Review of: "Infrared Spectroscopy (FT-NIR) and t-Distributed Stochastic Neighbor Embedding (t-SNE) as an Analytical Methodology for Rapid Identification of Tea Adulteration"

Andrei A. Bunaciu[1]

1 Independent researcher

Potential competing interests: No potential competing interests to declare.

The present comments are related to '**Infrared Spectroscopy (FT-NIR) and t-Distributed Stochastic Neighbor Embedding (t-SNE) as an Analytical Methodology for Rapid Identification of Tea Adulteration**". FT-NIR spectroscopy stands out because it offers a short analysis time, low operating cost, repeatability, accuracy of results, absence of preparatory steps for the material to be analyzed, non-destructiveness, and the fact that it does not use chemical solvents or toxic compounds, which are common in traditional analyses.

The objective of this study was to use Fourier Transform Near-Infrared Spectroscopy (FT-NIR) to characterize and identify the vibrational bands related to the main chemical groups present in the leaf, petiole, and stem of three types of medicinal plants used in the production of teas. The analyses were performed for Chamomile (*Matricaria recutita* L.), Ginseng (*Panax Ginseng*), and Quebra-pedrass (*Phyllanthus niruri*).

It is important to highlight that t-distributed Stochastic Neighbor Embedding (t-SNE) is a nonlinear statistical method used for visualizing high-dimensional data, widely applied in the analysis of complex data with a large amount of information, and generally presenting better performance than Principal Component Analysis (PCA), which is a linear statistical method aimed at reducing the number of dimensions while maintaining the variance of the data set.

The chemical composition observed among the different repetitions is quite similar, with bands that have approximately the same amplitude and the same wavenumbers.

This study demonstrated that the combination of FT-NIR spectroscopy and multivariate analysis techniques constitutes an analytical methodology capable of forming groups within which the spectroscopic and vibrational characteristics that define the chemical markers of the medicinal plants under study are preserved. This is important because spectra obtained from complex samples or those that do not belong to these groups clearly indicate possible adulteration, suggesting fraud. It was possible to differentiate samples obtained from leaves from those obtained from leaf stems and petioles (pieces).

Among the grouping methods used, t-SNE showed the best separation capacity between the samples, as the samples formed graphically well-defined and distant groups, making their identification unequivocal. The deconvolution of the spectra showed that a total of fifteen Gaussians was sufficient to characterize the main vibration frequencies of the

chemical groups present in the samples. These frequencies represent a variety of vibrational modes, such as bond vibrations, combinations, and overtone bands. The study constitutes an approach that has great potential for quality control by industries and government agencies.

I recommend this manuscript for publication in your journal.