The Cryosphere Discuss., 5, C1389–C1422, 2011
www.the-cryosphere-discuss.net/5/C1389/2011/

The Cryosphere Discussions

Interactive comment on "A statistical permafrost distribution model for the European Alps" *by* L. Boeckli et al.

L. Boeckli et al.

lorenz.boeckli@geo.uzh.ch

Received and published: 30 November 2011

Author's comments (AC)

We thank reviewer 1 for the valuable comments, which helped to improve the manuscript, and have addressed all of them in the revised version. Our point-by-point response to all comments and the resulting changes in the manuscript are described below.

The main changes of the manuscript include the sections "Introduction", "Background" and "Conclusion", which have been restructured and partly rewritten. Further, the former section 3 "Data" was placed after the section "Statistical Method". To address

the main comment, i.e. that the model was not actually applied in this contribution a sample map showing permafrost probabilities is presented in an additional figure. It is further emphasized more clearly that additional calibration and validation steps based on external data sources are necessary prior to presenting a final alpine-wide permafrost map, which will be presented in an additional paper.

General comments

1) The title of the article could be more precise to represent the content of this work and to avoid a misunderstanding as the reader would expect a permafrost distribution map as result. For the MS in its current stage an adequate title would be "A statistical approach to model the permafrost distribution in the European Alps" . . .

AC: We changed the title of the article accordingly: "A statistical approach to modelling permafrost distribution in the European Alps or similar mountain ranges."

Further the authors write to focus on the analysis of the explanatory variables, the development of statistical sub-models and their combination. . . . I expected firm scientific findings and aims that top the scientific request of the applied statistical methods. For the reviewer the scientific findings of this study are sparse and are not even stated in the conclusion. The formulation and the theory of the models and their real combination should be the focus of such a study. However, a potential scientific aim would be the estimation of the permafrost distribution in the European Alps. This includes the assessment of the model sensitivity with regard to uncertainties of the model input parameters (standard deviation and spatial resolution!).

AC: The objective of our contribution is to introduce a modelling approach designed to cater to the specific needs of permafrost distribution models for entire mountain regions. This objective is now expressed more clearly in section "Introduction" of the revised manuscript, and the conclusions better underline the importance of our con-

tribution with respect to this. We acknowledge that the original submission did not emphasize this contribution clearly enough. The modelling of permafrost distribution, however, (a) is complicated in a fundamental way by scaling issues and sparse data that only partly covers the diversity of relevant environmental conditions and (b) usually performed for smaller spatial extent, resulting in environmental gradients that may not suffice to detect certain patterns. In this context, we present (and we have formulated this more clearly in the revised version) three main results: (1) a suitable strategy and method for modelling and tackling the challenges named above under a); (2) the results of the statistical analysis, i.e. model parameters, some of which reveal surprising patterns; and (3) the insight that such models cannot be directly applied to an entire landscape but require later subjective adjustments.

In order to illustrate our modelling approach, an additional figure (Fig. 1) is added, which shows an example of a map based on the statistical model presented. The actual application of the model and the derivation of the required subjective adjustments are presented in detail in a companion article (Boeckli et al, to be submitted). A combined manuscript would be too long and the scope of the two papers, both in terms of methods and geographic transferability, is different.

We rewrote the conclusions to better reflect these issues (see also general comment 8). An assessment of the model sensitivity regarding the uncertainties of the model input parameter is presented in Boeckli et al. (to be submitted).

2) A concept is missing. This makes it difficult to follow the MS. I suggest to rename section 2 (Background) as section 1 (Introduction) and to move most of the text from section 1 (Introduction) to a new section 2 (Concept). The section "concept" has to explain which models are used (debris-, rock-), which method is applied (GLMM/nonlinear probability model, linear regression/linear probability model), and which is the outcome variable (binary variable, continuous variable). Further it has to define the statistical terminology (response variable, explanatory variable), to explain why scaling issues

are necessary, to introduce the methodology how the model is parameterized and calibrated, and to declare how the model output is tested.

AC: We changed the name of section 2 to "Conceptual background" and moved much of the text of section 1 (introduction) to this section 2. To do so, we partly rewrote section 2 and addressed all points listed by the reviewer. However, we avoid introducing general concepts such as response variable or explanatory variable, which we believe are known to most readers of The Cryosphere. Model evaluation is explained in section 3.5; we do not refer to this step in the earlier sections that introduce the conceptual background, because in this particular study the model evaluation is not part of the conceptual framework required for formulating the model.

3) References and theory of the statistical methods: In the current stage the MS include no information on the Generalized Linear Model, the Generalized Mixed-Effect Model (Dobson 2001, Crawley 2005), and contain only one reference for the Probit Model. To make this work comprehensible for the reader it would be necessary to establish an appendix where the major approach/theory of both the GLMM /GLM and the Probit Model is explained.

AC: In our opinion an appendix of the statistical methods that were used is not necessary because there is an ample literature that describes these methods. Some of the less mathematical key publications introducing these methods and the software implementation used here are referenced in our contribution (e.g., Hosmer and Lemeshow, 2000; Venables and Ripley, 2002). One recent publication that introduces the probit model and compares it to the logistic model has been added (Gelman and Hill, 2007). An additional reference has been included for cross-validation as well (Hand, 1997). On the other hand, we agree that the statistical methods used should be mentioned earlier in the text, which is now the case (Conceptual Background).

