<u>Response to - S. Kuehn – 2nd Review</u>

The manuscript also now indicates that the full data are accessible via the Pet Lab database at
https://pet.gns.cri.nz. An earlier version was also archived at EarthChem as a file submission.
Will this also be updated to the latest version of the dataset?
Yes it will, as soon as the manuscript is accepted for publication

Since the manuscript and dataset were modified in accordance with the Abbott et al best
practice document, those best practices should be cited in the manuscript.
Details and reference added into the manuscript Ln 97-99, and in SM Tables 1.1 and 1.2, we thank Dr
Kuehn for point out this oversight.

The data tables still contain many entries labeled "Below LOD." These still need to be replaced with
the actual analyzed concentrations. Note the concluding statement of Kirchmer (1994, Limits of
detection and accuracy in trace elements analysis): "When reporting data, particularly monitoring
data, the critical level, limit of detection, or limit of quantitation should not be used to censor data. To
avoid information loss and biased calculations of mean sample concentrations, all data should be
reported, together with an estimate of the uncertainty in the results. The critical level should be
provided separately as an aid in interpreting the reported results."
Unfortunately, the output from our LA-ICP-MS programme and iolite does not report these values,
we have however added text at the base of **SM Table 2** to show the LOD values for the elements
where "below LOD" is reported – we hope this is acceptable.

Line 212 – If chemistry has differences, how are two samples known to be the same tephra?
They are known to be the same through accurate 14C dating of the distal sites, the geochemistry
varies due to the potential different phases (and hence chemistries) of the eruptions captured by the
different sites. This is discussed further in the text lines 454 - 461 and presented in Figure 7.

Line 219 – Perhaps grain size and tephra samples would be better than glass-shard size and
glass samples? Unless density separates are performed to concentrate glass or the original
deposits are almost pure glass already, these samples technically aren't composed of just glass.
Edited to remove the definition "glass" line 221.

Line 438 – Do you mean two different magmatic liquids (and therefore glasses) formed due to
fractionation of different amounts of biotite?
Yes, text added lines 453-454 for clarity.

Lines 906-908 – Perhaps "most responsible for the separation in PC1 and PC2 space" ?
Changed

Line 943 – At the time of this re-review, www.tephranz.com resolves to a page with an "Oops! That
page can't be found" message. However, the header on this page does allow navigation to the data.
Thanks for checking this for us – we will look into this issue and hope to have the page up and
running again in time for the publication of the manuscript.

Excel reports a formula error upon opening the Excel file which contains the tables and
supplemental data. The problem may lie in the plots associated with supplemental Table 5.
This is now resolved – hopefully.

Additional comparative sets of analyses for VG-568 may be found in the following references. The
latter uses a high-precision approach with additional EPMA elements (Rb, Sr, Zr).
• Rowe et al., 2008, Catalog of Mount St. Helens 2004–2005 Tephra Samples with Majorand
Trace-Element Chemistry, U.S. Geological Survey Open File Report 2008-1131. (see
appendix in the pdf)
• Kuehn and Lyon, 2020, June Lake Tephra Dataset

(specifically see June Lake Glass DATA.xlsx)
We really appreciate these studies being brought to our attention and note that the values for Na2O even within these studies is variable. For example, VG-568 values in Rowe et al (2008) = 3.53 wt.% whereas in Kuehn and Lyon (2020) = 3.96 wt.%. Comparatively Kuehn and Lyon (2020) also present Na2O data for ATHO-G at 4.58 wt.% - high in comparison to those reported by Portnyagin et al (2020) discussed below. We feel that this issue with Na2O analysis is a community problem that cannot be resolved in the scope of our manuscript, but hope that the transparency in the text now allows that to be understood by future studies using the data. We also note, that very rarely is Na2O used in isolation for correlation purposes because of the known problems with analysing it.

Response to - M. Portnyagin - 2nd Review

Unfortunately, I have not received a detailed response to my previous comments.
These can be found online at: https://gchron.copernicus.org/preprints/gchron-2020-34/#discussion

On this iteration, I have checked only analytical techniques and tables. I see that the authors made some efforts towards publication. However, more detailed description of the analytical techniques revealed new problems.
We note the reviewer has only reviewed the analytical techniques and tables, and has not seen the rebuttals or changes made in the text therefore are not necessarily aware of the additional details and transparency added about some of the comments highlighted below. Regardless, we hope that the comments made below, and the links to where changes have been made in the text, allow the updated manuscript to be acceptable to the reviewer.

