HORIBA

Raman

Skin biomarkers & depth profile: PCA and PARAFAC correlations from HORIBA RZ-660 skin probe analyzer

Application Note

Cosmetics RA-108

Igor Carvalho[1], Felipe Siqueira[1], Hichem Kichou Ph.D[2]

[1]HORIBA Instruments Brazil, São Paulo, Brazil,

[2]HORIBA FRANCE SAS, 14 Boulevard Thomas Gobert - Passage Jobin Yvon, CS45002 - 91120 Palaiseau, France

Abstract: With the incoming regulatory rules for cosmetics and dermo cosmetics, it is imperative to properly identify and characterize the interactions between the cosmetics and skin biomarkers in a non-invasive method with fast and reliable results. Raman spectroscopy is an excellent and powerful tool for analyzing skin biomarkers and their depth profile because it's non-destructive, non-contact, fast, and it can measure numerous properties. Raman spectroscopy is also a selective technique to differentiate the characteristics of different skin profiles and integrated statistics tools, such as principal component analysis (PCA) and PARAFAC (Parallel Factor Analysis), which are powerful chemometric methods for breaking down complex, multi-dimensional (2D and 3D) data, leading to simpler, underlying chemical components and their concentrations, and helping identify and quantify different substances in mixtures, even when they overlap. This opens new possibilities for Raman spectroscopy to be used in clinical panels to address to the requirements of regulatory agencies.

Keywords: Raman spectroscopy, RZ-660 skin probe, PCA, PARAFAC, skin biomarkers, skin depth profile, chemometrics

Introduction

Raman spectroscopy is an essential tool for the characterization of skin biomarkers and depth profiles as it provides non-destructive, high-throughput, and versatile measurements. Skin biomarker identification and depth profiles can be acquired up to 200 microns in depth. Preliminary analyses were carried out in volunteers and consolidated into 6 packages of data treated by PCA and PARAFAC. Identification of adenosine, elastin, melatonin and serotonin were carried out on a LabRAM Soleil™ Raman spectrometer. Skin depth profiles (not aiming to reach such biomarkers) were also carried out with the RZ-660 Raman spectrometer, aiming to verify the performance from the RZ-660 in identifying the differences between the volunteers' skin profile.

Experimental and Results

The data shown in this note were obtained using a HORIBA LabRAM Soleil spectrometer equipped with 405, 532, 633, and 785 nm excitation lasers in backscattering geometry. In this note, the 785 nm laser was used for excitation, for the biomarker's identification, and a 660 nm laser for skin depth profile. The other acquisition parameters were (i) 1 s integration time, (ii) 600 gr/mm grating, (iii) 50x ELWD 0.6 NA objective, resulting in a spot size slightly below 2 µm, and (iv) 160 kW/cm^2 laser power density to limit heat-induced peak shifts (<1 cm^{-1} is acceptable when mainly relative peak intensities are considered). The skin depth profiles were carried out with RZ-660 Raman spectrometer, with (i) 3 s integration time, 3 accumulations with auto focus mode off (ii) 100 mW 660 nm laser, (iii) 830 gr/mm grating objective, 100x corrected air objective, 200 µm piezo Z-focus adjustment stage, nozzle with 300 µm thick CaF optical window for contact measurement with skin, with precise control of the incident skin layers, which is critical in these experiments.

