The Cryosphere Discuss., 5, C2220–C2239, 2012
www.the-cryosphere-discuss.net/5/C2220/2012/

The Cryosphere Discussions

TCD

5, C2220–C2239, 2012

Interactive Comment

Interactive comment on "Optimisation of quasi-3D electrical resistivity imaging – application and inversion for investigating heterogeneous mountain permafrost" *by* D. Schwindt and C. Kneisel

D. Schwindt and C. Kneisel

daniel.schwindt@uni-wuerzburg.de

Received and published: 13 May 2012

We would like to thank Andreas Kemna for his review and his comments and improvements on the manuscript. However, we do not agree with some of the referee's comments and his fundamental critique on the manuscript. The referee alleges arbitrariness in some cases (choice of final model, choice of inversion parameters) which is not the case, but even explained within the manuscript. The main aspect of the study was to test the ability of different setups for detecting the areal distribution of high resistive anomalies in a low resistive matrix and the "active layer thickness". This approach is –

like stated in the manuscript – mostly qualitative. As a consequence the choice of the final image was based on the best agreement between the distribution (location and extent) of resistive bodies between the initial model and the final image. This approach is, in our opinion, neither arbitrary nor incorrect but even in agreement with common geophysical practice. To cite the reviewer (RC12) the back-and-forth interpretation procedure is "…what geophysicists do for ages, i.e. setting up a model, simulating data and comparing inversion results for different inversion settings, data error characteristics etc.". Figure 6 shows, that the resistivity increase during the inversion process is less pronounced for the high resistive bodies than for the shallow depth layers above them, where anomalies appear after the third iteration. The quantitative approach constitutes a minor part of our contribution, but is the main issue in the referee's comments. Even though our results show, that a quantitative interpretability of quasi-3D images is, with regard to our approach and our "qualitative" stopping criterion, hard to assess, we do not want to criticize recent contributions aiming towards quantitative geophysics on permafrost related problems. Here quantitative approaches in permafrost studies for the interpretation of subsurface temperatures and permafrost characteristics were recently developed by Hauck et al. (2010, The Cryosphere) who estimate the volumetric content of rock, water, air and ice by the electrical resistivity and seismic velocity of the subsurface and by Krautblatter, Verleysdonk, Flores-Orozco and Kemna (2010, J. Geophy. Res.) who achieved the quantitative interpretation of ERT by employing a smoothness-constrained inversion with an empirically measured normal reciprocal error model. These approaches indeed exceed the scope of this contribution that is, as stated above, focused on the qualitative detection of permafrost. However, citing Krautblatter, Verleysdonk, Flores-Orozco and Kemna (2010, J. Geophy. Res.): "This approach, where the data are fitted to a well-defined degree based on an adequate data error description – in contrast to the widely used approach of just minimizing the data misfit – is essential for a quantitative interpretation of ERT results" it is questionable, if the referee's demands on quantitative ERT can be achieved using commercial software for inversion.

