Embracing Diversity: Bridging the Gap for Curly & Very Curly Formulations with In Vitro and Machine Learning Performance Prediction

Barros, Deborah [1*]; Conrado, Bruna[1]; Guimarães, Fernando[1]; Abreu,Celso [1]; Maxiqueira, Fatima [1]; Petali, Erika[1]; Santos, Fabiana[1]; Oliveira, Nathalia [2]

[1]L'Oréal Brazil Research and Innovation; Ilha do Fundão 00, Ilha do Bom Jesus – Cidade Universitaria - Via Projetada 1 do PAL 48212, n° 2100
[2]L'Oréal Research & Innovation, Chevilly-Larue, France

* Deborah Netto de Barros, L'Oréal Research and Innovation, Rua Cento e Treze 339, 21941-590, Rio de Janeiro, 55 21 984473738, deborah.barros@loreal.com

Abstract

This research aimed to develop an innovative and objective method for evaluating hair curls definition, addressing the need for inclusive methodologies that cater to the specific needs of curly consumers. An Atlas of Curls was created based on the evaluation of standardized images of hair swatches by curly consumers. Hair samples of brazilian type III and multi-ethnic hair were treated with various products and then polarized images were acquired by Rumba equipment to obtain parameters like STDev, volume and alignment coefficient. A paired analysis was conducted with volunteers who rated the images based on curl/ shape definition. The data generated was statistically analyzed following Bradley-Terry-Luce model to generate a visual Atlas with the ranked images. Combining the parameters measured and paired analysis results, a predictive mathematical model was generated using a Support Vector Machine (SVM) algorithm. Results demonstrated a strong correlation between consumer perception and measured parameters for type III-IV hair. However, type V hair requires further refinement due to diverse curl shapes. This novel method integrates consumer feedback, enabling more inclusive and accurate assessments of hair care products for various hair types.

Keywords: Curly hair; Very Curly hair; Instrumental Evaluation; Machine Learning.

Introduction.

Brazil is a country known for its rich diversity, and this is reflected in the variety of hair types present in its population. Regarding hair texture, all eight curls pattern are represented in Brazilian population, and over 50% of the population has curly hair types III-V [1,2].

The physicochemical differences between hair of different curvatures impact the relationship that a consumer has with their own hair. Individuals with different hair types may perceive signs of damage differently, for example [3]. Furthermore, cultural and geographical factors will also influence the needs and desires of consumers in relation to their hair.

According to our bibliographic review, it seems that traditional evaluation in vitro methods for hair care category have been predominantly centered on straight hair, side-lining unique needs of curly hair, and making these consumers feel underrepresented. Therefore, creating a specific method to ally consumer feedback with the best product performance is paramount for promoting inclusivity/diversity in our industry.

When it comes to hair styling, Image analysis has emerged as an effective tool for objective evaluation of hair product performance, allowing for quantification of parameters such as volume, frizz, and alignment coefficient. However, the mere measurement of these parameters fails to encompass the inherent subjectivity in individual perception of the desired end look.

Curl definition, for instance, presents a particular challenge in this context. The term "definition" can hold different meanings depending on the consumer's hair profile. The interpretation and expectation of the result vary considerably among individuals with different curl types, directly influencing the perceived effectiveness of the product. There is a clear need to integrate quantitative analysis methods with a deeper understanding of the subjective factors influencing end-user satisfaction. The correlation between measurable parameters and subjective user perception presents a challenge to be overcome in the development of hair products and technologies.

The aim of this study was to develop an innovative and objective in vitro Atlas of Curls based on human perception. To engage fast innovation, an AI algorithm with data acquired, thus predicting Atlas results for any new product to be developed. Figure 1 represents a Schematic diagram of the study stages.

