Review comments in bold, responses in plain text

Reviewer 1 Comments

This is an interesting and concise paper that proposes a compact method to evaluate the capacity of land surface models to represent the effect of snow inflation on the underlying soil. I have no doubt that the proposed metric, with some little changes proposed in the following, will be widely used. The paper is yet another illustration why the first author's recent passing away is a huge loss for the scientific community. The figures are all relevant and easily readable. Relevant scientific literature is appropriately referenced. No unnecessary detail clutters the simple and clear message of the paper.

We thank the reviewer for his comments. We agree that the simplicity of the metric is one of its strengths. We also appreciate the comments about Drew's passing and how his passing is a huge loss for the scientific community. We couldn't agree more.

This work should therefore be published after a few minor changes suggested below.

Specific remarks.

- Page 2, line 12: The primary motivation is certainly a good representation of soil temperatures. One could add, however, that wrong temperatures at the snow/soil interface, caused by wrong snow conductivity, can feed back on the snow pack itself via a modified snow metamorphism (in cases models do simulate snow metamorphism dependent on temperature or vertical temperature gradients).

This is a good point. We have added the following sentence: "Additionally, biases in the simulated temperature at the snow/soil interface can adversely affect the snow pack itself though the impact of these biases on snow metamorphism at the base of the snow pack."

- Page 2, line 26: There is a little incoherence that could be acknowledged: The theory presented here initially supposes a periodic (sine) air temperature signal; however, the theory is then limited to the "cooling season".

This is correct. We believe that the text is already relatively clear on this. E.g., we note: "The above theory is adapted to the cooling period of the year, defined here as October to March."

- Page 3, line 3 : "2m air temperature serves as a sufficient proxy as the two quantities tend to equilibrate towards each other, particularly in colder months of high latitude regions with low solar input": Yes and no: In some cases (strong inversion), temperature difference between the snow-air interface and the air at 2 m height can be substantial.

True. There is some error associated with differences (positive or negative) between air temperature at 2m and the temperature at the snow-air interface. We believe that the errors associated with this discrepancy have less of an impact on th metric than errors in the measurements themselves. But, we do now include this statement to acknowledge this point. "The actual land surface temperature is rarely observed *in situ*, but 2m air temperature serves as a sufficient proxy as the two quantities tend to equilibrate towards each other, particularly in colder months of high latitude regions with low solar input,

though in situations with strong inversions, the temperature difference between the snow-air interface and 2m height can be substantial (this is an acknowledged, yet unavoidable limitation)."

- Equation 3: Why not use immediately A_0 and A_z instead of introducing new variables A_air and A_soil which are not really used?

Certainly, one could use A_0 and A_z directly, but we feel that it is actually easier to understand what A_norm is the way it is presented so we have elected to maintain as in the original document.

- Equation 6: The general form of this equation, in particular the numerator of the right hand side, makes sense, but the specific form of the denominator does not. The denominator (which is a constant) should be chosen such that if snow depth is constant (i.e. all snow falls in October), the efficient snow depth is equal to this constant value. Therefore the denominator should read: \sum\limits_{n=1}^{M} n (or (M+1)*M/2, which is equivalent). For the case of the blue curve in figure 1, which is apparently S(i) = i*0.1 (with i=1 for October and i=6 for March), this would yield S_eff=0.266 m, which is less than the average depth of 0.35 m. This would make sense; in Figure 1, for the same case, S_eff is higher than the simple time average, which is incoherent. By the way, I have the impression that equation 6 is not what is plotted in Figure 1. In any case, the difference is only a constant factor, so this has no important effect on the results presented in the rest of the paper. But I think that the definition of S_eff should make immediate sense for simple cases. Right now, it does not.

Agreed. We have fixed the equation and replotted Figure 1, Figure 3, Figure 4, and Figure 5

- Equation 6: What would the results look like if the time period considered would be limited to the period before substantial snow melt occurs? In southerly areas, it can already melt in March. Does this introduce noise?

We tested with various limits to cooling season and the results are qualitatively similar.

- Page 5, line 12: Would it make sense, and would it change the results, to offset the snow depths by adding a positive constant corresponding to a slab of snow with equivalent thermal insulation as 20 cm of soil?

We do not think that this would add any value. There is an offset in the thermal insulation at 0 effective snow depth that represents the thermal offset between air temperatures and 20cm soil temperature. We feel that the way the results are presented now make this clear and that doing something like 'replacing' the snow with a slab of snow would reduce clarity.

- Page 6, line 9: Some models have a vertical 'soil' axis that comprises the snow. That is, 'soil' depth is not counted from the soil-snow interface downwards, but it starts at the snow-atmosphere interface. That could explain some very far off outliers.

We agree and we have already noted this in the discussion. "Incorrect snow heat transfer curves are symptomatic of model deficiencies. As an example, the land scheme in the Hadley Center models used here [MOSES2.2; Essery et al., 2001] applies a composite snow model where the top soil layer and snowpack share the same temperature [Slater et al., 2001], hence insulation is not properly accounted for and cold temperatures easily penetrate into the soil."

- Page 6, line 20: Yes, but the initial argumentation says that the metric presented here is valid in the case when there are no phase changes. (But the argument is correct nevertheless)

That's correct. We have elected to remove this statement because it is not a deficiency that is relevant to the snow insulation and therefore is outside the scope of this paper and has been noted previously in other studies.

- Equation 7 is not particularly elegant. It must be artificially limited to exclude values below 0. A more elegant definition could be: SHTM= SUM(MIN(A_{norm,obs,i},A_{norm,mod,i}))/SUM(MAX(A_{norm,obs,i},A_{norm,mod,i})) This would automatically yield values between 0 and 1 because A_norm is always >=0. Other rather natural and coherent forms for the RHS of equation 7 can be easily defined.

True. Alternative forms of this equation could be implemented, but this is the form that Dr. Slater implemented and we prefer to leave this as is. Alternative forms of the equation, while potentially more elegant, will not yield anything different in terms of results.

Reviewer 2 Comments

General comments: In this paper the authors develop a relatively straightforward diagnostic metric (SHTM - Snow Heat Transfer Metric) for establishing whether the heat transfer through the soil-snow layer is realistically simulated by a climate model. The diagnostic is based on the amplitude equation for the conductive heat flow which is integrated over the period when air temperature are below freezing to obtain the difference in the seasonal temperature amplitudes at some depth in the soil, and the effective snow depth which describes the insulating effect the snow layer over the accumulation season. The authors use observed air temperature, snow depth and soil temperatures at 20 cm from climate stations in Russia, Canada and the USA to obtain an estimate of the curve relating the effective snow depth to the normalized difference in temperature amplitude (Figure 3). There is considerable scatter around this curve which the authors describe as "noise" or "error". However, I suspect that the results shown in Figure 3 represent a number of different curves that reflect different snow-climate regions (e.g. Sturm et al. 1995) and soil properties (e.g. organic soils).

We agree and we have updated the text to make it clear that by noise / error, we are referring to observational error as well as the range of curves that arise due to different snow / soil regimes. We have rewritten to "The observations show the expected exponential shape and fit the underlying theory (5) well despite significant scatter in the data (Figure 3). The scatter is likely due to several things including measurement error, the range of conditions and snow regimes [Sturm et al. 1995], that occur across the landscape, including the timing of snowfall, the pattern of snow metamorphism, the properties and moisture content of the soil as well as uncertainty about the measurement locations of the observed data."

