"Predicting Performance in Cosmetics Formulations Through Algorithm-Driven Design of Experiment"

Cyrus Andriolo[1], Gizele Freitas[1] and Liliane Duarte[1]

[1] L'Oréal Research and Innovation, Rio de Janeiro, Brazil.

1. Introduction

Driven by consumer demand for enhanced efficacy and sustainability, the cosmetics industry is developing high-performance, eco-friendly ingredients. This trend is evident in the 190 new launches showcased at IN-COSMETICS GLOBAL 2024 — more new ingredients than ever before [1] — which offer numerous innovation opportunities for formulators. While these ingredients offer immense potential, two key questions arise: How can one effectively navigate this vast landscape to pinpoint ingredients with targeted properties? And, once identified, how can one understand the synergy between them and their optimal concentrations?

In a classical approach, a formulator might identify, for example, four promising raw materials and, wanting to investigate five different concentrations of each to find the best synergy between them. Which would require to make 625 unique formulations, consuming an extensive amount of resources and time.

With a smart Design of Experiments (DOE) [2], we can extract maximum information with a limited number of observations. Using machine learning algorithms its possible to explore formula data and predict outcomes, allowing to find the most relevant formula. In this way, in contrast with the classical approach, the DOE strategy requires just 21 combinations for that experiment and will yield the most relevant formula, considering the best concentration of those ingredients within the specified concentration range. Furthermore, an algorithm had been created, capable of predicting the results of any combination without requiring physical formulation.

In this work, we applied a two-stage Design of Experiments approach to improve a leave-on gel-cream (referred as gel-cream GC). This DOE approach made it possible to integrate and filter these innovations into product development, by ranking novel ingredients and evaluating multiple formulation objectives with a minimal number of experiments. Moreover, it leverages digital acceleration tools to generate an algorithm that predicts the results of compositions using the raw material combination selected without requiring physical formulation.

2. Materials and Methods

Ingredients selection: A rigorous selection process was undertaken to identify suitable candidates for formula improvement focused on 4 deliverables: Texture, conditioning, spreadability and formula footprint.

Formulation: All the prototypes deriving from Gel Cream GC were prepared using VMI Rayneri tool equipped with rotor/strator disperser with wide slots for viscous and fine emulsions. The concentration range of the selected ingredients was 0 till 4%.

Swatches: The swatches used in wet combing test were provided by the IHIP company. They were standardized virgin straight hair swatches with 2.7 grams of hair, they were shampooed prior to testing and let dry at room temperature.

Wet Combing: The analysis was performed in triplicate with straight hair swatches in a Texture Analyzer TA XT Plus equipment from Extralab. Hair support frame with comb holder with 30kg load-cell in compression mode with test speed of 5 mm/s. Calibration was made with height calibration of 2kg. Analysis was made by software Exponent v6.1.16.0.

Curls Definition: Was determined by the *in vitro* Atlas of Curls developed in IFSCC 2024 podium presentation SE 64 [3] on curly hair.

Texture evaluation: Visual evaluation by panelists through a binary analysis of 1 for positive visual texure and 0 for a negative one.

DOE Screening: A screening Design of Experiments (DOE) was conducted to evaluate 13 ingredients, categorized into three groups based on their function: five texturizers, four spreading agents, and four conditioning enhancers. These ingredients were incorporated into the algorithm at fixed concentrations, resulting in 18 formulations that were subsequently prepared and evaluated. Measurements of wet combing and curl definition were taken, and the software analyzed the factorial design effects to rank the ingredients within each group.

DOE Optimization: The four highest-performing ingredients — two texturizers, one spreading agent, and one conditioning enhancer — were selected for Design of Experiments optimization. Their concentrations varied between 0% and 4% and input into the optimization algorithm. Twenty-one formulations were then created and evaluated. The resulting curl definition and wet combing data were used to train a predictive model. This algorithm forecasts the performance of any combination of these four ingredients, including untested combinations through extrapolation.