However, our results on the quantitative potentialities of quasi-3D ERT are not contradictory to existing publications on 2D ERT. The reliability – or unreliability – of resistivity values applying ERT monitoring on a rock glacier has lately been discussed by Hilbich et al. (2009, Permafrost and Periglacial Processes). " ...including synthetic modeling, the depth of investigation index technique and the so-called resolution matrix approach" Hilbich et al. (2009) showed: "... that resistivity values of model regions with strong resistivity contrasts and highly resistive features are generally of critical reliability, and resistivity changes within or below the ice core of a rockglacier should therefore not be interpreted as a permafrost signal" and that "...no reliable information can be obtained for electrical properties of the highly resistive ice body, whereas in a zone with lower resistivities the reliability is much higher". These results coincide with the results from our study. Day-Lewis et al. (2005, J. Geophys. Res.) state: "Although qualitative insights into subsurface architecture are readily made from geophysical data, quantitative use of geophysical images to estimate values of hydrologic parameters suffers from limitations arising from imperfect and variable tomographic resolution. Geotomography tends to overpredict the extent and underpredict the magnitude of geophysical targets; moreover, the spatial structure of tomograms may only weakly reflect the true spatial structure of the subsurface". To quantify this "correlation loss" (Day-Lewis et al. 2005; J. Geophys. Res.), a number of approaches exist for image appraisal and data error estimation. However, this is not scope of our manuscript. Using a back-and-forth interpretation procedure allows for a direct comparison of input data (initial geocryologic model) and output data (inverted image). Conclusions on the capability of different setups for detecting certain structures can be drawn from the qualitative comparison of structures being imaged correctly in size and shape or not. The application of different methods for image appraisal and error estimates will not alter the fundamental results and conclusions of our study. An extensive analysis of data errors and image appraisal would in our view – due to the high number of scenarios used – exceed the scope of our already quite long contribution by far. Being no "ERT developers" but first of all users – like supposedly most permafrost researchers – we are more or less reliant upon the potentialities of commercial software. Some of the methods for image appraisal and error estimation are simply not available or currently in development (e.g. depth of investigation index, personal communication with M.H. Loke) in the software package (Res3Dinv) we use. However, we will exemplarily include information on model sensitivity and data errors to emphasize our results and conclusions. A further point in the referees comments points towards the references cited in the manuscript. We tried to focus the references – as far as possible – on recent and permafrost-relevant literature. Of course we are aware of the fact, that geophysical methods and the approach of synthetic data modeling (and also the back-and-forth interpretation procedure) was not invented by permafrost researchers but has a long tradition in geophysical research. Here we will try to improve the manuscript by including the references named by the referee. However, we submitted the manuscript to The Cryosphere – and not to a geophysical journal – as we aimed to put the focus of the manuscript on the growing number of permafrost researchers who apply geophysical methods for their studies. This should not be regarded as an excuse for incorrect or – to cite the referee – "arbitrary" investigations. Our data processing has been conducted with high accuracy and was based on logical approaches. In our opinion a number of comments on our manuscript given by the referee is beyond the scope of our contribution. We will try to consider the referees comments for improving our contribution.

Reply to the general comments

RC1) However, the methods and tools used in the paper to achieve the objectives (assessment of images, optimization procedure and stopping criterion in the inversion) are not state of the art in the field of quantitative ERT. Important aspects and approaches in quantitative ERT, which are well known and documented in the literature, are ignored: (1) On the one hand this holds for the different methods available for image appraisal, i.e. the assessment of the imaging results. Here a number of approaches based on the analysis of cumulated sensitivity, resolution matrix or model parameter uncertainty are at hand (which all account for the varying resolution of ERT in different regions of the model); however, the authors do not adopt any of them for a more differentiated analysis of their simulation results. References in this context include for example: Ramirez et al., 1995, J. Environ. Eng. Geophysics; Oldenburg & Li, 1999, Geophysics; Alumbaugh & Newman, 2000, Geophysics; Friedel, 2003, Geophys. J. Int.; Day-Lewis, Singha & Binley, 2005, J. Geophys. Res.; Nguyen et al., 2009, Near Surface Geophysics.

- First of all, an assessment of the quantitative interpretability of our results is not the main intention of our study. However, we are not sure if we understand this general critique correctly. Based on the comment #29 we suppose that the referee wants us to include an extensive analysis of data sensitivity and data error estimation. In our opinion this will exceed the scope of this contribution as the aim of the study was to make suggestions for an optimized, efficient data acquisition (cf. general reply above).

- For image appraisal we will exemplarily include information on model sensitivity and data errors.

RC2a) The other issue which is not properly accounted for in the manuscript is the influence of data error, its description in the inversion, and – closely related – the applied criterion to stop the inversion iterations on the imaging result. Instead of rigorously computing the smoothest image which fits the data to a predefined degree (depending on the data error level), the final model is chosen in a rather arbitrary way. Importantly, chosen images do not yet explain the data within the assumed errors – as revealed by Figure 6 –, which suggests problems of the used code to model/invert with the required accuracy. Also for the field data no information/analysis regarding the data errors is included. However, this is essential to avoid over-fitting and possible misinterpretation. Useful references addressing the issue of data errors in ERT and their description in the inversion scheme are for example: LaBrecque et al., 1996, Geophysics; Slater et al., 2000, J. Appl. Geophysics; Koestel et al., 2008, Water Resources Res.; Krautblatter et al., 2010, J. Geophys. Res.

- We will exemplarily include information on data error for some scenarios.

