Evaluation of Stability and Functional Packaging Effectiveness for Ferrous Ascorbate Tablets Coated with an Aqueous Moisture Barrier Film Coating Formulation

Authors: Jason Hansell, John Horan, Tom Mehaffey, Anne Fuller and Ali Rajabi-Siahboomi,

Colorcon, Inc. Harleysville, PA 19438, USA | AAPS Poster Reprint 2025

Challenge

Color is a critical quality attribute of solid oral dosage forms; it influences brand identity, reduces medication errors, and increases patient acceptability. The color matching formulation process is traditionally iterative and leverages the experience of the formulator to produce high quality matches. Ingredient selection is further complicated by the regulatory requirements of the target markets. The aim of this study was to develop an AI model that predicts the CIELAB color values of film coating systems based on raw material inputs and is flexible enough to be implemented with an integrated compliance check.

Method

A comprehensive and proprietary databank encompassing over55,000 unique film coating formulations was compiled (Colorcon Inc.). This dataset included more than 300 distinct raw materials, pigments, and opacifiers, and included broad coverage across the CIELAB color space. The AI model was optimized to predict LAB values, utilizing mean squared error (MSE) as the loss function. Color difference (ΔE) was used to validate the model prediction.

Color difference (ΔE) and mean squared error (MSE) are foundational methods for measuring distance between two points, and both utilize Euclidean distance for path determination. Color difference, expressed as;

$$\Delta E = \sqrt{[(L_1 - L_2)^2 + (a_1 - a_2)^2 + (b_1 - b_2)^2]}$$

quantifies perceptual differences between two colors in the CIELAB space. This mirrors the structure of the mean squared error, a loss function commonly used in machine learning, formulated as;

$$MSE = (1/n)\ \Sigma_i\ (Y_i - \hat{Y}_i)^2$$

where n represents the number of samples, Y_i the true values, and $\hat{Y}_i$ the predicted values. In the context of color prediction, the squared differences in ΔE and MSE both capture deviations between predicted and actual values, whether those are L, a, b coordinates or other features. Thus, minimizing MSE in model training is directly aligned with minimizing perceptible color difference, making it ideal for color prediction applications in pharmaceutical and nutraceutical film coating development.

Validation of the model's performance included a set of formulations separated from training data at the time of model development, and a final assessment of over 3,000 new formulations that were isolated from the training and validation data sets. The optimized model was then used to generate over30 million distinct colors to explore color space expansion for pharmaceutical and nutraceutical film coating applications.

Results

The ΔE value, calculated from the perceptually uniform CIELAB color space (Figure 1), represents the total color difference between two samples. Generally, a ΔE value below 2–4 is considered visually imperceptible or difficult for the human eye to detect, depending on the color and shade. The AI model achieved an average ΔE of 2.98 on independent test data, indicating its ability to predict colors within a perceptually acceptable range across various film coating systems and raw materials.

Figure 2 presents the model's capacity to analyze the effect of pigments and their concentration on film coating formulation color. For example, the use of yellow iron oxide in a polyvinyl alcohol (PVA) based film coating system illustrates how the model assesses pigment effects at different usage levels and their impact on specific color channels. These results show the model's effectiveness in understanding the quantitative contributions of ingredients to the color of film coating systems.

During a final assessment covering more than 3,000 newly formulated products, the model maintained consistent performance with an average ΔE of 2.99 (Figure 3), demonstrating generalizability over a wide variety of immediate release film coating formulations, which included a diverse set of base systems, opacifiers (such as TiO_2, $CaCO_3$), and various pigments.

Potential improvements have been identified, including further analysis of formulations with lower predictive accuracy. As shown in Figure 4, formulations featuring unique ingredient combinations with limited observations (<10) showed reduced performance, with a 1 ΔE increase compared to formulations with thousands of similar data points. This trend was particularly noticeable when the material's primary function was a pigment, as these influence color shade more significantly than other components.

Figure 1. CIELAB Color Space

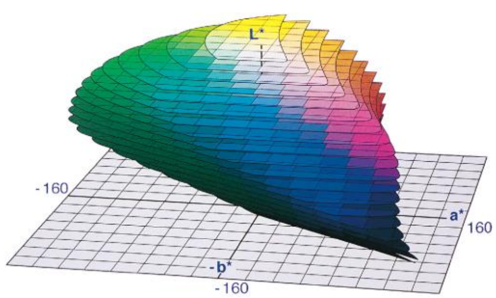


Figure 2. Predicted vs. Actual CIELAB values for yellow iron oxide film coating (test data).