Sensorial Evaluation: To complement the *in vitro* evaluation, *in vivo* analysis was performed. Formulas were applied to six human volunteers by expert hairdressers. Each volunteer's hair was divided into two sections. The reference formula was applied to one side, and the test formula to the other. The hairdressers evaluated texture, visual appearance, curl definition, and ease of detangling.

3. Results

3.1 DOE Screening – Ingredient selection

Gel-cream GC is a leave-in product designed to define curls. However, consumer feedback indicated opportunities for improvement in texture, detangling, and spreadability, without impacting in the curls definition. To address these three attributes, a two-stage DOE optimization strategy was employed.

The first stage of the methodology is the DOE screening that investigated 13 ingredients categorized into three functional groups: five texturizers, four spreading agents, and four conditioning enhancers. A design algorithm, using fixed ingredient concentrations, determined that 18 formulations were sufficient to extract maximum information. Curl definition was measured to ensure it wasn't negatively impacted, and wet combing was evaluated as a proxy for detangling and spreadability atributes. The resulting data were analyzed using factorial design analysis to rank the ingredients within each group (Figure 1).
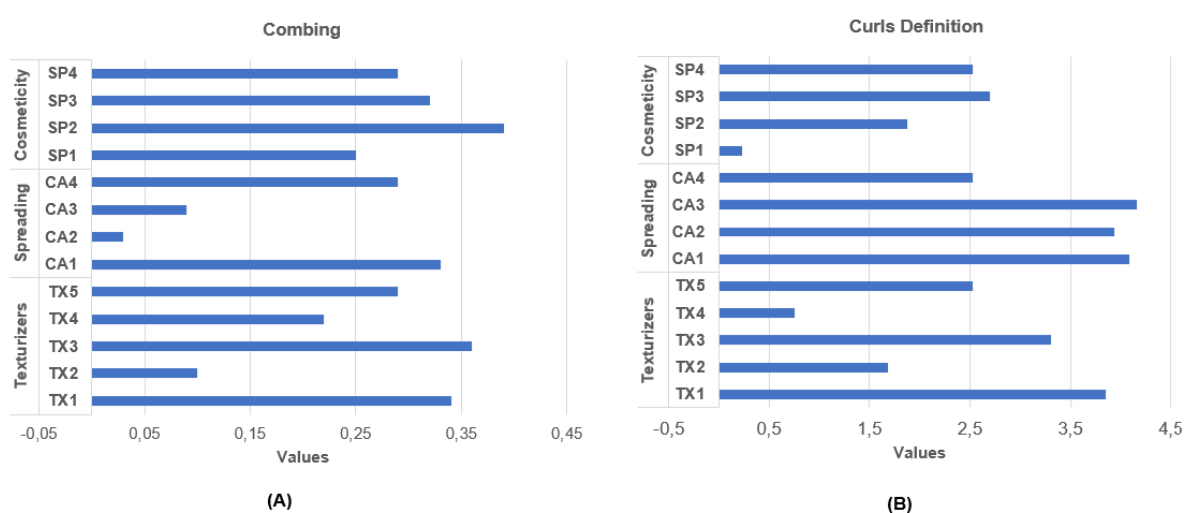

Figure 1. Ranking of (A) wet combing attribute for each group. (B) Curls definition attribute. Longer the bar, better is the atribute.

3.2 DOE Optimization – Concentration optimization predictive algorithm

The four highest-performing ingredients from the screening phase (two texturizers, one spreading agent, and one conditioning enhancer) were then selected for further optimization. Their concentrations were varied between 0% and 4%, and the optimization algorithm determined that 21 formulations were necessary for maximum information gain.