- As explained in the general reply above choice of the final model is not arbitrary but based on the best agreement between geocryologic and inverted model. Iterations of each Inversion were – qualitatively – analyzed regarding the evolution of resistivity values of resistive bodies, the “active layer” and the low resistive matrix. Hereby the evolution of areas above resistive structures where a high resistivity gradient exists was observed in particular. This approach was also used for choosing the final model for the field example where we can assume – based on field studies on humus thickness – quite homogeneous ground conditions for the uppermost meter. A prolonged inversion process results in distinctly increasing resistivity values in areas above the high resistive permafrost body and, as a result the distortion of results. This approach, based on the qualitative analysis of inversion results and knowledge of the geomorphological conditions of the study site states a simple, logical and solid method to avoid over-fitting and possible misinterpretations. Here we will include an additional figure.

RC2b) There are also some problems with the used terminology. For instance, the term “reliability” (e.g. in the Abstract) is not appropriate in the given context because it lacks proper definition. When is an inverted value reliable, when not? The relevant quantity here is “uncertainty”, which can be computed for each resistivity value of the inverted model (e.g. Alumbaugh & Newman, 2000, Geophysics). I also disagree with the use of the term “quasi-3D” (see list of specific comments below for more details on this).

- We do not exactly understand the point of this comment. The term reliable does not lack of proper definition at all. Based on the comparison of synthetic input-data and the final results after inversion clear implications on the reliability can be given, especially regarding the high number of scenarios used. Besides, the term reliable is used in at least three of the contributions cited by the referee in the general remark 1 (Oldenburg & Li, 1999, Geophysics ; Day-Lewis, Singha & Binley, 2005, J. Geophys. Res.; Ramirez et al., 1995, J. Environ. Eng. Geophysics) in a related context as it is used in our contribution. However, even though the uncertainty is a quantity that can be applied to infer the reliability of inversion results the terms are in our opinion not replaceable as both terms comprise a discrete meaning.

- We also do not agree with the referees view on the use of the term “quasi-3D”. In our opinion the differentiation between “real-3D” and “quasi-3D” is definitely of interest (discussed within the introduction P 3386, Line 3 -23). However, this will be discussed in detail in the list of specific comments.

RC2c) A final general remark is on the title: I don’t think that “optimization” is justified here. Optimization implies a systematic procedure where by some means one approach is preferred over a range of other approaches. What is done in the manuscript is rather a comparison and assessment of different approaches, but not an optimization. There are actually papers around which address the data acquisition strategy from an optimization point of view, including: Furman et al., 2004, Vadose Zone J.; Stummer et al., 2004, Geophysics; Furman et al., 2007, Geophysics.

- In this case we do not agree. Optimization can be regarded as the process of achieving the highest effectiveness and/or functionality/quality of a freely selectable application or process. The way of achieving this optimization is not part of the definition. To cite one of the references given by the referee (Furman et al., 2004, Vadose Zone Journal): “The optimization of an ERT survey, as with any type of optimization, should consider the set of decision parameters (in the case of ERT location of electrodes) that will provide the best results (most reliable and accurate resistivity image).” Our approach tries to optimize the application of quasi-3D imaging towards a time-efficient measurement (which first of all depends on the number of 2D arrays used to create the grid) in consideration of the best possible data quality. Two aspects are included in our contribution. First of all we tested the ability of setups using a high number of 2D arrays to detecting the areal distribution of high resistive structures applying different electrode spacings focused on the array types Wenner and dipole-dipole. In the second step we systematically reduce the number of arrays to find the most efficient grid setup weighting effectiveness (number of 2D arrays per quasi-3D grid) against data quality. In our opinion this process indeed does constitute an optimization process, as it aims to increase the effectiveness of data acquisition in the field.

Reply to specific comments

RC1) P 3384, lines 15-16: This statement only makes sense if the characteristic length scale of the heterogeneity (size of anomalies, correlation length of structures with statistical heterogeneity etc.) is related to the electrode separation.

- We do not exactly understand the point in this comment. The subsurface resistivity distribution at permafrost sites and the size of structures is first of all not known, which constitutes the reason for geophysical surveying. At permafrost sites subsurface resistivity values can vary on short distance. The geocryologic model used as initial model in this study tries to reproduce this heterogeneity. The fact that setups using 2 m and 3 m electrode spacing were capable of detecting this small scale heterogeneity is shown in the manuscript.