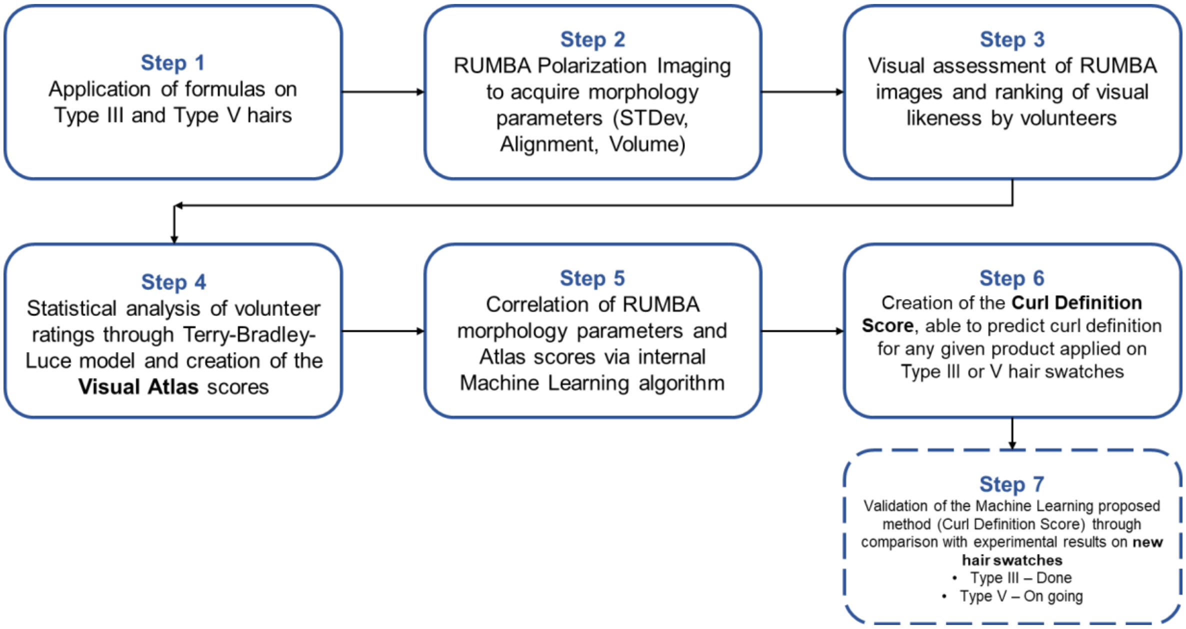

Figure 1 - Schematic diagram of the study steps.

Materials and Methods.

To achieve the goals proposed by our current work, several leave-on formulations were applied to Brazilian Natural Curl Pattern III and Multiethnic Curl Pattern V hair swatches, as provided by International Hair Importers.

A total of 20 leave-on hair formulations, both consisting of internal prototypes and market-available formulas, were tested on Type III hair, while a total of 29 formulations were tested on Type V hair. For each hair swatch and formulation pairing, the leave-on was applied according to the ratio 0.15g of product/ g of hair. The application technique was standardized across all samples to minimize variability. Each condition was tested in triplicate to account for potential variations.

Following application, the treated hair swatches were placed in a controlled oven set to standard room temperature and pressure and allowed to completely dry. Curliness, Volume and Alignment Coefficient were quantified using a polarization imaging device of the model RUMBA, as provided by Bossa Nova Technologies [4].

The visual assessment of Rumba images by volunteers was structured in two phases: the first focused on developing an atlas for hair types III-IV, and the second on hair types V+. Table I outlines the experimental design for each phase of the study.

Using an internal software, a paired analysis [5,6] was conducted with volunteers with curly hair profiles. During the analysis, a set of pictures was presented to each of them in pairs, the observers were asked to select the image that best answered the question posed to them. The test was conducted in a controlled environment. Images were displayed on a screen of 1920*1200 pixels positioned one meter away from the observer. Each observer's chair was adjusted to ensure the screen was at eye level. The room lights were kept off, and only the volunteer remained in the room during the test to further minimize distractions.

The cohesion analysis between the volunteer was done according to Spearman's correlation analysis which permits to face up the case where volunteer/observers' responses are not distributed normally.

After the exclusion of bad correlation observers, the images evaluated were ranked and grouped according to their probability of being chosen compared to another one, following Bradley-Terry-Luce model [7].

To create the visual Atlas, an expert defined the number of grades. The images with the worst and best ranking are chosen as the Atlas boundaries. To select the intermediate images, the following equation was employed.

$$s = \frac{M - m}{N - 1}$$

Equation 1 - Constant perceptive step calculation

Where S is the constant perceptive step between two images of the atlas, M and m are the maximum and minimum bounds of the atlas respectively and N is the number of grades.