Data from these 21 formulations, including curl definition, texture assessment, and wet combing, were used to train a predictive model. The figure 2 provides a multidimensional visualization of the response surface generated by the algorithm. Each large dot represents a tested formulation, color-coded to indicate performance. Smaller dots represent extrapolated

predictions for new, unmanufactured compositions. A Support Vector Classifier (SVC) model achieved a Learning estimated model average error of 2.723%, suggesting its effectiveness in constructing the predictive surface [4].
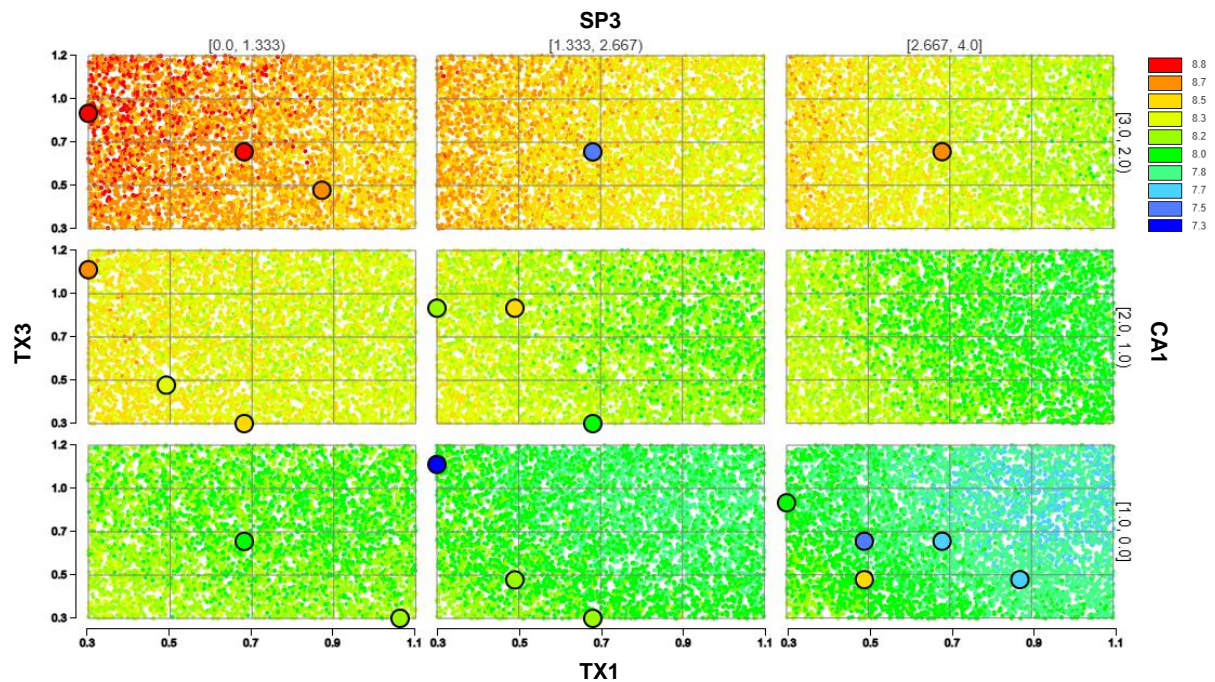

Figure 2. Multidimensional visualization of the surface that the algorithm creates to foresee the results of the attribute. The large dot represents a tested formulation, color-coded to indicate performance – warmer colors, better the result. Smaller dots represent extrapolated predictions for new, unmanufactured compositions.

