Anonymous Referee #2

Dear referee,

Thank you for reviewing our manuscript. We appreciate your comments and suggestions and have stated our comments and changes in the text below every comment.

Received and published: 31 March 2019

General comments:

The paper describes an application of electrical resistivity tomography to image structural features in the Cheb Basin, targeted to identify fluid-related structures. Its application of a large-scale survey in itself is quite novel, and the results agree well with borehole logs. Although the authors state that the main target is to image fluid-related structures, the paper really describes a more structural characterization of the Cheb basin by integrating large-scale resistivity, gravity, borehole, and geological information. While the geophysical data agrees well with the borehole logs, the contribution of the geophysics to the development of the geological model remains unclear, as the added benefit of the geophysical investigation is not clear. What also remains somewhat unclear is why the authors actually choose to use ERT? There are other, i.e. EM methods, that may be more suited for this kind of deep investigation of resistivity structures.

More generally, the logic of the paper should be improved. This is clear when considering the Figure ordering, referencing, and placing, where, e.g., Fig. 2 is referenced before Fig. 3, and Fig. 1 is about 5 pages after it has been referenced first.

Finding potential fluid- related structures, means having to characterize the area's structure, obviously. Having been successful in previous studies using geoelectrical methods in the first hundred meter (Flechsig et al. 2008 and Nickschick et al. 2015), we used an uncommon, large-scale setup (~7km profile). Of course, there is always the debate of which method to use. But, in general all EM methods including geoelectrics have as potential methods the same basic disadvantages.
We needed a method that is sensitive to fluid-induced effects at depths were borehole data does not exist (in general more than 200 meters). We focused on a depth scale of ~1000 m with a spatial resolution of 50-100 m. On a more regional scale, MT measurements (with a site spacing of 2 km) had been done (Munoz et al. 2018). However, these studies showed the problems for this area: High industrial/anthropogenic noise by lignite mining, agricultural usage with heavy machines, electrified railroads etc. Having to use farmland during crop growth season for a big part also does not allow using large coils amidst the fields and crossing the roads. In case of our setup and strategy, the specialty is an adapted statistical data processing which improves the signal-noise relation also for dipole-dipole measurements.

Specific comments:

One of the reasons for the limited benefit of the geophysics may perhaps be the large regularization factor that was applied to the resistivity inversion. This, in turn, led to a rather smooth resistivity model, which agrees well with the already existing borehole logs, but other than hinting to a basaltic intrusion, adds only limited new information. Perhaps more or an adapted data filtering may be required to help to achieve an acceptable Chiˆ2, while having a lower regularization factor. The authors are not providing any information on the sensitivity distribution or DOI index (e.g. Oldenburg & Li, 1999), which would allow to judge the reliability of the resistivity models particularly in depth. Providing a more thorough analysis and description of the resistivity models may help to improve the value of the geophysical data to the geological model development.

The chi^2 is actually acceptable, i.e. the data can be fitted within noise as Figure 7 shows. There is some misfit, but only in the large dipole separations that have large errors and low weight anyway. We are very positive that one could not derive a significantly better model without disregarding lithological information. Improving settings would just mean finding other, equivalent, models. We have tried many different regularization approaches and strengths, but ended up showing the smoothest (easiest) model that fits the data according to Occams razor. Please note that the deepest boreholes end at z=100m a.s.l. and our image is much deeper (note the different aspect ratios of Figs. 2 and 10/11).
We are now providing information on the sensitivity and DOI by alpha shading Figure 9 and 10 : see comments below.

Regarding the title, the authors state that they are investigating “fluid-related structures”. As resistivity depends on several factors, this relation from the resistivity model to fluids remains questionable. Especially given the geology of the study site, where the clay-rich Cyprus formation may well show the same response as a hydro-thermally altered rock formation.

