Response to reviewer 1:

We thank the reviewer for their supportive and insightful comments. In the first paragraph the reviewer mentions that this study does not include Crocus and SNOWPACK. These would be worthy to include in such a study, but would add the complication of differences in layering structure. Here, we have chosen an Eulerian (fixed layer) grid structure to allow better intercomparison of the internal snow physics process representation. Additionally, remote sensing retrieval applications are unlikely to be based on snowpack models of more than 5 layers, which rules out Crocus and SNOWPACK for operational, large-scale use.

The reviewer also mentions that this paper does not provide a recommendation as to the least worst model combinations. This is deliberate. The optimum choice depends on the end application (included in the discussion), but also could be totally different for an alternative site. We urge the snowpack and remote sensing community to undertake synergistic field campaigns to collect data at alternative sites with different snow conditions to allow this study to be repeated elsewhere.

With regards to the review points:

The main concern is that the reader should have a better idea of the order of magnitude of errors compared to those obtained when the RMT are driven by snow measurements (which can be considered as the reference level for RMSE and bias generally obtained in practice, under best conditions). There are a lot of papers from Finish, Canadian or American groups who provide such values. Do the scaled snow models reach the same mean level of RMSE when compared to measurements?

This is not a straightforward question to answer as even the snow measurements are usually scaled, and it will depend many things. The following paragraph has been added to the discussion:

Here, the lowest bias and RMSE for unscaled microstructure simulations were -6.9 to +6.9K and 4.2 to 12.2K respectively, but depended on microwave model, frequency and polarization. In an attempt to put these results into context, there are a number of studies that have quantified brightness temperature simulation errors for these models. These fall into different categories, depending on sensor characteristics, the source of the evaluation data (ground-based, airborne, satellite) and presence of ice lenses (Derksen et al. 2012), the treatment of the snow microstructure Picard et al. 2014), snow type, observation angle and the specific electromagnetic model Tedesco and Kim (2006) and the underlying substrate (Lemmetyinen et al., 2009, Derksen et al. 2014). Examples of unscaled field observations of microstructure compared with ground-based observations include the HUT simulations of Derksen et al. 2012 who found an RMSE of 10-34K and Rutter et al., 2014 who found a bias of 34-68K that was reduced to < 0.6K upon application of grain scale factors of 2.6-5.3. Scaling, or best-fit relationships were used by Durand et al., 2008 (mean absolute error 3.1K at V-pol and 9.3K at H-pol), Montpetit et al., 2013 (RMSE 8-20K), Brucker et al., 2011 (RMSE 1.5K), Picard et al., 2014 (RMSE 1-11K) and Roy et al., 2013 (RMSE 12-16K). However, in some cases the frequency-dependent results have been combined and in others kept separate.

End of page 13 and beginning of page 14: You must clearly mention that when considering SSA measurements for retrieving the optical diameter as inputs of RTM, one must scaled the measurements too.

This has been added to line 35: This is highlighted by the treatment of field observations of specific surface area to derive optical diameter. Even these require some form of scaling (e.g. Montpetit et al., 2013, Picard et al., 2014, Rutter et al., 2014).

All the figures are of very poor quality, making this revision very difficult, perhaps this is an issue with building the pdf for submission, but it should be considered prior to next revision; text of scales is really too small.

Please accept our apologies for this, we will endeavour to ensure good print quality. Figure 4 has been changed so that the two graphs are vertically aligned and larger. The font size on Figure 6 has been increased.

With regards to the specific comments:

I am not sure about the interest of giving all the equations in Section 2.3, especially if there are errors! See Eq. 7: (1-f) should be in numerator, i.e, factor of (es-eb).

We included these equations because a comparison between scattering coefficients has been missing in other publications and it highlights the empirical versus physics origin of these different models. Thank you for identifying this error, the typo has now been corrected.

Eq. 8 "it is given by the largest solution to…" largest of what?
Changed to 'the largest of the two solutions to the quadratic equation:'

Page 7 beginning of the Section 2.4: Why to combine HUT-MEMLS? not clear.
This has been changed to:

Interfacing of the various model inputs and outputs was enabled through the development of the ensemble framework, via a combination of shell script and Octave/Matlab code. The DMRT-ML model was run from the shell script, which subsequently calls an Octave / Matlab script to run HUT and MEMLS. HUT and MEMLS run alternately in this framework as the soil parameters (common between DMRT-ML and HUT) are used to calculate soil reflectivity in HUT, which is then used as the lower boundary condition in MEMLS.

Table 6 : The too much detailed analysis of the differences obtained here are probably not significant? A more synthetic analysis of results should be presented: mainly RMSE identical for the 3 models, large differences between years?

It is important to show the variation between models and years, and between frequencies and polarizations. Whilst the scaling may not have a large overall effect, in reality it improves the largest errors, but has a detrimental effect on the smallest

errors. It's important to highlight that scaling is not a universal solution, and we need a better understanding of the physics instead.

Again, page 11, line 17: "the lowest RMSE in brightness temperature is reduced, as shown in Table 8 " probably not statistically significant?
This phrase has been removed, and we focus only on individual models.

Page 12, line 28 : "For DMRT-ML, consideration of the stickiness is imperative, with the choice of microstructure evolution model of secondary importance. " This statement is in agreement with Roy et al. (2013).

Reviewer 2 asked for non-sticky simulations to be removed as this point can be considered a given. The Roy et al. (2013) reference is used as justification for this in section 2.3.1. Instead, stickiness values of 0.1 and 0.2 have been used.

The paper has been amended to say:
Roy et al., (2013) and Löwe and Picard (2015) showed that non-sticky representation in DMRT-ML is inappropriate. For this model ensemble, two DMRT-ML varieties have been chosen to capture the range of brightness temperatures simulated: "DMRT less sticky" (tau = 0.2) and "DMRT very sticky" (tau = 0.1). These two values represent reasonable values used by others (e.g. Tsang et al, 2007, Shih et al.,1997).

Page 13 "whereas the range in Brucker et al. (2011) was 0.25-0.4 for snow density between 100 and 400 kg m 3. " In the Brucker's paper, it is the reverse. The scaling factor noted Beta = 0.63, and this sentence refers to : A = [2/3 Beta (1 –f)]. Thus, when the density is 100 kg/m3, A = 0.4 and when the density = 400 kg/m3, A = 0.24.
We did not intend to imply an inverse relationship, merely that the density range of 100-400 kg m^-3 corresponds to a scaling range between 0.25 and 0.4. However, we will report the scaling range to be between 0.4-0.25 to remove ambiguity. 0.25 was reported in the paper, but 0.24 is calculated from the numbers given. Presumably limiting beta to two significant figures accounts for the difference here.

Figure 2: I cannot read on the this figure if ice crusts have occurred during these two years (i.e precipitations during winter with T>0)? This has certainly been observed if this was the case. Please clarify.
The following text has been added to section 2.5:

November rain events occurred in both years, and also in early December in 2011-2012. Layers with melt-freeze polycrystals and other melt forms were detected in snow pit observations during both seasons.

Figure 3 : A minor observation for the SWE measurements using the GMON: the authors observed a typical artefact of this instrument at the end of the season, when the soil is very wet due to the melting snow, leading to a strong anomaly in measurements (very high values in what is interpreted as SWE, while there is no more snow!).

This is an excellent point. The SWE > 0 points have been removed from Fig 3 when snow depth = 0, and the caption amended to reflect this processing to include:
Note that erroneous positive SWE observation points have been removed at the end of the season when snow depth is zero, as this is a sensor artifact related to soil moisture changes.

Table 4 could be interesting in percentage?

Percentages have been added to Table 4. The text has been amended to say:

The mean absolute error and mean relative difference for these simulations are presented in Table 4. SNT had the lowest bias (0.12mm) in 2011-2012, whereas SNI had the lowest bias (-0.14mm) in 2012-2013. Bulk grain diameter simulated by the microstructure models lead to a mean difference of between -53% and +45% relative to the observations.

Response to Reviewer #2: G. Picard

We thank Dr. Picard for his thorough and rigorous review of our paper. It is stated that the two main limitations of the study are inherent to the methodology and the paper is worth publishing regardless. However, we would like to respond to these comments on the limitations.

Firstly, it is true that the empirical models and models based on physics are given equal consideration. This was done to reflect the way that the models are used in the literature. Different models have different complexities due to their different intended applications. Here we intended to present this methodology as a way of identifying which processes need to be considered carefully in coupled snowpack-emission models. For processes that have little sensitivity, it doesn't matter which representation is chosen. For others with large sensitivity, it really does, and even the representations we do have are not up to scratch.

This leads on to the second limitation, that the paper does not address the physical reasons for the differences, merely which ones could be selected for further study. Investigating the physical reasons for the difference is very difficult to do within the framework of the study. It is touched upon with the snowpack component, which is computationally set up to do this thanks to Essery (2015), which brings all representations together. At present there is no equivalent electromagnetic model in the literature. This particular study is no substitute for rigorous intercomparisons of electromagnetic models such as Pan et al. (2016) and Löwe and Picard (2015). Instead, this study illustrates how future studies might be constructed. This would be substantially easier to do with a common electromagnetic model framework.

However, the potential for this study to be interpreted as 'empirical is better' is concerning. An empirical model may give a better performance at a specific site in a specific year but in no way should this be taken as a general recommendation from this paper. Nor does it help us understand why. It could be good enough for a specific application. It won't be good enough for all applications. Nor is it likely to be improved upon. A comment on this has been made in the discussion (please see later response to comment *P13L28)*

With regard to the parameter choice, we are satisfied that the literature demonstrates sufficiently that non-sticky parameterisations should never be used. We have removed the non-sticky simulations, and instead will look at the two reasonable values of stickiness applied in the literature, namely 0.1 and 0.2. This makes little difference to Figure 5, because it is dominated by the extreme stickiness = 0.1 range. It does, however, affect grain scale factor, Tb bias and RMSE, as can be expected. For the MEMLS simulations, we chose IBA as the physics-based electromagnetic model within MEMLS and the empirical scattering coefficient that is the default recommended parameterization. These are well documented and commonly used. As highlighted, there are a number of other representations of scattering coefficients (and even within the scattering coefficient model: IBA has 3 treatments of the effective permittivity and we have chosen only one). Wiesmann and Mätzler (1999) indicated that several empirical fits were available, which are versions included in the code. However, we used only the two recommended versions, commonly used in the literature, to avoid the discussion being dominated by inter-MEMLS comparisons or

disproportionately by empirical representations.

Specific comments:

P1L5: "(JIM)" is needed for reference at the end of the abstract.
Acronym has been replaced with Jules Investigation Model
P2L7: "only". Photogrammetry and altimetry seem to be good (better?) candidates.
Photogrammetry has great potential for catchment scales but isn't feasible for global scales. The text has been amended to: Microwave, altimetry or coarser-scale gravity satellite sensors offer the only feasible way to measure snow mass or depth on a global scale, with microwave observations spanning the longest time scale of these. However, microwave algorithms….
P2L9: "because of the because"
Removed 'because of the'
P2L30: "carried out Tedesco and Kim"
'by' added
P3L20: the information in parenthesis is not clear
Parenthesis removed, next sentence begins: As this earlier study did not incorporate snow microstructure changes, JIM was coupled with three microstructure evolution functions for this study,…
P5 Eq 6: add "i" in subscript to omega [change made]
P5L23: incidence-> zenith [change made]
P6 L24: 1) Is the Eq 10 only valid for average angle or this assumption is used to perform the integral ? 2) Please also check if this is the angle or the square sine/cosine of the angle that is averaged. 3) At last, I'd recommend to quote the original statement (in the reference paper) to avoid attributing this (possibly wrong) statement to your study. Or better, provide your own demonstration in Annex.
The actual calculation in MEMLS is a triple integral: an integration over the angular-dependent factor I sin^2 chi (Mätzler and Wiesmann, 1999, eqn 9) to find the azimuth-averaged phase function, then an integration over the zenith and scattering angles according to eqn 11a of Mätzler and Wiesmann, 1999, eqn 11a. However, this is not so relevant to the discussion in light of other changes, and the section is stripped back to just the definition of the phase function without the details of calculation.
P7 L3: "Eps_eff" should be defined, here.
Effective permittivity was not defined for DMRT-ML, however, here a precise reference for this has been given. The exact definition will be of limited interest and more suited to a rigorous intercomparison. The text has been amended to: It should be noted that the choice of oblate grains also affects the effective permittivity in I, represented by an empirical, density-dependent effective permittivity (Wiesmann and Mätzler, 1999, equations 45-47) for this case.
P7 Eq 13: "d_o" should be defined.
This has been added in the second sentence in the HUT section: These are nominally suited to different grain diameter (d_0) ranges, with some overlap between them.

No distinction is made here regarding microwave grain diameter vs maximal grain extent. This is left to the discussion section.
P8L1. It is not immediately clear where 189 comes from (3 times 63 I suppose). Maybe a slight reformulation would help.
The text has been amended to: Meteorological data are used to drive the 189 configurations of JIM (3 microstructural models for each of the 63 snowpack

P8L14: "there where" [where -> were]

P8L23: change of the state of the ground is not taken into account whereas it is an important factor in the arctic environment. Could you add information on this and provide some hints on the impact on the simulation performances ?

This is not considered important for this study because the permittivity measurements were so stable for the 1st December – 31st March period chosen. The evaluation was also constrained to a similar period, and the observations were very limited prior to December. The state of the ground is an interesting topic, and earlier stages of this study involved optimization of the soil permittivity, but raised many questions about the methodology. This will be revisited in the future as permittivity measurements are not commonplace but allowed us to simplify this study in this regard.

P9L25: this results seem dependent on the scaling factor for the precipitation. Another factor would give in different results, isn't it ? If yes, maybe better to remove this part.

The scaling factor is rather critical in terms of the snowpack modelling. It is essential to compensate for precipitation sensor undercatch as simulation of the correct temperature gradients rely on good simulation of snow depth. If these are wrong then the microstructural evolution functions (at least SNTHERM and SNICAR as these depend on the temperature gradient, whereas MOSES does not) will also be incorrect. The impact of the precipitation scaling factor itself on the microstructure is outside the scope of this study, but this modelling approach could be used to look at it.

P10L11 – L34 and Figure 5. The results on the spread presented in this part and in the figure depend on the choices of the parametrization and these choices seem to me unfair, i.e; orientated so that one model appears to behave a very different way from the others. This difference is however only the consequence of the choice 1) of considering the stickiness as a free parameter and 2) of the particular stickiness values. The very large value (ie. None-sticky case) is known to be unrealistic based on several recent studies. Note also that the Tsang's group usually uses values of 0.1 or 0.2. The fact that DMRT theory has two parameters to describe the microstructure while others (apparently) use only one, does not mean that any choice of these two parameters represent snow and are valid. Conversely, why only two parameterizations of Ks in MEMLS has been used while _12 difference ones are available in the code ?

