling and operational forecast, with Polar WRF, and AMPS, respectively.

The main results are as follows.

- The surface longwave radiative bias is significantly reduced over the Larsen C Ice Shelf when using the Morrison, the Thompson and the Milbrandt schemes, compared to WSM5 or WDM6.
- Importantly, the Morrison, the Thompson and the Milbrandt schemes, are also the schemes leading to better agreement with aircraft cloud measurements (occurrences of the liquid and ice phase, as well as values of the cloud mass mixing ratio), compared to WSM5 and WDM6.
- The Morrison, the Thompson and the Milbrandt schemes perform better than the WSM5 and WDM6 schemes because of their double-moment parameterisation for the ice phase. The latter are single-moment schemes for the ice crystals. A realistic ice parameterisation is essential to the simulation of supercooled liquid.
- The DeMott parameterisation (DeMott et al., 2010), which is not currently implemented in any of the WRF microphysics scheme, better accounts for the ice crystals number densities measured during both campaigns, when using as input the typical concentrations of out-of-clouds aerosols measured above 2000m, where primary ice production occurs (see Part 1). However, the INP alone cannot improve the simulation of the observed clouds.

- The model can simulate the average water vapour, and temperature distribution across the Peninsula, however biases in both fields can still explain the failure in simulating clouds when looking at specific flights (as opposed to the average fields). Moreover, larger bias in temperatures to the west of the Peninsula can explain the lesser ability of the simulations to reproduce the observed clouds.

5 - As WSM5 is the scheme used in the Antarctic Mesoscale Prediction System, the present work provide new results promoting the improvement of the current cloud scheme implementation in the operational model.

Future work will look at case studies focusing on specific flights at higher spatial and vertical resolution. This will also make use of the latest campaign for measuring Antarctic clouds in the eastern Weddell Sea in November-December 2015 (O'Shea et al., 2017). More investigation of the impact of smaller (temporal or spatial) scale temperature and water vapour biases on
10 mixed-phase clouds simulation will be needed. Often disregarded in simulation works performed over Antarctica not related to cloud studies, cloud schemes should be more systematically considered. Investigating Antarctic clouds, and their impact on the energy budget, is an important step to help quantify the role of atmospheric-driven processes in the evolution of the ice shelves, the glaciers, and the Antarctic ice mass balance, and importantly to improve the forecast for field operations.

Acknowledgements. The authors thank Tony Phillips, and Pranab Deb for providing the Antarctic topography adapted to WRF, and derived
15 from Fretwell et al. (2013). The study was funded by Natural Environment Research Council (NERC) under grant NE/K01305X/1. CL also thanks CNES for postdoctoral fellowship funding. The authors thank the two anonymous reviewers, who helped to improve the presentation, and the content of the manuscript.