Regarding the use of penalized quasi-likelihoods (PQL), explaining this statistical *quantity would go beyond the scope of this contribution. The statistical methods used (GLM, GLMM) are used as tools in this work, and further information is available in the references provided. PQLs are not directly and/or easily comparable to the ordinary least squares approach, which is also not applicable in the context of binary response variables. In order to avoid further confusion, we omit the reference to PQL in the revised paper because the mentioned software implementation does not provide other means (such as direct estimation with the maximum likelihood method) other than PQL.*

4) Debris model: Can you show that the same input parameters (MAAT, PISR, PRECIP) that discriminate intact from active rock glaciers can discriminate permafrost from non-permafrost sites that were found by other methods as the rock glacier inventory (see methods in next sentence)? I do not understand why the authors neglect important information that is provided by other methods such as borehole temperature (BH), ground surface temperature (GST), geophysical prospecting (GP), other indirect evidence (OIE), and maybe surface movements (SM) (Cremonese et al., 2011).

AC: To validate the discrimination of the debris model in other areas than rock glaciers is challenging, because there are hardly any observations available and the existing are strongly biased towards permafrost existence (e.g., boreholes that were drilled for permafrost research at locations with previously known permafrost presence). We are aware of the difficulty in transferring knowledge from rock glaciers to other types of terrain (cf. first and last bullet of conclusions) and this is the focus of the mentioned companion paper. However, current data does not permit extending the statistical analyses to other types of surface cover, as a large number of observations are required. This is a fundamental challenge to all statistical permafrost distribution modelling, and equally to the validation of physically–based numerical models. To clarify this, we added the following: "The other permafrost observations described by Cremonese et al. (2011) were not used for model calibration because, a) they are not sufficient in number to allow consistent statistical analysis; b) the integration and

homogenization of heterogeneous permafrost observations are subject to large uncertainty and subjectivity, and c) observations are strongly biased towards permafrost existence and less observation in non-permafrost conditions are available." Where possible, these data will be used for model evaluation, but this requires a more detailed examination of the data (Boeckli et al., to be submitted), which is beyond the scope of the present study.

5) Rock model: The established rock model might be based on a too sophisticated approach to satisfy the main objective of this study. In general, the reviewer favours an approach that avoids the need of scaling issues. The simplest approach would only consider the parameters ELEVATION, EXPOSTION, and maybe PRECIP.

AC: In general, the rock model is very simple: It uses MAAT and PISR as explanatory variables only. The scaling issue is necessary because PISR values that were used for model calibration differs from PISR values that are available for model prediction. The advantage of the described approach is that the rock model can be applied to every DEM, accounting for the corresponding grid size. We did not use elevation because the elevation of the 0°C isotherm is varying spatially in the Alps. Further, we used PISR instead of slope aspect because the slope angle affects the incoming solar radiation, which would result in the need to include an interaction term. And finally, aspect would have to be decomposed into, e.g., a sine and cosine component (as in Brenning and Trombotto, 2006) in order to use it in a meaningful way in a regression model. The "simplest approach" proposed by the reviewer would therefore effectively be more complex – i.e., involve more explanatory variables – than the approach chosen for this study.

6) Model combination and scaling issue: For the reviewer it seems that the two independent models are simply connected to together and that no conjoint parameter is found that is inverted from the combined model. Thus I deduce that the model combi-

nation is not obligatory and that this would easily allow to use standard methods for the interpolation. This methods could either be geometrical (Minimum Curvature Gridding) or geostatistical (Kriging)

AC: The two sub-models can indeed be applied separately. The offset terms delta_r and delta_d are not derived from each model but will be specified at a later stage based on published research and expert knowledge. These two terms express systematic differences between the observed proxy (e.g., rock glacier activity) and the predicted variable (permafrost presence), reflecting comment 4 of reviewer 1.

Interpolation methods would not be useful for our work because an extremely high point density of logger measurements across the entire Alps would be needed for this. On the other hand, there is also not much to be gained from using Universal Kriging not only because the necessary measurement density cannot be achieved but also because the linear regression model already explains more than 80% of the variance in our sample. As for the downscaling, geostatistical methods have been proposed in order to address the change of support problem, but precise knowledge of the residual semivariogram of our regression model over sub-pixel distances would be necessary in order to make operational use of these techniques. We make general reference to the geostatistical change of support problem in section 4.4 of the original submission, including a key reference in this field: Gotway and Young, 2002.

Overall, we are well aware of viable alternative approaches and chose the proposed approach very consciously. Linear regression and GLMM are the "natural" model choices for the data available to fit a rock model and a debris model, respectively, and the proposed approach for combining these apparently incompatible models is a necessary innovation.

7) Discussion - use and limitation of the model: I recommend that the authors rewrite this section as the current text is not adequate for a scientific discussion (e.g PSIR

calculated from a finer DEM is more accurate than from a courser one). The authors should discuss the impact of the adjustment offsets "delta" on the model result (state values) or to asses, how accurate a DEM must be for the two individual models to obtain a satisfying result for the prediction of the permafrost distribution.

AC: It is important to state that PISR estimates derived from ASTER GDEM are suboptimal to base the rock model on because this is not necessarily the case for the debris model. The impact of the adjustment term "delta" is not discussed here, because it was only treated as a theoretical concept so far. The determination as well as the impact is discussed in Boeckli et al. (to be submitted).