For HUT, why the numbers appearing in equations 13, 14 and 15 (which are no more than parameters of these equations recommended by some authors) are not freely changed in this study (e.g. +-20%) to reflect the treatment in DMRT ? These choices are subjective and have too much consequences on the conveyed message of this part. This is not critical for the paper as this has no impact on the major conclusions but it suggests that the models are very different, while I thing this is mostly due to the choice.

My recommendation is to narrow the range of stickiness in DMRT-ML and to explain why all the MEMLS parameterizations have not been used (or if possible/relevant to used them)

The recommendation here has been followed, with the removal of non-sticky DMRT and simulation of stickiness = 0.2. In fact, the DMRT-ML simulation range is dominated almost entirely by the impact of microstructure for the stickiness = 0.1 case. The range for stickiness = 0.2 only is much narrower. The new stickiness range does, however, have an impact on the mean bias and RMSE for DMRT-ML.

P12L12 – 15. References are needed to support the statements.

References added, and SNICAR description changed to state that it is an albedo model.

P12L25: The given reason is probably not valid, DMRT-ML predicts exactly the same propagation (within the layers) between the two polarizations (isotropic medium). This is the expected behavior for a random medium made of spheres and is indeed a sanity check used to verify the implementation. The difference in the terms of the phase function for H and V polarization in DMRT-ML are purely geometrical and the difference is canceled by the subsequent numerical integrations whereas in IBA MEMLS, it is canceled by analytical integration (or more precisely, it is removed by referring to the physical principles, the same used in DMRT-ML for the sanity check). The difference between TbH and TbV in DMRT-ML as well as in MEMLS and HUT is solely due to the interfaces. The scattering plays a role, but just because of the interaction between the volume and the interfaces, not because of difference of propagation in the volume.

A possible reason of the difference between DMRT-ML and the other models is because the initial choice of scaling (=none, with none-sticky) used in DMRT-ML makes the model really far from the observations, so that the scaling results in a general improvement.

This comment is largely accepted. However, with a narrower range of stickiness, the DMRT-ML results are improved further. The improvements in DMRT-ML seem to be across the board, but MEMLS involves more of a trade-off. This is probably an artifact of the methodology. Without scaling, MEMLS gives reasonable results at the higher frequency, yet the scaling improves the lower frequency results at the expense of the higher frequency. In this study, the results for the different microstructure models are combined with the different versions of the same model. If the results are broken down then taking MEMLS 2011-2012 bias as an example, the optimized Tb bias improves at all frequencies and polarizations with three exceptions: IBA-MOS, where the 19H bias is less than 1K worse than the non-optimised results, and EMP-SNI combination where the improvements at the lower frequency result in poorer higher frequency simulations at both polarizations. This is sufficient to have an effect on the overall results for MEMLS. The text in this section has been changed to:

In part, this may be due to the methodology of this study as the cost function is calculated per microstructure-electromagnetic model configuration, yet the bias and RMSE are presented for each electromagnetic model family. An individual contribution can influence the group in a non-intuitive way.

P13L28. Please distinguish here H and V polarizations.
Looking at it in this way led to some really interesting results, with higher V-pol difference than H-pol. Table 9 has been added, which looks at all the H and V maximum differences for each unscaled microstructure – electromagnetic model combinations for year 1. Broadly, the more scattering that takes place (using grain scale factor as a proxy), the bigger the impact of the snowpack parameterizations. This is intuitive, but has been demonstrated here.

This has been added to the results:
Overall, the snowpack parameterisations with MOSES microstructure and empirical MEMLS lead to a mean difference in the 36.5GHz brightness temperature of 11K at H-pol and 18K at V-pol. The maximum difference in 36.5GHz brightness temperature

was 33K at H-pol and 54K at V-pol for the 2011-2012 season. The maximum difference between H and V polarization for all unscaled microstructure-electromagnetic model combinations is demonstrated in Table 9. Large differences in the maximum brightness temperature difference as a result of the 63 snowpack configurations occurred for the SNT microstructure. Except for DMRT-ML less sticky and HUT with MOS or SNI microstructure, the V-pol difference is greater than the H-pol difference.

This has been added to the discussion:
Although empirical MEMLS driven by MOSES was chosen as an example to demonstrate the impact of parameterisations, this was purely because of the apparent consistency between the MOSES grain diameter converted to exponential correlation length and MEMLS simulations for 2011-2012 at this site. This is not a general endorsement of empirical models, as those based on physics are expected to be more universally applicable, but the specific application of these models will dictate the balance of accuracy versus simplicity. Extending the analysis beyond this example, snow parameterizations affect other unscaled model combinations to varying degrees. Microstructure scaling factors in Table 7 can be used as a proxy for the degree of scattering in the unscaled simulations. A higher scale factor acts to increase the simulated scattering, so for a scale factor <1, too much scattering occurs in the unscaled simulations. Snowpack parameterisations have a greater impact for a higher degree of scattering, larger at V-pol than H-pol. This is because scattering is already greater at H-pol so the spread in H-pol simulations as a result of snowpack parameterisations is suppressed by the existing level of scattering. The converse applies for high scaling factors (e.g. MOS with less sticky DMRT-ML).

Both conclusions and abstract have had the equivalent of this added:
The impact of snowpack parameterisation increases as the microwave scattering increases.

P14L8. Update the reference. [done]
Figure 7. I suggest to add (light) lines between related points. In addition, the legend is difficult to understand. Please add a reference for "cluster analysis" or use a simple wording to explain what it is. Remove the external reference for the value on the x-axis and if possible add a few words to make the Figure legend more self-sufficient.
It is not helpful to add lines between related points as the black dots (1st February) map onto the red dots (1st May) for each parameterization choice (0, 1, 2). These lines would be a series of superimposed vertical lines. To remove the external reference would make this plot very hard to read as 0,1,2 for snow densification corresponds to three different densification equations. These are separate to the 0, 1, 2 options for thermal conductivity, which relate to a different set of three equations. Similarly for fresh snow density and snow hydrology. However, the caption has been changed to:

Cluster analysis of brightness temperature simulated by MOSES microstructure and empirical MEMLS for 1st February 2012 (black dots) and 1st May 2012 (red dots) according to model parameterization choices. Brightness temperature simulations are split according to the different representations for each process representation. Values 0, 1, 2 relate to parameterizations given in Essery et al., 2013 and as described in section 2.1 Where distinct clusters occur that differ between parameterizations, this

indicates sensitivity to the parameterization.

Table 3. Ratio is well not defined and difficult to understand. Why not "Comparison of grain diameter simulated by different microstructure models. The mean and max ratio between pair of models is given in columns." ? Please reformulate. [Done, as suggested]

Table 5. Similarly "Ratio of mean brightness temperature ranges" is not clear. Also: would be "pair" more correct than "two" ? [Done, similar to the above]

Microstructure representation of snow in coupled snowpack and microwave emission models

Melody Sandells[1], Richard Essery[2], Nick Rutter[3], Leanne Wake[3], Leena Leppänen[4], and Juha Lemmetyinen[5]

[1]CORES Science and Engineering Limited, Burnopfield, UK
[2]Edinburgh University, Edinburgh, UK
[3]Northumbria University, Newcastle-upon-Tyne, UK
[4]Finnish Meteorological Institute, Arctic Research Centre, Sodankylä, Finland
[5]Finnish Meteorological Institute, Helsinki, Finland

Correspondence to: Melody Sandells (melody.sandells@coresscience.co.uk)

Abstract. This is the first study to encompass a wide range of coupled snow evolution and microwave emission models in a common modelling framework in order to generalise the link between snowpack microstructure predicted by the snow evolution models and microstructure required to reproduce observations of brightness temperature as simulated by snow emission models. Brightness temperatures at 18.7 and 36.5GHz were simulated by 1323 ensemble members, formed from 63 Jules Investigation Model snowpack simulations, three microstructure evolution functions and seven microwave emission model configurations. Two years of meteorological data from the Sodankylä Arctic Research Centre, Finland were used to drive the model over the 2011-2012 and 2012-2013 winter periods. Comparisons between simulated snow grain diameters and field measurements with an IceCube instrument showed that the evolution functions from SNTHERM simulated snow grain diameters that were too large (mean error 0.12 to 0.16mm), whereas MOSES and SNICAR microstructure evolution functions simulated grain diameters that were too small (mean error -0.16 to -0.24mm for MOSES, and -0.14 to -0.18mm for SNICAR). No model (HUT, MEMLS or DMRT-ML) provided a consistently good fit across all frequencies and polarizations. The smallest absolute values of mean bias in brightness temperature over a season for a particular frequency and polarization ranged from ~~0.9 to 7.2~~0.7 to 6.9K.

Optimal scaling factors for the snow microstructure were presented to compare compatibility between snowpack model microstructure and emission model microstructure. Scale factors ranged between 0.3 for the SNTHERM-Empirical MEMLS model combination (2011-2012), and ~~5.0 or greater when considering non-sticky particles in~~ 3.3 for DMRT-ML in conjunction with MOSES ~~or SNICAR~~ microstructure (2012-2013). Differences in scale factors between microstructure models were generally greater than the differences between microwave emission models, suggesting that more accurate simulations in coupled snowpack-microwave model systems will be achieved primarily through improvements in the snowpack microstructure representation, followed by improvements in the emission models. Other snowpack parameterisations in the snowpack model, mainly densification, led to a mean brightness temperature difference of 11K ~~when the JIM~~ at 36.5GHz H-pol and 18K at V-pol when the Jules Investigation Model ensemble was applied to the MOSES microstructure and empirical MEMLS emission model for the 2011-2012 season. The impact of snowpack parameterisation increases as the microwave scattering increases.

Consistency between snowpack microstructure and microwave emission models, and the choice of snowpack densification algorithms should be considered in the design of snow mass retrieval systems and microwave data assimilation systems.

1 Introduction

Global observations of the snow cover extent from optical and microwave satellite observations combined with in-situ data have shown a reduction in the spring snow cover (Brown et al., 2010; Brown and Robinson, 2011). Observed decline in snow cover extent during 2008-2011 exceeded that predicted by climate models (Derksen and Brown, 2012). Observations also indicate that duration of snow cover is also reducing, but cannot determine whether mass or volume of snow has changed.

Microwave, altimetry or coarser-scale gravity satellite sensors offer the only feasible way to measure snow mass or depth on a global scale, with microwave observations spanning the longest time scale of these. However, microwave algorithms such as those developed by Chang et al. (1987); Kelly (2009) can result in large errors because of the high sensitivity of applied forward models to parameterization of the snow microstructure (Davenport et al., 2012). In particular, the assumption of a fixed snow scatterer radius in the Chang et al. (1987) algorithm does not reflect the naturally changing snowpack structure. Errors in snow mass products derived from these algorithms mean that the products are difficult to use for evaluation of snow mass in climate models (Clifford, 2010) and unsuitable for assimilation into land surface models for streamflow forecasts (Andreadis and Lettenmaier, 2006). Development of the assimilation-based technique in GlobSnow allows changes in the snow microstructure to be taken into account through inversion of ground-based observations of snow depth and coinciding microwave brightness temperatures (Takala et al., 2011). Although more accurate than other global products, some errors remain, and the GlobSnow accuracy relies on the proximity and representativity of the ground stations (Hancock et al., 2013). In addition, the intermediate retrieval of the snow 'grain size' in GlobSnow is a parameter that also incorporates other land surface features, so is not a true representation of the snow effective diameter (Lemmetyinen et al., 2015).

Snowpack evolution models offer a way to estimate temporal changes in snow microstructural parameters and stratigraphy (e.g. Lehning et al., 2002; Brun et al., 1992). Intercomparison studies have shown large differences between snow evolution models driven by the same forcing data (Rutter et al., 2009). Given that the mass inputs were the same for the 33 snow models considered in the SNOWMIP2 study of Rutter et al. (2009), it is differences in the internal snow physics and model structure (layering assumptions) that result in the wide range of simulated depth and snow mass. Temperature, temperature gradient and density drive changes in the snow microstructure (e.g. Flanner and Zender, 2006), so it is likely that different snow physics assumptions in a coupled snowpack and emission model result in different thermal structures, microstructure parameterisations and ultimately different microwave extinction behaviour.

Theoretical differences between specific electromagnetic models have been examined in Löwe and Picard (2015); Pan et al. (2015), and other intercomparisons carried out by Tedesco and Kim (2006). These studies are useful for interpreting differences in electromagnetic model outputs for a snapshot profile of the snowpack properties. Given the dependence of microwave scattering on snow microstructure, a satellite retrieval system needs some quantification of microstructure. Snowpack evolution

modelling offers a means to quantify the metamorphic changes in snow microstructure. Indeed, snowpack evolution models have been coupled with microwave emission models to demonstrate the potential of this approach for snow remote sensing applications (Langlois et al., 2012; Andreadis and Lettenmaier, 2012; Brucker et al., 2011; Picard et al., 2009). These studies all examined the accuracy of a single snowpack model coupled with a single microwave emission model.

The purpose of this study is to inform future design of retrieval and assimilation systems where snowpack evolution models may be used to provide microstructural parameters for microwave emission models, by examining how particular snowpack and emission model choices lead to a variation in simulated brightness temperatures throughout the winter period, and evaluate how the simulated values compare to observations. The Jules Investigation Model (Essery et al., 2013) has been coupled with three widely used microwave emission models: The Dense Media Radiative Transfer Model - Multilayer (Picard et al., 2013), the Microwave Emission Model of Multi-Layer Snow (Wiesmann and Mätzler, 1999) and the Helsinki University of Technology Multilayer model (Lemmetyinen et al., 2010; Pulliainen et al., 1999). Snowpack microstructure metamorphism is represented here by three different options with differing complexity for grain diameter evolution (or equivalently the Specific Surface Area). These models are the grain growth models of SNTHERM (Jordan, 1991), SNICAR (Flanner and Zender, 2006) and MOSES (Essery et al., 2001). This allowed quantification of the seasonal variation in uncertainty in brightness temperature simulations from 1,323 coupled snowpack-emission model systems, as evaluated against ground based observations of brightness temperature.

The study approach, model descriptions and field measurements are given in section 2. Comparisons between simulations, and between simulations and observations are presented in section 3, and the implications for future approaches to the remote sensing of snow mass are discussed in section 4.