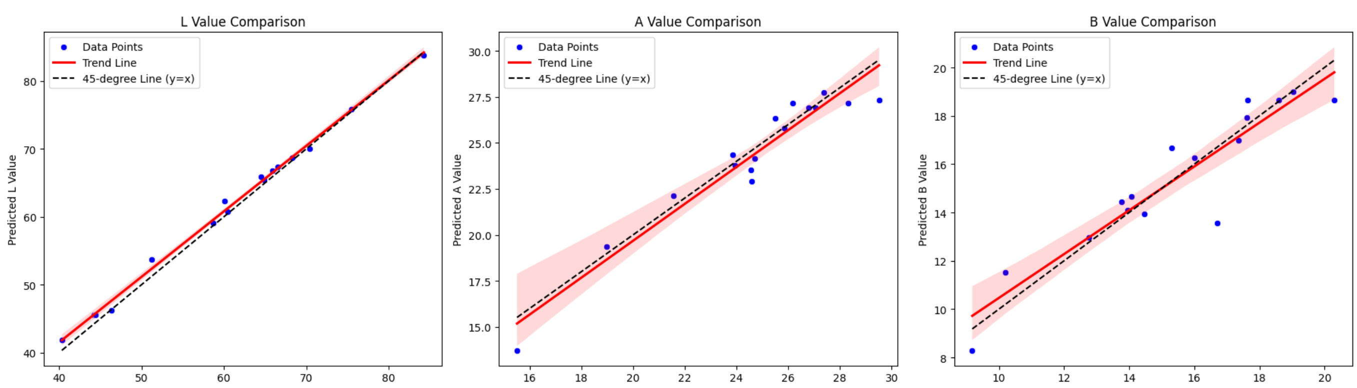


Figure 3. Distribution of color difference values between actual and predicted CIELAB values for 3000+ post-development formulations.

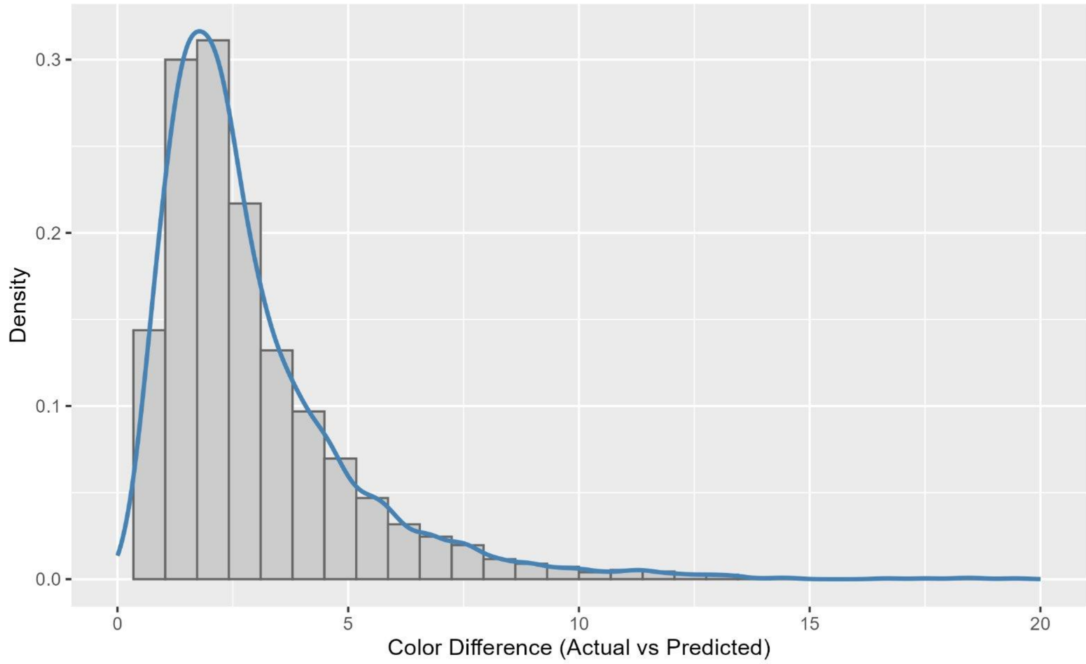
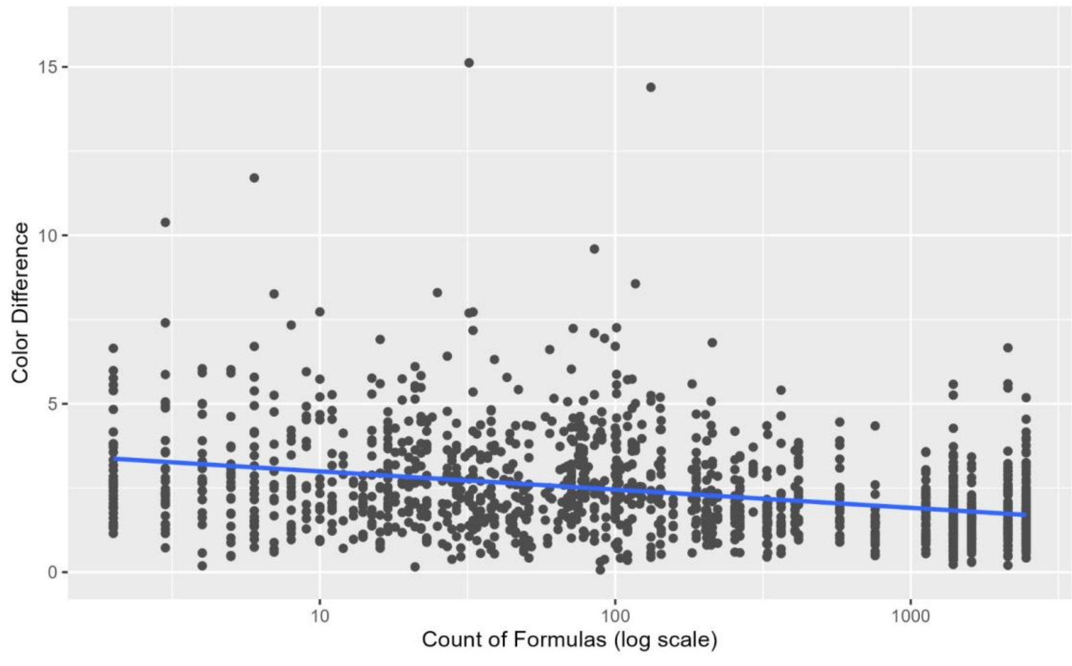