References

- Baumgardner, D., Jonsson, H., Dawson, W., O'Connor, D., and Newton, R.: The cloud, aerosol and precipitation spectrometer: a new instrument for cloud investigations, *Atmospheric Research*, 59-60, 251–264, doi:10.1016/s0169-8095(01)00119-3, 2001.
- Bigg, E. K.: The Supercooling of Water, *Proceedings of the Physical Society B*, 66, 688–694, doi:10.1088/0370-1301/66/8/309, 1953.
- 5 Boucher, O., Randall, D., Artaxo, P., Bretherton, C., Feingold, G., Forster, P., Kerminen, V.-M., Kondo, Y., Liao, H., Lohmann, U., Rasch, P., Satheesh, S., Sherwood, S., Stevens, B., and Zhang, X.: Clouds and Aerosols. In: *Climate Change 2013: The Physical Science Basis. Contribution of Working Group I to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change* [Stocker, T.F., D. Qin, G.-K. Plattner, M. Tignor, S.K. Allen, J. Boschung, A. Nauels, Y. Xia, V. Bex and P.M. Midgley (eds.)], 2013.
- Bromwich, D. H., Nicolas, J. P., Hines, K. M., Kay, J. E., Key, E. L., Lazzara, M. A., Lubin, D., McFarquhar, G. M., Gorodetskaya, I. V.,
10 Grosvenor, D. P., Lachlan-Cope, T., and van Lipzig, N. P. M.: Tropospheric clouds in Antarctica, *Reviews of Geophysics*, 50, RG1004, doi:10.1029/2011RG000363, 2012.
- Bromwich, D. H., Otieno, F. O., Hines, K. M., Manning, K. W., and Shilo, E.: Comprehensive evaluation of polar weather research and forecasting model performance in the Antarctic, *Journal of Geophysical Research: Atmospheres*, 118, 274–292, doi:10.1029/2012jd018139, <http://dx.doi.org/10.1029/2012JD018139>, 2013a.
- 15 Bromwich, D. H., Nicolas, J. P., Monaghan, A. J., Lazzara, M. A., Keller, L. M., Weidner, G. A., and Wilson, A. B.: Central West Antarctica among the most rapidly warming regions on Earth, *Nature Geoscience*, 6, 139–145, doi:10.1038/ngeo1671, 2013b.
- Brown, P. R. A. and Francis, P. N.: Improved Measurements of the Ice Water Content in Cirrus Using a Total-Water Probe, *Journal of Atmospheric and Oceanic Technology*, 12, 410–414, doi:10.1175/1520-0426(1995)012<0410:imotiw>2.0.co;2, 1995.
- Burrows, S. M., Hoose, C., Pöschl, U., and Lawrence, M. G.: Ice nuclei in marine air: biogenic particles or dust?, *Atmospheric Chemistry and Physics*, 13, 245–267, doi:10.5194/acp-13-245-2013, <http://dx.doi.org/10.5194/acp-13-245-2013>, 2013.
- 20 **Chou, M.-D. and Suarez, M. J.: A solar radiation parameterization for atmospheric studies., Tech. rep., NASA/GSFC, 1999.**
- Cohard, J.-M. and Pinty, J.-P.: A comprehensive two-moment warm microphysical bulk scheme. I: Description and tests, *Quarterly Journal of the Royal Meteorological Society*, 126, 1815–1842, doi:10.1002/qj.49712656613, 2000.
- Cooper, W. A.: Ice Initiation in Natural Clouds, *Meteorological Monographs*, 21, 29–32, doi:10.1175/0065-9401-21.43.29, 1986.
- 25 Deb, P., Orr, A., Hosking, J. S., Phillips, T., Turner, J., Bannister, D., Pope, J. O., and Colwell, S.: An assessment of the Polar Weather Research and Forecasting (WRF) model representation of near-surface meteorological variables over West Antarctica, *Journal of Geophysical Research: Atmospheres*, 121, 1532–1548, doi:10.1002/2015jd024037, <https://doi.org/10.1002/2015jd024037>, 2016.
- Dee, D. P., Uppala, S. M., Simmons, A. J., Berrisford, P., Poli, P., Kobayashi, S., Andrae, U., Balmaseda, M. A., Balsamo, G., Bauer, P., Bechtold, P., Beljaars, A. C. M., van de Berg, L., Bidlot, J., Bormann, N., Delsol, C., Dragani, R., Fuentes, M., Geer, A. J., Haimberger, L., Healy, S. B., Hersbach, H., Hólm, E. V., Isaksen, I., Kållberg, P., Köhler, M., Matricardi, M., McNally, A. P., Monge-Sanz, B. M., Morcrette, J.-J., Park, B.-K., Peubey, C., de Rosnay, P., Tavolato, C., Thépaut, J.-N., and Vitart, F.: The ERA-Interim reanalysis: configuration and performance of the data assimilation system, *Quarterly Journal of the Royal Meteorological Society*, 137, 553–597, doi:10.1002/qj.828, 2011.
- 30 Demott, P. J., Meyers, M. P., and Cotton, W. R.: Parameterization and Impact of Ice initiation Processes Relevant to Numerical Model Simulations of Cirrus Clouds., *Journal of Atmospheric Sciences*, 51, 77–90, doi:10.1175/1520-0469(1994)051<0077:PAIOII>2.0.CO;2, 1994.
- 35