RC2) P 3385, line 1: What is meant with “small-scale”? m, dm . . .?

- The term “small-scale” is not clearly defined but is related to a respective setting. Geophysical parameters of the subsurface can distinctly vary on a small scale between few meters and decimeters. We do not see the rationale to include a detailed description of the term small-scale.

RC3) P 3385, line 16: Dahlin and Loke (1997) is not a valid reference here (because they only apply 2D inversion). Other references in addition to Loke and Barker (1996) should be cited, in particular earlier work such as Park and Van (1991, Geophysics), Sasaki (1994, Geophysics), Ellis and Oldenburg (1994, Geophys. J. Int.). See e.g. references in Günther et al. (2006, Geophys. J. Int.)

- We agree. The list of references will be expanded; Dahlin and Loke (1997) will be deleted.

TCD
5, C2220–C2239, 2012

Interactive Comment

RC4) P 3385, lines 19-20: Put references in chronological order.

- The chronological order of references has been corrected throughout the manuscript

RC5) P 3385, lines 26-29: Of course there's many other works on 3D ERT. It would make sense here to cite a review chapter such as "Daily, W., Ramirez, A., Binley, A., and LaBrecque, D., 2005. Electrical resistance tomography – theory and practice, in Butler, D.K., Ed., Near-Surface Geophysics, Investigations in Geophysics, 13: Society of Exploration Geophysicists, Tulsa, 525-550" and/or "Binley, A., and Kemna, A., 2005. DC resistivity and induced polarization methods, in Rubin, Y., and Hubbard, S.S., Eds., Hydrogeophysics: Springer, 129-156

- The suggested citations will be included

RC6) P 3386, lines 3-23: I strongly disagree with the discussion in this paragraph and the proposed convention for the use of "quasi-3D". Actually the terminology is not at all unclear or inconsistently used among ERT developers: "2D" or "3D" refers to the parameterization in the underlying inversion, i.e., whether the resistivity distribution is allowed to vary in two or three dimensions, respectively. Of course, the resolution of model parameters in 2D or 3D requires appropriate data acquisition. So for instance nobody would expect to fully resolve a 3D image from a data set collected along a single linear array of surface electrodes. However, it is certainly not necessary – as long as isotropy is assumed (but that's beyond the scope of the paper) – to collect data using crossing (i.e. parallel and perpendicular, and perhaps diagonal too) surface lines. Parallel lines can be indeed sufficient because of the 3D measurement volume of a collinear 4-electrode array (i.e. the associated current and potential field lines do also penetrate the region outside the vertical plane defined by the collinear electrode array). The advantage of additional perpendicular acquisition lines may be the general improvement of the signal-to-noise ratio in the whole data set (by having more measurements) and perhaps a more homogeneous coverage of the subsurface (in terms of cumulated sensitivity). "Quasi-3D" refers to building a 3D image from independent

2D inversions, and not to running a 3D inversion of 3D data collected along several parallel acquisition lines. To avoid further confusion among ERT users, the paragraph (and actually the whole manuscript) needs to be revised in this respect. I also suggest differentiating between the inversion process and the data acquisition strategy here.

- In this case we do not agree. To start with the last point of the comment, we actually did differentiate between the inversion process and the data acquisition strategy. The following citation can be found in the introduction of our manuscript (Page 3386, Line 10-14): "Further differentiations must then be made concerning the inversion process. Papadopoulos et al. (2006) differentiate between the "quasi-3D" case, where inverted 2D models are merged to form a quasi-3D model and the "3D" case, where 2D profiles are merged and inverted afterwards, using a three dimensional inversion algorithm".