This is absolutely correct. This is why we added the drill log data to be able to relate the resistivity distribution to actual. However, it should be mentioned that these published older data (Dobes et al. 1986) are not clearly connected with information about depths of samples or logs. Being able to translate resistivity into actual geologic information was a basic need for us and for the scientific community interested in that area. To limit the effects of parameters such as salinity, porosity, clay content, and fluid conductivity, we have added information and a table with available parameters:

“In addition to this geological constraint, we regarded the results from Dobeš et al. (1986): Their report contains valuable petrophysical information from previous studies about the different stratigraphic units in and below the Cheb Basin which we have summarized in Tab. 1. The phyllitic-granitic basement is characterized by low porosities of less than 5% compared to the sedimentary deposits on top, which feature porosities of 15-30%. Resistivity, however, may vary drastically, depending on heterogeneities within the sediments and whether fluids such as mineral waters or CO2 are present or not. For this area, Bussert et al. (2017) provides additional information. Not only do they mention the occurrence of highly mineralized water in the central part of the HMF, their geophysical log of the HJB-1 drill reveals resistivities of 5-10 Ωm for the sediments of the Cypris formation and 10-20 Ωm for the topmost part of the weathered phyllites. They are about one order of magnitude lower than the values presented in Dobeš et al. (1986) - stressing the importance of regarding the occurrence or absence of fluids even more.”

Table 1. Petrological description of the stratigraphic layers of sediments and basements below the Cheb Basin, translated from Dobeš et al. (1986)

Name of stratigraphic unit	rock type	Porosity [%]	electrical resistivity [Ωm]	
			minimum-maximum	average
Vildštein	gravel, sand, clay	30.0	14-1600	350
Cypris	clay, silt, carbonates	14.5-21.5	50-1500	-
Main Seam	lignite, sand, clay	22	7-50	15
Lower Sand & Argillaceous	gravel, sandy clay	-	3-150 (depending on saturation)	7.5
Phylliitic basement	weathered phyllite	3.2	75-140	110
Phyllite basement	unweathered phyllite	1.0	500-1800	890
Granitic basement	granite	5.0	65-650 (weathered); > 650 for unweathered	-

Technical comments:

P1, Line 6: This is somewhat confusing. Why do you require a deep drilling program to study near-surface structures? Near-surface is perhaps a subjective phrase depending on the audience.

> Now we are confused. The sentence stated that the ICDP project “Drilling the Eger Rift” focuses on the possible connection between fluids (especially the ascending CO2 of mantle origin) and the swarm earthquakes. Within this ICDP project there are several projects that explore(d) the area and 5 drill holes up to 400 m. We have changed the “near-surface” part however, as you suggested.

P2, Lines 11-12: This sentence interrupts the flow here, as in the following sentence you provide more detail on the activities described before. Also, it might be worth adding what the open questions are.

> This sentence seems to have caused several issues, we have removed it to avoid confusion with our key questions.

P3, Lines 6-7: Why is a dipole-dipole array a "special investigation strategy"? I would describe this as a standard ERT array.

> Again, we have not expressed this very well. We meant special as “specific” not as extraordinary. The basic setup is a dipole-dipole array, but the measurement strategy is different to common measurements. We use a permanently placed array of single dipoles for the voltage registrations, a moving high power current source, and a subsequent data processing of the time series of voltage/current as input for data inversion.

P3, Lines 28-30: You should reference to Figure 3 here. Section “Geology and geodynamic activity”: This section is very detailed and can be shortened by focusing on the main processes that are causing the swarms and CO2 release.

> We are shortening this in the revised version. We thought it would help the reader to understand the multi-scale effect of the fluids/CO2, but that both referees prefer a shorter paragraph and thus we have shortened it.

P6, Lines 21 – 26: Since you refer to the results here, it would be good to also show them.

We would kindly ask to look at the references provided. Repeating existing data from other studies would not be appropriate.

P8, Line 8: You refer to Fig. 3 before Fig. 2. Please revise your order of figures, which doesn't seem very logical at the moment.