The authors then compare the ability of 13 CMIP5 climate models to replicate the observed relationship derived from the surface observations using all land grid points north of 55 deg N (Figure 4). The results show major differences between models but one question that crops up at this point is whether the somewhat limited spatial sample of observations (Fig. 2) influences this comparison. Repeating the analysis for grid points nearest the observations would answer this question. The large difference in SHTM between models is worrisome but we don't get any sense from the paper of the climatic consequences of a poor fit to the observed heat transfer relationship and how much of the poor fit is coming from representation of snowpack versus the specification of soil thermal properties.

The curves from the models are robust whether or not we sample at the same locations as the obs or globally, though obviously there is more scatter when sampling fewer points.

It is beyond the scope of this study to assess the climatic consequences of a poor representation of snow heat transfer. The one thing that is clear is that permafrost simulations will suffer if snow heat transfer is not represented accurately. To determine broader climatic consequences would require additional climate model sensitivity runs with a range of snow heat transfer representations / parameters.

Because the observations of soil temperature are not at the surface, it isn't possible to fully distinguish where in the snow/upper soil system a disagreement between models and observations might arise. However, the offset in Anorm at zero effective snow depth is an indication of the impact of soil heat transfer between the soil surface and 20cm depth. In obs, this value ranges from about 0.05 to 0.3. Many models exhibit a lower normalized temperature difference value at zero effective snow depth compared to obs, which means that they transfer heat through the soil too efficiently. The sources of a soil heat transfer bias are myriad and could be due to biases/errors in soil texture including organic matter, soil moisture and soil moisture phase, and soil thickness. Many to most of these models do not represent soil organic matter (which is highly insulative), so this is a potential source of bias in these models. We have added a paragraph to discuss this point.

Presumably one would not use a model with a poor SHTM metric for studies of the soil thermal regime or permafrost, but apart from that I'm not quite sure what the metric tell us. The metric would certainly have value in evaluating the performance of different versions of climate models and land surface schemes. One aspect of the paper that could be expanded on (topic for follow-on paper?) is the spatial variability in SHTM in observations and models. In conclusion this paper is a useful addition to the literature and a testament to Drew's ability to derive practical applications from complex processes.

The metric tells us that models with a poor SHTM metric are not correctly modeling the thermal insulation of snow and it should be used to assess the quality of different models and potentially identify what models are fit for purpose (e.g., as reviewer notes, a model with poor SHTM should likely not be utilized in permafrost studies). See below for more comments on spatial variability of SHTM, but basically we are not convinced that it is appropriate, without a lot more observations, to study the spatial variability of SHTM. Our goal here is to generate a constraint on the representation of snow heat transfer in models that can be applied globally.

Detailed comments:

- Page 1, line 30: Mudryk et al (2016) would be a useful reference to cite in this context as it specifically addresses the uncertainty issue in observational SWE datasets
Mudryk, L.R., C. Derksen, P.J. Kushner and R. Brown, 2015. Characterization of Northern Hemisphere snow water equivalent datasets, 1981–2010. Journal of Climate, 28:8037-8051.

Good suggestion. We have added the reference.

- Page 5: Observed data. The authors have a rather limited sample for characterizing the NH land area average amplitude used in eqn. (7). It would be instructive to provide the readers with some idea of the variability in Fig. 3 for a sample of the major snow climate (e.g. Sturm et al. 1995) and ecoclimatic regions.

It would certainly be good to be able to see how the curve differs for different snow climate regimes, but we are limited by the availability of collocated snow, air, and soil temperature data. As noted in the text, at least some of the significant scatter likely arises from the different snow climates that are sampled which lead to different snowpack densities across ecoclimatic regions. The best we can do is to note that some of the scatter is likely attributable to these factors and to further note that snowpacks with seasonal snowpack dynamics and average densities that lie outside our sampling could generate different curves.

- Figure 4: I suggest you use "scatter" rather than the statistical term "error"

Good point. We have modified to using the term 'scatter' rather than 'error' throughout the paper.

- Figure 5: the derivation of Figure 5 is not provided in the paper and there is no discussion of this Figure. This shows the CMIP5 ensemble close to the observations but this is a potentially misleading message.

We aren't clear what the reviewer thinks is missing. The derivation of the figure is in the figure caption and is pretty straightforward. It's true that the CMIP5 ensemble resembles the observations in terms of this diagnostic, which is only indicating that shallow snow depths are more common in both observations and models than deep ones. This seems uncontroversial. Not sure what additional discussion would be helpful.

- Figure 6: What about spatial variability in SHTM? Is this important? How does this vary between models? To what extent do the different geophysical fields used in models contribute to this variability i.e. how much of a model's behaviour in SHTM is related to representation of the snowpack versus specification of soil thermal properties?

There may be some spatial variability in the SHTM, but that is not really the point. In Figure 6, we are emphasizing that the metric is relatively insensitive to climate forcing since the values remain constant through time and with climate change. One needs quite a bit of data to create the functional relationship curves so at best one could potentially create a map of very large regions of SHTM scores, but since the underlying data generating the observed curve is quite sparse, it doesn't really make sense to make a map of SHTM.

It is not possible to identify from these standard CMIP5 model runs where the source of discrepancy between model behavior and the obs comes from. That said, one can infer that the offset of approximately 0.05-0.3 in the observed normalized temperature difference in Figure 4 at zero effective snow depth, reflects the impact of the soil. Models that have a low normalized temperature difference value at zero effective snow depth compared to obs likely transfer heat through the soil too efficiently. The sources of a soil heat transfer bias are myriad and could be due to biases/errors in soil texture including organic matter, soil moisture and soil moisture phase, and soil thickness.

Reviewer 3 comments

This is an interesting paper with a data-driven and process-focussed approach that typifies the recently deceased first author's work. I expect that it will prove to be useful for model evaluation; indeed, this method has already been suggested for inclusion in the methods for evaluating models in the upcoming Earth System Model – Snow Model Intercomparison Project (http://www.climatecryosphere.org/activities/targeted/esm-snowmip). The observed relations ship between "temperature amplitude difference" and "effective snow depth" shown in Figure 3 has a great deal of scatter. A lot of this scatter will come from genuine physical processes. It would be useful to have some

discussion of the influences of soil texture, soil moisture and freezing on the results. Without separating out these influences, it doesn't appear that this method could provide very strong constraints on models, but it is likely to still be useful because current models, as shown in Figure 4, have an even larger range.

We agree that the scatter in the observations can arise from multiple sources and in the revised version of the paper we better acknowledge this at the start of the Results and Discussion section: "The observations show the expected exponential shape and fit the underlying theory (5) well despite significant scatter in the data (Figure 3). The scatter likely arises from several sources including (1) the range of climate conditions and snow regimes [Sturm et al. 1995] that occur across the landscape, including the timing of snowfall and the pattern of snow metamorphism, (2) the properties and moisture content of the soil, and (3) uncertainties in the measurements themselves and the measurement locations of the observed data. It is not possible to distinguish which of these sources of uncertainty dominates the scatter seen in Figure 3."

Additionally, we note that we have already included discussion of the limitations of the SHTM metric and noted that small differences in a score are not necessarily indicative of a significant improvement. And, as the reviewer notes, current generation models show a wide range of performance for this metric. Our perspective is that even a weak constraint is better than no constraint and that if all models could be updated so that they lie somewhere near or within the gray shading in Figures 3/4, that would constitute a big improvement from an Earth System modeling perspective.