The final validation of the visual Atlas is done by the expert.

Study phase	Hair Tresses (Origin, Curls pattern, weight, length)	Number Of Images	Total number of comparisons during paired analysis	Volunteer's profile	Number of volunteers	Question posed
01	Brazilian, CP III, 2.7g, 27 cm	20	190	Curly hair type II-IV	26	Which image best represents a hair tress with well-defined curls?
02	Multi-ethnic, CP V+, 2g, 20 cm	29	406	Curly/ Very curly hair type V+	32	Which image shows a hair strand with the most defined shape?

Table I - Study's experiment design for Paired Analysis.

The raw data obtained from the RUMBA equipment and scores assigned to each image were compiled and analyzed using an internal machine learning software. A total of 6 algorithms were explored to determine the best fit for the dataset, which comprised of: Supply Vector Method, Random Forest, Neural Network, Linear Regression and Polynomial Regression (2nd and 8th degree).

The internal software is able to choose the best fit based on the calculation of the Root Mean Squared Error (RMSE), as denoted in Equation 2.

$$RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(y_i - p_i)^2}$$

Equation 2 - Root Mean Squared Error (RMSE).

Where n = number of data points, yi = observed experimental value, and pi = predicted value.

Following algorithm selection, the selected model was carried out to extrapolate RUMBA outputs for the three key parameters: Standard Deviation (STDev), Volume, and Alignment.
Additionally, the internal software is also able to predict the Curl Definition Score (CDS) for any given variation formula that exists within a data set, by extrapolation using the best fit method. This means that the impact of slight variations of raw materials and concentrations on the Curl Definition score can be predicted. However, due to confidentiality purposes, the exact algorithms and datasets used in this work cannot be further elaborated.

With means of validating the accuracy of the proposed extrapolation method provided by the internal software, an additional step was also carried out. From the 20 formulations used in the Atlas validation step, 18 new formulas with small variations of the original set were created, manufactured, and applied onto new hair swatches. Subsequently, the hair swatches were analyzed in RUMBA and the outputs were inserted in the internal software, for prediction of the CDS. Method accuracy was also assessed through RMSE, by comparing predicted values (before manufacturing the formulas) with real values, obtained after manufacturing, and applying the formulas.

All the tests involving image processing and output analysis via Machine Learning were carried out in an Intel Core i5 2.6GHz computer, with 15GB RAM and 8 processors. Average algorithm processing time in the internal software was 210 s for each condition.

Results.

Figure 2 presents the graphs generated with the scores of the ranked images after the paired analysis. The rectangles in the graph highlight images that were grouped together, meaning that there was no statistically significant difference between them. Some of the images assessed during the test are also illustrated below the graph.