The generated algorithm also shows the influence of each ingredient in the measured attribute by a graph bar (Figure 3) that ranging from -1 to 1, indicate both the magnitude and direction of its influence. A high absolute value signifies a strong influence, while the sign (positive or negative) indicates whether increasing the ingredient's concentration increases or decreases the attribute value. This analysis allowed validation of initial hypotheses, confirming that the ingredient predicted to have the greatest impact on a specific attribute was indeed the most influential. Based on these insights, the algorithm was used to identify the top-performing prototypes with the highest combined scores for curl definition, combing, and texture.
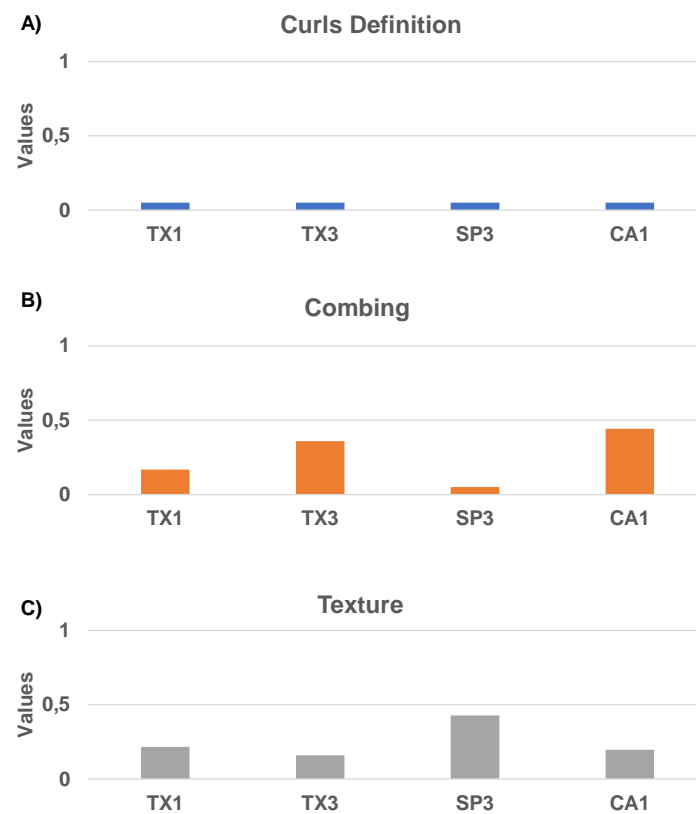
A)

Curls Definition

B)

Combing

C)

Texture

Figure 3. The linear correlation between variable and the attribute: a) Curls defintion, b) Combing and, c) Texture.

Two promising formulations were manufactured and evaluated, confirming the model's predictive accuracy (Table 1). The relative percent errors were below 7%, demonstrating good agreement between predicted and experimental results. However, Prototype 1 displayed superior results, exhibiting a 35% improvement in both combing and texture spreadability compared to the inicial formula Gel-Cream GC.

Table 1. Relative Percent Error of the model.

	Prototype 1	**Prototype 2**
Curls Defintion erro	4%	7%
Wet Combing erro	-7%	6%

Finally, Prototype 1 underwent sensory evaluation by expert hairdressers that applied the prototype on volunteers comparing it to a market benchmark confirming the potential of the formula, with strengths on curl definition, overall visual, and a positive "slippery" and easy-to-distribute texture.

4. Discussion

To improve Gel Cream GC we conducted an ingredient selection process based on reported benefits from literature and supplier data. Prioritizing sustainability, promising candidates were further screened for their eco-friendly attributes. This process resulted in 13 ingredients chosen for their potential to improve Gel Cream GC's texture spreadability, and overall cosmeticy. These ingredients were categorized into three functional groups: five texturizers, four spreading agents, and four conditioning enhancers.

A key challenge was to evaluate these ingredients while accounting for potential interactions between the different functional groups. Independent assessment of each group would be resource-intensive and might lead to suboptimal results, as a high-performing ingredient in one category could negatively affect another category. For instance, the best texturizer might negatively impact spreadability, or the top conditioning enhancer could alter the desired gel-cream texture. Therefore, a Design of Experiments (DOE) approach was employed to efficiently investigate the ingredient and identify optimal combinations.

The first stage DOE screening methodology determines the minimum number of formulations needed to extract maximum information, using the variance inflation factor (VIF) as a guide. The VIF measures the correlation between independent variables and reflects the precision of the model [5]. A VIF below 2.0 indicates high statistical confidence, while values below 2.5 are generally considered acceptable. While an initial DOE suggestion of 11 formulations yielded a VIF of 2.52, 18 formulations were chosen to achieve a lower VIF of 1.57, enhancing the statistical power of the experiment.

This initial screening phase evaluated the 13 ingredients across the three functional categories. Wet combing and curl definition were measured for the 18 formulations (Figure 1). Figure 1A illustrates the positive impact of various ingredients on wet combing, while Figure 1B highlights the ingredients contributing to curl definition. This screening phase successfully identified the top performers within each category.