It would be interesting to know if the results of this paper can be related to the performances of the same models in simulating permafrost extent, as discussed by Koven et al. (2013). The Hadley Centre model is identified as one in which soil temperatures under snow track air temperatures too closely because of the simplicity of the snow model used. The developers of this model are well aware of this limitation and have implemented a multi-layer snow model to address it; the model is described by Best et al. (2011) and its impacts on permafrost simulations by Chadburn et al. (2015).

We have added the following statement to acknowledge that this issue has been resolved in the Hadley Center model. "Note that the Hadley Centre model developers have addressed this limitation by implementing a multi-layer snow model [Best et al., 2011; Chadburn et al. 2015]."

The definition of effective snow depth in Equation (6) is curious and requires explanation. Why is it chosen so as to give an effective depth that is greater than the average depth for any month for the green line in Figure 1?

We have updated the equation as per the suggestion of reviewer 1 and redrawn the figure.

page 2, line 31 "the period over which the forcing is applied" is ambiguous. Something like "the frequency of the forcing" would be better.

To improve clarity, we modify the sentence to "The value of d is a function of the thermal diffusivity of the medium and the length of time that the forcing is applied."

page 4, line 3 The R parameter is an effective damping depth, not an effective thermal diffusivity.

Corrected.

Best, MJ, and 16 others, 2011. The Joint UK Land Environment Simulator (JULES), model description. Part 1: Energy and water fluxes. Geoscientific Model Development, 4, 677–699.
Chadburn, SE, EJ Burke, RLH Essery, J Boike, M Langer, M Heikenfeld, PM Cox and P Friedlingstein, 2015. Impact of model developments on present and future simulations of permafrost in a global land-surface model. The Cryosphere, 9, 1505–1521.

Process-level model evaluation: A Snow and Heat Transfer Metric

Andrew G. Slater[1], David M. Lawrence[2], Charles D. Koven[3]

[1] NSIDC/CIRES, University of Colorado, Boulder, 80303, USA
[2] National Center for Atmospheric Research, Boulder, 80305, USA
[3] Lawrence Berkeley National Laboratory, Berkeley, 94720, USA

Correspondence to: David Lawrence (dlawren@ucar.edu)

Abstract. Land models require evaluation in order to understand results and guide future development. Examining functional relationships between model variables can provide insight into the ability of models to capture fundamental processes and aid in minimizing uncertainties or deficiencies in model forcing. This study quantifies the proficiency of land models to appropriately transfer heat from the soil through a snowpack to the atmosphere during the cooling season (Northern Hemisphere: October-March). Using the basic physics of heat diffusion, we investigate the relationship between seasonal amplitudes of soil versus air temperatures due to insulation from seasonal snow. Observations demonstrate the anticipated exponential relationship of attenuated soil temperature amplitude with increasing snow depth and indicate that the marginal influence of snow insulation diminishes beyond an 'effective snow depth' of about 50cm. A Snow and Heat Transfer Metric (SHTM) is developed to quantify model skill compared to observations. Land models within the CMIP5 experiment vary widely in SHTM scores and deficiencies can often be traced to model structural weaknesses. The SHTM value for individual models is stable over 150 years of climate 1850-2005, indicating that the metric is insensitive to climate forcing and can be used to evaluate each model's representation of the insulation process.

1 Introduction

The current generation of land models are typically complex in nature and simulate a vast array of processes [Clark et al., 2015; Prentice et al., 2015]. Interdependencies within these models produce external and internal feedbacks that can operate

on various temporal and spatial scales. It is therefore imperative that such models be rigorously evaluated in order to interpret their performance, as well as to guide future development.

Verifying a model result against observations using statistics such as root mean squared error can provide useful information, but alone does not necessarily indicate whether the model is getting the right answers for the right reason [Abramowitz et al., 2008; Gupta et al., 2008]. For example, the impact of the forcing data can play an equal or greater role in results than some aspects of the model [Ménard et al., 2015]. Most global-scale "observationally based" data sets contain substantial uncertainty, especially in the relatively sparsely observed high latitudes [Anisimov et al., 2007; Decker et al., 2012; Slater, 2016; Mudryk et al. 2015], and coupled Earth System Models can produce biased surface meteorology, particularly in high latitude winters [Slater and Lawrence, 2013]. It therefore follows, that if one is trying to assess land model structural error (i.e., deficiencies in model architecture and/or parameterizations) it is preferable to reduce ambiguity by minimizing other sources of uncertainty, e.g. uncertainties in model parameters, initial conditions or forcing (in either land-only or coupled simulations).

In high latitude regions an important process is snow and soil heat transfer. The temperature and state of the soil, whether frozen or thawed, plays a pivotal role in the timing and magnitude of energy, mass and biogeochemical fluxes between the land and atmosphere [Hobbie et al., 2000; Dutta et al., 2006; Monson et al., 2006; Lawrence et al., 2012, 2015]. Concerns about the potential release of vast carbon stores currently locked in permafrost soils and its feedback on the climate has brought this topic to the fore [Schuur et al., 2015]. Model estimates of Arctic carbon fluxes are highly variable [Fisher et al., 2014; Koven et al., 2015] in part due to differences in simulation of soil temperatures and permafrost conditions. Essentially, we need to ensure soil temperature and moisture are correctly simulated before confidence in projected biogeochemical fluxes can be achieved. Additionally, biases in the simulated temperature at the snow/soil interface can adversely affect the snow pack itself though the impact of these biases on snow metamorphism at the base of the snow pack.

In cold mid-latitude and Arctic regions, snow forms an insulating barrier between the colder atmosphere and underlying soil during winter. The impact of snow on soil temperatures, particularly in permafrost regions, has been well documented in both observational [Sharratt et al., 1992; Zhang, 2005] and modeling studies [Luetschg et al., 2008; Lawrence and Slater, 2010; Ekici et al., 2015; Yi et al., 2015]. A variety of structures has been used to represent snow in models [Slater et al., 2001], with varying levels of simulated insulation [Koven et al., 2013]. The aim of this work is to define a metric that demonstrates processes-level model evaluation using the heat transfer mechanism from atmosphere to soil under the conditions of a seasonal snowpack.

2 Method

The theory of conductive heat flow under periodic forcing (e.g., the annual cycle) is demonstrated in many texts [Lunardini, 1981; Yershov, 1998].Taking an example of a semi-infinite medium forced at the surface with a periodic temperature wave with no phase change or mass exchange, we can compute the amplitude of this periodic wave at soil depth (z) as:

$$A_z = A_o\, e^{-\left(\frac{z}{d}\right)} \quad (1)$$

where A_z is the amplitude of the temperature wave at depth (z), A_o is the surface amplitude and d is the damping depth defined as the depth at which the $1/e$ of the surface amplitude occurs. The value of d is a function of the thermal diffusivity of the medium and the length of time that the forcing is applied.