8) Conclusion: This section reads such as a summary and lacks in scientific findings. For example, conclude which parameters are the most important in your model (e.g the model calibration), how accurate they must be (spatial resolution and the absolute value itself) and how sensitive their respond is to the predicted probability. The listing of steps that are needed to use the established approach for a map based product is rather an outlook as a conclusion, are obvious, and should not be a part of the conclusion.

AC: We agree with reviewer 1 and partly rewrote the conclusion (cf. comment 1). The listing of steps needed to transform this into a map is important to caution the reader who may be inclined to directly apply this in a further analysis. We have tried to reformulate this aspect more clearly that also directly relates to comment 4.

Particular comments

For all particular comments of reviewer 1 not specifically addressed below, the manuscript has been changed following his suggestions.

Abstract: 12: I suggest using model specific terminology such as "explanatory variable" later. You can write "parameters" or "input parameters" here.

AC: The terminology "explanatory variable" was kept because this term is common in statistical regression modelling (e.g. Crawley 2009) and the term is used in permafrost related literature (e.g. Etzelmueller et al., 2007; Lewkowicz and Bonnaventure, 2008). In statistics, "parameter" cannot be used for "explanatory variable", and its use in this context would be misleading.

Abstract 15: Delete the parenthesis - this is too much details.

AC: We have partly re-written the abstract to reflect better what the main results of this research are and this requires a certain level of detail, also with respect to model coefficients or quality.

Abstract 17: Is the rock model also a GLMM?

AC: The rock model is a linear regression (in the original submission: Abstract line 19, and p. 1429 l. 16). To make this clearer, the sentence was changed to: "The rock model is based on a linear regression and was calibrated with mean annual rock surface temperatures (MARST). The explanatory variables are MAAT and PISR."

Abstract 19: The root mean square error (RMS) is well known. Thus you can write "..a RMS of.."

AC: In our opinion, the acronym RMS as suggested by reviewer 1 is not explicit and needs to be established as well as the term "RMSE". We prefer to use the term "RMSE" as it is used in the software R.

Abstract 21: Here I would expect the statistics for the Alpine-wide permafrost distribu-

tion.

AC: The statistics are not presented in this manuscript as described above.

Introduction

p1421 13: You have to define somewhere which slope angles are related to the term "steep rock walls".

AC: We added the following sentence: "With the term "steep bedrock" we refer to terrain that a) is not or only marginally affected by a snow cover in wintertime, b) does not contain large amounts of blocks and/or debris, and c) is without vegetation coverage." The slope threshold that we define to separate steep rock faces from other terrain is 45° and was set manually.

Background

p1423 11: I miss the mention of the occurrence of the different types of permafrost found in the Alps. In which geomorphologic units does permafrost occur (bedrock/fissures, unconsolidated sediments/talus slopes, rock glacier, etc.) and what are typical textures (fine-grained) for permafrost in sediments?

AC: We consciously restricted the discussion to the two geomorphic units for which we have sufficient data.

p1423 11: 15-18: Use for example the term "creep behaviour " instead of movement feature. For the overestimation of the permafrost distribution only the creep behavior is crucial, not the cooling effect.

AC: We replaced "movement feature" with "creep behaviour of rock glacier". For the overestimation of the permafrost distribution the coarse block surface and the *related cooling effect are important, because this leads to an optimistic prediction (biased towards an overestimation of the permafrost distribution) in comparison with for example fine-grained surfaces.*

p1423 27: Do you mean the lateral variability of the rock surface temperature?

AC: Yes, we added ". . . .of MARST" to make this clearer.

p1424 1: Rewrite; It has been shown that MARST values can indicate... (Nötzli et al,?)

AC: See below.

p1424 2-3: Delete this sentence if it is not relevant for the next sentence. If yes, shorten this two sentences into one. At present this text section is very confusing.

AC: We agree and rewrote the last sentences from the former section "Background": "MARST values can indicate permafrost occurrence in the ground, but the extrapolation of MARST values to subsurface temperatures is affected by large uncertainties due to varying surface and subsurface conditions (Noetzli et al., 2008)."

Data

3.1 Response Variables

p1424 15: The different sources of the data need to be presented in a separate table. I suggest the following pattern: response variable—-country—-region–number of data–author/source rock glacier activity–CH——Ticino–1500———Frauenfelder et al., /.. MARST————- CH——Ticino–1500———Hasler et al., 2011/..

AC: Done, see Table 1.

p1425 5-10: The MARST values are computed from your observations and are then corrected by using a function of the MAT. Since a binary variable (permafrost found/not found) would satisfy the input parameter of the statistical model why do you use such a complicated way instead of using only the MARST? In the present approach you are interested to predict lateral variations of the permafrost, and not in vertical direction.

AC: We do not use a binary variable for our rock model because a linear regression can be used easily and for a binary regression we do not have enough observations. This step is necessary to homogenize all MARST to the same reference period. Otherwise inter-annual variations would strongly bias the regression model. Further, we are addressing both, vertical as well as lateral variations in permafrost occurrence.

3.2 Topographic and climate variables

p1426 1: Why do you center PRECIP? Do you need small values for the inversion, or do you need this for your initial model?

AC: We centred PRECIP because then it is possible to compare the model coefficients when including and excluding PRECIP as explanatory variable. We added the following sentence to the manuscript: "Centring PRECIP allows to directly compare the coefficients of the different models including and excluding PRECIP as explanatory variable."

p1426 8: Specify the term local horizons. Do you mean joints?

AC: Done: "...local horizons affecting the obstruction of solar irradiation. . ."

p1426 10: I assume that you use the ZAMG-MAAT for the predictions. So, how do you estimate the adjustment value for the predictions?