2   Models and methods

This study builds on the work of Essery et al. (2013), who incorporated many published snow model parameterisations (excluding microstructural representation) within a single model framework, the Jules Investigation Model (JIM), which is described in section 2.1. For this studyAs this earlier study did not incorporate snow microstructure changes, JIM was coupled with three microstructure evolution functions for this study, described in section 2.2 and three distinct snow emission models, detailed in section 2.3. Steps necessary to form the model ensemble, including assumptions about the representation of the soil are given in section 2.4. A description of the field site, driving and evaluation data for the simulations in this paper are presented in section 2.5.

2.1   Snow model parameterisation

Essery et al. (2013) developed the Jules Investigation Model (JIM), a system of 1701 snowpack evolution models to provide a systematic method and common framework to examine how the range of snowpack parameterisations used in land surface models impacts the simulation of snow parameters. Based on this work, a more computationally efficient version, the Factorial Snowpack Model , has been developed (Essery, 2015) , which that allows for 32 model configurations. JIM is based on an

Eulerian grid scheme (fixed layer structure), which requires mass redistribution between layers with precipitation events. An alternative approach is a Lagrangian grid scheme: a deforming layer structure that retains much of the same snow material throughout the season (e.g. Jordan, 1991; Brun et al., 1992; Lehning et al., 2002). For this paper, a subset of the original JIM members were selected as these were expected to influence the parameters important for microwave modelling. The subset includes variation in the representation of compaction, the density of newly-deposited snow, thermal conductivity and liquid water flow (snow hydrology). Table 1 summarizes the different approaches taken. Note that a variable fresh snow density (options 0 and 1) cannot be used if the snowpack has fixed density (compaction option 2), so there are only 63 model configurations in the model subset rather than 81. For all other snowpack parameterisations, option '1' from Essery et al. (2013) were used for albedo, surface exchange and snow fraction representations to form the JIM subset.

2.2 Microstructure evolution

JIM subset outputs were used to drive three microstructure models of differing complexity. SNTHERM (SNT) (Jordan, 1991) growth of snow grain diameter d is based on the rate of vapour transport through the snow (and therefore temperature gradient), which leads to the microstructure evolution function of dry snow in SNT as:

$$\frac{\partial d}{\partial t} = \frac{g1}{d} D_{eos} \left(\frac{1000}{P_a}\right) \left(\frac{T_s}{T_m}\right)^6 C_{kT_s} \left|\frac{\partial T_s}{\partial z}\right| \tag{1}$$

where $g1$ and D_{eos} are empirical constants, P_a is the atmospheric pressure, C_{kT_s} is the variation of saturation vapour pressure with snow temperature T_s, $T_m = 273.15K$ and $\frac{\partial T_s}{\partial z}$ is the temperature gradient. Grain growth under wet conditions is more rapid, with empirical constant $g2$ and is dependent on the liquid fractional volume, f_l by:

$$\frac{\partial d}{\partial t} = \frac{g2}{d} (f_l + 0.05) \qquad\qquad\qquad f_l < 0.09 \tag{2}$$

$$\frac{\partial d}{\partial t} = \frac{g2}{d} (0.14) \qquad\qquad\qquad f_l \geq 0.09 \tag{3}$$

SNICAR (SNI) microstructure evolution is a computationally efficient approximation to a model based on physics, and uses a look-up table for empirical parameters τ and κ, as described in Flanner and Zender (2006). These parameters are dependent on the snow density, temperature, and temperature gradient. The equation of microstructure evolution in SNI is based on snow specific surface area (SSA):

$$SSA(t) = SSA_0 \left(\frac{\tau}{t+\tau}\right)^{1/\kappa} \tag{4}$$

SSA per unit mass of ice (m^2 kg^{-1}) can then be converted to grain diameter with $D = 6/(\rho_i \, SSA)$ (Mätzler, 2002; Montpetit et al., 2012).

A third microstructure model, MOSES (MOS), parameterizes snow evolution as a function of grain radius r and snow age:

$$r(t+\Delta t) = \left[r(t)^2 + \frac{G_r}{\pi}\Delta t\right]^{1/2} - [r(t) - r_0]\frac{S_f \Delta t}{d_0} \tag{5}$$

where G_r is an empirical temperature-dependent grain area growth rate, S_f is the snowfall rate in time interval Δt and d_0 is a constant representing the mass of fresh snow required to reset the snow albedo to its maximum value.

Other microstructure parameterisations are available, namely the Crocus (Vionnet et al., 2012) and SNOWPACK (Lehning et al., 2002) microstructure evolution functions. It is not currently possible to couple these with the JIM model due to the Eulerian grid structure of JIM. Mass transfer between layers allows numerical averaging of concepts such as grain diameter and SSA, but not shape-dependent concepts such as dendricity and sphericity. Therefore the Crocus and SNOWPACK functions have not been included in this study.

2.3 Microwave emission models

The microwave models chosen for this application span a range of physical complexity in their representation of the snow. The Helsinki University of Technology (HUT) model (Lemmetyinen et al., 2010) is a semi-empirical model based on strong forward scattering assumptions, the Microwave Emission Model of Multi-Layer Snow (MEMLS) model (Wiesmann and Mätzler, 1999) is of intermediate complexity and contains the Improved Born Approximation (Mätzler, 1998), and the Dense Media Radiative Transfer Multilayer (DMRT-ML) model (Picard et al., 2013) is the most physically complex, being based on quasi-crystalline approximation with coherent potential (QCA-CP). Many other microwave emission models have been developed, such as Mie scattering approach of Boyarskii and Tikhonov (2000); Chang et al. (1976); Eom et al. (1983); strong fluctuation theory (Stogryn, 1986; Song and Zhang, 2007), distorted Born approximation (Tsang et al., 2000), the quasi-crystalline approximation (Grody, 2008), other QCA-CP models Rosenfeld and Grody (2000); Jin (1997) or the Numerical Method of Maxwell's equations in 3D (Xu et al., 2012). These references are not exhaustive but give an illustration of the range of models available. Here, we restrict the comparison to widely available multilayer models that span a range of complexity and whose computational efficiency is such that entire seasons can be simulated rapidly.

Of the models chosen, all are multiple layer and broadly require the same information i.e. they use layered information on snow temperature, density and layer thickness as input, but differ in their representation of the microstructure. They are all based on radiative transfer theory, which is governed by the following general equation:

$$\mu \frac{\partial \mathbf{T}_B\left(\theta_s, \phi_s, z\right)}{\partial z} = \kappa_a T(z) + \frac{1}{4\pi} \int\limits_{4\pi} \Psi(\theta_s, \phi_s; \theta_i, \phi_i) \cdot \mathbf{T}_B\left(\theta_i, \phi_i, z\right) d\Omega_i - \kappa_e \cdot \mathbf{T}_B\left(\theta_s, \phi_s, z\right) \tag{6}$$

where θ and ϕ are the ~~incidence~~ zenith and azimuth angles, $\mu = \cos\theta$, $\mathbf{T}_B$ is the brightness temperature vector, which we will assume here to consist of horizontally and vertically polarized brightness temperature only, κ_a is the absorption coefficient, κ_e is the extinction coefficient, which is a sum of the absorption coefficient and the scattering coefficient κ_s. The models differ in which two coefficients determine the third. In HUT, the derived coefficient is κ_s, whereas κ_e is derived in MEMLS and κ_a in DMRT-ML. Other differences between models include the representation of the phase function (single-stream model with separate up- and downwelling components in HUT, 6-stream in MEMLS and multiple streams in DMRT-ML), specification of the absorption coefficient and the numerical techniques applied to solve the radiative transfer equation (Lemmetyinen et al., 2010; Wiesmann and Mätzler, 1999; Picard et al., 2013; Mätzler and Wiesmann, 1999; Pan et al., 2015). Differences between

models are not restated here, but options chosen within each model leading to different model versions are stated in the following subsections.

2.3.1 DMRT-ML

DMRT-ML is based on a sticky hard spheres representation of the microstructure so that the scattering coefficient given by the quasi-crystalline approximation with coherent potential is given as:

$$\kappa_s = \frac{2}{9} k_0^4 a^3 f \left| \frac{\epsilon_s - \epsilon_b}{1 + \frac{\epsilon_s - \epsilon_b}{3E_{eff}(1-f)}} \frac{\epsilon_s - \epsilon_b}{1 + \frac{\epsilon_s - \epsilon_b}{3E_{eff}}(1-f)} \right|^2 \frac{(1-f)^4}{(1 + 2f - tf(1-f))^2} \tag{7}$$

where $k_0 = 2\pi/\lambda$ is the wave number, a is the radius of the spheres, f is the fractional volume of scatterers, ϵ_s is the permittivity of the scatterers, ϵ_b is the permittivity of the background and E_{eff} is the effective permittivity of the medium. t is related to the stickiness factor τ governing the potential of particles to coalesce. For non-sticky particles t=0 but for sticky particles, it is given by the largest ~~solution to~~ of the two solutions to the quadratic equation:

$$\frac{f}{12} t^2 - \left(\tau + \frac{f}{1-f} \right) t + \frac{1 + f/2}{(1-f)^2} = 0 \tag{8}$$

Whilst Löwe and Picard (2015) have shown that it may be possible to determine stickiness from micro-CT measurements of the snow, an appropriate value of stickiness is not known for the field observations used in this paper. Roy et al. (2013); Löwe and Picard (2015) that non-sticky representation in DMRT-ML is inappropriate. For this model ensemble, two DMRT-ML ~~varieties~~ configurations have been chosen to capture the range of brightness temperatures simulated: "DMRT ~~non-sticky~~less sticky" ($\tau =$ ~~10^6~~ 0.2) and "DMRT very sticky" ($\tau = 0.1$). These two values represent reasonable values used by others (e.g. Tsang et al., 2007; Shih et al., 1997).

2.3.2 MEMLS

Within MEMLS there are a suite of options for the calculation of the scattering coefficient. Two of the options within MEMLS were selected for this study to cover both empirical and theoretical approaches: "MEMLS empirical" and "MEMLS IBA". The empirical version of MEMLS used gives the scattering coefficient as:

$$\kappa_s = (9.2 p_{ec} - 1.23\rho + 0.54)^{2.5} (\nu/50)^{2.5} \tag{9}$$

where the correlation length p_{ec} is in mm, density ρ is in g cm^{-3} and frequency ν is in GHz. This is suitable for correlation lengths $0.05 < p_{ec} < 0.3$ mm, and density $0.1 < \rho < 0.4$ g cm^{-3}.

MEMLS IBA uses the Improved Born Approximation theory given in Mätzler (1998); Mätzler and Wiesmann (1999), where the scattering coefficient is given by the integral of the phase function for ~~an averaged~~ polarization angle χ

$$\kappa_s = \frac{1}{4\pi} \int\limits_{4\pi} f(1-f)(\epsilon_s - \underline{1}\epsilon_b)^2 K^2 I k_0^4 \sin^2 \chi \, d\Omega \tag{10}$$

Averaging of the polarization angle means that it is only dependent on the incidence angle θ, not the azimuth, so that:

$$\sin^2\chi = 1 - \frac{1}{2}\sin^2\theta$$

One further assumption applied to distinguish this MEMLS IBA configuration is that oblate grains are used rather than small spherical scatterers or thin spherical shells. This assumption governs the representation of the mean squared field ratio, K^2 as detailed in Mätzler and Wiesmann (1999). The microstructure length information is contained in I:

$$I = \frac{2p_{ec}}{\left(1 + 4\epsilon_{eff}k_0^2\sin^2(\theta/2)p_{ec}^2\right)^2} \tag{11}$$

It should be noted that the choice of oblate grains also affects the effective permittivity in I, represented by an empirical, density-dependent effective permittivity (Wiesmann and Mätzler, 1999, equations 45-47) for this case.

2.3.3 HUT

HUT has three options for the extinction coefficient. These are nominally suited to different grain size diameter (d_0) ranges, with some overlap between them. All three versions (termed "HUT H87", "HUT R04", "HUT K10") have been included in this version of the model ensemble. HUT H87 is based on the work of Hallikainen et al. (1987):

$$\kappa_e = 0.0018\nu^{2.8}d_0^{1.9} \tag{12}$$

This is nominally appropriate for frequency range ν=18-60GHz and $d_0 < 1.6$ mm.

The extinction coefficient in HUT H04, with a validity range of $1.3 < d_0 < 4$ mm was derived by Roy et al. (2004):

$$\kappa_e = 2\nu^{0.8}d_0^{1.2} \tag{13}$$

Kontu and Pulliainen (2010) gave the extinction coefficient for maritime snow, used here in the HUT K10 simulations as:

$$\kappa_e = 0.08\nu^{1.75}d_0^{1.8} \tag{14}$$

Scaling of the grain diameter by the relationship recommended in Kontu and Pulliainen (2010) has not been applied here as it was developed for snow microstructure observations rather than simulated snowpack microstructure.

2.4 Model framework

Interfacing of the various model inputs and outputs was enabled through the development of the ensemble framework, via a combination of shell script and Octave/Matlab code. Interfacing of the various model inputs and outputs was enabled through the development of the ensemble framework, via a combination of shell script and Octave/Matlab code. The DMRT-ML model was run from the shell script, which subsequently calls an Octave / Matlab script to run HUT and MEMLS. HUT and MEMLS run alternately in this framework as the soil parameters (common between DMRT-ML and HUT) are used to calculate soil reflectivity in HUT, which is then used as the lower boundary condition in MEMLS. Internal parallelization of the MATLAB

code of HUT-MEMLS means that a season-long simulation of one HUT-MEMLS combination with one grain scaling factor takes 9 minutes over eight cores. For the DMRT-ML FORTRAN code, external bash shell parallelization reduces execution time from 16 hours to circa 2 hours for one grain scale factor and two parameterisations of stickiness. Over 29 million individual brightness temperatures were simulated for this study.

5   For the purposes of this study, the effective sphere size in JIM, DMRT-ML and HUT are assumed to be identical i.e. $d_{hut} = 2 \times r_{dmrt} = d_{jim}$. This may not be a good assumption as the empirical extinction coefficient model used in HUT was based on observations of the maximum grain extent rather than effective diameter, which was almost impossible to measure at the time of the original work. The exponential correlation length in MEMLS (in mm) is calculated from the theoretical relationship to the effective grain diameter from JIM (in microns) as Montpetit et al. (2013); Mätzler (2002):

$$10 \quad p_{ec} = \frac{2}{3} \left(1 - \frac{\rho_s}{\rho_i} \right) \frac{d_{jim}}{1000} \qquad (15)$$

Figure 1 illustrates the flow of information in the model ensemble. Meteorological data are used to drive the 189 ~~flavours of JIM~~ configurations of JIM (3 microstructural models for each of the 63 snowpack parameterizations). The outputs from JIM are then reformatted for each of the three electromagnetic models. Table 2 gives a summary of the main differences in inputs between models. The electromagnetic model inputs are then used to drive the two DMRT-ML model versions (τ=0.1, τ

15 $= 10^6$), the two MEMLS model versions (empirical, IBA with oblate grains) and the three HUT versions (3 different extinction coefficient models). Meteorological and field data used to drive and evaluate the ensemble are described in the following section.