The above theory is adapted to the cooling period of the year, defined here as October to March. A vast portion of the terrestrial Northern Hemisphere becomes snow covered during October and remains so beyond March. The actual land surface temperature is rarely observed *in situ*, but 2m air temperature serves as a sufficient proxy as the two quantities tend to equilibrate towards each other, particularly in colder months of high latitude regions with low solar input, though in situations with strong inversions, the temperature difference between the snow-air interface and 2m height can be substantial (this is an acknowledged, yet unavoidable limitation). Observations (Section 3) show that the coldest air temperature is typically in January or February, while the coldest soil temperature at 20cm depth usually lags by a month. Confining data to the annual cooling period, the amplitude of air temperature is taken as

$$A_{air} = MAX(T_{air}) - MIN(T_{air}) \quad (2)$$

The soil temperature amplitude (A_{soil}) can be calculated annually in the same way. To elucidate the process of heat transfer and remove climatically driven factors, for example the large seasonal cycle associated with deep continental regions as compared to more moderate coastal locations, a normalized temperature amplitude difference is computed as

$$A_{norm} = \frac{A_{air} - A_{soil}}{A_{air}} \quad (3)$$

which ranges from 0 to 1. A_{norm} values near 0.0 indicate minimal difference in the annual cycle of air and soil temperatures, while a value close to 1.0 suggests soil temperatures essentially do not change over the cooling period. If we take A_{soil} to be A_z and A_{air} as A_o, then substitute equation (3) into equation (1) we arrive at the form where:

David Lawrence 3/8/2017 10:18 AM
Deleted: period over

David Lawrence 3/8/2017 10:18 AM
Deleted: which

$$A_{norm} = 1. - e^{-\left(\frac{z}{d}\right)} \quad (4)$$

Such theory pertains to an idealized case, but in reality the distance z would be affected by the quantity and temporal sequence of snow accumulation and the damping depth d will be impacted by snow density, soil inhomogeneity or phase change. Therefore, we propose a similar but more flexible approach:

$$A_{norm} = P + Q\left(1. - e^{-\left(\frac{S_{depth,eff}}{R}\right)}\right) \quad (5)$$

We introduce an offset, P, because even if there was absolutely no snow, the thermal properties of the soil, along with phase change, and imperfect heat transfer between the surface and near-surface air mean that the amplitude of soil temperature at a depth (e.g., 20cm) will likely be different to that of the air. A multiplier, Q, is applied to account for the temporal nature of snow accumulation. Our analysis pertains to locations with seasonal, rather than permanent, snow cover so even if 3m of snow accumulates by the end of March there is likely to be some cooling in soil temperature from atmospheric forcing in October. The R parameter is an effective damping depth of the snow and soil system. The density and morphology of the snow, along with soil moisture content and phase change play a role in determining R. The R value also tends to govern the marginal influence of additional snow insulation. To account for seasonal variations in the linear distance between the soil and the atmosphere (i.e., the snow depth z, in equation 4), we define an effective snow depth ($S_{depth,eff}$).

David Lawrence 3/10/2017 9:25 AM
Deleted: thermal diffusivity

$S_{depth,eff}$ describes the insulation impact of snow and is an integral value such that the mean snow depth (S) each month (m; 1-6) is weighted by its duration. The maximum duration (M=6) is the total cooling period of 6 months (Oct-Mar). The first snow depth value (S_1) is the October mean snow depth.

$$S_{depth,eff} = \frac{\sum_{m=1}^{M}\left(S_m * (M + 1 - m)\right)}{\sum_{m=1}^{M} m} \quad (6)$$

David Lawrence 2/25/2017 2:40 PM
Deleted: $\overline{\left(\sum_{m=1}^{M+1} m\right)}^{2}$

As shown in Figure 1, a season with an early snowfall will typically produce a higher effective snow depth compared to a linearly increasing snowpack with the same mean value. Similarly, if shallow snow persists for most of the winter but a large snowfall occurs in February, the effective snow depth will be lower than the linear case.

Given inputs of A_{norm} and $S_{depth,eff}$ (from observations or models) we can efficiently compute the values of the three parameters P, Q, R using a non-linear fitting method (e.g. the Levenberg-Marquardt (LM) algorithm [Press et al., 2007]). The relationship between A_{norm} and $S_{depth,eff}$ informs us about the heat transfer mechanism between the atmosphere and the

subsurface. To alleviate the problem of overfitting, the curve is fit using data that is sampled evenly across a range of effective snow depths with up to 35 data points per 5cm of $S_{depth,eff}$ up to 50cm (with the final category being 45cm or beyond). This sampling was performed 100 times, with the median of fit values taken; if the LM algorithm does not converge for a given sample, another sample is taken. Sampling also ensures that a wide variety of climatic regimes is used for characterizing the functional relationship seen in observations or a model.

David Lawrence 3/15/2017 12:25 PM
Deleted: 10

David Lawrence 3/15/2017 12:25 PM
Deleted: 1

David Lawrence 3/15/2017 12:25 PM
Deleted: 90

3 Data

Assessing the relation between A_{norm} and $S_{depth,eff}$ requires three pieces of data: screen level air temperature, 20cm soil temperature (as this is the most commonly observed depth) and snow depth. An analysis constraint is the need to use monthly mean values, as this is the most common output from large-scale land models.

A large network of hydrometeorological stations in Russia (and the Former Soviet Union) provide the three required variables as follows: air temperature was acquired from the NCDC Global Summary of the Day (https://data.noaa.gov/dataset/global-surface-summary-of-the-day-gsod), provided the snow depths measured at stations (as opposed to local transects) and information on soil temperatures is available in Sherstyukov and Sherstyukov, [2015]. Data in the USA is from the Natural Resources Conservation Service as part of the SNOTEL and SCAN networks (http://www.wcc.nrcs.usda.gov/). In Canada, Alberta's AgroClimate Information Service (ACIS; http://agriculture.alberta.ca/acis/) collates and distributes station data within the province.

For a given season, we use only sites with complete, quality controlled records from October to March; 2049 observed site-years are available. It is also recognized that the data has inherent limitations, for example it is unknown whether snow depths are measured at precisely the same location as the soil temperature measurements. Many measurements have been made on soils that have been historically altered, however our use of shallow (20cm) soil temperatures and focus on the winter period aims to minimizes possible artifacts due to disturbance.

Snow heat transfer in 13 models that participated in the CMIP5 experiment [Taylor et al., 2012; http://cmip-pcmdi.llnl.gov/cmip5/] is evaluated using the Historical (1850-2005) simulations. Soil columns within the various land models have different depths and layer thicknesses [Slater and Lawrence, 2013; their Figure 2] so a spline interpolation of each monthly mean temperature profile was used to estimate a 20cm value. Further details of the land models are available in Koven et al., [2013] and Slater and Lawrence, [2013].

All data is restricted to those instances where, during the cooling period (Oct-Mar), mean T_{air} is below -1°C, mean T_{soil} is below 2.5°C, A_{air} is greater than 10°C and $S_{depth,eff}$ is greater than 1cm and less than 150cm. Observed locations fitting the

David Lawrence 3/15/2017 12:24 PM
Deleted: 0.01

David Lawrence 3/15/2017 12:24 PM
Deleted: .5

above data criteria are shown in Figure 2. Model data meet the same criteria, but our intention is to test the ubiquitous physics of a process, thus model data is not restricted to the same locations as the observations and an order of magnitude more (or greater) grid cells from the models are available for sampling.

4 Results and Discussion

The observations show the expected exponential shape and fit the underlying theory (5) well despite significant scatter in the data (Figure 3). The scatter likely arises from several sources including (1) the range of climate conditions and snow regimes [Sturm et al. 1995] that occur across the landscape, including the timing of snowfall and the pattern of snow metamorphism, (2) the properties and moisture content of the soil, and (3) uncertainties in the measurements themselves and the measurement locations of the observed data. It is not possible to distinguish which of these sources of uncertainty dominates the scatter seen in Figure 3 and it remains possible that snowpack dynamics in regions outside the sampled data would generate slightly different relationships, though the shape of the curve would not likely change. Grey shading in Figure 3 shows the span of median scatter of all 100 fits to the observations with the black curve being the median fit value.