- To discuss the main point of this comment, we do not agree with the referee. First of all, an explanation of the reason of us using the term quasi-3D instead of simply professing to apply 3D ERT is discussed in detail in the introduction. Hereby we clearly differentiate – as shown above – between data acquisition and inversion process, as this states the main purpose of the differentiation. It is indeed correct, that quasi-3D referred to building a 3D image from 2D inversions. However, as discussed, the reason for us suggesting the term quasi-3D for 3D-inverted images based on merged parallel and perpendicular 2D arrays is to clearly differentiate from "real-3D" measurements where a complete three dimensional grid is built, with measurements being conducted along the x- and y-axis and diagonally. As we also apply "real-3D" measurements in our working group we find a differentiation necessary, as both methods have certain applicability. As the data acquisition techniques have been developed during the last years allowing for a faster data acquisition with an increased number of electrodes. As a result "real-3D" measurements will become more frequent. Based on the headlines of publications a differentiation between these two approaches is not possible. Thus we are assured, that a differentiation between 3D and quasi-3D would indeed pose a facilitation – if not for ERT developers but definitely for users. The fact that Wiederhold

TCD
5, C2220–C2239, 2012

(2005, Near Surface Geophysics, Investigations in Geophysics Vol. 13) used the term "poor man's 3D" to differentiate from "real 3D" seismics, that are often conducted in hydrocarbon exploration geophysics but are rare in near surface geophysics shows, that a differentiation between the different approaches is not redundant.

- We are not sure if the referee's opinion, that parallel lines are sufficient to reliably image the 3D subsurface resistivity distribution is universally valid. The 3D sensitivity of a collinear 4-electrode array depends on electrode spacing and the applied array type. Loke (2010, Tutorial: 2D and 3D electrical resistivity surveys) wrote: "Ideally there should be a set of survey lines with measurements in the x-direction, followed by another series of lines in the y-direction. The use of measurements in two perpendicular directions helps to reduce any directional bias in the data" and: "...to get a complete 3-D coverage, if the measurements are only made in the x-direction, the spacing between the lines should not be much more than the smallest electrode spacing used". This effect states the main part of our approach for optimizing quasi-3D imaging for permafrost studies. Testing the effect of enlarging the parallel spacing and reducing the number of perpendicular tie lines on the resulting 3D image.

RC7) P 3386, line 27: "This results in . . ." should be rephrased because it suggests a strict implication which is not at all the case.

- The sentence "This results in 1296 electrode positions..." is exemplarily to a grid size of 70 x 70 m using 36 electodes per array with an electrode spacing of 2m. This information is given in brackets as well as in the previous sentence (P2287, line 26/28) and, as a result, based on a simple calculation (36 x 36). So, in this case the information can be regarded as a strict implication. However, we will rephrase the sentence.

RC8) P 3387, lines 11-12: on "to optimize . . . by forward modelling": I think it is rather an investigation than an optimization; moreover it is not only forward modelling but also the inversion of simulated data, so something like "investigate . . . in numerical simulations" would be more appropriate.

TCD
5, C2220–C2239, 2012

Interactive Comment

- The use of the term optimization has been discussed in the general comments above (RC2c). For this case we will use the proposed phrasing.

RC9) P 3387, lines 16-17: “imaging” instead of “illustrating”; verb missing with respect to “resistivity values” (or positioning of “(1)”, “(2)” not correct).

- We do not see the mistake in this sentence, however we will rephrase.

RC10) P 3387, line 23: Note that the stopping criterion of the inversion is directly related to the issue of data errors.

- Information will be included

RC11) P 3388, lines 3-5: This statement should be rephrased and, if at all, only key references on this general topic should be given (only Scales and Snieder (2000) qualifies in this respect; however, it's better to refer to standard text books on the topic. The statement suggests that there are several studies dealing with geophysical forward modelling and inversion in recent years, while actually geophysics mainly CONSISTS of geophysical modelling and inversion and hence there's thousands of studies in the last 50 years (or more) on this. I suggest focussing on ERT only, where still a huge amount of literature exists.

- As stated earlier, we tried to focus the cited references on permafrost relevant literature without leaving out classical geophysical citations such as Scales & Snieder (2000). We do not see the point why the most recent permafrost relevant contributions on this topic (Hilbich et al. 2009; Fortier et al. 2008) should be deleted. However, we will implement the suggestions by the referee and will include further citations.

RC12) P 3388, lines 9-15: Again, these are all fundamental aspects of inversion theory in general and resistivity inversion in particular, which is not appropriately reflected by the selected references Hilbich et al. (2009) and Fortier et al. (2008). The “back-and-forth interpretation procedure” by Fortier et al. is what geophysicists do for ages, i.e. setting up a model, simulating data and comparing inversion results for different inversion settings, data error characteristics etc.