- DeMott, P. J., Prenni, A. J., Liu, X., Kreidenweis, S. M., Petters, M. D., Twohy, C. H., Richardson, M. S., Eidhammer, T., and Rogers, D. C.: Predicting global atmospheric ice nuclei distributions and their impacts on climate, *Proceedings of the National Academy of Sciences*, 107, 11 217–11 222, doi:10.1073/pnas.0910818107, <http://dx.doi.org/10.1073/pnas.0910818107>, 2010.
- Flato, G., Marotzke, J., Abiodun, B., Braconnot, P., Chou, S., Collins, W., Cox, P., Driouech, F., Emori, S., Eyring, V., Forest, C., Gleckler, P., Guilyardi, E., Jakob, C., Kattsov, V., Reason, C., and Rummukainen, M.: Evaluation of Climate Models. In: *Climate Change 2013: The Physical Science Basis. Contribution of Working Group I to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change* [Stocker, T.F., D. Qin, G.-K. Plattner, M. Tignor, S.K. Allen, J. Boschung, A. Nauels, Y. Xia, V. Bex and P.M. Midgley (eds.)], 2013.
- Fletcher: *The Physics of Rainclouds*, Cambridge University Press, 1962.
- 10 Fretwell, P., Pritchard, H. D., Vaughan, D. G., Bamber, J. L., Barrand, N. E., Bell, R., Bianchi, C., Bingham, R. G., Blankenship, D. D., Casassa, G., Catania, G., Callens, D., Conway, H., Cook, A. J., Corr, H. F. J., Damaske, D., Damm, V., Ferraccioli, F., Forsberg, R., Fujita, S., Gim, Y., Gogineni, P., Griggs, J. A., Hindmarsh, R. C. A., Holmlund, P., Holt, J. W., Jacobel, R. W., Jenkins, A., Jokata, W., Jordan, T., King, E. C., Kohler, J., Krabill, W., Riger-Kusk, M., Langley, K. A., Leitchenkov, G., Leuschen, C., Luyendyk, B. P., Matsuoka, K., Mouginit, J., Nitsche, F. O., Nogi, Y., Nost, O. A., Popov, S. V., Rignot, E., Rippin, D. M., Rivera, A., Roberts, J., Ross, N., Siegert, M. J., 15 Smith, A. M., Steinhage, D., Studinger, M., Sun, B., Tinto, B. K., Welch, B. C., Wilson, D., Young, D. A., Xiangbin, C., and Zirizzotti, A.: Bedmap2: improved ice bed, surface and thickness datasets for Antarctica, *The Cryosphere*, 7, 375–393, doi:10.5194/tc-7-375-2013, <http://dx.doi.org/10.5194/tc-7-375-2013>, 2013.
- Gorodetskaya, I. V., Kneifel, S., Maahn, M., Thiery, W., Schween, J. H., Mangold, A., Crewell, S., and Van Lipzig, N. P. M.: Cloud and precipitation properties from ground-based remote-sensing instruments in East Antarctica, *The Cryosphere*, 9, 285–304, doi:10.5194/tc- 20 9-285-2015, 2015.