The observational results demonstrate that the *marginal* influence of snow insulation relative to the annual cycle of air temperature diminishes beyond a $S_{depth,eff}$ of about 25cm. This phenomenon of insulation saturation has been previously noted in observations and models [Zhang, 2005; Lawrence and Slater, 2010].

Results from models are calculated the same as observations, with the median curve for each model shown on the same plot (Figure 4). The difference amongst model curves is considerable. Several models produce an exponential-like curve (e.g., CCSM4, GISS, MRI, NorESM), suggesting that they generally reflect the character of the observed relationship. Many models, however, do not reproduce the observed relationship. Models such as CanESM, GFDL, Hadley models, MIROC and MPI are more linear in their form. This group of poorer performing models all show A_{norm} values of less than 0.45 at $S_{depth,eff}$ of 20cm (compared to a median observed value of 0.65 at 20cm depth). A low A_{norm} value is produced when soil temperatures more closely track the changes in air temperature rather than being modulated by snow cover.

Note that the impact of heat transfer from the soil surface to 20cm depth can be inferred from the A_{norm} values at $S_{depth,eff}$ = 0cm (A_{norm} typically between 0.05 and 0.30 at $S_{depth,eff}$ = 0cm). Many models exhibit a lower normalized temperature difference value at $S_{depth,eff}$ = 0cm, which suggests that these models are transferring heat through the upper soil too efficiently. The potential sources of a soil heat transfer bias are myriad and could be due to biases/errors in soil texture, soil moisture, soil water phase, soil thickness, and vegetation. Most of these models, apart from CLM [Lawrence and Slater, 2008], do not represent the highly insulative soil organic matter, so this is a potential explanation of the common biases at $S_{depth,eff}$ = 0cm.

David Lawrence 3/10/2017 11:57 AM
Deleted: more

David Lawrence 3/10/2017 9:28 AM
Deleted: is

David Lawrence 2/25/2017 3:42 PM
Deleted: noise

David Lawrence 2/25/2017 3:42 PM
Deleted: noise

David Lawrence 3/8/2017 10:31 AM
Deleted: is due to

David Lawrence 2/25/2017 3:45 PM
Deleted: a

David Lawrence 2/25/2017 3:45 PM
Deleted: such as the

David Lawrence 2/25/2017 3:46 PM
Deleted: exact

David Lawrence 3/8/2017 10:26 AM
Deleted: ,

David Lawrence 3/8/2017 10:32 AM
Deleted: as well as

David Lawrence 3/8/2017 10:26 AM
Deleted: y

David Lawrence 3/8/2017 10:27 AM
Deleted: about

David Lawrence 2/25/2017 3:46 PM
Deleted: exact

David Lawrence 3/10/2017 9:24 AM
Deleted: error

David Lawrence 3/15/2017 12:24 PM
Deleted: 50

David Lawrence 3/15/2017 12:24 PM
Deleted: 40

David Lawrence 3/15/2017 12:24 PM
Deleted: 40

David Lawrence 3/15/2017 1:29 PM
Formatted: No widow/orphan control, Don't adjust space between Latin and Asian text, Don't adjust space between Asian text and numbers

Incorrect snow heat transfer curves are symptomatic of model deficiencies. As an example, the land scheme in the Hadley Center models used here [MOSES2.2; Essery et al., 2001] applies a composite snow model where the top soil layer and snowpack share the same temperature [Slater et al., 2001], hence insulation is not properly accounted for and cold temperatures easily penetrate into the soil. Conversely, the better performing models feature multi-layer snow packs that are more apt at emulating the nonlinear temperature profile of the snow pack. Note that the Hadley Centre model developers have addressed this limitation by implementing a multi-layer snow model [Best et al., 2011; Chadburn et al. 2015].

Despite the application of different coupled model component (e.g., atmosphere or ocean models) and initial conditions, both of which will influence terrestrial surface climate, the two Hadley Center climate models (HadGEM2-CC, HadGEM2-ES, both utilizing MOSES as their land model) produce essentially the same snow insulation curve. Similarly, CCSM4 and NorESM reproduce essentially the same curve, which is expected since both models utilize CLM4 as their land model. The fact that snow insulation curves are essentially the same for a particular land model, even when driven with different climatic forcings from their parent climate model, suggests that the form of the curve does indeed capture the functional land model behavior.

The steep gradient of A_{norm} at shallower $S_{depth,eff}$ is an important feature to capture as over 85% of seasonally snow covered land is estimated to have a $S_{depth,eff}$ less than 30cm. Reliable observed data of hemispheric scale snow depths or mass is not yet available, but the prevalence of shallow snow can be seen in estimated climatological $S_{depth,eff}$ derived from numerous reanalyses (ERA-Interim [Dee et al., 2011], MERRA [Rienecker et al., 2011], CFSR [Saha et al., 2010], JRA-55 [Kobayashi et al., 2015], some of which assimilate snow depths), products such as GlobSnow [Takala et al., 2011] and station interpolations [Foster and Davy, 1988; Brown and Brasnett, 2010] (Figure 5).

4.1 A Model Diagnostic: The Snow and Heat Transfer Metric (SHTM)

A land model should be able to capture the exponential relationship between A_{norm} and $S_{depth,eff}$, and it is useful to summarize the ability of a given model to do so as a compact metric. Here, we develop a Snow and Heat Transfer Metric (*SHTM*) that could be used within a broader land model analysis system such as the International Land Model Benchmarking project (ILAMB; http://www.ilamb.org; Luo et al., [2012]). The metric is designed to have a value from 0 to 1, and describes the departure of a model's snow insulation curve from the observed curve. As the marginal influence of snow insulation decreases after a $S_{depth,eff}$ value of about 25cm and most of the seasonal snow regions have a $S_{depth,eff}$ below 30cm (Figure 5), the SHTM value is only calculated over the range of 0 to 30cm. For each 1cm of $S_{depth,eff}$, up to 30cm, twice the difference in A_{norm} between the model and observational curves is obtained; a root mean square error of these values is then computed and subtracted from 1. The *SHTM* value is thus:

David Lawrence 2/25/2017 3:17 PM
Deleted: ... [1]

David Lawrence 3/10/2017 12:21 PM
Deleted:

David Lawrence 3/15/2017 12:23 PM
Deleted: 0.6

David Lawrence 3/15/2017 12:23 PM
Deleted: 0.50

David Lawrence 3/15/2017 12:23 PM
Deleted: 0.60

David Lawrence 3/15/2017 12:23 PM
Deleted: -0.60

David Lawrence 3/15/2017 12:23 PM
Deleted: 0.01

David Lawrence 3/15/2017 12:23 PM
Deleted: 0.60

$$SHTM = 1. - \sqrt{\overline{\left[2 * (Model\ A_{norm,t} - Observed\ A_{norm,t})\right]^2}} \quad (7)$$

The closer *SHTM* is to 1, the better the model is at reproducing the observed snow insulation curve; a lower limit of 0 is placed on the SHTM. An important feature of the S*HTM* is that it effectively isolates analysis of the snow insulation process and therefore should be robust across different climate forcing. We test that hypothesis here by calculating the *SHTM* for ten 15-year periods from 1850-2000 for each model. The mean 2m air temperature over the terrestrial Northern Hemisphere north of 55°N for the cooling period (Oct-Mar) was computed for each 15-year period for each model. As a measure of changing climates, the minimum and maximum of these 15-year averages was differenced per model, resulting in a span of 1.25°C to 4.15°C, with a mean of 2.15°C. Despite these climatic changes, the *SHTM* maintains a fairly constant value for each land model. Over the 150 years, all models return a *SHTM* standard deviation of less than 0.021, which is substantially less than the typical *SHTM* difference between models and indicates a general invariance to climate variability (Figure 6).