The four highest-performing ingredients—two texturizers (TX1 and TX3), one spreading agent (SP3), and one conditioning enhancer (CA1)—were selected for further optimization. The second stage DOE optimization was conducted, varying their concentrations between 0% and 4%. 21 formulations were created and evaluated to train a predictive model by Support Vector Classifier (SVC) which presented a learning estimated model average error of 2.723%. The model enables performance prediction based on ingredient combinations, including extrapolation to untested formulations. The large dots in Figure 2 represent the tested formulations,

color-coded to indicate performance. Smaller dots represent extrapolated predictions for new, unmanufactured compositions.

As shown in Figure 3, all selected ingredients positively impacted the targeted attributes, although not all at the same level. While the impact on curl definition was minimal across all ingredients (as expected, given that their primary function was not curl enhancement), the objective was to maintain, not negatively impact, curl definition while improving other attributes.

Based on the predictive model, two promising prototype formulations were selected for further evaluation. Table 1 compares the predicted and experimental performance of these prototypes, demonstrating the model's accuracy with prediction errors below 7%. Visual assessment of the texture also aligned with expectations.

Compared to the original Gel Cream GC, the optimized Prototype 1 formulation demonstrated a 35% improvement in wet combing, along with enhanced stability and a reduced environmental footprint, while maintaining curl definition. Finally, expert hairdressers compared Prototype 1 to a market benchmark, confirming its potential. The prototype excelled in curl definition and overall appearance with a positive slippery and easy-to-distribute texture. Allowing the formula to be a potential candidate for a consumer usage test.

5. Conclusion

This two-stage DOE (Design of Experiments) methodology provides a powerful and efficient approach for multi-objective formulation optimization. Requiring only 32 formulations compared to potentially 625 needed with a traditional method, this approach significantly reduces resource expenditure. Furthermore, the generated predictive algorithm offers a versatile tool for future product development, enabling prediction of critical formulation attributes such as texture, stability, and cosmeticity.

This study successfully leveraged this two-stage DOE approach to optimize Gel Cream GC, a leave-in curl defining product. This method could strategically screen a diverse set of ingredients and subsequently optimize the concentrations of the top performers, reaching a significantly improved formulation. The predictive model proved highly accurate, facilitating efficient exploration of the formulation and enabling the identification of optimal ingredient combinations. The optimized formulation demonstrated a 35% improvement in both wet combing and texture spreadability compared to the original Gel Cream GC, along with enhanced stability, while maintaining curl definition and minimizing environmental impact.

Sensory evaluation confirms the DOE effectiveness in delivering the key attributes investigated in order to select the Prototype 1 for further evaluation in consumer usage test.

This approach can be extended to evaluate other relevant attributes, such as viscosity, volume, frizz control, shine, and other properties of interest. By leveraging digital tools, this

methodology facilitates resource optimization and accelerates the development of sustainable, high-performing cosmetic products.

6. References

[1] Article: The Innovation Zone at in-cosmetics Global to Showcase Over 190 Recent Products Apr 9th, 2024 From Reed Exhibitions / in-cos Global

[2] ERIKSSON, L. et al. 2000, Design of experiments. **Principles and Applications, Learn ways AB, Stockholm**.

[3] Barros, D et, al. 2024 **Embracing Diversity: Bridging the Gap for Curly & Very Curly Formulations with In Vitro and Machine Learning Performance Prediction.** SE-64 Podium IFSCC 2024

[4] ZHOU, Yinzuo; CHEN, Weilun. 2025, A Link Prediction Algorithm Based on Support Vector Machine. **Available at SSRN 5137127**.

[5] MONTANEZ, John Joshua F. 2025, Detection of Soil Nitrogen Levels via Grayscale Conversion: A Full-Factorial Design of Experiment Approach. **International Journal of Computing**, v. 17, n. 1, p. 1-10.