- Grosvenor, D. P., Choularton, T. W., Lachlan-Cope, T., Gallagher, M. W., Crosier, J., Bower, K. N., Ladkin, R. S., and Dorsey, J. R.: In-situ aircraft observations of ice concentrations within clouds over the Antarctic Peninsula and Larsen Ice Shelf, *Atmospheric Chemistry and Physics*, 12, 11 275–11 294, doi:10.5194/acp-12-11275-2012, <https://doi.org/10.5194/acp-12-11275-2012>, 2012.
- Harig, C. and Simons, F. J.: Accelerated West Antarctic ice mass loss continues to outpace East Antarctic gains, *Earth and Planetary Science Letters*, 415, 134–141, doi:10.1016/j.epsl.2015.01.029, <http://dx.doi.org/10.1016/j.epsl.2015.01.029>, 2015.
- Herbert, R. J., Murray, B. J., Dobbie, S. J., and Koop, T.: Sensitivity of liquid clouds to homogenous freezing parameterizations, *Geophys. Res. Lett.*, 42, 1599–1605, doi:10.1002/2014gl062729, <http://dx.doi.org/10.1002/2014GL062729>, 2015.
- Hines, K. M. and Bromwich, D. H.: Development and Testing of Polar Weather Research and Forecasting (WRF) Model. Part I: Greenland Ice Sheet Meteorology, *Mon. Wea. Rev.*, 136, 1971–1989, doi:10.1175/2007mwr2112.1, <http://dx.doi.org/10.1175/2007MWR2112.1>, 2008.
- 30 Hong, S.-Y., Dudhia, J., and Chen, S.-H.: A Revised Approach to Ice Microphysical Processes for the Bulk Parameterization of Clouds and Precipitation, *Monthly Weather Review*, 132, 103, doi:10.1175/1520-0493(2004)132<0103:ARATIM>2.0.CO;2, 2004.
- Hoese, C. and Möhler, O.: Heterogeneous ice nucleation on atmospheric aerosols: a review of results from laboratory experiments, *Atmospheric Chemistry and Physics*, 12, 9817–9854, doi:10.5194/acp-12-9817-2012, <http://dx.doi.org/10.5194/acp-12-9817-2012>, 2012.
- Igel, A. L., Igel, M. R., and van den Heever, S. C.: Make It a Double? Sobering Results from Simulations Using Single-Moment Microphysics Schemes, *J. Atmos. Sci.*, 72, 910–925, doi:10.1175/jas-d-14-0107.1, <http://dx.doi.org/10.1175/JAS-D-14-0107.1>, 2015.
- 35 Khairoutdinov, M. and Kogan, Y.: A New Cloud Physics Parameterization in a Large-Eddy Simulation Model of Marine Stratocumulus, *Monthly Weather Review*, 128, 229, doi:10.1175/1520-0493(2000)128<0229:ANCPPI>2.0.CO;2, 2000.