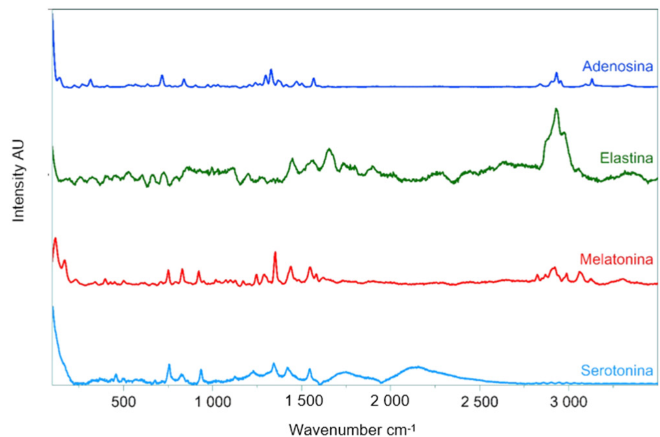

Figure 1: Skin biomarkers identified by Raman Spectroscopy

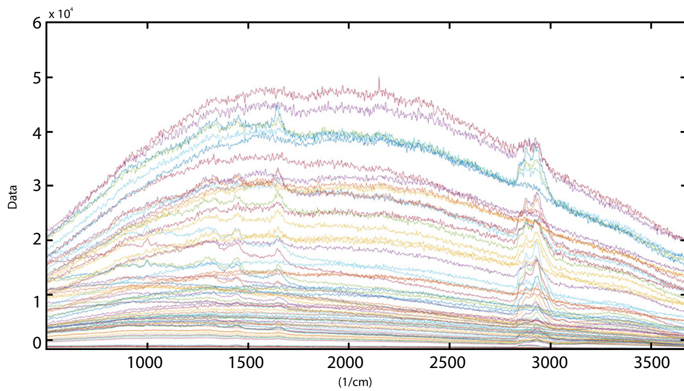

Figure 2: Raw data taken from skin depth profile.

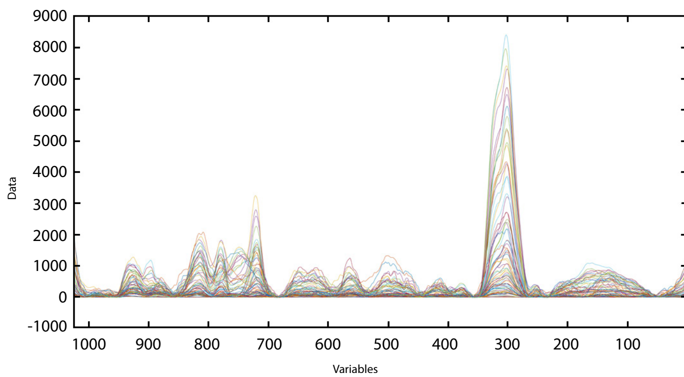

Figure 3: Treated data from skin depth profile.

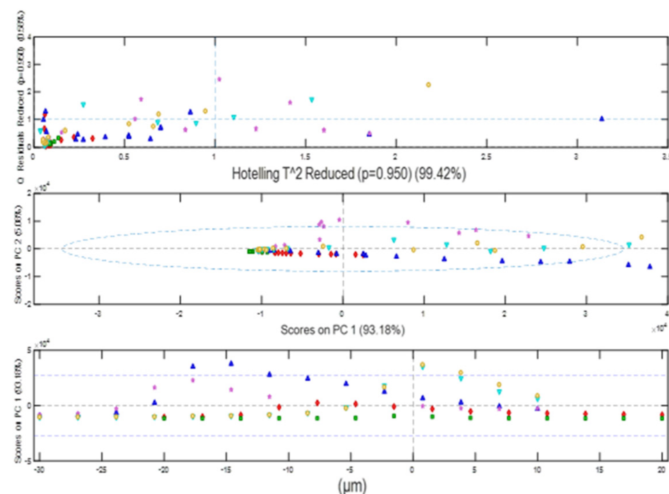

Figure 4: PCA diagnostics for the Raman Skin Analyzer as a function of focal depth.

Figure 4 shows the PCA diagnostics for the Raman Skin Analyzer as a function of focal depth for six data packages taken from volunteers. The top plot (reduced Q-residuals vs. reduced Hotelling T^2, 95% limits) shows that all spectra fall inside the confidence region, indicating a good model fit with no relevant outliers. The middle scores plot (PC1 vs. PC2; 93.2% and 5.1% of explained variance) forms a compact ellipse within the Hotelling contour, with strong overlap between subjects, meaning that the main variation is systematic and not subject-dependent. The bottom plot displays PC1 scores versus focal position (µm); PC1 changes consistently with depth and shows similar trends for all individuals, indicating that PC1 captures the instrument's focal response and that the analyzer discriminates focal planes in a stable and reproducible way.

