Review of HESS Technical Note:

Simple, exact and reliable way to extract soil water for stable isotope analysis
Jiří Kocum et al.

This paper introduces yet another method to dozens of highly operational approaches over the decades to soil porewater extraction of water for stable isotope analysis. Each method proponent claims a superior and more reliable approach (not true). This new method involves a low cost means of recirculating air through a heated (105 °C) closed-system soil sample vessel coupled with an inline 8 °C condensation coil to collect the evaporated soil water until completion (e.g. > 99.5 % water recovery). The author's aim is to demonstrate reliable isotopic data (recovery) and propose its use in field studies, though it may be less suited for high-throughput applications at only 4 samples per day. Nevertheless, the pilot results appear to be promising, and further testing and replication by others is warranted to find out all of the pros and cons of this appoach.

The authors should greatly temper their enthusiastic language about top performance because they only tested a few relatively easy porous materials with a high water content (20 %) – there is no performance information on different and low porewater content (<<5-10 %), nor on high organic matter content materials, or on low conductivity clays, etc. Stick to a basic description of pilot performance of the circulating air experiment, recognizing that you have not tested all possibilities. Moreover, until your system has been tested identically and independently in another laboratory, its should remain as a pilot proposal.

Title should be tempered to describe the method as giving pilot results – avoid adding qualitative judgement (simple, exact, reliable).

Recommendation: Major revision, with attention to explaining key details. Shorten by eliminating Section 4.3 (this has been reviewed countless times).

Comments

The authors incorrectly use the terms accuracy and precision in the manuscript, which is highly confusing. In terms of accuracy, it would be more appropriate to use the term "bias" which must be clearly defined as the delta change relative to H and O isotopic offset from the experimental starting water isotopic composition prior to the extraction and analysis. All methods have some form of bias (show stats), and most soil extraction methods show a positive bias due to lack of 100 % recovery or other factors.

Regarding precision – this is mess – they are reporting precision values that are extraordinarily low (for a laser or IRMS – ±0.06 permil system precision for $\delta^{18}O$ is frankly, impossible!) and are completely unrealistic from an over systems point of view. The authors should propagate all the sources of system uncertainty, including uncertainty of the primary reference waters (VSMOW. SLAP), laboratory water standards using on the Picarro, replicates of the experimental waters used, and replicated soil porewater extractions. A more realistic reporting of precision in this case is more likely to be inline with other methods. Avoid using hyperbolic terms like "better than" or "more accurate" than other methods – simply show comparative results factually.

They should also be clear that low Bias (their Accuracy) is only achievable on experimental and manipulated test samples – there is no guarantee of that "low bias" will be obtained on any

unknown field samples given the wide range of porosities, grain size and organic matter contents within a single core or soil samples. Urge caution making such sweeping statements based on a few material types.

Technical Comments

Many of the technical aspects are concerning where no justifications are given (pros or cons, possible issues):

- Why was 105 °C used? In the literature the T range is wider and higher. What would be the benefits or concerns with other T. Once one citation was chosen.
- What happens (during the baking process) if the soil sample, especially those of clay or high fine or organic content compact/matte and seals the inner exposure to drying?
- How do you know the extraction is complete for unknown samples? Is a secondary gravimetric water content test conducted?
- How do you prevent surface evaporation of the sample during handling?
- Silcone tubing is highly H_2O gas permeable (look it up) – why was PTFE tubing not used? There cold be water loss or gain through the silicone tubing.
- What does a system "blank" look like (recirculate for 5h or overnight with no sample) – any moisture collected? – you claim there is moist air at the start – so it cannot be zero. What would be its isotopic composition if condensed from ambient air? Does the system gain over a long blank time?
- Cooling system – why 8 °C and not a more effective cryo-coolant (< 0 °C)
- From the photo in Figure 1 – the sample boxes with clip on lids do not look very airtight to me – how was airtight ensured and demonstrated?
- Extraction time – this will depend on many factors like material and porosity etc.
- It would be helpful to know what a reasonable uncertainty target for this type of work is. For example, for most hydrological studies (and historically), 2 permil for 2H has been perfectly acceptable, as is 0.2 permil for 18O. Would this not be a more objective bar to compare with? This is never going to be in paleo-climate ice core territory.

The stated Picarro performance precision is absurd – its even lower than IRMS or what the manufacturer reports. Be realistic. No mention of well-known corrections for memory or drift are given (all adding to over uncertainly budget with or without).

Never use the term "signature" for delta "values". A signature is a representation of something else. Not the case here.

In 3.1 please add the statistical test results and p-values for significance. If you say the results were depleted – is this a mean value observation or a statistically defensible statement, rather odd when the bias is considerably less that the SD. Variance increased – statistically defensible? Strongly recommend to add fully propagated uncertainty to all reported values in the Tables and Figure 4.

Regarding the claim of moisture left in the system – was this determined gravimetrically or just visual inspection of droplets?

Define IRMS and IRIS

Section 4.3 is not relevant and should be deleted. Stick to reporting the experimental pilot results – its premature to compare this to other methods when you have not even compared this method and approach in another laboratory.

Shift or bias in Table 3?

N.O. is missing from the Author Contributions.