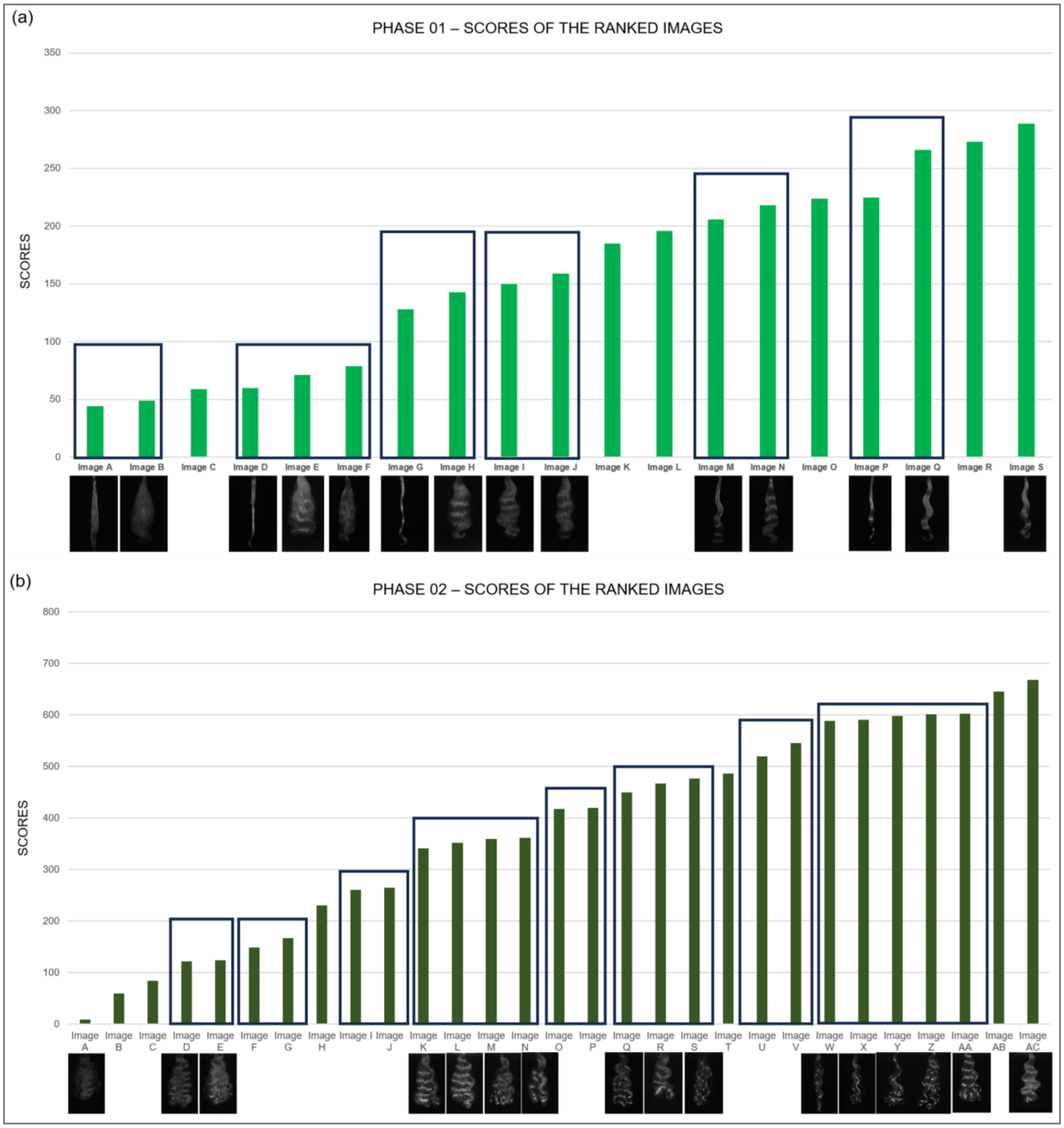

Figure 2 - scores of the Ranked Images (a) phase 01, (b) phase 02.

After expert evaluation of the ranked images, for Curls pattern III/IV target, the Atlas resulted in 6 grades, in which grade 1 represents the worst Curls Definition score according to volunteers' perception and 6 the best score.

For the test conducted with hair type V, it was observed that very different Curls shape results post swatch application leaded to the same score, the expert did not validate the atlas even after an attempt to arrange the images into 5 grades.

