The Cryosphere Discuss., 7, C3560–C3577, 2014
www.the-cryosphere-discuss.net/7/C3560/2014/

The Cryosphere
Discussions
Open Access

Interactive comment on "Large area land surface simulations in heterogeneous terrain driven by global datasets: application to mountain permafrost" *by* J. Fiddes et al.

J. Fiddes et al.

joel.fiddes@geo.uzh.ch

Received and published: 19 June 2014

AUTHORS REPLY TO REFEREE #1

We thank Anonymous Referee #1 for the effort in evaluating and commenting this manuscript. In the following text the referee's comments are marked "RC" and author comments "AC". All RC references refer to original manuscript. Comments have been subdivided (a,b,c,..) where appropriate. A corrected manuscript with additions in red is included as attachment to aid cross-referencing.

In summary, we agree that we can improve the clarity of the purpose of this manuscript. This recognition has been held prominent in completing revisions. In sum, this contri-

bution is primarily intended to showcase a new permafrost modelling approach that we believe has much wider applicability.

GENERAL COMMENTS

RC1: I am not completely happy with using the term Land-surface model for the GEOtop model used in this paper - this suggests a global or hemispheric application of the model, which is as far as I understand not the intended use of the model. On the other hand, if LSM are the potential target of the schemes presented in this study, why not using a real LSM to show and analyse their usefulness ?

AC1: GEOtop simulates energy fluxes at the land surface and we expect that other LSM such as Surfex could be used instead. We acknowledge that the term LSM has often been used in a context of global applicability and climate modeling, eg. p.5858 l.19 "It should be noted that this model is not an LSM in the conventional sense (e.g. Mosaic, CLM, NOAH, Koster and Suarez, 1992; Dai et al., 2003)", and have thus added p.5858 l.21: "In addition this model has not been designed for global or hemispheric application."

RC2: - a real application is missing: Figs. 6, 7 & 8 are not useful in this form/resolution (see below). One gets the impression, that the model was developed, but no time was left for real testing and applications. What do you intend to do with the model in future? You present the results as a test application but without real in-depth analysis or discussion of the permafrost distribution. It becomes not clear whether the permafrost application was one of your real aims or whether it is "just" an example and you plan to do other different examples in the future. This is important as you rely heavily on existing permafrost-relevant validation data (high-resolution GST data), which would be difficult to get for other applications. This model/data set seems to be tailored to your application, but you do not make use of it...

AC2: We have expanded the introduction to clarify the purpose of this study as follows:

“The main aim of this study is to establish this combined method as a proof-of-concept and perform an initial evaluation of its performance in the context of the ground thermal regime and specifically permafrost occurance in the European Alps, as a test case. That said, the aim is not to provide a best-possible result for e.g. permafrost (as the example subject of this study) but to provide a demonstration of this method using simple datasets. It is well known that precipitation bias is a common problem when using climate model or reanalysis data (e.g. Dai, 2006; Boberg et al., 2008) and a key driver of the energy and mass balance at the land surface. Therefore, an additional aim is to explore a simple method that may be useful in addressing precipitation bias using the parameter melt date (MD) of the snowpack. Specifically this manuscript will:

1. Conduct a test application of the combined schemes together with the LSM GEOtop to derive land surface / near-surface variables air temperature (TAIR), ground surface temperature (GST) and snow depth (SD) over a large area of the European Alps at a resolution of 30m , and additionally a derived permafrost estimate. 2. Evaluate the performance of the combined schemes against a large network of TAIR, GST and SD measurements in the Swiss Alps. 3. Demonstrate a simple bias correction method for the precipitation field. 4. Interpret results together with uncertainties in the model chain.”

RC3: - Figure 6: the plots are too small and details about the abbreviations in the legend, the location of the data point and its representativeness are missing. Too short explanation in the text (page 5864/5865 top): the reader is left alone to understand what Figure 6 is really showing. - Figure 7: why only a "visual comparison" ? What do we learn from this single comparison ? Where is it ? What is shown ? What is teh colour code ?

