This is a concisely written, well organized and matured manuscript reporting on a study that identifies and quantifies the controls of saturated and near saturated hydraulic conductivity of the soil – and their difference, here discussed as macropore conductivity– from available soil physical and structural properties, land-use information and certain climatic information. The study is based on a decently large data collection of tension infiltrometer-based measurements, collected from the literature.

The study is well justified, as roughly all similar estimation studies (pedotransfer studies) are solely concerned of the saturated hydraulic conductivity – lacking the distinction of matrix flow and macropore flow, and based on measurements obtained using techniques that do not distinguish between those two elements. The data set collected is a rather unique collection, and will expectedly remain a valuable information source beyond this single study.

The study is carefully worked out, using ample planning and design, and data resampling for instance, and the phrasing of the conclusions is moderate and mostly well supported, and I have no doubts that this is a publication-worthy study.

We would like to thank the reviewer for these kind words and also for the many thoughtful comments and suggestions, which will help us to produce an improved manuscript.

Nevertheless, I have three significant comments/suggestions that I believe should be elaborated on – followed by some relatively minor remarks.

Major comments:

1. *Compared to the cited PTF works, this study is unique due to the collection of in situ near-saturated conductivity data. Contrasting results are presented – which was expected. There is very brief reference in the very last paragraph to the differing data set; however, I think that the issue of laboratory measurements vs. field measurements and their differences that partially explain the differences in findings should be more emphasized than it currently is.*

Yes, we understand this comment, and will expand the text in the last paragraph to more strongly emphasize this aspect of the study.

2. *Table 4 presents method-wise averages, and methodology ended up being a selection criterion for the data subsets used in the analysis. I do not recall reading about any potential interaction between method used, and certain influential soil properties, e.g. texture. Is there a chance that some methods – whether for a practical reason or by accident – have been used on certain soil textures more than others? Is there a chance that this influenced the findings? A texture triangle with method-wise colors used for the soils may be helpful to support any related discussion. I can also imagine a similar interaction with the wetting sequence (wet to dry vs. dry to wet).*

Yes, this is a good point. We checked this by applying a chi-square test to a contingency table of either method or wetting sequence against soil texture class (similar to table 6 in the paper for land use and texture class). This showed some significant bias with texture for some of the lesser-used methods (perhaps arising by chance, due to small sample sizes), but none (with 97.3% confidence) for the two dominant methods in the database (steady-state, multiple tensions, piece-wise log-linear vs. single log-linear), whose significantly different K_{10} values are discussed in the paper. For wetting sequence, the chi-square test did reveal some bias (p<0.0001). Most of this is due to the fact that dry-to-wet sequences were applied more often than expected on fine-textured soils and less often on medium-textured soils, while the reverse is true for wet-to-dry sequences. Little bias with respect to wetting sequence was found for coarse and medium-fine textured soils.

We will discuss these issues in the revised version of the paper.

3. *While I have no doubts that climate may affect K10, and the offered potential explanation of what is often seen in nordic countries (short season -> tillage may be performed in sub-optimal conditions to prevent compaction) may potentially stand, I think the statement was made a little enthusiastically, without examining (or reporting on) any correlations between the climatic factor and other factors/variables that may influence the reported finding between K10 and mean annual temperature. For instance, is there any potential pattern in preference towards a given measurement technique in the North or the South? Is there any potential relationship between the climatic-location and soil type/texture/etc. that may be influential? Is there any potential correlation between the mean temperature (i.e. the climate) and the timing of measurements and the relation between the timing of measurements and the timing of tillage operations?*

Is there a balanced number of samples for the K10 analysis from a wide range of climatic zones, or are there perhaps a few influential samples that may affect the general picture a lot? Climate information may be strongly correlated with other factors (natural or not), and may mimic other effects; or a few dominant soils may drive the relations found. I think that elaboration on this aspect of the study is absolutely necessary.

First, any correlations between location and measurement technique are not relevant for the regression model for K_{10} which was built on data obtained using exactly the same method.

We did already show the correlations of the two climate variables with other variables (e.g. soil properties) in Figure 3. These are all very weak, with the exception of temperature and clay content. The trend between these two variables (low temperatures correlated with low clay content, *r*=0.3) would work in the opposite direction to the correlation found between temperature and K_{10} (low temperature with low K_{10}).