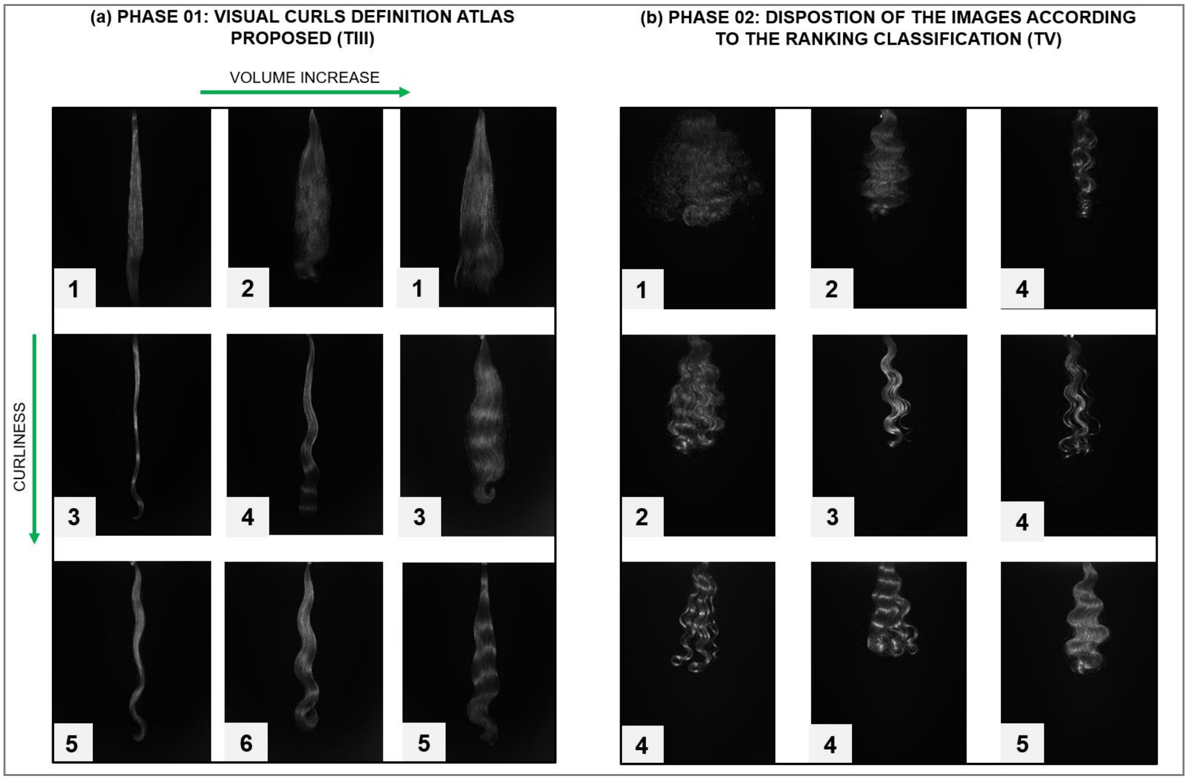

Figure 3- Phase 01: Visual Curls Definition Atlas Proposed TIII (a); Phase 02: Disposition of the Images According to the Ranking Classification TV (b).

Figures 4 represent the RMSEs of the Machine Learning models analyzed in the internal software, for the studies of Brazilian Natural Type III hair (a) and Multiethnic Type V hair (b).