AC3: Figure 6(a) and (b) give all IMIS (64) stations which are used to give SD ground truth, not just 1 point (c) is just an illustration at one point to show how the method works. We have reformatted this figure to include the effect of the correction on GST as discussed in AC1 of Review 2, and decided to drop the original third panel for

formatting reasons. We have expanded the caption to improve clarity:

“Snow depth (a) without and (b) with snow correction method evaluated at all available IMIS stations (64). GST (c) without and (d) with snow correction method evaluated at the same stations.”

These changes are also reflected in the main text:

"Figure 6 shows the effect of the snow bias correction factor on MASD and MAGST at valid IMIS stations (64). Figure 6(a) shows the large bias in precipitation inputs, particularly at sites with large snow accumulations (i.e MASD >50 cm). The effect of the snow correction method in successfully reproducing the spatial differentiation of precipitation quantities is shown in Fig. 6(b) by greatly reducing error and bias. Figure 6(c,d) show the improvement in MAGST as a result of the correction. The observed bias in Figure 6(d) demonstrates how even by fitting the melt-date, snow depths can be underestimated. This could be due to the fact that SWE is reproduced accurately but parameters governing density of the snow-pack or wind erosion are not correct. However, without SWE evaluation data at these stations, this is difficult to confirm."

Please note the original figure was mistakenly computed with an older version of Geotop which had an error in the energy-balance making the model too cold. This explains why a negative bias in MASD is now present in the figure 6 (d). As we fit the melt-date of the snow-pack this bias is possibly due to how densification is parameterised in the model. For example, it is possible for total SWE to be correct but snow depth to be underestimated by the model. These effects, however, were already explained in Section 5.2. Despite this model error the overall conclusion of the correction method remains valid.

Main text has been expanded as described in AC11, but the key message is that We have addressed the issue of precipitation inputs in this way as we see it as a primary source of uncertainty in this approach. The comparison in Figure 7 is discussed in AC/RC24.

RC4: conclusion: you should only mention those aspects in the Conclusion which you really assessed/discussed within the paper using your results. This is not always the case.

AC4: Done, addressed in A45, below.

DETAILED COMMENTS

RC1: p.5857: l15-17: "...we have an optional informed-scaling training routine, which regresses model results against input predictors after a training run in order to adjust the weighting of each input according to its significance, and in doing so improve the quality of the final results." –> this is much too vague, even if it is describing the method. But without some physical processes and/or explanations which variables/parameters are optimised, this is a meaningless statement. You can train everything, even if there is no meaningful relation between input and output! Please give more details.

AC1: p.5857: l.2-3: We have stated that further details are available in respective publication. It would be too bulky and repetitive to reproduce these methods in detail in this publication. The basic understanding of the manuscript does not suffer with these omissions.

RC2: l21: "a "sub-grid aware" aggregation...": nice term, but again too unspecific/vague..."aware" could mean anything in this context.

AC2: Changed to "...(3) efficient aggregation of simulated variables to coarse grids and...".

RC3: l22: "with fine-scale ground truth": which type of ground truth do you refer to here? Is that realistic that this kind of ground truth would be usually available in the context of your present paper?

AC3: Changed to "...(4) compare results and ground truth derived from similar scales." The scaling issue is now discussed more explicitly in the introduction: "Gubler2011 have shown that fine-scale variability of surface processes can be high in complex ter-

rain e.g. variation in soil moisture, ground cover, local shading can cause differences of as much as 3°C MAGST within a 10 m x 10 m grid. This underscores the importance of scale appropriate evaluation of models. There are many studies in the literature where models operating on grids of 10's –100's and in extreme cases, 1000's of metres are evaluated by point-scale measurements and this is known to pose a serious challenges to model evaluation (Randall 2003, Li 2005}. However, methods that provide simulation results over large areas capable of exploiting distributed site-scale ground-truth, are rare..

RC4: p.5858: l12: "The final output is the full set of scaled fluxes...": do you really mean (only) "fluxes" here ? or time series of meteorological variables, from which the fluxes can be calculated ?

AC4: We have changed to the recommended "time-series of meteorological variables".

RC5: l15-16: "...e.g. commonly used lapse rates or parameterisations": again too unspecific –> what are "commonly used parameterisations" ? That depends a lot on the model, temporal and spatial scales, processes included etc.

AC5: Changed to: "... to reference methods such as fixed lapse rates." More detail is readily available from the publication cited.