2.5 Data

Model runs for this study were performed for the Intensive Observation Area (IOA) of the Finnish Meteorological Institute

20 Arctic Research Centre (FMI-ARC). The site provides a wealth of forcing and evaluation data, including automated soil, snow and meteorological observations, ground-based microwave radiometry, and a programme of manual snow profile observations. Air temperature, solar radiation and precipitation observations from this site for the two seasons of simulations are shown in Figure 2. November rain events occurred in both years, and also in early December in 2011-2012. Layers with melt-freeze polycrystals and other melt forms were detected in snow observations during both seasons. Metadata and details on the me-

25 teorological instruments are given in Essery et al. (2016). Dual polarization microwave radiometers, including at frequencies of 18.7 and 36.5 GHz are situated on a 4m tower pointing inwards on the edge of a large clearing surrounded by a mainly pine forest. Further details about the IOA site are given in Lemmetyinen et al. (2016). Details on the manual snow profile observation programme are given by Leppänen et al. (2016).

Simulations were carried out for the winters of 2011-2012 and 2012-2013 as there ~~where~~ were 49 approximately bi-weekly

30 snow pit observations over these two years available for snowpack model evaluation. Snow samples from 31 of these pits were extracted and used to measure profiles of the snow specific surface area (SSA) with the IceCube instrument (Zuanon, 2013). A bulk grain diameter was calculated for the analysis from the SWE-weighted mean SSA, excluding layers without observations. For these two seasons, the real component of the soil permittivity measurements were available at 100MHz, measured at three

locations in the observation area with Delta-T Devices ML2x sensors, installed horizontally at a depth of approximately 2 cm beneath the organic surface layer. Mean measurements from the stable winter period (1st December to 31st March) were chosen as representative for the entire season, which resulted in values of soil permittivity of 4.4 in 2011-2012 and 4.6 in 2012-2014. For the JIM simulations in this paper, a scaling factor of 1.11 was applied to the 2011-2012 precipitation data, and a scaling factor of 1.06 was applied to the 2012-2013 data to match the measured snow accumulation on the ground better. These factors differ slightly from the values used in the 7-year consolidated dataset of Essery et al. (2016).

2.6   Simulation methodology

Choices in the snowpack evolution parameterisations made here lead to 189 unique JIM snowpack models. Modelled snowpack profiles of layer thickness, temperature, density and grain diameter were output daily at noon for this study. These were then applied to the seven microwave emission model combinations, resulting in 1,323 sets of brightness temperature simulations per day.

In order to illustrate and analyze the effects of assumptions regarding snowpack evolution and microwave scattering on simulated brightness temperatures over the course of winter season, the remainder of the paper will:

1. Present the range of brightness temperatures expected for any generic combination of snowpack and emission model.

2. Apply a range of scaling factors ($0.1 <= \Phi <= 5.0$) to simulated JIM snowpack diameters (equation) and calculate the degree of misfit between simulated and observed brightness temperatures using the following cost function (CF):

$$CF = \sum^{ndays} \sum^{\nu} \sum^{pol} \left(\frac{TB_{sim} - TB_{obs}}{2} \right)^2 \tag{16}$$

The cost function term is summed over the two polarizations (H and V pol) for the two frequencies (18.7 and 36.5 GHz) over the number of days (ndays) where observations and simulations are both available. Due to the observation schedule at the Sodankylä site, the noon 'observations' for comparison with the simulations were determined as the mean of the 10am and 2pm observations. If observations were missing from either or both of these times, the brightness temperature for that day was excluded from the cost function calculation. Optimal Φ were found from the minimisation of the cost function CF.

3. Isolate the effect of snowpack parameterizations on simulated brightness temperature by presenting simulation results grouped by parameterizations of densification, liquid water flow, initial snow density and thermal conductivity. This will determine which factors govern the spread in brightness temperature and are therefore important for the design of snow retrieval assimilation systems.

3   Results

Snow depth and snow water equivalent (SWE) simulated by the Jules Investigation Model is shown in Figure 3. There is a small difference between the automatic measurements and the manual field observations attributable to the spatial variability

of the snow and difference in measurement location. Ultrasonic snow depth measurements were on average 12mm deeper than the snow pit observations in 2011-2012 but were 29mm shallower than snow pit measurements in 2012-2013. SWE measured automatically by the gamma ray sensor had a mean value of 3.6mm SWE greater than the pit observations in 2011-2012 but 5.9mm less in 2012-2013.

5   Although the precipitation inputs were scaled due to known sensor undercatch problems, in 2011-2012 the SWE was under-estimated until the end of January, then overestimated until the melt period. Compared with snow pit observations, the SWE bias prior to 1st February was -14.2 mm. Between 1st February and 31st March, the SWE bias was ~~19.12~~ 19.1mm. From 1st April until the end of the season, the bias was ~~-24.51~~ -24.5 mm water equivalent. In 2012-2013, simulated SWE was overes-timated for most of the season, with a mean bias of ~~13.73~~ 13.7 mm compared with the snow pit observation. Simulated SWE

10 is relatively insensitive to the snow parameterisation in the accumulation period, but ~~3~~ three distinct model groups emerge in the melt period, which are due to the ~~3~~ three different representations of the liquid water flow. The snowpack model parame-terisations have a greater impact on the snow depth, which is to be expected as this is directly affected by the representation of densification and initial snow density.

    Figure 4 demonstrates the impact of the snow model parameterisations on snow grain diameter growth, as simulated with the

15 MOS, SNI and SNT microstructure evolution functions. Each microstructure model results in a spread of bulk grain diameter due to the 63 snowpack parameterisations, but in general the difference between microstructure models is greater than the difference due to snowpack parameterisations. The simulation range is greatest at the start and at the end of the season, when the snowpacks can be subject to the largest temperature gradients, or liquid-water dependent growth. In both years of simulation, the mid season bulk grain diameter is smallest with MOS, and largest with SNT. MOS and SNIs are similar in

20 magnitude, but SNT bulk grain diameters were approximately twice as large on average, with a mean ratio over the season of 1.9-2.2, as shown in Table 3. SNT bulk grain diameter was up to 3.2 times larger than MOS bulk grain diameter. Visual estimation of the snow grain diameter gave values that were always larger than all of the simulations. Measured SSA-derived bulk grain diameters generally lay in between the SNT simulations and the simulations with SNI and MOS. The mean absolute error and mean relative difference for these simulations are presented in Table 4. SNT had the lowest bias (0.12 mm) in 2011-

25 2012, whereas SNI had the lowest bias (-0.14 mm) in 2012-2013. Bulk grain diameter simulated by the microstructure models lead to a mean difference of between -53% and +45% relative to the observations.

    Simulation of mean and range of brightness temperature from the three emission models driven by all snowpack and mi-crostructure model combinations is shown in Figure 5. Note that excessively low brightness temperatures on 01 November 2011 were excluded from this figure as the snowpack for some JIM members was extremely thin with an unphysically high snow

30 density. In general, HUT with ~~3~~ three representations of extinction coefficient showed the smallest range of brightness temper-ature, whereas DMRT-ML (covering both ~~sticky and non-sticky spherical particle~~ very sticky and less sticky assumptions) had a much greater range, which was nearly as large as MEMLS (empirical representation and Improved Born Approximation with oblate grains). This is demonstrated by the ratio between the seasonal mean ranges of brightness temperature presented in Table 5, where the ranges compared with HUT had a ratio of greater than 1. MEMLS had a larger range than DMRT-ML, although

35 at 37GHz the difference was small. As illustrated in Figure 5, at 19GHz, the mean of DMRT-ML simulations were highest

and the mean of MEMLS simulations were generally lowest (with the exception of 19H in 2011-2012). At 37 GHz, horizontal polarization, HUT gives the highest mean brightness temperature in 2012-2013, whereas DMRT-ML gives the highest mean brightness temperature in 2011-2012 and at horizontal polarization (both years). MEMLS mean brightness temperatures are the lowest at 37GHz at both horizontal and vertical polarization in both years. All ranges exhibit a distinctive 'wedge' shape,

5 where the ranges generally increase throughout the season until the collapse of the range in the melt period.

Compared with the brightness temperature observations, no model gives a consistently better performance across both frequencies and both polarizations. This is illustrated in Table 6, where mean bias and root mean squared error (RMSE) for each season has been presented for each frequency and polarization combination. The lowest bias was less than ~~8K~~ 7K in magnitude, whereas the lowest RMSE for each frequency / polarisation was less than 13K. For both years, DMRT-ML gave the lowest bias

10 at 19H and MEMLS gave the lowest bias at 37GHz (V and H). At 19V DMRT-ML had the lowest bias in 2011-2012 whereas HUT had the lowest bias in 2012-2013. Figure 5 shows that the observed brightness temperature is generally within the range simulated by each of the three microwave emission models, with the exception of 19H in 2011-2012 (MEMLS and HUT) and 37H in 2012-2013 (HUT). End of season brightness temperature observations are not replicated in the simulations as the liquid water content of the snowpack model is currently decoupled from the electromagnetic snow model, so the simulations only

15 represent dry snow brightness temperature.

Table 7 indicates scaling factors that would need to be applied to the grain diameter in order to allow a particular microstructure evolution function to minimize the cost function given in equation 16 i.e. the best agreement with observed brightness temperature for all four frequency and polarization combinations. A scale factor of 1 suggests a perfect fit between snowpack microstructure and microwave microstructure. A scale factor of less than one indicates a snowpack grain diameter overesti-

20 mate, whereas a scale factor of greater than one is an underestimate. For SNT microstructure, a scale factor of less than one was required in 2011-2012 for all emission models with the exception of the ~~non-sticky~~ less sticky (τ=0.2) application of DMRT-ML. This indicates that the SNT microstructure resulted in grain diameters larger than that required by the emission models for that year. In 2012-2013 SNT microstructure required slight scaling to increase the grain diameter for HUT and for less sticky DMRT-ML, but downscaling for very sticky hard spheres in DMRT-ML and for MEMLS. With the exception of the application

25 to empirical MEMLS in 2011-2012, the SNI and MOS grain diameters were too small and required scaling upwards. A cost function minimum was achieved for empirical MEMLS driven by MOS microstructure with no scaling whatsoever in 2011-2012. ~~In contrast, large scaling factors were required with the assumption of non-sticky particles in DMRT-ML. For MOS and SNI in 2012-2013 the cost function was minimised by a scale factor of 5 (or greater), but even the grain diameters simulated with SNT were too small.~~ The pattern is consistent between years, with the greatest interannual difference in scale factor for

30 HUT~~and for DMRT-ML non-sticky microstructure from MOS~~.

Once the microstructure differences have been isolated through application of the optimal scale factor, ~~except for 37H in 2012-2013, the lowest RMSE in brightness temperature is reduced,~~ as shown in Table 8~~.~~, DMRT-ML bias and RMSE improved~~with the application of the optimal scale factor~~, with the exception of the small increase in 19V bias in 2011-2012. For MEMLS, improvements in bias and RMSE at ~~one frequency / polarisation~~ the lower frequency were at the expense of

35 ~~another frequency / polarisation~~ the higher frequency in both years. ~~This was also the case~~ The opposite occurred for HUT

in 2012-2013, whereas in 2011-2012 the bias and RMSE decreased at all frequencies and polarisations apart from a marginal (<0.04K) increase in RMSE at 19V.

Differences in brightness temperature also exist in the simulations due to the snowpack parameterisation (i.e. 63 JIM combinations). Empirical MEMLS with MOS microstructure in the 2011-2012 season was chosen as a test case to illustrate the

5 effects of snowpack parameterisation on the brightness temperature, because of the equivalence of snowpack and emission model microstructure (no scaling required). This subset of 63 simulations for 37H brightness temperature in 2011-2012 is shown in Figure 6. There is a seasonal dependence in the range, with model divergence from mid-January onwards. February 1st and May 1st were chosen for cluster analysis to determine which parameterisations caused the split in simulations, as shown in Figure 7.

10 Clear groupings of simulations in Figure 7, upper left, indicates that the snowpack densification parameterisation has a distinguishable effect on the simulation of brightness temperature. A physical representation of densification (parameterization = "0") gave the lowest brightness temperatures on the 1st February, but the highest by 1st May. In contrast, where no compaction is simulated i.e. snow density is constant throughout the season (parameterization = "2"), the opposite is true. An empirical representation of densification (parameterization = "1") results in brightness temperatures generally between those of the

15 physical, and of no densification. Thermal conductivity has no effect on the simulation of brightness temperature, whereas subtle differences are attributable to the fresh snow density value, and to the representation of snow hydrology. There is no discernible difference between fresh snow density parameterisation schemes "0" and "1", whereas "2" gives a different set of brightess temperatures. Snow hydrology has very little effect in the early season, but can lead to differences in the melt period. Overall, the snowpack parameterisations with MOSES microstructure and empirical MEMLS lead to a mean difference in

20 the 36.5GHz brightness temperature of 11K at H-pol and 18K at V-pol. The maximum difference in 36.5GHz brightness temperature was 33K at H-pol and ~~maximum difference of 33K in brightness temperature simulation at 37H~~ 54K at V-pol for the 2011-2012 season. The maximum difference between H and V polarization for all unscaled microstructure-electromagnetic model combinations is demonstrated in Table 9. Large differences in the maximum brightness temperature difference as a result of the 63 snowpack configurations occurred for the SNT microstructure. Except for DMRT-ML less sticky and HUT with MOS

25 or SNI microstructure, the V-pol difference is greater than the H-pol difference.

