I read this manuscript with great interest. Although field-based monitoring of DC resistivity has been around for several years, recent instrument developments have resulted in this being now widely available. Several manufacturers now offer the capability for relatively long-term monitoring of DC resistivity. However, the monitoring of induced polarization (IP) presents new challenges, and when one considers spectral IP (SIP) then many additional problems emerge. Accurate measurement of SIP in a controlled laboratory setting is not trivial. To translate this to the field and consider measurements over a period of time is immensely challenging. Several researchers have argued that there could be useful information captured in a measure of electrical polarization that can help improve our understanding of subsurface processes and properties. We still do not fully understand many of the controls of electrical polarization in geo materials but as we improve knowledge of such controls there will be, no doubt, a demand for SIP monitoring solutions. It is, therefore, good to see that the authors have taken steps to develop such a system and document their initial findings. An SIP measurement is easy to make. An accurate SIP measurement is not. Without a clear understanding of some of the issues that researchers need to consider (as outlined in this manuscript) there is a danger that unreliable measurements are obtained, and inaccurate interpretations made.

The manuscript essentially documents the authors’ attempts to implement a new system for temporal monitoring of SIP at a study site in Germany – the Selhausen rhizotron facility. The authors explain clearly their rationale for their design, explain how they conducted the measurements over a period of several months and the steps taken to assess data quality and filter data. They also show resultant images (and aspects of images) that result from the time-monitored data. The authors admit that their system is
I have a few general comments and then some specific comments.

GENERAL COMMENTS

Overall, the manuscript is clearly written and includes some carefully prepared illustrations. A few of the figures would benefit from revision (see specific comments later).

The title refers to “long term monitoring”. This is, in my view, an over-used and often inappropriate phrase. I’m not sure that monitoring over several months is really “long term”. I am being pedantic here but if the authors are to use such a term then they should define what they mean by it.

My main concern about the work detailed in the manuscript is that the field experiment is not well controlled. What I mean by that is that the authors collected data and produced images but we have no idea what processes occurred in the field setting. There is mention that SIP can tell us something about soil water and even plant roots but we are not given any information about the soils under study, nothing about the internal states and nothing about what plants are grown (if any) and their characteristics. It is not a controlled experiment and I believe that this was a major oversight since we are unable to assess whether the signals in the resultant images make sense. The authors show some time-lapse results but only report on precipitation as a complementary observation. If this is a “facility” then surely other states were observed. Why are they not reported? Even time-series of soil moisture at several depths would help. What is the difference
between plots 1, 2 and 3 in the facility? I realise that the authors wish to focus on instrumentation aspects and not be distracted by internal processes in the soil but without any independent observations many comments on influences are speculative. A much more appropriate setup (in my view) would have been to have ensure much greater control of the boundary conditions (infiltration, evaporation) and internal states (moisture, temperature, presence/absence of plant roots, etc.). Nevertheless, the findings are very useful. I just think that the authors need to provide more information to allow the reader to get a better understanding of properties and dynamics of states that are controlling the observed electrical responses.

The authors explain in their discussion that the system was originally developed for laboratory scale investigations. This helps explain why they use low (+/- 9V) voltages for the source. I think it would help if there is a bit more discussion on what challenges (if any) in moving to a higher voltage source since this would help transfer this approach to a broader range of studies/applications.

I understand the logic for selecting long dipole lengths to cope with the relatively low signal source. However, it seems that the shortest potential dipole is 3.5m, which seems very long to me given that you infer something about processes at very shallow depths. The images in Figure 13 are to depths of 1.2m. What exactly is your resolution capability? I think that at the very least you need to show some sensitivity maps. I also think that you need to be careful in discussing the inferred high resolution variation in (complex) resistivity given the electrode configuration and also discuss this further.