- King, J. C. and Turner, J.: *Antarctic Meteorology and Climatology* (Cambridge Atmospheric and Space Science Series), Cambridge University Press, 1997.
- King, J. C., Lachlan-Cope, T. A., Ladkin, R. S., and Weiss, A.: Airborne Measurements in the Stable Boundary Layer over the Larsen Ice Shelf, Antarctica, *Boundary-Layer Meteorology*, 127, 413–428, doi:10.1007/s10546-008-9271-4, 2008.
- 5 King, J. C., Gadian, A., Kirchaessner, A., Kuipers Munneke, P., Lachlan-Cope, T. A., Orr, A., Reijmer, C., Broeke, M. R., J. M., and Weeks, M.: Validation of the summertime surface energy budget of Larsen C Ice Shelf (Antarctica) as represented in three high-resolution atmospheric models, *Journal of Geophysical Research (Atmospheres)*, 120, 1335–1347, doi:10.1002/2014JD022604, 2015.
- Lachlan-Cope, T.: Antarctic clouds, *National Institute Polar Research Memoirs*, 29, 150–158, doi:10.1111/j.1751-8369.2010.00148.x, 2010.
- Lachlan-Cope, T., Listowski, C., and O’Shea, S.: The Microphysics of Clouds over the Antarctic Peninsula - Part 1: Observations, *Atmospheric Chemistry and Physics Discussions*, pp. 1–28, doi:10.5194/acp-2016-331, <http://dx.doi.org/10.5194/acp-2016-331>, 2016.
- 10 Lawson, R. P. and Gettelman, A.: Impact of Antarctic mixed-phase clouds on climate, *Proceedings of the National Academy of Sciences*, 111, 18 156–18 161, doi:10.1073/pnas.1418197111, <http://dx.doi.org/10.1073/pnas.1418197111>, 2014.
- Lim, K.-S. S. and Hong, S.-Y.: Development of an Effective Double-Moment Cloud Microphysics Scheme with Prognostic Cloud Condensation Nuclei (CCN) for Weather and Climate Models, *Mon. Wea. Rev.*, 138, 1587–1612, doi:10.1175/2009mwr2968.1, <http://dx.doi.org/10.1175/2009MWR2968.1>, 2010.
- 15 Lubin, D., Chen, B., Bromwich, D. H., Somerville, R. C. J., Lee, W.-H., and Hines, K. M.: The Impact of Antarctic Cloud Radiative Properties on a GCM Climate Simulation*, *Journal of Climate*, 11, 447–462, doi:10.1175/1520-0442(1998)011<0447:TIOACR>2.0.CO;2, 1998.
- Meyers, M. P., Demott, P. J., and Cotton, W. R.: New Primary Ice-Nucleation Parameterizations in an Explicit Cloud Model., *Journal of Applied Meteorology*, 31, 708–721, doi:10.1175/1520-0450(1992)031<0708:NPINPI>2.0.CO;2, 1992.
- 20 Milbrandt, J. A. and Yau, M. K.: A Multimoment Bulk Microphysics Parameterization. Part I: Analysis of the Role of the Spectral Shape Parameter., *Journal of Atmospheric Sciences*, 62, 3051–3064, doi:10.1175/JAS3534.1, 2005a.
- Milbrandt, J. A. and Yau, M. K.: A Multimoment Bulk Microphysics Parameterization. Part II: A Proposed Three-Moment Closure and Scheme Description., *Journal of Atmospheric Sciences*, 62, 3065–3081, doi:10.1175/JAS3535.1, 2005b.
- Morris, E. M. and Vaughan, D. G.: Spatial and temporal variation of surface temperature on the Antarctic Peninsula and the limit of viability of ice shelves, in: *Antarctic Peninsula Climate Variability: Historical and Paleoenvironmental Perspectives*, pp. 61–68, Wiley-Blackwell, doi:10.1029/ar079p0061, <http://dx.doi.org/10.1029/AR079p0061>, 2003.
- 25 Morrison, H., Curry, J. A., and Khvorostyanov, V. I.: A New Double-Moment Microphysics Parameterization for Application in Cloud and Climate Models. Part I: Description, *J. Atmos. Sci.*, 62, 1665–1677, doi:10.1175/jas3446.1, <http://dx.doi.org/10.1175/JAS3446.1>, 2005.
- Morrison, H., Thompson, G., and Tatarskii, V.: Impact of Cloud Microphysics on the Development of Trailing Stratiform Precipitation in a Simulated Squall Line: Comparison of One- and Two-Moment Schemes, *Monthly Weather Review*, 137, 991, doi:10.1175/2008MWR2556.1, 2009.
- 30 O’Shea, S. J., Choularton, T. W., Flynn, M., Bower, K. N., Gallagher, M., Crosier, J., Crawford, I., Fleming, Z., Listowski, C., Kirchaessner, A., Ladkin, R. S., and Lachlan-Cope, T.: In situ measurements of cloud microphysics and aerosol over coastal Antarctica during the MAC campaign, *Atmospheric Chemistry and Physics Discussions*, 2017, 1–47, doi:10.5194/acp-2017-212, <http://www.atmos-chem-phys-discuss.net/acp-2017-212/>, 2017.
- 35 O’Donnell, R., Lewis, N., McIntyre, S., and Condon, J.: Improved Methods for PCA-Based Reconstructions: Case Study Using the Steig et al. (2009) Antarctic Temperature Reconstruction, *Journal of Climate*, 24, 2099–2115, doi:10.1175/2010JCLI3656.1, <http://dx.doi.org/10.1175/2010JCLI3656.1>, 2011.