4   Discussion

The biggest difference to obtaining accurate simulations would be made by improving the microstructure evolution models within snowpack models because the optimal scale factors are generally larger between microstructure models than between emission models. SNTHERM grains tend to be too large for the emission models and generally require scaling down to smaller

30 values. SNICAR grains are in the mid-range and require a small amount of scaling, generally upwards to larger grains. MOSES grains are the smallest and generally require larger scale factors than SNICAR. These patterns are consistent, regardless of the electromagnetic radiative transfer model used. Differences between microstructure evolution models are so large because they were developed in models with different purposes. MOSES is a large scale land surface model, requiring snow grain size

for albedo calculations (Essery et al., 2001). SNICAR is a snow albedo model (Flanner and Zender, 2006). ~~SNICAR was also developed with an albedo focus.~~ SNTHERM, on the other hand, was developed to predict surface temperature and uses grain diameter in the simulation of liquid water flow as well as albedo (Jordan, 1991). SNICAR and MOSES grain sizes are closer to the SSA-derived grain diameter as a result. SNTHERM simulates a grain size that is closer in concept to the visual estimates of grain diameter than the other two models. The large spread when coupling snowpack evolution and microwave models, due to the differences in the modelling of snow microstructure is consistent with the wide range of studies that have investigated how to link snowpack observations of microstructure to the microstructure parameter required in electromagnetic models (e.g Kendra et al., 1998; Du et al., 2005; Tedesco et al., 2006; Liang et al., 2008; Durand et al., 2008; Brucker et al., 2011; Xu et al., 2012; Montpetit et al., 2013; Roy et al., 2013; Rutter et al., 2014; Picard et al., 2014).

Nevertheless there are differences between microwave emission models for a particular microstructure evolution model, and even differences within the same family of emission models. 'Improvement' in the microstructure for a particular model combination may lead to less accurate simulations at some frequencies and polarisations, which highlights that there is more to understand. ~~DMRT-ML is the only emission model of the three that has a polarisation-dependent phase function , which could partially explain why the reductions in~~ In part, this may be due to the methodology of this study as the cost function is calculated per microstructure-electromagnetic model configuration, yet the bias and RMSE are presented for each electromagnetic model family. An individual contribution can influence the group in a non-intuitive way.

Here, the lowest bias and RMSE ~~were nearly universal for DMRT-ML with application of a scale factor, but were frequency / polarization dependent for MEMLS or HUT.~~ for unscaled microstructure simulations were -6.9 to +6.9K and 4.2 to 12.2K respectively, but depended on microwave model, frequency and polarization. In an attempt to put these results into context, there are a number of studies that have quantified brightness temperature simulation errors for these models. These fall into different categories, depending on sensor characteristics, the source of the evaluation data (ground-based, airborne, satellite) and presence of ice lenses (Derksen et al., 2012), the treatment of the snow microstructure (Picard et al., 2014), snow type, observation angle and the specific electromagnetic model (Tedesco and Kim, 2006) and the underlying substrate (Lemmetyinen et al., 2009; Derksen et al., 2014). Examples of unscaled field observations of microstructure compared with ground-based observations include the HUT simulations of Derksen et al. (2012), who found an RMSE of 10-34K and Rutter et al. (2014) who found a bias of 34-68K that was reduced to < 0.6K upon application of grain scale factors of 2.6-5.3. Scaling, or best-fit relationships were used by Durand et al. (2008) (mean absolute error 3.1K at V-pol and 9.3K at H-pol), Montpetit et al. (2013) (RMSE 8-20K), Brucker et al. (2011) (RMSE 1.5K), Picard et al. (2014) (RMSE 1-11K) and Roy et al. (2013) (RMSE 12-16K). However, in some cases the frequency-dependent results have been combined and in others kept separate.

For DMRT-ML, consideration of the stickiness is imperative, ~~with the choice of microstructure evolution model of secondary importance. Only two extremes~~ . Two constant values were considered here: extremely cohesive, or ~~individual particles. The true value is likely to lie between these.~~ less sticky particles. Löwe and Picard (2015) have made progress in understanding ~~how to retrieve~~ stickiness from micro-CT data ~~, but more needs to be done in order to relate stickiness to other snow properties that are measurable in the field.~~ There are theoretical limits, based on snow density (Löwe and Picard, 2015, eqn. 35-36) but

in general stickiness is independent of diameter and of density, nor should a constant value be used, as was done here. Further research is needed in this regard.

For the HUT radiative transfer model, the optimum combinations of snowpack and microwave model are dependent on both models, and therefore the end application. The SNT microstructure is most closely matched to the microstructure of the Hallikainen et al. (1987) extinction model. Both were developed with a similar concept of microstructure. With MOS or SNI microstructure, Roy et al. (2004) would be most appropriate. Kontu and Pulliainen (2010) is more broadly applicable as the scale factor always lies between R04 and H87, regardless of the microstructure model. Therefore K10 may be better choice if a range of microstructure models are considered in a data assimilation retrieval scheme but with only one observation operator.

In the case of MEMLS, there are some differences between the empirical model, and IBA, but the microstructure model really matters. IBA is a more appropriate model for the larger SNT grains and endorses the recommendation of Mätzler and Wiesmann (1999) for IBA in the simulation of larger grains. The microstructural concept of MOS matches the microstructure of ~~the empirical~~ empirical MEMLS very well, with no scaling required in 2011-2012, although SNI is equally appropriate in 2012-2013.

There is little variation between years for the DMRT-ML (sticky) and MEMLS models, and a consistent pattern for HUT. Other studies have investigated the microstructural link between snowpack and microwave models. Wiesmann et al. (2000) found that the scale factor between exponential correlation length in MEMLS and grain diameter in SNTHERM for the Weissfluhjoch site in Davos, Switzerland, was 0.16. Applying equation 15 for snow of density 250 kg m^{-3}, the scale factor to relate the grain diameter of SNTHERM to the exponential length in MEMLS for the Sodankylä site would be 0.24 for IBA and 0.15 for empirical MEMLS for the 2011-2012 dataset. This is entirely consistent with the Wiesmann et al. (2000) study, in spite of the different locations and snowpack conditions.

Wiesmann et al. (2000) also reported a relationship for Crocus simulations, as did Brucker et al. (2011). At this stage it is not possible to make comparisons of this work with those studies as the Crocus evolution model has not been included in this study, due to the difficulty of applying these models to the Eulerian frame snowpack model scheme used here. These two studies are, however, consistent with each other. Wiesmann et al. (2000) found a snow type dependent scale factor of 0.3-0.4 between MEMLS correlation length and Crocus grain diameter, whereas the range in Brucker et al. (2011) was ~~0.25-0.4~~ 0.4-0.25 for snow density between 100 and 400 kg m^{-3}. The scaling factor between the SNOWPACK-derived correlation length and the correlation length of MEMLS was found to be 0.1 (Langlois et al., 2012) but again, a comparison with this work is not possible as the SNOWPACK grain evolution model has similar requirements to the Crocus microstructure model as they have a common origin.

When isolating the spread in brightness temperature due to snowpack parameterisations, this spread is largely due to the snowpack model representation of the densification process, with a variable impact throughout the season. After the microstructure model, snow compaction must be considered carefully in the design of a coupled snowpack-microwave model~~as it can account for differences in brightness temperature of the order of 30K~~. Liquid water flow representation in the snowpack model may become important in the melt period, particularly for a snowpack with mid-winter melt periods or if the snowpack model is used to provide information on SWE during melt when microwave observations cannot. If fresh snow is assumed to have a

constant density in a retrieval or assimilation system then that value will have an impact but is less important than compaction. Thermal conductivity has no discernable impact on the brightness temperature simulations so the choice of its representation is largely irrelevant for snow mass retrieval and assimilation systems.

Although ~~the~~ empirical MEMLS driven by MOSES was chosen as an example to demonstrate the impact of parameterisations, this was purely because of the apparent consistency between the MOSES grain diameter converted to exponential correlation length and MEMLS simulations for 2011-2012 at this site. This is not a general endorsement of empirical models, as those based on physics are expected to be more universally applicable, but the specific application of these models will dictate the balance of accuracy versus simplicity. Extending the analysis beyond this example, snow parameterizations affect other unscaled model combinations to varying degrees. Microstructure scaling factors in Table 7 can be used as a proxy for the degree of scattering in the unscaled simulations. A higher scale factor acts to increase the simulated scattering, so for a scale factor < 1, too much scattering occurs in the unscaled simulations. Snowpack parameterisations have a greater impact for a higher degree of scattering, larger at V-pol than H-pol. This is because scattering is already greater at H-pol so the spread in H-pol simulations as a result of snowpack parameterisations is suppressed by the existing level of scattering. The converse applies for high scaling factors (e.g. MOS with less sticky DMRT-ML).

Although the differences in scale factors between microstructure models are larger than the differences in scale factors between microwave models, this does not negate the need for developments in the microwave models. This is highlighted by the treatment of field observations of specific surface area to derive optical diameter as even these require some form of scaling (e.g. Montpetit et al., 2013; Picard et al., 2014; Rutter et al., 2014). Use of scale factors can improve brightness temperature accuracy at some frequency and polarisations, but may decrease the accuracy at others. The necessity of scale factors indicate the need for a deeper understanding in the role of microstructure in the microwave models. Much of this is discussed from a theoretical perspective by Löwe and Picard (2015). With a sticky hard sphere model of the microstructure, even if the stickiness is known, Löwe and Picard (2015) showed that a scale factor to relate the measured optical diameter to microwave diameter depends on the type of metamorphism the snow has been subjected to. Indeed, here, constant scale factors have been applied with no attempt to assess how these may change over the season. Nor do they account for the anisotropic nature of the snow, which adds to the complexity both in the modelling of the snowpack (Löwe et al., 2013) and in microwave scattering ~~(?)~~(Leinss et al., 2016). Some of the fundamental questions on how to relate snowpack and microwave microstructure may be addressed with a better microstructure descriptor of the snowpack rather than a single length scale, and would benefit from easy interchangability between different microwave models, and different snowpack evolution models. Ultimately a consistent microstructural treatment will be needed in both snowpack evolution and microwave models.

5 Conclusions

Future snow mass and depth retrievals systems may rely on snowpack models to provide snow microstructural parameters. To improve accuracy in seasonal simulations of brightness temperature, the largest gains will be achieved by improving the microstructural representation within snowpack models, followed by improvements in the emission models to use accurate

microstructural information and reduce bias and RMSE at all frequencies and polarizations simultaneously. For the design of retrieval systems with current capabilities, particular model combinations may be more suitable than others, and careful consideration must be given to snow compaction processes. Snow process representation becomes increasingly important as the snowpack scatters more. The future lies in a better and consistent treatment of snow microstructure in both snowpack and emission model developments.

5

10

References

Andreadis, K. M. and Lettenmaier, D. P.: Assimilating remotely sensed snow observations into a macroscale hydrology model, Advances in Water Resources, 29, 872–886, doi:10.1016/j.advwatres.2005.08.004, http://www.sciencedirect.com/science/article/pii/S0309170805002058, 2006.

5 Andreadis, K. M. and Lettenmaier, D. P.: Implications of Representing Snowpack Stratigraphy for the Assimilation of Passive Microwave Satellite Observations, Journal of Hydrometeorology, 13, 1493–1506, doi:10.1175/JHM-D-11-056.1, http://journals.ametsoc.org/doi/abs/10.1175/JHM-D-11-056.1, 2012.

Boyarskii, D. and Tikhonov, V.: The Influence of Stratigraphy On Microwave Radiation From Natural Snow Cover, Journal of Electromagnetic Waves and Applications, 14, 1265–1285, doi:10.1163/156939300X01201, http://dx.doi.org/10.1163/156939300X01201, 2000.

10 Brown, R., Derksen, C., and Wang, L.: A multi-data set analysis of variability and change in Arctic spring snow cover extent, 1967–2008, Journal of Geophysical Research: Atmospheres, 115, n/a–n/a, doi:10.1029/2010JD013975, http://onlinelibrary.wiley.com/doi/10.1029/2010JD013975/abstract, 2010.

Brown, R. D. and Robinson, D. A.: Northern Hemisphere spring snow cover variability and change over 1922–2010 including an assessment of uncertainty, The Cryosphere, 5, 219–229, doi:10.5194/tc-5-219-2011, http://www.the-cryosphere.net/5/219/2011/, 2011.

15 Brucker, L., Royer, A., Picard, G., Langlois, A., and Fily, M.: Hourly simulations of the microwave brightness temperature of seasonal snow in Quebec, Canada, using a coupled snow evolution–emission model, Remote Sensing of Environment, 115, 1966–1977, doi:10.1016/j.rse.2011.03.019, http://www.sciencedirect.com/science/article/pii/S0034425711000964, 2011.

Brun, E., David, P., Sudul, M., and Brunot, G.: A numerical model to simulate snow-cover stratigraphy for operational avalanche forecasting, Journal of Glaciology, 38, 13–22, 1992.

20 Chang, A. T. C., Gloersen, P., Schmugge, T., Wilheit, T. T., and Zwally, H. J.: Microwave emission from snow and glacier ice, Journal of Glaciology, 16, 23–39, http://adsabs.harvard.edu/abs/1976JGlac..16...23C, 1976.

Chang, A. T. C., Foster, J. L., and Hall, D. K.: Nimbus-7 SMMR Derived Global Snow Cover Parameters, Annals of Glaciology, 9, 39–44, 1987.

Clifford, D.: Global estimates of snow water equivalent from passive microwave instruments: history, challenges and future developments,
25 International Journal of Remote Sensing, 31, 3707–3726, doi:10.1080/01431161.2010.483482, http://www.tandfonline.com/doi/abs/10.1080/01431161.2010.483482, 2010.

Davenport, I., Sandells, M., and Gurney, R.: The effects of variation in snow properties on passive microwave snow mass estimation, Remote Sensing of Environment, 118, 168–175, doi:10.1016/j.rse.2011.11.014, http://www.sciencedirect.com/science/article/pii/S0034425711004093, 2012.

30 Derksen, C. and Brown, R.: Spring snow cover extent reductions in the 2008–2012 period exceeding climate model projections, Geophysical Research Letters, 39, n/a–n/a, doi:10.1029/2012GL053387, http://onlinelibrary.wiley.com/doi/10.1029/2012GL053387/abstract, 2012.

Derksen, C., Toose, P., Lemmetyinen, J., Pulliainen, J., Langlois, A., Rutter, N., and Fuller, M.: Evaluation of passive microwave brightness temperature simulations and snow water equivalent retrievals through a winter season, Remote Sensing of Environment, 117, 236–248, doi:10.1016/j.rse.2011.09.021, http://www.sciencedirect.com/science/article/pii/S0034425711003488, 2012.

35 Derksen, C., Lemmetyinen, J., Toose, P., Silis, A., Pulliainen, J., and Sturm, M.: Physical properties of Arctic versus subarctic snow: Implications for high latitude passive microwave snow water equivalent retrievals, Journal of Geophysical Research: Atmospheres, p. 2013JD021264, doi:10.1002/2013JD021264, http://onlinelibrary.wiley.com/doi/10.1002/2013JD021264/abstract, 2014.