RC6: l21: "it should be noted that this model is not an LSM in the conventional sense...": exactly, so why do you use this term for it ? This is misleading. I suggest to use the model name thorughout the text and make the connection to LSM's only where it is appropriate. (see also general comments)

AC6: see AC1.

RC7: l27: "...as well as freezing and thawing processes IN THE GROUND".

AC7: Done.

RC8: p.5859: l1: "...ground temperatures." –> do you mean subsurface temperature or ground surface temperature (GST) ?

AC8: Clarified in text: “...ground temperatures both at the surface and at depth.”

RC9: l1-2: for what purposes do you use this model in this application ?

RC9: This is clarified in the introduction: “conduct a test application of the combined schemes together with the LSM GEOtop to derive land surface variables air temperature (TAIR), ground surface temperature (GST) and snow depth (SD) over a large area of the European Alps at a resolution of 30 m, and additionally a derived an estimate of permafrost area.”.

RC10: l3-4: "..details specific to experiments...": unclear at this point...what experiments ?

AC10: Re-phrased sentence to: “Further details specifically relevant to this study are given in Section 4: Simulation experiments.”

RC11: p.5860: l1-2: Is this really possible for the "typical" LSM application ? Usually a data set as used in Schmid et al. 2012 will not be available, that is one has no way of knowing the correct melt date (MD) over a large area. Are you aiming in your paper at a spatially high-resolution case study at your field site (i-buttons or potentially PERMOS, where a lot of GST data are available) or at a generally applicable approach as stated in the introduction ? For the latter, I do not see how this snow correction method could be applied except by using remote sensing. But this, you did not treat in your study!

AC11: We have expanded this as stated above to explain further how/why this is intended to be used. The outlook is using remote sensing. It does not matter too much how one establishes a melt-date (GST or remote sensing), we aim to show its utility in providing a correction to a significantly uncertain, yet key input (precipitation). Text has been changed as follows:

p5856: “In addition it is well known that precipitation bias is a common problem with using climate model or reanalysis data (e.g. Dai, 2006; Boberg et al., 2008). Therefore,

an additional aim is to explore a simple method that may be useful in addressing precipitation bias using the parameter melt date (MD) of the snowpack.” p.5860 l.6: “The approach shown here could potentially be used together with satellite imagery in order to estimate snow fall bias based on MD. However, this manuscript evaluates the point-based performance of the new method without bias correction.” p5866 : "Two notes of caution with respect to this method are, (a) it is only valid currently at site scale, and (b) it relies on GST measurements. However, the approach shown here could potentially be used together with satellite imagery in order to estimate snowpack bias based on MD to enable scalability of the method. However this is beyond the scope of this manuscript."

RC12: l2-3: that would be a spatially distributed correction factor if you simulate large areas. How did you extra-/intrapolate that to the full grid in case of data sparsity (see comment above) ?

AC12: Remember this was not done in this study and the method was only presented in results of Figure 6. This has been clarified in the text as stated in AC11.

RC13: l19: "4-D-VAR assimilation scheme"

AC13: Changed.

RC14: p.5861: l5-8: "Landcover was derived..." –> that is a critical step for surface/subsurface studies in the Alps and should be explained here in more detail! It would also be good to get some indication of the validation and the corresponding uncertainty of these datasets.

AC14: The datasets are fully described together with uncertainty in reference given. We have added (l.7) “full details together with description of uncertainty are given in Boeckli et al. 2012a.”, to make this clear. In addition we have altered the introduction to reflect that this manuscript aims to demonstrate a method in a proof of concept fashion as opposed to generating a new result such as a permafrost map, as described in AC2.

RC15: l19: GST at IMIS stations: is that measured in the same way as for the other data sets ?

AC15: Changed to: “The dataset used covers years 1996–2011; GST is measured with a white temperature probe resting on the ground surface.” For completeness, the subsequent section has been modified to (p.5862 l.3): “Sensors measure GST a few cm below the terrain surface to avoid radiation effects.” The differences based on instrumentation are implicitly acknowledged by introducing data as discrete datasets.

RC16: p.5862: l.18-19: was attention paid to the so-called zero-curtain phase (non-changing surface/subsurface temperature data at the freezing point due to freezing/thawing)?

AC16: This criterion was only applied to wind direction measurements. Sentence change to: “Non-changing values beyond prescribed time limits were screened from wind direction data.”