- Peckham, S., Grell, G. A., McKeen, S. A., Barth, M., Pfister, G., Wiedinmyera, C., Fast, J. D., Gustafson, W. I., Zaveri, R., Easter, R. C., Barnard, J., Chapman, E., Hewson, M., Schmitz, R., Salzmann, M., and Freitas, S.: 2011: WRF-Chem Version 3.3 User's Guide. NOAA Technical Memo., 98 pp., 2011.
- 5 Powers, J. G., Manning, K. W., Bromwich, D. H., Cassano, J. J., and Cayette, A. M.: A Decade of Antarctic Science Support Through Amps, *Bull. Amer. Meteor. Soc.*, 93, 1699–1712, doi:10.1175/bams-d-11-00186.1, <http://dx.doi.org/10.1175/bams-d-11-00186.1>, 2012.
- Skamarock, W. C., Klemp, J. B., Dudhia, J., Gill, D. O., Barker, D. M., Huang, X.-Y., Wang, W., and Powers, J. G.: A description of the Advanced Research WRF Version 3, NCAR Technical Note NCAR/TN-475+STR, 113 pp, <http://dx.doi.org/10.1038/nature18645>, 2009.
- Stickney, T. M., Shedlov, M. W., and Thompson, D. I.: Goodrich total temperature sensors. Technical report 5755., Tech. rep., Goodrich Corporation, Burnsville, MN, 1994.
- 10 Thompson, G., Field, P. R., Rasmussen, R. M., and Hall, W. D.: Explicit Forecasts of Winter Precipitation Using an Improved Bulk Microphysics Scheme. Part II: Implementation of a New Snow Parameterization, *Monthly Weather Review*, 136, 5095, doi:10.1175/2008MWR2387.1, 2008.
- Turner, J., Lu, H., White, I., King, J. C., Phillips, T., Hosking, J. S., Bracegirdle, T. J., Marshall, G. J., Mulvaney, R., and Deb, P.: Absence of 21st century warming on Antarctic Peninsula consistent with natural variability, *Nature*, 535, 411–415, doi:10.1038/nature18645, <http://dx.doi.org/10.1038/nature18645>, 2016.
- 15 van Wessem, J. M., Reijmer, C. H., Lenaerts, J. T. M., van de Berg, W. J., van den Broeke, M. R., and van Meijgaard, E.: Updated cloud physics in a regional atmospheric climate model improves the modelled surface energy balance of Antarctica, *The Cryosphere*, 8, 125–135, doi:10.5194/tc-8-125-2014, <http://dx.doi.org/10.5194/tc-8-125-2014>, 2014a.
- Wesslén, C., Tjernström, M., Bromwich, D. H., de Boer, G., Ekman, A. M. L., Bai, L.-S., and Wang, S.-H.: The Arctic summer atmosphere: an evaluation of reanalyses using ASCOS data, *Atmospheric Chemistry and Physics*, 14, 2605–2624, doi:10.5194/acp-14-2605-2014, <http://dx.doi.org/10.5194/acp-14-2605-2014>, 2014.
- 20 Wilson, D. R. and Ballard, S. P.: A microphysically based precipitation scheme for the UK meteorological office unified model, *Q.J Royal Met. Soc.*, 125, 1607–1636, doi:10.1002/qj.49712555707, <http://dx.doi.org/10.1002/qj.49712555707>, 1999.
- Wouters, B., Martin-Español, A., Helm, V., Flament, T., van Wessem, J. M., Ligtenberg, S. R. M., van den Broeke, M. R., and Bamber, J. L.: Dynamic thinning of glaciers on the Southern Antarctic Peninsula, *Science*, 348, 899–903, doi:10.1126/science.aaa5727, 2015.
- 25 Yang, X., Pyle, J. A., and Cox, R. A.: Sea salt aerosol production and bromine release: Role of snow on sea ice, *Geophys. Res. Lett.*, 35, doi:10.1029/2008gl034536, <http://dx.doi.org/10.1029/2008GL034536>, 2008.