Du, J., Shi, J., and Wu, S.: A comparison of a second-order snow model with field observations., in: INTERNATIONAL GEOSCIENCE AND REMOTE SENSING SYMPOSIUM, vol. 4, p. 2649, 2005.

Durand, M., Kim, E., and Margulis, S.: Quantifying Uncertainty in Modeling Snow Microwave Radiance for a Mountain Snowpack at the Point-Scale, Including Stratigraphic Effects, IEEE Transactions on Geoscience and Remote Sensing, 46, 1753–1767, doi:10.1109/TGRS.2008.916221, 2008.

Eom, H. J., Lee, K. K., and Fung, A. K.: Microwave emission from an irregular snow layer, Remote Sensing of Environment, 13, 423–437, doi:10.1016/0034-4257(83)90011-1, http://www.sciencedirect.com/science/article/pii/0034425783900111, 1983.

Essery, R.: A factorial snowpack model (FSM 1.0), Geosci. Model Dev., 8, 3867–3876, doi:10.5194/gmd-8-3867-2015, http://www.geosci-model-dev.net/8/3867/2015/, 2015.

Essery, R., Best, M., and Cox, P.: MOSES 2.2 technical documentation, Tech. rep., Hadley Centre Technical Note, 2001.

Essery, R., Morin, S., Lejeune, Y., and B Ménard, C.: A comparison of 1701 snow models using observations from an alpine site, Advances in Water Resources, 55, 131–148, 2013.

Essery, R., Kontu, A., Lemmetyinen, J., Dumont, M., and Ménard, C. B.: A 7-year dataset for driving and evaluating snow models at an arctic site (Sodankylä, Finland), Geoscientific Instrumentation, Methods and Data Systems Discussions, doi:doi:10.5194/gi-2015-35, 2016.

Flanner, M. G. and Zender, C. S.: Linking snowpack microphysics and albedo evolution, Journal of Geophysical Research: Atmospheres, 111, n/a–n/a, doi:10.1029/2005JD006834, http://onlinelibrary.wiley.com/doi/10.1029/2005JD006834/abstract, 2006.

Grody, N.: Relationship between snow parameters and microwave satellite measurements: Theory compared with Advanced Microwave Sounding Unit observations from 23 to 150 GHz, Journal of Geophysical Research: Atmospheres, 113, n/a–n/a, doi:10.1029/2007JD009685, http://onlinelibrary.wiley.com/doi/10.1029/2007JD009685/abstract, 2008.

Hallikainen, M., Ulaby, F., and Van Deventer, T.: Extinction Behavior of Dry Snow in the 18-to 90-GHz Range, IEEE Transactions on Geoscience and Remote Sensing, GE-25, 737–745, doi:10.1109/TGRS.1987.289743, 1987.

Hancock, S., Baxter, R., Evans, J., and Huntley, B.: Evaluating global snow water equivalent products for testing land surface models, Remote Sensing of Environment, 128, 107–117, 2013.

Jin, Y.-Q.: Simulation of a multi-layer model of dense scatterers for anomalous scattering signatures from SSM/I snow data, International Journal of Remote Sensing, 18, 2531–2538, doi:10.1080/014311697217459, http://www.tandfonline.com/doi/abs/10.1080/014311697217459, 1997.

Jordan, R.: A One-Dimensional Temperature Model for a Snow Cover: Technical Documentation for SNTHERM.89., Tech. rep., 1991.

Kelly, R.: The AMSR-E Snow Depth Algorithm: Description and Initial Results, Journal of The Remote Sensing Society of Japan, 29, 307–317, doi:10.11440/rssj.29.307, 2009.

Kendra, J., Sarabandi, K., and Ulaby, F.: Radar measurements of snow: experiment and analysis, IEEE Transactions on Geoscience and Remote Sensing, 36, 864–879, doi:10.1109/36.673679, 1998.

Kontu, A. and Pulliainen, J.: Simulation of Spaceborne Microwave Radiometer Measurements of Snow Cover Using In Situ Data and Brightness Temperature Modeling, IEEE Transactions on Geoscience and Remote Sensing, 48, 1031–1044, doi:10.1109/TGRS.2009.2030499, 2010.

Langlois, A., Royer, A., Derksen, C., Montpetit, B., Dupont, F., and Goïta, K.: Coupling the snow thermodynamic model SNOWPACK with the microwave emission model of layered snowpacks for subarctic and arctic snow water equivalent retrievals, Water Resources Research, 48, n/a–n/a, doi:10.1029/2012WR012133, http://onlinelibrary.wiley.com/doi/10.1029/2012WR012133/abstract, 2012.

Lehning, M., Bartelt, P., Brown, B., Fierz, C., and Satyawali, P.: A physical SNOWPACK model for the Swiss avalanche warning: Part II. Snow microstructure, Cold Regions Science and Technology, 35, 147–167, doi:10.1016/S0165-232X(02)00073-3, http://www.sciencedirect.com/science/article/pii/S0165232X02000733, 2002.

Leinss, S., Löwe, H., Proksch, M., Lemmetyinen, J., Wiesmann, A., and Hajnsek, I.: Anisotropy of seasonal snow measured by polarimetric phase differences in radar time series, The Cryosphere, 10, 1771–1797, 2016.

Lemmetyinen, J., Derksen, C., Pulliainen, J., Strapp, W., Toose, P., Walker, A., Tauriainen, S., Pihlflyckt, J., Kärnä, J., and Hallikainen, M.: A Comparison of Airborne Microwave Brightness Temperatures and Snowpack Properties Across the Boreal Forests of Finland and Western Canada, IEEE Transactions on Geoscience and Remote Sensing, 47, 965–978, doi:10.1109/TGRS.2008.2006358, 2009.

Lemmetyinen, J., Pulliainen, J., Rees, A., Kontu, A., Qiu, Y., and Derksen, C.: Multiple-Layer Adaptation of HUT Snow Emission Model: Comparison With Experimental Data, IEEE Transactions on Geoscience and Remote Sensing, 48, 2781–2794, doi:10.1109/TGRS.2010.2041357, 2010.

Lemmetyinen, J., Derksen, C., Toose, P., Proksch, M., Pulliainen, J., Kontu, A., Rautiainen, K., Seppänen, J., and Hallikainen, M.: Simulating seasonally and spatially varying snow cover brightness temperature using HUT snow emission model and retrieval of a microwave effective grain size, Remote Sensing of Environment, 156, 71–95, doi:10.1016/j.rse.2014.09.016, http://www.sciencedirect.com/science/article/pii/S0034425714003605, 2015.

Lemmetyinen, J., Kontu, A., Pulliainen, J., Vehviläinen, J., Rautiainen, K., Wiesmann, A., Mätzler, C., Werner, C., Rott, H., Nagler, T., Schneebeli, M., Proksch, M., Schüttemeyer, D., Kern, M., and Davidson, M.: Nordic Snow Radar Experiment, Geoscientific Instrumentation, Methods and Data Systems Discussions, 2016, 1–23, doi:10.5194/gi-2015-29, http://www.geosci-instrum-method-data-syst-discuss.net/gi-2015-29/, 2016.

Leppänen, L., Kontu, A., Hannula, H.-R., Sjöblom, H., and Pulliainen, J.: Sodankylä manual snow survey program, Geoscientific Instrumentation, Methods and Data Systems, 5, 163–179, doi:10.5194/gi-5-163-2016, http://www.geosci-instrum-method-data-syst.net/5/163/2016/, 2016.

Liang, D., Xu, X., Tsang, L., Andreadis, K., and Josberger, E.: The Effects of Layers in Dry Snow on Its Passive Microwave Emissions Using Dense Media Radiative Transfer Theory Based on the Quasicrystalline Approximation (QCA/DMRT), IEEE Transactions on Geoscience and Remote Sensing, 46, 3663–3671, doi:10.1109/TGRS.2008.922143, 2008.

Löwe, H. and Picard, G.: Microwave scattering coefficient of snow in MEMLS and DMRT-ML revisited: The relevance of sticky hard spheres and tomography-based estimates of stickiness, The Cryosphere, 2015.

Löwe, H., Riche, F., and Schneebeli, M.: A general treatment of snow microstructure exemplified by an improved relation for thermal conductivity, The Cryosphere, 7, 1473–1480, doi:10.5194/tc-7-1473-2013, http://www.the-cryosphere.net/7/1473/2013/, 2013.

Mätzler, C.: Improved Born approximation for scattering of radiation in a granular medium, Journal of Applied Physics, 83, 6111–6117, doi:10.1063/1.367496, http://scitation.aip.org/content/aip/journal/jap/83/11/10.1063/1.367496, 1998.

Mätzler, C.: Relation between grain-size and correlation length of snow, Journal of Glaciology, 48, 461–466, doi:10.3189/172756502781831287, 2002.

Mätzler, C. and Wiesmann, A.: Extension of the Microwave Emission Model of Layered Snowpacks to Coarse-Grained Snow, Remote Sensing of Environment, 70, 317–325, doi:10.1016/S0034-4257(99)00047-4, http://www.sciencedirect.com/science/article/pii/S0034425799000474, 1999.

Montpetit, B., Royer, A., Langlois, A., Cliche, P., Roy, A., Champollion, N., Picard, G., Domine, F., and Obbard, R.: New shortwave infrared albedo measurements for snow specific surface area retrieval, Journal of Glaciology, 58, 941–952, doi:10.3189/2012JoG11J248, 2012.

Montpetit, B., Royer, A., Roy, A., Langlois, A., and Derksen, C.: Snow Microwave Emission Modeling of Ice Lenses Within a Snowpack Using the Microwave Emission Model for Layered Snowpacks, IEEE Transactions on Geoscience and Remote Sensing, 51, 4705–4717, doi:10.1109/TGRS.2013.2250509, 2013.

Pan, J., Durand, M., Sandells, M., Lemmetyinen, J., Kim, E., Pulliainen, J., Kontu, A., and Derksen, C.: Differences Between the HUT Snow Emission Model and MEMLS and Their Effects on Brightness Temperature Simulation, IEEE Transactions on Geoscience and Remote Sensing, PP, 1–19, doi:10.1109/TGRS.2015.2493505, 2015.

Picard, G., Brucker, L., Fily, M., Gallée, H., and Krinner, G.: Modeling time series of microwave brightness temperature in Antarctica, Journal of Glaciology, 55, 537–551, doi:10.3189/002214309788816678, 2009.

Picard, G., Brucker, L., Roy, A., Dupont, F., Fily, M., Royer, A., and Harlow, C.: Simulation of the microwave emission of multi-layered snowpacks using the Dense Media Radiative transfer theory: the DMRT-ML model, Geosci. Model Dev., 6, 1061–1078, doi:10.5194/gmd-6-1061-2013, http://www.geosci-model-dev.net/6/1061/2013/, 2013.

Picard, G., Royer, A., Arnaud, L., and Fily, M.: Influence of meter-scale wind-formed features on the variability of the microwave brightness temperature around Dome C in Antarctica, The Cryosphere, 8, 1105–1119, doi:10.5194/tc-8-1105-2014, http://www.the-cryosphere.net/8/1105/2014/, 2014.

Pulliainen, J., Grandell, J., and Hallikainen, M.: HUT snow emission model and its applicability to snow water equivalent retrieval, IEEE Transactions on Geoscience and Remote Sensing, 37, 1378–1390, doi:10.1109/36.763302, 1999.

Rosenfeld, S. and Grody, N.: Anomalous microwave spectra of snow cover observed from Special Sensor Microwave/Imager measurements, Journal of Geophysical Research: Atmospheres, 105, 14 913–14 925, doi:10.1029/1999JD900486, http://onlinelibrary.wiley.com/doi/10.1029/1999JD900486/abstract, 2000.

Roy, A., Picard, G., Royer, A., Montpetit, B., Dupont, F., Langlois, A., Derksen, C., and Champollion, N.: Brightness Temperature Simulations of the Canadian Seasonal Snowpack Driven by Measurements of the Snow Specific Surface Area, IEEE Transactions on Geoscience and Remote Sensing, 51, 4692–4704, doi:10.1109/TGRS.2012.2235842, 2013.

Roy, V., Goita, K., Royer, R., Walker, A., and Goodison, B.: Snow water equivalent retrieval in a Canadian boreal environment from microwave measurements using the HUT snow emission model, IEEE Transactions on Geoscience and Remote Sensing, 42, 1850–1859, doi:10.1109/TGRS.2004.832245, 2004.

Rutter, N., Essery, R., Pomeroy, J., Altimir, N., Andreadis, K., Baker, I., Barr, A., Bartlett, P., Boone, A., Deng, H., Douville, H., Dutra, E., Elder, K., Ellis, C., Feng, X., Gelfan, A., Goodbody, A., Gusev, Y., Gustafsson, D., Hellström, R., Hirabayashi, Y., Hirota, T., Jonas, T., Koren, V., Kuragina, A., Lettenmaier, D., Li, W.-P., Luce, C., Martin, E., Nasonova, O., Pumpanen, J., Pyles, R. D., Samuelsson, P., Sandells, M., Schädler, G., Shmakin, A., Smirnova, T. G., Stähli, M., Stöckli, R., Strasser, U., Su, H., Suzuki, K., Takata, K., Tanaka, K., Thompson, E., Vesala, T., Viterbo, P., Wiltshire, A., Xia, K., Xue, Y., and Yamazaki, T.: Evaluation of forest snow processes models (SnowMIP2), Journal of Geophysical Research: Atmospheres, 114, n/a–n/a, doi:10.1029/2008JD011063, http://onlinelibrary.wiley.com/doi/10.1029/2008JD011063/abstract, 2009.

Rutter, N., Sandells, M., Derksen, C., Toose, P., Royer, A., Montpetit, B., Langlois, A., Lemmetyinen, J., and Pulliainen, J.: Snow stratigraphic heterogeneity within ground-based passive microwave radiometer footprints: Implications for emission modeling, Journal of Geophysical Research: Earth Surface, pp. n/a–n/a, doi:10.1002/2013JF003017, http://onlinelibrary.wiley.com/doi/10.1002/2013JF003017/abstract, 2014.

Shih, S.-E., Ding, K.-H., Kong, J., Yang, Y., Davis, R., Hardy, J., and Jordan, R.: Modeling of millimeter wave backscatter of time-varying snow cover, Journal of Electromagnetic Waves and Applications, 11, 1289–1298, 1997.