We have reworked the figures. We are sorry for the order of the figures as this seems to be caused by a LaTeX error and floating figures, we apologize and fixed this.

P8, Line 13: These are good examples, but since you are referring to novel techniques, this list isn't exhaustive.

We deleted the word modern and added an "e.g." to make clear that this list is not exhaustive.

P8, Line 20: Although the practical reason is obvious to me, i.e. electrodes of the injection dipole need to be connected to each other, the theoretical reasoning is not as other arrays may achieve deeper penetration or higher resolution.

We agree, we have worded this poorly. We now added information that shows the practical reasons (agricultural usage, roads, total length of cables needed) but from the "theoretical" perspective we expected vertically oriented structures (faults, vertical fluid channels) and needed a high sensitivity towards that. Having to inject several Amperes of current over several kilometers would also be impractical in this noisy area with factories, streets and villages.

In the text you'll now find the paragraph:

"The data acquisition was performed using the dipole-dipole configuration (AB MN, with A and B being the current injection electrodes and M and N being the potential electrodes) which is, considering the cost-effect-relation for practical and theoretical reasons, most suitable for this large-scale ERT experiment. Transmitter and receiver units are physically separated on two lines reaching maximum dipole separations of 6.5 km (Fig. 1) while keeping the total length of required cables to a minimum as only neighbouring electrodes have to be connected. Considering crop growth in June in this rural area and traffic by agricultural farming machines in general, other arrays are not effective with large cable spreads of several kilometers. Furthermore, we expected vertically oriented features (faults, "fluid channels"), as seen in previous studies Nickschick et al. (2015), supporting the choice of using a dipole-dipole setup and achieving good results in previous studies at different location with a similar setup (Flechsig et al., 2010; Pribnow et al., 2003; Schmidt-Hattenberger et al., 2013)."

P12, Line 7: Do you mean that you assume that the signal is not distorted, hence has a very high signal-to-noise ratio?

Again, we have worded this poorly. Yes, we meant exactly that and have already fixed this.

"This provides correct results in case of a symmetric signal with an identical positive and negative amplitude, which is given in this case by controlling the source and assuming that the signal is not distorted by having a very high signal-to-noise ratio."

P12, Line 10: What is alpha?

Alpha is the rejection rate of samples after stacking, as is stated.

Inserted: ", i.e. the lowest and highest 10% of the amplitude distribution are removed before the computation of the mean (cf. Oppermann & Günther, 2016, Fig. 6)".

P14, Lines 1-2: Please clarify, what do you mean by this? Do you mean that you distinguished bad data points by their corresponding reciprocal error? Or do you mean that most of the bad data points have a good quality reciprocal measurement?

We meant that we do not have all data as reciprocal pair because certain current injections were not possible (vertical white columns) or certain voltage data could not be successfully retrieved. However, most of the missing data are available at least by one of the AB-MN or MN-AB combinations. See also more extensive reply to the other referee.

P14, Line 8: How do you deal with measurements that don't have a reciprocal measurement? Are you estimating an error model from the reciprocal data or are you assigning measurement errors otherwise?

As written, we estimated a percentage error of 5% and a voltage error of 2µV from the reciprocals. If a reciprocal pair is present, we took the current-weighted average. If only one was available, we took that value. The reciprocity analysis is an additional quality check compared to "traditional" surveys where only one measurement is carried out, we have in >70% of all dipole pairings another value to compare to decrease possible outliers or missing values.

P15, Line 9-10: If no error estimate is available I would suggest not including error weights in your inversion. Adding the BERT default is likely not your actual error model, and will have an impact on your inversion result.

Even an imperfect error model is better than no error model (i.e. assuming all data have equal quality independent on the voltage), since it is clear that measurements with large voltages (and low geometric factors) are more reliable than low-voltage measurements (with high geometric factors). This routine has been widely accepted in ERT. See also comment to the other reviewer and the new text about the background of the reciprocity and the interpretation of Figure 6b.