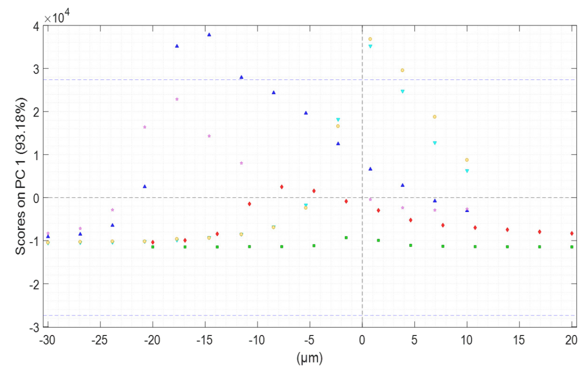

Figure 5: Evolution of the first principal component (PC1, 93.18% of explained variance) as a function of the axial position of the laser focus (µm).

Figure 5 shows the evolution of the first principal component (PC1, 93.18% of explained variance) as a function of the axial position of the laser focus (µm) for all volunteers and focal planes. PC1 behaves as a global focusing index of the Raman signal. As the probe is moved along the z-axis, the scores change systematically rather than randomly: Spectra acquired at similar depths group together, and the curves for different individuals follow the same trend. Around the nominal focus (vertical dashed line at 0 µm) the scores cross the central axis and then increase, indicating maximum coupling of the optical system and stronger, more structured Raman contribution. At more negative or more positive displacements, the scores move away from this central region and approach the horizontal control limits, consistent with loss of optimal focus and increasing dominance of background contribution. The overlap between subjects at each position, and the fact that all points remain within the horizontal dashed limits, indicate good reproducibility of the focal response and absence of outliers or instability. In practice, this means PC1 can be used as a robust scalar parameter to monitor and control focusing small changes in depth produce consistent and predictable variations in PC1, showing that the Raman RZ-660 Skin Analyzer discriminates axial positions in a stable and repeatable way across different individuals.

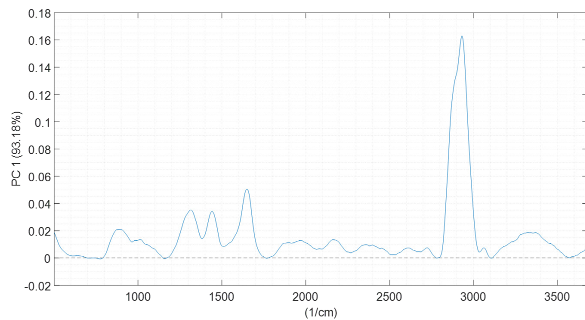

Figure 6: Loading vector of PC1 (93.18% of variance) for the Raman skin dataset.

Figure 6 shows the loading vector of PC1 (93.18% of variance) for the Raman skin dataset. The PC1 loading is dominated by well-defined positive bands in the fingerprint region and a very intense positive band around the CH-stretching region (~2,800–3,000 cm^{-1}), with only a small, almost flat, baseline contribution. This means that PC1 describes coordinated intensity changes of the true Raman bands rather than baseline drift or noise. Because all major skin-related bands have positive loadings, an increase in PC1 scores (as seen in the depth-profile plot) corresponds to a simultaneous increase of these Raman features—i.e., a stronger, more structured biochemical signal from the tissue (lipid/protein CH vibrations, amide and other fingerprint modes). Conversely, lower or negative scores indicate spectra where these bands are globally weaker and the spectrum is closer to background. In practical terms, PC1 behaves as a global “Raman signal content/focus” axis. Moving along PC1 scales the intensity of genuine tissue bands, which is exactly what is modulated when the probe goes in and out of optimal focus.

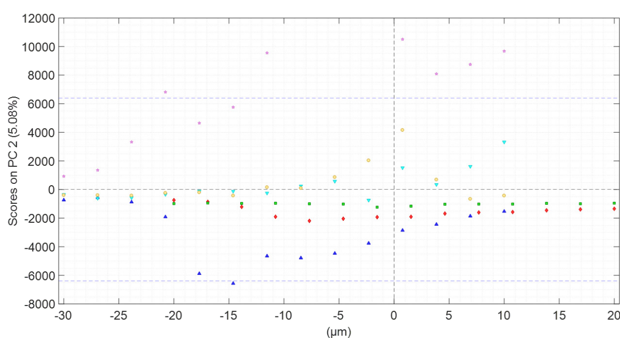

Figure 7: PC2 scores (5.08% of variance) as a function of depth for all subjects.