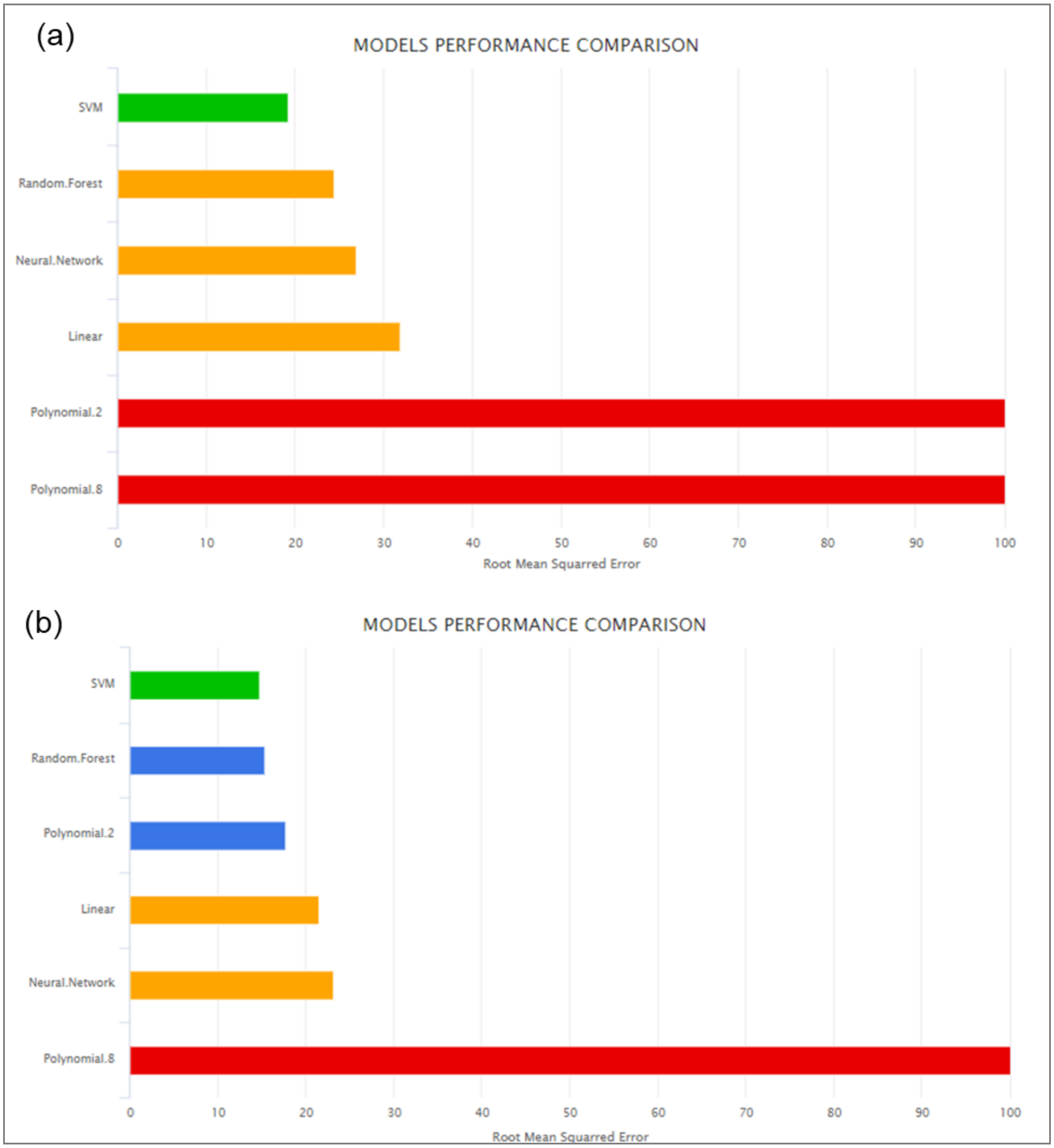

Figure 4 – Model comparison for the studies: Brazilian Natural Type III (a); Multiethnic Type V (b).

Following algorithm selection, the SVM model was carried out to extrapolate and correlate RUMBA outputs (Alignment, Integral and STDev) with the Atlas Scores. Results can be seen in Figure 5 for both hair types.