- We did not want to create the impression that the back-and-forth interpretation procedure was invented by permafrost researches who apply geophysical methods; nonetheless we tried to focus the references on permafrost-relevant and up-to-date literature. However, further references will be included.

RC13a) P 3389, lines 20-25: Please provide reference with an explanation of the "Wenner-Schlumberger" configuration;

- References will be provided

RC13b) should be "dipole-dipole";

- Will be changed throughout the document

RC13c) however, I'm not sure what the authors actually want to say here ("more sensitive to horizontal changes . . . and therefore a good tool for mapping vertical structures" is contradictory); actually the dipole-dipole array is likewise sensitive to resistivity changes in both horizontal and vertical directions.

- Aim of this paragraph is to discuss advantages and disadvantages of different array types used in this study, as a huge amount of literature, especially textbooks on Geophysics, is available on this subject the discussion is kept short. However, it is of course correct, that the dipole-dipole array is equally sensitive to horizontal and vertical resistivity variation. Nonetheless, aim was to compare properties of the Wenner and dipole-dipole arrays with the dipole-dipole array having advantages in the horizontal resolution towards the Wenner array and a slightly better vertical resolution for the Wenner array in comparison with the dipole-dipole array. To make this clearer the paragraph will be rephrased.

- In our opinion the cited sentence is not contradictory at all. Vertical structures could be defined as structures with the vertical extent exceeding the horizontal dimensions (e.g. walls). Horizontal structures (e.g. rock or sediment layers with different geophysical properties) can be defined as the opposite. While vertical structures – like walls – evoke a horizontal heterogeneity, horizontal structures – for example sediment layers – evoke vertical changes. As a consequence a tool that is sensitive to horizontal changes is capable of detecting vertical structures.

RC14) P 3389, line 27: I understand that "grid size" here only refers to the spatial extension of the grid (which is normally related to the electrode layout and thus electrode spacing and number of electrodes); however, grid size may also refer to the size of an individual grid cell or the number of total grid points (both being independent of electrode locations) – make clear what is meant here

- Correct, grid size refers to the spatial extent of the grid. The sentence will be rephrased to make the statement more clear.

RC15 a) P 3389, line 27 – P3390, line 4: I don't really understand the interdependencies between electrode spacing, number of electrodes etc. and grid specifications. If I understand it correctly, then electrode spacing is directly linked to the grid increments? If so, then the corresponding simulations differ systematically, with (unknown) effect on inversion behaviour, resolution of the final results etc. Why don't you use the same grid for all simulations? Then any grid effects on the imaging results could be ruled out. What is the idea of "roll-along" here (which actually is a data acquisition concept)? Why don't you just set up the grid as you need it?

- Grid size is related to electrode spacing and number of electrodes. As explained in P 3389, Line 29 the grid size of the synthetic models is adjusted for cables with 36 electrodes, based on the equipment used in our working group. The idea of the roll-along technique is related to the measurements in the field, where a y-distance of 168 m (using 36 electrodes with 2 or 3 m spacing) can only be achieved using a roll along technique. We will rephrase to avoid ambiguity.

- The geocryologic model created for this study tries to simulate the permafrost distribution of alpine/subalpine permafrost sites. The main idea of applying setups using different electrode spacings on the geocryologic model was to test the ability of these setups to resolving the given structures. For each simulation the same geocryologic model was used. Differences in the grid are restricted to its location inside the geocryologic model that was adjusted to cover all given structures.

RC16) P 3390, lines 11-15: Here I am a bit confused: I understood before that data along linear survey lines are modelled for a 3D resistivity distribution. Could you please specify which modelling code (and not only the type of modelling approach) was used? And you should explicitly state that the code is 3D to avoid confusion about the dimensionality in the modelling.

- Information will be included.

RC17) P 3390, lines 15-16: 3% is generally very optimistic for field measurements (actually even the numerical modelling error can easily amount to more than 3%), definitely for surveys in permafrost environments. Therefore this statement should be rephrased to avoid the impression that 3% is a typical value. Did you use different noise ensembles for the different data sets?

- The statement will be rephrased. We did not want to create the impression that 3% noise is typical for permafrost settings. However, a number of models were created and inverted using higher noise levels (3%, 5% and 8%). Based on the results we decided to use a relatively low amount of noise for the synthetic models.