Figure 7 plot shows the PC2 scores (5.08% of variance) as a function of depth for all subjects. Unlike PC1, the dependence on depth is weak and asymmetric: Most subjects remain clustered at moderately negative PC2 values across all z-positions, while one subject (magenta symbols) presents systematically positive scores, especially near and above the nominal focus. This indicates that PC2 captures inter-individual spectral differences (e.g., composition/skin condition or acquisition-day effects) superimposed on the global focusing trend described by PC1, rather than a pure focus effect itself. All measurements lie within the control limits, so these subject-specific variations are stable and do not behave as outliers.

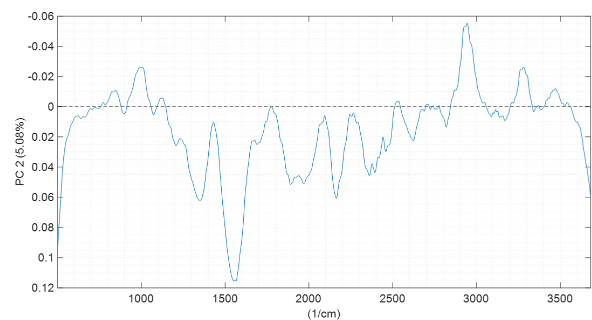

Figure 8: Loading vector of PC2 (5.08% of variance) for the Raman skin dataset.

Figure 8 shows the loading vector of PC2 (5.08% of variance) for the Raman skin dataset. PC2 displays a clear bipolar pattern, with alternating positive and negative bands across the fingerprint and CH-stretching regions. This means PC2 does not represent a uniform increase of the whole Raman spectrum (as PC1 does), but rather a redistribution of intensity between specific bands: spectra with positive PC2 scores have some bands enhanced while others are relatively suppressed, and the opposite occurs for negative scores. In practice, PC2 encodes subtle spectral contrasts linked to inter-individual differences or local biochemical variations (e.g. relative lipid/protein contributions or background structure), which explains why one subject separates along with PC2 while depth has only a secondary influence.

In the PCA, the first principal component explains about 93% of the variance and behaves as a global focusing index of the system. As we scan in depth, PC1 scores change in a systematic way for all volunteers, showing that the main source of variation is the strengthening and weakening of the true Raman bands as we move in and out of optimal focus. The corresponding PC1 loading is essentially a clean skin Raman spectrum, with strong fingerprint bands and a dominant CH-stretching peak, which confirms that PC1 represents the real biochemical signal rather than baseline or noise.

HORIBA

PC2, which accounts for about 5% of the variance, does not follow the in-depth trend. Instead, it separates one volunteer from the others, and reflects a contrast between specific bands in the spectrum. This tells us that secondary variation is subject-specific – for example small biochemical or background differences – and not an instability of the equipment. Importantly, all samples remain within the statistical limits of Hotelling T^2 and Q-residuals, so the model is stable and there are no problematic outliers.

With PARAFAC (from figure 9 to 11) we analyzed the same dataset using a three-way approch: subject × depth × spectrum. The diagnostics indicate that a two-component model is enough: The core consistency is around 98%, the residuals are small and structure-free, and there are no highly influential samples. The first PARAFAC component mirrors what we saw in PC1. In the spectral mode it looks like a typical skin Raman spectrum, and in the depth mode it shows a smooth profile with clear maxima at specific focal positions. In the sample mode, all volunteers have a strong contribution from this component. So, Component 1 represents the common biochemical signature of the tissue and its depth-dependent response.

Figures 9 to 11: PARAFAC three-way analysis: Subject × depth × spectrum indicates that a two-component model is enough with core consistency at around 98%, the residuals are small and structure-free, and there are no highly influential samples.

PCA – overall conclusion

PC1 (93% variance) acts as a global Raman signal / focus index: Scores change smoothly with depth and are highly reproducible across subjects. PC2 (5% variance) captures secondary, subject-specific spectral differences, not instability of the instrument. No relevant outliers: All spectra fall within Hotelling T^2 and Q-residual limits, indicating a stable and well-behaved optical response.