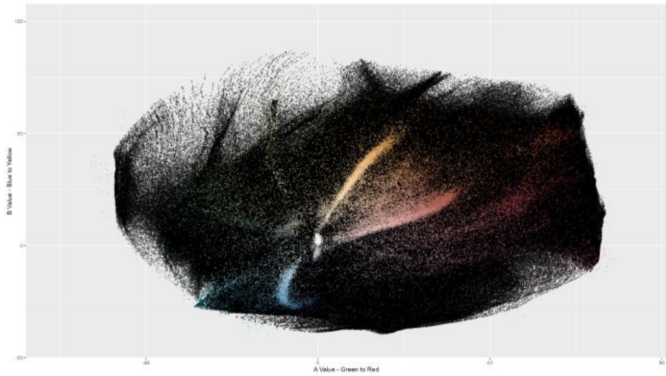
Figure 4. Impact of formulation sample size on model performance.



Conclusion

Building on the reliable model performance, an expanded color space encompassing 30 million theoretical formulations was developed. As illustrated in Figure 4, this approach notably broadened the available color space relative to the historical formulation database. This expanded color space brings value by offering new ways to address current challenges in pharmaceutical and nutraceutical product development—such as adapting to changing regulations and meeting evolving consumer preferences.

Figure 5. Historical color library (colored points) vs AI color library (black points) demonstrating expanded color space available through unique pigment combinations.



In this study, an AI-driven model was investigated and developed that can predict color from diverse film coating formulation compositions with low ΔE values. The model also allowed expansion of color space for regulated markets. This model not only accelerates product development and minimizes material waste from repeated testing but also ensures proactive adherence to changing regulatory landscapes.

References

Journal Article (Hill et al., 1997)

Hill, B., Roger, T., & Vorhagen, F. W. (1997). Comparative analysis of the quantization of color spaces on the basis of the CIELAB color-difference formula. *ACM Transactions on Graphics, 16*(2), 109–154. https://api.semanticscholar.org/CorpusID:13561124

Website / API Documentation (MSE from Keras)

Keras Team. (n.d.). *Mean squared error loss function*. Keras. Retrieved September 17, 2025, from https://keras.io/api/losses/regression_losses/

Contact your Colorcon representative or call:

North America
+1-215-699-7733

Europe/Middle East/Africa
+44-(0)-1322-293000

Latin America
+54-11-4617-8640

India
+91-832-6727373

China
+86-21-61982300

www.colorcon.com