RC18) P 3390, line 21: on “robust inversion”: this term refers to the use of the L1 norm to measure data misfit and/or model roughness in the objective function (which is being minimized). Please specify for which term the L1 norm was used. Following the argumentation in the text, it should only be applied to the measure of model roughness, not to the measure of data misfit. Is this correct?

- Correct, robust inversion was meant to refer to the L1 norm. The information addressed by the referee will be included.

RC19) P 3390, lines 22-24: How do you know that the inversion has indeed computed the smoothest possible model that explains the data within the given uncertainty (3%) always after 5 iterations (if at all)? For a non-linear inversion problem the inversion behaviour can vary strongly, particularly depending on the data errors (e.g. the used noise ensemble in the present case) and the resistivity distribution itself, exhibiting either slow or fast convergence. For quantitative ERT, the inversion process has to be completed, i.e. the smoothest model has to be computed that provides a data-error weighted RMS value of 1. With the applied approach you might arbitrarily over- or under-fit the data and thus produce inconsistent images. This is an absolutely crucial point if afterwards inverted resistivity values are interpreted quantitatively.

- As stated above, main intention of the study was to qualitatively test the ability of different setups to resolve given structures. Quantitative interpretation is of minor interest. As stated above, statements on the quantitative interpretation of our results will be relativized in the revised version of the manuscript. Of main interest in this study were the detection of active layer thickness, small scale heterogeneity with frozen and unfrozen ground in close distance and frozen ground thickness. As a consequence we chose the model that shows the best agreement to the geocryologic model. Adjusting the inversion parameters the models were inverted using far more than 5 iterations until the inversion converged. Main effect on the resulting quasi-3D image was that layers above high resistive structures were strongly biased with resistivity values increasing distinctly exceeding values given in the initial model by far. Resistivity values of high resistive structures do not come close to values given in the initial model throughout the inversion process. This effect is shown – for the first 5 iterations – in figure 6 and discussed in the corresponding chapter (Chapter 3.5, Page 3395).

RC20) P 3390, line 26: What does “1% accuracy” refer to? Is there any validation for this statement?

- 1% accuracy refers to the convergence limit used in the “incomplete Gauss-Newton” method that calculates an approximate solution of the least-squares method. Referring

TCD
5, C2220–C2239, 2012

to Loke (2010, Rapid 3-D Resistivity & IP inversion using the least-squares method) the accuracy of about 1% seems to provide a solution that is almost the same as obtained by the “standard Gauss-Newton” method. Information will be specified in the revised manuscript.

RC21) P 3391, line 5: “dipole-dipole” (check throughout the whole manuscript)

- Will be changed.

RC22) P 3391, lines 10-11: Using the 3rd iteration is totally arbitrary. On that basis a fair assessment of the resolving power of ERT is not possible. See comment #19 above.

- We do not agree with this point. As explained in the reply to comment #19 the choice of iteration is not arbitrary at all.

RC23) On the section “Reproduction of default resistivity values”: Based on the arbitrariness in the selection of the final model (see comment #19), the whole paragraph is pointless in my opinion (and to me the arbitrariness explains some of the inconsistencies observed regarding over /underestimation). All data sets need to be fitted exactly to the level of data noise. Different noise levels could then be considered to investigate the influence of the data noise on the recovery of specific resistivity values.

- As stated above the selection of the final model is not arbitrary. The effect of over/underestimation of resistivity values has not been observed in our study solely but also for example by Hilbich et al. (2009; Permafrost and Periglacial Processes) who applied the depth of investigation method to identify unreliable model regions. However, statements on the quantitative interpretation of ERT will be relativized.

RC24) P 3395, lines 15-22: The cited statement by Hauck and Vonder Mühll supports a very important misconception, again related to comment #19 above: The number of iterations is by no means a well-defined criterion to decide upon whether the inversion has converged, i.e. the optimization problem is solved. Whatever the number

TCD

5, C2220–C2239, 2012

Interactive Comment

of iterations is (which will normally depend on the iterative scheme used to solve the non-linear inverse problem, the accuracy in the computation of the Jacobian matrix if a gradient-based scheme is used, etc.), the mathematical task is to find the smoothest model which yields a well-defined data misfit (depending on the a-priori data noise level) (note that the task is not to minimize the data misfit! – which would never prevent from over-fitting the data, with the obvious problem mentioned in the last sentence of the citation by Hauck and Vonder Mühll). Normally this is achieved by, once a data-error weighted RMS value of 1 is reached, increasing again the value of the regularization parameter to indeed find the smoothest model. If a code is not doing this, then it is unclear which solution is actually computed and thus it can hardly be used for a quantitative ERT study!