AC: There is no adjustment value available for the predictions because only the DEM is available to calculate MAAT (no local elevation estimates).

p1426 15-16: You write "can be used", do you have used data from this techniques? Rewrite this sentence, and specify how accurate the DEM data have to be to allow reasonable calculation of PISR.

AC: This sentence was removed because it is not relevant for this study.

4. Statistical Methods

I suggest to use the singular and use either "Statistic Model", "Statistical Method" or "Statistical Modeling" for the heading.

AC: Title changed to "Statistical Method"

4.1 Theoretical framework

The section 4.1 "Theoretical framework" seems to be rather a concept than a theoretical framework. Here I would expect information on the Generalized Linear Model, the generalized mixed-effect model (Dobson 2001, Crawley 2005), and the probit model in general.

AC: Done, see general comment 3.

4.2 Model formulation

p1427 This section lacks in several issues such as clear description of the used symbols or the number of used symbols which makes it very difficult for the reader to follow the MS. This might be a result of the missing concept at the beginning of the MS. Explain what quantities where a "tilde" appear does mean. Further clearly formulate which equations are used for the rock- (eq. 3?) and the debris model (eq .7?). You might also introduce sub-sections such as "Rock model" and "Debris model".

AC: This is an important comment and helps us to improve the clarity of the model formulation and motivation. A paragraph has been added at the beginning of section 4 outlining the structure of this key section. The notation in the former section 3.1 could be simplified, and Eq. (2) could be omitted.

The use of the tilde throughout the manuscript is now briefly explained after its first occurrence in Eq. (2): "Throughout this work, model coefficients with a tilde refer to the temperature scale as in Eq. (2), while model coefficients at the probit scale will carry no tilde."

The revised manuscript states now more clearly that Eq. (5) (formerly Eq. (6)) is the probit model used as debris model, while the original version already expressed that Eq. (2) (formerly Eq. (3)) is the basis for the rock model.

1427 1: I would prefer the use of a permafrost definition where the formation of ice is possible (p=P(theta<0).

AC: Technically, the difference between "$\leq$" and "$<$" has probability 0 in the case of a continuous probability distribution as we have it here; in other words, it makes no difference. As a convention, "$\leq$" is preferred in probability theory, with some exceptions (survival analysis). In terms of permafrost and ground ice, "≤ 0" °C allows at least preservation during melting and it represents the definition of permafrost by Van Everdingen (1998).

1427: 9. Use "...that.. corresponds to the normal distribution" instead of "..being normally distributed".

AC: The entire sentence has been omitted.

1427: 10: Explain if you use different symbols or the same one for measurements and predictions in your equations. I would prefer to make the text more comprehensible and to omit the term ”prediction” or ”model prediction” at this point. You could introduce this term in line 19. The used symbols for the mean and the standard deviation are very confusing (sigma** 2_theta & theta dash). I suggest to use ”mu” & ”sigma**2” for the statistical moments of the ground temperature. To help the reader to understand your quantities I recommend to show the whole (well-known) formulation of cumulative distribution function.

AC: The entire sentence has been omitted. We think that the use of the Greek letter theta helps convey the fact that this variable represents the ground temperature. The use of sigma for standard deviation and sigma^2 for variances is well established. We omit subscripts wherever possible; the sigma^2_theta is not used any more in the revised manuscript because the corresponding sentence was deleted. Finally, the cumulative distribution function (in general and of the normal distribution) is in fact, as pointed out by the reviewer, a well-known function, which is why we prefer not to add an unnecessary formal representation involving an integral and the probability density function.

Equation 2: Leave the subscripts and add the formulation of the cumulative distribution function p=...=int((1/sigma*sqr(2*pi))*exp(-(x-mu)**2/(2*sigma**2)))dx.

AC: Equation has been omitted from revised manuscript; see also previous comment.

Equation 3: Define the quantity ”k” - number of observations? Are ”alpha” and ”delta tilde” scalars? If yes, the inversion algorithms might only be able to determine the sum of both quantities. How do you separate the value of the sum on the two quantities? The same problem might occur in equation 6.

AC: We inserted ”k” before ”explanatory variables” to clarify this. The distinction of alpha and delta is just formally. The separation is explained later, as mentioned in

the text. In brief, alpha is the usual intercept, delta is a ”manual” adjustment term that accounts for thermal offsets and similar effects, and this equation is not for direct coefficient estimation but a formal representation of our conceptual model.

1427: 17: Is the mean of your residuals indeed zero, or is this a request of the least square model? You can simply check this. You might better use ”random error” or ”disturbance term” to define ”epsilon tilde”.

AC: This is a theoretical model, not an empirical model, so at this point is not possible to check whether or not the mean is empirically close to 0. The formal assumption of normal distribution with mean 0 is necessary in order to get to the probit model, which uses the normal distribution function. The term ”residual error term” is well established in this context, although we acknowledge that similar terms are used in different fields. We believe that the terminology used here is sufficiently clear.

1427 18: Define exactly which quantity is the explanatory variable and which one is the coefficient.

AC: Done: ”. . ., x_i are the model”s explanatory variables and beta_i their coefficents, ...”

1427 19: State in which section the mentioned explanations will be found. *AC: Done, Sect. 3.3 in the revised manuscript.*

1427 19-21: The term ”predictive situation” is not a statistical/mathematical formulation. Use ”to make predictions ... to estimate/determine...”. You have to replace this at several places.