Table 1. WRF settings used for the simulations. The number in parenthesis indicates the scheme number (option) in the WRF settings.

Setting	Value
Number of domains	3
Domains size (px)	80/130/208
Resolution (km)	45/15/5
Number of vertical levels	30
Top pressure (hPa)	50
time step (s)	180/60/20
Cumulus param	on/on/off
LW radiation scheme	RRTM (1)
SW radiation scheme	Goddard (2)
Surface atmospheric layer	Eta similarity (2)
Land surface physics	Noah Land Surface model (2)
Planetary Boundary layer	Mellor-Yamada-Janjic (2)

Table 2. Microphysics schemes of WRF (version 3.5.1) used in this work with their predicted cloud variables. DM stands for Double-Moment scheme (see text for details). All prognosed hydrometeors variables are designated by letters as follows. c:clouds droplets; i: ice crystals; r:rain drops; s:snow crystals; g:graupel; h:hail. The Morrison scheme can be used as a Double-moment scheme for droplets only when WRF is used with WRF/Chem. See text for the references related to the cloud microphysics schemes.

Scheme	Mass	Number	Comment
WSM5	c,r,i,s	-	Used in the Antarctic Mesoscale Prediction System (AMPS)
WDM6	c,r,i,s,g	c,r	Upgrade of WSM5 to DM for c,r + predicted CCN
Morrison	c,r,i,s,g	r,i,s,g	Used in the Arctic System Reanalysis (ASR)
Thompson	c,r,i,s,g	r,i	State of the art parameterisation of snow
Milbrandt	c,r,i,s,g,h	c,r,i,s,g,h	DM for all hydrometeors + predicted CCN

Table 3. Characteristics of the ice phase, and liquid phase activation, for the microphysics schemes. T refers to the atmospheric temperature, and qc to the liquid water content. S_i (resp. S_w) is the saturation ratio with respect to ice (resp. liquid water). $r_{ice/snow}$ indicate the cut-off size for icy particles consider either as ice crystals (smaller particles) or snow (larger particles). INP parameterisations (INP param.) account for the various freezing processes presented in section 1: *imm* is immersion freezing; *dep* is deposition freezing; *cont* is contact freezing; *cond.* is condensation freezing; *hom.* is homogeneous freezing (considered as instantaneous, ie straight conversion of liquid to ice). IN/freezing parameterisations' references: (Fmod) is a modified version of Fletcher (1962) presented in Hong et al. (2004); (C) is Cooper (1986); (M) is Meyers et al. (1992) eq. 2.4; (M') is Meyers et al. (1992) eq. 2.6; (B) is Bigg (1953) for probabilistic freezing; (DeM) for Demott et al. (1994) for probabilistic freezing. CCN activation parameterisations: (K) is Khairoutdinov and Kogan (2000); (CP) is Cohard and Pinty (2000).

Scheme	Triggering of Ice formation	INP param.	$r_{ice/snow}$	Droplets/CCN
WSM5	$S_i > 1$ [T < 0°C & qc > 0] [T < -40°C & qc > 0]	(Fmod) <i>dep</i> (B) <i>imm</i> <i>hom</i>	250 μm	300 cm^{-3}
WDM6	same as WSM5			CCN (K)
Morrison	[T < -8°C & $S_w > 0.999$] or [$S_i > 1.08$] [T < -4°C & qc > 0] [T < -40°C & qc > 0]	(C) <i>dep, cond</i> (M') <i>cont</i> + (B) <i>imm</i> <i>hom</i>	125 μm	250 cm^{-3}
Thompson	[T < -12°C & $S_w > 1.$] or [$S_i > 1.25$] [T < 0°C & qc > 0] [T < -38°C & qc > 0]	(C) <i>dep, cond</i> (B) <i>imm</i> <i>hom</i>	200 μm	100 cm^{-3}
Milbrandt	[T < -5°C & $S_i > 1$] [T < -2°C & qc > 0] [T < -30°C & qc > 0] [T < -50°C & qc > 0]	(M) <i>dep, cond</i> (M') <i>cont</i> (DeM) <i>hom</i> <i>hom</i>	100 μm	CCN (CP)

Table 4. Monthly averaged shortwave (SW) and longwave (LW) surface radiative biases of daily averaged biases over Rothera, AWS14, and AWS15 for the two time periods of interest. The exponent gives the standard deviation (STD) of the daily biases. The number of “x” symbols as subscript tells how significant the difference is between WSM5 and each of the other three schemes (one, two, or three “x” mean statistical significance at the 90%, 95%, or 99% level, respectively). No symbol means that the difference is not significant. Statistically significant reductions in SW/LW biases are emphasised with bold characters.