Song, K. and Zhang, Y.: Snow-cover environmental monitoring and assessment in Northeast China using passive microwave emission models, Environmental Monitoring and Assessment, 140, 223–229, doi:10.1007/s10661-007-9862-4, http://link.springer.com/article/10.1007/s10661-007-9862-4, 2007.

Stogryn, A.: A Study of the Microwave Brightness Temperature of Snow from the Point of View of Strong Fluctuation Theory, IEEE Transactions on Geoscience and Remote Sensing, GE-24, 220–231, doi:10.1109/TGRS.1986.289641, 1986.

Takala, M., Luojus, K., Pulliainen, J., Derksen, C., Lemmetyinen, J., Kärnä, J.-P., Koskinen, J., and Bojkov, B.: Estimating northern hemisphere snow water equivalent for climate research through assimilation of space-borne radiometer data and ground-based measurements, Remote Sensing of Environment, 115, 3517–3529, doi:10.1016/j.rse.2011.08.014, http://www.sciencedirect.com/science/article/pii/S0034425711003166, 2011.

Tedesco, M. and Kim, E.: Intercomparison of Electromagnetic Models for Passive Microwave Remote Sensing of Snow, IEEE Transactions on Geoscience and Remote Sensing, 44, 2654–2666, doi:10.1109/TGRS.2006.873182, 2006.

Tedesco, M., Kim, E. J., Cline, D., Graf, T., Koike, T., Armstrong, R., Brodzik, M. J., and Hardy, J.: Comparison of local scale measured and modelled brightness temperatures and snow parameters from the CLPX 2003 by means of a dense medium radiative transfer theory model, Hydrological Processes, 20, 657–672, doi:10.1002/hyp.6129, http://onlinelibrary.wiley.com/doi/10.1002/hyp.6129/abstract, 2006.

Tsang, L., Chen, C.-T., Chang, A. T. C., Guo, J., and Ding, K.-H.: Dense media radiative transfer theory based on quasicrystalline approximation with applications to passive microwave remote sensing of snow, Radio Science, 35, 731–749, doi:10.1029/1999RS002270, http://onlinelibrary.wiley.com/doi/10.1029/1999RS002270/abstract, 2000.

Tsang, L., Pan, J., Liang, D., Li, Z., Cline, D., and Tan, Y.: Modeling Active Microwave Remote Sensing of Snow Using Dense Media Radiative Transfer (DMRT) Theory With Multiple-Scattering Effects, IEEE Transactions on Geoscience and Remote Sensing, 45, 990–1004, doi:10.1109/TGRS.2006.888854, 2007.

Vionnet, V., Brun, E., Morin, S., Boone, A., Faroux, S., Le Moigne, P., Martin, E., Willemet, J., and others: The detailed snowpack scheme Crocus and its implementation in SURFEX v7. 2, Geoscientific Model Development, 5, 773–791, 2012.

Wiesmann, A. and Mätzler, C.: Microwave Emission Model of Layered Snowpacks, Remote Sensing of Environment, 70, 307–316, doi:10.1016/S0034-4257(99)00046-2, http://www.sciencedirect.com/science/article/pii/S0034425799000462, 1999.

Wiesmann, A., Fierz, C., and Mätzler, C.: Simulation of microwave emission from physically modeled snowpacks, Annals of Glaciology, 31, 397–405, doi:10.3189/172756400781820453, 2000.

Xu, X., Tsang, L., and Yueh, S.: Electromagnetic Models of Co/Cross Polarization of Bicontinuous/DMRT in Radar Remote Sensing of Terrestrial Snow at X- and Ku-band for CoReH2O and SCLP Applications, IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, 5, 1024–1032, doi:10.1109/JSTARS.2012.2190719, 2012.

Zuanon, N.: IceCube, a portable and reliable instruments for snow specific surface area measurement in the field, in: International Snow Science Workshop Grenoble-Chamonix Mont-Blance-2013 proceedings, pp. 1020–1023, 2013.
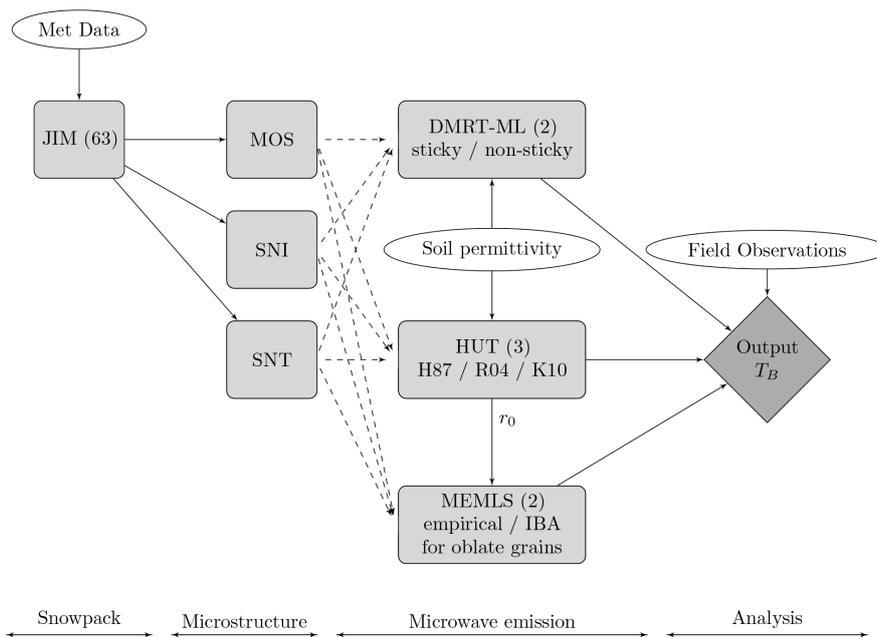
Figure 1. Flowchart showing flow of information from JIM snow evolution model outputs to outputs from the various microwave emission models.
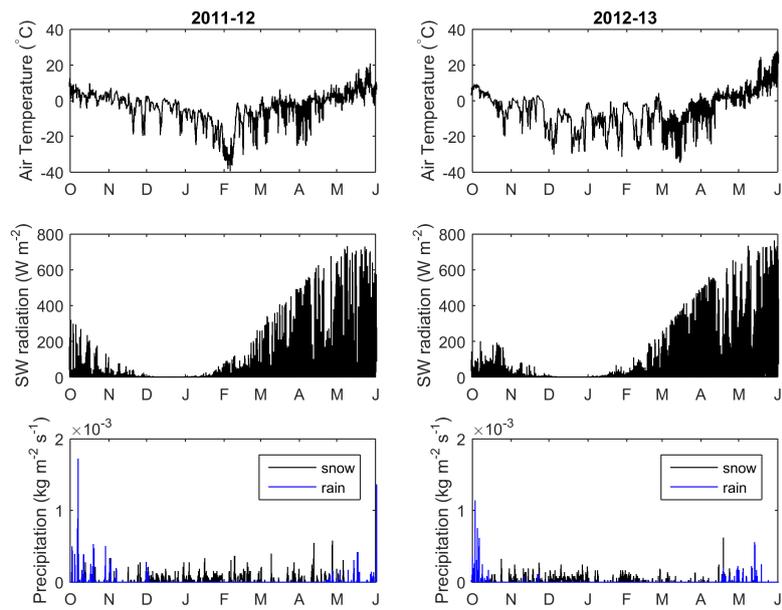
Figure 2. Air temperature, solar radiation and precipitation data measured at the Sodankylä site, used as inputs for the JIM simulations.
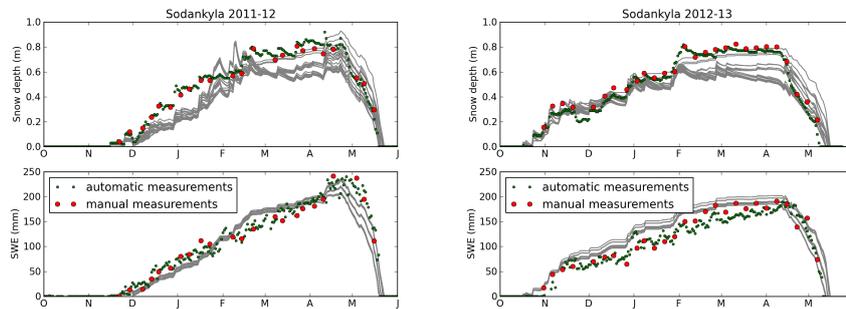
Figure 3. Snow depth and water equivalent simulated by the Jules Investigation Model subset used in this study. Grey lines indicate individual JIM subset member simulations. Note that erroneous positive SWE observation points have been removed at the end of the season when snow depth is zero, as this is a sensor artifact related to soil moisture changes.
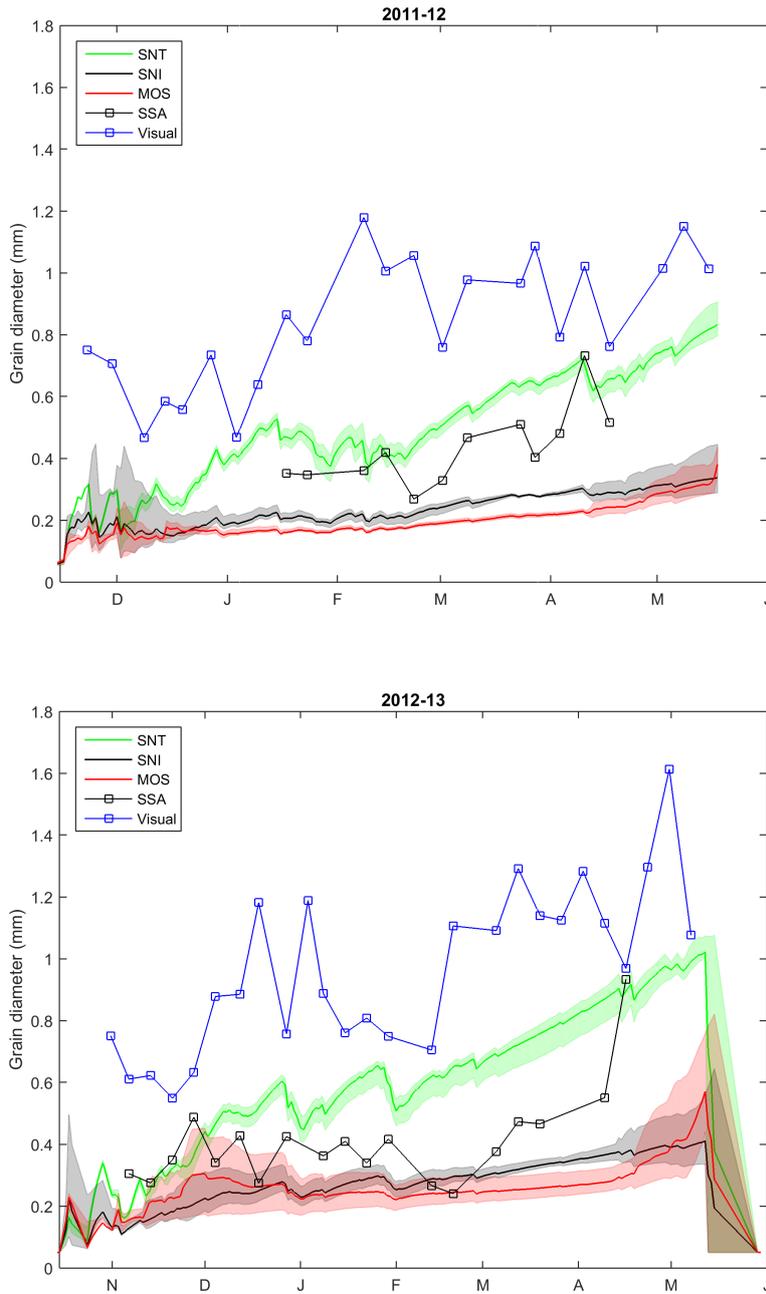
Figure 4. Bulk grain diameter evolution for the MOS, SNT and SNI microstructure evolution models and the spread in model results. Observations of bulk diameter were derived from macro-photography (Visual) and from SSA measurements from the IceCube instrument
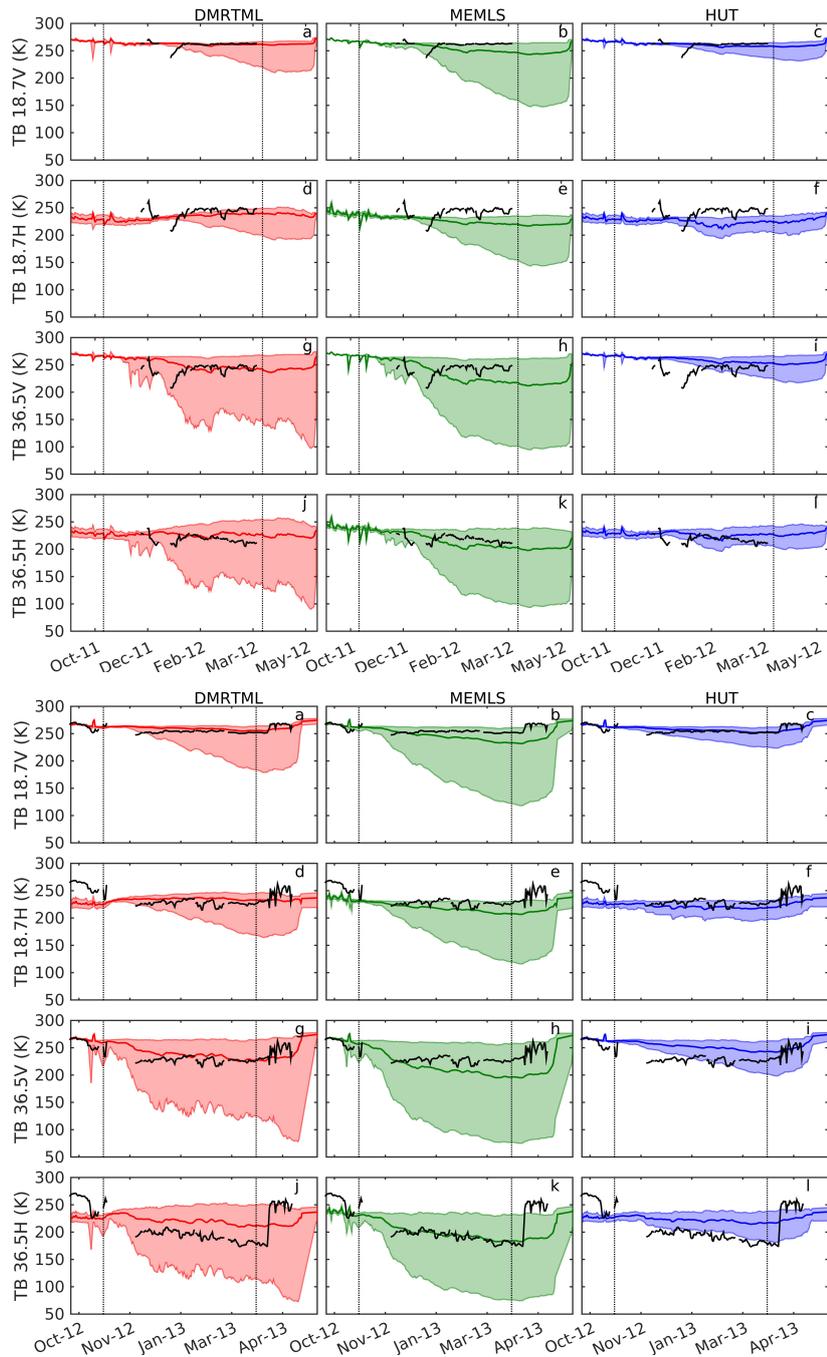
Figure 5. Range and mean of brightness temperature over the two winter seasons as simulated with the DMRT-ML, MEMLS and HUT models, driven by 63 JIM outputs and 3 microstructure evolution models. Black lines indicate the observed brightness temperatures. Vertical dashed lines enclose the period of analysis (1st November - 31st March).
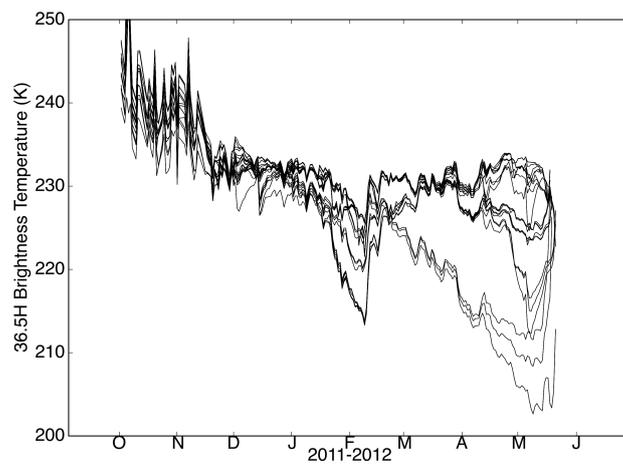
Figure 6. Variability in brightness temperature simulated with empirical MEMLS, driven by the MOSES microstructure model and 63 JIM snowpack outputs (no scaling of microstructure was required).
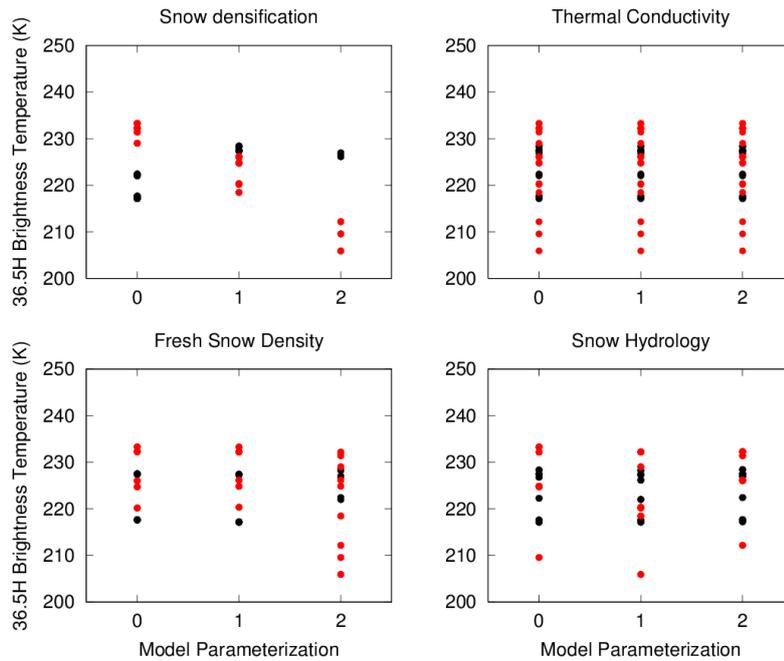
Figure 7. Cluster analysis of brightness temperature simulated by MOSES microstructure and empirical MEMLS for 1st February 2012 (black dots) and 1st May 2012 (red dots) according to model parameterization choices. Brightness temperature simulations are split according to the different representations for each process representation. Values 0, 1, 2 relate to parameterizations given in Essery et al. (2013) and as described in section 2.1. Where distinct clusters occur that differ between parameterizations, this indicates sensitivity to the parameterization.