Monitoring of SIP data will inevitably lead to periods of data collection with different acceptable measurement configurations (as illustrated in Figure 6). I think that this aspect needs some discussion. An image produced from 1500 measurements surely has different resolution capability than one based on 1000 measurements. While this is an issue for other monitoring techniques, SIP is particularly prone to dataset size dynamics.

SPECIFIC COMMENTS
Line 4. It would be useful to state the frequency range in the abstract since impedance spectroscopy means different things to different readers.

Line 4. Change “polarization” to “polarisation” to be consistent throughout.

Line 9. Define “long-term”.

Line 14. What “core”? I can work out what the authors mean but some readers may be confused by this.

Line 26. They may not always be “independent”.

Line 73. The Telford et al. (1990) reference is inappropriate regarding a statement on EIT. It would be unethical of me to propose citation of my own work but Binley and Slater (2020) [Resistivity and Induced Polarization. Theory and Applications to the Near-Surface Earth, Cambridge University Press] is a much more appropriate reference.
Line 87. Explain “small levels of systematic variations”.

Line 88. Explain “heuristics must be employed”.

Line 94. Don’t mix precision and accuracy.

Line 95. Explain “filtering are the only options” – you are implying that your (our) normal processing of errors is not useful in determining quantitative estimates of values, and yet you (we) have used this in this way many times.

Line 117. Make it clear that this is a field site, not a lab setup.

Line 118. “basement” is an odd and ambiguous term. Some explanation is needed. Perhaps a photo of the setup would help.

Figure 1. How far away is the electrode array from the edge? It looks to be about 5m. This needs to be covered so that we can understand any 3D effects.
Line 129. Remove “actual”. It is redundant.

Line 144. What “cable effects”?

Line 147. Comment on propagation of errors with superposition.

Line 161. The shunt resistor has been selected for this particular study. How would this be selected for others?

Line 172. It might be useful to reference the published work that explains the rationale for accepting the point electrode assumption here.

Line 176. So the electrode is at 10-15cm depth. Is this modelled as a buried electrode? If it is not (and assumed that the electrode is a point at the surface) then what are the implications?
Line 312. Typo in figure number.

Line 370. This is not specific to Hayley et al. (2007) – it is the standard approach that many adopt. All you are doing is correcting the cell values as anyone would do. Hayley actually proposed a more sophisticated method that corrected data not inverted values.

Line 377. Over what depths to you have soil temperature measurements?

Line 429. They increased but also decreased – what is the cause of this? You cannot justify this as due to soil drying. It looks more like an instrumentation fault to me.

Figure 11. I can’t follow what is being plotted here - there are two green dates. Some figure reworking needed and a clearer caption.

Figure 12. Delta phi ois not defined. I assume it is the phase error. So, for a given point in time do we have a range of delta phi (and RMS) for the range of frequencies? Is this what is being shown here? More explanation needed.

Line 469. This is an example of where information on plants being grown is needed. We cannot compare May and August events without knowing something about the state of the
plants.

Line 476. You don't give any information on variability in y.

Line 481. I don’t think that “intrinsic” is the right term here.

Line 481. Are the selected spectra typical? Or very carefully selected?

Figure 15. We need to see much more that rainfall!

Line 485. Are you showing an average? If so, why? It defeats the purpose of imaging.

Line 488. I agree but what this paper lacks is a clear statement of why this might be useful.
Line 495. This suggests that you need to plot some image appraisal information. You state just a bit earlier that the phase data gives information and now you say that some of this is unreliable.

SUMMARY

In summary, I think that this is a valuable piece of work that illustrates some of the challenges in monitoring SIP, and provides very useful insight into steps that are needed to assess and improve data quality. I sense that this will prove to me a useful reference article. My major concern is that the authors do not know what processes are operating in the soil (or at least they don’t show any evidence that they do) and so we do not know what is right and what is wrong. I also have some concerns about the inferred resolution of the resulting images. These issues can be addressed with relatively minor revision.

Andrew Binley

18-Feb-2022