RC17: l.20-23: did you validate only GST values ? Is that generally useful or just the specific focus of your application ?

AC17: No, we looked at TAIR and SD as well. GST is considered a generally important variable as it is a synoptic value of the surface energy balance.

RC18: p.5863: l6: "(10 times, 1979-1983 period)" –> why this spinup set-up ? is there a specific reason to use this period ?

AC18: Clarified: “Spin-up is performed over 50 yr (10 times 1979–1983). This is necessary to obtain soil temperatures at depth that reflect atmospheric conditions and are independent of their initial value.”

RC19: l8: "defined in Gubler et al. 2013": as far as I understood this study aimed at analysing a very specific region, which was also used as input data in the present study. But these data concern a very small region; would that not give a bias towards a good validation in exactly this region, but not necessarily a good performance within

all the other areas in the present study ?

AC19: The range of model parameter values in Gubler el al. (2013) were defined as the physically plausible range as based on literature review. As such there is no geographic connection. The Gubler et al. study was performed on artificial topographies and therefore also was an abstraction of reality – it aimed to test physically based model sensitivities and uncertainties, independent of location.

RC20: l8-10: mean annual values: but as you calculate also the daily values, why not validating them as well, at least some statistics of them ? At least if you really aim at some general applications as written in the introduction. Else, focus on your application (I would prefer that) and reduce the "generalisation focus" of the introduction.

AC20: We have now presented the study as application of our method and reduced generic focus in the introduction. We suggest that this approach may have more widespread use in the outlook only. For example this sentence is moved to conclusion:

“All inputs are derived from global datasets, suggesting that consistent application globally in heterogeneous and/or remote terrain is possible.”

RC21: p.5864: l10: Is mean annual snow depth really a good validation variable here ? As you mentioned yourself earlier, melt out and/or number of snow days are much more relevant and meaningful in this context. A strong under/overestimation in snow height changes MASD but may not be important for frozen ground, as along as MD is correct.

AC21: Both variables are strongly correlated and, given the strong bias, either is sufficient to reveal it.

RC22: l20: how do you know it’s precipitation input and not snow drift/parameterisation of snow metamorphism etc ?

AC22: The precipitation bias has been shown in Fiddes and Gruber (2014). Other processes mentioned by reviewer are likely to be a factor however what we see is a general bias affecting most sites irrespective of topographical situation.

RC23: l26: "...snow depths can be underestimated...": is there a discussion of the associated problems following somewhere ? performance depends on teh snow regime: if large, no problems associated, if small there are potential biases in winter through additional cooling of the subsurface

AC23: Additional text in Discussion p5866 l.24: “Bias associated with snow-based precipitation may have strong impacts on the ground thermal regime due to the thermal properties of the snowpack, duration of snowpack or even cooling effects of very shallow snowpacks where the albedo effect may dominate.”

RC24: p5865: l8-10: "Comparison of methods is only intended for...": so what is the aim of this rather arbitrary and purely visual comparison ? Why not comparing against ground truth ? Both models are based on temperature (GST, MAAT) estimates which are partly (Boeckli) calibrated with specific ground truth data...again with a bias towards the same data sets (PERMOS, ibuttons)...so what do you exactly learn from this comparison ?

AC24: Figure 4 is the closest to a comparison with “ground truth”, although GST is a valuable variable, it is not permafrost. The comparison in Figure 7 is intended to establish what one may call “face validity” (Rykiel 1996), that is demonstrate that similar spatial patterns result from this method as from established and trusted models. Clarified to: “Comparison of model results, despite differences in the definition of permafrost area and in observation periods, is intended to demonstrate the similarity of patterns resulting from both approaches (cf. “face validity”, Rykiel 1996).” In addition, approximate scale bars for distance, permafrost index and permafrost presence added to aid comparison. Caption has been modified.

RC25: l20-26: the error magnitudes are difficult to see in Fig. 8 - in this resolution the only thing one can clearly distinguish are the colours/sign of the bias.

AC25: We have reformatted figure to increase the size.

RC26: l26-27: "...fit magnitudes of precipitation...": do you have a reference for that ? L27: "...first stations...": what do you mean by "first" ?