EMPA
The authors provide more detailed description of their EMP analyses now. This description shows some fundamental problems, which were not obvious before. The major problem is that they used rhyolite glass VG568 for calibration of ALL elements, including Ti, Mn, Mg and Cl, which all occur in VG568 in concentrations below 0.1 wt%! Therefore, the authors obtained large variability of data for secondary standards, where these elements are more abundant. The best example is MgO in basaltic glass VGA99 (Figure in SM Table 3). The measured concentrations range from 4 to 6 wt%! Extremely poor accuracy. In essence, this figure illustrates that calibration of MgO on VG568 was not precise and varied between 7 sessions, as expected for such low concentration in standard and therefore imprecise peak positioning. Similar problems are evident for Mn, Ca, Ti, all minor elements poorly characterized in rhyolite.
See comments below where this is discussed in more detail by the reviewer.

The authors decided to use data for VG568 from Streck and Wacaster (2006) as reference values ("Streck" – not "Stracke" in tables)
Changes made to the spelling error throughout the text, tables, and figures.

Why? Simply because Streck and Wacaster reported some data for minor elements. These are not certified values. Moreover, these authors used 10nA and 5 um beam to analyze rhyolite glass, giving high electron beam density, and not surprisingly their Na2O is even lower than reported by Jarosevich et al. (1980).
(1) See comments about our decision to use the Streck and Wacaster 2006 data as our reference in original response to this reviewer, and the text lines 363-373 (added in first revision).
 (2) See also the additional supplementary data added in first revision (SM Table 6) where we compare the data from Streck and Wacaster 2006 with the data from Jarosewich et al (1980) and show there is a negligible difference between their values (apart from Na2O, which is discussed further below).

Na2O is also lower for secondary standard ATHO-G, and it must be higher as it has been discussed in literature already and mentioned in our previous reviews and by the editor.
See comments in more detail below where this issue is discussed in more detail by the reviewer.

I see that counting times for elements are not reported in tables. This was requested.
Counting times are reported on sheet SM Table 1 Analytical set up > SM Table 1.1(b) "EPMA machine set up" - this change was made during the first revision of the manuscript.

Thus, the data suffer from inadequate calibration and imprecise standard values. Normally, I would recommend rejection of manuscript which reported such inadequate EMPA technique. However, I understand that most glasses studied in this work have very low Ti, Fe, Mn, Mg and Cl, and the accuracy of the analyses reduced to 20 rel.% for some of these elements is comparable with single point precision. These elements seem not very important for the regional systematics. The analyses are not perfect but, I guess, still usable for correlation of tephras in this region.

I suggest the authors to do next iteration of their corrections.

Si, Al, Na, K should be re-calculated using reference data from Jarosevich et al. (1980) or even better ATHO-G, which is dry glass less affected by beam damage. Ti, Fe, Mg, Mn must be corrected using basaltic glass VGA99, where these elements are much more abundant. Cl can be kept as it is or slightly corrected using data from Portnyagin et al (2020) who reported average of c.2700 analyses obtained using Cl-rich scapolite as standard. What are these strange numbers written for "Cl" in Table SM6?

We are unsure what the reviewer is suggesting here for "correction". In the first iteration of this manuscript, the details of the calibration and data reduction process were outlined in detail. For clarity, transparency, and in a bid to reduce misunderstanding we list below our data handling process (which can be found in the text Sect. 2.3):
(1) EPMA is calibrated using a range of standards to determine peak positions the standards, crystals, count times and channels used are specific for each element and these details are list in SM Table 1, Table 1.1 (a) and (b) after "Best practises" published by Abbott et al., 2021. We believe this is possibly a source of confusion in the above comments?
(2) During the analysis VG568 is run as a primary standard, VG-A99 and ATHOG are run as secondary standards. For this data collection all samples are rhyolitic, therefore ATHOG provides the most appropriate secondary standard for this process.
(3) EPMA undertakes an online ZAF correction
(4) A secondary offline data reduction process is performed using the primary standard VG568 for all elements. This process involves calculating a correction factor for each element's variation from the reference value for VG568, then applying this correction factor back to all the data.
(5) Sample data are corrected for difference from 100 wt.% total with difference reported as "H_2O_D" to allow back calculation to the original values.