… "Therefore it can be used as a measure of data consistence and also to derive error models (Udphuay et al. 2011), however only if a statistically large number of data is available."

P15, Line 14: This is quite a large regularization parameter and will likely result in very smooth models. Did smaller values result in much higher misfits? Did you do a L-curve analysis?

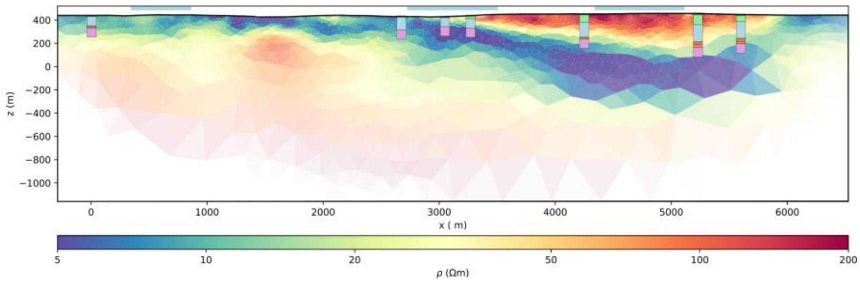
A L-curve analysis is not quite easy as the appearance of the L depends on the scaling and the range of the lambda values. We basically chose the lambda value high enough so that we could avoid artifacts (conservative approach or Occam's razor). A further reduction of lambda decreased the error only slightly and lead to more unrealistic structures in the model that was not helping the interpretation. We did a large number of different parameters with very similar results. Again, our model fits the lithological data very well for the first 300-400 meters.

P15, Line 19: This is only true if the outlier also has a high error, otherwise the high regularization factor is likely causing the smooth response.

The parts of the pseudosection that could not be fitted well are in areas of large dipole separation and thus high geometric factors and error levels (up to 20%, see above). Also, the model is in agreement from what can be derived from drill logs.

P17, Line 10: This would be more obvious if you add the sensitivity distribution, e.g. as shading.

We added an alpha shading based on the coverage for both the small and the large profiles (Figs. 9 and 10). Therefore, we also had to choose a different (rainbow-type) color map.
New Figure 10 (now Figure 8) base map:



P17, Line 11: Although most of them are not exactly on the line, could you add simplified logs to Fig. 9?

We did that as well. New Figure 9 (now Figure 7):

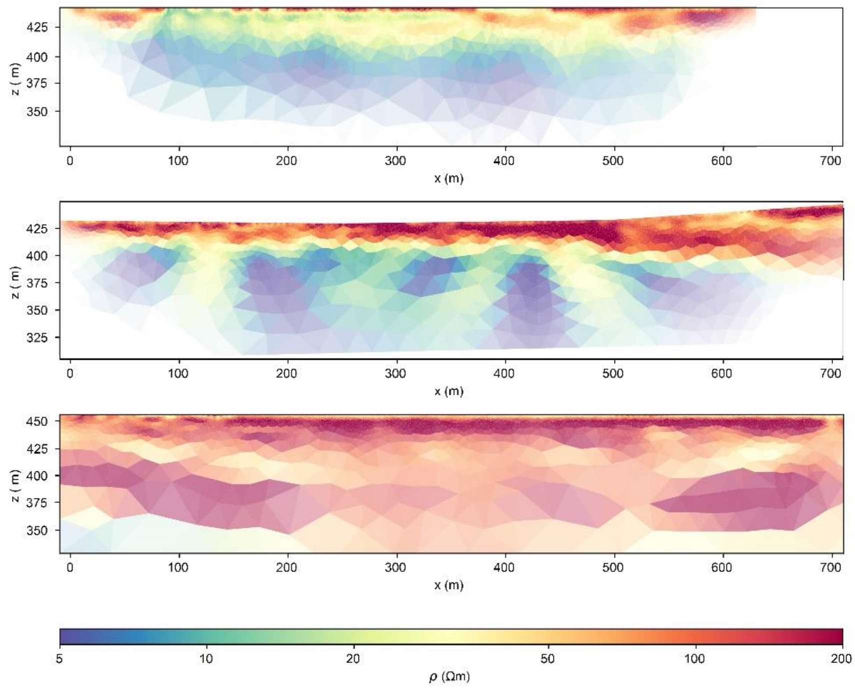


Figure 10: As for Figure 9, I suggest adding either the sensitivity distribution or calculating a depth-of-investigation index to quantify a "reliable" depth of your ERT models.