AC: A situation in which predictions are made is a predictive situation; this term is *widely used in the literature.*

Equation 4: It is unclear how this equation is derived. You might refer to other equations or introduce more details. Further, the equation seems to be non-conform with your permafrost definition (try theta= 0 C). You might use "$\geq$ " here.

AC: "> 0" was a mistake, this should be omitted; Phi is simply a function that takes a scalar as its argument. For a predicted theta tilde = 0 $^\circ$C we obtain a permafrost probability of 0.5, which makes sense given the fact that theta tilde is subject to uncertainty as expressed by its non-zero variance. A sentence explaining the origin of the negative sign has been added after Eq. (3) (formerly Eq. (4)).

p1428 1-6: This text rather fit to the section 4.1(Theoretical framework) or section 2. ("concept").

AC: Text has been reduced to avoid overlap with earlier sections.

p1428 20: This sentence is confusing, you might use "The relationship between the observed? temperature and the presence or absence of permafrost allows .."

AC: This sentence is in fact about the relationship between models, not between the mentioned variables. Rephrased to be more explicit.

4.3 Integration of continuous- and binary-response models

Heading: Use consistent terms for your models; rock/debris models or continuous/binary models. I seems that the model coefficients are also estimated within this section. Thus you might change the heading to "Integration and parameter estimation of the rock- and debris model".

AC: No, the coefficients of the model are not estimated in this section, but in section 5.1. Accordingly we did not change the title of this section.

p1429 5-9: Use "For .." instead of "In our case..". Avoid to use that much quantities. Write"..two models Mr (rock surface) and Md (debris surface) that are fitted separately." and leave the rest of the sentence. If the two models are fitted separately, is there at least one parameter that is inverted from both data sets? If no, why do you need to integrate these two models into one? You could then use two separate models and could easily use standard geostatistical methods (e.g. Kriging).

AC: We moved part of this sentence to the section "Conceptual background" and deleted it in this part of the manuscript. Regarding the geostatistical methods, see general comment 6.

p1429 13-15: I did not understand what implication the parameter "delta_d" has on the model. Do you use rock glacier velocities as discriminator for the presence of permafrost? I further expect that you are an expert for applying this statistical models. If "delta_d" is not used in your study it should be omitted from all equations. However leave "..but represent an expert-defined adjustment term".

AC: In the revised manuscript we state more explicitly that delta_d will be required for model application at a later stage, and that it is therefore important to allow for its inclusion in the model formulation.

p1429 17: Which are the same explanatory variables?

AC: Done, we included: "(MAAT, PISR)"

4.4 Scaling Issues

p1430 8: Do you mean resolution or the grid cell size?

AC: Yes, we used the term grid cell size instead of resolution.

1430 15: Use quotation marks for the term ”change of support”.

AC: No quotation marks needed, see Gotway and Young (2002)

1430 16: Why do you not apply wellknown statistical methods such as Kriging?

AC: See general comment 6.

p1431 1: see comment on p1427

See reply above regarding the expression ”predictive situation”.

p1431 6-9: What do you mean with ”possible”? - 2x

AC: The first occurrence has been changed to ”such as”; the second occurrence has been left deleted.

p1431 15: How do you determine N? - Bian and Butler (1999) suggest to sum up to the range of spatial autocorrelation to reduce errors by averaging dissimilar units.

AC: Here, the grid resolution is given a priori, e.g., by the resolution of ASTER-GDEM or Lidar DEM. We are not trying to identify an optimal N (and therefore resolution difference), we are just working with given data.

Equation 13: The average might be representative if the spatial correlation length is low (high scatter, random behavior). In the second case I suggest to use either the

median or Kriging.

AC: In the presence of autocorrelation there is no obvious reason why the median should have more desirable properties than the mean, apart from being a robust measure of location. Block kriging would possibly be an option if the semivariogram was known sufficiently well (Gotway and Young, 2002). However, this would be a computationally intensive process that is difficult to apply over an entire mountain range of the size of the Alps. The present approach furthermore allows us to arrive at a conservative solution that is satisfactory for this type of study.

1432 1: Why is this a ”conservative choice”?

AC: More detail is now provided to explain that under within-cell independence of fine-scale errors, the variance of the mean of these errors decreases with 1/N, while in the extreme case of ”perfect” positive autocorrelation it does not increase with increasing N. Hurlbert (1984, p. 201) is referenced to support this argument as well as to justify the use of this conservative estimate.

1432 5: Does your approach consider the Gaussian error propagation law?

AC: We earlier made the assumption of Gaussian residuals of the rock model, but Gaussianity is not required for this particular consideration here. No changes made.

1432 14: see comment on p1427 19-21

AC: See reply to the above comment.

4.5 Surface Types

Do you have used one of this approaches. If not, you can omit this section. If yes, this should be explained in the ”data” section.

AC: We added a new figure (Fig. 1) to the manuscript that is based on this approach. The section is not moved to section data, because we do not describe any new data here.

4.6 Model fitting and Assessment

The heading should be more precise. At present the text involve a heading such as ”Model implementation and accuracy”. I suggest to add information on how to parameterize the model and to use as potential heading ”Model parameterization and evaluation”. The systematic order would be implementation, parameterization, and accuracy.

AC: The first part of this paragraph was moved to the section ”Conceptual background”. The heading was changed to ”Model evaluation”.

p1433 1-4: This kind of sentence appears on various places of the MS, but is not necessary here.