Radiation bias	Microphysics scheme	Rothera		AWS 14		AWS 15	
		2010	2011	2010	2011	2010	2011
SW ^{STD} (W m ⁻²)	WSM5	15 ⁶⁸	49 ⁷⁶	-28 ⁶¹	53 ⁵²	-22 ⁵²	48 ⁵¹
	Morrison	20 ⁶³	52 ⁸⁴	-51 ⁶⁰ _x	-5 ⁶¹ _{xxx}	-43 ⁵² _x	-12 ⁵⁰ _{xxx}
	Thompson	46 ⁶⁷ _{xx}	70 ⁸⁹	-24 ⁶²	37 ⁵²	-24 ⁵¹	29 ⁵⁰ _x
	Milbrandt	7 ⁶³	48 ⁸²	-33 ⁶²	31 ⁵⁶ _x	-30 ⁵⁰	28 ⁵³ _x
LW ^{STD} (W m ⁻²)	WSM5	-28 ²⁵	-26 ²⁶	-11 ²⁸	-20 ²³	-10 ²⁹	-22 ²⁰
	Morrison	-22 ²²	-22 ²⁶	2 ²⁵ _{xx}	1 ²¹ _{xxx}	4 ²¹ _{xx}	1 ¹⁹ _{xxx}
	Thompson	-24 ²⁶	-25 ²⁷	0.5 ²⁶ _x	-6 ²¹ _{xxx}	3 ²⁵ _{xx}	-9 ²² _{xxx}
	Milbrandt	-19 ²⁰ _x	-19 ²³	3 ²⁶ _{xx}	-6 ²⁰ _{xxx}	5 ²⁴ _{xx}	-9 ²⁵ _{xx}

Table 5. Average ratio (%) of the number of occurrences of $LWC > 0.01 \text{ g kg}^{-1}$ in the simulations over the observations. The average is derived from the flight averages.

Region Year	WSM5	Morrison	Thompson	Milbrandt
West 2010	47	72	69	5
West 2011	54	49	4	88
East 2010	<1	7	3	3
East 2011	<1	215	130	105

Table 6. Statistics over the flight tracks on the mixed-phase ratio $f_m = \text{LWC}/(\text{LWC} + \text{SWC})$, for temperatures $T > -15^\circ$ (see text for details)

f_m	Observation	WSM5	Morrison	Thompson	Milbrandt
average	0.84	0.6	0.91	0.66	0.79
σ	0.08	0.24	0.1	0.20	0.10
max	0.94	0.9	1	0.92	0.95
min	0.60	0.07	0.61	0.21	0.56

Table 7. The median of the flight-averaged $\epsilon = |\text{INPparam} - \text{Obs}|/\text{Obs}$, and the corresponding median absolute deviation for different INP parameterisations, for the two different years, respectively. DeMott refers to (DeMott et al., 2010), and Fletcher to (Fletcher, 1962), see Table 3 for the other references.

	2010		2011	
	ϵ	$\Delta\epsilon$	ϵ	$\Delta\epsilon$
DeMott, variable n_{aer}	0.48	0.26	0.61	0.27
DeMott, $n_{\text{aer}}=0.1$	0.42	0.21	0.65	0.18
DeMott, $n_{\text{aer}}=1$	1.82	1.04	1.98	1.33
Fletcher	0.73	0.24	0.75	0.43
(Fmod)	26.7	4.50	21.10	12.16
(C)	1.56	1.08	2.00	1.47
(M), $S_w=1$	24.35	7.21	23.35	16.32
(M), $S_w=0.9$	4.65	1.40	4.38	3.48