Due to scatter in the data (both modeled and observed data), the S*HTM* can only be used as a broader indicator of model skill. Based on the standard deviation of scores and observational uncertainty, *SHTM* values within 0.05 of each other may not be significantly different and incremental improvements to the snow insulation schemes may not necessarily result in meaningful changes in the *SHTM* score. Additionally, the *SHTM* does not evaluate the ability of models to simulate snow accumulation or ablation; additional datasets and metrics are needed to assess these processes.

David Lawrence 2/25/2017 3:47 PM
Deleted: noise

David Lawrence 3/8/2017 10:24 AM
Deleted: ability

David Lawrence 3/8/2017 10:34 AM
Deleted: values

David Lawrence 3/8/2017 10:35 AM
Formatted: Font:Italic

5 Conclusions

Model evaluation is an important aspect of the overall modeling process: e.g., development, application and/or prediction. Land, hydrology and ecology models can range in complexity from simpler, more conceptually-based, formulations through to highly explicit representations that incorporate a multitude of processes. Regardless of their complexity, it is important that the underlying physical processes are represented correctly.

Deriving the relationship between a normalized temperature amplitude difference between air and soil, A_{norm}, and the seasonal effective snow depth, $S_{depth,eff}$, we have shown that observations are consistent with heat transfer theory. Analysis of observed results suggests that the marginal impact of snow insulation diminishes beyond a $S_{depth,eff}$ of about 25cm. Structural weaknesses in several models have been exposed by examining their ability to represent the atmosphere-land heat transfer process in the presence of snow. To quantitatively compare model performance, a snow and heat transfer metric, *SHTM*, was designed. The *SHTM* value per model changed little when calculated for different periods over 150 years of climate change,

suggesting it can fairly unambiguously provide an indication of whether model structural/parameter deficiencies exist by negating other areas of uncertainty, such as the model forcing data. The *SHTM* is a useful model diagnostic that can be added to existing land model analysis / benchmarking systems.

6 Author Contribution

AGS and DML conceived the idea. AGS conducted the analysis. AGS, DML, and CDK wrote the manuscript.

7 Acknowledgements

Among the authors, this work was supported by NSF PLR-1304152, PLR-1304220, NASA IDS project 1517544, DOE DE-FC03-97ER62402/A010 and DE-AC02-05CH11231. Code for this method is to be implemented within the ILAMB system. Data sources and acknowledgements are given in the text.

References

Abramowitz, G., R. Leuning, M. Clark, and A. Pitman (2008), Evaluating the Performance of Land Surface Models, J. Clim., 21(21), 5468–5481, doi:10.1175/2008JCLI2378.1.

Anisimov, O. A., V. A. Lobanov, S. A. Reneva, N. I. Shiklomanov, T. Zhang, and F. E. Nelson (2007), Uncertainties in gridded air temperature fields and effects on predictive active layer modeling, J. Geophys. Res.-EARTH Surf., 112(F2), doi:10.1029/2006JF000593.

Best, M.J., and 16 others, (2011), The Joint UK Land Environment Simulator (JULES), model description. Part 1: Energy and water fluxes. Geoscientific Model Development, 4, 677–699.

Brown, R., and B. Brasnett (2010), Canadian Meteorological Centre (CMC) Daily Snow Depth Analysis Data. © Environment Canada, 2010. Boulder, Colorado USA: National Snow and Ice Data Center. Digital media,

Chadburn, S.E., E.J. Burke, R.L.H. Essery, J. Boike, M. Langer, M. Heikenfeld, P.M. Cox and P. Friedlingstein, (2015), Impact of model developments on present and future simulations of permafrost in a global land-surface model. The Cryosphere, 9, 1505–1521.

Clark, M. P. et al. (2015), Improving the representation of hydrologic processes in Earth System Models, Water Resour. Res., 51(8), 5929–5956, doi:10.1002/2015WR017096.

Decker, M., M. A. Brunke, Z. Wang, K. Sakaguchi, X. Zeng, and M. G. Bosilovich (2012), Evaluation of the Reanalysis Products from GSFC, NCEP, and ECMWF Using Flux Tower Observations, J. Clim., 25(6), 1916–1944, doi:10.1175/JCLI-D-11-00004.1.

Dee, D. P. et al. (2011), The ERA-Interim reanalysis: configuration and performance of the data assimilation system, Q. J. R. Meteorol. Soc., 137(656, Part a), 553–597, doi:10.1002/qj.828.

Dutta, K., E. A. G. Schuur, J. C. Neff, and S. A. Zimov (2006), Potential carbon release from permafrost soils of Northeastern Siberia, Glob. Change Biol., 12(12), 2336–2351, doi:10.1111/j.1365-2486.2006.01259.x.

Ekici, A. et al. (2015), Site-level model intercomparison of high latitude and high altitude soil thermal dynamics in tundra and barren landscapes, Cryosphere, 9(4), 1343–1361, doi:10.5194/tc-9-1343-2015.

Essery, R. L. H., M. J. Best, and P. M. Cox (2001), MOSES 2.2 Technical Documentation, Hadley Centre technical note 30.

Fisher, J. B. et al. (2014), Carbon cycle uncertainty in the Alaskan Arctic, Biogeosciences, 11(15), 4271–4288.

Foster, D., and R. Davy (1988), Global Snow Depth Climatology, US Air Force Enviro. Tech. Appl. Cent., Scott Air Force Base, Illinois, USA.

Gupta, H. V., T. Wagener, and Y. Liu (2008), Reconciling theory with observations: elements of a diagnostic approach to model evaluation, Hydrol. Process., 22(18), 3802–3813, doi:10.1002/hyp.6989.

Hobbie, S. E., J. P. Schimel, S. E. Trumbore, and J. R. Randerson (2000), Controls over carbon storage and turnover in high-latitude soils, Glob. Change Biol., 6, 196–210, doi:10.1046/j.1365-2486.2000.06021.x.

Kobayashi, S. et al. (2015), The JRA-55 Reanalysis: General Specifications and Basic Characteristics, J. Meteorol. Soc. Jpn. Ser II, 93(1), 5–48, doi:10.2151/jmsj.2015-001.

Koven, C., W. Riley, and A. Stern (2013), Analysis of permafrost thermal dynamics and response to climate change in the CMIP5 Earth System Models, J. Clim., doi:10.1175/JCLI-D-12-00228.1.

Koven, C. D. et al. (2015), A simplified, data-constrained approach to estimate the permafrost carbon-climate feedback., Philos. Transact. A Math. Phys. Eng. Sci., 373(2054), doi:10.1098/rsta.2014.0423.

Lawrence, D. M., and A. G. Slater (2010), The contribution of snow condition trends to future ground climate, Clim. Dyn., 34(7–8), 969–981, doi:10.1007/s00382-009-0537-4.

Lawrence, D. M., A. G. Slater, and S. C. Swenson (2012), Simulation of Present-Day and Future Permafrost and Seasonally Frozen Ground Conditions in CCSM4, J. Clim., 25(7), 2207–2225, doi:10.1175/JCLI-D-11-00334.1.