We did it (see comment and new Figure 10 (now Figure 8) above).

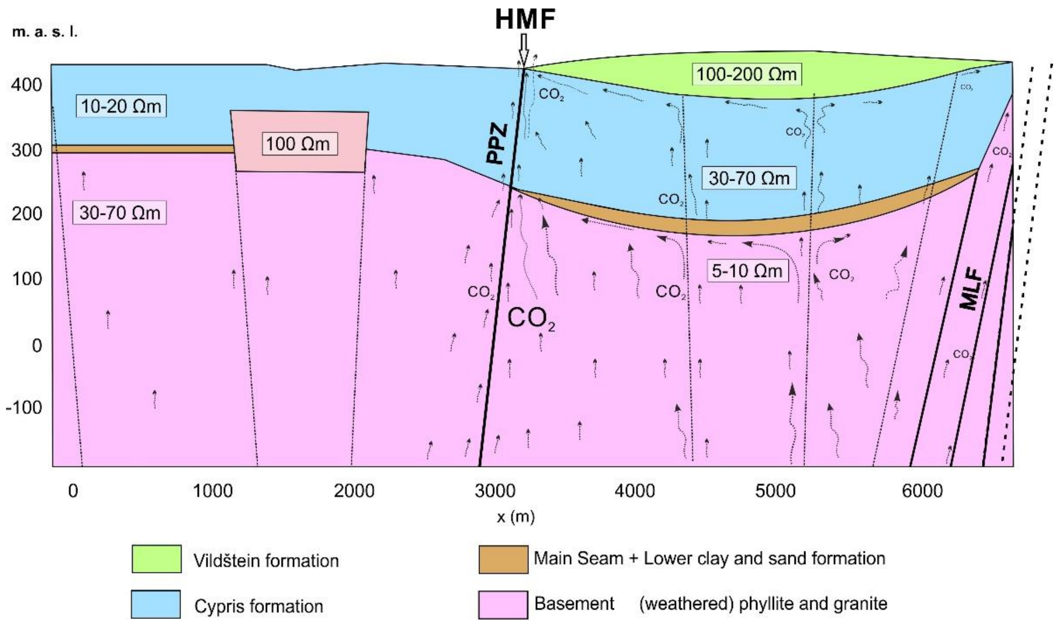
P 20, Line 11: Since you are referring to the gradient here, it might be worth plotting it as well.

As an exception we have decided to not include the horizontal gravity gradient here. Another plot for the gradient would overload the whole figure with additional, unnecessary information and the gradient could also be derived from the primary gravity curve, please take this as not to overwhelm the reader.

Figure 11: Other than the possible basaltic intrusion, what is the contribution of the ERT and gravity measurements to this model? Especially the PPZ doesn't seem to show an expression in the data.

We have decided to rework Figure 11. As you say, the impact of our survey is not visible at first glance so we changed that. We now have used the stratigraphic model but implemented the observed resistivity to show that even within the same lithologic unit it can change significantly which is new, especially for anything below 200-300 meters in the (unexpected deeply weathered/altered) basement. Having now the stratigraphy (colors) linked with the new resistivity distribution leads to our interpretation and that should be easier to grasp with the reworked figure.

New Figure 11:



P23, Lines 11-13: I don't think this conclusion is obvious from your data. Why couldn't it be related to a thickening of the Cyprus formation?

If this were the case, we would not observe a resistivity shift in the eastern part. It is very likely that here the Cypris formation is "dryer" than in the western half, which we have also underlined by adding references.