This paper presents a methodology to derive a nonlinear calibration curve for the relation between the observed signals and the temperature in the pure rotational Raman technique. The classical linear calibration curve is exact only if we use 2 single rotational Raman lines. In practice most lidar systems select 2 spectral bands with several Raman lines in each band and the linear calibration curve is no more exact. It is then valuable to try to improve the results taking into account nonlinear terms in the calibration curve. The application of nonlinear calibration curves to true lidar data show that it improves the accuracy of the comparison with external temperature data. However they are some weaknesses in the computation of uncertainties developed in Appendices A0 to A3 that makes the paper unpublishable in its present state and I recommend a major revision. The reasons for this recommendation are explained below.

The Formula (A2) giving the uncertainty on the ratio $Q$ between the two Raman lidar channels is not correct. As the signals in the two channels are independent of each other, the uncertainties should not be summed linearly but quadratically. The derivation of $\Delta T$ using Formulas (A3) to (A8) is therefore also not correct. Surprisingly the Formula A9 giving the uncertainty on $\Delta T/T$ is correct but I don't understand how it is possible to derive it from (A8). As a consequence there is an inconsistency in the experimental results on $\Delta T$ and $\Delta T/T$ presented on Figures 6 to 10 and 12 to 14. The ratio between $\Delta T$ and $\Delta T/T$ should be equal to the temperature $T$ that varies between 270 K and 205 K in the altitude range covered by the lidar. The ratio on the Figures seems to be more in the order of 120K, with for instance $\Delta T/T=0.005$ and $\Delta T=0.6K$. The same mistake exists also in Appendices A1 to A3.

Concerning the experimental results, the estimation of the temperature difference with the reference data CPAC is not affected by the uncertainty computation and can be considered as valid. It is clear that a nonlinear curve gives globally better results than the linear curve. It is especially true in the lower part of the atmosphere. However it is not so clear that it improves also the results above 8 km. In some cases the linear curve gives better results than the nonlinear ones, for instance at 9 km. Do the authors have an explanation for that and is it necessary to apply a nonlinear curve in the full tropospheric range or only in the lower part?