Lawrence, D. M., C. D. Koven, S. C. Swenson, W. J. Riley, and A. G. Slater (2015), Permafrost thaw and resulting soil moisture changes regulate projected high-latitude CO 2 and CH 4 emissions, Environ. Res. Lett., 10(9), 94011.

Lawrence, D.M. and A.G. Slater (2008), Incorporating organic soil into a global climate model, Clim. Dyn., 30, doi:10.1007/s00382-007-0278-1.

Luetschg, M., M. Lehning, and W. Haeberli (2008), A sensitivity study of factors influencing warm/thin permafrost in the Swiss Alps, J. Glaciol., 54(187), 696–704, doi:10.3189/002214308786570881.

Lunardini, V. J. (1981), Heat Transfer in Cold Climates, Van Nostrand Reinhold, New York.

Luo, Y. Q. et al. (2012), A framework for benchmarking land models, Biogeosciences, 9(10), 3857–3874, doi:10.5194/bg-9-3857-2012.

Ménard, C. B., J. Ikonen, K. Rautiainen, M. Aurela, A. N. Arslan, and J. Pulliainen (2015), Effects of Meteorological and Ancillary Data, Temporal Averaging, and Evaluation Methods on Model Performance and Uncertainty in a Land Surface Model, J. Hydrometeorol., 16(6), 2559–2576, doi:10.1175/JHM-D-15-0013.1.

Monson, R. K., D. L. Lipson, S. P. Burns, A. A. Turnipseed, A. C. Delany, M. W. Williams, and S. K. Schmidt (2006), Winter forest soil respiration controlled by climate and microbial community composition, Nature, 439(7077), 711–714, doi:10.1038/nature04555.

Mudryk, L.R., C. Derksen, P.J. Kushner and R. Brown, 2015. Characterization of Northern Hemisphere snow water equivalent datasets, 1981–2010. Journal of Climate, 28:8037-8051.

Prentice, I. C., X. Liang, B. E. Medlyn, and Y.-P. Wang (2015), Reliable, robust and realistic: the three R's of next-generation land-surface modelling, Atmospheric Chem. Phys., 15(10), 5987–6005, doi:10.5194/acp-15-5987-2015.

Press, W., S. Teukolsky, W. Vetterling, and B. Flannery (2007), Numerical Recipes: The Art of Scientific Computing, Cambridge University Press, New York.

Rienecker, M. M. et al. (2011), MERRA: NASA's Modern-Era Retrospective Analysis for Research and Applications, J. Clim., 24(14), 3624–3648, doi:10.1175/JCLI-D-11-00015.1.

Saha, S. et al. (2010), THE NCEP CLIMATE FORECAST SYSTEM REANALYSIS, Bull. Am. Meteorol. Soc., 91(8), 1015–1057, doi:10.1175/2010BAMS3001.1.

Schuur, E. A. G. et al. (2015), Climate change and the permafrost carbon feedback, Nature, 520(7546), 171–179, doi:10.1038/nature14338.

Sharratt, B., D. Baker, D. Wall, R. Skaggs, and D. Ruschy (1992), Snow Depth Required for Near Steady-State Soil Temperatures, Agric. For. Meteorol., 57(4), 243–251, doi:10.1016/0168-1923(92)90121-J.

Sherstyukov, A. B., and B. G. Sherstyukov (2015), Spatial features and new trends in thermal conditions of soil and depth of its seasonal thawing in the permafrost zone, Russ. Meteorol. Hydrol., 40(2), 73–78, doi:10.3103/S1068373915020016.

Slater, A. G. (2016), Surface Solar Radiation in North America: A Comparison of Observations, Reanalyses, Satellite, and Derived Products, J. Hydrometeorol., 401–420, doi:10.1175/JHM-D-15-0087.1.

Slater, A. G., and D. M. Lawrence (2013), Diagnosing Present and Future Permafrost from Climate Models, J. Clim., 26(15), 5608–5623, doi:10.1175/JCLI-D-12-00341.1.

Slater, A. G. et al. (2001), The representation of snow in land surface schemes: Results from PILPS 2(d), J. Hydrometeorol., 2(1), 7–25, doi:10.1175/1525-7541(2001)002<0007:TROSIL>2.0.CO;2.

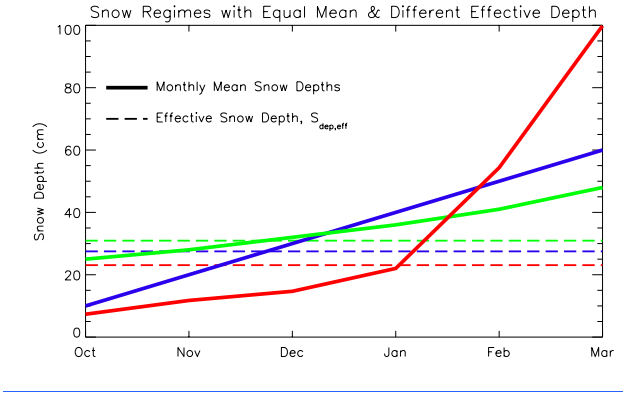
Takala, M., K. Luojus, J. Pulliainen, C. Derksen, J. Lemmetyinen, J.-P. Karna, J. Koskinen, and B. Bojkov (2011), Estimating northern hemisphere snow water equivalent for climate research through assimilation of space-borne radiometer data and ground-based measurements, Remote Sens. Environ., 115(12), 3517–3529, doi:10.1016/j.rse.2011.08.014.

Taylor, K. E., R. J. Stouffer, and G. A. Meehl (2012), AN OVERVIEW OF CMIP5 AND THE EXPERIMENT DESIGN, Bull. Am. Meteorol. Soc., 93(4), 485–498, doi:10.1175/BAMS-D-11-00094.1.

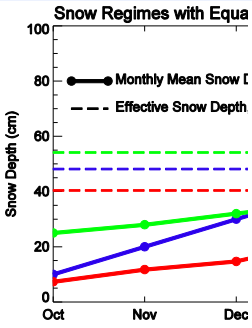
Yershov, E. (1998), General Geocryology, Cambridge University Press.

Yi, Y., J. S. Kimball, M. A. Rawlins, M. Moghaddam, and E. S. Euskirchen (2015), The role of snow cover affecting boreal-arctic soil freeze-thaw and carbon dynamics, Biogeosciences, 12(19), 5811–5829, doi:10.5194/bg-12-5811-2015.

Zhang, T. (2005), Influence of the seasonal snow cover on the ground thermal regime: An overview, Rev. Geophys., 43(4), doi:10.1029/2004RG000157.



David Lawrence 3/15/2017 11:57 AM



Deleted:

Figure 1: Effective snow depth ($S_{depth,eff}$, dashed lines) of three different snow regimes that have equal mean values over the period October-March. Earlier snowfall receives greater weighting as it represents greater insulation; data in green thus has the greatest effective depth.

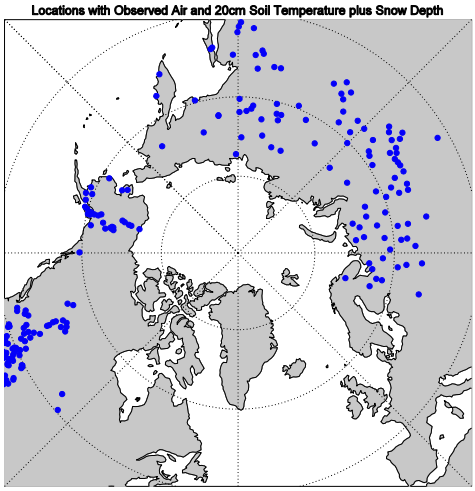


Figure 2: Locations with co-located observations of air temperature, 20cm soil temperature, and snow depth.