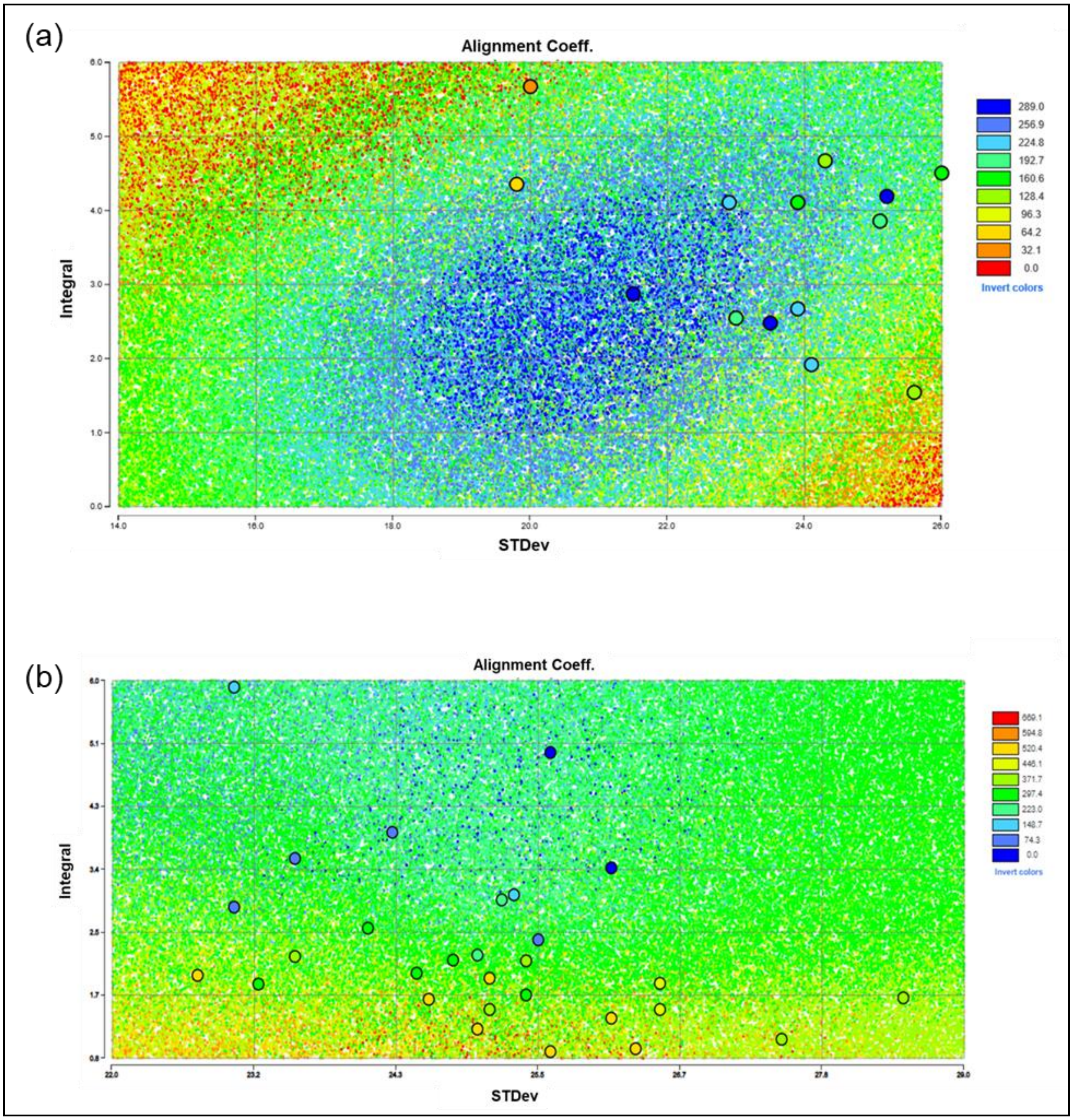

Figure 5** - **Correlation and extrapolation of RUMBA outputs with ATLAS score of Curl Definition for the studies**: **Brazilian Natural Type III (a); Multiethnic Type V (b).

To evaluate the reliability of the SVM-based extrapolation method on Type III swatches, we compared the predicted values against real-life results obtained from RUMBA measurements and their subsequent insertion in the software to calculate the ATLAS score. Visual assessments of treated hair on volunteers prediction and experimental results can be seen in Table II, alongside with the calculated RMSE.

Formula	Method-predicted Curl Definition Score	Real Curl Definition Score	Square Difference
Assay 1	216.8	240.0	537.9
Assay 2	228.7	234.1	27.9
Assay 3	230.7	237.3	40.1
Assay 4	221.3	228.0	44.8
Assay 5	225.9	255.4	847.7
Assay 6	231.2	231.6	0.0
Assay 7	228.2	210.0	330.8
Assay 8	245.6	249.2	11.9
Assay 9	240.3	243.2	7.4
Assay 10	219.8	224.8	27.5
Assay 11	250.1	263.6	193.8
Assay 12	236.5	237.1	0.2
Assay 13	214.8	228.9	198.8
Assay 14	221.1	230.7	92.1
Assay 15	226.2	223.2	9.0
Assay 16	231.6	186.6	2025.0
Assay 17	220.5	231.9	129.9
Assay 18	231.6	243.0	129.9
RMSE (according to Equation 1)	16.1		

Table II – Comparison between CDS values predicted by the SVM method and experimental results obtained through RUMBA and insertion in the internal software, for the Type III study.

Discussion.

According to the visual Atlas created for the hair target type III-IV (Fig 3.a), it is evident that Curl Definition is a balance between Volume and Curliness. Hair strands with low or high volume show the same probability of being selected if they present a similar level of curvature. The image that represents the best-case scenario for curl definition is the one where the strand displays intermediate volume and more pronounced curliness, highlighting the importance of considering human perception when evaluating the curl definition attribute.