PARAFAC – overall conclusion

A 2-component PARAFAC model with high core consistency (~98%) adequately describes the dataset. Component 1 reproduces the typical skin Raman spectrum and its smooth depth profile, confirming a dominant, common biochemical signal. Component 2 isolates localized subject/depth effects, in agreement with the PC2 behavior. PCA and PARAFAC converge to the same picture: The Raman Skin Analyzer shows a robust, reproducible response to depth, with secondary but consistent inter-individual variations.

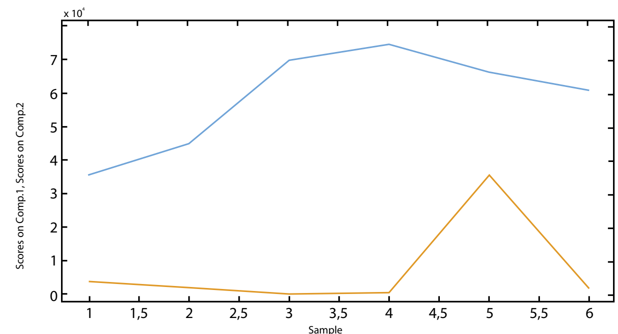

Figure 9

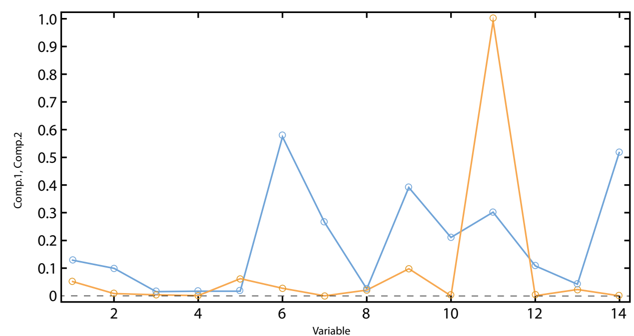

Figure 10

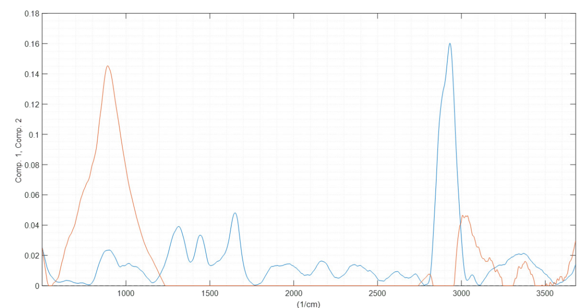

Figure 11

Conclusions and perspectives

The second component is much more localized. It is mainly expressed in one subject, and at specific depths, and it emphasizes a different set of bands in the spectrum. This matches the role of PC2 in the PCA: A secondary, inter-individual effect superimposed on the dominant focusing behavior. Taken together, PCA and PARAFAC converge to tell the same story: The Raman Skin Analyzer delivers a robust and reproducible response to depth, dominated by a single, physically meaningful component related to the true skin Raman signal, while secondary but consistent components capture subject-specific and localized spectral differences without indicating any instability of the instrument.

info.sci@horiba.com

www.horiba.com/scientific

USA:	+1 732 494 8660	**Italy:**	+39 06 51 59 22 1	**Singapore:**	+65 (6) 745-8300
France:	+33 (0)1 69 74 72 00	**Japan:**	+81(75)313-8121	**Taiwan:**	+886 3 5600606
Germany:	+49 (0) 6172 1396 0	**China:**	+86 (0)21 6289 6060	**Brazil:**	+55 (0)11 2923 5400
UK:	+44 (0)1604 542 500	**India:**	+91 (80) 4127 3637	**Other:**	+33 (0)1 69 74 72 00

This document is not contractually binding under any circumstances - Printed in France - 01/2026
©HORIBA FRANCE SAS, Société par actions simplifiée au capital de 7 075 168 €, dont le siège social est situé au 14 Boulevard Thomas Gobert, 91 120 Palaiseau, France, enregistrée au RCS de Evry sous le numéro 837 150 366
Imprimerie RGI - Parc Entreprise Esplanade - 6 rue Paul-Henri Spaak - 77462 Saint-Thibault des Vignes - Tél. 01 60 07 26 59