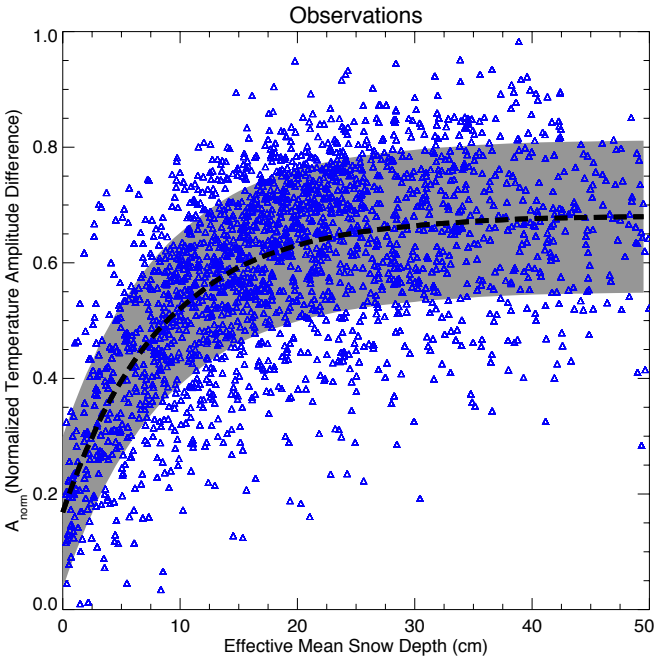


Figure 3: Observed relation between A_{norm} and effective mean snow depth ($S_{depth,\,eff}$) along with the resulting exponential fit (dashed line). The grey shading shows the median fit plus/minus the mean scatter (O_{err}) of all fits.

David Lawrence 3/10/2017 8:47 AM
Deleted: error

5

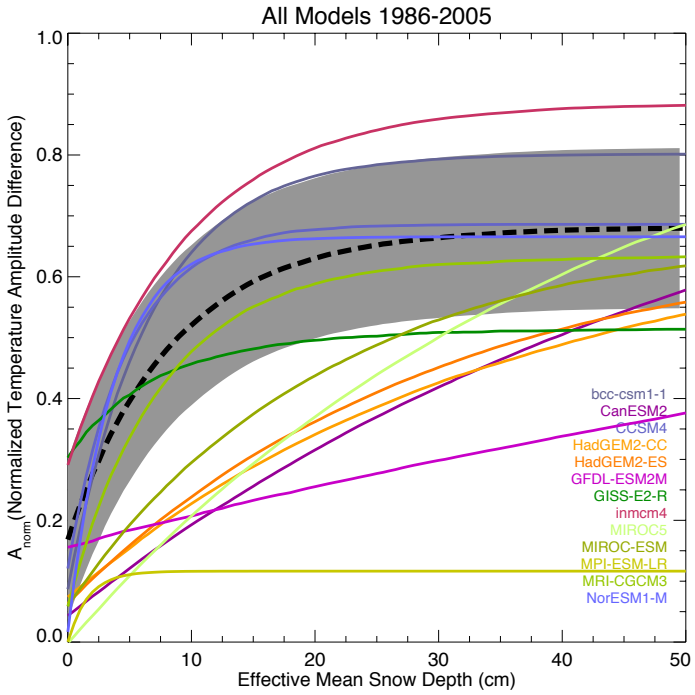


Figure 4: Model fits to equation (5), showing large differences in how heat is transferred though the snowpack to the soil. Data from the CMIP5 comparison period 1986-2005 is used. The dashed black line is the observational fit with grey shading representing the observational scatter as in Figure 3.

David Lawrence 3/10/2017 9:23 AM
Deleted: being its error

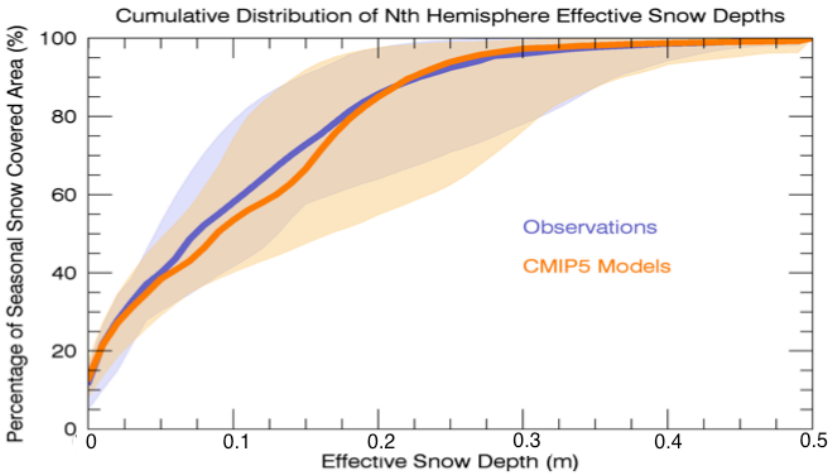


Figure 5: Cumulative distribution curves for climatological effective snow depths ($S_{depth,eff}$) as a percentage of all seasonally snow covered area in the Northern Hemisphere. The observed and model climatologies span a minimum of 10 years in the last quarter of the 20th century. The bold lines show the median values of the curves for both observations and models; shading shows the total range of curves. Note that the absolute snow covered area can be quite different amongst models and observational estimates.

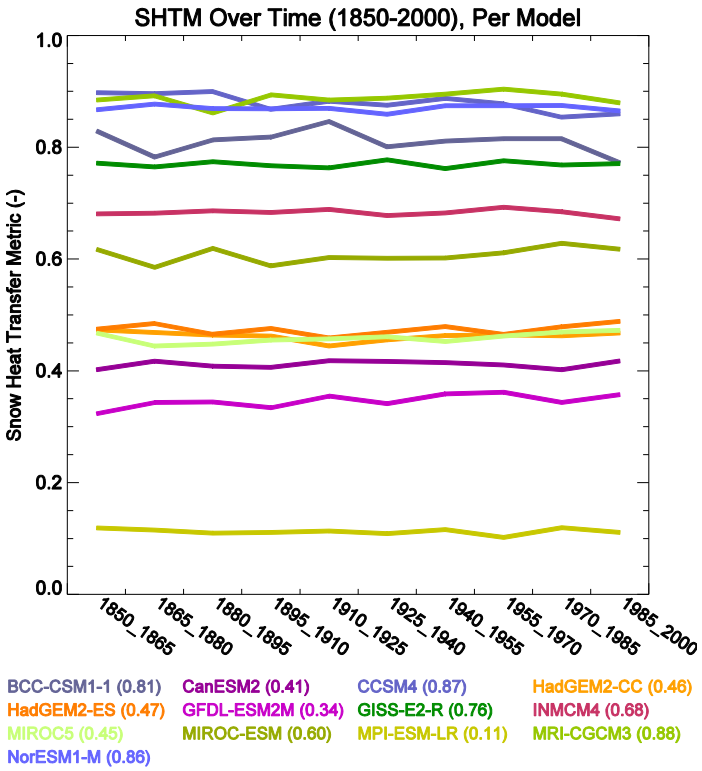


Figure 6: Mean value of the Snow and Heat Transfer Metric (SHTM) for each model. Values closer to 1.0 indicate better agreement with observations. SHTM values are relatively stable over 150 years of climate for each model.