It is possible to see from the results that the Support Vector Machine (SVM) method was the optimal model for the datasets of both hair types, as it demonstrated the lowest Root Mean Squared Error for the two studies. For the Type V study, Random Forest and 2nd Degree Polynomial also presented satisfactory results (Figure 4).

Figure 5.a showed that a strong correlation can be observed between Curl Definition and RUMBA outputs for hair target type III. By analyzing the graph generated by the Machine Learning software, it is possible to see well delimited zones of Curl Definition Score. A zone with the highest scores can be seen in the center, and it is outlined by zones with increasingly

lower scores, and therefore lower perceptions of Curl Definition. This finding suggests that a balance between the three RUMBA parameters play a significant role in defining the perception of curl definition in this hair type, as extreme values of alignment, integral and STDev are all associated with low scores according to the algorithm.

Contrary to the findings for type III hair, for the target hair type V, analyzing the image ranking (Fig 3.b), it's noticeable that hair strands with significantly distinct shapes from each other showed the same probability of being chosen. According to the expert's assessment, it wasn't possible to establish a relationship between the shape of the evaluated strands and the scores assigned after the paired analysis, rendering the creation of a visual atlas for this hair type unfeasible.

The results of the correlation and extrapolation of RUMBA outputs with Atlas score (fig 5) confirmed the behavior observed by the expert, no significant correlation was found between Atlas score and the output parameters for Curls Pattern type V. No clear zones of high and low score can be seen in a significant manner. Additionally, it is also possible to see the overlap of low and high scores, especially for the zones closest to the X axis: several points of the highest score (669.1) can be seen overlapping with intermediate (223.0) and low scores (74.3). In the case of this curl pattern, no individual parameter, nor a balance between them, seems directly impact the perception of Curl Definition. This suggests that additional factors might contribute to the perception of curl definition for this hair type and the algorithm needs to be further refined in order to generate reliable and consumer centric results. Further investigation utilizing a broader range of parameters and potentially different analytical approaches is necessary to develop a more representative model for Type V hair.

Conclusion.

With the experimental results encountered in this study, we have created a simple, cost-and-time effective method to incorporate the consumer's perception into the methodological development. To the best of our knowledge, no other work has proposed a method to correlate curl imaging with consumer perception, especially not for curly/ very curly hair. Due to its objectiveness, the method can also easily be extrapolated for other types of hair and other product performance parameters. Furthermore, it was possible to understand that Curl Definition is a complex attribute to be measured and with different meanings depending on the consumer's profile. For Type III-IV hair, we identify a clear relationship between Curl Definition and the parameters measured. However, for the Type V+, no clear relationship was found. This variation outlines the additional methodological challenge for type V+ hair, and importance of bringing consumer perception to the center of the development process.

Acknowledgments.

The authors wish to thank Margalith Harrar for the development of the software for paired analysis.

Conflict of Interest Statement.

L'Oréal Brazil provided the resources for this study.

References.

1. De la Mettrie R, Saint-Léger D, Loussouarn G, et al. Shape variability and classification of human hair: a worldwide approach. Hum Biol 2006; 79.
2. Loussouarn G, Garcel AL, Lozano I, Collaudin C, Porter C, Panhard S, Saint-Léger D, de La Mettrie R. Worldwide diversity of hair curliness: a new method of assessment. Int J Dermatol. 2007 Oct;46 Suppl 1:2-6. doi: 10.1111/j.1365-4632.2007.03453.x. PMID: 17919196.
3. EVANS, Trefor; SCHUELLER, Randy. Practical Modern Hair Science. 1. ed. Nova Iorque: Springer, 2011.
4. Fiber orientation measurement using polarization imaging, J. Cosmet. Sci, 62,85-100, March-April 2011.pdf (bossanovavision.com)
5. FRANCIS B., DITTRICH R., HATZINGER R., PENN R. Analysing partial ranks by using smoothed paired comparison methods: an investigation of value orientation in Europe. Applied Statistics, vol. 51, p. 319-336 (2002).
6. BROWN T, PETERSON G. *An Enquiry Into the Method of Paired Comparison*. GeneralTechnical Report (2009).
7. MATTHEWS J.N.S, MORRIS K.P. An application of Bradley-Terry models to the measurement of pain. Applied Statistics, vol. 44, p. 243-255 (1995).