AC: Sentence is moved to ”Conceptual background”.

p1433 4-6: How does the GLMM takes into account random inventory effects? Do you use weights for your input parameters? That rock glacier samples were taken randomly is also mentioned earlier and don’t need to state twice.

AC: In the estimation of model coefficients, random effects are taken into account by using a modified likelihood function for maximum likelihood estimation. This cannot be expressed in terms of simple weights. Please refer to the referenced literature for mathematical details on the estimation procedure and likelihood functions used in GLMMs.

p1433 7: Do you mean ”..the penalized quasi-likelihood method ..”?

AC: Yes, the penalized quasi-likelihood method was used, but to avoid further confusion we omit the reference to PQL in the manuscript (see general comment 3).

p1433 8-9: see comment 1-4.

AC: Sentence is moved to ”Conceptual background”.

p1433 10-12: Use only one, consistent, term for the ”debris model”. I do not understand how the AUROC is computed or what it does represent (reference!). ”ROC curve” can be omitted if the explanation of AUROC is clear.

AC: ”Probit model” changed to ”debris model”. The following reference was added to document the definition and interpretation of the ROC curve: Hosmer and Lemeshow (2000)

p1433 13-14: Do you also present results or AUROC”s without the random inventory effect? I assume that you again mean the random samples taken from the rock glaciers. - In this case, the term ”effect” seem to be not adequate.

AC: No, here again the random effects described above are meant; the term ”random effect” is therefore correct.

p1433 14: If the abbreviation ”cv” is not used again, leave it. For me it would be more important to see how the model adjust if all data are used. Did you also try this, and which values have your residuals.

AC: The term ”cv” is used in Table 4. The training-set AUROC was also reported in

Table 4, although our article focuses on the cross-validation AUROC. Cross-validation results in a bias-reduced measure of goodness-of-fit compared to the over-optimistic accuracies obtained when evaluating a model on the same data set on which is was fitted (trained).

p1433 15: You might write ”adapt”, ”adjust”, or ”focuses” instead of ”generalizes”.

AC: Changed to ”...how transferrable to independent test data sets the model is (Hand, 1997).”

p1433 21: see comment 10-12.

AC: Changed to ”rock model”.

p1433 22: Simply write RMS, this term is well-known.

AC: We used the term RMSE. See comment on Abstract 19.

5. Alpine-wide permafrost model

The heading should be more precise. The section mainly contains the model calibration but also an interpretation of results. I suggest to use the heading ”Model calibration” and an additional chapter ”Result” or ”Interpretation”.

AC: We decided to not split the model calibration and interpretation of each sub-model, because we think introducing an additional chapter does not contribute to the overall understanding of the manuscript (reader needs to jump forwards and backwards to follow the model calibration and interpretation). However, we added a small paragraph at the beginning of chapter 5 to make this clearer.

5.1 Debris model

p1434 1-20: This text include information on the model calibration/parameterization and should be moved to such a section.

AC: Done, see comment above.

p 1434 21-p1435 2: This text correspond to the analyses of coefficients of the final model and should be moved to an adequate section, e.g. ”Result” or ”Interpretation”.

AC: Done, see comment above.

p1434 5: If the input data should be consistent use the centroids for all rock glacier. It would be still possible to take random values along this line.

AC: Where rock glacier polygons were available, we chose to select a random location within each rock glacier because over the whole sample this will reflect ”average” conditions more accurately than a polygon centroid. In this context it is also important to note that polygon centroids can fall outside a polygon in the case of some non-convex polygon shapes. Using points outside a rock glacier to represent rock glacier topographic conditions would be misleading. No changes made.

p1434 11: Table 2 shows three different models that were not introduced up to now. This could be done in the section ”concept” or at the beginning of section 5. Why do you introduce exactly three models and how do you constrain your final model? Therefore you have to use a consistent methodology.

AC: We introduced this three model because we wanted to analyse the effect on precipitation in seasonal and in an annual context. This is now described at the beginning of Section 5.

p1434 12: Provide a reference for the Wald-test or briefly explain how it is derived.

AC: This is a standard testing procedure in the context of generalized linear models. References concerning logistic regression and generalized linear mixed models have been included earlier in the manuscript.

p1434 15: I do not understand this sentence.

AC: See general comments above concerning random effects.

Table 2: Put the term "in parenthesis" in parenthesis. Are the three models potential debris model? I prefer to simply write "residuals" or "errors" instead of "goodness of fit". State the units for the explanatory variables. I do not understand the term "inventorylevel standard deviation". You might use "sd" or "std" as abbreviations for " standard deviation".

AC: Parentheses added. "Goodness-of-fit" is the most accurate term here, not "error" or "residuals". It is common practice in mixed modelling to indicate the level at which residuals are summarized/reported, please see the referenced general literature on this topic.

Table 3: Use "Summary of statistic parameters and ..". The values are not well readable and should be aligned to columns - see comment on table 1.

AC: Done

5.2 Rock model

1435 7: It is not clear what the AIC criterion is. Rewrite this sentence and give a reference for this criterion. Further, which parameter shows the insignificance? For me

the AIC seems to be no good criterion since it shows almost the same values for the three models. In general there seems to be no effect to the residuals in the 3 models although Intercept, MAAT and SEASONAL changes significantly. Why do you fix PISR? You write at the beginning that this parameter is import to know (this parameter causes the scaling issue). AC: Reference to Gelman and Hall (2007) was added.