In case we have misunderstood the reviewers request, we have undertaken the secondary offline data reduction process using both the Streck and Wacaster (2006) and Jarosewich et al. (1980) reference values, and as suggested above, with different elements calibrated using different standards. We have kept Cl as it is (as suggested), SiO_2, Al_2O_3, Na_2O, and K_2O are calibrated using VG568, and TiO_2, FeO_t, MgO, MnO are calibrated using VG-A99. We note that Ca was not discussed in the comments above, but based on the concentration of this element and text in Portnyagin et al., (2020) we have

used VG-A99 for Ca. We also note, that where VG-A99 is used as the primary standard it would be a circular argument to look at the VG-A99 output therefore for the mixed calibrations, we have just presented ATHO-G values as a secondary standard.

For clarity:
Calibration 1 – VG568 for all using Streck and Wacaster 2006 reference values
Calibration 2 – VG568 for all using Jarosewich et al., 1980 reference values
Calibration 3 - The mix outlines above using Streck and Wacaster 2006 reference values
Calibration 4 - The mix outlines above Jarosewich et al., 1980 reference values
We have uploaded this as an additional file, but do not intend on having this as a supplementary addition to the published manuscript.

The results show that when VG568 is used as a primary standard the reference values from Streck and Wacaster (2006) result in values for our secondary standards (VG-A99) and (ATHO-G) that are more aligned to the reference values – the preferred range on GeoREM for VG-A99 and Jochum et al., 2006 values for ATHO-G. When the mixed values are used as detailed above (VG568 for SiO_2, Al_2O_3, Na_2O, and K_2O; and VG-A99 for TiO_2, FeO_t, MgO, MnO, and CaO) there are negligible differences between our original results and the output from the new results. See sheet "2nd std compare" and figures therein. We also note that the impact of these different reference values on the sample values are also negligible – see sheet "sample compare".

For these reasons, and the transparency in our methods, the publication of all the reference values used within our methods, and the details highlighted in the text we feel that it is unnecessary to recalibrate the entire dataset.

Na2O is definitely too low in the entire dataset because of too low accepted Na2O in primary standard after Streck and Wacaster. Correction using values from Jarosevich et al. (1980) for VG568 or GEOREM for ATHO-G data will bring the results to the lower range of accepted values. Lowe et al. (2017, page 8) – some co-authors of this manuscript- mentioned that GEOREM Na2O=3.75% for ATHO is likely too low. Portnyagin et al. (2020) suggested Na2O =4.1-4.2% in ATHO based on large set of EMP and LA-ICP-MS data. The authors may consider these results for further correction.
We recognise what is being pointed out here, and agree that depending on the reference data used Na_2O is one of the most impacted (e.g. for ATHOG; 3.85 wt.% (Streck and Wacaster 2006) vs. 4.1 wt.% (Jarosewich et al., 1979)). We have added a further note in the text (line 389-397) to make this concern more apparent, however, the data for our secondary standards are currently in agreement with the preferred value on GeoREM of 3.75 wt% (after Jochum et al., 2006). We feel that going into detail about this reference data issue is beyond the scope of this manuscript, but hope that the additional text, and the transparency in the methods used are now more acceptable?

LA-ICP-MS.
There is still poor agreement between Ti in EMP and La-ICP-MS data. The authors probably misunderstood that the data falling out from the error envelope must be excluded from tables. This was not done. The data for Ti in standards are missing (SM Table 4), but I assume that 20 rel % is acceptable deviation. Larger deviation is symptomatic of contamination of analysis by mineral phases. The data should be excluded from tables.
We have added Ti values back into the standard SM Table 4 as requested, and removed samples with Ti and TiO2 values which fall out of the error envelope (SM Table 6), the R^2 value for this relationship is now 0.71 (Figure SM 6.2.1).

Sc data for secondary standards is strongly affected by interference with SiO, as already said before and unevoidable at such high ThO/Th ratios. This is why they have up to 80 rel% (5 ppm) deviation from reference data for StHs6/80. This data should not be reported. The authors should admit that they likely had strong addition from SiO but because their samples have similar SiO2 content with ATHO, the data can still be usable. In other words, the relative contribution from SiO is similar for ATHO and unknown rhyolite samples. The data is however not usable for samples with lower SiO2

such as BHVO-2G or StHs60/8-G. Perhaps, other authors also reported ThO/Th>1%. It does not mean that this is good practice and that Sc data is not affected by oxides. It means that the authors run instrument with not optimal conditions and could generate poor results for some elements

Text, figures, and additional supplementary data (**SM Figure 6.2.5**) was added to highlight this problem in the first review of this manuscript, within which we comment on the impact SiO can have on Sc, and BaO can have on Eu. We also highlight our method to monitor and check this issue (lines 398 – 408). We have added in text to further highlight this issue, and add transparency to the data quality as suggested above (lines 403 to 406).