Table 1. Equations for options governing the representation of processes in the JIM model subset

Option	Description	Model Parameterization
Compaction: 0	Physical	$\frac{1}{\rho_s}\frac{d\rho_s}{dt} = \frac{M_s g}{\eta} + c_1 \exp[-c_2(T_m - T_s) - c_3 \max(0, \rho_s - \rho_0)]$
1	Empirical	$\rho_s(t+\delta t) = \rho_{max} + [\rho_s(t) - \rho_{max}]\exp(-\delta t/\tau_\rho)$
2	Constant	$\rho_s = 250 \text{ kg m}^{-3}$
Fresh snow density: 0	Empirical	$\rho_f = \max[a_f + b_f(T_a - T_m) + c_f U_a^{1/2}, \rho_{min}]$
1	Empirical	$\rho_f = \rho_{min} + \max[d_f(T_a - T_m + e_f)^{3/2}, 0]$
2	Constant	$\rho_f = 100 \text{ kg m}^{-3}$
Thermal conductivity: 0	Empirical	$\lambda_s = \lambda_a + (a_\lambda \rho_s + b_\lambda \rho_s^2)(\lambda_i - \lambda_a)$
1	Empirical	$\lambda_s = c_\lambda \left(\frac{\rho_s}{\rho_w}\right)^{n_\lambda}$
2	Constant	$\lambda_s = 0.265 \text{ W m}^{-1}\text{ K}^{-1}$
Maximum liquid water: 0	Empirical	$\frac{\gamma_{w,max}}{\rho_s} = r_{min} + (r_{max} - r_{min})\max\left(1 - \frac{\rho_s}{\rho_r}, 0\right)$
1	Constant	$\gamma_{w,max} = \rho_w\left(1 - \frac{\gamma_i}{\rho_i}\right)S_{wi}$
2	None	$\gamma_{w,max} = 0 \text{ kg m}^{-3}$

JIM variables are: snow density ρ_s, overlying snow mass M_s, snow temperature T_s, air temperature T_a, wind speed U_a, snow effective thermal conductivity λ_s, partial density of liquid water γ_w, partial density of ice γ_i. Other symbols represent constants, given in Essery et al. (2013)

Table 2. Electromagnetic model inputs, as a function of JIM snowpack model outputs

	JIM output	DMRT-ML input	MEMLS input	HUT input
Temperature	T_{jim} [K]	T_{jim} [K]	T_{jim} [K]	T_{jim} - 273.15 [°C]
Density	ρ_{jim} [kg m^{-3}]	ρ_{jim} [kg m^{-3}]	ρ_{jim} [kg m^{-3}]	$\frac{\rho_{jim}}{1000}$ [g cm^{-3}]
Layer size	Δz_{jim} [m]	Δz_{jim} [m]	$\frac{\Delta z_{jim}}{100}$ [cm]	$\rho_{jim}\Delta z_{jim}$ [mm$_{swe}$]
Microstructure	d_{jim} [μm]	d_{jim} [μm]	$\frac{2}{3}\left(1 - \frac{\rho_s}{\rho_i}\right)\frac{d_{jim}}{1000}$ [mm]	$\frac{d_{jim}}{1000}$ [mm]
Layer number	1 = base	1 = base	1 = base	1 = top
Soil permittivity	-	ϵ_{obs}	$r_{0,HUT}$	ϵ_{obs}

Note that for the purposes of the ensemble, DMRT-ML was adapted to allow the input of diameter rather than radius. HUT was adapted to ensure Fresnel reflectivity for a smooth soil surface, and to output the soil reflectivity $r_{0,HUT}$ at both polarizations for use in MEMLS simulations.

Table 3. Comparison ~~between~~ of grain ~~diameter ratio~~ diameters simulated by different microstructure models. ~~Mean ratios~~ The mean and ~~maximum grain diameter ratios for 2011-2012~~ max ratio between pairs of models is given in columns. Where the 2012-2013 values differ, these are given in parentheses.

	Mean	Max
SNI / MOS	1.2 (1.1)	1.4 (1.3)
SNT / MOS	2.2	3.1 (3.2)
SNT / SNI	1.9 (2.0)	2.5

Table 4. Mean absolute error (mm) between bulk grain diameter simulated with the microstructure models compared with observations derived from SSA measurements with IceCube. Smallest bias for each year is shown in bold. Percentages are given in parentheses.

	2011-2012	2012-2013
MOS	-0.24 (-53%)	-0.16 (-34%)
SNI	-0.18 (-40%)	**-0.14** (-31%)
SNT	**0.12** (32%)	0.16 (45%)

Table 5. Ratio of mean brightness temperature ranges simulated by two microwave emission models. The mean and max ratio between pairs of models is given in columns.

2011-2012	19V	19H	37V	37H	2012-2013	19V	19H	37V	37H
DMRTML / HUT	1.6	1.2	3.8	3.1	DMRTML / HUT	2.0	1.6	3.3	3.0
MEMLS / HUT	3.7	~~2.0~~ 1.9	4.5	~~3.3~~ 3.2	MEMLS / HUT	3.9	2.4	3.9	3.2
MEMLS / DMRTML	2.4	~~1.7~~ 1.5	1.2	~~1.1~~ 1.0	MEMLS / DMRTML	2.0	1.5	1.2	1.1

Table 6. Mean bias and RMSE in brightness temperature (K) simulated by DMRT-ML (~~very~~ sticky and ~~non-sticky~~less sticky), MEMLS (empirical and IBA-oblate) and HUT (H87, R04 and K10) forced by 189 JIM-Microstructure model combinations. Only days in the period 1st November to 31st March, where all four frequency / polarization measurements were available were included in the analysis. Bold values indicate the lowest bias / RMSE for each frequency / polarization.

	2011-2012	19V	19H	37V	37H	2012-2013	19V	19H	37V	37H
	DMRTML	~~0.9~~**0.7**	~~-5.3~~**-5.4**	~~12.7~~10.4	~~10.7~~8.7	DMRTML	~~6.0~~5.6	~~7.2~~**6.9**	~~30.0~~25.2	~~28.7~~24.5
Bias	MEMLS	-7.8	-16.3	**-6.9**	**-6.2**	MEMLS	-9.3	-11.4	**-0.9**	**1.5**
	HUT	-1.3	-18.8	20.6	7.6	HUT	**2.9**	-8.0	39.2	26.4
	DMRTML	**5.5**	~~11.3~~**11.4**	~~14.2~~**12.2**	~~13.8~~12.2	DMRTML	~~6.7~~6.4	~~10.0~~9.8	~~31.0~~26.2	~~31.4~~27.0
RMSE	MEMLS	11.3	20.4	~~13.0~~13.0	13.1	MEMLS	11.4	13.2	**7.0**	**7.0**
	HUT	6.1	22.4	21.1	**11.7**	HUT	**4.2**	**9.8**	40.2	28.9

Table 7. Optimal microwave microstructure scale factors dependent on snow microstructure evolution function for 2011-2012 simulations, based on minimization of cost function between 1st November and 31st March in each year.

		DMRTML non less	DMRTML sticky very	MEMLS IBA	MEMLS EMP	HUT H87	HUT R04	HUT K10
	SNT	2.0 1.1	0.6	0.5	0.3	0.9	0.5	0.7
2011-2012	MOS	2.3 3.3	1.6	1.7	1.0	2.6	1.4	2.2
	SNI	4.2 2.5	1.3	1.2	0.8	1.9	1.1	1.7
	SNT	2.3 1.3	0.7	0.7	0.5	1.2	1.1	1.1
2012-2013	MOS	5.0 3.1	1.7	1.6	1.1	3.2	2.7	2.9
	SNI	5.0 2.8	1.5	1.5	1.1	2.9	2.3	2.6

A value of 1.0 indicates that the snow grain diameter simulated by a particular form of the snow model may be used directly in the microwave model to give the best agreement with measured brightness temperature.

Table 8. Mean bias and RMSE in brightness temperature (K) simulated by DMRT-ML (sticky and non-sticky), MEMLS (empirical and IBA-oblate) and HUT (H87, R04 and K10) forced by 189 JIM-Microstructure model combinations, with optimal microstructure scale factors from Table 7 applied. Only days in the period 1st November to 31st March, where all four frequency / polarization measurements were available were included in the analysis. Bold values indicate the lowest bias / RMSE for each frequency / polarization.

		2011-2012	19V	19H	37V	37H	2012-2013	19V	19H	37V	37H
Bias	DMRTML		~~1.1~~ 1.0	**-5.1**	~~9.9~~ 6.6	~~8.6~~ 5.6		~~3.8~~ 3.7	5.6	**~~2.9~~ -0.3**	**~~4.3~~ 1.4**
	MEMLS		**-0.3**	-11.3	13.1	10.5		**-1.2**	**-5.1**	6.0	9.1
	HUT		-1.3	-18.8	19.1	**6.2** 6.2		-4.4	-14.3	18.0	7.5
RMSE	DMRTML		**~~5.5~~ 5.6**	**11.2**	~~11.4~~ 9.3	~~11.6~~ 9.5		4.7	~~8.9~~ 8.8	**~~5.9~~ 5.2**	**~~9.2~~ 7.5**
	MEMLS		5.7	15.5	14.0	12.7		**3.9**	**7.6**	9.3	11.7
	HUT		6.1	22.4	19.5	~~10.7~~ 10.7		6.2	15.4	19.8	11.2

Table 9. Maximum difference in brightness temperature in K (H-pol / V-pol) due to 63 snowpack parameterisations for each unscaled snow microstructure evolution function

	DMRTML less	DMRTML very	MEMLS IBA	MEMLS EMP	HUT H87	HUT R04	HUT K10
SNT	50 / 63	148 / 169	88 / 113	126 / 153	27 / 34	26 / 33	28 / 36
MOS	22 / 15	41 / 45	24 / 25	33 / 54	20 / 9	21 / 14	20 / 11
SNI	22 / 15	59 / 67	24 / 43	55 / 77	20 / 13	21 / 18	20 / 15