The AIC is a standard criterion to measure the fit of different models relative to each other, which is why it was included in Table 4. The AIC is introduced in the text, including a reference to the literature. The AIC is not an absolute value. In our view it is possible to say that the observed values are "almost the same", as suggested by the reviewer. The AIC includes a penalty term that accounts for the size of the model. Rock model 3 achieved a decrease in AIC relative to the smaller model 2 despite the mentioned penalty. One might therefore be inclined to choose model 3 over 2. In this study the smaller model 2 is, however, preferred in order to avoid extrapolating from the SEASONAL range of our samples to lower and especially higher SEASONAL values found in some parts of the Alps; this is explained in the manuscript text (section 5.2). Please refer to the statistical literature for more information on the AIC (e.g., Gelman and Hall, 2007; Hosmer and Lemeshow, 2000).

We are not sure if we can agree with the reviewer's interpretation of differences in coefficient estimates and residual standard deviations. On the one hand, the differences in MAAT coefficient estimates among the three models do not seem to be significant, as suggested by the reviewer (pairwise differences smaller than 0.11 compared to standard errors greater than 0.08). Differences in intercept estimates are due to the use of different sets of variables; this does not require interpretation or discussion (see statistical literature). On the other hand, it is not surprising that the residual standard errors (and RMSE, R^2) do not vary a lot among the models. The precipitation-related variables that were added/dropped simply contribute very little to the overall goodness-of-fit because their variation across the study area is limited, and because regional precipitation patterns cannot be expected to be a major control of rock glacier activity *levels, even though their influence may be statistically detectable.*

The coefficient of PISR for the different rock models was not fixed. The equality (up to the given number of decimal places) is coincidence and underlines the reliability of these estimates. No changes made because to our knowledge our manuscript does not suggest that PISR was fixed.

5.3 Scaling model and model combination

p1436 1:Which model do you mean? - The linear regression in Eq.16, or the rock model?

AC: The linear regression from Eq. 16. We added: "(Eq. 16)"

6. Discussion

6.1 Use and limitation of the model

1436 23: Do I need an alpine-wide permafrost model for regional application?

AC: Yes, for two reasons: (a) only this provides a sufficient data base for statistical analysis; and (b) this provides a way to transfer experiences in model application between provinces or countries.

1436 25-26: Note also other methods such as the Kriging or the Minimum Curvature Gridding.

AC: See comments above.

p1437 6: Explain why a large data set that is used for model calibration is not effected by variations in the elevation data.

AC: This sentence was replaced: "However, no better DEM is available at the moment for the entire Alps."

p1437 15-17: I expected that this was done in this study. If it is not possible to do this you should discuss the impact of the adjustment offsets delta on the model result.

AC: This could not be included in this manuscript but is described in Boeckli et al. (to be submitted).

6.2 Influence of precipitation

p1438 15-19: This should be explained earlier. How large are the uncertainties in the PRECIP data? - 30%, or more? How does this uncertainty effect your result?

AC: Uncertainties in precipitation data in the Alps are in general large due to the small number of high altitude weather stations and the fact that the precipitation field is less coherent over the Alps compared to flat terrain (Efthymiadis et al. 2006). In our regression model, only the regional pattern of precipitation is of interest due to the low spatial resolution of the data (ca 15 km). The effect of variations in precipitation is discussed in Sect. 5.1.

7. Conclusion

p1439 This section reads such as a summary and lacks in scientific findings. I recommend to completely rewrite this section and address the aims of the paper.

AC: We rewrote the conclusions partly. See general comment 8.

p1439 11-20: see comment above.

AC: See above.

p1439 14-18: I suggest that the model is also applicable with other data. Thus this information is not relevant for a conclusion. In contrast it is important to emphasize how sensitive your parameters are and to know which parameters are used for the two models.

AC: This sentence describes which variables are used to apply the debris model. No other data can be used to apply the statistical model, but a DEM with better spatial resolution can be used instead of ASTER GDEM to derive the terrain variables. The sensitivity of the relevant variables is discussed in section 5 and no further comment is needed here.

p1439 21-22: This seem to be one of your results and should not be questioned. Simply state that you have found this relation.

AC: In our view, stating this disagreement with earlier studies provides important information to assess the results and to judge the merit of their future scrutiny.

p1439 25-p1440 9: see comment on line 11-20. In a scientific paper it is not common to give instructions how to run a model.

AC: See general comment 4, this is important to put the results into perspective for potential application.

p1439 p1440 10-18: This text reads itself like an outlook. Points 2,4,5 are obvious and should be omitted.

AC: See general comment 4, this is important to put the results into perspective for potential application.

p1439 21-25: see comment above.

AC: See general comment 4, this is important to put the results into perspective for potential application.

Figures

Fig.1: Use a projection to avoid such large distortions. You might use UTM -zone 32. It is also necessary to add the names of the countries, optional the names of large cities.

AC: Done, UMT 32 is used and countries are labelled.

Fig.2: I don't understand your comment on the random effects.

AC: In the debris model we considered inventory-related random effects as described in Sect 5.1. Exploratory data analyses do not account for random effects. The statement in the figure caption is probably misleading because it is rather obvious; it was therefore removed.

Fig.3: Add a plot for the relation: residuals vs. PSIR. For the figures in the lower panel draw in or state the mean and the standard deviation.

AC: PISR is part of the regression that is used. Such diagnostic plots were examined during model development, but we are unable to include all such plots in the manuscript. Therefore we think it is not necessary to plot the residuals of the linear regression vs. PISR. For the figures in the lower panel, the mean and standard deviation are stated in the caption.