AC26: Precipitation patterns reference added (Frei and Schaer 1998) and sentence changed to: “i.e. greater north and south of the main alpine ridge and less in inner-alpine regions”. Removed “first”, now reads: “However, stations on the north-slope of the main Alpine chain appear to be modelled well.”

RC27: p5866: l13-14: "The exception being precipitation.": not clear what you mean

AC27: Error in text, removed this sentence.

RC28: l16: "climate models": do you mean atmospheric models ?

AC28: Changed to “atmospheric” models.

RC29: l19-20: "...and (b) it relies on GST...": if you have already GST measurements: do you really need the model to predict whether there is permafrost at these places?

AC29: These are two separate issues with GST measurements being used to demonstrate a potential correction method as discussed in AC11. Permafrost/GST results do not rely on these measurements.

RC30: 5867: l10-12: "...In addition there are advantages of the gridded ERA dataset over interpolated station data...": too vague, please explain! There is also a new high-resolution gridded data set for Switzerland for air temperature and precipitation (MeteoSuisse) , which you may want to mention/use.

AC30: (a) Expanded sentence reads p5867: l10-12:

“ In addition there are possibly advantages in the gridded ERA dataset over interpolated station data in terms of representing larger-scale, synoptic conditions.” (b) I think the point is missed slightly here, the central purpose is that we use a globally available source of meteorological data that is 'quasi-physically' downscaled based on atmospheric pressure levels, i.e. p5856 l.6:

"(ii) a method that scales gridded climate data necessary to drive an LSM, to the subgrid using atmospheric profiles. The philosophy behind these approaches is to develop methods that depend only on globally available datasets to derive local simulations in heterogeneous and/or remote regions." Assuming the reviewer refers to the gridded MeteoSwiss products based on station data, we identify the following limitations: (a) national level only (b) coarse resolution (c) statistical, based on interpolation of measurements (d) do not provide parameters such as downwelling longwave.

RC31: l24ff: "Landcover could however...": not clear to me: it seems to me a VERY IMPORTANT source of uncertainty ?! Especially in this kind of heterogeneous terrain ! If you take the results from Gubler et al. with GST variations of several degrees over a few metres, how do you get something reliable out of your scheme which is more than a typical MAAT-based estimate without this type of GST ground truth data as input data, like in your study ?

AC31: Again that this is a proof-of-concept method demonstration together with associated limitations (see AC2). In the future, we may (a) have better data in an application study, and (b) propagate remaining uncertainties as in Gubler et al. 2013. More specifically, in this study:

(a) We define a coarse debris, bedrock and vegetated surface. Which are prescribed to each sample based on the modal value. An alternative approach would be to create samples using landcover as a predictor alongside topographic predictors. We present a simple first working example here so have reduced the complexity somewhat. In addition, topography is a significant predictor of surface cover in its own right so this is in any case likely to be unnecessary. We purely use the parameter values for different sub/surface material from Gubler et al. (2013) Which are standard values from the literature. Not a result as such. Again you could go a lot further than this and have

detailed surface/subsurface mapping – but the aim is to have a simple, robust approach that can be applied with global datasets (ASTER DEM / LANDSAT SAVI) and does not rely on more detailed knowledge in this first demonstration of the method. (b) Again, we do not use GST as input. It is used in the presented bias correction scheme but not implemented in the main model results. (c) We think we go a long way beyond a MAAT based estimate. We present an approach which enables us to simulate a range of land surface variables not just GST (although that is the subject of this paper). The aim is not to present the best method to obtain a Permafrost estimate for the Alps, but a method which enable land surface simulation at any given point on the Earth's surface – with particular focus on remote/ ungauged areas – we believe that to be the value in the #scheme#.

RC32: 5868: l14-16: I do not understand: why/where do you have clay silt and peat sand in the mountains ? and there are many other process uncertainties on subgrid level, e.g. the effect of large boulders at the surface which are very important regarding the permafrost distribution

AC32: "clay, silt and rock" and "peat sand and gravel" are summary terms to report an example study.

RC33: l19-22: ok, but then you should do one example application yourself and show that it is feasible in terms of these uncertainty analyses...

AC33: This is beyond the scope of this work. This would not be a small addition to this manuscript, but essentially a whole now study – or likely several.