There was quite a balanced distribution of K_{10} across the range of the climate variables. This is already illustrated quite well in Figure 3. However, a few influential studies with a large number of entries might conceivably affect the regression, because in the bootstrapping only individual data entries are excluded for validation purposes rather than all entries belonging to a given study. We investigated this by grouping all the entries of mean annual temperature into four classes (<25th percentile, 25th to 50th percentile, 50th to 75th percentile and >75th percentile). The number of studies represented in these four groups was 12, 6, 3 and 13 respectively. In other words, the extremes of the distribution were very well represented by many different studies. It can also be noted that the studies in colder climates are not only from the Nordic countries: northern states in the USA and Canada are also very well represented!

The method we employed to select a regression model (bootstrapped regression in combination with the Akaike information criterion to select the best model from all possible models) should have prevented over-fitting (see also our response to referee 1). Of course, it is theoretically conceivable that temperature is correlated with some other variable which is not included in our analysis. However, we did write that the reasons for the correlation of temperature with K_{10} '*were not clear*', and that the reason we gave in the paper was only '*one possible explanation*'. Thus, we feel that we did not make this statement enthusiastically, but rather were quite careful not to overstate the case.

Minor comments:

P10850, L6-8 and L21: *It would probably be valuable towards the justification of this fitting approach if the distribution of the fitting error was also presented and briefly commented, not only the distribution of R2.*

The fittings were done manually, and the only goodness-of-fit statistic recorded in the database is the R^2 value, so information on the distribution of the error is not available. But we did not notice any systematic trend in the error, apart from a few cases where the initial decrease of K was more gradual than is given by the model assumption of a distinct air entry value. In the future, researchers could, if they were interested, re-fit to the raw data, which is still available of course.

P10850, L20: *this estimation at ´any˙ tension is understood within the available data range, is that correct?*

No, as it says in the paper, K could be estimated at any tension, but it would not be advisable to extrapolate far beyond the available data range (which is stored in the database). As was already mentioned (see table 3), for the regression equations developed in the paper, we did allow a small degree of extrapolation ($T_{max} \geq 80$ mm for estimating K_{10} and $T_{min} \leq 5$ mm for K_s).

P10851, L21-22: *the ´250 bootstrap samples˙ expression is misleading to me. I assume it refers to the test data sets separated by the 250 different bootstrap subset selections, which yield a variable number of test samples in most cases. The text could be adjusted.*

Ok, we understand that this was a little ambiguous. We will re-phrase this in the revised version.

P10851, L23-24: *A reason and justification should be given why these exclusions took place.*

The main reason is that it doesn't make sense to mix topsoil and subsoil data when land use is considered as a predictor variable, since land use will strongly affect topsoil, but not subsoil. We would have liked to develop separate PTFs for subsoil data, but the number of entries is too small. We were afraid that the two organic soils would bias regressions that included soil organic carbon as a predictor, since they represent extreme outliers. We will comment on this in the revised version.

P10852, L14 and on: *I understand the placing of the report on hysteresis results still in the materials and methods section. Yet, I wonder if it is not better placed as the first paragraph of the results section.*

Yes, both ways of writing this are justifiable, but on reflection, we agree, and will move this text to the Results and Discussion section (although as the fourth paragraph rather than the first).

P10853, L22: *I wonder to what extent shrinkage cracks play a role between Ksat and K10. Does significant shrinkage happen between saturation and -1kPa? In my opinion, at this moisture range, tillage voids and biopores are significantly more responsible for any differences, than the moisture-status dependent shrinkage is.*

We agree that shrinkage cracks will only open up in clay soils if water tensions become much larger than 1 kPa. However, if tension infiltrometer tests are conducted on a clay soil which is initially dry, any cracks present are very unlikely to close within the time-span of the test and so they will undoubtedly strongly affect infiltration rates, providing the applied tension is small enough.

P10855, L17: *perhaps add to the effect of tillage that it is primarily due to the removal of connected biopores.*

Yes, we agree that this is a likely explanation. We will mention this in the revised version.

P10855, L22: *regarding temporal variation: Is there information in the data set about the timing of tillage operations vs. the timing of measurements? Undoubtedly, the effect of such temporal variation is logical, but the BD vs. Ks relationship is expectable in a data set without a temporal dimension as well.*

Information on the timing of tillage events is very rarely reported in the literature, so we have no data on this in the database, although we have recorded the month(s) during which the measurements were made. Yes, we agree that a relationship between bulk density and K_s might be expected even without temporal variation. We will re-phrase this text to make this clearer.