Figures 1,3,4,6 can be show up in grey colors.

AC: It's hard to show Figure 1 and the new Figure in grey tones. Therefore we leave it *to the reader whether to print in colour or grey tones.*

References:

Akaike, H.: Likelihood and the Bayes procedure, Bayesian Statistics, Ed. J.M. Bernardo et al., Valencia: University Press. p.143-166. 1980.

Boeckli, L., Gruber, S., and Brenning, A.: Estimated permafrost distribution in the European Alps, The Cryosphere, to be submitted.

Barsch D.: Rock Glaciers: Indicators for the Present and Former Geoecology in High Mountain Environments. Springer-Verlag: Berlin. 1996.

Brenning, A. and Trombotto, D.:Logistic regression modeling of rock glacier and glacier distribution: Topographic and climatic controls in the semi-arid Andes Geomorphology, 81, 141 – 154. 2006

Crawley, M. J.: The R book. West Sussex, England, 0-978, 2009

Gelman, A. and Hill, J.: Data analysis using regression and multilevel/hierarchical models, vol. 648, Cambridge University Press: Cambridge, UK, 2007.

Gotway, C. and Young, L.: Combining incompatible spatial data, J. Am. Stat. Assoc., 97, 632–648, 2002.

Hand, D. J.: Construction and assessment of classification rules, Wiley Series in Probability and Statistics, John Wiley and Sons, Chichester, 1997. Harris, S. and Pedersen,

Hughes, P. D., Gibbard, P. L. and Woodward, J. C: Relict rock glaciers as indicators of Mediterranean palaeoclimate during the Last Glacial Maximum (Late Würmian) in northwest Greece, Journal of Quaternary Science, 18, 431–440, 2003.

Hurlbert, S.: Pseudoreplication and the design of ecological field experiments, Ecological monographs, 54, 187–211, 1984.

Lewkowicz, A. and Bonnaventure, P.: Interchangeability of mountain permafrost probability models, northwest Canada, Permafrost and Periglacial Processes, 19, 49–62, doi:10.1002/ppp.612, 2008.

Interactive comment on The Cryosphere Discuss., 5, 1419, 2011.

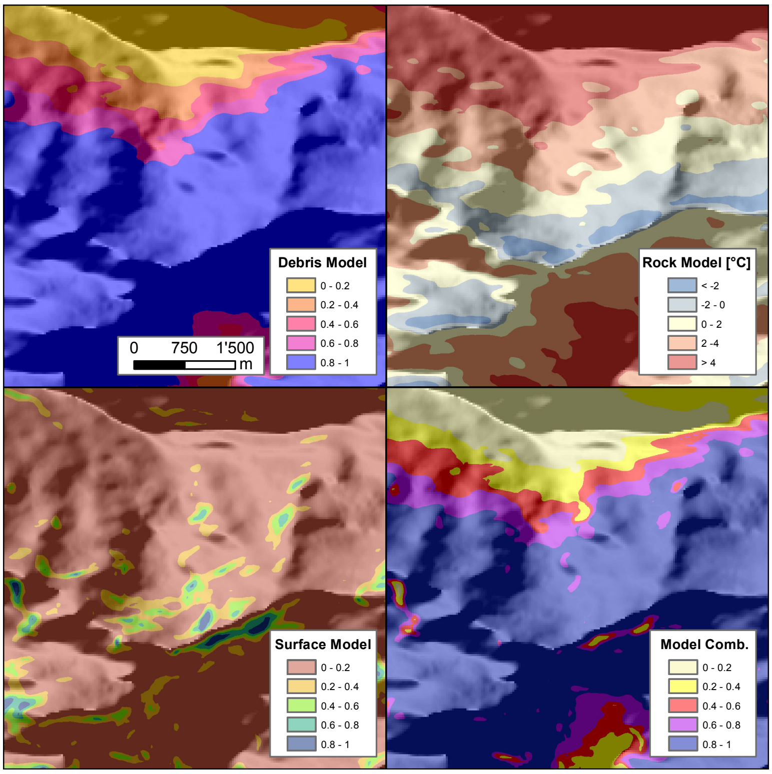

Fig. 1. Example of the application of the different models. Top left: Prediction of the debris model showing probabilities of permafrost occurrence. Top right: Predicted MARST values of the rock model. Bottom

Response Variable	Country	Region	N (intact / relict)	Source
RG status	AT, CH, FR, IT	Various regions	1625 / 3916	Cremonese et al. (2011)
RG status	CH	Entremont, Valais	115 / 137	Delaloye et al. (1998)
RG status	CH	Engadina, Graubünden	115 / 137	Frauenfelder et al. (2001); Frauenfelder (2005)
RG status	CH	Engadina, Graubünden	18 / 6	Hoelzle (1998)
RG status	CH	Aletsch region, Bern	11 / 13	Imhof (1998)
RG status	CH	Printse valley, Valais	115 / 137	Reynard and Morand (1998)
RG status	CH	Fletschhorn area, Valais	50 / 22	Frauenfelder (1998)
RG status	CH	Prealps, Vaud	0 / 25	Schoeneich et al. (1998)
MARST	AT, CH, FR, IT	Various regions	49	Cremonese et al. (2011)
MARST	CH	Matterhorn, Jungfraujoch	8	Hasler et al. (2011)

Fig. 2. Overview of data used for model calibration (RG Rock glacier; AT Austria, CH Switzerland, FR France, IT Italy).