RC34: 5869: l2-3: unclear to me what you mean

AC34: Reworded for clarity:

"In this study the PERMOS dataset is point-scale in both measurements and topographic properties upon which modelled results are based, as these properties have been measured locally and not extracted from the DEM."

RC35: l7-11: yes exactly, see my comments above

AC35: answered above.

RC36: l15: "driving fields"

AC36: Corrected.

RC37: l17-18: "like the Mattertal": why do you point out this psecific example if you do not focus on regions except the Engadine in the reminder of the paper ? The focus of the study is not clear to me.

AC37: It is just an example of a topographic precipitation barrier that exists in Switzerland.

RC38: l24-28: operational weather forecast models can still not simulate this type of sub-grid processes in mountain terrain...

AC38: no change required to our text.

RC39: l28: new paragraph after "...few years."

AC39: Done.

RC40: 5870: l17ff: Do you see evidences of these limitations in your validation experiments or are they "just" examples of a list of potential uncertainties/limitations without knowing the relative importance of them ? If you have any evidences (e.g. for the cold bias or the snow deficit) this would be useful to mention here.

AC40: We do mention the PERMOS2 results here and limitations we have observed with the 1-D approach, but we do not go so far as to conclusively show that differences are caused indeed by one or the other process as this would lead us into the testing and evaluation of GEOtop. This section is intended to identify limitations in the modelling procedure such as omitted physical processes. These do not require evidence as the effect of omission in model procedure is well known. We have strengthened the

number of references to make the argumentation more solid and extended the text in this section as follows: “Missing or inadequately described physical processes is a well known and common characteristic of most physical models (Arneth et al. 2012, Beven 1995), however, as testing of the physical model GOEtop is not the focus of this study we provide limited discussion on this topic.”

RC41: 5871: l4: "... as a test case.": I am not so sure you could really call it a test case. You present it as a test application but without real in-depth analysis or discussion of the results. It becomes not clear whether the permafrost application was one of your real aims or whether it is "just" an example and you plan to do other different examples in the future. This is important as you rely heavily on existing permafrost-relevant validation data (high-resolution GST data), which are difficult to get for other applications.

AC41: This point is addressed by point 2 of General comments.

RC42: l5-6: a word is missing towards the end of the sentence ?

AC42: Restructured sentence:

“However, the scheme is generic in that it is able to generate surface fields of any variable the LSM simulates.”

RC43: l12: "...that consider significant uncertainties in the model chain...": but this is at the cost of having introduced additional uncertainties due to the simplified approach ?

AC43: Yes. The new approach provides more flexibility as to choosing where computing power would bring the most benefit. The scenario we propose here is to accept some additional uncertainty through e.g., the lumped approach, but then to be able to estimate the uncertainty due to uncertain sub-grid patterns of snow redistribution (“would there be permafrost here in a wind-swept location?”).

RC44: l13-17: please rephrase this sentence: it is quite vague and difficult to understand

AC44: Rephrased sentence: “Such scenarios could be interpreted together with site specific knowledge to provide an improved quality of result, or a range of outcomes to be planned for in terms of uncertainty related to future conditions or other unknowns.”

RC45: l24: "due to biases in driving data": you did not show this in your analysis

AC45: TRUE – removed statement.

RC46: 5872: l6-7: "...such as changing sub-surface material properties...": this depends a lot on the subsurface properties which are generally only known at site level! You did not address or discuss that in your paper

AC46: We add confusion by using the term “sub-surface material properties”. We do not mean the medium that defines model parameters (eg. debris or bedrock). We have changed this term to “sub-surface properties” and is further clarified in the text by the example of ground ice-loss. The point here is that we are not limited to producing 2D maps with this approach. There is possibility to produce 4D results ie. 3 spatial dimensions of surface, xy to depth, z and through time, T.

RC47: l15-20: repetition to above

AC47: Addressed above.

RC48: 5877: Table 1: why do you give the values to the third decimal ? are these averaged values of the evaluation data set ? is it necessary in this context ?

AC48: These are parameter values that define our surface types in the model (GEOtop). They are quite sensitive hence high level of precision. These are generic values of natural materials taken from the literature. We have expanded caption to make this more clear:

“Description of surface and sub-surface parameters used in this study. These are generic values of natural materials obtained from the literature.”

Interactive comment on The Cryosphere Discuss., 7, 5853, 2013.