- This comment is first of all based on the quantitative interpretability of results. As stated above we do not agree with the referee in all points with regard to the aim of our study. As explained earlier, the choice of the final model was based on the best agreement with the initial geocryologic model. While adjusting the parameters used in the study, datasets were calculated until the inversion has converged applying different noise levels. The decision to calculate 5 iterations only was based on the analysis of these tests. The effect that resistivity values of high resistive structures of the inverted models differ from those of the initial geocryologic model appeared for all iterations.

- Figure 6 will be extended (cf. reply to RC25).

RC25) P 3395, line 23: On Figure 6: What would be interesting to see in addition to the RMS curve (as a function of iteration) are the curves of the model roughness and the regularization parameter. The model roughness curve will indicate whether there is already any convergence here (note that seeking the smoothest model implies convergence of the model roughness curve). However, Figure 6 reveals another problem, which I presume is a general problem of the whole study: The RMS value at the 5th iteration is approx. 12%! This means that the inversion is not able to explain the data to the desired degree, in turn indicating that the numerical modelling errors of the used code are much larger than the assumed data noise level of 3%. This raises the very critical question on the point of investigating the quantitative imaging capability of ERT with a tool that obviously has not the sufficient accuracy. To me the only solution here is to add noise to the data which is larger in amplitude than these modelling errors. Then of course all results would suffer from this relatively large noise level and the different comparisons might become pointless.

- Figure 6 will be extended. To make results clearer we will show the results for iterations 1 – 10. We will also try to implement the suggestions given by the referee.

RC26) P 3398, line 16: Res2Dinv? How was a 3D inversion performed with Res2Dinv?

- The inversion was not performed within Res2Dinv. "As described above, the 2D arrays were collated to a quasi-3D dataset and inverted using a true 3D inversion algorithm within RES2DINV/RES3DINV". Besides the information of inverting the datasets a second step – collating the 2D datasets – is given in this sentence. The collation of 2D datasets takes place in Res2Dinv, the inversion in Res3Dinv. As this process has been explained earlier in the manuscript (P 3390, line 17-20) the description in this section was kept short. The sentence will be changed to: As described above, the 2D arrays were collated to a quasi-3D dataset (RES2DINV) and inverted using a true 3D algorithm within RES3DINV.

RC27) P 3398, lines 17-18: On which basis were the inversion parameters adjusted? Sounds again arbitrary. Why didn't you use the settings of the numerical study?

- The fact that – applying inversion parameters used for the synthetic modeling – banding effects appeared in the quasi-3D images of the field data – most likely due to higher noise levels in the measured datasets – made an adjustment of the inversion parameters necessary.

RC28) P 3398, line 26: see comment #20

- Further information will be included.

TCD

5, C2220–C2239, 2012

Interactive Comment

RC29) P 3399: What is missing here, but absolutely necessary, is the analysis of the data error. Whatever approach is used, you have to come up with some estimate of the data error in order to avoid over-fitting the data in the inversion and thus avoid misinterpretation and invalid conclusions. Given the number of severe concerns outlined so far, I decided to not provide a detailed review of the Discussion and Conclusions sections at this point. All results are too much biased by the problems pointed out above, and many of the issues reappear here.

- Our approach to avoid over-fitting of datasets, in association with the choice of the final model is explained in our reply to referee comment RC2a. The datasets have been analyzed in detail for all iterations and also with regard to the model and data sensitivity. Fundamental misinterpretations and invalid conclusions can be precluded.

- In our opinion the results are not biased at all. As discussed in several occasions above our data processing has been based on a systematic approach that allows us for precluding fundamental misinterpretations. As mentioned, we will insert additional information exemplarily for some scenarios. We will also include a figure on the evolution of resistivity values during the inversion process in consideration of the RMS error as well as a figure on model sensitivity to further support our results and conclusions.

Interactive comment on The Cryosphere Discuss., 5, 3383, 2011.